Springer Protocols

Methods in Molecular Biology 543

DNA-Protein Interactions

Principles and Protocols

Third Edition

Edited by
Tom Moss
Benoît Leblanc

Humana Press

METHODS IN MOLECULAR BIOLOGY™

Series Editor
John M. Walker
School of Life Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For other titles published in this series, go to
www.springer.com/series/7651

DNA-Protein Interactions

Principles and Protocols

Third Edition

Edited by

Tom Moss*and Benoît Leblanc†

** Centre de Recherche en Cancérologie de l'Université Laval, Québec City, QC, Canada*
† Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

Humana Press

Editors
Tom Moss
Centre de Recherche en Cancérologie
de I´Université Laval
Québec City, QC
Canada

Benoît Leblanc
Département de Biologie
Faculté des Sciences
Université de Sherbrooke
Sherbrooke, QC
Canada

ISSN: 1064-3745 e-ISSN: 1940-6029
ISBN: 978-1-60327-014-4 e-ISBN: 978-1-60327-015-1
DOI: 10.1007/978-1-60327-015-1
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2008942788

Cover Background Image:
Chapter 22, Fig. 5A. Evaluation and exploitation of a 2-D crystallization experiment. Histidine-tagged yeast RNA polymerase I is incubated with nickel chelating lipids. (A) Low magnification electron microscopy image showing the organization of the protein complex into domains. The bar represents 5 μm

Cover Inset Images:
Chapter 1, Fig. 1. The structure of the minimal MetJ repressor-operator complex (pdb 1cma). The sequence of the consensus two met box operator (top strand) is shown alongside from 5' (top) to 3' (bottom). The repressor is shown as a ribbon representation with one subunit shaded. The corepressor molecules (AdoMet) are shown as ball and stick representations and the DNA operator is a framework model. (I am grateful to Prof. Simon Phillips and his colleagues for production of this figure.)
Chapter 9, Fig. 3. Summary of the results. Bottom: Space filling representation of the operator site showing the positions where ethylation results in complete inhibition of complex formation (highlighted in dark gray). MetJ dimers bind in the major groove with contacts between adjacent protein dimers above the minor groove at the symmetry dyad of the oligonucleotide shown

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

It is a rote of modern molecular biology classrooms to stress the importance of DNA–protein interactions by pointing out that species such as the bonobo and the human being share more than 95% of their genes – a level of similitude far exceeding that of, say, the translations of the *Odyssey* by Robert Fitzgerald and by Samuel Butler. The nonnegligible differences between our hirsute cousins and us, the argument goes, must find its explanation not so much in what genes we both use but in when and to what extent they are used. Hence the truly vital importance of understanding how regulatory proteins, be they active or passive, activating or inhibiting, interact with the genetic material to see it expressed in the proper way.

We say vital not only in the strictly intellectual sense of understanding the basic mechanisms of the living world (although that would seem to be a worthwhile goal in itself) but because our lives as individuals may well soon depend on an ever more precise understanding of gene expression. Following the discovery and widespread use of antibiotics in the mid-twentieth century, and the massive campaigns of vaccination that eradicated smallpox and helped control ancient scourges such as polio and tuberculosis, our species has recently become free, for the first time in its history, of the constant and urgent threat of infectious diseases. Next on the list of challenges to our health come problems associated with our recent extended lease on life, problems that are statistically rare in young people but tend to crop up in our later days. Type II diabetes, cardiovascular problems, neurodegeneracy, cancer, these are the new specters that haunt each of our visits to the doctor's office. These diseases have multiple etiologies and we have had some success in identifying and correcting some of their causes (be they environmental, related to our lifestyle or triggered by certain pathogens). In the long run, however, we understand that truly vanquishing them will demand a thorough understanding of the molecular mechanisms that turn cells immortal, that cause neurons to start dying, or that make cells insensitive to insulin. The same will hold true of multiple but less heralded pathological conditions associated with an improperly expressed genome, which are certainly as important to the person enduring them as cancer would be to the population at large.

Of course, the relevance of DNA–protein interactions is not limited to biology, medicine, or pharmacology. The word "biotechnology" coined by the Hungarian Karl Ereky in 1919 originally described the use of preexisting biological material for the creation of products and services but has come to signify much more in recent years. Gene splicing and the creation of transgenic species have fairly revolutionized agriculture, and in a world of roughly 7 billion mouths to feed, that qualifies as a life-altering realization in more ways than one. As we learn more and more about the interaction between DNA and the proteins that see to its proper expression in an ever-increasing number of species, we will be in a position to augment the benefits we reap from nature while at the same time minimizing our footprint on the environment. Whether this goal will be reached or not is another matter, but it is certainly easier to devise more efficient and environmentally friendly agribusiness strategies with a better understanding of how the world actually functions than relying on wild guesses and wishful thinking. Not content with modifying existing species, we have also taken the first steps in devising new life forms that may help,

for example, cleaning up oil spills; the success of such endeavors will rely heavily on our understanding of gene expression.

This is the third edition of *DNA–Protein Interactions: Principles and Protocols.* Each edition is separated from its predecessor by roughly 7 years, a period that appears to be sufficient for the field to come up with wild innovations in technology and its applications. New protocols are added to each successive edition, sometimes based on the novel use of ancient techniques, sometimes relying on new technological advances, but most of the time on some new integration of the two. These new ways to do things allow us to study where, when, and to what extent proteins interact with DNA, and that with an unprecedented precision and on a larger scale than ever. The current edition has retained updated versions of the basic techniques, the old-but-trusted ones that may never go out of fashion, but the impetus for this edition is the recent development of in vivo and genome-wide interaction techniques that have taken the domain by storm. The results obtained by years of in vitro techniques make everyone curious about whether they can be translated to an in vivo context, with DNA packaged in chromatin, and with several signaling cascades and protein actors all acting their part in unison. We are therefore happy to present several chapters on chromatin immunoprecipitation and other in vivo techniques. Along with new chapters on topological studies, photocrosslinking, FRET, and imaging techniques, we believe that this new edition will prove a worthy and practical addition to the series. As always, following the enlightened format of Methods in Molecular Biology, each protocol is rich in notes that authors think are important but which rarely make it into the spartan protocols that accompany most research papers. In the past, such notes were always our favorite part of these protocols, and we are glad to once again be able to exclaim "so that's how they do it" while perusing the notes section. It should be mentioned that some protocols featured in previous editions did not make it this time around, as they were either unchanged or perhaps less frequently used, but they are luckily still available on the SpringerProtocols.com database.

Although rapid progress in analytical capacities is something to be very proud of, we think it would be presumptuous to claim that the study of DNA–protein interactions has "finally come of age" following the recent (and sometimes spectacular) advances described in these pages. That is certainly not meant to belittle such amazing approaches as microarray-based gene expression analysis, atomic force microscopy studies, or ChIP-on-chip cartography, all of which were technological quantum leaps in the field and important enough to warrant the use of that overwrought cliché. It is meant, however, to emphasize the speed at which our fast-evolving technology allows us to ask new questions and to provide new answers. In a sort of biological application of Moore's law, we now live in a world where our ability to generate data on DNA–protein interactions and our ability to analyze it seems to increase exponentially, while the costs involved are getting low enough that experiments that were once the sole purview of big laboratories become available to most everyone. (Granted, we still have some work to do regarding the lowering of costs, but we are getting there.) The ingenuity of the scientific and engineering communities, driven by more and more demands from the medical and biotechnology fields, is bound to give rise to even more startling analytical technologies in the future, and personally I cannot wait to read further editions of this volume. I am sure that I will be as amazed then as I was this time around.

The current edition would not have been possible without the pioneering work of Geoff Kneale on the first edition, or without the expansion and diversification made by Tom Moss in the second edition. Both have my gratitude, and I thank Tom for bringing me aboard as coeditor for this third edition.

Have fun in the laboratory.

Benoît Leblanc

Contents

Contributors

Sankar Adhya • *Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA*

Blaine Bartholomew • *Department of Biochemistry and Molecular Biology, Southern Illinois University School of Medicine, Carbondale, IL, USA*

Nathalie Bastien • *Département de Pédiatrie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada*

Alain R. Bataille • *Institut de Recherches Cliniques de Montréal, Montréal, QC, Canada*

Simon Blouin • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada*

Philippe Bouvet • *Université de Lyon, Ecole Normale Supérieure de Lyon, Lyon, France*

Douglas Browning • *School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK*

Michael J. Bumbulis • *Department of Biology, Baldwin-Wallace College, Berea, OH, USA*

Stephen Busby • *School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK*

Peter D. Cary • *Institute of Biomedical and Biomolecular Sciences, School of Biological Sciences, University of Portsmouth, Portsmouth, UK*

David R. Chafin • *Ventana Medical Systems Inc., Tucson, AZ, USA*

David J. Clark • *National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA*

Benoit Coulombe • *Institut de Recherches Cliniques de Montréal, Montréal, QC, Canada*

Timothy D. Craggs • *School of Physics and Astronomy, University of St. Andrews, St. Andrews, UK*

Colyn Crane-Robinson • *Biophysics Laboratories, School of Biological Sciences, University of Portsmouth, Portsmouth, UK*

Corinne Crucifix • *Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch-Graffentaden, France*

Céline Domecq • *Institut de Recherches Cliniques de Montréal, Montréal, QC, Canada*

Klaus Dose • *Institut für Biochemie, Johannes Gutenberg-Universität, Mainz, Germany*

Anatoly I. Dragan • *Department of Biology, Johns Hopkins University, Baltimore, MD, USA*

Régen Drouin • *Département de Pédiatrie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada*

Sergei Druzhinin • *Howard Hughes Medical Institute, Waksman Institute, Piscataway, NJ, USA*

Richard H. Ebright • *Howard Hughes Medical Institute, Waksman Institute, Piscataway, NJ, USA*

Rosy Favicchio • *Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas, Heraklion, Crete, Greece*

Diane Forget • *Institut de Recherches Cliniques de Montréal, Montréal, QC, Canada*

Mayra Furlan-Magaril • *Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, México City, México*

Alexander A. Gall • *Cepheid, Bothell, WA, USA*

Luc Gaudreau • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada*

Manon Gaudreault • *Oncology and Molecular Endocrinology Research Center, Laval University, Québec City, QC, Canada*

Nicolas Gévry • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada*

Marie-Eve Gingras • *Oncology and Molecular Endocrinology Research Center, Laval University, Québec City, QC, Canada*

Sylvain L. Guérin • *Oncology and Molecular Endocrinology Research Center, Laval University, Québec City, QC, Canada*

Cathy V. Gunther • *Division of Biological Sciences, University of Missouri, Columbia, MO, USA*

Jean-François Harrisson • *Centre de Recherche en Cancérologie de l'Université Laval, Québec City, QC, Canada*

Jeffrey J. Hayes • *Department of Biochemistry and Biophysics, University of Rochester Medical Center, Rochester, NY, USA*

Mike Heilemann • *Physics Department, University of Bielefeld, Bielefeld, Germany*

Ling Chin Hwang • *Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, UK*

Indu Jagannathan • *Department of Biochemistry and Biophysics, University of Rochester Medical Center, Rochester, NY, USA*

Brenda F. Kahl • *Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO, USA*

Achillefs N. Kapanidis • *Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, UK*

Younggyu Kim • *Howard Hughes Medical Institute, Waksman Institute, Piscataway, NJ, USA*

G. Geoff Kneale • *Biophysics Laboratories, Institute of Biomedical and Biomolecular Sciences, University of Portsmouth, Portsmouth, UK*

Annie Kolb • *School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK*

Simon Labbé • *Département de Biochimie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada*

Daniel A. Lafontaine • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, QC,* Canada

Marc Larochelle • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada*

Luc Lebeau • *Laboratoire de Synthèse Bio-organique, URA 1386 du CNRS, Illkirch-Graffentaden, France*

Benoît Leblanc • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada*

Steeve Leclerc • *Oncology and Molecular Endocrinology Research Center, Laval University, Québec City, QC, Canada*

Maryse Lessard • *Oncology and Molecular Endocrinology Research Center, Laval University, Québec City, QC, Canada*

Konstantinos Lymperopoulos • *Department of Physics and IRC in Bionanotechnology, Clarendon Laboratory, University of Oxford, Oxford, UK*

Yuri L. Lyubchenko • *Department of Pharmaceutical Sciences, College of Pharmacy, University of Nebraska Medical Center, Omaha, NE, USA*

Iain W. Manfield • *Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK*

Rainer Meffert • *Institut für Biochemie, Johannes Gutenberg-Universität, Mainz, Germany*

Vladimir Mekler • *Howard Hughes Medical Institute, Waksman Institute, Piscataway, NJ, USA*

Jean-François Millau • *Département de Pédiatrie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada*

Tom Moss • *Centre de Recherche en Cancérologie de l'Université Laval, Québec City, QC, Canada*

Nikolai Naryshkin • *Howard Hughes Medical Institute, Waksman Institute, Piscataway, NJ, USA*

Peter E. Nielsen • *Department of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark*

Athanasios G. Papavassiliou • *Department of Biological Chemistry, Medical School, University of Athens, Athens, Greece*

Isabelle Paradis • *Département de Pédiatrie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada*

Marvin R. Paule • *Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO, USA*

J. Carlos Penedo • *School of Physics and Astronomy, University of St. Andrews, St. Andrews, UK*

Jim Persinger • *Department of Biochemistry and Molecular Biology, Southern Illinois University School of Medicine, Carbondale, IL, USA*

Björn Persson • *GE Healthcare Biosciences AB, Uppsala, Sweden*

Yury S. Polikanov • *Department of Pharmacology, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ, USA*

Gabriele Rathgeber • *Institut für Biochemie, Johannes Gutenberg-Universität, Mainz, Germany*

Christopher M. Read • *Biophysics Laboratories, School of Biological Sciences, University of Portsmouth, Portsmouth, UK*

Félix Recillas-Targa • *Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, México City, México*

Andrei Revyakin • *Howard Hughes Medical Institute, Waksman Institute, Piscataway, NJ, USA*

Héctor Rincón-Arano • *Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, México City, México*

François Robert • *Institut de Recherches Cliniques de Montréal, Montréal, QC, Canada*

Nigel Savery • *School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK*

Hans-Jochen Schäfer • *Institut für Biochemie, Johannes Gutenberg-Universität, Mainz, Germany*

Deborah B. Schulman • *Department of Biology, Lake Erie College, Painesville, OH, USA*

Patrick Schultz • *Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch-Graffentaden, France*

Carl Séguin • *Département d'Anatomie et Physiologie, Faculté de Médecine, Université Laval, Québec City, QC, Canada*

David R. Setzer • *Division of Biological Sciences, University of Missouri, Columbia, MO, USA*

Peter E. Shaw • *School of Biomedical Sciences, University of Nottingham, Nottingham, UK*

Luda S. Shlyakhtenko • *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, NE, USA*

Patrizia Spitalny • *Department of Microbiology, University of Regensburg, Regensburg, Germany*

Victor Y. Stefanovsky • *Cancer Research Centre and Department of Medical Biology, Laval University, Québec City, QC, Canada*

A. Francis Stewart • *BioInnovations Centre, Dresden University of Technology, Dresden, Germany*

Peter G. Stockley • *Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK*

Vasily M. Studitsky • *Department of Pharmacology, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ, USA*

Amy Svotelis • *Département de Biologie, Faculté des Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada*

Ian A. Taylor • *National Institute for Medical Research, The Ridgeway, Mill Hill, London, UK*

Michael Thomm • *Department of Microbiology, University of Regensburg, Regensburg, Germany*

François Vignault • *Département de Pédiatrie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, QC, Canada*

Christian Zwieb • *Department of Molecular Biology, The University of Texas Health Science Center at Tyler, Tyler, TX, USA*

Chapter 1

Filter-Binding Assays

Peter G. Stockley

Summary

Structural studies of DNA–protein complexes reveal networks of contacts between proteins and the phosphates, sugars and bases of DNA. Our understanding of these structures, especially at the usual level of resolution for complexes between proteins and DNA (1.5–3.0 Å), is not sufficient to be able to infer directly the relative contributions of each contact to the overall binding energy. A range of biochemical methods have therefore been developed in order to probe the relative affinities of proteins for particular DNA target sites in vitro. Of these, one of the easiest and most widely used is nitrocellulose filter-binding. By exploiting the differential adsorption to nitrocellulose of proteins and peptides compared to nucleic acids, it is possible to prepare equilibrium mixtures of the proteins and nucleic acids of choice and estimate the amount of complex formation by rapid filtration. The concentration dependence of binding yields estimates of the equilibrium dissociation constant and trivial variations allow access to kinetic and thermodynamic data. The use of this technique is illustrated here using results from our experiments with the *Escherichia coli* methionine repressor, MetJ.

Key words: Filter-binding, DNA–protein interactions, Affinity and kinetics, MetJ, Methionine repressor.

1. Introduction

Adsorption of biological samples to nitrocellulose in one form or another has been known for a very long time *(1)* but it was not until the 1960s that this property was exploited for the study of protein–nucleic acid complex formation, firstly for RNA–protein complexes *(2)*, and only then for DNA–protein interaction studies *(3)*. The principle of the technique is straightforward. Under a wide range of buffer conditions, nucleic acids pass freely

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_1

through nitrocellulose membrane filters, whereas proteins and their bound ligands are retained. Thus, if a particular protein binds to a specific DNA sequence, passage through the filter will result in retention of a fraction of the protein–DNA complex by virtue of the protein component of that complex. The amount of DNA retained can then be determined allowing a binding curve to be constructed. Although filtration results in a separation of an equilibrium mixture, it is relatively rapid and can be used to analyze both binding equilibria and kinetic behavior *(4–6)*, and, if the DNA samples retained on the filter and in the filtrate are recovered for further processing, the details of the specific binding site can be probed by interference techniques *(7, 8)*. The assay has also been used as the basis for fractionating the bound and unbound components in in vitro selection (SELEX) experiments *(9–11)*. The technique has a number of advantages over footprinting and gel retardation assays, although there are also some relative disadvantages, especially where multiple proteins are binding to the same DNA molecule. We have used the technique to examine the interaction between the *E. coli* methionine repressor, MetJ, and its various operator sites cloned into restriction fragments *[(12–14)*, *see* **Figs.1** and **2**, and also the Chapter entitled "Ethylation Interference Footprinting of DNA–Protein Complexes"]. Results from these studies will be used to illustrate

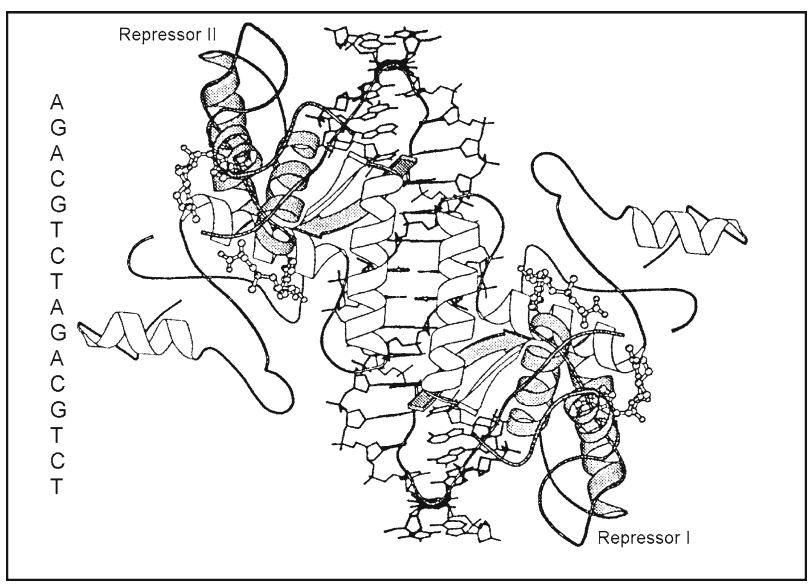


Fig. 1. The structure of the minimal MetJ repressor–operator complex (pdb 1cma). The sequence of the consensus two met box operator (*top strand*) is shown alongside from 5′ (*top*) to 3′ (*bottom*). The repressor is shown as a *ribbon* representation with one subunit *shaded*. The corepressor molecules (AdoMet) are shown as *ball* and *stick* representations and the DNA operator is a framework model. (I am grateful to Prof. Simon Phillips and his colleagues for production of this figure.).

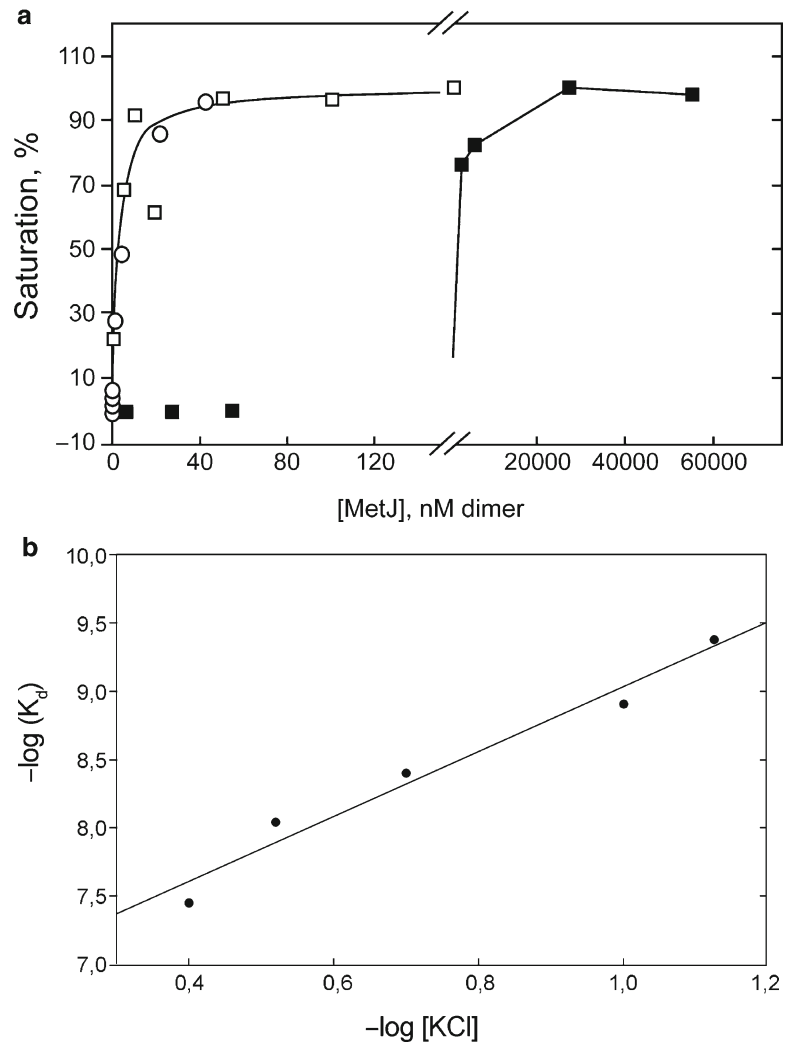


Fig. 2. Filter-binding assay of MetJ:operator interaction. (**a**) A plot of the fractional saturation with respect to repressor concentration of radiolabeled DNA fragments encompassing an operator consisting of two adjacent consensus met boxes cloned into a polylinker (*open square*); the same site cloned with anti-met boxes (sequences least likely to allow nonspecific binding by MetJ *(12)*; *open circle*) and the same polylinker lacking the insert, i.e., a nonspecific control (*filled square*). The *curve* drawn through the met box data is for a Hill plot model with weakly interacting sites. (**b**) A plot of log K_d vs. –log [KCl] determined by filter-binding assays of the operator containing fragment from (**a**), demonstrating the strong ionic strength dependence of the interaction. Figure redrawn from *(14)*.

the basic technique. The amount of DNA retained on the filter is most often determined using radioactively labeled samples. There are now also a number of nonradioactive alternatives that avoid the use of radioisotopes *(15, 16)*. In addition, an innovation to

allow simultaneous determination of the amount of uncomplexed DNA, by passing the filtrate from the nitrocellulose sheet directly through an adjacent DEAE membrane that retains all unbound DNA, allows improved quantitation of interactions in a higher throughput format *(17)*. Readers interested in these variations of the basic protocol described here should consult the primary references.

Before discussing the experimental protocols involved with such assays it is important to understand some fundamental properties of the filters being used. Nitrocellulose is a nitrated derivative of cellulose in which all the hydroxyl groups have been substituted. It therefore presents hydrophobic surfaces to protein molecules. The filters contain long pores through which the molecules must pass but these are much larger than the biological samples that get retained, so the physical basis of the technique is adsorption to the hydrophobic surface. In the past it has been assumed that this adsorption property is universal and that all proteins, even short peptides with molecular weights >2 kDa *(18)* will stick, although it was known that they do so with different efficiencies. More recent experiments with the *lac* repressor system, which was used to develop the assays initially *(4–6)*, have shown that complexes containing DNA loops with protein bound simultaneously to two targets sites in the same fragment have dramatically lower retention efficiencies than that of singly bound species *(19)*.

The cause of this change in relative retention of the same protein has been investigated in detail and provides an important caveat for all assays of protein–nucleic acid interactions using nitrocellulose filters *(20)*. It seems that proteins are "sticky" only if they have hydrophobic, and somewhat unstructured, regions on their surfaces. Lac repressor, in common with many other DNA-binding proteins, is relatively dynamic in the DNA-binding motif and careful experiments have established that it is the DNA-binding motif that allows the protein to stick to filters. Thus DNA-binding and filter-binding are competing for the same sites on the protein and when the DNA-binding sites in the repressor, which is normally a tetramer of 38 kDa subunits, are saturated the complex will not be retained by the filter and a naïve investigator would draw the wrong conclusion about the result. It is therefore important to validate these assays before spending a great deal of time on them. This is best done by showing that proteins of interest still bind to the filters even when saturated with excess target DNA. This allows the retention efficiency to be estimated. Tractable experimental values range from 30 to >95%. An example of the sort of results obtained with the MetJ repressor is shown in **Fig. 2a**.

2. Materials

2.1. Preparation of Radioactively End-Labeled DNA

1. Plasmid DNA carrying the binding site for a DNA-binding protein on a convenient restriction fragment (usually <200 bp).
2. Restriction enzymes and the appropriate buffers as recommended by the suppliers.
3. Phenol: Redistilled phenol equilibrated with 100 mM Tris–HCl, pH 8.0.
4. Chloroform.
5. Solutions for ethanol precipitation of DNA: 4 M NaCl and ethanol (absolute and 70%, v/v).
6. Alkaline phosphatase (AP), from calf intestine or shrimp.
7. AP reaction buffer, 10×: 0.5 M Tris–HCl, pH 9.0, 0.1 M $MgCl_2$, 0.001 M $ZnCl_2$.
8. TE buffer: 10 mM Tris–HCl, pH 8.0, 1 mM ethylenediaminetetra-acetic acid (EDTA).
9. Sodium dodecyl sulfate (SDS) 20% (w/v).
10. Ethylene diaminetetra-acetic acid (EDTA), 0.25 M, pH 8.0.
11. T_4 polynucleotide kinase (T_4 PNK).
12. T_4 PNK reaction buffer, 10×: 0.5 M Tris–HCl, pH 7.6, 0.1 M $MgCl_2$, 0.05 M dithiothreitol.
13. Radioisotope: [γ-^{32}P] ATP, typical specific activity >6,000 Ci/mmol. Use appropriate measures for protection against this radiation hazard.
14. 30% (w/v) Acrylamide stock (29:1 acrylamide: *N*,*N*′-methylene-bis-acrylamide).
15. TBE buffer, 5×: 450 mM Tris–borate, 5 mM EDTA, pH 8.3.
16. Polyacrylamide gel elution buffer: 0.3 M sodium acetate, 0.2% (w/v) SDS, 2 mM EDTA.
17. Polymerization catalysts: ammonium persulfate (10%, w/v) and *N*,*N*,*N*′,*N*′-tetramethylethylene diamine (TEMED).

2.2. Filter-Binding Assays

1. Nitrocellulose filters: We use HAWP (00024) filters from Millipore (Bedford, MA) but suitable filters are available from a number of other manufacturers, such as Schleicher and Schuell (Dassel, Germany). Filters tend to be relatively expensive. Some manufacturers produce sheets of membrane that can be cut to size and are thus less expensive.
2. Filter-binding buffer (FB): 100 mM KCl, 0.2 mM EDTA, 10 mM Tris–HCl, pH 7.6.

3. Binding buffer (BB): This is FB containing 50 μg/mL bovine serum albumin (BSA, protease and nuclease free; *see* **Note 1**).
4. Filtration manifold and vacuum pump: We use a Millipore 1225 Sampling Manifold (cat. no. XX27 025 50), which has 12 sample ports.
5. Liquid scintillation counter, vials, and scintillation fluid.
6. Siliconized glass test tubes.
7. TBE buffer: 89 mM Tris, 89 mM boric acid, 10 mM EDTA, pH 8.3.
8. Formamide/dyes loading buffer: 80% (v/v) formamide, 0.5× TBE, 0.1% (w/v) xylene cyanol, 0.1% (w/v) bromophenol blue.
9. Sequencing gel electrophoresis solutions and materials: 19% (w/v) acrylamide, 1% (w/v) bis-acrylamide, 50% (w/v) urea in TBE.
10. Acetic acid (10%, v/v).

3. Methods

3.1. Preparation of End-Labeled DNA

1. Digest the plasmid (20 μg in 200 μL) with the restriction enzymes used to release a suitably sized DNA fragment (usually <200 bp). Extract the digest with an equal volume of buffered phenol and add 2.5 vol of ethanol to the aqueous layer in order to precipitate the digested DNA. If preparing samples for interference assays, only one restriction digest should be carried out at this stage.
2. Add 50 μL 1× calf intestinal phosphatase (CIP) reaction buffer to the ethanol-precipitated DNA pellet (<50 μg). Add 1 U CIP and incubate at 37°C for 30 min followed by addition of a further aliquot of enzyme and incubate for a further 30 min. Terminate the reaction by adding SDS and EDTA to 0.1% (w/v) and 20 mM, respectively, in a final volume of 200 μL and incubate at 65°C for 15 min. Extract the digest with buffered phenol, then with 1:1 phenol:chloroform, and finally ethanol precipitate the DNA from the aqueous phase as above.
3. Redissolve the DNA pellet in 18 μL 1× T_4 PNK buffer. Add 20 μCi [γ-^{32}P] ATP and 10 U T_4 PNK and incubate at 37°C for 30 min. Terminate the reaction by phenol extraction and ethanol precipitation (samples for interference assays should be digested with the second restriction enzyme at this point).

Redissolve the pellet in nondenaturing gel loading buffer and electrophorese on a nondenaturing polyacrylamide gel.

4. After electrophoresis, separate the gel plates, taking care to keep the gel on the larger plate. Cover the gel with plastic wrap and, in the darkroom under the safe-light, tape a piece of X-ray film to the gel covering the sample lanes. With a syringe needle, puncture both the film and the gel with a series of registration holes. Locate the required DNA fragments by autoradiography of the wet gel at room temperature for several minutes (≈10 min). Excise slices of the gel containing the bands of interest using the autoradiograph as a guide. Elute the DNA into elution buffer overnight (at least) at 37°C. Ethanol precipitate the eluted DNA by adding 2.5 vol of ethanol, wash the pellet thoroughly with 70% (v/v) ethanol, dry briefly under vacuum, and rehydrate in a small volume (~50 μL) of TE. Determine the radioactivity of the sample by liquid scintillation counting of a 1-μL aliquot.

3.2. Filter-Binding Assays

3.2.1. Determination of the Apparent Equilibrium Constant

1. Presoak the filters in FB at 4°C for several hours before use. Care must be taken to ensure that the filters are completely "wetted." This is best observed by laying the filters carefully onto the surface of the FB using blunt-ended tweezers.
2. Prepare a stock solution of radioactively labeled DNA fragment in an appropriate buffer, such as FB. We adjust conditions so that each sample to be filtered contains roughly 20,000 cpm. Under these conditions the DNA concentration is <1 pM. Aliquot the stock DNA solution into plastic Eppendorf tubes. It is best at this stage if relatively large volumes are transferred in order to minimize errors caused by pipeting. We use 180 μL per sample. If the DNA-binding protein being studied requires a cofactor, it is best to add it to the stock solution at saturating levels so that its concentration is identical for every sample.
3. Prepare a serially diluted range of protein concentrations diluting into binding buffer (BB). A convenient range of concentrations for the initial assay is between 10^{-11} and 10^{-5} M protein.
4. Immediately add 20 μL of each protein concentration carefully to the sides of the appropriately labeled tubes of stock DNA solution. When the additions are complete centrifuge briefly (5 s) to mix the samples and then incubate at a temperature at which complex formation can be observed (37°C for MetJ). For each binding curve it is important to prepare two control samples. The first contains no protein in the 20 μL of BB and is filtered to determine the level of background retention. The second is identical to the first but is added to a presoaked filter in a scintillation vial (*see* **step 6**) and is

dried directly without filtering. This gives a value for 100% input DNA.

5. After an appropriate time interval to allow equilibrium to be established, re-centrifuge the tubes to return the liquid to the bottom of the tube and begin filtering.
6. The presoaked filters are placed carefully on the filtration manifold ensuring that excess FB is removed and that the filter is not damaged. Cracks and holes are easily produced by rough handling. The sample aliquot (200 μL) is then immediately applied to the filter, where it should be held stably by surface tension. Apply the vacuum. If further washes are used they should be applied as soon as the sample volume has passed through the filter. Remove the filter to a scintillation vial and continue until all the samples have been filtered.
7. The scintillation vials should be transferred to an oven at 60°C to dry the filters thoroughly (≈20 min) before being allowed to cool to room temperature and 3–5 mL of scintillation fluid added. The radioactivity associated with each filter can now be determined by counting on an open channel (*see* **Note 2**).
8. Correct the value for each sample by subtracting the counts in the background sample (no protein). Calculate the percentage of input DNA retained at each protein concentration using the value for 100% input from the unfiltered sample. Plot a graph of percentage retained vs. the logarithm of the protein concentration (e.g., **Fig. 2a**). The binding curve should increase from left to right until a plateau is reached. This is rarely at 100% of input DNA. The plateau value can be assumed to represent the retention efficiency, and for quantitative measurements the data points can be adjusted accordingly. There is not enough space here to describe in detail the form of the binding curve or how best to interpret the data. (For an authoritative yet accessible account *see* **ref. 21.**) For our purposes the protein concentration at 50% saturation can be thought of as the equilibrium dissociation constant (K_d) under these conditions.
9. Once an initial binding curve has been obtained, the experiment should be repeated with sample points concentrated in the appropriate region, i.e., the region where the percentage retained is changing most rapidly.

Control experiments with DNAs that do not contain specific binding sites should also be carried out to prove that binding is sequence specific (**Fig. 2a**). Highly diluted protein solutions appear to lose activity in our hands possibly because of nonspecific absorption to the sides of tubes, among other things. We therefore produce freshly diluted samples daily. BB can be stored at

4°C for several days without deleterious effect. Ideally binding curves should be reproducible. However, there is some variability between batches of filters and we therefore recommend filtering triplicate samples at every protein concentration and not switching lot numbers during the course of one set of experiments.

3.2.2. Kinetic Measurements

Kinetic analysis of the binding reaction depends on prior determination of the equilibrium binding curve, especially the concentration of DNA-binding protein required to saturate the input DNA. This information allows a reaction mixture to be set up containing a limiting amount of protein, for example, at a protein concentration that produces 75% retention. Both association and dissociation kinetics can be studied. The major technical problem arises because of the relatively rapid sampling rates that are required. However, it is almost always possible to adjust solution conditions such that sampling at 10 s intervals is all that is needed. Dissociation measurements often need to be made over periods of up to 1 h, whereas association reactions are usually complete within several minutes.

Dissociation

Repeat **steps 1** and **2** of Subheading 3.2.1 but do not aliquot the stock DNA solution. Add to this sample the appropriate concentration (i.e., which produces ~75% retention) of stock protein and allow to equilibrate. Add a 20-fold excess of unlabeled DNA fragment containing the binding site and begin sampling (≈200-μL aliquots) by filtration. Plots of radioactivity retained vs. time can then be analysed to derive kinetic constants. In the simplest case of a bimolecular reaction, a plot of the natural logarithm of the radioactivity retained at time t divided by the initial radioactivity vs. time yields the first-order dissociation constant from the slope. An important control experiment is to repeat the experiment with competitor DNA that does not contain a specific binding site to show that dissociation is sequence specific.

A variation of this experiment can be used in which the concentration of protein in the reaction mix is diluted across the range where most complex formation occurs such that the final protein concentration is well below the K_d. In this case it is necessary to prepare the initial complex in a small volume (≈50 μL) and then dilute 100× with BB, followed by filtering 500-μL aliquots.

Association

Set up a stock DNA concentration in a single test tube as above (Subheading 3.2.1, **steps 1** and **2**). Incubate both this DNA and the appropriate solution of protein at the temperature at which complexes form. Add the appropriate volume of protein (e.g., 200 μL) to the DNA stock solution (1,800 μL) and immediately begin sampling (10 × 200 μL aliquots).

3.2.3. Interference Measurements

Experiments of this type can be used to gain information about the site on the DNA fragment being recognized by the protein.

The principle is identical to that used in gel retardation interference assays but has the advantage that the DNA does not have to be eluted from gels after fractionation.

1. Modify the purified DNA fragment radiolabeled (~100,000 cpm) at a single site with the desired reagent; for example, hydroxyl radicals, which result in the elimination of individual nucleotide groups *(12)*, dimethyl sulfate, DMS *(11)*, which modifies principally guanines, or ethyl nitrosourea, ENU (*see* Chapter "Ethylation Interference Footprinting of DNA–Protein Complexes"), which ethylates the nonesterified phosphate oxygens. These techniques are covered in detail in Chapters "Hydroxyl Radical Footprinting of Protein–DNA Complexes," "Identification of Protein/DNA Contacts with Dimethyl Sulfate: Methylation Protection and Methylation Interference," and "Ethylation Interference Footprinting of DNA–Protein Complexes," respectively. The extent of modification should be adjusted so that any one fragment has no more than one such modification. This can be assessed separately in test reactions and monitored on DNA sequencing gels.
2. Ethanol precipitate the modified DNA, wash twice with 70% (v/v) ethanol and then dry briefly under vacuum. Resuspend in 200 µL FB. Remove 20 µL as a control sample. Add 20 µL of the appropriate protein concentration to form a complex and allow equilibrium to be reached. Filter as usual but with a siliconized glass test tube positioned to collect the filtrate. (The Millipore manifold has an insert for just this purpose.) Do not over dry the filter.
3. Place the filter in an Eppendorf tube containing 250 µL FB, 250 µL H_2O, 0.5% (w/v) SDS. Transfer the filtrate into a similar tube and then add SDS and H_2O to make the final volume and concentration the same as the filter-retained sample. Add an equal volume of buffer-saturated phenol to each tube, vortex, and centrifuge to separate the phases. Remove the aqueous top layers, re-extract with chloroform:phenol (1:1), and then ethanol precipitate. A Geiger counter can be used to monitor efficient elution of radioactivity from the filter, which can be re-extracted if necessary.
4. Recover all three DNA samples (control, filter-retained, and filtrate) after ethanol precipitation and, if necessary, process the modification to completion (e.g., piperidine for DMS modification, NaOH for ENU, and so on). Ethanol precipitate the DNA, dry briefly under vacuum, and then redissolve the pellets in 4 µL formamide/dyes denaturing loading buffer. At this stage it is often advisable to quantitate the radioactivity in each sample by liquid scintillation counting of 1-µL aliquots. Samples for sequencing gels should be adjusted to contain roughly equal numbers of counts in all three samples.

5. Heat the samples to 90°C for 2 min and load onto a 12% (w/v) polyacrylamide sequencing gel alongside Maxam–Gilbert sequencing reaction markers *(13)*. Electrophorese at a voltage that will warm the plates to around 50°C. After electrophoresis, fix gel in 1 L 10% (v/v) acetic acid for 15 min. Transfer the gel to 3MM paper and dry under vacuum at 80°C for 60 min. Autoradiograph the gel at –70°C with an intensifying screen or use a phosphorimaging screen.
6. Compare lanes corresponding to bound, free, and control DNAs for differences in intensity of bands at each position (*see* **Note 3**). A dark band in the "free fraction" (and a corresponding reduction in the intensity of the band in the "bound fraction") indicates a site where prior modification interferes with complex formation. This is interpreted as meaning that this residue is contacted by the protein or a portion of the protein comes close to the DNA at this point. (*See* Chapters "Identification of Protein/DNA Contacts with Dimethyl Sulfate: Methylation Protection and Methylation Interference" and "Ethylation Interference Footprinting of DNA–Protein Complexes" for more extensive discussions of interference experiments.)

3.3. Results and Discussion

Figure 1 shows a cartoon representation of the X-ray crystal structure of the minimal operator complex between wild-type methionine repressor, MetJ, and an oligonucleotide fragment encompassing two adjacent consensus met boxes (pdb 1cma). A typical filter-binding curve for MetJ binding to this operator site cloned into a plasmid polylinker and a fragment released and radiolabeled as described above (*see* **Subheading 3.2.1**), is shown in **Fig. 2a**, together with a similar curve for binding to the polylinker fragment lacking the specific operator site. Both curves show similar levels of saturation (~90%) but with dramatically different protein concentrations. The apparent K_d values determined from the point at which 50% input DNA was bound being $4.0(\pm 3.0) \times 10^{-9}$ M and 1.1×10^{-6} M repressor dimer for the specific and nonspecific complexes, respectively. The almost 300-fold difference in apparent affinity confirms the ability of the assay to discriminate between specific and nonspecific DNA-binding. Also shown in the figure is the binding curve for the same operator fragment flanked by specific DNA sequences designed to prevent additional repressor binding. This has the same affinity as the consensus operator in the polylinker, implying that all the interactions being assayed occur within this 16-bp operator site. Although the apparent equilibrium constant is derived from a technique that separates equilibrium mixtures, it is in very close agreement with a fluorescence anisotropy titration of a similar DNA fragment carrying a terminal fluorescein dye

(K_d 3.6(±0.7) × 10^{-9} M dimer; unpublished observation). Similar binding curves have been analyzed to produce Scatchard and Hill plots *(21)* in order to examine the co-operativity with respect to protein concentration *(14)*. However, such multiple binding events should also be studied by gel retardation assays which yield data about the individual complex species (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions").

The power of filter-binding assays is illustrated in **Fig. 2b** which shows a plot of the K_d values, determined at different salt concentrations, against the KCl concentration. Nitrocellulose filters are largely unaffected by changing buffer and salt concentrations, although the retention efficiency may vary slightly, and so it is facile to investigate the electrostatic contribution to complex formation. In this case, there is a steady decrease in operator affinity as the salt concentration is raised from 50 to 400 mM KCl. The plot of log K_d vs. –log [KCl] is roughly linear as expected and suggests that roughly two potassium ions are displaced from the DNA per repressor dimer during complex formation. Extrapolating these data suggests that the nonelectrostatic component of the observed binding free energy is ≈ –20 kcal/mol, i.e., equivalent to –10 kcal/mol per repressor dimer, or roughly 80% of the total for the formation of a dimer of repressor dimers along the DNA (–24.4 kcal/mol) *(14)*. The effects of altering the nature of the counterions or changing the temperature of complex formation are also easily investigated, the latter allowing thermodynamic analysis of the interaction. In this case, the data obtained with differing cations or at different ionic strengths suggest that the buffer conditions used to form the operator complex shown in **Fig. 1** do not alter the interaction of the protein and DNA, validating the interpretation of the structure seen in the crystal.

3.3.1. In Vitro Selection Experiments

In recent years in vitro selection experiments have been used to identify the range of preferred DNA target sequences by DNA-binding proteins (*see* Chapter "Identification of Nucleic Acid High Affinity Binding Sequences of Proteins by SELEX" and **refs. 9–11**). The technique depends on the separation of protein-bound DNA sequences from unbound, nonspecific, or low affinity sites. Filter-binding is an attractive option for this selection step because of the speed with which filtration and recovery of the bound fraction can be achieved. However, it is important to be aware that some minor DNA variants can be retained specifically by the filters, thus biasing the selected sequences. One way to avoid this and still retain the advantages of filter-binding is to alternate rounds of filter-binding with separation by gel retardation (*see* Chapter "Electrophoretic Mobility Shift Assays for

the Analysis of DNA–Protein Interactions"). A detailed discussion of the factors involved in such experiments is beyond the scope of this chapter and the reader is referred to more detailed descriptions (11).

4. Notes

There are a number of common problems in using this assay and here are answers to some of the most frequent:

1. None of the radioactivity is retained by the filter. This again can be caused by a variety of factors. Check that the preparation of DNA-binding protein is still functional (if other assays are available) or that the protein is still intact by SDS-PAGE. Check the activity/concentration of the cofactor if required. A common problem we have encountered arises because of the different grades of commercially available BSA. It is always advisable to use a preparation that explicitly claims to be nuclease and protease free.
2. All of the radioactivity is retained by the filter. This is a typical problem when first characterizing a system by filter-binding and can have many causes. Check that the filters being used "wet" completely in FB and do not dry significantly before filtration. Make sure that the DNA remains soluble in the buffer being used by simple centrifugation in a benchtop centrifuge. If the background remains high, add dimethyl sulfoxide to the filtering solutions. Classically 5% (v/v) is used but higher concentrations (~20%, v/v) have been reported with little, if any, effect on the binding reaction. We have experienced excessive retention when attempting to analyze the effects of divalent metal ions on complex formation, and in general it is best to avoid such buffer conditions.
3. Poor recoveries from the filter-retained samples in interference assays, or other problems in processing such samples further, can often be alleviated by addition of 20 μg of tRNA as a carrier during the SDS/phenol extraction step.

Acknowledgments

I am grateful to Dr. Iain Manfield for his helpful comments on the manuscript.

References

1. Elford, W.J. (1931). A new series of graded collodion membranes suitable for general bacteriological use, especially tint filterable virus studies. *J. Pathol. Bacteriol.* **34**, 505–521.
2. Nirenberg, M., and Leder, P. (1964). RNA codewords and protein synthesis: the effect of trinucleotides upon the binding of sRNA to ribosomes. *Science* **145**, 1399–1407.
3. Jones, O.W., and Berg, P. (1966). Studies on the binding of RNA polymerase to polynucleotides. *J. Mol. Biol.* **22**, 199–209.
4. Riggs, A.D, Bourgeois, S., Newby, R.F., and Cohn, M. (1968). DNA binding of the *lac* repressor. *J. Mol. Biol.* **34**, 365–368.
5. Riggs, A.D, Bourgeois, S., and Cohn, M. (1970). *lac* Repressor–operator interaction. III. Kinetic studies. *J. Mol. Biol.* **53**, 401–417.
6. Riggs, A.D., Suzuki, H., and Bourgeois, S. (1970). *lac* Repressor–operator interaction. I. Equilibrium studies. *J. Mol. Biol.* **48**, 67–83.
7. Siebenlist, U., and Gilbert, W. (1980). Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *Proc. Natl Acad. Sci. USA* 77, 122–126.
8. Hayes, J.J., and Tullius, T.D. (1989). The missing nucleoside experiment: a new technique to study recognition of DNA by protein. *Biochemistry* **28**, 9521–9527.
9. Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment – RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510.
10. Ellington, A.D., and Szostak, J.W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822.
11. Conrad, R.C., Giver, L., Tian, Y., and Ellington, A.D. (1996). *In vitro* selection of nucleic acid aptamers that bind proteins. *Methods Enzymol.* **267**, 336–367.
12. Phillips, S.E., Manfield, I., Parsons, I., Davidson, B.E., Rafferty, J.B., Somers, W.S., Margarita, D., Cohen, G.N., Saint-Girons, I., and Stockley, P.G. (1989). Cooperative tandem binding of Met repressor from *Escherichia coli*. *Nature* **341**, 711–715.
13. Old, I.G., Phillips, S.E.V., Stockley, P.G., and Saint-Girons, I. (1991). Regulation of methionine biosynthesis in the enterobacteriaceae. *Prog. Biophys. Mol. Biol.* **56**, 145–185.
14. He, Y.Y., Garvie, C.W., Elworthy, S., Manfield, I.W., McNally, T., Lawrenson, I.D., Phillips, S.E., and Stockley, P.G. (2002). Structural and functional studies of an intermediate on the pathway to operator binding by *Escherichia coli* MetJ. *J. Mol. Biol.* **320**, 39–53.
15. Christopoulos, T.K., Diamandis, E.P., and Wilson, G. (1991). Quantification of nucleic acids on nitrocellulose membranes with time-resolved fluorometry. *Nucleic Acids Res.* **19**, 6015–6019.
16. Czerwinski, J.D., Hovan, S.C., and Mascotti, D.P. (2005). Quantitative nonisotopic nitrocellulose filter binding assays: bacterial manganese superoxide dismutase–DNA interactions. *Anal. Biochem.* **336**, 300–304.
17. Wong, I., and Lohman, T.M. (1993). A double-filter method for nitrocellulose-filter binding: application to protein–nucleic acid interactions. *Proc. Natl Acad. Sci. USA* **90**, 5428–5432.
18. Ryan, P.C., Lu, M., and Draper, D.E. (1991). Recognition of the highly conserved GTPase center of 23S ribosomal RNA by ribosomal protein L11 and the antibiotic thiostrepton. *J. Mol. Biol.* **221**, 1257–1268.
19. Fickert, R., and Mullerhill, B. (1992). How *lac* repressor finds *lac* operator *in vitro*. *J. Mol. Biol.* **226**, 59–68.
20. Oehler, S., Alex, R., and Barker, A. (1999). Is nitrocellulose filter binding really a universal assay for protein–DNA interactions? *Anal. Biochem.* **268**, 330–336.
21. Wyman, J., and Gill, S.J. (1990). Binding and Linkage: Functional Chemistry of Biological Macromolecules. University Science Books, Mill Valley, CA, Chapter 2.

Chapter 2

Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions

Manon Gaudreault, Marie-Eve Gingras, Maryse Lessard, Steeve Leclerc, and Sylvain L. Guérin

Summary

Electromobility shift assay is a simple, efficient, and rapid method for the study of specific DNA–protein interactions. It relies on the reduction in the electrophoretic mobility conferred to a DNA fragment by an interacting protein. The technique is suitable to qualitative, quantitative, and kinetic analyses. It can also be used to analyze conformational changes.

Keywords: EMSA, Gel-shift, Mobility shift assay, Gel retardation.

1. Introduction

Several nuclear mechanisms involve specific DNA–protein interactions. To study such interactions, a simple, efficient, and widely used method, first described several years ago *(1, 2)*, is the electrophoretic mobility shift assay (EMSA, also known as gel mobility shift assay or gel retardation). Its ease of use, its versatility, and especially its high sensitivity (10^{-18} pmol of DNA *(2)*) make it a powerful method that has been successfully used in a variety of situations, particularly in gene regulation analyses, but also in studies of DNA replication, repair, and recombination. EMSA can be carried out in three simple steps (1) labeling of the DNA probe, (2) isolation of the labeled probe, and (3) preparation of the DNA–protein binding reaction and its analysis on a native polyacrylamide gel. Besides being useful for qualitative purposes, EMSA has

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_2

the added advantage of being suitable for quantitative and kinetic analyses *(3)*. Furthermore, because of its very high sensitivity, EMSA makes it possible to resolve complexes of different protein or DNA stoichiometry *(4)* or even to detect conformational changes.

1.1. Principle of the Method

EMSA is based on the simple rationale that molecules of differing size, molecular weight, and charge will have different electrophoretic mobilities in a nondenaturing gel matrix. In the case of a DNA–protein complex, the presence of a DNA-binding protein will cause the complex to migrate differently (usually more slowly) relative to the free DNA, and will thus cause a shift visible upon detection.

While the kinetic analysis of EMSA, which has recently been covered extensively elsewhere (*see* **ref. 5** and references therein), is not the focus of this practical work, it is useful to understand the basic theory underlying the method. A univalent protein, P, binding to a unique site on a DNA molecule, D, will yield a complex, PD, in equilibrium with the free components:

$$P + D \underset{k_d}{\overset{k_a}{\rightleftharpoons}} PD,$$

where k_a is the rate of association and k_d is the rate of dissociation. In the case of a strong interaction between protein and DNA, with $k_a > k_d$, two distinct bands are observed, corresponding to the complex PD and to the free DNA. However, because of the dissociation that inevitably occurs during electrophoresis, and because DNA released during electrophoresis can never catch up to the free DNA, a faint smear may be seen between the two major bands. In contrast, a weak DNA–protein interaction, with $k_a < k_d$, should produce a fainter band corresponding to the complex PD and a more intense smear. However, even weak DNA–protein complexes show distinct bands in EMSA, because they are stabilized in the gel matrix through cage effect *(6)* and/or molecular sequestration *(7)*. In both mechanisms, the dissociation of the complex is slower in the gel than in free solution, thereby contributing to complex stabilization. But whereas in the cage effect, the gel matrix prevents dissociated components P and D from diffusing out in the gel and thus favors re-formation of complex PD, molecular sequestration implies that the gel matrix isolates complex PD from competing molecules that would promote its dissociation.

As for a single DNA molecule bearing multiple binding sites for an individual protein, there will be as many complexes formed, as there are binding sites:

$$\begin{array}{ccc} \mathbf{2PD} + \mathbf{D} & \underset{k_{d1}}{\overset{k_a}{\rightleftharpoons}} & \mathbf{PD1} + \mathbf{P} \\ k_a \uparrow \downarrow k_{d2} & & k_a \uparrow \downarrow k_{d3} \\ \mathbf{PD2} + \mathbf{P} & \underset{k_{d4}}{\overset{k_a}{\rightleftharpoons}} & \mathbf{P2D} \end{array}$$

and this would result in three bands (D, P2D, and PD1-PD2, which migrate together).

Kinetics of more complicated situations, such as dimerizing protein complexes and multiple DNA–protein interactions, are too lengthy to go in details and are beyond the scope of this volume. We therefore refer the reader to some interesting and insightful papers *(4, 8)*, where these questions are addressed extensively.

1.2. Overview of the Procedure

Several components are required for EMSA and may influence the outcome of the procedure.

1.2.1. Nuclear Extracts

The choice of extract is governed by the objective of the study. Whole-cell or nuclear extracts are very useful in analyzing the regulatory elements of a DNA fragment such as a gene promoter. Partial proteins purification allows further characterization of a DNA–protein interaction and can be obtained by column chromatography on DNA–cellulose or heparin-Sepharose, or by SDS-polyacrylamide gel fractionation–renaturation (*(9)* and *see* **Note 1**). Purified or recombinant proteins give valuable information on protein interactions, competition, dimerization, or cooperativity. Whatever protein extract is used, the quality of its preparation is essential in EMSA (*see* **Notes 2** and **3**).

1.2.2. DNA Probe

Cloned DNA fragments of 50–400 base pairs (bp) in length or synthetic oligonucleotides of 20–70 bp both work very well in EMSA (*(10)* and *see* **Note 4**), and although double-stranded DNA is used more often, single-stranded DNA is also effective *(11)*. Whereas DNA fragments reveal more extensive regulatory sequences, oligonucleotides usually target fewer sites, thereby yielding more specific information; the two approaches can obviously complement one another. Detection of DNA–protein complexes usually involves labeling of the DNA (*see* **Note 5**), which is most commonly performed using a [^{32}P]-labeled deoxynucleotide. However, other, less hazardous methods are available (*see* **Note 5**), including labeling with ^{33}P *(12)*, digoxygenin *(9)*, or biotin *(13)*.

1.2.3. Gel Matrix, Temperature, and Voltage

Acrylamide (*see* **Note 6**) combines high resolving power with broad size-separation range and is the most widely used matrix, although use of less toxic, commercially available alternative matrices has been reported *(14–16)*. Because of its lower resolving power, agarose is sometimes used, alone or in combination with acrylamide, to study larger DNA fragments or multiprotein complexes *(17)*. Gel concentration is also important in EMSA (*see* **Note 7**); although lower concentrations will allow resolution of larger complexes, it may affect their stability (*see* **Fig. 1a** and **(7)**). The temperature at which the migration is carried out can also influences the migration of the DNA–protein complexes (**Fig. 1b** shows the influence of either room temperature (RT)

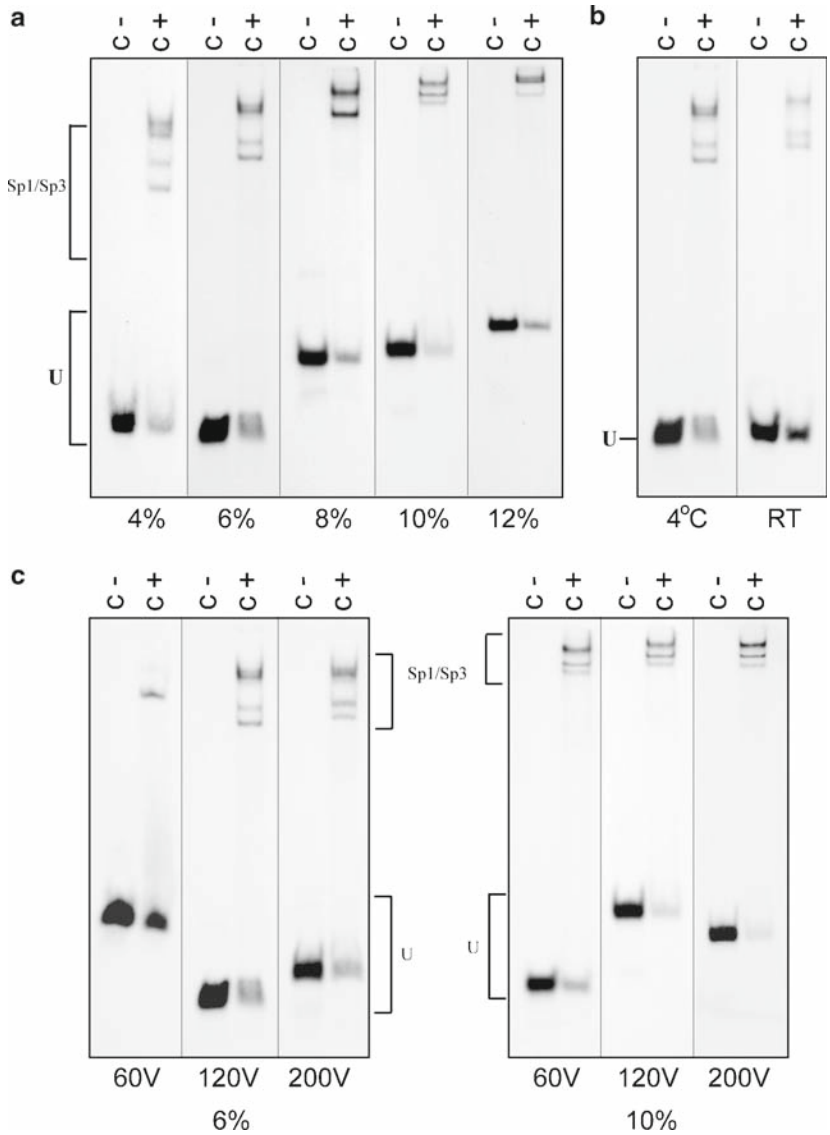


Fig. 1. Influence of the gel concentration, voltage, and running temperature on the formation of protein–DNA complexes in EMSA. A synthetic double-stranded oligonucleotide bearing the high-affinity binding site for the transcription factor Sp1 was 5′-end labeled and incubated with (C+) or without (C–) 5 µg crude nuclear extract from human epithelioid carcinoma HeLa (ATCC CCL 2) cells. The EMSAs were run at 4°C on native polyacrylamide gels whose concentration ranged from 4 to 12% (**a**), on a 6% gel runned at two different temperatures (4°C and room temperature (RT)) (**b**) or at three different voltages (60, 120, and 200 V) on either 6 and 10% polyacrylamide gels (**c**). The position corresponding to the Sp1 and Sp3 DNA–protein complexes (Sp1/Sp3) as well as that of the free probe (U) is indicated.

or 4°C on the formation of the Sp1 DNA–protein complex). However, among these many criteria, the voltage selected for the migration of the DNA–protein complexes into the polyacrylamide gel remains the most important. Indeed, reducing the voltage (which consequently increases the migration time) has

proved to substantially alter the stability of DNA–protein complexes in a gel-concentration-dependent manner (compare the results between the 6 and 10% gels on **Fig. 1c**; also *see* **Note 8**).

1.2.4. Buffer

Different low-ionic strength buffers can be used in EMSA (*(16)* and *see* **Note 9**), and they may include cofactors such as Mg^{2+} or cAMP, which may be necessary for some DNA–protein interactions *(17)*.

1.2.5. Nonspecific Competitors

To ensure specificity of the DNA–protein interaction, a variety of nonspecific competitors may be used. This is particularly important when using a crude protein extract, which contains nonspecific DNA-binding proteins. To avoid this problem, nonspecific DNA such as salmon sperm DNA, calf thymus DNA, or synthetic poly(dI:dC) is used (*(17)* and *see* **Notes 10** and **11**). Addition of nonionic or zwitterionic detergents *(18)* or nonspecific proteins (e.g., albumin *(19)*) may also increase specific DNA–protein interactions.

1.3. Alternative Procedures and Applications of the EMSA

Because EMSA often allows the detection of specific DNA-binding proteins in unpurified protein extracts *(20)*, the technique has been widely used to analyze crude cell or tissue extracts or partially purified extracts for the presence of protein factors implicated in transcription *(21–24)* and in DNA replication *(20, 25)*, recombination *(11)*, and repair *(26)*. The use of unlabeled competitor DNA fragments further aids in identification of DNA-binding proteins (*(10, 11, 20)* and **Fig. 2a**), and their purification can be monitored easily in EMSA (*(20, 23)* and **Fig. 3** as well as **Note 11**). Moreover, mutation or base deletion of residues on the labeled DNA probe is often an efficient approach to use when identifying the binding site of the protein of interest *(21, 23)*.

EMSA also yields invaluable data when purified or recombinant proteins are analyzed, since quantification and kinetic studies are rapidly achieved *(21, 25)*. Parameters of the DNA–protein interaction, such as association and dissociation constants and affinity, are accurately measured *(2, 3, 7, 21)*, and the effect of salt, divalent metals, protein concentration, and temperature of incubation on complex formation is directly observed (*(11, 28, 29)* and **Figs. 1** and **4**). EMSA has also greatly contributed to the elaboration of complex assembly models in transcription *(22, 31)*, DNA replication *(25, 32)*, and DNA repair *(26, 33)*.

Although EMSA is an informative and versatile method on its own, it becomes even more powerful when used in combination with other techniques. Methylation *(34)* and other forms of binding interference studies, where a partly methylated DNA probe is used, helps to define the exact position of the DNA binding

site of the protein *(21, 35)*. Immunological methods using specific antibodies, as in supershift experiments (*(23, 24)* and **Fig. 2b**), are also very helpful in identifying the protein component of the complexes. However, when analyzing large or multiprotein complexes, supershift may not be suitable, because supershifted complexes either are not distinct from the shifted ones or do not identify the different proteins involved. Therefore, other methods that resolve these problems have been developed, such as immunoblotting of EMSA gels *(36, 37)*, "Shift-Western blotting" *(38)*, and immunodepletion EMSA *(39)*. In addition, determination of the molecular weight of the DNA-binding protein(s) identified in EMSA can also be achieved by running a SDS-PAGE,

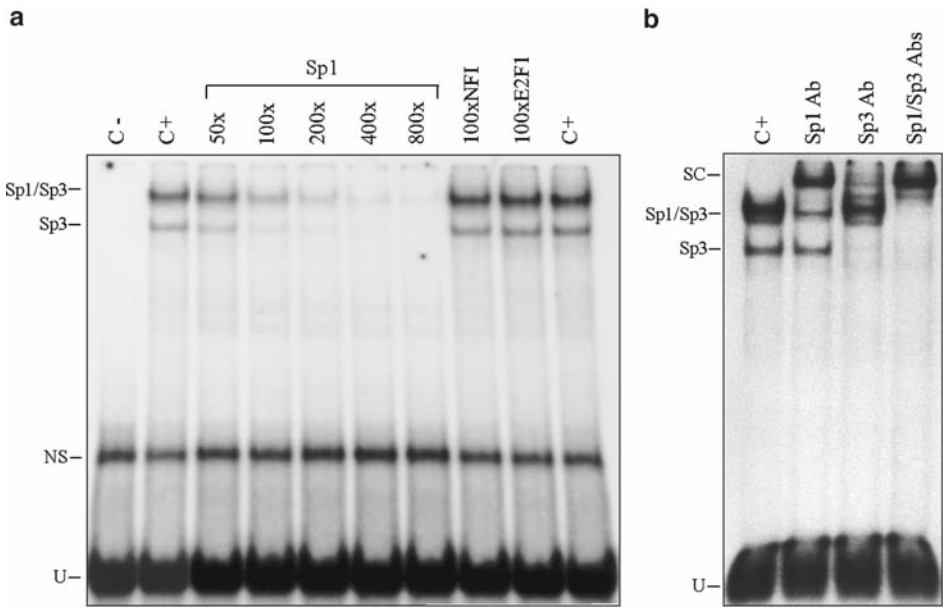


Fig. 2. (**a**) Competition in EMSA as a tool to evaluate the specific formation of DNA–protein complexes. A synthetic double-stranded oligonucleotide bearing the high-affinity binding site for the transcription factor Sp1 was 5′-end labeled and incubated with 3 µg of crude nuclear extract from primary cultured rabbit corneal epithelial cells (RCECs). Different concentrations (50-, 100-, 200-, 400- and 800-fold molar excesses) of an unlabeled double-stranded oligonucleotide bearing either the high affinity GC-rich binding site recognized by Sp1 or the target site for the unrelated transcription factors NFI and E2F1 (800-fold molar excess used) were added during the binding assays. Formation of DNA–protein complexes was evaluated on a native 8% polyacrylamide gel. Control *lanes* containing the labeled probe alone (C–) or incubated with proteins in the absence of any competitor (C+) have also been included. The position of the specifically retarded DNA–protein complexes (Sp1/Sp3 and Sp3) is indicated along with that of a nonspecific complex (NS). *U* unbound fraction of the labeled probe. (**b**) Supershift analyses in EMSA as a tool to unravel the identity of a DNA binding protein. The Sp1-labeled probe used in (**a**) was incubated with 3 µg of a crude nuclear extract prepared from primary cultured RCECs (**a**), either alone or in the presence of polyclonal antibodies directed against Sp1 and Sp3 (added individually or in combination). Formation of DNA/protein complexes was then monitored by EMSA as in (**a**). The position of the previously characterized Sp1 and Sp3 DNA/protein complexes is shown along with that of a supershifted complex (SC) resulting from the specific interaction of the antibodies with the Sp1 and Sp3–DNA complexes. The position of the remaining free probe (U) is indicated *(27)*.

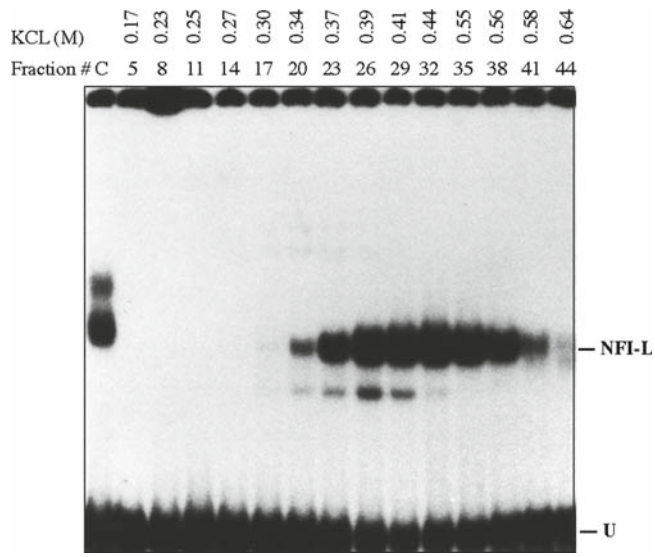


Fig. 3. Monitoring the enrichment of a nuclear protein by EMSA. Crude nuclear proteins (50 mg) of a rat liver extract were prepared and further purified on a heparin-Sepharose column. Nuclear proteins were eluted using a 0.1–1.0 M KCl gradient and fractions individually incubated with a 34-bp double-stranded synthetic oligonucleotide bearing the DNA sequence of the rat growth hormone promoter proximal silencer-1 element as the labeled probe. Both the concentration of KCl required to elute the proteins contained in each fraction, as well as the fraction number selected are indicated, along with the position of a major shifted DNA–protein complex corresponding to the rat liver form of the transcription factor NFI (termed NFI-L). *C* control *lane* in which the silencer-1 labeled probe was incubated with 5 μg crude nuclear proteins from rat liver; *U* unbound fraction of the labeled probe.

either immediately *(40)* or following ultraviolet cross-linking of the DNA to the protein(s) *(41)*. Another approach to detect unknown DNA-binding protein(s) consists to label the protein from any given DNA/protein complex with fluorescein following the EMSA. Briefly, the DNA-binding proteins complexed with labeled, double-stranded oligonucleotide probes are first resolved by EMSA. After locating their position by autoradiography, the shifted DNA-protein complexes are excised from the gel and incubated with 5-iodoacetamidofluorescein (5-IAF) for labeling of the protein component from the complexes. Proteins containing cysteine residues then become labeled with fluorescein through S-alkylation of the thiol residue of cysteine (Cys-SH) by 5-IA. The proteins are then separated by SDS-PAGE and both their number and molecular masses are analyzed by immunoblotting with fluorescein antibodies *(42)*.

With some modifications, the EMSA can also be used to estimate the binding capacity of a small ligand, such as drug candidates, and a receptor. A probe corresponding to the ligand modified with DNA is incubated with a sample containing the

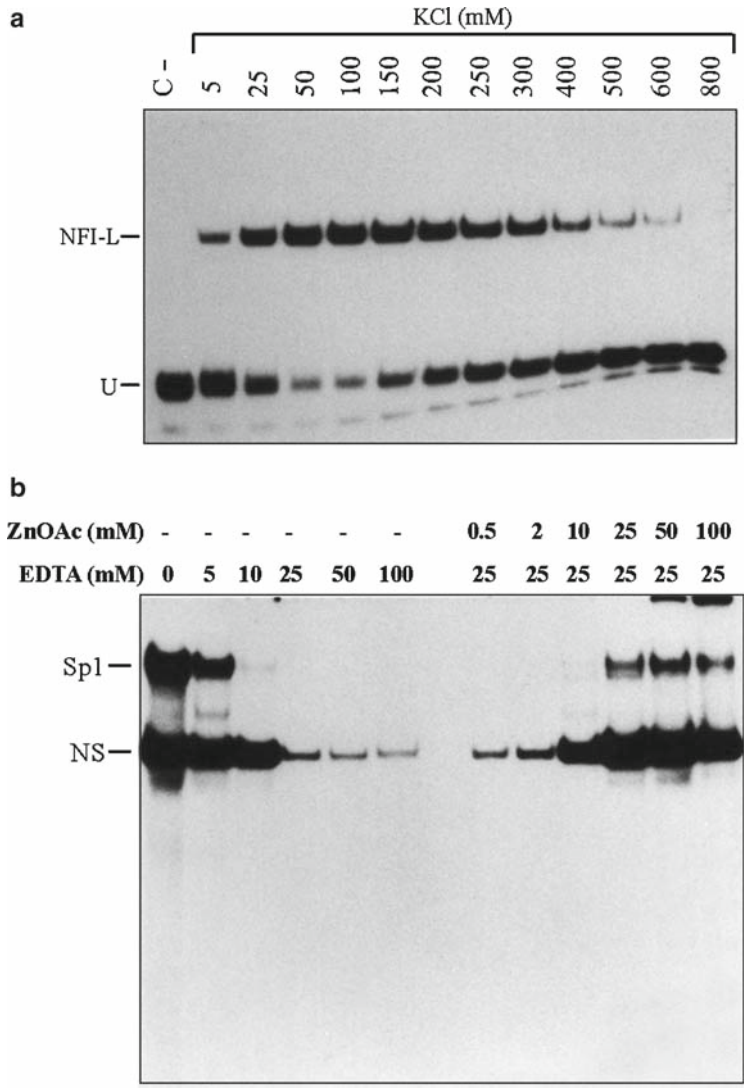


Fig. 4. (**a**) Salt-dependent formation of DNA–protein complexes in EMSA. A 5′-end labeled 35-bp synthetic double-stranded oligonucleotide bearing the NFI-L binding site identified in the 5′-flanking sequence of the human CRBP1 gene (and designated Fp5) was incubated in the presence of 1 µg of a heparin-Sepharose-enriched preparation of NFI-L and increasing concentrations of KCl (5–800 mM) using binding conditions similar to those described in this chapter. Formation of the Fp5/NFI-L DNA–protein complex was then resolved by electrophoresis on a 4% native polyacrylamide gel. Very little free probe (U) is observed in the presence of either 50 or 100 mM KCl, providing evidence that optimal binding of NFI-L to its target site in Fp5 is obtained at these salt concentrations (modified from *(28)*: reprinted with permission from Biotechniques, Copyright (2008)). (**b**) DNA binding properties of nuclear proteins revealed by EDTA chelation in EMSA. A double-stranded synthetic oligonucleotide bearing the sequence of the rat PARP US-1 binding site for the positive transcription factor Sp1 was 5′ end-labeled and incubated with 10 µg crude nuclear proteins from HeLa cells in the presence of increasing concentrations of EDTA (0–100 mM) under binding conditions identical to those described in this chapter. Formation of DNA–protein complexes was evaluated in EMSA on an 8% polyacrylamide gel. As little as 10 mM EDTA proved to be sufficient to chelate the zinc ions to totally prevent binding of Sp1 to the US-1 element. Similarly, reaction mixtures containing the US-1 labeled probe incubated with 10 µg nuclear proteins from HeLa cells in the presence of 25 mM final concentration of EDTA were supplemented with increasing concentrations (0.5–100 mM) of zinc acetate (ZnOAc) to evaluate the binding recovery for both Sp1 and the nonspecific DNA–protein complex (NS). A substantial proportion of the DNA binding capability of both the Sp1 and the NS proteins could be recovered upon further addition of 25 mM zinc acetate providing evidence that both factors likely interact with DNA through the use of a Zn-finger-containing DNA binding domain, a fact that was already known for Sp1 (modified from *(30)*: reprinted with permission from Eur. J. Biochem., Copyright (2008)).

receptor and separated by agarose gel electrophoresis. Interaction between the modified ligand and the receptor gives rise to a shifted complex. The concentration of the ligand in the sample could be determined by adding increasing concentrations of free ligand *(43)*. A variation of the traditional EMSA that can reveal the DNA binding capacity of any given protein consists in synthesizing and labeling this protein with ^{35}S and incubate it with unlabeled DNA. This method proved useful for studying a specific protein and for the detection of protein–DNA affinity constants when the protein concentration is unknown *(44, 45)*. The DNA–protein complex can also be separated by electrophoresis through uncoated capillaries or on poly(ethylene glycol)-modified glass microchannels instead of a polyacrylamide gel. Although less affordable, these methods, however, have the advantages of being computerized, very fast, and to require small amounts of samples *(46–48)*.

2. Materials

2.1. Probe Labeling

1. [γ-^{32}P] ATP. Caution: ^{32}P emits high-energy beta radiations. Refer to the rules established by your local control radioactivity agency for handling and proper disposal of radioactive materials and waste (*see* **Note 5**).
2. About 25–50 ng of DNA from a 30-bp double-stranded oligonucleotide. For a typical 70-bp probe derived from a subcloned promoter fragment, estimate the amount of the plasmid DNA that is required to end up with about 100–200 ng of the DNA fragment of interest following its isolation from the polyacrylamide gel (*see* **Note 4**).
3. Calf intestinal alkaline phosphatase (CIP) and 10× CIP reaction buffer (0.5 M Tris–HCl pH 9.0, 10 mM $MgCl_2$, 1 mM $ZnCl_2$, 10 mM spermidine).
4. T_4 polynucleotide kinase and 10× kinase buffer (0.5 M Tris–HCl pH 7.5, 0.1 M $MgCl_2$, 40 mM DTT, 1 mM spermidine, 1 mM EDTA).

2.2. Probe Isolation

1. Standard electrophoresis apparatus for agarose gel.
2. Stock solution of 10× TBE (0.89 M Tris, 0.89 M boric acid, 20 mM EDTA).
3. 1% (w/v) agarose diluted in 1× TBE supplemented with 1 µL of a 10 mg/mL solution of ethidium bromide. Caution: ethidium bromide is a suspected mutagenic agent (*see* **Note 12**).
4. Restriction enzyme(s) with corresponding buffer(s).
5. For DNA precipitation, a preparation of 1 mg/mL tRNA, a solution of 3 M NaOAc (pH 5.2), and a supply of dry ice.

6. Phenol/chloroform (Phenol saturated in 100 mM Tris pH 8.0).
7. Stock solution of 40% (w/v) acrylamide prepared in a ratio of 29:1 (w/w) acrylamide and *N*,*N*′-methylene bis-acrylamide. After complete dissolution of the components, the solution must be filtered on a Whatman paper and can be stored at room temperature. Caution: acrylamide is a potent neurotoxic agent (*see* **Note 6**).
8. Dialysis tubing (molecular weight cut off of 3,500 and flat width of 18 mm).
9. Saran Wrap®.
10. Autoradiography cassettes and films (Kodak XOmat AR).

2.3. Electrophoretic Mobility Shift Assay

1. Standard vertical electrophoresis apparatus for polyacrylamide gels (*see* **Note 13**).
2. A stock solution of 40% (w/v) acrylamide prepared in a 39:1 (w/w) ratio of acrylamide and *N*,*N*′-methylene bis-acrylamide. Caution: acrylamide is a potent neurotoxic agent (*see* **Note 6**).
3. A stock solution of 5× Tris–glycine (250 mM Tris, 12.5 mM EDTA, 2 M glycine) (*see* **Note 9**).
4. Extract (crude or enriched) containing cell or tissue nuclear proteins (*see* **Notes 2** and **3**).
5. A stock solution of 2× binding buffer (20 mM Hepes pH 7.9, 20% glycerol, 0.2 mM EDTA, 1 mM tetrasodium pyrophosphate (*see* **Note 3**) and 0.5 mM PMSF).
6. A stock solution of 6× loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, and 40% sucrose).
7. Whatman chromatographic paper (3MM) and Saran Wrap®.
8. Standard gel dryer.
9. Autoradiography cassettes and films (Kodak XOmat AR).

3. Methods

3.1. Probe Labeling

3.1.1. Labeling DNA Fragments Derived from a Subcloned Sequence

1. Select restriction enzymes that produce the shortest DNA fragment containing the sequence of interest. One of these restriction enzymes should produce a protruding 5′-end or blunt end to support labeling with T_4 polynucleotide kinase (*see* **Note 14**). Following the manufacturer's optimal enzymatic conditions, prepare a digestion mix with one of the restriction enzymes in 50 µL to linearize the vector. The initial amount of DNA should be calculated to end up

with at least 100–200 ng of DNA after double restriction enzyme digestion and further isolation of the DNA fragment from the polyacrylamide gel.

2. Before proceeding with dephosphorylation, make sure that digestion is complete by loading a sample (50–100 ng) on a 1% (w/v) agarose minigel. Once complete digestion of the plasmid DNA has been verified, add directly to the digestion reaction mix, 1 unit of CIP and 10 μL of 10× CIP buffer and fill to 100 μL with H_2O. Incubate at 37°C for 90 min.
3. To totally eliminate and inactivate CIP, transfer the reaction mix at 70°C for 10 min and perform both a phenol/chloroform and a chloroform extraction. Precipitate DNA with 1/10 volume of NaOAc 3 M pH 5.2 and 2 volumes of cold 90% ethanol. Allow DNA to precipitate on dry ice for 30 min then centrifuge for 15 min.
4. Resuspend DNA in 33 μL of H_2O and add 5 μL of 10× kinase buffer, 2 μL of T_4 polynucleotide kinase, and 100 μCi of [γ-^{32}P] ATP. Mix and incubate at 37°C for 2 h.
5. Following the labeling procedure, precipitate DNA and resuspend in 30 μL of H_2O. Keep a 2-μL sample and digest the remainder with the second restriction enzyme, following manufacturer's conditions.

3.1.2. Labeling Double-Stranded Synthetic Oligonucleotides

1. Anneal equal amounts of the complementary strands, heat the resulting DNA mix at 5°C over the specific melting temperature (T_M) of the oligo for 5 min and let cool at room temperature (RT). When DNA reaches RT, store at 4°C for a few hours prior to use.
2. Use 25–50 ng of the double-stranded oligonucleotide preparation and perform DNA labeling with T_4 polynucleotide kinase as described in **step 4** but using 30 μCi of [γ-^{32}P] ATP (*see* **Subheading 3.1.1**).

3.2. Probe Isolation

3.2.1. For a Typical 70-bp Probe Derived from a Subcloned Promoter Fragment

1. Rigorously clean and dry the polyacrylamide gel apparatus and its accessories prior to use. Gel plates should be cleaned using any good quality commercial soap and then rinsed with 95% ethanol. One plate can be treated with a coat of Sigmacote (chlorinated organopolysiloxane in heptane) to facilitate gel removal from the plates after running.
2. Prepare a 6% polyacrylamide gel *(49)* as follows; mix 2.5 mL of 10× TBE, 3.75 mL of 40% acrylamide (29:1) stock solution and H_2O to obtain 25 mL. Add 180 μL of 10% ammonium persulfate and 30 μL of TEMED. Carefully stir and pour the acrylamide solution between the plates. Place a comb and allow the gel to set for 30 min then mount the gel in the electrophoresis tank and fill the chamber with 1× TBE.

3. To the double digested DNA, add 10 μL of 6× loading buffer and load into two separate wells. For the 2-mL control sample from the single digestion, add 2-μL of loading buffer and load in a free well. Migration should be stopped when bromophenol blue, which is used as a migration marker, reaches two-third of the gel length.
4. Under radioactive protection, carefully disassemble the apparatus. Discard the running buffer as radioactive waste. Remove one plate and leave the gel on the remaining plate. Cover the gel with Saran Wrap and, in a dark room, place a film over it. It is very important to mark the exact position of the gel on the film as a reference. This can be achieved by drawing symbols on white lab tapes with [^{32}P]-labeled black ink. Once the labeled black ink has dried, stick the tapes (at least two of them) at different positions on the Saran Wrap-covered gel. Expose the film for 3 min and proceed with development.
5. If the digestion step with restriction enzymes is correct, two labeled bands resulting from the double digestion should appear on the autoradiogram (provided that each of the restriction enzymes selected initially cut the promoter-bearing recombinant plasmid only once). Using a razor blade, cut and remove from the film the lower band corresponding to the appropriate selected probe. Using the reference marks, place the film over the gel (still covered with Saran Wrap®). Use the hole on the film corresponding to the exact position of the probe and cut the probe-containing gel fragment using a scalpel blade.
6. Place the acrylamide fragment in a dialysis tubing closed at one end and add 1 mL of 1× TBE. Remove any remaining air bubble, close the other end and place the dialysis tubing in a standard electrophoresis tank for agarose gel filled with 1× TBE. Run at 100 V for 15 min.
7. Through the action of electrophoretic migration, the labeled probe will pass from the acrylamide fragment to the TBE solution contained in the dialysis tubing. DNA will concentrate on a tiny line along the dialysis tubing (on the cathode side). It must be removed by gently rubbing the tubing with a solid object. Other procedures may be selected for extracting the labeled probe from the polyacrylamide gel *(50)*. Using a Pasteur pipette, transfer the labeled probe-containing TBE from the dialysis tubing into three separate microcentrifuge tubes (about 300 μL each).
8. Repeat **steps 6** and 7 to make sure that all the probe has been eluted from the acrylamide fragment. At the end of the second elution, recover the TBE again into three other

microcentrifuge tubes. The probe is then purified by DNA purification columns (like the Elutip®-D columns, Whatman, Inc., Schleicher & Schuell, Sanford ME).

9. Precipitate the probe with 1/10 volume of NaOAc 3 M pH 5.2 and 2 volumes of cold 90% ethanol. Allow labeled DNA to precipitate on dry ice for 30 min.
10. Centrifuge and discard the supernatant and resuspend DNA in 50 µL of sterile H_2O. Pool every sample into one microcentrifuge tube and precipitate as in **step 9**.
11. Count the DNA pellet using a beta counter by Cerenkov counting or resuspend DNA in a small volume (100 µL) and count a 1-µL aliquot in scintillation liquid.
12. Resuspend labeled DNA in order to obtain 30,000 cpm/µL.

3.2.2. For a Double-Stranded Oligonucleotide Labeled Probe

1. Proceed as in Subheading 3.2.1 except that **steps 1–8** should be omitted and replaced by two sequential precipitations as described in **step 9** (*see* **Note 14**).

3.3. EMSA

1. Rigorously clean and dry the electrophoresis tank and its accessories prior to use and treat the glass plates as previously described for probe isolation (**Subheading 3.2.1**, **step 1)**.
2. For a typical 70-bp probe, prepare a 6% polyacrylamide gel (*see* **Note 7**) by mixing 2.5 mL of 10× Tris–glycine, 3.75 mL of 40% acrylamide (39:1) stock solution and H_2O to obtain 25 mL. Add 180 µL of 10% ammonium persulfate and 30 µL of TEMED. Carefully stir and pour the acrylamide solution between the plates (*see* **Note 15**). Use a comb that has 0.8-cm width teeth. Allow the gel to set for at least 2 h then mount gel in the electrophoresis tank and fill the chamber with 1× Tris–glycine (*see* **Note 8**). As soon as the gel is mounted and set, remove the comb and carefully wash the wells with running buffer.
3. Pre-run the gel at 4°C and 120 V until the current becomes invariant (this usually takes 30 min on average). This pre-running step ensures that the gel will be at a constant temperature at the moment of sample loading.
4. When the gel is ready for loading, prepare samples as follows. For each sample, mix 12 µL of 2× binding buffer, 1 µL of 1 mg/mL poly(dI:dC) (*see* **Notes 10**, **11**, and **17**) and 0.6 µL of 2 M KCl (*see* **Note 11**) then add 30,000 cpm of labeled probe. When possible, pool together invariant components in one microcentrifuge tube and then redistribute equal amounts into different tubes accounting for the different experimental conditions. Finally, add 1–10 µg nuclear proteins and add H_2O to a final volume of 24 µL. Mix gently each

tube and incubate at RT for 3 min. As a control, prepare a sample without nuclear extract and add 1 μL of 6× loading buffer containing bromophenol blue and xylene cyanol.

5. Load samples by changing the pipette tip for each sample.
6. Run at 120 V and let samples migrate until the free probe reaches the bottom of the gel (*see* **Note 8**). In the case of a 70-pb probe loaded on a 6% acrylamide gel, this means 5–6 h of migration.
7. After the gel is run, disassemble the apparatus and remove one of the glass plates, place a Whatman paper over the gel and carefully lift the gel off the remaining plate. Make sure that the gel is well fixed on the Whatman before lifting the gel to avoid gel breakage. Place a Saran Wrap® over the gel and dry at 80°C for 30 min.
8. Place an X-ray film over the gel in an autoradiography cassette and expose at –80°C overnight.

4. Notes

1. Very intense, large, or smeary shifted complexes usually result from multiple comigrating DNA–protein complexes that possess nearly identical electrophoretic mobilities in native polyacrylamide gels despite that the nuclear proteins they contain usually have distinctive molecular masses on denaturing SDS-PAGE *(51, 52)*. An attractive method that helps to distinguish between the nuclear proteins yielding these multiple, comigrating complexes is the SDS-polyacrylamide gel fractionation–renaturation procedure *(53)*. It allows recovery and enrichment of nuclear proteins that then become suitable for further analyses by EMSA, in addition to providing their precise molecular mass. Changing the concentration of the native acrylamide gel may also proved useful in resolving comigrating DNA–protein complexes, as can be appreciated in **Fig. 1a**.
2. When using crude nuclear extracts for detecting DNA–protein complexes in EMSA, their quality is obviously very critical. Whenever possible, nuclei purification procedures using sucrose pads *(54)* is preferred in order to eliminate contamination by cytosolic proteins, which most often also contain substantial amounts of proteases. Purifying nuclei on sucrose pads has generally yielded high-quality nuclear extract samples. However, they require large quantities of fresh tissue, which renders the approach inappropriate when limiting

amounts of small animal tissues such as spleen, pancreas, or prostate are available. In these cases, short, microprocedures adapted to prevent protease actions can also be performed *(55)*. Once the crude extract has been obtained, its quality must then be evaluated. An informative way to test this consists in assessing the DNA binding ability of the ubiquitously expressed positive transcription factor Sp1. We have found this transcription factor to be particularly sensitive to the action of proteases *(56)*. Little or no Sp1 binding to its high-affinity binding site (5′-GATCATAT-CTGCGGGGCG-GGGCAGACACAG-3′) *(57)* used as a probe in EMSA is usually indicative of poor-quality nuclear extracts. Although such an assay is clearly invaluable when crude extracts are obtained from established tissue-culture cells, some caution must be observed when they are prepared from whole animal tissues since not all organs express Sp1 at the same level *(30, 56, 58)*.

3. The use of crude extracts prepared from whole animal tissues in EMSA is still restricted partly because of the numerous enzymatic activities they support, such as proteases and deacetylases, that strongly interfere with the EMSA's sensitivity. Degradation of nuclear proteins by endogenous proteases can be prevented by the further addition of protease inhibitors to the buffers used. However, whole animal tissue extracts are also often contaminated with considerable amounts of highly active endogenous phosphatases. Tissues such as liver, kidney, and bone have been reported to be rich in these enzymes *(59)*, some of which substantially decrease the sensitivity of the EMSA by removing the [^{32}P]-labeled phosphate from the DNA probe. Although addition of phosphatase inhibitors, such as tetrasodium pyrophosphate or sodium fluoride, to the reaction buffer can efficiently prevent dephosphorylation, we have found that the same can also be achieved by simply reducing either the temperature at which the binding reaction is normally performed (30 min of incubation at 4°C) or the time allowed for the DNA–protein interaction to occur (as low as 1 min of incubation at 22°C) *(29)*. Alternatively, probes labeled by fill-out of unpaired 5′ termini using T_4 DNA polymerase or the Klenow fragment of DNA polymerase I and an appropriate [α-^{32}P] dNTP may be used (*see also* **Note 13**).
4. When double-stranded oligonucleotides are selected as labeled probes in EMSA, we recommend their size to be in the range of 20–70 bp. When working with subcloned DNA sequences, optimal signal strength and resolution can be achieved using fragments of 50–250 bp. Although larger fragments may be used, they require longer migration time

in order to efficiently resolve the potential DNA–protein complexes. Furthermore, larger labeled probes are likely to bind an increased number of nuclear proteins, which might complicate result's interpretation.

5. Handling [γ-^{32}P] ATP requires that special care be taken when labeling the DNA probes used in EMSA. The reader is referred to standard procedures and guidelines on manipulation of radioactive materials in effect at its own research facility. Alternative procedures for nonradioactive probe labeling have been reported for EMSA analyses *(9, 13, 60–62)*.
6. Acrylamide is a potent neurotoxic compound that is easily absorbed through skin. Wearing gloves and a mask to avoid direct contact with the skin or inhalation is therefore required when manipulating acrylamide dried or in solution. Similar care should also be taken with polyacrylamide gels, as they might still contain low levels of unpolymerized acrylamide. The acrylamide solution is light-sensitive and should be kept away from direct light. It is worth noting that acrylamide and bis-acrylamide are slowly converted to acrylic and bis-acrylic acid, respectively, upon prolonged storage. To avoid the use of acrylamide, alternative nontoxic gel matrices are available, with resolution properties comparable to those of polyacrylamide *(14, 15)*. The use of agarose gels containing a nontoxic synergistic gelling and sieving agent (Synergel™) that helps improve the resolution of DNA–protein complexes has also been reported *(16)*.
7. The concentration of the polyacrylamide gel used in EMSA is primarily dictated by both the size of the labeled probe selected and the resolution of the DNA–protein complexes obtained. It can vary from 4% with large labeled DNA fragments (of over 150 bp in length) up to 12% with synthetic oligonucleotides (we even used 14% gels on some specific occasion). Two (or more) closely migrating DNA–protein complexes that would normally appear as a single diffuse, smeary complex on a 4% gel can usually be resolved on a 8% gel. However, although increasing the gel concentration clearly improved the resolution of most DNA–protein complexes, formation of others turned out to be unstable under high gel concentration therefore interfering with their detection in EMSA.
8. Formation of DNA–protein complexes is highly dependent on the voltage selected for their migration into the polyacrylamide gel *(7)*. We have found that reducing the migration time by running the EMSA at voltage lower than 120 V (usually corresponding to 10 mA for a single gel) on a 6% polyacrylamide gel rendered very unstable the formation of most DNA–protein complexes and consequently resulted in our inability to detect them (*see* **Fig. 1c**).

9. Although we feel DNA–protein interactions are best revealed using the Tris–glycine buffer system, some complexes may turn out to be undetectable under such conditions. The alternative use of other running buffer systems with varying ionic strength such as Tris–acetate, pH 7.5 or TBE, pH 8.0 *(34, 63)* is advisable in order to explore all kind of DNA–protein binding kinetics.
10. Nonspecific DNA–protein interactions are usually prevented by the addition to the reaction mix of 1–5 μg of a nonspecific competitor DNA. Although this is clearly very effective when crude nuclear extracts are used, such high concentrations of nonspecific competitor DNA were found to compete even for specific DNA–protein complexes when enriched preparations of nuclear proteins are used in EMSA *(64)*. The more enriched the nuclear protein of interest, the lower the amount of nonspecific competitor required. For example, we routinely use 1–2 μg poly(dI:dC) with crude nuclear proteins, 250 ng when the nuclear extract is enriched on heparin-Sepharose column, and no more than 25–50 ng with purified or recombinant proteins.
11. The signal strength of a shifted DNA–protein complex can be substantially increased by favoring the odds for the interaction between the protein of interest and its target sequence. This can easily be achieved with enriched preparations of nuclear proteins either by increasing the amount of the labeled probe selected, or by decreasing the concentration of poly(dI:dC), or both. Furthermore, the DNA binding ability of some nuclear proteins proved to be highly dependent on the amount of salt (usually provided by KCl) that is present in the reaction mix. Transcription factors such as NFI and Sp1 best interact with their respective target sequence in the presence of 100 mM and 150 mM KCl, respectively *(28)*. It is therefore useful to evaluate which KCl concentration best allows binding of nuclear proteins to a specific DNA target probe.
12. Ethidium bromide is a suspected carcinogen that also possesses a moderate toxicity. Wearing gloves is required when manipulating solutions that contain this DNA dying agent. Decontamination of ethidium bromide-containing solutions can be achieved using either hypophosphorous acid or potassium permanganate (*see* **ref. 49** for an overview and detailed protocols).
13. Nearly all-vertical electrophoresis apparatus can be used in order to perform EMSA analyses. Although some EMSA protocols involve gel migration at room temperature, we recommend the polyacrylamide gel to be run at 4°C. With some apparatus, this can be easily achieved using cooling units specially designed for these pieces of equipment. However, for those which are not equipped with such cooling apparatus, simply run the gel in a cold room.

14. Although 5′ end-labeling of the selected DNA fragment is best done using polynucleotide kinase, very efficient labeling can also be accomplished using alternative procedures, such as filling 5′-protruding ends using the Klenow fragment of *E. coli* DNA polymerase I *(49)*, a particularly attractive alternative when crude nuclear extracts rich in various phosphatases are used (in the event that no phosphatase inhibitors are used in the binding buffer). Larger DNA segments can also be efficiently labeled by PCR.
15. Chemical synthesis of any oligonucleotide yields a substantial proportion of intermediate products of progressively decreasing length, which is particularly true for larger oligos, since the efficiency of synthesis normally ranges between 98.5 and 99% for each single nucleotide addition. For a 40-nucleotide oligo, this means that 60% of the synthesized products are of the correct length and that the remaining 40% range in size between 1 and 39 nucleotides. Further HPLC, OPC column, or gel purification is recommended before proceeding to the annealing of the oligos. The loss of a few bases on either side of the resulting double-stranded oligonucleotide might be sufficient to prevent protein binding and therefore result in a reduced ability to detect such a DNA–protein complex in EMSA.
16. We have found that thickness of the native polyacrylamide gel profoundly affects resolution of the shifted DNA–protein complexes: the thinner the gel, the better the resolution. We currently load our sample reactions on 0.75-mm thick gels.
17. Adding high concentrations of neutral osmolytes such as betaine glycine, triethylene glycol, or methylglucoside in reaction mixture for EMSA may help prevent dissociation of already formed specific DNA–protein complexes *(65, 66)*.V

Acknowledgments

The experimental data included in this chapter were supported by a grant from the National Science and Engineering Research Council of Canada (NSERC) (grant #138624-06) to S.L.G.

References

1. Garner, M.M. and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**, 3047–3060.
2. Fried, M. and Crothers, D.M. (1981). Equilibria and kinetics of lac repressor–operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**, 6505–6525.
3. Gerstle, J.T. and Fried, M.G. (1993). Measurement of binding kinetics using the gel

electrophoresis mobility shift assay. *Electrophoresis* **14**, 725–731.

4. Fried, M.G. and Daugherty, M.A. (1998). Electrophoretic analysis of multiple protein–DNA interactions. *Electrophoresis* **19**, 1247–1253.
5. Cann, J.R. (1998). Theoretical studies on the mobility-shift assay of protein–DNA complexes. *Electrophoresis* **19**, 127–141.
6. Cann, J.R. (1989). Phenomenological theory of gel electrophoresis of protein–nucleic acid complexes. *J. Biol. Chem.* **264**, 17032–17040.
7. Vossen, K.M. and Fried, M.G. (1997). Sequestration stabilizes lac repressor–DNA complexes during gel electrophoresis. *Anal. Biochem.* **245**, 85–92.
8. Cann, J.R. (1997). Models of mobility-shift assay of complexes between dimerizing protein and DNA. *Electrophoresis* **18**, 1092–1097.
9. Suske, G., Gross, B. and Beato, M. (1989). Non-radioactive method to visualize specific DNA–protein interactions in the band shift assay. *Nucleic Acids Res.* **17**, 4405.
10. Laniel, M.A., Bergeron, M.J., Poirier, G.G. and Guérin, S.L. (1997). A nuclear factor other than Sp1 binds the GC-rich promoter of the gene encoding rat poly(ADP-ribose) polymerase in vitro. *Biochem. Cell Biol.* **75**, 427–434.
11. Kironmai, K.M., Muniyappa, K., Friedman, D.B., Hollingsworth, N.M. and Byers, B. (1998). DNA-binding activities of Hop1 protein, a synaptonemal complex component from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **18**, 1424–1435.
12. Wolf, S.S., Hopley, J.G. and Schweizer, M. (1994). The application of 33P-labeling in the electrophoretic mobility shift assay. *Biotechniques* **16**, 590–592.
13. Ludwig, L.B., Hughes, B.J. and Schwartz, S.A. (1995). Biotinylated probes in the electrophoretic mobility shift assay to examine specific dsDNA, ssDNA or RNA–protein interactions. *Nucleic Acids Res.* **23**, 3792–3793.
14. Ramanujam, P., Fogerty, S., Heiser, W. and Jolly, J. (1990). Fast gel electrophoresis to analyze DNA–protein interactions. *Biotechniques* **8**, 556–563.
15. Vanek, P.G., Fabian, S.J., Fisher, C.L., Chirikjian, J.G. and Collier, G.B. (1995). Alternative to polyacrylamide gels improves the electrophoretic mobility shift assay. *Biotechniques* **18**, 704–706.
16. Chandrasekhar, S., Souba, W.W. and Abcouwer, S.F. (1998). Use of modified agarose gel electrophoresis to resolve protein–DNA complexes for electrophoretic mobility shift assay. *Biotechniques* **24**, 216–218.
17. Revzin, A. (1989). Gel electrophoresis assays for DNA–protein interactions. *Biotechniques* **7**, 346–355.
18. Hassanain, H.H., Dai, W. and Gupta, S.L. (1993). Enhanced gel mobility shift assay for DNA-binding factors. *Anal. Biochem.* **213**, 162–167.
19. Zhang, X.Y., Asiedu, C.K., Supakar, P.C. and Ehrlich, M. (1992). Increasing the activity of affinity-purified DNA-binding proteins by adding high concentrations of nonspecific proteins. *Anal. Biochem.* **201**, 366–374.
20. Murakami, Y., Huberman, J.A. and Hurwitz, J. (1996). Identification, purification, and molecular cloning of autonomously replicating sequence-binding protein 1 from fission yeast *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. USA* **93**, 502–507.
21. Zhang, W., Shields, J.M., Sogawa, K., Fujii-Kuriyama, Y. and Yang, V.W. (1998). The gut-enriched Kruppel-like factor suppresses the activity of the CYP1A1 promoter in an Sp1-dependent fashion. *J. Biol. Chem.* **273**, 17917–17925.
22. Tyree, C.M., George, C.P., Lira-DeVito, L.M., Wampler, S.L., Dahmus, M.E., Zawel, L. and Kadonaga, J.T. (1993). Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes Dev.* **7**, 1254–1265.
23. Gille, J., Swerlick, R.A. and Caughman, S.W. (1997). Transforming growth factor-alpha-induced transcriptional activation of the vascular permeability factor (VPF/VEGF) gene requires AP-2-dependent DNA binding and transactivation. *EMBO J.* **16**, 750–759.
24. Roy, A.L., Du, H., Gregor, P.D., Novina, C.D., Martinez, E. and Roeder, R.G. (1997). Cloning of an inr- and E-box-binding protein, TFII-I, that interacts physically and functionally with USF1. *EMBO J.* **16**, 7091–7104.
25. Ng, J.Y. and Marians, K.J. (1996). The ordered assembly of the phiX174-type primosome. I. Isolation and identification of intermediate protein–DNA complexes. *J. Biol. Chem.* **271**, 15642–15648.
26. Wakasugi, M. and Sancar, A. (1998). Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proc. Natl Acad. Sci. U. S. A.* **95**, 6669–6674.
27. Gaudreault, M., Vigneault, F., Leclerc, S. and Guérin, S.L. (2007). Laminin reduces expression of the human alpha6 integrin subunit gene by altering the level of the transcription factors Sp1 and Sp3. *Invest. Ophthalmol. Vis. Sci.* **48**(8), 3490–3505.
28. Robidoux, S., Guérin, S.L., Eskild, W., Kroepelin, C.F. and Hansson, V. (1992). Salt-dependent formation of DNA/protein

complexes in vitro, as viewed by the gel mobility shift assay. *Biotechniques* **13**, 354–357.

29. Laniel, M.A. and Guérin, S.L. (1998). Improving sensitivity of the electrophoretic mobility shift assay by restricting tissue phosphatase activities. *Biotechniques* **24**, 964–970.
30. Potvin, F., Roy, R.J., Poirier, G.G. and Guérin, S.L. (1993). The US-1 element from the gene encoding rat poly(ADP-ribose) polymerase binds the transcription factor Sp1. *Eur. J. Biochem.* **215**, 73–80.
31. Dasgupta, A. and Scovell, W.M. (2003). TFIIA abrogates the effects of inhibition by HMGB1 but not E1A during the early stages of assembly of the transcriptional preinitiation complex. *Biochim. Biophys. Acta* **1627**, 101–110.
32. Jing, D., Beechem, J.M. and Patton, W.F. (2004). The utility of a two-color fluorescence electrophoretic mobility shift assay procedure for the analysis of DNA replication complexes. *Electrophoresis* **25**, 2439–2446.
33. Smider, V., Hwang, B.J. and Chu, G. (2006). Electrophoretic mobility shift assays to study protein binding to damaged DNA. *Methods Mol. Biol.* **314**, 323–344.
34. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G. and Smith, J.A. (eds.) (1992). *Short Protocols in Molecular Biology.* Wiley, New York, NY.
35. Bergeron, M.J., Leclerc, S., Laniel, M.A., Poirier, G.G. and Guérin, S.L. (1997). Transcriptional regulation of the rat poly(ADP-ribose) polymerase gene by Sp1. *Eur. J. Biochem.* **250**, 342–353.
36. Granger-Schnarr, M., Lloubes, R., de Murcia, G. and Schnarr, M. (1988). Specific protein–DNA complexes: immunodetection of the protein component after gel electrophoresis and Western blotting. *Anal. Biochem.* **174**, 235–238.
37. Osborn, M.T., Herrin, K., Buzen, F.G., Hurlburt, B.K. and Chambers, T.C. (1999). Electrophoretic mobility shift assay coupled with immunoblotting for the identification of DNA-binding proteins. *Biotechniques* **27**, 887–890, 892.
38. Demczuk, S., Harbers, M. and Vennström, B. (1993). Identification and analysis of all components of a gel retardation assay by combination with immunoblotting. *Proc. Natl Acad. Sci. USA* **90**, 2574–2578.
39. Dyer, R.B. and Herzog, N.K. (1995). Immunodepletion EMSA: a novel method to identify proteins in a protein–DNA complex. *Nucleic Acids Res.* **23**, 3345–3346.
40. Yamamoto, H. (1997). DNA mobility shift assay coupled with SDS-PAGE for detection of DNA-binding proteins. *Biotechniques* **22**, 210–211.
41. Williams, M., Brys, A., Weiner, A.M. and Maizels, N. (1992). A rapid method for determining the molecular weight of a protein bound to nucleic acid in a mobility shift assay. *Nucleic Acids Res.* **20**, 4935–4936.
42. Adachi, Y., Chen, W., Shang, W.H. and Kamata, T. (2005). Development of a direct and sensitive detection method for DNA-binding proteins based on electrophoretic mobility shift assay and iodoacetamide derivative labeling. *Anal. Biochem.* **342**, 348–351.
43. Funabashi, H., Ubukata, M., Ebihara, T., Aizawa, M., Mie, M. and Kobatake, E. (2007). Assessment of small ligand–protein interactions by electrophoretic mobility shift assay using DNA-modified ligand as a sensing probe. *Biotechnol. Lett.* **29**, 785–789.
44. Hope, I.A. and Struhl, K. (1985). GCN4 protein, synthesized in vitro, binds HIS3 regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast. *Cell* **43**, 177–188.
45. Filion, G.J., Fouvry, L. and Defossez, P.A. (2006). Using reverse electrophoretic mobility shift assay to measure and compare protein–DNA binding affinities. *Anal. Biochem.* **357**, 156–158.
46. Ronai, Z., Wang, Y., Khandurina, J., Budworth, P., Sasvari-Szekely, M., Wang, X. and Guttman, A. (2003). Transcription factor binding study by capillary zone electrophoretic mobility shift assay. *Electrophoresis* **24**, 96–100.
47. Chuang, Y.J., Huang, J.W., Makamba, H., Tsai, M.L., Li, C.W. and Chen, S.H. (2006). Electrophoretic mobility shift assay on poly(ethylene glycol)-modified glass microchips for the study of estrogen responsive element binding. *Electrophoresis* **27**, 4158–4165.
48. Fraga, M.F., Ballestar, E. and Esteller, M. (2003). Capillary electrophoresis-based method to quantitate DNA–protein interactions. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **789**, 431–435.
49. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning, A Laboratory Manual.* Cold Spring Harbor, New York, NY.
50. Harvey, M., Brisson, I. and Guérin, S.L. (1993). A simple apparatus for fast and inexpensive recovery of DNA from polyacrylamide gels. *Biotechniques* **14**, 942–948.
51. Larouche, K., Leclerc, S., Giasson, M. and Guérin, S.L. (1996). Multiple nuclear regulatory proteins bind a single cis-acting promoter element to control basal transcription of the human alpha 4 integrin gene in corneal epithelial cells. *DNA Cell Biol.* **15**, 779–792.
52. Leclerc, S., Eskild, W. and Guérin, S.L. (1997). The rat growth hormone and human cellular retinol binding protein 1 genes share homolo-

gous NF1-like binding sites that exert either positive or negative influences on gene expression in vitro. *DNA Cell Biol.* **16**, 951–967.

53. Ossipow, V., Laemmli, U.K. and Schibler, U. (1993). A simple method to renature DNA-binding proteins separated by SDS-polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **21**, 6040–6041.
54. Graves, B.J., Johnson, P.F. and McKnight, S.L. (1986). Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* **44**, 565–576.
55. Roy, R.J., Gosselin, P. and Guérin, S.L. (1991). A short protocol for micro-purification of nuclear proteins from whole animal tissue. *Biotechniques* **11**, 770–777.
56. Robidoux, S., Gosselin, P., Harvey, M., Leclerc, S. and Guérin, S.L. (1992). Transcription of the mouse secretory protease inhibitor p12 gene is activated by the developmentally regulated positive transcription factor Sp1. *Mol. Cell. Biol.* **12**, 3796–3806.
57. Dynan, W.S. and Tjian, R. (1983). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**, 79–87.
58. Saffer, J.D., Jackson, S.P. and Annarella, M.B. (1991). Developmental expression of Sp1 in the mouse. *Mol. Cell. Biol.* **11**, 2189–2199.
59. McComb, R.B., Bowers, G.N. and Posen, S. (1979). *Alkaline Phosphatase.* Plenum, New York, NY.
60. Ruscher, K., Reuter, M., Kupper, D., Trendelenburg, G., Dirnagl, U. and Meisel, A. (2000). A fluorescence based non-radioactive electrophoretic mobility shift assay. *J. Biotechnol.* **78**, 163–170.
61. Zhang, N., Xu, Y., Zhang, Z. and Xiong, W. (2003). A nonradioactive method for detecting DNA-binding activity of nuclear transcription factors. *J. Huazhong Univ. Sci. Technol. Med. Sci.* **23**, 227–229.
62. Jing, D., Agnew, J., Patton, W.F., Hendrickson, J. and Beechem, J.M. (2003). A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. *Proteomics* **3**, 1172–1180.
63. Roder, K. and Schweizer, M. (2001). Running-buffer composition influences DNA–protein and protein–protein complexes detected by electrophoretic mobility-shift assay (EMSA). *Biotechnol. Appl. Biochem.* **33**, 209–214.
64. Larouche, K., Bergeron, M.J., Leclerc, S. and Guérin, S.L. (1996). Optimization of competitor poly(dI–dC).poly(dI–dC) levels is advised in DNA–protein interaction studies involving enriched nuclear proteins. *Biotechniques* **20**, 439–444.
65. Sidorova, N.Y. and Rau, D.C. (2000). The dissociation rate of the EcoRI–DNA-specific complex is linked to water activity. *Biopolymers* **53**, 363–368.
66. Sidorova, N.Y., Muradymov, S. and Rau, D.C. (2005). Trapping DNA–protein binding reactions with neutral osmolytes for the analysis by gel mobility shift and self-cleavage assays. *Nucleic Acids Res.* **33**, 5145–5155.

Chapter 3

DNase I Footprinting

Benoît Leblanc and Tom Moss

Summary

The association of proteins with the DNA double helix can interfere with the accessibility of the latter to nucleases. This is particularly true when using bulky nucleases such as DNAse I. The DNAse I footprinting method was developed to make use of this phenomenon in the study of DNA–protein interactions; it consists in comparing the pattern of fragments produced by the partial digestion of DNA in the absence of a protein to that produced by partial digestion of DNA in the presence of a protein. Normally, when the two sets of fragments are separated side by side on a gel, the fragments produced in the presence of the protein will feature blank regions (indicating protection) and/or enhanced cleavage sites (indicating increased availability). This technique can furthermore reveal if multiple sites for a DNA-binding protein are present on a same fragment, and allow the comparison of their respective affinities.

Key words: DNAse I, Footprint, Footprinting, Protection.

1. Introduction

DNase I footprinting was developed by Galas and Schmitz in 1978 as a method to study the sequence-specific binding of proteins to DNA *(1)*. In the technique, a suitable uniquely end-labeled DNA fragment is allowed to interact with a given DNA-binding protein and the complex is then partially digested with DNase I. The bound protein protects the region of DNA with which it interacts from attack by the DNase. Subsequent molecular-weight analysis of the degraded DNA by electrophoresis and autoradiography identifies the region of protection as a gap in the otherwise continuous background of digestion products; for example, *see* **Fig. 1**. The technique can be used to determine the site of interaction

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_3

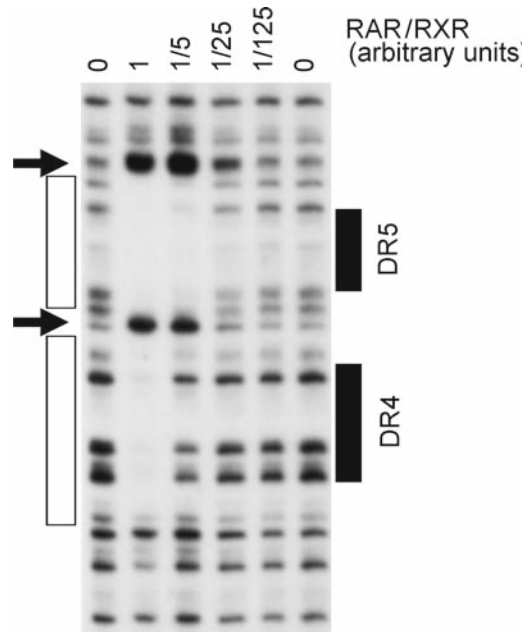


Fig. 1. Example of DNase I footprinting. A DNA fragment containing a mutated version of the mouse RARβ2 gene promoter, labeled downstream of the initiation site, was incubated with decreasing amounts of recombinant RAR/RXR heterodimers and partially digested with DNase I. The direct repeat (DR) elements bound by the proteins are indicated by *solid boxes*. Protected regions are indicated by *empty boxes*. DNase I-hypersensitive sites, which are more frequently cleaved in the presence of the proteins than in their absence, are indicated by *arrowheads*. Note that the heterodimers bind to direct repeats separated by 5 nucleotides (DR5) with roughly five times more affinity than to direct repeats separated by 4 nucleotides (DR4), a preference that can be easily visualized by this technique.

of most sequence-specific DNA-binding proteins but had been most extensively applied to the study of transcription factors. Because the DNase I molecule is relatively large as compared to other footprinting agents (*see* Chapters "Hydroxyl Radical Footprinting of Protein–DNA Complexes" and "The Use of Diethyl Pyrocarbonate and Potassium Permanganate as Probes for Strand Separation and Structural Distortions in DNA" on the use of hydroxy radicals and diethylpyrocarbonate), its attack on the DNA is relatively easily sterically hindered. Thus, DNase I footprinting is the most likely of all the footprinting techniques to detect a specific DNA–protein interaction. This is clearly demonstrated by our studies on the transcription factor xUBF, a factor binding DNA with relatively low affinity (*see* **Fig. 2**).

DNase I footprinting can not only be used to study the DNA interactions of purified proteins but also be used as an assay to identify proteins of interest within a crude nuclear extract *(2)*. Thus, it can serve much the same function as an EMSA analysis (*see* Chapter 2 "Electrophoretic Mobility Shift Assays for the Analysis

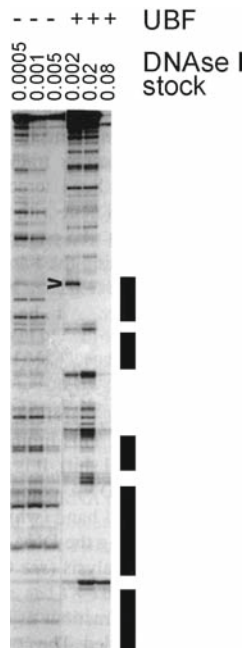


Fig. 2. Course of digestion with increasing amounts of DNase I. Here xUBF was footprinted on the *Xenopus* ribosomal promoter using a 5′ end-labeled fragment. The *numbers* above the tracks refer to the DNase I dilution (in units/μL) employed and *minus* and *plus* to the naked and complexed DNAs, respectively. The predominant footprints are indicated by *solid boxes*. A visible double-stranded fragment, which could be mistaken for a hypersensitive site, is marked by an *arrowhead*.

of DNA–Protein Interactions") in following a specific DNA-binding activity through a series of purification steps. Because DNase I footprinting can often be used for proteins that do not perform well in EMSA experiments (the above-mentioned UBF is an example of a protein more easily studied by footprinting than by EMSA), it has more general applicability. However, because of the need of an excess of protein and the visualization of the footprint by a partial digestion ladder, the technique requires considerably more material than would an EMSA assay and cannot distinguish individual components of heterogeneous DNA–protein complexes.

DNase I (E.C.3.1.4.5) is a protein approximately 40 Å in diameter. It binds in the minor groove of the DNA and cuts the phosphodiester backbone of both strands independently *(3)*. Its bulk helps to prevent it from cutting the DNA under and around a bound protein. However, a bound protein will also usually have other effects on the normal cleavage by DNase I (*see* **Figs. 1** and **2**). It is also not so uncommon to observe a change in the pattern of DNase cleavage without any obvious extended protection (*see* **Fig. 2**).

Unfortunately, DNase I does not cleave the DNA indiscriminately, some sequences being very rapidly attacked while others remain unscathed even after extensive digestion *(4)*. This results in a rather uneven "ladder" of digestion products after electrophoresis, something which limits the resolution of the technique (as seen in the naked DNA tracks in **Figs. 1** and **2**). However, when the protein-protected and the naked DNA ladders are run alongside each other, the footprints are normally quite apparent. To localize the position of the footprints, sequencing ladders originating from the same nucleotide that is labeled in the footprinted DNA fragment should accompany the naked and the protected tracks (*see* **Note 1**). As a single end-labeled fragment allows one to visualize interactions on only one strand of DNA at a time, it is informative to repeat the experiment with the same fragment labeled on the other strand. DNA fragment can be conveniently 5′ labeled using T_4 DNA polynucleotide kinase and 3′ labeled using the Klenow or the T_4 DNA polymerases (in fill out reactions) or terminal transferase *(5)*. To analyze both strands of the DNA duplex side by side and make direct comparisons between the two, two footprinting reactions should be run in parallel. Both would use the same DNA fragment labeled at the same end, but one would label the 5′-strand and the other 3′-strand.

DNase I footprinting requires an excess of DNA-binding activity over the amount of DNA fragment used. The higher the precent occupancy of a site on the DNA, the clearer a footprint will be observed. It is therefore important not to titrate the available proteins with too much DNA. This limitation can, in part, be overcome when a protein also generates a gel-shift. It is then feasible to fractionate the partially DNase-digested protein–DNA complex by nondenaturing gel electrophoresis and to excise the shifted band (which is then a homogeneous protein–DNA complex) before analyzing the DNA by denaturing gel electrophoresis as in the standard footprint analysis.

Footprinting crude or impure protein fractions usually requires that an excess of a nonspecific competitor DNA be added to the reaction. The competitor binds nonspecific DNA-binding proteins as effectively as the specific labeled target DNA fragment and hence, when present in sufficient excess, leaves the main part of the labeled DNA available for the sequence specific protein. Homogeneous and highly enriched protein fractions usually do not require the presence of a nonspecific competitor during footprinting. When planning a footprinting experiment, it is prerequisite to start by determining the optimal concentration of DNase I to be used. This will be a linear function of the amount of nonspecific competitor but more importantly (although less reproducibly) it will also be a function of the amount and purity of the protein fraction added. As a general rule, more DNase I will be required if more protein is present in the binding reaction,

whether or not this protein binds specifically. Thus, very different DNase concentrations may be required to produce the required degree of digestion on naked and protein-bound DNA. A careful titration of the DNase concentration is therefore essential to optimize the detection of a footprint and can even make the difference between the detection of a given interaction or lack thereof.

The following protocol was developed to study the footprinting of the *Xenopus* ribosomal transcription factor xUBF, which is a rather weak DNA-binding protein with a rather broad specificity. It can be broadly applied and has been used successfully for studies on the human retinoic acid receptor alpha and the yeast Ace1 transcription factor, among others. We recommend that the reader also refer to **ref.** *6* for more information on the quantitative analysis of protein–DNA interactions by footprinting.

2. Materials

1. Binding buffer (2×): 20% glycerol, 0.2 mM EDTA, 1 mM dithiothreitol (DTT), 20 mM HEPES pH 7.9, and 4% poly (vinyl alcohol) (*see* **Note 2**).
2. poly(dAdT)-poly(dAdT) (Sigma-Aldrich #P0883): 1 mg/mL solution in TE. Keep at –20°C (*see* **Note 3**).
3. End-labeled DNA fragment of high specific activity (*see* **Note 1**).
4. Cofactor solution: 10 mM $MgCl_2$, 5 mM $CaCl_2$.
5. DNase I stock solution: a standardized vial of DNase I (Sigma-Aldrich #D4263) is dissolved in 50% glycerol, 135 mM NaCl, 15 mM CH_3COONa pH 6.5 at a concentration of 10 units/μL. This stock solution can be kept at –20°C indefinitely (*see* **Note 4**).
6. 1 M KCl solution.
7. Reaction Stop solution: 1% sodium dodecyl sulfate (SDS), 200 mM NaCl, 20 mM EDTA pH 8.0, and 40 μg/mL tRNA (*see* **Note 5**).
8. 10× TBE buffer: 900 mM Tris–borate pH 8.3 (108 g/L Tris base and 55 g/L boric acid, 20 mM EDTA).
9. Loading buffer: 95% formamide, 0.05% xylene cyanol, 0.05% bromophenol blue.
10. 6% Acrylamide, 7 M urea, and 1× TBE sequencing gel.
11. Phenol: chloroform: isoamyl alcohol (25:24:1, v/v) (Invitrogen #15593-031).

12. Sequenase DNA sequencing kit (USB #70770) or an equivalent.
13. A sequencing primer of appropriate sequence, as defined in **Subheading 3, step 12**.
14. Fixing solution: 10% ethanol, 10% acetic acid.
15. Blotting paper (Whatman 3MM or the equivalent).

3. Methods

The footprinting reaction is performed in three stages (1) binding of the protein to the DNA; (2) partial digestion of the protein–DNA complex with DNase I and recovery of these fragments; and (3) separation of the fragments on a DNA sequencing gel.

1. Each binding reaction is performed in a total volume of 50 µL containing 25 µL of 2× binding buffer, 0.5 µL of 1 mg/mL poly(dAdT)-poly(dAdT), 2–3 ng of end-labeled DNA fragment (approximately 15,000 cpm) (*see* **Note 6**), an adjustable volume of protein fraction, and enough 1 M KCl to bring the final KCl concentration to 60 mM. The maximum volume of the protein fraction that can be used will often be determined by how much salt it contains. Each reaction is performed in a 1.5-mL Eppendorf tube. The following work chart, **Table 1**, can be used to facilitate the experiment.
2. Incubate on ice for 20 min (binding reaction).
3. During the binding reaction, prepare the many DNase I working dilutions in 1.5-mL Eppendorf tubes. We suggest diluting an aliquot of the stock solution in water, on ice, to concentrations ranging from 0.0005 to 0.1 Kunitz units/µL. A good

Table 1
Work chart for DNAse I footprinting reactions

Tube #	Labeled DNA (15,000 cpm)	2× Binding buffer	Poly(dAdT)–poly(dAdT) 1 mg/mL	Protein fraction	1 M KCl (to 60 mM final)	H_2O	Final volume
	µL	25 µL	0.5 µL	µL	µL	µL	50 µL
1							
2							
3							
4							
⋮							

range of useful dilutions would consist in aliquots at 0.0005, 0.001, 0.002, 0.004, 0.008 for the digestion of naked DNA, and 0.005, 0.01, 0.02, 0.04, and 0.08 for the digestion of protein–DNA complexes (*see* **Note 7**).

4. Set up the digestion reaction. You should have three micropipettes on hand: one set at 50 μL, another set at 5 μL, and the last one set at 100 μL. Pipette tip boxes should be nearby and kept open (as **step 5** will require to move quickly). The cofactor solution and the reaction stop solutions should both be at room temperature (R/T), in open tubes, and within easy reach. The DNase I dilutions should also be within easy reach, kept on ice, with the tube lids open. Have a watch or a stopwatch on the bench (not on your wrist). Since the DNA fragment is radioactive, all the operations should be performed behind a plexiglass shield.
5. Once the 20 min of the binding reaction is over, transfer the reaction tubes, eight at a time, to a rack at R/T. Each of the eight tubes will be processed one after the other, at 15-s intervals. Start with the first tube and (a) add 50 μL of cofactor solution; (b) add 5 μL of the appropriate DNase I dilution; (c) cap the tube and move on to the next one. (Roughly 15 s should have elapsed, as you can judge from the watch on the bench.)
6. After having processed the eighth tube, you go back to the first one and stop its digestion reaction by adding 100 μL Reaction Stop solution. As previously, move on to each subsequent tube every 15 s. The total digestion time for each of the eight tubes will have been 2 min (*see* **Note 8**).
7. After all the reactions have been processed, extract each of them with phenol:chloroform.
8. Add at least 1 volume of isopropanol to precipitate the nucleic acids.
9. Microcentrifuge for 15 min at approximately 15,000 × *g*. Remove the supernatant with a micropipette and keep it in a fresh tube in case the nucleic acids did not precipitate. Check the presence of a radioactive pellet with a Geiger counter. (The pellet might come unstuck and could be found floating in the supernatant. If that occurs, centrifuge again.)
10. Add 100 μL ice cold, 70% ethanol. Microcentrifuge for 2 min as above. Remove the supernatant, checking with the Geiger counter that you did not pick up the pellet (the supernatant at this step should not be radioactive, or only very slightly). Air dry the pellets or briefly dry in a vacuum dessicator. Do not overdry.
11. Resuspend each pellet in 5 μL loading buffer, vortex, and centrifuge briefly. The samples are now ready to be loaded

on a sequencing gel, but can be kept at –20°C and run at a later time.

12. A sequencing ladder should be run in parallel with the samples on the sequencing gel **(step 13)**. Traditionally, this ladder has mostly been a G + A chemical degradation ladder generated by the Maxam and Gilbert method *(5)*. Briefly, approximately 200,000 cpm of end-labeled DNA are diluted into 30 µL H_2O (no EDTA). Two microliters of 1 M piperidine formate (pH 2.0) are added and the solution is incubated at 37°C for 15 min; 150 µL of 1 M piperidine are then added directly and the solution is incubated at 90°C for 30 min in a well-sealed tube (we use a microtube in a thermal cycler). After incubation, the sample is transferred into a 1.5-mL Eppendorf tube and 20 µL of 3 M CH_3COONa and 500 µL of ethanol are added. DNA is precipitated at –80°C for 10–20 min. DNA is centrifuged (10,000 × *g*, 10 min), resuspended in 200 µL of 0.3 M CH_3COONa, and precipitated again with 500 µL of ethanol. It is centrifuged again as above and resuspended in 200 µL of H_2O. It is then lyophilized to remove the last traces of piperidine. The dry pellet is finally resuspended in gel loading buffer; about 5,000 cpm are used per track. However, a much better resolution can be achieved by using the four tracks (A, C, G, and T) of a sequencing reaction performed according to the Sanger protocol *(7)*. This can be done using home-made reagents or a commercial kit such as the one listed in **Subheading 2**. For the sequencing ladder and the footprinting ladders to be properly aligned, care must be taken in using a sequencing primer that aligns precisely with the labeled base of the DNA fragment that is footprinted (*see* **Note 9**). Although such a sequencing reaction is performed with ^{35}S instead of ^{32}P, the intensity of its signal after autoradiography is quite acceptable and in the range of that of the footprint ladders.
13. Pre-run a standard 6% polyacrylamide sequencing gel until hot (roughly 30 min) before loading the samples. Wash the wells thoroughly before loading. Denature the samples for 2 min at 95°C and load with a micropipette. (We find that regular tips, wedged at the top of the wells, work just as well as specialty elongated tips and do not clog as easily.) Run the gel hot to keep the DNA denatured (*see* **Note 10**). The gel should be run until the xylene cyanol reaches two-third of its way to the bottom.
14. After the run, fix the gel with 10% ethanol and 10% acetic acid. It is not necessary to immerse the gel in the solution; laying it flat in or near a sink and covering it with a thin layer of solution will suffice. (The gel plate can be put in a

large tray if no large sink is available.) The solution can be replenished from time to time for about 5–10 min. The gel should then be covered with a wet piece of blotting paper and washed with gently running water for about 3–4 min, which will help remove the urea. The paper is gently removed (make sure the gel does not stay attached to it or start tearing up!) and replaced with a new, dry piece of blotting paper. The gel is then dried in a gel drier.

15. Expose the gel overnight to an autoradiography film or with in phosphorimager cassette.

4. Notes

1. Single-stranded breaks in the end-labeled DNA fragment must be avoided as they give false signals indistinguishable from genuine DNase I cleavage and hence can mask an otherwise good footprint. It is therefore advisable, when running the gel, to keep one lane for the undigested footprinting probe which should give only one band. Radiochemical nicking is a theoretical threat to the probe's integrity, but in practice it has never caused us any problem.
2. This binding buffer has been shown to work well for the transcription factor NF-1 *(6)*. In our hands it has also worked well for several other factors, such as the human retinoid receptors RAR and RXR, the yeast Ace1 factor, and *Xenopus* and human UBF. Glycerol and poly (vinyl alcohol) (an agent used to reduce the available volume of water and hence concentrate the binding activity) are not mandatory. In fact, very simple binding buffers can be used; the original footprinting conditions for the binding of the *lac* repressor to the *lac* operator were 10 mM cacodylate buffer pH 8.0, 10 mM $MgCl_2$, 5 mM $CaCl_2$, and 0.1 mM DTT *(1)*. Particular conditions of binding might be required depending on the protein studied.
3. Poly-dIdC is another popular nonspecific general competitor. Its only drawback is that it may act as a specific competitor for proteins favoring GC-rich stretches. Another nonspecific competitor can be sheared genomic DNA.
4. Note that the usual unit for DNase I activity is the *Kunitz* unit; the name "Kunitz" should not be mistaken for 1,000 units. The standardized vials give very reproducible results. Glycerol in the buffer will keep the enzyme from freezing.
5. Do not be tempted to use too much carrier tRNA, as an excess of it causes a very annoying fuzziness of the gel bands,

preventing resolution of closely packed individual bands. If attempting to omit the carrier altogether, make sure that the labeled DNA does precipitate by using the Geiger counter. Alternatively, 10 μg glycogen can be used as a carrier.

6. The use of 5′-end labeling with polynucleotide kinase in the presence of crude protein extracts can sometimes lead to a severe loss of signal because of the presence of phosphatases. In these cases 3′-end labeling by "filling out" with Klenow or T_4 DNA polymerase is to be preferred.
7. Ranges of dilutions should be tested empirically with the experimenter's material. Initial experiments should be only concerned with defining the appropriate range of dilutions that give the best looking footprint. In **step 3**, we suggested a series of twofold dilutions; a series of fivefold digestions would also be acceptable (and require fewer lanes on a gel, as well as less material) to identify the best-suited range. Note that naked DNA or DNA–protein complexes using highly purified proteins will always require less DNase than when crude protein fractions are used.
8. The 15-s interval are quite convenient once one gets the hang of the technique. Intervals of 10 s can be considered for the most daring, but they increase the chance of making a mistake and do not really save time. For beginners, intervals of 30 s should be considered (and so four tubes could be processed at a time instead of eight).
9. The chemical G + A ladder has the advantage of using the same labeled DNA fragment as the footprinting reaction; however, it requires the use of toxic reagents and its interpretation can be ambiguous since it does not differentiate between purines. The four ladders generated by the dideoxynucleotide method allow a full read of the region of interest, use mostly innocuous reagents, and have a long shelf life since they are labeled with ^{35}S. For them to be aligned with the footprint ladders, the only thing to pay attention to is that the 5′ end of the sequencing primer be at the same position as the labeled nucleotide of the footprinting probe. For example, let us imagine a DNA fragment cut with *Eco*RI and labeled with T_4 kinase. The labeled fragment (and so the one that is visible on the gel) would look like this: 5′ *AATTC-NNNNNN…3′, where the asterisk represents the radioactive label. To match such a footprinting probe, the sequencing reaction should be performed with a primer starting at the same base: 5′ AATTCNNNNNN 3′.
10. Sequencing gels are not denaturing unless run hot (7 M urea produces only a small reduction in the T_m of DNA). A double-stranded form of the full-length DNA fragment is

therefore often seen on the upper part of the autoradiogram (*see* **Fig. 2**), especially at low levels of DNase I digestion, and can often be misinterpreted as a hypersensitive cleavage site. Comparison to an undigested track of DNA should clarify the point.

References

1. Galas, D.J. and Schmitz, A. (1978). DNAse footprinting: a simple method for the detection of protein–DNA binding specificity. *Nucleic Acids Res.* **5**, 3157–70
2. Rousseau, S., Renaud, J., and Ruiz-Carrillo, A. (1989). Basal expression of the histone H5 gene is controlled by positive and negative cis-acting sequences. *Nucleic Acids Res.* **17**, 7495–511
3. Suck, D., Lahm, A., and Oefner, C. (1988). Structure refined to 2A of a nicked DNA octanucleotide complex with DNase I. *Nature.* **332**, 464–8
4. Drew, H.R. (1984). Structural specificities of five commonly used DNA nucleases. *J. Mol Biol.* **176**, 535–57
5. Maxam, A.M. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560
6. Brown, T. (1987). Analysis of RNA by Northern and slot blot hybridisation. In *Current Protocols in Molecular Biology*, F. Ausubel, R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith and K. Struhl, eds., Green Publishing Associates, New York, NY
7. Sanger, F. and Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**, 441–8

Chapter 4

Exonuclease III Footprinting on Immobilized DNA Templates

Patrizia Spitalny and Michael Thomm

Summary

DNA footprinting is a widely used method to locate the binding sites of protein on the DNA. It is based on the observation that a protein bound to DNA protects it from degradation by an enzyme or chemical reagent.

Exonuclease III is a suitable probe to analyze the boundaries of a protein when it is necessary to eliminate any excess unbound DNA from the reaction to avoid background problems. In combination with biotin-labeled DNA that is bound to streptavidin-coated magnetic particles, information on the precise position of a DNA bound protein is available within a few hours. The position of the archaeal RNA polymerase at different stages of transcription in the *Pyrococcus furiosus* in vitro transcription system was analyzed by this method.

Key words: Exonuclease III, Immobilized DNA, Archaea, In vitro transcription, Stalled transcription complexes.

1. Introduction

To easily determine the position of the RNA polymerase in various registers of transcription within a few hours, we used Exonuclease III (ExoIII) as a footprinting probe in combination with DNA immobilized on magnetic particles. This method can also be used for the detection of the boundaries of any DNA binding protein recognizing a specific sequence.

Apart from having RNase H, 3′-phosphatase and AP-endonuclease activities ExoIII catalyzes the stepwise removal of nucleotides from the 3′-hydroxyl termini of double stranded

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_4

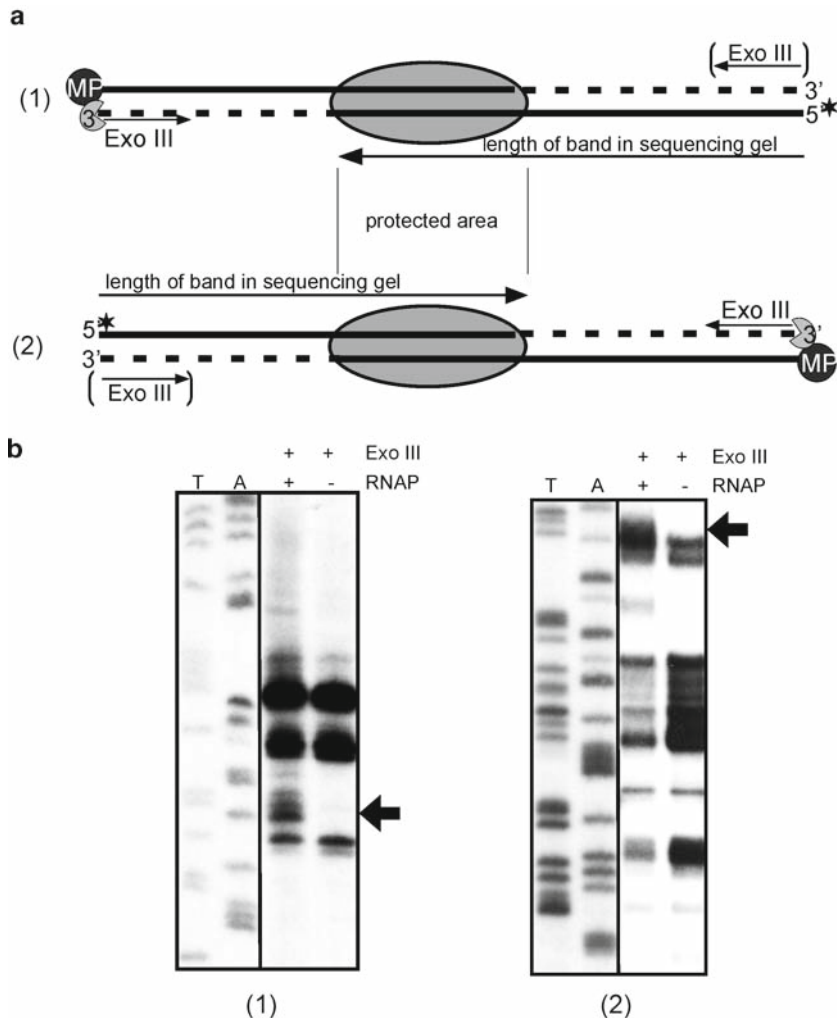


Fig. 1. Exonuclease III footprintig procedure. (**a**) Schematic representation of ExoIII digestion of immobilized end-labeled DNA. To identify both boundaries of the DNA bound protein either the template strand (1) or the RNA-like strand (2) is radioactively labeled, determined by the position of the magnetic particle (MP). ExoIII hydrolyzes the DNA from both 3′ ends to the binding site of the protein (*dotted line*). Only the resulting length of the labeled strand is detected on a sequencing gel as shown in (**b**). The *arrows* indicate the position of the bands representing the upstream (1) or downstream (2) boundary of the protein on the DNA.

DNA *(1)*. A specifically bound protein blocks the action of ExoIII and leaves double stranded DNA only in the region protected by the protein (*see* **Fig. 1**). Any free DNA is fully digested, an important advantage of ExoIII over other footprinting probes. Especially in reactions that do not allow DNA–protein binding to saturation level it is necessary to avoid a high background signal.

The DNA bound to magnetic particles with a biotinylated 5′ terminus allows fast and easy isolation of DNA–protein complexes and is a suitable tool for the precise detection of the band marking one boundary of the protein. Only the free 5′-terminus can be radioactively labeled leaving only one strand detectable on

the sequencing gel for each reaction. To determine the complete binding region for one protein on the DNA the same reaction has to be conducted twice with either the RNA-like strand (**Fig. 1a** (1)) or the template strand (**Fig. 1a** (2)) being attached to the magnetic particle. The ExoIII footprint is seen as band that marks the 3′-boundary of the protein on the labeled strand (**Fig. 1b**).

For the interpretation of the results it is important to know that ExoIII may have an effect on the DNA–protein complex. A steric hindrance caused by the size of both proteins has to be taken into account when the maximum extension of the binding site is estimated. With weak DNA–protein interactions ExoIII digestion activity can run into the region bound by the protein. This may lead to a smaller protection site. To analyze whether such an action occurs a kinetics experiment of ExoIII digestion should be performed.

The activity of ExoIII is partially dependent on the helical structure of DNA *(2)* and it displays a sequence dependence (C > A = T > G) *(3)* that leads to additional pausing sites detectable as bands on the gel. To identify the band marking a protein boundary a control reaction without protein has to be done for every reaction (**Fig. 1b**). Subtracting the control lane from the lane with protein leaves only the signal caused by the bound protein.

ExoIII footprinting has been used to follow the movement of *E. coli* RNAP during transcription *(4, 5)* and also the translocation of pol II transcription complexes has been analyzed by this method *(6)*. In our studies we followed the movement of the archaeal RNA polymerase from *Pyrococcus furiosus* along the DNA during the transition from initiation to elongation of transcription *(7)*. RNA synthesis was paused at distinct positions from +5 to +20 by the use of a C-less cassette until the desired stop position. The paused ternary complexes were isolated by the use of a magnetic particle separator and subjected to ExoIII digestion to detect the boundaries of the archaeal RNA polymerase at each defined position.

2. Materials

2.1. Preparation of Immobilized and End-Labeled DNA Templates

1. Streptavidin magnetic particles (Roche, Basel, Switzerland). Can be stored for many years at 6–8°C.
2. Magnetic particle separator (Caution: keep away from computer discs, etc. because the strong magnetic field can cause damage to magnetic storage material).
3. DNA template from a PCR reaction with a primer 5′end-labeled with biotin.

4. Wash-buffer A: 10 mM Tris–HCl pH 7.5, 1 mM EDTA, 100 mM NaCl.
5. Binding buffer B: 50 mM Tris–HCl pH 7.5, 1 mM EDTA, 100 mM NaCl.
6. Wash-buffer C: 10 mM Tris–HCl pH 7.5, 1 mM EDTA, 1 M NaCl.
7. TE 10: 0.1.
8. [α-^{32}P] ATP (220 TBq/mmol) (Hartmann Analytic, Braunschweig, Germany) (Caution: ^{32}P is a radionuclide. Wear a lab coat and gloves at any time. Use shielding).
9. T_4 Polynucleotide kinase (New England Biolabs, Ipswich, USA).

2.2. Sequencing Gel

1. Sequencing gel apparatus (Sequi-Gen GT System 21 × 50 cm, Bio-Rad, Hercules, USA).
2. Sigmacote (Sigma, St. Louis, USA).
3. 6% acrylamide solution: for 1 L add 420 g urea, p.a. (Merck, Darmstadt, Germany) to 200 mL of Rotiphorese Gel 30, 37.5% acrylamide: 1% bisacrylamide (Roth, Karlsruhe, Germany), (Caution: acrylamide and bisacrylamide are neurotoxins and should be handled with gloves) and 100 mL 10× TBE, adjust to 1 L.
4. 10× TBE: 900 mM Tris, 900 mM boric acid, 20 mM EDTA.
5. 10% (w/v) Ammonium persulfate (Merck, Darmstadt, Germany) (*see* **Note 1**).
6. TEMED 99%, p.a. (Roth, Karlsruhe, Germany).
7. Running buffer: 1× TBE.

2.3. Archaeal In Vitro Transcription and Preparation of Isolated Ternary Transcription Complexes

1. Transcription buffer: 40 mM HEPES, 0.1 mM EDTA, 1 mM DTT, 275 mM KCl, and 3 mM $MgCl_2$ (*see* **Note 2**).
2. 660 ng RNA polymerase purified from *P. furiosus (8)*.
3. 63 ng recombinant TBP and 125 ng recombinant TFB from *P. furiosus* expressed and purified from *E. coli* as described previously *(9)*.
4. End-labeled immobilized DNA fragment.
5. NTPs: 40 μM ATP, 40 μM GTP, and 2 μM UTP.
6. Wash-buffer WB: 40 mM HEPES, 0.1 mM EDTA, 1 mM DTT, 275 mM KCl, and 3 mM $MgCl_2$, 0.5% (w/v) *N*-lauroylsarcosine (NLS) and 40 μM GTP (*see* **Note 3**).

2.4. Exonuclease III Digestion

1. Exonuclease III (New England Biolabs, Ipswich, USA). Can be stored for many months at –20°C.
2. Reaction buffer (Exo RB): 40 mM KCl, 2 mM $MgCl_2$, 100 mM Tris–HCl pH 8.5, and 1 mM DTT.

3. 0.5 M EDTA.
4. Loading buffer: 98% formamide, 10 mM EDTA, and 0.1% each bromphenol blue and xylene cyanol.

3. Methods

3.1. Preparation of Immobilized and End-Labeled DNA Templates

1. The template used for our experiments was cloned into pUC19 and contained a C-less cassette from the transcription start site to the position that was probed. Amplify the DNA template by PCR in a 100 μL reaction using one primer with a 5′ biotin modification. Subsequent PCR purification is not necessary.
2. Wash 50 μL of the magnetic particles with 90 μL of wash-buffer A, discard the supernatant and repeat the washing step tree times.
3. Resuspend the washed pellet in 90 μL of binding buffer B and add 50 μL of the PCR product.
4. Incubate 30 min at room temperature. Make sure by gently shaking that the beads do not settle.
5. Remove supernatant and repeat **steps 3** and **4** adding the other 50 μL of PCR product (*see* **Note 4**).
6. Remove supernatant and resuspend in 90 μL of wash-buffer C. Let stand for 1 min, remove the supernatant and repeat the washing step.
7. Resuspend the pellet with the immobilized DNA template in 90 μL TE 10: 0.1 (*see* **Note 5**).
8. Label 40 μL of the immobilized DNA at the free 5′-end (*see* **Note 6**) by using 4 μL T4 polynucleotide kinase and 4 μL fresh [α-^{32}P] ATP.

3.2. Sequencing Gel

1. The instructions refer to the use of the Bio-Rad Sequi-Gen GT System 21 × 50 cm but can as well be used for any other gel system.
2. Thoroughly wash the front and back glass plate with soap and water. Rinse with deionized water and dry. Wet plates with 96% ethanol and wipe dry with lint-free paper towel. Apply a film of Sigmacote on the front plate. After the film dried wipe with lint-free paper towel.
3. Assemble the gel with 0.4-mm thick spacers and place it flat on the lab bench.
4. Prepare 80 mL of 6% acrylamide solution in a 100-mL beaker, add 20 μL of TEMED and 400 μL of 10% ammonium persulfate and mix thoroughly.

5. Put the gel solution immediately into a syringe, avoid bubbles, and inject it through the injection port of the caster at the bottom of the Sequi-Gen GT cell.
6. When the gel solution reaches the top of the gel plates insert the 16 wells forming comb. Be careful to avoid bubbles.
7. The gel should polymerize for 1–2 h (*see* **Note 7**).
8. Prepare running buffer by diluting 100 mL of 10× TBE with 900 mL of deionized water.
9. Remove the caster. Excess acrylamide should be removed from around the comb with a razor blade before the comb is removed.
10. Assemble the gel unit and pour the running buffer into the upper and lower reservoirs.
11. Connect the gel system to a power supply. The gel can be run at 2,000 V, 50 W, and 25 mA. Maintain a gel temperature of ~50°C (*see* **Note 8**).
12. After electrophoresis transfer the gel to Whatman paper and expose the gel to an X-ray film overnight at –70°C using an intensifying screen. Alternatively use an phosphoimaging plate for 1–2 h to visualize the bands.

3.3. Archaeal In Vitro Transcription and Preparation of Isolated Ternary Transcription Complexes

1. For the in vitro transcription reaction add 3 μL of immobilized end-labeled DNA containing a C-less cassette, 660 ng of purified *P. furiosus* RNA polymerase, 63 ng of purified recombinant *P. furiosus* TBP and 125 ng of purified recombinant *P. furiosus* TFB and NTPs (40 μM ATP, 40 μM GTP, and 2 μM UTP) to transcription buffer in a total volume of 25 μL.
2. Incubate in a thermocycler 3 min at 70°C.
3. Isolate paused transcription complexes at room temperature by the use of the magnetic particle separator. Remove supernatant, wash with 30 μL of wash-buffer WB, and resuspend the pellet in 25 μL of reaction buffer for exonuclease III digestion (Exo RB).

3.4. Exonuclease III Digestion

1. Subject the isolated ternary complexes resuspended in Exo RB to ExoIII digestion by adding 100 units of the enzyme.
2. Incubate 15 min at 37°C.
3. Stop the reaction by adding 1.5 μL of EDTA and 13 μL of loading buffer. Denature at 95°C for 3 min. Under these conditions the DNA is released from the magnetic particles.
4. To avoid partly blocked wells due to the presence of magnetic particles they were pelleted. Twelve microliters of the supernatant were loaded onto the sequencing gel. A sequencing ladder was loaded to identify the length of the bands.

4. Notes

1. Ammonium persulfate is very reactive and hygroscopic by nature. It starts to break down almost immediately when dissolved in water. It is recommended to prepare it freshly every time. Alternatively it can be stored in 1.5 mL aliquots at –20°C for several months. Storage at 6–8°C is possible for a few days.
2. Prepared as stock solution the transcription buffer can be easily adjusted to the volume of proteins in the transcription reaction. At –20°C it can be stored for many months.
3. In our experiments GTP (40 μM) was added to the wash-buffer WB to stabilize paused transcription complexes. All used templates had a GTP as last incorporated nucleotide before the first C position. GTP in the WB prevented the RNA polymerase from backtracking while paused.
4. Take an aliquot of the supernatant in each step before and after the binding reaction. Analyze it on an agarose gel to make sure the binding reaction works to saturation. After the first binding reaction there should be no DNA left in the supernatant. After the second binding step there should still be some DNA detectable on the agarose gel.
5. The immobilized templates can be stored at 6–8°C for many months. Do not freeze them because they will accumulate and settle more quickly so the handling will become difficult.
6. Depending on which strand is attached to the magnetic particle on the 5′-end either the coding or the noncoding strand can selectively be labeled. The specifically labeled DNA can easily be purified from the labeling reaction by the use of the magnetic particle separator. Resuspended in water it was used in the in vitro transcription reaction.
7. The polymerization time should not exceed a few hours to avoid drying in the region around the comb. The sequencing gel can be stored at 6–8°C for up to 2 days. Put a paper towel soaked with 1× TBE over the comb. Wrap the top of the gel with plastic wrap.
8. With the used gel apparatus the upper buffer chamber is attached to the back of the glass plate and acts in cooling the entire gel area during electrophoresis. Alternatively use an aluminum plate for other gel systems. Uniform heating prevents sample mobility artifacts and greatly reduces cracking of plates.

References

1. Rogers, S.G., and Weiss, B. (1980). Exonuclease III of Escherichia coli K-12, an AP endonuclease. *Methods Enzymol.* **65**, 201–211.
2. Richardson, C.C., Lehman, I.R., and Kornberg, A. (1964). A deoxyribonucleic acid phosphatase-exonuclease from Escherichia coli. II. Characterization of the exonuclease activity. *J. Biol. Chem.* **239**, 251–258.
3. Linxweiler, W., and Hörz, W. (1982). Sequence specificity of exonuclease III from E. coli. *Nucleic Acids Res.* **10**, 4845–4859.
4. Straney, D.C., and Crothers, D.M. (1987). A stressed intermediate in the formation of stably initiated RNA chains at the Escherichia coli lac UV5 promoter. *J. Mol. Biol.* **193**, 267–278.
5. Metzger, W., Schickor, P., and Heumann H. (1989). A cinematographic view of Escherichia coli RNA polymerase translocation. *EMBO J.* **8**, 2745–2754.
6. Samkurashvili, I., and Luse, D.S. (1998). Structural changes in the RNA polymerase II transcription complex during transition from initiation to elongation. *Mol. Cell. Biol.* **18**, 5343–5354.
7. Spitalny, P., and Thomm, M. (2003). Analysis of the open region and of DNA–protein contacts of archaeal RNA polymerase transcription complexes during transition from initiation to elongation. *J. Biol. Chem.* **278**, 30497–30505.
8. Hethke, C., Geerling, A.C., Hausner, W., de Vos, W.M., and Thomm, M. (1996). A cell-free transcription system for the hyperthermophilic archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.* **24**, 2369–2376.
9. Hausner, W., Wettach, J., Hethke, C., and Thomm, M.J. (1996). Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. *J. Biol. Chem.* **271**, 30144–30148.

Chapter 5

Hydroxyl Radical Footprinting of Protein–DNA Complexes

Indu Jagannathan and Jeffrey J. Hayes

Summary

This unit details the use of hydroxyl radicals to characterize protein–DNA interactions. This method may be used to assess the exact location of contacts between a protein and its cognate DNA and details of the complex structure. We describe several methods to prepare DNA templates for footprinting and ways to avoid many of the pitfalls associated with the use of hydroxyl radical footprinting. In addition, we describe in detail one example of the application of this technique.

Key words: Hydroxyl radical footprinting, Footprinting, Protein–DNA interactions.

1. Introduction

Footprinting is a powerful technique allowing regions of DNA in contact with sequence-specific DNA binding proteins to be determined. Traditionally, such experiments have been accomplished with the endodeoxyribonuclease, DNase I. Enzymatic reagents such as DNase I digestion can reveal many details about the location, affinity, and extent of protein–DNA interactions and can be used in complicated extracts containing multiple reagents and factors *(1)*. However, finer details of protein–DNA interactions and DNA structure within protein–DNA complexes can be obtained by footprinting with the hydroxyl radical, an efficient chemical nuclease. Because of the small size (about the size of a water molecule) and lack of sequence specificity of this diffusible reactive species, very fine details of DNA conformation and protein–DNA contacts can be resolved. These features are reflected

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_5

in the rates of DNA cleavage at individual nucleotide positions. Cleavage is effected by the abstraction of a hydrogen atom by a hydroxyl radical from a carbon within a deoxyribose in the minor groove of the DNA *(2)*. This leaves a carbon-centered radical, which rapidly disintegrates to leave 5′ and 3′ monophosphoester termini on either side of a one nucleoside gap. Typically, hydroxyl radicals are conveniently produced in solution by chemical means, although radical production by gamma photon photolysis of water has been employed for footprinting *(3, 4)*. Hydroxyl radicals can be efficiently produced in solution in virtually any salt conditions and in the presence of any soluble protein–DNA assembly via the one electron reduction of hydrogen peroxide by an iron(II)EDTA complex to produce ·OH and OH^-:

$$[Fe(II)EDTA]^{2-} + H_2O_2 \rightarrow \bullet OH + OH^- + [Fe(III)EDTA]^{1-}$$

The oxidation of Fe(II) is readily reversible by mild reducing agents such as sodium ascorbate, thus effectively recycling the iron complex until the peroxide is depleted (*see* **ref.** *5* for a complete review of hydroxyl radical mediated DNA cleavage).

Note that the hydroxyl radical cleavage procedure has the drawback that the complexes must be prepared in relatively pure form and buffer solutions must be free of radical scavengers. Common buffer components such as glycerol >0.2%, Tris buffer in excess of 10 mM, and carbohydrates must be avoided. Here we describe the basic methodology for footprinting using chemically generated hydroxyl radicals and detail one example.

2. Materials

2.1. Radioactive End-Labeling of a DNA Fragment

1. Plasmid containing the DNA fragment of interest or oligonucleotides that can be annealed to produce a double-stranded DNA fragment with the cognate site for the protein (see other alternative strategies for generation of the DNA fragment, below).
2. 10× T4 polynucleotide kinase (PNK) buffer (supplied with enzyme).
3. [γ-^{32}P]ATP, 6,000 Ci/mmol.
4. T4 polynucleotide kinase (PNK) 10,000 units/ml (NEB).
5. 2.5 M ammonium acetate.
6. 3 M sodium acetate, pH 5.2.
7. 95% Ethanol, –20°C.
8. 70% Ethanol, –20°C.

9. 10% sodium dodecyl sulfate (SDS) stock solution.
10. Alkaline phosphatase, calf intestinal (Roche Applied Science).
11. TE buffer: 10 mM Tris–HCl pH 8.0 and 1 mM EDTA.

2.2. Maxam–Gilbert G-Specific Reaction

1. 10× G-specific reaction buffer: 0.5 M Na cacodylate and 10 mM EDTA.
2. Dimethylsulfate (DMS) (Sigma).
3. G-reaction stop buffer: 1.5 M Na acetate, 1 M β-mercaptoethanol and 1 μg/μl sonicated calf thymus DNA.
4. Piperidine (neat, 10 M stock) (Sigma-Aldrich).

2.3. Footprinting of Protein–DNA Interactions with Hydroxyl Radicals

1. 5 g $Fe(II)(NH_4)_2SO_4 \cdot (H_2O)_6$, reagent grade, (Sigma-Aldrich) (store in dark).
2. L-Ascorbic acid (sodium salt), reagent grade, (Sigma-Aldrich) (store dry).
3. H_2O_2, 30% solution, store at 4°C.
4. TE (10 mM Tris–HCl, pH 8.0, 1 mM EDTA).
5. Stock solution of 20 mM sodium ascorbate, store frozen at –20°C.
6. Stock solution of 1 mM Fe(II) EDTA; store frozen at –20°C.
7. Solution of 0.15% H_2O_2, freshly made.
8. Stop solution I: 50% glycerol and 10 mM EDTA.
9. Microcentrifuge filtration devices (similar to Series 8,000 originally made by Lida Manufacturing Corporation now part of Thermo-Fisher).
10. 10 mM Tris–HCl pH 8.0 and 0.1% SDS.
11. Microcentrifuge pestles, can be obtained from Stratagene. 95% and 70% ethanol solutions, cooled to –20 °C.
12. 3 M sodium acetate.
13. 0.7% agarose gels made with 0.5× TBE.
14. Pronase or Proteinase K solution; 25 mg/ml stock.

2.4. Sequencing Gel Analysis

1. Solid urea; molecular biology grade.
2. 5× TBE.
3. 40% Acrylamide (19:1 acrylamide:bis-acrylamide).
4. 20% ammonium persulfate (APS).
5. *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED).
6. Formamide loading buffer: 100% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol FF, and 1 mM EDTA.

2.5. 6× SDS Gel Loading Solution

1. 30% glycerol, 0.25% bromophenol blue, 0.25% xylene cyanol FF, 0.4% SDS. Store at 4°C.

2.6. PCR Reaction Setup

1. 10× PCR buffer (no $MgCl_2$) (from polymerase manufacturer).
2. 1.5 mM $MgCl_2$.
3. Plasmid DNA template 0.5 μg.
4. Cold primer 0.5 μg.
5. ^{32}P labeled primer (0.5 μg labeled) 10 μl.
6. dNTPs 2.5 mM each.
7. VENT/Taq polymerase, 1 μl.

Final volume 100 μl

3. Methods

The basic protocol includes the preparation of radiolabeled DNA, some details of complex formation (although the exact details will depend on the particular protein of interest), the cleavage reaction, and analysis of cleavage products by sequencing gel electrophoresis. Our procedure includes an optional nucleoprotein gel electrophoresis for fractionation of the complexes before sequencing gel analysis of the cleavage patterns.

Proper sample preparation is critical for the hydroxyl radical footprinting method. Complexes are prepared with DNA containing a unique radioactive end-label. Alternatively, 5′ end-label can be incorporated after the cleavage reaction. Note that unlike DNase I or other nucleases that require metal-cofactors, chelation agents such as EDTA or ATP do not affect the chemical system used to generate ·OH. Similarly, typical salts, even in high concentrations do not reduce the efficiency of cleavage *(5)*. For radioactively end-labeled complexes, it is recommended that each sample contain at least 5,000 cpm. Samples containing less radioactivity can be used depending on the losses during sample preparation, but at least 2,000 cpm should be loaded in each lane of the sequencing gel. Note that the samples to be treated with hydroxyl radicals prior to resolving nucleoprotein complexes on preparative 0.7% agarose gels should contain >20,000 cpm (see below).

3.1. DNA Fragment Preparation and Labeling

The objective is to prepare a DNA fragment containing a radiolabel at either the 5′ or 3′ end of one of the two strands. We have used at least three methods for this preparation.

3.1.1. Method 1

The easiest and most traditional is to radiolabel the DNA fragment as it is excised from a purified plasmid.

1. Cut ~5 μg of a plasmid containing the ultimate DNA fragment of interest a selected restriction enzyme according to the manufacturer's specifications. Typically, an enzyme that cuts the plasmid only once, about 30–100 bp from the site of interest is used to allow for proper resolution on sequencing gels (see below) (*see* **Notes 1** and **2**).
2. For samples to be treated with PNK to radiolabel the 5′ end of the DNA, treat the plasmid with alkaline phosphatase according to manufacturer's instructions and proceed to **step 3**. For DNA to be filled in with DNA polymerase Klenow fragment (Klenow) and radiolabeled dNTPs, proceed to **step 5**.
3. Remove the phosphatase by phenol/chloroform extraction. Adjust the volume of the reaction to 200 μl with TE containing 0.2% SDS; add 200 μl of water-saturated phenol/chloroform (50:50) and vortex vigorously, then spin for 2 min in the microfuge. Carefully remove and keep the aqueous layer – it is ok to leave a small amount of aqueous layer behind. Add 160 μl of TE back to the organic layer and repeat the extraction, combining the two aqueous layers in a separate tube. Add 40 μl of 3 M NaOAc and 1 ml of ice-cold 95% ethanol, vortex and microfuge at top speed to precipitate the DNA. Carefully remove the supernatant with a pipette, taking care to not disturb the pellet. Add 200 μl of 0.3 M NaOAc, vortex, then repeat the precipitation with 550 μl of EtOH. Carefully wash the final pellet with 500 μl of ice-cold 70% EtOH (do not vortex), spin the tube for 2 min in the microfuge, and slowly remove 95% of the supernatant with a narrow-bore transfer pipette. Dry the tube completely in a Speed Vac then resuspend the DNA in 20 μl of TE for the kinase reaction.
4. Radiolabel the free 5′ ends in the plasmid by treating with polynucleotide kinase and ~5 μl of [γ ^{32}P]-ATP according to manufacture's instructions and according to safe radioactivity practices. Terminate the reaction with 200 μl of 2.5 M NH_4OAc, then precipitate with 700 μl of ice-cold EtOH as described above. Note that the supernatant should be carefully discarded into the liquid ^{32}P waste. Resuspend the pellet in 180 μl TE by gentle vortexing and collect the liquid in the bottom of the tube with a quick microfuge spin. Adjust the solution to 0.3 M sodium acetate by the addition of 20 μl of 3 M sodium acetate stock and reprecipitate the DNA by the addition of 700 μl of –20°C 95% ethanol as described above, again discarding the supernatant into the radioactive (hot) waste. Wash the final pellet and dry in a Speed Vac as described above. Resuspend the pellet in 25 μl of TE and transfer to a new eppendorf tube. Discard the old tube in the hot waste and proceed to **step 6**.

5. Alternatively, radiolabel the free 3′ ends in the plasmid by filling in with Kenow fragment of DNA polymerase I and ~5 µl of the appropriate α [32-P]-dNTP according to manufacture's instructions and according to safe radioactivity practices. Terminate the reaction with 200 µl of 2.5 M NH_4OAc, then precipitate with 550 µl of ice-cold EtOH as described above. Note that the supernatant should be carefully discarded into the liquid ^{32}P waste. Resuspend the pellet in 180 µl of TE by gentle vortexing and collect the liquid in the bottom of the tube with a quick microfuge spin. Adjust the solution to 0.3 M sodium acetate by the addition of 20 µl of 3 M sodium acetate stock and reprecipitate the DNA by the addition of 550 µl of –20°C 95% ethanol as described above, again discarding the supernatant into the hot waste. Wash the final pellet and dry in a Speed Vac as described above. Resuspend the pellet in 25 µl of TE and transfer to a new eppendorf tube. Discard the old tube in the hot waste and proceed to **step 6**.
6. Digest the sample with a second restriction enzyme, typically a single-cutter, to liberate the DNA fragment of interest according to manufacture's instructions such that the resulting fragment is 100–350 bp in size. Multiple cutters can be used for the second restriction enzyme cut but the DNA fragment of interest must be resolved from the second radiolabeled fragment (in the case of a single-cutter first enzyme) on the preparative acrylamide gel. Note that DNA fragments larger than about 400 bp are inefficiently recovered from acrylamide gels **(step 7)**.
7. After the digestion (1 h) add 5 µl of SDS-gel loading solution and apply the sample to a preparative 5% polyacrylamide gel with 1/2× TBE both in the gel and as a running buffer. Run the gel for 1.5 h (actual time may vary depending on size of fragments) to separate the radiolabeled fragment from the other longer (or shorter) radiolabeled fragment(s). Note that if you used two single-cutter enzymes then the unwanted fragment will be large and contain the remainder of the original plasmid (*see* **Note 3**).
8. Expose the wet gel to autoradiographic film for 30 s to 2 min and, using the film as a guide, cut out the appropriate band with a razorblade. Recover the DNA from the gel slice by passive elution (crush and soak) or by electroelution (*see* **Note 4**).
9. Precipitate the eluted DNA. Carrier DNA (~0.1 µg) can be added to improve recovery. Resuspend the DNA in TE buffer at no more than ~10 kcpm/µl. Store at 4 °C.

3.1.2. Method 2

This method involves using PCR to generate the fragment of interest. Note that radiolabeling of primers typically results in less specific activity incorporated into the nucleic acid compared to

Subheading 3.1.1. Nevertheless, this method does not require appropriately positioned restriction sites and does not require large amounts of the parent fragment.

1. Design two primers to PCR the DNA fragment from an available template containing the DNA binding site of interest. See above for criteria for selection of the fragment size and position of the binding site.
2. Radiolabel one of the two primers. Typically, ~0.5 μg of primer is radiolabeled via the procedure mentioned above. However, employ ~2× more γ[^{32}-P]ATP, precipitate the primer as described, rinse and dry. Resuspend in 10 μl of TE.
3. Carry out a PCR reaction containing the radiolabeled primer and the cold primer.
4. Isolate the radiolabeled fragment as described above.

3.1.3. Method 3

A similar method may be used to generate an appropriate binding target for the protein of interest directly from synthesized oligonucleotides. Two complementary oligonucleotides are prepared, one radiolabeled as above, then annealed together. The annealed product is then isolated from a preparative nondenaturing gel. While the simplest approach, this method is limited due to the size of the oligonucleotides that can be annealed and may be subject to background bands in the final analysis due to the less than perfect chemical synthesis used to generate the oligonucleotides.

1. Oligonucleotides. The length of the oligonucleotides depends on the size of the binding site but should be >40 nt each. This allows for efficient recovery and resolution of shorter cleavage products on the sequencing gel as oligonucleotides <15 nt tend not to precipitate efficiently in ethanol.
2. Radiolabel the 5′ end of one oligonucleotide as per above procedure.
3. Anneal the radiolabeled oligonucleotide with a 1.5-fold excess of the unlabeled complementary oligonucleotide. Mix both together in TE, incubate at 90°C for 10 min, turn off the heating block then allow to slowly cool to room temperature.
4. Isolate the radiolabeled annealed product from a preparative acrylamide gel (*see* **Note 3** regarding background in the oligo-only method).
5. For constructing longer DNA substrates, two primers can be ligated in the presence of a splint for correct alignment. The ligated primers can be used for further annealing thus forming the necessary DNA substrate *(6)*.

3.2. Hydroxyl Radical Footprinting Reaction

Typically, footprinting reactions involve purified proteins mixed with a radiolabeled DNA fragment containing the cognate site

for the protein in a buffer compatible with binding activity and with the chemical generation and cleavage activity of hydroxyl radicals. In practice this means that often common reagents such as glycerol that are good scavengers of hydroxyl radicals must be omitted or minimized in the buffers. However, as mentioned above, the cleavage reaction is compatible with a wide range of salts.

1. Prepare the protein–DNA complex in a solution containing ≤0.2% glycerol and ≤10 mM Tris–HCl. Most salts and EDTA may be present at any concentration *(5)*. Any volume may be used; typically the protein–DNA binding reaction is prepared at seven-tenths of the final reaction volume with no correction for dilution that occurs upon addition of the cleavage reagents. However, if there is concern, additional binding buffer can be added by making the sodium ascorbate 3× in binding buffer (see below).
2. Prepare ~500 µl working stocks of 1 mM Fe(II)/2 mM EDTA, 10 mM sodium ascorbate, and 0.12% H_2O_2 (1:250 dilution of the 30% stock) from the concentrated stocks immediately before the cleavage reaction (*see* **Note 5**).
3. The cutting reaction is initiated by placing a drop each of the iron/EDTA solution (one-tenth volume), 10 mM sodium ascorbate (one-tenth volume) and the 0.12% H_2O_2 (one-tenth volume) on the inner wall of a 1.5-ml Eppendorf tube that contains the nucleoprotein binding reaction (seven-tenth volume). This allows the three reagents to be mixed very briefly and then immediately after the addition of the peroxide solution, the reagents are pushed down to the solution containing the nucleoprotein complex using a pipette tip. The whole reaction is then mixed thoroughly by pipetting. For example, for digestion of a 35 µl solution of protein–DNA complexes, 5 µl of each of the reagents would be mixed together on the side of the tube and then rapidly mixed with the sample.
4. The free radical reactions are incubated at room temperature for 30 s to 2 min and then the cleavage reaction is quenched. The method used to terminate the reaction depends on the subsequent steps to be followed. If complexes are to be fractionated by preparative nucleoprotein gel electrophoresis, stop the reaction with the addition of glycerol to 5% and the entire reaction volume transferred to the gel as described in **step 6**. If the cleaved DNA is to be prepared for sequencing gel electrophoresis, the reaction is adjusted to 0.3 M sodium acetate and precipitated with ice-cold ethanol as above. Note that the presence of ethanol effectively quenches the radical cleavage reaction. The reaction may also be quenched by adjusting the solution to 50 mM thiourea *(7)*. Note that the extent of cleavage of the DNA can be adjusted by changing either

the incubation time for cleavage, or reducing the Fe[EDTA] concentration or adjusting the H_2O_2 concentration. Under optimal conditions approximately two-third to three-fourth or more of the original DNA fragment should remain *uncut* after digestion (*see* **Note 6**).

5. Purify the DNA for sequencing gel electrophoresis. Complexes containing protein concentrations in excess of 0.1 mg/ml should be digested with proteinase K or Pronase ~0.5 mg/ml final concentration for 2–3 h at 37–42°C in the presence of 0.2% SDS, by the addition of one-tenth volume of a 10× stock solution. The sample volume is then adjusted to 180 µl with TE, 0.1% SDS, and extracted with one volume of buffer-equilibrated phenol. The aqueous phase is then adjusted to 0.3 M sodium acetate by the addition of 20 µl of 3 M sodium acetate stock and the DNA precipitated by the addition of 550 µl of –20°C 95% ethanol. The DNA pellet is collected by centrifugation at the ~14,000 rpm in a benchtop microfuge for 20–30 min at room temperature and the supernatant carefully removed with a transfer micropipette. The DNA is then resuspended in 0.3 M sodium acetate/TE and reprecipitated with ethanol to remove additional salt and SDS, The final pellet gently washed with 400 µl of ice-cold 70% ethanol. However, typical preparations of reconstituted protein–DNA complexes with pure proteins contain protein concentrations that do not require protease treatment and phenol extractions. These may be precipitated directly with ethanol/salt in the presence of SDS as described. The DNA pellet is then dried and analyzed as in **step 9**.
6. If digested complexes are to be fractionated by nucleoprotein gel electrophoresis, the reaction is stopped by addition of one-tenth volume 25 mM EDTA/50% glycerol. The sample is then immediately loaded onto a 0.7% preparative agarose gel (45 mM Tris borate, pH 8.3) with an applied voltage of 5 V/cm and the samples electrophoresed for 3 h. Typically, all visual tracking dyes such as bromophenol blue or xylene cyanol are omitted from complex samples but are loaded in adjacent wells to gage the progress of the gel. The bromophenol blue dye is typically run to about one-half to two-third of the way through a standard 10–12 cm horizontal agarose gel.
7. The gel is wrapped tightly in Saran wrap and exposed to autoradiographic film at 4°C. We find that the gel must be exposed for 1.5–3 h before sufficiently dark bands are obtained (overnight exposure might be required in some cases). It is crucial at this point to mark the film to be able to precisely align the gel with the exposure after developing. We typically use phosphorescent markers available from Stratagene which are placed directly on the surface of the plastic wrap. Alternatively,

one can simply cut one corner from the gel, wrap it in Saran wrap, then carefully trace around the edges of the gel with a marker right on the film. Care must be taken so that the film does not shift position as the gel is prepared for exposure (the gel can be fastened at the corners with scotch tape). The gel may also be aligned by preparation of radioactive "marker dye" Some older radioactive nucleotides are mixed with concentrated bromophenol blue dye and spotted into the gel with a micropipette. Although the dye diffuses somewhat, the original point of placement usually can be visualized on the gel and the exposure. It is also possible to prepare a "radioactive marker pen" by dipping the end of a marker in some old source radioactivity and then writing directly onto the Saran wrap.

8. Cut out the digested protein–DNA complex and naked DNA bands and isolate the radiolabeled DNA. This can be accomplished by electroelution of the isolated band or by standard "freeze and squeeze" procedures using a small filtration unit that fits into the top of a microfuge tube (pore size 0.4 μm). If the latter is used, we find it improves recovery to reextract the plug of agarose in the filter with 0.5 ml of TE, 0.2% SDS solution either overnight at room temperature or with constant gentle mixing on a rotator for 2 h at 37°C. Radiolabeled digested DNAs are recovered by adjusting the salt concentration to 0.3 M sodium acetate and precipitation with –20°C ethanol via standard protocols.
9. The DNA is analyzed by sequencing gel electrophoresis. The digested DNA is typically resuspended in 10–20 μl of TE, depending on how many radioactive counts are used in the experiment. This allows the radioactivity to be more quantitatively resuspended by vortexing. Note that typically, much of the precipitated, dried DNA is stuck to the inner wall of the microfuge tube and care must be taken to resuspend all the radioactive material. This can be done by vigorous vortexing and/or by rubbing the droplet of TE against the inner wall of the tube. Ideally, approximately one-fourth of this solution (~5 μl) is transferred to a fresh microfuge tube and dried to completion in a Speed Vac. This sample is then easily resuspended in ~3 μl of 100% formamide containing 1 mM EDTA and 0.005% tracking dyes, denatured at 90°C for 2–5 min and loaded into the well of a sequencing gel. The gel is run and exposed to autoradiographic film or a phosphoimager plate. The extent of cleavage in the hydroxyl radical treated samples and the level of background in the control samples are evaluated (*see* **Notes 6–9**). The cleavage patterns are then analyzed by phosphorimage and densitometric scanning.

3.3. Example of Hydroxyl Radical Footprinting

The Msh2/3 complex binds DNA containing mismatched bases and other unusual secondary structures *(8)*. To determine the extent of interaction of purified Msh2/3 with a stem-loop structure, two oligonucleotides were annealed to generate the CA4 loop DNA All oligonucleotides were obtained PAGE purified (from IDT, Coralville, IA). Approximately 1 μg of either top strand or bottom strand oligonucleotides were radioactively end-labeled with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP as described above. The oligonucleotides were then precipitated and isolated on 10% polyacrylamide gels (1× TBE), the band carefully excised to eliminate oligonucleotides differing by >1 nucleotide in length. The labeled DNA was passively eluted into 600 μl extraction solution (0.4% SDS in TE), and fragments of gel removed by filtration with 0.45-μm filters

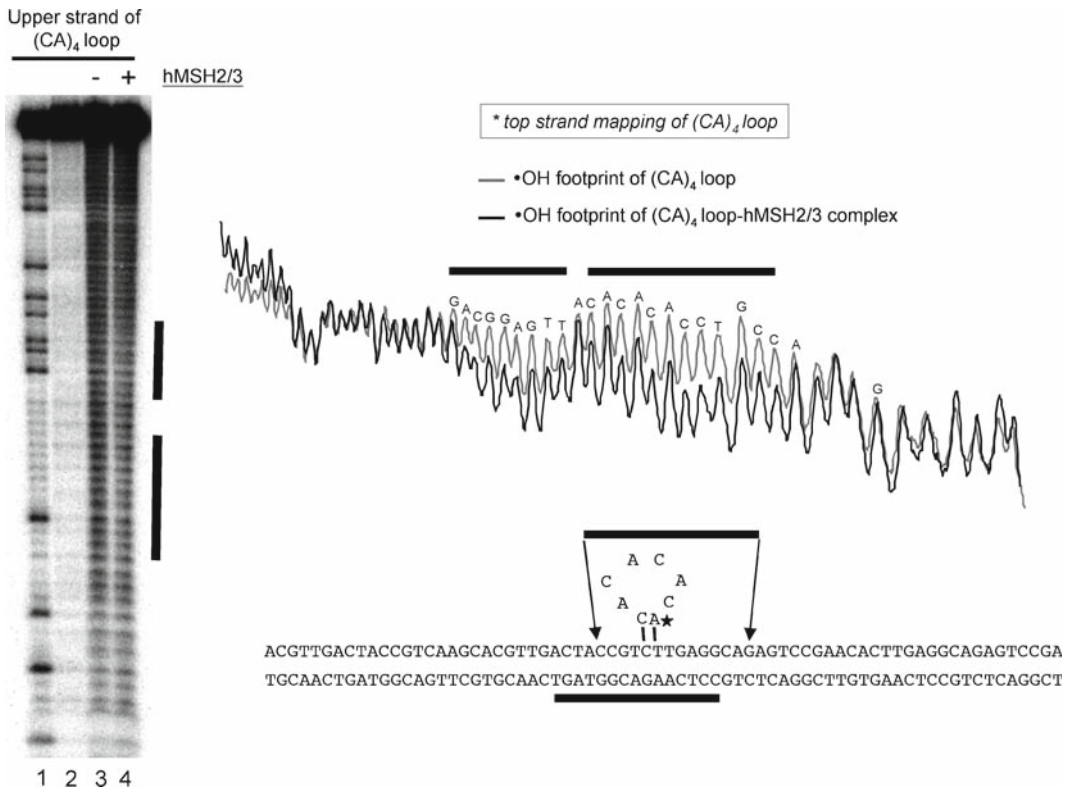


Fig. 1. Footprinting of the Msh2/3 complex bound to a CA4 loop. The hMsh2/3 complex was prepared and bound to annealed oligonucleotides as described in the text and in **ref.** *8*. The gel shows a G-reaction marker (*lane 1*), untreated control DNA (*lane 2*), and DNA digested in the absence or presence of the hMsh2/3 complex (*lanes 3* and *4*, respectively). *Black bars* indicate regions in which cleavage is somewhat diminished in the presence of the protein. In this experiment the *top strand* contained a radiolabel at the 5′ end of the strand. The *right half* of the figure shows the scans of *lanes 3* and *4* (*gray* and *black lines*, respectively) and, below, a scheme of the location of protections for this experiment and an analogous experiment in which the radiolabel was located on the *bottom strand*. The *star* corresponds to the single A residue for which cleavage is not diminished in presence of the protein in the center of the binding site (see gap between *black lines*).

(SPIN-X, COSTAR). Oligonucleotides were precipitated then finally resuspended in 60 µl of TE buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA). Labeled top or bottom strands were annealed with 2.5 µg of either cold bottom or top strands, respectively, by incubation at 83°C for 5 min in TE buffer, then gradual cooling to 25°C. The annealed samples were PAGE purified on 10% polyacrylamide/1× TBE gels as described above. The final annealed structure is shown in **Fig. 1**, right.

Approximately 2 ng of the annealed double-stranded oligonucleotides was incubated without or with ~100 ng hMSH2/3 protein for 30 min at RT in binding buffer in a total volume of 20 µl, then treated with 2 µl of each of the hydroxyl radical reagents for 2 min at room temperature, and the reaction quenched by the addition of 3 µl of 50% glycerol. Digested DNA was precipitated and total radioactivity recovered was determined by Cerenkov counting (Beckman, LS600SC). About 7,000 cpm from each digest was loaded on 10% polyacrylamide sequencing gels and electrophoresed for 1 h 5 min at 60 W, then the dried gels analyzed by Phosphoimager using Imagequant software (Molecular Dynamics). The patterns were further compared by densitometric scanning of the phosphorimager files using the program Imagequant. Although the protection by the protein is not readily apparent in the gel image, the scans reveal a clear footprint due to the Msh2/3 complex.

4. Notes

1. A restriction enzyme that cuts the plasmid more than once can be used here but after cleavage with a second enzyme to liberate the fragment (**Subheading 3.1.1**, **step 6)** one must be able to resolve the fragment of interest from other radiolabeled products; also when multiple ends are available for labeling, the efficiency of radiolabel incorporation into the fragment of interest can be decreased. Thus, carefully planned strategy must be employed.
2. It is important to choose a restriction enzyme that generates a 5′ overhang after cleavage so that either the kinase or fill-in reactions can proceed efficiently.
3. When isolating a radiolabeled DNA fragment for the first time, it is a good idea to apply size markers to the gel in a lane that can be cut away and stained with ethidium bromide.
4. The radiolabeling should produce a DNA fragment with 200,000 to 1 million cpm in total. This will be evident by a dark band after a maximum of 2 min exposure. If a dark band

is not observed, stain the preparative gel to determine if the DNA fragment was recovered efficiently from the precipitation.

5. Efficiency of radical production is greatly reduced in acidic solutions. Thus, it is important to employ sodium ascorbate. Prepare sodium ascorbate solutions from dry crystals that do not appear to have become moist or reddish in color. We have found that 1 M solutions of this reagent are stable for >1 year when stored at –70°C and the 10 mM working stocks are stable for >3 months when stored frozen. Likewise the working stocks of the Fe(II)EDTA complex are stable for extended periods if kept frozen. Note that on formation of insoluble μ-hydroxy iron complexes, the stock should be discarded. Although solutions of 30% H_2O_2 are stable for several years if stored properly at 4°C, contamination of this solution with transition metals can cause a loss of peroxide. In addition, the diluted working stock solution of hydrogen peroxide is relatively unstable on the bench top and should be discarded immediately after the reaction.
6. *Evaluation of the extent of cleavage*. Since the hydroxyl radical yields an even pattern of cleavage at every nucleotide position, the intensity of individual bands on the gel is somewhat less than that obtained with DNase I. Thus, it is sometimes difficult to evaluate the extent of cleavage. A common problem is the lack of apparent cleavage on the gel and due to underexposure of the gel. Thus it is essential that a control sample of undigested DNA of equivalent cpm to the experimental samples be loaded onto the gel. The gel is exposed to the point that the "background" within the control lane is barely visible and compared to the signal in the experimental lane. The bands in the latter should be ≥10× darker than those in the control lanes. If not, see below. In addition, the full-length bands from hydroxyl radical treated samples should be about 75–80% the intensity of the untreated control. This may be evaluated by a 10–15 min exposure of the top of the gel.
7. *Determination of background*. The hydroxyl radical cleavage pattern within DNAs containing a high amount of inherent background cleavage may be difficult to see. The amount of background cleavage may be judged by running a mock-treated sample as described above and direct evaluation of the autoradiograph. Alternatively, after a brief exposure, the full-length band may be cut out of the sequencing gel and the total counts within this band determined by scintillation. The total counts within this band should be ≥95% of the counts loaded in the lane.
8. *Background remediation*. (a) High background in native DNA fragments or DNA produced by PCR may be due to several

problems. Note that labeled DNA should be stored in the presence of EDTA if possible and at <~10,000 cpm/μl to avoid autoradiolysis. In addition, DNA subjected to treatment with Fe(II)EDTA should never be stored as a precipitate immediately after the digestion; the DNA should be processed at least through the rinse step. Another source of background comes from the use of old formamide or excessive heating before loading the sequencing gel. DNA should not be heated in formamide at temperatures greater than 95°C and the formamide should be stored frozen at –20°C between uses. These problems are manifested as elevated background cleavage at purine residues within the DNA. (b) Background in DNA fragments generated by PCR (**Subheading 3.1.2**) can be due to excessive PCR cycling that can result in depletion of dNTPs in the final cycles and incomplete extension of one of the strands. This will show up on the sequencing gels as a background near the full-length product. If this occurs, increase dNTP concentrations to 10 mM and/or reduce the number of cycles used. (c) Background in the oligo-only method (**Subheading 3.1.3**) can come from remnants of the chemical synthesis method such as incomplete deprotection or modified bases. In the case of background using this method, the individual oligonucleotides can be incubated for 2 min in hot (95°C) formamide then isolated from preparative denaturing acrylamide gels. The location of the full-length band on these gels can be determined by radiolabel or UV shadowing on a TLC plate. (d) *Buffer components.* A lack of cleavage may be due to the presence of radioprotectants within solutions. Buffer components such as glycerol in excess of ~0.2% and Tris–HCl in excess of 10 mM should be avoided.

9. *Band smearing.* "Smeary" bands may be due to several problems. First, a common mistake is to use too much carrier in the preparation of samples. For example, an excess of tRNA added as carrier will cause a smear in the region 120–150 bp on the sequencing gel. Most samples will have sufficient DNA for quantitative precipitation without addition of carrier. Another source of "fuzziness" is loading too much of the final sample into the gel pocket. The sample typically contains relatively insoluble contaminants from the agarose gel. If these are loaded into the well, the heated sample has a very viscous character and yields a poorly resolved lane. If necessary, remove 20 μl of the 25 total resuspension in TE after a brief microfuge spin of this material to sediment insoluble contaminants.

Acknowledgments

This work supported by NIH grant 52426. We are grateful to Zungyoon Yang who performed the footprinting experiment.

References

1. Tullius, T. D. (1989). Physical studies of protein-DNA complexes by footprinting. *Ann. Rev. Biophys. Biophys. Chem.* **18**, 213–37.
2. Pogozelski W. K., and Tullius T. D. (1998). Oxidative strand scission of nucleic acids: routes initiated by hydrogen abstraction from the sugar moiety. *Chem. Rev.* **98**, 1089–108.
3. Hayes J., Tullius T. D., and Wolffe A. P. (1989). A protein–protein interaction is essential for stable complex formation on a 5 S RNA gene. *J. Biol. Chem.* **264**, 6009–12.
4. Sclavi B., Woodson S., Sullivan M., Chance M. R., and Brenowitz M. (1997). Time-resolved synchrotron X-ray "footprinting", a new approach to the study of nucleic acid structure and function: application to protein–DNA interactions and RNA folding. *J. Mol. Biol.* **266**, 144–59.
5. Tullius T. D., Dombroski B. A., Churchill M. E., and Kam L. (1987). Hydroxyl radical footprinting: a high-resolution method for mapping protein–DNA contacts. *Methods Enzymol.* **155**, 537–58.
6. Chafin D. R., Vitolo J. M., Henricksen L. A., Bambara R. A., and Hayes J. J. (2000). Human DNA ligase I efficiently seals nicks in nucleosomes. *EMBO J.* **19**, 5492–501.
7. Tullius T. D., and Dombroski B. A. (1985). Iron(II) EDTA used to measure the helical twist along any DNA molecule. *Science* **230**, 679–81.
8. Owen B. A., Yang Z., Lai M., et al. (2005). (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat. Struct. Mol. Biol.* **12**, 663–70.

Chapter 6

The Use of Diethyl Pyrocarbonate and Potassium Permanganate as Probes for Strand Separation and Structural Distortions in DNA

Brenda F. Kahl and Marvin R. Paule

Summary

Diethyl pyrocarbonate (DEPC) and potassium permanganate are useful reagents for detecting DNA distortions, especially melted regions. Unlike most other footprinting methods, these reagents can detect such distortions even within the regions of protein–DNA complexes normally protected in other footprinting techniques. Further, reactions are very robust, so that distorted regions can be detected even under conditions where efficiency of DNA–protein complex formation is not high. DEPC reacts with bases that are fully or partially unstacked in DNA, in the preferential order adenosine > guanine >> cytosine. Permanganate reacts strongly with thymine in unstacked regions of DNA, and exhibits only very weak reaction with guanine, cytosine, or adenine. The combination of both reagents gives excellent coverage of all sequence regions of DNA. Because reaction requires unstacking, the two reagents detect both melted regions and regions that are unstacked because of other distortions such as bending. Permanganate has the additional advantage that it can be utilized in living cells.

Key words: DNA, Footprinting, Melted DNA, DNA bending, Dithylpyrocarbonate, Permanganate.

1. Introduction

In the search for methods to explore the interaction between proteins and DNA, a plethora of footprinting techniques have been developed and many of which are discussed elsewhere in the present work. Most footprinting techniques are based on the simple premise of specific DNA regions being protected from the reagent by the bound protein or molecule of interest. However,

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_6

a number of studies over the past decade have revealed remarkable distortion of the DNA molecule, including bending and strand separation, in response to the bound protein. These distortions are often within the classical footprint, but are rarely detected by these classical techniques. Thus, their detection requires alternative approaches. Unlike enzymatic methods, the chemical probes diethyl pyrocarbonate (DEPC) and potassium permanganate can access and react with the entire sequence of the DNA, distinguishing DNA distortions "under the foot" of a typical footprinting experiment. Combining the information obtained from these chemical probing techniques with the spatial information afforded by traditional footprinting gives an in-depth account of the various ways proteins and other molecules interact with DNA.

DEPC and potassium permanganate are useful probes because of their preferential reactivity with single-stranded over double-stranded DNA. In addition, unlike typical footprinting techniques where near saturation of the DNA with protein is necessary to observe a clean footprint, potassium permanganate and DEPC probing does not require most of the DNA to be in complex with the protein or molecule of interest. Due to the sensitivity of these chemical probes, it is possible to observe reactive bases even when the population of single-stranded DNA is very small.

The mechanism by which DEPC modifies bases has been investigated, but is still not clearly understood *(1–3)*. DEPC predominantly reacts with purine residues, but may react weakly with cytosine residues as well. DEPC modifies DNA by an out-of-plane attack on several of the nucleophilic centers in purines, leading to the scission of the glycosidic bond. Double stranded B-form DNA does not undergo modification by DEPC because the close stacking of neighboring bases occludes access to the out-of-plane surfaces. However, under conditions where the conformation deviates from B-form (such as strand separation or bending), purines in the sequence become more accessible to modification by DEPC. Carbethoxylation of the imidizole ring N-7 produces strand scission under alkaline conditions. Thus, DEPC is commonly used to detect purines which are present in melted or distorted DNA sequences. While DEPC can react with both purines, it shows a marked preference for adenines over guanines in most instances.

Potassium permanganate reacts with double bonds, oxidizing them to vicinal diols. In nucleic acids, the base thymine is oxidized most vigorously, while reaction with C, G, and A is minimal. The mechanism behind this preferential reactivity is believed to arise from an out-of-plane attack on the 5,6-double bond of the thymine ring *(4–9)*. Although the ring is still intact, the loss of aromaticity resulting from insertion of hydroxyl groups on the 5- and 6- carbons leads to a reduction in hypochromicity.

Treatment of the vicinal diol with strong base leads to ring opening and cleavage of the phosphodiester backbone. As with DEPC, stereochemical hindrance from base stacking prohibits reactivity of double stranded B-form DNA. In DNA, which is denatured or is altered from B-form, the thymine ring becomes susceptible to modification by potassium permanganate. Because of their base preferences, the combined use of DEPC and potassium permanganate allows complete analysis of both GC- and AT-containing sequences in DNA.

DEPC and potassium permanganate modification have been used to detect a number of distorted DNA structures in both prokaryotic and eukaryotic cells, including open complex formation during transcription *(10–14)*, steps in promoter clearance *(15–17)*, elongation *(17–19)*, and termination *(20, 21)*, RNA–DNA hybrid structures in transcription elongation complexes *(22, 23)*, drug binding to DNA *(24–26)*, chromatin positioning *(27, 28)*, recombination events *(29–31)*, and single-stranded binding protein binding domains *(32–34)*. On DNA alone, these reagents can reveal sequence-dependent distortions *(35–37)*, negatively supercoiled DNA *(38–40)*, cruciform DNA structures *(41, 42)*, DNA hairpins such as those found in triplet expansion diseases like fragile X-syndrome *(43, 44)*, and Z-DNA, H-DNA, or triplex DNA *(45–48)*. **Figures 1** and **2** give examples of the

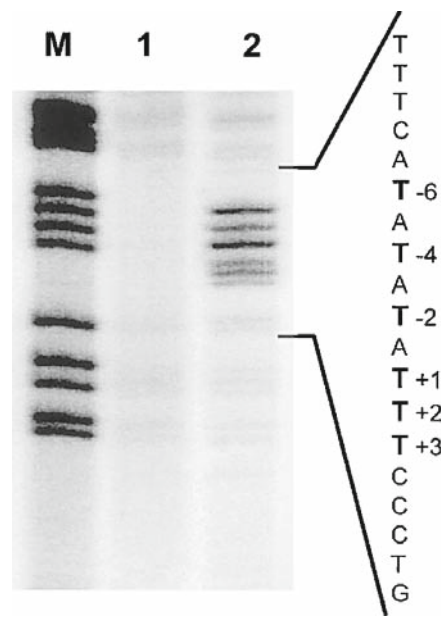


Fig. 1. Potassium permanganate sensitive sites when an open promoter complex is formed on the coding strand by RNA polymerase I from *Acanthamoeba castellanii*. *Lane M* contains a G + A Maxam–Gilbert sequencing ladder, *lane 1* contains DNA exposed to $KMnO_4$ treatment in the absence of any proteins, and *lane 3* contains DNA with the addition of proteins necessary for melting of DNA. *Numerical designations* refer to the transcription start site. Hypersensitive sites in *bold* denote regions of strand separation.

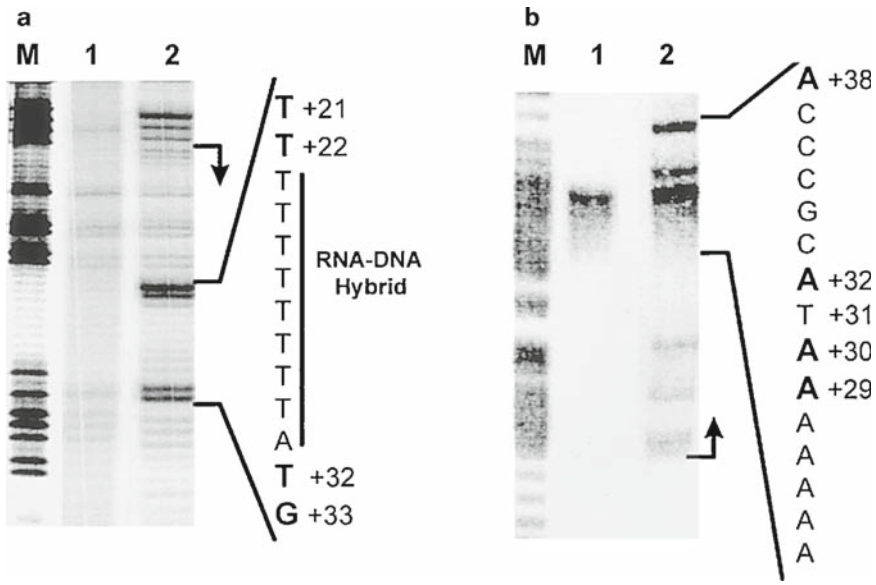


Fig. 2. Potassium permanganate and DEPC probing of stalled transcription complex. (**a**) Potassium permanganate probing of the coding strand when RNA polymerase I is paused at +31 from the transcription start site. *Lane M* contains G + A Maxam–Gilbert sequencing ladder, *lane 1* contains DNA alone with $KMnO_4$ treatment, and *lane 2* displays the permanganate sensitive sites when RNA polymerase from *Acanthamoeba castellanii* is paused at +31 from the transcription start site. The hypersensitive sites in *bold* at positions +21, +22 and +32, +33 define the transcription bubble on the coding strand. Hyposensitive thymidines in the transcription bubble depict protection due to the formation of an RNA–DNA hybrid. (**b**) DEPC probing on the noncoding strand when RNA polymerase I is paused at +31 from the transcription start site. *Lanes* are the same as in (**a**). DEPC sensitive sites in *bold* define the leading edge of the transcription bubble on the noncoding strand. Unreactive *A*s in the region may be due to protein interference or the slightly less reactive nature of DEPC compared to $KMnO_4$.

data which can be obtained. **Figure 1** shows one of the more common uses of potassium permanganate, the analysis of open promoter complex formation. **Figure 2** shows results for both DEPC and potassium permanganate acting on a stalled transcription elongation complex, showing both the melted transcription bubble and the unreactive RNA–DNA hybrid.

Virtually any sequence that deviates from B-form DNA is a candidate for probing with DEPC and potassium permanganate. Since DEPC and potassium permanganate only modify the susceptible bases without cleaving the phosphodiester backbone, further steps need to be taken to visualize the positions of the modified bases. There are three different methods commonly used: The first, which can only be used for experiments performed in vitro, utilizes 5′ or 3′ end-labeled DNA fragments. After treatment with DEPC or potassium permanganate, the DNA is treated with piperidine to cleave the phosphate backbone on the 3′ side of the modified nucleotide. This procedure is useful when comparing the results from multiple footprinting

techniques, since the same fragment of labeled DNA may be utilized with each type of experiment. This method also works well when examining short tracts of DNA or DNA which does not amplify well in a thermocycler. The other two methods, primer extension and thermocycle amplification, are used when working in vivo (potassium permanganate only) or with circular pieces of DNA in vitro. Thermocycle amplification is particularly beneficial when working with limited amounts of DNA (in the low nanogram range) or when the ratio of protein–DNA complex to DNA is low. **Subheadings 2** and **3** are general guidelines for probing with DEPC and potassium permanganate. Care and attention to detail is necessary for all of these methods as footprinting with these reagents is generally somewhat more difficult than with many other footprinting reagents. There is usually a need to optimize conditions for each particular application, and guidelines for this can be found in **Subheading 4**.

2. Materials

1. Potassium permanganate, DEPC, piperidine, and 2-mercaptoethanol can be purchased from Sigma-Aldrich. DEPC, piperidine, and 2-mercaptoethanol should be stored at 4°C and used with caution in a fume hood. Other reagents listed below should be of the highest quality. Low adhesion microcentrifuge tubes (siliconized) can be obtained from USA Scientific (Ocala, FL).
2. The following stock solutions can be made, filter sterilized, divided into aliquots, and stored at –20°C until ready for use:
   - 1 M HEPES pH 7.9.
   - 1 M $MgCl_2$.
   - 0.5 M dithiothreitol (DTT).
   - 2 M KCl.
   - 10 mg/mL Bovine Serum Albumin (BSA), DNase free.
   - 0.5 M EDTA pH 8.0.
   - 3 M Sodium acetate pH 5.2.
   - 5 mg/mL linear polyacrylamide (acrylamide polymerized without *N,N′*-methylenebisacrylamide).
   - 10 mg/mL Proteinase K.
   - 10% SDS.
   - DEPC Stop Buffer: 0.2% SDS, 0.6 M sodium acetate pH 5.2.

   - TE: 10 mM Tris–HCl pH 8.0, 1 mM EDTA.
   - Electrophoresis loading buffer: 0.1% bromophenol blue, 0.1% xylene cyanol, 10 mM EDTA, 80% deionized formamide.
3. A 300 mM stock solution of $KMnO_4$ can be made by heating 2.73 g of $KMnO_4$ in 50 mL of deionized water. The solution can be stored at room temperature in a brown bottle for 1 month.
4. 10× reaction buffer, for example: 200 mM HEPES pH 7.9, 100 mM $MgCl_2$, 1 mM DTT, 1 mg/mL BSA, 100 mM KCl. The choice of this buffer will depend upon the conditions needed to form the complex under study.
5. Protein dilution buffer, for example: 20 mM HEPES pH 7.9, 1 mM EDTA, 10% glycerol. Choice will dependent on protein used.
6. 1 M piperidine: 10% v/v in water, freshly made.
7. 0.6 M sodium acetate.
8. 0.6 M sodium acetate, 20 mM EDTA.
9. 10× neutralization buffer: 0.5 M HEPES pH 7.9, 0.1 M $MgSO_4$, 2 mM DTT.
10. dNTP mix: 5 mM of each dNTP.

3. Methods

Experiments using DEPC or potassium permanganate follow similar protocols. The only exceptions are the addition of 2-mercaptoethanol to quench reactions using potassium permanganate, the duration of the DNA modification reactions, and the necessity to purify the modified DNA away from DEPC before the cleavage step.

3.1. In Vitro Experiments on Linear DNA Fragments

1. End-labeled DNA (5′ or 3′) should be separated from excess radiolabeled precurser either by agarose gel electrophoresis or by size exclusion column chromatography, and stored in TE.
2. In a 1.5-mL siliconized microcentrifuge tube, add 4 μL 10× reaction buffer, 40,000 cpm of DNA, and sterile deionized water to a volume of 20 μL. Proteins, diluted in an appropriate buffer, are added to the reaction to give a final volume of 40 μL (*see* **Note 1**). Incubate for the desired amount of time to form DNA–protein complexes.
3. Modifying the susceptible bases (*see* **Notes 2–4**):

- *$KMnO_4$ treatment*. Add freshly diluted potassium permanganate to give the appropriate concentration and incubate for desired time period. For example, adding 2 μL of 100 mM $KMnO_4$ (approximately 9 mM final concentration) for 2 min seems to work well for detecting melted DNA in transcription initiation complexes.
- *DEPC treatment*. Add 1 μL of DEPC to each tube. Mix by briefly vortexing and repeat vortexing every 5 min for 15 min. Vortexing is necessary because DEPC is sparingly soluble in aqueous solutions, so all reactions are run at essentially saturating DEPC and the concentration cannot be altered significantly.

4. Stopping the reactions:
   - *$KMnO_4$ treatment*. Quench the reaction by adding 3 μL of 2-mercaptoethanol, vortex and place on ice. Add 45 μL of 0.2% SDS, 2 mg/mL proteinase K and incubate at 50°C for 1 h. Add 90 μL of 0.6 M sodium acetate, 300 μg/mL linear polyacrylamide and 2.5 volumes 95% ethanol. Mix and centrifuge for 30 min at 14,000 rpm (16,000g) in a microfuge. Remove supernatant and wash with 150 μL of 70% ethanol. Centrifuge 5 min as above. Remove supernatant and dry the pellet on medium heat for 5 min in a Speed Vac.
   - *DEPC treatment*. Stop the reaction by adding an equal volume of DEPC Stop Buffer and phenol–$CHCl_3$ extract. Add 5 μL of 5 mg/mL linear acrylamide, and precipitate with 2.5 volumes of ethanol as above. After centrifugation, rinse with 70% ethanol and centrifuge again. Remove supernatant and dry the pellet on medium heat for 5 min in a Speed Vac.
5. Alkaline cleavage: Suspend the pellet in 50 μL of 1 M piperidine (10% v/v) and incubate at 90°C for 30 min. Place a lead weight on top of the tubes or use tube locks to prevent the lids from opening.
6. Place tubes on ice to cool and centrifuge briefly. Add 50 μL of 0.6 M sodium acetate, 300 μg/mL linearized polyacrylamide, and 250 μL of 95% ethanol. Mix and centrifuge for 30 min as above. Wash the pellet with 150 μL of 70% ethanol. Spin samples for 5 min. Remove supernatant.
7. To remove residual piperidine, add 30 μL of sterile deionized water to each sample and dry on medium heat in a Speed Vac (*see* **Note 5**).
8. Add 5 μL of electrophoresis loading buffer, vortex samples for 30 s and heat samples at 95°C for 3 min. Place samples on ice.
9. Load samples on sequencing gel and analyze by standard methods (*see* **Notes 6** and 7).

3.2. Treatment of DNA In Vivo with $KMnO_4$ and Purification

Potassium permanganate has been used to modify DNA in vivo, followed by analysis of modifications by primer extension or PCR. DEPC cannot be used for in vivo experiments because of its low solubility. The procedure for in vivo modification is fairly straightforward, however, certain nutrient rich media can quench permanganate. To avoid this problem, use minimal medium or increase the permanganate concentration so that the reaction mixture does not turn brown in less than 1 min. For some experiments, one can dilute the culture in minimal medium just prior to treatment with potassium permanganate. In vivo modification of mammalian cell cultures usually requires the removal of the growth medium just prior to treatment.

1. To 10 mL of diluted bacterial or yeast culture, add the appropriate amount of $KMnO_4$ (typically in the low mM range, depending upon the medium, approximately 10–20 mM for most media) for the desired amount of time (10 s to 5 min) in a shaking water bath. To quench treatment, remove samples from the water bath and pour immediately into prechilled Corex tubes and add 2-mercaptoethanol until the purple color disappears. Centrifuge to pellet the cells in a cold Sorvall SS34 rotor for 5 min at 5,000 rpm (3,000g). Discard the supernatant.
2. For mammalian cells grown to subconfluence on plastic growth dishes, remove growth medium and wash twice with phosphate buffered saline or minimal growth medium. Add the desired concentration of potassium permanganate (usually 2–20 mM) for the necessary period of time (10 s to 5 min). Stop the permanganate reaction by washing cell monolayers twice with phosphate buffered saline containing 2% 2-mercaptoethanol and once with phosphate buffered saline. Cells are then harvested with a rubber policeman or cell scraper.
3. Plasmid and genomic DNA can be isolated by a variety of standard methods *(49–51)*. Purified modified DNA should be adjusted to a final concentration of approximately 15 ng/μL and be free of contaminants that interfere with extension reactions. Extractions involving phenol should be repeated 3–4 times or until there is no contaminating interphase. Modifications to the DNA can then be visualized by PCR amplification detailed in **Subheading 3.4**. An alternative method for identifying modified genomic DNA is the use of ligation mediated PCR (LMPCR) *(52, 53)*.

3.3. Primer Extension Analysis of DNA Treated In Vitro or In Vivo

Since DNA modified in vivo or when in circular form is not end-labeled, primer extension is the method of choice to analyze the sites of modification. Modified bases result in extension stop sites because they block the elongating DNA polymerase. For in vitro studies, follow **steps 2–4** of **Subheading 3.1**, substituting

20–500 ng of purified DNA in place of radiolabeled DNA. For in vivo studies follow the steps of **Subheading 3.2**.

1. To the isolated, modified DNA (20–500 ng for in vitro and 500 ng for in vivo studies), add 0.3–0.5 × 10^6 cpm of 5′ end-labeled primer and dilute to 36 μL with distilled water.
2. Add 4 μL of 0.01 M NaOH to each reaction and mix well.
3. Denature DNA by heating to 95°C for 2 min.
4. Add 5 μL of 10× neutralization buffer and mix.
5. Hybridize primer to DNA by heating sample for 3 min at or just under the calculated T_m of the primer.
6. Add 5 μL of a solution containing all four dNTPs at a concentration of 5 mM each.
7. Add 0.5–1.0 unit of the Klenow fragment of DNA polymerase I and mix gently. Incubate tube for exactly 10 min at 50°C.
8. Quench by adding an equal volume (~50 μL) of 0.6 M sodium acetate, 20 mM EDTA, and place on ice.
9. Precipitate DNA by adding 300 μL of 95% ethanol, mix and centrifuge for 30 min. Wash the pellet with 150 μL of 70% ethanol. Centrifuge for 5 min, remove supernatant, and dry the pellet.
10. Suspend the pellet in 5 μL of electrophoresis loading buffer and run on a normal sequencing gel. Analyze by standard techniques (*see* **Notes 8** and **9**).

3.4. PCR Amplification of DNA Treated In Vivo or In Vitro

1. In a 0.65-mL microcentrifuge tube, add the following:
   - 5 μL 10× reaction buffer supplied with the thermostable polymerase
   - 2 μL of dNTP mix
   - 0.5 × 10^6 cpm end-labeled primer
   - Distilled water to a final volume of 49.5 μL.
2. Program thermocycler. For example:
   - 1 Round:
     - 2 min at 95°C, pause and add 0.5 μL of thermostable polymerase (2.5 units/μl)
     - 30 s at T_m of primer
     - 30 s at 72°C
   - 15–20 Rounds:
     - 1 min at 95°C
     - 30 s at T_m of primer
     - 30 s at 72°C
   - 1 Round:
     - 5 min at 72°C.

3. Precipitate DNA by adding 50 μL of 0.6 M sodium acetate, 300 μg/mL of linear polyacrylamide, and 250 μL of 95% ethanol. Mix and centrifuge for 30 min at 14,000 rpm (16,000g) in a microfuge. Wash the pellet with 150 μL of 70% ethanol, centrifuge for 5 min. Remove the supernatant and dry the pellet on medium heat for 5 min in a Speed Vac.
4. Suspend the pellet in 5 μL of loading buffer and run on sequencing gel. Analyze by standard techniques (*see* **Notes 8** and **9**).

4. Notes

1. False positive results can occur from the presence of nucleases in any of the proteins being tested. A necessary control is to incubate each protein with the DNA in the absence of further treatment with the modifying reagent. The DNA isolated from these reactions is run through the remainder of the analysis procedure to reveal any digestion of the DNA by contaminating nucleases.
2. Certain sequences of DNA are sensitive to DEPC and potassium permanganate treatment even in the absence of proteins. It is important to run a control lane of DNA to obtain a background level of sensitive sites.
3. To optimize reaction conditions, a titration of potassium permanganate for varying amounts of time may be necessary. Too little potassium permanganate results in no signal, and too much potassium permanganate can result in a high background. 2–5 mM potassium permanganate is typical for in vitro experiments, but for in vivo experiments where the medium may quench the reagent, concentrations up to 200 mM can be used. Times of reaction have been varied from 10 s up to 5 min, but in our hands there is much less difference in the results obtained with different reaction times than with different potassium permanganate concentrations. Thus, one can set up the experiment for a convenient period of time.
4. Potassium permanganate and DEPC react with proteins as well as DNA, which can impair their function. Thus, negative experiments may result from protein denaturation rather than a lack of DNA modification by the protein.
5. It is important to remove all the piperidine from the DNA following cleavage. If smeared bands are found on the gel, try doing more than one round of drying in the Speed Vac by redissolving the pellet in 30 μL of deionized water and drying as described in **step 7** of **Subheading 3.1**.

6. To obtain the maximum amount of information from the DNA of interest, perform separate experiments with either the template or the RNA-like strand radiolabeled.
7. A Maxam and Gilbert sequencing ladder of the DNA being analyzed run adjacent to the probing reactions is useful to identify specific modified sites.
8. For primer extension and PCR amplification reactions, several factors can affect the observed signal. A loss of signal can be due to (a) improper primer sequence; (b) annealing temperature higher than T_m ; (c) contaminants present in reaction; and (d) high concentration of magnesium ion.
9. Extra bands or smearing can occur if the annealing temperature is suboptimal and allows mispriming. Nonspecific hybridization can also occur if the radiolabled primer has undergone extensive decay. Freshly labeled primer reduces the risk of mispriming events. Supercoiled DNA can cause sequence induced stopping of the DNA polymerase. Linearizing the plasmid before fill-in or amplification can reduce improper extension.

References

1. Leonard, N.J., McDonald, J.J., and Reichmann, M.E. (1970). Reaction of diethyl pyrocarbonate with nucleic acid components I: Adenine. *Proc. Natl Acad. Sci. USA* **67**, 93–98.
2. Leonard, N.J., McDonald, J.J., Henderson, R.E.L., and Reichmann, M.E. (1971). Reaction of diethyl pyrocarbonate with nucleic acid components: Adenosine. *Biochemistry* **10**, 3335–3342.
3. Mendel, D., and Dervan, P.B. (1987). Hoogsteen base pairs proximal and distal to echinomycin binding sites on DNA. *Proc. Natl Acad. Sci. USA* **84**, 910–914.
4. Hayatsu, H., and Ukita, T. (1967). The selective degredation of pyrimidines in nucleic acids by permanganate oxidation. *Biochem. Biophys. Res. Commun* **29**, 556–561.
5. Howgate, P., Jones, A.S., and Tittensor, J.J. (1968). The permanganate oxidation of thymidine. *J. Chem. Soc. C* **1**, 275–279.
6. Iida, S., and Hayatsu, H. (1971). The permanganate oxidation of thymidine. *J. Biophys. Acta* **240**, 370–375.
7. Rubin, C.M., and Schmid, C.W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res.* **8**, 4613–4619.
8. Akman, S.A., Doroshow, J.H., and Dizdaroglu, M. (1990). Base modifications in plasmid DNA caused by potassium permanganate. *Arch. Biochem. Biophys* **282**, 202–205.
9. Akman, S.A., Forrest, G.P., Doroshaw, J.H., and Dizdaroglu, M. (1991). Mutation of potassium permanganate- and hydrogen peroxide-treated plasmid pZ189 replicating in CV-1 monkey kidney cells. *Mutat. Res* **261**, 123–130.
10. Sasse-Dwight, S., and Gralla, D. (1988). Probing the Escherichia coli glnALG upstream activation mechanism in vivo. *Proc. Natl Acad. Sci. USA* **85**, 8934–8938.
11. Sasse-Dwight, S. and Gralla, D. (1989). KMnO4 as a probe for lac promoter DNA melting and mechanism in vivo. *Proc. Natl Acad. Sci. USA* **264**, 8074–8081.
12. Grimes, E., Busby, S., and Minchin, S. (1991). Differential thermal energy requirement for open complex formation by Escherichia coli RNA polymerase at two related promoters. *Nucleic Acids Res.* **19**, 6113–6118.
13. Suh, W.C., Ross, W., and Record, M.T., Jr. (1992). Two open complexes and a requirement for magnesium to open the lambda-P-R transcription start site. *Science* **259**, 358–361.

14. Lofquist, A.K., Li, H., Imboden, M.A., and Paule, M.R. (1993). Promoter opening (melting) and transcription initiation by RNA polymerase I requires neither nucleotide β,γ hydrolysis nor protein phosphorylation. *Nucleic Acids Res.* 21, 3233–3238.
15. Kassavetis, G.A., Blanco, J.A., Johnson, T.E., and Geiduschek, E.P. (1992). Formation of open and elongating transcription complexes by RNA polymerase III. *J. Mol. Biol.* **226**, 47–58.
16. Wong, C., and Gralla, J.D. (1992). A role for the acidic repeat region of transcription factor sigma 54 in setting the rate and temperature dependence of promoter melting in vivo. *J. Biol. Chem.* **267**, 24762–24768.
17. Kainz, M., and Roberts, J. (1992). Structures of transcription elongation complexes in vivo. *Science* **255**, 838–841.
18. Ohlsen, K.L., and Gralla, J.D. (1992). Melting during steady-state transcription of the RRNB P-1 promoter in vivo and in vitro. *J. Bacteriol.* 174, 6071–6075.
19. Li, B., Weber, J.A., Chen, Y., Greenleaf, A.L., and Gilmour, D.S. (1996). Analysis of promoter-proximal pausing by RNA polymerase II on the hsp70 heat shock gene promoter in a Drosophila nuclear extract. *Mol. Cell. Biol.* **16**, 5433–5443.
20. Hartvig, L., and Christiansen, J. (1996). Intrinsic termination of T7 RNA polymerase mediated by either RNA or DNA. *EMBO J.* **15**, 4767–4774.
21. Komissarova, N., and Kashlev, M. (1997). Transcriptional arrest: Escherichia coli RNA polymerase translocates backward, leaving the 3′ end of the RNA intact and extruded. *Proc. Natl Acad. Sci. USA* 94, 1755–1760.
22. Lee, D.N., and Landick, R. (1992). Structure of RNA and DNA chains in paused transcription complexes containing Escherichia coli RNA polymerase. *J. Mol. Biol.* **228**, 759–777.
23. Carles-Kinch, K., and Kreuzer, K.N. (1997). RNA–DNA hybrid formation at bacteriophage T4 replication origin. *J. Mol. Biol.* **266**, 915–926.
24. Jeppesen, C., and Nielsen, P.E. (1988). Detection of intercalation-induced changes in DNA structure by reaction with diethyl pyrocarbonate or potassium permanganate. *FEBS Lett.* **231**, 172–176.
25. Fox, K.R., and Grigg, G.W. (1988). Diethyl pyrocarbonate and permanganate provide evidence for an unusual DNA conformation induced by binding of the antitumour antibiotics bleomycin and phleomycin. *Nucleic Acids Res.* **16**, 2063–2075.
26. Bailly, C., Gentle, D., Hamy, F., Purcell, M., and Waring, M.J. (1994). Localized chemical reactivity in DNA associated with the sequence-specific bisintercalation of echinomycin. *Biochem. J.* **300**, 165–173.
27. Michelottic, G.A., Michelotti, E.F., Pullner, A., Duncan, R.C., Eick, D., and Levens, D. (1996). Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. *Mol. Cell. Biol.* **16**, 2656–2669.
28. Hershkovitz, M., and Riggs, A.D. (1997). Ligation-mediated PCR for chromatin-structure analysis of interphase and metaphase chromatin. *Methods* **11**, 253–263.
29. Chiu, S.K., Rao, B.J., Story, R.M., and Radding, C.M. (1993). Interactions of three strands in joints made by RecA protein. *Biochemistry* **32**, 13146–13155.
30. Voloshin, O.N., and Camerini-Otero, R.D. (1997). The duplex DNA is very underwound in the three-stranded RecA protein-mediated synaptic complex. *Genes Cells* **2**, 303–314.
31. Plug, A.W., Peters, A.H.F.M., Keegan, K.S., Hoekstra, M.F., De Boer, P., and Ashley, T. (1998). Changes in protein composition of meiotic nodules during mammalian meiosis. *J. Cell Sci.* **111**, 413–423.
32. Duncan, R., Bazar, L., Michelotti, G., Tomonaga, T., Krutzch, H., Avigan, M., and Levens, D. (1994). A sequence-specific, single-strain binding protein activates the far upstream element of c-myc and defines a new DNA-binding motif. *Genes Dev.* **4**, 465–480.
33. Sun, W., and Godson, G.N. (1998). Structure of the Escherichia coli primase/single-strand DNA-binding protein/ phage G4ori-c complex required for primer RNA synthesis. *J. Mol. Biol.* **276**, 689–703.
34. Godson, G.N., Mustaev, A.A., and Sun, W. (1998). ATP cross-linked to Escherichia coli single-strand DNA-binding protein can be utilized by the catalytic center of primase as initiating nucleotide for primer RNA synthesis on phage G4ori-c template. *Biochemistry* **37**, 3810–3817.
35. McCarthy, J.G., and Rich, A. (1991). Detection of an unusual distortion in A-tract DNA using potassium permanganate effect of temperature and distamycin on the altered conformation. *Nucleic Acids Res.* **19**, 3421–3430.
36. Matyasek, R., Fulnecek, J., Fajkus, J., and Bazdek, M. (1996). Evidence for a sequence-directed conformation periodicity in the genomic highly repetitive DNA detectable with single-strand-specific chemical probe potassium permanganate. *Chromosome Res.* **4**, 340–349.

37. Epplen, J.T., Kyas, A., and Maeueler, W. (1996). Genomic simple repetitive DNAs are targets for differential binding of nuclear proteins. *FEBS Lett.* **389**, 92–95.
38. Herr, W. (1985). Diethyl pyrocarbonate: A chemical probe for secondary structure in negatively supercoiled DNA. *Proc. Natl Acad. Sci. USA* **82**, 8009–8013.
39. Voloshin, O.N., Mirkin, S.M., Lyamichev, V.I., Belotserkovskii, B.P., and Frank-Kamenetskii, M.D. (1988). Chemical probing of homopurine–homopyrimidine mirror repeats in supercoiled DNA. *Nature* **333**, 475–476.
40. Bentin, T., and Nielsen, P.E. (1996). Enhanced peptide nucleic acid binding to supercoiled DNA: Possible implications for DNA "breathing" dynamics. *Biochemistry* **35**, 8863–8869.
41. Furlong, J.C., and Lilley, D.M.J. (1986). Highly selective chemical modification of cruciform loops by diethyl pryrocarbonate. *Nucleic Acids Res.* **14**, 3995–4007.
42. Scholten, P.M., and Nordheim, A. (1986). Diethyl pyrocarbonate: A chemical probe for DNA cruciforms. *Nucleic Acids Res.* **14**, 3981–3993.
43. Balagurumoorthy, P., and Brahmachari, S.K. (1994). Structure and stability of human telomeric sequences. *J. Biol. Chem.* **269**, 21858–21869.
44. Nadel, Y., Weisman-Shomer, P., and Fry, M. (1995). The fragile X syndrome single strand d(CGG)n nucleotide repeats readily fold back to form unimolecular hairpin structures. *J. Biol. Chem* **270**, 28970–28977.
45. Jiang, H., Zacharia, W., and Amirhaeri, S. (1991). Potassium permanganate as an in situ probe for B–Z and Z–Z junctions. *Nucleic Acids Res.* **19**, 6943–6948.
46. Woelfl, S., Wittig, B., and Rich, A. (1995). Identification of transcriptionally induced Z-DNA segments in the human c-myc gene. *Biochim. Biophys. Acta* **1264**, 294–302.
47. Glover, J.N.M., Farah, C.S., and Pulleyblank, D.E. (1990). Structural characterization of separated H-DNA conformers. *Biochemistry* **29**, 11110–11115.
48. Haner, R., and Dervan, P.B. (1990). Single-strand DNA triple-helix formation. *Biochemistry* **29**, 9761–9765.
49. Huibregtse, J.M., and Engelke, D.R. (1991). Direct sequence and footprint analysis of yeast DNA by primer extension. *Methods Enzymol.* **194**, 550–562.
50. Holmes, D.S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**, 193–197.
51. Owen, R.J., and Borman, P. (1987). A rapid biochemical method for purifying high molecular weight chromosomal DNA for restriction enzyme analysis. *Nucleic Acids Res.* **15**, 3631.
52. Mueller, P.R., and Wold, B.J. (1989). In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246**, 780–786.
53. Garrity, P.A., and Wold, B.J. (1992). Effects of different DNA polymerases in ligation-mediated PCR: Enhanced genomic sequencing and in vivo footprinting. *Proc. Natl Acad. Sci. USA* **89**, 1021–1025.

Chapter 7

Uranyl Photofootprinting

Peter E. Nielsen

Summary

The uranyl-(VI) cation (UO_2^{2+}) forms strong complexes with accessible phosphates of nucleic acid (DNA and RNA) backbones. Upon excitation with long wavelength ultraviolet light (λ = 300–420 nm), uranyl ions bound to backbone phosphates oxidize proximal sugars and induce nucleic acid backbone cleavage. Thus the uranyl(VI) ion functions as a very specific and efficient photochemical probe for identifying ligand(protein)-phosphate contacts in nucleic acid complexes as well as potential (high affinity) cation (e.g., Mg^{2+})-binding sites in folded nucleic acids. Finally, the cleavage modulation of duplex DNA reflects helix conformation in terms of minor groove width, due to preferential affinity/oxidation efficiency for such regions of the DNA helix.

Key words: Uranyl, Nucleic acid photocleavage, DNA footprinting, DNA conformation, Metal ion binding.

1. Introduction

It has long been known that the uranyl-(VI) ion (UO_2^{2+}) forms strong complexes with various inorganic and organic anions, including phosphates, and that the photochemically excited state of this ion is a very strong oxidant *(1)*. For instance, uranyl-mediated photooxidation of alcohols has been studied in detail *(2, 3)*. It is also widely recognized that uranyl chemistry and photophysics/photochemistry are very complex. Thus monomeric UO_2^{2+} is only present at low pH (pH ~ 2), whereas polynuclear species and various "hydroxides," which often precipitate form aqueous solution, form at higher pH *(4)*.

In spite of this complexity we have found that uranyl-mediated photocleavage of DNA can be used to probe for accessibility of

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_7

the phosphates in the DNA backbone *(5–8)*. Thus uranyl is a sensitive probe for protein–DNA-phosphate contacts *(5, 6)* as well as for DNA conformation in terms of DNA minor groove width *(7–11)* and helix flexibility/deformability *(12)*. Furthermore, binding sites for divalent metal ions in folded DNA *(13)* or RNA *(14, 15)* can be studied by uranyl photocleavage.

The systems that have so far been analyzed by uranyl-mediated DNA photocleavage include the λ-repressor/O_R1 operator complex *(5)*, *E. coli* RNA polymerase/deoP1 promoter transcription initiation open complex *(6)*, transcription factor IIIA (TFIIIA)/Xenopus 5S internal control region (ICR) complex *(16)*, catabolite regulatory protein (CRP)/operator DNA complex and the CRP/RNA polymerase/deoP2 promoter initiation complex *(17)*, bent kinetoplast DNA *(7)*, and triplex DNA *(18)*. Furthermore, we have found that some small ligand–DNA complexes (exemplified by mitramycin *(19)*, distamycin *(20)*, and spermine *(21)*) may also be studied by uranyl photofootprinting. Finally, divalent metal ion-binding sites in an RNA polymerase–promoter open complex *(6)*, a four-way DNA "Holliday junction" *(13)*, a hammer head ribozyme *(14)*, and in yeast $tRNA^{phe}$ *(15)* have been analyzed by uranyl photocleavage. This technique takes advantage of competition of low-affinity cleavage (binding) sites by a metal ion chelator like citrate.

The molecular mechanism for uranyl-mediated photocleavage of DNA is not fully understood, but we have shown that uranyl binds to the phosphates of DNA and oxidizes the proximal deoxyriboses most likely via a direct electron transfer mechanism *(22)*. The main products are 3′-phosphate and 5′-phosphate termini in the DNA, and the free nucleobases are liberated in the process *(19)*. Since uranyl binding is to the phosphate groups of the DNA, very limited sequence dependence of the photocleavage is seen.

2. Materials

1. Uranyl nitrate ($UO_2(NO_3)_2$), analytical grade: 100 mM stock solution in H_2O. This solution was found to be stable for photofootprinting purposes for at least 12 months and was diluted to working concentrations immediately prior to use (*see* **Note 1**).
2. ^{32}P-Endlabeled DNA restriction fragments (*see* **Note 2**).
3. Buffer for formation of protein–DNA complex (*see* **Note 3**).
4. Na-acetate, 0.5 M, pH 4.5.
5. Ethanol, 96%.

6. Ethanol, 70%.
7. Calf thymus DNA, 2 mg/ml.
8. Gel-loading buffer 80% formamide in TBE buffer, 0.05% bromophenol blue, 0.05% xylene cyanol.
9. TBE buffer: 90 mM Tris-borate, 1 mM EDTA, pH 8.3.
10. Polyacrylamide gel: 8% acrylamide, 0.3% bis-acrylamide, 7 M urea, TBE buffer. Size: 0.2 mm × 60 cm × 20 cm.
11. *l*-Repressor: 1 μg/μl (*see* **Note 4**).
12. Buffer for λ-repressor footprinting: 40 mM Tris–HCl, pH 7.0, 2.5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1 mM EDTA, 200 mM KCl.
13. DNase I: 1 mg/ml in 10 mM Tris–HCl, pH 7.4, 1 mM $MgCl_2$.
14. X-ray film: Agfa Curix RP1.
15. Philips TL 40W/03 fluorescent light tube which fits into standard (20 W) fluorescent light tube sockets if the transformer is changed to 40 W (*see also* **Note 5**).

3. Methods

3.1. Protocol

A typical uranyl photoprobing experiment is performed as follows:

1. Form the complex to be analyzed by mixing the ^{32}P-endlabeled DNA fragment (>20,000 cpm/sample) (*see* **Note 2**) with the DNA-binding ligand in 90 μl of the buffer (*see* **Note 3**) (containing 0.5 μg calf thymus DNA carrier) in a 1.5-ml polypropylene microfuge tube at the desired temperature.
2. Dilute the 100 mM uranyl stock solution to 10 mM in H_2O.
3. Add 10 μl of this to the sample and mix well (*see* **Note 6**).
4. Place the sample in a thermostated heating/cooling block if the temperature is critical.
5. Irradiate the sample for 30 min at 420 nm by placing the open microfuge tube directly under the fluorescent light tube (*see* **Note 5**).
6. Add 20 μl of 0.5 M Na-acetate, pH 4.5, to prevent coprecipitation of uranyl (which will interfere with subsequent gel analysis), and precipitate the DNA by addition of 250 μl 96% ethanol.

7. Place the sample on dry ice for 15 min (or overnight at –20°C) and centrifuge 30 min at 20,000 × *g*.
8. Wash the pellet with 100 µl 70% ethanol, dry in vacuo, and redissolve in 4–10 µl 80% formamide gel-loading buffer.
9. Heat the sample at 90°C for 5 min.
10. Load 10,000 cpm on a polyacrylamide sequencing gel (0.2–0.4 mm thick, 60-cm long) and run the gel at 2,000 V. A sequence ladder (e.g., A + G) is run in parallel (see example).
11. Visualize radioactive bands by autoradiography overnight (or longer) at –70°C using amplifying screens.
12. Quantitate the results by densitometric scanning of the autoradiograms, if desired (*see* **Note 7**).

3.2. A Typical Analysis

Figure 1 shows a footprinting experiment of the complex between λ-repressor and the O_R1 operator DNA using uranyl and DNase I. Quantitative analysis of these results by densitometric scanning, and displaying the results on the DNA sequence (**Fig. 2a)** reveals that mainly four regions of the O_R1 operator are protected from photocleavage by uranyl. The results, furthermore, show that these regions coincide with those protected from attack by hydroxy radicals as well as the phosphates indicated by ethylation interference and X-ray crystallography to be involved in protein binding *(23)*. However, the footprinting patterns are not identical as discussed later. Display of uranyl footprinting results on a DNA double helix model is very informative and in the case of the λ-repressor/O_R1 complex shows (**Fig. 2b**) that the repressor binds to one face of the helix and that the phosphate contacts are positioned mainly in regions across the major groove of the DNA thus in full accord with binding of the α_3-recognition protein helix of each repressor subunit in the major groove.

3.3. Comparison of Uranyl Photoprobing to Other Techniques

The results obtained by uranyl photofootprinting are comparable to those obtained by hydroxyl radical (EDTA(FeII)) footprinting *(24, 25)* and ethylation interference experiments *(26)*.

Because the uranyl ion binds to the phosphates of the DNA, a uranyl photofootprinting experiment will report which phosphates of the DNA backbone are accessible to the uranyl, and which are therefore not involved in contacts with the bound ligand. Hydroxyl radical footprinting reports accessibility of the deoxyriboses of the DNA backbone and in the cases studied with both uranyl and hydroxyl radical probing (λ-repressor *(5, 24)*, RNA polymerase *(6, 27)*, and TFIIIA *(14, 28)*), the footprint obtained by uranyl involves fewer nucleobase positions than that obtained by EDTA(FeII).

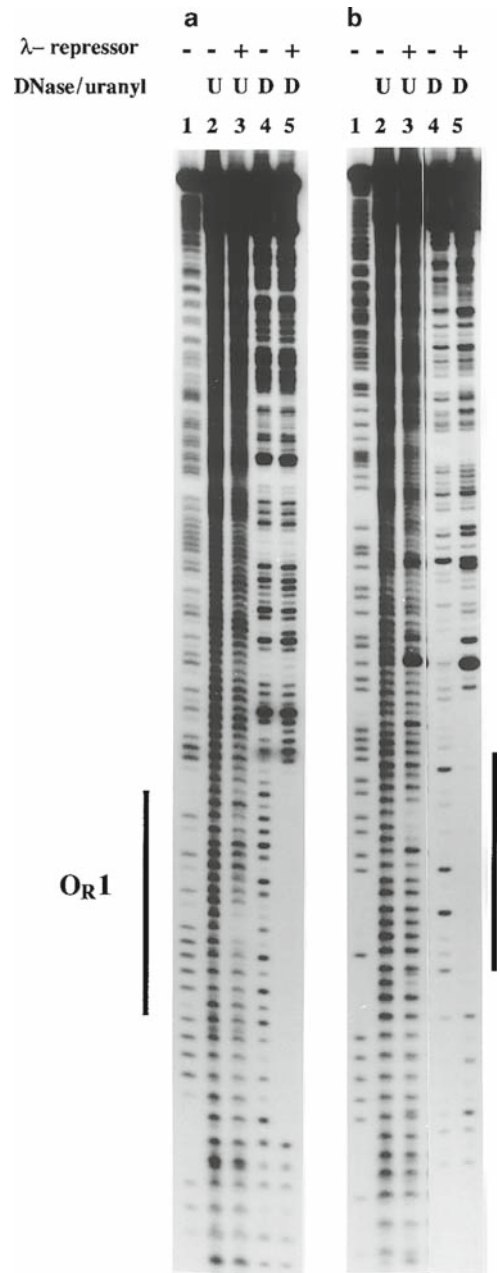


Fig. 1. Uranyl photofootprint and DNase I footprint of the λ-repressor/O_R1 operator complex (*see* **ref.** *5* for details). The O_R1 operator sequence was cloned into the *Bam*/*Hin*dIII site of pUC19 and the 225-base pair *Eco*RI/*Pvu*II fragment labeled with ^{32}P at the 3′- or 5′-end of the *Eco*RI site was used in the experiments. *Lanes 1* and *3* are controls without added λ-repressor (0.7 μg/sample). *Lanes 1* and *2* are uranyl photofootprints, while *lanes 3* and *4* are DNase I (0.5 μg/ml, 5 min at RT) footprints. *Lanes S* are A + G sequence reactions obtained by treating the DNA with 60% formic acid for 5 min at RT and subsequent piperidine treatment. The samples were analyzed on an 8% polyacrylamide gel and run at 2,500 V. The gel was subjected to autoradiography for 16 h at –70°C using intensifying screens.

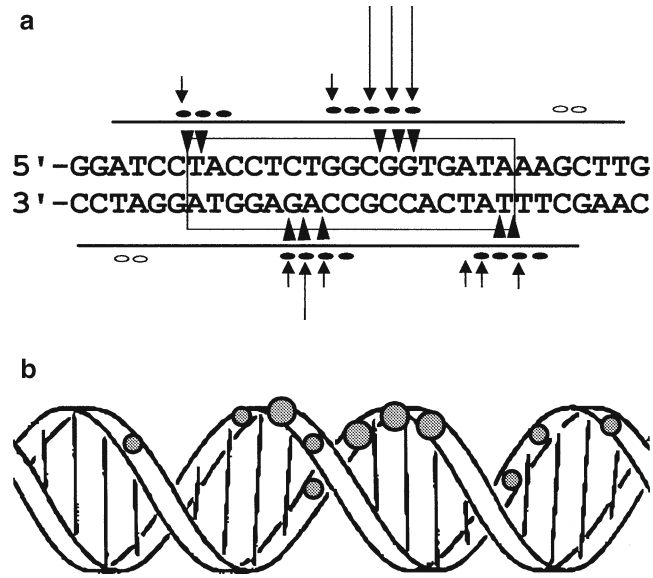


Fig. 2. (**a**) O_R1-operator sequence (*box*) showing uranyl photofootprint (*arrows*), EDTA/FeII footprint (*dots* (*23*)) ethylation interference (*arrow heads* (*24*)) and DNase I footprint (*brackets*). (**b**) Display of the uranyl photofootprint on O_R1-helix model. The size of the *dots* signifies degree of protection.

Interference probing by phosphate ethylation using ethyl nitroso urea also reports the involvement of individual phosphate groups in protein–DNA interaction. However, this is an interference technique and therefore only phosphate groups which are indispensable for complex formation are detected. Thus for small complexes – e.g., the λ-repressor – O_R1 complex-ethylation interference and uranyl results are virtually equivalent, whereas for larger complexes – e.g., RNA polymerase–promoter-complexes – only part of the contacts detected by uranyl photofootprinting are picked up by ethylation interference. In conclusion, ethylation interference, hydroxyl radical, and uranyl footprinting complement each other. Hydroxyl radicals reflects the accessibility of individual deoxyriboses, uranyl reports on the accessibility of individual phosphates, and ethylation interference reports on phosphates which are indispensable for complex formation.

Furthermore, both hydroxyl radicals and uranyl are able to sense variations in DNA conformation, and for both probes groove width has been implicated as the determinant factor. Generally speaking, hydroxyl radical cleavage is more intense as the minor groove widens, whereas uranyl photocleavage becomes more intense as it narrows. Macroscopically bent DNA contains a number of A-tracts in phase with the DNA helical pitch. These A-tracts all exhibit a significantly narrowed minor groove which is sensed by both uranyl and hydroxyl radicals. However, the bending is most likely confined to the A-tract junctions. Interestingly,

identically positions mixed A/T tracts do not induce macroscopic bending of the DNA helix, although such A/T tracts according to uranyl photoprobing possess a narrow minor groove very similar to A-tracts *(12)*, whereas they are not recognized by hydroxyl radicals as having a narrowed minor groove. It has been proposed that the explanation for this apparent contradiction resides in the mechanism of action of the uranyl ion, which requires actual complexation with the DNA which may affect the DNA helix conformation. Therefore, a DNA helix which is poised for minor groove narrowing, i.e., for which the energy difference between a "normal" and a narrow minor groove is rather small, may upon binding of uranyl (or other ligands such as distamycin and similar minor groove DNA binders) which prefer binding to such a narrow minor groove conformation adopt this conformation by an induced fit mechanism *(12)*. Thus while hydroxy radicals only probe existing (static) DNA conformations, the uranyl may in addition probe energetically close and thus dynamically possible DNA conformations, which may be very important for ligand (protein) binding in general *(12)*.

Finally, UO_2^{2+} being a divalent cation may to some extent mimic Mg^{2+} in terms of high-affinity binding sites in protein–nucleic acid complexes and in folded nucleic acids. Consequently, such high-affinity binding sites for the UO_2^{2+} ion can result in hypersensitive cleavage sites. In this respect the uranyl photoprobing can be compared to oxidative cleavage of nucleic acids by the Fe^{2+}/Fe^{3+} ion redox pair *(29, 30)*. Uranyl being a photochemical technique has the added – but so far unexplored – potential to be used for temperature and kinetic studies.

4. Notes

1. Uranyl acetate ($UO_2(CH_3COO)_2$) gives identical results, but the 100 mM stock solution in this case has to be made 50 mM in HCl in order to be stable.
2. ^{32}P-endlabeled DNA fragments of ~50–300 base pairs in length are prepared by standard techniques *(31)*: Typically, the plasmid containing the protein-binding site is opened by a restriction enzyme that cleaves at a distance of 20–50 base pairs from the binding site. (This distance is important since the best resolution is obtained in the 20–70 base interval and the bands of uranyl cleaved DNA fragments become "fuzzy" above ~ 100 bases). The plasmid is labeled either at the 3′-end with α-^{32}P-dNTP and the Klenow fragment of DNA polymerase, or at the 5′-end (after dephosphorylation with alkaline

phosphatase) with γ^{32}P-ATP and polynucleotide kinase. The plasmid is then treated with a second restriction enzyme cutting 50–300 base pairs from the labeling site, and the DNA fragment containing the protein-binding site is purified by gel electrophoresis in 5% polyacrylamide, TBE buffer. The DNA fragment is extracted from the excised gel slice with 0.5 M NH_4-acetate, 1 mM EDTA (16 h, RT) and precipitated by addition of 2 vols. of 96% ethanol. The pellet is washed with 70% ethanol and dried.

3. *Choice of Buffer*. The choice of an optimal buffer for a uranyl photofootprinting experiment is crucial for a successful result. In particular, the pH of the medium is important. The uranyl-mediated photocleavage of DNA is extremely dependent on pH being most efficient at pH ~ 6, less efficient at pH 5 and 7, and virtually absent at pH 8 *(22)*. Furthermore, as the pH is lowered, a strong modulation of the sequence dependence of the cleavage is observed. In fact, this modulation reflects the conformation of the DNA *(7–10)*. Thus for photofootprinting where an even cleavage is warranted, buffers of pH 6.5–7.0 are advantageous whereas buffers of pH 6.0–6.5 should be chosen for studies of DNA structure. If information about metal ion-binding sites is desired, citrate (0.1–1 mM) should be included in the buffer.

   The composition of the buffer and the buffer capacity is also of importance. Since the uranyl solution is acidic, it should be checked if addition of uranyl changes the pH of the medium. Furthermore, uranyl-mediated photocleavage of DNA is most efficient in acetate of formate buffers, less efficient in Tris–HCl, very inefficient in Hepes or Pipes buffer, and virtually absent in phosphate buffers (uranyl phosphate precipitates). The ionic strength (as Na^+) is of minor importance and the cleavage is not affected by the presence of Mg^{2+} or DTT either. Finally, the uranyl photoreaction is not influenced by the temperature (0–70°C) *(31)*. Within these constraints a buffer that allows protein–DNA binding must be chosen.

4. λ-Repressor was prepared according to ref. *32* using an overproducer plasmid, pAE305 in *E. coli*.

5. Light source. Any light source emitting at 300–420 nm can be used. This could be the Philips TL 40W/03 tube emitting at 420 ± 30 nm. Alternative fluorescent light tubes are Philips TL 20W/12 (300 nm) or TL 20W/09N (365 nm). Lamps emitting below 300 nm are not recommended due to absorption by the DNA bases at these wavelengths. The fluorescent light tubes suggested in this paper are not very powerful, but quite sufficient for footprinting and they are inexpensive and do not require sophisticated power supplies. However, if shorter irradiation times are required, uranyl photofootprinting experiments

are quite adequately performed with pyrex-filtered light from high-pressure Hg-lamps, Xenon lamps, or lasers of the appropriate wavelength (300–420 nm).

6. Order of mixing. It is important that the uranyl be added last since the uranyl–DNA complex is very stable (K_a is estimated to 10^{10} M^{-1} *(31)*). Uranyl–DNA aggregates which precipitate often form, without this affecting adversely the outcome of the footprinting reaction. Conversely, if uranyl is added prior to the DNA-binding ligand, the ligand will only have limited access to the DNA.

   It is also extremely important that dilution of the uranyl stock solution is performed immediately prior to use since uranyl solutions are not stable at pH ≥ 2.

7. Examples of densitometric scanning and quantification of footprinting experiments can be found in *(8, 33, 34, 35)*.

References

1. Burrows, H.D., & Kemp, T.J. (1974). The photochemistry of the uranyl ion. *Chem. Soc. Rev.* **3**, 138–165.
2. Azenha, M.E.D.G., Burrows, H.D., Furmosinho, S.J., & Miguel, M.G.M. (1989). Photophysics of the excited uranyl ion in aqueous solutions. Part 6. Quenching effects of aliphatic alcohols. *J. Chem. Soc. Faraday Trans.* **85**, 2625–2634.
3. Cunningham, J., & Srijaranai, S. (1990). Sensitized photo-oxidations of dissolved alcohols in homogeneous and heterogeneous systems. Part 1. Homogeneous photo-sensitization by uranyl ions. *J. Photochem. Photobiol. A. Chem.* **55**, 219–232.
4. Greenwood, N.N., & Earnshaw, A. (1986). Chemistry of the Elements. p. 1478. Pergamon Press, Oxford.
5. Nielsen, P.E., Jeppesen, C., & Buchardt, O. (1988). Uranyl salts as photochemical agents for cleavage of DNA and probing of protein-DNA contacts. *FEBS Lett.* **235**, 122–124.
6. Jeppesen, C., & Nielsen, P. (1989). Uranyl mediated photofootprinting reveals strong E. coli RNA polymerase-DNA backbone contacts in the +10 region of the deoP1 promoter open complex. *Nucleic Acids Res.* **17**, 4947–4956.
7. Nielsen, P.E., Møllegaard, N.E. & Jeppesen, C. (1990). Uranyl photoprobing of conformational changes in DNA induced by drug binding. *Anticancer Drug Des.* **5**, 105–110.
8. Nielsen, P.E., Møllegaard, N.E., & Jeppesen, C. (1990). DNA conformational analysis in solution by uranyl mediated photocleavage. *Nucleic Acids Res.* **18**, 3847–51.
9. Bailly, C., Møllegaard, N.E., Nielsen, P.E., & Waring, M.J. (1995). The influence of the 2-amino group of guanine on DNA conformation. Uranyl and DNase I probing of inosine/diaminopurine substituted DNA. *EMBO J.* **9**, 2121–2131.
10. Sönnichsen, S.H., & Nielsen, P.E. (1996). Enhanced uranyl photocleavage across the minor groove of all $(A/T)_4$ sequences indicates a similar narrow minor groove conformation. *J. Mol. Recognit.* **9**, 219–227.
11. Møllegaard, N.E., & Nielsen, P.E. (2003). Increased temperature and 2-Methyl-2,4-pentadiol change the DNA structure of both curved and uncurved adenine/thymine-rich sequences. *Biochemistry.* **42**, 8587–8593.
12. Møllegaard, N.E., Lindemose, S., & Nielsen, P.E. (2005). Uranyl photoprobing of nonbent A/T- and bent A-tracts. A difference of flexibility?. *Biochemistry.* **44**, 7855–7863.
13. Møllegaard, N.E., Murchie, A.I.H., Lilley, D.M.J., & Nielsen, P.E. (1994). Uranyl photoprobing of a four-way DNA junction. Evidence for specific metal ion binding. *EMBO J.* **13**, 1508–1513.
14. Bassi, G.S., Møllegaard, N.E., Murchie, A.I., von Kitzing, E., & Lilley, D.M. (1995). Ionic interactions and the global conformations of the hammerhead ribozyme. *Nat. Struct. Biol.* **2**, 45–55.
15. Nielsen, P.E., & Møllegaard, N.E. (1996). Sequence/structure selective thermal and

photochemical cleavage of yeast-tRNAPhe by UO22+. *J. Mol. Recognit.* **9**, 228–232.

16. Nielsen, P.E., & Jeppesen, C. (1990). Photochemical probing of DNA complexes. *Trends Photochem. Photobiol.* **1**, 39–47.
17. Møllegaard, N.E., Rasmussen, P.B., Valentin-Hansen, P., & Nielsen, P.E. (1993). Characterization of promoter recognition complexes are formed by CRP and CytR for repression and by CRP and RNA polymerase for activation of transcription on the *E. coli deoCp2 promoter. J. Biol. Chem.* **268**, 17471–17477.
18. Nielsen, P.E. (1992). Uranyl photofootprinting of triple helical DNA. *Nucleic Acids Res.* **20**, 2735–2739.
19. Nielsen, P.E., Cons, B.M.G., Fox, K.R., & Sommer, V.B. (1990). Uranyl photofootprinting. DNA structural changes upon binding of mithramycin In: Molecular Basis of Specificity in Nucleic Acid Drug Interactions (Eds. B. Pullman & J. Jortner) vol. 23, pp. 423–432. The Jerusalem Symposium on Quantum Chemistry and Biochemistry. Kluwer, Dordrecht.
20. Møllegaard, N.E., & Nielsen, P.E. (1997). Uranyl photoprobing of DNA structures and drug-DNA complexes In: Drug-DNA Interaction Protocols (Ed. K.R. Fox) pp. 43–50. Humana Press, New Jersey.
21. Lindemose, S., Nielsen, P.E., & Møllegaard, N.E. (2005). Polyamines preferentially interact with bent adenine tracts in double-stranded DNA. *Nucleic Acids Res.* **33**, 1790–1803.
22. Nielsen, P.E., Hiort, C., Buchardt, O., Dahl, O., Sönnichsen, S.H., & Nordèn, B. (1992). DNA binding and photocleavage by uranyl(VI) ($UO2^{2+}$) salts. *J. Am. Chem. Soc.* **114**, 4967–4975.
23. Schultz, S.C., Shields, G.C., & Steitz, T.A. (1991). Crystal structure of a CAP-DNA complex: The DNA is bent by 90°. *Science.* **253**, 1001–1007.
24. Tullius, T.D., & Dombroski, B.A. (1986). Hydroxyl radical "footprinting": high-resolution information about DNA protein contacts and application to lambda repressor and Cro protein. *Proc. Natl. Acad. Sci. USA.* **83**, 5469–5473.
25. Burkhoff, A.M., & Tullius, T.D. (1987). The unusual conformation adopted by the adenine tracts in kinetoplast DNA. *Cell* **48**, 935–943.
26. Siebenlist, U., Simpson, R.B., & Gilbert, W. (1980). E. coli RNA polymerase interacts homologously with two different promoters. *Cell* **20**, 269–281.
27. O'Halloran, T.V., Frantz, B., Shin, M.K., Ralston, D.M., & Wright, J.G. (1989). The MerR heavy metal receptor mediates positive activation in a topologically novel transcription complex. *Cell* **56**, 119–129.
28. Vrana, K.E., Churchill, M.E.A., Tullius, T.D., & Brown, D.D. (1988). Mapping functional regions of transcription factor TFIIIA. *Mol. Cell. Biol* **8**, 1684–1696.
29. Berens, C., Streicher, B., Schroeder, R., & Hillen, W. (1998). Visualizing metal-ion-binding sites in group I introns by iron(II)-mediated Fenton reactions. *Chem. Biol.* **5**, 163–175.
30. Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A., & Mustaev, A. (1996). Mapping of catalytic residues in the RNA polymerase active center. *Science* **273**, 107–109.
31. Maniatis, T., Fritsch, E.F., & Sambrook, J. (1982). Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
32. Amann, E., Brosins, J., & Ptasne, M. (1983). Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli. Gene* **25**, 167–178.
33. Jeppesen, C., & Nielsen, P.E. (1989). Photofootprinting of drug-binding sites on DNA using diazo- and azido-9-aminoacridine derivatives. *Eur. J. Biochem* **182**, 437–444.
34. Dabrowiak, J.C., Kissinger, K., & Goodisman, J. (1989). Quantitative footprinting analysis of drug-DNA interactions: Fe(III) methidium-propyl-EDTA as a probe. *Electrophoresis* **10**, 404–412.
35. Wittberger, D., Berens, C., Hammann, C., Westhof, E & Schroeder, R. (2000). Evaluation of uranyl photocleavage as a probe to monitor ion binding and flexibility in RNAs. *J. Mol. Biol.*, **300**, 339–52.

Chapter 8

Identification of Protein/DNA Contacts with Dimethyl Sulfate: Methylation Protection and Methylation Interference

Peter E. Shaw and A. Francis Stewart

Summary

Sequence-specific protein/DNA contacts direct most transcription factors to binding sites within the promoters of genes they regulate. Several chemical probes, such as dimethyl sulfate, have been used to obtain information on these sites of interaction. Protection and interference patterns frequently correspond to highly conserved positions within binding sites and are often specific for a given transcription factor or family of factors. The methods described here can be used to identify sites within a DNA sequence that are bound by nuclear factors or to characterise the contacts made by a purified factor or recombinant protein in vitro. As methylation protection is the in vitro equivalent of in vivo genomic footprinting, a direct comparison between in vivo and in vitro footprints can be made.

Key words: Maxam and Gilbert sequencing, Gel electrophoresis, DNA-binding proteins, Gel retardation.

1. Introduction

Dimethyl sulfate (DMS) is an effective and widely used probe for sequence-specific protein/DNA interactions. It is the only probe routinely used both for in vitro (methylation protection, methylation interference) and in vivo (DMS genomic footprinting) applications since it rapidly reacts with DNA at room temperature and readily penetrates intact cells *(1)*. DMS methylates predominantly the 7-nitrogen of guanine and 3-nitrogen of adenine. Thus reactivity with G residues occurs in the major groove and

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_8

with A residues occurs in the minor groove. In standard Maxam and Gilbert protocols *(2)*, the methylated bases are subsequently converted to strand breaks and displayed on sequencing gels.

Methylation protection and interference are essentially combinations of the gel retardation assay *(3, 4)* with the DMS reaction of the Maxam and Gilbert sequencing procedure. Protein/DNA interactions are reflected either as changes in DMS reactivities caused by bound protein (methylation protection) or as selective protein binding dictated by methylation (methylation interference).

In methylation protection, protein is first bound to DNA that is uniquely end-labelled and the complex is reacted with DMS. DMS reactivities of specific residues are altered by bound protein either by exclusion, resulting in reduced methylation, increased local hydrophobicity, resulting in enhanced methylation or by local DNA conformational changes, such as unwinding, resulting in altered reactivity profiles *(5–7)*. After the DMS reaction, free DNA is separated from protein bound DNA by gel retardation and both DNA fractions are recovered from the gel. Methylated residues are converted into strand scissions and the free and bound DNA fractions are compared on a sequencing gel. A complete analysis requires the examination of both strands. This is accomplished by preparing two DNA probes, each uniquely labelled at one end, and carrying both probes through the protocols. A binding site characterised by methylation protection will therefore appear as a cluster of altered DMS reactivities.

In methylation interference *(8, 9)*, DNA is first reacted with DMS, purified, and then presented to protein. Under the reaction conditions used methylation is partial, yielding approximately one methylated base per DNA molecule. Thus the protein is presented with a mixture of DNA molecules that differ with respect to the positions of methyl groups. Some methyl groups will interfere with protein binding since they lie in or near the binding site. Gel retardation separates the mixture into two fractions: free DNA, which, as long as DNA is in excess over binding activity, represents the total profile of methylation reactivity and bound DNA, which will not contain any molecules with methyl groups incompatible with binding. Both free and bound DNA fractions are recovered, methylated residues are converted to strand scissions and the fractions are compared on a sequencing gel. The binding site is observed as the absence of bands in the bound sample corresponding to the positions where methylation interferes with binding.

It is obvious that these two uses of DMS may not deliver identical results. For example, **Fig. 1** presents a comparison obtained from experiments with the serum response element (SRE) in the human c-*fos* promoter and its binding factors SRF and p62^{TCF}/Elk-1 *(10)*. Since the use of DMS in vivo for genomic footprinting

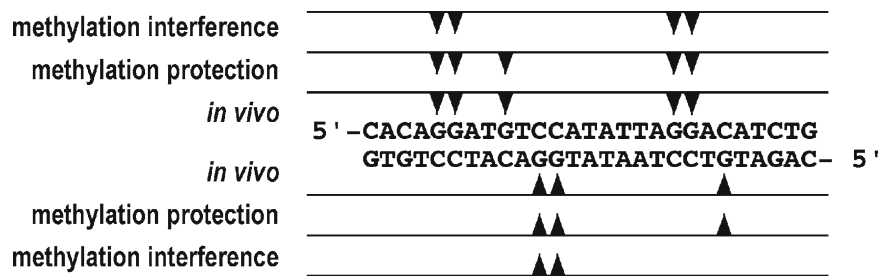


Fig. 1. Comparison of methylation interference and protection patterns formed by factors binding at the c-*fos* serum response element (SRE) in vitro and in vivo. G residues identified by methylation interference *(9)*, methylation protection and in vivo genomic footprinting *(7)* are indicated. An additional G on both the *upper* and *lower* strands is implicated in the protein/DNA interaction by methylation protection.

is limited to the equivalent of methylation protection, a direct comparison between in vivo and in vitro footprints requires the use of methylation protection rather than the more widely used methylation interference assay. These two techniques are, however, very similar in practical terms and thus are presented together. Both techniques rely on pre-established conditions that permit a protein/DNA complex to be resolved in a gel retardation assay (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions) and on the chemical DNA sequencing methodology of Maxam and Gilbert, for which the reader is advised to consult ref. *2* and chapter "In Cellulo DNA Analysis (LMPCR Footprinting)" for a detailed treatment.

2. Materials

1. Dimethyl sulfate (Merck), analytical grade.
2. Piperidine (Sigma), analytical grade: use freshly made 1:10 dilution in double-distilled water.
3. Phenol/chloroform 50% v/v, buffered with 50 mM Tris–HCl pH 8.
4. NA45 paper (Schleicher and Schuell) or equivalent, if available.
5. 3MM paper (Whatman).
6. Gel retardation equipment (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions").
7. Electro-blotting apparatus for Western transfer, e.g. Bio-Rad Trans-Blot.

8. Standard DNA sequencing gel electrophoresis equipment.
9. Vacuum gel drier (optional).
10. TBE buffer: 108 g Tris base, 55 g boric acid, 40 mL of 0.5 M EDTA pH 8.3 per litre for 10× stock; working solution is 89 mM Tris base, 89 mM boric acid, 2 mM EDTA.
11. NA45 elution buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, 1 M NaCl).
12. Salmon testis DNA or calf thymus DNA, sheared, as carrier (Sigma; dissolved at 3 mg/mL in 10 mM Tris–HCl pH 8.0, 1 mM EDTA).
13. Sequencing loading buffer (90% formamide, 10 mM EDTA 0.1% w/v bromophenol blue, 0.1% w/v xylene cyanol blue).
14. Gel retardation loading buffer (20% Ficoll, 20 mM EDTA, 0.1% w/v bromophenol blue).
15. 2× DMS buffer (120 mM NaCl, 20 mM Tris–HCl pH 8.0, 20 mM $MgCl_2$, 2 mM EDTA).
16. DMS stop buffer (1.5 M NaAc pH 7.0, 1 M 2-mercaptoethanol), store frozen.
17. X-ray film (e.g. Kodak X-Omat) or imaging plate for phosphorimager.

3. Method

3.1. Methylation Protection

1. Incubate 300,000 cpm of uniquely end-labelled DNA probe and a corresponding amount of protein together, as previously optimised for gel retardation analysis (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions, **Note 1**) in a total volume of 100 μL.
2. Add 1 μL of DMS and incubate at room temperature (the incubation time depends on the length of the DNA probe and is empirical: as a guide 200 bp fragment $t = 1.5$ min; for 50 bp oligo duplex $t = 3.0$ min).
3. Add one-tenth volume of 250 mM DTT, mix gently, add one-tenth volume of gel retardation loading buffer, mix gently and load onto a 2-mm-thick retardation gel in 1× TBE (or alternative buffer) and run as optimised for analytical gels. However, the load should be spread over up to ten times more well area (*see* **Note 5**).
4. After electrophoresis, separate the glass plates carefully so that the gel adheres to one plate and cover the gel with cling film. Expose to X-ray film long enough to reveal complexes

clearly, i.e. 6 h to overnight. The alignment of the film to the gel must be reliably marked.

5. Put the developed film on a light box. Remove the gel from the cling film and re-align it on the X-ray film. Cut pieces of NA45 paper (or equivalent) sufficiently large to cover individual complexes in the gel yet small enough to fit into 1.5-mL tubes when rolled up. Wet the paper pieces in retardation gel running buffer and, with the help of tweezers, position one over each complex of interest in the gel, as visualised from the underlying film. Also position a similar sized piece of NA45 paper over (some of) the uncomplexed DNA. NA45 paper can be labelled with pencil before wetting (*see* **Note 6**).
6. Carefully cover the gel and paper pieces with two sheets of 3MM paper wetted in 1× TBE (or alternative gel running buffer from **step 3**). Lay a scotch bright pad from the electroblotting apparatus on top of the paper and turn the gel over. Carefully remove the second glass plate, cover the other side of the gel with 3MM paper and scotch bright as before, and insert the package into an electro-transfer apparatus as described in the manufacturer's instructions with the NA45 paper towards the anode. Transfer in 1× TBE (or alternative buffer) at 80 V for 1.5 h (*see* **Note 2**).
7. Stop the transfer, unpack the gel carefully with the NA45 paper on top and transfer each piece to a labelled 1.5-mL tube containing 600 μL of elution buffer (check that the radioactivity has transferred to the paper). Incubate at 70°C for 1 h.
8. Remove NA45 paper from each tube, check that the radioactivity has eluted into the buffer (do not expect quantitative elution but at least 50% should come off) (*see* **Note 3**), add 20 μg carrier DNA, extract with phenol/chloroform and precipitate with 1 volume of isopropanol (*see* **Note 4**). Wash the precipitate once in 70% ethanol and dry under vacuum.
9. Dilute piperidine 1:10 in water and add 50 μL to each pellet (*see* **Note 11**). Vortex briefly and incubate at 90°C for 30 min (tubes must be clamped or weighted down to prevent the lids opening) then dry under vacuum. Take up the samples in 100 μL of water and repeat the drying process.
10. Measure the Cerenkov counts in each tube, then re-dissolve the samples in water (e.g. 10 cpm/μL) and transfer equivalent counts (1,000 cpm in each case would be optimal (*see* **Note 7**)) from each into fresh tubes. Dry down and re-dissolve in 5 μL sequencing loading buffer (*see* **Note 8**).
11. Prepare and pre-electrophorese a standard sequencing gel (5–12% acrylamide, depending on probe length) (*see* **Note 9**). Denature probes at 95°C for 5 min, snap cool in ice and

load onto the gel. Run the gel until optimal separation of sequence is achieved (*see* **Note 10**).

12. Stop electrophoresis, remove the gel from the tank and lift off one glass plate. Fix the gel in 20% ethanol, 10% acetic acid for 10 min. Drain briefly and then overlay the gel with two sheets of 3MM paper and carefully peel it off the glass plate. Cover the gel surface with cling film and dry on vacuum gel drier (*see* **Note 12**). Expose the dry gel to X-ray film with intensifying screens as necessary, or to imaging plate.

3.2. Methylation Interference

1. Mix 300,000 cpm of end-labelled probe, 100 μL of 2× DMS buffer and water to 200 μL. Add 2 μL of DMS and incubate at RT (the same guidelines as given in **step 2** of **Subheading 3.1**. apply for the reaction time). Stop the reaction by the addition of 50 μL cold DMS stop mix and precipitate with 850 μL cold ethanol. Re-dissolve in 200 μL cold 0.3 M NaAc pH 7.0, add 700 μL cold ethanol and re-precipitate. Wash twice in 80% ethanol and re-dissolve the probe in water or binding buffer (about 20,000 cpm/μL).
2. Incubate the probe with protein for gel retardation as previously optimised for gel retardation analyses of the complexes in question (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions") in a total volume of 100 μL.
3. Add one-tenth volume of gel retardation loading buffer, mix gently and load onto a 2-mm-thick retardation gel in 1× TBE (or alternative buffer) and run as optimised for analytical gels. However, the load should be spread over up to ten times more well area (*see* **Note 5)**.
4. Continue with **step 4** and all subsequent steps as described for methylation protection (**Subheading 3.1**).

4. Notes

1. To have sufficient counts to complete the procedure, proceed with at least ten times the amount of material required for a simple gel retardation analysis, i.e. at least 300,000 cpm.
2. It is also possible to use a semi-dry electro-transfer apparatus (e.g. Bio-Rad Trans-Blot SD) to transfer the DNA from the gel retardation gel onto NA45 paper. In this case the transfer time and potential are both reduced.
3. In some instances it may prove difficult to elute the DNA from the NA45 paper, in which case raising the salt concentration

or the temperature may improve elution. (Extending the incubation time does not seem to help.) If not, the batch of NA45 may be to blame or it is even conceivable that the DNA–protein complex in question is adsorbed too tightly onto the paper. It is not possible to phenol extract the NA45 paper in order to remove bound protein/DNA.

4. Retain the isopropanol supernatants until you are sure the samples have precipitated quantitatively. Add more carrier DNAs if required.
5. A common difficulty with these methods is the persistence of contaminants that accompany DNA after the preparative retardation gel. These contaminants interfere with the migration of DNA on the sequencing gel producing blurred and distorted patterns. In order to minimise this problem it is worth ascertaining the load limit of the retardation gel so that the protein/DNA complex will not smear but be well resolved and therefore concentrated in the gel before elution.
6. As an alternative to electro-elution from the gel onto NA45 paper, corresponding gel slices can be excised and the DNA recovered with the following protocol *(11)*. Incubate gel slice overnight at 37°C in 0.5 mL 0.5 M ammonium acetate, 0.1% SDS, 1 mM EDTA, 10% methanol, and 25 μg proteinase K. Remove gel fragment (check that radioactivity has eluted into the buffer), extract with 1 volume of phenol/chloroform (1:1) and precipitate DNA with ethanol. Re-dissolve DNA in 200 μL of 1 mM Tris–HCl, 0.1 mM EDTA pH8.0, add 0.5 μg carrier DNA and spermine to 0.5 mM final concentration, incubate on ice for 10 min and recover DNA by centrifugation. Re-dissolve DNA in 1 mM Tris–HCl, 0.1 mM EDTA pH8.0 and 1 M NaOAc, dilute with 3 volumes of 1 mM Tris–HCl, 0.1 mM EDTA pH8.0 and precipitate with ethanol.
7. It is similarly advisable to load as little material onto the sequencing gel as practicable. With the advent of the phosphorimager the lower limit for the sequencing gel is well under 1,000 cpm/lane.
8. Bromophenol blue and, to a lesser extent, xylene cyanol blue have been observed to cause exclusion distortions in the sequencing gels and either can be excluded from the sequencing loading buffer if this occurs in a critical part of the gel.
9. If the end-labelled DNA fragment is relatively long and multiple binding sites are to be resolved, a gradient or wedge sequencing gel can be used in **step 11** of **Subheading 3.1**.
10. An appropriate complement for the final result is to perform the Maxam and Gilbert G + A reactions on the end-labelled

probe. On the sequencing gel these reactions should provide unambiguous sequence information and, in case difficulties are encountered, clues as to the steps that are problematic.

11. The strand scission protocol described here should not convert methylated A residues into strand breaks. It is often observed, however, that breakages at As occur with reasonable efficiency. The following modification will produce efficient cleavage at both G and A residues. After the preparative retardation gel, re-suspend the dried, purified DNA in 30 μL of 10 mM sodium phosphate pH 6.8, 1 mM EDTA. Incubate for 15 min at 92°C. Then add 3 μL of 1 M NaOH and incubate for 30 min at 92°C, followed by 320 μL of 500 mM NaCl, 50 μg/mL of carrier DNA and 900 μL of ethanol. Chill and centrifuge to pellet the radioactivity. Wash the pellet in 70% ethanol and dry. Proceed as above from **step 10**.
12. It is not essential to dry down the sequencing gel as, after one glass plate has been removed, it can be covered with cling film and exposed to X-ray film at –70°C with one screen. This alternative should only be considered if the signal is sufficiently strong or if a gel drier is not available.

References

1. Church, G.M., and Gilbert, W. (1984). Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–1995.
2. Maxam, A., and Gilbert, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
3. Fried, A., and Crothers, D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**, 6505–6525.
4. Garner, M.M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of protein to specific DNA regions: application to components of the E. coli lactose operon regulatory system. *Nucleic Acids Res* **9**, 3047–3059.
5. Gilbert, W., Maxam, A., and Mirzabekov, A. (1976). Contacts between the LAC repressor and DNA revealed by methylation. In: Control of Ribosome Biosynthesis, Alfred Benzon Symposium IX (eds. N.O. Kjelgaard and O. Maaloe), pp. 139–148. Academic, New York, NY.
6. Johnsrud, L. (1978). Contacts between Escherichia coli RNA polymerase and a lac operon promoter. *Proc. Natl Acad. Sci. USA* **75**, 5314–5318.
7. Herrera, R.E., Shaw, P.E., and Nordheim, A. (1989). Occupation of the c-fos serum response element in vivo by a multi-protein complex is unaltered by growth factor induction. *Nature* **340**, 68–70.
8. Siebenlist, U., and Gilbert, W. (1980). Contacts between E. Coli RNA polymerase and an early promoter of phage T7. *Proc. Natl Acad. Sci. USA* 77, 122–126.
9. Shaw, P.E., Schröter, H., and Nordheim, A. (1989). The ability of a ternary complex to form over the serum response element correlates with serum inducibility of the human c-fos promoter. *Cell* **56**, 563–572.
10. Shaw, P.E., and Saxton, J. (2003). Ternary complex factors: prime nuclear targets for mitogen-activated protein kinases. *Int. J. Biochem. Cell Biol* **35**, 1210–1226.
11. Treisman, R. (1986). Identification of a protein binding site that mediates transcriptional response of the c-fos gene to serum factors. *Cell* **46**, 567–574.

Chapter 9

Ethylation Interference Footprinting of DNA–Protein Complexes

Iain W. Manfield and Peter G. Stockley

Summary

Structural studies of DNA–protein complexes reveal networks of contacts between proteins and the phosphates, sugars and bases of DNA. A range of biochemical methods, termed chemical footprinting, aim to determine the functional groups on DNA which are protected in solution by bound protein against modification or where chemical pre-modification interferes with subsequent protein binding. One of these approaches, termed ethylation interference footprinting, reveals which backbone phosphate groups are contacted by protein and the positions where the DNA–protein interface is so tight that the modification cannot be accommodated. This chapter describes the steps necessary to perform an ethylation interference experiment, including modification of DNA using ethylnitrosourea, fractionation of the products based on their affinities for a DNA-binding protein and analysis of the "bound" and "free" fractions to reveal sites critical for complex formation. This is illustrated using results from our experiments with the *Escherichia coli* methionine repressor, MetJ.

Key words: Ethylation interference, Ethylnitrosourea, Footprinting, MetJ, Methionine repressor, DNA–protein interactions.

1. Introduction

Structural studies of double-stranded DNA–protein complexes have shown that specific sequence recognition is accomplished in two ways, either directly by the formation of hydrogen bonds to base-pair edges from amino acid side chains located on a DNA-binding motif, such as a helix–turn–helix, or indirectly as a result of sequence-dependent distortions of the DNA conformation *(1)*, and often by a combination of these two types of interaction. These contacts occur in the context of oriented complexes between

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_9

macromolecules that juxtapose the specific recognition elements. As part of these processes, proteins make a large number of contacts to the phosphodiester backbone of DNA, as was predicted from biochemical assays of the ionic strength dependence of DNA binding.

Contacts to phosphate groups can be inferred by the ethylation interference technique *(2)*. Ethylnitrosourea (EtNU) reacts with DNA to form, principally, phosphotriester groups at the non-esterified oxygens of the otherwise phosphodiester backbone. Minor products are the result of the reactions of EtNU with oxygen atoms in the nucleotide bases themselves (*see* **Note 1**). Under alkaline conditions and at high temperature, the backbone can be cleaved at the sites of modification to form a population of molecules carrying either 3′-OH or 3′ ethylphosphate groups.

The length of the ethyl group (~4.5 Å) means that at a number of positions along a DNA molecule encompassing the binding site for a protein, complex formation will be inhibited by the presence of such a modification. At other modified sites, outside the binding site, no interference with protein binding will be observed. Addition of the DNA-binding protein to a randomly ethylated DNA sample, followed by some procedure to separate the complexes formed from unbound DNA, fractionates the DNA into those molecules still able to bind protein with high affinity and those for which ethylation has lowered the affinity (**Fig. 1**). In practice, modification at different sites produces molecules with a spectrum of affinities for the protein. It is therefore not possible to prove conclusively that a particular phosphate is contacted by the protein, but only that ethylation at that site interferes with complex formation.

Only occasionally are large amounts of highly purified protein readily available for in vitro biochemical assay of DNA-binding activity. Often only small amounts of crude nuclear extracts are available. In many commonly used assays, complex formation could not easily be detected in such situations. For example, using DNase I or hydroxyl radical footprinting, a high level of binding site occupancy is required for a footprint to be detected. Fractional occupancy is readily detected by gel retardation of complexes but offers only limited characterisation of the details of the protein–DNA interaction. Interference techniques, such as the ethylation and hydroxyl radical interference techniques *(3)*, do allow the molecular details of complex formation to be studied even when only small amounts of crude protein are available *(4)*. Whatever the level of saturation, DNA fragments modified at sites reducing the affinity of protein for DNA are less likely to form complexes. The bound fraction on gel retardation assays will therefore always give an indication of the sites that do not interfere with complex formation when modified. The groups on the DNA recognised by the protein can then be inferred. Ethylation can also be used to analyse RNA–protein complexes, e.g. *(5)*.

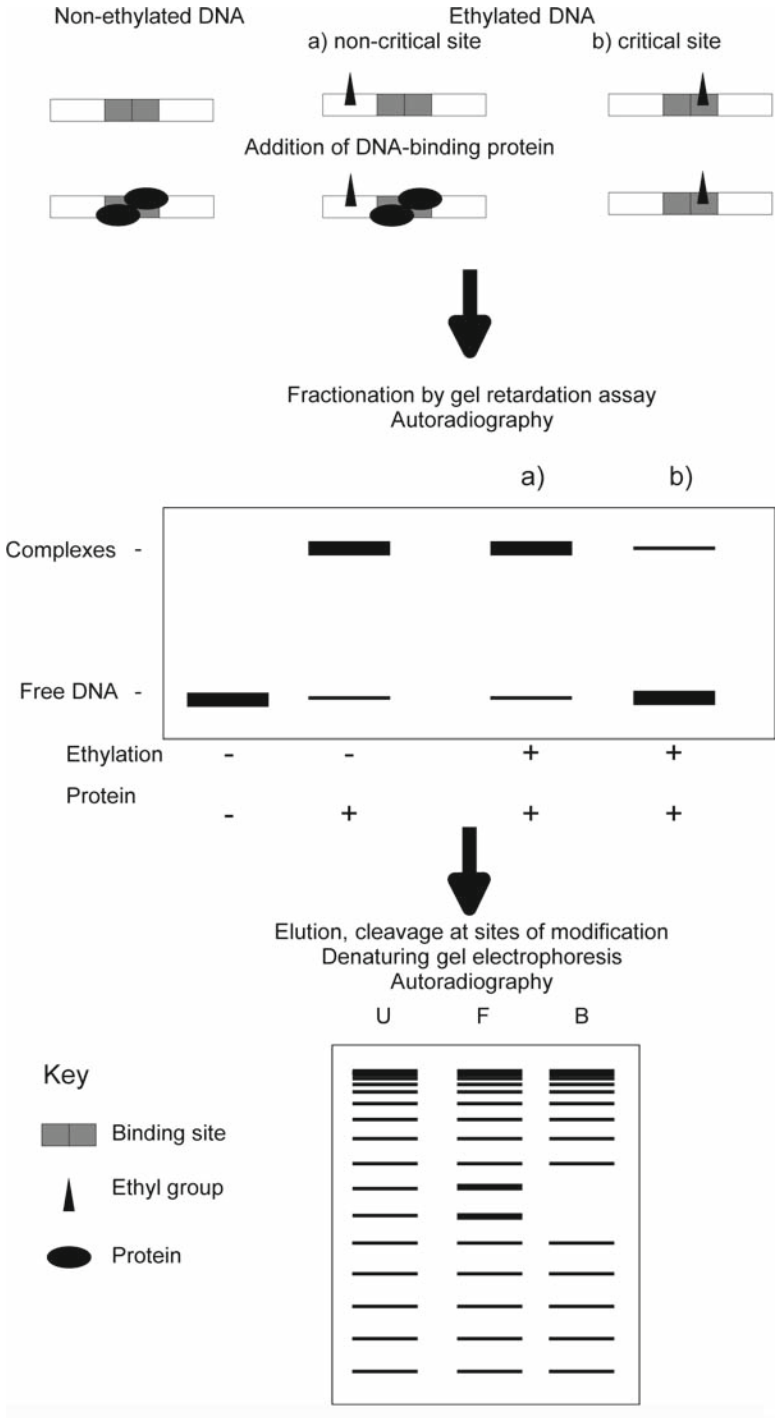


Fig. 1. Diagrammatic outline of the ethylation interference experiment. The *upper* section shows an end-labelled DNA randomly ethylated (ethyl groups represented by *triangles*) at sites outside or within a protein binding site (*grey boxes*), and subsequent complex formation with the protein (represented by *ovals*). The expected mobility of each species on non-denaturing gels is shown in the *middle* section. The pattern of cleavage products that might be expected after recovery of the DNA fragments (assuming random ethylated throughout their length) from the gel retardation assay is shown in the *lower* section. The distribution of products from unfractionated, ethylated DNA is represented by "U" and the products from "free" and "bound" fractions in gel shift assays by "F" and "B". Sites critical for complex formation are under-represented (or *blank*) in the "bound" fraction and these sites are correspondingly over-represented in the "free" fraction. An example of the final autoradiograph for a MetJ:operator complex is shown in Fig. 2.

We have used the ethylation interference technique to probe a number of DNA–protein complexes *(6, 7)* and illustrate the technique here with results from experiments probing the interaction of the *E. coli* methionine repressor, MetJ, with its binding site in vitro *(7)*. Binding sites for MetJ consist of two or more immediately adjacent copies of an 8-bp site with the consensus sequence 5′-dAGACGTCT-3′, termed a "met box." X-ray crystallography has been used to determine the structure of the MetJ dimer, its complex with the corepressor, *S*-adenosyl methionine (AdoMet), and a complex of the holorepressor with a 19-mer oligonucleotide containing two consensus met boxes *(8, 9)* (*see* Chapter "Filter-Binding Assays"). The structure of the protein–DNA complex in the crystal reveals two MetJ dimers (one per met box), binding to the DNA by insertion of a β-ribbon into the major groove. The general features of this model have been corroborated by the results of ethylation interference experiments and by the data from a range of other footprinting techniques *(7)*.

2. Materials

2.1. Preparation of Radioactively End-Labelled DNA

1. Plasmid DNA carrying the binding site for a DNA-binding protein on a convenient restriction fragment (usually <200 bp).
2. Restriction enzymes and the appropriate buffers as recommended by the suppliers.
3. Phenol: Redistilled phenol equilibrated with 100 mM Tris–HCl, pH 8.0. Similar products can be purchased (e.g. Sigma, P4557).
4. Chloroform.
5. Solutions for ethanol precipitation of DNA: 4 M NaCl and ethanol (absolute and 70% v/v).
6. Alkaline phosphatase (AP), from calf intestine or shrimp.
7. AP reaction buffer, 10×: 0.5 M Tris–HCl, pH 9.0, 0.1 M $MgCl_2$, 0.001 M $ZnCl_2$.
8. TE buffer: 10 mM Tris–HCl, pH 8.0, 1 mM ethylenediaminetetra-acetic acid (EDTA).
9. Sodium dodecyl sulfate (SDS) 20% w/v.
10. Ethylenediaminetetra-acetic acid (EDTA), 0.25 M, pH 8.0.
11. T_4 polynucleotide kinase (T_4PNK).
12. T_4PNK reaction buffer, 10×: 0.5 M Tris–HCl, pH 7.6, 0.1 M $MgCl_2$, 0.05 M dithiothreitol.
13. Radioisotope: γ-[^{32}P]-ATP, typical specific activity >6,000 Ci/mmol. Use appropriate measures for protection against this radiation hazard.

14. 30% w/v acrylamide stock (29:1 acrylamide:*N*,*N*′-methylene-*bis*-acrylamide).
15. TBE buffer, 5×: 450 mM Tris-borate, 5 mM EDTA, pH 8.3.
16. Polyacrylamide gel elution buffer: 0.3 M sodium acetate, 0.2% w/v SDS, 2 mM EDTA.
17. Polymerisation catalysts: Ammonium persulfate (10% w/v) and *N*,*N*,*N*′,*N*′-tetramethylethylene diamine (TEMED).

2.2. Ethylation Modification

1. End-labelled DNA in TE: 250,000 cpm are required per ethylation reaction (roughly 20 ng DNA fragment) to be performed. Ethylation of more DNA allows a range of protein concentrations to be used when protein–DNA binding reactions are prepared. Standard safety procedures should be used when handling radiolabelled DNA (e.g. work behind Lucite shields).
2. Cacodylate buffer: 50 mM sodium cacodylate (used without adjustment of the pH, which is usually close to 8.0). Cacodylate buffer contains arsenic and therefore should be handled with caution. It is only necessary to prepare small volumes (ca. 10 mL) of this solution. Passing the solution through a 0.45-μm filter is the preferred method of sterilisation.
3. Ethylnitrosourea (EtNU): This reagent is readily synthesized *(10)* or can be purchased from a commercial source (e.g. Sigma, St. Louis, MO, listed as *N*-nitroso-*N*-ethylurea, cat. no. N 8509). The solid material should be stored at –20°C and allowed to warm to room temperature before use. EtNU should be handled in a fume hood, and contaminated waste stored there until disposal. Wear two pairs of latex gloves when handling samples containing EtNU.
4. tRNA (1 mg/mL).
5. Solutions for ethanol precipitation of DNA: 4 M NaCl and ethanol (absolute and 70% v/v).

2.3. Fractionation of Modified DNA

1. Complex buffer: for MetJ; 10 mM Tris–HCl, pH 7.4, 150 mM NaCl, 1 mM *S*-adenosyl methionine *p*-toluene sulfonate salt (AdoMet). Add glycerol to this buffer to 10% v/v.
2. Purified DNA-binding protein or protein extract.
3. Non-denaturing gel acrylamide stock solution: 30% w/v (37:1 acrylamide:*N*,*N*′-methylene-*bis*-acrylamide).
4. Electrophoresis buffer stock solution: 1.0 M Tris–HCl, pH 8.0, solid AdoMet.
5. Ammonium persulfate, APS (10% w/v).
6. TEMED.
7. Glass plates: 150 × 150 × 1.5 mm.

8. Peristaltic pump capable of recirculating buffer at 5–10 mL/min.

2.4. DNA Recovery

1. X-ray film.
2. Autoradiography cassette.
3. X-ray film developer and fixer.
4. Plastic wrap.
5. Scalpel blade.
6. Syringe needle.
7. Polyacrylamide gel elution buffer: 0.3 M sodium acetate, 0. 1% w/v SDS, 1 mM EDTA.
8. Solutions for ethanol precipitation of DNA: 4 M NaCl; ethanol (absolute and 70% v/v) and tRNA (1 mg/mL).

2.5. Phosphotriester Cleavage and Denaturing Gel Electrophoresis

1. Sodium phosphate, 10 mM pH 7.0, 1 mM EDTA.
2. 1.0 M NaOH (freshly prepared).
3. Acetic acid solutions: 1.0 M and 10% v/v.
4. Solutions for ethanol precipitation of DNA: 4 M NaCl; ethanol (absolute and 70% v/v) and tRNA (1 mg/mL).
5. Sequencing gel loading buffer: 80% v/v formamide, 0.5× TBE, 0.1% w/v xylene cyanol, 0.1% w/v bromophenol blue.
6. Acrylamide stock solutions for sequencing gel: 19% w/v acrylamide, 1% w/v *N*,*N*′-methylene-*bis*-acrylamide, 50% w/v urea in TBE.
7. TBE (1×): 90 mM Tris, 90 mM boric acid, 1 mM EDTA, pH 8.3.

2.6. Maxam–Gilbert Chemical DNA Sequencing Reactions

There is insufficient space to cover these methods in detail here, but extensive information, materials, methods, and trouble-shooting guides are readily available in published literature *(11, 12)*.

3. Methods

3.1. Preparation of End-Labelled DNA

1. Digest the plasmid (e.g. 10 μg of the plasmid in a reaction volume of 200 μL) with one of the pair of restriction enzymes used to release a suitably sized DNA fragment (usually <200 bp). Extract the digest with an equal volume of buffered phenol, and then add one-tenth volume NaCl (4 M) and 2.5 volumes of ethanol to the aqueous layer to precipitate the DNA.
2. Add 50 μL 1× AP reaction buffer to the ethanol-precipitated DNA pellet (<50 μg). Add 1 U AP, incubate at 37°C for 30

min, followed by addition of a further aliquot of enzyme, and incubate for a further 30 min. Terminate the reaction by adding SDS and EDTA to 0.1% and 20 mM, respectively, in a final volume of 200 μL, and incubate at 65°C for 15 min. Extract the digest with an equal volume of buffered phenol, then with 1:1 phenol:chloroform, and finally ethanol precipitate the DNA from the aqueous phase by addition of 2.5 volumes of ethanol.

3. Redissolve the DNA pellet (~2.5 μg) in 18 μL 1× T_4PNK buffer. Add 20 μCi γ-[^{32}P]ATP and 10 U T_4PNK, and incubate at 37°C for 30 min. Terminate the reaction by phenol extraction (followed by ethanol precipitation and a second restriction enzyme digest) or by addition of non-denaturing gel loading buffer, and electrophoresis on a non-denaturing polyacrylamide gel. We use 12% (w/v) polyacrylamide gels (19:1, acrylamide:*bis*-acrylamide) with 1× TBE as the electrophoresis buffer.
4. After electrophoresis, locate the required DNA fragments by autoradiography of the wet gel for around 30 min (as described in **Subheading 3.4**). Excise slices of the gel containing the bands of interest using the autoradiograph as a guide. Elute the DNA into 500 μL elution buffer overnight (at least) at 37°C. Ethanol precipitate the DNA, wash the pellet thoroughly with 70% v/v ethanol, dry under vacuum, and rehydrate in a small volume of TE buffer (e.g. 50 μL). Determine the radioactivity of the sample by liquid scintillation counting of a 1 μL aliquot.

3.2. Ethylation Reaction

1. Dispense the required volume containing ~250,000 cpm of radiolabelled DNA solution into an Eppendorf tube, add cacodylate buffer to a final volume of 100 μL, and heat to 50°C in a heating block. Prepare the minimum volume of EtNU-saturated ethanol (at 50°C) for the required reactions, add 100 μL of this to the DNA, mix, and incubate at 50°C for 60 min (*see* **Notes 1** and **2**).
2. Add to the sample 5 μL of 4 M NaCl, 2 μg of tRNA, and 150 μL of ethanol. Mix and place at –20°C for 60 min or in a dry ice/ethanol bath for 15 min. Pellet the DNA by centrifugation in a microfuge (15 min at 13,000 × *g* and 4°C) and remove the supernatant (store separately to be destroyed by incineration). Add 500 μL of 70% ethanol, mix thoroughly, recentrifuge, and remove the supernatant. Dry the pellet briefly under vacuum.

3.3. Fractionation of DNA by Gel Retardation Assay

The following procedure allows the separation of free and bound DNA by means of gel retardation. The precise conditions will depend on the protein under investigation (*see* **Notes 3**–**7**).

1. Mix 10 mL of non-denaturing acrylamide stock solution, 1.5 mL of 1.0 M Tris–HCl, pH 8.0, 29.5 mL of distilled water, 0.2 mL of APS, 15 μL of TEMED, and 1.5 mg of AdoMet (for the MetJ example), and pour into gel frame. Insert the well former and leave to polymerise for 1–2 h.
2. When polymerised, insert the gel into the tank and connect peristaltic pump tubing such that buffer is pumped in both directions (i.e. top to bottom and bottom to top reservoirs). Pre-electrophorese gel for 30 min at 100 V.
3. Redissolve the pellet of ethylated DNA in complex buffer plus glycerol (*see* Notes 4 and 5). Set aside 10% of the ethylated DNA sample, which will be used as an unfractionated control to indicate the variation in modification at each residue. Dry this sample under vacuum and store at –20°C until **step 1** of **Subheading 3.5**. Add DNA-binding protein to a concentration that would saturate unmodified DNA and incubate at 37°C for 15 min to allow complex formation (N.B. exact conditions will vary depending on the protein being studied).
4. Load the DNA–protein complex solution onto the gel and electrophorese into the gel at 250 V for 2 min. Reduce the voltage to 100 V and continue electrophoresis until a small amount of bromophenol blue dye loaded into an unused lane has reached the bottom of the gel.

3.4. DNA Recovery

1. After electrophoresis, remove one plate, and wrap the gel and remaining plate in clear plastic wrap film (we use Saran Wrap™). Cut a piece of X-ray film large enough to cover the lanes used on the gel and fix it to the gel firmly with masking tape. Using a syringe needle, make a number of holes through the film and gel that will serve to orient the developed film with respect to the gel.
2. Place the assembly in an autoradiography cassette and leave at 4°C overnight.
3. Take the film off the gel and develop as usual. Align the film and gel using the holes created previously. Using a syringe needle, make a series of holes into the gel around the fragments of interest using the bands on the film as a guide. Remove the film and excise the marked regions of polyacrylamide. Place gel fragments (~10 × 5 mm) in Eppendorf tubes, add 600 μL of gel elution buffer, and incubate at 37°C overnight (at least).
4. Transfer 400 μL of the eluate to a fresh tube, add 2 μg of tRNA and 1 mL of ethanol, mix, and place at –20°C for 60 min. Pellet DNA by centrifugation in a microfuge for 15 min (13,000 × *g*, 4°C). Discard the supernatant (check for absence of radioactivity), wash the pellet with 500 μL of 70% ethanol, recentrifuge, and discard the supernatant. Dry the pellet briefly

under vacuum. If the DNA does not pellet readily, incubate the sample at –20°C for longer or recentrifuge at 4°C.

3.5. Phosphotriester Cleavage and DNA Sequencing

1. Redissolve each pellet in 15 μL of sodium phosphate buffer and add 2.5 μL of 1 M NaOH. Seal tube with plastic film (e.g. Parafilm) and incubate at 90°C for 30 min. Centrifuge samples briefly to collect any condensation. Add 2.5 μL of 1.0 M acetic acid, 2 μL of 4 M NaCl, 1 μg of tRNA, and 70 μL of ethanol. Leave samples at –20°C for 60 min. Pellet DNA as described in **step 4** of Subheading 3.4.
2. Redissolve the pellet in 4 μL of sequencing gel loading buffer. Heat to 90°C for 2 min and load samples onto a 12% w/v polyacrylamide sequencing gel alongside Maxam and Gilbert sequencing reaction markers. Electrophorese at a voltage that will warm the plates to around 50°C. After electrophoresis, fix the gel in 1 L of 10% (v/v) acetic acid for 15 min. Transfer the gel to 3MM paper and dry under vacuum at 80°C for 60 min. Autoradiograph the gel at –70°C with an intensifying screen.
3. Compare lanes corresponding to bound, free and control DNA for differences in intensity of bands at each position; this is best done using densitometry (*see* **Note 8**). A dark band in the "free fraction" (and a corresponding reduction in the intensity of the band in the "bound fraction") indicates a site where ethylation interferes with complex formation. This is interpreted as meaning that this residue is contacted by the protein or a portion of the protein comes close to the DNA at this point. For 5′ end-labelled DNA, the ethylation reaction products migrate slightly more slowly than the Maxam–Gilbert chemical sequencing products (*see* **Notes 9 & 10**, and also **Fig. 2**).

3.6. Results and Discussion

The result of an ethylation interference experiment with MetJ is shown in **Fig. 2** along with densitometer traces showing quantitative comparisons of the distribution of products in bound and free fractions (**Fig. 3**). Visual examination of the autoradiograph shows that ethylation at 5′-pG2 results in total exclusion of such fragments from protein–DNA complexes. Densitometry indicates that ethylation at other sites also inhibits complex formation to varying degrees. These data can be interpreted in terms of the MetJ–DNA crystal structure *(9)*. The site at which ethylation completely inhibits complex formation corresponds to the phosphate 5′ to the guanine at position 2 of each met box. The crystal structure shows a contact to this phosphate from the N-terminus of the repressor B-helix. Indeed, at the centre of the two met box operator complex, sequence-dependent distortions of the oligonucleotide away from the canonical B-DNA structure result in displacement of this 5′ G2 phosphate by up to 2 Å in the

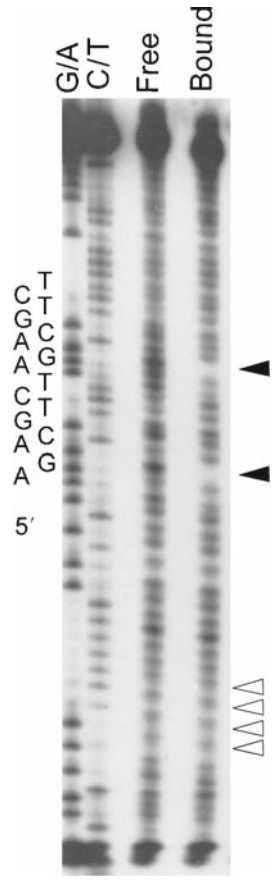


Fig. 2. Ethylation interference of the MetJ:DNA interaction. Samples for denaturing gel electrophoresis were prepared following the methods given here. "G/A" and "C/T" are the products of the purine- and pyrimidine-specific Maxam–Gilbert chemical cleavage reactions, respectively. "Free" and "bound" are the DNA fractions that did not and did form complexes with MetJ, respectively. The sequence of the MetJ binding site is indicated alongside the autoradiograph. Electrophoresis is from *top* to *bottom*. The phosphates, ethylation of which interferes most strongly with complex formation, are indicated by *solid triangles*. Small, *open triangles* indicate the minor reaction products of the cleavage reaction, which, for small fragments, are resolved on these gels.

direction of the protein. Variant oligonucleotides with less tendency to undergo this distortion are bound significantly less well than the consensus site *(13)*. Thus, this site of complete inhibition of complex formation corresponds to an important contact between the DNA and a secondary structural element in the protein, which presumably is unable to adjust to the presence of a bulky ethyl group. The other sites of ethylation interference effects are contacted by amino acid side chains and peptide backbone groups in extended loops of the repressor. It might be expected that side chains and loops can be flexible enough to reorient in order to reduce steric hindrance between the protein and the ethyl group, thus explaining these partial interference effects. Similar good

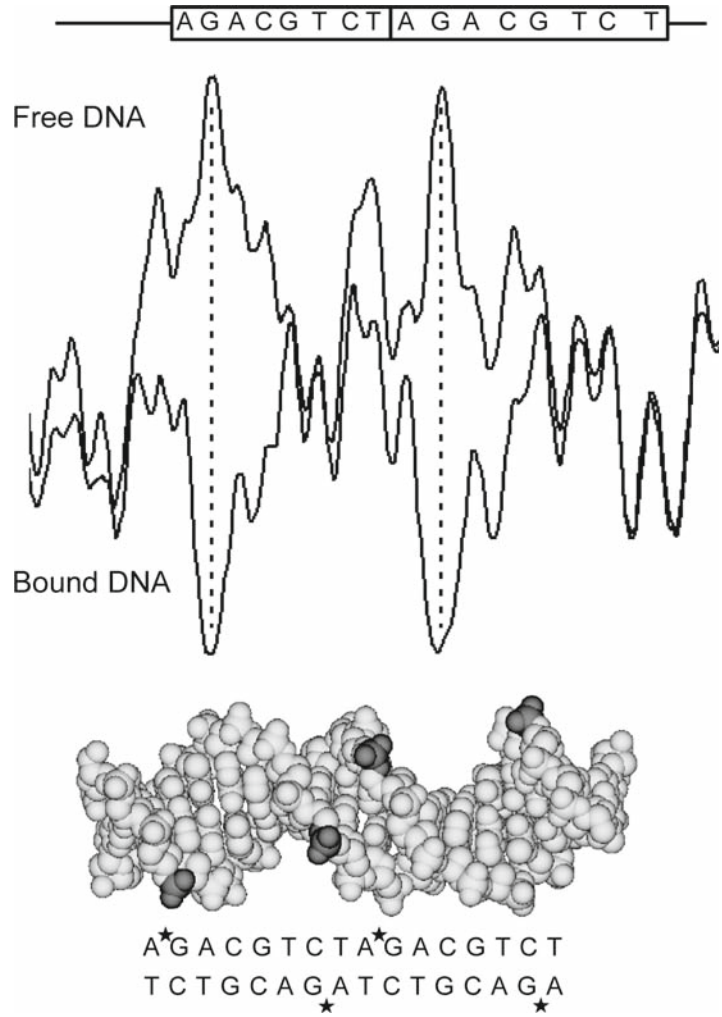


Fig. 3. Summary of the results. *Top:* Densitometer traces of ethylated DNA from the free and bound fractions. DNA from the free and bound fractions is represented by the *upper* and *lower curves*, respectively. The sequence of the *top* strand of each binding site is indicated above the *curves*. Note that the greatest inhibitory effect is at the second residue (G2) of each *box* and around five sites in each *box* show an interference effect. *Bottom:* Space filling representation of the operator site showing the positions where ethylation results in complete inhibition of complex formation (highlighted in *dark grey*). MetJ dimers bind in the major groove with contacts between adjacent protein dimers above the minor groove at the symmetry dyad of the oligonucleotide shown.

correlations between the contacts identified by ethylation interference experiments and those seen in crystals have been demonstrated in other systems, such as phage 434 repressor *(14)*.

MetJ shows a high level of sequence specificity and therefore clear interference effects are observed. However, we have observed that other proteins with low-sequence specificity show only relatively weak interference effects *(6)*. Densitometry and quantitation may be necessary in these cases to resolve these

effects from the background of cleavage from fragments shifted by non-specifically bound proteins (*see* **Note 8**). Additionally, proteins binding to DNA asymmetrically, for example to sites which do not show the dyad symmetry common to prokaryotic repressor proteins, will show different interference effects on each strand, again providing valuable information on the structure of the protein–DNA complex in solution *(6)*.

4. Notes

1. *Modification at secondary (non-phosphate) sites.* Early work on the reaction of alkylating agents with DNA showed that a number of products are obtained. With EtNU, phosphotriester groups comprise 60–65% of the reaction products *(10)*. The remaining products are the result of reactions at base oxygen groups with relative abundance in the order thymine O2 = guanine O6 > thymine O4 >> cytosine O2 for double-stranded DNA *(15)*. To our knowledge, the effects of such modifications on DNA binding by proteins have not been addressed in the literature on ethylation interference experiments. However, some prediction of the effects might be made based on knowledge of the structure of DNA. Thymine O2 and cytosine O2 are in the minor groove, and guanine O6 and thymine O4 are in the major groove. For sequence-specific DNA-binding proteins interacting with DNA via the major groove, it would be expected that the presence of an ethyl group would inhibit complex formation, but that a modification in the minor groove would be less inhibitory.
2. *Level of modification of DNA by EtNU.* A similar intensity of each band (subject to the position-dependent variation of reactivity observed with EtNU) is the required level of modification. We have used a single batch of "home-produced" EtNU for all our ethylation interference experiments. The conditions of time and temperature used with this batch give an appropriate level of probe modification. There may be variations in the reactivity of EtNU from different sources leading to under- or over-modification. The correct modification conditions can readily be determined by performing a test ethylation on a small amount of DNA, such as 20,000 cpm, followed by alkaline cleavage (omitting the gel fractionation step), sequencing gel electrophoresis, and autoradiography. Over-modification will produce a bias toward short fragments.
3. *Choice of fractionation method.* The original report of the use of the phosphate ethylation reaction used the filter binding assay to fractionate protein-bound and protein-free DNA *(2)*.

Filter-bound DNA is then eluted by washing the filter in a high salt buffer containing SDS. For other experiments, we find that TE + 0.1% SDS efficiently elutes DNA from nitrocellulose filters. The rapid recovery of DNA from filters is an advantage of using the filter binding assay compared with the gel retardation assay. Despite this, fractionation by gel retardation assay has proven to be by far the most popular method in the literature. The advantage of the gel retardation assay is that the presence and amounts of multiple complexes can be determined and recovered separately, which is not possible by filter binding. Parallel binding reactions with ethylated and unethylated DNA and separation on non-denaturing polyacrylamide gels readily demonstrate any differences in the mobility of complexes formed with each DNA sample.

4. *Cofactor requirements.* High affinity DNA binding in our system (MetJ) is dependent on the presence of *S*-adenosyl methionine at millimolar concentrations. This cofactor is present in binding reactions and is included in the gel mix, but it would be prohibitively expensive to include it in the electrophoresis buffer. This does not seem to affect the results obtained by this technique. Electrophoresis for extended times does deplete the lower region of the gel of corepressor leading to some complex dissociation. Another feature specific to this system is the hydrolysis of the corepressor to presumably inactive products. For this reason, gels were not left to polymerise for more than 2 h.

In other systems where a cofactor is not required or in which the cofactor is cheap and/or stable over extended periods, the conditions used for the gel retardation assay fractionation step should be optimised by a consideration of the specific features of the system under study. For further details, consult other chapters of this book.

5. *Effects of salt precipitates.* During ethanol precipitation of DNA, salt is often also precipitated. This white crystalline precipitate is readily distinguished from the almost clear nucleic acid pellets. The interaction of DNA-binding proteins with their recognition sites is strongly ionic strength-dependent, and therefore the presence of a high concentration of salt following ethanol precipitations will inhibit complex formation in addition to any ethylation interference effects. After the cleavage reaction at modified sites, another ethanol precipitation is performed. The presence of a large amount of salt at this stage will prevent complete dissolution of the pellet and will interfere with subsequent electrophoresis. A dark background between each band was often observed on autoradiographs. It is believed that this is caused by the presence of salt in the sample. Reprecipitation as described below helps to reduce this problem.

To remove any salt precipitate, the pellet can be dissolved in a small volume of TE (e.g. 100 μL) and precipitated by addition of 2 volumes of ethanol without addition of further amounts of salt. The DNA can be pelleted as described in **step 4** of **Subheading 3.4**.

6. Recommended controls for the gel retardation assay fractionation step: With a pre-modification reaction such as this, it is important to perform appropriate control reactions especially for the gel retardation step. For such binding reactions, we use 20,000 cpm unmodified DNA in the presence and absence of MetJ at the concentration used in the binding reaction with ethylated DNA samples. The specific activity of protein samples may vary from batch to batch. The control reaction outlined above will confirm that the protein sample used is active for DNA binding. In gel retardation assays, the mobility of complexes is a function of a number of properties of the system, such as charge and molecular weight of the protein, stoichiometry of the complex, and bending of the DNA induced by binding of the protein. Demonstration that the mobility of complexes formed with ethylated and unethylated DNA is the same is probably good evidence that there are no significant differences between the complexes.
7. The protein concentration used in the fractionation step dictates how many interfering sites are reported. Because the ethyl groups at different sites affect protein binding to differing degrees, increasing the protein concentration can mask any weak interference effects such that only the most strongly interfering sites will be detected. The control reaction with unmodified DNA will show the level of binding site saturation, and therefore indicate the level of discrimination between strongly and weakly interfering sites that can be expected. Using a range of protein concentrations in the binding reaction with aliquots of the ethylated DNA should therefore allow the strength of the inhibitory effect at each site to be placed in rank order. This is valuable structural information, since it might be expected that the strongest effects will be observed at sites that are in closest contact to the protein in the complex.
8. Quantitative data can be obtained by exposing dried sequencing gels to storage phosphor screens and analysing images using the software available with the phosphorimager. Alternatively, densitometry of images can be performed by scanning autoradiographs and saving the image as a TIFF file for analysis using software such as the ImageJ program. In both cases a 1D line scan gives images such as that shown in **Fig. 3** allowing comparison of peak heights from bound, free and unfractionated samples. In practice, the short distance between

bands corresponding to each fragment makes further quantitation, of peak areas, impractical. However, the software used should be able to provide the *x*, *y* coordinates describing each curve and these can be used in spreadsheet software to produce presentation quality images.

9. The presence of multiple cleavage products at each phosphate: The products of the cleavage reaction at phosphotriester groups carry either 3′-OH or 3′-ethylphosphate groups. For large fragments on low percentage polyacrylamide gels, these two species are not resolved. However, for short fragments on high percentage gels, two bands are observed at each residue. In practice, this does not produce problems with data analysis.
10. *Troubleshooting*. We have experienced few problems with this technique. Most problems have been associated with the specific properties of the proteins we have studied. However, it is possible to envisage a number of potential problems and explanations for these, and remedies are presented here.

Of the available structures of DNA-binding proteins complexed to DNA fragments, there are none in which the protein does not make some contacts to the phosphodiester backbone. Thus, it is expected that because of the size of the ethyl group, an interference effect will always be observed. In the event that no inhibition of complex formation is observed, it should be confirmed that ethylation has occurred by performing a titration of the ethylation reaction as discussed in **Note 2**.

It is possible that an increase in free DNA is observed on the gel retardation assay but that no cleavage products are observed on the sequencing gel, although the full length DNA is present. This could be caused by an error with the buffer used to resuspend the eluted DNA pellet, the NaOH solution used to perform strand scission, or the temperature of the reaction, all of which can be readily checked.

References

1. Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F., and Sigler, P.B. (1988). Crystal-structure of Trp repressor operator complex at atomic resolution. *Nature* **335**, 321–329.
2 Siebenlist, U., and Gilbert, W. (1980). Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *Proc. Natl Acad. Sci. USA* 77, 122–126.
3. Hayes, J.J., and Tullius, T.D. (1989). The missing nucleoside experiment – A new technique to study recognition of DNA by protein. *Biochemistry* **28**, 9521–9527.
4. Alimov, A.P., Langub, M.C., Malluche, H.H., and Koszewski, N.J. (2003). Sp3/Sp1 in the parathyroid gland: Identification of an Sp1 deoxyribonucleic acid element in the parathyroid hormone promoter. *Endocrinology* **144**, 3138–3147.

5. Jossinet, F., Paillart, J.C., Westhof, E., Hermann, T., Skripkin, E., Lodmell, J.S., Ehresmann, C., Ehresmann, B., and Marquet, R. (1999). Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding. *RNA* **5**, 1222–1234.
6. Manfield, I.W., Reynolds, L.A., Gittins, J., and Kneale, G.G. (2000). The DNA-binding domain of the gene regulatory protein AreA extends beyond the minimal zinc-finger region conserved between GATA proteins. *Biochim. Biophys. Acta* **1493**, 325–332.
7. Phillips, S.E., Manfield, I., Parsons, I., Davidson, B.E., Rafferty, J.B., Somers, W.S., Margarita, D., Cohen, G.N., Saint-Girons, I., and Stockley, P.G. (1989). Cooperative tandem binding of Met repressor of *Escherichia coli*. *Nature* **341**, 711–715.
8. Rafferty, J.B., Somers, W.S., Saint-Girons, I., and Phillips, S.E.V. (1989). 3-Dimensional crystal-structures of *Escherichia coli* Met repressor with and without corepressor. *Nature* **341**, 705–710.
9. Somers, W.S., and Phillips, S.E.V. (1992). Crystal-structure of the Met repressor-operator complex at 2.8 Angstrom resolution reveals DNA recognition by beta-strands. *Nature* **359**, 387–393.
10. Jensen, D.E., and Reed, D.J. (1978). Reaction of DNA with alkylating-agents – Quantitation of alkylation by ethylnitrosourea of oxygen and nitrogen sites on poly dA-dT including phosphotriester formation. *Biochemistry* **17**, 5098–5107.
11. Maxam, A.M., and Gilbert, W. (1980). New method for sequencing DNA. *Proc. Natl Acad. Sci. USA* **74**, 560–564.
12. Maxam, A.M., and Gilbert, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
13. He, Y.Y., Garvie, C.W., Elworthy, S., Manfield, I.W., McNally, T., Lawrenson, I.D., Phillips, S.E., and Stockley, P.G. (2002). Structural and functional studies of an intermediate on the pathway to operator binding by *Escherichia coli* MetJ. *J. Mol. Biol.* **320**, 39–53.
14. Bushman, F.D., Anderson, J.E., Harrison, S.C., and Ptashne, M. (1985). Ethylation interference and x-ray crystallography identify similar interactions between 434 repressor and operator. *Nature* **316**, 651–653.
15. Singer, B. (1976). All oxygens in nucleic acids react with carcinogenic ethylating agents. *Nature* **264**, 333–339.

Chapter 10

Site-Directed Cleavage of DNA by Protein–Fe(II) EDTA Conjugates Within Model Chromatin Complexes

David R. Chafin and Jeffrey J. Hayes

Summary

The nucleosome and other chromatin complexes are examples of complicated protein–DNA assemblies that are not easily studied by traditional structural methods. Site-directed cleavage of DNA is a method for mapping the location of interaction of a specific site in a protein such as a linker histone within a large complex such as the nucleosome. In this chapter we describe the application of the site-directed cleavage method, employing linker histones site-specifically modified with the chemical cleavage reagent Fe(II) (EDTA-2-aminoethyl) 2-pyridyl disulfide (ebr). Addition of hydrogen peroxide and a reducing agent to the complex containing the modified protein leads to the production of hydroxyl radicals from the iron center, resulting in cleavage of DNA backbones in the vicinity of the modified residue. The cleavages can then be mapped and ascribed to a particular location within the nucleosome, allowing the binding site of the protein within this structure to be determined.

Key words: Site-directed cleavage, EPD, Cleavage mapping, Linker histones, Nucleosomes, Hydroxyl radical cleavage of DNA.

1. Introduction

The assembly of eukaryotic DNA into nucleosomes and higher-order structures within chromatin allows the efficient use of genomic DNA for nuclear processes such as replication, transcription, recombination, and repair *(1)*. Thus, to completely understand molecular mechanisms associated with these processes, an understanding of the details of the protein–DNA interactions comprising chromatin is required. For example, two

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_10

fundamental questions still lingering concerning chromatin are the location interactions of the core histone tail domains and the site of contact of the globular domain of linker histones with the nucleosome *(2, 3)*. Both the core histone tail domains and the

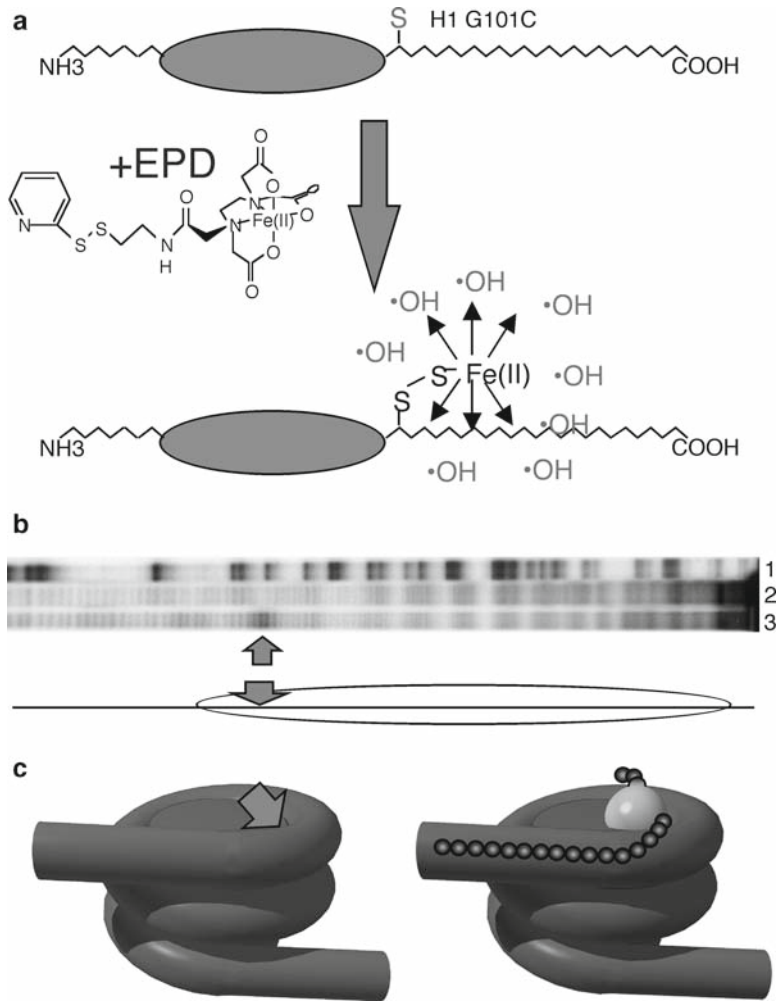


Fig. 1. Directed cleavage of nucleosome DNA by an EPD-modified linker histone. (**a**) Scheme showing site of modification of the linker histone used in this experiment (*see* **ref. 12**) and structure of EPD before the modification reaction (*see* **ref. 9**). (b) Specific cleavages by H1 G101C-EPD within a nucleosome. Nucleosomes were bound by the modified linker histone and cleavage carried out as described and cleavages analyzed via sequencing gel electrophoresis. *Lane 1* shows a G-reaction marker; *lane 2*, cleavage reaction with mock-modified linker histone; *lane 3*, cleavage reaction carried out with the modified linker histone. The *arrows* indicate the location of modification-dependent cleavages. The schematic indicates the location of the nucleosome (*oval*) on the DNA (*black line*). (**c**) The location of the cleavages indicated in (**b**) is indicated on the *left side* of the figure. A plausible model for linker histone binding to the nucleosome is shown on the *right*. The *large sphere* indicates the globular domain as shown by the *oval* in (**a**), while the *smaller spheres* indicate the N- and C-terminal tail domains.

linker histones play crucial roles in the stability and organization of chromatin but participate in poorly understood interactions *(4–7)*.

A useful strategy to map interactions of specific protein domains with DNA involves targeted cleavage of the DNA. This strategy generally involves site-specific attachment of a chemical nuclease to the protein of interest, formation of the complex, followed by localized cleavage of the DNA component *(8–14)*. Of note, the reverse process can also be accomplished with site-directed cleavage of a protein backbone via cleavage reagents tethered to specific sites on DNA *(15)*. Using the former approach, we previously mapped the location of the globular domain of a linker histone when bound to a single nucleosome *(12, 16)*. Histones are in general particularly good candidates for the cysteine-dependent site-directed cleavage approach since most are naturally devoid of cysteine residues. This method, first introduced by the Fox and Ebright groups, takes advantage of the unique nucleophilic properties of the cysteine sulfhydryl, which can be placed at rationally selected locations with the protein of interest. The single sulfhydryl group within these proteins is then modified with a bifunctional reagent that contains a disulfide exchange moiety and an iron(II)-based DNA cleavage reagent (**Fig. 1a**) *(8, 9)*. These proteins are then assembled into the complex of interest and DNA cleavage is initiated by addition of hydrogen peroxide to produce hydroxyl radicals. Hydroxyl radicals are produced by the one-electron reduction of hydrogen peroxide by the tethered iron(II)EDTA complex to produce ·OH and OH^-:

$$[Fe(II)EDTA]^{2-} + H_2O_2 \rightarrow \cdot OH + OH^- + [Fe(III)EDTA]^{1-}$$

The oxidation of Fe(II) to Fe(III) is readily reversed by the relatively innocuous reducing agent sodium ascorbate, while the OH^- is absorbed by the buffer *(17)*. Hydroxyl radicals are highly reactive diffusible species that, upon encountering a DNA molecule, will abstract a hydrogen atom from the deoxyribose unit in the phosphodiester backbone, leaving a carbon-centered radical, which rapidly decays to leave primarily two phosphomonoester termini on either side of a one nucleotide gap at the original location of the attack. Importantly, this cleavage occurs regardless of the identity of the base sequence at the attack position. Thus the location of the cleavage event is easily determined to single nucleotide resolution by monitoring the target DNA for strand breaks on sequencing gels.

2. Materials

2.1. Construction of Cysteine-Substituted Protein

2.1.1. Point Mutation by PCR

1. Oligonucleotide primers: Two primers are complementary to the 5′ and 3′ ends of the coding sequence to be amplified. In addition, if the codon to be altered is located more than ~10–15 nucleotides from the end of the coding sequence one additional primer is needed which must contain the sequence substitutions for the altered codon flanked by 12–15 nucleotides of complementary sequence on each side. Alternatively, a primer pair for the quick change (whole plasmid) PCR can be employed. Store at –20°C.
2. 10× Stock containing all four dNTPs at 10 mM concentration each.
3. A source of clean reliable 18 MΩ water, free of chemical contaminants.
4. 10× PCR buffer can be obtained commercially from the supplier of the PCR enzyme of choice.
5. Vent or Taq DNA polymerase can be obtained from commercial sources.

2.1.2. Ligation and Transformation of PCR Insert into DH5α or BL21 Cells

1. DH5α or BL21 cells can be obtained commercially or prepared in competent form; store at –70°C.
2. Luria broth (LB), sterile.
3. 1,000× stock of ampicillin (100 mg/mL).
4. LB-agar plates containing 0.1 mg/mL ampicillin.

2.1.3. Overexpression and Purification of Mutant Histone Proteins

1. 100× (0.2 M) stock of isopropyl β-D-thiogalactopyranoside (IPTG).
2. Luria broth (LB), sterile.
3. 10 mg/mL lysozyme solution.
4. A saturated room temperature solution of Phenylmethylsulfonylfluoride (PMSF).
5. Triton-X 100 detergent.
6. A 2 M solution of NaCl.
7. A 50% (v/v) slurry of Bio-Rex 50–100 mesh chromatography resin (BioRad).
8. 10 mM Tris–Cl of pH 8.0, 1 mM EDTA solutions containing 0.5, 0.6, 1.0, and 2.0 M NaCl.

2.2. Reduction and Modification of Cysteine-Substituted Proteins with EPD

1. 1 M stock of DTT (Dithiothreitol), made fresh.
2. A 50% slurry of Bio-Rex 100–200 mesh chromatography resin (BioRad).

3. 10 mM Tris–Cl, pH8.0 solutions containing 0.5, 0.6, 1.0, and 2.0 M NaCl.
4. 0.3 M Stock of EPD (synthesized according to *(8, 9)*). Alternatively, iodoacetamido-1,10 phenanthrolineCu^{+2} can be employed in place of EPD (Molecular Probes, Eugene Oregon, USA).
5. Disposable 10-mL plastic chromatography columns (BioRad).

2.3. Radioactive End Labeling of a Purified DNA Restriction Fragment

1. Linear DNA fragment with convenient restriction sites on either end, previously phosphatased.
2. 10× T4 polynucleotide kinase buffer (supplied with enzyme).
3. [γ-^{32}P]dATP 6,000 Ci/mmol.
4. T4 Polynucleotide Kinase: 10,000 units/mL (Promega).
5. 2.5 M ammonium acetate.
6. –20°C 95% Ethanol.
7. –20°C 70% Ethanol.
8. 10% SDS stock solution.
9. Alkaline phosphatase (Roche).

2.4. In Vitro Reconstitution of Nucleosomes

1. 10 mM Tris–Cl of pH 8.0 solutions containing 1 mM EDTA, and 1.2, 1.0, 0.8, and 0.6 M NaCl.
2. TE Buffer.
3. Stock of 6,000–8,000 molecular weight cutoff dialysis tubing.
4. Stock of sonicated CT DNA, approximately 1–2 mg/mL.
5. Stock of 5 M NaCl.
6. Source of purified core histone proteins H2A/H2B and H3/H4, ours are purified from chicken erythrocyte blood or produced recombinantly (*see* ref 18–20 and below).

2.5. Maxam–Gilbert G-Specific Reaction

1. 10× G-specific reaction buffer containing 0.5 M NaCacodylate, 10 mM EDTA.
2. Dimethylsulfate (DMS) (neat).
3. G reaction stop buffer containing 1.5 M NaAcetate, 1 M β-Mercaptoethanol, 0.004 μg/μL sonicated Calf thymus DNA.
4. Piperidine (neat, 10 M stock).

2.6. Chemical Mapping of Protein–DNA Interactions with EPD

1. 0.7% agarose made with 0.5× TBE (Note: treat all solutions with chelex 100 resin (BioRad) to remove adventitious redox-active metals).
2. Histone dilution buffer containing (10 mM Tris–Cl of pH 8.0, 50 mM NaCl).

3. Stock solution of 20 mM sodium ascorbate.
4. Stock solution of 1 mM Fe(II)EDTA.
5. Solution of 0.15% H_2O_2, freshly made.
6. Stop solution containing 50% glycerol, 10 mM EDTA.
7. Microcentrifuge filtration devices (Series 8000 can be obtained from Lida Manufacturing Corporation).
8. Stock solution of 10 mM Tris–Cl of pH 8.0, 0.1% SDS.
9. Microcentrifuge pestles can be obtained from Stratagene.
10. Ice-Cold 95% and 70% ethanol solutions.
11. Stock solution of 3 M sodium acetate.

2.7. Sequencing Gel Analysis

1. Solid urea: molecular biology grade.
2. Stock solution of 5× TBE.
3. Stock solution of 40% acrylamide (19:1 acrylamide:Bis-acrylamide).
4. Stock solution of 20% APS (Ammonium persulfate).
5. Stock solution of TEMED (*N*,*N*,*N*′,*N*′-tetramethylethylenediamine).
6. Formamide-loading buffer (100% Formamide + 0.05% bromophenol blue + xylene cyanol).

3. Methods

3.1. Overexpression and Purification of Single Cysteine-Substituted Proteins

The following methods work well for incorporating a single amino acid substitution into any protein of interest. Standard three or four primer PCR methods are used. Alternatively, the "quick-change" method may be used to swap in a cysteine codon. The protocol also describes generic expression and purification of a histone protein for use in the site-directed cleavage technique. Other purification schemes may be substituted depending on the protein of interest.

1. Standard PCR methods are used to amplify a DNA fragment containing a cysteine codon in place of the wild-type codon. If the codon to be changed is near the end of the amplified coding region, then only two primers are necessary with the change incorporated into one of these "parent" primers. If more central to the sequence, then a three-primer technique is used with the change incorporated into an internal primer, amplified with one of the parent primers and then the shorter amplified fragment used as a primer with the remaining parent

primer and the original DNA as the template. Finally, if this method fails, two complementary internal primers with the intended change are used to amplify overlapping short fragments using the appropriate parent primers and then these two fragments are combined with the parent primers and the entire insert amplified without additional template added.

2. Ligate the insert containing the single cysteine substitution into the appropriate expression vector. We typically use the pET expression system (Novagen). Both DNAs must be digested with the same restriction endonucleases. Incubate equimolar amounts of insert DNA and pET3d DNA in 1× T4 ligation buffer. Add 400 units of T4 DNA ligase (Bio-Labs) and incubate at 4°C overnight (*see* **Note 1**).
3. Transform DH5α cells with a small amount of the ligation sample. Plate the transformation on LB-ampicillin plates and incubate at 37°C overnight. Also transform a negative (minus insert) control.
4. If there are more colonies on the + insert plate, inoculate several tubes containing 3–5 mL of LB-ampicillin medium with several colonies and grow at 37°C. Isolate the DNA from these cultures by standard DNA mini-prep techniques (*see* **Note 2**).
5. Digest part of the isolated plasmid (typically 1/4 to 1/2) with the original restriction endonucleases used for ligation to liberate the DNA fragment corresponding to the original insert. Stop the digestion with native-gel-loading buffer containing 0.2% SDS.
6. Determine which colonies contain a single insert by loading the digests onto a 1.8% agarose gel along with appropriate size markers. The plasmids that contain correct inserts can be used to transform BL21 cells for overexpression.
7. Next, transform an undigested amount of a plasmid containing the insert into BL21 cells in the same manner as for the DH5α cells (see earlier).
8. Inoculate 200 mL of LB-ampicillin medium with one BL21 colony from the LB-ampicillin plate from **Step 6**.
9. Grow the culture in the absence of IPTG at 37°C to an optical density of 0.6 at 595-nm wavelength light. Add IPTG to a final concentration of 0.2 mM and return the culture to 37°C for approximately 2–4 h (*see* **Note 3**).
10. After 2–4 h pellet the bacteria by centrifugation at 4,000 × *g*. for 15 min.
11. Decant the supernatant and resuspend the pellet in 5–10 mL of TE buffer.

12. Add 10 mg/mL lysozyme to a final concentration of 0.2 mg/mL. Then add Triton-X 100 to a final concentration of 0.2% and incubate for 30 min at room temperature.
13. Dilute the bacteria twofold with 2 M NaCl to a final concentration of 1 M NaCl. Transfer the bacteria to oakridge centrifuge tubes on ice.
14. Sonicate the bacterial slurry for 6 min total in two 3-min sonications (*see* **Note 4**).
15. Pellet the cell debris by centrifugation at 10,000 × *g*. for 30 min at 4°C.
16. Add PMSF to a final concentration of 1×. Dilute the supernatants twofold with TE buffer to bring the NaCl concentration to 0.5 M.
17. Linker histones and most other proteins will bind directly to the Bio-Rex beads. However, core histone proteins must first be incubated with their partner proteins before they will bind to the chromatography matrix (i.e., H2A with H2B) (*see* **Note 5**).
18. Incubate the diluted supernatant with 12.5 mL of a 50% suspension of Bio-Rex 50–100 mesh beads for 4 h at 4°C with constant agitation.
19. After 4 h collect the beads in a plastic 10-mL disposable chromatography column. Collect the flow through fraction in a 50-mL conical tube and save (may be frozen).
20. Wash the column with 2–3 column volumes of 10 mM Tris–Cl of pH 8.0 containing 0.6 M NaCl. Collect the first 10 mL of the wash fraction in a 15-mL conical tube and freeze.
21. Elute the bound proteins with two separate single column volume elutions of 10 mM Tris–HCl of pH 8.0 containing 1.0 M NaCl. Collect the 1.0-M elution fractions in separate 15-mL conical tubes and freeze.
22. After elution, the column was washed with one column volume of 10 mM Tris–Cl of pH 8.0 containing 2.0 M NaCl. Collect the 2.0 M elution fraction in a 15-mL conical tube and freeze.
23. Check 10 μL of each fraction for protein content by SDS–PAGE.

3.2. Reduction and Modification of Cysteine-Substituted Proteins

3.2.1. Reduction of Cysteine-Substituted Proteins

1. Incubate the purified protein of interest in a 15-mL conical tube with 50 mM DTT final concentration for 1 h on ice (*see* **Note 6**).
2. Dilute the protein sample twofold with TE (10 mM Tris–Cl of pH 8.0, 1 mM EDTA) to reduce the NaCl concentration to 500 mM NaCl.

3. Add 0.8 mL of a 50% slurry of Bio-Rex (BioRad) 100–200 mesh chromatography resin and incubate with rotation for 2 h at 4°C.
4. Pour slurry into a 10-mL plastic, disposable chromatography column and collect the flow through fraction. These are commercially available from BioRad or other manufacturers.
5. Wash the column with 3–5 column volumes of buffer containing 10 mM Tris–Cl, pH 8.0, and 0.5 M NaCl. Immediately remove 20 μL of the freshly eluted sample into a separate Eppendorf tube for analysis later on a 12% SDS–protein gel and immediately freeze the 0.5-mL sample to insure that the protein remains reduced. Aliquoting the sample in this manner insures that the sample does not need to be thawed for analysis.
6. An intermediate wash of the column with buffer containing 0.6 M NaCl is performed to remove proteins that are less well bound due to partial degradation. Aliquots of these samples are obtained in the same manner as the previous wash step.
7. Linker histone proteins can be eluted with 0.5-mL aliquots of the same buffer except with 1.0 M NaCl. Typically, five separate 1.0-M NaCl elution steps are performed and collected separately. As previously, 5 μL of the fractions is aliquoted for SDS/PAGE analysis and the samples are frozen immediately. A final elution with buffer containing 2.0 M NaCl buffer will ensure that all the protein has been eluted from the column.
8. Check the protein content of each aliquot obtained from the elution fractions on a 12% SDS–polyacrylamide gel. After separation, incubate protein gel in enough Coomassie blue stain to cover protein gel (45% methanol, 10% acetic acid, and 2.5 mg/mL Coomassie brilliant blue R250). Stain for approximately 1 h at room temperature and destain with 45% methanol, 10% acetic acid until the background of the gel is clear.

3.2.2. Modification of Cysteine-Substituted Proteins with EPD

1. Thaw fraction containing the reduced protein to be modified with EPD. Working as quickly as possible, add a 1.1-fold molar excess of EPD to 60 μL of reduced protein. Incubate for 1 h at room temperature in the dark (*see* **Note 7**).
2. At this point there are several options for treatment of the modified protein. First, if the modified protein is a core histone to be used in standard salt-dialysis reconstitution of nucleosomes, the protein maybe used directly as unreacted reagent will be removed. However, if the protein is to be complexed directly with DNA for directed cleavage mapping, the excess reagent must be removed. In the case of linker histones, removal of excess cleavage reagent can be accomplished by one more round of Bio-Rex chromatography, identical to that

presented except that 60 μL of the 50% slurry is added to the protein. In addition, the slurry is poured into a column made from a blue 1-mL pipet tip fitted with glass wool at the opening. Wash and elute as earlier except that all elution volumes are scaled according to the resin amount. Aliquots for protein analysis are exactly the same size as previously indicated (*see* **Note 8**).

3. The extent of EPD modification of the available cysteine sulfhydrys can be determined by subsequent modification with an excess of (FM) fluorescein maleimide or ^{14}C-NEM (*N*-[ethyl-1-^{14}C]-maleimide). Typically FM can be added in large excess directly to a small sample of modified protein before removal of excess reagent. A control should also be done with unmodified protein. The extent of modification can be determined by running the FM-reacted proteins on SDS–PAGE followed by direct illumination on a long-wave UV light box. Alternatively, add 0.25–0.5 μCi of ^{14}C-NEM to each protein aliquot made from the elution fractions of the Bio-Rex column. Ten minutes later, add 2 volumes of 2× protein-loading buffer to the labeling samples and separate the proteins on a 12% polyacrylamide gel. Stain and destain the gel as earlier and dry the gel onto a piece of Whatman filter paper. Visualize the labeled proteins by exposing the dried gel to ultrasensitive Bio-Max autoradiography film. Be sure to expose the gel to the emulsion side of this film.
4. A protein gel at this step performs two functions. (1) determine which fractions contain the protein of interest and (2) determine the extent of modification with the DNA cleavage reagent.

3.3. Radioactive End Labeling of a Purified DNA Restriction Fragment

1. Treat approximately 5 μg of plasmid DNA or ~0.5 μg of a purified DNA fragment with the appropriate restriction endonuclease in the manufacturer's buffer.
2. Precipitate the DNA by adjusting the solution to 0.3 M sodium acetate and addition of 2.5 volumes of cold ethanol.
3. Resuspend the DNA in phosphatase buffer and treat with alkaline phosphatase for 1 h at 37°C according to manufacturer's instructions.
4. Adjust the solution to 0.1% SDS, phenol extract the solution, and then precipitate the aqueous phase twice with ethanol and sodium acetate. Note that large amounts of phosphatased DNA can be prepared in advance.
5. Resuspend the DNA in 10 μL TE and add 2.5 μL of 10× T4 polynucleotide kinase buffer.
6. Add 25–50 μCi of ^{32}P-γdATP and adjust volume to 24 μL with water.

7. Start the reaction by adding 10 units of T4 polynucleotide kinase and incubate for 30 min at 37°C.
8. Stop the kinase with 200 μL of 2.5 M ammonium acetate (NH_4OAc) and 700 μL of cold 95% ethanol.
9. Pellet the DNA in a microcentrifuge for 30 min at room temperature.
10. Wash the DNA pellet briefly with cold 70% ethanol and dry the DNA in a speedvac concentrator.
11. Dissolve the DNA in 34 μL of TE buffer.
12. Digest the DNA fragment with a second restriction endonuclease that liberates the fragment of interest and yields fragments that can be easily separated by on a gel.
13. Apply the sample to a 6% native polyacrylamide gel.
14. After separation, wrap the gel in Saran wrap and expose the gel to film for 1 min, which is sufficient to detect the specific band containing the labeled fragment. The use of fluorescent markers allows the alignment of the gel (see earlier).
15. Excise the band of interest from the polyacrylamide gel and place into a clean Eppendorf tube. Crush the acrylamide gel slice with an Eppendorf pestle and add 700 μL of TE buffer. The labeled fragment will elute overnight with passive diffusion.
16. Split the sample equally into two Series 8000 Microcentrifuge Filtration Devices and spin for 30 min in a microcentrifuge.
17. Precipitate the DNA and dissolve in TE buffer, pH 8.0. Add enough TE buffer so that the labeled DNA is approximately 1,000 cpm/μL (*see* **Note 9**).

3.4. Reconstitution of Nucleosomes by Salt Step Dialysis

We describe a generic method for the reconstitution of large quantities of homogenous nucleosomes containing radioactively end-labeled DNA for site-directed cleavage analysis *(13)*. Nucleosomes can be reconstituted from virtually any piece of DNA ~150 bp or longer. However, mapping of site-directed cleavages must be done with DNA fragments containing nucleosome positioning sequences such as the *Xenopus borealis* 5S rRNA gene or the 601-based sequences to restrict the histone octamer to one translational position *(14–16)*. Reconstituted nucleosomes bind linker histone in a physiologically relevant manner provided sufficient linker DNA is present – optimal linker histone binding requires nucleosomes centrally positioned on 180-bp DNA fragments.

1. Add approximately 8 μg of unlabeled calf thymus DNA, 200,000–400,000 cpm of singly labeled *Xenopus borealis* 5S ribosomal DNA, purified chicken erythrocyte core histone pro-

tein fractions (H2A/H2B and H3/H4), (*see* **Note 10**), 160 μL of 5 M NaCl (2.0 M final), and TE buffer to a final of 400 μL.

2. Place the reconstitution mixture into a 6–8 kDa MW cut-off dialysis bag. All subsequent dialysis steps are for 2 h at 4°C against 1 L of dialysis buffers unless specified. The first dialysis buffer is 10 mM Tris–Cl of pH 8.0, 1.2 M NaCl, 1 mM EDTA with subsequent dialyses with fresh buffer containing 1.0, 0.8, and then 0.6 M NaCl. The procedure is completed with a final dialysis against TE buffer overnight. Nucleosomes at this stage can be used for gel shift experiments where EDTA does not interfere.
3. For DNA cleavage experiments with EPD, two additional dialysis steps are required. The first dialysis with the reconstitutes is against 10 mM Tris–Cl of pH 8.0 several hours to remove the EDTA. A second dialysis against fresh 10 mM Tris–Cl of pH 8.0 removes trace amounts of EDTA and prepares the samples for chemical mapping with EPD (*see* **Note 11**).

3.5. Maxam–Gilbert G-Specific Reaction

The G-specific reaction used in the Maxam–Gilbert sequencing method provides an easy and quick method to identify the exact location of bases within any known sequence on sequencing gels. It is used here to determine the sites of DNA to base-pair resolution. Since this method is not generally used anymore, the steps are outlined in detail as follows.

1. Add approximately 20,000 cpm of singly labeled DNA (same DNA used to reconstitute nucleosomes).
2. Add 20 μL of 10× G-specific reaction buffer.
3. Add water to a final volume of 200 μL.
4. Start by adding 1 μL of straight dimethylsulfate (DMS) to the tube. Mix immediately and spin briefly in a microfuge (do this in a hood; be careful not to get any DMS on your skin or on standard laboratory gloves; store DMS in a tightly capped brown glass bottle at 4°C).
5. Add 50 μL of G reaction stop solution and mix immediately.
6. Precipitate the DNA.
7. Dissolve the DNA in 90 μL of H_2O.
8. Add 10 μL of piperidine and incubate at 90°C for 30 min.
9. Dry the DNA solution in a speedvac to completion.
10. Dissolve the DNA in 20 μL of water and repeat the drying step. Repeat this step one more time.
11. Dissolve DNA in 100 μL of TE buffer and store at 4°C.

3.6. Site-Directed Hydroxyl Radical Cleavage of DNA

3.6.1. Binding Single Cysteine-Substituted Linker Histone Proteins to Reconstituted Nucleosomes

1. The exact amount of each mutant linker histone protein needed to stoichiometrically bind the nucleosome needs to be empirically determined. Increasing amounts of the linker histone are titrated to a fixed amount of reconstituted nucleosomes (typically 5,000 cpm), and analysis carried out via a gel shift procedure *(20)*. This is typically scaled up tenfold for the site-specific cleavage reaction.
2. Add 5% glycerol final to the binding reaction.
3. Add 50 mM NaCl final to the binding reaction (*see* **Note 12**).
4. Incubate the binding reactions for 15 min at room temperature.
5. Separate the complexes on a 0.7% agarose, 0.5× TBE gel. After drying the gel, expose to autoradiography film and determine the amount of protein necessary for good complex formation. (*see* ref 20)

3.6.2. Site-Directed Hydroxyl Radical Mapping of Linker Histone–DNA Interaction

1. Scale up the binding reaction to include 40,000–50,000 cpm of labeled reconstituted nucleosomes and add enough modified mutant linker histone to form H1-nucleosome complexes (*see* **Note 13**).
2. Add glycerol to 0.5% final concentration (*see* **Note 14**).
3. Add sodium ascorbate to a final concentration of 1 mM.
4. Add H_2O_2 to a final concentration of 0.0075%.
5. Incubate for 30 min at room temperature in the dark.
6. The procedure for stopping the reaction depends on whether one wishes to isolate the complex of interest from other complexes and/or naked DNA in the sample on a preparative agarose gel or just isolate the DNA directly by ethanol precipitation for analysis on sequencing gels. If the latter, proceed to **step 19**. If the former, proceed to **step 7** (*see* **Note 14**).
7. After 30 min, add 1/10th volume of 50% glycerol, 10 mM EDTA solution.
8. Load samples immediately onto a running (90 V) preparative 0.7% agarose/0.5× TBE gel.
9. Separate the samples so that the H1-nucleosome complexes are well resolved from tetramer and free DNA bands.
10. Next, wrap the gel tightly with Saran wrap so that the gel cannot move within the plastic. Lay fluorescent markers onto various portions of the gel for alignment purposes (can be obtained from Stratagene) or otherwise accurately mark the position of the gel on the film.
11. Expose the wet gel for several hours at 4°C.
12. Next, develop the autoradiograph and overlay onto the wet gel, lining up the fluorescent markers.

13. Cut and remove the agarose containing the H1-nucleosome complexes or bands of interest and place them into Series 8000 Microcentrifuge Filtration Devices.
14. Freeze the filtration tubes containing the agarose plugs on dry ice for 15 min.
15. Spin down the agarose in a microfuge at maximum speed for 30 min at room temperature. The fluid from the agarose matrix will be collected in the 2-mL centrifuge tube surrounding the filtration device.
16. Gently remove the agarose plug from the bottom of the filtration device and place into a clean Eppendorf tube. Save the centrifugation devices for use later.
17. Using a microcentrifuge pestle, crush the agarose pellet and add 500 μL of 10 mM Tris–Cl of pH 8.0, 0.1% SDS and continue to crush the agarose.
18. After the agarose is crushed into tiny pieces, place all samples at 4°C overnight or for several hours.
19. Place the crushed agarose into the same centrifugation device and pellet. Spin down the agarose in a microfuge at maximum speed for 30 min at room temperature.
20. Combine identical samples from both spins and precipitate the DNA.
21. Dissolve the DNA in 15 μL of TE buffer.

3.7. Sequencing Gel Analysis of H1°C-EPD Cleavage

1. Add equal numbers of counts from each sample, including the G-specific reaction, to clean Eppendorf tubes.
2. Place the sample into a speedvac concentrator and dry to completeness.
3. Dissolve the sample in 4 μL of formamide-loading buffer.
4. Heat the samples to 90°C for 2 min to denature.
5. Place sample directly onto ice to prevent renaturation.
6. Separate samples on a 6% polyacrylamide/8 M urea sequencing gel running at constant 2,000 V.

3.7.1. Example of Site-Directed Cleavage of Nucleosomal DNA by EPD

An example of a linker histone site-directed DNA cleavage reaction is presented in **Fig. 1b**. A singly end-labeled DNA fragment was incorporated into nucleosomes via the salt dialysis procedure detailed earlier. Labeled mononucleosomes were bound by an EPD-modified linker histone containing a single cysteine substitution for the lysine residue at position 101, referred to as H1 G101C-EPD (*see* **ref.** *12*). After allowing directed hydroxyl radical cleavage for 30 min, the protein–DNA complexes were separated on a 0.7% agarose gel, and the labeled DNA fragments corresponding to the H1-nucleosome complexes were purified.

The DNA from this complex was purified and then analyzed on a sequencing gel (**Fig. 1b**). A schematic of the 5S mononucleosome is shown below the gel. The horizontal black line represents the DNA fragment that contains the nucleosome positioning element. The oval indicates the ~150 bp region wrapped into the nucleosome. Importantly, the DNA is both translationally and rotationally oriented on the surface of the core histones such that bases on the "inside" and "outside" the superhelical gyres are predetermined. The gel shows that H1-G101C-EPD cleaves the DNA as indicated by the blue arrows. The cleavage site, at +63 within the 5S sequence *(12)*, corresponds to a site on the inside of the top superhelical gyre of the nucleosome, near where the DNA exits the nucleosome core (**Fig. 1c**, left). A model for placement of the H1 on the nucleosome based on these data is shown in **Fig. 1c**, right.

4. Notes

1. Many ligation procedures are available from primary literature or commercial sources. Ligation of two DNA fragments occurs more rapidly at room temperature or 37°C if the base-pair overlap is sufficiently stable.
2. Many DNA mini-prep procedures are described in detail in Sambrook et al. *(21)*. The DNA isolated for the techniques described here was done using the alkaline-lysis mini-prep kit from Qiagen Inc.
3. Before proceeding, it is recommended that a small amount of the culture be checked for overexpression of the protein of interest. This can be done by removing 0.5 and 0.2 mL of the culture before and after induction by IPTG, respectively, spinning down the cells, and adding 1× SDS–PAGE gel-loading solution directly to the pellet. The pellet is then incubated at 95°C for 15 min and loaded onto SDS–PAGE with standard markers.
4. Sonication techniques tend to increase the temperature of the sample quickly which could induce proteolysis of the proteins. The sample must therefore be cooled before and during sonication. Allow several minutes between sonication runs to keep the sample as cold as possible.
5. Core histones do not bind to BioRex 70 strongly as individual proteins. However, we have found that when allowed to heterodimerize they bind to the column and elute off consistently in 1 M NaCl *(18)*. This characteristic could be due

to the fact that the core histones are completely unfolded when separated from each other *(20)*.

6. Other reducing agents such as β-mercapto-ethanol may be used but these tend to form a small proportion of β-mercaptol-protein disulfides if the concentration of reducing agent becomes low. DTT is the reducing reagent of choice given its greater reducing power.
7. From this point forward scrupulously avoid introduction of all cysteine-reactive reagents (thiols, maleimides, haloalkyls) as these will destroy the linkage between the protein and the EPD.
8. In some cases the conditions of complex preparation (such as an extended dialysis procedure) may not be conducive to retention of activity of the tethered compound. Thus it is possible to modify the cysteine after complex formation; however, one needs to maintain the cysteine in a reduced state, completely remove the reducing agent before modification, and then remove the excess reagent by some means before the cleavage reaction.
9. Storing labeled DNA in a concentrated form is not advised as autodegradation of the DNA takes place. DNA can be stored for several weeks at approximately 5,000–10,000 cpm/μL.
10. A complication of the in vitro reconstitution procedure is that purified histone proteins are often obtained in two fractions, H2A/H2B and H3/H4 *(18, 19)*. Thus, in addition to total histone mass, the ratio between these two substituents must be empirically adjusted to yield maximum octamer–DNA complexes *(18)*.
11. Note, it is especially difficult to completely remove iron from the high salt solutions used for dialysis. One solution we found was to include high-quality highly polymerized DNA (50–100 ng/mL calf thymus or salmon sperm) in the dialysis buffers to ensure that metal ions from the solution do not accumulate on the sample DNA.
12. Several methods can be used for the incorporation of linker histones into reconstituted mononucleosomes. The method described here involves direct addition of linker histones to mononucleosomes in 50 mM NaCl. Linker histones are folded in low salt solutions in the presence of DNA *(1)*. Indeed, we find that linker histones can be directly mixed to nucleosomes in either 5 or 50 mM NaCl solutions, and these proteins then bind in a physiologically relevant manner *(20)*.
13. Note that Tris buffers >~20 mM are fairly good radical scavengers and thus Tris should be kept≤10 mM for

optimal cleavage. In some cases, utilizing buffers that do not scavenge such as phosphate may increase cleavage signals. On the other hand having such high Tris concentrations may actually help keep the cleavages localized.

14. In cases where it is desirable to separate cleaved native complexes on nucleoprotein gels after the reaction, the cleavage can be terminated by addition of EDTA and glycerol to 5%. Note that free EDTA binds the metal several orders of magnitude better than the tethered chelator and glycerol is an effective radical scavenger.

Acknowledgments

This work was supported by NIH grant GM52426.

References

1. Wolffe A. P. (1998). *Chromatin Structure and Function*. 3rd edn Academic Press, San Diego.
2. Crane-Robinson C. (1997). Where is the globular domain of linker histone located on the nucleosome?. *Trends Biochem Sci* **22**, 75–77.
3. Zhou Y. B., Gerchman S. E., Ramakrishnan V., Travers A., and Muyldermans S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature* **395**, 402–405.
4. van Holde K. E. (1989). *Chromatin*. Springer, New York.
5. Zheng C., and Hayes J. J. (2003). Structures and interactions of the core histone tail domains. *Biopolymers* **68**, 539–546.
6. Carruthers L. M., Bednar J., Woodcock C. L., and Hansen J. C. (1998). Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher-order chromatin folding. *Biochemistry* **37**, 14776–14787.
7. Brown D. T., Izard T., and Misteli T. (2006). Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. *Nat Struct Mol Biol* **13**, 250–255.
8. Pendergrast P. S., Chen Y., Ebright Y. W., and Ebright R. H. (1992). Determination of the orientation of a DNA binding motif in a protein-DNA complex by photocrosslinking. *Proc Natl Acad Sci USA* **89**, 10287–10291.
9. Ermacora M. R., Delfino J. M., Cuenoud B., Schepartz A., and Fox R. O. (1992). Conformation-dependent cleavage of staphlyococcal nuclease with a disulfide-linked iron chelate. *Proc Natl Acad Sci USA* **89**, 6383–6387.
10. Chen Y., and Ebright R. H. (1993). Phenylazide-mediated photocrosslinking analysis of Cro-DNA interaction. *J Mol Biol* **230**, 453–460.
11. Lavoie B. D., Shaw G. S., Millner A., and Chaconas G. (1996). Anatomy of a flexer-DNA complex inside a higher-order transposition intermediate. *Cell* **85**, 761–771.
12. Hayes J. J. (1996). Site-directed cleavage of DNA by a linker histone–Fe(II) EDTA conjugate: localization of a globular domain binding site within a nucleosome. *Biochemistry* **35**, 11931–11937.
13. Heilek G. M., and Noller H. F. (1996). Site-directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S5. *Science* **272**, 1659–1662.
14. Flaus A., Luger K., Tan S., and Richmond T. J. (1996). Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc Natl Acad Sci USA* **93**, 1370–1375.
15. Miller G., and Hahn S. A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex. Nat Struct Mol Biol **13**, 603–610.

16. Chafin D. R., Lee K. M., and Hayes J. J. (1999). Site-directed chemical probing of histone-DNA interactions. *Methods Mol Biol* **119**, 27–43.
17. Tullius T. D. (1988). DNA footprinting with hydroxyl radical. *Nature* **332**, 663–664.
18. Hayes J. J., and Lee K. M. (1997). In vitro reconstitution and analysis of mononucleosomes containing defined DNAs and proteins. *Methods* **12**, 2–9.
19. Simon R. H. and Felsenfeld G. (1979). A new procedure for purifying histone pairs H2A + H2B and H3 + H4 from chromatin using hydroxylapatite. *Nucleic Acids Res* **6**, 689–696.
20. Hayes J. J., and Wolffe A. P. (1993). Preferential and asymmetric interaction of linker histones with 5S DNA in the nucleosome. *Proc Natl Acad Sci USA* **90**, 6415–6419.
21. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning, A laboratory manual. 2nd Editon. Cold Spring Harbor Laboratory Press, USA.

Chapter 11

Identification of Nucleic Acid High-Affinity Binding Sequences of Proteins by SELEX

Philippe Bouvet

Summary

A technique is described for the identification of nucleic acid sequences bound with high affinity by proteins or by other molecules suitable for a partitioning assay. Here, a histidine-tagged protein is allowed to interact with a pool of nucleic acids and the protein–nucleic acid complexes formed are retained on a Ni-NTA matrix. Nucleic acids with a low level of recognition by the protein are washed away. The pool of recovered nucleic acids is amplified by the polymerase chain reaction and is submitted to further rounds of selection. Each round of selection increases the proportion of sequences that are avidly bound by the protein of interest. The cloning and sequencing of these sequences finally completes their identification.

Key words: SELEX, Binding sequences, Sequence recognition, Nucleic acids ligands.

1. Introduction

The interactions of nucleic acids with proteins are involved in numerous biological functions. Most of these interactions involve specific contacts between nucleic acid and protein with variable binding affinities. The identification of nucleic acid recognition sequence of a specific protein is often the first step to undertake the study of the biological function of this protein. Over the last 15 years, the SELEX procedure (Systematic Evolution of Ligands by EXponential enrichment) has been used to identify high-affinity nucleic acids ligands for different proteins. This method was first described for the selection of DNA and RNA target of nucleic acid binding proteins *(1, 2)* but has then been used for the

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_11

selection of nucleic acid ligands for any kind of targets *(3)*. This methodology uses the power of genetic selection techniques combined with the advantage of in vitro biochemical experiments. It is a rapid in vitro technique and is relatively easy to implement in all laboratories. This procedure should accelerate and simplify nucleic-acid/protein interaction studies.

This process involves few simple steps as described in **Fig. 1**. The procedure consists of an enrichment of individual oligonucleotide molecules from complex mixtures of nucleic acid sequences by repeated rounds of selection. First, ligand sequences that bind to the target protein are selected. Then the bound and the unbound molecules are partitioned. In a final step, the selected sequences are amplified by PCR. This cycle of selection–amplification (round) is repeated until enrichment is obtained for nucleic acid sequences that bind to the protein with high affinity.

The strategy is designed to determine the optimal binding nucleic acid sequences, also called "aptamer" *(4)*. However, it

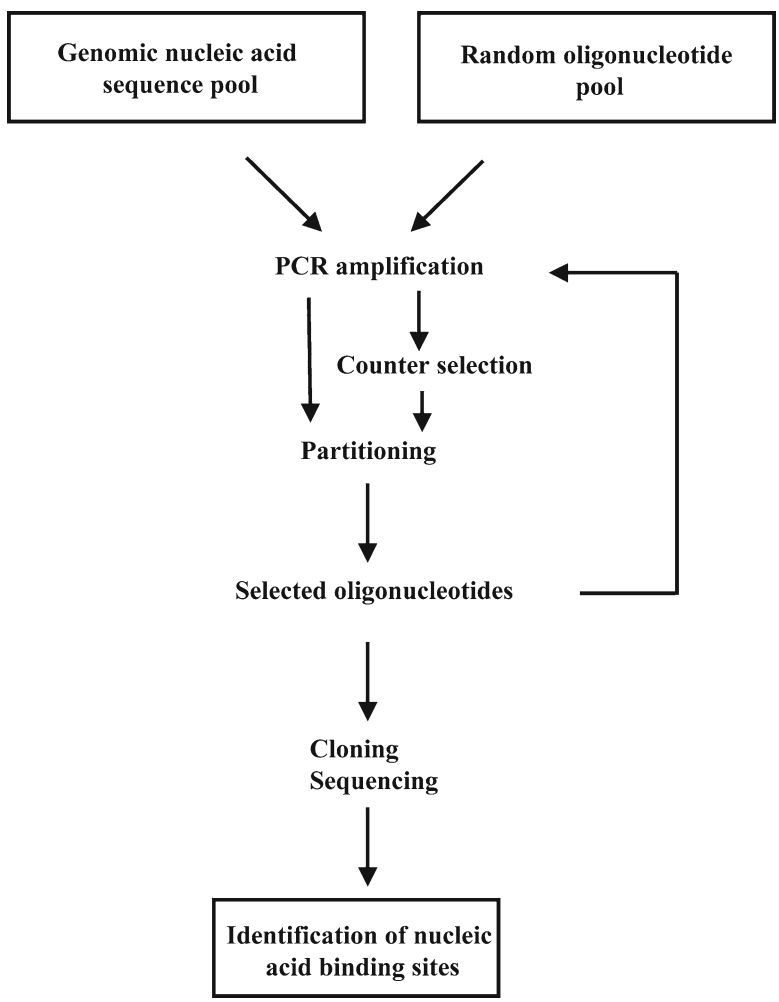


Fig. 1. Schematic representation of the different steps of SELEX.

should be noted that high-affinity nucleic acid ligand could be isolated even for partitioning agents which are not nucleic-acid-binding proteins, and the isolated SELEX sequences may not be related to the real binding site of the protein in vivo. Therefore, if this strategy is used to determine the nucleic acid recognition sequence of a protein, one must realize that the most difficult part of this study will be the analysis of the isolated SELEX sequences and the demonstration that there are relevant for the in vivo function of the studied protein.

Numerous protocols for the SELEX procedure have been used successfully by different laboratories. In fact, each step of this procedure can be optimized *(5)* and modified in function of the characteristics of the nucleic-acid-binding protein that is studied *(6)*. In the initial SELEX protocol, the aptamer libraries were made with chemically synthesized nucleic acid molecules. In genomic SELEX, genome-derived nucleic acid sequences (DNA or RNA) that bind with affinity to specific protein are identified *(7, 8)*. More recently, selection procedures using microfluidic devices such as capillary chromatography and microarrays have also been developed *(9)*. Several automated procedures have also been proposed *(10–12)*, but these technologies cannot yet be routinely used in any molecular biology laboratory.

We will provide here a basic typical protocol that has been used successfully by several laboratories *(13–17)* to identify nucleic acid ligands for RNA- and DNA-binding proteins and that can be very easily implemented in any laboratory to perform the selection procedure in just a few days. Typically, one round of selection can be done in 1 day. In this protocol, we propose to use an oligonucleotide template that contains a random sequence of 25 nucleotides. This oligonucleotide has been used in many studies for the identification of RNA- and DNA-binding proteins *(13–15, 18)*, but it can be replaced by any other DNA template with variable random nucleotide length that would be more suitable for the protein that is being studied. We are providing a detailed protocol for the selection of aptamers of RNA-binding proteins. In the case of DNA-binding proteins, **steps 6** and **18** of **Subheading 3** should be omitted. For each step, detailed descriptions of the strategic choices that should be made and of modifications that can be introduced in the protocol will be given in **Subheading 4**.

2. Materials

1. The following synthetic DNA template has been used with success by several laboratories: 5′ TGGGCACTATTTATAT CAAC (N25) AATGTCGTTGGTGGCCC 3′ with these

flanking primers: T7 5′-CGCGGATCCTAATACGACTCACTATAGGGGCCACCAACGACATT-3′ and Rev 5′-CCCGACACCCGCGGATCCATGGGCACTATTTATATCAAC-3′. The T7-Xba primer 5′ GGTCTAGATAATACGACTCACTATAGGGG 3′ and Rev-HIII primer 5′ ACCGCAAGCTTATGGGCACTATTTATAT 3′ can be used for the final PCR amplification and will allow an oriented cloning (*Xba*I and *Hin*dIII) of the PCR product in a cloning vector like pBluescript (Stratagene).

2. Thermocycler.
3. *Taq* polymerase.
4. Partitioning matrix (to be chosen in function of the studied protein).
5. Nucleic acid electrophoresis system.
6. NT2 buffer: 50 mM Tris–HCl (pH 7.4), 150 mM NaCl, 0.05% NP 40, 1 mM $MgCl_2$.
7. Binding buffer (BB): 50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 20 mM KCl, 1 mM DTT, 0.05% NP 40, 1 mM $MgCl_2$, 2.5% polyvinyl alcohol (PVA), 1 mM EGTA, 50 μg/mL poly(A), 2 μL/mL vanadyl ribonucleoside complex (VRC), 0.5 μg/mL tRNA, 125 μg/mL BSA.
8. 5× Reverse transcription buffer: 250 mM Tris–HCl (pH 8.5), 40 mM $MgCl_2$, 5 mM DTT, 250 μg/mL BSA, 150 mM KCl.
9. 1× Transcription buffer: 40 mM Tris–HCl (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT.
10. Gel shift buffer (GSB): 100 mM Tris–HCl (pH 7.4), 4 mM $MgCl_2$, 200 mM KCl, 20% glycerol, 1 mM dithiothreitol, 0.5 mg/mL tRNA, 4 μg/mL BSA.
11. Polyacrylamide gel shift: 8% polyacrylamide (acrylamide:bis, 60:1) containing 5% glycerol in 0.5× TBE buffer (0.045 M Tris-borate, 1 mM EDTA).

3. Methods

1. About 10 pmol of synthetic template DNA (N25) (*see* **Note 1 and 2**) is amplified by PCR in a 100-μL reaction in a 500-μL test tube (*see* **Note 3**). Add 2 μL of a mix of all four dNTPs (10 mM each) and 500 ng of each primer. One unit of *Taq* polymerase is added just before the start of the amplification procedure. If the thermocycler does not possess a hot cover, the reaction mixture is overlaid with two drops of mineral oil.

2. Set up the thermocycler with the following cycle conditions: denaturation 1 min at 94°C, annealing 1 min at 50°C, elongation 1 min at 72°C for 25 cycles, then finish with an elongation of 10 min at 72°C. At the end of the PCR reaction, the reaction can be left at 4°C without further purification.
3. Analyze 5 µL of the PCR reaction on a 3% agarose gel (made with 1× TAE). Run in parallel a commercial DNA ladder which gives characteristic bands around 100 pb. The PCR reaction should give a nice signal at 108 bp.
4. Add 100 µL of phenol:chloroform (1:1) to the PCR reaction and mix vigorously for 1 min. After a 5-min centrifugation at 21,000 × *g*, the upper aqueous phase is extracted one more time with 1 volume of phenol:chloroform. Ten microliters of 3 M sodium acetate (NaOAc) of pH 5.0 and 300 µL of cold ethanol are added. Allow DNA precipitation for at least 15 min at –20°C.
5. The PCR product is recovered by centrifugation (15 min at 21,000 × *g*), washed with 70% ethanol, dried and resuspended in 10 µL of sterile water.
6. In vitro transcription. This step should be omitted for a SELEX with a DNA-binding protein. One microgram of PCR product from **step 5** is incubated in 1× transcription buffer with 0.5 mM of each rNTP, 1 unit of RNasin, and 20 units of T_7 RNA polymerase. The reaction is allowed for 1 h at 37°C. The DNA template is eliminated by addition of 1 unit of RNase-free DNase for ten more minutes at 37°C. After two phenol:chloroform extractions, the RNA is purified through a G50 column (to remove most of unincorporated nucleotides) then precipitated with 0.1 volume of 3 M NaAc of pH 5 and 2 volumes of 100% EtOH for 15 min at –20°C. The RNA is pelleted for 15 min at 21,000 × *g*, washed with 70% EtOH, dried and resuspended in 20 µL of RNase-free water. One microliter of the transcription reaction is loaded on a 3% agarose gel to check the quality of the RNA, and the RNA concentration can then be determined through UV absorption at 260 nm.
7. Preparation of the partitioning matrix. The nature of this matrix will depend on the protein that is used (*see* **Note 4**). We will provide here a detailed protocol for a selection procedure with a histidine-tagged protein. Other strategies are mentioned in **Note 4**. Take 2 µL of Ni-NTA agarose beads (Qiagen), and wash twice with 500 µL of sterile water to remove all storage buffer (by a 15-s centrifugation in a bench top centrifuge). Then wash the beads twice with 500 µL of NT2 buffer. During the last wash, the solution is divided into two test tubes (tubes A and B). Centrifuge and eliminate 150 µL of supernatant.

8. Add about 1 pmol of purified histidine-tagged protein to tube A. Incubate 30 min at 4°C on a roller to allow binding of the protein on the Ni-NTA beads.
9. Centrifuge tube A for 15 s, remove supernatant, then wash the beads twice with 500 µL of NT2 buffer to remove all unbound protein.
10. Centrifuge tubes A and B, and remove most of the NT2 supernatant to leave about 10 µL of buffer above the beads. Ni-NTA beads must be visible at the bottom of the tube.
11. Add 100 µL of BB buffer in tube B (*see* **Note 5**).
12. Add 15 µg of nucleic acid from **step 5** (for a DNA-binding protein) or from **step 6** (for an RNA-binding protein) in tube B (*see* **Note 6**). Incubate 5 min at room temperature. Centrifuge tube B for 15 s. Remove and save supernatant. This is the counter selection (*see* **Note 7**).
13. Add supernatant from **step 12** to the tube that contains the protein (tube A) (*see* **Note 8**). Incubate 5 min at room temperature. Centrifuge tube A for 15 s, then remove and discard supernatant.
14. Add 1 mL of NT2 buffer (*see* **Note 9**) in tubes A and B. Mix well by inverting the tubes three times. Centrifuge 15 s to pellet the Ni-NTA beads. Remove the supernatant as much as possible.
15. Repeat this wash four more times. At the last wash, transfer the nucleic acid/protein complex in a new test tube (*see* **Note 10**).
16. After the last wash, leave 100 µL of NT2 buffer in each tube. Add 100 µL of sterile water and 200 µL of phenol:chloroform (1:1). Vortex 30 s and spin 5 min at full speed. Repeat this extraction one more time.
17. Recover the upper aqueous phase, and add 2 µL of 1 M $MgCl_2$, 20 µL of 3 M NaOAc of pH 5, and 700 µL of 100% EtOH. Precipitate for at least 30 min at –20°C, spin 30 min at 21,000 × *g*. Wash the pellet with 70% EtOH, dry and resuspend it in 13 µL of sterile water.
18. This step (reverse transcription) should be omitted if the SELEX is performed with a DNA oligonucleotide. To each tube A and B, add 100 ng of Rev primer (in 1 µL), 2 µL of a dNTP mix (each dNTP at 10 mM), 4 µL of 5× reverse transcription buffer, 30 units of RNasin (Promega), and 25 units of AMV reverse transcriptase (Boehringer, Mannheim). The reaction is set up for 5 min at 55°C, then 1 h at 42°C. Five microliters of this reverse transcription reaction is used directly, without further purification, for the next PCR amplification.
19. Add to the recovered nucleic acid of **step 17** or **18** the reagent necessary for the PCR reaction, as described in **step 1**. Include a control reaction without oligonucleotide template

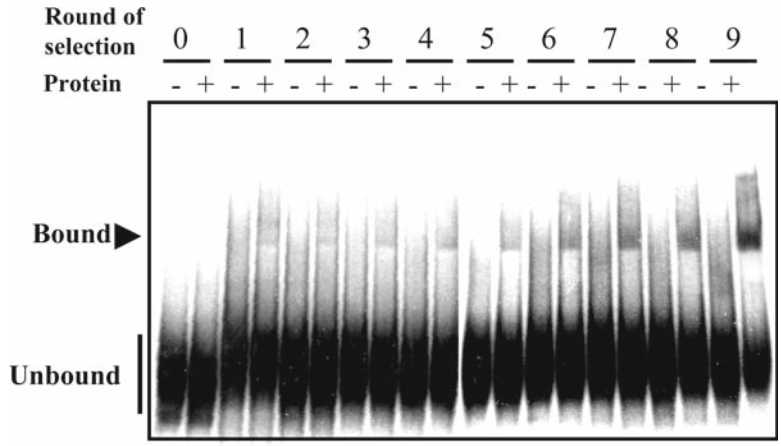


Fig. 2. Example of enrichment of the selected sequence after different rounds of amplification. An aliquot of each oligonucleotide pool after each round of amplification was used for in vitro transcription with [α-^{32}P] CTP and used for gel shift assay. About 10 fmol of [^{32}P] RNA was incubated with (+) or without (–) protein (10 nM) for 15 min at room temperature in 20 μL of RNA binding buffer (GSB). Then the mixture was directly loaded on an 8% polyacrylamide gel assay. This example corresponds to the selection that was performed with nucleolin *(15)*.

to make sure that the PCR reaction is performed in good conditions. No PCR product should be obtained with tube B (*see* **Note 10**).

20. Purify the PCR product obtained with tube A as described in **steps 4** and **5** and repeat several rounds of selection.
21. After several rounds of selection (usually between four and ten) check for an enrichment of the oligonucleotide pool in high-affinity ligand for the target protein (for example, *see* **Fig. 2** and ref. *14)* (also *see* **Note 11**).
22. The final PCR can be performed with the primers T7-Xba and Rev-HIII (*see* **Note 1**). After gel purification, the PCR products are digested with the restriction enzymes *Xba*I and *Hin*dIII, and the oligonucleotides are cloned in an adequate vector (for example pBluescript). Individual clones are selected and inserts are sequenced using standard methodology (*see* **Note 12**).

4. Notes

1. A custom-made random oligonucleotide can be easily synthesized using standard chemistry. The variable sequence is flanked by fixed sequences at its 5′ and 3′ extremities.

Full-length oligonucleotides should be gel purified before their amplification by PCR. The length of the random sequence may vary from a few nucleotides to as much as a hundred. If a simple DNA-binding site is expected, a random sequence of as little as 20 nucleotides should be sufficient. A library with a short random sequence has also the advantage of being more likely to contain all possible random sequences and therefore to allow the selection of the best binding sequence. Sequencing of a few random sequences should be ideally performed to ensure that the synthesis of the random sequences has not been biased by a preferential incorporation of one deoxynucleotide. If the composition of the random sequence is severely biased, this should be corrected by modifying the percentage of addition of nucleotides accordingly during the synthesis of the random sequence. If a binding site is already known for the protein, the SELEX procedure can be used to determine nucleotides important for binding affinity and specificity. In this case, an oligonucleotide containing a degenerated sequence within the known binding site can be synthesized *(20)*. Libraries that contain genomic sequences can also be used to identify potential natural binding sites *(21, 22)*.

2. The random sequence is flanked by fixed sequences (17–20 nt) to allow PCR amplification with the corresponding primers (*see* **Note 1** for an example of primer sequences). It is important to check that the flanking sequences are not a binding site for the protein. This can easily be done by performing a binding assay between the random pool and the studied protein.
3. The PCR reaction can be modified to allow the production of single-stranded oligonucleotides, the incorporation of modified nucleotides or random mutations, etc. If the PCR reaction produces aberrant products (higher molecular weight DNA products, smear, etc.) several tests reactions (with various amount of primers, number of cycle) can be realized.
4. Several methods of partitioning can be used in function of what is available to the researcher *(19)*. If the target is a tagged recombinant protein (GST, Histidine, or any other tag) nucleic acid/protein complexes can be recovered by classical affinity chromatography as described in this protocol. If the protein is pure but not tagged, filtration of the binding reaction mixture through nitrocellulose filter allows separation of the bound and unbound molecules *(25)*. An alternative method of partitioning uses gel shift analysis. For this method, labeled oligonucleotides are preferentially used to easily identify the nucleic acid/protein complex. Shifted

oligonucleotides are eluted from the gel and used for PCR amplification.

5. The composition of the binding buffer should be adapted to the protein that is being studied. The addition of nucleic acid competitors like tRNA, or homopolymers like poly(A) or poly (dI–dC), might be necessary in some selection experiments to reduce the nonspecific binding of the protein to the random oligonucleotide. Preliminary tests of interaction of the random pool with the protein can be performed with nucleic acid competitors to determine the best selection conditions and the concentration of these competitors that need to be added if required.
6. The amount of oligonucleotide present in the binding reaction should be in large excess over the protein. This ensures an efficient competition between ligands for the protein. The ratio oligonucleotide/protein is often comprised within the range of 10 and 1,000. The volume of the binding reaction should also be determined in function of the diversity of the library. A large binding reaction volume might be required if one wants to test all possible sequences (4^n, where n is the number of random nucleotides) present in the initial library.
7. The interaction of the random oligonucleotide pool with the partitioning matrix, without the protein (called counter selection or negative selection) (**Fig. 1**), is important to remove from the random oligonucleotide pool molecules with high affinity for the partitioning matrix. This counter selection is not necessary if the DNA bound to the protein is recovered using gel shift since only the shifted band will be used for the next cycles. The counter selection can be performed during the first rounds of selection and could be omitted for the next cycles.
8. In most published experiments, the SELEX procedure is performed with purified recombinant proteins. However, the nucleic-acid-binding specificity and affinity can sometimes be the result of interactions between the protein and other cellular polypeptides. The SELEX procedure can be performed with crude cell extracts that contain the protein of interest or multiprotein complexes if the partitioning procedure allows a specific recovering of the protein target *(23, 24)*. Epitope-tagged protein can be expressed in cells, or added to a cell extract and used for the SELEX. For some SELEX experiments it could be also interesting to use truncated protein with only the nucleic-acid-binding domain. In some case, this can significantly reduce nonspecific binding of the random oligonucleotide pool with the protein and therefore reduce the number of rounds necessary for the isolation of specific ligands.

9. The stringency of the binding and washing buffer can be increased if necessary. This could be done by increasing the salt concentration or by adding 0.5–1.0 M urea. Usually, between 1 and 10% of the initial oligonucleotide pool bind to the target protein. Preliminary binding tests should determine the optimal buffer stringency to allow a binding which falls within this range. Buffer stringency can also be increased during the cycling process if no substantial enrichment is observed.
10. It has been sometimes observed that nucleic acids bind poorly to some plastic tubes. This binding is however sufficient to increase the nonspecific binding and to give a PCR product after the amplification reaction. This problem could be overcome by using siliconized test tubes, or by transferring the nucleic acid/protein complex in a new test tube during the washing procedure.
11. After several rounds of selection (usually between four and ten) it is important to check for an enrichment of the oligonucleotide pool in high-affinity ligands for the target protein before proceeding to the cloning and sequencing steps of these selected sequences. This can be done by doing an interaction between the protein and labeled oligonucleotide pools of each round of selection (5′ labeling for DNA oligonucleotide or in vitro labeled transcription for RNA oligonucleotide). An example is shown in **Fig. 2**. This example corresponds to the selection that was performed with nucleolin *(15)*. An aliquot of each oligonucleotides pool after each round of amplification was used for in vitro transcription with [α-^{32}P] CTP and used for gel shift assay. About 10 fmol of [^{32}P] RNA was incubated with or without 10 nM of protein for 15 min at room temperature in 20 µL of RNA binding buffer (GSB, *see* Material & Methods **Note 10**). The mixture was then directly loaded on an 8% polyacrylamide gel (acrylamide:bis, 60:1) containing 5% glycerol in 0.5× TBE buffer. The gel was then dried and subjected to autoradiography. In this particular experiment, the selection procedure was stopped after nine rounds of selection when a significant enrichment of bound oligonucleotides was obtained. Subsequent sequencing of selected molecules showed that about 20% of them contained a consensus motif. If no enrichment is observed, more rounds of selection can be performed in the same experimental conditions or with higher stringency (*see* **Note 9**).
12. The number of individual clones that need to be sequenced to identify a consensus binding site might vary from one experiment to the other. But, in general, if an enrichment of the selected sequences has been detected during the different rounds of selection (*see* **Note 11**) it might be possible

to detect a consensus binding site in as few as 20 individual sequences. For the determination of a more precise consensus motif, more sequences will be needed. It might also happen that exactly the same selected sequence over the full length of the initial random region is found several times in the sequenced clones, and it is therefore not possible to identify the consensus motif recognized by the protein since they have all the same sequence. It might indicate that the selection has been too strong, and in this case it will be required to perform the sequencing from the previous round of selection.

Acknowledgments

The work in the authors' laboratory is supported by grants from the CNRS, ANR N° BLAN07-2_190263, and Association pour la Recherche sur le Cancer (ARC).

References

1. Oliphant, A.R., Brandl, C.J., and Struhl, K. (1989). Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein. *Mol. Cell. Biol* **9**, 2944–2949.
2. Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510.
3. Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995). Diversity of oligonucleotide functions. *Ann. Rev. Biochem* **64**, 763–797.
4. Ellington, A.D., and Szostak, J.W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822.
5. Irvine, D., Tuerk, C., and Gold, L. (1991). SELEXION. Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *J. Mol. Biol* **222**, 739–761.
6. Stoltenburg, R., Reinemann, C., and Strehlitz, B. (2007). SELEX – a (r)evolutionary method to generate high-affinity nucleic acid ligands*Biomol. Eng* **24**, 381–403.
7. Kim, S., Shi, H., Lee, D. K., and Lis, J. T. (2003). Specific SR protein-dependent splicing substrates identified through genomic SELEX. *Nucleic Acids Res* **31**, 1955–1961.
8. Shtatland, T., Gill, S. C., Javornik, B. E., Johansson, H. E., Singer, B. S., Uhlenbeck, O. C., Zichi, D. A., and Gold, L. (2000). Interactions of Escherichia coli RNA with bacteriophage MS2 coat protein: genomic SELEX. *Nucleic Acids Res.* **28**, E93.
9. Mosing, R. K., and Bowser, M. T. (2007). Microfluidic selection and applications of aptamers.*J. Sep. Sci* **30**, 1420–1426.
10. Cox, J.C., and Ellington, A.D. (2001). Automated selection of anti-protein aptamers. *Bioorg. Med. Chem* **9**, 2525–2531.
11. Cox, J. C., Hayhurst, A., Hesselberth, J., Bayer, T. S., Georgiou, G., and Ellington, A. D. (2002). Automated selection of aptamers against protein targets translated in vitro: from gene to aptamer. *Nucleic Acids Res.* **30**, e108.
12. Hybarger, G., Bynum, J., Williams, R.F., Valdes, J.J., and Chambers, J.P. (2006). A microfluidic SELEX prototype. *Anal. Bioanal. Chem* **384**, 191–198.
13. Bouvet, P., Matsumoto, K., and Wolffe, A. P. (1995). Sequence-specific RNA recognition by the Xenopus Y-box proteins. An essential

role for the cold shock domain. *J. Biol. Chem* **270**, 28297–28303.

14. Clouaire, T., Roussigne, M., Ecochard, V., Mathe, C., Amalric, F., and Girard, J. P. (2005). The THAP domain of THAP1 is a large C2CH module with zinc-dependent sequence-specific DNA-binding activity. *Proc. Natl. Acad. Sci. USA* **102**, 6907–6912.
15. Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F., and Bouvet, P. (1996). Nucleolin is a sequence-specific RNA-binding protein: characterization of targets on pre-ribosomal RNA. *J. Mol. Biol* **260**, 34–53.
16. Triqueneaux, G., Velten, M., Franzon, P., Dautry, F., and Jacquemin-Sablon, H. (1999). RNA binding specificity of Unr, a protein with five cold shock domains. *Nucleic Acids Res* **27**, 1926–1934.
17. Tsai, D.E., Harper, D.S., and Keene, J.D. (1991). U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids Res* **19**, 4931–4936.
18. Harper, D.S., Fresco, L.D., and Keene, J.D. (1992). RNA binding specificity of a Drosophila snRNP protein that shares sequence homology with mammalian U1-A and U2-B proteins. *Nucleic Acids Res* **20**, 3645–3650.
19. Gopinath, S.C. (2007). Methods developed for SELEX. *Anal. Bioanal. Chem* **387**, 171–182.
20. Bartel, D. P., Zapp, M. L., Green, M. R., and Szostak, J. W. (1991). HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67**, 529–536.
21. Gao, F. B., Carson, C. C., Levine, T., and Keene, J. D. (1994). Selection of a subset of mRNAs from combinatorial 3′ untranslated region libraries using neuronal RNA-binding protein Hel-N1. *Proc. Natl. Acad. Sci. USA* **91**, 11207–11211.
22. Singer, B. S., Shtatland, T., Brown, D., and Gold, L. (1997). Libraries for genomic SELEX. *Nucleic Acids Res* **25**, 781–786.
23. Pollock, R., and Treisman, R. (1990). A sensitive method for the determination of protein-DNA binding specificities. *Nucleic Acids Res* **18**, 6197–6204.
24. Ringquist,S.,Jones,T.,Snyder,E.E.,Gibson,T., Boni, I., and Gold, L. (1995). High-affinity RNA ligands to Escherichia coli ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. *Biochemistry* **34**, 3640–3648.
25. Tuerk, C., Eddy, S., Parma, D., and Gold, L. (1990). Autogenous translational operator recognized by bacteriophage T4 DNA polymerase. *J. Mol. Biol* **213**, 749–761.

Chapter 12

Identification of Sequence-Specific DNA-Binding Proteins by Southwestern Blotting

Simon Labbé, Jean-François Harrisson, and Carl Séguin

Summary

We describe a Southwestern blotting method for characterization of both DNA-binding proteins and their specific sites. Proteins are first separated on a sodium dodecyl sulfate (SDS) polyacrylamide gel, then renatured in SDS-free buffer and transferred by electroblotting to an immobilizing membrane, and detected by their ability to bind radiolabeled DNA. The protein(s) interacting with the labeled DNA is visualized by autoradiography. This technique was used in our laboratory to visualize the metal regulatory consensus sequence-binding protein MTF-1 in L cell crude nuclear extracts.

Key words: DNA-binding proteins, Binding site, Detection, Blotting assay, Metal transcription factor-1, Sequence specificity, Transcription activator.

1. Introduction

Southwestern blotting was first described by Bowen et al. *(1)* and was used to identify DNA-binding proteins that specifically interact with a chosen DNA fragment in a sequence-specific manner. In this technique mixtures of proteins such as crude nuclear extracts or partially purified protein preparations are first fractionated on an SDS denaturing gel, after which the gel is equilibrated in an SDS-free buffer to remove the detergent, and the proteins are then transferred by electroblotting to an immobilizing membrane. During this transfer, the proteins can renature and DNA-binding proteins may be detected by their ability to bind radiolabeled DNA. Fractionation of crude nuclear extracts

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_12

on an SDS gel followed by protein blotting and analysis of sequence specificity directly on the blot combines the advantages of a high-resolution fractionation step with the rapid analysis of a large number of different DNA-binding proteins.

The successful identification of specific DNA-binding proteins by this technique largely depends on the renaturation of the proteins after their separation by SDS/PAGE. The ease with which renaturation can be achieved after treatment with SDS varies from protein to protein. Some DNA-binding proteins may be inefficiently renatured and thus be unable to bind DNA once there are adsorbed onto membranes. In addition, any multimeric protein that requires a combination of subunits to bind DNA will be missed. Proteins requiring a cofactor(s) in order to show their ability to specifically interact with DNA will also be hard to detect.

Another important point to keep in mind when using this procedure relates to the fact that site-specific protein–DNA interactions may be obscured by the large number of nonspecific DNA-binding proteins present in a crude nuclear extract. Conditions for DNA binding, such as pH, ionic strength, and divalent cation requirement, as well as the type and amount of nonspecific DNA added to the buffers, should be optimized for each protein under investigation. Special attention should also be given to known contaminants frequently present in crude nuclear extracts, which copurify with sequence-specific DNA-binding proteins (*see* **Note 14**). Nonetheless, over the last few years, Southwestern blotting techniques have been used to identifying and characterizing several specific DNA-binding proteins (for examples, *see* **Fig. 1** and refs. *2–13)* including both histone and nonhistone proteins, as well as RNA-binding proteins *(14–16)*. In addition, preparative Southwestern blots have been used to define DNA sequences recognized by a DNA-binding protein *(17)*, while Southwestern screening of a cDNA library has been utilized to isolate several sequence-specific DNA-binding proteins *(18, 19)*, and a Southwestern chemistry-based ELISA was developed for quantitative assessment of the transformation state of the aryl hydrocarbon receptor *(20)*.

For another detailed description of the Southwestern methodology, readers are referred to the recent publication of Siu et al. *(21)*.

2. Materials

1. Crude nuclear extracts or protein chromatographic fractions.
2. Tris–HCl 1.5 M, pH 8.8 (all solutions are made with double distilled or Milli Q water).

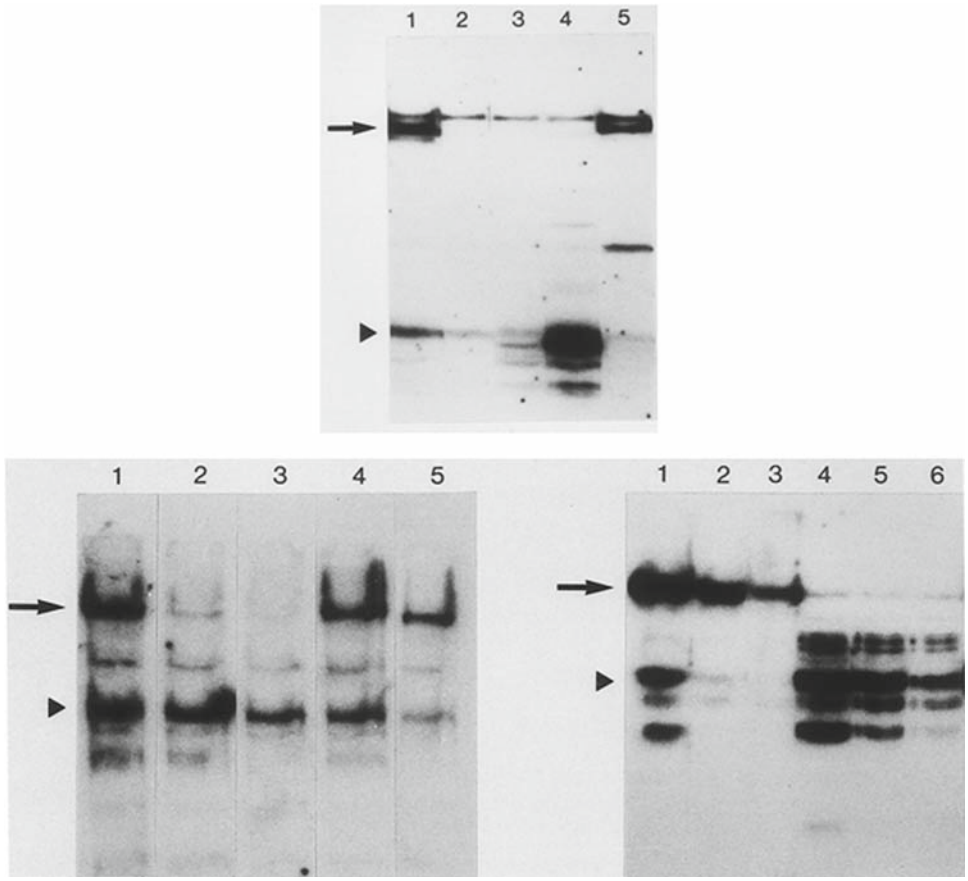


Fig. 1. Southwestern analysis of mouse L-50 cell crude nuclear extracts. Approximately 200 μg of extract was resolved by SDS/PAGE, transferred to a PVDF membrane, and incubated with radiolabeled DNAs. *Upper panel:* Binding of labeled oligos corresponding to control and mutant metal regulatory elements (MREs) of the mouse metallothionein 1 gene promoter. Lanes: 1, control MREd; 2, Mutant-a; 3, Mutant-b; 4, Mutant-c; 5, Mutant-d. *Left lower panel:* Analysis of the binding specificity of the MREd oligo to nuclear proteins using MREd as the probe and cold MREd (*lanes 1–3*) or Mutant-a (*lanes 4* and *5*) as competitor. *Right lower panel:* Competition experiments using heterologous probe and nonspecific nucleic acids as competitors. Probes: Lanes, 1–3, MREd; 4–6, a 309-bp-pBR322 *Msp*I fragment. Nonspecific nucleic acid competition cocktail. Lanes: 1 and 4, 0 μg; 2 and 5, 0.8 μg/mL; 3 and 6, 4 μg/mL. The *arrow* corresponds to M_r 108,000 and the *arrowhead* indicates the position of M_r 45,000. (Reproduced from ref. *26* with permission from Oxford University Press).

3. Tris–HCl 0.5 M, pH 6.8.
4. Sodium Dodecyl Sulfate (SDS) 10% and 2-*bis* mercaptoethanol.
5. Tris–Glycine buffer 10×: 250 mM Tris base, 1.92 M Glycine, pH 8.3. To make this up dissolve 30 g Tris base, 144 g Glycine in 1 L of water (*see* **Note 1**).
6. Electrophoresis running buffer: Tris–Glycine buffer 2× containing 1% SDS.
7. Transfer buffer: Tris–Glycine buffer 1×.
8. Binding buffer 1×: 20 mM Hepes of pH 7.9, 5 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol (DTT) (*see* **Notes 2** and **3**).

9. Sample loading buffer 2×: 125 mM Tris of pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-*bis* mercaptoethanol, 0.025% (w/v) bromophenol blue. Make 1-mL aliquots and store at –20°C.
10. Acrylamide (50:1 acrylamide:*N*′,*N*′-methylene *bis*-acrylamide in water) (*see* **Note 4**). Make up a stock solution at 50% (w/v) from 250 g of acrylamide and 5 g of bisacrylamide with water to bring the final volume to 500 mL, and then stir until dissolved. Deionize the acrylamide solution with a mixed-bed resin *(22)* and store at 4°C. N.B.: *Acrylamide is a potent neurotoxin. Wear gloves and a facemask when handling the dry powder.*
11. Ammonium persulfate, 10% (w/v) in water (prepared weekly, and stored in aliquots at –20°C), and *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED).
12. Methanol and isobutanol.
13. Blocking buffer – 5%: 5% (w/v) nonfat dry milk, 0.01% (v/v) Antifoam A emulsion (Sigma) in 1× binding buffer.
14. Blocking buffer – 0.25%: 0.25% (w/v) nonfat dry milk, 0.01% (v/v) Antifoam A emulsion in 1× binding buffer.
15. End-labeled oligonucleotides (oligos) or DNA fragments of high specific activity. Standard procedure *(22)* can be followed in the preparation of DNA samples (*see* **Note 5**).
16. Poly(dA–dT)·poly(dA–dT), poly(dI–dC)·poly(dI–dC), calf thymus DNA, salmon sperm DNA, *E.coli* DNA.
17. ^{14}C-labeled protein standards (5 μCi/mL, Amersham) and prestained protein standards such as the Kaleidoscope Prestained Standards from BioRad.
18. Vertical electrophoresis apparatus with a central cooling core (e.g., Protean II slab cell, BioRad), and glass plates (*see* **Note 6**).
19. Electroblotting apparatus with cooling coil (e.g., Trans-Blot transfer cell, BioRad).
20. Electrophoresis dc power supply capable of 200 V and 2 A (e.g., model 200/2.0 of BioRad).
21. Recirculating water chiller apparatus.
22. Immobilon®-PVDF membranes (Millipore).
23. Whatman 3-MM paper.

3. Methods

The Southwestern procedure is conducted in three stages: crude nuclear protein extracts or purified or partially purified protein preparations are separated by electrophoresis on an

SDS–polyacrylamide gel, transferred to PVDF membranes, and assessed for their ability to bind to an oligo corresponding to a specific DNA *cis*-acting regulatory element, or to a DNA promoter fragment.

1. Prepare an 8% acrylamide–bisacrylamide (50:1) separating gel (*see* **Note 4**). To make this, mix 25 mL 1.5 M Tris–HCl of pH 8.8, 15.7 mL acrylamide, and 57.2 mL water. Filter through a 0.45-μm membrane and then add 0.5 mL SDS 10%, 0.5 mL APS 10%, and 35 μL TEMED. Mix gently by inversion and pour the gel. Slowly add isobutanol over the acrylamide solution to make a flat surface, and allow the gel to set for at least 1 h. Remove the isobutanol, wash thoroughly with water, and add a 4% acrylamide–bisacrylamide (30:0.8) stacking gel. Mix 6.25 mL 0.5 M Tris–HCl of pH 6.8, 3.35 mL acrylamide–bisacrylamide (30:0.8), 15 mL water. Filter and add 0.25 mL SDS 10%, 0.25 mL APS 10%, and 17 μL TEMED. Mix gently, pour over the separating gel, add a comb, and allow the gel to set for at least 1 h. Carefully wash the wells with running buffer using a syringe.
2. Dilute equal volumes of protein preparation and loading buffer 2×. We typically use 100–400 μg of crude nuclear extracts prepared according to Dignam *(23)* in presence of a cocktail of four protease inhibitors, i.e., leupeptin, PMSF, antipain, and chemostatin A. Instead of dealing with aliquots of different protease inhibitors, Roche's Complete Protease Inhibitor Cocktail Tablets is a modern and convenient alternative. Lower amounts are required when using purified or partially purified protein preparations. Mix loading buffer and samples just before loading to avoid SDS precipitation. We do not boil or denature the protein preparations (*see* **Note 7**). Include in a separate lane 15 μL of ^{14}C-labeled protein markers, and run in electrophoresis electrode buffer (Tris–Glycine buffer 2×, 10% SDS) overnight at 75 V constant. Do not pre-run the gel (*see* **Note 8**). We perform the electrophoresis at 4°C and connect the cooling coil provided with the apparatus to a recirculating water chiller.
3. Electrophoresis gels should be pre-equilibrated in transfer buffer prior to commencement of electrophoretic transfer. Pre-equilibration will aid the removal of SDS and contaminating electrophoresis salts from the proteins, and will facilitate subsequent renaturation to the native conformation (*see* **Note 9**). In addition, the pretransfer step will allow any changes in the size of the gel due to the swelling or shrinking to occur at this stage rather than during the transfer. Thus, after the gel is run, dismantle the apparatus, remove the gel from the plates, and soak it in transfer buffer for 60 min with gentle agitation. We usually perform this operation in the cold room (*see* **Note 8**). Wear clean disposable gloves.

4. Layer the gel onto 3-MM paper prewetted with transfer buffer. Always keep the gel wet by regularly pouring buffer over it.
5. Prewet the PVDF membrane in methanol for few seconds and transfer it to water.
6. Start mounting the protein blotting sandwich by successively placing inside the holding cassette the first foam sponge pad provided with the blotting transfer apparatus, and the 3-MM paper on which the gel was placed (*see* **Fig. 2** and **Note 10**).
7. Layer the membrane over the gel.
8. Carefully remove all the air bubbles that could be trapped between the gel and the membrane. We use a 10-mL glass pipette that we gently roll over the membrane.
9. Put a second prewetted (with transfer buffer) 3-MM paper and repeat **step 8**. Put the other foam pad, close the cassette, and introduce it into the transfer apparatus filled with cold transfer buffer. Be sure that the gel and the membrane are facing toward the positive electrode. Place the tank transfer unit on stirring plate and connect it to a circulating chiller apparatus (*see* **Note 11**).
10. It is important that a stirring bar be placed inside the transfer cell and that the transfer buffer be stirred during the course of the experiment. This will help to maintain uniform conductivity and temperature during electrophoretic transfer. Failure to do so will result in poor transfer of proteins and may pose a safety hazard.
11. Transfer at 50 V constant for 3 h.
12. After the transfer is completed, dismantle the apparatus, remove the membrane from the gel, and successively place it in: (1) 200 mL of binding buffer in a plastic dish for 15 min with gentle agitation, (2) 250 mL blocking buffer – 5% for a minimum of 1 h with agitation, (3) 150 mL blocking buffer – 0.25% containing 10^6 cpm/mL of probe, overnight

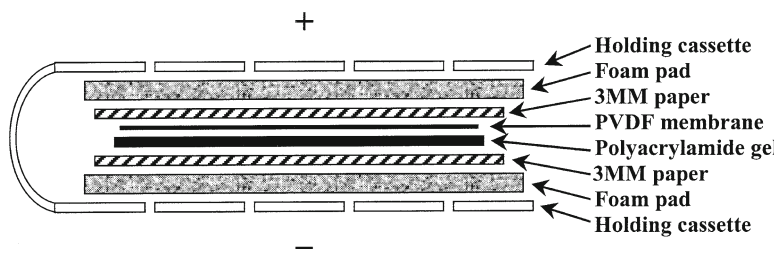


Fig. 2. Schematic representation of the protein blotting sandwich assemblage used in Southwestern experiments. The orientation of the sandwich in relation to the cathode (+) and the anode (–) is indicated.

with agitation (*see* **Note 12**). We never let the membrane dry between the transfer and the binding steps.

13. The gel can be transferred in 50% methanol for subsequent silver staining and visualization of the proteins remaining on the gel after transfer (*see* **Note 13**).
14. Wash the membrane by successively soaking it with agitation in: (1) blocking buffer – 0.25%, 30 s; (2) blocking buffer – 0.25%, 20 min; (3) blocking buffer – 0.25%, 20 min; (4) binding buffer 1×, 20 min. Recover wash liquids since they contain ^{32}P.
15. Let the membrane dry on 3-MM paper at room temperature for 1 h, and autoradiograph it with an X-ray film.

4. Notes

1. Tris–Glycine buffers should not be pH adjusted.
2. DTT should be stored at −20°C, thawed on ice, and added when the solution is ice cold.
3. Several DNA-binding proteins require zinc ions to efficiently bind to DNA and, as a precaution, we routinely add 5 μM $ZnCl_2$ to all solutions (loading, electrophoresis, transfer, binding, and wash buffers).
4. The percentage of gel and the ratio between acrylamide and bisacrylamide will depend on the M_r of the protein of interest. We use a ratio of 50:1 to increase resolution of the high molecular weight protein species. A ratio of 30:0.8 or 29:1 may be more appropriate for average protein species. Plates are cleaned with phosphate-free soap and rinsed with deionized water, Milli Q water, and ethanol 95%. We usually do not silanize the plates with dichlorodimethylsilane as we observed that it could interfere with subsequent steps.
5. Synthetic oligos or DNA fragments can be conveniently 5′ labeled with T4 kinase *(22)*. Radiolabeled probe with high specific activity increases the sensitivity of detection and reduces background *(24)*. A nonradioactive Southwestern procedure, using chemiluminescent detection, has also been described *(25)*.
6. The procedure can indifferently be carried out on standard electrophoresis and transfer apparatus or on small-sized apparatus such as the Mini-protean II and mini-Trans-Blot transfer cell of BioRad *(26)*.
7. We do not boil samples before electrophoresis in order to avoid irreversibly denaturing proteins or affect the DNA-binding properties of more sensitive protein species *(3)*.

8. Depending on the biochemical properties of the protein under investigation and of its stability at room temperature, it is recommended to perform the whole procedure at 4°C. Particular conditions of temperature may need to be determined for an optimal binding of different factors to DNA.
9. Although the extent of renaturation is not known, this wash by removing SDS from the proteins permits functional recognition of the DNA-binding domain. Triton X-100 *(1, 27)*, urea *(2, 4, 6, 8)*, or guanidium HCl *(7, 28)* can be included in the transfer buffer to promote the removal of SDS and to facilitate subsequent renaturation of the proteins. Low molecular weight proteins (≤ 10 kDa) may diffuse out of gels more readily during the pre-equilibration step. To avoid this, one can change the pre-equilibration buffer several times during a shorter pre-equilibration period. Methanol (20%) can be added to the transfer buffer *(4, 28, 29)*. However, this solvent may affect the DNA-binding properties of the proteins. The choice of renaturation procedure and buffers should be determined empirically and will depend on the factors studied.
10. Each layer of the sandwich is thoroughly prewetted with buffer and then carefully positioned on top of the previous layer, taking care to avoid trapping any air bubbles which would distort the resulting transfer.
11. Placing the Trans-Blot cell in the cold room is an inadequate means of controlling transfer buffer temperature. Efficient heat removal is obtained by connecting the transfer cell to a recirculating water chiller. Transfer is a function of molecular weight, with the largest proteins being transferred more slowly. The precise transfer conditions have to take into account the molecular weight range of the proteins under investigation and the composition of the gel.
12. It is important to block any remaining free binding sites on the membrane in order to reduce nonspecific binding of the radiolabeled substrate, which would otherwise lead to a high background in subsequent steps. The binding buffer should thus be designed to optimize the specific binding of the radiolabeled probe, while keeping nonspecific binding to a minimum. Nonspecific binding can be reduced by the addition of an excess of unlabeled DNA such as poly(dI–dC)·poly(dI–dC) or poly(dA–dT)·poly(dA–dT), or sheared *E. coli*, salmon sperm, or calf thymus DNA. Nonspecific binding can also be reduced by increasing salt concentrations in the binding reaction and/or in the wash buffer. In addition, the use of lipid-free BSA instead of nonfat dry milk has been reported to optimize the signal-to-noise ratio *(30)*. While specific high-affinity DNA–protein interactions should not be competed

out with nonspecific DNA, on the addition of large excess of cold nonspecific DNA even specific complexes will dissociate. Thus, the amount of nonspecific competitor DNA that needs to be added and the concentration of salt to be used in the buffers to remove nonspecific signals will depend on the specificity of the interaction. The specificity and affinity of the binding should be further examined by adding unlabeled specific control or mutated competitor DNA to the blocking and/or binding solutions. Alternatively, a nonfunctional mutant DNA molecule can be used as the radiolabeled substrate (*see* **Fig. 1**) *(26)*.

13. The DNA-binding domain of some proteins may not retain its ability to bind selectively with its DNA response element after immobilization of the protein on the membrane. Furthermore, the membrane will not necessarily retain all the proteins that will have migrated out the gel. It is thus useful to determine which proteins have been transferred to the membrane. Detection of the transferred proteins may be achieved after autoradiography by staining the membrane with either Coomassie brilliant blue R-250, Amido black, or Red Ponceau. Proteins remaining on the gel after transfer can be visualized by silver staining.
14. Crude nuclear extracts prepared from human HeLa and mouse L cells contain three common nonspecific DNA-binding proteins that often contaminate preparations of affinity-purified factors, poly (ADP–ribose) polymerase (PARP) which has an M_r of 116,000 (as well as a typical proteolytic fragment of M_r 60,000) *(31)*, the Ku antigen which consists of two polypeptides of M_r 70,000 and 80,000 *(32)*, and replication protein A (RP-A) of M_r 74,000 *(33)*. The heat shock protein hsp70 (M_r 70,000) that can stick to proteins that are improperly folded may also cause problems. Thus, when Southwestern analyses are performed with affinity-purified nuclear proteins, it is recommended to be suspicious of polypeptides of M_r 60,000; 70,000–74,000; 80,000; and 110,000 and to perform some control experiments. For instance, in the case of PARP, if a polypeptide with an M_r in the range of this protein is detected in the Southwestern procedure, a Western analysis could be performed on an aliquot of the protein preparation using anti-PARP antibodies available commercially (Roche Diagnostics/Boehringer). Alternatively, a chromatographic step with Red Agarose (BioRad) which binds PARP with high affinity can be performed to get rid of this possible contaminant *(31)*. However, one should bear in mind that PARP has recently been shown to act as a coactivator for a number of transcription factors, including AP-1 *(34)*, AP-2 *(35)*, TEF-1 *(36)*, and NF-κB *(37)*, and

thus its detection using a specific DNA probe corresponding to the binding site of a known transcription factor may be functionally relevant.

Acknowledgments

This work was supported by a Grant from the Conseil de recherches en sciences naturelles et en génie du Canada to CS.

References

1. Bowen B., Steinberg J., Laemmli U. K., and Weintraub H. (1980). The detection of DNA-binding proteins by protein blotting. *Nucleic Acids Res* **8**, 1–20.
2. Jack R. S., Brown M. T., and Gehring W. J. (1983). Protein blotting as a means to detect sequence-specific DNA-binding proteins. *Cold Spring Harb Symp Quant Biol* **47**, 483–491.
3. Miskimins W. K., Roberts M. P., McClelland A., and Ruddle F. H. (1985). Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferrin receptor gene. *Proc Natl Acad Sci USA* **82**, 6741–6744.
4. Silva C. M., Tully D. B., Petch L. A., Jewell C. M., and Cidlowski J. A. (1987). Application of a protein-blotting procedure to the study of human glucocorticoid receptor interactions with DNA. *Proc Natl Acad Sci USA* **84**, 1744–1748.
5. Wegenka U. M., Buschmann J., Lutticken C., Heinrich P. C., and Horn F. (1993). Acute-phase response factor, a nuclear factor binding to acute-phase response elements, is rapidly activated by interleukin-6 at the posttranslational level. *Mol Cell Biol* **13**, 276–288.
6. Kwast-Welfeld J., Debelle I., Walker P. R., Whitfield J. F., and Sikorska M. (1993). Identification of a new cAMP response element-binding factor by Southwestern blotting. *J Biol Chem* **268**, 19581–19585.
7. Villafuerte B. C., Zhao W., Herington A. C., Saffery R., and Phillips L. S. (1997). Identification of an insulin-responsive element in the rat insulin-like growth factor-binding protein-3 gene. *J Biol Chem* **272**, 5024–5030.
8. Wu J., Jiang Q., Chen X., Wu X. H., and Chan J. S. (1998). Identification of a novel mouse hepatic 52 kDa protein that interacts with the cAMP response element of the rat angiotensinogen gene. *Biochem J* **329**, 623–629.
9. Chen A., and Davis B. H. (1999). UV irradiation activates JNK and increases alphaI(I) collagen gene expression in rat hepatic stellate cells. *J Biol Chem* **274**, 158–164.
10. Ashizawa M., Miyazaki M., Abe K. et al. (2003). Detection of nuclear factor-κB in IgA nephropathy using Southwestern histochemistry. *Am J Kidney Dis* **42**, 76–86.
11. Scassa M. E., Guberman A. S., Ceruti J. M., and Canepa E. T. (2004). Hepatic nuclear factor 3 and nuclear factor 1 regulate 5-aminolevulinate synthase gene expression and are involved in insulin repression. *J Biol Chem* **279**, 28082–28092.
12. Xing C., LaPorte J. R., Barbay J. K., and Myers A. G. (2004). Identification of GAPDH as a protein target of the saframycin antiproliferative agents. *Proc Natl Acad Sci USA* **101**, 5862–5866.
13. Wei C. C., Guo D. F., Zhang S. L., Ingelfinger J. R., and Chan J. S. D. (2005). Heterogenous nuclear ribonucleoprotein F modulates angiotensinogen gene expression in rat kidney proximal tubular cells. *J Am Soc Nephrol* **16**, 616–628.
14. Katsu Y., Yamashita M., and Nagahama Y. (1997) Isolation and characterization of goldfish Y box protein, a germ-cell- specific RNA-binding protein. *Eur J Biochem* **249**, 854–861.
15. Johnston K. A., Polymenis M., Wang S., Branda J., and Schmidt E. V. (1998). Novel regulatory factors interacting with the promoter of the gene encoding the mRNA cap binding protein (eIF4E) and their function in growth regulation. *Mol Cell Biol* **18**, 5621–5633.

16. Hamann S., and Stratling W. H. (1998). Specific binding of Drosophila nuclear protein PEP (protein on ecdysone puffs) to hsp70 DNA and RNA. *Nucleic Acids Res* **26**, 4108–4115.
17. Keller A. D., and Maniatis T. (1991). Selection of sequences recognized by a DNA binding protein using a preparative Southwestern blot. *Nucleic Acids Res* **19**, 4675–4680.
18. Singh H., Clerc R. G., and LeBowitz J. H. (1989). Molecular cloning of sequence-specific DNA binding proteins using recognition site probes. *Biotechniques* **7**, 252–261.
19. Stuempfle K. J., and Floros J. (1997). Caution is advised when cDNA expression libraries are screened by Southwestern methodologies. *Biotechniques* **22**, 260–264.
20. Fukuda I., Nishiumi S., Yabushita Y., et al. (2004). A new Southwestern chemistry-based ELISA for detection of aryl hydrocarbon receptor transformation: application to the screening of its receptor agonists and antagonists. *J Immunol Methods* **287**, 187–201.
21. Siu F. K., Lee L. T., and Chow B. K. (2008). Southwestern blotting in investigating transcriptional regulation. *Nature Protoc* **3**, 51–58.
22. Sambrook J., Fritsch E. F., and Maniatis T. (1989). *Molecular cloning. A laboratory manual.* 2nd edn. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
23. Dignam J. D. (1990). Preparation of extracts from higher eukaryotes. *Methods Enzymol* **182**, 194–203.
24. Handen J. S., and Rosenberg H. F. (1997). An improved method for Southwestern blotting. *Front Biosci* **2**, c9–c11.
25. Dooley S., Welter C., and Blin N. (1992) Nonradioactive Southwestern analysis using chemiluminescent detection. *Biotechniques* **13**, 540–543.
26. Séguin C., and Prévost J. Detection of a nuclear protein that interacts with a metal regulatory element of the mouse metallothionein 1 gene. *Nucleic Acids Res* **16**, 10547–10560.
27. Melkonyan H., Hofmann H. A., Nacken W., Sorg C., and Klempt M. (1998). The gene encoding the myeloid-related protein 14 (MRP14), a calcium- binding protein expressed in granulocytes and monocytes, contains a potent enhancer element in the first intron. *J Biol Chem* **273**, 27026–27032.
28. Dhawan P., Chang R., and Mehta K. D. (1997). Identification of essential nucleotides of the FP1 element responsible for enhancement of low density lipoprotein receptor gene transcription. *Nucleic Acids Res* **25**, 4132–4138.
29. Schaufele F., Cassill J. A., West B. L., and Reudelhuber T. (1990). Resolution by diagonal gel mobility shift assays of multisubunit complexes binding to a functionally important element of the rat growth hormone gene promoter. *J Biol Chem* **265**, 14592–14598.
30. Papavassiliou A. G., and Bohmann D. (1992). Optimization of the signal-to-noise ratio in south-western assays by using lipid-free BSA as blocking reagent. *Nucleic Acids Res* **20**, 4365–4366.
31. Mazen A., Ménissier-de Murcia J., Molinete M., et al. (1989). Poly(ADP-ribose) polymerase: a novel finger protein. *Nucleic Acids Res* **12**, 4689–4698.
32. Mimori T., Hardin J. A., and Steitz J. A. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J Biol Chem* **261**, 2274–2278.
33. Wold M. S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA- binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**, 61–92.
34. Andreone T. L., O'Connor M., Denenberg A., Hake P. W., and Zingarelli B. (2003). Poly(ADP-Ribose) polymerase-1 regulates activation of activator protein-1 in murine fibroblasts. *J Immunol* **170**, 2113–2120.
35. Kannan P., Yu Y., Wankhade S., and Tainsky M. A. (1999). PolyADP-ribose polymerase is a coactivator for AP-2-mediated transcriptional activation. *Nucleic Acids Res* **27**, 866–874.
36. Butler A. J., and Ordahl C. P. (1999). Poly(ADP-ribose) polymerase binds with transcription enhancer factor 1 to MCAT1 elements to regulate muscle-specific transcription. *Mol Cell Biol* **19**, 296–306.
37. Nakajima H., Nagaso H., Kakui N., Ishikawa M., Hiranuma T., and Hoshiko S. (2004). Critical role of the automodification of poly(ADP-ribose) polymerase-1 in nuclear factor-{kappa} B-dependent gene expression in primary cultured mouse glial cells. *J Biol Chem* **279**, 42774–42786.

Chapter 13

Footprinting DNA–Protein Interactions in Native Polyacrylamide Gels by Chemical Nucleolytic Activity of 1,10-Phenanthroline–Copper

Athanasios G. Papavassiliou

Summary

Various methodologies have been developed for the detection of DNA-binding activities and the identification of the "footprints" of a protein on DNA. The most widely used footprinting techniques employ reagents such as deoxyribonuclease I (DNase I) and dimethyl sulfate (DMS) for protection analysis in solution. Nevertheless, these techniques have several disadvantages, and although these may be bypassed by coupling the footprinting reaction with an electrophoretic mobility-shift assay (EMSA), the size and the sequence specificity of DNase I and DMS as well as the problem of protein exchange during the footprinting reaction pose significant limitations. These limitations can be circumvented by combining the advantages of EMSA, with the subsequent exposure of the resolved DNA–protein complex(es) to the chemical nuclease 1,10-phenanthroline–copper ion (OP–Cu) while they are still embedded in the polyacrylamide matrix (*in-gel assay*).

Key words: DNA-binding activity, DNA–protein complex, Electrophoretic mobility-shift assay, In-gel footprinting, Chemical nuclease, 1,10-Phenanthroline–copper.

1. Introduction

The existence of cell-type specific promoter and enhancer elements has been known for several years. However, the mechanisms responsible for the remarkable specificity of such elements, in comparison to the ubiquitously active promoters and enhancers of "house-keeping" genes and DNA tumor viruses, have remained elusive until recently. Although transfection and mutagenesis experiments have taught us a great deal about the structure

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_13

of cell type-specific cis-acting elements, the breakthrough in understanding the molecular basis for the differential activity of these elements has come from the analysis of their recognition by specific DNA-binding proteins.

Several techniques have been developed for the detection of cell type-specific DNA-binding activities and the identification of sequence-specific contacts ("footprints") of a protein on DNA. Many of these techniques involve forming DNA–protein complexes (by incubating a asymmetrically labeled double-stranded DNA fragment containing the region of interest with a crude or partly purified protein extract), exposing the complex to enzymatic or to chemical reagents that can cleave or modify the DNA and determining which bases are protected from attack when the protein(s) is bound. The most widely used reagents are deoxyribonuclease I (DNase I, an endonuclease) and dimethyl sulfate (DMS). Footprinting experiments with DNase I are performed using parallel reactions with free DNA and with DNA–protein complexes, and the nuclease is allowed to digest DNA only to a limited extent *(1)*. The DNA is then purified and denatured, and the single-stranded end-labeled DNA fragments are resolved on a sequencing gel and autoradiographed. Comparing the digestion pattern of the DNA–protein complexes revealed on the autoradiograph with that of free DNA shows a band-free region (footprint) where the bound protein(s) has prevented access of the enzyme to DNA (*see* Chapter "DNase I Footprinting"). In a similar analysis, the DNA is allowed to react mildly with DMS, which methylates primarily deoxyguanosine residues and renders their phosphodiester linkages labile under conditions of Maxam–Gilbert chemistry (*see* Chapter "Identification of Protein/DNA Contacts with Dimethyl Sulfate: Methylation Protection and Methylation Interference"). The binding of a protein(s) to a specific DNA region will result in a protection of the corresponding bases from chemical modification *(2)*.

The suitability of the above assays in determining the binding sequences of proteins on DNA is hindered by several disadvantages. First, the clarity of the footprint is highly dependent on the extent of occupancy of the binding site(s) (i.e., a "clear" footprint is observed only if all DNA molecules are involved in complexes). Unfortunately, this is not always easy to achieve, especially when the concentration and/or purity of the specific binding protein(s) is not satisfactory. Second, DNA–protein complexes formed in crude extracts may often be heterogeneous in terms of both binding specificity and kinetic stability. Therefore, direct footprinting in solution will not correspond to a single species, but instead reflect an "integral" of the multiple equilibria operating over the entire region of interest (i.e., the protection pattern will actually represent a composite of more than one complex, with complexes having a very low dissociation rate dominating the footprint). Finally,

two different proteins that recognize the same sequence within the probe are most likely to yield indistinguishable footprints. These drawbacks may be overcome by coupling treatment with a footprinting reagent in solution with the electrophoretic mobility-shift assay (EMSA; also known as gel retardation assay, *see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions") *(3–5)*. In this approach, the protein and DNA molecules are incubated together, and the equilibrated reaction mixture is exposed to DNase I or DMS as before. DNA–protein complexes are subsequently isolated from the free probe by electrophoresis in a nondenaturing polyacrylamide gel. While the negatively charged free DNA migrates rapidly toward the anode, once it is bound by a specific protein its mobility decreases *(3, 4)*. Following the separation of the free and bound DNA species, the corresponding bands are cut out of the gel, and the DNA eluted and analyzed on a sequencing gel. The region(s) of protection evident in the DNA derived from the complexed fraction indicates the binding site *(5)*. Since the complexes are separated from contaminating unbound DNA fragments, their footprints will be free of background cutting, and thus considerably more evident. Similar considerations apply when more than one complex can be formed on the fragment. As long as the DNA-binding proteins differ in their molecular masses and charges, they will cause altered electrophoretic mobilities of the corresponding complexes and, hence, different migrations in the native polyacrylamide gel. These complexes can be isolated and run in individual lanes on the sequencing gel. Thus, the exposure of the binding reaction to footprinting reagents, in combination with the fractionation offered by mobility-shift gels, permits the identification of the regions of DNA bound by protein in different complexes, even if a low percentage of the initial DNA molecules have been complexed.

Although by employing the EMSA one can substantially increase the sensitivity of DNase I or DMS footprinting experiments in solution, several additional problems have still to be faced:

(a) DNAse I is a relatively bulky molecule (MW 30,400) that cannot cleave the DNA in the immediate vicinity of a bound protein because of steric hindrance. As a result, the region(s) protected from cutting extends beyond the actual protein-binding site.

(b) The nonrandom nature of DNA cleavage by DNase I makes it impossible to assess the involvement in protein binding of nucleotides that lie in an area of the fragment resistant to the endonucleolytic activity of this enzyme [e.g., tracts of A and T residues, or TpA (as opposed to ApT) dinucleotide islands scattered within or adjacent to the binding site], so that binding sites or parts of binding sites may not be detected.

(c) The primary site of reaction of DMS with B-DNA is the N-7 atoms of guanine bases that are located in the major groove.

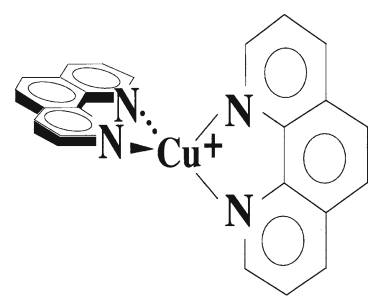
Thus, those guanines in close proximity with the protein will be protected from methylation. However, if a protein primarily makes contacts with a DNA sequence in the minor groove, or if there are no guanine residues in a major groove-binding site, DMS will not reveal these interactions.

(d) In many instances, particularly when a complex has a relatively high "off"-rate, the bound protein can dissociate from the protected DNA fragment and reassociate to other DNA fragments that have already been nicked by DNase I, or modified by DMS. In this case, the DNA-cleavage pattern derived from the complexed fraction will closely resemble that of the uncomplexed DNA, rendering it difficult to observe a footprint. The limitations imposed by the size and the sequence, or base specificity of the aforementioned footprinting reagents, as well as the problem of protein exchange from the binding site(s) during treatment, are circumvented by merging the advantages inherent in the EMSA, with the *subsequent* exposure of the gel (hence of the resolved complexes while embedded in the polyacrylamide matrix) to a chemical DNA-scission reagent, namely 1,10-phenanthroline–copper ion (OP–Cu) *(6)*.

1.1. OP–Cu as a Footprinting Agent

1.1.1. Chemistry of DNA Cleavage

OP–Cu (**Fig. 1**) is an efficient chemical nuclease that cleaves the phosphodiester backbone of nucleic acids at physiological pH and temperature by oxidation of the deoxyribose (DNA) or ribose (RNA) moiety *(7)*. The kinetic scheme of the reaction is summarized in **Fig. 2**. The first step is the formation of the 1,10-phenanthroline–cupric ion coordination complex, under conditions that favor the 2:1 stoichiometry ($[OP]_2Cu^{2+}$). The DNA-scission process is initiated by adding a reducing agent, usually 3-mercaptopropionic acid (a thiol), to the aerobic reaction mixture containing the target DNA. Under these conditions, the 2:1 cupric complex is reduced to the 2:1 cuprous complex ($[OP]_2Cu^{+}$) that is in turn oxidized by molecular oxygen to generate hydrogen



1,10-phenanthroline·Cu(I)

Fig. 1. Structure of 1,10-phenanthroline complexed with copper (I) ion (OP–Cu).

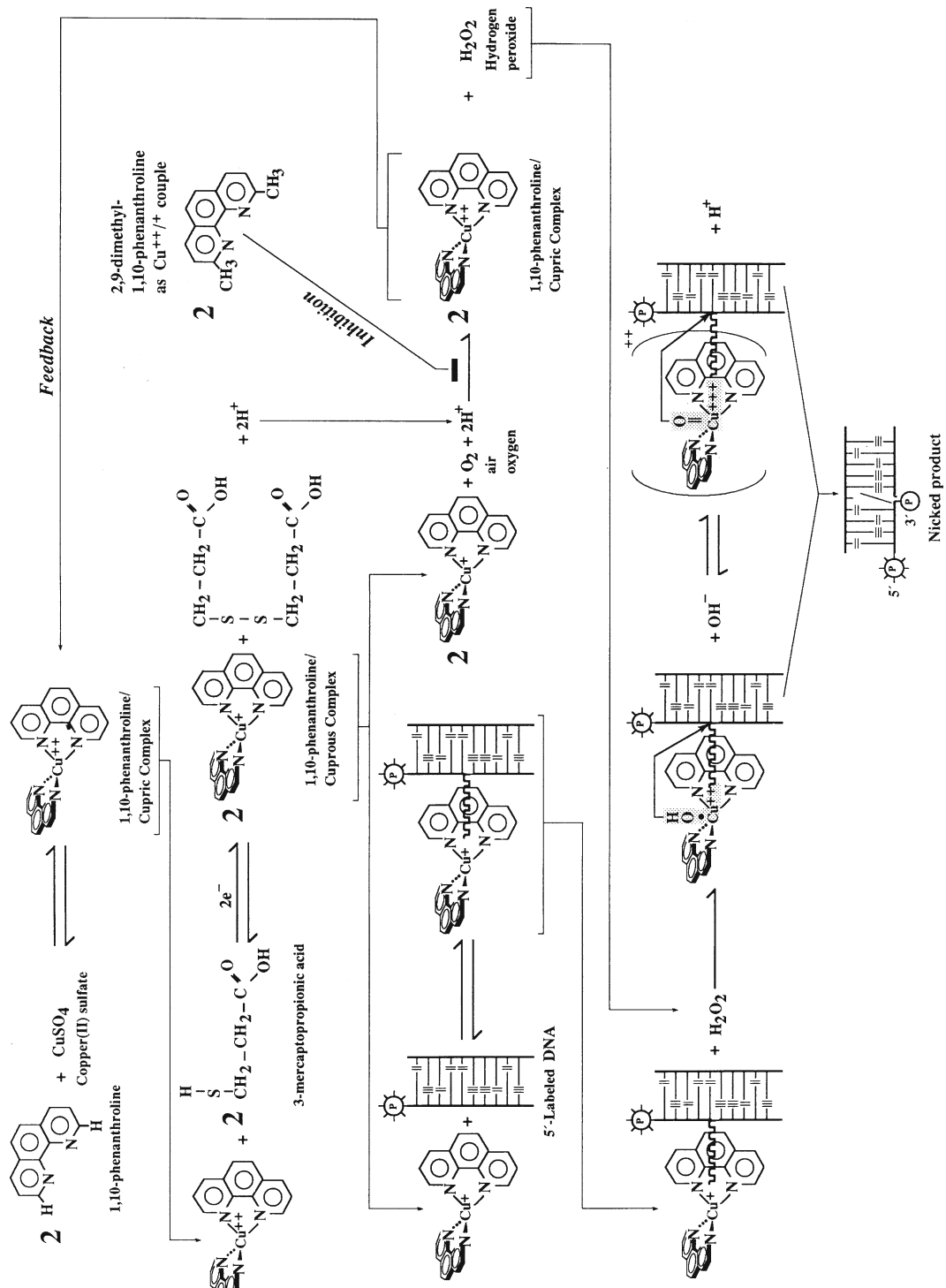


Fig. 2. Schematic representation of the kinetic mechanism for the nuclease activity of 1,10-phenanthroline–copper ion.

peroxide. Hydrogen peroxide is an essential co-reactant for the chemical nuclease activity and can be generated as described above or added exogenously *(8)*. The tetrahedral cuprous complex, present at the steady-state concentration defined by the experimental conditions (note the feedback mechanism in Fig. 2), then binds reversibly to the minor groove of DNA to form a central intermediate through which the reaction is funneled *(9)*. The DNA-bound cuprous complex undergoes in situ an one-electron oxidation by hydrogen peroxide to form a short-lived, highly reactive DNA-bound copper-oxo species that can be written either as a hydroxyl radical coordinated to the cupric ion or as a copper–oxene structure (**Fig. 2**). This species then attacks the H1′-deoxyribose protons of nucleotides, which are accessible in the minor groove; this reaction initiates a series of reactions culminating in cleavage of the phosphodiester backbone *(9)*. Reaction rates at any given sequence position depend on the stability of the intermediate formed between DNA and $(OP)_2Cu^+$ and on the orientation and proximity of the copper-oxo species relative to the C1′-deoxyribose hydrogen in the minor groove. Because both criteria are met satisfactorily in B-DNA sequences, the tetrahedral cuprous complex prefers B-DNA as its substrate. Such stereoelectronic interactions are less efficient in the broad minor groove of A-DNA and not possible in Z-DNA, which has practically no minor groove; as a result, A-DNA is cleaved at 25–33% of the rate with which B-DNA is cleaved and Z-DNA is not cleaved at all *(9)*. The products of the strand-scission event include the free base, DNA fragments bearing 5′- and 3′-phosphorylated termini, and the deoxyribose oxidation product 5-methylene-2-furanone *(10)*. The DNA-chain cleavage reaction can be efficiently quenched by adding to the mixture 2,9-dimethyl-1,10-phenanthroline (2,9-dimethyl-OP). This phenanthroline derivative can also chelate copper ions to form a minor groove-associated cuprous complex (thus competing with $[OP]_2Cu^+$), but the reduction potential of the Cu^{2+}/Cu^+ couple is too positive to allow significant nuclease activity under normal assay conditions *(11)*.

1.1.2. OP–Cu Footprinting Following EMSAs

In as much as the structural and functional properties of DNA are not altered by entrapment in a polyacrylamide gel matrix *(6)*, the small size and the ready diffusibility of all reaction components in solid supports permit the coupling of OP–Cu footprinting with the EMSA to study DNA–protein interactions *(12, 13)*. In this method, the DNA-binding reaction is performed as usual, electrophoresed under established, nondenaturing conditions, and the entire mobility-shift gel is immersed in a footprinting reaction mixture containing 1,10-phenanthroline, cupric ion, and 3-mercaptopropionic acid. Following the reaction quench with 2,9-dimethyl-OP, footprints are obtained after elution of

the radioactive free and protein-bound DNA cleavage products from the mobility-shift gel and analysis on a sequencing gel (**Fig. 3**). Because the nuclease activity of $(OP)_2Cu^+$ produces 3′-phosphorylated and 5′-phosphorylated ends as cleavage products, sequencing gels can be accurately calibrated with the Maxam–Gilbert sequencing reactions.

1.2. Advantages of OP–Cu Over Other Footprinting Agents

1.2.1. General Considerations

The nuclease activity of $(OP)_2Cu^+$ bears several advantages as a footprinting reagent relative to protection analyses using DNase I or DMS as a probe. First, the $(OP)_2Cu^+$ chelate is a small molecule (compared to DNase I) that permits cleavage closer to the edge of the DNA sequence protected by protein binding, and therefore, a more precise definition of it. Second, since the scission chemistry involves attack on the deoxyribose moiety, $(OP)_2Cu^+$ is able to cut at all sequence positions regardless of base. However, the intensity of cutting (rate of cleavage) does depend on local sequence, with attack at adenines of TAT triplets being most preferred [*(14)*, *see* **Fig. 3**]. Interestingly, a preference for C-3′,5′-G steps, rather than T-3′,5′-A steps, is observed at a phenanthroline to copper ratio of 1:1, which strongly favors formation of the $OPCu^+$ complex *(15)*. Nevertheless, the cutting patterns obtained with $(OP)_2Cu^+$ are usually sufficiently well-defined to identify protected regions, even though this endonucleolytic agent exhibits some degree of sequence specificity in its rate of cleavage of naked DNA. Third, because $(OP)_2Cu^+$ binds to the minor groove of DNA it will reveal minor-groove interactions. Since the binding of the coordination complex should be restricted to three base pairs, the complex is more sensitive to local, protein-induced conformational changes than DNase I, which by possessing an extended minor groove-binding site, may be unable to sense. In this context, the complex will also detect binding in the major groove when its approach to its minor groove-binding site is sterically blocked, or if the interaction of the protein in the major groove alters the minor groove geometry so that the tetrahedral coordination complex binds poorly (both being frequent features of DNA–protein interactions). Furthermore, due to the difference in their respective mechanisms of cleavage, DNase I and $(OP)_2Cu^+$ probe different aspects of the structure of a DNA–protein complex. DNase I cleavage relies on the accessibility of a particular phosphodiester bond, and thus protection is indicative of an interaction *on the outer face of the DNA helix*. In contrast, protection from $(OP)_2Cu^+$-mediated cleavage is most likely caused by the inhibition of its binding to the minor groove, and implies that *a portion of the protein occupies at least the minor groove*. Finally, in contrast to other chemical nucleases such as ferrous-EDTA (introduces single-stranded nicks in DNA through the generation of diffusible hydroxyl radicals; *see* Chapter “Hydroxyl Radical Footprinting of

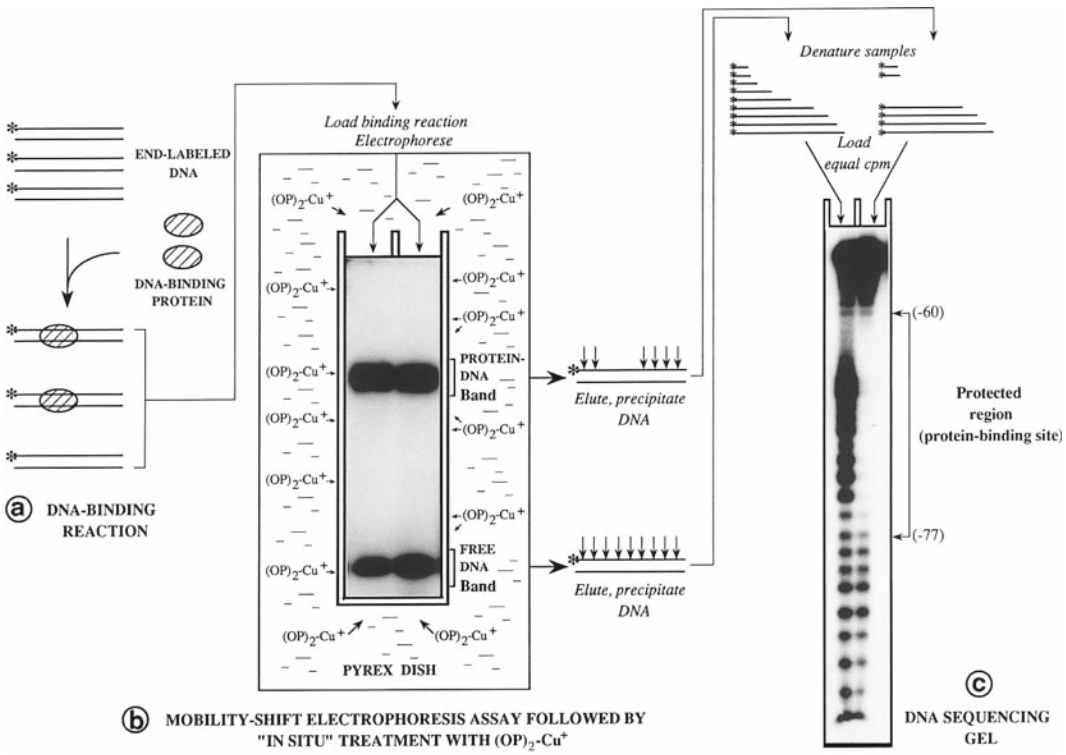


Fig. 3. Outline of the combined electrophoretic mobility-shift/*in gel* OP–Cu footprinting assay. (**a**) DNA restriction fragments containing a protein-binding site(s) are labeled with ^{32}P at a unique end and incubated with a crude or partially purified extract containing the DNA-binding protein(s) of interest, under optimized binding conditions. (**b**) After equilibration of the DNA-binding reaction, the free and bound DNA fragment populations are separated by electrophoresis through a nondenaturing polyacrylamide gel; the gel is then transferred into a buffer-containing *Pyrex dish*, and the retarded and unretarded DNA species are exposed in situ to the nuclease activity of $(OP)_2Cu^+$. The two DNA fractions are subsequently located by autoradiography of the wet gel, excised and eluted from the gel matrix, precipitated, and recovered in formamide buffer. (**c**) Samples are heat-denatured, and equal amounts of radioactivity from the two fractions are electrophoresed on a denaturing polyacrylamide gel (DNA sequencing gel) and autoradiographed. In the sample prepared from the *free DNA band*, bands will appear in the gel corresponding to positions of protein binding. For the sample(s) prepared from the *protein-DNA band(s)*, bands will appear at all positions *except* those bound by the protein(s) (*protected region*). The particular example depicts the OP–Cu mapping of a DNA–protein complex formed between bacterially expressed LFB1 (a liver-specific transcription factor) and an oligonucleotide bearing its binding site within the –95 to –54 region of the α1-antitrypsin promoter. *Arrowheads* connected by *line* demarcate the footprinted site. The enhanced cleavage observed within the protein-binding site in the free DNA sample is due to the presence of repeated TA elements in this sequence (*see* **Subheading 1.2.1**).

Protein–DNA Complexes"), the nucleolytic activity of OP–Cu is not inhibited by glycerol, a free radical scavenger, which is present in most protein storage buffers.

1.2.2. Benefits of OP–Cu Footprinting Within Mobility-Shift Gels

The major advantage of the combined OP–Cu footprinting procedure arises from the topography of treatment: Preformed DNA–protein complexes are exposed to the chemical nuclease *within the gel* (i.e., not prior but subsequent to an electrophoretic

mobility-shift experiment). This characteristic of the technique makes it ideal for protection analysis of kinetically labile complexes *(16)*. At least three factors account for the latter. The first is that the background cleavage is greatly reduced by the separation of unbound DNA from the DNA–protein complex(es) pool. The second factor is the so-called "caging effect" *(3, 4)*. The gel matrix forms "cage-like" compartments that prevent a dissociated protein from diffusing away from the DNA, so that by enhancing reassociation the apparent affinity constant will be higher than the true value. The protein could also interact with the gel matrix, thereby orienting its diffusion toward reassociation. Whatever the mechanism(s), the increase in stability of the complex contributed by the gel leads to a more efficient blockage of the access of the $(OP)_2Cu^+$ chelate to the protein-binding DNA segment. The third factor comes from the nature and the site of action of the cupryl intermediate through which the reaction is funneled, and acts synergistically with the previous one. Since this highly reactive oxidative species is generated near the surface of the DNA (in situ), diffusible radicals, if formed at all, will have a short or restricted diffusive path and, therefore, will be unable to achieve a fast equilibrium distribution along the DNA polymer. Consequently, protein-binding sites exposed during multiple dissociation events will most of the time escape the nucleolytic attack and hence remain intact.

In addition to the fact that discrete complexes with defined stoichiometries and a wide range of kinetic stabilities can be mapped simultaneously, the in situ OP–Cu footprinting procedure is superior to oligonucleotide binding competition assays in the analysis of multiple complexes frequently obtained in electrophoretic mobility-shift experiments employing unfractionated extract preparations. For example, multiple retarded bands can arise from protein–protein interactions between a non-DNA-binding transcription factor(s) and a specific DNA-binding protein, or from two proteins binding to distinct DNA sequences in a cooperative manner *(12)*. Although both complexes would be abolished by competition with oligonucleotides, these possibilities can be readily distinguished by direct footprinting within the gel.

1.3. Additional Applications and Outlook

The nucleolytic activity of $(OP)_2Cu^+$ in a polyacrylamide matrix has been also demonstrated to be a viable means of gaining insight into the interactions of RNA-binding proteins with their recognition sequences *(17, 18)*. Application of OP–Cu in this context may be invaluable toward defining structural perturbations in RNA on protein binding and mapping the binding domains of various proteins. Because of the preferential nucleolytic activity of $(OP)_2Cu^+$ toward single-stranded bulge and loop RNA regions (double-stranded stem regions can be cut at elevated concentrations of the chemical nuclease), hypersensitive sites may be

obtained on footprinting an RNA–protein complex following a gel retardation assay. Such sites would imply an unwinding of a helical structure on protein binding or perturbations in the minor groove accessibility of the bound RNA molecule.

The *in gel* OP–Cu footprinting methodology has already expanded the "tool box" available to investigators wishing to explore the structure and the function relationships of nucleic acid–protein complexes, and emerging improvements in the chemical mechanism (e.g., DNA-strand scission by the coordination complex of OP with a non-redox-active metal) as well as future modifications will likely make this technology even more efficient and broadly useful.

2. Materials

2.1. Analytical and Preparative EMSA

2.1.1. Solutions

1. A variety of binding and gel buffers are commonly employed in EMSA (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions"). A suitable binding buffer is: 10 mM Hepes pH 7.9, 10% glycerol, 0.1 mM EDTA, 0.5 mM tetrasodium pyrophosphate, 0.5 mM PMSF. The most common gel buffers are Tris–glycine: 50 mM Tris, 2.5 mM EDTA, 0.4 M glycine; 0.5 × TBE: 45 mM Tris, 45 mM boric acid, 1 mM EDTA or Tris–acetate: 6.7 mM Tris–HCl, pH 7.5, 3.3 mM sodium acetate, 1 mM EDTA.
2. Ammonium persulfate (10%; w/v): Weigh out 1 g of ammonium persulfate, and put it in a sterile plastic tube containing 10 mL of distilled, deionized water. Vortex vigorously until the salt is completely dissolved. Filter through a 0.22-μm membrane filter. This solution may be stable for a period of a few days at 4°C, but it is recommended that you prepare it freshly for each new gel. Ammonium persulfate is extremely destructive to tissue of the mucous membranes and upper respiratory tract, eyes, and skin. Inhalation may be fatal. Exposure can cause gastrointestinal disturbances and dermatitis. Wear gloves, safety glasses, respirator, and other protective clothing and work in a chemical fume hood. Wash thoroughly after handling.
3. Dye-containing binding buffer: 0.05% (w/v) bromophenol blue in 1× optimized binding buffer (store at 4°C after filtering).

2.1.2. Reagents/Special Equipment

1. Highly purified duplex DNA fragment labeled *exclusively* at one of its four ends (5′ or 3′); use standard procedures for

unique labeling *(19)*. All necessary precautions should be observed to minimize exposure to ionizing radiation during labeling and isolation of the probe. Consult the institutional environmental health and safety office for further guidance in the appropriate use of radioactive materials.

2. Reagents employed in the optimized binding reaction.
3. 16–18 × 16–18 cm front and back glass gel electrophoresis plates: The plates must be absolutely clean before use. Wash them with warm soapy water; then holding them by the edges, rinse several times first in tap water and then in deionized water. Finally, rinse with ethanol, and let them air dry. Using a pad of Kimwipes, siliconize the inner side of the back plate with a 2% dimethyldichlorosilane solution in 1,1,1-trichloroethane in a chemical fume hood (this product is particularly toxic; gloves, safety glasses, respirator, and other protective clothing should be worn when handling it).
4. 0.3-cm Spacers.
5. Electroresistant plastic tape (e.g., 3M yellow electrical tape).
6. 0.22- and 0.45-μm filters (Millipore, Bedford, MA).
7. *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED; Bio-Rad, Richmond, CA). TEMED is extremely destructive to tissue of the mucous membranes and upper respiratory tract, eyes, and skin. Inhalation may be fatal. Prolonged contact can cause severe irritation or burns. Wear gloves, safety glasses, respirator, and other protective clothing and work in a chemical fume hood (TEMED is also flammable!). Wash thoroughly after handling.
8. 3-mm Gel comb with 10-mm wide teeth.
9. 10-mL syringe and 18-gauge needle.
10. 100- to 200-mL Hamilton syringe.
11. Additional reagents and equipment: Powdered acrylamide and *N*,*N*′-methylene-bis-acrylamide (Bio-Rad); plenty of binder clamps (fold-back spring clips); razor blades; polyacrylamide gel electrophoresis apparatus; constant current power supply; peristaltic pump for recirculating electrophoresis buffer (if required); siliconized 1.5-mL Eppendorf microcentrifuge tubes; spatula. Acrylamide and *N*,*N*′-methylene-bis-acrylamide are potent neurotoxins and are absorbed through the skin. Their effects are cumulative. Wear gloves and a face mask when weighing these substances and when handling solutions containing them. Although polyacrylamide is considered to be nontoxic, it should be handled with care because of the possibility that it might contain small quantities of unpolymerized acrylamide.

2.2. DNA Chemical Cleavage (Footprinting) Reactions Within the Gel

2.2.1. Solutions

1. 10 mM Tris–HCl, pH 8.0 (store at room temperature after autoclaving).
2. MPA solution (58 mM 3-mercaptopropionic acid): Add 100 μL of neat 3-mercaptopropionic acid (Sigma-Aldrich) to a sterile 50-mL conical tube containing 19.9 mL of distilled, deionized water. Mix by vortexing. 3-Mercaptopropionic acid is toxic and causes burns in contact with skin and eyes; wear gloves and handle accordingly. Store the liquid reagent in a place protected from light. Dilute immediately prior to use.
3. OP solution (40 mM 1,10-phenanthroline): Weigh out 80 mg of 1,10-phenanthroline monohydrate (Sigma-Aldrich or G.F. Smith), and dissolve (by vortexing and shaking vigorously for 2 min) in 10 mL of absolute ethanol in a sterile 50-mL conical tube. Wear gloves and dust mask when weighing this reagent. Store the powdered reagent in a place protected from light. Prepare just prior to use.
4. Cu^{2+} solution (9 mM $CuSO_4$): Weigh out 72 mg of anhydrous copper (II) sulfate (Sigma-Aldrich or Mallinckrodt, Chesterfield, MO), and dissolve (by vortexing for 1 min) in 50 mL of distilled, deionized water in a sterile 50-mL conical tube. Powdered copper (II) sulfate is irritating to eyes, respiratory system, and skin; wear gloves and eye/face protection when weighing this chemical. Store the powdered chemical sealed in a dry place. Prepare just prior to use.
5. $(OP)_2Cu^+$-STOP solution (28 mM 2,9-dimethyl-OP): Weigh out 127 mg of 2,9-dimethyl-1,10-phenanthroline (Neocuproine) monohydrate (Sigma-Aldrich or G.F. Smith), and dissolve (by vortexing vigorously for 2 min) in 20 mL of absolute ethanol in a sterile 50-mL conical tube. Wear gloves and dust mask when weighing this reagent. Store the powdered reagent in a place protected from light. Prepare just prior to use.

2.2.2. Reagents/Special Equipment

1. 20 × 20 cm Pyrex dish (available in most supermarkets); wash the dish with detergent, water, and then with ethanol. Rinse with deionized water and dry with tissues.
2. Sterile 50-mL conical tubes.
3. Additional equipment: Protective gloves; glass or plastic beaker; 20-mL glass pipette; vacuum aspirator.

2.3. Isolation of Free and Complexed DNA Fractions

2.3.1. Direct Elution from the Polyacrylamide Gel Matrix

Solutions

1. Gel elution buffer: 0.5 M ammonium acetate, pH 7.5 (promotes diffusion of the DNA out of the gel matrix and is readily soluble in ethanol in the subsequent precipitation step), 1 mM EDTA, 0.1% sodium dodecyl sulfate (SDS; w/v; effectively denatures any contaminating DNAase activity). For improved recovery of DNA fragments smaller than 60 bp, the buffer should also include 10 mM magnesium chloride. This stock

solution can be stored at room temperature protected from light for several months.

2. Phenol/chloroform/isoamyl alcohol (25:24:1; v/v): Mix 25 vol of phenol [redistilled under nitrogen and equilibrated with 100 mM Tris–HCl, pH 8.0, 1 mM EDTA in the presence of 0.1% (w/v) 8-hydroxyquinoline] with 24 vol of chloroform and 1 vol of isoamyl alcohol. Phenol can be stored at 4°C in dark (brown) bottles for up to 2 months. It is highly corrosive and can cause severe burns. Any areas of skin that come in contact with phenol should be rinsed with a large volume of water or PEG 400 and washed with soap and water (*do not use ethanol!*). Chloroform is irritating to the skin, eyes, mucous membranes, and respiratory tract. It is also a carcinogen and may damage the liver and the kidneys. Wear gloves, protective clothing, safety glasses, and respirator when handling these substances and carry out all manipulations in a chemical fume hood.
3. Chloroform/isoamyl alcohol (24:1; v/v): Mix (in a chemical fume hood and wearing gloves, safety glasses, and respirator) 24 vol of chloroform with 1 vol of isoamyl alcohol. This organic mixture can be stored at room temperature in dark (brown) bottles indefinitely.
4. 70 and 90% (v/v) ethanol.
5. Sequencing-gel loading buffer: 90% (v/v) deionized formamide, 0.5× TBE (*see* Subheading 2.3.2), 0.025% (w/v) xylene cyanol FF, and 0.025% (w/v) bromophenol blue. Store at –20°C after filtering. (Preparation of deionized formamide: Combine 200 mL of formamide with 5 g of AG501-X8 [D] [Bio-Rad] in a 250-mL Erlenmeyer flask. Cover the mouth of the flask with Parafilm, and gently stir the mixture at room temperature for 30 min. Avoid aeration of the formamide when stirring the mixture. Filter the solution through a coarse sintered glass funnel, and store the deionized formamide at –20°C. Formamide is a teratogen; take all safety precautions to avoid contact during the above manipulations.)

Reagents/Special Equipment

1. Plasticwrap, such as Saran Wrap® or cling film.
2. Old X-ray film covered with plasticwrap.
3. Glass stirring rod.
4. Small adhesive labels (or Scotch Tape®).
5. Radioactive ink: Mix a small amount of ^{32}P with waterproof black drawing ink, to a concentration of ~200 cps (on a Geiger counter) per μL.
6. Fiber-tip pen.
7. Kodak X-Omat AR film.

8. Lightproof cardboard film holder.
9. Aluminum foil.
10. Lab marking pen.
11. Sharp scalpel.
12. Fine-tip waterproof marking pen.
13. Single-edged disposable razor blades.
14. 18-Gauge syringe needles or sterile forceps.
15. Sterile 3-mL syringes attached to a shortened 18-gauge needle (broken with pliers).
16. Siliconized 1.5-mL Eppendorf microcentrifuge tubes.
17. 22-Gauge syringe needle.
18. Siliconized capless 0.5-mL Eppendorf microcentrifuge tubes.
19. Conformable self-sealing tape (e.g., Parafilm).
20. 1-mL Sterile syringes.
21. 0.22-μm Syringe filters (Millipore).
22. Ice-cold absolute ethanol.
23. Glycogen (Sigma-Aldrich or other supplier).
24. Drawn-out Pasteur pipettes.
25. Elutip™-d mini-columns (Schleicher & Schuell) and low and high salt solutions as recommended by manufacturer.
26. Additional equipment: Geiger counter; all equipment for autoradiography; low-speed centrifuge (Beckman J6B or Sorvall RC3); microcentrifuge (Eppendorf or equivalent); 37–42°C shaking incubator; vacuum centrifuge (e.g., Speed-Vac, Savant, Hicksville, NY); scintillation counter or Bioscan Quick Count for ^{32}P; waterbath at 68°C.

2.3.2. Electrotransfer of the Entire Gel and Elution from NA-45 Membrane

Solutions

1. 10 mM EDTA, pH 7.6 (store at room temperature after autoclaving).
2. 0.5 M NaOH.
3. 1× TBE buffer: 89 mM Tris base, 89 mM borate, 2.5 mM EDTA. To prepare 5 L of 5× TBE buffer, dissolve (stirring for at least 1 h) 272.5 g of ultrapure Tris base, 139.1 g of boric acid, and 23.3 g of disodium EDTA dihydrate in 4.5 L of distilled, deionized water, and make up to a final volume of 5 L. It is not necessary to adjust the pH of the resulting solution, which should be about 8.3. Store at room temperature; this stock solution is stable for many months, but it is susceptible to the formation of a precipitate and should occasionally be inspected visually.

4. 20 mM Tris–HCl, pH 8.0, 0.1 mM EDTA (store at room temperature after autoclaving).
5. NA-45 membrane elution buffer: 1.0 M NaCl, 20 mM Tris–HCl, pH 8.0, 0.1 mM EDTA (store at room temperature after autoclaving).
6. Solutions in **items 2–5** of "Solutions" under Subheading 2.3.1.

Reagents/Special Equipment

1. NA-45 membrane sheets (0.45-mm pore size; Schleicher & Schuell).
2. Filter paper [Whatman (Clifton, NJ) 3MM or equivalent].
3. Clean glass plate.
4. **Items 1**, **2**, and **4–7** of "Reagents/Special Equipment" under **Subheading 2.3.1**.
5. Metal cassette.
6. **Items 10–14**, **16**, and **22–24** of "Reagents/Special Equipment" under **Subheading 2.3.1**.
7. Additional equipment: Protective gloves; electrotransfer unit [high current (2–3 A) power supply, e.g., Bio-Rad or Hoefer, San Francisco, CA]; Kimwipes; Geiger counter; all equipment for autoradiography; microcentrifuge (Eppendorf or equivalent); 55°C waterbath with agitation; vacuum centrifuge (e.g., SpeedVac, Savant); scintillation counter or Bioscan Quick Count for ^{32}P; waterbath at 68°C.

2.4. Preparation of the G + A Sequencing Ladder

2.4.1. Solutions

1. Carrier DNA stock solution: salmon sperm DNA extracted sequentiallywithphenol/chloroform/isoamylalcohol(25:24:1; v/v) and chloroform/isoamylalcohol (24:1; v/v), precipitated with ethanol, resuspended in distilled, deionized water at 1 mg/mL, and sonicated to an average chain length of 200 bp.
2. 1.0 M aqueous piperidine: Add 100 μL of concentrated piperidine (reagent grade; BDH, London, England) into 0.9 mL of distilled, deionized water. Piperidine is somewhat hard to pipette; when diluting rinse the micropipet tip thoroughly by repeated pipetting and then mix the solution well by vortexing. Make dilution in a chemical fume hood just prior to use.
3. 1% SDS: Prepare a 10% (w/v) solution of SDS in distilled, deionized water (wear dust mask when weighing powdered SDS); heat at 68°C to assist dissolution (do not autoclave). Dilute 1:10 with distilled, deionized water. Store at room temperature.
4. Solution in **item 5** of "Solutions" under **Subheading 2.3.1**.

2.4.2. Reagents/Special Equipment

1. Siliconized 1.5-mL Eppendorf microcentrifuge tubes.
2. 20,000 cpm of the end-labeled DNA fragment used in the preparative EMSA.
3. 88% Aqueous formic acid.
4. Conformable self-sealing tape (e.g., Parafilm).
5. 1-Butanol (*n*-butylalcohol).
6. Drawn-out Pasteur pipette.
7. Additional equipment: Wet ice; 37°C waterbath with agitation; thermostated heating block at 90°C; lead weight; microcentrifuge (Eppendorf or equivalent); vacuum centrifuge (e.g., SpeedVac, Savant); waterbath at 68°C.

2.5. Analysis of the Chemical Cleavage Products on a DNA Sequencing Gel

2.5.1. Solutions

1. 40% (w/v; 19:1) acrylamide: *N*,*N*′ methylene-bis-acrylamide solution.
2. Solution in **item 3** of "Solutions" under **Subheading 2.3.2**.
3. Solution in **item 2** of **Subheading 2.1.1**.
4. Fixing solution: 10% (v/v) acetic acid, 10% (v/v) methanol.

2.5.2. Reagents/Special Equipment

1. Urea (enzyme grade).
2. 34 × 40 cm front and back glass plates: treat the plates as described in **Subheading 2.1.3**.
3. 0.04-cm Spacers.
4. 0.4-mm Custom-ordered sample comb with 5-mm lanes spaced on 10-mm centers.
5. 10-mL Syringe with a 22-gauge needle.
6. Calibrated glass capillaries with finely drawn tips or disposable flat-capillary pipette tips (National Scientific Supply Company, Inc., San Rafael, CA).
7. 30-mL Syringe with a bent 20-gauge needle.
8. Backing paper (Whatman No. 1 or equivalent).
9. Plasticwrap, such as Saran Wrap or cling film.
10. Filter paper (Whatman 3MM or equivalent).
11. Kodak X-Omat AR film.
12. Large metal autoradiography cassette.
13. Intensifying screen and film adapted for sensitive ^{32}P autoradiography.
14. Additional reagents and equipment: TEMED; plenty of binder clamps (fold-back spring clips); sequencing gel electrophoresis apparatus; power supply delivering high voltage (2,500–3,000 V) (e.g., Bio-Rad); dry-block heater at 90°C; wet ice; aluminum plate of an appropriate size; razor blade; plastic tank at gel dimensions; 10-mL glass pipette; Kimwipes; gel dryer; all equipment for autoradiography.

3. Methods

3.1. Analytical EMSA

1. Perform a preliminary EMSA (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions" and Note 1) to identify conditions for the formation of the complex(es) to be footprinted. Because no universal binding and/or gel system is likely to be found for the study of all DNA–protein interactions, it may be necessary to optimize conditions for formation and adequate resolution of the DNA–protein complex(es) of interest (*see* **Note 1**).
2. If crude or partly fractionated extracts are employed, ascertain the DNA-binding specificity of the resolved complex(es) by performing an analytical competition binding assay (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions" and **Note 2**).
3. It is advisable, prior to proceeding to the more laborious preparative gel retardation/in situ footprinting assay, to perform an additional analytical experiment and obtain a qualitative estimation of the dissociation rates of pre-equilibrated protein–DNA complexes of interest (*see* **Note 3**). Although the $(OP)_2Cu^+$-mediated cleavage reactions in the gel are relatively insensitive to the kinetic stability of the DNA–protein complex(es) under investigation (*see* **Subheading 1.2.2**), this information can be used in adjusting the exposure time to the chemical nuclease (*see* **step 6** in **Subheading 3.3** and **Note 7**), thereby enhancing the clarity of the expected footprint.

3.2. Preparative EMSA

1. Assemble 16–18 × 16–18 cm front and back glass plates and 0.3-cm spacers for casting a preparative mobility-shift polyacrylamide gel (3-mm-thick polyacrylamide gels are preferable, because they are easier to load and give sharper bands than 1.5-mm-thick gels). The plates must be scrupulously clean and free of grease spots to avoid trapping air bubbles while pouring the gel; it is highly recommended to use one glass plate (preferably the back) that has been siliconized on the inner side for ease of removal after electrophoresis is completed. Taking particular care (to prevent leakage), seal the entire length of the two sides and the bottom of the plates with electrician's plastic tape.
2. Prepare, filter (through a 0.45-μm filter), and degas (by applying vacuum) 100 mL of the acrylamide gel solution found during optimization of the analytical assay (**Subheading 3.1, step 1**). Because of the ready permeability of the gel matrix to all reagent and quenching solutions used for the subsequent chemical treatment and the lack of diffusible radicals mediating the DNA-scission reaction, the $(OP)_2Cu^+$

in situ footprinting technique is compatible with a broad spectrum of gel porosities [ranging from 3.5 to 6% (w/v), with an acrylamide to *N*,*N*′-methylene-bis-acrylamide molar ratio of 19:1–80:1] and gel/running-buffer compositions [from glycerol-containing/low-ionic-strength (pH 7.5–7.9) to high-ionic-strength TBE (pH 8.3) or Tris–glycine (pH 8.5) buffer systems].

3. Add to the solution 0.8 mL of 10% ammonium persulfate and 75 μL of TEMED, and swirl the mixture gently.
4. Slowly pour the acrylamide gel mix between the glass plates, and quickly insert a 3-mm comb bearing 10-mm wide teeth. Allow the gel to polymerize (lying flat or nearly flat, to avoid undesirable hydrostatic pressure on the bottom) at room temperature for about 45 min.
5. After polymerization is complete, remove the electrical tape from the bottom of the gel (by cutting with a razor blade), and clamp the gel into place on the electrophoresis apparatus. Fill both chambers of the electrophoresis tank with the buffer used for preparation of the acrylamide gel mix **(step 2)**, carefully remove the comb, and immediately rinse the sample wells with reservoir buffer using a 10-mL syringe with an 18-gauge needle.
6. Prior to assembling the preparative binding reaction, preelectrophorese the gel for 60 min at 20 mA, with or without buffer recirculation between the two compartments, depending on the nature of the gel/running-buffer system used (low- or high-ionic strength). This removes any excess persulfate and unpolymerized acrylamide. Prerunning of the gel should be done at the temperature at which the binding reaction will be performed (known from **step 1** in **Subheading 3.1**).
7. In a siliconized 1.5-mL Eppendorf tube, scale up the optimized analytical reaction 5- to 10-fold, depending on the relative proportion of the DNA–protein complex(es) obtained. If the detected specific DNA-binding activity(ies) (**Subheading 3.1, step 1**) represents <1% of the total label input, the amount of radioactive probe in the scaled reaction should be at least 250,000 cpm (*see also* **Note 4**).
8. Turn off the electric power. Using a 100- to 200-mL Hamilton syringe, load the preparative binding reaction onto one or two wells (depending on the total volume of the sample) in the middle of the gel. Raise the tip of the needle as the sample is loaded into the well. Do not attempt to expel all of the sample from the syringe since this almost always produces air bubbles that blow the sample out of the well. If the

glycerol concentration in the binding buffer is low (<5%), it is important to load the well gently to prevent dilution (*see also* **Note 5**). Avoid adding bromophenol blue to the binding reaction prior to loading because this dye can rapidly disrupt some DNA–protein complexes. Instead, you may load just dye-containing binding buffer in one of the adjacent lanes to monitor the progress of electrophoresis.

9. Run the gel at 25–35 mA (it may be necessary to adjust the voltage occasionally if a constant power supply is not available) a time sufficient to allow migration of the free DNA probe to ~2 cm from the bottom of the gel. Provided the same plate size has been used in establishing the optimal electrophoresis conditions, this can be monitored by the migration of the tracking dye, in correlation with the position of the free probe on the autoradiogram obtained from the optimized analytical assay (**Subheading 3.1, step 1**). If electrophoresis is performed at room temperature, the glass plates should be allowed to become only slightly warm, since excess heating may perturb the equilibrated complexes, or even cause protein denaturation; decrease the current if the plates become any hotter.
10. Following electrophoresis, detach the glass plates from the gel apparatus, and using a spatula carefully remove the spacers and pry the glass plates apart, taking extreme care not to distort or tear the gel, which should remain attached to only one of the plates (the nonsiliconized front plate).

3.3. DNA Chemical Cleavage (Footprinting) Reactions Within the Gel

1. Wear protective gloves and wash your fingers extensively in a beaker containing deionized water to remove the talc powder. Immerse the whole gel, still attached to the lower plate (with the gel facing up), in a 20 × 20 cm scrupulously clean Pyrex dish (*never use a plastic tray!*), containing 200 mL of 10 mM Tris–HCl, pH 8.0. Loosen it on its supporting glass plate [omit this step if using Tris–glycine-containing or low-percentage/low-ionic-strength polyacrylamide gels (e.g., a 3.5–4% gel), which are very sticky and extremely difficult to manipulate without fracturing].
2. Prepare the MPA, OP, and Cu^{2+} solutions (*see* **Note 6**).
3. In a sterile 50-mL conical tube, transfer 1 mL of the freshly made OP solution. To this, add 1 mL of the freshly prepared Cu^{2+} solution, and wait 1 min while pipetting up and down (the mixture should turn light blue, indicating efficient formation of the $[OP]_2Cu^{2+}$ chelate). Add 18 mL of distilled, deionized water and vortex the tube. This is the OP/Cu^{2+} solution (1,10-phenanthroline to copper ratio of ~4.5:1).

4. Add the OP/Cu^{2+} solution (20 mL) to the gel equilibrating in the 200 mL buffer, and shake the Pyrex dish while laying it on an even horizontal surface to distribute evenly.
5. Initiate the chemical nuclease reaction by adding the MPA solution (20 mL); distribute evenly by quickly shaking the Pyrex dish as before. The gel will turn brownish. The appearance of a dark brown precipitate indicates the presence of impurities in the original $CuSO_4$ solution, which will interfere with the cascade leading to DNA-strand scission. It is, therefore, crucial for the assay to use copper (II) sulfate of the best available analytical grade.
6. Incubate for a period of 8–30 min *without shaking* (do not disturb the equilibrated complexes; the small size and the ready diffusibility of all reaction components within the gel matrix are sufficient for a productive attack on the target DNAs). To obtain an intelligible and homogeneous cleavage pattern of all DNA species in the gel, the exact time of chemical treatment has to be adjusted for each particular case, based on the considerations discussed in **Note 7.**
7. During the last 5 min of the incubation period, prepare the $(OP)_2Cu^+$-STOP solution.
8. Quench the reaction by adding the $(OP)_2Cu^+$-STOP solution (20 mL), and wait for 2 min while shaking the Pyrex dish (*see* **Note 8**). The gel will turn yellowish, which is diagnostic for the quality of 2,9-dimethyl-OP, and hence for efficient termination of the chemical nuclease action.
9. Using a 20-mL glass pipette, aspirate (staying away from the corners of the gel) all the liquid from the Pyrex dish, and carefully rinse the gel (still on the glass plate) with four changes of deionized water. Remove the plate with the gel on it from the Pyrex dish. It is not necessary to take any specific precautions in dispensing the original mixture and the washing material, because all reaction components are oxidatively destroyed. Immerse the Pyrex dish in household bleach for 1 h at room temperature followed by extensive washing down the drain with tap water.

3.4. Isolation of Free and Complexed DNA Fractions

3.4.1. Direct Elution from the Polyacrylamide Gel Matrix

1. Smoothly wrap the gel and plate with plasticwrap or, preferably, peel off the gel onto a suitable backing material (old X-ray film covered with plasticwrap is best) and wrap it with plasticwrap. Using a glass stirring rod as a rolling pin, remove any air bubbles trapped under the plasticwrap, being careful not to disturb the shape of the gel.
2. To aid accurate subsequent alignment of gel and film, trace three corners of the plasticwrap covering the gel with small adhesive labels (or pieces of Scotch Tape), marked with

radioactive ink spots; use an almost equal amount of radioactivity in each ink dot to that in the protein-bound DNA fraction(s) (this can be monitored by a Geiger counter). Use a fiber-tip pen to apply ink of the desired activity to the sticky labels; let the ink dots dry completely before exposing to X-ray film.

3. In the darkroom, tape the sealed gel to a piece of Kodak X-Omat AR film. Enclose the assembly in a lightproof cardboard film holder, exerting an even gentle pressure on the "sandwich," and wrap the entire packet with aluminum foil to ensure a light-tight environment.
4. Expose the film at 4°C for 1–3 h [the length of exposure time depends on the relative abundance of the specific complex(es)] to assess the position of the retarded (bound) and unretarded (free) DNA fragments. The energy of the β particles produced by the ^{32}P decay is sufficient to penetrate several millimeters thickness of hydrated gels, without significant absorption (quenching by the gel is <40%), thus allowing the direct autoradiographic detection of [^{32}P]-labeled DNA embedded in the gel matrix.
5. Develop the film, and using a lab marking pen, encircle the position of the complex(es) to be mapped as well as that of the unretarded probe. Any DNA released by dissociation during the run will trail just above the free DNA band as a smear; do not include this region in your marking.
6. Using a sharp scalpel, cut out from the X-ray film the marked rectangles containing the autoradiographic images of the free and bound probe.
7. Line up the radioactive ink spots on the film with the corresponding markings at the three corners of the plasticwrap. With a fine-tip waterproof marking pen, mark the position of the free and bound probe on the plasticwrap, using the periphery of the rectangular holes on the film as a template.
8. Remove the film, and cut through the marks on the plasticwrap with a disposable razor blade for each species. Separate the polyacrylamide slices from the rest of the gel (and from the plasticwrap), using either 18-gauge single-use syringe needles or sterile forceps, and transfer them onto a piece of plasticwrap. It is desirable to keep the size of the polyacrylamide strips to a minimum (*see* **Note 4**).
9. Crush the gel slices by extruding them from a sterile 3-mL syringe barrel through a shortened 18-gauge needle (broken with pliers) into a siliconized 1.5-mL Eppendorf tube by low-speed centrifugation (5 min at 2,500 × *g*) in a swinging bucket rotor. Alternatively, punch a small hole by forcing a 22-gauge sterile needle through the bottom of a siliconized

capless 0.5-mL Eppendorf tube, place this tube into another siliconized capped 1.5-mL Eppendorf tube, put the gel slice in the upper tube, and spin at 12,000 × *g* in a microcentrifuge (minus rotor cover) for 1 min. The gel will be crushed through the hole into the lower tube.

10. To each tube add enough gel elution buffer to cover the gel paste, and mix well by vortexing. The volume of the buffer added depends on the size of gel slice, but as a guide, 0.5–0.6 mL is used for a slice 10 × 3.5 × 3 mm.
11. Seal each tube with conformable self-sealing tape, and allow the DNA fragments to diffuse out by incubating at 37–42°C for 10–16 h in a shaking incubator.
12. Vortex the tubes vigorously, and pellet the gel paste by centrifuging at room temperature for 1 min in a microcentrifuge (12,000 × *g*).
13. Using a micropipet, pipette off the supernatant solution, taking care to avoid polyacrylamide pieces, and transfer it to a 1-mL sterile syringe.
14. Remove any remaining tiny pieces of polyacrylamide by slowly passing the supernatant through a 0.22-μm syringe filter into a fresh siliconized Eppendorf tube (do not use polystyrene tubes to collect the filtrate, since they cannot withstand the subsequent organic extractions). The eluted yield of DNA fragments should be >90%.
15. Extract the filtered supernatant sequentially with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1; v/v) and chloroform/isoamyl alcohol (24:1; v/v), to eliminate contaminating proteins that might distort DNA fragment migration during subsequent electrophoresis. In both steps, mix the contents of the tube thoroughly by vortexing for 30 s, and centrifuge at 12,000 × *g* (microcentrifuge) for 5 min at room temperature to separate the organic and aqueous phases (*see* **Note 9**).
16. With a micropipet, transfer the aqueous phase (no more than 0.55 mL) to a fresh siliconized Eppendorf tube. Add ~2 vol of ice-cold absolute ethanol (*no additional salt is required!*), vortex well, and precipitate the radioactive DNA fragments by chilling the tube at –20°C for a minimum time of 60 min. Although it is generally not necessary to add carrier to aid precipitation (the small acrylamide polymers released from the crushed gel slice will suffice), it is recommended to precipitate the DNA in the presence of glycogen (10 μg/sample, added prior to ethanol) to improve even further the recovery of DNA.
17. Recover DNA by centrifugation at 12,000 × *g* for 30 min in a microcentrifuge (4°C). Carefully aspirate off the ethanol supernatant with a drawn out Pasteur pipette, taking care not

to disturb the faintly visible radioactive pellet (its presence can be monitored by a Geiger counter, and its location identified from the position of the tube in the rotor) (*see also* **Note 10**).

18. Remove traces of salt trapped in the precipitate (which interfere with subsequent electrophoresis), by rinsing the pellet twice with 1 mL of 70 and 90% (v/v) ethanol, respectively, centrifuging each time at 12,000 × *g* (microcentrifuge) for 2 min at 4°C. For both washes, invert the tube gently several times; *do not vortex.*
19. Dry the pellet for 5 min in a vacuum centrifuge.
20. Measure each pellet by Cerenkov counting to determine radioactivity (1,500–2,000 cpm is sufficient for an overnight exposure with intensifying screen).
21. Resuspend the pellets (by heating at 68°C for 2 min, vortexing vigorously, and repeatedly pipetting) in sequencing-gel loading buffer, so that 5 μL will contain equal Cerenkov cpm from the free and bound DNA fractions. It is important to equalize the number of cpm/μL in the two fractions in order to compare their cleavage patterns accurately. If the sequencing gel is not to be run immediately, you can store the DNA samples at –70°C.

3.4.2. Electrotransfer of the Entire Gel and Elution from NA-45 Membrane

If the EMSA was performed using Tris–glycine-containing or low-percentage/low-ionic-strength polyacrylamide gels, which behave like a poorly set "Jello" and are, therefore, extremely difficult both to manipulate for autoradiography and to handle as polyacrylamide strips in the subsequent DNA-elution steps, it is highly recommended to transfer the entire gel electrophoretically onto a sheet of NA-45 membrane (DEAE cellulose in membrane form). Following electroblotting, the NA-45 membrane is exposed to X-ray film, the bands corresponding to free and bound species are cut out, and DNA is eluted. The remaining steps in the procedure, beginning with organic extractions of the eluates, are identical to those described in **steps 15–21** of **Subheading 3.4.1**:

1. Cut a piece of NA-45 membrane and four pieces of filter paper to the exact size of the gel; cut the membrane between liner sheets wearing gloves.
2. To increase binding capacity, wash the membrane for 10 min in 10 mM EDTA, pH 7.6, and for 5 min in 0.5 M NaOH, followed by several rapid washes in distilled, deionized water; let the membrane soak in 1× TBE buffer.
3. Remove the plate with gel from the Pyrex dish and place it on a flat surface. Carefully lay two pieces of prewet (in 1× TBE buffer) filter paper onto the surface of the gel, making sure no air bubbles are trapped between the filter paper and the gel.

4. Slowly and with extreme care, lift the gel (adhered to the filter paper), and place it (with the gel facing up) on a clean glass plate. Wet the gel with a thin layer of 1× TBE buffer.
5. Wearing gloves, lay the wet membrane sheet over the gel, again being careful not to trap air bubbles beneath the membrane.
6. Complete the "sandwich" by placing the two remaining pieces of prewet (in 1× TBE buffer) filter paper on top of the membrane.
7. Insert the "sandwich" of filter paper/gel/membrane/filter paper into a gel-holder cassette, and load the assembly into one of the center slots in a (wet) transfer apparatus (any of the commercially available electroblot units are suitable), with the NA-45 membrane positioned between the gel and the anode (positive electrode).
8. Fill the transfer apparatus with 1× TBE buffer (precooled at 4°C), and transfer the chemically cleaved double-stranded DNA fragments electrophoretically from the gel to the NA-45 membrane. Electroblotting is performed at 4°C for 3 h, at either 20 V (~1 V/cm), if small DNA fragments (40–90 bp) are being transferred, or at 35 V (~2 V/cm), if fragments >100 bp have been employed in the EMSA.
9. When transfer is completed, turn off the power, remove the gel "sandwich," lift the membrane sheet away from the gel while wearing gloves, and rinse it in 20 mM Tris–HCl, pH 8.0, 0.1 mM EDTA to remove residual polyacrylamide. *Do not let the membrane dry!*
10. Place the wet membrane (with transferred DNA face up) on the surface of a used piece of X-ray film wrapped in plasticwrap, and cover it with a tightly drawn layer of plasticwrap. With a pad of Kimwipes, push out any trapped air bubbles under the plasticwrap. Efficient transfer can be monitored by checking with a Geiger counter.
11. Follow **steps 2** and **3** in **Subheading 3.4.1**. Use a metal cassette instead of a cardboard film holder to expose the membrane to X-ray film.
12. Autoradiograph the membrane at 4°C for 15–45 min (about one-fourth the time required for the wet gel).
13. Follow **steps 5–8** in **Subheading 3.4.1.** Using sterile forceps, transfer the wet NA-45 membrane strips into siliconized 1.5-mL Eppendorf tubes.
14. Add to each NA-45 membrane strip 0.6 mL of NA-45 membrane elution buffer, and spin for a few seconds in a microcentrifuge to submerge the whole strip.
15. Incubate at 55°C for 2–3 h in a waterbath with agitation.

16. Vortex the tubes vigorously, and pellet the NA-45 membrane strips by centrifuging at room temperature for 10 s in a microcentrifuge.
17. Remove the buffer, and place it in a fresh siliconized 1.5-mL Eppendorf tube. Monitor paper for loss of radioactivity; typically, ~90% of the membrane-bound DNA is released with this technique.
18. Follow **steps 15–21** in **Subheading 3.4.1**.

3.5. Preparation of the G + A Sequencing Ladder

Provided the sequence of the DNA probe is known, you may perform at this stage a Maxam–Gilbert guanine- and adenine-specific modification/cleavage reaction (G + A sequencing ladder) of the end-labeled DNA fragment used in the gel retardation assay. This reaction will be co-electrophoresed with the DNA samples eluted from the free and the bound fractions, to identify nucleotides protected from chemical cleavage (protein-contact sites) in the final stage of the footprinting analysis (**Subheading 3.6**). Below is a fast version (requiring only 1 h) of this otherwise time-consuming reaction:

1. In a siliconized 1.5-mL Eppendorf tube mix successively
   - 20,000 cpm of the end-labeled DNA fragment used in the EMSA
   - 1.5 μL of carrier DNA stock solution
   - Distilled, deionized water to 10 μL
2. Chill the tube on ice and add 1.5 μL of 88% aqueous formic acid.
3. Incubate at 37°C for 14 min in a waterbath.
4. Chill again on ice and add 150 μL of freshly prepared 1.0 M aqueous piperidine. Close the tube and wrap the cap tightly with a conformable self-sealing tape.
5. Heat at 90°C for 30 min in a thermostated heating block, with the wells filled with water. It is necessary to put a lead weight on top of the tube to prevent it from popping open as pressure builds up inside.
6. Cool the tube on ice. Remove the conformable tape and spin for a few seconds in a microcentrifuge; transfer to a fresh siliconized Eppendorf tube.
7. Add 1 mL of 1-butanol. Vortex vigorously until only one phase is obtained.
8. Mark the position of the tube in the rotor, and spin at 12,000 $\times g$ for 2 min in a microcentrifuge (room temperature).
9. Carefully remove and discard the supernatant using a drawn-out Pasteur pipette, taking care not to disturb the tiny pellet or the area of the tube where the pellet should be located.

10. Add 150 μL of 1% SDS and vortex the tube. Add 1 mL of 1-butanol. Mix well by repeatedly inverting the tube. This step removes remaining traces of piperidine trapped in the precipitate that interfere with the subsequent electrophoretic separation.
11. Resediment the precipitate by centrifuging at 12,000 × *g* for 2 min in a microcentrifuge, and remove the supernatant as in **step 9**.
12. Spin for a few seconds in a microcentrifuge to collect any traces of liquid at the bottom of the tube, and carefully remove it using a micropipet.
13. Dry the pelleted DNA for 5 min in a vacuum centrifuge.
14. Resuspend the samples (as in **step 21** of **Subheading 3.4.1**) in 5 μL of sequencing-gel loading buffer, and store at –70°C until ready to load onto the sequencing gel.

3.6. Analysis of the Chemical Cleavage Products on a DNA Sequencing Gel

Visualization of the length(s) on the DNA affected by protein(s) binding specifically to it [that is the protein-binding site(s), or footprint(s) left by the protein(s) on the DNA] requires electrophoretic fractionation of the single-stranded fragments resulting from the chemical nuclease attack in a denaturing polyacrylamide gel of the type employed in DNA sequencing, followed by autoradiography. The location of the footprint(s) in the known DNA sequence is identified by including the sequencing marker G + A track prepared in **Subheading 3.5.**

1. Assemble and pour a 34 × 40 × 0.04 cm 6–15% (w/v) sequencing polyacrylamide gel, containing 1× TBE buffer and 8.3 M urea *(20)*. The percentage of acrylamide depends on the size of the DNA fragments to be separated as well as on the size and location (relative to the labeled end) of the suspected protein-binding site(s). As for the preparative EMSA, you should siliconize the inside surface of the back glass plate to aid pouring into gel mold and removal at the end of electrophoresis. To avoid dispersing of radioactivity across lanes (which might produce significant errors in a subsequent densitometric analysis of the free and bound DNA chemical cleavage patterns on the autoradiogram; *see* **Note 12**), it is recommended to use a 0.4-mm custom-made sample comb with 5-mm lanes and 5-mm spacing. You may wrap the polymerized gel in plasticwrap and keep it at room temperature until use (it can be stored as such for up to 36 h).
2. Attach the gel apparatus to the gel electrophoresis tank. Fill both the top and bottom electrode chambers with 1× TBE buffer, and remove the well-forming comb. Check that wells are free from "tails" of polyacrylamide adhering to sides,

which may lead to uneven loading of samples and consequently band-shape distortion.

3. Pre-electrophorese the gel for 45–60 min before loading the samples. This removes persulfate ions and heats the gel. Prerunning of the gel is performed at constant temperature (approximately 55°C), which is most easily achieved by application of constant power (~50–70 W). If the surface temperature becomes too high (>65°C), the glass plates will crack.
4. Thaw the DNA samples (if frozen), heat-denature them (including the G + A sequencing ladder) at 90°C in a dry-block heater for 5 min, and quick-chill in wet ice.
5. Disconnect the power supply, and immediately prior to applying the samples, thoroughly rinse out (using a 10-mL syringe with a 22-gauge needle) the wells of the gel with the upper reservoir TBE buffer; this prevents streaking of the DNA samples caused by urea that has diffused into the wells.
6. Using calibrated glass capillaries with finely drawn tips or, preferably, disposable flat-capillary pipette tips, load (as quickly as possible) 5 μL of each sample (plus the G + A sequencing ladder) onto the wells of the sequencing gel, sweeping the sample evenly from side to side. An untreated naked DNA sample (i.e., not subjected to the gel retardation assay) should always be diluted in sequencing-gel loading buffer, heat-denatured, and co-electrophoresed with the treated samples to verify the integrity of the DNA, since single-strand nicks can mask protein binding sites or even create artificial ones.
7. Remove all bubbles from the bottom of the gel (they may prevent even migration of the samples) using a 30-mL syringe with a bent 20-gauge needle.
8. Run the gel under pre-electrophoresing conditions (constant power, ~50–70 W), taking care not to overheat the glass plates. Uneven migration of fragments ("smiling") caused by an uneven gel temperature can be avoided by clamping an aluminum plate to the front glass plate. It is customary to electrophorese the samples until the bromophenol blue marker dye is about 3–5 cm from the bottom of the gel, but longer electrophoresis times may be required to obtain single-band resolution in the area of the footprint(s). The location of this region(s) depends on the distance between the radioactive label and the suspected protein-binding site(s) as well as on the length of the DNA fragment. Make use of available tables in the literature (referring to the migration of oligodeoxynucleotides in sequencing gels in relation to marker dyes) to determine how long to run your gel in order to achieve the desired electrophoretic resolution in

the region of the expected footprint(s) *(19)*; this will allow you to discern differences between the chemical cleavage patterns of the free DNA and that derived from the complexed fraction(s).

9. After completion of electrophoresis, remove the gel from the apparatus, and with the aid of a razor blade, slowly lift the siliconized plate. The thin polyacrylamide sheet will stick to the unsiliconized plate. Fix the gel for 15–20 min by gently immersing it (still attached to the lower plate) in a tank containing enough fixing solution; this removes excess urea that would otherwise crystallize out.
10. Carefully remove the plate bringing the gel on it from the tank, and lay it on a flat surface. Place a prewet (in fixing solution) sheet of backing paper (cut slightly bigger than gel dimensions) over the gel, press it gently down on the gel, roll out any air pockets (using a 10-mL glass pipette), and peel it off patiently and with extreme care together with the gel attached.
11. Cover the gel surface (but not the back of the filter paper) with a tightly drawn layer of plasticwrap. With a Kimwipe, push out any trapped air bubbles under the plasticwrap that might interfere with good uniform contact among the film, gel, and screen. Add two sheets of filter paper next to the backing paper as a support pad, and put the "sandwich" into a gel dryer (paper pad closest to vacuum source).
12. Dry the gel under vacuum at 80°C for 45–60 min. Do not release the vacuum before the gel is completely dried (sequencing gels with acrylamide concentrations >10% are susceptible to fracturing).
13. In the dark room, place a sheet of Kodak X-Omat AR or similar film against the plastic-covered face of the gel. Depending on the sensitivity required, it may be advisable to preflash the X-ray film, that is, expose the film to a millisecond flash of light prior to placing it in contact with the sample in order to achieve a background optical density of 0.15 (A_{540}) above that of the unexposed film. This increases sensitivity (the time of exposure to the radioactivity-generated light needed to blacken the film) and linearity of the film response (blackening above the background proportional to the amount of radioactivity), which are both essential if a densitometric analysis of the chemical cleavage products is to be performed (*see* **Notes 11 & 12**). Preflashing requires a photographic flash unit appropriately fitted with filters and adjusted as described *(21)* and is only useful if an intensifying screen is used, see next step.
14. Autoradiograph in a metal cassette containing a single calcium tungstate intensifying screen (place the flashed face of film

toward intensifying screen) at –70 to –80°C (to enhance effect of screen) for 12–16 h. Shorter or longer periods of time may be also required, since the clarity of the footprint depends on the intensity of the bands in the autoradiograph.

15. Immediately remove the film from the cassette and develop it, preferably in an automatic processing machine. If re-exposing the gel, it is necessary to let the cassette warm up before inserting a second film (cold cassettes will quickly collect moisture from the air).
16. Compare the chemical cleavage pattern of the naked DNA to that of the DNA–protein complex(es). The position of a band in the gel corresponds to the distance between the label and the point at which the DNA has been cleaved by the chemical reagent. Accordingly, bands at the bottom of the gel represent the smallest end-labeled DNA fragments, increasing in size as one reads up the gel until the pattern terminates abruptly in a strongly labeled band, corresponding to the uncleaved full-length probe (*see* **Fig. 3**). The protected region(s) (indicating sequence-specific protein binding) appears as an area resistant to cleavage (footprint), resulting in an almost complete absence of fragments (gap) arising from within the protein-binding site(s) in the chemical cleavage pattern of the complexed DNA. The nucleotides exhibiting protection are identified by aligning the bands in the cutting pattern of the free DNA with positions (bonds) in the sequence of the co-electrophoresed Maxam–Gilbert marker G + A track. In this comparison, it is necessary to note that, regardless of the end of the DNA fragment labeled in the experiment (5′ or 3′), the obtained set of products from the chemical cleavage reaction matches the mobilities of the G + A sequencing fragments exactly. This is a consequence of the identical 3′ and 5′ ends generated by both chemistries at the cleavage points *(22)*.

4. Notes

1. *Optimization of the analytical EMSA*. This is generally achieved by assessing binding-reaction parameters and gel electrophoresis conditions. Furthermore, success in interpretation of results from the coupled gel retardation/in situ OP-Cu footprinting assay depends critically on some properties of the DNA fragment used in the initial binding reaction. (a) *Properties of the DNA fragment used in the binding reaction*. An EMSA employing crude extracts works

best with short DNA fragments, since these reduce nonspecific interactions of proteins in the extract with sequences flanking the specific binding site(s) and are able to detect binding of large proteins more readily. Optimal sizes range between 100 and 150 bp, with the putative protein-binding site(s) located at approximately the center or at least 20–25 bp away from the radioactive labeled end. If a 20–25 bp synthetic oligonucleotide is to be used as a probe, it is advisable to design it in a way that it can be readily subcloned into the polylinker region of a suitable vector, and then labeled and released as a 40- to 45-bp restriction fragment, in order to obtain the desired single-base resolution within and around the expected footprint(s). Although for EMSAs the DNA fragment may be labeled at all ends, the subsequent footprinting analysis requires the DNA fragment to be radioactively labeled (to a high specific activity) at the 3′- or 5′-end of *one* of the two strands. Klenow enzyme-labeled probes are preferable to kinased probes because some protein extracts contain substantial phosphatase activities. Finally, the labeled probe should be un-nicked, because the resulting fragments may obscure the cleavage pattern obtained after the chemical attack in the footprinting analysis. Therefore, sufficient care should be taken to minimize nuclease activities during all steps of preparation, labeling, and isolation. To this end, we have found that purification of singly end-labeled probes from native polyacrylamide gels by "crush-and-soak" methods (similar to that described in **Subheading 3.4.1**) results in less damage to DNA than electroelution. (b) *Binding-reaction parameters.* These include binding-buffer composition (pH, ionic strength, metal ion content, and presence or absence of nonionic detergents and/or stabilizer polycations), amount of crude or partly fractionated extract or purified protein, concentration of labeled DNA probe, type and amount of bulk carrier DNA, and temperature and duration of incubation. Specifically, the following considerations should be evaluated: The optimal ratio of protein to DNA for the assay is best determined by titrating a fixed concentration of the radioactively labeled DNA fragment with increasing amounts of crude or partly fractionated extracts, or purified protein. Frequently, as protein concentration increases, binding passes through a maximum. Note, however, that increasing amounts of protein to a fixed concentration of DNA will not necessarily increase the yield of specific complex(es) seen. This is because of the fact that, whereas a given preparation of any DNA-binding protein(s) tends to be fully active in nonspecific binding, it is typically only fractionally active in site-specific binding activity [the

apparent fractional activity varying from 5 to 75%, depending on the particular protein(s) and the individual sample]. Therefore, too much protein, particularly with crude extract preparations, leads to occlusion of the binding site(s) by proteins interacting with DNA in a sequence-independent manner. This problem can be minimized by raising simultaneously the concentration of bulk carrier (competitor) DNA [typically of the order of 250- to 5,000-fold (w/w) excess over binding-site DNA]; this increases the occupancy of the binding site(s) by sequestering nonspecifically bound proteins, including nonspecific DNA-binding nucleases that may degrade the end-labeled DNA during the binding incubation. Bear in mind, however, that although some proteins are able to locate their target binding site(s) in the presence of vast excesses of nonspecific natural DNAs (sonicated salmon sperm or calf thymus DNA), other proteins cannot tolerate natural DNA carriers, but bind readily in the presence of an excess of synthetic polynucleotides, such as poly d(I–C)·d(I–C) or poly d(A–T)·d(A–T). It is noteworthy in the latter case that the efficacy of competition for nonspecific binding can vary among different batches from the same vendor. On the other hand, if too much carrier DNA is added, it will compete for the specific factor(s) of interest, and the level of complex(es) will decrease. Finally, provided an adequate resolution of DNA-bound species is obtained for a fixed concentration of competitor, increasing the amount of probe increases the fraction of DNA driven into complex(es), until the limit set by the binding constant(s) is reached. (c) *Gel electrophoresis conditions.* Gel parameters, such as percentage of acrylamide, degree of crosslinking, and pH and type of gel/running buffer system (high- or low-ionic strength) dramatically affect the size, aggregation state, and stability of DNA–protein complexes, hence their abundance and quality of separation. Accordingly, it may be necessary to try more than one gel fractionation/buffer system to obtain sufficient formation of the DNA–protein complex(es) of interest. The electrophoresis time has to be optimized for the complex(es) studied and the separation required (if more than one complex has to be mapped). The most promising conditions can then be applied in the subsequent preparative EMSA.

2. *Binding competition analysis.* If competitor DNA is identical to and in relatively large excess over the labeled DNA, >90% of the radioactive signal should be eliminated from complexes corresponding to protein(s) that interact with the binding-site DNA in a sequence-specific manner. Complexes unaffected or only slightly affected by the addition of competitor are thought to arise from nonspecific, low-affinity

binding by abundant proteins that are present in excess to binding-site DNA; DNA derived from these complexes after the in situ chemical cleavage reactions can serve as a negative control in the footprinting analysis, because its cutting pattern will closely resemble that of the free (unbound) probe.

3. *Assaying relative dissociation rates of DNA–protein complexes.* To follow dissociation kinetics, complexes are allowed to form under optimal reaction conditions and, at time zero, exposed to a large excess of an agent that does not perturb their stability (commonly 100- to 250-fold mass excess of nonspecific competitor DNA or, preferably, of the same DNA fragment unlabeled). The "scavenger" molecules sequester the protein(s) as it dissociates from its specific binding site(s), hence preventing it from rebinding. Analysis by the EMSA of aliquots at various times (from a few seconds to 2 h) after quenching the reaction with the sequestering agent, shows the amount of free DNA increasing while the level of protein-bound label diminishes. The experiment can be designed in a way that individual reactions can be started and quenched at different times, such that all reach the point at which they will be applied to the gel more or less simultaneously and electrophoresed for the same period of time. Since dissociation of typical DNA–protein complexes is a first-order process (i.e., independent of the concentration of complexes), the results of the analytical study are also applied to the subsequent preparative assay.
4. *Trailing of the bands during electrophoresis.* Glycerol-containing binding buffers tend to cause significant trailing at the edges of the bands during electrophoresis. If you noticed this trailing under the optimized conditions in the analytical assay, you may substitute glycerol for Ficoll [2.5% (w/v); Type 400, Pharmacia] in the preparative binding buffer. The presence of Ficoll (a copolymer of sucrose and epichlorohydrin), while not interfering with the thermodynamic and kinetic parameters of the binding reaction, gives rise to straight bands in the gel (by tending not to spread so much due to surface tension effects when loading the sample), thus minimizing the size of the polyacrylamide strips in the subsequent DNA-elution steps (**Subheading 3.4.1**).
5. *Loading the preparative polyacrylamide gel.* If you have problems with the sample not sinking to the bottom of the well (which may be caused by substantial differences between the binding buffer and the electrophoresis buffer, and/or the large volume of the preparative reaction), you can preload the well(s) of the gel with binding buffer or, alternatively, load your sample with the power supply running at 10–15 mA (*Wear dry plastic gloves and use only plastic tips if you do this!*).

6. (a) Use the recommended suppliers to obtain the liquid and powdered reagents. The care given to the preparation of reagents is crucial. In particular, the water used must be of the highest quality. Our laboratory uses only water purified with a Milli-Q system (Millipore), which removes virtually all organics, ions, and bacteria. Such precautions help prevent spurious reactions of impurities with the reagents. (b) You may try to footprint the complexes using a 1:1 ratio of 1,10-phenanthroline to copper, if your suspected protein-binding site(s) or the adjacent regions are particularly rich in 5′-CG-3′ elements (*see* **Subheading 1.2.1**). In this case, prepare the following solutions: 100 mM 3-mercaptopropionic acid, 5 mM OP, and 5 mM $CuSO_4$.
7. *Optimizing the time of in situ chemical treatment.* The length of incubation period is determined by several factors, among which the following are of particular importance (a) Kinetic stability of the complex(es) (**Subheading 3.1, step 3**). In principle, the higher the dissociation rate of the complex(es), the lower the time of exposure to the chemical nuclease. However, because of the gel "caging effect" and the in situ (on the DNA surface) funneling of the reaction (*see* **Subheading 1.1.1**), this rule of thumb is applicable only for DNA–protein complexes with half-lives either $\ll$ 1 min or $\gg$ 60 min; provided the complex(es) is relatively abundant (see below), incubation times <8 min and >30 min, respectively, should be used in these cases. (b) Relative abundance of the complex(es). The aim of the reaction is to generate, on average, about one chemical cleavage event per DNA strand. Assuming that the cleavage process is governed by Poisson statistics, the product oligonucleotides will statistically be the result of a single cleavage ("single-hit kinetics") when the concentration of the full-length labeled strand is ~50–70% of its original value (i.e., ~50–70% of the DNA fragments should be left uncleaved). Accordingly, when the abundance of the target complex(es) is very low (i.e., a bound DNA to free DNA ratio $\ll$ 1), early termination of the reactions will be critical for high-"off"-rate complexes, beneficial for intermediate-stability complexes, and safe for extremely stable complexes. (c) Temperature of incubation. If optimization of the DNA-binding reaction and, consequently, of the EMSA requires that both be performed at low temperature (4°C), you should carry out the chemical nuclease treatment at low temperature as well. However, the amount of dissolved air oxygen in the reaction mixture under these conditions is considerably higher (increases with decreasing temperature); since molecular oxygen catalyzes a rate-limiting step that generates in situ hydrogen peroxide (an essential

co-reactant for the chemical nuclease activity; *see* **Fig. 2**), its presence in more than stoichiometric amounts will shift the subsequent equilibria toward the right side, leading to increased rates of DNA-strand scission. To compensate for this accelerated cleavage kinetic, you should decrease the time of exposure to the chemical nuclease. Inversely, incubation times should be longer than originally established (or you may add hydrogen peroxide exogenously) for chemical treatments performed at bench temperature during hot summer days in non-airconditioned rooms. (d) Concentration of reagents. Increasing the concentration of OP while holding the concentrations of copper (II) sulfate and 3-mercaptopropionic acid constant increases the overall rate of DNA-strand scission, without significantly affecting the sequence preferences of cleavage and the resulting fragment size distribution *(15)*. On the other hand, substituting 3-mercaptopropionic acid by the same concentration of ascorbic acid reduces the overall rate of DNA cleavage both at low (1:1) and high (>4.5:1) 1,10-phenanthroline–copper molar ratios *(15)*. These observations are of practical significance to estimating the time of chemical treatment, particularly when extremely labile or extraordinary stable complexes are being mapped. A higher concentration of OP and a short incubation time might be used in the former case, whereas ascorbate (as the reducing agent) and prolonged treatments are recommended in the latter. (e) Presence of dithiothreitol (DTT). DTT (often employed in EMSAs to maintain the activity and/or stability of DNA-binding proteins) slows the DNA cleavage rate because this chemical sequesters the copper necessary for the oxidative cleavage to occur. If excessive amounts of DTT (i.e., >1 mM) have been used either in the preparative binding reaction or for casting the gel in which the $(OP)_2Cu^+$ cleavage reaction is to be performed, longer incubation times or increased $(OP)_2Cu^+$ concentrations should be used to restore the cleavage efficiency.

8. 2,9-Dimethyl-OP is a redox inert analog of OP which acts as a Cu^+-specific chelator; it will bind all the metal ion, preventing further oxidative chemistry and DNA cleavage.
9. Phase inversion during organic extractions of the gel- or membrane-eluted DNA samples. Because of the high-salt content of the elution buffers, the aqueous phase (which normally forms the upper layer) may sometimes be dense enough to form the lower layer. If this is the case, the aqueous phase can be easily identified by monitoring the eluted radioactivity with a Geiger counter or by following the strong yellow color of the organic phase (contributed by

8-hydroxyquinoline that is added to phenol during equilibration as an antioxidant; *see* **item 2** of "Solutions" under **Subheading 2.3.1**).

10. Excess acrylamide in the gel-eluted DNA samples. If the DNA pellet is highly contaminated with acrylamide monomers or other impurities from the gel matrix (a problem sometimes encountered when employing low-percentage mobility-shift gels and is usually apparent from the formation of an excessive turbidity during the ethanol-precipitation step), then you can further purify it by passing through a pre-equilibrated Elutip™-d mini-column (an ion-exchange column) according to the manufacturer's instructions. Adsorption to and desorption from the column can be followed with a Geiger counter. Alternatively, you may follow the experimental strategy developed by Ragnhildstveit et al. *(23)*.
11. Appearance of the bands on the autoradiogram. If the band sharpness or shape in the chemical cleavage patterns on the autoradiogram suffers (e.g., narrowing of the bands toward the smallest end-labeled DNA fragments), the DNA samples contain residual proteins, salts, or acrylamide contaminants. These faults should be expected if uneven running and retardation of marker dyes is observed during electrophoresis, and can be avoided if careful attention is paid during **steps 15–18** in **Subheading 3.4.1**. If necessary, the number of organic extractions (**Subheading 3.4.1, step 15**) and/or ethanol washes (**Subheading 3.4.1, step 18**) can be increased; acrylamide contaminants are efficiently removed by the use of Elutip™-d mini-columns (*see* **Note 10**). Note, however, that smeared or fuzzy bands without apparent electrophoretic problems may also arise from scattering during autoradiography; use a light-tight metal film cassette with a particularly effective closure mechanism, capable of firmly fixing filter/gel and film, and with intensifying screen in place (those produced by Wolf or Picker International Health Care Products, Highland Heights, OH, are best in that).
12. Wondering about the footprint: scanning the autoradiogram. If the effects of protein binding on the cleavage rate of the chemical nuclease are not clearly discernible by the eye, or when partial protection is obtained, you may analyze the ladders quantitatively by subtracting the cleavage pattern of the DNA derived from the bound fraction from that derived from the free fraction. This involves calculating the probability of cleavage at each bond (which is related to the amount of radioactivity, or intensity, in the corresponding band in the cutting pattern) and, finally, for each lane, the average number of cuts in the DNA strand, with the aid of automated laser densitometers linked to a computer

(available from, for instance, Bio-Rad or LKB-Pharmacia). Because quantification of band intensity is limited within the linear response of Kodak X-Omat AR films to radioactivity (bands with an optical density >0.15 and <1.8 absorbance units), it is essential not to use overexposed films for this type of analysis. Moreover, the film should be free of scratches, fingerprints, or other blemishes that will appear as optical signals indistinguishable from ^{32}P, thus interfering with the calculations. A difference probability plot along the entire length of the binding site(s) and surrounding DNA can be obtained, revealing the protein-binding site(s) much more clearly and reliably than can be done by comparison by eye of the chemical cleavage patterns. If still in doubt, however, you may analyze the footprint on the other DNA strand.

Acknowledgments

This work has profited greatly from discussions of the author with the students and teachers participating in the 1991–1996 EMBO Courses on "DNA–Protein Interactions".

Further Reading

Sigman, D. S., Kuwabara, M. D., Chen, C. H., and Bruice, T. W. (1991). Nuclease activity of 1,10- phenanthroline–copper in study of protein–DNA interactions. *Methods Enzymol.* **208**, 414–433.

Papavassiliou, A. G. (1995). Chemical nucleases as probes for studying DNA–protein interactions. *Biochem. J.* **305**, 345–357.

References

1. Schmitz, A. and Galas, D. J. (1978). The interaction of RNA polymerase and lac repressor with the *lac* control region. *Nucleic Acids Res.* **6**, 111–137.
2. Humayun, Z., Kleid, D., and Ptashne, M. (1977). Sites of contact between λ operators and λ repressor. *Nucleic Acids Res.* **4**, 1595–1607.
3. Garner, M. M. and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**, 3047–3060.
4. Revzin, A. (1989). Gel electrophoresis assays for DNA–protein interactions. *Biotechniques* **7**, 346–355.
5. Topol, J., Ruden, D. M., and Parker, C. S. (1985). Sequences required for in vitro transcriptional activation of a *Drosophila* hsp 70 gene. *Cell* **42**, 527–537.
6. Kuwabara, M. D. and Sigman, D. S. (1987) . Footprinting DNA–protein complexes *in situ* following gel retardation assays using 1,10-phenanthroline–copper ion: *Escherichia coli* RNA polymerase-*lac* promoter complexes. *Biochemistry* **26**, 7234–7238.
7. Sigman, D. S., Graham, D. R., D'Aurora, V., and Stern, A. M. (1979). Oxygen-dependent cleavage of DNA by the 1,10-phenanthroline–cuprous complex. Inhibition of *Escherichia coli* DNA polymerase I. *J. Biol. Chem.* **254**, 12269–12272.
8. Marshall, L. E., Graham, D. R., Reich, K. A., and Sigman, D. S. (1981). Cleavage of DNA by the 1,10-phenanthroline–cuprous complex. Hydrogen peroxide requirement: primary and

secondary structure specificity. *Biochemistry* **20**, 244–250.

9. Goyne, T. E. and Sigman, D. S. (1987). Nuclease activity of 1,10-phenanthroline–copper ion. Chemistry of deoxyribose oxidation. *J. Am. Chem. Soc.* **109**, 2846–2848.
10. Pope, L. M., Reich, K. A., Graham, D. R., and Sigman, D. S. (1982). Products of DNA cleavage by the 1,10-phenanthroline copper complex. Identification of *E. coli* DNA polymerase I inhibitors. *J. Biol. Chem.* **257**, 12121–12128.
11. Tamilarasan, R., McMillin, D. R., and Liu, F. (1989). Excited-state modalities for studying the binding of copper phenanthrolines to DNA, in *Metal-DNA Chemistry* (Tullius, T. D., ed.), ACS Symposium Series **402**, Washington, DC, pp. 48–58.
12. Kakkis, E. and Calame, K. (1987). A plasmacytoma-specific factor binds the c-*myc* promoter region. *Proc. Natl Acad. Sci. USA* **84**, 7031–7035.
13. Flanagan, W. M., Papavassiliou, A. G., Rice, M., Hecht, L. B., Silverstein, S., and Wagner, E. K. (1991). Analysis of the Herpes Simplex Virus type 1 promoter controlling the expression of UL38, a true late gene involved in capsid assembly. *J. Virol.* **65**, 769–786.
14. Veal, J. M. and Rill, R. L. (1988). Sequence specificity of DNA cleavage by bis(1,10-phenanthroline)copper(I). *Biochemistry* **27**, 1822–1827.
15. Veal, J. M., Merchant, K., and Rill, R. L. (1991). The influence of reducing agent and 1,10-phenanthroline concentration on DNA cleavage by phenanthroline + copper. *Nucleic Acids Res.* **19**, 3383–3388.
16. Papavassiliou, A. G. and Silverstein, S. J. (1990). Interaction of cell and virus proteins with DNA sequences encompassing the promoter/regulatory and leader regions of the Herpes Simplex Virus thymidine kinase gene. *J. Biol. Chem.* **265**, 9402–9412.
17. Darsillo, P. and Huber, P. W. (1991). The use of chemical nucleases to analyze RNA– protein interactions: the TFIIIA–5S rRNA complex. *J. Biol. Chem.* **266**, 21075–21082.
18. Papavassiliou, A. G. (1993). *In situ* (OP)-2-Cu^{+} mapping of electrophoretically resolved RNA–protein complexes. *Anal. Biochem.* **214**, 331–334.
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Sambrook, J., Fritsch, E., and Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual, Second Edition*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
21. Laskey, R. A. (1980). The use of intensifying screens or organic scintillators for visualizing radioactive molecules resolved by gel electrophoresis. *Methods Enzymol.* **65**, 363–371.
22. Maxam, A. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
23. Ragnhildstveit, E., Fjose, A., Becker, P. B., and Quivy, J. P. (1997). Solid phase technology improves coupled gel shift/footprinting analysis. *Nucleic Acids Res.* **25**, 453–454.

Chapter 14

Determination of a Transcription Factor-Binding Site by Nuclease Protection Footprinting onto Southwestern Blots

Athanasios G. Papavassiliou

Summary

DNA-transcription factor interactions in eukaryotic systems have been documented by a broad gamut of biochemical techniques including deoxyribonuclease I (DNase I) footprinting and Southwestern (SW) assays. In spite of their wide applicability, each of these approaches provides only partial information about DNA-protein complexes. DNase I footprinting identifies the extent and location of the binding site within the DNA but does not yield information about the protein(s) involved. On the other hand, the SW assay can reveal the relative size of active protein species in crude extracts, facilitating their identification, but fails to localize their binding site within the probing DNA sequence. Coupling SW and in situ (on-blot) DNase I footprinting methodologies has the dual potential of accurately determining the molecular mass of individual DNA-binding transcription factors and precisely mapping their cognate binding sites.

Key words: DNA–protein interaction, DNA-binding transcription factor, Southwestern (SW) assay, Nuclease protection, Deoxyribonuclease I (DNase I), on-blot DNase I footprinting.

1. Introduction

The interaction of cell-type specific or inducible transcription factors with regulatory DNA sequences in gene promoters or enhancers is a pivotal step in genetic reprogramming during cell proliferation and differentiation and in response to extracellular stimuli. The study of these interactions and the characterization of the factors involved are, therefore, a critical aspect of gene control. Transcription factor-DNA interactions in eukaryotes have been demonstrated by a wide variety of biochemical approaches

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_14

including deoxyribonuclease I (DNase I) and chemical nuclease footprinting *(1–3)* (*see* Chapters "DNase I Footprinting," "Hydroxyl Radical Footprinting of Protein-DNA Complexes," "The Use of Diethyl Pyrocarbonate (DEPC) and Potassium Permanganate as Probes for Strand Separation and Structural Distortions in DNA," and "Uranyl Photofootprinting"), methylation protection *(4)* (*see* Chapter "Identification of Protein/DNA Contacts with Dimethyl Sulfate: Methylation Protection and Methylation Interference"), electrophoretic mobility shift *(5, 6)* (*see* Chapter "Electrophoretic Mobility Shift Assays for the Analysis of DNA-Protein Interactions"), and Southwestern (SW) assays *(7)* (*see* Chapter "Identification of Sequence-Specific DNA-Binding Proteins by Southwestern Blotting"). Despite their broad applicability, these techniques provide only partial information about the DNA-protein system under investigation. The first three techniques identify either the site(s) of transcription factor binding within the DNA (size and location of nucleotide stretches or atoms on individual bases) or the complexity of the binding pattern (stoichiometry), but do not yield information about the protein(s) involved. On the other hand, the SW assay reveals the relative molecular mass of renatureable (on a membrane support) active species in heterogeneous protein mixtures facilitating their identification, but fails to localize the exact target element within the probing DNA sequence.

Combining SW and DNase I (but also chemical nuclease and methylation protection, *see below*) footprinting methodologies has the dual potential for accurately determining the size of individual DNA-binding transcription factors and precisely mapping their cognate binding sites *(8)* (**Fig. 1**). In addition to allowing the detection of only fractional binding (footprints in solution can be obtained only when the binding site[s] is almost completely occupied), coupling in situ DNase I footprinting with the SW technique offers the advantage of both resolving the protein component and mapping the binding site of complexes formed by either different factors recognizing an identical sequence within the DNA probe or by two factors interacting in a noncooperative manner with adjacent but distinct sequences. However, this is dependent on the factor being able to bind to DNA as a monomer or a homodimer. If the active form is not a single species (i.e., a heterodimeric or heteromeric complex is required to reconstitute the binding activity), specific DNA binding will not be detected and the procedure will not be applicable. The most critical stage of the coupled assay lies within its first part, namely the SW procedure, and concerns the ability of a transcription factor to efficiently renature into its active form (at least in terms of DNA-binding capacity) on the membrane. Many transcription factors are composed of domains with distinct structural conformations, which aids the process of renaturation. However, since the protein surface immobilized on the membrane poses an

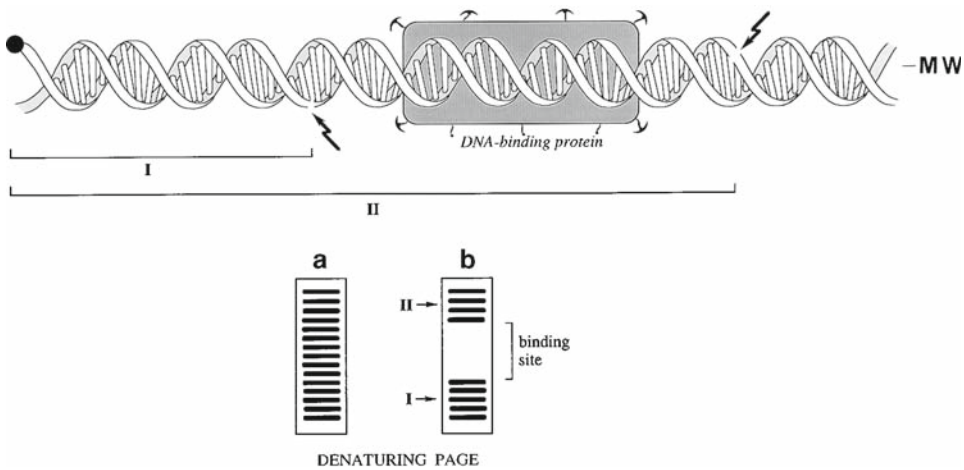


Fig. 1. Rationale in a combined SW – DNase I footprinting procedure. The "*hammer*"*-shaped* extensions from the DNA-binding protein indicate immobilization on the blotting membrane. The *black-filled sphere* indicates the end label in one strand of the DNA footprint probe. *MW* denotes the molecular size of the membrane-blotted DNA-binding protein (estimated by comparing the position of the active membrane area to the mobilities of coelectroblotted protein MW standards). *Arrows mark* representative sites of DNase I attack. (**a**) DNase I digestion pattern of free (in solution) probe; (b) DNase I digestion pattern of protein-complexed (membrane-bound) probe.

impediment on the refolding process, the likelihood of successful renaturation increases with increased size of the DNA-binding transcription factor(s) under study. As a result of the increased kinetic stability of a membrane-immobilized DNA-protein complex (reversible binding to even low-affinity proteins is enhanced because excess unbound DNA has been washed out and hence is not present to compete), reaction parameters such as the size of the DNA probe, the concentration of DNase I, and digestion time are no longer determined by the dissociation rate of the complex, a normal limitation on DNase I protection assays performed in solution.

The fidelity of the combined analytical assay is demonstrated in **Fig. 2**. Evidently, the structural and functional properties of DNA are not altered by entrapment on the blotting membrane surface (i.e., the probe exhibits identical pr otection pattern and sequence-dependent reactivity with DNase I as it does in solution). Therefore, this coupled assay provides a fast and reliable method which will allow the user who has identified specific regulatory regions in the gene of interest to begin characterizing in detail the transcription factor(s) that bind to them in a certain cellular milieu.

Modifications of the presented DNase I protection analysis on SW blotting membranes (in situ) substitute the enzymatic probe for either the chemical nuclease 1,10-phenanthroline-copper ion (OP-Cu) *(9)* or dimethyl sulfate (DMS) *(10)*. Both protocols have the additional advantage of providing information on the nature (i.e., specific -protected versus nonspecific – unprotected) of several membrane-immobilized DNA-protein species often

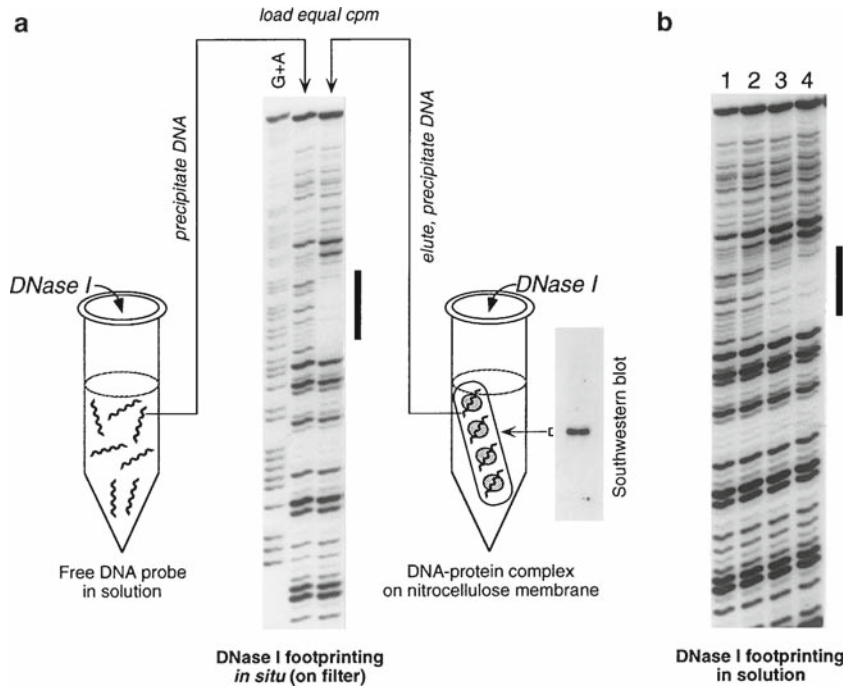


Fig. 2. (**a**) Schematic outline of the manipulations involved in the combined SW – DNase I footprinting procedure. In the example presented, a crude lysate of bacterial cells overexpressing the proto-oncoprotein c-Jun (a component of the transcription factor AP-1) was subjected to a quantitative SW assay utilizing as probe a 134-bp DNA restriction fragment (5′-end labeled in the coding strand) encompassing the AP-1-binding sequence of the human collagenase promoter [5′(-72)TGAGTCA3′(-66)]. (**b**) DNase I footprinting reactions of the same fragment in solution performed with increasing amounts (*lanes 2–4*) of the c-Jun preparation (*lane 1*, free DNA probe). The footprinted region (*solid bar*) includes in both cases ~14 bases centered around the AP-1-binding motif *(20)*.

observed in an SW assay. The molecular and functional properties of DNase I (i.e., its relatively large size, mode of target searching and binding to cleave, and requirement for Mg^{2+} (which often stabilizes both specific and nonspecific complexes)) highly reduce its potential to detect these differences. In addition, these methodologies are useful in rapidly confirming the binding specificity of a protein isolated by screening a cDNA expression library with recognition-site DNA.

2. Materials

2.1. SW Blotting

2.1.1. Solutions

1. Sodium dodecyl sulfate (SDS; 20%; w/v): Dissolve 100 g of SDS (Sigma-Aldrich) in distilled, deionized water to a 0.5 L final volume; heat at 68R°C to assist dissolution (do not autoclave). Store at room temperature in a clear bottle. A respirator or dust mask should be worn when handling powdered SDS.

2. Lower (separating) gel buffer (4R× stock): 1.5 M Tris–HCl of pH 8.8, 0.4% (w/v) SDS. Filter through a 0.22-μm membrane filter. Store at 4R°C.
3. Upper (stacking) gel buffer (4R× stock): 0.5 M Tris–HCl of pH 6.8, 0.4% (w/v) SDS. Filter through a 0.22-μm membrane filter. Store at 4R°C.
4. Acrylamide/*bis*-acrylamide gel mixture: 30% (w/v) acrylamide, 0.8% (w/v) *N,N*R′-methylene-*bis*-acrylamide (Bio-Rad, Richmond, CA). Filter through a 0.22-μm membrane filter. Store at 4R°C in the dark. Powdered acrylamide and *N,N*R′-methylene-*bis*-acrylamide are potent neurotoxins and are absorbed through the skin. Their effects are cumulative. Wear gloves and a facemask when weighing these substances and when handling solutions containing them (work in a chemical fume hood). Although polyacrylamide is considered to be nontoxic, it should be handled with care because of the possibility that it might contain small quantities of unpolymerized acrylamide.
5. Ammonium persulfate (10%; w/v): Dissolve (by vigorous vortexing) 1 g of ammonium persulfate (Bio-Rad) in 10 mL of distilled, deionized water. Filter through a 0.22-μm membrane filter. Store at 4R°C; make fresh weekly. Ammonium persulfate is extremely destructive to tissue of the mucous membranes and upper respiratory tract, eyes, and skin. Inhalation may be fatal. Exposure can cause gastrointestinal disturbances and dermatitis. Wear gloves, safety glasses, respirator, and other protective clothing and work in a chemical fume hood. Wash hands thoroughly after handling.
6. SDS-polyacrylamide gel electrophoresis (PAGE) running buffer (10R× stock): 0.25 M Tris base (Sigma-Aldrich), 1.92 M glycine (Sigma-Aldrich). It is not necessary to adjust the pH. The proportions of Tris base and glycine give pH 8.3. Store at room temperature in a large vessel (carboy). Dilute to 1R× with distilled, deionized water; then add SDS to a final concentration of 0.1% (w/v).
7. SDS-PAGE sample buffer (4R× stock): 0.5 M Tris–HCl of pH 6.8, 8% (w/v) SDS, 40% (v/v) glycerol (Sigma-Aldrich), 0.4% (w/v) bromophenol blue (Sigma-Aldrich). Store in aliquots at −20R°C. Before use, make a working solution of SDS-PAGE sample buffer by diluting the 4R× stock buffer and adding 2-mercaptoethanol (Sigma-Aldrich) to a final concentration of 5% (v/v) (*see* **Note 1**). 2-mercaptoethanol is harmful if inhaled or absorbed through the skin. High concentrations are extremely destructive to the mucous membranes, upper respiratory tract, skin, and eyes. Use only in a chemical fume hood. Gloves, safety glasses, and respirator should be worn.

8. Western transfer buffer: 50 mM Tris base, 380 mM glycine, 0.1% (w/v) SDS, 20% (v/v) methanol (*see* **Note 2**). There is no need to adjust the pH of this buffer by the addition of acid or alkali. Store at room temperature in a large vessel.
9. Dithiothreitol (stock): Prepare a stock solution of 0.5 M dithiothreitol (Sigma-Aldrich) in distilled, deionized water and store in aliquots at –20R°C. (*see* earlier safety Note for 2-mercaptoethanol).
10. Phenylmethylsulfonyl fluoride (PMSF; stock): Prepare a 100-mM stock solution of PMSF in absolute ethanol and store in aliquots at –20R°C. PMSF is extremely destructive to the mucous membranes of the respiratory tract, the eyes, and the skin. It may be fatal if inhaled, swallowed, or absorbed through the skin. It is a highly toxic cholinesterase inhibitor. Therefore it should be used in a chemical fume hood and gloves and safety glasses should be worn during handling.
11. 1 M KCl.
12. 0.5 M $MgCl_2$.
13. SW blocking/renaturation buffer: 3% (w/v) nonfat dried milk (or 5% (w/v) lipid-free BSA; *see* Note 3a), 25 mM HEPES.KOH of pH 7.5, 50 mM KCl, 6.25 mM $MgCl_2$, 1 mM dithiothreitol (freshly added from the 0.5 M stock solution immediately before use), 10% (v/v) glycerol, 0.1% (v/v) Nonidet P-40 (IGEPAL CA630, Sigma-Aldrich), 0.2 mM PMSF (freshly added from the 100 mM stock solution just prior to use). Store at 4R°C. Prepare also some SW blocking/renaturation buffer minus dried milk (or BSA).
14. Polyvinyl alcohol (stock): Prepare a 10% (w/v) stock solution of polyvinyl alcohol (Sigma-Aldrich P8136; average MW = 10,000) in distilled, deionized water and store at –20R°C.
15. SW binding/washing buffer: 12.5 mM HEPES.KOH of pH 7.5, 50 mM KCl, 6.25 mM $MgCl_2$, 0.5 mM dithiothreitol (freshly added from the 0.5 M stock solution immediately before use), 2% (w/v) polyvinyl alcohol, 10% (v/v) glycerol, 0.05% (v/v) Nonidet P-40, 0.2 mM PMSF (freshly added from the 100 mM stock solution just prior to use). Store at 4R°C (*see* **Note 4b**).
16. Phenol/chloroform/isoamyl alcohol (25:24:1; v/v): Mix just prior to use 25 vol. of phenol with 24 vol. of chloroform and 1 vol. of isoamyl alcohol. Phenol is highly corrosive and can cause severe burns. Any areas of skin that come in contact with phenol should be rinsed with a large volume of water or PEG 400 (Sigma-Aldrich) and washed with soap and water (*do not use ethanol!*). Chloroform is irritating to the skin, eyes, mucous membranes, and respiratory tract. It is also a carcinogen and may damage the liver and kidneys.

Wear gloves, protective clothing, safety glasses, and respirator when handling these substances and carry out all manipulations in a chemical fume hood.

17. Chloroform/isoamyl alcohol (24:1; v/v): Mix 24 vol. of chloroform with 1 vol. of isoamyl alcohol. This organic mixture can be stored at room temperature in dark (brown) bottles indefinitely.
18. Nonspecific competitor DNA (stock solution): Dissolve salmon/herring sperm or calf thymus DNA in distilled water, deproteinize it by sequential phenol/chloroform/isoamyl alcohol (25:24:1; v/v) and chloroform/isoamylalcohol (24:1; v/v) extractions, sonicate to reduce the mean DNA length to 100–200 bp, precipitate the DNA with ethanol, then resuspend it at 1 mg/mL in distilled, deionized water.

2.1.2. Reagents/Special Equipment

1. *N,N,N*R′,*N*R′-tetramethylethylenediamine (TEMED; Bio-Rad). TEMED is extremely destructive to tissue of the mucous membranes and upper respiratory tract, eyes, and skin. Inhalation may be fatal. Prolonged contact can cause severe irritation or burns. Wear gloves, safety glasses, respirator, and other protective clothing and work in a chemical fume hood (TEMED is also flammable!). Wash hands thoroughly after handling.
2. Protein extract of interest (e.g., whole cell-free extract, crude nuclear extract, or partially purified extract), as concentrated as possible.
3. Prestained nonradioactive MW standards (Bio-Rad), or [^{14}C]-methylated protein MW markers (Amersham Biosciences).
4. Nonfat dried milk powder (e.g., Cadbury's Marvel or Carnation), or fatty acid-free BSA (fraction V).
5. Phenol: redistilled (under nitrogen) phenol equilibrated with 100 mM Tris–HCl of pH 8.0, 1 mM EDTA in the presence of 0.1% (w/v) 8-hydroxyquinoline. Phenol can be stored at 4R°C in dark (brown) bottles for up to 2 months. See relevant safety Note in **Subheading 2.1.1**.
6. Absolute ethanol.
7. Salmon/herring sperm or calf thymus DNA (Sigma-Aldrich or other supplier).
8. Singly ^{32}P end-labeled DNA probe bearing the binding site(s) of interest (*see* **Note 5**). All necessary precautions should be observed to minimize exposure to ionizing radiation during labeling and isolation of the probe; work behind protective screens whenever possible.

9. Radioactive ink: Mix a small amount of ^{32}P with waterproof black drawing ink, to a concentration of ~200 cps (on a Geiger counter) per µL.
10. 0.22-µm membrane filters (Millipore, Bedford, MA).
11. Mini-slab gel electrophoresis apparatus, giving 0.5–1.0-mm thick mini-gels (e.g., Bio-Rad Mini Protean system), and accompanying equipment.
12. High current (2–3 A) power supply (e.g., Bio-Rad or Hoefer, San Francisco, CA) and electroblotting apparatus for Western transfer (e.g., Bio-Rad Trans-Blot); additional equipment for Western transfer *(11)*.
13. Nitrocellulose (NC) membrane: suitable membranes comprising unsupported or supported nitrocellulose are available from a number of manufacturers, such as Schleicher & Schuell (BA85, 0.45 µm), Millipore (Immobilon-NC), and Amersham Biosciences (Hybond-C or -C extra).
14. Plastic trays.
15. Forceps.
16. Plasticwrap such as cling film or Saran Wrap.
17. X-ray film (e.g., Kodak X-Omat AR, Rochester, NY).
18. Additional equipment: Protective gloves/glasses/respirator; 4R°C shaking air incubator; sonicator; 25R°C shaking air incubator; Geiger counter; Kimwipes; all equipment for autoradiography.

2.2. Exposure of SW Blots to DNase I Treatment

2.2.1. Solutions

1. Eppendorf tube siliconization solution: 2% (v/v) dimethyldichlorosilane in *1,1,1*-trichloroethane. Eppendorf tubes should be silanized by briefly immersing the opened tubes in a beaker containing this solution, pouring off excess solution, and allowing the tubes to dry in air at room temperature. Dimethyldichlorosilane is particularly toxic. Gloves, safety glasses, respirator, and other protective clothing should be worn when handling it and should only be used in a chemical fume hood.
2. **Items 1**, **12** and **15–17** of **Subheading 2.1.1**.
3. DNase I (stock solution): Dissolve DNase I in 50% glycerol (in distilled, deionized water) to a concentration of 2.5 mg/mL. Store frozen in 10-µL aliquots at –20R°C or –70R°C. This stock is stable indefinitely.
4. 1 M $CaCl_2$.
5. DNase I reaction buffer: 10 mM $MgCl_2$, 5 mM $CaCl_2$. Store at room temperature.
6. DNase-STOP solution A: 20 mM HEPES.KOH of pH 7.5, 20 mM EDTA, 0.5% (w/v) SDS. Store at 4R°C.

7. 5 M NaCl.
8. DNase-STOP solution B: 60 mM HEPES.KOH of pH 7.5, 0.6 M NaCl, 60 mM EDTA, 1.5% (w/v) SDS. Store at 4R°C.
9. Proteinase K (stock solution): Dissolve Proteinase K in TE (10 mM Tris–HCl of pH 7.4, 1 mM EDTA) to a concentration of 2.5 mg/mL. Store in aliquots at –20R°C.
10. Probe elution buffer: 20 mM HEPES.KOH of pH 7.5, 0.3 M NaCl, 3 mM EDTA, 0.2% (w/v) SDS, 50 μg/mL Proteinase K.
11. Glycogen (stock solution): Prepare a stock solution of 10 mg/mL glycogen (Sigma-Aldrich G0885) in distilled, deionized water and store in aliquots at –20R°C; glycogen is used as a carrier to promote the precipitation of nucleic acids.
12. Ice-cold 80% (v/v) ethanol.
13. Formamide loading buffer: 90% (v/v) deionized formamide, 0.5R× TBE (*see later*), 0.025% (w/v) xylene cyanol FF (Sigma-Aldrich), 0.025% (w/v) bromophenol blue. Store at –20R°C after filtering. Formamide is a teratogen; take all safety precautions to avoid contact during manipulations involving this reagent.
14. TBE buffer (5R× stock): 445 mM Tris base, 445 mM borate, 12.5 mM EDTA. Dissolve (under stirring for at least 1 h) 272.5 g of ultrapure Tris base, 139.1 g of boric acid, and 23.3 g of EDTA.(Na_2) dihydrate in 4.5 L of distilled, deionized water. Make up to a final volume of 5 L. It is not necessary to adjust the pH of the resulting solution which should be around 8.3. Store at room temperature in a large vessel. This stock solution is stable for many months, but it is susceptible to the formation of a precipitate and should be inspected visually from time to time.
15. Fixing solution: 10% (v/v) acetic acid, 10% (v/v) methanol.

2.2.2. Reagents/Special Equipment

1. DNase I (DPFF grade; Worthington, Freehold, NJ).
2. Proteinase K.
3. Absolute ethanol (at –20R°C).
4. Single-edged disposable razor blades.
5. Forceps.
6. Siliconized 0.5-mL Eppendorf microcentrifuge tubes.
7. Siliconized 1.5-mL Eppendorf microcentrifuge tubes.
8. Drawn-out Pasteur pipets.
9. Whatman (Clifton, NJ) 3MM paper.
10. **Items 16** and **17** in **Subheading 2.1.2**.

11. Intensifying screen and film adapted for sensitive ^{32}P autoradiogroaphy.
12. Additional equipment: beaker; sharp-tip pencil; wet ice; scintillation vials/counter; vortexer; electronic timer; microcentrifuge (Eppendorf or equivalent); vortexing shaker set at 37R°C; Geiger counter; SpeedVac concentrator (Savant, Hicksville, NY); thermostatted heating block at 95R°C; plastic tank at gel dimensions (for gel fixing); vacuum gel dryer; all equipment for autoradiography.

3. Methods

3.1. SW blotting

The SW protocol involves four steps: electrophoretic separation of proteins by SDS-PAGE, electroblotting of the gel-fractionated proteins onto nitrocellulose (NC) membrane, probing of the blocked blot with the desired DNA probe, and detection of bound DNA by autoradiography of the washed, wet membrane.

1. Prepare and load the sample(s) and protein MW markers onto a standard SDS-polyacrylamide gel *(11, 12)* and electrophorese at an appropriate voltage until the bromophenol blue dye reaches the bottom of the gel. An 8–10% polyacrylamide-separating gel is capable to resolve throughout most of the known DNA-binding transcription factor size-range. Since the strength of the final signal obtained by the SW technique is proportional to the quantity of protein electrophoresed on the gel, best results are obtained by running the maximum amount of extract that does not overload the gel. For typical 0.5–1 mm thick protein mini-gels, this is usually 30–150 μg of whole cell-free, crude nuclear, or partially purified extract protein per lane (*see* **Note 1**).
2. Remove the electrophoresed gel from the glass plates. Assemble a Western blot sandwich, place it in the electrophoresis tank containing an appropriate volume of Western transfer buffer, and electroblot the proteins in the electrophoresed gel onto an NC membrane according to standard Western blotting protocols *(11, 13)*. Bear in mind that the best protein transfer is usually achieved by longer transfer times (i.e., 30 V [40 mA] overnight at 4R°C) (*see* **Note 2**).
3. When transfer is complete, gently peel the NC membrane off the gel, place it in a plastic tray, and gently wash for 10 min with 20–30 mL of SW blocking/renaturation buffer omitting dried milk (or BSA).

4. Decant the solution and replace it with a sufficiently large volume of SW blocking/renaturation buffer to completely immerse the membrane (usually 30–50 mL).
5. Incubate overnight at 4R°C with gentle rocking or shaking to block nonspecific binding sites on the membrane and to allow renaturation of the filter-immobilized proteins (*see* **Note 3**).
6. Using forceps, transfer the membrane to a fresh plastic tray and gently wash for 5 min with 20–30 mL of SW binding/washing buffer. The membrane can be stored in this buffer at 4R°C for up to 1 day before incubation with the DNA probe.
7. Immerse the membrane (using forceps) in a fresh plastic tray containing a radioactive probe mixture consisting of:
   - SW binding/washing buffer (*see* **Note 4**).
   - 5–10 μg (specific activity ~10_7 c.p.m./μg) of an asymmetrically [^{32}P]-labeled DNA fragment bearing the recognition site(s) for the sequence-specific DNA-binding factor(s) of interest (*see* **Note 5**).
   - 20 μg/mL nonspecific competitor DNA (*see* **Note 6**).
     To make the probe concentration as high as possible, the volume of SW binding/washing buffer should be the minimum needed to cover the membrane fully (smallest volumes are achieved if the membrane is sealed in a plastic bag). *Work behind perspex or glass shields!*

◦ Incubate for 2–4 h at room temperature with gentle rocking or shaking. Longer incubation times at a lower temperature may result in a better signal from a poorly binding factor.

◦ Carefully remove the radioactive probe mixture and dispose of it safely (*work behind perspex or glass shields!*).

◦ Using forceps, transfer the membrane to a plastic tray and wash for 10 min with 100 mL of SW binding/washing buffer on a platform shaker (room temperature).

◦ Repeat **step 10** two to three times or until the radioactive level of the membrane no longer falls appreciably between washes (this can be monitored by a Geiger counter) (*see* **Note 7**).

◦ Place the wet NC sheet between two layers of tightly drawn cling film. With a pad of Kimwipes, push out any trapped air bubbles under the cling film.

◦ Expose to X-ray film at 4R°C to locate regions of radioactive signal (protein[s]-bound probe). Exposure times of 1–3 h are usually sufficient to detect the radioactive species on the membrane. (*see* **Note 8**).

3.2. Exposure of SW Blots to DNase I Treatment

The in situ footprinting reaction is done in four stages: localization and excision of the areas on the NC sheet corresponding to protein-bound probe, partial digestion of the individual strip-associated and control (free in solution) DNAs with DNase I, extraction of the protein-bound DNA from the strip immobilized protein-DNA complex, and analysis of the free and complexed DNA digestion products on a DNA-sequencing gel.

1. Following autoradiography, align the NC sheet with the autoradiogram (*see* **Note 8**), mark with a sharp-tip pencil the precise position of radioactive signal(s) (this will permit determination of the relative molecular weight of the detected DNA-binding factor[s]), and cut the strip(s) corresponding to radioactive signal(s) with a clean, sharp razor blade.
2. Using forceps, uncover the strip from the cling film and immediately transfer it into a siliconized 0.5-mL Eppendorf tube containing 200 μL of SW binding/washing buffer.
3. Allow the strip to equilibrate for 15 min on ice. Meanwhile, thaw an aliquot of the 2.5 mg/mL DNase I stock solution.
4. Bring the tube to room temperature, place it in a scintillation vial, and determine c.p.m. of probe retained on the strip by Cerenkov counting.
5. Transfer an equal amount of radioactivity of free probe to a separate, siliconized 0.5-mL Eppendorf tube containing 200 μL of SW binding/washing buffer, and subject it to the same manipulation as its membrane-associated counterpart (**step 3**).
6. Prepare an appropriate dilution of DNase I in ice-cold distilled, deionized water. Mix thoroughly by inversion and gentle vortexing.
7. Add to both samples 200 μL of DNase I reaction buffer and mix by flicking. Let the tubes stand for 1 min at room temperature. It is not necessary to close the caps on the tubes until addition of the DNase-STOP solutions (**step 10**).
8. Add dilute DNase I to a final concentration of 25 ng/mL and quickly distribute by flicking. It is helpful -especially when footprinting several strips- to have all necessary items (pipets, buffers, timer, DNase I) in close proximity. The smoother this procedure goes, the better (and more reproducible!) your footprint(s) will be.
9. Incubate for 1 min at room temperature (*see* **Note 9**). As with solution footprinting protocols, the exposure time for the free DNA control reaction can be titrated to achieve the cutting intensity profile of the membrane-bound form. Nevertheless, we have found (employing probes of various

lengths) that the combination of digestion time and DNase I concentration generates sufficient cleavage for a good signal-to-noise ratio.

10. Following treatment, remove the strip (with forceps) from the tube and rapidly immerse it in 500 µL of ice-cold DNase-STOP solution A. Leave on ice for 2 min (*do not vortex!*). Terminate the control reaction (free probe) by adding 200 µL of ice-cold DNase-STOP solution B; vortex thoroughly, spin briefly, transfer into a siliconized 1.5-mL Eppendorf tube, and proceed directly to **step 15**.
11. Place the strip (using forceps) in a siliconized 0.5-mL Eppendorf tube containing 200 µL of probe elution buffer. Spin briefly in a microcentrifuge to submerge the entire strip.
12. Incubate the tube on a vortexing shaker at 37°C for 2 h.
13. Add 100 µL of distilled, deionized water and vortex the tube vigorously (2 min).
14. Microcentrifuge for 2 min to pellet the NC strip, and transfer the supernatant into a siliconized 1.5-mL Eppendorf tube. A Geiger counter can be used to monitor efficient recovery of radioactivity; typically, >90% of the membrane-bound probe is released by this process.
15. Extract once with phenol/chloroform/isoamyl alcohol (25:24:1; v/v) and once with chloroform/isoamyl alcohol (24:1; v/v). In both cases, mix by vortexing (for 10 s) and spin for 5 min.
16. Transfer the aqueous (top) layer to a fresh siliconized 1.5-mL Eppendorf tube. Precipitate the DNA with cold (–20°C) ethanol in the presence of 10 mM $MgCl_2$ (added from the 0.5 M stock solution; $MgCl_2$ aids in the recovery of small DNA fragments) and 10 µg carrier glycogen; mix by inversion, spin at 12,000 R× *g* for 15 min.
17. Remove and discard the supernatant with a drawn-out Pasteur pipet, being careful not to aspirate the DNA (the bottom of the tube can be held to a Geiger counter to check that the DNA pellet remains). Add 800 µL of ice-cold 80% ethanol and rinse the pellet by gently rolling the microcentrifuge tube. Respin for 3 min.
18. Remove and discard the supernatant using a drawn-out Pasteur pipet, taking extreme care not to disturb the tiny, whitish pellet or the area of the tube where the pellet should be located (the DNA pellet frequently adheres only loosely to the walls of the tube). Dry the pellet in a SpeedVac rotary concentrator. Do not allow the drying procedure to continue long past the point of dryness because the sample may be difficult to resuspend.

19. Dissolve the pellet in 6–8 μL of formamide loading buffer; pipet the loading buffer onto the upper, inside surface of tube and tap the tube to drop the droplet onto the DNA pellet. Vortex briefly at high speed for ~15 s and microcentrifuge for 30 s to collect all of the solution to the bottom of the tube.
20. Transfer to fresh siliconized 1.5-mL Eppendorf tubes and determine the total radioactivity recovered by Cerenkov counting each sample for 1 min in a scintillation counter.
21. Heat samples to 95°C for 3 min to denature DNA and immediately chill in wet ice.
22. Spin briefly to bring the liquid to the bottom of the tubes. Samples can be electrophoresed immediately or stored at –70°C for no more than 24 h after footprinting.
23. Load 1,500–2,000 cpm of each DNA digestion product (adjust volumes accordingly if necessary; it is essential that a consistent volume of sample be loaded on each lane) onto a pre-electrophoresed, 5–10% denaturing urea (sequencing) gel (*see* **Notes 9b** and **10a**). Electrophorese in 1R× TBE buffer at 60–70 W constant power (for a 34 R× 40 cm, 0.4 mm-thick gel) until the marker dye fronts have migrated the appropriate distance in order to visualize the DNA region of interest *(11)*. To determine the location of the transcription factor-binding site(s), the DNase I digests are electrophoresed alongside a Maxam-Gilbert G+ A sequencing ladder prepared from the end-labeled footprint probe *(14)*. See Chapter "Footprinting DNA–Protein Interactions in Native Polyacrylamide Gels by Chemical Nucleolytic Activity of 1,10-Phenanthroline-Copper" for a fast protocol for preparing such a ladder.
24. Disassemble gel apparatus, carefully lift off one glass plate, and soak the gel (still on the second glass plate) in fixing solution for 15 min (*see* **Note 10b**).
25. Drain briefly, overlay the gel with two sheets of 3-MM paper, and carefully peel it off the glass plate. Cover the gel surface with plasticwrap and dry under vacuum at 80°C for approximately 1 h.
26. Expose the dry gel to X-ray film overnight at –70°C with an intensifying screen (a piece of paper placed between the gel and the film will prevent spurious exposure of the film due to static electricity). Several different exposures may be required to obtain suitable band densities.
27. Compare lanes corresponding to free and protein-bound DNAs to identify the region(s) of protection; the region(s) of the DNA fragment that is bound by the factor appears as a blank stretch (footprint) in the otherwise continuous background of digestion products.

4. Notes

1. The diversity of properties characteristic of DNA-binding transcription factors imposes an empirical determination of the conditions under which the sample(s) is prepared for SDS-PAGE. In some cases the reducing agent (2-mercaptoethanol) should be omitted from the sample buffer and in others the SDS concentration should be lowered to 0.5%. Furthermore, some DNA-binding proteins may not withstand the sample boiling before loading on the gel.
2. The presence of methanol in the Western transfer buffer may cause a problem during the electrophoretic transfer of some bulky DNA-binding proteins (gel shrinkage reduces pore size).
3. a. The commercially available nonfat dried milk preparations from various suppliers contain large amounts of protein kinase/phosphatase activities; these activities can potentially interfere with binding of the DNA probe to transcription factors whose DNA-binding capacity is known -or suspected- to be subject to regulation by inducible phosphorylation/dephosphorylation events. If this is the case, substitute nonfat dried milk in the SW blocking/renaturation buffer for the recommended grade of lipid-free BSA; this particular grade contains only trace amounts of the aforementioned activities and should be preferred as blocking agent *(15, 16)*. Moreover, the use of lipid-free BSA results in an even background throughout and enhances the specific signal-to-noise ratio in the DNA-probing step *(17)*.

   b. By manipulating the conditions for renaturation of the membrane-immobilized proteins (e.g., by incorporating a cycle of protein denaturation and renaturation *(16)*), the method may be extended to the analysis of proteins resolved in two-dimensional gels *(18)*. This offers a powerful and convenient means for studying cell cycle/type/stimulus-dependent DNA-transcription factor interactions and their regulatory roles in gene activity.
4. a. Inclusion of PVA (a molecular crowding agent [volume excluder]) in the buffer decreases the amount of small ions/water available for hydration of any probe dissociated from the immobilized protein matrix and renders the aqueous environment unfavorable for the unbound DNA. Consequently, the effective concentration of the probe is increased and interactions with low binding constants are stabilized.

   b. It may be important in some cases to supplement this buffer with $ZnSO_4$ (Aldrich, Milwaukee, WI; final

concentration 10 μM) if the transcription factor(s) under study is known -or suspected- to contain a zinc-finger domain(s).

5. a. The DNA chosen for probing the protein blot can be a *cis*-acting regulatory (promoter/enhancer) DNA restriction fragment in the size range of 125–250 bp, with the putative transcription factor-binding sites located no less than 20–25 bp from the labeled end. This is to ensure that the region of DNA to be investigated for the presence of footprints is capable of being accurately resolved on a sequencing gel.

   b. A prerequisite for the subsequent DNase I footprinting analysis is the use of a DNA probe that has been labeled on only one strand of the DNA duplex. The labeling of only one strand of a promoter/enhancer restriction fragment can be achieved in a number of ways, such as using T4 polynucleotide kinase and [γ-^{32}P] ATP (5′-end labeling), the large (Klenow) fragment of *E. coli* DNA polymerase I and [α-^{32}P] dNTPs (3′-end labeling ["filling out"]), or the polymerase chain reaction (PCR) amplification *(19)*. Preparation of radiolabeled DNA employing any of these methodologies requires about 8 h. A combination of 5′- and 3′-end labeled DNA probes allows both strands to be analyzed side by side from the same end of the DNA duplex.

   c. For optimal sensitivity in the SW procedure, the probe should be of as high a specific activity as possible and highly purified. The latter can be assured by using a nucleic acid-specific, ion-exchange column, previously the Elutip™-d (Schleicher & Schuell) or the NACS prepac cartridge (BRL, Gaithersburg, MD) were used *(11)*. It is recommended not to store the pure, labeled DNA probe longer than 2–4 days, because the radiation creates nicks in the DNA which will appear as additional bands in the sequencing gel.

6. Synthetic alternate copolymers, such as poly[dI-dC]. poly[dI-dC] or poly[dA-dT]. poly[dA-dT], at similar final concentrations may be more suitable competitors for some DNA-binding transcription factors.

7. Longer washing times at room temperature are detrimental due to dissociation of the bound probe, but longer washing times with cold (4R°C) SW binding/washing buffer can reduce background without significant signal loss. To reduce possible low-specificity DNA-protein complexes, the final wash can be performed in cold SW binding/washing buffer with higher salt concentration (i.e., 100–200 mM KCl).

8. If prestained nonradioactive protein MW standards have been used, the edges of the plasticwrap should be marked with pieces of tape labeled with radioactive (or fluorescent) ink (let the ink dots dry completely before exposing to X-ray film!). These marks allow the autoradiogram to be aligned with the protein size markers on the NC sheet (**Subheading 3.2**, **step 1**), facilitating calculations on the relative molecular weight of the specific DNA-bound protein species (obtained radioactive signal[s]). If [^{14}C]-methylated protein MW markers have been used, their position will be apparent on the X-ray film without the need to use the radioactive (or fluorescent) ink procedure.
9. a. Although longer digestion times do not enhance background cutting -uncomplexed DNA is minimized- they may lead to substantial deviations from the required "single-hit kinetics" (i.e., on average, each DNA molecule is nicked at most once; this corresponds to nicking approximately 30–50% of the DNA molecules).
   b. Intense bands due to uncleaved, full-length probe should be visible at the top of each lane in the DNA-sequencing gel. This aids in determining whether single-hit kinetics are operative and whether equal amounts of total radioactivity are loaded in each lane.
10. a. If the DNA probe is relatively long (i.e., >175 bp) and multiple transcription factor-binding sites are to be resolved, a gradient or wedge-sequencing gel can be used.
    b. Wedge-shaped gels must be soaked in fixing solution, followed by 5% glycerol for 10 min prior to drying.

References

1. Galas, D. J. and Schmitz, A. (1978). DNase footprinting: A simple method for the detection of protein-DNA binding specificities. *Nucleic Acids Res.* **5**, 3157–3170.
2. Tullius, T. D. and Dombroski, B. A. (1986). Hydroxyl radical "footprinting": High- resolution information about DNA-protein contacts and application to λ repressor and Cro protein. *Proc. Natl. Acad. Sci. USA* **83**, 5469–5473.
3. Kuwabara, M. D. and Sigman, D. S. (1987). Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: *Escherichia coli* RNA polymerase-*lac* promoter complexes. *Biochemistry* **26**, 7234–7238.
4. Johnsrud, L. (1978). Contacts between *Escherichia coli* RNA polymerase and a *lac* operon promoter. *Proc. Natl. Acad. Sci. USA* **75**, 5314–5318.
5. Garner, M. M. and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**, 3047–3060.
6. Fried, M. and Crothers, D. M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**, 6505–6525.
7. Miskimins, W. K., Roberts, M. P., McClelland, A., and Ruddle, F. H. (1985). Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferrin receptor gene. *Proc. Natl. Acad. Sci. USA* **82**, 6741–6744.

8. Smith, S. E. and Papavassiliou, A. G. (1992). A coupled Southwestern – DNase I footprinting assay. *Nucleic Acids Res.* **20**, 5239–5240.
9. Polycarpou-Schwarz, M. and Papavassiliou, A. G. (1993). Probing of DNA-protein complexes immobilized on protein-blotting membranes by the chemical nuclease 1,10-phenanthroline(OP)-cuprous ion. *Methods Mol. Cell. Biol.* **4**, 22–26.
10. Polycarpou-Schwarz, M. and Papavassiliou, A. G. (1993). Distinguishing specific from non-specific complexes on Southwestern blots by a rapid DMS protection assay. *Nucleic Acids Res.* **21**, 2531–2532.
11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
12. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
13. Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
14. Maxam, A. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
15. Papavassiliou, A. G., Bohmann, K., and Bohmann, D. (1992). Determining the effect of inducible protein phosphorylation on the DNA-binding activity of transcription factors. *Anal. Biochem.* **203**, 302–309.
16. Polycarpou-Schwarz, M. and Papavassiliou, A. G. (1995). Protein-DNA interactions revealed by the Southwestern blotting procedure. *Methods Mol. Cell. Biol.* **5**, 152–161.
17. Papavassiliou, A. G. and Bohmann, D. (1992). Optimization of the signal-to-noise ratio in south-western assays by using lipid-free BSA as blocking reagent. *Nucleic Acids Res.* **20**, 4365–4366.
18. Moreland, R. B., Montross, L., and Garcea, R. L. (1991). Characterization of the DNA-binding properties of the polyomavirus capsid protein VP1. *J. Virol.* **65**, 1168–1176.
19. Lakin, N. D. (1993). Determination of DNA sequences that bind transcription factors by DNA footprinting, in *Transcription Factors: A Practical Approach* (Latchman, D.S., ed.), IRL Press, Oxford, pp. 27–47.
20. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P., and Karin, M. (1987). Phorbol ester-inducible genes contain a common *cis* element recognized by a TPA-modulated *trans*-acting factor. *Cell* **49**, 729–739.

Chapter 15

Use of a Reporter Gene Assay in Yeast for Genetic Analysis of DNA–Protein Interactions

David R. Setzer, Deborah B. Schulman, Cathy V. Gunther, and Michael J. Bumbulis

Summary

We describe methods for the genetic analysis of a DNA–protein interaction from any species. The DNA-binding domain of the protein of interest is expressed in yeast cells as a fusion with a known transcriptional activation domain, and the target binding site is used as an artificial upstream activation sequence (UAS) in an engineered promoter driving expression of a reporter gene, such as β-galactosidase. Expression of the reporter gene is dependent upon specific, high-affinity interaction between the DNA-binding domain of the artificial activator and the synthetic UAS. Error-prone PCR is used to introduce mutations into either member of this interacting pair, and homologous recombination is used to return the mutagenized sequences to their proper sequence contexts in vivo. Altered expression of the reporter gene is then used as a screen or selection for mutations conferring the desired phenotype, such as reductions or increases in the stability of the DNA–protein complex. Following identification of the relevant mutations, the mutant protein or binding site can be subjected to further analyses to confirm the expected biochemical basis of the selected phenotype. This approach has been used extensively in the analysis of the TFIIIA-5S rRNA gene interaction from both *Xenopus laevis* and *Schizosaccharomyces pombe*.

Key words: DNA binding, Transcription factors, Error-prone PCR, Homologous recombination, Genetic analysis, TFIIIA, 5S RNA, One hybrid, Yeast.

1. Introduction

Understanding the underlying structural and physicochemical basis for the recognition of specific DNA sequences by regulatory proteins is a goal of modern biochemical genetics. A general method for the rapid identification of mutant molecules altered

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_15

in the affinity and/or specificity of such interactions could be a powerful tool in achieving this goal. Conventional genetic approaches for obtaining and analyzing interesting mutant forms of specific DNA-binding proteins are often infeasible because of the genetic intractability of the species being studied or as a result of difficulties in identifying relevant and specific phenotypes associated with alterations in the interaction under investigation. High-resolution genetic analysis of DNA–protein interactions is particularly problematic in multicellular eukaryotes. We have devised an approach which makes use of the modularity in structure and function of eukaryotic transcription factors *(1)*, the power of the polymerase chain reaction (PCR) to generate specific DNA fragments with defined levels of mutagenesis in vitro *(2, 3)*, and the recombinogenic potential of *S. cerevisiae (2, 4)* to carry out a high-resolution genetic analysis of the sequence-specific DNA-binding properties of *Xenopus* Transcription Factor IIIA (TFIIIA) *(5)*. We have also applied similar methods to the analysis of TFIIIA from *Schizosaccharomyces pombe*. Our approach is a variation on yeast "one-hybrid" methods and should applicable to the study of many DNA–protein interactions.

1.1. Outline of the Approach

In its simplest form, the method we describe here involves the construction and introduction of two plasmids into an appropriate yeast strain (**Fig. 1**). The first, called the reporter plasmid, places the *E. coli* β-galactosidase gene under control of the core promoter of the *S. cerevisiae* iso-1-cytochrome c (*CYC1*) gene, but the upstream activator sequence (UAS) normally required for expression from the *CYC1* promoter has been deleted and replaced by a DNA fragment containing the cognate recognition site for the DNA-binding protein to be analyzed. The low levels of β-galactosidase that are expressed from this plasmid in vivo result in strains that are white on X-gal indicator plates. In the second plasmid, the expression construct, the sequence encoding the VP16 activation domain from herpes simplex domain is fused in-frame to a sequence encoding the DNA-binding domain of the protein of interest. Expression of the DNA-binding domain-VP16 fusion protein is under control of the constitutive glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter of *S. cerevisiae*. The reporter and expression plasmids carry different selectable markers (*URA3* and *TRP1*, for example) so that both can be selected and maintained in an appropriate yeast strain (ura3$^-$trp1$^-$, for example). When both plasmids are introduced into a single yeast cell, and if the DNA-binding domain of the protein of interest binds with sufficiently high affinity and specificity to its recognition site in the reporter construct, display of the VP16 activation domain in the vicinity of the core *CYC1* promoter will activate transcription of the β-galactosidase reporter gene. On X-gal indicator plates, such a strain will be

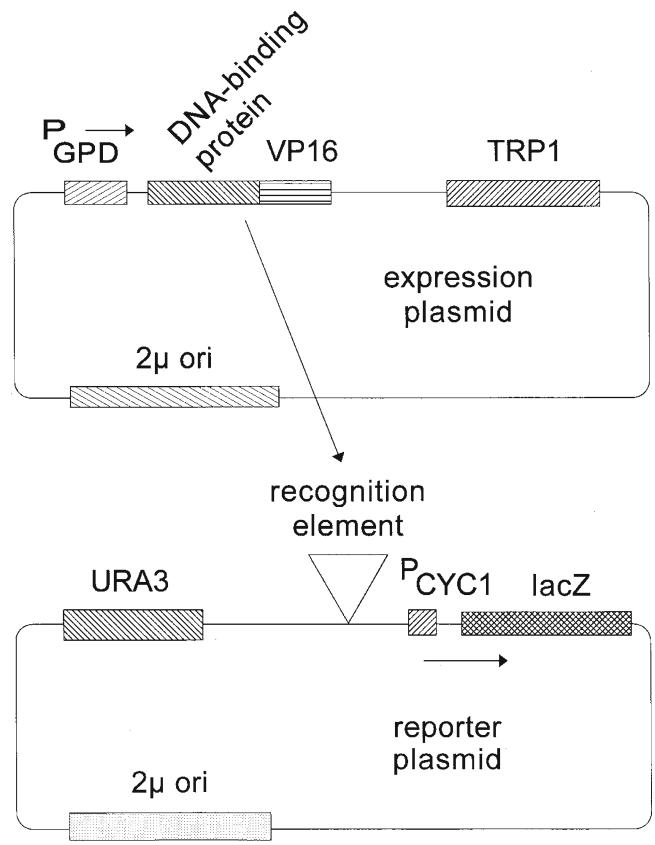


Fig. 1. Schematic representation of generic expression and reporter plasmids derived from pG1 and pΔSS, respectively. Only the plasmid components functional in yeast cells are shown; the plasmids also contain colE1 origins and β-lactamase genes for replication and selection in *E. coli*.

blue. Thus, this blue phenotype can be used as a marker for high-affinity interaction of the DNA-binding domain of interest with its recognition sequence. Mutations in either the DNA-binding protein or the DNA sequence to which it binds may adversely affect binding, resulting in white or light blue colonies, or may increase the stability of binding, resulting in dark blue colonies (*see* **Notes 1–5**).

We generate randomly mutated sequences encoding either the protein of interest or its cognate recognition site using error-prone PCR in vitro. Introduction of these mutated sequences into their appropriate contexts in either the expression or reporter plasmids is achieved with technical ease via homologous recombination in vivo following transformation (**Fig. 2**). Unusually long oligonucleotide primers are used for error-prone PCR (about 60–70 nucleotides, but *see* **Note 7**). The 3′ portion of these primers

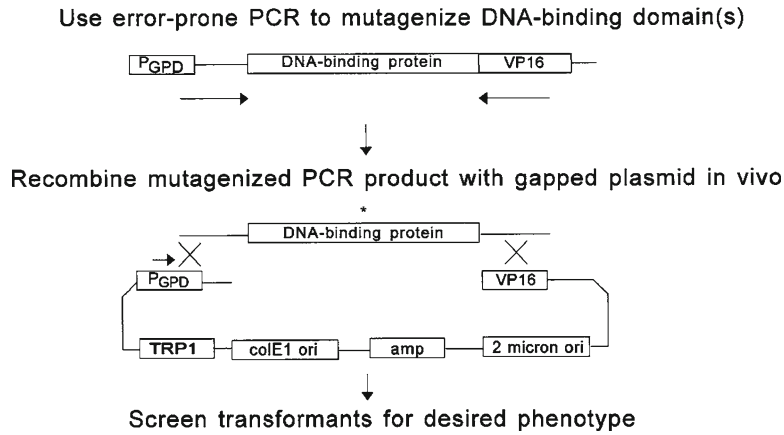


Fig. 2. Error-prone PCR in vitro and homologous recombination in vivo to mutagenize the DNA-binding protein of interest and introduce the mutagenized fragment into the expression vector. The *asterisk* represents a mutation introduced during error-prone PCR. In the example shown here, the substrate used for error-prone PCR is the expression plasmid itself, derived in this case from pG1, and the entire DNA-binding protein is subjected to mutagenesis. It is also possible to target only a portion of the DNA-binding protein for mutagenesis or to use a different plasmid as a substrate for PCR, provided the expression plasmid and amplification primers are appropriately designed.

anneals to a template plasmid at sites flanking the sequence to be mutagenized. The primers also contain 5′ sequences identical to those that flank the ends of a linear version of the plasmid construct into which the mutated sequences are to be introduced. Error-prone PCR is used to synthesize a population of mutant DNA fragments containing the sequence of interest flanked by the sequences defined by the long amplification primers. In parallel, the plasmid into which the mutagenized PCR product is to be inserted is linearized or gapped by digestion at one or two sites, respectively, such that the unique ends of the PCR product correspond to sequences near the ends of the linear or gapped plasmid. When cotransformed into competent yeast cells, the linear plasmid and mutant PCR products undergo homologous recombination in vivo to produce a circular plasmid product in which the mutagenized fragment is integrated into the target plasmid at or between the restriction sites used in linearization/gapping. Successful recombination events can be clonally selected on appropriate media using the marker gene (*URA3* or *TRP1*, for example) on the target plasmid. If the yeast strain used for transformation already contains the second plasmid component of the system, then resulting colonies containing both reporter and expression plasmids can be selected and screened subsequently by replica plating on indicator plates to identify mutants resulting in altered phenotypes (white, light blue, or dark blue). After further

tests to ensure authenticity of the mutant phenotype and that the mutant protein or DNA sequence is likely to be of interest, the mutant plasmid is recovered in *E. coli* and the mutation identified by DNA sequence analysis. If desired, the mutant DNA–protein interaction can be subjected to detailed biochemical or further genetic analysis.

2. Materials

2.1. Initial Design and Testing of the System

1. *S. cerevisiae* strain BJ2168 *(6)* or other haploid strain with appropriate genotype (stable mutant alleles of genes used as selectable markers in expression and reporter plasmids) (*see* **Note 1**).
2. pΔSS *(7)* or other plasmid to be used as parent for reporter plasmid construction (*see* **Note 2**).
3. pG1 *(7)* or other plasmid to be used as parent for expression plasmid construction (*see* **Note 3)**.
4. pSJT-1193-CRF1 *(8)* or other source of DNA encoding the VP16 activation domain (*see* **Note 4**).
5. Source of DNA encoding the DNA-binding protein or DNA-binding domain of interest.
6. Source of DNA that includes the recognition sequence for the DNA-binding protein or DNA-binding domain of interest.
7. Complete medium (C) agar plates lacking appropriate nutrients to permit selection of yeast strains containing reporter and expression plasmids. These will include C -uracil, C -tryptophan, and C -uracil -tryptophan for systems making use of derivatives of pΔSS and pG1. Procedures for preparation of liquid C medium and C agar are described by Rose et al. *(9)*. Our specific procedures are (*see* **Note 6**):
   a. Dissolve the following in 1 L water: 20 g dextrose, 20 g Bacto-agar, 1.7 g yeast nitrogen base without amino acids and without ammonium sulfate, 5 g ammonium sulfate, and 0.5 g amino acid mixture (*see* **item 7b**). Autoclave to sterilize and use to pour approximately forty 100-mm plates.
   b. The amino acid mixture used to prepare C agar plates contains 0.2 g arginine, 0.2 g histidine, 0.5 g lysine, 0.4 g methionine, 0.2 g phenylalanine, 0.4 g tryptophan, 2.0 g threonine, 0.4 g tyrosine, 0.5 g serine, 0.2 g adenine, and 0.1 g uracil. For selective plates, the appropriate combination of nutrients (uracil and tryptophan, e.g., to select for plasmids containing *URA3* and *TRP1* markers) should be omitted from the mixture.

8. SSX agar plates lacking appropriate nutrients as described in **item 7**, but also containing 40 μg/mL X-gal, prepared as follows:
   a. Dissolve the following in 900 mL water: 1.7 g yeast nitrogen base without amino acids and without ammonium sulfate, 5 g ammonium sulfate, 20 g dextrose, 14 g agar, 0.5 g appropriate amino acid mixture (*see* **item 7b**). Autoclave to sterilize.
   b. Cool to 48° and add aseptically: 1 mL 40 mg/mL X-gal prepared in *N,N*-dimethylformamide, 100 mL 10× phosphate buffer (*see* **item 8c**).
   c. 10× phosphate buffer is prepared by mixing in 1 L water the following: 136.1 g KH_2PO_4 (1 M), 19.8 g $(NH_4)_2SO_4$ (0.15 M), and 42.1 g KOH (0.75 N). Adjust the pH to 7.0 and autoclave to sterilize.
9. Standard reagents and methods for subcloning DNA fragments into plasmids.

2.2. Error-Prone PCR

1. Plasmid(s) or other source(s) of DNA containing the sequence encoding the protein of interest and/or the DNA sequence recognized by the protein of interest.
2. Oligonucleotide primers that contain, at their 5′ ends, approximately 50 nucleotides of sequence identity to the site immediately adjacent to the end of the linear plasmid into which the PCR product is to be inserted. The 3′ 15–20 nucleotides of these primers should have sequence identity with the parts of the template plasmid that define the DNA sequence to be amplified and mutagenized (*see* **Fig. 2** and **Note 7**).
3. *Taq* DNA polymerase.
4. Stock solutions of 100 mM $MgCl_2$ and 100 mM $MnCl_2$.
5. 10× stock solution of *Taq* PCR buffer, lacking $MgCl_2$: 100 mM Tris-HCl of pH 9.0, 500 mM KCl, 1% Triton X-100 (Promega).
6. Individual stock solutions of dATP, dGTP, dCTP, and dTTP, each at a concentration of 10 mM. In addition, individual stock solutions of the same, each at a concentration of 2 mM.
7. Thermal cycler.

2.3. Yeast Transformation and Homologous Recombination

1. *S. cerevisiae* strain BJ2168 or other appropriate strain (*see* **Note 1**).
2. If the DNA-binding protein is to be mutagenized, BJ2168 containing the reporter plasmid, and BJ2168 containing the parent of the reporter plasmid (pΔSS, for example).
3. If the DNA recognition site is to be mutagenized, BJ2168 containing the expression plasmid and BJ2168 containing the parent of the expression plasmid (pG1, for example).

4. Linearized or gapped plasmid to be used as the target for integration of the PCR-generated DNA fragment.
5. Crude or purified product of the error-prone PCR.
6. C agar plates lacking the relevant nutrients for selection and maintenance of both reporter and expression plasmids (uracil and tryptophan for pΔSS- and pG1-derived plasmids).
7. Sterile stock solution of 100 mM lithium acetate, 1 mM EDTA, 10 mM Tris–Cl of pH 8.0.
8. Sonicated salmon sperm DNA of about 10,000 bp average length, denatured by heating to 100°C for 5 min at a concentration of about 5–10 mg/mL. Commercially available salmon sperm DNA should be extracted with phenol:chloroform and precipitated prior to use.
9. 60% (w/v) Polyethylene glycol (PEG) (average molecular weight of 3,350 Da). To prepare this solution, autoclave 3 g solid PEG in a sealable tube. The PEG will melt during sterilization and then solidify at room temperature. Many sterile PEG aliquots can be prepared simultaneously. On the day of use, add 2.7 mL sterile solution from **step 7** to one of these aliquots, heat to 65°, and mix vigorously to dissolve (final concentration of PEG is 60% in a volume of about 5 mL; adjust the volume to a final of 5 mL by adding additional solution from **step 7**, if necessary). This quantity of solution is sufficient for approximately 25 transformations.
10. YEPD medium: per liter of water, add 10 g Bacto-yeast extract, 20 g Bacto-peptone, and 20 g glucose.

2.4. Screening for Mutants

1. SSX agar plates lacking both uracil and tryptophan, and containing 40 μg/mL X-gal.
2. Liquid C-trp or C-ura medium for selection of only one of the two plasmids in BJ2168.
3. C agar plates lacking uracil and tryptophan individually, as well as plates lacking both.
4. Standard *E. coli* strain for plasmid transformation, propagation, and isolation, along with reagents for distinguishing reporter and expression plasmids by restriction endonuclease analysis.
5. 2% (w/v) sodium dodecyl sulfate (SDS).
6. Acid-washed glass beads (0.45 mm diameter) from Sigma prepared by washing overnight in 3N HCl and then rinsing repeatedly in water. Prewashed beads are also available (Sigma-Aldrich G8772).
7. Reagents for protein concentration determination using the BCA method (Pierce) *(18)*.

8. If possible, antibodies to the DNA-binding protein of interest and/or the activation domain used in the construction of the expression plasmid; reagents for Western blotting.

2.5. Analysis of Mutants

1. Z buffer for determination of β-galactosidase activity: 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 40 mM β-mercaptoethanol, with a final pH of 7.0.
2. Other reagents for determination of β-galactosidase activity: 0.2% sarkosyl in Z buffer; 4 mg/mL O-nitrophenol-β-D-galactoside (ONPG) prepared in Z buffer; 1 M Na_2CO_3.
3. Reagents for DNA sequence determination.

3. Methods

3.1. Initial Design and Testing of the System

Details of the construction of appropriate reporter and expression constructs will depend upon the specific features of the plasmids and clones to be used. It is therefore impossible to describe a step-by-step protocol for use in every case, but standard recombinant DNA methods should suffice for preparation of the desired plasmids. We will briefly outline the steps necessary for construction and testing of reporter and expression constructs prepared in pΔSS and pG1, respectively.

3.1.1. Construction of the Reporter Plasmid

A DNA fragment containing one or more copies of the DNA sequence recognized by the protein of interest must be subcloned upstream of the *CYC1* core promoter in pΔSS. The only unique restriction site in pΔSS that is suitable for insertion of such a fragment is an Xho I site. A DNA fragment with Xho I-compatible ends and containing one or more copies of the relevant DNA sequence should be subcloned into the Xho I site of pΔSS. Most typically, this fragment would be either a restriction fragment from another plasmid or a PCR product digested to produce Xho I-compatible ends. It is possible, and probably desirable, to obtain pΔSS derivatives with single or multiple inserts of the DNA sequence of interest in either orientation. The number and orientation of insert fragments must be diagnosed by some means, typically including restriction endonuclease mapping using enzymes that cut asymmetrically within the insert fragment to determine orientation, PCR with primers flanking the insert site to determine number of inserts, or DNA sequence analysis to determine orientation, or the number of inserts if the insert fragment is not too long. An alternative to subcloning the insert fragment into the Xho I site of pΔSS is to use homologous recombination as described in **Subheading 3.3** to integrate a PCR-generated

DNA fragment into Xho I-digested pΔSS. In this case, the PCR fragment should be produced under high-fidelity conditions; even so, the inserted fragment should be sequenced in the resulting plasmid to ensure that no mutations were introduced during amplification. If homologous recombination is used to generate the reporter plasmid, judicious choice of sequences in the long primers used for PCR can be used to regenerate either or both of the Xho I sites at the end of the insert, or even to introduce novel restriction endonuclease recognition elements. This may facilitate the introduction of multiple copies of the DNA-binding site into pΔSS.

3.1.2. Construction of the Expression Plasmid

One or more DNA fragments encoding an in-frame fusion of the DNA-binding protein of interest and a transcriptional activation domain must be introduced into pG1 downstream of the GPD promoter. The unique Sal I site in pG1 is probably the most convenient site for doing this. Fusion of the DNA-binding protein of interest and the VP16 activation domain can be done directly in pG1 or, perhaps more conveniently, in a smaller, simpler plasmid vector and then subcloned as a unit into pG1. Subsequent mutagenesis (**Subheading 3.2**) can be more directly targeted to the DNA-binding protein rather than to the activation domain if a unique restriction site (more precisely, one that does not occur elsewhere in the plasmid outside of the sequence encoding the DNA-binding protein) can be engineered at the junction of the DNA-binding protein and the activation domain. It is also important to note that some DNA-binding proteins, particularly transcriptional activator proteins acting through the RNA polymerase II core machinery, may contain endogenous transcriptional activation domains that will function in *S. cerevisiae*; in that event, transcriptional activity in the absence of the VP16 activation domain may be observed. The existence of an intrinsic activation domain in the protein of interest might obviate the need to prepare a fusion construct, but one must be careful in the subsequent analysis to distinguish mutations affecting DNA binding from those affecting transcriptional activation directly. The VP16 activation domain coding sequence followed by a polyadenylation signal from the herpes simplex thymidine kinase gene can be excised on a Kpn I-Hind III fragment of approximately 760 bp from the plasmid pSJT-1193-CRF1 *(8)*. At the Kpn I cleavage site, the reading frame for fusion to VP16 is XXG-GTA-CCX, but other plasmids in which the VP16 reading frame is shifted relative to the Kpn I cleavage site have also been constructed *(8)*. The Kpn I-Hind III fragment from this family of constructs is suitable for preparing a fusion protein in which the VP16 domain is at the C-terminus. Depending upon what is known about the polarity of DNA binding by the protein of interest, this may or may not be desirable. Construction of N-terminal fusions may

be preferable in some cases (*see* **Note 8**); these can be made by taking advantage of one of a number of vectors intended for the construction of libraries for use in two-hybrid screens, e.g., pACT-II, pGAD-GH, and pB42-AD from Clontech. One must be cautious in the choice of vector, however (*see* **Note 3**). In the case of *Xenopus* TFIIIA binding to the *Xenopus* 5S rRNA gene, some of these vectors (including pGAD10) result in very low levels of protein expression and no detectable transcription activation. pGAD10 makes use of the alcohol deydrogenase (ADH) promoter to drive protein expression, and we have experienced similar difficulties with the ADH promoter in other, but not all, plasmid contexts. Also, see **Note 4** concerning choice of activation domains vis-a-vis sensitivity of the genetic assay.

3.1.3. In Vivo Assay of the Reporter and Expression Constructs

For each reporter plasmid constructed and the parent vector (pΔSS, as a control), two strains derived from BJ2168 should be prepared, one containing the expression plasmid in addition to the reporter, and another containing the parent plasmid from which the expression plasmid was derived (pG1), in combination with the reporter. The different selectable markers on these two plasmids (URA3 and TRP1) allow their simultaneous maintenance in BJ2168 ($ura3^{-}trp1^{-}$) by selecting for growth in medium lacking uracil and tryptophan. The requisite strains should be constructed by sequentially transforming BJ2168 with the reporter and expression plasmids. Methods for transformation of BJ2168 (or other yeast strains) are described in detail in **Subheading 3.3**. One need only adjust the protocol to reflect the nutritional requirements of the strain being transformed and the selectable marker on the plasmid being introduced. Thus, the doubly transformed strain would be selected on C -ura -trp plates.

Colonies of strains containing both reporter and expression plasmids can be replica plated, spotted, or streaked onto SSX -ura -trp plates containing 40 μg/mL X-gal. Colony color is assessed at an empirically determined time after robust colony growth has occurred. For analysis of the *Xenopus* TFIIIA-5S rRNA gene interaction, this was done typically after 2–3 days of growth at 30° and an additional 2–3 days at room temperature. For the system to be exploited successfully, one must be able to distinguish reproducibly the color of strains containing both the expression and reporter plasmids from that of all the other control strains (lacking either expression of the fusion protein containing the DNA-binding domain(s) of interest or the cognate recognition site in the reporter construct, or both). If this is not the case, it may be possible to correct the problem by manipulation of parameters as described in **Note 5**. Of course, it is also possible that the particular interaction being studied will not be amenable to analysis with this method; among other reasons, this could result from a low-affinity/specificity interaction, or from the

existence of endogenous yeast factors that interact with the binding site introduced into the reporter plasmid, resulting in high levels of transcriptional activity in the absence of the interaction being targeted for study.

3.2. Error-Prone PCR

The DNA-binding protein or its recognition site can be subjected to random mutagenesis using error-prone PCR. In the following protocol, we assume that the DNA-binding protein is targeted for mutagenesis, but the procedure can be adapted readily for mutagenesis of the recognition site.

1. Set up a 50-μL polymerase chain reaction mixture containing 10–50 ng of plasmid DNA containing the sequence encoding the region to be mutagenized. This can be the expression plasmid itself (*see* **Note 9**) or another plasmid containing the sequence of interest. In addition, add 5 μL 10× PCR buffer lacking $MgCl_2$, long amplification primers (*see* **Subheading 2.2**, **Fig. 2**, and **Note 7**) to a final concentration of 0.3-μM each, three deoxynucleoside triphosphates to a final concentration of 1-mM each, the fourth deoxynucleoside triphosphate to a final concentration of 0.2 mM, $MgCl_2$ to a final concentration of 3 mM, $MnCl_2$ to a final concentration of 0.05 mM, and 1 U *Taq* DNA polymerase (*see* **Note 10**).
2. Amplify using a thermal cycler for 25 cycles, with each cycle being 94° for 1 min, 42° for 2 min, and 72° for 1 min. After 25 cycles, use a final extension step of 72° for 7 min (*see* **Note 10**).
3. Use the crude or purified PCR product in a yeast transformation with linearized/gapped target plasmid as described later.

3.3. Yeast Transformation and Homologous Recombination

1. Prepare a stock of linearized or gapped target plasmid by digesting to completion with one or two restriction endonucleases that result in ends corresponding to the site at which integration of the mutagenized DNA fragment is to occur. As an example, with an expression plasmid derived from pG1, this might be a double digest with Sal I and an enzyme recognizing the fusion junction between the DNA-binding protein and the VP16 activation domain (*see* **Note 11**). A stock of this gapped plasmid can be prepared in 10 mM Tris–Cl of pH 8.0, 1 mM EDTA at a concentration of 10–100 ng/μL and stored at –20°.
2. Prepare competent yeast cells:
   a. Grow a 50-mL culture of the yeast strain containing the reporter plasmid (if mutations in the DNA-binding protein are to be analyzed) or expression plasmid (if mutations in the DNA recognition site are to be analyzed) overnight at 30° until the OD_{600} = 0.5–1.0.

b. Pellet the yeast cells by spinning for 5 min in a clinical centrifuge (1,000 × *g*) at room temperature.

c. Resuspend the cells in 10 mL 100 mM lithium acetate, 1 mM EDTA, 10 mM Tris–Cl of pH 8.0. Pellet again as in **step 2b**.

d. Resuspend again in 10 mL 100 mM lithium acetate, 1 mM EDTA, 10 mM Tris–Cl of pH 8.0.

e. Incubate at 30° for 30 min and pellet again as in **step 2b**.

f. Resuspend in 500 μL 100 mM lithium acetate, 1 mM EDTA, 10 mM Tris–Cl of pH 8.0. Place on ice until used for transformation.

3. Combine 50 μL competent yeast cells with 40 μL of the error-prone PCR mixture (**Subheading 3.2**) or purified PCR product (*see* **Notes 12** and **13**), 25 μg sonicated, denatured salmon sperm DNA, and 100 ng gapped plasmid (*see* **Notes 11** and **12**). The final volume should be 100 μL.
4. Incubate at 30° for 30 min without agitation.
5. Add 0.2 mL 60% PEG (3,350 mol wt) in 100 mM lithium acetate, 1 mM EDTA, 10 mM Tris–Cl of pH 8.0 and incubate further at 30° for 60 min.
6. Incubate at 42° for 15 min.
7. Centrifuge for 1 min at room temperature in a microcentrifuge at 14,000 rpm (16,000 × *g*).
8. Remove the supernatant and resuspend in 100 μL sterile water.
9. Plate half of the resuspended transformed cells on each of two C agar plates lacking the appropriate nutrients for selection of both expression and reporter plasmids (e.g., C -ura -trp).
10. Incubate at 30° for 2–3 days, until robust colony growth is obtained.

3.4. Screening for Mutants

1. Initial screen.

a. Replica plate the colonies obtained in **step 10** of **Subheading 3.3** onto selective SSX agar plates (e.g., SSX -ura -trp) containing 40 μg/mL X-gal.

b. Place replica plates at 30° for 2–3 days and then at room temperature for an additional 2–3 days.

c. Identify potentially interesting mutants by color. These could include light-blue, white, or dark-blue colonies; in each case, the color should differ from that exhibited by colonies containing wild-type versions of both the expression and reporter plasmids.

2. Pick candidates for further study from the initial selective plates and respot on selective C agar plates. Grow for about 2 days at 30°. Replica plate onto indicator media and score colony color again after growth as in **step 1**. If the phenotype is consistent with that seen initially, proceed to **step 3**.
3. Enrich for cells containing only the mutagenized plasmid.
   a. Pick one or more colonies with an interesting phenotype from selective C agar plates without X-gal indicator. Grow in 2 mL liquid C -trp medium (if the expression plasmid has been mutagenized; grow in C -ura medium if the reporter plasmid has been mutagenized) overnight at 30°.
   b. Prepare a 1:200 dilution of this overnight culture in liquid medium and plate 10 μL of this diluted culture on a C -trp plate (assuming throughout that the expression plasmid has been mutagenized).
   c. After 3 days of growth at 30°, replica plate onto C -ura -trp as well as onto C -trp.
   d. Grow overnight at 30° and compare the colonies that grew on the two replica plates. Choose several that grew on C -trp but not on C -ura -trp. These are likely to have lost the reporter plasmid but not the expression plasmid. Use these to inoculate 3-mL liquid cultures of C -trp and grow overnight at 30°.
4. Isolate the mutated plasmid.
   a. From a 3-mL overnight culture, pellet yeast cells in a 1.5-mL microcentrifuge tube. Do two sequential spins with 1.5-mL each for 1-min each.
   b. Resuspend the cell pellet in 200 μL breaking buffer: 10 mM Tris–HCl of pH 8.0, 100 mM NaCl, 1 mM EDTA, 2% Triton X-100, 1% SDS.
   c. Add 0.3 g glass beads and 200 μL phenol: chloroform: isoamyl alcohol (25:24:1).
   d. Mix vigorously for ~10–20 min.
   e. Spin 5 min in a microcentrifuge.
   f. Remove aqueous phase (upper) and extract it again with 200 μL chloroform.
   g. Spin 5 min in microcentrifuge.
   h. Recover aqueous phase (upper) and allow to gas off in a fume hood for 10 min or more.
   i. Precipitate by adding 1 μL of 20 μg/μL glycogen, NaCl to a final concentration of 0.2 M, and 2 volumes of ethanol. Cool in dry ice. Pellet by spinning 10 min at top speed in a microcentrifuge.

j. Rinse the pellet with 70% ethanol. Spin again in a microcentrifuge for 5 min to pellet.

k. Dry under vacuum.

l. Redissolve in 50–100 µL 10 mM Tris of pH 8.0. The pellet may be relatively large, colored yellow/brown, and somewhat difficult to dissolve.

m. Store at –20°.

5. Use ~5 µL of the isolated plasmid to transform *E. coli*. Pick individual colonies and perform standard plasmid minipreps. Analyze the plasmids thus obtained using appropriate restriction digests to ensure that the plasmid isolated is the mutagenized plasmid (e.g., the expression plasmid) and not the other plasmid component of the system (e.g., the reporter plasmid).

6. Analyze isolated plasmids by retransformation.

   a. Use the purified plasmid to retransform the appropriate yeast strain (already containing the reporter plasmid) and obtain colonies on selective C -ura -trp plates of cells with both reporter and expression plasmids.

   b. Analyze transformants by replica plating onto indicator plates as described in **step 1**. If the phenotype observed is consistent with that seen initially, then one can conclude that the phenotype is plasmid dependent. If the expression plasmid was mutagenized and it is the DNA-binding protein that is under analysis, continue with **step 7**. If, instead, the reporter plasmid was mutagenized and is being studied, one can move directly to **Subheading 3.5**.

7. Western blot screen to ensure the expressed fusion protein is full length and expressed at normal levels.

   a. Inoculate 30 mL of C -ura -trp medium with a mutant strain and grow overnight at 30°.

   b. Pellet the cells in a clinical centrifuge at (1,000 × *g*) for 5 min. Resuspend the pellet in 1 mL of 2% SDS.

   c. Add 1 g glass beads and vortex vigorously for 15 min.

   d. Transfer the mixture to another tube and spin in a microcentrifuge at top speed (16,000 × *g*) for 15 s. Transfer the supernatant to another tube.

   e. Determine protein concentration in the extract using the BCA method (reagents from Pierce) *(18)* with bovine serum albumin as the standard.

   f. Using standard Western blot methods, analyze 20 µg total protein from each mutant and compare to equivalent amounts of protein from a control strain expressing

wild-type DNA-binding protein as well as a second, negative control strain containing the parent plasmid from which the expression construct was derived, but which lacks coding sequence for the DNA-binding protein. The primary antibody used in the Western blot should be specific for the DNA-binding protein of interest; if such antibodies are unavailable, it may be possible to obtain, from commercial sources, antibodies to the activation domain used in the fusion. The point of the Western blot exercise is to exclude from further consideration those mutants whose steady-state level of expression is different from that of wild-type protein or which are truncated as a result of chain-termination or frame-shift mutations. Some truncation mutants may be interesting, but recall that the activation domain will be completely absent from truncated proteins if the activation domain was fused to the C-terminus of the DNA-binding protein. We have found that a substantial fraction of potential loss-of-function mutants (white phenotype) of TFIIIA are truncation mutants, and a Western blot screen was important in removing them from further, more laborious analysis *(5)*.

3.5. Analysis of Mutants

3.5.1. Quantitative Determination of β-Galactosidase Activity

We have found that relative DNA-binding affinities (but *see* **Note 5**) of various mutant forms of TFIIIA can be predicted with reasonable precision from measurements of β-galactosidase activity in vivo in yeast reporter strains expressing the mutant protein fused to the activation domain of VP16. Whether this will prove to be generally true remains to be seen, but β-galactosidase activities can be readily measured using standard methods and may prove informative.

1. Inoculate liquid C -ura -trp medium with the yeast strain to be analyzed. In addition, prepare a similar culture of a control strain containing the same reporter plasmid but with the parent vector (pG1, for example) of the expression construct, rather than the expression construct itself. This control will permit determination of β-galactosidase activity in the absence of the DNA–protein interaction of interest. Grow the cultures overnight at 30° until the OD_{600} = 0.3–1.0.
2. Pellet cells from 4 mL culture by spinning at 1,000 × *g* for 5 min. Resuspend in 4 mL Z buffer. Pellet again, and resuspend in 4 mL fresh Z buffer.
3. Remove a 1-mL aliquot and determine the optical density at 600 nm.
4. In a 1.5-mL microcentrifuge tube, add 200 μL of the cell suspension to 400 μL Z buffer containing 0.2% sarkosyl. As a

"blank," add 200 μL Z buffer to 400 μL Z buffer containing 0.2% sarkosyl.

5. Incubate at 30° for 30 min to equilibrate temperature before adding substrate.
6. Add 150 μL ONPG at a concentration of 4 mg/mL in Z buffer. Mix briefly and note the start time. Incubate at 30° until a yellow color becomes apparent. Depending on the strain, this could take a few seconds or more than an hour.
7. Add 400 μL Na_2CO_3 and note the elapsed time since ONPG was added.
8. Transfer the reaction mixture to a microcentrifuge tube and spin at 16,000 × g for 10 s.
9. Measure the absorbance at 420 nm of the supernatant from **step 8** relative to that of the control processed in parallel.
10. Calculate the number of units of β-galactosidase activity according to the following formula:

$$\frac{(\text{absorbance at } 420 \text{ nm} \times 1{,}000)}{(\text{assay duration in min}) \times (\text{culture volume analyzed in mL}) \times (\text{OD600 of culture})}$$

11. Normalize the final β-galactosidase activity relative to that of the negative control culture (*see* **Note 14**).

3.5.2. Sequence Determination

Ultimately, one must determine the sequence of the mutant DNA-binding protein or recognition site in the expression or reporter plasmid, respectively. The details of the sequencing strategy to be used will depend upon the specific system under analysis and the technology available to the investigator. We do recommend that the complete set of screens described in **Subheading 3.4** be completed prior to sequence analysis, since the latter is typically laborious, expensive, or both.

3.5.3. Biochemical Analysis of Mutants

Until correlations between in vivo phenotypes and DNA-binding affinity and/or specificity can be verified in a particular system, the approach described here can be applied best as a method for generating interesting mutants that can be subjected to further biochemical analysis. In most cases, this will require expression and purification of the mutant DNA-binding protein under analysis, preferably without fusion to the artificial activation domain. Again, the details of how this can be done will depend upon the specific interaction being studied. We recommend, however, that the necessity of further sequence manipulation, including subcloning into other plasmid vectors, be taken into account in the initial design and construction of the yeast expression vector.

4. Notes

1. The genotype of strain BJ2168 is reported to be *MATa ura3–52 trp1 leu2 pep4–3 prc1–407 prb1–1122 gal2 (6)*. A variety of alternative strains should also be suitable for use. The only relevant genotypes for the approach used here pertain to the nutritional markers used for plasmid selection. If *HIS3* were to be used as a reporter, then it would also be important that the parent strain be *his3*$^-$ or that the chromosomal *HIS3* locus be placed under control of the hybrid promoter being used in the analysis. Clearly, the mutant chromosomal alleles of the nutritional markers should be stable and not give rise to revertants at measurable frequencies. BJ2168 also contains mutations in three vacuolar proteases (the *PEP4*, *PRC1*, and *PRB1* genes), but it is unlikely that these are relevant to the approach described here. In fact, we have made use of other strains that presumably contain wild-type alleles at these loci without detectably altering the results obtained using the *Xenopus* TFIIIA/5S rRNA gene interaction.
2. A variety of reporter gene constructs probably can be used successfully. Plasmid-based reporters could be constructed using various parent plasmids containing different selectable markers, and alternative reporter genes might also be chosen. In addition, the use of a chromosomally integrated reporter gene is possible; in fact, we have successfully used *HIS3* as a chromosomal reporter gene in this fashion to study the *Xenopus* TFIIIA-5S rRNA gene interaction. Advantages of β-galactosidase as a reporter include the fact that it permits a range of phenotypes to be scored and that its activity is easily quantified with a simple spectrophotometric assay. *HIS3*, on the other hand, is probably much more sensitive, permitting detection of weaker DNA–protein interactions or use of a single-copy reporter gene when β-galactosidase reporters on multicopy plasmids may be required. It also allows selection of rare gain-of-function mutants that might occur at such a low frequency that identifying them would be difficult or impossible using a screening protocol like that required with β-galactosidase as a reporter. While *HIS3* expression does not provide a wide dynamic range of phenotypes in a single-plate assay, this limitation can be overcome partially by analyzing growth on plates containing variable 3-aminotriazole (3-AT) concentrations. 3-AT is an inhibitor of the *HIS3* gene product, and can therefore be used to adjust the level of *HIS3* expression required to permit growth.

   We have not investigated the use of different core promoter elements other than that derived from *CYC1*. Even

for *CYC1*, however, levels of reporter gene expression potentially can be manipulated by adjusting the number of copies and orientation, relative to the core promoter, of the DNA target sequence for the binding protein under study. We have found that both orientation and number of TFIIIA-binding sites in the promoter can have substantial effects on reporter gene activity when a TFIIIA-VP16 fusion protein is expressed. In our case, the orientation dependence could be explained readily on the basis of the known polarity of TFIIIA binding to the 5S rRNA gene *(10)*. If similar information is available for the DNA-binding protein of interest, it may be possible to design reporter constructs rationally. Otherwise, it is probably advisable to generate multiple constructs for testing. It is possible that homodimers binding to symmetrical sites may exhibit little or no orientation dependence. With TFIIIA, we found that increasing the number of binding sites from one to two in the reporter gene promoter resulted in a doubling of β-galactosidase activity *(5)*. For many natural pol II activators, the effect of having multiple binding sites is synergistic *(11)*, so it is possible that even larger effects will be observed when the number of binding sites for other DNA-binding proteins is increased. We have not investigated the effects of varying spacing between the artificial upstream activator sequence and the core promoter, but others have shown that transcriptional activity is only moderately affected by alterations in the wild-type spacing between these two elements *(12–14)*. Nonetheless, it is possible that attaining maximal activity in any particular system may require optimizing the spatial relationship between the DNA recognition site of interest and the core promoter. In the end, it may be necessary or desirable to adjust several of these parameters to obtain a level of reporter gene activity that is phenotypically detectable but that also provides for reasonable sensitivity in the detection of mutants that weaken the DNA–protein interaction under analysis.

3. A variety of alternative expression plasmids are available, including several from commercial sources that may be preferable to pG1 in some respects (*see* **Subheading 3.1.2**). We have used vectors other than pG1 and, in some cases, these vectors are better engineered for the easy introduction of DNA fragments encoding the DNA-binding protein of interest while producing an in-frame fusion to an N-terminal activation domain. Choice of expression vector may be affected by the selectable marker used, the presence of an activation domain in the vector and whether the domain is to be N- or C-terminally fused to the DNA-binding domain, and the strength of the promoter driving expression of the

fusion protein. It is also possible that regulated expression of the fusion protein will be desirable, in which case a regulated promoter, like those of the *GAL1* or *CUP1* genes, might be chosen over the constitutive GPD promoter of pG1. We have, in fact, used the metal-inducible *CUP1* promoter for this purpose.

4. The VP16 activation domain we have used *(8)* is probably the best studied activation domain acting to stimulate transcription by RNA polymerase II and is probably also the strongest one in common use. Depending on the system being studied (*see* **Note 5**) and the choice of expression plasmid, an alternative activation domain might be chosen. We have used both the GAL4 activation domain and the synthetic domain B42 in addition to that of VP16 to study the *Xenopus* TFIIIA/5S rRNA gene interaction.
5. It may be necessary to adjust the system's parameters to achieve a level of transcriptional activity that affords a reasonable sensitivity of the assay to variations in DNA-binding strength. Naively, one might expect that the affinity of the relevant DNA–protein interaction would be the most important parameter. Remarkably, our experience in analyzing the *Xenopus* TFIIIA-5S rRNA gene interaction suggests that the level of reporter gene activity is determined not by equilibrium binding affinity, but instead by the kinetic stability of the DNA–protein complex *(16)*. In most cases, kinetic stability and equilibrium binding affinity will covary, but analysis of unusual mutants in TFIIIA that result in large effects on kinetic stability but only minor effects on equilibrium binding affinity reveals that the kinetic parameter is the most relevant one in determining reporter gene activity. Potentially, activity may be affected not only by the affinity or kinetic stability of the DNA–protein interaction, but also by a number of other variables, many of which can be manipulated experimentally to yield a system with optimal sensitivity to variations in the DNA-binding properties of the protein of interest. These could include, for example, the number, spacing, functional synergy between, and orientation of the binding sites in the reporter construct (*see* **Note 2**), the level of expression of the fusion protein (*see* **Note 3**), the fraction of protein that is available for binding (effective in vivo concentration), the copy number of the reporter gene (*see* **Note 2**), the choice of reporter gene (see **Note 2**), and the choice of activation domain (*see* **Note 4**).
6. We have provided recipes for preparation of complete (C) media from individual amino acids and other required components. Synthetic "drop-out" mixes are also available from commercial sources and simplify media preparation

considerably. We routinely use such media lacking uracil (Y1501), or lacking histidine, leucine, tryptophan, and uracil (Y2001) from Sigma-Aldrich.

7. The length of oligonucleotide primers we recommend for error-prone PCR and recombination in vivo is based on our success with primers of this length for mutagenizing TFIIIA and introducing the resulting mutant DNA sequences into an appropriate target vector. In other applications involving homologous recombination to construct plasmids in yeast cells, we have used significantly shorter primers successfully. In fact, it is possible that shorter primers would work equally well *(16, 17)* in this application, but we have not empirically demonstrated that to be the case.
8. If the DNA-binding domain of the protein of interest interacts with its binding site in an asymmetric manner, then the fused activation domain may be presented in two different ways depending on the orientation of the binding site in the reporter gene promoter and whether it is fused to the N- or C-terminus of the DNA-binding protein. In the case of the *Xenopus* TFIIIA/5S rRNA gene interaction, it was important that the activation domain be promoter-proximal in order to obtain high-level transcriptional activation. Either N- or C-terminal activation domain fusions could be made active, but only in combination with the appropriate orientation of binding site in the reporter construct. It is difficult to know how common this phenomenon will prove to be, but it clearly has the potential to impact on the successful use of a particular set of reporter and expression constructs.
9. If the wild-type expression plasmid is used as the template for error-prone PCR, one should keep in mind that this wild-type plasmid will be carried over into the subsequent transformation reaction. This should not generally be a problem unless, for some reason, a low background of wild-type plasmid cannot be tolerated. One circumstance in which this might prove problematic is if very low efficiencies of recombination are obtained, so that most transformants recovered are derived from the template plasmid originally used for PCR and not from integration of PCR products into the gapped target vector. If necessary, the product of the error-prone PCR can be purified prior to transformation.
10. The extent of mutagenesis can be adjusted by manipulating the concentrations of $MgCl_2$and $MnCl_2$ in the polymerase chain reaction. The concentrations given resulted in about 1–2 amino acid substitutions per 344 amino acids when used to mutagenize *Xenopus* TFIIIA. In general, the mutation frequency was higher at greater concentrations of $MgCl_2$ and $MnCl_2$, with the highest mutation rate being observed at

the highest concentrations tested (4.5 mM $MgCl_2$, 0.5 mM $MnCl_2$). The lowest mutation rate was observed at "standard" PCR conditions (1.5 mM $MgCl_2$, no $MnCl_2$). Our evidence would also suggest that there is a substantially greater rate of mutation at positions where the low-concentration deoxynucleoside triphosphate should be incorporated. For this reason, we recommend performing a series of reactions in which each of the four deoxynucleoside triphosphates is represented in turn as the low-concentration nucleotide. One should also note that the conditions suggested for thermal cycling are intended only as a starting point; it may be necessary or desirable to adjust these conditions, particularly the annealing temperature, to reflect differences in primer T_m and specificity. It is worth noting that anecdotal evidence suggests that mutagenesis using error-prone PCR is not truly random. In our hands, for example, certain mutations have been recovered multiple times as independent events; with truly random mutagenesis of a sequence of the length we have targeted for mutagenesis, the odds of this occurring by chance are very, very small. Consequently, one should not imagine that the method samples all possible mutations at equivalent frequencies.

11. A potential problem in this protocol is the recircularization of the target plasmid without integration of the PCR-generated fragment. This can result in a high background of apparent loss-of-function mutants because no protein is being produced from the expression plasmid. To minimize this problem, we recommend several steps. If possible, use two different restriction endonucleases to generate the linear, gapped plasmid. Second, treat the cut plasmid with phosphatase to dephosphorylate the 5′ termini generated by restriction digestion. Third, gel purify the gapped plasmid to eliminate any small fragment that was removed by the restriction digestion as well as to eliminate any uncut molecules that might remain as a result of incomplete cleavage.

12. Although we have not exhaustively optimized the quantities of gapped plasmid and error-prone PCR product that should be used to maximize recovery of colonies containing plasmids with integrated PCR products, we can suggest some guidelines. First, transformation efficiency (number of transformants per unit of DNA) is higher when lower amounts of DNA are used. In practice, there is a trade-off between efficiency and obtaining a reasonable absolute number of transformants. We have found that performing transformations with 50–100 ng gapped plasmid is a reasonable compromise. Second, the fraction of plasmids that contain a recombined insert after cotransformation with the gapped plasmid and

the mutagenized PCR product is affected by the molar ratio between these two components in the transformation reaction (*see* **Fig. 3**). In general, insert-to-plasmid ratios of 10 or higher are the most efficient, although reasonable results can certainly be obtained at lower ratios if the amount of PCR product is limiting.

13. We have frequently used the crude, unpurified products of the error-prone PCR reaction in cotransformations with the gapped expression plasmid to generate and study mutant forms of TFIIIA without encountering significant problems *(5)*. In some related applications, however, we have noted that a high background of colonies can be produced as a consequence of poorly characterized recombination events that we believe derive from the production and integration of aberrant PCR products into the gapped target plasmid. It is most likely that these are small DNA fragments, perhaps "primer dimers." If these are produced at anything other than a very low frequency and are capable of recombining with the target plasmid, they can complicate the identification of loss-of-function mutants in the DNA-binding protein of interest, since they

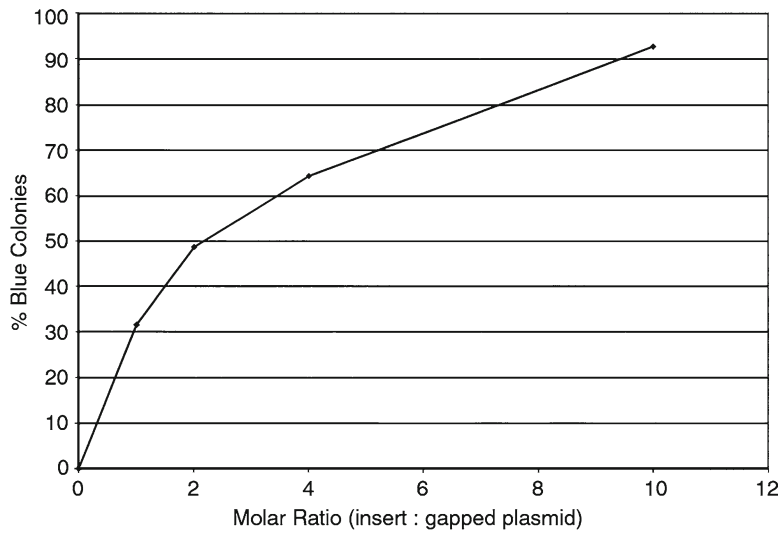


Fig. 3. Dependence of the recovery of recombinant plasmids on the ratio of "insert" to "gapped plasmid" during transformation. Gapped expression plasmid and PCR-generated insert (as defined in the text) were combined in various ratios during transformation of a yeast strain containing a β-galactosidase reporter plasmid, as illustrated in **Fig. 1**. Following recombination in vivo and selection of transformants on C –ura –trp plates, colonies were scored for β-galactosidase activity by replica plating to SSX –ura –trp plates. Only those colonies that had undergone productive homologous recombination produced a functional activator protein that resulted in blue colonies on the indicator plate. Thus, the fraction of colonies that were blue was indicative of the frequency of productive homologous recombination and the generation of a functional expression plasmid.

result in the production of white colonies on X-gal plates. If initial studies in a new system suggest that a similar problem occurs with the use of the crude PCR product, then we recommend purifying the mutagenized PCR product prior to recombination/transformation. Simply gel isolating a product of the correct length should eliminate any problems derived from the presence of unincorporated primers, primer dimers, or other aberrant products of the PCR.

14. We recommend performing β-galactosidase assays on three independent yeast cultures for each strain being analyzed. In addition, we suggest repeating these triplicate assays three times on different days for a total of nine independent determinations. Normalization of activity to a standard seems to be important in correcting day-to-day variability in the determination of β-galactosidase activity.

References

1. Brent, R. and Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729–36.
2. Muhlrad, D., Hunter, R., and Parker, R. (1992). A rapid method for localized mutagenesis of yeast genes. *Yeast* **8**, 79–82.
3. Kuipers, O.P. (1996). Random mutagenesis by using mixtures of dNTP and dITP in PCR. *Methods Mol. Biol.* **57**, 351–356.
4. Ma, H., Kunes, S., Schatz, P.J., and Botstein, D. (1987). Plasmid construction by homologous recombination in yeast. *Gene* **58**, 201–216.
5. Bumbulis, M.J., Wroblewski, G., McKean, D., and Setzer, D.R. (1998). Genetic analysis of Xenopus transcription factor IIIA. *J. Mol. Biol.* **284**, 1307–1322.
6. Jones, E.W. (1990). Vacuolar proteases in yeast *Saccharomyces cerevisiae*. *Methods Enzymol.* **185**, 372–386.
7. Schena, M. and Yamamoto, K.R. (1988). Mammalian glucocorticoid receptor derivatives enhance transcription in yeast. *Science* **241**, 965–967.
8. Triezenberg, S.J., Kingsbury, R.C., and McKnight, S.L. (1988). Functional dissection of VP16, the trans-activator of *Herpes simplex* virus immediate early gene expression. *Genes Dev.* **2**, 718–729.
9. Rose, M.D., Winston, F. and Hieter, P. (1990). *Methods in yeast genetics: a laboratory course manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
10. Smith, D.R., Jackson, I.J., and Brown, D.D. (1984) Domains of the positive transcription factor specific for the Xenopus 5S RNA gene. *Cell* **37**, 645–652.
11. Carey, M., Lin, Y.-S., Green, M.R., and Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature* **345**, 361–364.
12. Struhl, K. (1982). The yeast his3 promoter contains at least two distinct elements. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7385–7389.
13. Brent, R. and Ptashne, M. (1984). A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene. *Nature* **312**, 612–665.
14. Guarente, L. and Hoar, E. (1984). Upstream activation sites of the CYC1 gene of Saccharomyces cerevisiae are active when inverted but not when placed downstream of the "TATA" box. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7860–7864.
15. Oldenburg, K.R., Vo, K.T., Michaelis, S., and Paddon, C. (1997). Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* **25**, 451–452.
16. Brady, K.L., Ponnampalam, S.N., Bumbulis, M.J., and Setzer, D.R. (2005). Mutations in TFIIIA that increase stability of the TFIIIA-5S rRNA gene complex: unusual effects on the kinetics of complex assembly and dissociation. *J. Biol. Chem.* **280**, 26743–26750.
17. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., and Cullin, C. (1993). A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic Acids Res.* **21**, 3329–3330.
18. Smith, P.K. (1985). Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76–85.

Chapter 16

Chromatin Immunoprecipitation in Mammalian Cells

Amy Svotelis, Nicolas Gévry, and Luc Gaudreau

Summary

The ensemble of the genes in the mammalian genome is organized into a structure of DNA and proteins known as chromatin. The control of gene expression by the proteins that bind to chromatin regulates many cell processes, such as differentiation and proliferation. Transcription of protein-encoding genes in mammalian cells is performed by the concerted action of the RNA polymerase II holoenzyme, transcription factors, co-activator complexes that bind to the promoter areas of genes. In addition, different proteins can interact with these complexes and chromatin to create a repressive state. In order to fundamentally understand transcriptional control, it is important to define the areas that these proteins will bind. Classical laboratory techniques unable to provide distinct locations of these factors have now been replaced by the chromatin immunoprecipitation (ChIP) assay. The ChIP technique allows us to isolate chromatin along with its associated proteins from cells and analyse the binding sites of specific proteins and complexes at high resolution.

Key words: Chromatin, Chromatin immunoprecipitation (ChIP), Gene, Transcription, Transcription factor, Activator, Repressor, Protein–DNA interaction.

1. Introduction

The specific regulation of gene expression is a very important process in cellular proliferation, differentiation, and response to stimuli in multicellular organisms. Transcription of a gene will occur with the concerted action of several complexes that bind to the regulatory region of genes and the 5′ crucial region in the recognition of a gene, the core promoter *(1)*. In addition to the regulation of the initiation of transcription, chromatin structure also plays an important role in the regulation of expression.

The regulation of gene expression in eukaryotic cells is controlled by the presence of *cis*-acting regulatory sequences in areas

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_16

surrounding the gene that are bound specifically by proteins involved in transcriptional regulation, called transcription factors *(2)*. Several factors that bind *cis*-acting sequences can recruit not only the RNA Polymerase II (Pol II) complex, but complexes that will change the structure of chromatin at promoters in order to regulate transcription. Transcriptional activators have been shown to interact with and recruit chromatin remodelling complexes to promoters, inducing either the displacement or modification of nucleosomes *(3)*. Repressors can also bind these regulatory sequences and prevent the binding of the transcriptional machinery or can also act indirectly by interfering with the action of activators by either binding to activators or to enhancers in order to prevent the binding of activators to their targets *(4, 5)*. Classical laboratory techniques such as gel shifts provided us with some information concerning the location of these proteins in chromatin, but the conception of the chromatin immunoprecipitation assay has revolutionized the study of the chromatin landscape.

1.1. Principle of Chromatin Immunoprecipitation

A major conundrum in molecular biology has been the identity and specificity of proteins that are involved in both the modification of chromatin and gene expression, such as histones, their variants, elements of chromatin remodelling complexes, and the transcriptional machinery or transcription factors. Recent technology now permits us to understand and uncover existing relationships via the isolation of chromatin from living cells. In order to study the proteins present in chromatin as well as the specific areas to which they associate, we must first break up the chromatin into workable sections and 'fish out' our protein of interest. Then we can determine to which areas our protein associates by analysing the DNA with specific primers and quantitative PCR analysis. This method, called chromatin immunoprecipitation, is now a widely used technique that, although finicky, is essential in determining areas of DNA to which a specific protein binds.

1.2. Overview of the Method

First, protein-DNA interactions are crosslinked chemically using formaldehyde, solidifying any relationships in the cell. Chromatin is isolated from the cells and fragmented by sonication to create 200–500 bp fragments. Using antibodies specific for the DNA-binding protein to be analysed, DNA fragments are immunoprecipitated with the help of agarose beads. DNA is purified and then analysed by quantitative PCR analysis (*see* **Fig. 1**).

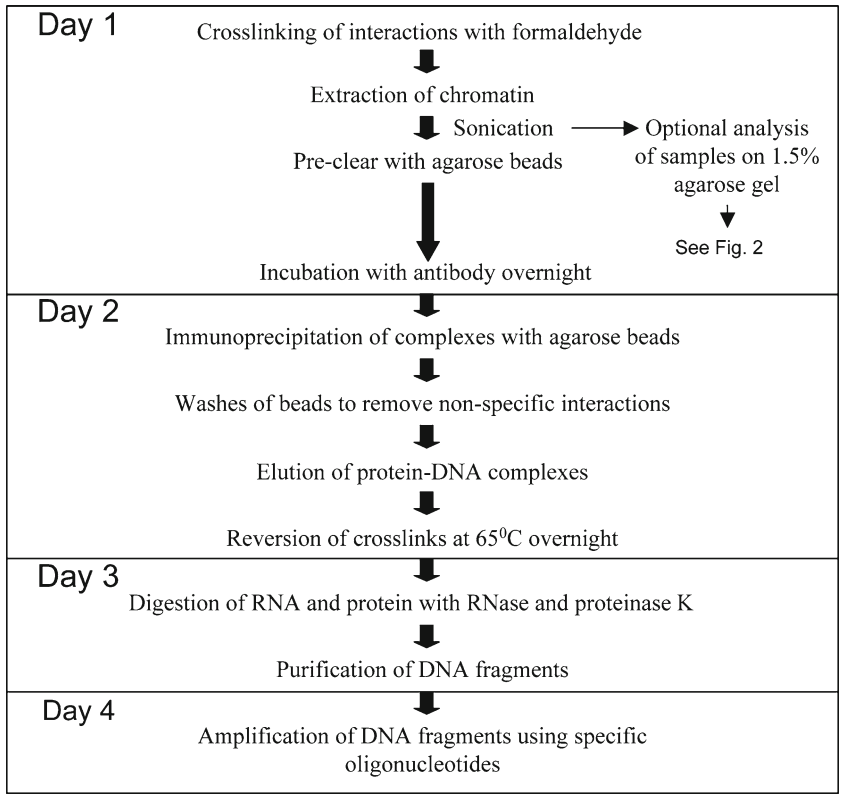


Fig. 1. Layout of a typical ChIP assay. The different steps of the ChIP assay are outlined in this figure, providing a typical and ideal timeline for the undertaking of this experiment.

2. Materials

2.1. Crosslink of DNA–Protein Complexes with Formaldehyde

1. Prepare fresh 1.1% PBS-formaldehyde mix from 37% formaldehyde and PBS of pH 7.4.
2. Prepare a 100× protease inhibitor cocktail stock solution in 90% ethanol: 0.2 mM Pepstatin A, 72 μM Leupeptin, and 26 μM Aprotinin.
3. Prepare Buffer A to a final concentration of: 100 mM Tris–HCl (pH 9.4). Add 1 mM PMSF, 10 mM DTT and 1× protease inhibitor cocktail before use.

2.2. Cellular Lysis and Chromatin Preparation

1. Prepare a stock solution of Buffer I of: 0.25% Triton, 10 mM EDTA, 0.5 mM EGTA, 10 mM HEPES-KOH (pH 6.5). Store at 4°C and add 1 mM PMSF and 1× protease inhibitor cocktail to the required volume for the ChIP before use.
2. Prepare a stock solution of Buffer II of: 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM HEPES-KOH (pH 6.5). Store at 4°C and add 1 mM PMSF and 1× protease inhibitor cocktail to the required volume for the ChIP before use.

3. Prepare SDS lysis buffer of: 1% SDS, 10 mM EDTA, 50 mM Tris–HCI of pH 8.1. Add 1 mM PMSF and 1× protease inhibitor cocktail before use.
4. Prepare a stock solution of IP dilution buffer of: 0.01% SDS, 1.1% Triton, 1.2 mM EDTA, 16.7 mM Tris of pH 8.1, 167 mM NaCl. Store at 4°C and add 1 mM PMSF and 1× protease inhibitor cocktail to the required volume for the ChIP before use.

2.3. Sample Preparation and Pre-clear

2.3.1. Preparation of Agarose Beads

1. Prepare 40 µL of protein A or G (or mix A/G) agarose beads (50% slurry) per IP (*see* **Note 1** and **Table 1**).
2. Wash beads three times with the IP dilution buffer (2,200 × *g*) for 3 min at 4°C.
3. Resuspend beads to 50% slurry with the IP dilution buffer.

Table 1
Affinities of antibodies to protein A/G. The different affinities of antibodies from different species to protein A/G, with the '+' indicating a positive affinity, and the number of '+' indicating the strength of the affinity

Species	Affinity for protein A	Affinity for protein G	Subclass	Affinity for protein A	Affinity for protein G
Mouse	++	++	IgG1	+	++++
			IgG2a	++++	++++
			IgG2b	+++	+++
			IgG3	++	+++
Rat	+/−	++	IgG1	−	+
			IgG2a	−	++++
			IgG2b	−	++
			IgG2c	+	++
Chicken	−	+			
Cow	++	++++			
Goat	−	++			
Guinea pig	++++	++			
Hamster	+	++			
Horse	++	++++			
Pig	+++	+++			
Sheep	+/−	++			
Rabbit	++++	+++			

4. Add 2 μg of sonicated salmon sperm DNA per IP to the prepared beads (*see* **Note 2**).

2.4. Washes and Elution of Immunoprecipitated DNA–Protein Complexes

1. Wash buffers should ideally be chilled at to 4°C for the washes. All the wash buffers can be prepared as stock solutions and stored at 4°C. Most of the solutions required to make the wash solutions can be prepared, autoclaved, and stored at room temperature, except for the LiCl which is preferably stored at 4°C.
2. For the TSE-150: 0.1% SDS, 1% Triton, 2 mM EDTA, 20 mM Tris–HCl of pH 8.1, 150 mM NaCl.
3. For the TSE-500: 0.1% SDS, 1% Triton, 2 mM EDTA, 20 mM Tris of pH 8.1, 500 mM NaCl.
4. For the LiCl detergent: 0.25 M LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris of pH 8.1.
5. For the TE 1×, prepare 10× TE, sterilize by autoclaving, and store at room temperature.

3. Method

This ChIP method is performed over several days and proper time management is crucial in the success of this experiment, as any 'shortcuts' can lead to variability in results! For mammalian cells, the amount of cells required varies per cell type, from 2×10^6 to 10×10^6 cells per ChIP. Therefore, it is important to verify the literature for the specific cell type you are using. This protocol is designed for cells cultured in 150-mm dishes, so modify the volumes for the following steps as required.

3.1. Crosslink of DNA–Protein Complexes with Formaldehyde

1. Remove culture medium from cells and wash cells twice with PBS 1×.
2. Add 20 mL PBS-formaldehyde mix (1.1%) per 150-mm dish and incubate 10 min at room temperature (*see* **Note 3**).
3. Remove formaldehyde and wash cells twice with cold PBS, removing all traces of PBS after the second wash.
4. Add 1 mL Buffer A to cells, collect cells using a cell scraper, and incubate for 20 min at 30°C (*see* **Note 4**).
5. Centrifuge for 6 min at 1,000 × *g* at 4°C (*see* **Note 5**).

3.2. Cellular Lysis and Chromatin Preparation

1. Resuspend pellet in 1 mL of buffer I (*see* **Note 6**).
2. Centrifuge for 6 min at 1,000 × *g* at 4°C.
3. Resuspend pellet in 1 mL of buffer II and centrifuge for 5 min at 1,000 × *g* at 4°C.

4. Resuspend pellet in 100 µL per immunoprecipitation (IP) experiment in SDS lysis buffer and incubate on ice for 10 min (*see* **Notes 6** and **8**).

3.3. Sonication of Lysate

1. Sonicate samples 3× 10 s at amplitude 60, including 3 min intervals on ice between each sonication (*see* **Note 9 Fig. 2**).
2. Centrifuge for 10 min at full speed (15,000 × *g*) at 4°C.
3. Keep 5% of the volume of 1 IP for analysis of results (='input').

3.4. Sample Preparation and Pre-clear

1. To the sonicated samples, add 10 volumes of IP dilution buffer (include in this total the volume of IP dilution buffer added before sonication) with a final concentration of 1 mM PMSF and 1× protease inhibitor cocktail.
2. Pre-incubate samples with 40 µL of beads (50% slurry) per IP and 1 mg/mL BSA (final) or 15–20 µL of pre-immune serum per IP for 2–4 h at 4°C on a rotator (*see* **Note 10**).

3.5. Immunoprecipitation

1. Centrifuge preincubated samples at 1,000 × *g* for 2 min at 4°C.
2. Transfer 1 mL of supernatant per IP into a microtube and add the desired antibody (*see* **Notes 11** and **12**).
3. Incubate overnight at 4°C on a rotator.
4. Add 40 µL of prepared beads (at 50% slurry) per IP and incubate for 2–4 h at 4°C on a rotator.

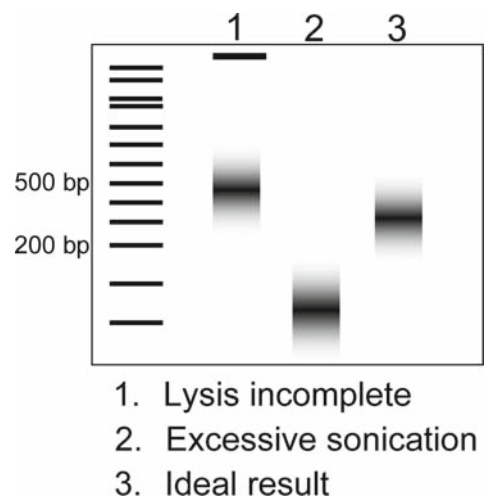


Fig. 2. Possible migration patterns following agarose gel electrophoresis of DNA obtained from the ChIP assay. *Lane 1* shows a migration pattern that suggests incomplete cell lysis, with a band of slowly migrating genomic DNA at the top of the lane. *Lane 2* demonstrates DNA that has been oversonicated and can lead to aberrant results. *Lane 3* represents the ideal migration pattern, with purified DNA sonicated to fragments between 200 and 400 bp.

3.6. Washes and Elution of Immunoprecipitated DNA–Protein Complexes

1. Centrifuge for 3 min at 2,200 × *g* at 4°C.
2. Wash beads with 800 µL of wash buffer, incubate at 4°C for 5 min on a rotator, and centrifuge for 3 min at 2,200 × *g* at 4°C using the following:
   - two washes with TSE-150
   - three washes with TSE-500
   - two washes with LiCl detergent
   - two washes with TE 1×
3. After the final wash, remove all traces of TE and add 200 µL of 1% SDS/0.1 M $NaHCO_3$.
4. Vortex and incubate 20 min at 65°C, vortexing every 2 min.
5. Centrifuge at 4,200 × *g* for 2 min and transfer the supernatant to a new microtube.
6. To reverse the DNA–protein crosslinks, incubate 4–6 h or overnight at 65°C (*see* **Note 13**).

3.7. DNA Preparation

1. Digest samples with RNase A (20 µg) for 30 min at 37°C.
2. Add 100 µg of proteinase K to digest proteins and incubate for 2–3 h at 37°C.
3. Perform two phenol/chloroform extractions and one chloroform extraction (*see* **Note 14**):
   - Add 1 volume, vortex, centrifuge at full speed for 5 min
   - Transfer supernatant to a new tube and repeat
4. Add 2 volumes of 100% ethanol and 1/10 volume of 3 M NaOAc (*see* **Note 15**).
5. Incubate overnight at –20°C.
6. Centrifuge for 10 min at full speed.
7. Wash once with 70% ethanol, remove all traces of ethanol, resuspend in 50 µL of sterile water, and store at –20°C.
8. Dilute input DNA to 1%.

3.8. Analysis of Results

1. DNA immunoprecipitated during the ChIP can be analysed using quantitative real-time PCR. Design primers that will amplify specific target regions of 80–150 bp. The size of the amplicon can vary to accommodate ideal amplification conditions (*see* **Note 16**).
2. The resulting threshold amplification levels of immunoprecipitated samples (CT^{IP}) or of samples without antibody ($CT^{no\ Ab}$) are compared to the input DNA (CT^{input}) by subtraction ($CT^{no\ Ab\ or\ IP} - CT^{input} = \Delta CT$) (*see* **Notes 17** and **18**).
3. For a relative DNA level, we calculate $2^{-\Delta CT}$ for all samples. Background levels can be eliminated by subtracting the no Ab result from the ChIP results.

4. Notes

1. The salmon sperm DNA blocks non-specific interactions with the beads.
2. The pre-clear is important to eliminate a large amount of non-specific interactions of the samples to the beads.
3. The incubation time in formaldehyde can vary depending on the degree of crosslink you require. For example, a 10-min incubation time is perfect for crosslinking chromatin-associated proteins to DNA, yet to crosslink transcription factors the crosslink reaction can be extended.
4. Alternatively, the crosslink reaction can be stopped by adding 1/20 volume of 2.5 M glycine to plates to quench formaldehyde following the crosslink and incubating for 5 min. The pellet can either be lysed in Buffer I and II followed by SDS lysis buffer, or the pellet can directly be lysed with SDS lysis buffer following this step. The lysis protocol can vary according to cell type.
5. Pellets can be frozen in liquid nitrogen and conserved at –80°C at this point.
6. Cellular pellets of the same type/treatment can be pooled at this point, using a total of 1 mL of buffer I
7. The incubation in the lysis buffer can be prolonged if necessary to properly lyse cells.
8. IP dilution buffer can be added after incubation if the volume is not sufficient for sonication (ideal volume for sonication is 500 μL).
9. The level of sonication can vary depending of the cell type and the apparatus (the apparatus used in this protocol was a Branson Digital Sonifier 450). In order to determine the ideal sonication time and amplitude, it is important to test before embarking on the full ChIP assay. Test different sonication settings by preparing cells up to this point and verifying the level of DNA fragmentation on agarose gel. In order to do so, collect 5% input and proceed to a rapid DNA preparation. Digest with 20 μg RNase and 100 μg proteinase K and load on a 1.2% agarose gel (if the volume is superior to that which you can load on a gel, proceed to a DNA precipitation mentioned in the section 'DNA preparation'). The ideal size of DNA fragments following sonication is between 200 and 500 bp. Adjust sonication settings accordingly.
10. It is important to determine which beads are required for the specific antibodies you will be using.

11. It is important here to include a positive control (e.g. for histones, H3 is an interesting positive control) as well as a negative control, the supernatant incubated with no antibody or non-specific antibody. Thus, it is essential to plan ahead when preparing cells for the ChIP assay that you will have the required amount for all the control situations.
12. The concentration of antibody required here is determined by the supplier of your antibody or by testing any homemade antibody. In general, 2–4 μg of antibody per ChIP is sufficient. The concentration of the antibody should be empirically determined.
13. DO NOT FORGET to reverse the crosslinks in the input samples collected in **step 3** of **Subheading 3.3** by adding 200 μL of elution buffer to the input samples and incubating overnight with ChIP samples.
14. It is VERY IMPORTANT to remove all traces of phenol from your samples, as any contamination will disturb the amplification of DNA by qPCR. Alternatively, a commercial DNA purification kit can also be used, preventing any possible phenol contamination.
15. 20 μg of glycogen can be added in order to aid in the precipitation and visualization of the DNA pellet.
16. If it is required to analyse a region with a greater resolution, design primers that will allow the amplicons to encompass a large area yet overlap one another by 20–50 bp. The freeware Primer3 available on the internet at http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi is useful in creating primers that are compatible with qPCR standards. The annealing temperature of 60°C is ideal for qPCR reactions.
17. This eliminates the difference in efficiency between the different primers.
18. Alternatively, input DNA can be diluted to 0.01, 0.1, and 1% to produce a standard curve that will give precisely the % input of the IP.

References

1. Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**, 466–469.
2. Chen, L. (1999). Combinatorial gene regulation by eukaryotic transcription factors. *Curr. Opin. Struct. Biol.* **9**, 48–55.
3. Szutorisz, H., Dillon, N., and Tora, L. (2005). The role of enhancers as centres for general transcription factor recruitment. *Trends Biochem. Sci.* **30**, 593–599.
4. Mitchell, P.J., Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371–378.
5. Villard, J. (2004). Transcription regulation and human diseases. *Swiss Med. Wkly.* **134**, 571–579.

Chapter 17

Sequential Chromatin Immunoprecipitation Protocol: ChIP–reChIP

Mayra Furlan-Magaril, Héctor Rincón-Arano, and Félix Recillas-Targa

Summary

Chromatin immunoprecipitation has been widely used to determine the status of histone covalent modifications and also to investigate DNA-protein and protein-protein associations to a particular genomic location in vivo. Generally, DNA regulatory elements nucleate the interaction of several transcription factors in conjunction with ubiquitous and/or tissue-specific cofactors in order to regulate gene transcription. Therefore, it has become relevant to determine the cohabitation of several proteins in a particular developmental stage and cell type. Furthermore, multiple post-translational histone modifications can be analyzed on the same genomic location with the aim of deciphering the combinatorial pattern of histone modifications associated to specific transcriptional stages during cell commitment. Here we describe the ChIP–reChIP assay that represents a direct strategy to determine the in vivo colocalization of proteins interacting or in close contact in a chromatinized template on the basis of double and independent rounds of immunoprecipitations with high-quality ChIP grade antibodies.

Key words: Chromatin, Protein–DNA interaction, Immunoprecipitation, Histone modifications, GATA-1, YY1, Telomeric position effect.

1. Introduction

The chromatin status and the in vivo association of transcription factors to DNA regulatory regions or even RNA are essential to understand the transcriptional regulatory circuits that orchestrate cellular processes such as cell proliferation and differentiation among others. Chromatin immunoprecipitation has proven to be a powerful assay to characterize DNA-protein and protein-protein interactions in vivo and determine their association to a particular genomic region. However, ChIP–reChIP assays allow

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_17

us to determine whether two or more proteins are close enough in a chromatin context or whether they are interacting, probably forming a complex, in a localized genomic region *(1–3)*. It is worthwhile to mention that this method does not allow DNA-binding sequence determination, for which other in vitro techniques such as DNase I footprinting or electrophoretic mobility shift assays are needed. Although ChIP–reChIP can suggest the cohabitation of two proteins in the genome and their probable physical interaction, direct protein-protein contacts cannot be addressed by sequential ChIP assay. For such aims, standard Co-IP or in vitro pull-down assays should be used.

Sequential chromatin immunoprecipitation is based on the principle that chromatin and associated transcription factors and/or cofactors are first immunoprecipitated with a specific antibody, and the eluted material is then subjected to another immunoprecipitation with a second antibody. Importantly, this second antibody recognizes a protein that is suspected to be located near (or forms a direct or indirect complex with) the first immunoprecipitated antigen. Then, the recovered DNA is purified and constitutes the genomic template for PCR amplification using specific primers for the genomic region of interest. At this point, three important controls are needed:

1. In theory, the order of the antibodies used for immunoprecipitation could be inverted and still be expected to deliver the same PCR amplification product. It is however possible not to obtain any immunoprecipitation enrichment depending on the order in which the antibodies were used: this could suggest that one of the two antibodies is not able to access its epitope due to sterical restrains, probably originating from the associated proteins in the complex. It could be that the components under study have different relative abundance in the region of interest *(2)*.
2. Negative genomic controls for PCR amplification are also important, i.e., a set of primers should be designed within a genomic sequence in which we do not expect any interaction of the proteins under study, resulting in a clear absence of PCR amplification product. For example, to evaluate a given regulatory element, a set of primers coming from a distal region (like an exon) or located on a distinct chromosome could represent a proper control.
3. To determine and confirm the specificity of the ChIP–reChIP assay, it is useful to perform the sequential immunoprecipitation with the antibodies under study and an irrelevant antibody (i.e., nonspecific IgG).

Sequential chromatin immunoprecipitation has several potential applications and can be useful to complement indirect assays at the protein and chromatin structure levels.

1.1. Chromatin Status

The applicability of sequential ChIP has recently been assayed in embryonic stem (ES) cells. ES cells are pluripotent cells present in the early mouse embryo in which the chromatin structure is considered to be greatly dynamic and globally relaxed, permitting its pluripotency and contributing to the highly sophisticated process of cell differentiation into any lineage *(4, 5)*. Consequently, ES cells are subjected to intense chromatin reprogramming that involves histone modifications and chromatin movement inside the nucleus *(6, 7)*. It has been shown through sequential ChIP that, on particular locations of the ES cell genome, there are "bivalent chromatin domains." These refer to the coexistence of repressive marks (H3K27me3) and active marks (H3K4me3) on the same genomic region *(8)*. This coexistence is proposed to prepare the genome of ES cells in order to make rapid epigenetic decisions on the transcriptional status (activation or repression).

The analysis of sequential chromatin immunoprecipitation assays can also be performed in a quantitative manner through the use of Real-Time PCR or duplex-PCR *(8–10)*. In our laboratory, we have systematically used radioactive duplex-PCR to analyze the colocalization of antagonistic histone covalent modifications on stably integrated transgenes in cultured cells, in the presence or absence of chromatin insulators (**ref.** ***10*** and *see* **Fig. 1**). Duplex-PCR allows the normalization of immunoprecipitation enrichments and the comparison, in the same genomic location, of the amount of specific histone post-translational modifications *(11)*. Particular attention is needed in the design of these control primers, since they should correspond to known genomic regions where the chromatin state is contrary to the histone mark under study. For example, if we analyze the H3 hyperacetylation over the studied region (corresponding to an open histone mark) the control primers should amplify a known region of heterochromatin (corresponding to a closed histone mark) and vice versa. This strategy allows the determination of the relative abundance of a particular group of histone marks and their cohabitation in the same genomic regions.

1.2. In Vivo Determination of Protein Complex Formation

One of the principal uses of sequential chromatin immunoprecipitation is the determination of multiple interactions between different proteins and DNA. As mentioned this experimental strategy does not allow the determination of the contacted sequences, or the order of the interactions, or their number. For example, we have recently determined that the GATA-1 and YY1 transcription factors colocalize in vivo at the chicken α-globin enhancer (**ref.** ***2*** and *see* **Fig. 2**). Whole extracts coimmunoprecipitations and pull-down experiments were further required to demonstrate a physical contact between them. These results do not preclude the possibility of interaction by other factors, and in particular that of specific cofactors with the capacity to recruit

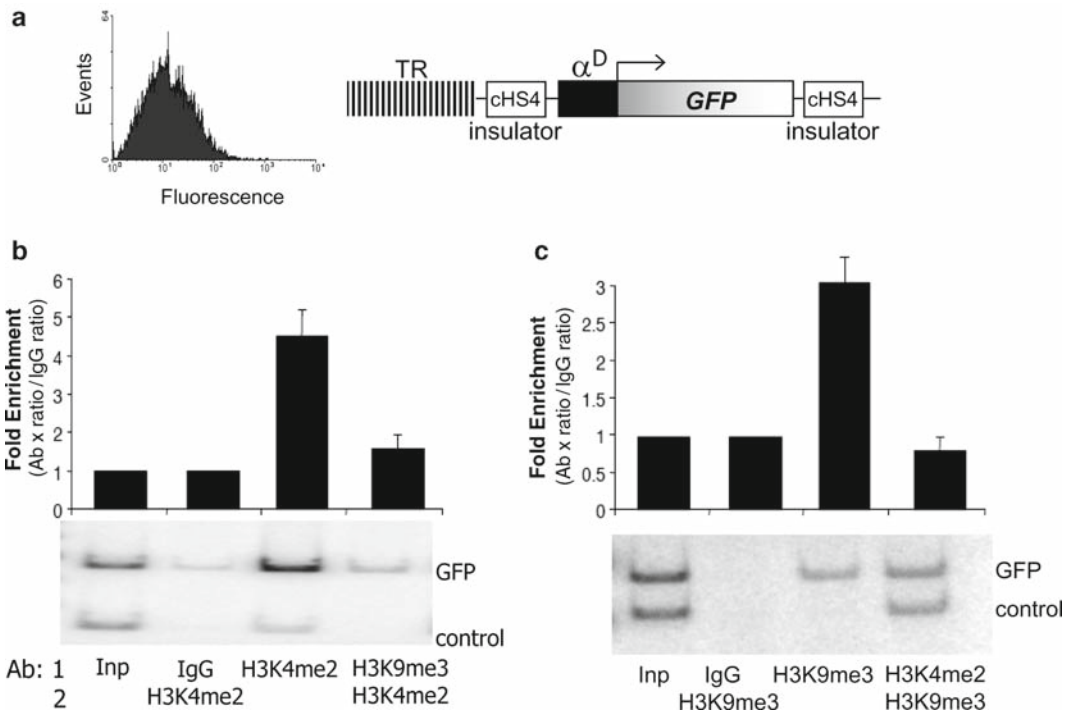


Fig. 1. Re-ChIP analysis by duplex PCR. In this experiment the coexistence of H3K4me2 and H3K9me3 in the body of a *GFP* reporter gene was analyzed. HD3 cells were stably transfected with the plasmid shown in the diagram. The plasmid contains *GFP* reporter gene under chicken α^D globin gene promoter; additionally, the transgene is flanked by the 5′HS4 β-globin insulator and contains a telomeric repeat fragment. This plasmid was used to evaluate the protection against telomeric position effect of the 5′HS4 insulator *(10)*. (**a**) Fluorescent cytometry analysis of the *GFP* expressing clone used for re-ChIP experiments. (**b**) Re-ChIP using H3K4me2 for the first round of IP. The control region used corresponds to a heterochromatic region upstream of the β-globin gene domain in which this mark is absent. As can be seen in the graph and gel, there is no re-ChIP signal even if the H3K4me2 alone is present. (**c**) Re-ChIP using H3K9me3 for the first round of IP. The control region used corresponds to the chicken folate receptor hypersensitive site HSA known to be in an open chromatin conformation that lacks H3K9me3. There is no re-ChIP signal even the H3K9me3 is present. The previous experiments show that the two marks are not localized in the same region. In consequence there are two different cell populations, one with each histone modification.

chromatin-remodeling activities. This kind of association has been recently demonstrated for the nuclear protein NLI/Ldb1 (the human homolog of the *Drosophila melanogaster* protein Chip) which was not known to possess enzymatic or DNA-binding activity *(3, 12)*. Sequential chromatin immunoprecipitation demonstrated that NLI/Ldb1 is able to form a complex with LMO2, SCL, and GATA-1 that contribute to β-globin regulation, but most importantly that facilitate long-range interaction between the β-globin locus control region and the target globin gene by chromatin loop formation *(3)*. This sequential ChIP experiments were complemented with a ChIP assay followed by a chromosome conformation capture (3C) assay demonstrating that NLI

is responsible for long-distance interactions *(3)*. These are two examples supporting the experimental contribution of the ChIP–reChIP assay to the in vivo determination of protein location that could also be complemented with other methodologies like the 3C assay or the widely used ChIP followed by microarrays allowing the identification of specific proteins on a genome-wide scale or their associated DNA sequence *(1, 3, 9, 13, 14)*.

Based on the vast amount of data showing that genomic regulatory elements are functional through the formation of multiple peptide complexes and the variety of histone modifications on a single nucleosome, ChIP–reChIP assays represent a very useful experimental tool to decipher the diversity of combinations of regulatory protein and chromatin modifications in the genome.

2. Materials

2.1. Cell Culture

1. HD3 cells are chicken erythroblasts arrested in a CFU stage via infection with the avian erythroblastosis virus (AEV). HD3 cells are cultured in DMEM (GIBCO) supplemented with 8% Fetal Bovine Serum (Invitrogen), 2% Chicken Serum, and 1× Penicillin/Streptomycin (Invitrogen) at 37°C, 5% CO_2 in a 150-mm dish.

2.2. ChIP–reChIP

1. PBS: NaCl, 137 mM; KCl, 2.7 mM; Na_2HPO_4, 10 mM; and $KHPO_4$, 2 mM of pH 7.2.
2. PBS/2% FBS/PMSF, 1 mM.
3. Formaldehyde 11% (25 mL): 7.45 mL 37% formaldehyde, 0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA, and 50 mM HEPES.
4. Glycine 2.5 M: 18.767 g in 100 mL sterile water.
5. Lysis buffer: 1% SDS, 5 mM EDTA, 50 mM Tris–HCl of pH 8.1. Supplement with protease and deacetylase inhibitors (SIGMA) before using as follows: 1 mM PMSF, 100 μM leupeptin, 0.3 μM aprotinin, 1 μM pepstatin A, 10 μM bestatin, and 1 mM sodium butyrate.
6. Dilution buffer: 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl of pH 8.1, 150 mM NaCl. Supplemented with protease inhibitors before using.
7. Wash buffer I: 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris–HCl of pH 8.1. Supplemented with protease inhibitors before using.
8. Wash buffer II: 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris–HCl of pH 8.1. Supplemented with protease inhibitors before using.

9. Wash buffer III: 0.25 M LiCl, 1% NP40, 1% Deoxycholate, 1 mM EDTA, 10 mM Tris–HCl of pH 8.1. Supplemented with protease inhibitors before using.
10. TE: 10 mM Tris–HCl of pH 8.0, 2 mM EDTA. Supplemented with protease inhibitors before using.
11. Washing beads buffer: 0.1 mM $NaHCO_3$, 1% SDS. Add 10 μg/μL proteinase K and 10 μg/μL RNase A before use.
12. Proteins A and G (Amersham). Once beads (proteins A + G) are preadsorbed with BSA and salmon sperm DNA, store them at 4°C for up to 6 months.
13. Bovine Serum Albumin Fraction V (BSA).
14. Salmon sperm DNA.
15. MiniElute purification Kit (Qiagen).
16. Phenol: Chloroform:Isoamyl alcohol (25:24:1).
17. Antibodies against H3K4me2 and H3K9me3 (Dr. Thomas Jenuwein, IMP Vienna, Austria), rabbit α-IgG (Zymed) (**Fig. 1**), antibodies against GATA-1 (H-200 Santa Cruz Biotechnology) and YY1 (H-414 Santa Cruz Biotechnology) (**Fig. 2**).
18. Dithiothreitol (DTT).

2.3. Duplex PCR and Acrylamide Gel

1. Taq Polymerase (Qiagen).
2. dNTPs (Roche).
3. Specific oligonucleotides.
4. [α-^{32}P] dCTP (Amersham).
5. 40% Acrylamide.
6. TBE buffer.
7. Ammonium persulfate (APS) and TEMED.
8. Phosphor screen (Amersham).
9. Typhoon Scanner.
10. Image Quant program.

3. Methods

The purpose of ChIP–reChIP assay is to evaluate the coexistence of two different factors or histone modifications in a given genomic region. ChIP assay alone is unable to distinguish between two factors coexisting in the same DNA region and two different

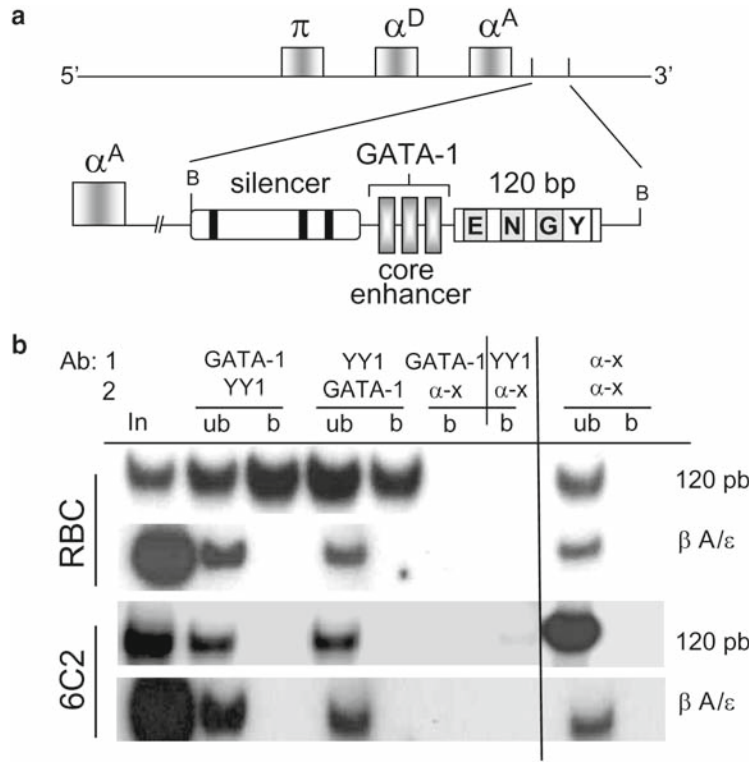


Fig. 2. Re-ChIP analysis by qualitative PCR. In this experiment we analyze the coexistence of GATA-1 and YY1 in a 120-bp region of the 3′enhancer of the chicken α-globin gene domain. These factors have been shown to modulate enhancer's activity in an erythroid-stage-specific manner *(2)*. (**a**) Diagram of the α-globin gene domain and the 3′enhancer element. The core enhancer presents three GATA-1 binding sites. In addition, a 120-bp conserved sequence downstream of the core enhancer has an EKLF (*E*), NF-E2 (*N*), and a fourth GATA-1 (*G*) and YY1 binding site (*Y*). (**b**) Re-ChIP of GATA-1 and YY1. As can be seen, both factors are binding the 120 bp in the mature red blood cells (RBC) stage while in 6C2 cells, which represent a pre-erythroblastic stage in erythroid differentiation, they are not. These results show that binding of both factors to the 120 bp is stage dependent, and these correspond to their influence on enhancer action over the globin genes *(2)*. Control region corresponds to the βA/ε enhancer in the β-globin gene domain.

cell populations with each one. A ChIP–reChIP experiment consists basically in performing a ChIP assay with two rounds of immunoprecipitation before DNA recovery. For simplicity in the following protocol, we will refer to antibodies A and B for each of the putative factors to analyze. Two examples are presented (**Figs. 1** and **2**) in which different factors are analyzed in distinct genomic contexts.

3.1. ChIP–reChIP Assay

3.1.1. Day 1

Before starting, add protease and deacetylase inhibitors (1 mM PMSF, 100 μM leupeptin, 0.3 μM aprotinin, 1 μM pepstatin A, 10 μM bestatin, and 1 mM sodium butyrate) to lysis and dilution buffers.

1. Grow cells to an 80–90% confluence in a 150-mm dish.
2. Recover the cells in a 50-mL Falcon tube.
3. Count the cells to ensure an adequate and comparable number per assay (a minimum of 10×10^6 cells are needed per antibody). Chromatin for ten different antibodies is obtained from a cell culture performed in two 150-mm semiconfluent HD3 dishes ($10–20 \times 10^6$ cells) (*see* **Note 1**).
4. Centrifuge 5 min ($200 \times g$) at 4°C.
5. Wash twice with 20 mL PBS/2% FBS/PMSF 1 mM. Gently resuspend the pellet with a 25-mL pipette. Centrifuge each time 5 min ($200 \times g$) at 4°C.
6. Resuspend the pellet with 30 mL PBS. Crosslink by incubating cells with 1% formaldehyde 10 min at room temperature (add 3 mL of a formaldehyde 11% stock solution and mix gently by flipping the tube three times).
7. Stop crosslink by adding glycine to a 125 mM final concentration. (1.5 mL of a 2.5 M glycine stock solution). Gently mix by flipping the tubes.
8. Centrifuge 5 min ($200 \times g$) at 4°C.
9. Wash cells twice with 35 mL of cold PBS. Centrifuge each time 5 min ($200 \times g$) at 4°C.
10. Resuspend cells in 1 mL lysis buffer and transfer to a 15 mL Falcon tube. Incubate 10 min on ice and proceed to chromatin sonication.
11. Sonicate chromatin to obtain fragments of a desired size depending on the expected PCR product (*see* **Note 2**). To obtain HD3 chromatin fragments between 300 and 500 bp sonicate with 35% amplitude giving 12 pulses of 30-s each (Cole and Palmer ultrasonic processor). Be sure to keep the tip of the sonicator at the bottom of the tube without touching the tube walls while giving the pulse. Put the tube on ice for two minutes between pulses.
12. Transfer chromatin into a 1.5 mL tube and centrifuge ($15{,}000 \times g$) for 10 min at 4°C.
13. Transfer the supernatant into a 15 mL Falcon tube. Take a 50 μL aliquot as INPUT and store it at –20°C.
14. Dilute chromatin 1:10 with dilution buffer (*see* **Note 3**).
15. Incubate chromatin with 50 μL of previously blocked beads (*see* **step 20**) for 2 h at 4°C in constant rotation (*see* **Note 4**).
16. Centrifuge ($1{,}000 \times g$) for 5 min at 4°C.

17. Divide chromatin into 1 mL aliquots, one aliquot per antibody. This chromatin can be stored at 4°C for up to 4 months.
18. For the first round of IP take four aliquots: (1) Antibody A, (2) Control IgG, (3) Antibody B, and (4) Control IgG (*see* **Note 5**).
19. Add 4 μg of antibody to each sample and incubate overnight at 4°C with constant rotation (*see* **Note 6**).
20. Simultaneously, prepare Sepharose beads as follows: Take 0.2 g of protein A and 100 μL of protein G in a final volume of 2 mL of 9:1 dilution buffer:lysis buffer. Preadsorb beads with 100 μg/mL BSA and 500 μg/mL salmon sperm DNA and rotate overnight at 4°C.

3.1.2. Day 2

Before starting add protease inhibitor to the lysis and dilution buffers as before.

21. Remove the preabsorption mix and wash the beads twice with 2 mL of dilution buffer. Finally resuspend the beads in 2 mL 9:1 dilution buffer:lysis buffer. Store adsorbed beads at 4°C for up to 6 months.
22. Add 50 μL of preadsorbed beads per sample.
23. Rotate samples for 2–4 h at 4°C.
24. Add protease inhibitors to the washing buffers I, II, III, and TE.
25. Centrifuge beads (800 × *g*) for 1 min in a conventional table centrifuge (Eppendorf). Store 50 μL of supernatant as "unbound" at –20°C and remove the rest with a pipette (*see* **Note 7**).
26. Wash sequentially with 1 mL of washing buffers I, II, and III, rotating 10 min at 4°C. Centrifuge (800 × *g*) for 1 min between washes. Remove supernatant carefully with a pipette.
27. Wash twice with 1 mL TE as before.
28. Resupend beads in 75 μL TE/10 mM DTT (*see* **Note 8**).
29. Elute immunocomplexes by incubating 30 min at 37°C.
30. Centrifuge (800 × *g*) for 2 min and transfer the supernatant into a clean 1.5 mL tube.
31. Dilute sample 20 times (to a final volume of 1.5 mL) with dilution buffer.
32. For the second round of IP use: (1) Antibody B, (2) Antibody B, (3) Antibody A, (4) Antibody A. Take two new chromatin aliquots and incubate with (5) Antibody A and (6) Antibody B (*see* **Note 9**). Add 4 μg of antibody to each sample. Incubate overnight at 4°C with constant rotation.

3.1.3. Day 3

33. Proceed in the same way as before **(steps 21–26)**.
34. Take the INPUT and UNBOUND samples out of the freezer.
35. Add 150 μL of washing beads buffer to each sample plus proteinase K and RNase A (5 μg/μL). Incubate for at least 4 h or overnight at 65°C to reverse the crosslink.

3.1.4. Day 4

36. Extract DNA using the MiniElute purification Kit (Qiagen), following the manufacturer's instructions (*see* **Note 10**).
37. Elute DNA in 50 μL of sterile water. Dilute INPUT and UNBOUND samples up to 300 μL. Proceed to PCR reactions.

The obtained data could be analyzed using different approaches. Here we describe two of them: a semiquantitative one in which results are normalized using a control region in a duplex PCR reaction (**Fig. 1**) and a qualitative one in which no duplex PCR is performed and amplification signal of the immunoprecipitated (bound) chromatin is compared with the signal obtained using unbound sample as template (**Fig. 2**).

3.2. Semiquantitative Duplex PCR

In order to quantify the enrichment of a factor or histone mark (in this case antibodies A and B) in a given genomic region, a duplex PCR is performed. One PCR product corresponds to the region of interest and the other PCR product corresponds to a control region in which the evaluated mark is not expected, in order to normalize the data (**Fig. 1**).

1. Design the oligonucleotides to amplify the experimental and control regions (*see* **Note 11**).
2. Standardize radioactive duplex PCR with INPUT DNA. Be sure that the Experimental/Control ratio is near 1 and that amplification is at the linear range (*see* **Note 12**). Perform PCR reactions as indicated in **Tables 1** and **2** (these conditions were the ones used for the experiment shown in **Fig. 1**).
3. Once the radioactive duplex PCR conditions are established, proceed to perform reactions using the ChIP DNA as template. In this case: (1) Input, (2) IgG + A, (3) A, (4) B + A, (5) Input, (6) IgG + B, (7) B, and (8) A + B.
4. Run samples in a 6% acrylamide gel in 0.5% TBE buffer at 150 V for 1 h and 45 min (*see* **Note 13**).
5. Dry the gel on 3MM paper for 2 h.
6. Expose gel on the phosphor screen or film for 2 h.
7. Reveal image on Typhoon Scan and analyze the results in Image Quant program using the volume report tool (*see* **Note 14**).

Table 1
Radioactive PCR

Reagent	Amount
DNA (Input or ChIP sample)	3 µL (of the 50 µL)
dNTPs (10 mM)	0.1 µL
dCTP [α-^{32}P] 10 mCi/mL (Amersham)	0.05 µL
Experimental oligos forward + reverse (pool of 5-pmol each)	0.5 µL
Control oligos forward + reverse (pool of 5-pmol each) (*see* **Note 15**)	0.5 µL
Buffer 10×	2 µL
Taq Polymerase	0.05 µL
H_2O	Up to 20 µL

Table 2
PCR cycle conditions

Cycle step	Conditions
Denaturation	95°C for 5 min
Denaturation	95°C for 50 s
Annealing	T_m for 50 s
Extension	72°C for 50 s
Extension	72°C for 5 min
Hold	25°C

8. Once you have all values normalize data by applying the following equation:

$$\text{Enrichment} = \frac{(\text{Experimental} / \text{Control})\text{Ab}}{(\text{Experimental} / \text{Control})\text{IgG}}$$

For each antibody you should use its corresponding IgG. For example, to normalize A + B antibodies samples, use IgG + B. To normalize B + A antibodies samples, use IgG + A; and to normalize sample A and sample B, use input value.

9. Graph data (**Fig. 1**).

3.3. Qualitative PCR

In this case PCR reactions are performed independently in the experimental and control genomic regions, and the signal obtained using the immunoprecipitated DNA as a template is compared with the signal obtained with the unbound DNA as template (**Fig. 2**). Specifications for PCR reactions and visualization are the same as described earlier.

4. Notes

1. The number of cells needed varies according to each cell type. Adjust cell number in order to obtain a good amount of chromatin (for at least ten independent antibodies) per experiment by measuring protein concentration (*see* **Note 3**). Keep cell number constant in each experiment as well as incubation and sonication conditions. Chromatin can be stored at 4°C for up to 4 months.
2. In order to standardize sonication conditions make a pulse curve by taking a sample every 2 pulses (1–15 pulses). Treat chromatin with proteinase K and RNase A (500 μg/mL) for a minimum of 3 h at 65°C. Purify DNA by phenol-chloroform extraction followed by precipitation. Load DNA in a 1% agarose gel and visualize by Ethidium Bromide staining. Choose conditions in which the right size of chromatin fragments is generated.
3. Diluting 1:10 with dilution buffer makes a 10 mL solution in which 1 mL is used per antibody starting from 10–20 × 10^7 cells. If the cell type is changed, quantify protein concentration and dilute with dilution buffer + protease inhibitors to have a 1 μg/μL solution. Store it at 4°C for up to 6 months. Take 400 μL (400 μg) per antibody.
4. Chromatin preclearance with adsorbed beads reduces unspecific background. This step is not essential: if beads and IgG amplification signals are low, it can be skipped.
5. The IgG must be from the same species in which the antibody was generated.
6. 4 μg of antibody is an optimum amount to assure efficient immunoprecipitation. In order to reduce antibody, use 3 and 2 μg and assay for signal. Depending on the antibody 2 μg could be sufficient.
7. The aliquot of unbound material can be used for a comparison with the bound material (chromatin that has specifically bound the antibody and is now precipitated with the beads) (**Fig. 2**). Beads that now carry immunocomplexes settle at the bottom of the tube. When removing the supernatant, be

careful to avoid detaching the beads. When the washing buffers are added, beads will dislodge on their own. Do not vortex.

8. 10 mM DTT is used to elute immunocomplexes from the beads. This can be done in 75–200 μL TE; consider these volumes to dilute 20 times after incubation.
9. Re-ChIP assay makes it possible to know whether two given factors localize at the same genomic region. Is important to include independent immunoprecipitation for each one as controls. In the second round of IP include antibodies A and B in two separate chromatin samples.
10. DNA extraction could be done with phenol-chloroform and be followed by precipitation.
11. When designing experimental and control oligonucleotides consider that the control PCR product must come from a genomic region in which the analyzed factor is not expected. Also be sure both have similar T_m and that the amplification products sizes are adequate to separate them in an acrylamide gel.
12. In order to compare the amplification signal from the genomic region of interest against the control genomic region, the signal obtained with input DNA must be similar for both PCR products. Consider conditions in which Experimental/Control ratio of 0.8–1.2 is obtained. To have a real value of amplification signal, be sure that the PCR cycles used fall into the linear range. To do so, make an amplification curve using increasing amounts of input template (50, 100, 300, 500 ng). Amplification product has to increase linearly. If not, reduce PCR cycles until linear range is achieved.
13. Load samples in the following order: (1) Input, (2) A, (3) IgG + A, (4) B + A (5) Input, (6) B, (7) IgG + B, (8) A + B.
14. ImageQuant (Molecular Dynamics) is one of the programs available to analyze the obtained data. Take the square tool and make a little square around bands. Keep square area constant. Finally display the volume report and save data as an Excel file. Proceed to do the calculation. Other quantification tools can be used.
15. When performing qualitative analysis, control and experimental oligonucleotides are used in separate PCR reactions.

Acknowledgments

We would like to thank Georgina Guerrero for her excellent technical assistance. This work was supported by grants from the Dirección General de Asuntos del Personal Académico-UNAM

(IN209403 and IN214407), Consejo Nacional de Ciencia y Tecnología, CONACyT (42653-Q and 58767), and the Third World Academy of Sciences (TWAS, Grant 01–055 RG/BIO/LA). MF-M and HR-A were fellowship recipients from CONACyT.

References

1. Hatzis, P. and Talinidis, I. (2002). Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol. Cell* **10**, 1467–1477.
2. Rincón-Arano, H., Valadez-Graham, V., Guerrero, G., Escamilla-Del-Arenal, M. and Recillas-Targa, F. (2005). YY1 and GATA-1 interaction modulate the chicken 3'-side α-globin enhancer activity. *J. Mol. Biol* **349**, 961–975.
3. Song, S.H., Hou, C. and Dean, A. (2007). A positive role for NLI/Ldb1 in long-range β-globin locus control region function. *Mol. Cell.* **28**, 810–822.
4. Surani, M.A., Hayashi, K. and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747–762.
5. Spivakov, M. and Fisher, A.G. (2007) Epigenetic signatures of stem-cell identity. *Nat. Rev. Genet* **8**, 263–271.
6. Misteli, T. (2007) Beyond the sequence: cellular organization of genome function. *Cell* **128**, 787–800.
7. Schneider, R. and Grosschedl, R. (2007). Dynamics and interplay of nuclear architecture, genome organization and gene expression. *Genes Dev* **21**, 3027–3043.
8. Bernstein, B.E., Mikkersen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Werning, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L. and Lander, E.S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326.
9. Mikkersen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Werning, M., Jaenisch, R., Nusbaum, C., Lander, E.S. and Bernstein, B.E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560.
10. Rincón-Arano, H., Furlan-Magaril, M. and Recillas-Targa, F. (2007). Protection against telomeric-position effects by the chicken cHS4 β-globin insulator. *Proc. Natl. Acad. Sci. USA* **104**, 14044–14049.
11. Escamilla-Del-Arenal, M. and Recillas-Targa, F. (2008). GATA-1 modulates the chromatin structure and activity of the chicken α-globin enhancer. *Mol. Cell. Biol* **28**, 575–586.
12. Dorsett, D. (1999). Distant liaisons: long range enhancer-promoter interaction in *Drosophila*. *Curr. Opin. Genet. Dev.* **9**, 505–514.
13. Lee, T.I., Johstone, S.E. and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat. Protocol* **1**, 729–748.
14. Euskirchen, G.M., Rozowsky, J.S., Wei, C.-L., Lee, W.H., Zhang, Z.D., Hartman, S., Emanuelsson, O., Stola, V., Weissman, S., Gerstein, M.B., Ruan, Y. and Snyder, M. (2007). Mapping of transcription factor binding regions in mammalian cells by ChIP: comparison of array- and sequencing-based technologies. *Genome Res* **17**, 898–909.

Chapter 18

Profiling Genome-Wide Histone Modifications and Variants by ChIP–chip on Tiling Microarrays in *S. cerevisiae*

Alain R. Bataille and François Robert

Summary

Chromatin immunoprecipitation coupled to DNA microarray has become a widely used method to study transcription factors and chromatin structure. Here, we provide a detailed protocol for the localization of the variant histone Htz1 in the *S. cerevisiae* genome. This protocol can easily be adapted to fit other purposes such as profiling histone modifications.

Key words: Chromatin immunoprecipitation, Location analysis, ChIP on chip, DNA microarrays, Chromatin, Histone modifications, Histone variants.

1. Introduction

Chromatin is an important player in the regulation of nuclear processes such as gene transcription, DNA replication, recombination, and DNA repair. This is possible because chromatin is a very plastic structure in vivo. Chromatin can be remodeled, that is to say that its level of condensation and the position of the nucleosomes can be modulated via the action of cellular factors. In addition, the protein moiety of the nucleosomes, the histones, are subject to a plethora of post-translational modifications that can occur either prior to or after chromatin assembly. Finally, many nonallelic variants exist for most histone genes. The incorporation of these variant into chromatin is another way for the cell to modify chromatin structure and function.

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_18

Tools to study chromatin in vitro have been around for a long time, but the study of chromatin in vivo has long been a challenge. A first generation of techniques to study chromatin in vivo was making use of nucleases *(1–3)*. These assays were based on the principle that nucleosomes protect DNA from the action of nucleases. Information about the composition and the epigenetic state of the nucleosomes cannot be obtained from those assays. The development of antibodies against specific histone modifications, coupled to a technology called Chromatin immunoprecipitation (ChIP), has enabled the possibility to study the epigenetic state of specific DNA loci in vivo *(4)*. More recently, the ChIP assay has been coupled to DNA microarrays, enabling the study of chromatin in large scale *(5, 6)*.

In a ChIP experiment, an immunoprecipitation is made from crosslinked chromatin fragments, and the coimmunoprecipitated DNA is monitored by quantitative or semiquantitative PCR. The enrichment of a given locus in the immunoprecipitated material, compared to a control sample, is read as an indication that the immunoprecipitated protein was bound to that locus in vivo. The control sample is often a nonimmunoprecipitated extract or a mock immunoprecipitation. When the immunoprecipitation is performed using an epitope-tagged protein, the best control is perhaps an immunoprecipitation done on a control cell line that carries no tag *(7)*. When monitoring histone modifications or the presence of a histone variant, however, these kinds of controls are not appropriate. This is because nucleosome density is not uniform along the genome. As a consequence, a high enrichment at a specific locus using an antibody directed against a given histone modification could be the sole consequence of a higher nucleosome density at that locus. Inversely, a low enrichment may be cause by a depletion of nucleosomes at the locus. To circumvent the problem, an immunoprecipitation performed against total nucleosomes is usually used as a control. This normalizes for nucleosome density and a more accurate measure of the modification state of chromatin is therefore achieved.

Here we detail a protocol used to map the variant histone H2A.Z (called Htz1 in yeast) across the Saccharomyces cerevisiae genome. In that experiment, a yeast strain was engineered to express a myc-tagged version of Htz1 and an HA-tagged version of H2B. Immunoprecipitations using anti-myc and anti-HA antibodies are performed from the same crosslinked extract, and both derived DNAs are labeled and hybridized on the same DNA microarray. The H2B (HA) IP is therefore used to normalize nucleosome density. A similar strategy was originally used and published in *(8)*.

2. Materials

2.1. Growing and Crosslinking Cells

1. 50 mL of fresh logarithmic growth yeast culture in YPD. The strain we use has the following genotype (*MATa ura3-1, leu2-3,112, ade2-1, his3-11,15, trp1-1, can1-100, HTZ1::3xMyc, HTB1::3xHA*).
2. 37% Formaldehyde.
3. Ice-cold TBS.

2.2. Breaking Cells

Lysis buffer (50 mM HEPES-KOH of pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 1 mM PMSF, 1 mM Benzamidine, 10 μg/mL Aprotinin, 1 μg/mL Leupeptin, 1 μg/mL Pepstatin).

1. Acid-washed glassbeads (400–600 μm) (Sigma).
2. Bead beater.

2.3. Immunoprecipitation

1. Dynabeads Pan Mouse IgG (Dynal).
2. Anti-Myc antibody (9E10 from Santa Cruz Biotech.).
3. Anti-HA antibody (F-7 from Santa Cruz Biotech.).
4. Hematology Chemistry Mixer 346 (Fisher) or equivalent.
5. Magnet system MPC-S (Dynal).
6. Lysis buffer (Same as for breaking the cells).
7. PBS + BSA (PBS supplemented with 0.5% Bovine Serum Albumin).
8. Wash buffer (10 mM Tris-HCl of pH 8.0, 250 mM LiCl, 0.5% NP40, 0.5% Na-deoxycholate, 1 mM EDTA).
9. TE (10 mM Tris of pH 8.0, 1 mM EDTA).

2.4. DNA Purification

1. TE/SDS (10 mM Tris of pH 8.0, 1 mM EDTA, 1% SDS).
2. TE (10 mM Tris of pH 8.0, 1 mM EDTA).
3. DNase-free RNaseA solution (10 mg/mL in water).
4. 10% SDS.
5. Proteinase K solution (20 mg/mL from Invitrogen).
6. Glycogen (20 mg/mL from Roche).
7. Phenol/chloroform/isoamyl alcohol (25:24:1 from Sigma).
8. 5 M NaCl.
9. 100% freezer-cold ethanol.
10. 70% freezer-cold ethanol.

2.5. Amplification and Labeling

1. T_4 DNA polymerase (NEB).
2. Buffer NEB2.

3. BSA (10 mg/mL from New England Biolabs).
4. dNTP mix (20-mM each of dATP, dCTP, dGTP, and dTTP).
5. 3 M NaOAc of pH 5.2 (Sigma).
6. 0.1 M NaOAc of pH 5.2.
7. Glycogen (20 mg/mL from Roche).
8. Phenol/chloroform/isoamyl alcohol (25:24:1 from Sigma).
9. Freezer-cold 100% ethanol.
10. Freezer-cold 70% ethanol.
11. 5× ligase buffer (Invitrogen).
12. Annealed linker (see protocol later).
13. T_4 DNA ligase high concentration (Invitrogen).
14. 10× ThermoPol buffer (New England Biolabs).
15. 5-(3-aminoallyl)-2'deoxyuridine-5'-triphosphate (aa-dUTP) (Ambion).
16. aa-dUTP dNTP mix (5 mM each of dATP, dCTP, dGTP, 3 mM dTTP, 2 mM aminoallyl-dUTP).
17. Monoreactive Cy-5 NHS Ester (GE Healthcare).
18. Monoreactive Cy-3 NHS Ester (GE Healthcare).
19. Oligo 1 (GCGGTGACCCGGGAGATCTGAATTC) (40 μM stock).
20. Oligo 2 (GAATTCAGATC) (40 μM stock).
21. *Taq* DNA polymerase (Invitrogen).
22. *Pfu* DNA polymerase (Fermentas).
23. Qiaquick PCR purification kit (Qiagen).
24. Phosphate wash buffer (80% ethanol/5 mM KPO_4 of pH 8.5 buffer).
25. Phosphate elution buffer (4 mM KPO_4 of pH 8.5 buffer).
26. 0.1 M Na_2CO_3 of pH 9.0 (from 1 M stock solution made fresh every few weeks to a month).

2.6. Hybridization

1. Ozone-free enclosure.
2. Microarray hybridization chamber (Agilent).
3. Hybridization oven with rotator slide rack (Agilent).
4. *S. cerevisiae* tiling array (Agilent).
5. Microarray gasket slide (Agilent).
6. Hybridization solution (0.05 mM Mes NaOH of pH 6.9, 30% Formamide, 650 mM NaCl, 0.5% sodium N-Lauroyl-sarcosinate, 6 mM EDTA, 250 ng/slide salmon sperm DNA, 80 μg/slide yeast tRNA).

7. Wash solution A (6× SSPE, 0.005% N-Lauroylsarcosine).
8. Wash solution B (0.06× SSPE).
9. Heat blocks at 95°C and 40°C.
10. Three staining dishes.
11. Slide rack.
12. Lab rotator orbital shaker (Barnstead) or equivalent.

3. Methods

3.1. Growing and Crosslinking Cells

In a ChIP experiment, cells are first treated with formaldehyde. Formaldehyde is applied directly to living cells and acts very quickly, which allows for the design of very sharp time course experiments if needed. Virtually any growth condition can be used depending on experimental needs. We typically use 50 mL of exponentially growing cells (i.e., OD_{600}: 0.6–1.0), which represents around 3×10^8–5×10^8 cells. The crosslinking time is one of the first factors that should be optimized if initial attempts to ChIP a protein fail. It has been reported that excessive exposure to formaldehyde can make the subsequent sonication steps more difficult.

1. Grow 50 mL of yeast in YPD medium until OD_{600} reaches 0.6–1.0.
2. Add 1.4 mL 37% formaldehyde directly to the flask or transfer the content of the flask into a 50-mL conical tube containing 1.4 mL formaldehyde.
3. Incubate 30 min at room temperature with agitation.
4. Centrifuge into 50-mL conical tubes for 5 min at 3,000 rpm (2047 g) in a refrigerated (4°C) table top centrifuge.
5. Pour off the supernatant (discard as chemical waste) and wash the cells by resuspending into 40 mL of ice-cold TBS (vigorous shaking is required) and centrifuge again as in **step 4**.
6. Repeat **step 5** for a total of two washes.
7. Using the remaining liquid, resuspend the cell pellet and transfer into a 1.5-mL screw cap tube.
8. Centrifuge 1 min at maximum speed, remove the supernatant with a pipette, and snap freeze in liquid nitrogen.
9. Samples can either be used directly or stored at –80°C.

3.2. Breaking Cells

Yeast cells are broken by vigorous agitation in the presence of glass beads (often referred to as bead beating). We do not recommend using regular vortexers since they tend to break rapidly due

to overuse. We use a bead beater from Biospec Products sold in an 8-tube or 48-tube format.

1. Thaw cell pellet on ice.
2. Resuspend in 700 µL lysis buffer.
3. Add the equivalent of a 0.5-mL PCR tube of glass beads.
4. Screw cap tightly and secure tubes to the bead beater.
5. Beat for 5 min, and then put into ice-water bath for 5 min.
6. Repeat **step 5** three more times (*see* **Note 1**).
7. Punch a hole at the bottom using an 18G1 needle (do not go all the way through. Stop as soon as liquid or air is going through) and set up over a 2-mL tube.
8. Loosen the cap and spin 3–4 s. This should allow the material to transfer into the 2-mL tube while the beads stay trapped into the top tube. This operation may not work with all types of centrifuge. We use an Eppendorf 5417C with a 30-slot rotor.
9. Resuspend the chromatin pellet with its own supernatant and transfer into a new 1.5-mL tube (*see* **Note 2**).

3.3. Preparation of Soluble Chromatin Fragments

Just like for cell breakage, there is a lot of variation in the efficiency of the sonication due to the type of sonicator used, but also the type and age of the probe used. This step therefore needs to be optimized, which can be done by looking at the size of the sheared DNA on an agarose gel. We suggest omitting the formaldehyde treatment when optimizing the sonication in order to remove the need for reversing the crosslinking reaction. The protocol described here was optimized for a Fisher dismembrator 100 equipped with a microprobe.

1. Sonicate the sample four times for 20 s at power 4, 1.5 (for an output of 7 Watts). Put the tube on ice for at least a minute between every round of sonication (*see* **Note 3**).
2. Centrifuge 5 min at maximum speed and transfer the supernatant (which contains solubilized chromatin fragments) to a new 1.5-mL tube. From now on, this will be referred to as the whole-cell extract (WCE).

3.4. Immunoprecipitation

Immunoprecipitation can be done with a great variety of systems. Both monoclonal and polyclonal antibodies have been shown to work for ChIP. All sorts of matrices such as agarose, Sepharose, or magnetic beads were used as well. The protocol described here uses magnetic beads coupled to an anti-myc and anti-HA antibody.

1. Set up a tube with 500 µL of WCE and 30 µL of anti-Myc-coupled-magnetic beads (*see* **Note 4**).

2. Incubate overnight at 4°C with agitation (we use a Hematology Chemistry Mixer 346 from Fisher).
3. Wash the beads twice with 1 mL lysis buffer, twice with 1 mL lysis buffer supplemented with 360 mM NaCl, twice with 1 mL Wash buffer, and once with 1 mL TE (*see* **Note 5**).
4. After the last wash (TE), centrifuge briefly and remove the last bit of liquid using a P200 pipette. This should leave a dry bead pellet.

3.5. DNA Purification

The crosslink between DNA and the proteins needs to be reversed. This is done here by the action of Tris which, with heat, attacks the chemical bonds generated by the formaldehyde and liberates free DNA. The DNA is purified by a succession of RNase A and proteinase K treatments, followed by phenol extractions and ethanol precipitation.

1. To the bead pellets (from **subheading 3.4**, **step 4**), add 50 µL of TE/SDS, vortex, and incubate overnight at 65°C (use an oven instead of a water bath in order to minimize evaporation and condensation in the cap of the tubes).
2. Centrifuge 3 min at 3,000 rpm (300 g) and transfer the supernatant to a new tube.
3. Add 350 µL of a mixture (345 µL TE, 3 µL RNase A, 2 µL glycogen), mix by vortexing, and incubate 2 h at 37°C.
4. Add 15 µL of 10% SDS and 7.5 µL of proteinase K solution, mix by vortexing, and incubate two more hours at 37°C.
5. Extract twice with 400 µL phenol/chloroform/isoamyl alcohol.
6. Add 12–16 µL 5 M NaCl (200 mM final concentration), vortex; add 1 mL ethanol, vortex, and centrifuge 20 min at maximum speed at 4°C.
7. Pour off liquid and wash with 70% ethanol.
8. Centrifuge 5 min and pour off liquid.
9. Centrifuge briefly, remove the last bit of liquid with a pipette, and let dry 1 min (Do not overdry the pellet).
10. Resuspend the pellet (which should be very small but still visible) in 50 µL TE.
11. Store at –20°C.

3.6. Amplification and Labeling

The immunoprecipitated DNA fragments are amplified and labeled by ligation-mediated PCR (LM-PCR) where aminoallyl-modified dUTP is incorporated in the product. The labeling is done post-PCR using monoreactive Cy-dye NHS esters that will react specifically with the aminoallyl-modified dUTP (*see* **Note 6**). The source of reagents is indicated here as a suggestion (*see* **Note 7**).

3.6.1. Blunting

1. Transfer 40 µL of immunoprecipitated DNA to a 0.5-mL PCR tube. Keep the sample on ice.
2. Add 70 µL of blunting mix (11 µL 10× NEB2 buffer, 0.5 µL BSA, 0.5 µL dNTPs, 0.2 µL T_4 DNA pol, and 57.8 µL ice-cold water).
3. Mix by pipetting and incubate 20 min at 12°C (we use a thermocycler block).

3.6.2. Ligation

1. Place on ice and add 12 µL of (11.5 µL 3 M NaOAc and 0.5 µL glycogen).
2. Vortex and add 120 µL of phenol/chloroform/isoamyl alcohol.
3. Vortex hard and centrifuge 5 min at maximum speed.
4. Transfer 110 µL to a new 1.5-mL tube and add 230 µL ethanol.
5. Vortex and centrifuge 20 min at 4°C.
6. Pour off supernatant and add 500 µL 70% ethanol.
7. Vortex and centrifuge 5 min at 4°C.
8. Pour off supernatant, centrifuge briefly, and remove the remaining liquid with a pipette.
9. Let air dry 1 min, add 25 µL ice-cold water, and put on ice for about 30 min.
10. Vortex, centrifuge briefly, and put on ice.
11. Add 25 µL of ligase mix (8 µL water, 10 µL 5× ligase buffer, 6.7 µL annealed linker, 0.5 µL T_4 DNA ligase) (*see* **Note 8** for linker preparation).
12. Mix by pipetting and incubate overnight at 16°C (a water-bath is fine for this step).

3.6.3. PCR

1. Add 6 µL of 3 M NaOAc, vortex, and add 130 µL EtOH.
2. Vortex to mix and centrifuge 20 min at 4°C, maximum speed.
3. Pour off supernatant and add 500 µL 70% ethanol.
4. Vortex and centrifuge 5 min at 4°C, maximum speed.
5. Pour off supernatant centrifuge briefly and remove remaining liquid with a pipette.
6. Let air dry for a minute (do not overdry the pellet), and resuspend in 25 µL ice-cold water by pipetting up and down and rinsing the side of the tube (the pellet at this stage is usually diffused on the side of the tube).
7. Leave on ice for about 30 min, vortex, centrifuge briefly, and put on ice.

8. Add 15 μL of labeling mix (4 μL 10× ThermoPol buffer, 7.75 μL water, 2 μL aa-dUTP dNTP mix, 1.25 μL oligo 1).
9. Mix by pipetting and transfer to a new 0.5-mL PCR tube.
10. Put in thermocycler and start the program (see later).
11. When temperature reaches 55°C, add 10 μL of enzyme mix (8 μL water, 1 μL 10× ThermoPol buffer, 1 μL *Taq* polymerase, 0.01 μL *Pfu* polymerase) and mix by pipetting (*see* **Note 9** for PCR program).
12. Purify PCR reactions on Qiaquick PCR purification kit. Use manufacturer protocol, but replace buffer PE with the Phosphate wash buffer and buffer EB with the Phosphate elution buffer. Do the wash step twice. Elute twice with 30 μL phosphate elution buffer (60 μL total volume).
13. Dry the eluates in a speed-vac.
14. Resuspend the pellets in 4.5 μL fresh 0.1 M Na_2CO_3 of pH 9.0 buffer.
15. Add 4.5 μL of the appropriate Cy-dye (we use Cy5 for the IP and Cy3 for the control) (*see* **Note 10**).
16. Incubate 1 h in the dark at room temperature.
17. Add 35 μL 0.1 M NaOAC of pH 5.2.
18. Add 250 μL buffer PB (Qiagen) and purify on Qiaquick PCR purification columns, using the manufacturer's protocol and buffers. Do the wash step twice. Elute twice with 30 μL buffer EB (60 μL total volume) (*see* **Note 11**).
19. Dry the eluates in a speed-vac (*see* **Note 12**).
20. Store the dry colored pellets at –20°C in the dark.

3.7. Hybridization

Hybridization is done in Agilent hybridization chambers. This hybridization technology ensures a homogenous hybridization of the slide surface. All the process and especially the washes and scanning are done in an ozone-free chamber to minimize the degradation of the Cy-dyes (*see* **Note 13**). More details concerning Agilent chamber manipulation and Agilent hybridization technology can be found in the Agilent microarray hybridization chamber user guide (provided with each chamber or in their web site).

3.7.1. Hybridization

1. Resuspend Cy3-labeled sample in 4 μL of water and transfer to the corresponding Cy5-labeled sample (*see* **Note 14**).
2. Add 450 μL of fresh hybridization solution to each combined sample (volumes from 400 μL to 500 μL will work) and mix by vortexing (*see* **Note 15**).
3. Spin briefly and denature samples for 3 min in a 95°C heat block.

4. Spin briefly and transfer to a 40°C heat block for another 10 min.
5. Assemble the gasket slide and the base of the hybridization chamber, and check for misalignment.
6. Dispense the sample on the gasket slide starting close to one end and slowly moving to the opposite end without touching the surface of the gasket slide (avoid bubbles).
7. Carefully cover with the array and check for misalignment (*see* **Note 16**).
8. Assemble the hybridization chamber upper part.
9. Slip the clamp assembly onto the chamber and tighten the thumbscrew so that it is fully hand tight.
10. Vertically rotate the slide to wet the coverslip and ensure that the bubbles are moving freely (*see* **Note 17**).
11. Incubate in rotating oven at 40°C, 20 rpm, overnight (don't forget to balance the rotor both horizontally and vertically).

3.7.2. Wash

1. Pour buffer A to a standard Staining dish with a slide rack (around 250 mL) and to another dish that will be used as "disassembly dish" (*see* **Note 18**).
2. Remove hybridization chamber from the oven, disassemble all metal parts, submerge slide-gasket slide sandwich into "disassembly dish" and separate the slide from and gasket slide using plastic tweezers (*see* **Note 19**).
3. Briefly rinse the slide in "disassembly dish" and transfer to the slide rack in washing dish with buffer A.
4. Proceed to next slide (*see* **Note 20**).
5. Wash with buffer A on orbital shaker for 5 min (start the timer when the last slide has been transferred).
6. Transfer the slide rack to a dish filled with buffer B and wash on orbital shaker for another 5 min. Slowly remove the rack from the liquid (*see* **Note 21**).
7. Store the slides in dark and proceed to scanning rapidly.

4. Notes

1. The amount of beating that is required may vary depending on the equipment, the yeast strain, the age of the culture, and the sample/beads ratio used. Cell breakage can be easily monitored by looking at cells under a light microscope.

2. Alternatively, the supernatant can be discarded and the chromatin pellet resuspended into 700 μL of fresh lysis buffer. For proteins with a prominent soluble form (not associated with chromatin), this may prevent unbound proteins from competing for antibody during the subsequent immunoprecipitation step.
3. When doing multiple samples, sonicate each sample once and go back to the first one. Rinse the probe with distilled water between every sample in order to avoid crosscontamination. Avoid contact between the probe and the side of the tube. This would reduce the efficiency of the sonication.
4. The beads should be prepared the day prior use the following way. Take enough beads to get 50 μL per immunoprecipitation. Put into a conical plastic tube and centrifuge 3 min at low speed to pellet the beads. Wash twice with PBS + BSA and incubate them overnight at 4°C with agitation in the presence of anti-Myc antibody (we use 2 μg of 9E10 antibody per immunoprecipitation reaction) or anti-HA antibody (we use 2 μg of F-7 antibody per immunoprecipitation reaction.) and enough PBS + BSA to ensure a good agitation. The next day (prior to use) wash twice in lysis buffer to remove nonbound antibodies and resuspend beads in 30 μL of lysis buffer per immunoprecipitation.
5. The washes can be done quickly using a magnetic system (MCP-S from Dynal). Put the tubes to the magnetic system and wait a few seconds for the beads to get captured by the magnet. By keeping the tubes in the system, remove the liquid by aspiration or simply by inverting the tubes. Add the next washing solution, close the tubes, remove the magnet, and agitate to resuspend the beads. Put the magnet back on and remove the liquid as earlier. Process as so for the necessary steps.
6. This method is less expensive and yields better results than the direct incorporation of Cy-dye-modified dNTPs during PCR. We found that dye swap experiments are not necessary with this protocol.
7. These reagents represent a combination that gives good results in our laboratory. While it probably does not matter for most of other reagent, we found that using ligase buffers which contains polyethylene glycol and the high concentration ligase makes an improvement.
8. The annealed linker is prepared by annealing oligos 1 and 2 the following way. Mix 250 μL 1 M Tris of pH 7.9, 375 μL oligo 1, and 375 μL oligo 2. Make 50-μL aliquots in 1.5-mL tubes. Place in a 95°C heat block for 5 min. Transfer tubes to a 70°C heat block with water in wholes. Place the block

at room temperature and let it cool down to 25°C. Transfer the block to 4°C and allow it to cool down overnight. Store annealed linkers at –20°C. The T_m of the linker is very low. Always keep the linkers on ice and use ice-cold buffers when manipulating it to avoid deannealing.

9. Use the following PCR program for amplification: step 1: 4 min at 55°C; step 2: 5 min at 72°C; step 3: 2 min at 95°C; step 4: 30 s at 95°C; step 5: 30 s at 55°C; step 6: 1 min at 72°C; step 7: Go back to step 4, 31 more times; step 8: 4 min at 72°C; step 9: 4°C forever.
10. Resuspend new tubes of the Cy3 and Cy5 dyes in 73 µL of DMSO (kept in the dark, under desiccant and vacuum). Store at –80°C.
11. The color should be visible on the filter and the eluate should have a faint but visible taint of blue/pink.
12. Optionally DNA amount and dye incorporation can be estimated by spectrometry prior speed-vac using 50 µL of the undiluted elution (which is recovered from the cuvet after measurement). For this we measured the absorbance at 260, 320, 550, 650 nm for Cy3 and 260, 320, 650, 750 nm for Cy5 and apply the following formulas:

$$\text{DNA}(\mu\text{g}) = (A_{260} - A_{320}) \times 50\mu\text{g/mL} \times \text{Volume (mL)} \times \text{dilution factor}$$

$$\text{Base/Cy3} = (((A_{260} - A_{320}) - ((A_{550} - A_{650}) \times 0.08)) \times 150{,}000)/((A_{550} - A_{650}) \times 6{,}600)$$

$$\text{Base/Cy5} = (((A_{260} - A_{320}) - ((A_{650} - A_{650}) \times 0.05)) \times 250{,}000)/((A_{650} - A_{750}) \times 6{,}600)$$

$$\text{Dye molecules/100bp} = 100/(\text{base/dye})$$

Expect about 5 µg of DNA (typically between 4 and 6 µg) and an incorporation of 1.5 dyes per 100 bp (typically between 1.2 and 1.8 dyes per 100 bp).

13. Cyanide dyes (Cy5 particularly) are sensitive to even low level of atmospheric ozone. Working under an ozone-free enclosure is especially important when the dyes are dry and thus directly exposed to atmospheric ozone (after the last wash). Alternatively Agilent sells a solution saturated with ozone scavengers that will help protect the slide from ozone during the scan.
14. Samples are easily resuspended in a small volume, but a larger volume can be used (adjust water volume in hybridization solution in consequence).
15. The hybridization solution is made fresh each time. While it probably does not matter for most of other reagent, we

found that using Sodium N-Lauroylsarcosinate from Fluka makes an improvement compared to other brands.

16. The active face of the array is the face reading "Agilent"; thus, this face should face the gasket slide while you are facing the side reading the slide number.
17. It is ideal to obtain just one "mixing" bubble; however, small bubbles are fine as long as they can move freely between the slides. Static bubbles at the edge of the gasket can be dislodged by gently tapping the corner of the assembled chamber on a firm surface.
18. We use a glass staining dish for the disassembly dish, but black plastic ones for the subsequent washes in order to protect the slides from the light as much as possible.
19. Insert one end of the tweezers between the slide and the gasket slide in the barcode area and then gently turn the tweezers to separate the slides. Let the gasket slide drop to the bottom of the dish while you are still maintaining the array with the tweezers. To transfer the slide into the rack it is easier and safer to manipulate the slide by holding it at the barcode area with your hand.
20. Up to ten slides can be processed in parallel in the same slide rack.
21. It is important to slowly remove the slides from the liquid. The surface of the slides will come out dry and usually do not need to be centrifuged or air blowed.

References

1. Wu, C. (1989). Analysis of hypersensitive sites in chromatin. *Methods Enzymol.* **170**, 269–289.
2. Bellard, M., Dretzen, G., Giangrande, A. and Ramain, P. (1989). Nuclease digestion of transcriptionally active chromatin. *Methods Enzymol.* **170**, 317–346.
3. Lutter, L.C. (1989). Digestion of nucleosomes with deoxyribonucleases I and II. *Methods Enzymol.* **170**, 264–269.
4. Kuo, M.H. and Allis, C.D. (1999). In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. *Methods* **19**, 425–433.
5. Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. *Cell* **117**, 721–733.
6. Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D.K. and Young, R.A. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517–527.
7. Drouin, S. and Robert, F. (2006). Genome-wide location analysis of chromatin-associated protein by ChIP on CHIP: Controls Matter. *Methods* (in press).
8. Guillemette, B., Bataille, A.R., Gévry, N., Adam, M., Blanchette, M., Robert, F. and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol.* **3**, e384.

Chapter 19

Nucleosome Mapping

Nicolas Gévry, Amy Svotelis, Marc Larochelle, and Luc Gaudreau

Summary

The basic repeating unit of chromatin, the nucleosome, is known to play a critical role in regulating the process of gene transcription. The positioning of nucleosomes on a promoter is a significant determinant in its responsiveness to gene-inducing signals. For example, positioning and subsequent mobilization of nucleosomes can regulate the access of various DNA factors to underlying DNA templates. Several mechanisms have been proposed to direct the process of nucleosome displacement such as chemical histone modifications, ATP-dependent remodelling, and the incorporation of histone variants. Thus, rather than being an inert molecular structure, chromatin is highly dynamic in response to the transcription process. In this section, we describe two methodologies that allow the determination of exact nucleosome positioning within specific gene regions.

Key words: Transcription, Chromatin, Nucleosomes, LM-PCR; Nucleosome density assay.

1. Introduction

In eukaryotes, nuclear chromosomes are packaged by proteins into a condensed structure called chromatin. Chromatin is the physiological template for nuclear processes involving genomic DNA, including gene transcription. The organization of DNA into chromatin is separated into several levels, achieved via the use of several different proteins and complexes. The first level of chromatin organization is defined by the nucleosome, the basic repeating unit of chromatin, composed of an octamer of the four core histones (H2A, H2B, H3, and H4) around which 147 bp of DNA are wrapped *(1)*. Polynucleosomal arrays are visualized as "beads on a string" in microscopy where nucleosome particles are the "beads" connected by 20–100 bp of linker DNA *(2)*.

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_19

Packaging genomic DNA into chromatin presents a physical barrier for regulatory proteins such as transcription factors and general transcriptional machinery to have access to the underlying promoter DNA region *(3)*.

In order to bypass the physical barrier chromatin presents, several mechanisms have been proposed to alter chromatin structure during transcription. These modifications can lead either to nucleosome sliding, nucleosome eviction, or to nucleosome modification. Nucleosome stability is regulated by the combined effects of nucleosome-positioning sequences, histone chaperones, ATP-dependent nucleosome remodelling complexes, post-translational modifications, and histone variants *(4–7)*. These mechanisms are dedicated to repressing or facilitating events that occur on the DNA template and maintaining the nucleosomal structure organization. The high degree of precision in setting up a chromatin template and small changes in nucleosome position can have dramatic consequences on recruitment of regulatory proteins and on transcription *(8–10)*. Recently, it has become evident that the association of regulatory proteins, the position of nucleosomes, and their composition are highly dynamic *(3)*. Thus rather than being an inert molecular structure upon which transcription occurs, chromatin structure also plays a more specific role in transcriptional regulation.

Over the last few years, chromatin research efforts have made it increasingly apparent that chromatin structure imposes profound and ubiquitous effects on almost all DNA-related processes, namely transcription. Thus, the characterization of nucleosome positions is of profound interest in the study of how genes are regulated. The techniques described in this section will help to analyse the consequences that the dynamic nature of nucleosomes has on chromatin-based processes such as transcription. These techniques take advantage of the enzyme micrococcal nuclease (MNase), which preferentially digests the internucleosomal (linker) DNA region of chromatin before nucleosome-protected DNA.

2. Materials

2.1. Cell Preparation

1. PBS 1×: 137 mM NaCl, 27 mM KCl, 43 mM KH_2PO_4, and 14 mM $Na_2HPO_4{\cdot}7H_2O$. For the 1× PBS dilute 10× PBS 1:9 in ddH_2O.
2. Formaldehyde, 37%.

3. 100× Protease inhibitor cocktail: 0.2 mM Pepstatin A, 72 μM Leupeptin, and 26 μM Aprotinin in 90% ethanol.
4. Prepare fresh 1.1% PBS-formaldehyde mix from 37% formaldehyde and PBS of pH 7.4.
5. Buffer A: 10 mM Hepes of pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10% Glycerol, 1× protease inhibitor cocktail, and 1 mM DTT.
6. Buffer B: 3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, 1× protease inhibitor cocktail, and 1 mM PMSF.
7. MNase buffer: 50 mM Tris–HCl of pH 7.4, 25 mM KCl, 4 mM $MgCl_2$, 12.5% Glycerol, 1 mM $CaCl_2$.
8. 2.5 M glycine.
9. Buffer Z: 1 M sorbitol, 50 mM Tris–HCL of pH 7.4.
10. Zymolyase solution: 10 mg/mL in Buffer Z. Prepare fresh for each experiment.
11. NP buffer: 1 M Sorbitol, 50 mM NaCl, 10 mM Tris–HCL of pH 7.4, 5 mM $MgCl_2$, 1 mM $CaCl_2$.
12. 250 mM Spermidine.
13. β-Mercaptoethanol (14.3 M).
14. 10% NP-40.

2.2. MNase Digestion and DNA Extraction

1. Micrococcal nuclease (MNase) (USB, Cleveland, OH, USA).
2. Stop solution: 100 mM EDTA, 100 mM EGTA.
3. RNAse A: DNAse-free stock at a concentration of 10 mg/mL.
4. Proteinase K (from a 10 mg/mL stock).
5. Glycogen (from a 20 mg/mL stock).
6. 100% Ethanol.
7. 80% Ethanol.
8. 3 M Sodium acetate of pH 5.5.
9. 5× DNA-loading buffer: 25 mg xylene cyanol (0.25%), 4 g sucrose (40%); adjust volume to 10 mL with H_2O.
10. 50× TAE: 2 M Tris–HCL, 2 M acetic acid, and 0.05 M EDTA (pH 8.0). For the 0.5× TAE, dilute 1:100 in ddH_2O.
11. 10 mM Tris–HCl of pH 7.5.
12. Ethidium bromide (10 mg/mL).
13. QIAquick PCR purification kit (Qiagen).

2.3. Nucleosome Density Assay (NDA)

1. qPCR commercial master mix.
2. 10 mM dNTP.
3. Gene-specific oligonucleotides.

2.4. High-Resolution Ligation-Mediated PCR (LM-PCR) Analysis

1. Oligonucleotides for linker:

   Upper strand (US1): 5′-GCGGTGACCCGGGAGAGATCT-GAAT TC-3′

   Lower strand: 5′-P-GAATTCAGATCT-3′.
2. T_4 DNA ligase.
3. Gene-specific oligonucleotides (GS1 and GS2).
4. 10 mM dNTP.
5. *Taq* enzyme (Invitrogen).
6. [α-^{32}P] dATP.
7. 2× formamide DNA-loading buffer: 20 mM EDTA of pH 8.0, 0.05% xylene cyanol, prepared in 95% formamide.

3. Method

3.1. Preparation of Cells

3.1.1. Mammalian Cells

This protocol is described for cultured cells grown in 150-mm dishes, containing around 4×10^7 cells per dish (*see* **Note 1**).

1. Remove the medium and wash the cells once with PBS.
2. Add 20 mL of 1.1% PBS-formaldehyde mix at room temperature.
3. Incubate for 15 min at room temperature.
4. Add glycine to final concentration of 125 mM.
5. Wash the cells twice with cold 1× PBS.
6. Harvest cells using a cell scraper, spin down at 500 × *g* for 5 min at 4°C, and discard supernatant (*see* **Note 2**).
7. Resuspend the cell pellet in 1 mL of cold buffer A.
8. Add Triton X-100 to a final concentration of 0.1%.
9. Incubate cells on ice for 10 min (*see* **Note 3**).
10. Centrifuge at 1,300 × *g* for 5 min at 4°C.
11. Wash the pellet with 1 mL of buffer A and centrifuge at 1,300 × *g* for 5 min at 4°C.
12. Resuspend the pellet in 1 mL of cold buffer B and incubate for 30 min on ice (*see* **Note 4**).
13. Centrifuge at 1,700 × *g* for 5 min at 4°C.
14. Resuspend the pellet in MNase buffer to obtain an OD_{260} of 0.2 per 5 µL sample and keep on ice. OD_{260} is measured by diluting 5 µL of chromatin in 100 µL 1 M NaOH.

3.1.2. Yeast Cells

1. Inoculate 5–10 mL of the desired culture media with the desired cells (*see* **Note 5**).

2. Inoculate 450 mL of media in a 2-L flask with the required amount of preculture in order to obtain the proper OD_{600} for the following morning. Grow the cells at 30°C in a shaking incubator set at 225 rpm overnight (*see* **Note 5**).
3. When you have reached an OD_{600} of around 0.8, add formaldehyde to a final concentration of 2%. Incubate 30 min at 30°C in a shaking incubator.
4. Stop the crosslink reaction by adding glycine to a final concentration of 125 mM.
5. Transfer the culture into a centrifuge bottle and spin for 10 min at 4,500 × *g*.
6. Wash the cells once with 50 mL of ddH20 and centrifuge at 3,000 × *g* for 5 min.
7. Resuspend each cell pellet (from 450 mL of cells) in 39 mL of Buffer Z. Add β-mercaptoethanol to a final concentration of 10 mM. Vortex cells to complete resuspension.
8. Add 1 mL Zymolyase and incubate at 30°C in a shaking incubator at 125 rpm for 30–35 min (*see* **Note 6**).
9. Meanwhile, make NP buffer with β-mercaptoethanol NP-40 and spermidine: 5 mL of NP buffer with of 500 μM spermidine, β-mercaptoethanol to a final concentration of 1 mM and 0.075% NP-40.
10. Following zymolyase digestion, pellet cells for 10 min at 5,000 × *g* at 4°C. Remove supernatant (*see* **Note 7**).
11. Resuspend pellet in NP buffer to obtain an OD_{260} of 0.2 per 5 μL sample (keep on ice). OD_{260} is measured by diluting 5 μL of chromatin in 100 μL 1 M NaOH.

3.2. MNase Digestion

1. Prepare Eppendorf tubes with different amounts of micrococcal nuclease from 0 to 200 U/mL (e.g. 0, 25, 50, 100, 150, 200 U/mL) (*see* **Note 8**).
2. Add 500 μL of the resuspended cells at a concentration of 0.2 OD_{260}.
3. Incubate for 10 min at 37°C (*see* **Note 9**).
4. Stop the reaction by the addition of equal volume of stop solution.
5. Reverse the crosslinked DNA by incubation at 65°C for at least 4 h (or overnight).

3.3. DNA Isolation

1. Add 2 μL of RNase A and incubate at 37°C for 1 h.
2. Add 2 μL of proteinase K and incubate at 37°C for 2 h.
3. Add one volume of phenol/chloroform/isoamylalcohol solution. Vortex and centrifuge at full speed for 5 min. Repeat this extraction twice, always keeping the aqueous phase.

4. Extract once by adding one volume of chloroform. Vortex and centrifuge at full speed for 5 min.
5. Precipitate the DNA by adding 1 μL of glycogen, 1/10 volume sodium acetate, and two volumes of 100% ethanol.
6. Vortex and incubate overnight at 20°C.
7. Centrifuge at 15,000 × *g* 15 min at 4°C.
8. Wash the precipitated DNA with 1 mL of 80% ethanol and centrifuge at 15,000 × *g* for 5 min.
9. Air dry the pelleted DNA for 10 min at room temperature.
10. Resuspend the DNA in 20 μL of H_2O.
11. Prepare a 1.2% agarose gel for a horizontal migration with 0.5× TAE.
12. Add DNA-loading buffer to the digested DNA (*see* **Note 10**).
13. Electrophorese at 100 V until the xylene cyanol dye reaches one-quarter the length of the gel.
14. Stain for 15 min with two gel volumes of 0.75 μg/mL ethidium bromide in water and destain for 1 h using several changes of water.
15. Examine the micrococcal digestion under a UV transilluminator and use the best MNase titration (the one containing the most mononucleosomal DNA) and cut the region of the gel corresponding to the mononucleosome-sized DNA (*see* **Note 11**).
16. Use QIAquick PCR columns to purify the DNA according to the manufacturer's protocol.
17. Elute in 50 μL of sterile ddH_2O.

3.4. Analysis

3.4.1. Analysis by Nucleosome Density Assay

The nucleosome density assay (NDA) can be used to accurately map the position of nucleosomes at a region of interest. This approach has been used very successfully in yeast and mammalian cells, and has the advantage of being extremely rapid and accurate *(11, 12)*. Furthermore, this assay provides a high-quality estimate of the "fuzziness" or stability of a nucleosome over a given area. The mononucleosome-length DNA is amplified by quantitative PCR with primers amplifying ~100 bp (±10 bp) amplicons overlapping by ~70 bp, thus leading to a resolution of the assay of 30 bp (**Fig. 1**).

The NDA assay is based on the fact that highly positioned nucleosomes will protect the corresponding DNA region from MNase digestion, leading to higher PCR amplification rates compared to regions with diffuse nucleosome positioning or devoid of nucleosomes. Amplicons located within a perfectly positioned nucleosome will yield a maximal histone density that can be defined as 100%. On the other hand, if one primer of the amplicon

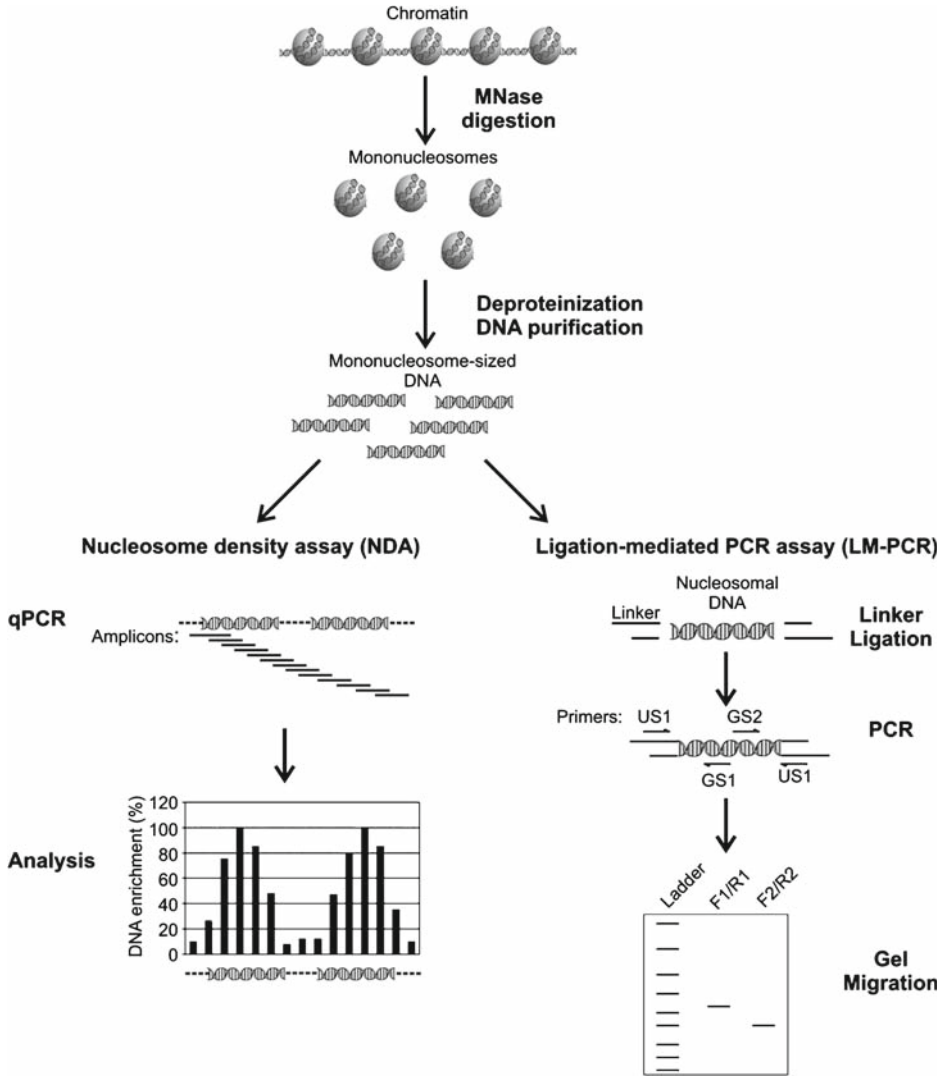


Fig. 1. General outline of the nucleosome mapping procedures.

is outside the nucleosome (in the linker region) the nucleosome density will be low. Thus, a positioned nucleosomal region will generate peaks and valleys, with the valleys corresponding to linker regions between nucleosomes. Furthermore, randomly positioned nucleosomes region should display a relatively constant level of nucleosome density that is around 20–40% of the maximal level. Finally, a region depleted in nucleosomes should display low density over the entire region (**Fig. 1**).

1. Set up quantitative PCR reaction using the method described by the commercial mix (*see* **Note 12**).

2. The resulting threshold amplification level of genomic DNA (gDNA) samples (CT) is compared to the mononucleosome DNA (mono) by subtraction ($CT^{gDNA} - CT^{mono} = \Delta CT$).

For a relative DNA level, we calculate $2^{-\Delta CT}$ for all samples.

3.4.2. Analysis by High-Resolution LM-PCR

This method allows determining subtle changes in nucleosome position that cannot be detected by the NDA assay *(8)*. The disadvantage of this technique is that we need to know to some extent the gross position of nucleosomes. Briefly, we need to ligate linker oligonucleotides to both ends of the mononucleosome-sized DNA that will be used with specific oligonucleotides of the target sequence (*see* **Fig. 1**).

1. Dissolve the oligonucleotides in sterile nuclease-free ddH_2O to a concentration of 100 μM.
2. Combine 20 μM of each oligonucleotide in 250 mM Tris–HCl of pH 7.7 and heat for 5 min at 95°C.
3. Incubate the mixture at 90°C for 4 min and then at 70°C for 10 min. Slowly cool the annealed oligos to 4°C (*see* **Note 13**).
4. Prepare the ligation reaction using mononucleosome-sized DNA (around 1 μg), 25 μM linker, and 5 U of DNA ligase in a total volume of 30 μL.
5. Incubate the reaction overnight at 16°C.
6. To stop the ligation reaction incubate at 70°C for 5 min.
7. To each tube, add two volumes of ethanol 100%, 1/10 volume of sodium acetate, and 1 μL of glycogen.
8. Vortex and allow to precipitate at 20°C overnight.
9. Decant supernatant and wash the pellet in 1 mL 80% ethanol.
10. Dissolve DNA in 100 μL of 10 mM Tris–HCl of pH 7.5.
11. Set up a PCR reaction using 2 μCi of [α-32P] dATP with 2 μL of DNA, 200 nM each of the Upper strand (F1) and specific gene oligonucleotides, and 200 μM dNTP (*see* **Note 14**).
12. Denature PCR products (2–4 μL) along with 10-bp ladder mixed in 2× loading buffer for 5 min at 95°C (*see* **Notes 15** and **16**).
13. Analyse the reactions using gel electrophoresis on an 8% denaturing polyacrylamide sequencing gel.
14. Transfer the gel onto Whatman paper and dry using a gel dryer.
15. Expose the dried gel on a phosphoimager screen for at least 2 h and analyse the frontiers of the nucleosome targeted.

4. Notes

1. This protocol for mammalian cells has been used successfully with the MCF-7 cell line. However, if you are using different cell line some optimization will need to be to increase the efficiency and the quality of this chromatin extraction. Critical parameters are cell number, the amount of lysis buffer, and lysis time.
2. At this point the samples can be snap frozen in liquid nitrogen and kept in a −80°C freezer for extended storage.
3. The incubation time with 0.1% Triton can be increased to 20–30 min. One characteristic of incomplete digestion is the presence high molecular weight DNA in higher MNase concentrations.
4. If necessary this time of incubation can be increased, but usually 30 min is sufficient.
5. The culture time and the amount of media for the preculture and overnight culture need to be optimized with the yeast strain that you will use. Information concerning the generation time will be important in planning your yeast culture.
6. Formation of the spheroplast can be followed by spectrophotometer (OD_{600}) by adding 40 μL of cells in 1 mL of 1% SDS every 10–15 min. Spheroplast formation is completed when the OD is less than 5% of the starting value.
7. You need to aspirate carefully the supernatant since spheroplast pellets have a fluffy appearance. We generally try to remove almost all the supernatant while losing only a small amount of the spheroplasts.
8. Usually this range of MNase is sufficient to obtain mononucleosome-sized fragments, but may need to be optimized for the particular cell type and experiment. Also, MNase activity can vary according to company and lot.
9. If you perform many digestions in same experiment, it is important to start the reaction at different time intervals to achieve the same digestion time. The digestion time has to be determined experimentally and can be different for each cell type and each MNase preparation.
10. It is important to use a dye in the DNA-loading Buffer that will not interfere with the visualization of the DNA. The best dye to use is xylene cyanol, since it migrates around 8–10

kb as compared to the bromophenol blue which migrates around the mononucleosome-sized DNA.

11. Since mononucleosome-sized DNA migrates at a low molecular weight and ethidium bromide migrates in the opposite direction of the DNA, it is sometime difficult to visualize the DNA with a gel that already contains ethidium bromide. Best results are obtained with staining after the migration.
12. Use the freeware Primer3 on the internet (http://www.frodo.wi.mit.edu) to design the oligonucleotides for the NDA assay, using an annealing temperature around 60°C.
13. Step cool to 37°C for 15–20 min and incubate for 15–20 min at room temperature. The annealed oligonucleotides can be used immediately in the ligation reaction or cooled further to 4°C. For extended storage, keep at 20°C.
14. Nucleosome-specific primers should be 20–25 nucleotides in length and have a GC content of 40–60%. This will ensure that the primers will effectively anneal to the template at an annealing temperature around 62–65°C. Primers should not be able to fold back and form intramolecular hydrogen bonds, and sequences at the 3 end of your primers should not be able to anneal to the 3 end of the adaptor primers. There should be no more than three G and C in the last six positions at the 3 end of the primer.
15. For preliminary and optimization experiments, you may use a small vertical gel electrophoresis system (such as the apparatus used for western blots) with a non-denaturing acrylamide gel. Furthermore, you may also omit the use of radioactivity and stain the gel with ethidium bromide as described in **Subheading 3.3**.
16. DNA sequencing reaction can be used as radioactive ladder to obtain the exact frontier of the nucleosome.

Acknowledgements

The authors would like to thank Benoît Leblanc for his help in the artwork included in **Fig. 1**.

References

1. Luger, K. (2003). Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev* **13**, 127–135.
2. Olins, A. L., and Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science* **183**, 330–332.
3. Mellor, J. (2005). The dynamics of chromatin remodeling at promoters. *Mol Cell* **19**, 147–157.
4. Eberharter, A., and Becker, P. B. (2004). ATP-dependent nucleosome remodelling: factors and functions. *J Cell Sci* **117**, 3707–3711.
5. Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693–705.
6. Rando, O. J., and Ahmad, K. (2007). Rules and regulation in the primary structure of chromatin. *Curr Opin Cell Biol* **19**, 250–256.
7. Workman, J. L. (2006). Nucleosome displacement in transcription. *Genes Dev* **20**, 2009–2017.
8. Guillemette, B., Bataille, A. R., Gevry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol* **3**, e384.
9. Lomvardas, S., and Thanos, D. (2001). Nucleosome sliding via TBP DNA binding in vivo. *Cell* **106**, 685–696.
10. Lomvardas, S., and Thanos, D. (2002). Modifying gene expression programs by altering core promoter chromatin architecture. *Cell* **110**, 261–271.
11. Metivier, R., Penot, G., Hubner, M. R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003). Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**, 751–763.
12. Sekinger, E. A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol Cell* **18**, 735–748.

Chapter 20

In Cellulo DNA Analysis (LMPCR Footprinting)

Régen Drouin, Nathalie Bastien, Jean-François Millau, François Vigneault, and Isabelle Paradis

Summary

The in cellulo analysis of DNA protein interactions and chromatin structure is very important to better understand the mechanisms involved in the regulation of gene expression. The nuclease-hypersensitive sites and sequences bound by transcription factors often correspond to genetic regulatory elements. Using the Ligation-mediated polymerase chain reaction (LMPCR) technology, it is possible to precisely analyze these DNA sequences to demonstrate the existence of DNA–protein interactions or unusual DNA structures directly in living cells. Indeed, the ideal chromatin substrate is, of course, found inside intact cells. LMPCR, a genomic-sequencing, technique that map DNA single-strand breaks at the sequence level of resolution, is the method of choice for in cellulo footprinting and DNA structure studies because it can be used to investigate any complex genomes, including human. The detailed conventional and automated LMPCR protocols are presented in this chapter.

Key words: Footprints, Ligation-Mediated Polymerase Chain reaction, Polymerase chain reaction DNA polymerase, Living cell, Deoxyribonuclease I, Dimethylsulfate, Ultraviolet light C, DNA–protein interaction.

1. Introduction

The in cellulo analysis of DNA protein interactions and chromatin structure can provide several critical information regarding regulation of gene expression. For example, DNA sequences spanned by nuclease-hypersensitive sites or bound by transcription factors often correspond to genetic regulatory elements. Using the Ligation-Mediated Polymerase Chain Reaction (LMPCR) technology, it is possible to map DNA sequences to

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_20

demonstrate the existence of DNA–protein interactions or unusual DNA structures directly in living cells. LMPCR analyses can thus be used as a primary investigative tool to identify the regulatory sequences involved in gene expression. Once specific promoter sequence sites shown to be bound by transcription factors in living cells, it is often possible to establish the identity of these factors simply by comparison with the consensus binding sites of known factors such as SP1, AP-1, NF-1, and so forth. The identity of each factor can then be confirmed by performing the well-established Chromatin immunoprecipitation (ChIP) (*see* Chap. "Chromatin Immunoprecipitation in Mammalian Cells") technique, or by using silencing RNA directed against the protein of interest.

The native state of a gene and most of the special DNA structures are unavoidably lost when DNA is cloned or purified *(1–4)*. Hence, the commonly used in vitro (purified DNA) footprinting and Electrophoretic Mobility Shift Assay (EMSA) cannot demonstrate that a given DNA–protein interaction occurs within the cells of interest. It is clear that gene promoters are best studied in their natural state in the living cell. Thus, it is not surprising that in cellulo (living cell) DNA footprinting is one of the most accurate predictors of the state of gene transcriptional activities *(1, 3, 5)*. LMPCR is the method of choice for in cellulo footprinting and DNA structure studies because it can be used to investigate complex animal genomes, including human. However, the quality and usefulness of the information obtained from any in cellulo DNA analysis depends on three parameters: (1) the integrity of the native chromatin substrate used in the experiment, (2) the structural specificity of the chromatin probe, and (3) the sensitivity of the assay. The ideal chromatin substrate is, of course, found inside intact cells. However, a near-ideal chromatin substrate can still be found in permeabilized cells, allowing the application of a wider range of DNA cleavage agents, including DNase I.

In cellulo footprinting assesses the local reactivity of probing agents on living cells DNA as compared to that on purified DNA (*see* **Figs. 1–4**). Two steps are necessary for in cellulo footprinting analysis: (1) the treatment of purified DNA and intact cells with a given probing agent followed, if necessary, by DNA damage conversion into single-strand breaks and (2) the analysis of amplified DNA fragments on a sequencing gel. A comparison is then made between the modification frequency obtained in vitro and in cellulo. For example, each guanine residue of purified DNA has a near-equivalent probability of being methylated by dimethylsulfate (DMS) and thus, the cleavage pattern of in vitro modified DNA appears on a sequencing gel as a ladder of bands of roughly equal intensity. However, in presence of DNA-binding proteins (transcription factors), all guanine residues do not show the same accessibility to DMS in living cells (**Fig. 1**). Thus, differences between banding patterns obtained from in vitro and in cellulo

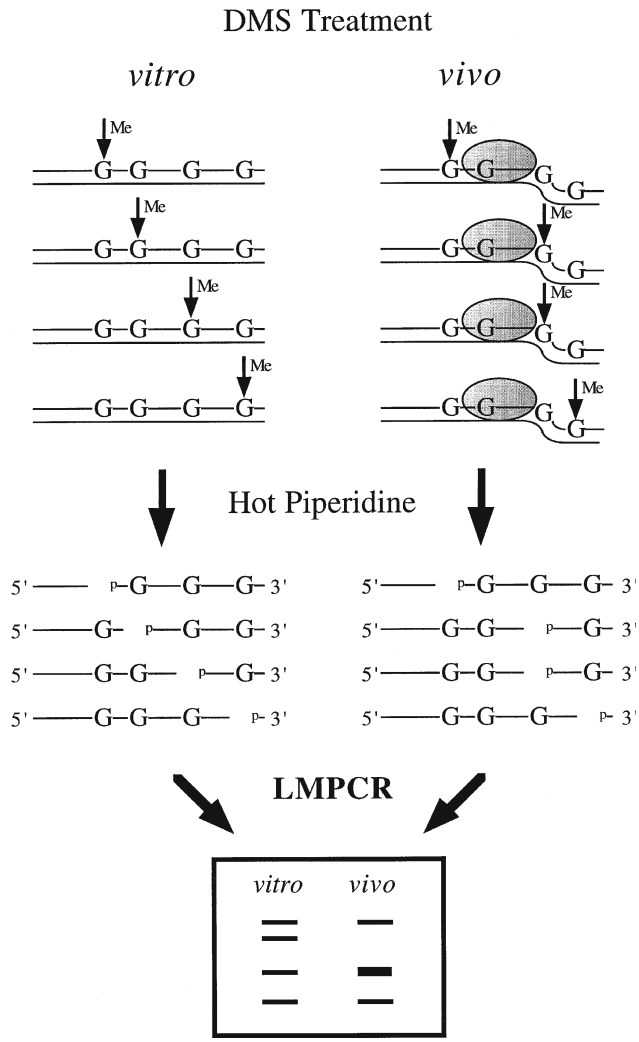


Fig. 1. Overall scheme for in vivo DNA analysis using DMS. The methylation of guanine residues following DMS treatment of purified DNA (in vitro) and cells (in vivo) is shown with vertical arrows and methylated residues (Me). When purified DNA is treated with DMS, every guanine residue has a similar probability of being methylated. However, the guanine residue in intimate contact with a sequence-specific DNA-binding protein illustrated by the dotted oval is protected from DMS methylation, whereas the guanine residue localized close to the boundary of the DNA–protein contact that modifies DNA structure, allowing a better accessibility to DMS, is methylated more frequently. The methylated guanine residues are cleaved by hot piperidine leaving phosphorylated 5′ ends. On the sequencing ladder following LMPCR, guanine residues that are protected from methylation appear as missing or less intense bands when compared with the sequencing ladder from the same DNA sequence obtained after DMS treatment of purified DNA. On the other hand, guanine residues that undergo enhanced DMS methylation appear as darker bands in the sequencing ladder relative to the purified DNA control.

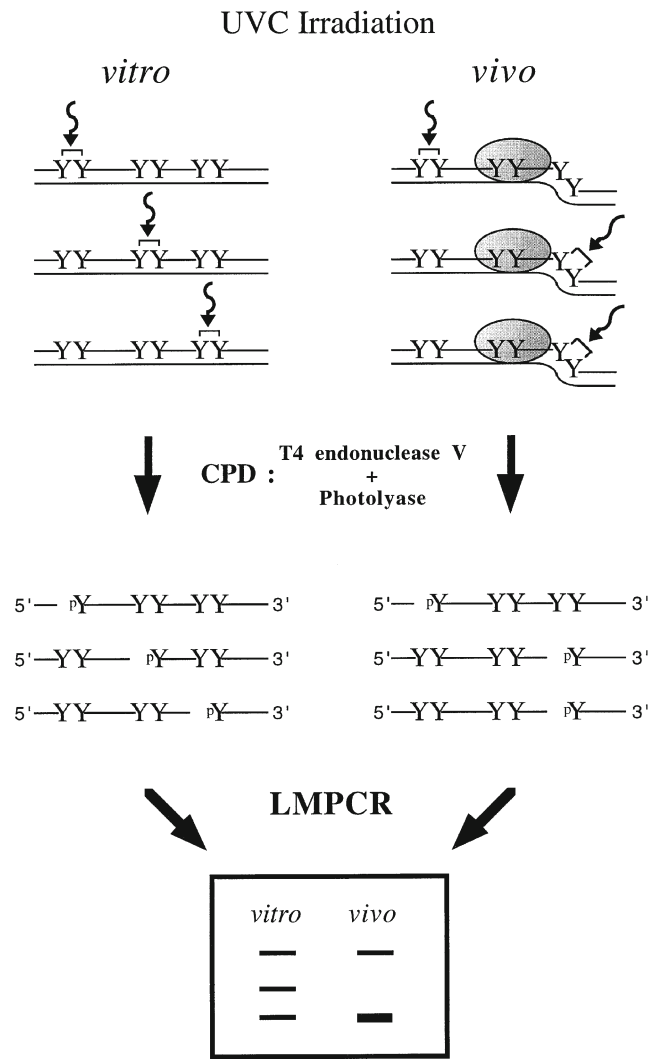


Fig. 2. Overall scheme for in vivo DNA analysis using UVC and CPD formation. The CPD formation following UVC exposure of purified DNA (in vitro) and cells (in vivo) is shown with curved arrows and brackets linking two adjacent pyrimidines (Y). When purified DNA is irradiated with UVC, the frequency of CPD formation at dipyrimidine sites is determined by the DNA sequence. However, the presence of a sequence-specific DNA-binding protein illustrated by the dotted oval as well as DNA structure can prevent (negative photofootprint) or enhance (positive photofootprint) CPD formation. The CPD are cleaved by T_4 endonuclease V digestion and photolyase photoreactivation leaving phosphorylated 5′ ends. On the sequencing ladder following LMPCR, the negative photofootprints appear as missing or less intense bands when compared with the sequencing ladder from the same DNA sequence obtained after UVC irradiation of purified DNA. On the other hand, positive photofootprints appear as darker bands in the sequencing ladder relative to the purified DNA control.

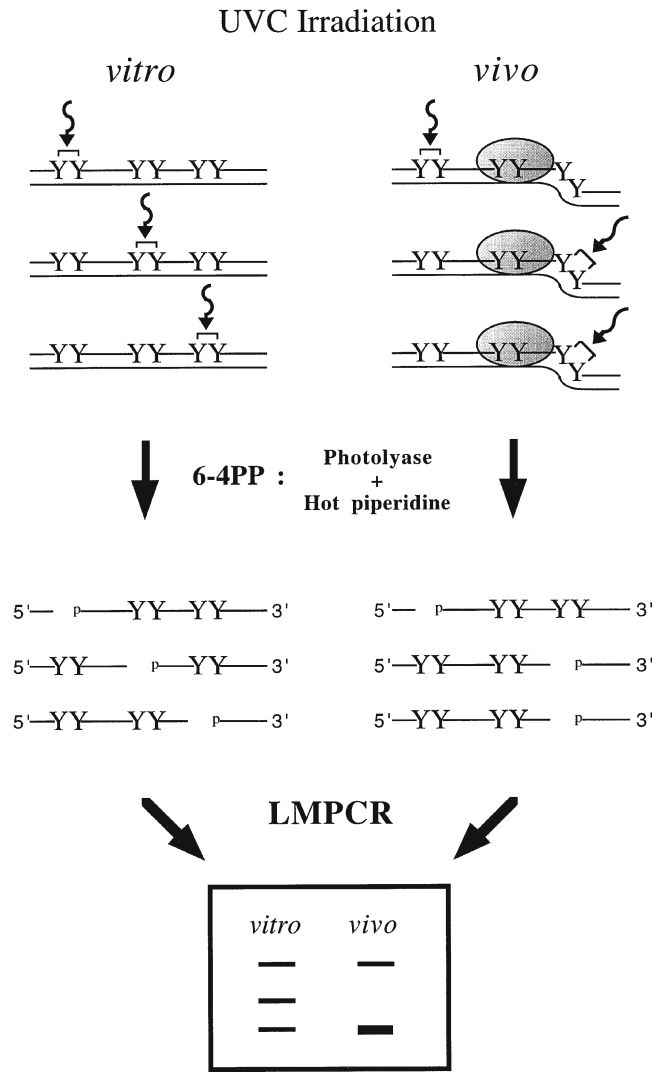


Fig. 3. Overall scheme for in vivo DNA analysis using UVC and 6-4PP formation. The 6-4PP formation following UVC exposure of purified DNA (in vitro) and cells (in vivo) is shown with curved arrows and brackets linking two adjacent pyrimidines (Y). When purified DNA is irradiated with UVC, the frequency of 6-4PP formations at dipyrimidine sites is determined by the DNA sequence. However, the presence of a sequence-specific DNA-binding protein illustrated by the dotted oval as well as DNA structure can prevent (negative photofootprint) or enhance (positive photofootprint) 6-4PP formation. First, CPD are photoreactivated by photolyase and then 6-4PP are cleaved by hot piperidine treatment leaving phosphorylated 5′ ends. On the sequencing ladder following LMPCR, the negative photofootprints appear as missing or less intense bands when compared with the sequencing ladder from the same DNA sequence obtained after UVC irradiation of purified DNA. On the other hand, positive photofootprints appear as darker bands in the sequencing ladder relative to the purified DNA control.

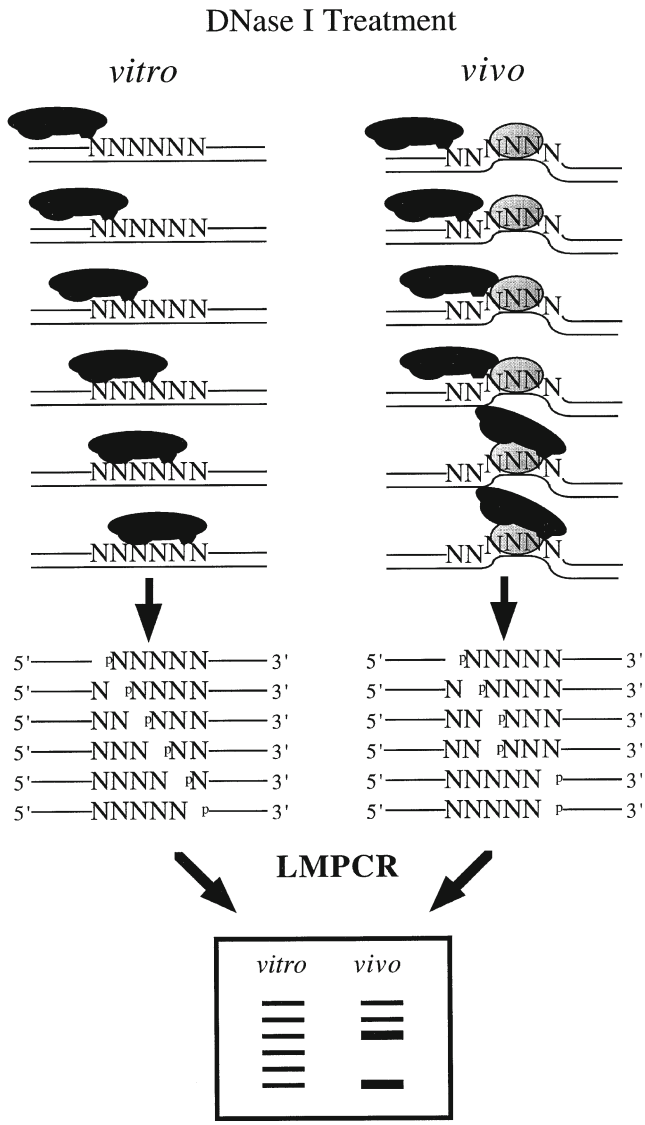


Fig. 4. Overall scheme for in vivo DNA analysis using DNase I. The DNase I enzyme (the solid black) digestion of purified DNA (in vitro) and cells (in vivo) is shown. When purified DNA is digested with DNase I, the cleavage pattern shows that sites of the nucleotide sequence have similar probabilities of being cleaved. However, the presence of a sequence-specific DNA-binding protein illustrated by the dotted oval as well as DNA structure can prevent (protection) or enhance (hypersensitive) DNase I cleavage. The DNase I cleavage leaves phosphorylated 5′ ends. On the sequencing ladder following LMPCR, DNA sequences that are protected from DNase I cleavage appear as missing or less intense bands when compared with the sequencing ladder from the same DNA sequence obtained after DNase I digestion of purified DNA. On the other hand, hypersensitive sites that undergo enhanced DNase I cleavage appear as darker bands in the sequencing ladder relative to the purified DNA control.

modified DNA can be used to infer protein-binding sites in living cells. It is always advisable to validate such interpretations using more than one footprinting agent.

The measure of in cellulo footprints has historically been problematic due to the dilute nature of target sequences and the complexity of higher eukaryotes genomes. The development of LMPCR, an extremely sensitive and specific technique, resolved this problem. The LMPCR technique quantitatively maps DNA single-strand breaks having phosphorylated 5′ ends within single-copy DNA sequences. It was first developed by Mueller and Wold *(6)* for DMS footprinting and subsequently, Pfeifer and colleagues adapted it to DNA sequencing *(7)*, methylation analyses *(1, 7)*, DNase I footprinting *(5)*, nucleosome positioning *(5)*, and UV footprinting (photofootprinting) *(4, 8)*. LMPCR can be combined with a variety of DNA-modifying agents used to probe the chromatin in cellulo. No single technique can provide as much information on DNA–protein interactions and DNA structures existing within the living cells as LMPCR can.

1.1. General Overview of LMPCR

Genomic-sequencing techniques, such as developed by Church and Gilbert *(9)*, can be used to map strand breaks in mammalian genes at the nucleotide resolution. However, by incorporating an exponential amplification step, LMPCR (outlined in **Fig. 5**) is advantageously more sensitive. It uses 20 times less DNA to obtain a nucleotide-resolution banding pattern and allows shorter autoradiographic exposure times than this technique. The unique aspect of LMPCR is the blunt-end ligation of an asymmetric double-stranded linker (5′ overhanging to avoid self-ligation or ligation in the wrong direction) onto the 5′ end of each cleaved blunt-ended DNA molecule *(6, 7)*. The blunt end is created by the primer extension (PE) of a gene-specific primer (primer 1 in **Fig. 5**) until a strand break is reached. Because the generated breaks are randomly distributed along the genomic DNA and thus have 5′ ends of unknown sequence, the asymmetric linker adds a common and known sequence to all 5′ ends. This then allows exponential PCR amplification using the longer oligonucleotide of the linker (linker-primer) and a second nested gene-specific primer (primer 2 in **Fig. 5**). LMPCR preserves the quantitative representation of each fragment in the original population of cleaved molecules *(10–13)*. After 22 cycles of PCR, the DNA fragments are size fractionated on a sequencing gel, allowing quantification on a phosphorimager *(14–17)*. Thus, the band intensity pattern obtained by LMPCR directly reflects the frequency distribution of 5′-phosphoryl DNA breaks along a 200-bp (or more if a sequencer is used) sequence adjacent to the nested primer.

Three variations of the LMPCR technique do exist. Pfeifer and colleagues *(7)* took advantage of electroblotting DNA onto

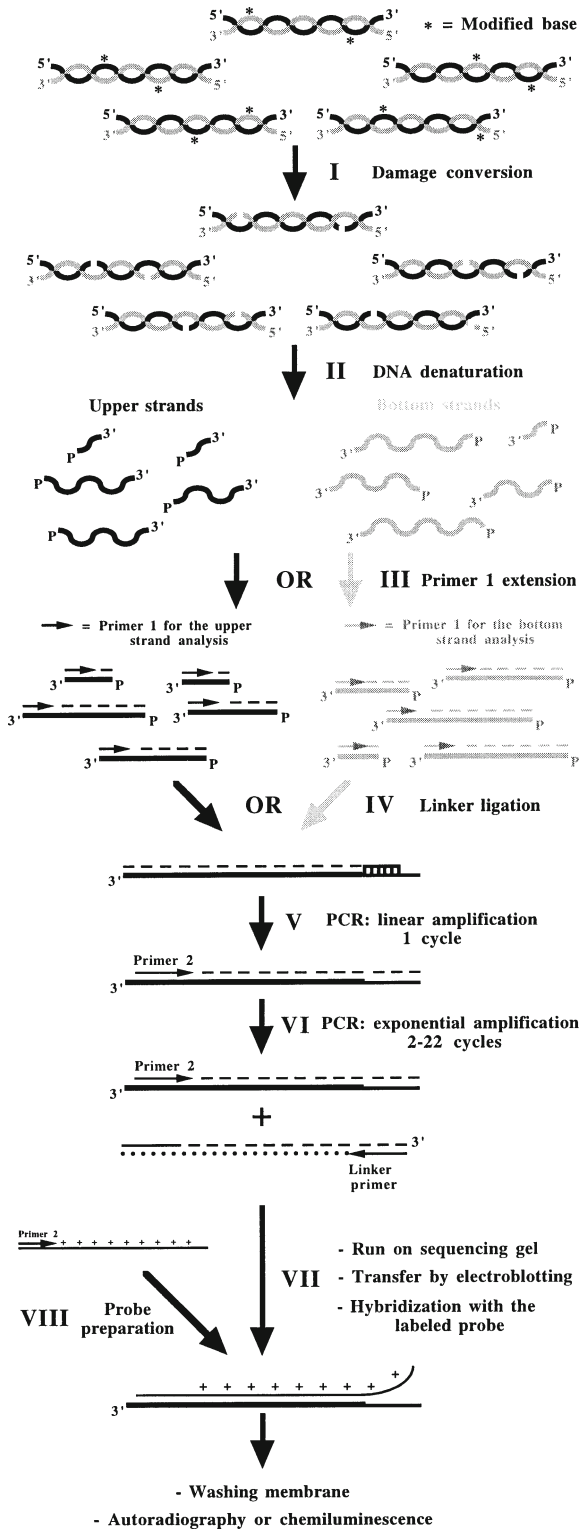


Fig. 5. Outline of the LMPCR procedure. **Step I**: specific conversion of modified bases to phosphorylated single-strand breaks; **Step II**: denaturation of genomic DNA; **Step III**: annealing and extension of primer 1 (although both strands can be studied, each LMPCR protocol only involves the analysis of either the nontranscribed strand or the transcribed strand); **Step IV**: ligation of the linker; **Step V**: first cycle of PCR amplification, this cycle is a linear amplification because only the gene-specific primer 2 can anneal; **Step VI**: cycles 2–22 of exponential PCR amplification of gene-specific fragments with primer 2 and the linker primer (the longer oligonucleotide of the linker); **Step VII**: separation of the DNA fragments on a sequencing gel, transfer of the sequence ladder to a nylon membrane by electroblotting, and visualization of the sequence ladder by hybridization with a labeled single-stranded probe; **Step VIII**: preparation and isotopic or nonisotopic labeling of single-stranded probe.

a nylon membrane followed by hybridization with a gene-specific probe to reveal the sequence ladders. This probe is typically ^{32}P-radiolabeled, but some have successfully used digoxigenin to get rid of the radioactivity issue (see detailed protocol in the second edition of *Methods in molecular biology*, *(18)*). On the other hand, Mueller and Wold *(6)* used a nested third radiolabeled primer for the last one or two cycles of the PCR amplification step. It is worthwhile to note that this last technique was recently employed to analyze LMPCR-amplified DNA fragments using sequencer devices *(19, 20)*. In this chapter, we describe two LMPCR protocols routinely used in the laboratory: one derived from the Pfeifer and colleagues' protocol and one based on the Mueller and Wold alternative sequencer method.

We also describe three probing methods generally combined with LMPCR to reveal in cellulo DNA protein interactions: dimethylsulfate (DMS), ultraviolet (UV), and DNase I **(Figs. 1–4**, **Table 1)**. These modifying agents provide complementary information and each has its associated advantages and drawbacks **(Table 2)**. To best characterize DNA protein interactions, it is often necessary to use two or even all three of these methods. Treatments with any probing agent must produce either strand breaks or modified nucleotides that can be converted to DNA strand breaks with a 5′-phosphate **(***see* **Figs. 1–4, Table 3)**. These protocols may also be adapted to footprinting with other probing agents, such as $KMnO_4$ and

Table 1
Purposes of the three main in cellulo footprinting approaches

Approaches	Goals
Dimethylsulfate (DMS)	Localizes in cellulo DNA–protein contacts located in the major groove of the DNA double helix Can detect special DNA structures
UV irradiation (UVB or UVC)	Localizes in cellulo DNA–protein interactions and shows how DNA structure is affected in the presence of transcription factors Can detect special DNA structures Can show evidence of positioned nucleosomes
DNase I	Localizes in cellulo DNA–protein contacts Precisely maps in cellulo DNase I-hypersensitive sites Shows evidence of nucleosomes and their positions; can differentiate core DNA from linker DNA

Table 2
Advantages and drawbacks of the three main In Cellulo footprinting approaches

Approaches	Advantages	Drawbacks
DMS	Treatment is technically easy to carry out; the DMS is a small molecule that penetrates very easily into living cells with little disruption	Requires guanines; therefore is sequence dependent Does not detect all DNA–protein interactions
UV irradiation (UVB or UVC)	Treatment is technically easy to carry out; UV light penetrates through the outer membrane of living cells without disruption Detects many DNA–protein interactions Very sensitive to particular DNA structures	Requires two adjacent pyrimidines; therefore is sequence dependent The interpretation of the results is sometime difficult; to differentiate between DNA–protein interactions and special DNA structures can be very difficult
DNase I	Little sequence dependency No conversion of modified bases required Detects all DNA–protein contacts Very sensitive to particular DNA structures	Technically difficult to carry out; reproducibility is often a problem DNase I is a protein which can penetrate in living cells only following membrane permeabilization, thus causing some cell disruption

Table 3
Mapping schemes used with the three main in cellulo footprinting approaches

Approaches	Strand breaks	Modified bases	Conversion of modified bases to DNA single-strand breaks
DMS	Few	Guanine: methylated guanines at N7 position Adenine: to a much lesser extent, methylated adenines at N3 position	Hot piperidine
UV irradiation (UVB or UVC)	Very few	Cyclobutane pyrimidine dimers 6-4 photoproducts	T4 endonuclease V followed by photolyase Photolyase followed by hot piperidine
DNase I	Yes	None	No conversion is required

OsO_4 (*see* Chap. "The Use of Diethyl Pyrocarbonate (DEPC) and Potassium Permanganate as Probes for Strand Separation and Structural Distortions in DNA" and **ref**. *21)* although a detailed description is beyond the scope of the present chapter.

Table 4
Exponential amplification steps using cloned *Pfu* DNA polymerase or *Taq* DNA polymerase

Cycles	Denaturation (*T* in °C for *D* in s)		Annealing (*T* is the T_m of the oligonucleotide for *D* in s)	Polymerization (*D* in s) *T* is the same for all cycles: 75°C for *Pfu* and 74°C for *Taq*
	Pfu	Taq	Pfu or Taq	–
0	–	93 for 120	–	–
1	98 for 300	98 for 150	T_m for 180	180
2	98 for 120	95 for 60	T_m for 150	180
3	98 for 60	95 for 60	T_m –2°C for 120	180
4	98 for 30	95 for 60	T_m –3°C for 120	180
5	98 for 20	95 for 60	T_m –4°C for 90	150
Repeat cycle 5, 13 more times (add 5 s per cycle for annealing and polymerization)				
6	98 for 20	95 for 60	T_m –3°C for 240	240
7	98 for 20	95 for 60	T_m –2°C for 240	240
8	98 for 20	95 for 60	T_m –1°C for 240	240
9	98 for 20	95 for 60	T_m for 120	600

Note: temperature (*T*) and duration (*D*) of the denaturation, annealing, and polymerization steps

1.2. In Cellulo Dimethylsulfate (DMS) Footprint Analysis (Fig. 1)

DMS is a small highly reactive molecule that easily diffuses through the outer cell membrane and into the nucleus. It preferentially methylates the N7 position of guanine residues via the major groove and, to a lesser extent, the N3 position of adenine residues via the minor groove. The most significant technical advantage of in cellulo DMS footprinting is that DMS can be simply added to the cell culture medium (*see* **Table 2** for advantages and drawbacks). Each guanine residue of in vitro DNA displays the same probability of being methylated by DMS. Because DNA inside living cells forms chromatin and is often found associated with a number of proteins, it is expected that its reactivity toward DMS will differ from that of in vitro DNA. **Figures 6** and **7** show in cellulo DMS treatment patterns compared to the treatment of in vitro DNA. Proteins in contact with DNA either decrease accessibility of specific guanines to DMS (protection) or, as frequently observed at the edges of a footprint, increase reactivity (hyper-reactivity) *(1)*. Hyper-reactivity can also indicate a greater DMS accessibility of special in cellulo DNA structure *(23)*.

Hot piperidine cleaves the glycosylic bond of methylated guanines and adenines, leaving a ligatable 5′-phosphate *(24)*.

Genomic footprinting using DMS reveals DNA–protein contacts located in the major groove of the DNA double helix (**Table 1**). However, it should be noted that in cellulo DNA analysis studies using DMS alone may not detect some DNA–protein interactions *(25)*. First, no DNA–protein interactions are detected in the absence of guanine residues. Second, some proteins do not affect DNA accessibility to DMS. Third, certain weak DNA–protein contacts could be disrupted because of the high reactivity of DMS. Thus, when using DMS, it is often important to confirm results with alternative in cellulo footprinting approaches *(25, 26)*.

1.3. In Cellulo UV Footprint (Photofootprint) Analysis (Figs. 2 and 3)

UVC (200–280 nm) and UVB (280–320 nm) can also be used as probing agents for in cellulo footprinting *(4, 8, 27–29)*. When cells are subjected to UVC (254 nm) or UVB, two major classes of lesions are introduced into DNA at dipyrimidine sites (CT, TT, TC, and CC): cyclobutane pyrimidine dimers (CPD) and the pyrimidine (6-4) pyrimidone photoproducts (6-4PP) *(30)*. CPD are formed between the 5,6 bonds of any two adjacent pyrimidines, whereas a stable bond between positions 6 and 4 of two adjacent pyrimidines characterizes 6-4PP. 6-4PP are formed at a rate 15–30% of that of CPD *(31)* and are largely converted to their Dewar valence isomers by direct secondary photolysis (photoisomerization) *(31)*. In living cells, the photoproduct distribution is determined both by sequence context and chromatin structure *(32)*. In general, CPD and 6-4PP appear to form preferentially in longer pyrimidine runs. Because DNA absorbs directly UVB and UVC and because cells are exposed during a short period of time to high UV intensities, there are relatively few perturbations of other cellular processes and secondary events that could modify the chromatin structure or release DNA–protein interactions. Thus, UV irradiation is probably one of the least disruptive footprinting methods and hence truly reflects the in cellulo situation (**Table 2**). As for DMS, DNA-binding proteins influence the distribution of UV photoproducts in a significant way *(27)*. When the photoproduct spectrum of in vitro irradiated DNA is compared with that obtained after irradiation in cellulo, differences become apparent. The photoproduct frequency within sequences bound by sequence-specific DNA-binding proteins is suppressed or enhanced in comparison to in vitro DNA *(2, 4, 8)*. Effects of chromatin structure may be significant in regulatory gene regions that bind transcription factors (**Fig. 6**). Mapping of CPD at the single-copy gene level can reveal positioned nucleosomes because CPD are modulated in a 10-bp periodicity within nucleosome core DNA *(33, 34)*. 6-4PP form more frequently in linker DNA than in core DNA *(35)*.

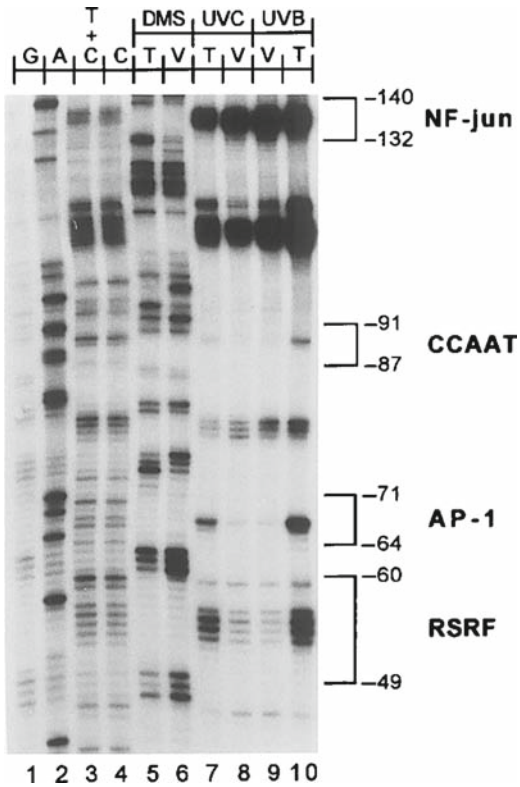


Fig. 6. LMPCR analysis of methylated guanines and CPD along the nontranscribed strand of the c-jun promoter following DMS treatment, and UVB and UVC irradiation, respectively. The membrane was hybridized with an isotopic [^{32}P]-dCTP-labeled probe. *Lanes 1–4*: LMPCR of DNA treated with chemical cleavage reactions. These lanes represent the sequence of the *c-jun* promoter analyzed with JD primer set (*22*). *Lanes 5–6*: LMPCR of DMS-treated naked DNA (T: in vitro) and fibroblasts (V: in vivo) followed by hot piperidine treatment. *Lanes 7–10*: LMPCR of UVC- and UVB-irradiated naked DNA (T) and fibroblasts (V) followed by T_4 endonuclease V/photolyase digestions. On the *right*, the consensus sequences of transcription-factor-binding sites are delimited by brackets. The numbers indicate their positions relative to the major transcription initiation site.

Photofootprints reveal variations in DNA structure associated with the presence of transcription factors or other proteins bound to DNA. UV light has the potential to reveal all DNA–protein interactions provided when a dipyrimidine sequence on either DNA strand within a putative binding sequence. Because UV footprints can be seen outside protein-binding sites, UV light should not be used as the only in cellulo footprinting agent. The precise delimitations of the protein–DNA contact are difficult to determine with the simple in cellulo UV probing method.

The distribution of UV-induced CPD and 6-4PP along genomic DNA can be mapped at the sequence level by LMPCR following conversion of these photoproducts into ligatable 5′-phosphorylated single-strand breaks. CPD are enzymatically converted by cleavage with T_4 endonuclease V followed by UVA (320–400 nm) photoreactivation of the overhanging pyrimidine

using photolyase (**Fig. 2**) *(8)*. Because the 6-4PP and their Dewar isomers are hot alkali-labile sites, they can be cleaved by hot piperidine (**Fig. 3**) *(2)*. Generally we simply measure the CPD distribution. Performing 6-4PP mapping is of interest only if no other alternative footprinting method is available.

1.4. In Cellulo DNase I Footprint Analysis (Fig. 4)

DNase I treatment of permeabilized cells gives clear footprints when the DNase I-induced breaks are mapped by LMPCR *(5)*. As with DMS and UV, footprint analyses are obtained by comparing in cellulo DNase I digestion patterns to patterns obtained from the in vitro DNA digestion (**Fig. 7**). When compared to in

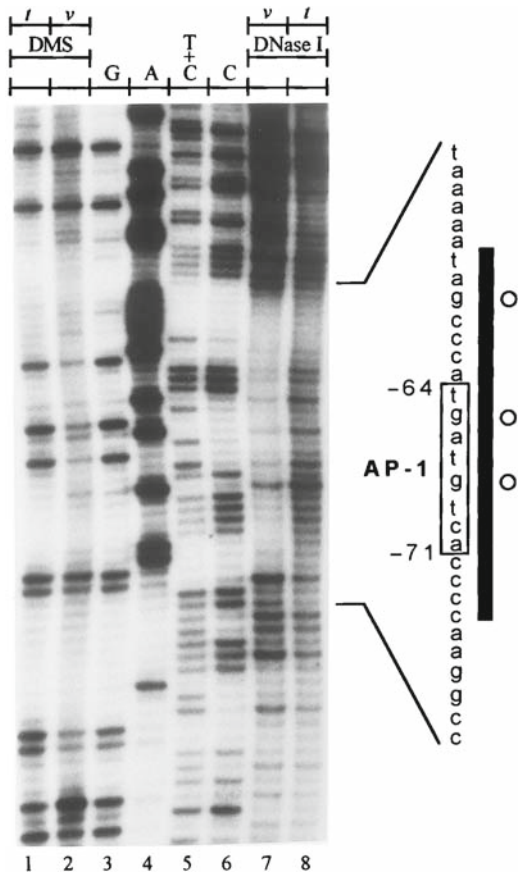


Fig. 7. LMPCR analysis of methylated guanines and DNA strand breaks along the transcribed strand of the c-jun promoter following DMS treatment and DNase I digestion, respectively. The membrane was hybridized with an isotopic [^{32}P]-dCTP-labeled probe. *Lanes 1–2*: LMPCR of DMS-treated purified DNA (t: in vitro) and fibroblasts (v: in vivo) followed by hot piperidine treatment. *Lanes 3–6*: LMPCR of DNA treated with chemical cleavage reactions. These lanes represent the sequence of the *c-jun* promoter analyzed with JC primer set (*22*). *Lanes 7–8*: LMPCR of DNase I-digested permeabilized fibroblasts (v) and purified DNA (t). As a reference, a small portion of the chemically derived sequence is shown on the right of the autoradiogram, the AP-1-like binding sequence is enclosed by a box, and the numbers indicate its position relative to the major transcription initiation site. *Open circles* represent guanines that are protected against DMS-induced methylation (negative DMS footprints) in vivo. The *black bar* shows the protected sequence against DNase I-induced cleavage in vivo. Thus, in vivo DNase I footprinting analysis delimits much better the DNA–protein interactions.

vitro DNA, permeabilized cells show protected bands at protein–DNA interaction sequences and DNase I-hypersensitive bands in regions of higher-order nucleoprotein structure *(5)*. Compared to DMS, DNase I is less base selective, more efficient at detecting minor groove DNA–protein contacts, provides more information on chromatin structure, displays larger and clearer footprints, and better delimits the boundaries of DNA–protein interactions (**Fig. 7**). The nucleotides covered by a protein are almost completely protected on both strands from DNase I nicking, allowing a better delimitation of the boundaries of DNA–protein contacts. However, it should be underlined that the relatively bulky DNase I molecule cannot cleave the DNA in the immediate vicinity of a bound protein because of steric hindrance. Consequently, the regions protected from cutting can extend beyond the real DNA–protein contact site. On the other hand, when DNA is wrapped around a nucleosome-sized particle, DNase I cutting activity is increased at 10-bp intervals and usually no footprint is observed (**Tables 1** and **2**).

The DNase I is a relatively large 31-kD protein and cannot penetrate cells without previous cell-membrane permeabilization. Cells can be efficiently permeabilized by L-α-lysophosphatidylcholine (lysolecithin) *(5)* or Nonidet P40 *(36)* (see detailed protocol on the second edition of *Methods in molecular biology*, *(18)*). It has been shown that cells permeabilized by lysolecithin remain intact, replicate their DNA very efficiently, and show normal transcriptional activities *(37, 38)*. There are numerous studies showing that lysolecithin-permeabilized cells maintain a normal nuclear structure to a greater extent than isolated nuclei, because the chromatin structure can be significantly altered during the nuclear isolation procedures *(5)*. Indeed, DNase I footprinting studies using isolated nuclei can be flawed because transcription factors are lost during the isolation of nuclei in polyamine containing buffers *(5)*. Even though other buffers may be less disruptive, transcription factors can still be lost during the isolation procedure, leading to the complete or partial loss of footprints.

DNase I digestion of DNA leaves ligatable 5′-phosphorylated breaks, but the 3′-ends are free hydroxyl groups. Pfeifer and colleagues *(5, 39)* observed that these 3′-OH ends can be used as short primers and extended by the DNA polymerases during the Prime Extention (PE) and/or PCR steps of LMPCR, thereby reducing significantly the overall efficiency of LMPCR and giving a background smear on sequencing gels. To avoid the nonspecific priming of these 3′-OH ends, three alternative solutions have been applied: (1) blocking these ends by the addition of a dideoxynucleotide *(5, 39)*, (2) enrichment of fragments of interest by extension product capture using biotinylated gene-specific primers and magnetic streptavidin-coated beads *(22, 40–42)*, and (3) performing primer 1 hybridization and PE at a higher temperature (52–60 vs. 48°C, and 75 vs. 48°C, respectively, using a thermostable

DNA polymerase such as *Vent* exo⁻ and *Pfu* exo⁻ *(3, 43–46)*). Although effective, the first two alternatives involve additional manipulations that are time consuming. Because of its simplicity, we select primer 1 with higher T_m (52–60°C) and use the *Pfu* exo⁻ or *Vent* exo⁻ DNA polymerase for the PE.

1.5. Choice of DNA Polymerases for LMPCR

LMPCR involves the PCR amplification of differently sized genomic DNA fragments. During the LMPCR procedure, DNA polymerases are required for two steps: PE and PCR amplification. For the PE step, the best DNA polymerase would be one that (1) is thermostable and very efficient, (2) has no terminal transferase activity, (3) is able to efficiently polymerize about 0.5 kb of DNA even when the DNA is very GC rich, and (4) is able to polymerize through any DNA secondary structures. For the PCR step, the best DNA polymerase would be (1) thermostable, (2) very efficient, (3) able to amplify indiscriminately a mixture of DNA fragments of different lengths (between 50 and 500 bp) and of varying GC-richness (from 5 to 95%), and (4) able to efficiently resolve DNA secondary structures. We find that *Pfu* exo⁻ and *Vent* exo⁻ are the best enzymes for the PE and PCR steps of LMPCR *(45, 47)*. In this chapter, LMPCR protocols using *Pfu* exo⁻ DNA polymerase for PE and PCR steps will be described in detail. However, because the *Vent* exo⁻ polymerase is frequently used for the PE step and *Taq* DNA polymerase or *Vent* exo⁻ polymerase for the PCR amplification, an alternative protocol using these polymerases will also be included.

2. Materials

Nanopure H_2O should be used in making any buffers, solutions, and dilutions, unless otherwise specified.

2.1. DNA Purification (for 10^7–10^8 Cells)

1. Any type of cells (*i.e.*, fibroblasts, lymphocytes, etc.).
2. Trypsin-EDTA (Wisent).
3. Buffer A: 300 mM sucrose, 60 mM KCl, 15 mM NaCl, 60 mM Tris–HCl of pH 8.0, 0.5 mM spermidine, 0.15 mM spermine, and 2 mM EDTA. Store at –20°C.
4. Buffer A + 1% Nonidet P40 Substitute (Fluka). Store at –20°C.
5. Buffer B: 150 mM NaCl and 5 mM EDTA of pH 7.8.
6. Buffer C: 20 mM Tris–HCl of pH 8, 20 mM NaCl, 20 mM EDTA, and 1% SDS.
7. Proteinase K from *Tritirachium album* (Roche Applied Science).

8. RNase A from bovine pancreas (Roche Applied Science).
9. Phenol, equilibrated, pH 8 (USB Corporation).
10. Chloroform.
11. 5 M NaCl.
12. Precooled absolute ethanol (–20°C).
13. Precooled 80% ethanol (–20°C).

2.2. Chemical Cleavage for DNA-Sequencing Products

1. K_2PdCl_4 solution: 10 mM K_2PdCl_4 (potassium tetrachloropalladate(II), Aldrich) and 100 mM HCl, pH 2.0 (adjusted with NaOH). Store at –20°C.
2. K_2PdCl_4 stop: 1.5 M sodium acetate of pH 7.0 and 1 M β-mercaptoethanol. Store at –20°C.
3. Dimethylsulfate (DMS Store at –20°C, 99+%, Fluka). Considering its toxic and carcinogenic nature, DMS should be manipulated in a well-ventilated hood. DMS is stored under nitrogen at 4°C and should be replaced every 12 months. DMS waste is detoxified in 5 M NaOH.
4. DMS buffer: 50 mM sodium cacodylate and 1 mM EDTA of pH 8. Store at 4°C.
5. DMS stop: 1.5 M sodium acetate of pH 7.0 and 1 M β-mercaptoethanol. Store at –20°C.
6. Hydrazine (Hz, anhydrous, Aldrich). Considering its toxic and carcinogenic potentials, Hz should be manipulated in a well-ventilated hood. Hz is stored under nitrogen at 4°C in an explosion-proof refrigerator, and the bottle should be replaced at least every 6 months. Hz waste is detoxified in 3 M ferric chloride.
7. Hz stop: 300 mM sodium acetate of pH 7.0 and 0.1 mM EDTA. Store at 4°C.
8. 5 M NaCl.
9. 3 M Sodium acetate of pH 7.0.
10. Precooled absolute ethanol (–20°C).
11. Precooled 80% ethanol (–20°C).
12. Dry ice or –80°C freezer.
13. Piperidine (99+%, 10M, Fluka or Sigma): diluted to 2 M with H_2O just before use. Cap immediately to minimize evaporation and keep on ice. Considering its toxic and carcinogenic potentials, piperidine should be manipulated in a well-ventilated hood. Piperidine 10 M is stored at 4°C under nitrogen atmosphere.
14. Teflon tape.
15. Lock caps.
16. 3 M Sodium acetate of pH 5.2.

17. 20 µg/µL Glycogen (Roche Applied Science).
18. Vacuum concentrator.

2.3. Template Preparation: PCR Products

2.3.1. PCR Amplification

1. 5× *Taq* buffer: 50 mM Tris–HCl of pH 8.9, 200 mM NaCl, and 0.05% [w/v] gelatin (*see* **Note 1**).
2. Two primer 2 (50 pmol/µL), one for each strand of the DNA fragment to be amplified, distant from 150 to 450 bp.
3. *Taq* DNA polymerase PCR product mix: 2× *Taq* buffer, 4 mM $MgCl_2$, 0.4 mM of each dNTP, 10 pmol of each primer 2 and 3 U *Taq* DNA polymerase (5 U/µL, Roche Applied Science).
4. *Taq* DNA polymerase stop: 1.56 M sodium acetate of pH 5.2 and 60 mM EDTA.
5. Precooled absolute ethanol (–20°C).
6. Precooled 80% ethanol (–20°C).
7. Dry ice or –80°C freezer.
8. 5× Neutral loading buffer: 0.25% bromophenol blue, 0.25% xylene cyanol FF, and 30% glycerol. Store at 4°C.

2.3.2. Purification and Quantification of PCR Products

1. Agarose.
2. 1× TAE buffer: 40 mM Tris base, 20 mM glacial acetic acid, and 1 mM EDTA of pH 8.0.
3. DNA size standard (100 bp, Invitrogen).
4. Ethidium bromide.
5. Glass wool.
6. 3 M sodium acetate of pH 7.0.
7. Precooled absolute ethanol (–20°C).
8. Precooled 80% ethanol (–20°C).
9. Dry ice or –80°C freezer.
10. Low DNA mass ladder (Invitrogen).
11. 5× Neutral loading buffer: 0.25% bromophenol blue, 0.25% xylene cyanol FF, and 30% glycerol. Store at 4°C.

2.4. Treatment of Purified DNA and Living Cells with Modifying Agents

2.4.1. DMS Treatment

1. 0.2% DMS (99 + %, Fluka) freshly prepared in serum-free medium.
2. Trypsin-EDTA (Wisent).
3. Hank's balanced salt solution (HBSS, Wisent), cold.
4. DMS buffer: 50 mM sodium cacodylate and 1 mM EDTA of pH 8. Store at 4°C.
5. DMS stop: 1.5 M sodium acetate of pH 7.0 and 1 M β-mercaptoethanol. Store at –20°C.
6. Buffer A: 300 mM sucrose, 60 mM KCl, 15 mM NaCl, 60 mM Tris–HCl of pH 8.0, 0.5 mM spermidine, 0.15 mM spermine, and 2 mM EDTA. Store at –20°C.

7. Buffer A + 1% Nonidet P40 substitute (Fluka). Store at –20°C.
8. Buffer B: 150 mM NaCl and 5 mM EDTA of pH 7.8.
9. Buffer C: 20 mM Tris–HCl of pH 8, 20 mM NaCl, 20 mM EDTA, and 1% SDS.
10. Proteinase K from *Tritirachium album* (Roche Applied Science).
11. RNase A from bovine pancreas (Roche Applied Science).
12. Phenol, equilibrated, pH 8 (USB Corporation).
13. Chloroform.
14. 5 M NaCl.
15. Precooled absolute ethanol (–20°C).
16. Precooled 80% ethanol (–20°C).
17. Dry ice or –80°C freezer.

2.4.2. UVC (254 nm) and UVB Irradiation

1. Germicidal lamp (254 nm UVC, Philips G15 T8, TUV 15W) or UVB light (Philips, FS20T12/UVB/BP).
2. UVX digital radiometer and the appropriate probe (Ultraviolet Products, Upland, CA).
3. 0.9% NaCl.
4. 150-mm Petri dishes.
5. UV irradiation buffer: 150 mM KCl, 10 mM NaCl, 10 mM Tris–HCl of pH 8.0, and 1 mM EDTA.
6. Buffer A: 300 mM sucrose, 60 mM KCl, 15 mM NaCl, 60 mM Tris–HCl of pH 8.0, 0.5 mM spermidine, 0.15 mM spermine, and 2 mM EDTA. Store at –20°C.
7. Buffer A + 0.5% Nonidet P40 substitute (Fluka). Store at –20°C.
8. Buffer A + 1% Nonidet P40 substitute (Fluka). Store at –20°C.
9. Scraper.
10. 5 M NaCl.
11. Precooled absolute ethanol (–20°C).
12. Precooled 80% ethanol (–20°C).
13. Dry ice or –80°C freezer.

2.4.3. DNase I Treatment

1. 0.5 mg/mL Deoxyribonuclease I (DNase I, Worthington Biochemical Corporation).
2. Hank's balanced salt solution (HBSS,Wisent).
3. Solution I: 150 mM sucrose, 80 mM KCl, 35 mM HEPES of pH 7.4, 5 mM $MgCl_2$, and 0.5 mM $CaCl_2$.
4. Solution I + 0.05% L-α-Lysophosphatidylcholine (L-a-Lysolecithin, Sigma).
5. Solution II: 150 mM sucrose, 80 mM KCl, 35 mM HEPES of pH 7.4, 5 mM $MgCl_2$, and 2 mM $CaCl_2$.

6. Scraper.
7. Conical tubes, 15 mL.
8. Buffer B: 150 mM NaCl and 5 mM EDTA of pH 7.8.
9. Buffer C: 20 mM Tris–HCl of pH 8.0, 20 mM NaCl, 20 mM EDTA, and 1% SDS.
10. Proteinase K from *Tritirachium album* (Roche Applied Science).
11. RNase A from bovine pancreas (Roche Applied Science).
12. Phenol, equilibrated, pH 8 (USB Corporation).
13. Chloroform.
14. 5 M NaCl.
15. 20µg/µl Glycogen (Roche Applied Science).
16. Precooled absolute ethanol (–20°C).
17. Precooled 80% ethanol (–20°C).
18. Dry ice or –80°C freezer.

2.5. Conversion of Modified Bases to DNA Single-Strand Breaks

1. Piperidine (99 + %, 10 M, *see* **Subheading 2.2**, **item 13**).
2. Teflon tape.
3. Lock caps.
4. 3 M Sodium acetate, pH 5.2.
5. 20 µg/µL Glycogen (Roche Applied Science).
6. Precooled absolute ethanol (–20°C).
7. Precooled 80% ethanol (–20°C).
8. Dry ice of –80° freezer.
9. Vacuum concentrator.

2.5.1. DMS-Induced Base Modifications

2.5.2. UV-Induced Base Modifications

CPD

1. 10× Dual buffer: 500 mM Tris–HCl of pH 7.6, 500 mM NaCl, and 10 mM EDTA.
2. T_4 endonuclease V enzyme (Trevigen). The saturating amount of T_4 endonuclease V enzyme can be estimated by digesting UV-irradiated genomic DNA with various enzyme quantities and separating the cleavage products on alkaline agarose gel *(48)*. The saturating amount of the enzyme is the one next to the minimum quantity that produces the maximum cleavage frequency as evaluated on the alkaline agarose gel.
3. T_4 endo V mix: 2× dual buffer, 2 mM DTT (1,4-dithiothreitol, Roche Applied Science), 0.2 mg/mL BSA (nuclease-free bovine serum albumin, Roche Applied Science), and a saturating amount of T_4 endonuclease V.
4. *E. coli* photolyase enzyme (Trevigen). The saturating amount of photolyase can be estimated by photoreactivating UV-irradiated genomic DNA with various enzyme quantities, digestion with T_4 endonuclease V, and separating the cleavage products

on alkaline agarose gel *(48)*. The saturating amount of photolyase is the next to the minimum enzyme quantity which produces no cleavage following T_4 endonuclease V digestion as evaluated on the gel. Because photolyase is light sensitive, all steps involving photolyase should be carried out rapidly.

5. Photolyase mix: 1× dual buffer, 1.1 M DTT, 0.1 mg/mL BSA, and a saturating amount of photolyase.
6. UVA black light blue (Sankyo Denki 350 nm).
7. Plastic film (plastic wrap).
8. 1% SDS solution.
9. Phenol, equilibrated, pH 8 (USB Corporation).
10. Chloroform.
11. 5 M NaCl.
12. Precooled absolute ethanol (–20°C).
13. Precooled 80% ethanol (–20°C).
14. Dry ice of –80°C freezer.

1. Piperidine (99+%, 10 M, *see* **Subheading 2.2**, **item 13**).
2. Teflon tape.
3. Lock caps.
4. 3 M Sodium acetate, pH 5.2.
5. 20 µg/µL Glycogen (Roche Applied Science).
6. Precooled absolute ethanol (–20°C).
7. Precooled 80% ethanol (–20°C).
8. Dry ice of –80° freezer.
9. Vacuum concentrator.

2.6. Ligation-Mediated Polymerase Chain Reaction Technology

2.6.1. Primer Extension (Steps II and III, Fig. 5)

1. A gene-specific primer (primer 1) is used to initiate PE. The primer 1 used in the first-strand synthesis is a 15–22 bp oligonucleotide and has a calculated melting temperature (T_m) of 50–60°C. They are selected using a computer program (Oligo 4.0 software, National Biosciences, Rychlik and Rhoads 1989) and optimally, their T_m, as calculated by a computer program (GeneJockey software), should be about 10°C lower than that of subsequent primers (*see* **Note 2**) *(49)*. The first-strand synthesis reaction is designed to require very little primer 1 with a lower T_m so that this primer does not interfere with subsequent steps *(11–13, 50)*. The primer 1 concentration is set at 0.5 pmol/µL in H_2O.
2. Thermocycler (Biometra or PTC™, MJ research, Inc.).
3. Polymerase extension mix:
   (a) If *Pfu* exo$^-$ DNA polymerase (also named cloned *Pfu*) is used:

– 10× *Pfu* exo⁻ buffer: 200 mM Tris–HCl of pH 8.8, 20 mM $MgSO_4$, 100 mM NaCl, 100 mM $(NH_4)_2SO_4$, 1% (v/v) Triton X-100, and 1 mg/mL nuclease-free BSA (*see* **Note 1**).
– *Pfu* exo⁻ extension mix: 1.25 pmole primer 1, 0.25 mM of each dNTP, 1× *Pfu* exo⁻ buffer, and 1.5 U *Pfu* exo⁻ (2.5 U/μL, Stratagene).

(b) If *Vent* exo⁻ polymerase is used:

– *Vent* exo⁻ extension mix: 1.25 pmole primer 1, 0.25 mM of each dNTP, 1× *Vent* buffer (10×, New England BioLabs), and 0.75 U *Vent* exo⁻ polymerase (2.0 U/μL New England BioLabs).

2.6.2. Ligation (Step IV, Fig. 5)

1. The DNA molecules that have a 5′-phosphate group and a double-stranded blunt end are suitable for ligation. A DNA linker with a single blunt end is ligated directionally onto the double-stranded blunt end of the extension product using T_4 DNA ligase. This linker has no 5′ phosphate and is staggered to avoid self-ligation and provide directionality. Also, the duplex between the 25-mer (L25: 5′ GCGGTGACCCGGGAGATCTGAATTC) and 11-mer (L11: 5′ GAATTCAGATC) is stable at the ligation temperature, but denatures easily during subsequent PCR reactions *(6, 50)*. The linker 20 pmol/μL is prepared in aliquots of 500 μL by annealing in 250 mM Tris–HCl of pH 7.7, 120 mM $MgCl_2$ and 20 pmol/μL each of the 25-mer and 11-mer (the stock is at 60 pmol/μL), heating at 95°C for 3 min, transferring quickly at 70°C, cooling gradually to room temperature and stored at 4°C overnight. Linker is stored at –20°C and thawed on ice before use.
2. Ligation mix: 33 mM DTT, 1.1 mM ATP, 16.6 μg/mL BSA, 48.9 mM Tris–HCl of pH 7.4, 100 pmol linker, and 3.25 U T_4 ligase (1 U/μL, Roche Applied Science).
3. T_4 ligase stop mix: 7.2 M Ammonium acetate, 4.2 mM EDTA of pH 8.0, and 0.67 μg/μL glycogene (Roche Applied Science).
4. Precooled absolute ethanol (–20°C).
5. Precooled 80% ethanol (–20°C).
6. Dry ice or –80°C freezer.

2.6.3. Polymerase Chain Reaction (Steps V and VI, Fig. 5)

1. At this step, gene-specific fragments can be exponentially amplified because primer sites are available at each target fragment ends (*i.e.*, primer 2 on one end and the longer oli-gonucleotide of the linker on the other end). Primer 2 may or may not overlap with primer 1. The overlap, if present, should not be more than seven to eight bases *(11–13, 50)*. The primer 2 and the linker primer (L25) are diluted in H_2O to give 50 pmol/μL and 60 pmol/μL, respectively.

2. Thermocycler (Biometra or PTC, MJ research, Inc.).
3. Polymerase amplification reaction
   (a) If *Pfu* exo$^-$ polymerase is used:
      – 10× *Pfu* exo$^-$ buffer: 200 mM Tris–HCl of pH 8.8, 20 mM $MgSO_4$, 100 mM NaCl, 100 mM $(NH_4)_2SO_4$, 1% (v/v) Triton X-100, and 1 mg/mL nuclease-free BSA (*see* **Note 1**).
      – *Pfu* exo$^-$ amplification mix: 2× *Pfu* exo$^-$ buffer, 0.5 mM of each dNTP, 10 pmol of L25 (linker primer), 10 pmol of primer 2, and 3.5 U of *Pfu* exo$^-$ DNA polymerase (2.5 U/μL, Stratagene).
   (b) If *Taq* DNA polymerase is used:
      – 5× *Taq* buffer: 50 mM Tris–HCl of pH 8.9, 200 mM NaCl, and 0.05% [w/v] gelatin (*see* **Note 1**).
      – *Taq* amplification mix: 2× *Taq* buffer, 4 mM $MgCl_2$, 0.5 mM of each dNTP, 10 pmol L25 (linker primer), 10 pmol primer 2 and 3 U *Taq* DNA polymerase (5 U/μL, Roche Applied Science).
   (c) If *Vent* exo$^-$ is used:
      – *Ven*t exo$^-$ amplification mix: 2× *Vent* exo$^-$ buffer (10×, New England BioLabs), 0.5 mM of each dNTP, 10 pmol of L25 (linker primer), 10 pmol of primer 2 and 2.5 U *Vent* exo$^-$ polymerase (2.0 U/μL New England BioLabs).

2.7. LMPCR-Amplified DNA Fragment Analysis

2.7.1. Conventional Radioactive Method

Precipitation, Gel Electrophoresis, and Electroblotting (Step VII, Fig. 5)

1. Stop mix:
   (a) If *Pfu* exo$^-$ is used for amplification:
      – *Pfu* exo$^-$ stop mix: 1.56 M sodium acetate of pH 5.2 and 20 mM EDTA.
   (b) If *Taq* DNA polymerase is used for amplification:
      – *Taq* stop mix: 1.56 M sodium acetate of pH 5.2 and 60 mM EDTA.
   (c) If *Vent* exo$^-$ is used for amplification:
      – *Vent* exo$^-$ stop mix: 1.56 M sodium acetate of pH 5.2 and 20 mM EDTA.
2. Precooled absolute ethanol (–20°C).
3. Precooled 80% ethanol (–20°C).
4. Dry ice or –80°C freezer.
5. Formamide-loading dye: 94% formamide, 2 mM EDTA of pH 7.7, 0.05% xylene cyanole FF, and 0.05% bromophenol blue *(11–13)*.
6. 66 cm Long × 34.5 cm wide sequencing gel apparatus (Owl Scientific).

7. Spacers (0.4-mm thick).
8. Plastic well-forming comb (0.4-mm thick, BioRad).
9. 5× (0.5 M) Tris-Borate-EDTA (TBE) buffer: 500 mM Tris, 830 mM boric acid, and 10 mM EDTA of pH 8.3. Use this stock to prepare 1× (100 mM) TBE buffer.
10. 8% Polyacrylamide. To prepare 1 L: 77.3 g acrylamide, 2.7 g *bis*-acrylamide, 420.42 g urea, and 200 mL of 0.5 M TBE dissolved in H_2O. Polyacrylamide solution should be kept at 4°C.
11. Gel preparation: mix 100 mL of 8% polyacrylamide with 1 mL of 10% ammonium persulfate (APS), and 30 µL of *N,N,N′,N′*-tetra-methylethylenediamide (TEMED). This mix is prepared immediately before pouring the solution between the glass plates. Without delay, take the gel mix into a 50-mL syringe and inject the mix between the plates, maintaining a steady flow. During pouring, the plates should be kept at a 30° angle and tilted to the side into which the mix is injected. Any air bubbles should be avoided and removed if they form. The gel should be left to polymerize for a minimum of 2 h before use. If the gel is to be left overnight, 45 min after pouring, place a moistened paper tissue over the comb and cover the upper end of the assembly with a plastic film to prevent the gel from drying out.
12. Power supply (Bio-Rad PowerPac 3000).
13. Electroblotting apparatus (HEP3, Owl Scientific Inc.) used according to the manufacturer's instructions.
14. Whatman 3-MM Chr paper (Fisher Scientific).
15. Plastic film (plastic wrap).
16. Whatman 17-MM Chr paper (Fisher Scientific).
17. Nylon membrane, positively charged (Roche Applied Science).
18. Power supply (Bio-Rad, model 200/2.0).
19. UVC (254 nm) germicidal lamp.
20. UVX digital radiometer and appropriate probe (Ultraviolet Products, Upland, CA).

Preparation of Single-Stranded Hybridization Probes **(Step VIII, Fig. 5)**

1. 5× *Taq* buffer: 50 mM Tris–HCl of pH 8.9, 200 mM NaCl, and 0.05% [w/v] gelatin (*see* **Note 1**).
2. dNTP (dATP, dGTP, dTTP) mix (200 µM of each) diluted 1:10 in H_2O. This mix is changed every 2 weeks.
3. Isotopic labeling mix: 1× *Taq* buffer, 2 mM $MgCl_2$, 133 pg/µL DNA template (PRC products), 0.5 pm/µL primer 2, 1 µL of 1:10 dNTP mix, 5 U *Taq* DNA polymerase (5 U/µL, Roche Applied Science), and 1.85 MBq α-[^{32}P] dCTP (0.37 MBq/µL, PerkinElmer Life Sciences Inc).

4. 7.5 M Ammonium acetate.
5. 20 μg/μL glycogen (Roche Applied Science).
6. Precooled absolute ethanol (–20°C).
7. Geiger counter.
8. TE buffer of pH 8.0: 10 mM Tris–HCl of pH 8.0 and 1 mM EDTA of pH 7.8.
9. Hybridization buffer: 250 mM sodium phosphate of pH 7.2, 1 mM EDTA, 7% SDS, and 1% BSA.

Hybridization (Step VII, Fig. 5), Washing, and Autoradiography

The hybridization is performed in a rolling 8 cm diameter × 22 cm long borosilicate glass hybridization tubes in a hybridization oven (Techne). The nylon membrane is soaked in 100 mM TBE and placed in the tube using a 25-mL pipet, so that the membrane sticks completely to the wall of the hybridization tube. Following hybridization and washing, the membrane is placed in an autoradiography cassette Fujifilm EC-DW (Christie Group Ltd) and exposed to Kodak X-ray film (BiomaxMR, 35 × 43 cm, Kodak Scientific Imaging Film) with intensifying screens (35 × 43 cm, Fisher Scientific, cat. no. FB-IS-1417) at –80°C.

1. Hybridization buffer: 250 mM sodium phosphate of pH 7.2, 1 mM EDTA, 7% SDS, and 1% BSA.
2. Single-stranded hybridization probe diluted in 6–7 mL of hybridization buffer.
3. Washing buffer I: 20mM sodium phosphate of pH 7.2, 1 mM EDTA, 0.25% BSA, and 2.5% SDS.
4. Washing buffer II: 20 mM sodium phosphate of pH 7.2, 1 mM EDTA and 1% SDS.
5. Plastic film (plastic wrap).
6. Kodak X-ray film (BiomaxMR, 35 × 43 cm, Kodak Scientific Imaging Film).
7. Autoradiography cassette Fujifilm EC-DW (Christie Group Ltd).
8. Intensifying screens (35 × 43 cm, Fisher Scientific).

2.7.2. Fluorescent Method Using Sequencer

Fluorescent Labeling

1. At this step, in order to remove excess of free primer 2, an exonuclease digestion is performed, followed by a fluorescent labeling extension of all DNA fragments amplified during the LMPCR step. For this purpose a third primer (primer 3) fluorescently labeled is used to perform five consecutive extension cycles. This primer is located right after the 3′ end of the primer 2. The primer 3 is diluted in H_2O to give 1 nmol/mL.
2. Thermocycler (Biometra or PTC™, MJ research, Inc.).
3. Exonuclease I (20.000 U/mL, New England Biolabs)

4. 5× *Taq* buffer: 50 mM Tris–HCl of pH 8.9, 200 mM NaCl, and 0.05% [w/v] gelatin (*see* **Note 1**).
5. Exonuclease I mix: 1× *Taq* buffer, 667 U/mL Exonuclease I.
6. 1 μM primer 3, fluorescent labeled (Li-Cor Bioscience).
7. Labeling stop mix: 1.56 M sodium acetate of pH 5.2 and 60 mM EDTA.
8. Precooled absolute ethanol (–20°C).
9. Precooled 80% ethanol (–20°C).
10. Dry ice or –80°C freezer.
11. Formamide-loading dye: 94% formamide, 2 mM EDTA of pH 7.7, 0.05% xylene cyanole FF, and 0.05% bromophenol blue *(11–13)*.

Sequencing Gel

1. 66-cm-long sequencing gel apparatus (LI-COR Bioscience).
2. Spacers (0.2-mm thick, LI-COR Bioscience).
3. Plastic well-forming comb (0.2-mm thick, LI-COR Bioscience).
4. 5× (0.5 M) Tris-Borate-EDTA (TBE) buffer: 500 mM Tris, 830 mM boric acid, and 10 mM EDTA of pH 8.3. Use this stock to prepare 1× (100 mM) TBE buffer.
5. 8% polyacrylamide. To prepare 1 L: 77.3 g acrylamide, 2.7 g *bis*-acrylamide, 420.42 g urea, and 200 mL of 0.5 M TBE dissolved in H_2O. Polyacrylamide solution should be kept at 4°C.
6. Gel preparation: mix 50 mL of 8% polyacrylamide with 0.5 mL of 10% ammonium persulfate (APS) and 15 μL of *N,N,N′,N′*-tetra-methylethylenediamide (TEMED). This mix is prepared immediately before pouring the solution between the glass plates. Without delay, take the gel mix into a 50-mL syringe and inject the mix between the plates, maintaining a steady flow. During pouring, the plates should be kept at a 30° angle. Any air bubbles should be avoided and removed if they form. The gel should be left to polymerize for a minimum of 2 h before use. If the gel is to be left overnight, 45 min after pouring, place a moistened paper tissue over the comb and cover the upper end of the assembly with a plastic film to prevent the gel from drying out.
7. DNA 4300 sequencer (LI-COR Bioscience).

3. Methods

3.1. DNA Purification (for 10^7–10^8 Cells)

1. Detach cells using trypsin (if needed) and sediment the cell suspension by centrifugation in 50-mL conical tubes.
2. Resuspend the cells in 2–8 mL of buffer A.

3. Add one volume (2–8 mL) of buffer A containing 1% Nonidet P40 Substitute.
4. Incubate at 4°C for 5 min.
5. Sediment nuclei by centrifugation at 4,500 × *g* for 15 min at 4°C.
6. Remove the supernatant. Resuspend nuclei in 5–10 mL of buffer A by gentle vortexing. Resediment nuclei at 4,500 × *g* for 15 min at 4°C.
7. Remove supernatant. It is recommended to leave a small volume (100–500 µL) of buffer A to facilitate resuspension of nuclei.
8. Dilute the nuclei in 1–2 mL of buffer B.
9. Add an equivalent volume of buffer C and proteinase K to a final concentration of 450 µg/mL.
10. Incubate at 37°C for 3 h (*see* **Note 3**).
11. Add RNase A to a final concentration of 150 µg/mL.
12. Incubate at 37°C for 1 h.
13. Purify DNA by extraction with one volume phenol (one or two times as needed), one volume phenol:chloroform (one or two times as needed), and one volume chloroform. Phenol extraction and phenol-chloroform extraction should be repeated if the aqueous phase is not clear.
14. Precipitate DNA in 200 mM NaCl and two volumes of precooled absolute ethanol. Ethanol should be added slowly. Make sure to mix very gently.
15. Recover DNA by spooling the floating DNA filament with a micropipet tip. If DNA is in small pieces or not clearly visible, recover DNA by centrifugation (4,500 × *g* for 15 min at 4°C), but expect RNA contamination (*see* **Note 4**). RNA contamination does not cause any problems for LMPCR. RNase digestion can be repeated if needed.
16. Wash DNA once with 10 mL of 80% ethanol.
17. Centrifuge the DNA (4,500 × *g* for 10 min at 4°C).
18. Remove supernatant and air dry DNA pellet.
19. Dissolve DNA in water at an estimated concentration of 60–100 µg/mL. The quantity of DNA can be estimated based upon the number of cells that were initially used for DNA purification. About 6 µg of DNA should be purified from 1×10^6 cells.
20. Carefully measure DNA concentration by spectrophotometry at 260 nm. Alternatively, DNA can be measured by fluorometry after staining with DAPI. Only double-stranded DNA concentration must be measured; be careful if there is RNA contamination (*see* **Note 5**).

3.2. Chemical Cleavage for DNA-Sequencing Products

In cellulo DNA analysis using LMPCR requires complete DNA-sequencing ladders from genomic DNA. Base-specific chemical modifications are performed according to Iverson and Dervan *(51)* for the A reaction and Maxam and Gilbert for the G, T + C, and C reactions. DNA from each of these base modification reactions is processed by LMPCR concomitantly with the analyzed samples and loaded in adjacent lanes on the sequencing gel to allow the identification of the precise location and sequence context of footprinted regions. The chemical modifications induced by DMS, Hz, and K_2PdCl_4 and cleaved by piperidine destroy the target base. Therefore, one must bear in mind that when analyzing a chemical-sequencing ladder, each band corresponds to a DNA fragment ending at the base preceding the one read. In this section, we will describe the chemical sequencing of genomic DNA. The cleavage protocol works optimally with 10–50 μg of genomic DNA per microtube. The required amount of DNA is ethanol precipitated and the pellet is air dried. For each base-specific reaction, we usually carried out the treatment in three microtubes containing 50 μg of genomic DNA for three different incubation times with the modifying agent in order to obtain low, medium, and high base-modification frequencies.

3.2.1. A Reaction

1. Add 160 μL of H_2O to dissolve the DNA pellet.
2. Add 40 μL of K_2PdCl_4 solution, mix carefully, and keep on ice.
3. Incubate at room temperature for 5, 10, or 15 min.
4. Add 50 μL of K_2PdCl_4 stop.
5. Add 750 μL of precooled absolute ethanol.

3.2.2. G Reaction

1. Add 5 μL of H_2O to dissolve the DNA pellet and mix.
2. Add 200 μL of DMS buffer and 1 μL of DMS, carefully mix, and keep on ice.
3. Incubate at room temperature for 30, 45, or 60 s.
4. Add 50 μL of DMS stop.
5. Add 750 μL of precooled absolute ethanol.

3.2.3. T + C Reaction

1. Add 20 μL of H_2O to dissolve the DNA pellet and mix.
2. Add 30 μL of Hz, carefully mix, and keep on ice.
3. Incubate at room temperature for 120, 210, or 300 s.
4. Add 200 μL of Hz stop.
5. Add 750 μL of precooled absolute ethanol.

3.2.4. C Reaction

1. Add 5 μL of H_2O to dissolve the DNA pellet and mix.
2. Add 15 μL of 5 M NaCl and 30 μL of Hz, carefully mix, and keep on ice.
3. Incubate at room temperature for 120, 210, or 300 s.

4. Add 200 μL of Hz stop.
5. Add 750 μL of precooled absolute ethanol.

All samples are processed as follows:

6. Mix samples well and place on dry ice for 15 min or at –80°C until the samples are frozen.
7. Centrifuge for 15 min at 16,000 × *g*.
8. Remove supernatant, recentrifuge for 1 min, and remove all the liquid using a micropipet.
9. Carefully dissolve pellet in 405 μL of H_2O.
10. Add 45 μL of 3 M sodium acetate of pH 7.0.
11. Add 1 mL of precooled absolute ethanol and mix well.
12. Leave on dry ice for 15 min or at –80°C until the samples are frozen.
13. Centrifuge for 15 min at 16,000 × *g*.
14. Remove supernatant.
15. Wash with 1 mL of precooled 80% ethanol and centrifuge for 5 min at 16,000 × *g*.
16. Remove the supernatant, spin quickly, remove the liquid with a micropipette, and air dry pellet.
17. Dissolve pellet in 50 μL H_2O. Add 50 μL of freshly prepared 2 M piperidine and mix well.
18. Secure caps with Teflon tapes and lock the caps with "lock caps."
19. Incubate at 80°C for 30 min.
20. Pool all three microtubes of the same chemical reaction in a new 1.5-mL microtube.
21. Add 105 μL H_2O, 45 μL of 3 M sodium acetate of pH 5.2, 1 μL of glycogen, and 1 mL of precooled absolute ethanol and mix well.
22. Leave on dry ice for 15 min or at –80°C until the samples are frozen.
23. Spin 15 min at 16,000 × *g*.
24. Take out the supernatant and wash twice with 1 mL of precooled 80% ethanol, respin for 5 min, and remove all the liquid using a micropipet.
25. Add 200 μL of H_2O and remove traces of remaining piperidine by drying the sample in a Speedvac concentrator.
26. Dissolve DNA in H_2O to a concentration of 0.5 μg/μL.
27. Determine the DNA strand break frequency by running the samples on a 1.5% alkaline agarose gel *(48)*. The size range of the fragments should span 100–500 bp.

3.3. Template Preparation: PCR Products

3.3.1. PCR Amplification

1. To 100 ng of purified genomic DNA in H_2O, add 50 µL of the *Taq* DNA polymerase PCR product mix and mix.
2. Cycle 35 times at 95°C for 1 min (97°C for 3 min for the first cycle), 61–73°C (1–2°C below the calculated T_m of the primer with the lowest T_m) for 2 min, and 74°C for 3 min. The last extension should be done for 10 min.
3. Add 25 µL of *Taq* DNA polymerase stop.
4. Add 400 µL of precooled absolute ethanol and mix well.
5. Leave 15 min on dry ice or at –80°C until the samples are frozen and spin 5 min at 16,000 × *g*.
6. Wash once with 1 mL of precooled 80% ethanol.
7. Spin 5 min at 16,000 × *g*.
8. Air dry DNA pellets.
9. Resuspend DNA pellets in 12 µL H_2O and add 3 µL of 5× neutral loading buffer.

3.3.2. Purification and Quantification of PCR Products

1. Load 15 µL of PCR products per well along with a 100-bp DNA size standard.
2. Migrate the PCR products on a 1% neutral agarose gel.
3. Stain the gel with ethidium bromide and photograph on a UV transilluminator. Recover the band containing the DNA fragment of expected molecular weight using a clean scalpel blade. Minimize the size of the slice by removing as much extraneous agarose as possible.
4. Crush the slice and put it in a 0.6-mL microtube pierced at the bottom, and containing a column of packed dry glass wool (*see* **Note 6**).
5. Insert the 0.6-mL microtube containing the column in a 1.5-mL microtube and spin 15 min at 7,000 × *g*. Transfer the flowthrough to a new 1.5-mL microtube. If there is still some agarose remaining, respin 15 min at 7,000 × *g*.
6. Add 50 µL of H_2O to wash the column of any remaining DNA by spinning 8 min at 7,000 × *g*. Pool all the flowthrough contents in one 1.5-mL microtube.
7. Complete the volume to 405 µL with H_2O and add 45 µL of 3 M sodium acetate of pH 7.0 and 1 mL of precooled absolute ethanol to precipitate DNA. Mix well and leave 15 min on dry ice or at –80°C until the samples are frozen. Spin 15 min at 16,000 × *g*.
8. Wash once with 1 mL of precooled 80% ethanol and spin 5 min at 16,000 × *g*.
9. Air dry DNA pellet.

10. Dissolve DNA pellets in 104 μL H_2O.
11. On a 1.5% neutral agarose gel, load aliquots of 1 and 3 μL of the DNA template dissolved in 1× neutral loading buffer along with a quantitative low DNA mass ladder.
12. Stain the gel with ethidium bromide and photograph on a UV transilluminator. The DNA concentration of the aliquots is estimated by comparison with the low DNA mass ladder band intensities, and H_2O is added to obtain a final concentration of template DNA of 10 ng/μL. The DNA template is aliquoted and stored at –20°C.

3.4. Treatment of Purified DNA and Cells with Modifying Agents

3.4.1. DMS Treatment

1. If cells are grown to confluence as a monolayer, replace the culture medium with a freshly prepared serum-free medium containing 0.2% DMS and incubate at room temperature for 6 min. If cells are grown in suspension, sediment the cells by centrifugation and remove the cell culture medium. The cells are diluted in a freshly prepared serum-free medium containing 0.2% DMS and are then incubated at room temperature for 6 min.
2. Remove the DMS-containing medium and quickly wash the cell monolayer with 10 mL of cold HBSS. Sediment cells by centrifugation if they are treated in suspension and remove the DMS-containing medium and wash the cells with 10 mL of cold HBSS.
3. Detach cells using trypsin for cells grown as monolayer.
4. Nuclei are isolated and DNA purified as described in **Subheading 3.1**.
5. Purified DNA obtained from the same cell type is treated as described in **Subheading 3.2.2**. Usually, a DMS treatment of 45 s should give a break frequency corresponding to that of the in cellulo treatment described in this section. This DNA is the in vitro treated DNA used to compare with DNA DMS-modified in cellulo (*see* **Notes 5** and 7).

3.4.2. UVC (254 nm) and UVB Irradiation

1. If cells are grown as monolayer in Petri dishes, replace cell culture medium with cold 0.9% NaCl. If cells are grown in suspension, sediment the cells by centrifugation and remove the cell culture medium. The cells are diluted in cold 0.9% NaCl at a concentration of 1×10^6 cells/mL (*see* **Note 8**) and, to avoid cellular shielding, a thin layer of the cell suspension is placed in 150-mm Petri dishes.
2. Expose the cells to 0.5–2 kJ/m² of UVC (254 nm UV) or 25–100 kJ/m² of UVB. The cells should be exposed on ice with uncovered Petri dishes. The UV intensity is measured using a UVX digital radiometer and the appropriate probe.

3. Remove the 0.9% NaCl by aspiration for cells grown as monolayer in Petri dishes or by sedimentation for cell suspensions.
4. If cells were irradiated in suspension, follow the procedure described in **Subheading 3.1** to isolate nuclei and purify DNA. After DNA purification, DNA is dissolved in H_2O at a concentration of 0.2 μg/μL. For cells cultured in Petri dishes, add in each dish 8 mL of buffer A containing 0.5% Nonidet P40 Substitute.
5. Incubate at 4°C for 5 min.
6. Scrape the cells and transfer them in a conical 50-mL tube.
7. Wash the dishes once with 8 mL of buffer A + 0.5% Nonidet P40 Substitute.
8. Continue from **step 5** of **Subheading 3.1**. After DNA purification, DNA is dissolved in H_2O at a concentration of 0.2 μg/μL.
9. Expose purified DNA to the same UVC or UVB dose as the cells. Purified DNA should be irradiated on ice and diluted in the UV irradiation buffer at a concentration of 60–75 μg/mL (*see* **Note 7**). Purified DNA should be obtained from the same type of cells as the type irradiated in cellulo (*see* **Note 9**). This DNA is used as control DNA to compare with DNA UV modified in cellulo (*see* **Notes 7** and **8**).
10. Following UV irradiation, DNA is ethanol precipitated and DNA is resuspended in H_2O at a concentration of 0.2 μg/μL.

3.4.3. DNase I Treatment

Genomic footprinting with DNase I requires cell permeabilization (*see* **Note 10**). Cells grown as a monolayer can be permeabilized while they are still attached to the Petri dish or in suspension following trypsinization. Here, we will describe cell permeabilization using lysolecithin applied to monolayer cell cultures (alternatively, the cells can be permeabilized using Nonidet P40 Substitute, *see* detailed protocol in *(18)*). For monolayer cultures, cells are grown to about 80% of confluency. For cells in suspension, cells are diluted at a concentration of approximately 1×10^6 cells/mL. To permeabilize the vast majority of cells in suspension, they must not be clumped and must not form aggregates during the permeabilization step and subsequent DNase I treatment. To achieve this, we gently flick the microtubes during permeabilization and DNase I treatment and keep the cell concentration below 2×10^6/mL.

1. For cells in monolayers, wash the cells with 5 mL solution I and permeabilize the cells by treating them with 4 mL of solution I + 0.05% lysolecithin at 37°C *(52)*.

2. Add 55 μL of DNase I (0.5 mg/mL). Incubate at 37°C for 20 min. DNase I concentration and incubation times may have to be adjusted for different cell types.
3. After 8 min, detach cells using scraper and transfer in a 15-mL tube. Replace at 37°C for the rest of the 20-min incubation.
4. Centrifuge 1 min at 450 g.
5. Resuspend the pellet in 500 μL buffer B.
6. Add 500 μL buffer C.
7. Add 450 μg/mL Proteinase K and incubate at 37°C for 3 h.
8. Add RNase A to a final concentration of 200 μg/mL and incubate at 37°C for 1 h.
9. Purify DNA by phenol-chloroform extraction (*see* **Subheading 3.1**, **step 13**).
10. Precipitate DNA in 200 mM NaCl, 1 μL glycogen, and two volumes of precooled absolute ethanol.
11. Recover DNA by centrifugation (4,500 × *g* for 15 min at 4°C), but expect RNA contamination. RNA contamination does not cause any problems for LMPCR. RNase A digestion can be repeated if needed.
12. Remove supernatant and wash DNA once with 10 mL of precooled 80% ethanol.
13. Centrifuge the DNA (4,500 × *g* for 10 min at 4°C). Remove supernatant and air dry DNA pellet.
14. Dissolve DNA in H_2O and measure DNA concentration (*see* **Subheading 3.1**, **step 20**).
15. To obtain purified DNA controls (*see* **Notes 7** and **9**), digest 40 μg of purified DNA in solution II with 10–20 ng/mL of DNase I at room temperature for 10–20 min. Stop the reaction by adding 400 μL of phenol. Extract once with phenol-chloroform and once with chloroform. Dissolve DNA in H_2O at a concentration of 0.2 μg/μL.

3.5. Conversion of Modified Bases to DNA Single-Strand Breaks

When purified DNA or living cells are treated with DMS or UV, DNA base modifications are induced **(Table 3)**. These modifications must be converted to single-strand breaks before running LMPCR. Following UV exposure, CPD and 6-4PP are converted individually because they use different conversion procedures **(Table 3)**. On the other hand, DNase I digestion directly generates DNA strand breaks suitable for LMPCR without any conversion procedures. Before running LMPCR, the DNA strand break frequency must be determined by running the samples on a 1.5% alkaline agarose gel *(49)*. The size range of the fragments should span 200–2,000 bp (*see* **Note 7**).

3.5.1. DMS-Induced Base Modifications (See Fig. 1)

1. Dissolve DNA (10–50 µg) in 50 µL H_2O, add 50 µL of 2 M piperidine, and mix well.
2. Samples are processed as described in **Subheading 3.2**, **steps 18–27**.
3. Dissolve DNA in H_2O to a concentration of 0.2 µg/µL.

3.5.2. UV-Induced Base Modifications

CPD (See Fig. 2)

1. To specifically cleave CPD, dissolve 10 µg of UV-irradiated DNA in 50 µL H_2O and add 50 µL of T_4 endo V mix.
2. Incubate at 37°C for 1 h.
3. Perform the photolyase digestion to remove the overhanging dimerized base that would otherwise prevent ligation *(8)*; add 10 µL of the photolyase mix.
4. Leaving their caps opened, cover the microtubes with a plastic film to prevent UVB-induced damage and place open ends 2–3 cm from a UVA black light blue for 1 h.
5. Add 200 µL of 1% SDS and 100 µL H_2O and mix well.
6. Extract DNA using one volume (400 µL) phenol, one volume phenol:chloroform, and one volume chloroform.
7. To precipitate DNA, add 20 µL of 5 M NaCl, 1 mL of precooled absolute ethanol and mix well.
8. Leave 15 min on dry ice or at –80°C until the samples are frozen and spin 15 min at 16,000 × *g*.
9. Wash once with 1 mL of precooled 80% ethanol.
10. Spin 5 min at 16,000 × *g*.
11. Air dry the pellet and dissolve DNA in H_2O to a concentration of 0.2 µg/µL.

6-4PP (See Fig. 3)

1. Dissolve DNA (10–50 µg) in 50 µL of H_2O, add 50 µL of 2 M piperidine, and mix well.
2. Samples are processed as described in **Subheading 3.2**, **steps 18–27**.
3. Dissolve DNA in H_2O to a concentration of 0.2 µg/µL.

3.6. Ligation-Mediated Polymerase Chain Reaction Technology

The PE, ligation, and PCR steps are carried out in 0.6-mL microtubes and a thermocycler is used for all incubations.

3.6.1. Primer Extension (Steps II and III, Fig. 5) (See Note 11)

1. Mix 0.5–2 µg of genomic DNA with the appropriate polymerase extension mix (*Pfu* exo$^-$ or *Vent* exo$^-$ extension mix) for a final volume of 30 µL
2. Denature DNA at 98°C. Incubate the samples at the annealing temperature for 4 min, then incubate at 75°C for 10 min. Finally, the samples are cooled to 4°C.

3.6.2. Ligation (Step IV, Fig. 5)

1. To the primer extension reaction, add 45 µL of the ligation mix and mix well.
2. Incubate at 18°C for a minimum of 2 h.
3. Precipitate DNA by adding 30 µl of T_4 ligase stop mix, 300 µL of precooled absolute ethanol and mix well.
4. Leave 15 min on dry ice or at –80°C until the samples are frozen and spin 15 min at 16,000 × *g*.
5. Wash once with 500 µL of precooled 80% ethanol.
6. Spin 5 min at 16,000 × *g*.
7. Remove supernatant and air dry DNA pellets.
8. Dissolve DNA in 50 µL H_2O.

3.6.3. Polymerase Chain Reaction (Steps V and VI, Fig. 5) (See Note 11)

1. Add 50 µL of the appropriate polymerase amplification mix (*Pfu* exo⁻, *Vent* exo⁻, or *Taq* amplification mix) and mix.
2. Cycle 22 times as described in **Table 4**. The last extension should be done for 10 min to fully extend all DNA fragments.

3.7. LMPCR-Amplified DNA Fragment Analysis

The PCR-amplified fragments are separated by electrophoresis through an 8% polyacrylamide/7 M urea gel, 0.4-mm thick and 66 cm long, then transferred to a nylon membrane by electroblotting *(11–13)*.

3.7.1. Conventional Radioactive Method

Precipitation, Gel Electrophoresis, and Electroblotting (Step VII, Fig. 5)

1. Add 25 µL of the appropriate stop mix: *Pfu* exo⁻, *Taq*, or *Vent* exo⁻ stop mix, depending of the polymerase mix used for amplification.
2. Add 400 µL of precooled absolute ethanol and mix well.
3. Leave 15 min on dry ice or at –80°C until the samples are frozen and spin 15 min at 16,000 × *g* at 4°C.
4. Wash once with 500 µL of precooled 80% ethanol.
5. Spin 5 min at 16,000 × *g*.
6. Air dry DNA pellets.
7. Dissolve DNA pellets in 7.5 µL of formamide-loading dye in preparation for sequencing gel electrophoresis. For the sequence samples G, A, T + C, and C, it is often advisable to dissolve DNA pellets in 15 µL of formamide-loading dye.
8. Prerun the 8% polyacrylamide gel in 100 mM TBE, until the temperature of the gel reaches 50°C.
9. To denature DNA, heat the samples at 95°C for 2 min and then keep them on ice prior to loading.
10. Wash the wells of the gel using a syringe.
11. Load an aliquot of 1.5–2 µL.
12. Run the gel at the voltage and power necessary to maintain the temperature of the gel at 50°C. This will ensure that the DNA remains denatured.

13. Stop the gel when the green dye (xylene cyanole FF) reaches 1–2 cm from the bottom of the gel.
14. Separate the glass plates using a spatula, and then remove one of the plates by lifting it carefully. The gel should stick to the less treated plate (*see* **Note 12**).
15. Cover the lower part of the gel (approximately 40–42 cm) with a clean Whatman 3-MM Chr paper, carefully remove the gel from the glass plate and cover it with a plastic film.
16. On the bottom plate of the electroblotter, individually layer three sheets of Whatman 17-MM Chr paper presoaked in 100 mM TBE and squeeze out the air bubbles between the paper layers by rolling with a bottle.
17. Add 150 mL of 100 mM TBE on the top layer and place the gel quickly on the Whatman 17-MM Chr papers before TBE is absorbed. Remove all air bubbles under the gel by gently rolling a 25-mL pipet.
18. Remove the plastic film and cover the gel with a positively charged nylon membrane presoaked in 100 mM TBE, remove all air bubbles by gently rolling a 25-mL pipet, then cover with three layers of presoaked Whatman 17-MM Chr paper, and squeeze out air bubbles with rolling bottle. Paper sheets can be reused several times.
19. Place the upper electrode onto the paper.
20. Electrotransfer for 45 min at 2 A. The voltage should be at approximately 10–15 V.
21. UV-crosslink (1,000 J/m^2 of UVC) the blotted DNA to the membrane, taking care to expose the DNA side of the membrane. If probe stripping and rehybridization are planned, keep the membrane damp.

Preparation of Single-Stranded Hybridization Probes (Step VIII, Fig. 5)

The [^{32}P]-dCTP-labeled single-stranded probe is prepared by 30 cycles of repeated linear primer extension using *Taq* DNA polymerase. Primer 2 (or primer 3, *see* **Note 13**) is extended on a double-stranded template which can be a plasmid or a PCR product. The latter is produced by using two opposing primers 2 separated by a distance of 150–450 bp. Alternatively, any pair of gene-specific primers suitable for amplifying a DNA fragment containing a suitable probe sequence (*see* **Note 13**) can be employed.

1. Prepare 150 µL of the isotopic labeling mix.
2. Cycle 30 times at 95°C for 1 min (97°C for 3 min for the first cycle), 60–68°C for 2 min, and 74°C for 3 min.
3. Transfer the mixture to a conical 1.5-mL microtube with screw cap containing 50 µL of 10 M ammonium acetate, 1 µL of glycogen, and 400 µL of precooled absolute ethanol.

4. Mix well, leave 5 min at room temperature, and spin 5 min at 16,000 × *g*.
5. Transfer the supernatant into a new 1.5-mL microtube. Using a Geiger counter, compare the counts per minute between the pellet (probe) and the supernatant; counts from the probe should be equal or superior to the counts from the supernatant for optimal results.
6. Dissolve the probe in 100 µL of TE buffer.
7. Add the probe to 6–8 mL of hybridization buffer and keep the probe at 65°C.

Hybridization **(Step VII, Fig. 5)**, Washing, and Autoradiography

1. Prehybridize the membrane with 20 mL of hybridization buffer at 65°C for 20 min.
2. Decant the hybridization buffer and add the single-stranded hybridization probe in 6–8 mL of hybridization buffer.
3. Hybridize at 65°C overnight.
4. Wash the membrane with prewarmed (65°C) washing buffers. The membrane is placed into a tray on an orbital shaker. Wash with buffer I for 10 min and with buffer II three times for about 10 min each time.
5. Wrap the membrane in plastic film. Do not let the membrane become dry if stripping and rehybridization are planned after exposure of the film.
6. Expose membrane to X-ray films with intensifying screen at –80°C. The exposure time depends of the cpm count evaluated with a Geiger counter. Nylon membranes can be rehybridized if more than one set of primers has been included in the primer extension and amplification reactions *(11–13)*. Probes can be stripped by soaking the membranes in boiling 0.1% SDS solution twice for 5–10 min each time.

3.7.2. Fluorescent Method Using Sequencer

Labeling Extension

1. Add 3 µL of exonuclease I mix.
2. Incubate on PCR at 37°C for 30 min and at 76°C for 20 min.
3. Add 1 µL of primer 3 fluorescently labeled.
4. Cycle five times as described in **Table 5**. The last extension should be done for 10 min to fully extend all DNA fragments.
5. Add 25 µL of labeling stop mix.
6. Add 300 µL of precooled absolute ethanol and mix well.
7. Leave 15 min on dry ice or at –80°C until the samples are frozen and spin 15 min at 16,000 × *g* in a centrifuge at 4°C.
8. Wash once with 500 µL of precooled 80% ethanol.
9. Spin 5 min at 16,000 × *g* at 4°C.
10. Air dry DNA pellets.

Table 5
PCR extension labeling program for *Taq* DNA polymerase

Cycles	Denaturation (*T* in °C for DF in s)	Annealing (*T* is the T_m of the oligonucleotide for *D* in s)	Polymerization (*D* in s) *T* is the same for all cycles: 74°C
0	95 for 120		
1	95 for 45	T_m –2°C for 120	180
Repeat cycle 1, 3 more times			
2	98 for 45	T_m –2°C for 120	600

11. Dissolve DNA pellets in 15 µL of formamide-loading dye in preparation for sequencing gel electrophoresis. For the sequence samples G, A, T + C, and C, it is often advisable to dissolve DNA pellets in 30 µL of formamide-loading dye.
12. Incubate 2 min at 95°C and then keep on ice before loading on the sequencing gel.

Sequencer Electrophoresis

The labeled DNA fragments are separated by electrophoresis through an 8% polyacrylamide/7 M urea gel, 0.2-mm thick and 66–cm long using a LI-COR DNA 4300 sequencer (*see* **Note 14**).

1. Prerun the 8% polyacrylamide gel 25 min in order to set the temperature at 47°C. Running buffer is 100 mM TBE. Before loading the samples, wash the wells thoroughly using a syringe.
2. Load an aliquot of 1–1.5 µL.
3. Run the gel with a constant power of 100 W during 11 h.

4. Notes

1. Originally, *Pfu* exo$^-$ and *Taq* buffers were prepared using KCl which was, however, shown to stabilize secondary DNA structures, thus preventing an optimal polymerization *(53)*. The use of NaCl prevents, to some extent, the ability of DNA to form secondary structures. This is particularly helpful when GC-rich regions of the genome are being investigated.
2. Primers should be selected to have a higher T_m at the 5′ end than in the 3′ end. This higher annealing capacity of the 5′-end lowers false priming, thus allowing a more specific extension and less background *(54)*. A guanine or a cytosine

residue should also occur at the 3′ end. This stabilizes the annealing and facilitates the initiation of the primer extension. It is important that the selected primer does have long runs of purines or pyrimidines, does not form loops or secondary structure, and does not anneal with itself. If primer dimerization occurs, less primer will be available for annealing and polymerization will not be optimal. The purity of the primers is verified on a 20% polyacrylamide/7 M urea gel (to prepare 500 mL: dissolve 96.625 g acrylamide, 3.375 g *bis*-acrylamide, and 210.21 g urea in 100 mM TBE); if more than one band is found, the primer is reordered. The primers are also tested in a conventional PCR to prepare the template for the probe synthesis (*see* **Note 13**).

3. The genomic DNA used for LMPCR needs to be very clean and undegraded. Any shearing of the DNA during preparation and handling before the PE must be avoided. After an incubation of 3 h, if clumps of nuclei are still visible, proteinase K at a final concentration of 450 μg/mL should be added and the sample reincubated at 37°C for another 3 h.
4. If no DNA can be seen, add glycogen (1–2 μg) to the DNA solution and put the DNA on dry ice or in −80°C freezer until the samples are frozen and centrifuge the DNA (5,000 × *g* for 20 min at 4°C). This should help DNA recovery but increases the probability of RNA contamination.
5. Because in cellulo DNA analysis is based on comparison of DNA samples modified in cellulo with DNA control modified in vitro, given the quantitative characteristic and high sensitivity of LMPCR technology, the DNA concentrations should be as accurate as possible. Indeed, it is critical to start LMPCR with similar amounts of DNA in every sample to be analyzed. The method to evaluate DNA concentration should measure only double-stranded nucleic acids. RNA contamination does not affect LMPCR, although it can however interfere with the precise measurement of the DNA concentration.
6. The bottom of a capless 0.6-mL microtube can be easily pierced with a heated needle. It is important to emphasize that the hole should be made as small as possible for the column to efficiently retain agarose. The pierced microtube is packed with wetted glass wool. Three successive centrifugation steps of 1 min each at 16,000 × *g* are necessary to compact and dry the glass wool. The water is recovered in a capless 1.5-mL microtube. If glass wool is found with the effluent, the column should be discarded. A final 5 min centrifugation at 16,000 × *g* should be carried out to ensure the glass wool is fully compacted and dry. The glass wool column is stored at room temperature in a new capless

1.5-mL microtube and covered with a plastic film to protect the column from dust. In this way, the column can be stored indefinitely until it is used.

7. The DNA break frequency is even more critical than the DNA concentration. For DMS and UV, the base-modification frequency determines the break frequency following conversion of the modified bases to single-strand breaks, whereas for DNase I, the frequency of cleavage is exactly the break frequency. The break frequency must be similar among the samples to be analyzed. It should not average more than one break per 150 bp for in cellulo DNA analysis, the optimal break frequency varying from one break per 200 bp to one break per 2,000 bp. When the break frequency is too high, we typically observe dark bands over the bottom half of the autoradiogram and very pale bands over the upper half, reflecting the low number of long DNA fragments. In summary, to make the comparison of the in cellulo modified DNA sample with an in vitro DNA control easily interpretable and valid, the amount of DNA and the break frequency must be similar between the samples to be compared. On the other hand, it is not so critical that the break frequency of the sequence ladders (G, A, T + C, and C) be similar to that of the samples to be studied. However, to facilitate sequence reading, the break frequency should be similar between the sequence reactions. It is often necessary to load less DNA for the sequence ladders.
8. If the cell density is too high, multiple cell layers will be formed and the upper cell layer will obstruct the lower ones. This will result in an inhomogeneous DNA photoproduct frequency.
9. It is imperative that the in vitro DNA samples used as DNA control and the in cellulo samples come from the same cell type. For instance, differing cytosine methylation patterns of genomic DNA from different cell types affect photoproduct formation *(2, 17)* and give altered DNase I cleavage patterns *(5)*.
10. A nearly ideal chromatin substrate can be maintained in permeabilized cells. Nonionic detergents such as lysolecithin *(52)* and Nonidet P40 Substitute *(36)* permeabilize the cell membrane sufficiently to allow the entry of DNase I. Conveniently, this assay can be performed with cells either in a suspension or in a monolayer. One concern is that permeabilized cells will lyse after a certain amount of time in a detergent; thus, care must be taken to monitor cell integrity by microscopy during the course of the experiment. A further difficulty with the permeabilization technique concerns the relatively narrow detergent concentration range over which the assay can be performed. Each cell type appears to require

specific conditions for the detergent cell permeabilization. Furthermore, the DNase I concentration must be calibrated for each cell type to produce an appropriate cleavage frequency. Optimally, the in cellulo DNase I protocol works better if the enzyme has cleaved the DNA backbone every 1.5–2 kb. Cutting frequencies greater than 1 kb are associated with higher LMPCR backgrounds because the number of 3′-OH ends is much higher, making the suppression of the extension of these ends more difficult.

11. So far with the method using the automated Sequencer, we have tested and used only the *Pfu* exo$^-$ for primer extension and the *Taq* DNA polymerase for the PCR amplification.
12. To facilitate sequencing gel removal following migration, it is crucial to siliconize the inner face of both glass plates prior to pouring the gel. For security, cost effectiveness, efficiency and saving time, we recommend to treat the glass plates with RAIN-AWAY™ solution (Wynn's Canada, product no. 63020). We apply 0.75 mL on one plate and 1.5 mL on the other before each utilization as specified by the manufacturer. In this way, the gel is easier to pour and will tend to stick on the less siliconized plate.
13. If a third primer (primer 3) is used to make the probe, it should be selected from the same strand as the amplification primer (primer 2), just 5′ to primer 2 sequence and with no or no more than seven to eight bases of overlap on this primer, and have a T_m of 60–68°C. As first reported by Hornstra and Yang *(44, 55, 56)*, we use the primer 2 employed in the amplification step and we produce the probe from PCR products. Such probes cost less (no primer 3) and are more convenient (the preparation of the PCR products permits the testing of primers).
14. To have well-shaped wells, we treat the upper part of both glass plates with a silane-acetic acid solution. We mix 50 μL of silane with 50 μL of 10% acetic acid. Before each utilization, using a cotton-tip applicator, we apply 100 μL of this solution on the plate areas where we place the comb.

Acknowledgments

This work was supported by the Canadian Genetic Diseases Network (MRC/NSERC NCE program) and the Canada Research Chair. R. Drouin holds the Canada Research Chair in "Genetics, Mutagenesis and Cancer."

References

1. Pfeifer, G. P., Tanguay, R. L., Steigerwald, S. D., and Riggs, A. D. (1990). In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1, *Genes Dev.* **4**, 1277–1287.
2. Pfeifer, G. P., Drouin, R., Riggs, A. D., and Holmquist, G. P. (1991). In vivo mapping of a DNA adduct at nucleotide resolution: detection of pyrimidine (6-4) pyrimidone photoproducts by ligation-mediated polymerase chain reaction, *Proc. Natl. Acad. Sci. U S A.* **88**, 1374–1378.
3. Chen, C. J., Li, L. J., Maruya, A., and Shively, J. E. (1995). In vitro and in vivo footprint analysis of the promoter of carcinoembryonic antigen in colon carcinoma cells: effects of interferon gamma treatment, *Cancer Res.* **55**, 3873–3882.
4. Tornaletti, S. and Pfeifer, G. P. (1995). UV light as a footprinting agent: modulation of UV-induced DNA damage by transcription factors bound at the promoters of three human genes, *J. Mol. Biol.* **249**, 714–728.
5. Pfeifer, G. P. and Riggs, A. D. (1991). Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR, *Genes Dev.* **5**, 1102–1113.
6. Mueller, P. R. and Wold, B. (1989). In vivo footprinting of a muscle specific enhancer by ligation mediated PCR, *Science.* **246**, 780–786.
7. Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B., and Riggs, A. D. (1989). Genomic sequencing and methylation analysis by ligation mediated PCR, *Science.* **246**, 810–813.
8. Pfeifer, G. P., Drouin, R., Riggs, A. D., and Holmquist, G. P. (1992). Binding of transcription factors creates hot spots for UV photoproducts in vivo, *Mol. Cell. Biol.* **12**, 1798–1804.
9. Church, G. M. and Gilbert, W. (1984). Genomic sequencing, *Proc. Natl. Acad. Sci. U S A.* **81**, 1991–1995.
10. Pfeifer, G. P. (1992). Analysis of chromatin structure by ligation-mediated PCR, *PCR Methods Appl.* **2**, 107–111.
11. Pfeifer, G. P. and Riggs, A. D. (1993). Genomic footprinting by ligation mediated polymerase chain reaction. In *PCR Protocols: Current Methods and Applications* (White, B., Ed.), pp 169–181, Humana, Totowa, NJ.
12. Pfeifer, G. P. and Riggs, A. D. (1993). Genomic sequencing, *Methods Mol. Biol.* **23**, 169–181.
13. Pfeifer, G. P., Singer-Sam, J., and Riggs, A. D. (1993). Analysis of methylation and chromatin structure, *Methods Enzymol.* **225**, 567–583.
14. Gao, S., Drouin, R., and Holmquist, G. P. (1994). DNA repair rates mapped along the human PGK1 gene at nucleotide resolution, *Science.* **263**, 1438–1440.
15. Tornaletti, S. and Pfeifer, G. P. (1994). Slow repair of pyrimidine dimers at p53 mutation hotspots in skin cancer, *Science.* **263**, 1436–1438.
16. Rodriguez, H., Drouin, R., Holmquist, G. P., O'Connor, T. R., Boiteux, S., Laval, J., Doroshow, J. H., and Akman, S. A. (1995). Mapping of copper/hydrogen peroxide-induced DNA damage at nucleotide resolution in human genomic DNA by ligation-mediated polymerase chain reaction, *J. Biol. Chem.* **270**, 17633–17640.
17. Drouin, R. and Therrien, J. P. (1997). UVB-induced cyclobutane pyrimidine dimer frequency correlates with skin cancer mutational hotspots in p53, *Photochem. Photobiol.* **66**, 719–726.
18. Drouin, R., Therrien, J. P., Angers, M., and Ouellet, S. (2001). In vivo DNA analysis. In *Methods in Molecular Biology* (Moss, T., Ed.), pp 175–219, Humana, Totowa, NJ.
19. Dai, S. M., Chen, H. H., Chang, C., Riggs, A. D., and Flanagan, S. D. (2000). Ligation-mediated PCR for quantitative in vivo footprinting, *Nat. Biotechnol.* **18**, 1108–1111.
20. Dai, S. M., O'Connor, T. R., Holmquist, G. P., Riggs, A. D., and Flanagan, S. D. (2002). Ligation-mediated PCR: robotic liquid handling for DNA damage and repair, *Biotechniques.* **33**, 1090–1097.
21. McLellan, J. A. (2001). Osmium tetroxide modification and the study of DNA-protein interactions, *Methods Mol. Biol.* **148**, 121–34.
22. Rozek, D. and Pfeifer, G. P. (1993). In vivo protein-DNA interactions at the c-jun promoter: preformed complexes mediate the UV response, *Mol. Cell. Biol.* **13**, 5490–5499.
23. Cartwright, I. L. and Kelly, S. E. (1991). Probing the nature of chromosomal DNA-protein contacts by in vivo footprinting, *Biotechniques.* **11**, 188–190, 192–184, 196 passim.
24. Maxam, A. M. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages, *Methods Enzymol.* **65**, 499–560.

25. Chin, P. L., Momand, J., and Pfeifer, G. P. (1997). In vivo evidence for binding of p53 to consensus binding sites in the p21 and GADD45 genes in response to ionizing radiation, *Oncogene.* **15**, 87–99.
26. Angers, M., Drouin, R., Bachvarova, M., Paradis, I., Marceau, F., and Bachvarov, D. R. (2000). In vivo protein-DNA interactions at the kinin B(1) receptor gene promoter: no modification on interleukin-1 beta or lipopolysaccharide induction, *J. Cell. Biochem.* **78**, 278–296.
27. Becker, M. M. and Wang, J. C. (1984). Use of light for footprinting DNA in vivo, *Nature.* **309**, 682–687.
28. Pfeifer, G. P. and Tornaletti, S. (1997). Footprinting with UV irradiation and LMPCR, *Methods.* **11**, 189–196.
29. Pfeifer, G. P., Chen, H. H., Komura, J., and Riggs, A. D. (1999). Chromatin structure analysis by ligation-mediated and terminal transferase-mediated polymerase chain reaction, *Methods Enzymol.* **304**, 548–571.
30. Cadet, J., Anselmino, C., Douki, T., and Voituriez, L. (1992). Photochemistry of nucleic acids in cells, *J. Photochem. Photobiol. B.* **15**, 277–298.
31. Mitchell, D. L. and Nairn, R. S. (1989). The biology of the (6-4) photoproduct, *Photochem. Photobiol.* **49**, 805–819.
32. Holmquist, G. P. and Gao, S. (1997). Somatic mutation theory, DNA repair rates, and the molecular epidemiology of p53 mutations, *Mutat. Res.* **386**, 69–101.
33. Gale, J. M., Nissen, K. A., and Smerdon, M. J. (1987). UV-induced formation of pyrimidine dimers in nucleosome core DNA is strongly modulated with a period of 10.3 bases, *Proc. Natl. Acad. Sci. U S A.* **84**, 6644–6648.
34. Gale, J. M. and Smerdon, M. J. (1990). UV induced (6-4) photoproducts are distributed differently than cyclobutane dimers in nucleosomes, *Photochem. Photobiol.* **51**, 411–417.
35. Mitchell, D. L., Nguyen, T. D., and Cleaver, J. E. (1990). Nonrandom induction of pyrimidine-pyrimidone (6-4) photoproducts in ultraviolet-irradiated human chromatin, *J. Biol. Chem.* **265**, 5353–5356.
36. Rigaud, G., Roux, J., Pictet, R., and Grange, T. (1991). In vivo footprinting of rat TAT gene: dynamic interplay between the glucocorticoid receptor and a liver-specific factor, *Cell.* **67**, 977–986.
37. Miller, M. R., Castellot, J. J., Jr., and Pardee, A. B. (1978). A permeable animal cell preparation for studying macromolecular synthesis. DNA synthesis and the role of deoxyribonucleotides in S phase initiation, *Biochemistry.* **17**, 1073–1080.
38. Contreras, R. and Fiers, W. (1981). Initiation of transcription by RNA polymerase II in permeable, SV40-infected or noninfected, CVI cells; evidence for multiple promoters of SV40 late transcription, *Nucleic Acids Res.* **9**, 215–236.
39. Tanguay, R. L., Pfeifer, G. P., and Riggs, A. D. (1990). PCR-aided DNaseI footprinting of single copy gene sequences in permeabilized cells, *Nucleic Acids Res.* **18**, 5902.
40. Tormanen, V. T., Swiderski, P. M., Kaplan, B. E., Pfeifer, G. P., and Riggs, A. D. (1992). Extension product capture improves genomic sequencing and DNase I footprinting by ligation-mediated PCR, *Nucleic Acids Res.* **20,** 5487–5488.
41. Tornaletti, S., Bates, S., and Pfeifer, G. P. (1996). A high-resolution analysis of chromatin structure along p53 sequences, *Mol. Carcinog.* **17**, 192–201.
42. Szabo, P. E., Pfeifer, G. P., and Mann, J. R. (1998). Characterization of novel parent-specific epigenetic modifications upstream of the imprinted mouse H19 gene, *Mol. Cell. Biol.* **18**, 6767–6776.
43. Garrity, P. A. and Wold, B. J. (1992). Effects of different DNA polymerases in ligation-mediated PCR: enhanced genomic sequencing and in vivo footprinting, *Proc. Natl. Acad. Sci. U S A.* **89**, 1021–1025.
44. Hornstra, I. K. and Yang, T. P. (1994). High-resolution methylation analysis of the human hypoxanthine phosphoribosyltransferase gene 5′ region on the active and inactive X chromosomes: correlation with binding sites for transcription factors, *Mol. Cell. Biol.* **14**, 1419–1430.
45. Angers, M., Cloutier, J. F., Castonguay, A., and Drouin, R. (2001). Optimal conditions to use Pfu exo(⁻) DNA polymerase for highly efficient ligation-mediated polymerase chain reaction protocols, *Nucleic Acids Res.* **29**, E83.
46. Rouget, R., Vigneault, F., Codio, C., Rochette, C., Paradis, I., Drouin, R., and Simard, L. R. (2005). Characterization of the survival motor neuron (SMN) promoter provides evidence for complex combinatorial regulation in undifferentiated and differentiated P19 cells, *Biochem J.* **385**, 433–443.
47. Vigneault, F. and Drouin, R. (2005). Optimal conditions and specific characteristics of Vent exo⁻DNA polymerase in ligation-mediated polymerase chain reaction protocols, *Biochem. Cell. Biol.* **83**, 147–165.
48. Drouin, R., Gao, S., and Holmquist, G. P. (1996). Agarose gel electrophoresis for DNA damage analysis. In *Technologies for Detection*

of DNA Damage and Mutations (Pfeifer, G.P., Ed.), pp 37–43, Plenum, New York.

49. Drouin, R., Rodriguez, H., Holmquist, G. P., and Akman, S. A. (1996). Ligation-mediated PCR for analysis of oxidative damage. In *Technologies for Detection of DNA Damage and Mutations* (Pfeifer, G.P., Ed.), pp 211–225, Plenum, New York.
50. Mueller, P. R. and Wold, B. (1991). Ligation-mediated PCR: applications to genomic footprinting, *Methods*, 20–31.
51. Iverson, B. L. and Dervan, P. B. (1987). Adenine specific DNA chemical sequencing reaction, *Nucleic Acids Res.* 15, 7823–7830.
52. Zhang, L. and Gralla, J. D. (1989). In situ nucleoprotein structure at the SV40 major late promoter: melted and wrapped DNA flank the start site, *Genes Dev.* 3, 1814–1822.
53. Fry, M. and Loeb, L. A. (1994). The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure, *Proc. Natl. Acad. Sci. U S A.* **91**, 4950–4954.
54. Rychlik, W. (1993). Selection of primers for polymerase chain reaction. In *PCR Protocols: Current Methods and Applications* (White, B., Ed.), pp 31–40, Humana, Totowa, NJ.
55. Hornstra, I. K. and Yang, T. P. (1992). Multiple in vivo footprints are specific to the active allele of the X-linked human hypoxanthine phosphoribosyltransferase gene 5′ region: implications for X chromosome inactivation, *Mol. Cell. Biol.* 12, 5345–5354.
56. Hornstra, I. K. and Yang, T. P. (1993). In vivo footprinting and genomic sequencing by ligation-mediated PCR, *Anal. Biochem.* 213, 179–193.

Chapter 21

Atomic Force Microscopy Imaging and Probing of DNA, Proteins, and Protein–DNA Complexes: Silatrane Surface Chemistry

Yuri L. Lyubchenko, Luda S. Shlyakhtenko, and Alexander A. Gall

Summary

Despite their rather recent invention, atomic force microscopes are widely available commercially. AFM and its special modifications (tapping mode and noncontact operation in solution) have been successfully used for topographic studies of a large number of biological objects including DNA, RNA, proteins, cell membranes, and even whole cells. AFM was also successfully applied to studies of nucleic acids and various protein–DNA complexes. Part of this success is due to the development of reliable sample preparation procedures. This chapter describes one of the approaches based on chemical functionalization of mica surface with 1-(3-aminopropyl) silatrane (APS). One of the most important properties of APS-mica approach is that the sample can be deposited on the surface in a wide range of ionic strengths, in the absence of divalent cations and a broad range of pH. In addition to imaging of dried sample, APS-mica allows reliable and reproducible time lapse imaging in aqueous solutions. Finally, APS mica is terminated with reactive amino groups that can be used for covalent and ionic attachment of molecules for AFM force spectroscopy studies. The protocols for the preparation of APS, functionalization with APS mica and AFM probes, preparation of samples for imaging in air and in aqueous solutions, and force spectroscopy studies are outlined. All these applications are illustrated with a few examples.

Key words: Atomic force microscopy (AFM), Force spectroscopy, Surface chemistry, Mica functionalization, Silanes, Silatranes, DNA local structures, Protein–DNA complexes.

1. Introduction

Atomic Force Microscopy (AFM; another name is Scanning Force Microscopy, SFM) and its predecessor, Scanning Tunneling Microscope (STM), are relatively new tools with enormous potential

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_21

importance to structural biology. The prototype STM instrument was conceived by Binnig et al. *(1)*, an invention for which Binnig and Rohrer were awarded the 1986 Nobel Prize in Physics. AFM is a direct descendent of this early instrument *(2)*, but is capable of imaging nonconducting, as well as conducting surfaces. Unfortunately, early artifactual STM images of a graphite surface mimicking DNA molecules *(3, 4)* have seriously undermined hopes for immediate success for applications of both STM and AFM to structural biology. A serious practical limitation to the application of AFM to structural and conformational studies of DNA and its complexes with proteins and other biological macromolecules has been sample preparation. The macromolecules must be tethered to the substrate surface in order to avoid resolution-limiting motion caused by the sweeping tip during scanning. The breakthrough in reliable and reproducible imaging of DNA with AFM was made in early 1990s due to development of methods for sample preparation. Several of such methods were developed simultaneously in a number of laboratories *(5–14)*. A major feature of these methods is the use of a specially prepared surface that holds (usually electrostatically) the sample in place during scanning. We *(8, 14–18)* initially worked out a procedure for chemical modification of mica. A weak cationic surface is obtained if aminopropyltriethoxy silane (APTES) is used to functionalize the mica surface with amino groups (AP-mica). AP-mica allowed us to routinely perform visualization of DNA with AFM, achieving resolution as good as that of traditional electron microscopy (EM). We have also shown that AP-mica holds nucleic acids under physiological conditions, allowing for the ability to image *in situ*. Remarkably, we have achieved resolution for DNA in solution exceeding that of EM for dried DNA samples. Recently, we have developed an improved surface chemistry utilizing a more hydrolytically stable silatrane such as 3-aminopropylsilatrane (APS) instead of silanes. This method provides a more reproducible, robust surface for topographic, *(19–32)* as well as for force spectroscopy AFM applications *(33–37)*. The scheme for synthesis of APS is presented in **Fig. 1a**.

The method of functionalization of mica is based on covalent attachment of 3-aminopropyl silatrane (APS) to the surface of mica *(24)*. As illustrated by the scheme in **Fig. 1b**, the silicon residues of APS are bound covalently to exposed hydroxyl groups of the freshly cleaved mica resulting in a strong covalent attachment of alkylamino residues to the surface, giving it properties similar to an anion exchange resin used in affinity chromatography. The amino group after protonation in water solution becomes positively charged in a rather broad range of pH (aliphatic amines have a pK of around 10.5). Therefore, DNA, which is a negatively charged polymer, should adhere to this surface strongly. APS-mica preparation requires a very low concentration of APS

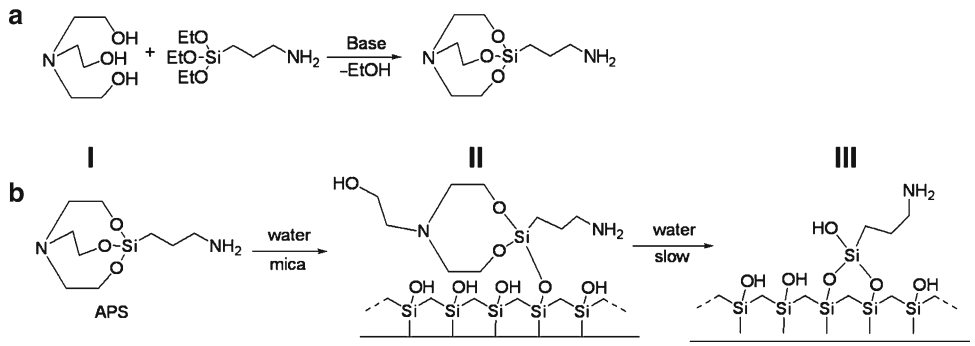


Fig. 1. (**a**) Scheme for the synthesis of APS. Chemical formula for APS is 1-(3-aminopropyl)silatrane (2,8,9-trioxa-5-aza-1-silabicyclo[3.3.3]undecane) (CAS [17869-27-1]). Molecular weight: 232.36. Molecular formula: $C_9H_{20}N_2O_3Si$. (**b**) Scheme for reaction of 3-aminopropylsilatrane (I, APS) with hydroxyl groups on a silicon surface. The initial adduct II may react with a second surface OH group forming III in a reversible equilibrium.

(~150 μM) allowing to prepare a smooth surface, so DNA and DNA–protein complexes can be visualized unambiguously (*see* **Figs. 2a, b** respectively).

The features of this procedure of sample preparation are as follows *(26)*:

1. Binding of DNA to APS-mica is insensitive to the type of buffer and presence of Mg^{2+} or other di- and miltivalent cations; hence, sample preparation can be done in a variety of conditions.
2. Deposition can be done in a wide variety of pH and over a wide range of temperatures.
3. Once prepared, samples are stable and do not absorb any contaminants for months with minimal precautions for storing.
4. As low as 2–10 ng of DNA is sufficient for preparation of one sample.
5. APS-mica is terminated with primary amines that can be used for covalent attachment of biomacromolecules.

These characteristics of APS-mica were crucial for routine imaging nucleic acids of various conformations (e.g., *(19, 20, 22, 24, 27, 38, 39)*); nucleoprotein complexes of different types (e.g., *(25, 26, 30–32)*), protein assemblies such as amyloid fibrils, spheres, and toroids *(33, 35, 37, 40–47)*; and for AFM force spectroscopy studies *(33–35, 47, 48)*.

We have shown that amino-terminated surface (APS-treated mica and AFM tips) can be functionalized further using molecules containing amine-reactive groups such as glutaraldehyde *(33, 34)* and hydroxysuccinamides *(49)*. This property of APS-functionalized surfaces was used for covalent attachment of biomolecules

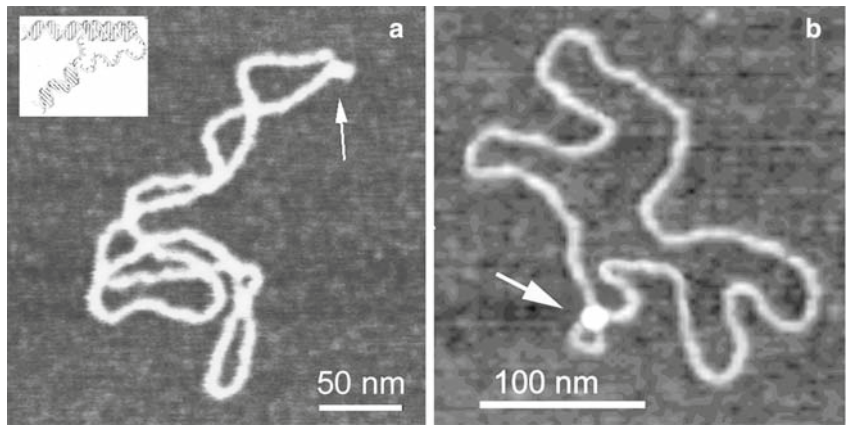


Fig. 2. AFM of images of supercoiled DNA with extruded H-DNA local structure indicated with an arrow (**a**) (*see* **ref.27** for details) and looped structure formed by the interaction of SfiI enzyme (indicated by arrow) with two recognition sites in the DNA template (**b**). Specifics for SfiI-DNA complexes are described in **ref.32.**

of various kinds for probing intermolecular interaction. Moreover, the use of relatively long flexible linkers made it possible to perform the force spectroscopy analysis at a single molecule level. **Subheading 3.3** outlined the specifics for the preparation of surfaces with covalently attached functionalized PEG linkers.

2. Materials

2.1. Mica Substrate

Any type of commercially available mica sheets (green or ruby mica) can be used. Asheville-Schoonmaker Mica Co (Newport News, VA) supplies thick and large (more than 5 × 7 cm) sheets suitable for making the substrates of different sizes.

2.2. Chemicals

1. 3-Aminopropyltriethoxy silane (Fluka, Chemika-BioChemika, Switzerland; Aldrich, USA; United Chemical Technology, USA) can be used without additional purification.
2. Sodium metal and triethanolamine were purchased from Sigma-Aldrich.
3. Xylene solvent was purchased from VWR.
4. Argon gas supplied in cylinders of reagent grade.
5. Bifunctional NHS-PEG-MAL (bf PEG, *N*-hydroxysuccinimide-polyethylene glycol)-maleimide, 3,400 MW (Laysan Bio, Inc., Arab, AL).
6. Dimethylsulfoxide (DMSO, Sigma-Aldrich Inc.)

2.3. Water

Double glass distilled or deionized water was filtered through 0.2-μm filter.

2.4. Minor Equipment

1. A vacuum cabinet or dessicator for storing the samples. A unit Gravity Convention Utility Oven (VWR) is recommended.
2. UV ozone tip cleaning unit for UV irradiation of the AFM tips before the functionalization. The ProCleaner™ system (BioForce Nanosciences, Inc., Ames, IA) or CL1000 UV Crosslinker (UVP, LLC, Upland, CA) is recommended.

3. Methods

3.1. Synthesis of APS

1-(3-Aminopropyl)silatrane (CAS [17869-27-1]) was prepared by a method briefly described previously *(24)*. (Silatrane refers to 2,8,9-trioxa-5-aza-1-silabicyclo[3.3.3]undecane. However, silatranes are generally accepted to contain an N->Si coordinate bond.)

A catalytic amount of sodium metal (5 mg) is added to 15.0 mL (16.8 g, 0.11 mol) of triethanolamine (Aldrich) in a 250 mL round-bottom flask under argon or nitrogen atmosphere and allowed to form a solution (sodium catalyst can be replaced with the equivalent amount of sodium ethoxide or sodium hydroxide in 1 mL of ethanol). A rubber balloon is then attached to the flask via a rubber septum and a needle to allow hydrogen to escape without building up pressure. Moderate heat (up to 100°C) can be applied to accelerate the process, but the mixture should be allowed to cool to room temperature before the next step. An equivalent amount of (3-aminopropyl)triethoxysilane (26.4 mL or 25.0 g, 0.11 mol) is added to the mixture; then the flask is placed into a 60°C water bath and connected to the vacuum line to absorb ethanol released in the reaction. This reaction can be simply performed on a rotary evaporator. At the end of the reaction, the mixture loses ~17 g of ethanol and is turned into a solid. This process can take more than 24 h, but mechanical or occasional manual stirring can reduce time to 1–2 h. Evaporation at the end with 150 mL of xylenes at 60°C can also accelerate the process and help crystallization. Vacuum-dried 1-(3-aminopropyl) silatrane obtained by this method can be used directly for AFM, or for better results and higher stability the product can be purified by crystallization. Minute amounts of sodium hydroxide in the product practically do not affect the performance of the reagent, or change the pH of stock solutions of APS used for AFM. Recrystallization from xylenes provides 20 g (80% yield) of colorless solid. The synthesized APS has the following

characteristics: m.p. 91–94°C (open capillary tube); lit. (US Patent 3,118,921) m.p. 87.2–87.9°C (sealed capillary tube). ^{1}H NMR (DMSO-d_6), ppm: 0.08–0.14 (2H, m, $SiCH_2$); 1.1 (2H, br. s, NH_2); 1.28–1.37 (2H, m, CH_2); 2.37 (2H, t J = 7.2 Hz, NCH_2); 2.77 (6H, t J = 5.9 Hz, NCH_2); 3.59 (6H, t J = 5.9 Hz, OCH_2) *(50)*.

3.2. Mica Functionalization with APS

1. Prepare 50 mM APS stock solution in water and store it in refrigerator. The shelf life of the stock solution is not less than 6 months.
2. Prepare working APS solution for mica modification dissolving the stock solution in 1:300 ratio in water; it can be stored at room temperature for several days.
3. Cleave mica sheets of needed sizes (typically 1 × 3 cm) to make them as thin as 0.05–0.1 mm, place them in appropriate plastic tubes, pour working APS solution to cover the mica sheet, and leave on the bench for 30 min. Depending on the size of the mica strip, the plastic disposable 3-mL cuvettes or plastic 15-mL tubes are suitable for these purposes.
4. After 30 min, discard the reagent and rinse the mica with water under the water stream or replacing the water in the tube 3–4 times.
5. Remove the mica sheets and dry them under Argon stream. The strips are ready for the sample preparation. Note, however, as prepared, the APS mica sheets can be stored under Argon for several days. This procedure allows one to obtain a weak cationic surface with rather uniform distribution of the charge. **Fig. 2a** shows uniform distribution of DNA fragments over the surface.

3.3. Covalent Attachment of Bifunctional PEG Linker to APS Surfaces

This procedure describes the surface functionalization steps of aminoterminated APS surfaces that are required for AFM force spectroscopy applications. References *(33, 34, 36, 37, 49)* are good examples for a few applications with the use of APS and other silatranes. The following procedure describes specifics for covalent attachment of bifunctional PEG (bf-PEG) terminated with maleimide and NHS groups. The latter provides covalent bond reacting with the surface-terminated amines allowing the use of maleimide termini for covalent attachment of molecules containing thiol groups *(34, 36)*. A typical force curve with clearly identified rupture event is shown in **Fig. 3**.

1. Prepare 1.67 mM solution of bf-PEG in DMSO, 10 mM of pH 3.7, TCEP hydrochloride (Hampton Research Inc, Lynchburg, VA), and 20 mM β-mercaptoethanol.

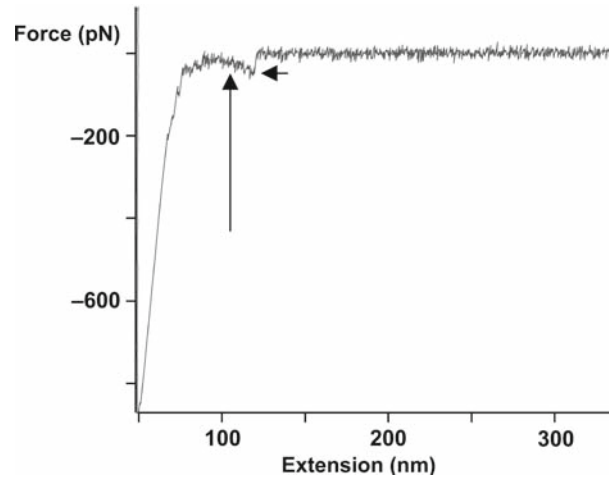


Fig. 3. Force curve illustrating the single molecule detection of interaction of two α-synuclein molecules at conditions facilitating the misfolding and aggregation of this protein *(35)*. The event for the rupture of the protein dimer spontaneously formed upon approaching tip to the surface is indicated with a short arrow. The long arrow points to the region of the extension curve corresponding to stretching of the PEG linker *(36.* The protein was immobilized on APS-functionalized mica and silicone nitride tips via PEG linker.

2. Apply 0.1 mL of the bf-PEG solution to APS-mica and allow the reaction to proceed for 3 h.
3. Rinse the mica three times with 1 mL of DMSO followed by a thorough rinse with deionized water and use for the covalent attachment of your sample. The following protocol for immobilization of α-synuclein via formation covalent bonds between maleimide moiety of immobilized PEG with cystein of the protein can be used as a template for development of the procedure for immobilizing the needed molecule.

3.3.1. Procedure for Immobilization of α-Synuclein on Mica

1. Prepare α-synuclein solution in the concentration 19 nM in water with 0.25 mM TCEP.
2. Place 10 μL of the protein solution on the PEG-activated mica sheets for 1 h.
3. Rinse the mica with DMSO for three times to remove NHS-PEG-MAL solution from the slide followed by a thorough rinse with water.
4. Rinse with HEPES buffer (pH 7.0) few times, deposit 50 μL of α-mercaptoethanol solution, and leave for 10 min.
5. Rinse with HEPES buffer (pH 7.0) and mount the mica sample on the microscope stage.

3.3.2. Procedure for Immobilization of α-Synuclein on the AFM Tip

1. Clean silicon nitride tips with 95% ethanol, immerse tips in 95% ethanol for 30 min, rinse in water and dry in argon flow, place tips under the UV lamp for 30 min, immerse tips in APS solution for 30 min.
2. *NHS-PEG-MAL treatment*: remove tips from APS solution, rinse with water, dry in argon flow, and immerse the tips in NHS-PEG-MAL DMSO solution (*see* **Subheading 3.3.1**.) for 3 h.
3. *Protein attachment*: mix 10μL of 10 mM TCEP stock solution with 37 μL of stock protein (500 nM solution) and allow to stay for ~15 min; rinse the tips with DMSO for three times to remove NHS-PEG-MAL solution from the tips, rinse the tips with water many times to remove DMSO and immerse the tips in TCEP/protein/buffer solution for 1 h, rinse tips in HEPES buffer (pH 7.0), immerse in α-mercaptoethanol solution for 10 min, rinse the tips in HEPES buffer (pH 7.0), and leave in this buffer.

3.4. Sample Preparation for Imaging in Air

1. Prepare the solution of the sample (DNA, RNA, protein-DNA complex) in appropriate buffer. DNA concentration should be between 0.1 and 1 μg/mL depending on the size of the molecules.
2. Place 5–10 μl of the solution in the middle of APS-mica substrate (usually 1 × 1 cm squares) for 2–3 min.
3. Rinse the surface thoroughly with water (2–3 mL per sample) to remove all buffer components. 5–10 mL plastic syringes are useful for rinsing. Attach an appropriate plastic tip instead of a metal needle.
4. Dry the sample by blowing with clean argon gas. The sample is ready for imaging. It is recommended 30 min drying of the sample in vacuum cabinet prior imaging especially if the humidity is high. Store the samples in vacuum cabinets or desiccators filled with argon. They remain unchanged for months.

3.5. Sample Preparation for Imaging in Liquid

1. Prepare the solution (DNA, RNA, nucleoprotein complexes) and preincubate it for 10–20 min to allow the temperature to equilibrate. Recommended concentration of DNA is 0.1–1 μg/mL depending on the size of the molecules.
2. Mount a piece of APS-mica on the microscope stage, and follow the steps under **Subheading 3.7**.

3.6. Procedure for Imaging in Air

The procedure for imaging in air is straightforward: mount the sample and do imaging. Although both contact and intermittent (tapping) modes can be used, the latter is preferable and allows one to get images of DNA and DNA–protein complexes routinely. Our experience is mostly limited to MultiMode system (Veeco, CA) running various NanoScope III, IIIa, and IV

controllers and MFP 3D microscope (Asylum Research, Santa Barbara, CA). However, the samples prepared on APS-mica were imaged on other commercially available instruments, for example, the microscopes manufactured by Molecular Imaging (Agilent, Tempe, AZ). With MultiMode system any types of probes designed for intermittent imaging modes (Tapping mode for the Veeco systems) can be used. Typical tapping frequency of 240–380 kHz and scanning rate of 2–3 Hz allow one to obtain stable images.

3.7. Procedure for Imaging in Aqueous Solutions

The capability of AFM to perform scanning in liquid is its most attractive feature for numerous biological applications allowing imaging at conditions closed to physiological ones. In addition, this mode of imaging permits one to eliminate undesirable resolution-limiting capillary effect typical for imaging in air (e.g., *(12, 16, 51)*). As a result, images of DNA filaments as thin as ~3 nm were obtained in water solutions *(18)*, and helical periodicity was observed when dried DNA samples were imaged in propanol *(52)*. In addition to APS-mica, our previously developed procedure with the use of AP-mica can be used as a substrate for imaging in liquid *(14)*; note that the first images of DNA in fully hydrated state were obtained by the use of AP-mica *(15)*. This type of imaging is recommended in cases when dynamics is studied. The following procedure is described for the use of MultiMode AFM (Veeco, Santa Barbara, CA) and can be easily adapted for other systems.

1. Install an appropriate tip in the holder designed for imaging in liquid (fluid cell). Use Si_3N_4 big triangle with thick legs or small triangle with thin legs cantilevers *(18)*.
2. Mount the APS mica sheet on the stage of the microscope. Mica pieces of 1 × 1 cm are sufficient for NanoScope design of fluid cell. Wrap the stage of the scanner for MM AFM with Parafilm to prevent potential shortage due to accidental leak of the fluid underneath of the mica sheet.
3. Attach the head of the microscope with installed fluid cell and make appropriate adjustments of the microscope.
4. Approach the sample to the tip manually, leaving ~50–100 μm gap between the tip and the surface.
5. Inject the sample solution or appropriate solvent with 1-mL plastic syringe. Take ~50 μl of the sample with 1-mL syringe. Use 200-μl plastic tips with capillary ends instead of metal needle. Cut both ends of the tip to fit to the syringe and the diameter of inlet hole of the fluid cell. The use of the second syringe attached to the outlet of fluid cell allowing for pulling out the solution is recommended for manipulating with small volume of solution.

6. Change the position of the mirror to maximize the signal on the photodetector.
7. Find a resonance peak. Typically it is a quite broad peak in the range of 7–12 kHz for the MultiMode system. Follow the recommendations given in the manual for the fluid cell operation how to find the peak.
8. Minimize the drive amplitude. The values vary from tip to tip, but amplitude as low as 20 nm or less provides good quality pictures (*see* **Notes 1–5**).
9. Allow the microscope to approach the sample and engage the tip.
10. Adjust the values for the set point voltage and drive amplitude to improve the quality of images and start the data acquisition using the continuous mode scanning.

3.8. AFM Force Spectroscopy

These specifics are made for MFP 3D system (Asylum Research, Santa Barbara, CA). Modify the protocol for your system.

3.8.1. Spring Constant Measurements

1. Install the tip into the AFM head and put a piece of freshly cleaved mica on the stage.
2. Based on thermal profile estimate the resonance frequency of the tip, type in start and end frequency. Do auto tune.
3. Approach to the surface.
4. When on the surface, click stop and change from AC mode to Contact mode.
5. Go to force menu. Click continue force and slowly move the red bar down (or arrow) until you see characteristic force curve. Click stop.
6. Find the linear slope on the force curve, put marks on the linear part of the force curve as far away from each other as possible. Press the arrow button in either direction to make sure that both marks are moving in the same direction (they should be on the same curve: approach or retraction, and when they move in the same direction, it indicates that they are).
7. Click the deflection button to see deflection value in air.
8. Withdraw and go to thermal and do thermal. Wait for several seconds to collect more curves then stop.
9. Do the fit to get the spring constant value. Use this value for the next step.

3.8.2. Force Measurements

Before you start force measurements you have to do tune and approach as it is described earlier. When you are on the surface do the following:

1. Type in the designated box the measured spring constant and measure the deflection of the tip in liquid- the same way as it

is described in the sections earlier. The deflection will be different from the value measured in air.

2. Set up the parameters for scan rate, trigger, dwell time (if any) and start collecting force curves.

4. Notes

1. *Silatrane synthesis.* In **Subheading 3.1**, sodium catalyst can be replaced with the equivalent amount of sodium ethoxide or sodium hydroxide in 1 mL of ethanol.
2. *DNA concentration.* This parameter depends on the length of molecules. If the molecules are as small as several hundred base pairs, the concentration ca. up to 1 μg/mL is recommended to avoid intermolecular crossing. To the contrary, low DNA concentration is recommended for large DNA molecules. For example, concentration of lambda DNA (~48 kb) ~0.01 μg/mL allows one to get images of individual DNA molecules *(15, 17, 52)*.
3. *DNA preparation.* Very small amount of DNA is needed. Typically 1–10 ng of DNA is sufficient for the AFM preparation of plasmid DNA (~3 kb long). Because one band of DNA in agarose gel usually contains 100 ng of DNA, DNA extracted from the gel can be sufficient for preparation of a set of samples. The following procedures can be used for purification of DNA extracted from the gel.
   - *Electrophoretical deposition of DNA bands onto DEAE paper.* The strips of the paper are placed into the slot 3–5 mm below the band and the DNA is electrophoresed onto the paper for 5–10 min (the time can be determined by a direct examination the gel under a UV transilluminator). The DNA is extracted from the paper by elution in 2 M NaCl solution for 30 min at 37°C followed by ethanol precipitation.
   - *Extraction from the gel.* The procedure is based on the use of the extraction kit UltraClean15 and the protocol provided (MoBio Laboratories, Inc., Solana Beach, CA). Generally, the purification consists of melting of the slice of agarose gel, immobilizing DNA on the sorbent, washing off all contaminants, and eluting DNA from the matrix with a low-salt buffer. At least one step of the ethanol precipitation is needed to remove UV-absorbing low molecular weight materials. A similar procedure can be applied to purification of the sample extracted from polyacrylamide gel *(53)*.

4. *Imaging conditions.* It was recommended to operate the instrument at the lowest possible drive amplitude. This recommendation is based on the following considerations. An oscillating tip provides rather large energy to the sample. A total energy deposited into the sample by oscillating tip can be as high as 10^{16}–10^{17} J at amplitude oscillations of 30 nm *(18)*. However, this value is almost three orders of magnitude lower if the microscope is operated at amplitudes as low as ~3 nm. Such imaging conditions allow one to minimize effect of the tip on the sample, to prevent damaging the tip, and get images with high contrast. In addition, such conditions considerably simplify the study with AFM dynamic processes such as segmental DNA mobility, local structural transitions in supercoiled DNA *(18, 20)* and conformational transitions of DNA Holliday junctions including branch migration *(26, 38, 39, 54)*. This capability is illustrated in **Fig. 4**, in which a set of time lapse images of a segment of supercoiled DNA containing a cruciform is shown.
5. *Alternative procedures for AFM sample preparation.* Among other techniques applied to AFM studies of DNA, the method based on using divalent cations *(7, 12, 55)* allowed to get images of a number of nucleoprotein complexes. In this approach, the mica surface is treated with multivalent ions (e.g., Mg^{2+}) to increase its affinity to DNA, which is held in place strongly enough to permit reliable imaging by AFM. An alternative is to deposit the sample in the buffer containing a multivalent ion. This cation-assisted procedure of sample preparation was used for studies of the process of DNA degradation with nuclease *(56)* and interaction of DNA with a site-bound enzyme. EcoP15I *(57)*. The mechanism of this technique remains unclear; detailed protocol depends on the

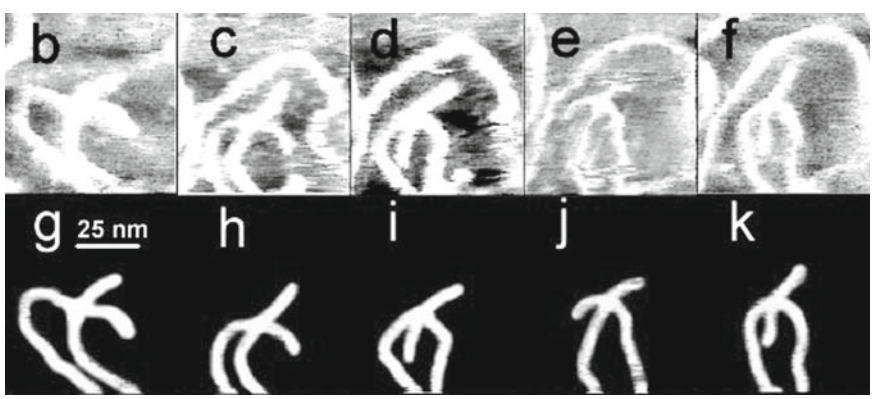


Fig. 4. Time lapse images illustrating dynamics of cruciforms formed within supercoiled DNA. Sections of the plasmid containing the cruciform are shown. Frames (**b–f**) correspond to times 11 min (**b**), 44 min (**c**), 60 min (**d**), 134 min (**e**), and 155 min (**f**). Traces of the same molecule on black background are shown in frames (**g–k**) (*see* **ref.** *39* for details).

system studied, the type of the cation used, and efficiency of DNA deposition is buffer sensitive *(55, 58)*. In some cases, a special class of tips (electron-beam-deposited tips) is required for reliable imaging *(59)*.

Acknowledgments

This work was supported by the grants GM 54991 (NIH) and PHY-06155909 (NSF) to YLL.

References

1. Binnig, G., Rohrer, H., Gerber, C., and Weibel, E. (1982). Surface Studies by Scanning Tunneling Microscopy, *Phys Rev Lett* **49**, 57–60.
2. Binnig, G., Quate, C. F., and Gerber, C. (1986). Atomic force microscope, *Phys Rev Lett* **56**, 930–933.
3. Clemmer, C. R., and Beebe, T. P., Jr. (1991). Graphite: a mimic for DNA and other biomolecules in scanning tunneling microscope studies, *Science* **251**, 640–642.
4. Heckl, W. M., and Binnig, G. (1992). Domain walls on graphite mimic DNA, *Ultramicroscopy* 42–44(Pt B), 1073–1078.
5. Bustamante, C., Vesenka, J., Tang, C. L., Rees, W., Guthold, M., and Keller, R. (1992). Circular DNA molecules imaged in air by scanning force microscopy, *Biochemistry* **31**, 22–26.
6. Yang, J., Takeyasu, K., and Shao, Z. (1992). Atomic force microscopy of DNA molecules, *FEBS Lett* **301**, 173–176.
7. Vesenka, J., Guthold, M., Tang, C. L., Keller, D., Delaine, E., and Bustamante, C. (1992). Substrate preparation for reliable imaging of DNA molecules with the scanning force microscope, *Ultramicroscopy* 42–44(Pt B), 1243–1249.
8. Lyubchenko, Y. L., Gall, A. A., Shlyakhtenko, L. S., Harrington, R. E., Jacobs, B. L., Oden, P. I., and Lindsay, S. M. (1992). Atomic force microscopy imaging of double stranded DNA and RNA, *J Biomol Struct Dyn* **10**, 589–606.
9. Allen, M. J., Dong, X. F., O'Neill, T. E., Yau, P., Kowalczykowski, S. C., Gatewood, J., Balhorn, R., and Bradbury, E. M. (1993). Atomic force microscope measurements of nucleosome cores assembled along defined DNA sequences, *Biochemistry* **32**, 8390–8396.
10. Hegner, M., Wagner, P., and Semenza, G. (1993). Immobilizing DNA on gold via thiol modification for atomic force microscopy imaging in buffer solutions, *FEBS Lett* **336**, 452–456.
11. Mou, J., Czajkowsky, D. M., Zhang, Y., and Shao, Z. (1995). High-resolution atomic-force microscopy of DNA: the pitch of the double helix, *FEBS Lett* **371**, 279–282.
12. Bustamante, C., Erie, D., and Keller, D. (1994). Biochemical and structural applications of scanning force microscopy, *Curr Opin Struct Biol* **4**, 750–760.
13. Bustamante, C., Rivetti, C., and Keller, D. J. (1997). Scanning force microscopy under aqueous solutions, *Curr Opin Struct Biol* **7**, 709–716.
14. Lyubchenko, Y. L., Gall, A. A., and Shlyakhtenko, L. S. (2001). Atomic force microscopy of DNA and protein-DNA complexes using functionalized mica substrates, *Methods Mol Biol* **148**, 569–578.
15. Lyubchenko, Y., Shlyakhtenko, L., Harrington, R., Oden, P., and Lindsay, S. (1993) Atomic force microscopy of long DNA: imaging in air and under water, *Proc Natl Acad Sci U S A* 90, 2137–2140.
16. Lyubchenko, Y. L., Jacobs, B. L., Lindsay, S. M., and Stasiak, A. (1995). Atomic force microscopy of nucleoprotein complexes, *Scanning Microsc* 9, 705–724; discussion 724–707.
17. Lyubchenko, Y. L., Blankenship, R. E., Gall, A. A., Lindsay, S. M., Thiemann, O., Simpson, L., and Shlyakhtenko, L. S. (1996). Atomic force microscopy of DNA, nucleoproteins and cellular complexes: the use of functionalized substrates, *Scanning Microsc Suppl* 10, 97–107; discussion 107–109.

18. Lyubchenko, Y. L., and Shlyakhtenko, L. S. (1997). Visualization of supercoiled DNA with atomic force microscopy in situ, *Proc Natl Acad Sci U S A* **94**, 496–501.
19. Shlyakhtenko, L. S., Potaman, V. N., Sinden, R. R., Gall, A. A., and Lyubchenko, Y. L. (2000). Structure and dynamics of three-way DNA junctions: atomic force microscopy studies, *Nucleic Acids Res* **28**, 3472–3477.
20. Lyubchenko, Y. L., Shlyakhtenko, L. S., Potaman, V. P. and Sinden, R. R. (2002). Global and local DNA structure and dynamics. Single molecule studies with AFM, *Microsc Microanal* **8**, 170–171.
21. Yodh, J. G., Woodbury, N., Shlyakhtenko, L. S., Lyubchenko, Y. L., and Lohr, D. (2002). Mapping nucleosome locations on the 208–12 by AFM provides clear evidence for cooperativity in array occupation, *Biochemistry* **41**, 3565–3574.
22. Kato, M., Hokabe, S., Itakura, S., Minoshima, S., Lyubchenko, Y. L., Gurkov, T. D., Okawara, H., Nagayama, K., and Shimizu, N. (2003). Interarm interaction of DNA cruciform forming at a short inverted repeat sequence, *Biophys J* **85**, 402–408.
23. Potaman, V. N., Bissler, J. J., Hashem, V. I., Oussatcheva, E. A., Lu, L., Shlyakhtenko, L. S., Lyubchenko, Y. L., Matsuura, T., Ashizawa, T., Leffak, M., Benham, C. J., and Sinden, R. R. (2003). Unpaired structures in SCA10 (ATTCT)n.(AGAAT)n repeats, *J Mol Biol* **326**, 1095–1111.
24. Shlyakhtenko, L. S., Gall, A. A., Filonov, A., Cerovac, Z., Lushnikov, A., and Lyubchenko, Y. L. (2003). Silatrane-based surface chemistry for immobilization of DNA, protein-DNA complexes and other biological materials, *Ultramicroscopy* **97**, 279–287.
25. Lushnikov, A. Y., Brown, B. A.,2nd, Oussatcheva, E. A., Potaman, V. N., Sinden, R. R., and Lyubchenko, Y. L. (2004). Interaction of the Zalpha domain of human ADAR1 with a negatively supercoiled plasmid visualized by atomic force microscopy, *Nucleic Acids Res* **32**, 4704–4712.
26. Lyubchenko, Y. L. (2004). DNA structure and dynamics: an atomic force microscopy study, *Cell Biochem Biophys* **41**, 75–98.
27. Tiner, W. J., Sr., Potaman, V. N., Sinden, R. R., and Lyubchenko, Y. L. (2001). The structure of intramolecular triplex DNA: atomic force microscopy study, *J Mol Biol* **314**, 353–357.
28. Kato, M., McAllister, C. J., Hokabe, S., Shimizu, N., and Lyubchenko, Y. L. (2002). Structural heterogeneity of pyrimidine/purine-biased DNA sequence analyzed by atomic force microscopy, *Eur J Biochem* **269**, 3632–3636.
29. Dahlgren, P. R., Karymov, M. A., Bankston, J., Holden, T., Thumfort, P., Ingram, V. M., and Lyubchenko, Y. L. (2005). Atomic force microscopy analysis of the Huntington protein nanofibril formation, *Dis Mon* **51**, 374–385.
30. Lonskaya, I., Potaman, V. N., Shlyakhtenko, L. S., Oussatcheva, E. A., Lyubchenko, Y. L., and Soldatenkov, V. A. (2005). Regulation of poly(ADP-ribose) polymerase-1 by DNA structure-specific binding, *J Biol Chem* **280**, 17076–17083.
31. Lushnikov, A. Y., Potaman, V. N., and Lyubchenko, Y. L. (2006). Site-specific labeling of supercoiled DNA, *Nucleic Acids Res* **34**, 111–117.
32. Lushnikov, A. Y., Potaman, V. N., Oussatcheva, E. A., Sinden, R. R., and Lyubchenko, Y. L. (2006). DNA strand arrangement within the SfiI-DNA complex: Atomic force microscopy analysis, *Biochemistry* **45**, 152–158.
33. McAllister, C., Karymov, M. A., Kawano, Y., Lushnikov, A. Y., Mikheikin, A., Uversky, V. N., and Lyubchenko, Y. L. (2005). Protein interactions and misfolding analyzed by AFM force spectroscopy, *J Mol Biol* **354**, 1028–1042.
34. Kransnoslobodtsev, A. V., Shlyakhtenko, L. S., Ukraintsev, E., Zaikova, T. O., Keana, J. F., and Lyubchenko, Y. L. (2005). Nanomedicine and protein misfolding diseases, *Nanomedicine* **1**, 300–305.
35. Lyubchenko, Y. L., Sherman, S., Shlyakhtenko, L. S., and Uversky, V. N. (2006). Nanoimaging for protein misfolding and related diseases, *J Cell Biochem* **99**, 53–70 [Figure featured on journal cover].
36. Krasnoslobodtsev, A. V., Shlyakhtenko, L. S., and Lyubchenko, Y. L. (2007). Probing Interactions within the synaptic DNA-SfiI complex by AFM force spectroscopy, *J Mol Biol* **365**, 1407–1418.
37. Shlyakhtenko, L. S., Yuan, B., Emadi, S., Lyubchenko, Y. L., and Sierks, M. R. (2007). Single-molecule selection and recovery of structure-specific antibodies using atomic force microscopy, *Nanomedicine* **3**, 192–197.
38. Lushnikov, A. Y., Bogdanov, A., and Lyubchenko, Y. L. (2003). DNA recombination: Holliday junctions dynamics and branch migration, *J Biol Chem* **278**, 43130–43134 [Figure featured on journal cover].
39. Mikheikin, A. L., Lushnikov, A. Y., and Lyubchenko, Y. L. (2006). Effect of DNA supercoiling on the geometry of Holliday junctions, *Biochemistry* **45**, 12998–13006.

40. Liu, R., McAllister, C., Lyubchenko, Y., and Sierks, M. R. (2004). Residues 17-20 and 30-35 of beta-amyloid play critical roles in aggregation, *J Neurosci Res* **75**, 162–171.
41. Liu, R., McAllister, C., Lyubchenko, Y., and Sierks, M. R. (2004). Proteolytic antibody light chains alter beta-amyloid aggregation and prevent cytotoxicity, *Biochemistry* **43**, 9999–10007.
42. Liu, R., Yuan, B., Emadi, S., Zameer, A., Schulz, P., McAllister, C., Lyubchenko, Y., Goud, G., and Sierks, M. R. (2004). Single chain variable fragments against beta-amyloid (Abeta) can inhibit Abeta aggregation and prevent abeta-induced neurotoxicity, *Biochemistry* **43**, 6959–6967.
43. Emadi, S., Liu, R., Yuan, B., Schulz, P., McAllister, C., Lyubchenko, Y., Messer, A., and Sierks, M. R. (2004). Inhibiting aggregation of alpha-synuclein with human single chain antibody fragments, *Biochemistry* **43**, 2871–2878.
44. Yamin, G., Munishkina, L. A., Karymov, M. A., Lyubchenko, Y. L., Uversky, V. N., and Fink, A. L. (2005). Forcing nonamyloidogenic beta-synuclein to fibrillate, *Biochemistry* **44**, 9096–9107.
45. Watson, D., Castano, E., Kokjohn, T. A., Kuo, Y. M., Lyubchenko, Y., Pinsky, D., Connolly, E. S., Jr., Esh, C., Luehrs, D. C., Stine, W. B., Rowse, L. M., Emmerling, M. R., and Roher, A. E. (2005). Physicochemical characteristics of soluble oligomeric Abeta and their pathologic role in Alzheimer's disease, *Neurol Res* **27**, 869–881.
46. Uversky, V. N., Yamin, G., Munishkina, L. A., Karymov, M. A., Millett, I. S., Doniach, S., Lyubchenko, Y. L., and Fink, A. L. (2005). Effects of nitration on the structure and aggregation of alpha-synuclein, *Brain Res Mol Brain Res* **134**, 84–102.
47. Uversky, V. N., Kabanov, A. V., and Lyubchenko, Y. L. (2006). Nanotools for megaproblems: probing protein misfolding diseases using nanomedicine modus operandi, *J Proteome Res* **5**, 2505–2522.
48. Dahlgren, P. R., Bankston, J., Holden, T., Karymov, M. A., Thumfort, P., Ingram, V. M., and Lyubchenko, Y. L. (2005). Atomic force microscopy analysis of the Huntington protein nanofibril formation, *Nanomedicine* **1**, 52–57.
49. Riener, C. K., Stroh, C. M., Ebner, A., Gall, A. A., Klampfl, C., Romanin, C., Lyubchenko, Y. L., Hinterdorfer, P., and Gruber, H. J. (2003). A simple test system for single molecule recognition force microscopy, *Anal Chim Acta* **479**, 59–75.
50. *US Patent 3,118,921.*
51. Hansma, H. G., and Hoh, J. H. (1994). Biomolecular imaging with the atomic force microscope, *Annu Rev Biophys Biomol Struct* **23**, 115–139.
52. Lyubchenko, Y. L., Oden, P. I., Lampner, D., Lindsay, S. M., and Dunker, K. A. (1993). Atomic force microscopy of DNA and bacteriophage in air, water and propanol: the role of adhesion forces, *Nucleic Acids Res* **21**, 1117–1123.
53. Potaman, V. N., Lushnikov, A. Y., Sinden, R. R., and Lyubchenko, Y. L. (2002). Site-specific labeling of supercoiled DNA at the A + T rich sequences, *Biochemistry* **41**, 13198–13206.
54. Shlyakhtenko, L. S., Potaman, V. N., Sinden, R. R., and Lyubchenko, Y. L. (1998). Structure and dynamics of supercoil-stabilized DNA cruciforms, *J Mol Biol* **280**, 61–72.
55. Bezanilla, M., Manne, S., Laney, D. E., Lyubchenko, Y. L., and Hansma, H. G. (1995). Adsorption of DNA to mica, silylated mica, and minerals: characterization by atomic force microscopy, *Langmuir* **11**, 655–659.
56. Bezanilla, M., Drake, B., Nudler, E., Kashlev, M., Hansma, P. K., and Hansma, H. G. (1994). Motion and enzymatic degradation of DNA in the atomic force microscope, *Biophys J* **67**, 2454–2459.
57. Crampton, N., Yokokawa, M., Dryden, D. T., Edwardson, J. M., Rao, D. N., Takeyasu, K., Yoshimura, S. H., and Henderson, R. M. (2007). Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping, *Proc Natl Acad Sci U S A* **104**, 12755–12760.
58. Hansma, H. G., and Laney, D. E. (1996). DNA binding to mica correlates with cationic radius: assay by atomic force microscopy, *Biophys J* **70**, 1933–1939.
59. Kasas, S., Thomson, N. H., Smith, B. L., Hansma, H. G., Zhu, X., Guthold, M., Bustamante, C., Kool, E. T., Kashlev, M., and Hansma, P. K. (1997). *Escherichia coli* RNA polymerase activity observed using atomic force microscopy, *Biochemistry* **36**, 461–468.

Chapter 22

Two-Dimensional Crystallisation of Soluble Protein Complexes

Patrick Schultz, Corinne Crucifix, and Luc Lebeau

Summary

This method aims at providing structural information on protein or nucleoprotein complexes by high-resolution electron microscopy. The objective is to promote the self-assembly of the macromolecules into two-dimensional crystals in order to use electron crystallography methods. When combined with observations in the frozen hydrated states and dedicated image processing software these methods can provide detailed 3-D models of the complex. The 2-D crystals of soluble nucleoprotein complexes are formed on lipid monolayers spread at the air–water interface. The macromolecule of interest is targeted to the monolayer by either electrostatic or ligand-induced interactions with the hydrophilic head group of the lipid. Upon interaction with the lipids, the nucleoprotein complex is concentrated at the vicinity of the lipid layer whose in-plane mobility facilitates the contacts between macromolecules and the formation of ordered arrays.

Key words: Protein complexes, Nucleoprotein complexes, High-resolution electron microscopy, 2-D crystals, Self-assembly.

1. Introduction

Structural data of biological macromolecules provide invaluable insights into the interactions of proteins with nucleic acids. Atomic resolution obtained by X-ray diffraction or NMR studies ultimately describes the fold of the polypeptide chains and the chemical contacts between proteins and DNA. However, biologically active multi-protein complexes are often difficult to decipher to atomic details because of their heterogeneity, low cellular abundance, or difficulty to crystallize in three dimensions

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_22

(3-D). Recent progresses in electron microscopy, including better instruments (field emission guns, cold stages), improved specimen preservation in the frozen hydrated state, and dedicated image processing software, provide the possibility to calculate detailed 3-D molecular envelopes that are complementary to X-ray crystallography. Depending on the degree of organisation of the molecule to be studied, two general methods emerged to determine the structure of biological complexes from electron microscopy images. In the first family of methods the specimen is organised into a regular array in which the position of each molecule can be determined from the position of its neighbours. These approaches which include two-dimensional (2-D) crystals, helical arrays or icosahedral symmetry, yield the highest resolution, and in the case of membrane proteins organised 2-D arrays atomic models could be derived from electron microscopy data *(1, 2)*. The second category of methods aims at recording individual views of the macromolecule of interest, and the massive use of image analysis programmes will identify the angular relationships between images. It is out of the scope of this volume to describe the electron crystallography, cryo-electron microscopy, and image analysis methods required to determine a 3-D model of a protein complex, and these aspects are described in detail elsewhere *(3–5)*. This chapter will focus on the formation of 2-D crystals of soluble proteins for high-resolution structure determination by electron microscopy and on the decoration of 2-D streptavidin crystals by biotinylated nucleoprotein complexes. It provides the necessary information to set up, screen, and evaluate crystallisation experiments.

The remarkable achievements in the field of membrane proteins 2-D crystals prompted the pioneering work of Kornberg and collaborators which aimed at transposing the crystallisation mechanisms occurring in a lipid bilayer to soluble proteins *(6)*. The aim is to target the protein of interest to a lipid surface self-assembled as a monomolecular film at the air–water interface. The buffer-exposed hydrophilic part of the lipid molecule can be naturally charged *(7, 8)* or chemically modified *(9, 10)* to interact with the protein. Consequently the protein is concentrated on the lipid plane and adopts a few orientations relatively to the lipid plane. If the lipid molecules are free to diffuse in the monolayer plane the system can evolve towards the 2-D crystallisation of the macromolecule. Finally the lipid–protein film is transferred onto a solid support and processed for electron microscopy observations. High-resolution structural data were obtained using these methods as demonstrated by the analysis of streptavidin 2-D crystals which revealed structural information down to 3 Å in projection *(11)*.

Two categories of lipid derivatives can be used according to the specificity of the interaction to be established with the protein complex *(12)*. In one situation a charged surface will be created

by using lipids with positive or negative charges that will interact with the surface potential of the protein *(8)*. The characteristic behaviour of the protein in ion-exchange chromatography gives a hint on which type of charged lipids can be used. Valuable lipids for introducing charges into a monolayer are phosphatidyl serine which carries a negative charge or stearyl amine and alkyl trimethyl ammonium which are positively charged. In the other situation the lipid film is derivatised by a ligand which specifically interacts by the protein of interest. The ligand will be chemically grafted to the lipid moiety through a linker whose length modulates the accessibility of the protein. The grafted molecule can be a natural ligand of the protein such as dinitrophenol for specific anti-DNP antibodies *(6)*, novobiocin for gyrase B subunit *(9)* (**Fig. 1**), or

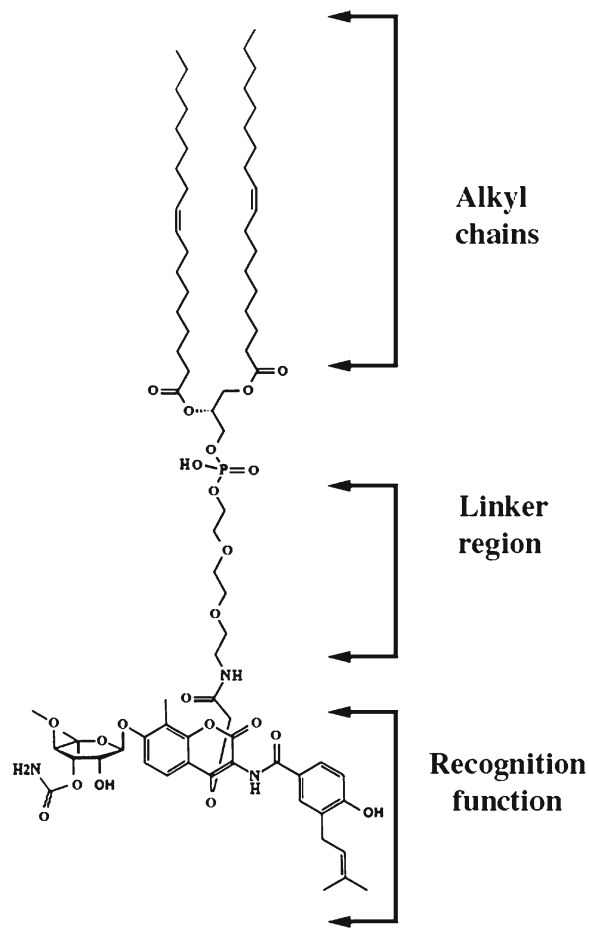


Fig. 1. Schematic representation of a lipid molecule used for 2-D crystallisation of the B subunit of DNA gyrase. The hydrophobic alkyl chain with a cis double bond confers fluidity and stability to the spread lipid layer. The linker region provides accessibility of the ligand for the protein of interest. The recognition function, here a novobiocin molecule, determines the interaction properties between the lipid and the protein.

biotin for streptavidin *(13)*. Lipid molecules were also designed to interact with specific tags (such as poly-histidines) introduced genetically into the sequence of the protein of interest *(14–17)*. In this case, the polar headgroup of the lipid molecule carries a nitrilotriacetate moiety which chelates nickel ions and interacts with the histidine tag.

Stability and fluidity of the spread lipid layer are mainly provided by the hydrophobic part of the lipid (**Fig. 1**). The lipid layer has to be in a fluid phase at the incubation temperature since crystallisation of the lipid chains was shown to prevent protein organisation probably by lowering the 2-D mobility of lipids *(13, 17)*. In most cases a cis double bond in the alkyl chains provides sufficient fluidity. In addition the lateral cohesion of the lipid molecules has to be strong enough to allow the spreading of a stable monolayer and prevent the solubilisation of the lipid as micelles or liposomes. A double alkyl chain containing 18 carbon atoms (dioleyl) generally fulfils these requirements.

When protein–lipid interactions occur the proteins are rapidly concentrated close to the lipid layer and are likely to be partly oriented. In the case of a specific interaction with a functionalised lipid, the macromolecules are tethered to the lipid film by a unique site, and the extent of orientation is determined by the length and flexibility of the linker region *(9)* (**Fig. 1**). Increased concentration, possible preferential orientation, and in-plane mobility facilitate contacts between macromolecules which results in their increased organisation when an ordered network of interactions is established. However, a dead end can be reached when the macromolecules interact too strongly or too rapidly one with each other. As a consequence 2-D aggregates will form which may be evidenced by closely packed, non-crystalline macromolecular assemblies.

The method requires limited amounts of protein since a single experiment requires about 300 ng of protein. However a quantity of 1 mg is more realistic since any new project implies systematic trials starting without a priori knowledge of the factors affecting the lipid–protein interactions. When a specific interaction of the protein with the lipid is involved, the purity grade of the biological sample appears to be of less crucial importance than when charged lipid are used *(17, 18)*.

The 2-D crystallisation of transient nucleoprotein complexes may however be hampered by the coexistence of multiple subcomplexes leading to a heterogeneous particle population. To tackle this problem, it is possible to immobilise biotin-labelled double-stranded DNA molecules or nucleoprotein complexes onto a 2-D streptavidin crystal formed in contact of a biotinylated lipid layer *(19)* (**Fig. 2)**). The biotin-binding sites of the streptavidin tetramer that are not engaged in an interaction with the biotinylated lipids can be used to recruit biotinylated DNA

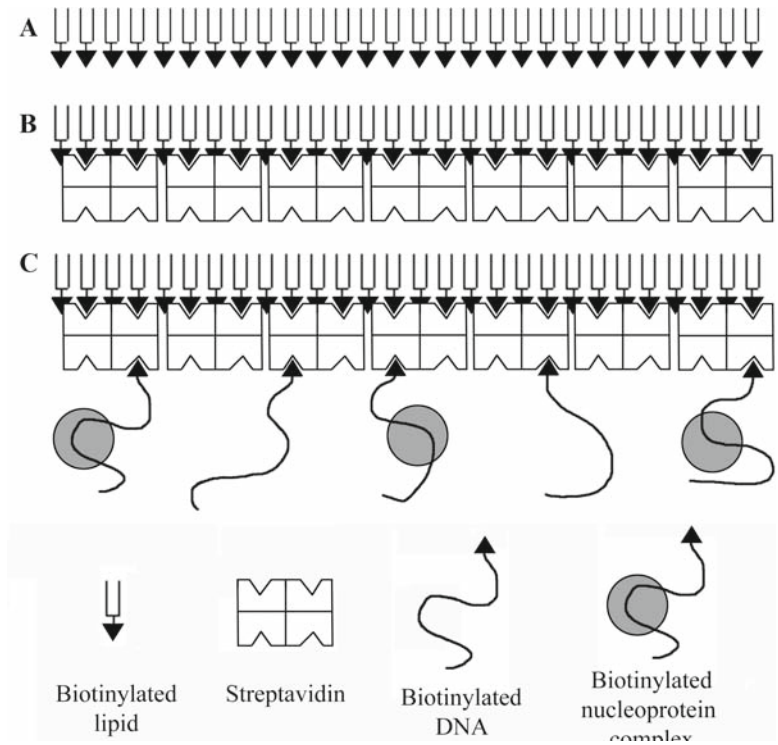


Fig. 2. Schematic description of the proposed method to immobilise DNA. (**A**) A lipid layer, functionalised with biotin, is spread at the air–water interface. (**B**) Upon addition of streptavidin, the tetrameric molecule interacts with the biotin moiety grafted onto the lipid and forms 2-D crystals. (**C**) Two biotin-binding sites per streptavidin molecule are still accessible and can interact with biotinylated DNA molecules or nucleoprotein complexes.

molecules onto the 2-D crystals. In this case, the objective is not to crystallise the nucleoprotein complexes but to tether them to the 2-D crystal, and therefore the image analysis strategies will be similar to those used for single particles *(4, 5, 19)*.

2. Materials

1. A lipid suitable to interact with the protein of interest.
2. Teflon® or Nylon® blocks in which cylindrical wells 4 mm in diameter and 1 mm deep are drilled such that each well can contain about 15 μL of aqueous solution (**Fig. 3**).
3. Standard Cu or Cu/Rh 300 mesh electron microscopy grids.
4. Mica sheets, 2.5 by 9 cm in size.

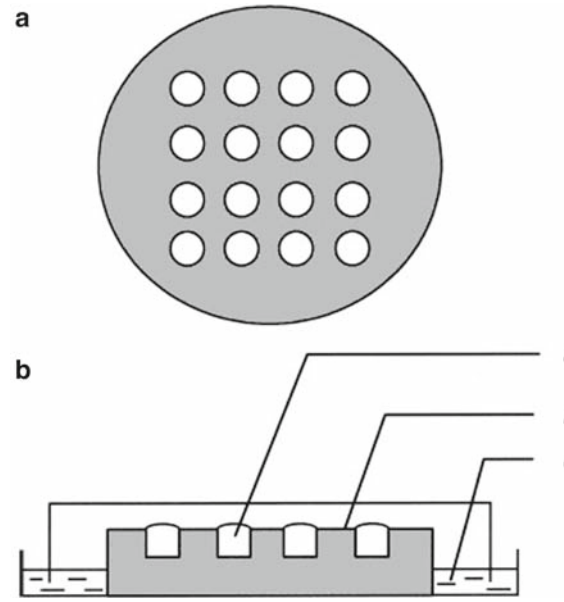


Fig. 3. Design of a Teflon block for 2-D crystallisation experiments. (**A**) A Teflon cylinder 4-cm in diameter is truncated into 1-cm thick slices and 16 wells, 4 mm in diameter and 1 mm deep, are drilled into the block such that each well can contain about 15 μL of solution. (**B**) During the crystallisation experiments performed in the wells (a) the Teflon Block (b) is placed into a humid chamber consisting of a reverted Petri dish containing some buffer (c).

5. A carbon evaporator.
6. A 2% uranyl acetate solution.
7. A control electron microscope.
8. An optical diffraction bench.

3. Methods

3.1. Preparation of Electron Microscopy Supports

The lipid–protein assemblies will have to be transferred onto a standard electron microscopy grid. The grids need to be coated by a thin hydrophobic carbon foil to support the assemblies and to allow the adsorption of the hydrophobic lipid alkyl chains.

1. The mica sheets are freshly cleaved to create a clean and flat surface.
2. The mica sheets are placed into a carbon evaporator and a 10- to 50 nm thick carbon film is evaporated under vacuum onto the cleaved face of the sheets.

3. Electron microscopy grids are placed on a supporting filter paper below the surface of a water bath. The size of the filter paper matches that of the mica sheet and will hold a total of 75–100 grids.
4. The carbon foil is floated at the cleaned air–water interface by dipping the mica sheet in the water bath with an angle of about 30°.
5. Finally, the carbon foil is gently lowered onto the grids by removing the water with a vacuum pump.

It was observed with some systems, such as streptavidin, that the contact of the lipid layer with the carbon foil interferes with the quality of protein arrays *(20)*. Holey carbon grids can then be used to transfer the crystals without interactions with a carbon surface, the lipid–protein layer being spread over the holes. Here follows a protocol to prepare holey carbon foils *(21)*.

6. The surface of an optical microscope glass slide is extensively cleaned by boiling in an aqueous detergent solution and extensive rinsing with demineralised water (H_2Od).
7. The slide is immersed in a 0.1% Triton X400 solution for 30 min, briefly rinsed with H_2Od to remove the excess of detergent and left to dry. This will result in a clean hydrophobic surface.
8. The slide is placed on a pre-cooled aluminium block to allow minute water droplets to form on the surface by condensation. The size of the droplets depends on the hygrometry of the room and the condensation time.
9. One millilitre of a cellulose acetate or cellulose butyrate solution (0.4%, w/v in ethyl acetate) is poured with a pipette over the surface; excess solution is removed from one end of the slide by touching a filter paper and left to dry. Upon drying, the cellulose forms a thin foil around the water droplets thus forming holes.
10. At this point, an optical microscope can be used to check holes size and distribution.
11. The slide is then immersed 30 min in a 0.5% (w/v) sodium dodecyl sulfosuccinate solution to peel off the cellulose foil.
12. Electron microscopy grids are deposited on a supporting filter paper just below the surface of a water bath.
13. The holey cellulose foil is floated on a clean air–water interface and deposited onto the grids.
14. A 50-nm thick carbon film is evaporated onto the cellulose-coated grids.
15. The cellulose foil is dissolved by placing the grids on an ethyl acetate-soaked filter paper.

3.2. Crystallisation Experiments

1. The lipids are best stored as a dry powder under an argon atmosphere at –80°C. A mother solution at a concentration of 10 mg/mL is produced by solubilising the lipids in an organic solvent such as a 1:1 mixture of chloroform:hexane. This solution can be stored under argon up to 1 year at –20°C. The working lipid solution is at a concentration of 0.5–1 mg/L in an organic solvent. All solutions are stored in 2-mL glass vials with Teflon caps to prevent solvent evaporation.
2. The Teflon wells have to be cleaned prior use for crystallisation experiments to remove residues of proteins or lipids. The Teflon support can be dipped into a sulfochromic acid solution for 1 h, rinsed 10 times with H_2Od, dipped for 1 h into H_2Od and rinsed again 3 times with H_2Od. Alternatively the support can be rinsed 10 times with methanol to eliminate proteins, 10 times with a chloroform:methanol 2:1 or hexane:methanol 9:1 solution to remove lipids and rinsed 10 times with hexane to remove fatty acids. Finally the support is brought into contact with a filter paper to remove the excess of H_2Od or organic solvent without wiping to avoid electrostatic charging, and allowed to dry in a dust-free chamber.
3. Incubations are performed in a humid chamber to prevent buffer evaporation (**Fig. 3b**). The Teflon block is placed in a reverted Petri dish containing some buffer with an opening in the top to let air in and out during removal of the lid.
4. In each well, 10 μL of buffer is added (**Fig. 4b**).
5. To spread the lipid at the air–buffer interface, 1 μL of the lipid solution at a concentration of 0.5–1 mg/mL is placed on the top of the drop of buffer with a micropipette. At this moment it can be observed that the surface tension of the drop is released (**Fig. 4c**).
6. The organic solvent is allowed to evaporate for 5 min.
7. The protein solution (5 μL) is injected into the aqueous phase (**Fig. 4d**). The final protein concentration is generally set between 20 and 200 μg/mL.
8. The incubation chamber is closed and if oxidation is a problem air is replaced by argon.
9. The incubation time will vary from one system to another but is generally in the range of 1–36 h. Most experiments can be performed at room temperature, but longer incubation times at 4°C may improve crystal quality in some cases.

3.3. Feed Back Loops

1. If the protein is not concentrated on the lipid film it is advisable to act on the lipid region involved in protein recognition, on its environment or on the buffer composition. In the case of a specific lipid, the linker may be too short to allow

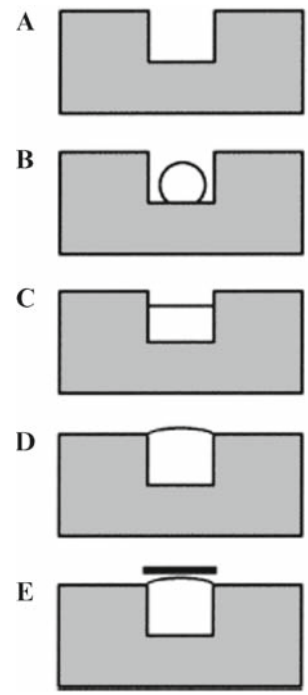


Fig. 4. Set up of a crystallisation experiment. In each well (**A**), a 10-μL drop of buffer is placed (**B**). Since the Teflon well is hydrophobic, the drop does not wet the surface. Upon addition of 1 μL of the lipid solution at a concentration of 0.5–1 mg/mL the surface tension of the drop is released (**C**). After evaporation of the organic solvent 5 μL of the protein solution is injected into the well (**D**) and is allowed to interact with the lipid layer. The resulting lipid–protein assembly is transferred to a carbon-coated electron microscopy grid placed on the top of the drop (**E**).

the ligand to be recognised by the protein. Alternatively the surface potential created by the lipid layer may have a repulsive effect on the protein, and it may be of importance to modify the environment by addition of a dilution lipid. Finally the ionic strength of the buffer may be too high and screen electrostatic interactions between the protein and the charged lipid (*see* **Note 1**).

2. When the protein tends to form close-packed arrays which do not evolve towards organised protein patches it is advisable to reduce the kinetics of protein concentration either by increasing the viscosity of the medium by adding glycerol (up to 40%) or by reducing the temperature or the protein concentration. The specific or charged lipid can also be diluted with a neutral lipid to reduce the local concentration of ligand or the charge density of the surface (*see* **Note 2**).
3. The experiment has further to be evaluated in terms of macromolecular organisation. Higher degrees of order are

recognised visually during the electron microscopy inspection of the specimen by the appearance of patches of ordered arrays (**Fig. 5a**, **b**). Once larger crystalline areas are obtained, electron micrographs are recorded and the extent of order is evaluated by optical diffraction. A large number of parameters such as the pH, the ionic strength, the buffer composition, the presence of divalent cations, the protein concentration, the presence of glycerol, the incubation temperature, or the incubation time can be modified to improve crystal order (*see* **Note 3**). At this stage the homogeneity of the specimen suspension may be crucial.

4. In the case of streptavidin, the method for preparing the sample for electron microscopy and in particular the transfer mode proved to be essential to recover a large number of highly ordered crystals *(20)*. More generally, the manipulation of one-molecule-thick assemblies during transfer to the electron microscope is likely to introduce at least part of the defaults affecting 2-D crystals such as rotational and translational distortions, fragmentation, and other disorders (*see* **Notes 4–6**).
5. An improvement of the interpretable resolution once the specimen diffracts to about 1/2 nm^{-1} will probably need a change in the method of specimen preservation from negatively stained to frozen hydrated samples *(22)*.
6. To calculate a 3-D model, astigmatism-free and well-focused images of the oriented or crystallised macromolecules have to be recorded under minimal exposure conditions and generally at low temperature. These images are analysed to calculate a noise-free image representing a projection of the macromolecular densities (**Fig. 5c**). Since the particles are adsorbed on a planar surface, tilted views are then recorded to recover the information normal to the lipid plane. The images are then processed and the different views are merged into a 3-D model *(3)* (**Fig. 5d**).

3.4. Decoration of 2-D Streptavidin Crystals

1. 2-D crystals of streptavidin were grown on biotinylated phospholipids as described earlier. Ten microlitres crystallisation buffer (20 mM Tris of pH 7.6, 40 mM NaCl) is placed in the Teflon well, and the lipid film is formed at the surface of the droplet by adding 1 μL undiluted biotinylated lipids at a concentration of 0.5 mg/mL in chloroform/hexane (1:1, v/v). The streptavidin preparation is diluted to a concentration of 300 μg/mL in crystallisation buffer, 5 μL of this suspension is injected in the sub-phase, and the protein–lipid layer was incubated for 1 h at 18°C.
2. The 2-D streptavidin crystals are transferred onto a holey carbon foil and washed 3 times by placing the electron microscopy

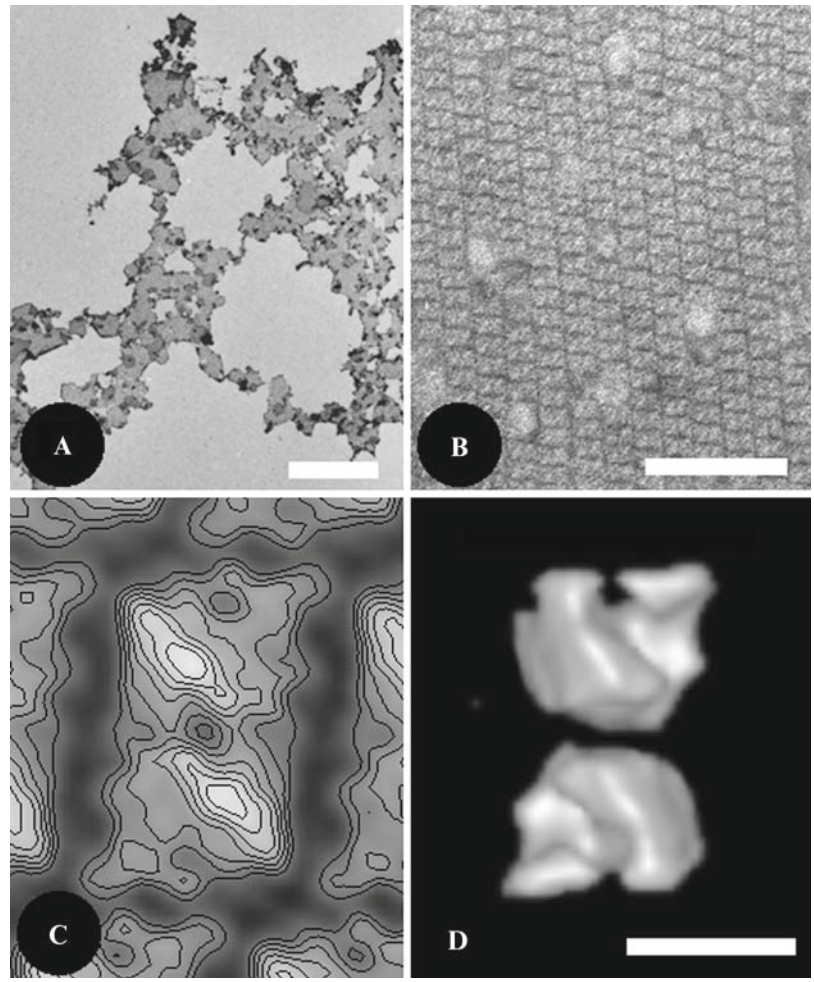


Fig. 5. Evaluation and exploitation of a 2-D crystallisation experiment. Histidine-tagged yeast RNA polymerase I is incubated with nickel-chelating lipids. (**A**) Low-magnification electron microscopy image showing the organisation of the protein complex into domains. The bar represents 5 μm (**B**) A higher magnification reveals ordered RNA polymerase arrays. The bar represents 50 nm (**C**) A noise-free image is obtained by averaging multiple molecular images. The stain excluding protein densities are in *white* and represented as *lines* of equal densities. (**D**) A 3-D model of the protein complex can be calculated by combining several views of the macromolecule obtained by tilting the crystals in the microscope. The bar represents 10 nm in (**C**) and (**D**).

grid on a 40-μL drop of crystallisation buffer to eliminate the unbound streptavidin molecules.

3. The support is then placed for 15 min on a 12-μL drop of a solution containing 0.8 pmol of biotinylated DNA or of nucleoprotein complexes.
4. The decorated 2-D crystals are rinsed 3 times with 60 μL drops of a solution containing 10 mM Tris of pH 7.4, 500 mM NaCl and can be processed for electron microscopy.

3.5. Electron Microscopy

1. The 2-D crystals are transferred to the electron microscopy grid through hydrophobic contacts between the lipid chains and the carbon foil (**Fig. 4e**). This is simply done by placing the grid over the well for 1–2 min. The grid is then withdrawn and prepared for observation (*see* **Notes 7 and 8**).
2. To be visualised by electron microscopy, the specimen can be contrasted by creating a mould of heavy atoms around the proteins, a process named negative staining. The transferred specimen placed on the grid held by forceps is washed with a drop of buffer that is quickly removed. The buffer is then replaced by a drop of a 1–2% aqueous solution of uranyl acetate, and the grid is dried by touching a piece of filter paper with the edge of the grid (*see* **Notes 9 and 10**). Alternatively, the specimen can be prepared for cryo-electron microscopy in which the specimen is transferred into the electron microscope in a frozen-hydrated state. The transferred 2-D crystals are rinsed twice with distilled water, blotted for 5–10 s with filter paper, and plunged into an ethane slush cooled with liquid nitrogen.
3. The crystallisation experiments have then to be evaluated in terms of protein concentration on the lipid layer and degree of organisation (*see* **Notes 11 and 12**). When the specimen is large enough (>50 kDa) individual molecules can be identified visually during electron microscopy inspection. To ascertain that the specimen is specifically bound to the lipid layer and not in a non-specific way to the carbon foil, it is useful to locate breaks in the lipid layer and to evidence a difference in binding efficiency with the underlying carbon foil. A frequently observed intermediate step in specimen ordering is the formation of symmetry-related oligomers which arise when a particular set of protein–protein interactions is energetically favoured. The formation of oligomers probably favours further organisation since interactions between symmetry-related surfaces will propagate forming linear polymers or 2-D crystals. Once larger crystalline areas are obtained, electron micrographs are recorded and the extent of order is evaluated by optical diffraction.

4. Notes

1. It is useful to check the specificity of the protein–lipid interactions. In the case of charged lipids, the protein binding should be reduced by increasing the ionic strength. In the case of functionalised lipids, the amount of transferred protein should

diminish by adding some competing ligand in solution. Note that in the case of nickel-chelating lipids, it was observed that addition of small amounts of imidazole prevented the non-specific aggregation of the protein and allowed to select the specific interaction with the poly-histidine tag *(17)*.

2. In order to better evaluate the organisational state of the molecule in the crystallisation experiment, it is useful to control its shape and size by direct adsorption of the sample on a carbon foil and negative staining. Such an experiment will also give an insight into the aggregation state of the protein in solution.
3. Detergents should be avoided in the incubation buffer since they may solubilise the lipid layer.
4. Another method of specimen transfer is the loop method *(23)*. A loop is formed with a thin Pt/Pd wire (0.075-mm diameter). The inside diameter must be slightly larger than the outside diameter of the electron microscopy grid. The loop is then lowered onto the drop, and the entire loop makes contact with the drop surface at the same time. This can be observed through drop deformation. The loop should not go through the monolayer and into the sub-phase so that no excess sub-phase is picked up. The loop is then gently and carefully raised and lowered onto a glow-discharged grid. The grid is held with forceps and is parallel to the film in the loop. The transfer is made by hydrophilic contacts between the carbon foil and the crystal. The film is then broken by tilting the loop to increase the angle between the film and the grid.
5. In some cases it was observed that the grid side on which the carbon foil was deposited affected crystal transfer *(17)*. This effect may be related to the surface roughness of the carbon foil and of the grid *(24)*.
6. To strengthen the crystals, 1 μL of a 0.5% glutaraldehyde solution can be added to the incubation drop before placing the electron microscopy grid in order to cross-link the specimen.
7. A good macroscopic indication that proteins bind to the lipid layers and that the transfer is efficient is obtained by visual inspection of the carbon surface after transfer. The originally hydrophobic grid becomes hydrophilic as assessed by the change in its wetting properties.
8. Storage of the carbon-coated grids in hexane atmosphere may provide higher reproducibility in the specimen transfer step by preventing adsorption of contaminating material.
9. Do not use phosphate buffers or buffers with high ionic strength which precipitate uranyl salts.

10. Other heavy metal solutions can be used for negative staining such as sodium phototungstate or ammonium molybdate.
11. The appearance of vesicular structures is often an indication for a too large excess of lipids. The working lipid solution should then be diluted.
12. To remove excess lipids, a detergent solution at low concentration can be used *(23)*. Care must be taken during this step as the drop might migrate to both sides of the grid and interfere with the staining process.

References

1. Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., and Downing, K.H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899–929.
2. Kuhlbrandt, W., Wang, D.N., and Fujiyoshi, Y. (1994). Atomic model of plant light-harvesting complex by electron crystallography. *Nature*. **367,** 614–621.
3. Amos, L.A., Henderson, R., and Unwin, P.N. Three-dimensional structure determination by electron microscopy of two-dimensional crystals. Prog. Biophys. Mol. Biol. **39**, 183–231.
4. van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., and Patwardhan, A. (2000). Single-particle electron cryo-microscopy: towards atomic resolution. *Q. Rev. Biophys.* **33**, 307–369.
5. Frank, J. (2002). Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 303–319.
6. Uzgiris, E.E., and Kornberg, R.D. (1983). Two-dimensional crystallization technique for imaging macromolecules, with application to antigen–antibody–complement complexes. *Nature*. **301**, 125–129.
7. Schultz, P., Celia, H., Riva, M., Darst, S.A., Colin, P., Kornberg, R.D., Sentenac, A., and Oudet, P. (1990). Structural study of the yeast RNA polymerase A. Electron microscopy of lipid-bound molecules and two-dimensional crystals. *J. Mol. Biol.* **216**, 353–362.
8. Darst, S.A., Ribi, H.O., Pierce, D.W., and Kornberg, R.D. (1988). Two-dimensional crystals of *Escherichia coli* RNA polymerase holoenzyme on positively charged lipid layers. *J. Mol. Biol.* **203**, 269–273.
9. Lebeau, L., Regnier, E., Schultz, P., Wang, J.C., Mioskowski, C., and Oudet, P. (1990). Two-dimensional crystallization of DNA gyrase B subunit on specifically designed lipid monolayers. *FEBS Lett.* **267,** 38–42.
10. Ribi, H.O., Reichard, P., and Kornberg, R.D. (1987) Two-dimensional crystals of enzyme-effector complexes: ribonucleotide reductase at 18-A resolution. *Biochemistry*. 26, 7974–7979.
11. Avila-Sakar, A.J., and Chiu, W. (1996). Visualization of beta-sheets and side-chain clusters in two-dimensional periodic arrays of streptavidin on phospholipid monolayers by electron crystallography [published erratum appears in *Biophys. J.*71, 517]. *Biophys. J.* **70**, 57–68.
12. Lebeau, L., Schultz, P., Célia, H., Mésini, P., Nuss, S., Klinger, C., Olland, S., Oudet, P., and Nioskowski, C. (1996). Specifically designed lipid assemblies as tools for two-dimensional crystallization of soluble macromolecules, in *Handbook of Nonmedical Applications of Liposomes 11* (Barenllolz, Y. and Lasic, D.D., eds.) CRC Press, Boca Raton, pp. 133–186.
13. Darst, S.A., Ahlers, M., Meller, P.H., Kubalek, E.W., Blankenburg, R., Ribi, H.O., Ringsdrof, H., and Kornberg, R.D. (1991). Two-dimensional crystals of streptavidin on biotinylated lipid layers and their interactions with biotinylated macromolecules. *Biophys. J.* **59**, 387–396.
14. Celia, H., Jontes, J.D., Whittaker, M., and Milligan, R.A. (1996). Two-dimensional crystallization of brush border myosin I. *J. Struct. Biol.* **117**, 236–241.
15. Venien-Bryan, C., Balavoine, F., Toussaint, B., Mioskowski, C., Hewat, E.A., Helme, B., and Vignais, P.M. (1997). Structural study of the response regulator HupR from rhodobacter capsulatus. Electron microscopy of two-dimensional crystals on a nickel-chelating lipid. *J. Mol. Biol.* **274**, 687–692.
16. Kubalek, E.W., Le Grice, S.F., and Brown, P.O. (1994). Two-dimensional crystallization of

histidine-tagged, HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid. *J. Struct. Biol.* **113**, 117–123.

17. Bischler, N., Balavoine, F., Milkereit, P., Tschochner, H., Mioskowski, C., and Schultz, P. (1998). Specific interaction and two-dimensional crystallization of histidine tagged yeast RNA polymerase I on nickel-chelating lipids. *Biophys. J.* **74**, 1522–1532.
18. Mosser, G., and Brisson, A. (1991). Structural analysis of two-dimensional arrays of cholera toxin B-subunit. *J. Electron Microsc. Tech.* **18**, 387–394.
19. Crucifix, C., Uhring, M., and Schultz, P. (2004). Immobilization of biotinylated DNA on 2-D streptavidin crystals. *J. Struct. Biol.* **146**, 441–451.
20. Kubalek, E.W., Kornberg, R.D., and Darst, S.A. (1991). Improved transfer of two-dimensional crystals from the air/water interface to specimen support grids for high-resolution analysis by electron microscopy. *Ultramicroscopy.* **35**, 295–304.
21. Fukami, A., and Adachi, K. (1965). A new method of preparation of a self-perforated micro plastic grid and its application. *J. Electron Microsc.* **14**, 112–118.
22. Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowall, A.W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228.
23. Asturias, F.J., and Kornberg, R.D. (1995). A novel method for transfer of two-dimensional crystals from the air/water interface to specimen grids. EM sample preparation/lipid-layer crystallization. *J. Struct. Biol.* **114**, 60–66.
24. Schmutz, M., and Brisson, A. (1996). Analysis of carbon film planarity by reflected light microscopy. *Ultramicroscopy.* **63**, 263–272.

Chapter 23

Assays for Transcription Factor Activity

Douglas Browning, Nigel Savery, Annie Kolb, and Stephen Busby

Summary

Transcription factors interact at promoters to modulate the transcription of genes. This chapter describes three in vitro methods that can be used to monitor their activity: transcript assays, abortive initiation assays, and potassium permanganate footprinting. These techniques have been developed using bacterial systems, and can be used to study the kinetics of transcription initiation, and hence to unravel regulatory mechanisms.

Key words: Gene regulation, RNA polymerase, Activator proteins, In vitro transcription, Abortive initiation, Potassium permanganate footprinting.

1. Introduction

Transcription activator proteins bind to specific sequences near promoters, and, when bound, they stimulate the initiation of transcription by RNA polymerase *(1)*. This chapter is concerned with in vitro methods for measuring the transcription activation function of this important class of proteins. In most cases, these methods are applied to activators that have been substantially purified, and for which the target sequences are known and the binding sites characterized (other chapters in this volume cover methods for locating and investigating the binding sites for such activators). Here we are concerned with the measurement of the products of activation. Since *Escherichia coli* transcription activators have been studied more than any others, we will take these as the paradigm, though these techniques are now being applied to all organisms for which in vitro systems have been developed.

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_23

The starting point of the methodology was the observation, made in the early 1970s, that purified *E. coli* RNA polymerase could initiate transcription at promoters in purified DNA *(2)*. With improvements in RNA methodology, it was found that the transcription start site in vitro, in many cases, was the same as in vivo. Further, at a number of promoters, interactions with specific transcription activators could be demonstrated and factor-dependent transcription in vitro occurred. The literature is now full of instances of factor-dependent transcription initiation with purified proteins and DNA, setting the scene for studies on the mechanism of transcription activation. However, as with all in vitro techniques, it is worth noting that the conditions found in the plastic tube differ greatly from those in vivo and that, for some instances, the reconstituted transcription system simply may not work.

Three principal methods can be used to monitor transcription activation: transcript assays, abortive initiation, and potassium permanganate footprinting to monitor open complex formation (**Fig. 1**). In transcript assays, RNA polymerase is incubated with a DNA template carrying the promoter of interest either with or without the transcription activator protein. Radioactively labelled nucleotides are then added and RNA polymerase molecules that have formed transcriptionally competent complexes start to make RNA. RNA polymerase molecules then "run" to a suitably placed downstream terminator (or to the end of the fragment), thus making labelled RNA of a discrete length that can be monitored and quantified *(3)*. In abortive initiation assays, nucleotide precursors corresponding just to the start of the message are added. RNA polymerase forms transcriptionally competent complexes at the chosen promoter, but elongation is prevented because some of the four nucleoside triphosphates are withheld *(4)*. The result is that RNA polymerase is "trapped" at the promoter and can only synthesize a short oligonucleotide, that is released as the polymerase cycles between a number of conformations in the abortive complex (it is termed abortive because a longer transcript cannot be made). The consequence of the cycling is that each polymerase molecule, trapped at a single promoter, synthesizes the oligonucleotide continuously, and the appearance of the product can be measured directly. The rate of product formation will be dependent on the number of promoters that are occupied by polymerase in an open complex. This, in turn, will depend on the activity of the transcription factor under study. Usually, transcript assays are used to locate transcription starts and to monitor factor activity qualitatively, whereas abortive initiation assays are exploited for quantitative and kinetic work.

Although the appearance of a transcript makes a good assay for the activity of a transcription factor, there may be situations where the transcript cannot be detected. In these cases, the best

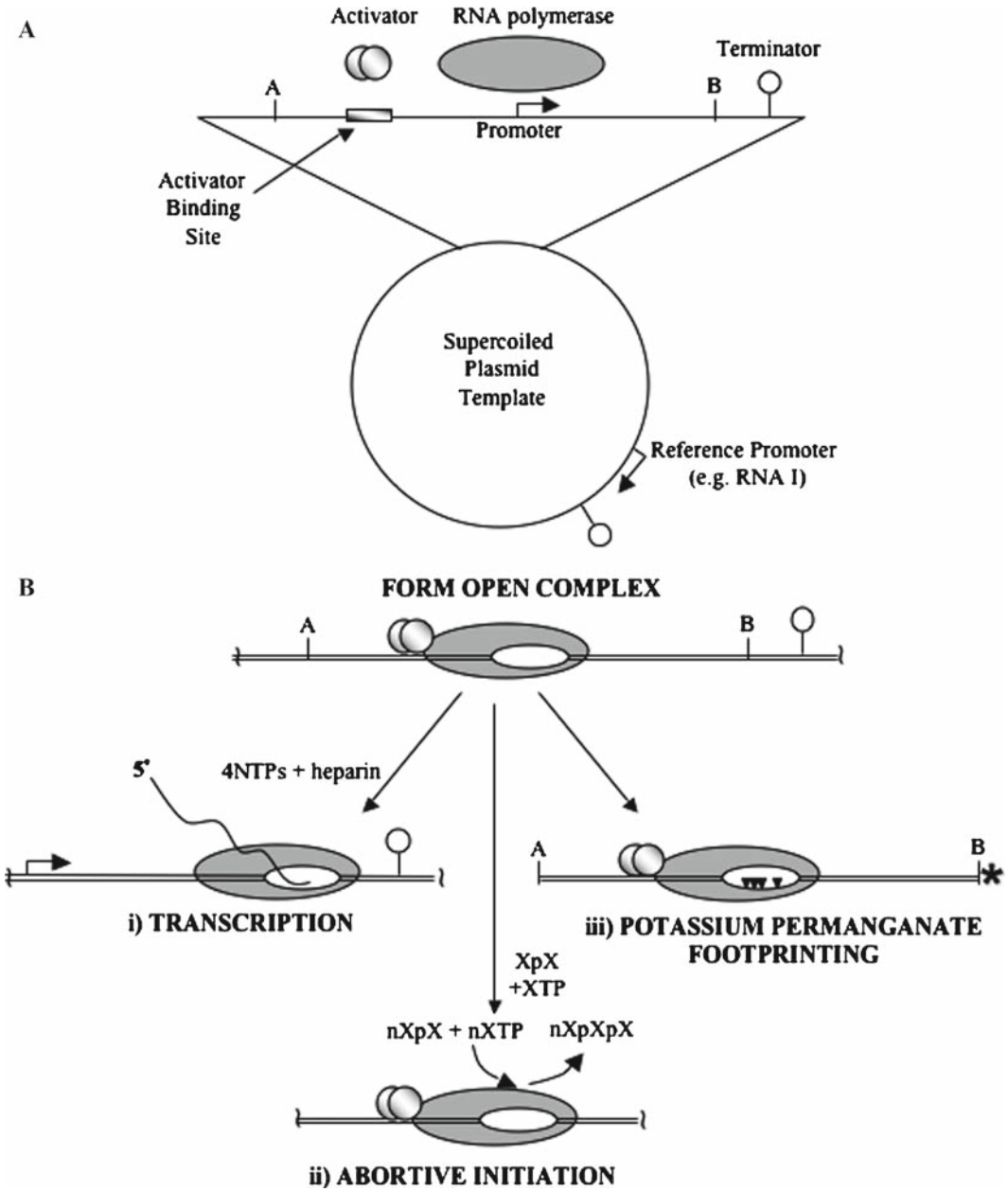


Fig. 1. Overview of techniques discussed in this chapter. (**A**) A plasmid carrying a test promoter, cloned on a restriction fragment, upstream of a terminator. The location of the reference RNA I promoter is shown. (**B**) The various techniques described: (i) transcription assays **(Subheading 3.1)**, (ii) abortive initiation **(Subheading 3.2)**, (iii) potassium permanganate footprinting **(Subheading 3.3)**.

strategy is to attempt to monitor the opening of the DNA strands by RNA polymerase directly. The transcription activator will promote open complex formation by RNA polymerase, resulting in unwinding of the DNA duplex around the promoter –10 sequence

and transcription start site. Many chemical reagents can be used to monitor this unwinding, but one of the simplest is potassium permanganate, that preferentially attacks non-base-paired thymines. Thus, potassium permanganate footprinting provides a simple method for checking that the activator is functional and provides information on the size of the region of unwinding in the open complex *(5, 6)*.

2. Materials

1. *E. coli* RNA polymerase. This enzyme can be purified *(7, 8)* or can be purchased from Epicentre Biotechnologies. The enzyme can also be reconstituted from individual subunits that have been overproduced using overexpression vectors *(9, 10)*. Purity can be easily checked by denaturing polyacrylamide gel electrophoresis and activity can be verified using standard templates. Preparations of RNA polymerase are usually stored at –20°C in buffer containing 50% glycerol.
2. Transcription factors. These must be purified, at least partially, away from any nuclease activities. Transcription assays can be used to monitor the purification.
3. Plasmids and DNA fragments. The DNA used for transcription assays should be free from nicks and RNase activity. Short DNA sequences carrying different promoters can be constructed using standard recombinant DNA methodology and cloned into plasmid vectors such that a strong transcription terminator is located downstream of the promoter *(11)*. Transcript assays can be performed directly on such recombinant plasmids, provided they have been purified, for example by using caesium chloride gradient centrifugation *(12)*. Linear, promoter-containing fragments can be generated by PCR or purified by a variety of methods from restricted plasmid DNA by polyacrylamide or agarose gels *(12)*. Although any DNA fragment can be used (*see* **Note 1**), fragments around 200–1,000 bp are most desirable. Stock solutions of most short fragments will be adjusted to around 20 μg/mL, the concentration being checked by gel electrophoresis.
   Linear DNA fragments for potassium permanganate footprinting are prepared from high-quality recombinant plasmid, purified using caesium chloride gradient centrifugation *(12)*. As only one end of the DNA molecule is to be labelled with ^{32}P, the fragment must be prepared by sequential digestion and dephosphorylation. 100 μg of plasmid DNA is linearized using a suitable restriction enzyme that generates 3′ recessed

ends (e.g. *Hin*dIII) and then treated with calf intestinal alkaline phosphatase (Invitrogen). After phenol/chloroform extraction and ethanol precipitation, the promoter fragment is released by digesting with a second restriction enzyme (preferably that generates 5′ recessed ends, e.g. *Aat*II). The fragment is then purified from either polyacrylamide or agarose gels prior to labelling *(12)*.

4. Transcription buffer: 20 mM Tris–HCl of pH 8, 100 mM NaCl, 5 mM $MgC1_2$, 0.1 mM EDTA, 1 mM DTT, 50 μg/mL nuclease-free BSA, 5% glycerol. This is a standard 1× buffer for many in vitro transcription assays and can be prepared as a 10× stock. There are many variations of this and the literature must be checked for any particular instance. For all reagents for transcription experiments, it is best to use water that has been treated with 0.3% diethylpyrocarbonate.
5. Heparin (Sigma). Make up as a 10 mg/mL stock solution in water.
6. Nucleotides. [α-^{32}P] UTP from Perkin Elmer can be used in conjunction with the four nucleoside triphosphates (Roche Diagnostics) for transcription assays. Most workers use final concentrations of 200 μM ATP/CTP/GTP, 10 μM UTP, 0.5–5.0 μCi of [α-^{32}P] UTP per reaction and 100 μg/mL heparin to prevent reinitiation. Typically, an 8× stock NTP + heparin solution is prepared containing 80 μM UTP, 1.6 mM of the three other NTPs, and 800 μg/mL heparin in 2× Transcription buffer. A "hot" NTP + heparin mix is then made by diluting this 1:1 with [α-^{32}P] UTP in water.

   [γ-^{32}P] ATP (MP Biomedical), specific activity 7,000 Ci/mmol, is used to end label linear DNA fragments for potassium permanganate footprinting. Each end-labelling reaction (volume 20 μL) contains 166 μCi of [γ-^{32}P] ATP, and the final concentration of linear DNA fragment is ~400 nM. Labelling is catalysed by T4 polynucleotide kinase (NEB) and unincorporated nucleotide is removed by passing the reaction mixture down a Sephadex G50 spin column. Alternatively, PCR can be used to generate a labelled fragment from a suitable template, by using an unlabelled oligo primer and a second primer that has been 5′ end labelled with [γ-^{32}P] ATP.
7. Dinucleotides. These can be bought from IBA Gmbh (Gottingen, Germany) and used without further purification as 10 mM stock solutions in water. The choice of the appropriate dinucleotide for priming abortive initiation assays is discussed in **Subheading 3.2.1**.
8. Transcription stop solution: 80% deionized formamide, 0.1% xylene cyanol FF, 0.1% bromophenol blue, 20 mM EDTA in the standard gel running buffer, 1× TBE.

9. Sequencing gels for analysis of RNA transcripts and permanganate DNA cleavage products. Standard 6% polyacrylamide gels containing 6–8 M urea and run in TBE *(12)* can be used to separate transcripts. A phosphor screen is used to detect the products.
10. Gel running buffer, TBE. This is usually made up in large volumes and kept as a 5× stock solution. To make up 1 L of 5× stock use 54 g Tris base, 27.5 g boric acid, and 20 mL of 0.5 M EDTA.
11. Whatman 3-MM paper. Cut into strips 20 cm in length for chromatography of abortive products.
12. Chromatography buffer: 18:80:2 (v/v) water/saturated ammonium sulphate/isopropanol.
13. RNA size markers. These are usually generated from transcripts of previously well-characterized DNA fragments. Alternatively, sequence ladders can be used.
14. 0.1 M EDTA.
15. Footprinting buffer: 20 mM HEPES of pH 8.0, 50 mM potassium glutamate, 5 mM $MgCl_2$, 1 mM DTT, 500 μg/mL bovine serum albumin. This is the standard 1× buffer and can be prepared as a 10× stock and stored at –20°C.
16. 200 mM $KMnO_4$ solution. Make up fresh and vortex well to ensure that all the $KMnO_4$ solid dissolves.
17. Permanganate stop solution: 3 M ammonium acetate, 0.1 mM EDTA of pH 8, 1.5 M β-mercaptoethanol. Make up fresh each time.
18. Phenol/chloroform/isoamyl alcohol (25:24:1) (Fluka).
19. 20 mg/mL Glycogen stock solution (Roche Diagnostics).
20. Ice-cold 100 and 70% ethanol (store at –20°C).
21. 10 M Piperidine stock (Sigma).
22. 3 M Sodium acetate, pH 5.2 *(12)*.
23. Gel fix solution. Ethanol (20%), acetic acid (10%).
24. Loading buffer: 95% (v/v) high-quality deionized formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF.

3. Methods

3.1. Transcript Assays

1. The first step is the binding reaction. Purified transcription factor and supercoiled plasmid or linear template DNA (*see* **Note 1**) is mixed gently in 1× transcription buffer (*see* **Note 2**)

in a final volume of 8 μL and incubated for 5 min at 37°C (*see* **Note 3**). It is important to include any cofactor required by the transcriptional activator. For example, the cyclic AMP receptor protein (CRP) requires cAMP in the transcription buffer for activity. Some other transcription activators require covalent modification, such as phosphorylation. The active form must be used.

2. Next 4 μL of RNA polymerase, diluted in 1× transcription buffer, is added and mixed in gently. The binding reaction is then incubated for a further 5–20 min at 37°C. Typically, the incubations are performed in a final volume of 12 μL with a template concentration of 0.5–5 nM, a range of transcription factor concentrations from 5 to 50 times the promoter concentration and up to 100 nM RNA polymerase.
3. The second step is the transcription reaction. Add 4 μL of 'hot' NTP + heparin mix (*see* **Note 4**) to each 12 μL binding reaction, mix gently, and incubate at 37°C for 5 min (*see* **Note 5**). For each individual DNA molecule, the RNA polymerase may or may not have reached an open complex depending on the activity of the transcription factor. At molecules where an open complex has formed, the polymerase will then "run" from the promoter to the downstream terminator (or to the end of the fragment) making a discrete sized RNA product. Since only one molecule of polymerase can occupy a promoter at any time, and since the inclusion of heparin prevents further initiation, the amount of any particular transcript will be directly proportional to the amount of open complex formation (*see* **Notes 6** and **7**).
4. Terminate the reactions by adding 12 μL of transcription stop solution. The samples can be stored for short periods on ice, or for longer periods at –20°C, until ready for loading on a sequencing gel.
5. Heat the samples for 2 min at 90°C and load 8 μL on a sequencing gel, together with size markers, and perform the electrophoresis. We routinely use an ECPS 3,000/150 electrophoresis power pack (Pharmacia), running the gel for 2–3 h at 60 W constant power. After running, the gel is dried, exposed to a phosphor screen, and the screen is scanned using a Molecular Imager FX scanner (Bio-Rad). Bands due to transcription initiation at the promoter under study can be identified and their size determined from calibrations. An example is shown in **Fig. 2** (*see* **Notes 8** and **9**).

3.2. Abortive Initiation Assays

1. Choose the nucleotides to be employed in the assay. Typically, this is done by selecting a dinucleotide appearing in the sequence anywhere from position –4 to +2 and using the next nucleotide as the labelled precursor. For example,

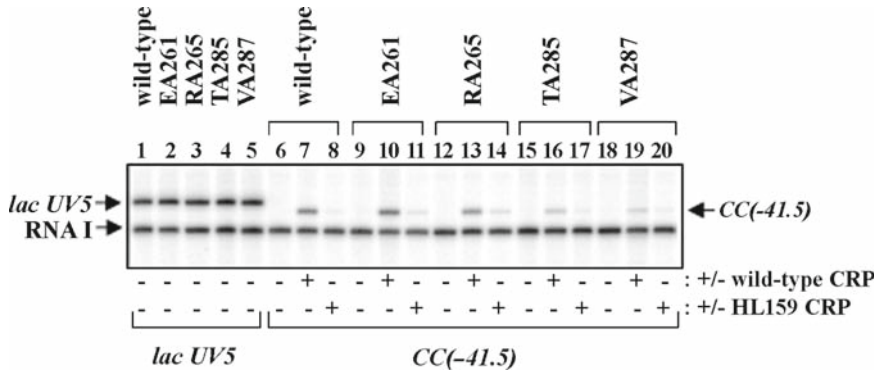


Fig. 2. In vitro transcription from plasmid carrying a CRP-independent promoter, *lacUV5* (*lanes 1–5*), or a CRP-dependent promoter, *CC(–41.5)* (*lanes 6–20*), cloned upstream of a transcription terminator. The figure shows the transcripts produced using purified RNA polymerase containing wild-type or mutant α subunits (EA261, RA265, TA285, and VA287 as indicated). Prior to the addition of RNA polymerase, CRP or CRP carrying the HL159 substitution (that interferes with the CRP-RNA polymerase interaction) was added to the reaction mixtures as indicated. The position of transcripts initiated at the *lacUV5* or *CC(–41.5)* promoters, and the position of the plasmid-encoded *RNA I* transcript are indicated (figure taken from *(13)*).

at the *E. coli galP1* promoter *(14)*, the sequence at the transcription start is 5′-TCATA-3′ with the central A as +1. The dinucleotide CpA and [α-^{32}P] UTP can be used to give the product, ^{32}P-labelled CpApU. It is important to ensure that no extended products can form (*see* **Note 10**).

2. Before the assay is performed, set up a series of Whatman 3-MM paper chromatograms (typically 20 cm long). Spot the origins with 20 µL of 0.1 M EDTA to ensure that product formation ceases the moment that the samples are loaded on the chromatogram.
3. Set up the standard assay, using concentrations of reagents as for the transcript analysis experiment. In a typical starting experiment, excess RNA polymerase and transcription factor will be premixed with DNA and incubated long enough to reach complete open complex formation. The experiment will be started by the addition of nucleotides. The final reaction mix will contain, for example, 0.5–5 nM promoter DNA, 100 nM RNA polymerase, 0.5 mM dinucleotide, and 0.05 mM UTP with 2.5 µCi [α-^{32}P] UTP in 100 µL. Run experiments both with and without the transcription factor and perform a control with no DNA.
4. At different times after addition of the [α-^{32}P] UTP, remove 15 µL aliquots and spot at the origin of the chromatogram. Aliquots can be taken every minute.
5. Develop the chromatogram using chromatography buffer. After the solvent front has progressed 20 cm, remove the

chromatogram and dry. Cut the paper into 5-mm slices and count each slice for Cerenkov radiation to locate the bands resulting from product and unincorporated UTP. For each time point, determine the number of counts incorporated into the product ($cpm_{product}$), and the number of counts in the unincorporated UTP (cpm_u). From the ratio of counts in the product to the total counts ($cpm_{product} + cpm_u$) the amount of product at each time-point can be deduced. Alternatively, the products can be analysed and quantified using a phosphorimager. In this case, it is sufficient to spot 2-μL aliquots onto the chromatogram, and the transcription reactions can be scaled down 3- to 5-fold (*see* **Note 11**).

A plot of product formed versus time should be linear, and from the slope, the rate of product formation can be deduced (*see* **Note 12**). The rate of product formation per promoter (TON, turn over number) can then be calculated from the molar amount of DNA fragment that was used in the experiment. A control run without the transcription factor will give factor-independent activity and will allow the effect of the activator to be quantified. Assuming that the rate of product formation in the presence of the activator reflects 100% occupancy at the promoter, the occupancy in the absence of the activator can be calculated (from the ratio of the TON values in the absence and the presence of the activator) (*see* **Note 13**).

6. The analysis can then be taken a stage further, for examples, *see* **refs. 15–18**. In the earlier experiment, RNA polymerase is preincubated with the DNA template prior to the addition of substrate, so product formation is linear from zero time. However, if the reaction is started by the addition of polymerase, the plot of product formation versus time shows a lag where the RNA polymerase "installs" itself at the promoter: This lag time (τ) can be measured and is a function of the initial binding of polymerase to the promoter and subsequent isomerizations to the open complex (*see* **Fig. 3**)

   The interaction of holoenzyme with promoters involves at least two steps: a rapid and reversible binding to promoter DNA, characterized by an association constant K_B, which leads to the "closed" inactive complex; followed by a conformation change to the "open" complex characterized by the rate constant k_f. For most promoters, the reverse of open complex formation is extremely slow. Thus, according to McClure *(4)*, the measured lag time (τ) is related to the enzyme concentration [RNP] by the relation,

$$\tau = 1 / k_f + 1 / (K_B k_f [\mathrm{RNP}]).$$

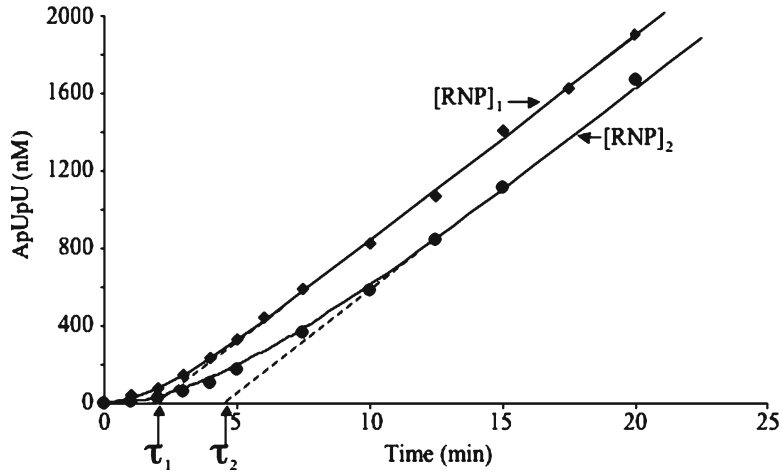


Fig. 3. Lag plots of a CRP-dependent promoter in the presence of different concentrations of RNA polymerase (50 nM $[RNP]_1$ and 16.7 nM $[RNP]_2$). A 15-μL sample of 4 nM DNA fragment in standard buffer containing CRP HL159 and cAMP was preincubated for 10 min at 37°C with 5 μL of a mixture containing 3 mM ApU and 300 μM UTP with 0.75 μCi [α–^{32}P]-UTP. At time 0, 10 μL of a prewarmed RNA polymerase solution was added (150 or 50 nM) and the reaction carefully mixed. At the indicated times, 2 μL portions of the reactions were removed for product quantification. The normalized quantity of ApUpU product was fitted using the Fig-P program according to the equation $Y = Vt - V\tau(1 - e^{-t/\tau})$, where V is the final steady-state velocity (mol of ApUpU per mol of promoter per minute). Care was taken to run the reaction until $t = 5\tau$ and to check that the final slope V was in agreement ±15% with the value of the TON, determined after preincubation of promoter and holoenzyme as described in **Subheading 3.2.5**.

To make a kinetic analysis for a promoter, it is necessary to perform the assays with a range of different RNA polymerase concentrations (typically 5–200 nM: for kinetic analysis, RNA polymerase should always be present in significant excess over promoter DNA). The lag time (τ) is measured in each case and is plotted as function of the reciprocal of the RNA polymerase concentration. This plot can be extrapolated to infinite RNA polymerase concentrations (the intersection with the Y-axis) to give the reciprocal of k_f, and K_B can be deduced from the intercept of the τ plot with the X-axis. Alternatively, K_B can be calculated from the ratio of the lag time at infinite enzyme concentration and the slope of the straight line. Data are normally fitted using a computer program, such as Enzfitter or Fig-P (*see* **Notes 14** and **15**). **Figure 4** shows τ plots resulting from experiments to analyse the contribution of the different activation regions of the *E. coli* cyclic AMP receptor protein on transcription activation at a target promoter.

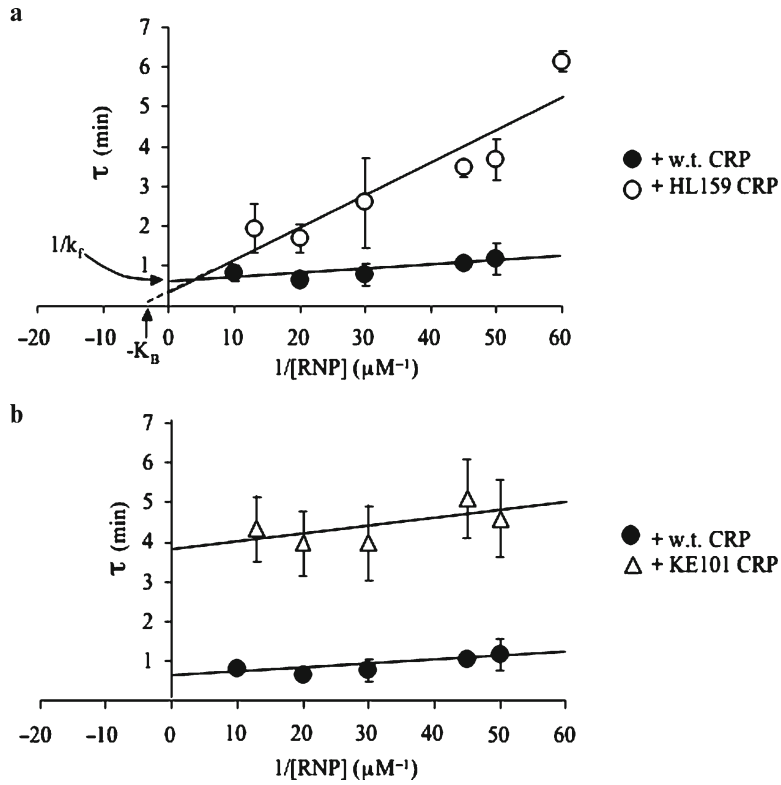


Fig. 4. Tau plots comparing the effects of different substitutions in CRP on transcription activation at a CRP-dependent promoter. The lag time (τ) before linear production of ApUpU is plotted against the reciprocal of RNA polymerase concentration. Plot (**a**) compares CRP carrying the HL159 substitution (which inactivates Activating Region 1 and decreases K_B) with wild-type CRP. Plot (**b**) compares CRP carrying the KE101 substitution (which inactivates Activating Region 2 and decreases k_f) with wild-type CRP. $K_B k_f$ and k_f can be calculated from the slope and intercept of each plot, respectively. Each data point represents the average of three independent assays and the error bars show one standard deviation either side of the mean (data taken from *(18)*).

3.3. Potassium Permanganate Footprinting

1. Purified transcription factor, RNA polymerase, and linear ^{32}P end-labelled template (*see* **Subheading 2.3** and **Note 16**) are mixed in 1× footprinting buffer in a final volume of 20 μL (*see* **Note 17**). Typically the template concentration is 1–4 nM and RNA polymerase is present at a final concentration of 50 nM. Reactions are left at 37°C for 30 min to allow open complex formation to take place.
2. One microliter of freshly prepared 200 mM $KMnO_4$ is added and reactions are incubated for a further 4 min at 37°C (*see* **Note 18**). This allows the $KMnO_4$ to attack the non-base-paired thymines within the –10 region due to RNA-polymerase-induced duplex unwinding.

3. Reactions are stopped by the addition of 50 μL of permanganate stop solution and then made up to a volume of 220 μL with sterile deionized water. Samples are extracted with 200 μL of phenol/chloroform/isoamyl alcohol, and the DNA in the aqueous phase (~200 μL) is precipitated by the addition

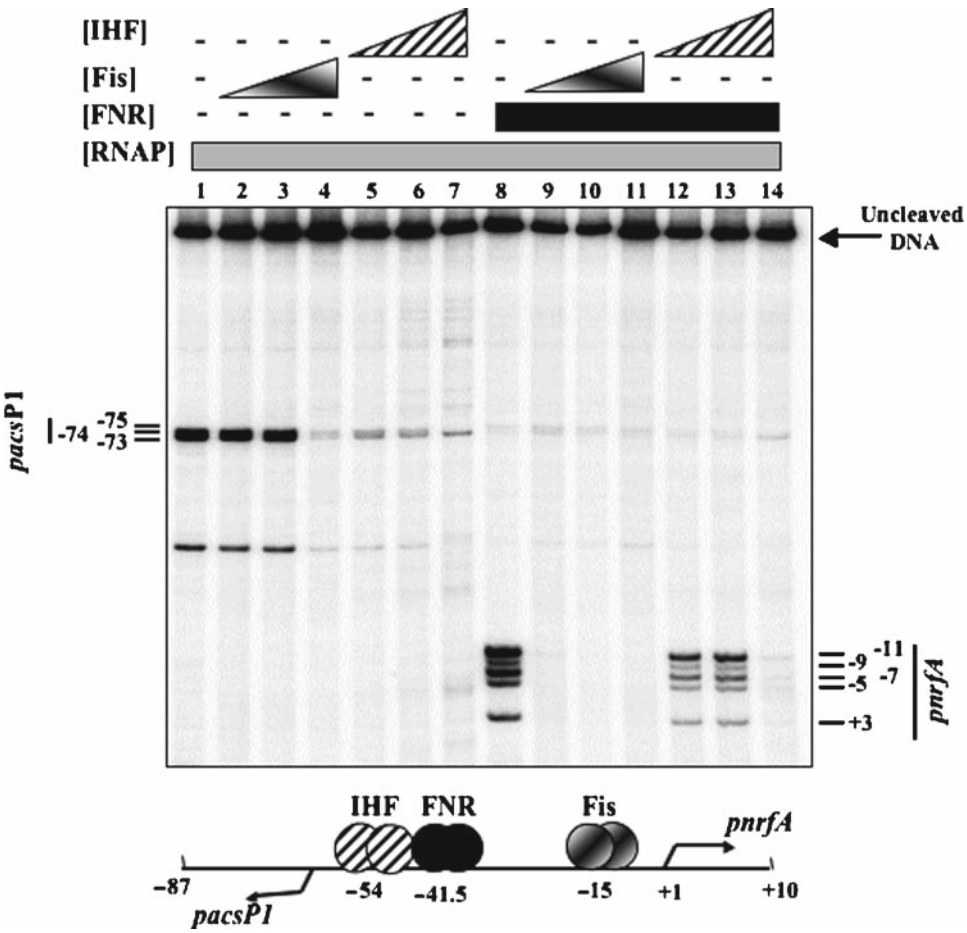


Fig. 5. In vitro potassium permanganate footprints using the *pnrf97* fragment, which carries the *E. coli pnrf* and *pacsP1* promoters. The figure shows the cleavage sites obtained during potassium permanganate footprinting when ^{32}P end-labelled *pnrf97* promoter fragment was incubated with purified Fis, IHF, and FNR proteins in various combinations. All reactions contained *E. coli* RNA polymerase and the location of cleavage sites is shown. Note that in the absence of RNA polymerase no cleavage products were detected (not shown). The sketch at the bottom of the figure shows the organization of the *pnrf97* fragment, which carries *E. coli nrf* sequences from position –87 to +10, the location of the *pnrf* and *pacsP1* promoters and the binding sites for Fis, IHF, and FNR (figure adapted from *(19)*).

of 1 μL 20 mg/mL glycogen and 400 μL of ice-cold 100% ethanol.

4. After 15 min at –70°C, the samples are centrifuged at 4°C for 15 min at 20,000 × *g*. The DNA pellet is washed with 600 μL of ice-cold 70% ethanol and centrifuged at 4°C for 10 min. The supernatant is discarded and the pellet dried under vacuum.
5. Resuspend the DNA pellet in 40 μL of 1 M piperidine and incubate at 90°C for 30 min. As piperidine is toxic, all steps involving piperidine must be carried out in a fume hood. Add 10 μL of 3 M sodium acetate of pH 5.2, 1 μL 20 mg/mL glycogen, and 70 μL sterile deionized water. Precipitate the DNA by adding 400 μL of ice-cold 100% ethanol and place samples at –70°C for 15 min.
6. Collect the DNA by centrifugation at 4°C for 15 min at 20,000 × *g*. Wash the pellet once with 600 μL of ice-cold 70% ethanol and centrifuge at 4°C for a further 10 min.
7. Dry the pellet under vacuum and resuspend the DNA in 8 μL of loading buffer.
8. Heat the samples for 2 min at 90°C and load 4 μL on a sequencing gel, together with size markers (*see* **Note 19**). Electrophoresis is carried out as in **Subheading 3.1.5** and the gel is fixed for 15 min in 10% methanol, 10% acetic acid. Once the gel is dried, bands are visualized using a phosphor screen and imager. It is possible to identify RNA-polymerase-induced unwinding of the –10 region and the effect that activator and repressor proteins have on transcription initiation (*see* **Notes 20** and **21**). An example is shown in **Fig. 5**.

4. Notes

4.1. Transcript assays

1. Transcript assays are generally performed on templates with the transcription start of interest positioned 50–150 bp upstream from a transcription terminator or the end of the fragment. Often a longer fragment that carries more than one promoter will be chosen: longer transcripts can be sized by running the sequence gels further. Individual transcripts can be identified by using families of fragments that are truncated from one end. Transcription assays can be performed using both relaxed and supercoiled DNA. Reference promoters can be used to aid in the quantification of transcripts (e.g. if *colE1* plasmid derivatives are used as vectors for the promoter under study, the 107 nucleotide transcript from the RNA I promoter can be used, *see* **Figs 1** and **2**).

2. A number of alternative buffer systems can be used and the final choice is largely a matter of trial and error. An alternative system is 40 mM Tris of pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, and 100 μg/mL acetylated bovine serum albumin. In some cases the effects of substituting different anions or cations may be significant *(20)*. Many recent studies have used glutamate-containing buffers to enhance DNA binding of different factors.
3. Care should be taken to avoid introducing RNase contamination during protein and DNA purifications, in the preparation of solutions and in the handling of plasticware. If necessary, commercially available RNase inhibitors can be added to the transcription reactions to counteract low levels of nuclease contamination.
4. The inclusion of heparin in assays ensures a single round of transcript formation. However, multi-round assays can be performed by omitting the heparin. This can be useful when working with promoters where the open complex is sensitive to heparin.
5. The transcription step of the protocol is very fast and complete in minutes. Sometimes a doublet band is seen corresponding to a particular transcript. This is often due to 'hesitation' by the polymerase at the end of the transcript. The relative intensities of the doublet can depend on temperature or the length of time of the elongation step. In some cases, multiple bands are caused by ambiguity in the starting base of the transcript. This can be resolved by working with [α-^{32}P]-labelled initiating nucleotide, for example *see* **ref. 21**.
6. The kinetics of open complex formation can be monitored using transcript assays. After addition of polymerase, take aliquots at different times and add to the 'hot' NTP + heparin mix. Since heparin blocks reinitiation, the amount of transcript from that sample will be proportional to the amount of open complex formed at that time, for an example *see* **ref. 22**. Typically, the half-time for open complex formation ranges from 20 s to 30 min. In principle it is possible to make these measurements at different polymerase concentrations and make the τ plot analysis, as for abortive initiation, but, in practice, this is extremely difficult and the abortive initiation assay is preferable.
7. Elongation can be studied by performing open complexes and then adding nucleotide precursors one by one. 3′ *O*-methyl *(23)* or deoxy *(24)* derivatives of nucleotides can be used to freeze elongation complexes at particular lengths.

8. Transcript analysis assays provide a simple method for monitoring the effects of transcription factors and their cofactors. However, it can also be exploited to investigate effects of conditions (e.g. temperature, salt, etc.) on open complex formation. It is important to note that changes may affect elongation rather than transcription initiation. This can be checked simply by preforming open complexes and then altering the conditions. In our experience, the elongation step is usually unaltered by changes in the assay conditions, and differences reflect changes at one or other step in the formation of the open complex *(25)*.
9. Many, but not all, promoters are active in transcript assays and there is no way of predicting whether a particular activator will or will not work in vitro. Many workers find that such experiments produce more bands than "ought" to be seen. In particular, some run-off transcripts made with purified fragments as templates exceed the size of the template fragment *(3)*. This results from RNA polymerase molecules failing to stop when reaching the end of the fragment, turning around, and continuing to transcribe the opposite strand. This effect can be partially circumvented by lowering NTP concentrations or decreasing the temperature. Another problem may arise because any DNA sequence will contain a number of potential transcription starts that are normally not used in vivo, since the competition for polymerase in vivo favours stronger promoters. In vitro conditions are such that there is little discrimination against weak promoters, for example *see* **ref. 26**. If the appearance of bands from these weak promoters "spoils" the results, they can be reduced by using higher salt concentrations or lower concentrations of polymerase to increase specificity.

4.2. Abortive Initiation Assays

10. A great feature of the abortive initiation assay is that it can be performed on promoters carried by both circular DNA and linear fragments: the dinucleotide primer picks out one promoter from others. Obviously, there is more chance of interference from other promoters with longer DNAs. Thus, if working with circular plasmid, it is prudent to test the reaction using plasmid either with or without the insertion carrying the promoter under study. It may be possible to reduce interfering signals from the vector by altering the dinucleotide used. Some primers can be used without being completely specific for the promoter tested. For instance, CpA and UTP gives the trinucleotide CpApU at *gal*P1, but also the longer oligonucleotides CpApUpU and CpApUpUpU starting from *gal*P2, which can be separated on the chromatogram *(27)*.

11. The abortive initiation assay is tedious because of the chromatographic analysis of the products, which takes 2–3 h. One way to accelerate the procedure is to replace radioactive UTP with a fluorescent analogue, UTP-γ-ANS (1-naphthylamine-5-sulphonic acid UTP). The assay can then be measured fluorometrically by following the increase in light emission caused by the release of the pyrophosphate-ANS moiety each time a unit is incorporated *(28)*. A considerable advantage of this method is that it allows the continuous monitoring of product formation. A disadvantage is that the fluorescent label may alter the kinetics, although, to date, this has not been reported.
12. In some cases, product formation may never become linear with respect to time. Assuming that there are no contaminating nucleases, this is likely to be a result of the consumption of nucleoside triphosphates, which reduces the reaction velocity. This can be overcome by lowering the dinucleotide concentration. Ideally any time course needs to be run for at least 5 times τ.
13. Before starting any kinetics, it is advisable to check chosen combinations of primer and nucleotide for specificity and for product formation: a TON value of <10/min is useless for kinetic studies. Some promoters give no abortive cycling reaction, whilst others may give homopolymer synthesis caused by slippage in the enzyme's active site *(29)*, rendering the abortive initiation assay useless.
14. The most powerful use of abortive initiation is to determine the microscopic rate constants of individual steps during transcription initiation. Measurements of these rates in the absence or presence of a transcription factor can provide mechanistic information about the enzymology of activation. However, the method relies on a number of assumptions that are true for most, but not all, promoters *(4)*. First, active RNA polymerase must be present in significant (i.e. >5×) excess over the promoter DNA; second, the isomerization from closed to open complex must be essentially irreversible over the time course of the experiment; and third, in order for the equation in **Subheading 3.2** to hold true, the closed complex must be in rapid equilibrium with free polymerase and DNA.
15. In different situations, transcription activators can affect K_B *(16)*, k_f *(15)*, or TON *(16)*. In a small number of cases, transcription factors have no effect on abortive initiation parameters. In such instances the activator cannot be intervening at the level of open complex formation, but must be affecting later steps of the transcription process, for example *see* **ref. 30**. Such situations can be analysed by single or multiple rounds

of transcript assays. Note that in some complex cases (e.g. overlapping promoters), microscopic rate parameters cannot be deduced from abortive initiation assays *(31)*.

4.3. Potassium Permanganate Footprinting

16. Potassium permanganate footprinting can also be carried out using supercoiled plasmid template. Reactive thymine bases can be detected using primer extension with a ^{32}P end-labelled oligonucleotide. Potassium permanganate can also be used to detect unwinding at promoters in vivo in whole cells, with target promoters on the chromosome or on multicopy plasmids *(6)*.
17. A typical experiment will include reactions with DNA alone, DNA plus RNA polymerase and DNA plus RNA polymerase and activator protein. For some DNA sequences, a limited amount of DNA cleavage occurs in the absence of added proteins. This is due to intrinsic distortion and must be taken account of when data are analysed.
18. Incubation times with potassium permanganate can be varied from 1 to 4 min depending on the intensity of cleaved fragment obtained from pilot experiments.
19. A suitable DNA marker can be obtained by treating the end-labelled DNA fragment with formic acid and piperidine (Maxam–Gilbert sequencing protocol) to generate an "A + G" DNA ladder *(12)*.
20. The cleavage of non-base-paired thymines is also dependent on which DNA strand has been end labelled. As RNA polymerase interacts intimately with the non-template stand during promoter unwinding, this may protect some unpaired thymine bases from permanganate modification. DNA fragments that have been end labelled on the template strand are not protected in this manner *(6)*.
21. This type of footprinting is very sensitive as it results in a signal on a clear background. In one application, the cleavage pattern of a DNA fragment carrying two promoters was examined in the presence of RNA polymerase. The analysis showed that, for each individual DNA molecule, an open complex could form at one promoter or the other, but not at both promoters simultaneously *(32)*.

References

1. Browning, D. and Busby, S. (2004). The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65.
2. Losick, R. and Chamberlin, M., eds. (1976). *RNA Polymerase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
3. Zubay, G. (1980). The isolation and properties of CAP, the catabolite gene activator. *Methods Enzymol.* **65**, 856–877.
4. McClure, W. (1980). Rate-limiting steps in RNA chain initiation. *Proc. Natl Acad. Sci. U. S. A.* **77**, 5634–5638.

5. Sasse-Dwight, S. and Gralla, J. (1989). KMnO4 as a probe for lac promoter DNA melting and mechanism *in vivo*. *J. Biol. Chem.* **264**, 8074–8081.
6. Chan, B., Minchin, S., and Busby, S. (1990). Unwinding of the duplex DNA during transcription initiation at the *Escherichia coli* galactose operon overlapping promoters. *FEBS Lett.* **267**, 46–50.
7. Burgess, R. and Jendrisak, J. (1975). A procedure for the rapid, large-scale purification of *Escherichia coli* DNA-dependent RNA polymerase involving Polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* **14**, 4634–4638.
8. Hager, D., Jun Jin, D., and Burgess, R. (1990). Use of mono Q high resolution ionic exchange chromatography to obtain highly pure and active *Escherichia coli* RNA polymerase. *Biochemistry* **29**, 7890–7894.
9. Tang, H., Severinov, K., Goldfarb, A., and Ebright, R. (1995). Rapid RNA polymerase genetics: one-day, no-column preparation of reconstituted recombinant *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. U. S. A.* **92**, 4902–4906.
10. Fujita, N. and Ishihama, A. (1996). Reconstitution of RNA Polymerase. *Methods Enzymol.* **273**, 121–130.
11. Kolb, A., Kotlarz, D., Kusano, S., and Ishihama, A. (1995). Selectivity of the *Escherichia coli* RNA polymerase $E\sigma^{38}$ for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res.* **23**, 819–826.
12. Sambrook, J., Fritsch, E., and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Second Edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
13. Savery, N., Lloyd, G., Kainz, M., Gaal, T., Ross, W., Ebright, R., Gourse, R., and Busby, S. (1998). Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *EMBO J.* **17**, 3439–3447.
14. Herbert, M., Kolb, A., and Buc, H. (1986). Overlapping promoters and their control in *Escherichia coli: the* gal case. *Proc. Natl Acad. Sci. U. S. A.* **83**, 2807–2811.
15. Hawley, D. and McClure, W. (1982). Mechanism of activation of transcription initiation from the λ P_{RM} promoter. *J. Mol. Biol.* **157**, 493–525.
16. Malan, T., Kolb, A., Buc, H., and McClure, W. (1984). Mechanism of CRP-cAMP activation of *lac* operon transcription initiation: activation of the P1 promoter. *J. Mol. Biol.* **180**, 881–909.
17. Leirmo, S. and Gourse, R. (1991). Factor independent activation of *E. coli* rRNA transcription (I) Kinetic analysis of the role of the upstream activator region and supercoiling on transcription of the *rrnB* P1 promoter *in vitro*. *J. Mol. Biol.* **220**, 555–568.
18. Rhodius, V., West, D., Webster, C., Busby, S., and Savery, N. (1997). Transcription activation at Class II CRP-dependent promoters: the role of different activating regions. *Nucl. Acids Res.* **25**, 326–333.
19. Browning, D.F., Grainger, D.C., Beatty, C.M., Wolfe, A.J., Cole, J.A., and Busby, S.J. W. (2005). Integration of three signals at the *Escherichia coli nrf* promoter: a role for Fis protein in catabolite repression. *Mol. Microbiol.* **57**, 496–510.
20. Leirmo, S., Harrison, S., Cayley, D., and Burgess, R. (1987). Replacement of potassium chloride by potassium glutamate dramatically enhances protein DNA interactions *in vitro*. *Biochemistry* **26**, 2095–2101.
21. Hsu, L. (1996). Quantitative parameters for promoter clearance. *Methods Enzymol.* **273**, 59–71.
22. Chan, B. and Busby, S. (1989). Recognition of nucleotide sequences at the *Escherichia coli* galactose operon *P1* promoter by RNA polymerase. *Gene* **84**, 227–236.
23. Straney, D. and Crothers, D. (1985). Intermediates in transcription initiation from the *E. coli lac* UV5 promoter. *Cell* **43**, 449–459.
24. Krummel, B. and Chamberlin, M. (1992). Structural analysis of ternary complexes of *E. coli* RNA polymerase: individual complexes halted along different transcription units have distinct and unexpected biochemical properties. *J. Mol. Biol.* **225**, 221–237.
25. Grimes, E., Busby, S., and Minchin, S. (1991). Different thermal energy requirement for open complex formation by *Escherichia coli* RNA polymerase at two related promoters. *Nucleic Acids Res.* **19**, 6113–6118.
26. Ponnambalam, S., Spassky, A., and Busby, S. (1987). Studies with the *Escherichia coli* galactose operon regulatory region carrying a point mutation that simultaneously inactivates the two overlapping promoters. Interactions with RNA polymerase and the cyclic AMP receptor protein. *FEBS Lett.* **219**, 189–196.
27. Goodrich, J. and McClure, W. (1992). Regulation of open complex formation at the galactose operon promoters. Simultaneous interaction of RNA polymerase, *gal*

repressor and CAP/cyclic AMP. *J. Mol. Biol.* **224**, 15–29.

28. Bertrand-Burggraf, E., Lefevre, J. F., and Daune, M. (1984). A new experimental approach for studying the association between RNA polymerase and the *tet* promoter of pBR322. *Nucleic Acids Res.* **12**, 1697–1706.
29. Qi, F., Liu, C., Heath, L., and Turnbough, C. (1996). *In vitro* assay for reiterative transcription during transcriptional initiation by *Escherichia coli* RNA polymerase. *Methods Enzymol.* **273**, 71–85.
30. Menendez, M., Kolb, A., and Buc, H. (1987). A new target for CRP action at the *malT* promoter. *EMBO J.* **6**, 4227–4234.
31. Gussin, G. (1996). Kinetic analysis of RNA polymerase-promoter interactions. *Methods Enzymol.* **273**, 45–59.
32. El-Robh, M. and Busby, S. (2002). The *Escherichia coli* cAMP receptor protein bound at a single target can activate transcription initiation at divergent promoters: a systematic study that exploits new promoter probe plasmids. *Biochem. J.* **368**, 835–843.

Chapter 24

Ultraviolet Crosslinking of DNA–Protein Complexes via 8-Azidoadenine

Rainer Meffert, Klaus Dose, Gabriele Rathgeber, and Hans-Jochen Schäfer

Summary

The synthesis of 8-azido-2′-deoxyadenosine-5′-triphosphate is described. The photoreactive dATP analog was characterized by thin layer chromatography and UV spectroscopy. Its photoreactivity upon UV irradiation was studied. After incorporation of this dATP analog by nick translation into DNA containing the *tet* operator sequence the investigation of the interactions between *tet* operator DNA and Tet repressor becomes possible. Photocrosslinking of protein to DNA was demonstrated by the reduced migration of the DNA protein crosslinks in SDS polyacrylamide gel electrophoresis.

Key words: 8-azido-2′-deoxyadenosine-5′-triphosphate, Nick translation, **tet** Operator, Tet repressor, Photoaffinity crosslinking, DNA protein crosslinks.

1. Introduction

In biological systems photoreactive derivatives have been widely applied to study specific interactions of receptor molecules with their ligands by photoaffinity labeling *(1–3)*.

$$\text{Receptor} + \text{Ligand} \rightleftharpoons [\text{Receptor} \cdot \text{Ligand}]$$

Receptors are generally proteins like enzymes, immunoglobulins, or hormone receptors, for example. The ligands, however, differ widely in their molecular structure (e.g., sugars, amino acids, nucleotides, or oligomers of these compounds).

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_24

The advantage of photoaffinity labeling compared with affinity labeling, or chemical modification with group-specific reagents is that photoactivatable nonreactive precursors can be activated at will by irradiation (**Fig. 1**). These reagents do not bind covalently to the protein unless activated. On irradiation of the precursors, highly reactive intermediates are formed that react indiscriminately with all surrounding groups. Therefore, after activation a photoaffinity label – interacting at the specific binding site – can label all the different amino acid residues of the binding area. Today aromatic azido compounds are mostly used as photoactivatable ligand analogs. They form highly reactive nitrenes upon irradiation because of the electron sextet in the outer electron shell of these intermediates (**Fig. 2**).

In addition to the azido derivatives photoreactive precursors forming radicals or carbenes on irradiation can be used as photoaffinity labels as well. All these intermediates (nitrenes, for example) vehemently try to complete an electron octet (**Fig. 3**).

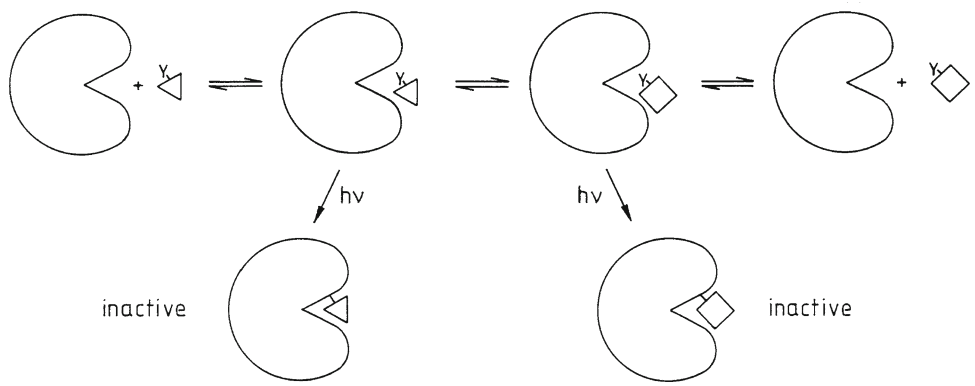


Fig. 1. Photoaffinity labeling of receptor proteins (e.g., enzymes) by photoactivatable ligand analogs (e.g., substrate analog/product analog). In the dark (upper line) the biological interactions of the protein with the ligand analog can be studied. On irradiation (lower line) the protein (enzyme) is labeled and inactivated by the substrate analog/product analog.

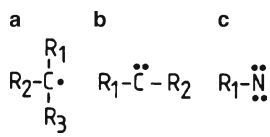


Fig. 2. Highly reactive photogenerated intermediates: radical (**A**), carbene (**B**), nitrene (**C**).

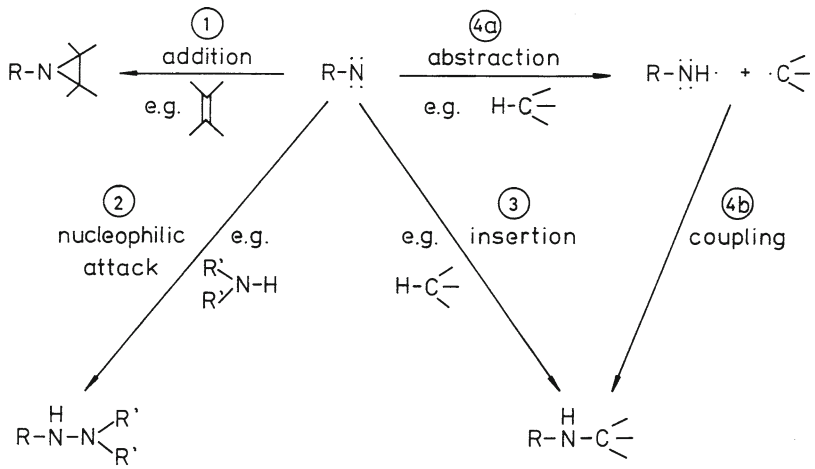


Fig. 3. Reactions of nitrenes. Cycloaddition to multiple bonds forming three-membered cyclic imines (1), addition to nucleophiles (2), direct insertion into C–H bonds yielding secondary amines (3), hydrogen atom abstraction followed by coupling of the formed radicals to a secondary amine (4a, 4b).

To produce covalent crosslinks between proteins and DNA, various methods have been applied *(4–11)*: UV irradiation, γ-irradiation, chemical methods, and even vacuum or extreme dryness. Besides these methods, photoaffinity labeling and photoaffinity crosslinking are helpful tools for the study of specific interactions between proteins and deoxyribonucleic acids. To date several successful attempts have been made to photocrosslink proteins to nucleic acids using different photoactivatable deoxynucleotides. 5-Bromo-, 5-iodo-, 5-azido-, and 5-[N-(p-azidobenzoyl)-3-aminoallyl]-2′-deoxyuridine-5′-mon10-phosphate *(12–19)*, 4-thio-2′-deoxythymidine-5′-monophosphate *(20)*, exo-N-{2-[N-(4-azido-2,5-difluoro-3-chloropyridine-6-yl)-3-aminopropionyl]aminoethyl}-2′ -deoxycytidine-5′-monophosphate *(21, 22)*, 5-{N-[N-(4-azido-2,5-difluoro-3-chloropyridine-6-yl)-3-aminopropionyl]trans-3-aminopropenyl-1}-2′ deoxyuridine-5′ -monophosphate *(21)*, and the dATP analogs N^6-[4-azidobenzoyl-(2-aminoethyl)]-2 -deoxyadenosine-5 -monophosphate *(23)*, N^6-{4-[3-(trifluoromethyl)-diazirin-3-yl]benzoyl-[2-aminoethyl]}-2 -deoxyadenosine-5′ -monophosphate *(23)* and 8-azido-2′-deoxyadenosine-5′-monophosphate *(24, 25)* have been incorporated into deoxyribonucleic acids to bind DNA covalently to adjacent proteins (for review *see* **ref.** *26)*.

Here we describe the synthesis of 8-azido-dATP (8-N_3dATP), its incorporation into DNA by nick translation, and the procedure to photocrosslink azido-modified DNA to proteins *(24, 25)*.

2. Materials

2.1. Synthesis of 8-N_3dATP

1. dATP (disodium salt, Boehringer Mannheim, Mannheim, Germany).
2. Potassium acetate buffer: 1 M, pH 3.9.
3. Bromine.
4. Sodium disulfite ($Na_2S_2O_5$).
5. Ethanol.
6. DEAE-Sephadex A-25.
7. Triethylammonium bicarbonate buffer: 0.7 M, pH 7.3.
8. Dimethylformamide.
9. Hydrazoic acid: 1 M in benzene.
10. Triethylamine.

2.2. Characterization of 8-N_3dATP

Silica gel TLC plates F254 (Merck, Darmstadt, Germany), Cellulose F TLC plates (Merck), isobutyric acid/water/ammonia (66:33:1 v/v), n-butanol/water/acetic acid (5:3:2 v/v).

2.3. Preparation of Azido-Modified DNA

1. DNA (e.g., pBR 322 or pWH 106).
2. Deoxyribonucleotides (dATP, dGTP, dCTP, dTTP, α-[^{32}P]-dCTP).
3. DNase I (*E. coli,* 2,000 U/mg) in 0.15 M NaCl, 50% glycerol.
4. Tris(hydroxymethyl)aminomethane–HCl buffer, Tris–HCl: 50 mM, pH 7.2.
5. Magnesium sulfate ($MgSO_4$).
6. Bovine serum albumin.
7. DNA polymerase I (*E. coli*), (Roche Applied Science or similar).
8. Ethylenediaminetetraacetic acid disodium salt (EDTA).
9. Sephadex A-25.

2.4. Photocrosslinking

UV lamp (e.g., Mineralight handlamp UVSL 25 (UV Products) at position "long wave") emitting UV light at wavelengths of 300 nm and longer.

3. Methods

3.1. Synthesis of 8-N_3dATP

The synthesis of 8-N_3dATP (**Fig. 4**) is performed principally by analogy to the synthesis of 8-N_3ATP *(27)* (*see* **Note 1**). In the first step bromine exchanges the hydrogen at position 8

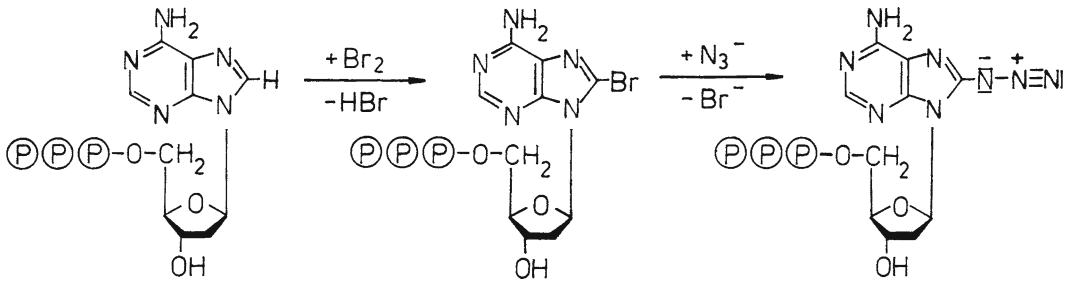


Fig. 4. Synthesis of 8-N_3dATP.

of the adenine ring. Then the bromine is substituted by the azido group.

1. Dissolve 0.2 mmol (117.8 mg) of dATP in 1.6 mL of potassium acetate buffer (1 M, pH 3.9) and add 0.29 mmol (15 µL) of bromine. Keep the reaction mixture in the dark at room temperature for 6 h (the absorption maximum shifts from 256 to 262 nm; *see* **Note 2**).
2. Reduce excessive bromine by addition of traces of (ca. 5 mg) $Na_2S_2O_5$ until the reaction mixture looks colorless or pale yellow. Pour the reaction mixture into 20 mL of cold ethanol (10°C) and allow to stand for at least 30 min at 70°C in the dark.
3. Collect the precipitated deoxynucleotide by centrifugation and redissolve the residue in 0.5 mL of double-distilled water. Further purification is achieved by ion exchange chromatography over DEAE-Sephadex A-25 column (50 × 2 cm) with a linear gradient of 1,000 mL each of water and triethylammonium bicarbonate (0.7 M, pH 7.3).
4. Combine the fractions containing 8-bromo-dATP (8-BrdATP) (main peak of the elution profile) and dry the solution by lyophilization. 8-BrdATP is obtained as the triethylammonium salt. The expected yield should be 65% (spectroscopically).
5. Dissolve 0.1 mmol (87.3 mg) of dried 8-BrdATP (triethylammonium salt) in 3 mL of freshly distilled dimethylformamide (*see* **Notes 3** and **4**). Add a dried solution of 0.8 mmol (34.4 mg) of hydrazoic acid (HN_3) in 800 µL of benzene and 0.8 mmol (111.3 µL) of freshly distilled triethylamine. Keep the reaction mixture in the dark at 75°C for 7 h (the absorption maximum shifts from 262 to 280 nm).
6. Evaporate the solvents under vacuum and redissolve the residue in 1 mL of water. Further purification is achieved by ion exchange chromatography over DEAE-Sephadex A-25 as

described in **Subheading 3.1.3**. **Figure 5** shows the elution profile of the chromatography (*see* **Notes 4 and 5**).

7. Combine the fractions containing 8-N_3dATP and dry the solution by lyophilization. 8-N_3dATP is obtained as the triethylammonium salt. Yield: 30% (spectroscopically). 8-N_3dATP can be stored at –20°C in the dark (*see* **Notes 6–8**), freeze-dried, or frozen in aqueous solution, pH 7.0.

3.2. Characterization of 8-N_3dATP

1. Thin layer chromatography. TLC is carried out on silica gel plates F254 or cellulose F plates. The development is performed in either isobutyric acid/water/ammonia (66:33:1 v/v) or n-butanol/water/acetic acid (5:3:2 v/v).
2. UV absorbance. Record the UV absorbance spectrum of 8-N_3dATP. It shows a maximum at 280 nm. The UV absorbance of 8-N_3dATP is pH dependent (*see* **Note 9**).
3. Photoreactivity. The photoreactivity of 8-N_3dATP is tested twice by different methods. It can either be demonstrated by the spectroscopic observation of the photolysis (**Fig. 6;** *see* **Note 10**) or by the ability of the photolabel to bind irreversibly to cellulose on thin layer plates on UV irradiation (Mineralight handlamp UVSL 25) prior to the development of the chromatogram. After development, most of the irradiated label is detected at the origin of the chromatogram in contrast to the nonirradiated control, which has completely migrated.

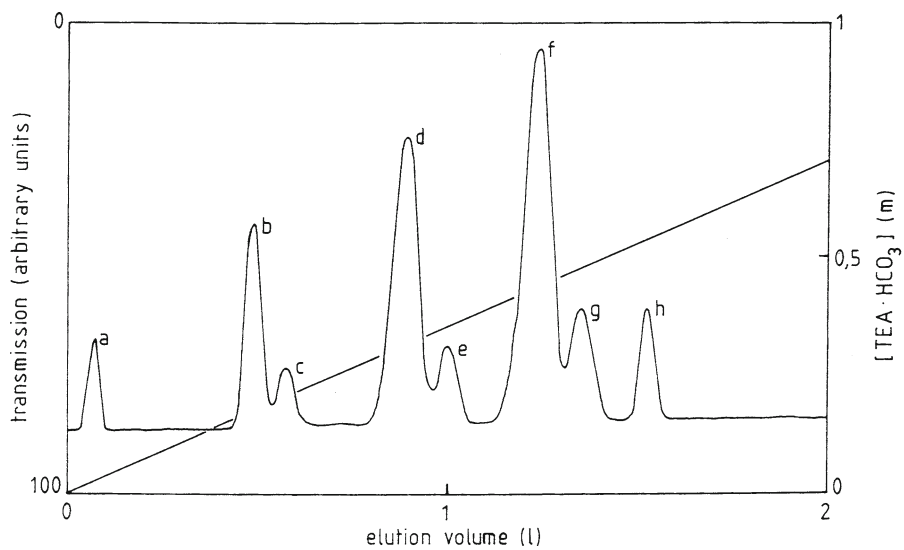


Fig. 5. Elution profile (anion exchange chromatography on DEAE-Sephadex A 25; elution buffer: linear gradient of 1,000 mL each of water and 0.7 M triethylammonium bicarbonate (pH 7.3), of the reaction products of 8-N_3dATP synthesis: front (**a**), 8-N_3dAMP (**b**), 8-BrdAMP (**c**), 8-N_3dADP (**d**), 8-BrdADP (**e**), 8-N_3dATP (**f**), 8-BrdATP (**g**), and probably a higher phosphorylated 8-azidoadenosine derivative (**h**).

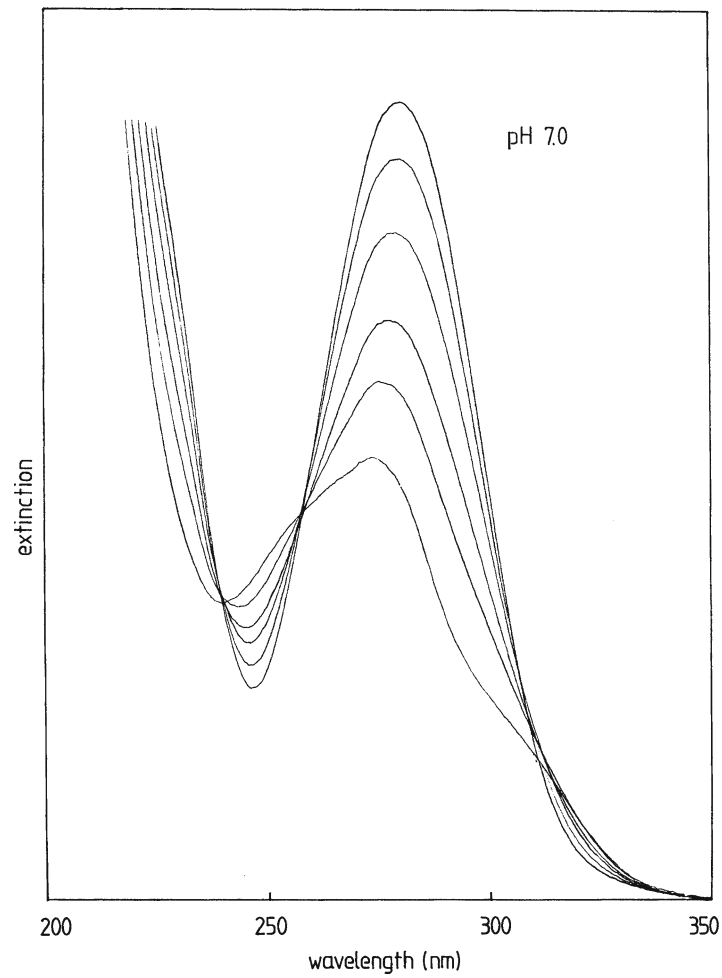


Fig. 6. Change of the optical absorption spectrum of 8-N_3dATP on UV irradiation in Tris–HCl buffer (0.01 M, pH 7.0, 20°C). The irradiation time between two subsequent absorption spectra was 2 min initially. It was increased up to 10 min toward the end of photolysis. The final spectrum was taken after 30 min of irradiation. During the photolysis the absorbance at 280 nm decreased; two new absorbance maxima at 248 and 305 nm are formed.

3.3. Preparation of Azido-Modified DNA

Azido-modified and [^{32}P]-labeled DNA are prepared by nick translation. For this purpose the detailed and exact composition of the reaction medium depends strongly on the size as well as on the amount of the DNA to be modified. The optimal ratio of DNA, DNase I, and DNA polymerase I (Kornberg enzyme) has to be tested out in preliminary experiments (*see* **Notes 11** and **12**).

Here we describe the well-tested reaction conditions for the modification of plasmid pBR322 (4,363 bp). The preparation of azido-modified pWH106 (4,970 bp) can be performed analogously.

1. Add 17.3 pmol of pBR322 to a mixture of 50 nmol of dGTP, 50 nmol of dTTP, 50 nmol of 8-N_3dATP, and 500 pmol of dCTP; prepare on ice.
2. Add 370 kBq of α-[^{32}P]-dCTP (110 TBq/mmol) and 20 pg of DNase I (freshly prepared out of a stock solution of 1 mg of DNAse I in 1 mL of 0.15 M NaCl/50% glycerol).
3. Adjust the reaction medium to an end concentration of 50 mM Tris–HCl of pH 7.2, 10 mM $MgSO_4$, and 50 mg/mL of bovine serum albumin (standard reaction volume: 100 µL).
4. Start the nick translation reaction by adding 100 U of DNA polymerase I from *E. coli* (*see* **Subheading 2.3.7**).
5. Incubate for 1 h at 15°C in the dark.
6. Stop the reaction by adding EDTA (final concentration: 20 mM).
7. Separate the excess deoxyribonucleotides not incorporated during the nick translation from photoreactive [^{32}P]-labeled pBR322 by gel filtration over Sephadex A-25 column using a 1-mL syringe.
8. Precipitate photoreactive pBR322 by adding twice the volume of cold ethanol and redissolve the residue in double-distilled water. Store the aqueous solution at –20°C in the dark.

Nonphotoreactive DNA (control) can be prepared analogously replacing the 8-N_3dATP by 50 nmol of dATP.

3.4. Photocrosslinking

1. Prepare 20–30 µL aqueous solutions containing the photoreactive DNA (0.5 pmol) and the protein (1–25 pmol) planned to be crosslinked (*see* **Notes 13** and **14**).
2. Incubate the reaction mixture for 10 min at 37°C in the dark.
3. Expose the sample to UV irradiation (*see* **Notes 15** and **16**). The irradiation times can be chosen in a range from 1 s to 60 min (*see* **Note 17**).
4. Keep the solutions in the dark before and after photolysis (*see* **Note 6**).

3.5. Analysis of DNA–Protein Adducts

Analysis of DNA–protein adducts can be done, for example, by polyacrylamide gel electrophoresis of the irradiated samples followed by autoradiography. SDS polyacrylamide gel electrophoresis should be performed immediately after photocrosslinking according to Laemmli *(28)* with some variations. After the addition of 20 mg/mL of bromophenol blue the samples are loaded onto an SDS polyacrylamide gel of 5% polyacrylamide (separating gel) with an overlay of 3.5% polyacrylamide (stacking gel) containing 1% of SDS, respectively. After the electrophoretic separation the gels are silver stained according to Adams and Sammons *(29)*,

dried, and exposed to X-ray films at –70°C. The autoradiograms are then developed. A quantitative determination of the DNA–protein adducts is possible by densitometric measurement of the autoradiogram *(30)*. **Figure 7** shows a typical result on photocrosslinking of *Eco*RI-digested plasmid pWH106 with a specific interacting protein (Tet repressor).

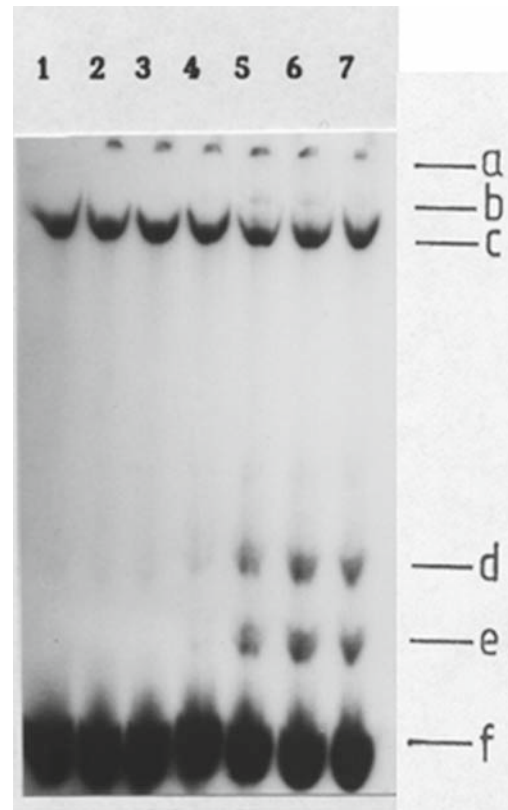


Fig. 7. Photocrosslinking of proteins to DNA (pWH106). Autoradiogram of a denaturing 5% SDS-polyacrylamide gel electrophoresis showing photocrosslinking of Tet repressor to azido-activated 187-bp and 3,848-bp fragments of pWH106 (radioactive labeled by ^{32}P). Each 187-bp fragment contains two *tet* operator sequences, the 3,848-bp fragment none. UV irradiation of azido-modified 187-bp fragment in the presence of Tet repressor results in a reduced migration of the 187-bp fragment because of covalent crosslinking of the DNA to one or two Tet repressor dimers. In each of *lanes 1–7,* 0.06 pmol pWH106 (cleaved by *Eco*RI) and 20 pmol Tet repressor were applied. *Lane 1*: photoactive fragments of pWH 106 without protein (30′ UV); lane 2: nonphotoactive fragments of pWH106 with Tet repressor (30′ UV); *lanes 3–7*: photoactive fragments of pWH106 with Tet repressor (0′, 1′, 4′, 10′, 30′ UV). Fractions: origin of sample loading (**a**); traces of 3,848-bp fragment covalently crosslinked (unspecifically) to Tet repressor (**b**); 3,848-bp fragment (no Tet repressor bound) (**c**); 187-bp fragment covalently crosslinked to two Tet repressor dimers (**d**); 187-bp fragment covalently crosslinked to one Tet repressor dimer (**e**); 187-bp fragment (no Tet repressor bound) (**f**).

Another possibility to detect the DNA–protein adducts is the application of the nitrocellulose filter binding assay according to Braun and Merrick *(31)*.

4. Notes

1. Experiments to synthesize α-[^{32}P] or U-[^{14}C]-labeled 8-N_3dATP by starting the synthesis with α-[^{32}P] or U-[^{14}C] dATP, respectively, failed. This is most probably because of the formation of bromine radicals induced by radioactive irradiation. These radicals could react unspecifically with the deoxyribonucleotide, suppressing the very specific electrophilic substitution of the hydrogen in position 8 of the adenine ring by the bromine ion.
2. Do not stop the reaction of dATP with bromine before the end of the given 6 h even if the absorption maximum is next to 262 nm after 1 or 2 h (otherwise a significant reduction of the yield of 8-BrdATP may occur; *see* **Subheading 3.1.1**).
3. 8-BrdATP is obtained as triethylammonium salt, which is soluble in dimethylformamide in contrast to the alkali salts of this nucleotide. This is advantageous for the following substitution of bromine by the azido group yielding 8-N_3dATP (*see* **Subheading 3.1.5**).
4. The exchange reaction of bromine by the azido group requires absolute dryness. However, the formation of 8-N_3dAMP and 8-N_3dADP is usually observed, due to a limited hydrolytic cleavage of 8-N_3dATP (*see* **Fig. 5**; *see also* **Subheadings 3.1.5** and **3.1.6**).
5. Besides the three azidoadenine deoxyribonucleotides, minor amounts of unreacted 8-bromoadenine deoxyribonucleotides are eluted as well (*see* **Fig. 5**; *see also* **Subheading 3.1.6**).
6. Because of the photoreactivity of azido compounds, samples containing 8-N_3dATP should always be kept in the dark if possible. However, short exposure of azido compounds to normal day light in our laboratory never falsified the results obtained (*see* **Subheading 3.1.7**).
7. 8-N_3dATP can be stored frozen at –20°C in aqueous solution, pH 7.0, in the dark for at least 2 years without significant loss of photoreactivity as demonstrated by subsequent photocrosslinking experiments (*see* **Subheading 3.1.7**).
8. Exclude dithiothreitol from any buffers or other solutions that contain 8-N_3dATP (*see* **Subheading 3.1.7**). It is well known that dithiothreitol reduces azido groups to the

corresponding amines *(32)*. In addition, the UV absorbance of dithiothreitol resembles that of 8-N_3dATP because of the formation of disulfide bonds by oxidation on storage in aqueous solution. This results in a less efficient photocrosslinking rate by the UV irradiation.

9. The UV absorption of 8-N_3dATP shows a maximum at 280 nm. The absorbance at 280 nm increases with decreasing pH value (*see* **Subheading 3.2.2**). A second absorption maximum at 219 nm shifts to 204 nm in acidic solution. Both effects are a result of the protonation at N^1 of the purine ring *(33)*. The UV absorption spectrum of 8-N_3dATP resembles that of 8-N_3ATP *(27)*.
10. Take into account that the photolysis of 8-N_3dATP (to test its photoreactivity) is pH dependent (*see* **Subheading 3.2.3**). Exhaustive irradiation in neutral solution yields new absorption maxima at 248 and 305 nm, whereas in acidic or basic solution the destruction of the purine ring is observed as indicated by the disappearance of the absorbance between 240 and 310 nm (data not shown).
11. 8-N_3dATP should prefer the *syn* conformation (**Fig. 8**) by analogy with 8-azidoadenine nucleotides because of the bulky substituent in position 8 of the purine ring *(34)*. This seems to be contradicted by our results indicating that DNA polymerase I *(E. coli)* accepts 8-N_3dATP for a successful nick translation. It has been suggested that the enzyme only interacts with 2-deoxynucleoside triphosphates in the *anti* conformation *(35)*. This discrepancy may be explained in two ways: First, the steric requirements for the binding of 8-N_3dATP by DNA polymerase I are less restrictive than assumed *(36)*, or second, 8-N_3dATP interacts in the *anti* conformation with the binding site of the enzyme. This could be demonstrated for the interaction of 8-N_3ATP with the F_1ATPase from mitochondria *(37)*.

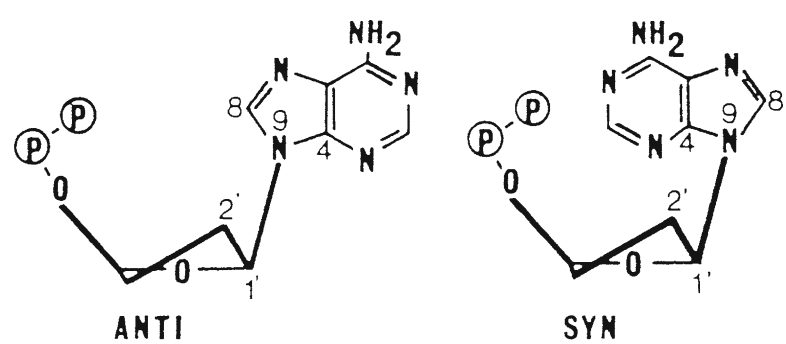


Fig. 8. Conformation of adenine nucleotides *(23)*.

12. The preparation of azido-modified and [^{32}P]-labeled DNA by nick translation is critically dependent on the ratio of DNA, DNase I, and DNA polymerase I in the reaction medium (*see* **Subheading 3.3**). Higher concentrations of DNase I on the one hand result in a very efficient incorporation rate of the azido-modified and radioactive labeled deoxynucleotides, but on the other hand, the degradation of the DNA probes by DNase I has to be evaluated. Application of too small amounts of DNase I results in inefficient incorporation of the photoactivatable deoxynucleotides and in an insufficient photocrosslinking to the interacting proteins during subsequent irradiation of the azido-modified DNA.
13. Tris–HCl buffer, 50 mM, pH 7.2, can be used instead of double-distilled water without any significant effect on the photocrosslinking efficiency (*see* **Subheading 3.4.1**).
14. The amount of protein planned to be photocrosslinked to photoreactive DNA can be varied over a wide range (*see* **Subheading 3.4.1**). Too high an excess of proteins, however, should be avoided because the absorbance maximum of proteins at 280 nm will lead to inefficient crosslinking rates.
15. One way to expose the samples to UV light is to deposit the probes (typically 30–50 μL) in plastic wells (normally used for RIA or ELISA tests). The UV lamp is positioned directly above the samples; thus, more than one probe can be irradiated simultaneously (*see* **Subheading 3.4.3**).
16. The emitted light of the UV lamp used for photocrosslinking should not contain light of shorter wavelengths than 300 nm because of the possibility of photodamaging DNA or protein. For example the Mineralight handlamp UVSL 25 (long wave) emits UV light of mainly 366 nm. The small portion of UV light of wavelengths between 300 and 320 nm emitted allows the photoactivation of the azido group without any significant photodamage of DNA or protein (*see* **Subheading 3.4.3**).
17. UV irradiation times for photocrosslinking can be chosen over a wide range (*see* **Subheading 3.4.3**). Optimal UV fluence rates (flux per unit area) have to be tested out. In our experiments (using the Mineralight handlamp UVSL 25 fixed in a position resulting in a fluence rate of 4 J/m^2/s at the position of the sample) first slight amounts of DNA–protein adducts are detected after irradiation times of 10–30 s; irradiation periods longer than 15–20 min do not improve the yield of photocrosslink products.

Acknowledgments

The authors thank Dr. Marianne Schüz (Mainz) for editing the manuscript. This work was supported by the Bundesministerium für Forschung und Technologie (07QV8942) and by the Deutsche Forschungsgemeinschaft (Scha 344/1–3).

References

1. Bayley, H. and Knowles, J. R. (1977). Photoaffinity labeling. *Methods Enzymol.* **46**, 69–114.
2. Bayley, H. (1983). Photogenerated reagents in biochemistry and molecular biology. In *Laboratory Techniques in Biochemistry*, vol.12 (Work, T.S. and Burdon, R.H., eds.) Elsevier, Amsterdam.
3. Schäfer, H. J. (1987). Photoaffinity labeling and photoaffinity crosslinking of enzymes. In *Chemical Modifications of Enzymes, Active Site Studies* (Eyzaguirre, J., ed.), Ellis Horwood, Chichester, UK, pp. 45–62.
4. Smith, K. C. (1962). Dose dependent decrease in extractability of DNA from bacteria (by UV-light). *Biochem. Biophys. Res. Commun.* **8**, 157–163.
5. Shetlar, M. D. (1980). Cross-linking of proteins to nucleic acids by UV-light. *Photochem. Photobiol. Rev.* **5**, 105–197.
6. Welsh, J. and Cantor, C. R. (1984). Protein-DNA crosslinking. *Trends Biochem. Sci.* **9**, 505–508.
7. Ekert, B., Giocanti, N., and Sabattier, R. (1986). Study of several factors in RNA-protein crosslink formation induced by ionizing radiations within 70 S ribosomes of *E. coli* MRE 600. *Int. J. Radiat. Biol.* **50**, 507–525.
8. Lesko, S. A., Drocourt, J. L., and Yang, S. U. (1982). DNA-protein- and DNA interstrand crosslinks induced in isolated chromatin by H_2O_2 and Fe-EDTA-chelates. *Biochemistry.* **21**, 5010–5015.
9. Wedrychowsky, A., Ward, W. S., Schmidt, W. N., and Hnilica, L. S. (1985). Chromium-induced cross-linking of nuclear proteins and DNA. *J. Biol. Chem.* **260**, 7150–7155.
10. Summerfield, F. W. and Tappel, A. L. (1984). Cross-linking of DNA in liver and testis of rats fed 1,3-propanediol. *Chem. Biol. Interact.* **50**, 87–96.
11. Dose, K., Bieger-Dose, A., Martens, K.-D., Meffert, R., Nawroth, T., Risi, S., Steinborn, A., and Vogel, M. (1987). Survival under space vacuum–biochemical aspects. *Proc. 3rd Eur. Symp. Life Sci. Res. in Space (ESA SP-271)* 193–195.
12. Lin, S. Y. and Riggs, A. D. (1974). Photochemical attachment of lac repressor to bromodeoxyuridine-substituted lac operator by UV radiation. *Proc. Natl. Acad. Sci. U S A* **71**, 947–951.
13. Evans, R. K., Johnson, J. D., and Haley, B. E. (1986). 5-Azido-2′-deoxyuridine-5′-triphosphate: a photoaffinity labeling reagent and tool for the enzymatic synthesis of photoactive DNA. *Proc. Natl. Acad. Sci. U S A* **83**, 5382–5386.
14. Bartholomew, B., Kassavetis, G. A., Braun, B. R., and Geiduschek, E. P. (1990). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *EMBO J.* **9**, 2197–2205.
15. Lee, D. K., Evans, R. K., Blanco, J., Gottesfeld, J., and Johnson, J. D. (1991). Contacts between 5 S DNA and Xenopus TFIIIA identified using 5-azido-2′-deoxyuridine-substituted DNA. *J. Biol. Chem.* **266**, 16478–16484.
16. Blatter, E. E., Ebright, Y. W., and Ebright, R. H. (1992). Identification of an amino acid-base contact in the GCN4-DNA complex by bromouracil-mediated photocrosslinking. *Nature* **83**, 650–652.
17. Willis, M. C., Hicke, B. J., Uhlenbeck, O. C., Cech, T. R., and Koch, T. H. (1993). Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. *Science* **262**, 1255–1257.
18. Hicke, B. J., Willis, M. C., Koch, T. H., and Cech, T. R. (1994). Telomeric protein-DNA point contacts identified by photo-cross-linking using 5-bromodeoxyuridine. *Biochemistry* **33**, 3364–3373.
19. Pingoud, V., Geyer, H., Geyer, R., Kubareva, E., Bujnicki, J. M., and Pingoud, A. (2005).

Identification of base-specific contacts in protein-DNA complexes by photocrosslinking and mass spectrometry: a case study using the restriction endonuclease *Ssoll*. *Mol. Biosyst.* **1**, 135–141.

20. Bartholomew, B., Braun, B. R., Kassavetis, G. A., and Geiduschek, E. P. (1994). Probing close DNA contacts of RNA-polymerase III transcription complexes with the photoactive nucleoside 4-thiodeoxythymidine. *J. Biol. Chem.* **269**, 18090–18095.
21. Dezhurov, S. V., Khodyreva, S. N., Plekhanova, E. S., and Lavrik, O. I. (2005). A new highly efficient photoreactive analogue of dCTP. Synthesis, characterization, and application in photoaffinity modification of DNA binding proteins. *Bioconjug. Chem.* **16**, 215–222.
22. Maltseva, E. A., Rechkunova, N. I., Gillet, L. C., Petruseva, I. O., Schärer, O. D., and Lavrik, O. I. (2007). Crosslinking of the NER damage recognition proteinsXPC-HR23B, XPA and RPA to photoreactive probes that mimic DNA damages. *Biochim. Biophys. Acta* **1770**, 781–789.
23. Zofall, M. and Bartholomew, B. (2000). Two novel dATP analogs for DNA photoaffinity labeling. *Nucleic Acids Res.* **28**, 4382–4390.
24. Meffert, R. and Dose, K. (1988). UV-induced cross-linking of proteins to plasmid pBR322 containing 8-azidoadenine 2′-deoxyribonucleotides. *FEBS Lett.* **239**, 190–194.
25. Meffert, R., Rathgeber, G., Schäfer, H.-J., and Dose, K. (1990). UV-induced cross-linking of Tet repressor to DNA containing tet operator sequences and 8-azidoadenines. *Nucleic Acids Res.* **18**, 6633–6636.
26. Bartholomew, B., Tinker, R. L., Kassavetis, G. A., and Geiduschek, E. P. (1995). Photochemical cross-linking assay for DNA tracking by replication proteins. *Methods Enzymol.* **262**, 476–494.
27. Schäfer, H. J., Scheurich, P., and Dose, K. (1978). Eine einfache Darstellung von 8-N_3ATP: ein Agens zur Photoaffinitätsmarkierung von ATP-bindenden Proteinen. *Liebigs Ann. Chem.* **1978**, 1749–1753.
28. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680.
29. Adams, L. D. and Sammons, D. W. (1981). A unique silver staining procedure for color characterization of polypeptides. *Electrophoresis* **2**, 155–165.
30. Westermeier, R., Schickle, H., Thesseling, G., and Walter, W. W. (1988). Densitometrie von Gelelektrophoresen. *GIT Labor-Medizin* **4/88**, 194–202.
31. Braun, A. and Merrick, B. (1975). Properties of UV-light-mediated binding of BSA to DNA. *Photochem. Photobiol.* **21**, 243–247.
32. Staros, J. V., Bayley, H., Standring, D. N., and Knowles, J. R. (1978). Reduction of aryl azides by thiols: implications for the use of photoaffinity reagents. *Biochem. Biophys. Res. Commun.* **80**, 568–572.
33. Koberstein, R., Cobianchi, L., and Sund, H. (1976). Interaction of the photoaffinity label 8-azido-ADP with glutamate dehydrogenase. *FEBS Lett.* **64**, 176–180.
34. Vignais, P. V. and Lunardi, J. (1985). Chemical probes of the mitochondrial ATP synthesis and translocation. *Ann. Rev. Biochem.* **54**, 977–1014.
35. Czarnecki, J. J. (1978). Phd Thesis, University of Wyoming, Laramie, WY.
36. Englund, P. T., Kelly, R. B., and Kornberg, A. (1969). Enzymatic synthesis of DNA: binding of DNA to DNA polymerase. *J. Biol. Chem.* **244**, 3045–3052.
37. Garin, J., Vignais, P. V., Gronenborn, A. M., Clore, G. M., Gao, Z., and Bäuerlein, E. (1988). ^{1}H-NMR studies on nucleotide binding to the catalytic sites of bovine mitochondrial F_1-ATPase. *FEBS Lett.* **242**, 178–182.

Chapter 25

Static and Kinetic Site-Specific Protein–DNA Photocrosslinking: Analysis of Bacterial Transcription Initiation Complexes

Nikolai Naryshkin, Sergei Druzhinin, Andrei Revyakin, Younggyu Kim, Vladimir Mekler, and Richard H. Ebright

Summary

Static site-specific protein–DNA photocrosslinking permits identification of protein–DNA interactions within multiprotein–DNA complexes. Kinetic site-specific protein–DNA photocrosslinking – involving rapid-quench-flow mixing and pulsed-laser irradiation – permits elucidation of pathways and kinetics of formation of protein–DNA interactions within multiprotein–DNA complexes. We present detailed protocols for application of static and kinetic site-specific protein–DNA photocrosslinking to bacterial transcription initiation complexes.

Key words: Structure, Kinetics, Protein–DNA interaction, Transcription initiation, RNA polymerase, Promoter, Open complex, Phosphate, Phosphorothioate, Phenyl-azide photoactivatible crosslinking agent, Quench-flow rapid mixing, Pulsed-laser UV irradiations.

1. Introduction

1.1. Static Site-Specific Protein–DNA Photocrosslinking

In published work, we have developed a procedure for site-specific protein–DNA photocrosslinking to define positions of proteins relative to DNA in protein–DNA and multiprotein–DNA complexes *(1–6)*. The procedure has four parts (**Fig. 1**):

(1) Chemical *(7–9)* and enzymatic *(10)* reactions are used to prepare a DNA fragment containing a photoactivatible crosslinking agent and an adjacent radiolabel incorporated at a single, defined DNA phosphate (with a 9.7-Å linker between the photoreactive atom of the crosslinking agent and the

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_25

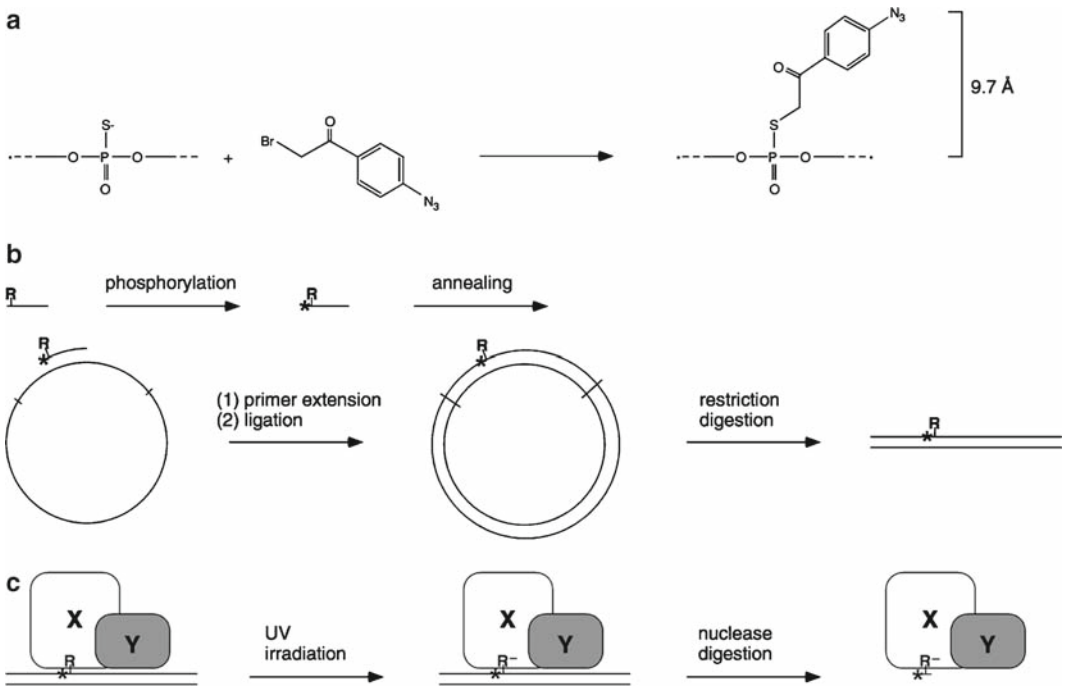


Fig. 1. Site-specific protein–DNA photocrosslinking *(1–6)*. (**a**, **b**) Chemical and enzymatic reactions are used to prepare a full-length-promoter DNA fragment with a phenyl-azide photoactivatible crosslinking agent (R) and an adjacent radioactive phosphorus (*) incorporated at a single, defined site. Based on the chemistry of incorporation, the maximum distance between the site of incorporation and the photoreactive atom is 9.7 Å; the maximum distance between the site of incorporation and a crosslinked atom is ~11 Å. (**c**) UV irradiation of the derivatized protein–DNA complex initiates crosslinking. Nuclease digestion eliminates uncrosslinked DNA and converts crosslinked, radiolabeled DNA to a crosslinked, radiolabeled 3–5 nucleotide "tag".

phosphorus atom of the phosphate, and with an ~11 Å maximum "reach" between potential crosslinking targets and the phosphorus atom of the phosphate).

(2) The multiprotein–DNA complex of interest is formed using the site-specifically derivatized DNA fragment, and the multiprotein–DNA complex is UV irradiated, initiating covalent crosslinking with proteins in direct physical proximity to the photoactivatible crosslinking agent.

(3) Extensive nuclease digestion is performed, eliminating uncrosslinked DNA and converting crosslinked DNA to a crosslinked, radiolabeled 3–5 nucleotide "tag."

(4) The "tagged" proteins are identified.

The procedure is performed in systematic fashion, with preparation and analysis of at least ten derivatized DNA fragments, each having the photoactivatible crosslinking agent incorporated at a single, defined DNA phosphate (typically each second DNA phosphate – each 12 Å – on each DNA strand spanning the region of interest *(1–6, 11–15)*).

The results of the procedure define the translational positions of proteins relative to the DNA sequence. Plotted on a three-dimensional representation of a DNA helix, the results also define the rotational orientations of proteins relative to the DNA helix axis, and the groove orientations of proteins relative to the DNA major and minor grooves *(1–5, 11–13)*.

The procedure has been validated in experiments with three multiprotein–DNA complexes for which crystallographic structures are available: i.e., the TBP-DNA complex, the TBP-TFIIA-DNA complex, and the TBP-TFIIB-DNA complex *(1, 16–20)*. In each case, there was a one-for-one correspondence between sites at which strong crosslinking was observed and sites that in the crystallographic structure were within 11 Å of crosslinked proteins *(1, 16–20)*. The procedure also has been applied to multiprotein–DNA complexes for which crystallographic structures are not available *(1–6, 11–15)*, including bacterial transcription complexes containing 6–8 distinct polypeptides and having molecular masses of 450–400 kDa *(5)*, archaeal transcription complexes containing 14 distinct polypeptides and having molecular masses of 400 kDa *(12–14)*, and eukaryotic transcription complexes containing 16–27 distinct polypeptides and having molecular masses of 800–1,700 kDa *(2, 4, 15)*.

The procedure is related to a procedure developed by Geiduschek and coworkers *(21–24)*, *see also (25–30)*, but offers important advantages. First, since the photoactivatible crosslinking agent is incorporated into DNA chemically, it can be incorporated at a single, defined site. (In the procedure of Geiduschek and coworkers, this is true only at certain DNA sequences.) Second, since the photoactivatible crosslinking agent is incorporated on the DNA phosphate backbone, it can be incorporated at any nucleotide: A, T, G, or C. Third, since the photoactivatible crosslinking agent is incorporated on the DNA phosphate backbone, it probes interactions both in the DNA minor groove and in the DNA minor groove.

1.2. Kinetic Site-Specific Protein–DNA Photocrosslinking

The procedure for site-specific protein–DNA photocrosslinking summarized in **Subheading 1.1** can be combined with quench-flow rapid mixing and pulsed-laser flash UV irradiation in order to permit analysis of kinetics of formation or breakage of protein–DNA interactions within protein–DNA and multiprotein–DNA complexes ("kinetic site-specific protein–DNA photocrosslinking"; S.D. and R.H.E, in preparation). Kinetic site-specific protein–DNA photocrosslinking involves four main steps (**Figs. 1** and **2**):

(1) A DNA fragment containing a phenyl-azide photoactivatible crosslinking agent and an adjacent radiolabel incorporated at a single, defined DNA phosphate (prepared as summarized in **Subheading 1.1**) is mixed with the protein(s) of interest at time = 0 using a quench-flow rapid mixer.

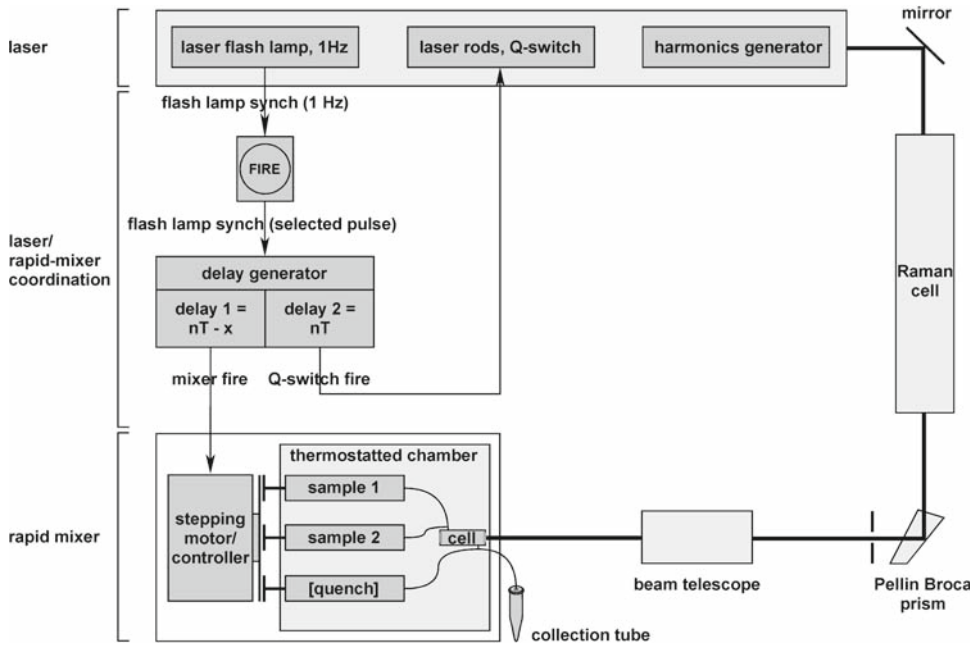


Fig. 2. Apparatus for kinetic site-specific protein–DNA photocrosslinking (S.D. and R.H.E., in preparation). The apparatus consists of a nanosecond pulsed Nd-YAG laser with integral third harmonic generator; a quench-flow rapid mixer; electronics for coordination of the laser and rapid mixer; and a Raman cell, Pellin-Broca prism, 355-nm laser optics (first mirror), and 309-nm laser optics (all other optics). To minimize variations in energy output, the laser flash lamp is allowed to operate continuously, and firing of the rapid mixer and firing of the laser Q-switch are coordinated with the timing of the laser flash lamp. Upon selection of a synch pulse from the laser flash lamp, the rapid mixer is fired at time = nT–x (where *n* is an integer, *T* is the flash lamp period, and *x* is the desired reaction time) resulting in mixing, the Q-switch is fired at time = *nT* resulting in UV irradiation, and the rapid mixer is fired again at time = *nT* + *y* (where *y* is the desired post-UV irradiation, prequenching reaction time; <1 ms in this work) resulting in quenching.

(2) The sample is UV-irradiated at time = x using a single pulse from a pulsed laser, initiating covalent crosslinking with protein(s) in direct physical proximity to the phenyl-azide photoactivatible crosslinking agent.

(3) The sample is mixed with a quench solution at time = x + ~1 ms, terminating covalent crosslinking (by inactivating photogenerated reactive species and by dissociating complexes).

(4) Crosslinked protein(s) are identified (performed as summarized in **Subheading 1.1**), and yield(s) of crosslinked protein(s) are quantified.

The procedure provides a "snapshot" or "motion-picture frame" of protein–DNA interactions at one position in DNA at one point in time. By performing a series of experiments with UV irradiation at times x_1, x_2, …x_n, one obtains a series of "motion-picture frames" and thus obtains a "cinematographic" record of protein–DNA interactions at one position in DNA. By performing such experiments systematically, using a set of DNA fragments derivatized at different single positions in DNA, one is able to

define the full pathway and kinetics for formation or breakage of protein–DNA interactions within a protein–DNA or multiprotein–DNA complex.

1.3. Bacterial Transcription Initiation Complexes

Escherichia coli RNA polymerase holoenzyme (RNAP) consists of two copies of an α subunit (37 kDa), one copy of a β subunit (150 kDa), one copy of a β′ subunit (160 kDa), one copy of an ω subunit (10 kDa), and one copy of a σ subunit (70 kDa for the principal σ subunit species, σ^{70}) *(31, 32)*. RNAP is a molecular machine that carries out a complex series of reactions in transcription initiation *(31, 32)*. Formation of a catalytically competent transcription initiation complex involves at least three steps: (a) RNAP binds to promoter DNA to yield an RNAP-promoter closed complex; (b) RNAP clamps tightly onto promoter DNA, to yield an RNAP-promoter intermediate complex; and (c) RNAP unwinds ~14 bp of promoter DNA surrounding the transcription start, rendering accessible the genetic information in the template strand of DNA, to yield an RNAP-promoter open complex.

In published work, we have used static site-specific protein–DNA photocrosslinking to define the structural organization of the RNAP-promoter open complex *(5, 32)*. We constructed a set of 120 derivatized DNA fragments, each containing a photoactivatible crosslinking agent incorporated at a single, defined position of the *lacPUV5* promoter (positions –95 to +25 relative to the transcription start site). For each derivatized DNA fragment, we formed the RNAP-promoter open complex, isolated the complex using nondenaturing polyacrylamide gel electrophoresis, UV irradiated the complex *in situ* – in the gel matrix – and identified crosslinked polypeptides. We performed experiments both with wild-type RNAP and with RNAP derivatives having discontinuous β and β′ subunits ("split-β RNAP" and "split-β′ RNAP"; reconstituted *in vitro* from recombinant α, recombinant σ^{70}, and sets of recombinant fragments of β and β′; *see (33, 34)*). Use of split-β and split-β′ RNAP permitted unambiguous assignment of crosslinks to β and β′ (which were not well resolved in SDS-polyacrylamide gel electrophoresis) and permitted rapid, immediate mapping of crosslinks to segments of β and β′ (e.g., N-terminal segment, central segment, or C-terminal segment) (**Fig. 3**).

In current work, we are using kinetic site-specific protein–DNA photocrosslinking to define the pathway and kinetics of formation of the RNAP-promoter open complex. We are analyzing a set of more than 30 derivatized DNA fragments, each containing a photoactivatible crosslinking agent incorporated at a single, defined position of the *lacPUV5* promoter (with the sites being chosen to report key RNAP-promoter interactions spanning the RNAP-promoter interface).

In this chapter, we present protocols for use of static site-specific protein–DNA photocrosslinking to define the structural

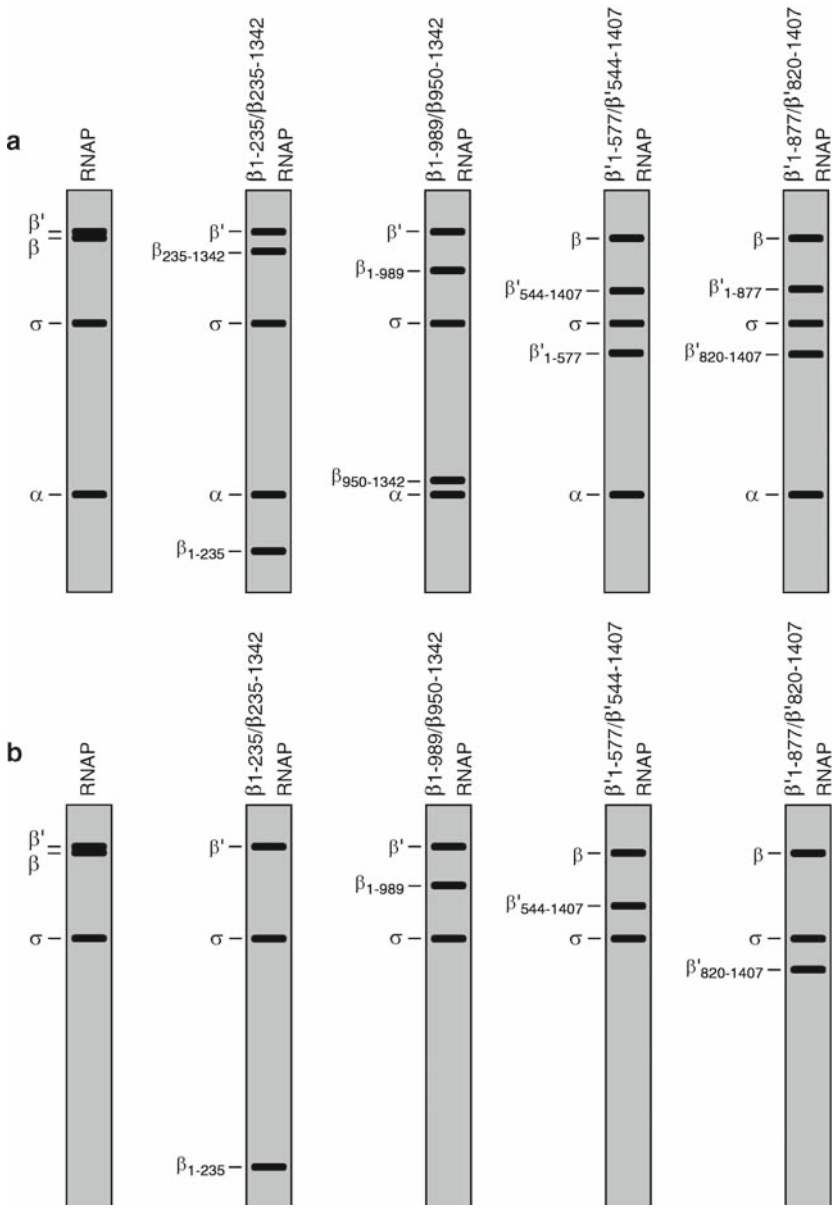


Fig. 3. Use of split-subunit RNAP derivatives permits unambiguous assignment of crosslinks to RNAP subunits and permits rapid mapping of crosslinks to segments of RNAP subunits (5), *see (33, 34)*. (**a**) Subunit compositions of RNAP, two split-β RNAP derivatives, and two split-β′ RNAP derivatives (idealized Coomassie-stained SDS-PAGE gels). (**b**) Results of site-specific protein–DNA photocrosslinking experiments using the RNAP derivatives of panel A and a DNA fragment derivatized at a site close to or in contact with residues 1–235 of β, residues 820–1,407 of β′, and σ^{70} in the RNAP-promoter complex (idealized autoradiographs of SDS-PAGE gels).

organization of the RNAP-promoter complex and for use of kinetic site-specific protein–DNA photocrosslinking to define the pathway and kinetics of formation of the RNAP-promoter open complex. In addition, we present support protocols for preparation of wild-type RNAP, split-β RNAP, and split-β′ RNAP.

2. Materials

2.1. Preparation of Derivatized DNA Fragment: Chemical Reactions

1. Azidophenacyl bromide (Sigma)
2. Tetraethylthiuram disulfide/acetonitrile (Applied Biosystems)
3. dA-CPG, dC-CPG, dG-CPG, T-CPG (1 μmol, 500 Å) (Applied Biosystems)
4. dA, dC, dG, T β-cyanoethylphosphoramidites (Applied Biosystems)
5. Reagent kit for oligodeoxyribonucleotide synthesis (0.02 M iodine) (Applied Biosystems)
6. Denaturing loading buffer (0.3% bromophenol blue, 0.3% xylene cyanol, 12 mM EDTA, in formamide)
7. 0.5× TBE (45 mM Tris-borate of pH 8.3, 1 mM EDTA)
8. TE (10 mM Tris–HCl of pH 7.6, 1 mM EDTA)
9. 50 mM triethylammonium acetate, pH 7.0 (Prime Synthesis)
10. 1 M potassium phosphate, pH 7.0
11. 3 M sodium acetate, pH 5.2
12. 100% ethanol (store at –20°C)
13. 70% ethanol (store at –20°C)
14. Dichloromethane (anhydrous) (Applied Biosystems)
15. Acetonitrile (anhydrous) (Applied Biosystems)
16. Acetonitrile (HPLC grade) (Fisher)
17. Formamide (Sigma)
18. 12% polyacrylamide (29:1 acrylamide:bisacrylamide), 8 M urea, 0.5× TBE slab gel (10 × 7 × 0.075 cm)
19. Oligonucleotide purification cartridge (OPC) (Applied Biosystems)
20. LiChrospher 100 RP-18 reversed-phase HPLC column (5 μm) (Merck)
21. Autoradiography-intensifying screen (Sigma)
22. 254-nm germicidal lamp
23. ABI392 DNA/RNA synthesizer (Applied Biosystems)
24. Varian 5000 HPLC (Varian)
25. L-3000 diode-array HPLC UV detector (Hitachi)
26. Speedvac evaporator (Thermo Scientific)

2.2. Preparation of Derivatized DNA Fragment: Enzymatic Reactions

1. Derivatized oligodeoxyribonucleotide (**Subheading 3.1**)
2. M13mp2*(ICAP-UV5)* or M13mp2*(ICAP-UV5)*-rev ssDNA (*see* **Notes 1 and 2**)

3. T4 polynucleotide kinase (10 U/μl) (New England Biolabs, cat #M0201S)
4. T4 DNA polymerase (3 U/μl) (New England Biolabs, cat #M0203S)
5. T4 DNA ligase (5 U/μl) (Roche Applied Science, cat #799009)
6. HaeIII (40 U/μl)(Roche Applied Science, cat #1336029)
7. PvuII (40 U/μl)(Roche Applied Science, cat #899216)
8. [γ^{32}P]-ATP (10 mCi/mL, 6,000 Ci/mmol) (Perkin Elmer)
9. 100 mM ATP (GE Life Sciences)
10. 100 mM dNTPs (GE Life Sciences)
11. Downstream primer (5′-CGGTGCGGGCCTCTTCGCTATTAC-3′)
12. 10× phosphorylation buffer (500 mM Tris–HCl of pH 7.6, 100 mM $MgCl_2$, 15 mM β-mercaptoethanol)
13. 10× annealing buffer (400 mM Tris–HCl of pH 7.9, 500 mM NaCl, 100 mM $MgCl_2$)
14. 10× digestion buffer (100 mM Tris–HCl of pH 7.9, 500 mM NaCl, 100 mM $MgCl_2$) (*see* **Note 3**)
15. Elution buffer (0.5 M ammonium acetate, 10 mM magnesium acetate of pH 7.5, 1 mM EDTA)
16. Denaturing loading buffer (0.3% bromophenol blue, 0.3% xylene cyanol, 12 mM EDTA, in formamide)
17. Nondenaturing loading buffer (0.3% bromophenol blue, 0.3% xylene cyanol, 30% glycerol, in water)
18. 0.5× TBE (45 mM Tris-borate of pH 8.3, 1 mM EDTA)
19. TE (10 mM Tris–HCl of pH 8.0, 1 mM EDTA)
20. Low-EDTA TE (10 mM Tris–HCl of pH 8.0, 0.1 mM EDTA)
21. 0.5 M EDTA, pH 8.0
22. 10% SDS
23. 100% ethanol (store at −20 °C)
24. 70% ethanol (store at −20°C)
25. 12% polyacrylamide (29:1 acrylamide:bisacrylamide), 8 M urea, 0.5× TBE slab gel (10 × 7 × 0.075 cm)
26. 7.5% polyacrylamide (29:1 acrylamide:bisacrylamide), 0.5× TBE slab gel (10 × 7 × 0.15 cm)
27. CHROMA SPIN + TE-10 spin column (Clontech)
28. CHROMA SPIN + TE-100 spin column (Clontech)
29. Spin-X centrifuge filter (0.22 μm, cellulose acetate) (Fisher)
30. PicoGreen dsDNA quantitation kit (Invitrogen, cat #P7589)

31. Disposable scalpels (VWR)
32. Autoradiography markers (Stratagene)
33. Light box (VWR)
34. Speedvac evaporator (Savant)

2.3. Preparation of RNAP and RNAP Derivatives

1. *E. coli* strain XL1-blue (Stratagene, cat #200249)
2. *E. coli* strain BL21(DE3) pLysS (Novagen, cat #69388-3)
3. Plasmids encoding RNAP subunits (*see* **Table 1**)
4. Plasmids encoding fragments of RNAP subunits (*see* **Table 2**)
5. LB broth (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl; autoclave- sterilized)
6. TYE agar containing 200 μg/mL ampicillin and 35 μg/mL chloramphenicol (10 g/L tryptone, 5 g/L yeast extract,

Table 1
Plasmids encoding RNAP subunits

Plasmid	Relevant characteristics	Source
pHTT7f1-NHα	Ap^R; ori-pBR322; ori-f1; ϕ*10P-rpoA(H6,Nter)*[a]	*(35, 36)*
pMKSe2	Ap^R; ori-pBR322; *lacP-rpoB*	*(37)*
pT7β′	Ap^R; ori-pBR322; ϕ*10P-rpoC*	*(38)*
pHTT7f1-	Ap^R; ori-pBR322; ori-f1; ϕ*10P-rpoD*	*(35, 36)*
pT7ω	Ap^R; ori-pBR322; ori-f1; ϕ*10P-rpoZ*	*(39)*

[a] *rpoA(H6,Nter)* is a derivative of *rpoA* having a non-native hexahistidine coding sequence immediately after the *rpoA* start codon

Table 2
Plasmids encoding fragments of RNAP subunits

Plasmid	Relevant characteristics	Source
$pβ_{1–235}$	Ap^R Km^R; ori-pBR322; *lacP-rpoB(1–235)*	*(33)*
$pβ_{235–1,342}$	Ap^R; ori-pBR322; *lacP-rpoB(235–1,342)*	*(33)*
$pβ_{1–989}$	Ap^R; ori-pBR322; *lacP-rpoB(1–989)*	*(33)*
$pβ_{950–1,342}$	Ap^R; ori-pBR322; ϕ*10P-rpoB(950–1,342)*	*(33)*
$pβ'_{1–577}$	Ap^R; ori-pBR322; ori-f1; *lacP-ϕ10P-rpoC(1–577)*	*(34)*
$pβ'_{544–1,407}$	Ap^R; ori-pBR322; ori-f1; *lacP-ϕ10P-rpoC(544–1,407)*	*(34)*
$pβ'_{1–877}$	Ap^R; ori-pBR322; ori-f1; *lacP-ϕ10P-rpoC(1–877)*	*(34)*
$pβ'_{820–1,407}$	Km^R; ori-pBR322; ori-f1; ϕ*10P-rpoC(820–1,407)*	*(34)*

8 g/L NaCl, 15 g/L agar; autoclave- sterilized without antibiotics; supplemented with antibiotics after cooling to 55°C; poured into sterile 100 × 15 mm Petri plates at ~25 mL/plate)

7. TYE agar containing 200 μg/mL ampicillin and 20 μg/mL tetracycline (10 g/L tryptone, 5 g/L yeast extract, 8 g/L NaCl, 15 g/L agar; autoclave- sterilized without antibiotics; supplemented with antibiotics after cooling to 55°C; poured into sterile 100 × 15 mm Petri plates at ~25 mL/plate)
8. TYE agar containing 40 μg/mL kanamycin and 35 μg/mL chloramphenicol (10 g/L tryptone, 5 g/L yeast extract, 8 g/L NaCl, 15 g/L agar; autoclave- sterilized without antibiotics; supplemented with antibiotics after cooling to 55°C; poured into sterile 100 × 15 mm Petri plates at ~25 mL/plate)
9. TYE agar containing 40 μg/mL kanamycin and 20 μg/mL tetracycline (10 g/L tryptone, 5 g/L yeast extract, 8 g/L NaCl, 15 g/L agar; autoclave-sterilized without antibiotics; supplemented with antibiotics after cooling to 55°C; poured into sterile 100 × 15 mm Petri plates at ~25 mL/plate)
10. 100 mg/mL ampicillin (filter- sterilized) (Sigma)
11. 35 mg/mL chloramphenicol in ethanol (filter- sterilized) (Sigma)
12. 40 mg/mL kanamycin (filter- sterilized) (Sigma)
13. 20 mg/mL tetracycline in methanol (filter- sterilized) (Sigma)
14. 1 M IPTG (filter- sterilized) (Roche Applied Science)
15. Buffer A (20 mM Tris–HCl of pH 7.9, 500 mM NaCl, 5 mM imidazole)
16. Buffer B (20 mM Tris–HCl of pH 7.9, 6 M guanidine chloride, 500 mM NaCl)
17. Buffer C (40 mM Tris–HCl of pH 7.9, 300 mM KCl, 10 mM EDTA, 1 mM PMSF, 1 mM DTT)
18. Buffer D (50 mM Tris–HCl of pH 7.9, 6 M guanidine chloride, 10 mM $MgCl_2$, 0.01 mM $ZnCl_2$, 1 mM EDTA, 10 mM DTT, 10% glycerol)
19. Buffer E (50 mM Tris–HCl of pH 7.9, 200 mM KCl, 10 mM $MgCl_2$, 0.01 mM $ZnCl_2$, 1 mM EDTA, 5 mM β-mercaptoethanol, 20% glycerol)
20. Buffer F (50 mM Tris–HCl of pH 7.9, 5% glycerol)
21. α storage buffer (50 mM Tris–HCl of pH 7.9, 200 mM KCl, 10 mM $MgCl_2$, 1 mM EDTA, 5 mM β-mercaptoethanol, 20% glycerol)
22. 2× SDS-loading buffer (63 mM Tris–HCl of pH 6.8, 2% SDS, 5% β-mercaptoethanol, 25% glycerol, 0.3% bromophenol blue)

23. SDS-running buffer (25 mM Tris, 250 mM glycine of pH 8.3, 0.1% SDS)
24. Destaining solution (10% acetic acid, 50% methanol, 40% water)
25. 100 mM phenylmethanesulfonyl fluoride (PMSF) in ethanol (Sigma)
26. 2% lysozyme (Sigma, cat #L-6876) (~50,000 U/mg)
27. 10% sodium desoxycholate (Sigma)
28. 10% *n*-octyl-β-D-glucopyranoside (Sigma)
29. Triton X-100 (Sigma)
30. 2 M imidazole (pH adjusted to 8.0 with 10 M HCl) (Sigma)
31. Glycerol (Fisher)
32. Trichloroacetic acid (Aldrich)
33. Coomassie Brilliant Blue G-250 (Bio-Rad)
34. Acetone (Aldrich)
35. 10% polyacrylamide (37.5:1 acrylamide:bisacrylamide), 0.1% SDS, slab gel (10 × 7 × 0.075 cm)
36. Prestained protein molecular-weight markers (7–210 kDa) (Bio-Rad)
37. Protein Assay Kit (Bio-Rad, cat #500-0002)
38. Ni:NTA-agarose (Qiagen)
39. Dialysis membranes (10 kDa molecular-weight cutoff) (VWR, cat #25223-821)
40. Dialysis-membrane closures (VWR)
41. Collodion dialysis bags (10 kDa molecular-weight cutoff) (Schleicher & Schuell)
42. Nanosep-30 K centrifugal concentrators (VWR)
43. Econo-Pac 20-mL chromatography columns (Bio-Rad)
44. Chromatography column frits (1.5 × 0.3 cm) (Bio-Rad)
45. 15-mL culture tubes (autoclave sterilized) (VWR)
46. Culture-tube stainless steel closures (autoclave-sterilized) (VWR)
47. 2.8 L triple-baffled Fernbach flask (autoclave-sterilized) (Bellco Glass, Inc., cat #2551-02800)
48. 30-mL polypropylene copolymer centrifuge tube with cap (VWR, cat #21010-567)
49. 250-mL polypropylene copolymer centrifuge bottle with cap (VWR, cat #21020-028)
50. 1-L polypropylene copolymer centrifuge bottle with cap (VWR, cat #21020-061)
51. 200-mL steel beaker (VWR)

52. Branson 450 sonicator (VWR)
53. Sorvall RC-3B centrifuge (Thermo Scientific)
54. Sorvall RC-5B centrifuge (Thermo Scientific)

2.4. Static Photocrosslinking

1. Cystamine dihydrochloride (Sigma)
2. Acryloyl chloride (Aldrich)
3. Acrylamide (Bio-Rad)
4. TEMED (BioRad)
5. 10% ammonium persulfate (freshly made)
6. SurfaSil siliconizing agent (Pierce)
7. Derivatized promoter DNA fragment (**Subheading 3.2**)
8. RNAP or RNAP derivative (**Subheading 3.3**)
9. DNase I (126 U/μl) (Sigma, cat #D7291)
10. Micrococcal nuclease in nuclease dilution solution (50 U/μl) (GE Life Sciences, cat#27-0584)
11. Nuclease dilution solution (5 mM $CaCl_2$, 0.1 mM PMSF, 50% glycerol)
12. 2× DTT-free transcription buffer (50 mM HEPES-HCl of pH 8.0, 200 mM KCl, 20 mM $MgCl_2$, 10% glycerol)
13. Nondenaturing loading buffer (0.3% bromophenol blue, 0.3% xylene cyanol, 30% glycerol)
14. 5× SDS-loading buffer (300 mM Tris–HCl of pH 6.8, 10% SDS, 20 mM EDTA, 25% β-mercaptoethanol, 0.1% bromophenol blue, 50% glycerol)
15. 0.5× TBE (45 mM Tris-borate of pH 8.0, 1 mM EDTA)
16. SDS-running buffer (25 mM Tris, 250 mM glycine of pH 8.3, 0.1% SDS)
17. 10% SDS
18. 1 M DTT (freshly made)
19. 0.1 mM PMSF (Sigma)
20. 0.22 μg/mL heparin (Sigma, cat #H-3393) (grade I–A, from porcine intestinal mucosa, ~170 USP U/mg)
21. 7–15% gradient polyacrylamide (37.5:1 acrylamide:bisacrylamide) Tris–HCl slab gel (Bio-Rad, cat #161-0902)
22. Prestained protein molecular-weight markers (7–210 kDa) (Bio-Rad)
23. Silicone rubber heating mat (200 W, 120 VAC; 25 × 10 cm) (Cole-Parmer, cat #P-03125-40)
24. Variable voltage controller (Cole-Parmer, cat #P-01575-10)
25. Digital thermometer (Cole-Parmer, cat #P-91000-00)
26. Thermocouple probe (needle, 0.7-mm diameter) (Cole-Parmer, cat #P-08505-92)

27. Large binder clips (width = 5 cm) (Staples)
28. Filter unit (22-μm pore size, 250 mL) (Millipore)
29. 50-mL Büchner funnel with glass frit (10-μm pore size) (Fisher)
30. 500-mL separating funnel (Fisher)
31. X-ray exposure holder with intensifying screen (Kodak)
32. Light box (VWR)
33. Rayonet RPR-100 photochemical reactor equipped with 16 RPR-3500 Å tubes (Southern New England Ultraviolet)
34. Speedvac evaporator (Savant)

2.5. Kinetic Photocrosslinking

1. Derivatized promoter DNA fragment (**Subheading 3.2**).
2. RNAP or RNAP derivative (**Subheading 3.3**).
3. DNase I (126 U/μl) (Sigma).
4. Micrococcal nuclease (2,000 U/μl) (New England Biolabs).
5. Nuclease dilution solution: 5 mM $CaCl_2$, 0.1 mM PMSF, and 50% glycerol.
6. OmniCleave endonuclease (200 U/μl) (Epicentre).
7. OmniCleave endonuclease dilution buffer (Epicentre).
8. DTT- and glycerol-free transcription buffer: 25 mM HEPES-HCl of pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, and 0.1% Tween-20.
9. DTT- and glycerol-free, BSA-containing transcription buffer: 25 mM HEPES-HCl of pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 0.1% Tween-20, 100 μg/mL BSA (prepared immediately before use; *see* **Subheading 3.5.1**).
10. Urea/NaI solution: 5 M urea and 0.5 M NaI
11. Quench solution: 5 M urea, 0.5 M NaI, 0.5 M β-mercaptoethanol (prepared immediately before use; *see* **Subheading 3.5.1**).
12. 0.1 M PMSF in ethanol (Sigma).
13. Criterion precast 4–20% gradient Tris–HCl gel (Bio-Rad).
14. Precision-Plus protein dual-color standards (10–250 kDa) (Bio-Rad).
15. 5× SDS-loading buffer: 300 mM Tris–HCl of pH 8.3, 10% SDS, 20 mM EDTA, 25% β-mercaptoethanol, 0.1% bromophenol blue, and 50% glycerol.
16. SDS-running buffer: 25 mM Tris, 250 mM glycine of pH 8.3, and 0.1% SDS.
17. Tween-20 (Bio-Rad).
18. Bovine serum albumin (BSA; Sigma).
19. Urea (Fisher).
20. NaI (Fisher).

21. β-mercaptoethanol (14.2 M; Bio-Rad).
22. Methanol (Fisher).
23. Argon (compressed, high-purity; Airgas).
24. 50-mL polypropylene centrifuge tubes (Falcon, Fisher).
25. 5-mL Luer-Lok disposable syringes (VWR).
26. 1-mL tuberculin disposable syringes (VWR).
27. Hamilton 50-μl Luer-Lok syringe (Hamilton).
28. Linagraph direct print paper (Kodak).
29. Microprocessor-controlled, stepping-motor-driven rapid-quench-flow system equipped with 20-μl sample chamber with acrylic OP1 optical windows(RQF-3 with modifications; KinTek, Austin, TX) (*see* **Note 21**).
30. Nanosecond pulsed Nd-YAG laser equipped with third-harmonic generator (Powerlite 7,010-DS/TS/QS; 1 Hz; Continuum, Santa Clara CA) (*see* **Note 22**).
31. Raman cell (RC-1; Light Age, Somerset NJ).
32. Pellin-Broca prism (CVI Laser, Albuquerque NM).
33. 355- and 309-nm laser optics (CVI Laser, Albuquerque NM).
34. Laser optics mounts (Thorlabs, Newton NJ).
35. Optical table (4′ × 6′; Newport, Irvine CA).
36. Laser calorimeter (AC25UV; Scientech, Boulder CO).
37. Laser energy meter (Astral AD30; Scientech, Boulder CO).
38. Laser safety goggles (OD ≥ 6 at 1,064 nm; OD ≥ 7 at 190–532 nm; L332CB; Uvex Safety, Inc., Smithfield RI).
39. Digital delay/pulse generator (DG535; Stanford Research Systems, Sunnyvale CA).
40. Circulating water bath (Isotemp 1016D; Fisher).
41. Multimode imager (Typhoon 9,400; GE Life Sciences).

3. Methods

3.1. Preparation of Derivatized DNA Fragment: Chemical Reactions

3.1.1. Preparation of Phosphorothioate Oligodeoxyribonucleotide

1. Perform 24 standard cycles of solid-phase β-cyanoethyl-phosphoramidite oligodeoxyribonucleotide synthesis to prepare CPG-linked precursor containing residues 3–26 of desired oligodeoxyribonucleotide. Use the following settings: cycle, 1.0 μmol CE; DMT, on; end procedure, manual.
2. Replace iodine/water/pyridine/tetrahydrofuran solution (bottle 15) by tetraethylthiuram disulfide/acetonitrile solution.

Perform one modified cycle of solid-phase β-cyanoethylphosphoramidite oligodeoxyribonucleotide synthesis to add residue 2 and phosphorothioate linkage. Use the following settings: cycle, 1.0 μmol sulfur; DMT, on; end procedure, manual.

3. Replace tetraethylthiuram disulfide/acetonitrile solution (bottle 15) by iodine/water/pyridine/tetrahydrofuran solution. Place collecting vial on the DNA synthesizer. Perform one standard cycle of solid-phase β-cyanoethylphosphoramidite oligodeoxyribonucleotide synthesis to add residue 1. Use the following settings: cycle, 1.0 μM CE; DMT, on; end procedure, CE.
4. Remove collecting vial, screw cap tightly, and deblock by incubating 8 h at 55°C. Transfer sample to 6-mL polypropylene round-bottom tube, place tube in Speedvac, and spin 20 min with Speedvac lid ajar and with no vacuum (allowing evaporation of ammonia). Close Speedvac lid, apply vacuum, and dry.
5. Detritylate and purify ~75 nmol on OPC according to supplier's protocol.
6. Dry in Speedvac.
7. Resuspend in 100 μl TE. Remove 2 μl aliquot, dilute with 748 μl TE, and determine concentration from UV absorbance at 260 nm (molar extinction coefficient = 240,000 AU/M cm).
8. To confirm purity of oligodeoxyribonucleotide, mix aliquot containing 1 nmol oligodeoxyribonucleotide with equal volume of formamide. Apply to 12% polyacrylamide (29:1 acrylamide:bisacrylamide), 8 M urea, 0.5 × TBE slab gel (10 × 7 × 0.075 cm). As marker, load in adjacent lane 5 μl denaturing loading buffer. Electrophorese 30 min at 25 V/cm. Disassemble gel, place on intensifying screen, and view in dark using 254-nm germicidal lamp. Oligodeoxyribonucleotide should appear as dark shadow against green background and should migrate more slowly than bromophenol blue. If purity is >95%, proceed to next step.
9. Divide remainder of sample into 50-nmol aliquots, transfer to 1.5-mL siliconized polypropylene microcentrifuge tubes, dry in Speedvac, and store at –20°C (stable for at least 2 years).

3.1.2. Derivatization of Oligodeoxyribonucleotide [All Steps Carried Out Under Subdued Lighting (See Note 4)]

1. Dissolve 10 mg (42 μmol) azidophenacyl bromide in 1 mL chloroform. Transfer 100 μl aliquots (4.2 μmol) to 1.5-mL siliconized polypropylene microcentrifuge tubes, and dry in Speedvac. Wrap tubes with aluminum foil, and store desiccated at 4°C (stable indefinitely).

2. Resuspend 50 nmol aliquot of phosphorothioate oligodeoxyribonucleotide (**Subheading 3.1.1**) in 50∝ l water, and resuspend 42 μmol aliquot of azidophenacyl bromide in 220 μl methanol.
3. Mix 50 μl (50 nmol) phosphorothioate oligodeoxyribonucleotide solution, 5 μl 1 M potassium phosphate (pH 7.0), and 55 μl (1 μmol) azidophenacyl bromide solution in 1.5 mL siliconized polypropylene microcentrifuge tube. Incubate 3 h at 37°C in the dark.
4. Precipitate derivatized oligodeoxyribonucleotide by adding 11 μl 3 M sodium acetate (pH 5.2), and 275 μl ice-cold 100% ethanol. Invert tube several times, and place at –80°C for 30 min. Centrifuge 5 min at 13,000 × *g* at 4°C. Remove supernatant, wash pellet with ice-cold 70% ethanol. Air dry 15 min at RT. Store at –20°C (stable for at least 1 year).

3.1.3. Purification of Derivatized Oligodeoxyribonucleotide [All Steps Carried Out Under Subdued Lighting (See Note 4)]

1. Resuspend derivatized oligodeoxyribonucleotide in 100 μl 50 mM triethylammonium acetate (pH 7.0).
2. Analyze 5 μl aliquot by C18 reversed-phase HPLC to confirm efficiency of derivatization reaction. Use LiChrospher 100 RP-18 C18 reversed-phase HPLC column (5 μm), with solvent A = 50 mM triethylammonium acetate (pH 7.0), 5% acetonitrile; solvent B = 100% acetonitrile; and flow rate = 1 mL/min. Equilibrate column with ten column volumes solvent A before loading sample. After loading sample, wash column with six column volumes solvent A, and elute with 50 min gradient of 0–70% solvent B in solvent A. Derivatized and underivatized oligodeoxyribonucleotides elute at ~25% solvent B and ~16% solvent B, respectively (*see* **Notes 5** and **6**).
3. If derivatization efficiency is ≥80%, purify remainder of sample using procedure of **step 2**, collecting peak fractions (*see* **Notes 5** and **6**).
4. Pool peak fractions, divide into 1 mL aliquots, and dry in Speedvac. Store desiccated at –20°C in the dark (stable for at least 1 year).
5. Resuspend one aliquot in 100 μl TE. Remove 5 μl, dilute with 495 μl water, and determine concentration from UV absorbance at 260 nm (molar extinction coefficient = 242,000 AU/M cm).
6. Divide remainder of derivatized-oligodeoxyribonucleotide/TE solution from **step 5** into twenty 5 pmol aliquots and one larger aliquot, dry in Speedvac, and store desiccated at –20°C in the dark (stable for at least 1 year).

3.2. Preparation of Derivatized DNA Fragment: Enzymatic Reactions

3.2.1. Radiophosphorylation of Derivatized Oligodeoxyribonucleotide [All Steps Carried Out Under Subdued Lighting (See Note 4)]

1. Resuspend 5 pmol derivatized oligodeoxyribonucleotide in 12 µl water. Add 2 µl 10× phosphorylation buffer, 5 µl [γ^{32}P] ATP (50 µCi), and 1 µl (10 U) T4 polynucleotide kinase. Incubate 15 min at 37°C. Terminate reaction by heating 5 min at 65°C (*see* **Note 7**).
2. Add 15 µl water.
3. Desalt radiophosphorylated derivatized oligodeoxyribonucleotide into TE using CHROMA SPIN + TE-10 spin column according to supplier's protocol.
4. Immediately proceed to next step, or, if necessary, store radiophosphorylated derivatized oligodeoxyribonucleotide solution at –20°C in the dark (stable for up to 24 h).

3.2.2. Annealing, Extension, and Ligation of Radiophosphorylated Derivatized Oligodeoxyribonucleotide [All Steps Carried Out Under Subdued Lighting (See Note 4)]

1. In 1.5-mL siliconized polypropylene microcentrifuge tube, mix 34 µl radiophosphorylated derivatized oligodeoxyribonucleotide, 1 µl 10 µM downstream primer, 1∝ l 1 µM M13mp2 *(ICAP-UV5)* ssDNA (for analysis of crosslinks to template DNA strand) or M13mp2 *(ICAP-UV5)*-rev ssDNA (for analysis of crosslinks to nontemplate DNA strand), and 4 l 10× annealing buffer.
2. Heat 5 min at 65°C (*see* **Note 7**). Transfer to 500-mL beaker containing 200 mL water at 65°C, and place beaker at room temperature to permit slow cooling (65–25°C in ~60 min).
3. Add 1∝ l 25 mM dNTPs, 1 µl 100 mM ATP, 1 µl (3 U) T4 DNA polymerase, and 1 µl (5 U) T4 DNA ligase. Perform parallel reaction without ligase as "no-ligase" control.
4. Incubate 15 min at room temperature, followed by 35 min at 37°C. Terminate reaction by adding 1 l 10% SDS.
5. Desalt into TE using CHROMA SPIN + TE-100 spin column according to supplier's protocol. Immediately proceed to next step.

3.2.3. Digestion and Purification of Derivatized DNA Fragment [All Steps Carried Out Under Subdued Lighting (See Note 4)]

1. In 1.5-mL siliconized polypropylene microcentrifuge tube, mix 40 µl product from **Subheading 3.2.2**, 4.5 µl 10× digestion buffer, 0.25 µl (10 U) HaeIII, or 0.25 µl (10 U) PvuII (*see* **Note 8**). Incubate 1 h at 37°C.
2. Perform parallel reaction using 40 µl "no-ligase" control from **Subheading 3.2.2**.
3. Mix 3 µl aliquots of reaction of **step 1** and of "no-ligase" control reaction of **step 2**, each with 7 µl denaturing loading buffer. Heat 5 min at 65°C, and then apply to 12% polyacrylamide (29:1 acrylamide:bisacrylamide), 8 M urea, 0.5× TBE slab gel (10 × 7 × 0.075 cm). As marker, load in adjacent lane 5 µl denaturing loading buffer. Electrophorese 30 min at 25 V/cm. Dry gel, expose to X-ray film 1 h at room tem-

perature, and process film. Estimate ligation efficiency by comparing reaction and "no-ligase" control lanes. If ligation efficiency is ≥80%, proceed to next step. If not, repeat the steps in **Subheadings 3.2.1** and **3.2.2**.

4. Mix remainder of reaction of **step 1** (42 µl) with 10 µl 50% glycerol. Apply to *nondenaturing* 7.5% polyacrylamide (29:1 acrylamide:bisacrylamide), 0.5× TBE slab gel (10 × 7 × 0.15 cm). As marker, load in adjacent lane 5 µl *nondenaturing* loading buffer. Electrophorese at 25 V/cm until bromophenol blue reaches bottom of the gel.
5. Remove one glass plate, and cover gel with plastic wrap. Attach two autoradiography markers to gel. Expose to X-ray film for 60–90 s at room temperature, and process film. Cut out portion of film corresponding to derivatized DNA fragment. Using light box, superimpose cut-out film on gel, using autoradiography markers as alignment reference points. Using disposable scalpel, excise portion of gel corresponding to derivatized DNA fragment.
6. Place excised gel slice in 1.5-mL siliconized polypropylene microcentrifuge tube, and crush with 1-mL pipette tip. Add 300 µl elution buffer, centrifuge 5 s at 5,000 × *g*, and incubate 12 h at 37°C.
7. Transfer supernatant to Spin-X centrifuge filter, centrifuge 1 min at 13,000 × *g* at room temperature in fixed-angle microcentrifuge.
8. Transfer filtrate to 1.5-mL siliconized polypropylene microcentrifuge tube. Precipitate derivatized DNA fragment by addition of 1 mL ice-cold 100% ethanol. Invert tube several times, and place at –20°C for 30 min. Centrifuge 5 min at 13,000 × *g* at 4°C in fixed-angle microcentrifuge. Remove and dispose of supernatant, wash pellet with 500 µl ice-cold 70% ethanol, and air dry 15 min at room temperature.
9. Resuspend in 30 µl low-EDTA TE. Determine radioactivity by Cerenkov counting. Remove 1 µl aliquot, and determine DNA concentration using PicoGreen dsDNA quantitation kit according to supplier's protocol. Calculate specific activity (expected specific activity ~5,000 Ci/mmol).
10. Store derivatized DNA fragment at 4°C in the dark (stable for ~1 week).

3.3. Preparation of RNAP and RNAP Derivatives

3.3.1. Preparation of Hexahistidine-Tagged Recombinant α Subunit

1. Transform *E. coli* strain BL21(DE3) pLysS with plasmid pHTT7f1-NHα. Plate to TYE agar containing 200 µg/mL ampicillin and 35 µg/mL chloramphenicol, and incubate 12 h at 37°C.
2. Inoculate single colony into 5 mL LB containing 200 µg/mL ampicillin and 35 µg/mL chloramphenicol in 15-mL

culture tube with culture-tube stainless steel closure, and shake vigorously for 12 h at 37°C. Transfer to 15-mL polypropylene centrifuge tube, and centrifuge 5 min at 3,000 × *g* at room temperature. Discard supernatant, wash cell pellet twice with 5 mL LB, and resuspend cell pellet in 5 mL LB in 15-mL polypropylene centrifuge tube.

3. Inoculate into 1 L LB containing 200 μg/mL ampicillin and 35 μg/mL chloramphenicol in 2.8-L Fernbach flask, and shake vigorously at 37°C until OD_{600} = 0.6. Add 1 mL 1 M IPTG, and shake vigorously for an additional 3 h at 37°C.
4. Transfer culture to 1-L polypropylene copolymer centrifuge bottle. Harvest cells by centrifugation 20 min at 5,000 × *g* at 4°C.
5. Resuspend cell pellet in 100 mL buffer A at 4°C. Transfer into 200 mL steel beaker, and place beaker on ice. Lyse cells with four 40-s sonication pulses at 25% maximum sonicator output (2 min pause between each pulse).
6. Transfer lysate to 250-mL polypropylene copolymer centrifuge bottle. Centrifuge 15 min at 15,000 × *g* at 4°C. Collect supernatant.
7. Transfer supernatant to 250-mL glass beaker. Add 35 g ammonium sulfate, and stir 20 min on ice.
8. Transfer suspension to 250-mL polypropylene copolymer centrifuge bottle. Centrifuge 20 min at 15,000 × *g* at 4°C. Discard supernatant.
9. Resuspend pellet in 28 mL buffer B containing 5 mM imidazole. Transfer to 30-mL polypropylene copolymer centrifuge tube, and rock gently 30 min at 4°C. Centrifuge 15 min at 15,000 × *g* at 4°C.
10. Load supernatant onto 5 mL Ni:NTA-agarose column preequilibrated with 25 mL buffer B containing 5 mM imidazole (prepared by pouring 10 mL Ni:NTA-agarose suspension into a 20-mL Econo-Pac column, removing snap-off tip at bottom of column, allowing liquid to drain, and then placing a frit on the top of column bed). Collect flow through, and reload onto column. Wash column with 50 mL buffer B containing 5 mM imidazole, and 25 mL buffer B containing 10 mM imidazole. Elute column with 15 mL buffer B containing 20 mM imidazole, 15 mL buffer B containing 30 mM imidazole, 15 mL buffer B containing 40 mM imidazole, and 15 mL buffer B containing 150 mM imidazole. Collect 5 mL fractions.
11. Transfer 10 μl aliquot of each fraction to 1.5-mL siliconized polypropylene microcentrifuge tube, add 90 μl water and 100 μl 10% trichloroacetic acid. Place on ice 20 min. Centrifuge 5 min at 13,000 × *g* at room temperature. Discard supernatant. Wash pellet with 500 μl acetone, and air dry for

15 min. Dissolve pellet in 5 µl water, add 5 µl 2× SDS-loading buffer, heat 3 min at 100°C, and apply to 10% polyacrylamide (37.5:1 acrylamide:bisacrylamide), 0.1% SDS, slab gel (10 × 7 × 0.075 cm). As marker, load into adjacent lane 5 µl prestained protein molecular-weight markers. Electrophorese in SDS-running buffer at 25 V/cm until bromophenol blue reaches bottom of gel. Stain gel by gently shaking for 5 min in 50 mL 0.2% Coomassie Brilliant Blue G-250 in destaining solution. Destain by gently shaking for 1 h in 100 mL destaining solution.

12. Pool fractions containing homogeneous α (typically fractions with buffer B containing 40–150 mM imidazole). Dialyze using 10-kDa molecular-weight-cutoff dialysis membrane against two 1 L changes of α storage buffer for 16 h at 4°C.
13. Determine protein concentration and total protein amount using Bio-Rad Protein Assay according to supplier's protocol.
14. After dialysis, measure volume and transfer to 30-mL polypropylene copolymer centrifuge tube. Add 3 g ammonium sulfate per 10 mL, and rock gently 20 min at 4°C. Centrifuge 20 min at 15,000 × *g* at 4°C.
15. Remove and discard 10 mL of supernatant. Resuspend pellet in remaining supernatant. Divide into 50 µl aliquots, and transfer to 1.5-mL siliconized polypropylene microcentrifuge tubes. Centrifuge aliquots 5 min at 13,000 × *g* at 4°C. Store at –80°C (stable for at least 1 year). Expected yield: 20–30 mg (250–500 µg/aliquot). Expected purity: > 99%.

3.3.2. Preparation of Crude Recombinant RNAP Subunits and Subunit Fragments

1. Transform plasmid encoding RNAP subunit or subunit fragment into *E. coli* strain BL21(DE3) pLysS (for plasmids with ϕ*10P*- or *lacP*- ϕ*10P*-based expression; **Tables 1** and **2**) or *E. coli* strain XL1-blue (for plasmids with *lacP*-based expression; **Tables 1** and **2**). Plate transformants of BL21(DE3) pLysS to TYE agar containing 200 µg/mL ampicillin (40 µg/mL kanamycin for plasmid $p\beta'_{820-1,407}$) and 35 µg/mL chloramphenicol, and incubate 12 h at 37°C. Plate transformants of XL1-blue to TYE agar containing 200 µg/mL ampicillin (40 µg/mL kanamycin for plasmid $p\beta_{1-235}$) and 20 µg/mL tetracycline, and incubate 16 h at 37°C.
2. Inoculate single colony into 5 mL LB containing antibiotics at concentrations specified in **step 1** in 15-mL culture tube with stainless steel closure, and shake vigorously for 12 h at 37°C. Transfer to 15-mL polypropylene centrifuge tube, and centrifuge 5 min at 3,000 × *g* at room temperature. Discard supernatant, wash cell pellet twice with 5 mL LB, and resuspend cell pellet in 5 mL LB.

3. Inoculate into 1 L LB containing antibiotics at concentrations specified in **step 1** in 2.8-L Fernbach flask, and shake vigorously at 37°C until OD_{600} = 0.6. Add 1 mL 1 M IPTG, and shake vigorously for an additional 3 h [transformants of BL21(DE3) pLysS] or 5 h (transformants of XL1-blue) at 37°C.
4. Transfer culture to 1-L polypropylene copolymer centrifuge bottle. Harvest cells by centrifugation 20 min at 5,000 × *g* at 4°C.
5. Resuspend cell pellet in 10 mL buffer C containing 0.2% sodium desoxycholate and 0.02% lysozyme in 30-mL polypropylene copolymer centrifuge tube at 4°C. Place tube on ice. Lyse cells with five 30-s sonication pulses at 25% maximum sonicator output (2-min pause between each pulse).
6. Centrifuge 20 min at 15,000 × *g* at 4°C. Discard supernatant.
7. Resuspend pellet in 10 mL buffer C containing 0.2% *n*-octyl-β-D-glucopyranoside (0.5% Triton X-100 for preparation of σ^{70}) and 0.02% lysozyme at 4°C. Sonicate as in **step 5**. Centrifuge 20 min at 15,000 × *g* at 4°C. Discard supernatant.
8. Resuspend pellet in 10 mL buffer C containing 0.2% *n*-octyl-β-D-glucopyranoside (0.5% Triton X-100 for preparation of σ^{70}) at 4°C. Sonicate as in **step 5**. Centrifuge 20 min at 15,000 × *g* at 4°C. Discard supernatant.
9. Resuspend pellet in 10 mL buffer C at 4°C. Place tube on ice, and sonicate 10 s at 25% maximum sonicator output. Divide into 500 µl aliquots, and transfer to 1.5–mL siliconized polypropylene microcentrifuge tubes. Centrifuge 5 min at 13,000 × *g* at 4°C. Discard supernatant.
10. Add 100 µl ice-cold buffer C containing 10% glycerol to each aliquot. Store at –80°C (stable for at least 2 years). Expected yield: 50–100 mg (1.5–3 mg/aliquot). Expected purity: 50–90%.

3.3.3. Reconstitution of RNAP and RNAP Derivatives

1. Thaw aliquots containing purified α subunit (from **Subheading 3.3.1**, **step 15**) and crude recombinant RNAP subunits and subunit fragments (from **Subheading 3.3.2**, **step 10**) by placing on ice for 10 min. Centrifuge 30 s at 13,000 × *g* at 4°C. Discard supernatants.
2. Resuspend each pellet in 500 µl buffer D. Incubate 30 min at 4°C, rocking gently. Centrifuge 5 min at 13,000 × *g* at 4°C.
3. Transfer supernatant to new 1.5-mL siliconized polypropylene microcentrifuge tubes at 4°C. Determine protein concentrations using Bio-Rad Protein Assay according to supplier's protocol (expected concentrations: 3–6 mg/mL).

4. Prepare core reconstitution mixture by combining in 1.5-mL siliconized polypropylene microcentrifuge tube 30 μg N-terminally hexahistidine-tagged α, 300 μg β (or 170 μg $\beta_{1–235}$ and 800 μg $\beta_{235–1,342}$; or 700 μg $\beta_{1–989}$ and 300 μg $\beta_{950–1,342}$), 500 μg β′ (or 400 μg $\beta'_{1–577}$ and 500 μg $\beta'_{544–1,407}$; or 700 μg $\beta'_{1–877}$ and 330 μg $\beta'_{820–1,407}$), and 70 μg ω, and diluting with buffer D to a total protein concentration 450 μg/mL.
5. Prepare σ^{70} reconstitution mixture by adding 250 μg σ^{70} to 1.5-mL siliconized polypropylene microcentrifuge tube and diluting with buffer D to a total protein concentration 1,500 μg/mL.
6. Dialyze core and σ^{70} reconstitution mixtures separately in collodion dialysis bags against two 1 L changes of buffer E for 16 h at 4°C.
7. Transfer core and σ^{70} reconstitution mixtures to separate 2.0-mL siliconized polypropylene microcentrifuge tubes. Centrifuge 5 min at 13,000 × *g* at 4°C. Combine supernatants in a single, new 2.0-mL siliconized polypropylene microcentrifuge tube.
8. Incubate 45 min at 30°C. Centrifuge 10 min at 13,000 × *g* at 4°C.

3.3.4. Purification of RNAP and RNAP Derivatives

1. During incubation of **step 8** of **Subheading 3.3.3**, place 200 μl Ni:NTA-agarose in 2.0-mL siliconized polypropylene microcentrifuge tube, and centrifuge 2 min at 13,000 × *g* at 4°C. Remove supernatant.
2. Resuspend Ni:NTA-agarose in 1 mL buffer F containing 5 mM imidazole at 4°C. Centrifuge 2 min at 13,000 × *g* at 4°C. Remove supernatant. Repeat two times.
3. Add supernatant of **step 8** of **Subheading 3.3.3** to Ni:NTA-agarose from **step 2**. Incubate 45 min at 4°C, rocking gently. Centrifuge 2 min at 13,000 × *g* at 4°C. Discard supernatant.
4. Resuspend in 1.5 mL buffer F containing 5 mM imidazole at 4°C. Rock gently 15 s at 4°C. Centrifuge 2 min at 13,000 × *g* at 4°C. Discard supernatant. Repeat two times.
5. Resuspend in 250 μl buffer F containing 150 mM imidazole. Rock gently 2 min at 4°C. Centrifuge 2 min at 13,000 × *g* at 4°C.
6. Transfer supernatant to Nanosep-30 K centrifugal concentrator. Centrifuge at 13,000 × *g* at 4°C until sample volume is reduced to ~50 μl (~15 min).
7. Transfer sample to 1.5-mL siliconized polypropylene microcentrifuge tube. Add, in order, 1 μl 0.1 M β-mercaptoethanol and 50 μl glycerol; mix well; and store at –20°C (stable for at least 1 month).

8. Determine protein concentration using Bio-Rad Protein Assay according to supplier's protocol. Expected yield: 100 μg. Expected purity > 90%.

3.4. Static Photocrosslinking (see Note 9)

3.4.1. Synthesis of N,N′-Bisacryloylcystamine (BAC) (See Note 10)

1. Acryloyl chloride is highly toxic. Therefore, all manipulations in this section must be performed in a fume hood.
2. Dissolve 4.0 g (18 mmol) cystamine dihydrochloride in 40 mL 3 M NaOH (120 mmol). Dissolve 4.3 mL (54 mmol) acryloyl chloride in 40 mL chloroform. Mix solutions in 500-mL flask (*see* **Note 11**). (Two phases will form: an upper, aqueous phase; and a lower, organic phase.) Place flask on plate stirrer, and stir 3 min at room temperature, followed by 15 min at 50°C.
3. Discontinue stirring. Immediately transfer reaction mixture to 500-mL separating funnel, allow phases to separate (~2 min), and transfer lower, organic phase to 250-mL beaker.
4. Place on ice for 10 min. Collect precipitate by filtration in Büchner funnel.
5. Transfer precipitate to 250-mL beaker with 30 mL chloroform at room temperature. Place beaker on plate stirrer, and stir 1 min at room temperature, followed by 5 min at 50°C. Place on ice for 10 min, and collect precipitate (BAC) by filtration in Büchner funnel.
6. Transfer precipitate to 50-mL polypropylene centrifuge tube. Seal tube with Parafilm, pierce seal several times with syringe needle, place tube in vacuum desiccator, and dry under vacuum 16 h at room temperature. Expected yield: 1.5–1.9 g.

3.4.2. Preparation of Polyacrylamide:BAC Gel

1. Prepare 20% acrylamide:BAC (19:1) stock solution by dissolving, in order, 19 g acrylamide and 1 g BAC in 80 mL water in 200-mL beaker at room temperature. Place on plate stirrer and stir 10 min at 60°C (*see* **Note 12**). Adjust volume to 100 mL with water. Allow solution to cool to room temperature. Filter stock solution using 0.22-μm filter unit, and store at room temperature in the dark (stable for at least 2 months).
2. Mix 9 mL 20% acrylamide:BAC (19:1) stock solution, 1.8 mL 10× TBE, and 25.2 mL water. Add 180 μl TEMED and 90 μl freshly prepared 10% ammonium persulfate (*see* **Note 13**). Immediately pour into slab gel assembly with siliconized notched glass plate (27 × 16 × 0.1 cm) (*see* **Note 14**). Insert comb and heat slab gel assembly to ~60°C by positioning task lamp with 60-W tungsten bulb 2 cm from the outer glass plate (*see* **Note 15**). Allow 10–20 min for polymerization. (The polyacrylamide:BAC gel is stable for up to 72 h at 4°C.)

3.4.3. Formation and Isolation of RNAP-Promoter Open Complex [All Steps Carried Out Under Subdued Lighting (See Note 4)]

1. Place polyacrylamide:BAC slab gel in electrophoresis apparatus, clip 10 × 25 cm silicone heating mat to directly to outer glass plate of the slab gel assembly with four large binder clips, and pour 0.5× TBE buffer in upper and lower reservoirs.
2. Prerun gel for 2 h at 20 V/cm.
3. Prewarm electrophoresis unit by placing in 37°C cabinet for 3 h.
4. During 37°C prewarming of **step 3**, dilute RNAP or RNAP derivative to 180 μg/mL (400 nM) in buffer F containing 1 mM β-mercaptoethanol and 50% glycerol.
5. Immediately after 37°C prewarming of **step 3**, add the following, in order, to a 1.5-mL siliconized polypropylene microcentrifuge tube: 2 μl 5 nM derivatized DNA fragment (~5,000 Ci/mmol), 5 μl 2× DTT-free transcription buffer, 2 μl water, and 1 μl 180 μg/mL (400 nM) RNAP or RNAP derivative (all at room temperature).
6. Incubate 20 min at 37°C in the dark.
7. During incubation of **step 6** apply voltage to gel: 16 V/cm. Wash wells of gel carefully with 0.5× TBE to remove unpolymerized acrylamide and BAC. (*Caution*: Care must be exercised to avoid electrocution.) Connect heating mat to variable voltage controller. Monitor gel temperature at 5-min intervals by inserting thermocouple probe into the gel for 5 s (and removing immediately thereafter). Maintain gel temperature at 37°C, adjusting heater voltage as necessary (typically 12–14 V).
8. After completing incubation of **step 6**, immediately add 1 μl 0.22 μg/mL heparin (prewarmed to 37°C), mix, and immediately apply sample to gel (*see* **Note 16**). Load 5 μl nondenaturing loading buffer into adjacent lane. (*Caution*: Care must be exercised to avoid electrocution.) Continue electrophoresis 20 min at 16 V/cm. Monitor gel temperature at 5-min intervals by inserting thermocouple probe into the gel for 5 s (and removing immediately thereafter). Maintain gel temperature at 37°C, adjusting heater voltage as necessary (typically 12–14 V).
9. Immediately proceed to next step (**Subheading 3.4.4**).

3.4.4. UV Irradiation of RNAP-Promoter Open Complex

1. Remove gel with both glass plates in place (*see* **Note 17**), and mount vertically in a Rayonet RPR-100 photochemical reactor equipped with 16 RPR-3,500 Å tubes.
2. Immediately UV irradiate 3 min (17 mJ/mm^2 at 350 nm) (*see* **Note 18**). (*Caution*: Care must be exercised to avoid injury to eyes and skin.)
3. Immediately proceed to next step (**Subheading 3.4.5**).

3.4.5. Identification, Excision, and Solubilization of Portion of Gel Containing RNAP-Promoter Open Complex

1. Remove one glass plate, and cover gel with plastic wrap (leaving other glass plate in place). Attach two autoradiography markers. Expose to X-ray film for 1.5 h at room temperature (*see* **Note 19**). Process film.
2. Cut out portion of film corresponding to RNAP-promoter complex of interest. Using light box, superimpose cut-out film on gel, using autoradiography markers as reference points. Using disposable scalpel, excise portion of gel corresponding to RNAP-promoter complex. Transfer excised gel slice to 1.5-mL siliconized microcentrifuge tube.
3. Solubilize gel slice by adding 10 µl 1 M DTT (~0.4 M final) and heating 5 min at 37°C (*see* **Note 20**).
4. Immediately proceed to next step (**Subheading 3.4.6**).

3.4.6. Nuclease Digestion

1. During X-ray film exposure of **step 1**, **Section 3.4.5**, dilute DNaseI and micrococcal nuclease with ice-cold nuclease dilution solution to final concentration of 10 U/µl.
2. Transfer 10 µl to new 1.5-mL siliconized polypropylene microcentrifuge tube, and add 1 µl 200 mM $CaCl_2$, 1 µl 0.1 mM PMSF, 0.5 µl (5 U) micrococcal nuclease, and 0.5 µl (5 U) DNase I. Incubate 20 min at 37°C. Terminate reaction by adding 3 µl 5× SDS-loading buffer and heating 5 min at 100°C.
3. Immediately proceed to next step (**Subheading 3.4.7**).

3.4.7. Analysis

1. Apply entire sample (16 µl) to 7–15% gradient polyacrylamide (37.5:1 acrylamide:bisacrylamide) slab gel. As marker, load into adjacent lane 5 µl prestained protein molecular-weight markers. Electrophorese in SDS-running buffer at 25 V/cm until bromophenol blue reaches bottom of gel.
2. Dry gel, and autoradiograph or perform storage-phosphor imaging.
3. Identify bands corresponding to crosslinked proteins.

3.5. Kinetic Photocrosslinking

3.5.1. Rapid-Quench-Flow Mixing and Pulsed-Laser UV Irradiation [Steps 15–26 Carried Out Under Subdued Lighting (See Note 4)]

1. Start laser, and allow laser to operate for 1 h. Open 355-nm laser port, and check laser pulse energy using laser calorimeter and laser energy meter. Adjust laser pulse energy to 80 mJ/cm (a pulse energy that minimizes damage to optical windows of Raman cell and sample chamber). (*Caution*: Care must be exercised to avoid injury. All contact with the laser beam must be avoided. All work with laser must be performed by persons having laser-safety training and wearing appropriate protective gear, including appropriate laser safety goggles, laboratory coat, and gloves. Persons not having laser-safety training and/or not having appropriate protective gear must be excluded from the room. The room must be windowless or have fully

blocked windows. The entrance to the room must have a warning light that operates when the laser operates and must have a supply of appropriate laser goggles, for all wavelengths and all energies used.)

2. Connect circulating water bath to rapid-quench-flow system, and allow to operate at 37°C for 1 h
3. Start rapid-quench-flow-system controller, and allow to operate for 10 min.
4. Using a nail, puncture hole in lid of each of 16 1.5-mL siliconized polypropylene microcentrifuge tubes (in order to permit insertion of exit line rapid-quench-flow system through lid, allowing collection of sample without splashing).
5. Add 11 mL of DTT- and glycerol-free transcription buffer to a 50-mL polypropylene centrifuge tube, and equilibrate at 37°C using a water bath. Add and dissolve 1.1 mg BSA (yielding DTT- and glycerol-free, BSA-containing transcription buffer).
6. Add 5 mL of urea/NaI solution to a 50-mL polypropylene centrifuge tube. Add 176 µl 14.2 M β-mercaptoethanol (yielding quench solution; *see* **Note 23**). Equilibrate at 37°C using a water bath.
7. Set the left and right syringe-load valves of the rapid-quench-flow system to the load position. Using two Luer-Lok 5-mL disposable syringes, load the two drive syringes of the rapid-quench-flow system with DTT- and glycerol-free, BSA-containing transcription buffer. (To remove air bubbles from the drive syringes while loading, work the solution back and forth rapidly several times.)
8. Set the center syringe-load valve of the rapid-quench-flow system to the load position. Using a Luer-Lok 5-mL disposable syringe, load the third drive syringe of the quench-flow device with 5 mL of the quench solution.
9. Set the center syringe-load valve of the rapid-quench-flow system to connect the loading and firing lines. Fill Luer-Lok 50-µl Hamilton syringe with the quench solution; attach syringe to loading port of the third drive syringe of the rapid-quench-flow system; and flush connection lines, flow cell, and exit line.
10. Set all valves of the rapid-quench-flow system to the fire position. Set the rapid-quench-flow-system drive plate to the ready position.
11. Set the sample-load valves of the rapid-quench-flow system to the flush position. Flush sample loops, flow cell, and exit line with 1 mL of methanol and then with 1 mL of water. Dry flushed areas for 1 min using compressed argon.

12. Using the Linagraph direct-print paper, verify that the laser beam passes through the optical windows of the sample chamber. Adjust optics, if necessary. Using the laser calorimeter and laser energy meter, quantify laser pulse energy at front optical window of the sample chamber. If necessary, adjust laser pulse energy to 1 mJ.
13. Establish electrical connections between the digital delay/pulse generator and the power supplies of the laser and the rapid-quench-flow system.
14. Prepare 40 nM solution of RNAP in 170 μl of DTT- and glycerol-free, BSA-containing transcription buffer (*see* **Note 24**). Fill a 1-mL tuberculin-slip-tip disposable syringe with the resulting solution.
15. Prepare 1 nM solution of derivatized promoter DNA fragment in 170 μl of DTT- and glycerol-free, BSA-containing transcription buffer (*see* **Note 24**). Fill a 1-mL tuberculin-slip-tip disposable syringe with the resulting solution.
16. Set the sample-load valves of the rapid-quench-flow system to the sample-load position. Attach syringes with RNAP and DNA solutions to the sample-load lines. Fill each load line to the tubing connector of the sample-load valve.
17. Program the rapid-quench-flow-system controller, entering the following settings: preflash volume, 40 μl; postflash volume, 240 μl; delay time after flash, 0; and preflash reaction time, x_1 (where x_1 is the desired time, in seconds, between rapid-quench-flow mixing and pulsed-laser UV irradiation).
18. Program the digital delay/pulse generator, entering the following setting: delay time, x_1 (where x_1 is the desired time, in seconds, between rapid-quench-flow mixing and pulsed-laser UV irradiation).
19. Perform test firing. Press "G" button on the rapid-quench-flow-system controller to enable system. Press "fire" button on the laser controller to fire.
20. Flush and dry the sample loops, the flow cell, and the exit line as in **step 11**.
21. Using the 50-μl Hamilton syringe connected to the loading port of the third drive syringe in **step 9**, draw 20 μl of quenching solution out of the flow-cell connection tube.
22. Turn the sample-load valve to the sample-load position, and load the RNAP and DNA samples, one at a time, into the sample loops, bringing the meniscus of each sample to the 10-μl mark on each sample loop. (Sample loops should be precalibrated and premarked.)
23. Insert the exit line through the hole in the lid of a 1.5-mL siliconized polypropylene microcentrifuge tube from **step 4**.

24. Fire system (as described for test firing in **step 19**). Remove the collection tube with the expelled sample (~140 µl).
25. Repeat **steps 16–18** and **20–23** to perform 16 separate experiments, each with a different preflash reaction time, x_2 through x_{16}, on the rapid-quench-flow system and delay time, x_2 through x_{16}, on the delay/pulse generator (*see* **Note 25**).
26. Immediately proceed to the next step (**Subheading 3.5.2**).

3.5.2. Nuclease Digestion

1. Dilute DNase I, and micrococcal nuclease with ice-cold nuclease dilution solution to a final concentration of 50 U/µl.
2. Dilute OmniCleave endonuclease with ice-cold dilution buffer to a final concentration of 50 U/µl.
3. Prepare digestion mixture, containing: 20 µl of 200 mM $CaCl_2$, 20 µl of 200 mM $MgCl_2$, 4 µl of the solution from **step 1**, and 4 µl of the solution from **step 2**.
4. Transfer 34 µl of each sample from **Subheading 3.5.1** to a new 1.5-mL siliconized polypropylene microcentrifuge tube, and add 2 µl of the digestion mixture from **step 2**. Incubate for 15 min at 37°C. Terminate reaction by adding 9 µl of 5× SDS-loading buffer and heating for 5 min at 100°C.
5. Immediately proceed to the next step (**Subheading 3.5.3**).

3.5.3. Analysis

1. Apply the entire sample (45 µl) to a Criterion precast 4–20% Tris–HCl gradient slab gel. As a marker, load 5 µl of Precision-Plus protein dual-color standards into the adjacent lane. Electrophorese in SDS-running buffer at 25 V/cm until the bromophenol blue is ~0.5 cm from the bottom of the gel.
2. Dry gel, and perform storage-phosphor imaging.
3. Identify bands corresponding to crosslinked proteins.
4. Quantify intensities of bands corresponding to crosslinked proteins.
5. Plot intensities vs. time, using data from at least three independent determinations.
6. Perform curve fitting to extract kinetic parameters as follows: For crosslinked species for which intensities exhibit a monoexponential rise to a maximum, fit data to:

$$y = A[1 - \exp(-k_{obs}t)] + C_0,$$

where y is intensity, t is the time in seconds between rapid-quench-flow mixing and pulsed-laser UV irradiation DNA and RNAP, k_{obs} is the pseudo-first-order rate constant, C_0 is the intensity at $t = 0$, and A is the intensity at $t = \infty$.

For crosslinked species for which intensities exhibit a biexponential rise to a maximum, fit data to:

$$y = A_1[1 - \exp(-k_{1,\text{obs}}t)] + A_2[1 - \exp(-k_{2,\text{obs}}t)] + C_0,$$

where y is intensity, t is the time in seconds between rapid-quench-flow mixing and pulsed-laser UV irradiation DNA and RNAP, $k_{1,\text{obs}}$ and $k_{2,\text{obs}}$ are pseudo-first-order rate constants for processes1 and 2, and where A_1 and A_2 are the intensities at $t = \infty$ of processes 1 and 2.

4. Notes

1. M13mp2 *(ICAP-UV5)* carries the *lacP(ICAP-UV5)* promoter, a derivative of the *lacP* promoter having a consensus DNA site for CAP and a consensus –10 element *(5)*. M13mp2 *(ICAP-UV5)*-rev carries the *lacP(ICAP-UV5)* promoter in an orientation opposite to that in M13 mp2 *(ICAP-UV5) (5)*. M13mp2 *(ICAP-UV5)*-rev was prepared from M13mp2 *(ICAP-UV5)* by deleting the PvuII-PvuII segment corresponding to positions –217 to –125 of *lacP(ICAP-UV5)* and inverting the PvuII-PvuII segment corresponding to positions –124 to +145 of *lacP(ICAP-UV5)*.
2. M13mp2 *(ICAP-UV5)* and M13mp2 *(ICAP-UV5)*-revss DNAs carry the nontemplate and template strands of *lacP(ICAP-UV5)*, respectively. M13mp2 *(ICAP-UV5)* and M13mp2 *(ICAP-UV5)*-rev ssDNAs are prepared as in *(10)*.
3. The specified 10× digestion buffer is for PvuII and HaeIII. Use 10× digestion buffer recommended by supplier – omitting DTT; *see (40)* for other restriction enzymes.
4. Fluorescent light and daylight must be excluded. Low to moderate levels of incandescent light (e.g., from a single task lamp with 60-W tungsten bulb) are acceptable.
5. The derivatized oligodeoxyribonucleotide tolerates exposure to the Hitachi L-3,000 diode-array HPLC UV detector. The derivatized oligodeoxyribonucleotide can be identified unambiguously by monitoring the UV-absorbance spectrum from 200 to 350 nm. The derivatized oligodeoxyribonucleotide exhibits an absorbance peak at 260 nm, attributable to DNA, and a shoulder at 300–310 nm, attributable to the azidophenacyl group.
6. The derivatization procedure yields two diastereomers in an approximately 1:1 ratio: one in which azidophenacyl is

incorporated at the sulfur atom corresponding to the phosphate O1P, and one in which azidophenacyl is incorporated at the sulfur atom corresponding to the phosphate O2P (*see* ref. *7, 8)*. Depending on oligodeoxyribonucleotide sequence and HPLC conditions, the two diastereomers may elute as a single peak or as two peaks (e.g., at 24 and 25% solution B). In most cases, no effort is made to resolve the two diastereomers, and experiments are performed using the unresolved diastereomeric mixture. This permits simultaneous probing of protein–DNA interactions in the DNA minor groove (probed by the O1P-derivatized diastereomer) and the DNA major groove (probed by the O2P-derivatized diastereomer).

7. Phenyl azides are unstable at temperatures above 70°C. Avoid heating above 70°C.
8. HaeIII digestion, which yields a DNA fragment corresponding to positions –141 to +63 of *lacP(ICAP-UV5)*, is used for preparation of DNA fragments derivatized between positions –80 and –1, inclusive. *Pvu*II digestion, which yields a DNA fragment corresponding to positions –124 to +145 of *lacP(ICAP-UV5)*, is used for preparation of DNA fragments derivatized between positions +1 and +80 inclusive. [Use of DNA fragments with >60 bp between the site of derivatization and the nearest DNA-fragment end eliminates "nonspecific" crosslinking from the subpopulation of complexes having RNAP bound at a DNA-fragment end *(41)* rather than at the promoter.]
9. The procedures described are for "in-gel" static photocrosslinking, in which complexes are prepared in solution, isolated by nondenaturing polyacrylamide gel electrophoresis, and UV irradiated *in situ* in the polyacrylamide gel matrix *(2, 5, 6)*. In-gel procedures facilitate preparation, isolation, and analysis of homogeneous complexes. For alternative procedures, involving "on-bead" static photocrosslinking, *see (4)*. For alternative procedures, involving "in-solution" static photocrosslinking, *see (1, 5)*.
10. BAC is a disulfide-containing analog of bisacrylamide *(42–44)*. Polyacrylamide:BAC gels can be solubilized by addition of reducing agents *(42–44)*. The synthesis of BAC in this chapter is adapted from *(42)*.
11. Acryloyl chloride reacts violently with water. Add acryloyl chloride in 0.5-mL portions, waiting 30 s between successive additions.
12. BAC is substituted for bisacrylamide on a mole-equivalent, not mass-equivalent, basis *(42–44)*. The solubility of BAC in water is increased by adding acrylamide before adding BAC and by performing additions at 60°C.

13. TEMED and ammonium persulfate concentrations are critical variables in preparation of polyacrylamide:BAC gels *(43, 44)*. (Use of nonoptimal TEMED and ammonium persulfate concentrations in preparation of polyacrylamide:BAC results in difficulties in subsequently solubilizing gels.)
14. Siliconize notched glass plate by applying 30 μl SurfaSil siliconizing agent and spreading evenly with Kimwipe.
15. Heating during polymerization yields polyacrylamide:BAC gels that are maximally solubilizable upon addition of reducing agents *(43, 44)*. Heat the glass plates of the gel assembly evenly. (If necessary, use two task lamps.) Avoid heating above 70°C, as this can result in formation of bubbles and/or detachment of gels from the glass plates.
16. Do not add loading buffer to the reaction mixture. The reaction mixture is sufficiently dense for loading (due to the presence of glycerol).
17. UV irradiation is performed with both glass plates in place. The glass plates exclude wavelengths <300 nm, minimizing photodamage to protein and DNA. It is important to verify that the plates exhibit absorbances of 1.5 AU at 320 nm (e.g., by sacrificing a glass plate and placing a piece in the cuvette holder of a UV/Vis spectrophotometer). Glass plates purchased from Aladin, Inc. (Aladin Enterprises, Inc., 1,255 23rd Avenue, San Francisco CA 94,122 USA) have performed satisfactorily.
18. For in-gel UV irradiation of RNAP-promoter intermediate complex, prechill photochemical reactor 15 min in 15°C cabinet.
19. Do not use tight-fitting X-ray autoradiography cassettes, which can squeeze and distort the gel on the glass plate during exposure. The Kodak X-ray exposure holder with intensifying screen has performed satisfactorily.
20. 2–4 M β-mercaptoethanol can be substituted for 1 M DTT.
21. For rapid mixing, we use a commercial microprocessor-controlled, stepping-motor-driven rapid-quench-flow system modified to incorporate a 20 μl sample chamber with OP1 acrylic optical windows (KinTek RQF-3 with modifications; *see (45, 46)* and **Fig. 2**). The dead time for mixing is ~10 ms.
22. For laser UV irradiation, we use a nanosecond pulsed Nd-YAG laser with third harmonic generator (Continuum Powerlite 7,010-DS/TS/QS), a Raman cell (Light-Age RC-1), a Pellin-Broca prism, 355-nm laser optics (first mirror), and 309-nm laser optics (all other optics) (*see (47, 48)* and **Fig. 2**). This system is able to generate up to 5 mJ per 6 ns pulse at 309 nm (a wavelength selected to be close to the absorbance maximum of the azidophenacyl photoactivatible

crosslinking agent and well separated from the absorbance maxima of DNA and protein).

23. Laser UV irradiation of phenyl-azide photoactivatible crosslinking agents reportedly produces only short-lived reactive species (lifetimes ~5 ms in heptane; lifetimes <<1 ms in the presence of nucleophiles; *(49–51)*). Nevertheless, to ensure effectively immediate inactivation of reactive species and effectively immediate termination of crosslinking, we quench reactions immediately after laser UV irradiation by mixing with a quench solution of 5 M urea, 0.5 M NaI, and 0.5 M β-mercaptoethanol. The dead time for quenching is ~10 ms.
24. Use of a 40-fold mole excess of RNAP over promoter DNA, ensures that kinetics of RNAP-DNA association are pseudo-first order.
25. Each set of 16 experiments requires ~1 h. When performing successive sets of experiments, freshly made solutions should be prepared for each set of experiments.

Acknowledgments

The basic protocol for preparation of derivatized DNA fragments was developed by T. Lagrange, the basic protocol for preparation of RNAP was developed by H. Tang and K. Severinov, and the basic protocol for in-gel static photocrosslinking was developed by T.-K. Kim. We thank K. Severinov for plasmids; T.-K. Kim, T. Lagrange, D. Reinberg, and K. Severinov for discussion; and a Howard Hughes Medical Institute Investigatorship; and National Institutes of Health grant GM41376 to R.H.E. for financial support.

References

1. Lagrange, T., Kim, T.K., Orphanides, G., Ebright, Y., Ebright, R.H., and Reinberg, D. (1996). High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex. *Proc. Natl. Acad. Sci. U S A* **93**, 10620–10625.
2. Kim, T.-K., Lagrange, T., Wang, Y.-H., Griffith, J., Reinberg, D., and Ebright, R.H. (1997). Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proc. Natl. Acad. Sci. U S A* **94**, 12268–12273.
3. Lagrange, T., Kapanidis, A., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* **12**, 34–44.
4. Kim, T.-K., Ebright, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**, 1418–1421.
5. Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R.H. (2000). Structural organization

of the RNA polymerase-promoter open complex. *Cell* **101**, 601–611.

6. Kim, T.-K., Lagrange, T., Naryshkin, N., Reinberg, D., and Ebright, R.H. (2000). Site-specific protein-DNA photocrosslinking. In *Protein-DNA Interactions: A Practical Approach*, Travers, A. and Buckle, M. eds. IRL, Oxford, pp. 319–335.
7. Fidanza, J., Ozaki, H., and McLaughlin, L. (1992). Site-specific labeling of DNA sequences containing phosphorothioate diesters. *J. Am. Chem. Soc.* **114**, 5509–5517.
8. Yang, S.-W. and Nash, H. (1994). Specific photocrosslinking of DNA-protein complexes: identification of contacts between integration host factor and its target DNA. *Proc. Natl. Acad. Sci. U S A* **91**, 12183–12187.
9. Mayer, A. and Barany, F. (1995). Photoaffinity cross-linking of TaqI restriction endonuclease using an aryl azide linked to the phosphate backbone. *Gene* 153, 1–8.
10. Sambrook, J. and Russell, D. (2001). *Molecular Cloning: A Laboratory* . Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
11. Wang, Y. and Stumph, W. (1998). Identification and topological arrangement of Drosophila proximal sequence element (PSE)-binding protein subunits that contact the PSEs of U1 and U6 small nuclear RNA genes. *Mol. Cell. Biol.* **18**, 1570–1579.
12. Bartlett, M., Thomm, M., and Geiduschek, E. (2000). The orientation of DNA in an archaeal transcription initiation complex. *Nat. Struct. Biol.* 7, 782–785.
13. Bartlett, M., Thomm, M., and Geiduschek, E. (2004). Topography of the euryarchaeal transcription initiation complex. *J. Biol. Chem.* 279, 5894–5903.
14. Renfrow, M., Naryshkin, N., Lewis, M., Chen, H.-T., Ebright, R.H., and Scott, R. (2004). Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. *J. Biol. Chem.* 279, 2825–2831.
15. Chen, B., Mandal, S., and Hampsey, M. (2004). High-resolution protein-DNA contacts for the yeast RNA polymerase II general transcription machinery. *Biochemistry* 43, 12741–12749.
16. Kim, Y., Geiger, J., Hahn, S., and Sigler, P. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365, 512–520.
17. Kim, J., Nikolov, D., and Burley, S. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365, 520–527.
18. Geiger, J., Hahn, S., Lee, S., and Sigler, P. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* 272, 830–836.
19. Tan, S., Hunziker, Y., Sargent, D., and Richmond, T. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381**, 127–134.
20. Nikolov, D., Chen, H., Halay, E., Usheva, A., Hisatake, K., Lee, D.K., Roeder, R., and Burley, S. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**, 119–128.
21. Bartholomew, B., Kassavetis, G., Braun, B., and Geiduschek, E.P. (1990). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *EMBO J.* **9**, 2197–2205.
22. Bartholomew, B., Kassavetis, G. and Geiduschek, E.P. (1991). Two components of *Saccharomyces cerevisiae* transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. *Mol. Cell. Biol.* **11**, 5181–5189.
23. Braun, B., Bartholomew, B., Kassavetis, G., and Geiduschek, E.P. (1992). Topography of transcription factor complexes on the *Saccharomyces cerevisiae* 5S RNA gene. *J. Mol. Biol.* **228**, 1063–1077.
24. Kassavetis, G., Kumar, A., Ramirez, E., and Geiduschek, E.P. (1998). Functional and structural organization of Brf, the TFIIB-related component of the RNA polymerase III transcription initiation complex. *Mol. Cell. Biol.* **18**, 5587–5599.
25. Bell, S. and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134.
26. Coulombe, B., Li, J., and Greenblatt, J. (1994). Topological localization of the human transcription factors IIA, IIB, TATA box-binding protein, and RNA polymerase II-associated protein 30 on a class II promoter. *J. Biol. Chem.* **269**, 19962–19967.
27. Gong, X., Radebaugh, C., Geiss, G., Simon, S., and Paule, M. (1995). Site-directed photocross-linking of rRNA transcription initiation complexes. *Mol. Cell. Biol.* **15**, 4956–4963.
28. Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E., and Wolffe, A. (1996). An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres. *Science* **274**, 614–617.
29. Robert, F., Forget, D., Li, J., Greenblatt, J., and Coulombe, B. (1996). Localization of

subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter. *J. Biol. Chem.* **271**, 8517–8520.

30. Forget, D., Robert, F., Grondin, G., Burton, Z., Greenblatt, J., and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. *Proc. Natl. Acad. Sci. U S A* **94**, 7150–7155.
31. Record, M.T., Reznikoff, W., Craig, M., McQuade, K., and Schlax, P. (1996). *Escherichia coli* RNA polymerase (Eσ70), promoters, and the kinetics of the steps of transcription initiation. In *Escherichia coli and Salmonella*, Vol. 1, Neidhart, F., ed. ASM, Washington, DC, pp. 792–820.
32. Ebright, R. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J. Mol. Biol.* **304**, 687–698.
33. Severinov, K., Mustaev, A., Severinova, E., Bass, I., Kashlev, M., Landick, R., Nikiforov, V., Goldfarb, A., and Darst, S. (1995). Assembly of functional *Escherichia coli* RNA polymerase containing β subunit fragments. *Proc. Natl. Acad. Sci. U S A* **92**, 4591–4595.
34. Severinov, K., Mustaev, A., Kukarin, A., Muzzin, O., Bass, I., Darst, S., and Goldfarb, A. (1996). Structural modules of the large subunits of RNA polymerase. Introducing archaebacterial and chloroplast split sites in the β and β′ subunits of Escherichia coli RNA polymerase. *J. Biol. Chem.* **271**, 27969–27974.
35. Tang, H., Severinov, K., Goldfarb, A., and Ebright, R.H. (1995). Rapid RNA polymerase genetics: one-day, no-column preparation of reconstituted recombinant Escherichia coli RNA polymerase. *Proc. Natl. Acad. Sci. U S A* 92, 4902–4906.
36. Tang, H., Kim, Y., Severinov, K., Goldfarb, A., and Ebright, R.H. (1996). *Escherichia coli* RNA polymerase holoenzyme: rapid reconstitution from recombinant α, β, β′, and σ subunits. *Methods Enzymol.* 273, 130–134.
37. Martin, E., Sagitov, V., Burova, E., Nikiforov, V., and Goldfarb, A. (1992). Genetic dissection of the transcription cycle. A mutant RNA polymerase that cannot hold onto a promoter. *J. Biol. Chem.* **267**, 20175–20180.
38. Zalenskaya, K., Lee, J., Chandrasekhar, N.G., Shin, Y.K., Slutsky, M., and Goldfarb, A. (1990). Recombinant RNA polymerase: inducible overexpression, purification and assembly of Escherichia coli rpo gene products. *Gene* **89**, 7–12.
39. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A., Niu, W., Ebright, Y., Levy, R., and Ebright, R. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**, 599–614.
40. Staros, J., Bayley, H., Standring, D., and Knowles, J. (1978). Reduction of aryl azides by thiols: implications for the use of photoaffinity reagents. *Biochem. Biophys. Res. Commun.* **80**, 568–572.
41. Melancon, P., Burgess, R., and Record, M. (1983). Direct evidence for the preferential binding of *Escherichia coli* RNA polymerase holoenzyme to the ends of deoxyribonucleic acid restriction fragments. *Biochemistry* **22**, 5169–5176.
42. Hansen, J.N. (1976). Electrophoresis of ribonucleic acid on a polyacrylamide gel which contains disulfide cross-linkages. *Anal. Biochem.* **76**, 37–44.
43. Hansen, J.N. (1980). Chemical and electrophoretic properties of solubilizable disulfide gels. *Anal. Biochem.* **105**, 192–201.
44. Hansen, J.N. (1981). Use of solubilizable acrylamide disulfide gels for isolation of DNA fragments suitable for sequence analysis. *Anal. Biochem.* **116**, 146–151.
45. Langowski, J., Urbanke, C., and Scuppe, E. (1984). Construction of a microprocessor-controlled pulsed quench-flow apparatus for the study of fast chemical and biochemical reactions. *Anal. Biochem.* **142**, 91–97.
46. Johnson, K. (1995). Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. *Methods Enzymol.* **249**, 38–61.
47. Hockensmith, J., Kubasek, W., Vorachek, W., and von Hippel, P. (1986). Laser cross-linking of nucleic acids to proteins: methodology and first application to the phage T4 DNA replication system. *J. Biol. Chem.* **261**, 3512–3518.
48. Hockensmith, J., Kubasek, W., Vorachek, W., Evertsz, E., and von Hippel, P. (1991). Laser cross-linking of protein-nucleic acid complexes. *Methods Enzymol.* **208**, 211–236.
49. Shields, C., Chrisope, D., Schuster, G., Dixon, A., Poliakoff, M., and Turner, J. (1987). Photochemistry of aryl azides: detection and characterization of a dehydroazepine by time-resolved infrared spectroscopy and flash photolysis at room temperature. *J. Am. Chem. Soc.* **109**, 4723–4726.
50. Li, Y.-Z., Kirby, J., George, M., Poliakoff, M., and Schuster, G. (1988). 1,2-Didehy-

droazepines from the photolysis of substituted aryl azides: analysis of their chemical and physical properties by time-resolved spectroscopic methods. *J. Am. Chem. Soc.* **110**, 8092–8098.

51. Marcinek, A., Leyva, E., Whitt, D., and Platz, M. (1993). Evidence for stepwise nitrogen extrusion and ring expansion upon photolysis of phenyl azide. *J. Am. Chem. Soc.* 115, 8609–8612.

Chapter 26

Use of Site-Specific Protein–DNA Photocrosslinking of Purified Complexes to Analyze the Topology of the RNA Polymerase II Transcription Initiation Complex

Diane Forget, Céline Domecq, and Benoit Coulombe

Summary

A method for the photocrosslinking of proteins to DNA in purified complexes is described. It makes use of the juxtaposition of a limited number of photoreactive nucleotides with a limited number of radiolabeled nucleotides at a specific location in a DNA fragment. Protein–DNA complexes are submitted to an electrophoretic mobility shift assay that is then irradiated with UV light in order to crosslink the proteins to DNA. The specific complexes are localized on the gel, purified, and processed for the identification of the crosslinked polypeptides.

Key words: Photocrosslinking, Protein–DNA complexes, Photoreactive nucleotides, Multisubunit complexes.

1. Introduction

Site-specific protein–DNA photocrosslinking has proved to be the method of choice to analyze the molecular organization and topology of large nucleoprotein complexes such as those involved in the transcription reaction by mammalian RNA polymerase II (RNAPII). A principal advantage of the method is that it yields low-resolution structural information on large multisubunit complexes.

The initiation of mRNA synthesis requires the formation of a preinitiation complex containing RNAPII and the general initiation factors TFIID (or TBP), TFIIB, TFIIE, TFIIF, and TFIIH

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_26

on promoter DNA *(1)*. Because many general initiation factors and RNAPII are multisubunit proteins, the preinitiation complex can comprise up to 50 polypeptides. Neither X-ray crystallography nor NMR, which can resolve the structure of complexes containing protein bound to small pieces of DNA, could provide any detailed structural information on this complex. This method has also the advantage of presenting sufficient flexibility to allow the rapid analysis of complexes assembled under various conditions. In the past, our laboratory has used crosslinking probes carrying a photoreactive nucleotide at specific positions along promoter DNA to localize the components of the transcription machinery in the preinitation complex *(2–7)*. Nevertheless, a possible limitation of protein–DNA photocrosslinking relates to the specificity of preinitiation complex assembly. Because most general initiation factors and RNAPII have an affinity for any DNA (although lower than that for promoter DNA), it is often difficult to set up conditions that will systematically and exclusively allow for the formation of specific complexes on our various photoprobes. In order to circumvent this problem, we developed a method for crosslinking proteins to DNA in purified complexes. The overall procedure is summarized in **Fig. 1**. Briefly, complexes are first assembled by mixing transcription factors with a photoprobe that juxtaposes one (or a few) photoreactive nucleotide(s) with one (or a few) radiolabeled nucleotide(s) at a specific location in the promoter. The complexes are submitted to an electrophoretic mobility shift assay (EMSA) in a native gel that is then irradiated with UV light in order to crosslink the proteins to DNA. The specific complexes are then localized on the gel, purified by cutting out the gel slices, and processed for the identification of the crosslinked polypeptides. Because this procedure helps to reduce to a minimum the nonspecific crosslinking signals due to aggregation, it has allowed us to define with more precision the topological organization of the RNAPII preinitiation complex *(8)*.

2. Materials

1. Buffer A (10×): 300 mM Tris–HCl of pH 8.0, 500 mM KCl, and 70 mM $MgCl_2$, freshly prepared.
2. Bovine serum albumin (BSA) solution: Prepare a 25 mg/mL solution of BSA in deionized distilled water. Store in aliquots at 20°C. Dilute with water to 10 mg/mL prior to use.
3. dNTP mix: 10 mM each dATP, dCTP, dGTP, and dTTP in 1× buffer A, freshly prepared.

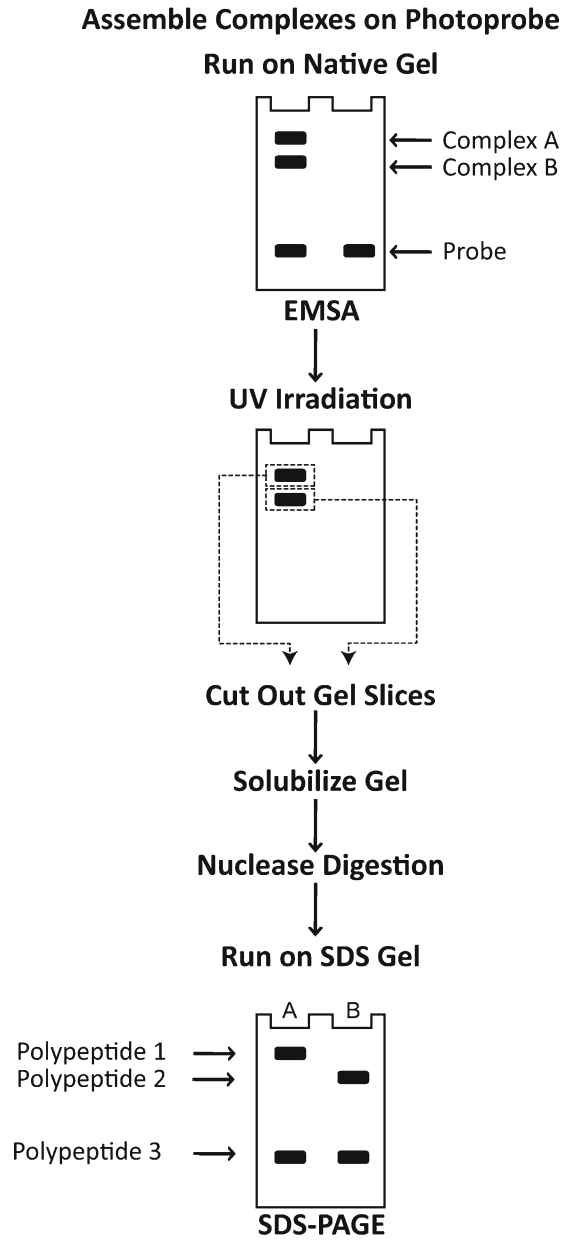


Fig. 1. Overall representation of the procedure for in-gel site-specific protein–DNA photocrosslinking.

4. 6× gel-loading solution: 0.25% bromophenol blue, 0.25% xylene cyanol, and 30% glycerol in deionized distilled water.
5. 1× TBE buffer: The TBE buffer is prepared as a 5× stock by mixing 54 g Tris-base, 27.5 g boric acid, and 20 mL EDTA 0.5 M (pH 8.0) in 1 L of deionized distilled water.

6. Buffer G: 12 mM Hepes of pH 7.9, 60 mM KCl, 0.12 mM EDTA, 8 mM $MgCl_2$, 50 ng/mL BSA, 5 mM β-mercaptoethanol, and 12% glycerol.
7. Poly(dIdC.dIdC) stock: Prepare a 25 mg/mL solution of poly(dIdC.dIdC) in deionized distilled water. Store in aliquots at 20°C.
8. Buffer ND: 20 mM Hepes of pH 7.9, 100 mM KCl, 0.2 mM EDTA, 0.2 mM EGTA, 0.4 mM DTT, and 20% glycerol.
9. DNase mix: A solution containing 0.5 U/μL DNase I and 30 mM $CaCl_2$.
10. Acid mix: Mix equal volume of 5% acetic acid and 30 mM $ZnCl_2$, freshly prepared.
11. 5× Loading buffer: 80 mM Tris–HCL of pH 6.8, 12.5% glycerol, 2.5% SDS, 0.9 M β-mercaptoethanol, 0.2% bromophenol blue.

3. Methods

3.1. Day 1: Preparation of the Photoprobes

The first step of the procedure is the synthesis of the photoprobes. A schematic representation is illustrated in **Fig. 2**. In this example, one photoreactive nucleotide is placed at position +1 and three radiolabeled guanosines at positions –1, –3, and –4 of the adenovirus major late promoter.

The site-specific incorporation of the photoreactive nucleotide (*see* **Note 1**) and the radiolabeled nucleotide is directed through the annealing of a primer, named the specific primer, with a single-stranded DNA template containing the promoter DNA. The promoter is flanked by two restriction sites (in this example *Sma*I). A second primer, named the upstream primer, is annealed a few base pairs upstream of the *Sma*I site. After annealing, the photoreactive and radiolabeled nucleotides are incorporated by primer extension using T_4 DNA polymerase with limiting amounts of dNTPs (*see* **Note 2**). After the labeling step, the extension reaction is completed by the addition of an excess of cold dNTPs (*see* **Note 3**) and nicks at the 5′ end of the primers are repaired by the addition of T_4 DNA ligase. The photoprobe is generated by digestion with the restriction enzyme and gel purified (*see* **Fig. 3** for an example of a gel on which the products of a photoprobe synthesis reaction have been separated).

1. Mix 500 ng of single-stranded (ss) DNA (approximately 0.5 pmol) with 40 ng (approximately 5 pmol each) of both the specific and the upstream primers. Add 1 μL of 10× buffer A and complete to 10 μL with deionized distilled water.
2. Mix well and incubate for 5 min at 90°C.
3. Incubate for 30 min at room temperature.

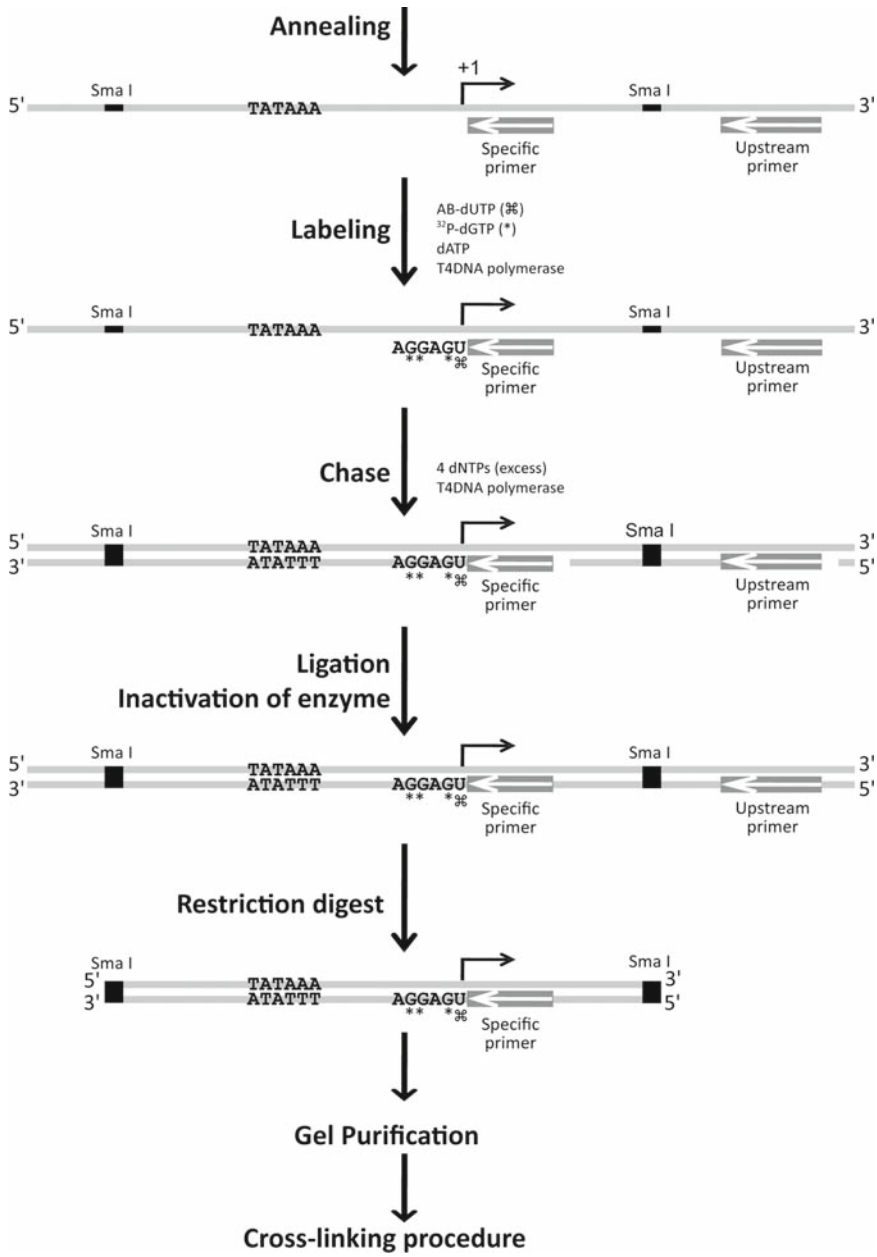


Fig. 2. Schematic design for the synthesis of the photoprobes. In this example, the photoprobe contains one photonucleotide (U) at position +1, and 3 radiolabeled guanosines (G) at positions -1, -3, and -4 of the adenovirus major late promoter.

4. From this point on, all manipulations must be carried out under reduced light conditions (*see* **Note 4**). Add 0.5 µL BSA (10 mg/mL), 1 µL AB-dUTP (*see* **Note 1**), 20 µCi of the appropriate (α^{32}P)-dNTP (3,000 Ci/mmol) (α^{32}P-dGTP for the example in **Fig. 2**), 5–10 U of T_4 DNA polymerase, and 1 µL 10× buffer A. Complete to a final volume of 20 µL with deionized distilled water.

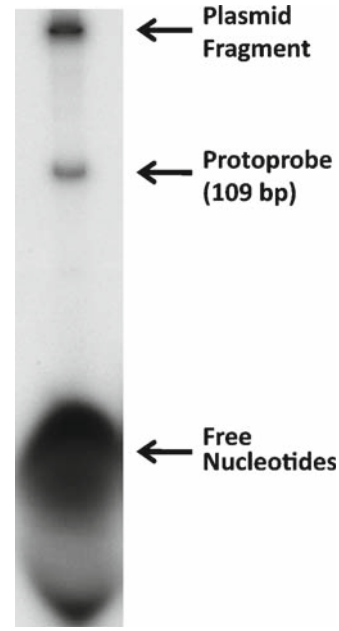


Fig. 3. An autoradiogram of a gel used for the purification of the photoprobes. The position of the DNA fragment carrying both the photoreactive and radiolabeled nucleotides at a specific location is shown

5. Incubate for 30 min at room temperature.
6. Add 5 µL of dNTP mix.
7. Incubating for 5 min at room temperature.
8. Incubate for 20 min at 37°C.
9. Add 5 U of T_4 DNA ligase and ATP (1 mM final concentration).
10. Incubate for 1 h at room temperature.
11. Incubate at 65°C for 20 min in order to inactivate the ligase.
12. Add 10–20 U of restriction enzyme (*Sma*I in the example shown in **Fig. 2**).
13. Incubate for 90 min at the temperature recommended by the supplier of restriction enzymes (25°C for *Sma*I).
14. Add 5 µL of a 6× gel-loading solution.
15. Load on a native 8% polyacrylamide gel (40:1 acrylamide:bis) in 1× TBE buffer.
16. Run at 150 V for about 1 h in 1× TBE buffer.
17. Remove the glass plates containing the gel from the gel box.
18. Separate the glass plates and leave the gel on one of them.
19. Wrap the gel/glass plate in plastic wrap.

20. Wrap the entire package in an aluminum foil.
21. Move to a dark room (*see* **Note 5**).
22. Place a Kodax X-OMAT AR film on a clean bench.
23. Remove the foil and place the gel on the film with the glass plate facing up (i.e., gel side down).
24. Expose 2–5 min.
25. During the exposition time, mark the film using a sharp tool by tracing the contour of the glass plate (this will be helpful later for the localization of the photoprobe in the gel).
26. Remove the gel and rewrap it with the foil.
27. Develop the film (*see* **Note 6**).
28. Using a scalpel, cut the film so that the square piece containing the band corresponding to the photoprobe is removed. This operation leaves the film with a window at the position of the photoprobe.
29. Superimpose the film on the gel by taking advantage of the marks made in **step 25**, and mark the square corresponding to the photoprobe on the plastic wrap using a pen.
30. Cut out the gel slice containing the photoprobe using a clean scalpel.
31. Cut the gel slice in small pieces (six to eight fragments).
32. Place the gel fragments in a 1.5-mL microcentrifuge tube and add 10 mM Tris–HCL (pH 7.9) in order to completely submerge the gel (about 125–200 μL).
33. Incubate overnight at room temperature.
34. Collect the liquid containing the probe.
35. Purify the probe on a Micro-spin S-200 HR column (Amersham) to remove any salts and other putative contaminants.
36. Count 1 μL aliquot of the photoprobe solution by liquid scintillation, and dilute the probe to the appropriate count number with deionized distilled water.
37. The probe is now ready for use and can be stored in the dark at 4°C for 1–2 weeks (*see* **Note 7**).

3.2. Day 2: In-Gel Protein–DNA Photo-Crosslinking

The gel-purified photoprobe is used for preinitiation complex assembly with the purified transcription factors (TBP, TFIIA, TFIIB, TFIIE, TFIIF, and TFIIH) and RNAPII. The protein–DNA complexes are submitted to an electrophoresis in a native polyacrylamide:bac gel (*see* **Notes 8** and **9**) in order to isolate specific complexes. The gel is immediately irradiated with UV light to crosslink the proteins to DNA. Specific complexes are identified by autoradiography. Examination of the autoradiogram permits identification of the preinitiation complexes that

Fig. 4. An autoradiogram of protein–DNA complexes electrophoresis in a native gel. The complexes were assembled with calf thymus RNAPII, TFIIB, TFIIF, and TFIIE in either the presence (+) or the absence (–) of TBP on photoprobe –39/–40. Two complexes, A and B, are resolved using electrophoretic mobility shift assay (EMSA).

assembled on photoprobes (*see* **Fig. 4**). For the RNAPII initiation complex, assembly of specific complexes on promoter in the native gels can be assessed by comparing reactions with photoprobes containing a wild-type or a mutated TATA box and/or by comparing reactions performed in either the presence or the absence of TBP (*see* **Note 10**). The band corresponding to each complex assembled on photoprobes is excised, solubilized, and treated with DNase I and S1 nuclease. Enzymatic treatments permit to liberate polypeptides that are covalently attached to a short piece of DNA carrying one to four radiolabeled nucleotides. After separation of the photocrosslinked polypeptides by SDS-PAGE, the gel is dried and exposed to X-ray film. Examination of the autoradiogram permits identification of the protein(s) that interact with a particular site. The photocrosslinked polypeptides can be identified according to their molecular weight (*see* **Fig. 5**).

1. Prechill the 4.5% polyacrylamide:bac gel and the 0.5× TBE reservoir containing 2 mM $MgCl_2$ by placing them in a 4°C cabinet for 3 h.

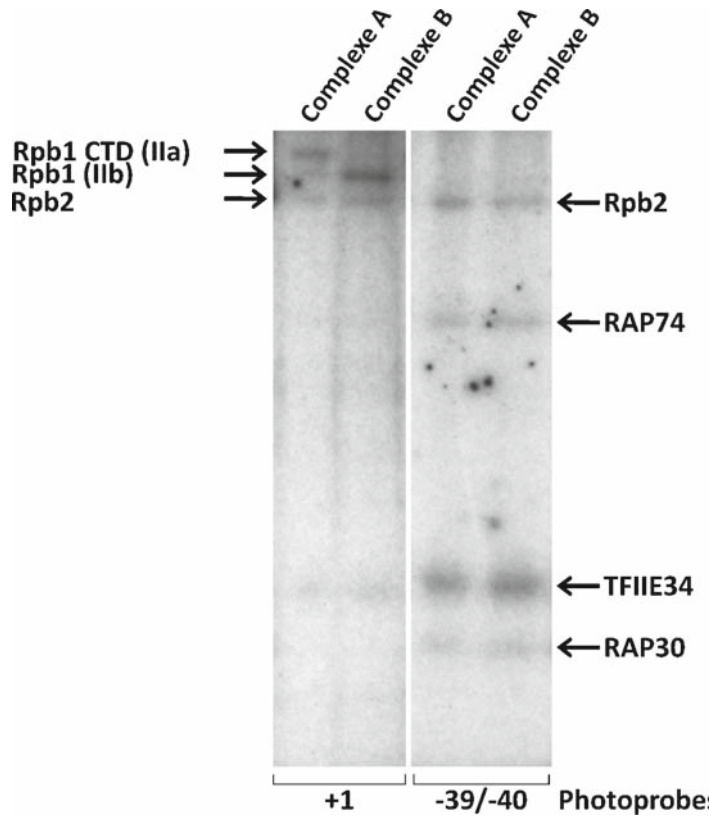


Fig. 5. An autoradiogram of an SDS-PAGE gel showing the crosslinked polypeptides in complexes A and B. Complexes A and B were assembled on photoprobes –39/–40 and +1. No difference is observed in the polypeptides that crosslink to photoprobe –39/–40, the form of RPB1 that crosslinked to position +1 varies when complex A is compared to complex B. The difference in the molecular weight of RPB1 indicates that the IIa form of RNAPII (e.g., with a hypophosphorylated CTD) is found in complex A, whereas the IIb form (e.g., without the CTD due to its proteolysis during purification) is present in complex B.

2. Mix the proteins (80–600 ng each) and complete the volume to 20 µL of buffer G (*see* **Note 11**). Add 1 µL of diluted poly(dIdC.dIdC) (*see* **Note 12**) and 6,000 cpm of the photoprobe. The final volume is 21 µL (*see* **Note 13**).
3. Mix well and incubate for 30 min at 30°C.
4. Load on a 4.5% acrylamide:bac gel in a 0.5× TBE reservoir containing 2 mM $MgCl_2$.
5. Run in a 4°C cabinet for 60 min at 400 V.
6. Remove one glass plate and irradiate the protein–DNA complexes in the gel for 10 min using with UV light (*see* **Notes 14** and **15**).
7. Cover the gel with Whatman paper.
8. Remove the other glass plate.

9. Cover the gel with a plastic wrap.
10. Expose the gel to a phosphorimager screen overnight and print the image.
11. Using a scalpel, cut the print so that the square piece containing the band corresponding to the photoprobe is removed. This operation leaves the film with a window at the position of the complex of interest.
12. Superimpose the print on the gel.
13. Cut out the gel slices containing the complex of interest using a clean scalpel.
14. Place the excised gel in a 1.5-mL microcentrifuge tube taking care to avoid carrying any pieces of Whatman paper and plastic wrap.
15. Add 10 µL of 1 M DTT to solubilize gel slices (*see* **Note 16**).
16. Incubate 10 min at 37°C.
17. Add 40 µL of ND buffer.
18. Incubate at 37°C for 25 min.
19. Add 13 µL of DNase mix.
20. Incubate at 37°C for 20 min.
21. Add 3.9 µL of 10% SDS.
22. Incubate at 95°C for 3 min.
23. Add 5.2 µL of acid mix.
24. Add 2.6 µL of S1 nuclease at 80,000 U/mL.
25. Incubate at 37°C for 20 min.
26. Add 15 µL of 5× loading buffer to stop the reaction.
27. Boil the samples for 5 min.
28. Resolve the photocrosslinked polypeptides by SDS-PAGE gel (run at 30 mA in the stacking gel and at 55 mA in the separating gel) (*see* **Note 17**).
29. Transfer the gel to Whatman paper and dry.
30. Expose the dried gel to X-ray film using an intensifying screen (*see* **Note 18**).

4. Notes

1. The nucleotide derivative we use, namely 5-(N-(p-azidobenzoyl)-3-aminoallyl)-dUTP (AB-dUTP or N_3R-dUTP) (*see* Chap. "Site-Directed DNA Cross-Linking of

Large Multi-Subunit Protein–DNA Complexes"), possesses a side chain that places a photoreactive nitrene in the major groove of the DNA helix 10 Å away from the DNA backbone (*see* **ref. 10***)*. For this reason the crosslinking of a polypeptide to the photoprobe does not require a direct interaction of the polypeptide with the DNA helix. The amount of AB-dUTP to be added to the reaction is determined empirically for each preparation of the photoreactive nucleotide and is generally between 0.5 and 2 μL (often 1 μL).

2. The specific primer must be designed in such a manner that T_4 DNA polymerase only adds a few nucleotides. In the example shown in **Fig. 2**, the incorporation during the site-specific labeling is restricted to positions −4 to +1 by omitting dCTP from the reaction. The success of this step can be monitored by analysis of the reaction products on a sequencing gel.
3. The addition of dNTP in large excess is crucial because it is necessary to limit the incorporation of radiolabeled and photoreactive nucleotides during the extension of the photoprobes.
4. The use of a standard dark room is not necessary. As a rule, we find that conditions providing just enough light to be able to work are acceptable.
5. A conventional red light can be used.
6. An example of the autoradiogram of a gel used for photoprobe purification is shown in **Fig. 3**. The position of the band corresponding to the photoprobe can be easily identified because the size of the fragment generated by the restriction enzyme is known.
7. Fresh probes (less than a week old) give the best results.
8. Bac is a disulfide-containing analog of bis-acrylamide (*see* **refs. 11–13**). Its chemical synthesis is essentially as described by Naryshkin et al. (*see* **ref. 14***)*. We have used *N*,*N*′ -Bis(acryloyl) cystamine from Sigma (product number A4929).
9. The polyacrylamide:bac gels for the EMSA were prepared as follows. A 20% acrylamide:bac (19:1) stock solution is prepared by dissolving 19 g of acrylamide and 1 g of bac in 80 mL of water in a 200-mL beaker and stirring for 30 min at 60°C (the solubility of bac in water is increased by adding the acrylamide before adding the bac and by performing the addition at 60°C). The volume is then adjusted to 100 mL with water and the solution allowed to cool down to room temperature prior to filtering through 0.22-μm filter unit and storing at 4°C in the dark (stable for a few weeks). The gel is assembled using one glass plate that has

been siliconized by applying 100 µl of Surfacil siliconizing agent and spreading evenly with a Kimwipe. A 4.5% polyacrylamide:bac gel in 0.5× TBE buffer is prepared by mixing 11.25 mL of an acrylamide:bac (19:1) stock solution, 5 mL of 5× TBE, 100 µL $MgCl_2$, and 33.7 mL water and preheating the slab gel assembly and the gel mixture at 50°C in an incubator for 30 min. Polymerization is initiated by adding 250 µl TEMED and 125 µl freshly prepared 10% ammonium persulfate, and the gel poured immediately into the slab gel. Allow 20 min for polymerization at 50°C (the TEMED and ammonium persulfate concentrations are critical variables in the preparation of polyacrylamide:bac gels). The polyacrylamide:bac is stable for up to 72 h at 4°C, but it is better to use it freshly

10. Because some of the general transcription factors and RNAPII bind nonspecifically to DNA, it is important to discriminate between specific and nonspecific assembly. Assembly of these complexes is promoter specific because both a mutation in the TATA box (TATAAA to TAGAGA; not shown) and the omission of TBP in the assembly mixture (**Fig. 4**, compare + and –) abolish the formation of the two complexes.
11. In the preinitiation complex assembly, we use recombinant human proteins in following quantities: 80 ng of TFIIB, 300 ng of RAP30, 600 ng of RAP74, 160 ng of TFIIE34, 380 ng of TFIIE56, 300 ng of calf thymus RNAPII, and 80 ng of recombinant yeast TBP. The amounts of the different protein factors should be optimized for each different combination of proteins and for each protein preparation.
12. The poly (dI.dC-dI.dC) stock should be diluted just prior to use. The exact dilution should be determined experimentally in order to favor specific vs. nonspecific signals without adversely affecting the intensity of the specific signals.
13. The EMSA were performed as described previously by Wolner & Gralla (*see* **ref. 15***)*
14. Irradiation time with UV light should be optimized by performing a time course with the particular system to be used. We use a Hoefer UVC 500 Ultraviolet Crosslinker with 254-nm bulbs.
15. From this point on, normal light conditions can be used.
16. 1 M DTT can be substituted by 2–4 M β-mercaptoethanol.
17. Detailed procedures for SDS-PAGE electrophoresis have been described (*see* **ref. 6***)*.
18. The use of BioMax (Kodak) screens is recommended.

References

1. Coulombe, B. and Burton, Z.F. (1999). DNA bending and wrapping around RNA polymerase: a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* **63**, 457–478.
2. Robert, F., Forget, D., Li, J., Greenblatt, J. and Coulombe, B. (1996). Localization of subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter. *J. Biol. Chem.* **271**, 8517–8520.
3. Forget, D., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J. and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. *Proc. Natl Acad. Sci. U S A.* **94**, 7150–7155.
4. Robert, F., Douziech, M., Forget, D., Egly, J.M., Greenblatt, J., Burton, Z.F. and Coulombe, B. (1998). Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell.* **2**, 341–351.
5. Douziech, M., Coin, F., Chipoulet, J.M., Arai, Y., Ohkuma, Y., Egly, J.M. and Coulombe, B. (2000). Mechanism of promoter melting by the *Xeroderma pigmentosum* complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-crosslinking. *Mol. Cell. Biol.* **20**, 8168–8177.
6. Langelier, M.F., Forget, D., Rojas, A., Porlier, Y., Burton, Z.F. and Coulombe, B. (2001). Structural and functional interactions of transcription factor (TF) IIA with TFIIE and TFIIF in transcription initiation by RNA polymerase II. *J. Biol. Chem.* **276**, 38652–38657.
7. Dion, V. and Coulombe, B. (2003). Interactions of a DNA-bound transcriptional activator with the TBP-TFIIA-TFIIB-promoter quaternary complex. *J. Biol. Chem.* **278**, 11495–11501.
8. Forget, D., Langelier, M.F., Thérien, C., Trinh, V. and Coulombe, B. (2004). Photocross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. *Mol. Cell. Biol.* **24**, 1122–1131.
9. Bartholomew, B., Kassavetis, G.A., Braun, B.R. and Geiduschek, E.P. (1990). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *EMBO J.* **9**, 2197–2205.
10. Hansen, J.N. (1976). Electrophoresis of ribonucleic acid on a polyacrylamide gel which contains disulfide cross-linkages. *Anal. Biochem.* **76**, 37–44.
11. Hansen, J.N., Pheiffer, B.H. and Boehnert, J.A. (1980). Chemical and electrophoretic properties of solubilizable disulfide gels. *Anal. Biochem.* **105**, 192–201.
12. Hansen, J.N. (1981). Use of solubilizable acrylamide disulfide gels for isolation of DNA fragments suitable for sequence analysis. *Anal. Biochem.* **116**, 146–151.
13. Naryshkin, N., Kim, Y., Dong, Q. and Ebright, R.H. (2001). Site-specific protein-DNA photocrosslinking. Analysis of bacterial transcription initiation complexes. *Methods Mol. Biol.* **148**, 337–361.
14. Wolner, B.S. and Gralla, J.D. (2000). Roles for non-TATA core promoter sequences in transcription and factor binding. *Mol. Cell. Biol.* **20**, 3608–3615.
15. Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual.* Cold spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.

Chapter 27

Site-Directed DNA Crosslinking of Large Multisubunit Protein–DNA Complexes

Jim Persinger and Blaine Bartholomew

Summary

Several methods have been developed to site-specifically incorporate photoreactive nucleotide analogs into DNA for the purpose of identifying the proteins and their domains that are in contact with particular regions of DNA. The synthesis of several deoxynucleotide analogs that have a photoreactive group tethered to the nucleotide base and the incorporation of these analogs into DNA are described. In a second approach, oligonucleotide with a photoreactive group attached to the phosphate backbone is chemically synthesized. The photoreactive oligonucleotide is then enzymatically incorporated into DNA by annealing it to a complementary DNA template and extending with DNA polymerase. Both approaches have been effectively used to map protein–DNA interactions in large multisubunit complexes such as the eukaryotic transcription or ATP-dependent chromatin remodeling complexes. Not only do these techniques map the binding sites of the various subunits in these complexes, but when coupled with peptide mapping also determine the protein domain that is in close proximity to the different DNA sites. The strength of these techniques is the ability to scan a large number of potential sites by making combinations of different DNA probes and is facilitated by using an immobilized DNA template for synthesis.

Key words: Chromatin, Nucleosomes, RNA polymerase, Transcription, DNA photoaffinity labeling, Protein–DNA interaction, DNA modification.

1. Introduction

Site-specific DNA photoaffinity labeling is a useful technique for mapping interactions of proteins with DNA in complex systems such as the RNA polymerase III (Pol III) transcription complex having at least 25 different proteins *(1–3)* and ATP-dependent chromatin remodeling complexes *(4)*. This technique allows

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_27

probing of protein–DNA interactions across large stretches of DNA and can be done in relatively crude extracts. The regions or domains of the protein contacting DNA can be identified by peptide mapping of the photoaffinity labeled protein *(5)*. Our discussion of DNA photoaffinity labeling will focus on (a) synthesis of photoreactive nucleotide analogs, (b) the manner in which the photoreactive nucleotide is incorporated into DNA, and (c) experimental details of DNA photoaffinity labeling.

Some of the advantages of this approach are: (a) detailed mapping of protein interactions with DNA in large multisubunit protein–DNA complexes and (b) the ability to use crude protein extracts potentially containing important auxiliary factors that may be lost upon purification. Solid-phase DNA probe synthesis allows for the synthesis of multiple probes in a single day where in the past this process would have taken several days. The synthesis of modified analogs for dATP, dCTP, and dTTP allows for incorporation at nearly all positions in DNA *(6–8)*. The photoreactive moiety can be changed on these nucleotides to place the more photoreactive phenyl diazirine into DNA to better target all potential protein surfaces *(9)*. These photoreactive groups have short half-lives of less than a nanosecond to ~5 μs and can be used for kinetic analysis of changes in specific protein–DNA contacts. The 4-thiothymidine nucleotide has also been used for zero-distance crosslinking of protein to DNA and may be useful for probing very close protein–DNA contacts *(10)*.

2. Materials

2.1. Synthesis of Modified Nucleotides

1. Para-azidobenzoic acid (4-ABA) (Molecular Probes).
2. Succinimidyl esters of 4-azidobenzoic acid, 4-azido-2,3,5,6-tetrafluorobenzoic acid, 4-benzoylbenzoic acid (Molecular Probes).
3. 5-[N-(3-aminoallyl)]-deoxyuridine triphosphate (5 aa-dUTP) (Fermentas).
4. dCTP and *p*-azidophenacyl bromide(Sigma).
5. Ethylene diamine, dicyclohexylcarbodiimide, anhydrous dioxane (99 + %), ethyl ether, and sodium metabisulfite (Aldrich).
6. DEAE-Sephadex A-25 resin (GE Life Sciences).
7. Glycine, glycyl glycine, and glycylglycyl glycine (Sigma).
8. pH indicator strip (Panpeha, Schleicher, and Schull).
9. Polyethyleneimine (PEI)-cellulose TLC plates (J.T. Baker, with fluorescence indicator).
10. TE: 10 mM Tris–HCl of pH 8.0, 1 mM EDTA.

2.2. Immobilized DNA Templates

1. Buffer A: 10 mM Tris–HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM NaCl, and 1 mM DTT.
2. Buffer B: 1 M LiCl, 10 mM Tris–HCl (pH 8.0), 1 mM EDTA, and 0.1% SDS.
3. Plasmid DNA pTZ1 containing the *SUP4* $tRNA^{Tyr}$ gene with promoter-up mutation inserted into pGEM1 [*7*].
4. Magnetic separation stand for DNA bead isolation (Promega).
5. M-280 Streptavidin Dynabeads (Dynal).
6. Buffer C: 2 M NaCl, 10 mM Tris–HCl (pH 7.5), 1 mM EDTA (pH 8.0).
7. Buffer D: 30 mM Tris–HCl (pH. 8.0), 50 mM KCl, 7 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 0.05% Tween 20.
8. Five-inch (12.5 cm) polystrene chromatography columns with 45–90 μm filter, (Evergreen Scientific). These disposable columns are ideal for the 2.5-ml spin columns.
9. Bio-11-dUTP and Bio-14-dATP (Sigma).

2.3. DNA Probe Synthesis

1. Buffer E: 150 mM Tris–HCl (pH 8.0), 250 mM KCl, 35 mM $MgCl_2$, 5 mM 2-mercaptoethanol, and 0.25% Tween 20.
2. Storage buffer F: 50 mM potassium phosphate of pH 7.0, 5 mM 2-mercaptoethanol, and 50% glycerol.
3. Storage buffer G: 50 mM KCl, 10 mM Tris–HCL of pH 7.5, 0.1 mM EDTA, 5 mM 2-mercaptoethanol, 200 μg/ml BSA, and 50% glycerol.
4. Site-specific oligonucleotides and upstream oligonucleotide (50-nmol scale synthesis).
5. 4-Thiothymidine triphosphate (Trilink Biotechnologies).
6. Exonuclease-free version of the Klenow fragment of DNA Polymerase I (GE Life Sciences, 5 units/μl) is diluted to 0.25 units/μl with Storage Buffer F.
7. T_4 DNA ligase (New England Biolabs, high concentration form, 2,000 units/μl) is diluted to ~300 units/μl with Storage Buffer G containing 5 mM 2-mercaptoethanol.
8. TE: 10 mM Tris–HCl of pH 8.0 and 1 mM EDTA.
9. PBS: phosphate-buffered saline solution of pH 7.4 containing.
10. T_4 DNA polymerase from New England Biolabs comes stored in 100 mM potassium phosphate of pH 6.5, 10 mM 2-mercaptoethanol, and 50% glycerol.

2.4. Photoaffinity Labeling

1. Buffer H: 100 mM Tris–HCl (pH 8.0), 25 mM MgCl2, and 250 mM NaCl.
2. Buffer I: 100 mM NaCl, 40 mM Tris–HCl (pH 8.0), 5 mM MgCl2, 1 mM EDTA, 20% glycerol, 10 mM 2-mercaptoethanol, 0.5 mM PMSF, 1 μg/ml pepstatin, and 1 μg/ml leupeptin.
3. Zinc acetate solution: 0.5 M glacial acetic acid, 12.5 mM zinc acetate.
4. 5× DB: 10% SDS, 25% 2-mercaptoethanol, 0.3 M Tris–HCl (pH 6.8), 0.4% bromophenol blue.
5. 3 NTP mix: 2 μl of 100 mM ATP, UTP, and CTP (Boehringer Mannheim), 20 μl Buffer H, 35 μl Buffer I, and 43 μl of sterile deionized water.

2.5. Peptide Mapping

1. Formic acid (99%, Sigma).
2. Diphenylamine (ACS 99 + %, Aldrich).
3. Cyanogen bromide (97%, Aldrich).
4. Centricon 30 (Millipore).

3. Methods

Synthesis of modified nucleotides and oligonucleotides followed by incorporation into double-stranded DNA is done with indirect lighting using 40-W incandescent lamps.

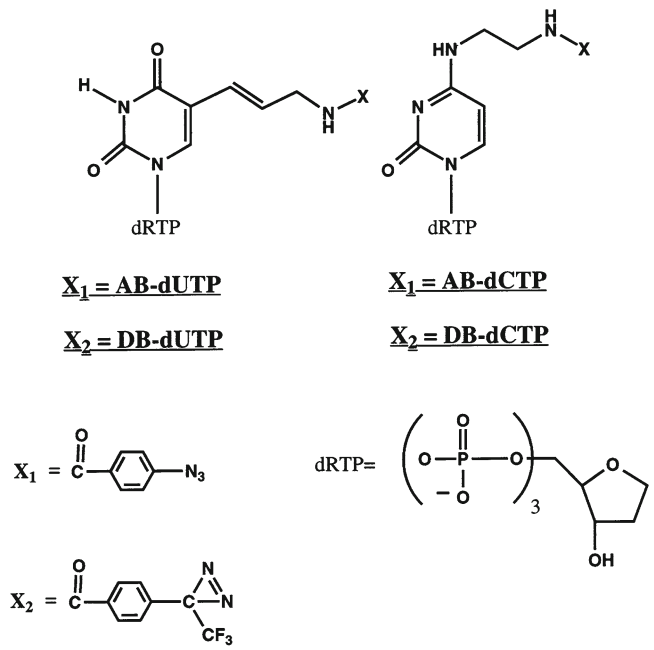
3.1. Synthesis of Modified Nucleotides

We have used a variety of modified nucleotides to probe the RNA polymerase III transcription complex and ATP-dependent chromatin remodeling complexes SWI/SNF and ISW2. The following section contains procedures for the synthesis of some of the commonly used nucleotides. Not included in this description is the synthesis of N^6-[4-azidobenzoyl-(2-aminoethyl)]-2′-deoxyadenosine-5′-triphosphate (AB-dATP) or its precursor N^6-(2-aminoethyl)-2′-deoxyadenosine-5′-triphosphate (N^6-dATP) *(8)*. Synthesis of this nucleotide analog requires multiple steps and starts with dAMP being converted into N^1-dAMP followed by its rearrangement into N^6-dAMP.

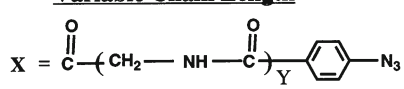
3.1.1. AB-dUTP

AB-dUTP (**Fig. 1**) is synthesized as follows:

1. Add 100 μl of 100 mM 4-azidobenzoic acid *N*-hydroxysuccinimide solution (ABA-NHS) in dimethylformamide (DMF) to 100 μl of 20 mM 5-aa-dUTP in 100 mM sodium borate of pH 8.5 *(6)*. The reaction is incubated at 25°C for 4 h, and the pH of the reaction is checked with pH indicator strips (pH 8.5). Any excess precipitation that forms can be eliminated by addition of DMF.



Variable Chain Length



	dUTP analogs	dCTP analogs	Chain Length
Y= 0	AB-dUTP	AB-dCTP	[~10.0 Å]
Y= 1	ABG-dUTP	ABG-dCTP	[~14.3 Å]
Y= 2	ABG_2-dUTP	ABG_2-dCTP	[~18.6 Å]
Y= 3	ABG_3-dUTP	ABG_3-dCTP	[~22.9 Å]

Fig. 1. Structures of each of the different photoreactive nucleotide analogs have been shown. The complete IUPAC names for each of the nucleotides is given in the text.

2. The coupling reaction is stopped by the addition of 200 µl of sterile deionized water and the product applied to a 0.7 × 8 cm (1.6 ml) DEAE Sephadex A-25 column equilibrated in 100 mM TEAB (pH 8.0). The column is washed with 5 ml of the same buffer at a flow rate of 5 ml/h, and eluted with a 30-ml linear gradient of 0.1–1.5 M TEAB (pH 8.0), and 500 µl fractions are collected.
3. Every third fraction from the column is evaporated to dryness by vacuum centrifugation and resuspended in 250 µl of sterile deionized water. The samples are dried down and this process is repeated for a second time.

4. The fractions are resuspended in 50 µl of sterile deionized water and 2 µl of each analyzed on PEI-cellulose TLC plates (*see* **Note 2**). The plates were developed with 1 M LiCl and visualized with a UV light source (254 nM). The reported R_f values for 5-aa-dUTP and AB-dUTP are 0.54 and 0.098, respectively *(6)*.
5. All the fractions containing AB-dUTP are combined and TEAB is removed by repeated drying and resuspension in sterile deionized water. The final product is resuspended in 200 µl of TE. An estimated extinction coefficient of AB-dUTP at 270 nm is 10.3×10^3/M/cm at pH 8.0 (based on the sum of the extinction coefficients of ABA and 5-aa-dUTP at the indicated pH and wavelength). A concentrated stock of AB-dUTP is stable at –80°C for several years, and a 0.2-mM working stock can be stored at –20°C wrapped in foil (*see* **Notes 1**, **3**, and **4**).

3.1.2. Varied Tether Length Nucleotides

The tether of AB-dUTP is 9–10 Å in length and places the photoreactive group near the edge of the major groove of DNA. We have synthesized different dUTP and dCTP analogs with varying tether lengths by addition of glycine residues into the tether *(11)*. Synthesis of these nucleotide analogs is similar to that of AB-dUTP and is as follows (**Fig. 1**).

Para-azidobenzoic acid (4-ABA) was esterified with N-hydroxysuccinimide (NHS) using the coupling reagent dicyclohexylcarbodiimide (DCI), and the product was recrystallized from anhydrous dioxane and ethyl ether (1:1) *(12)*.

1. A typical reaction contained 28 mmol ABA and 28 mmol NHS in 50 ml of anhydrous dioxane (99 + %).
2. The solution is cooled in ice and DCI in 15 ml dioxane is added and stirred for ~24 h at room temperature.
3. Dicyclohexyl urea is removed by centrifugation and the supernatant was evaporated to dryness by vacuum centrifugation.

   The ABA-NHS is coupled to glycine, or (Gly-Gly) or (GlyGlyGly) to make ABA derivatives with glycine, Gly-Gly, or Gly-Gly-Gly attached to the carboxylic group of 4-ABA.
4. The reaction is started on ice and contains 2 mmol of glycine or the Gly-Gly or Gly-Gly-Gly and 4 mmol of sodium bicarbonate in 4 ml of deionized water to which is added 2 mmol of ABA-NHS in 8 ml of dioxane with constant stirring.
5. The reaction proceeds for 10–15 min on ice and is then transferred to room temperature and left with stirring for an additional 24 h.
6. Any insoluble material is removed from the reaction by centrifugation.
7. The pH of the reaction is lowered to 2 with concentrated HCl to precipitate the product.

8. Products are washed with deionized water.
9. The products are esterified with N-hydroxysuccinimide as described for ABA, **Subheading 3.1.1**, except that dimethyl sulfoxide is used instead of dioxane for the ABG_3-NHS because of the limited solubility of this compound.
10. The DMSO or dioxane is removed by vacuum centrifugation and the product is recrystallized from dioxane/isopropyl alcohol (1:1). Any residual solvent is removed by vacuum centrifugation. These products are coupled to 5-aa-dUTP in the same fashion as described for AB-dUTP, **Subheading 3.1.1**.

3.1.3. Varied Photochemistry Nucleotides

We have also varied the photoreactive group attached to 5-aa-dUTP to contain either a phenyldiazirine, tetrafluoroaryl azide, or a benzophenone group to optimize for nonselective crosslinking *(9)*. The coupling reactions of 5-aa-dUTP to the NHS esters of 4-azido-2,3,5,6-tetrafluorobenzoic acid, 4-benzoylbenzoic acid (commercially available from Molecular Probes), and 4-[3–9trifluoromethyl)diazirin-3-yl]benzoic acid (synthesized as described in *(13)*) are similar to that for the synthesis of AB-dUTP (**Fig. 1**).

3.1.4. Synthesis of dCTP Analogs

The synthesis of dCTP nucleotides begins with the synthesis of N^4-aminoethyl deoxycytidine triphosphate (daeCTP) by a bisulfate-catalyzed transamination reaction *(7)*.

1. A bisulfite amine solution is made by adding dropwise 2 ml of freshly distilled ethylene diamine to 4.0 ml of concentrated HCl and 3.5 ml of deionized water on ice.
2. Next, sodium metabisulfite (1.895 gm) is added and the pH is adjusted to 5.0 with concentrated HCl.
3. Next, 100 μl of 1 mg/ml hydroquinone in ethanol is added to the reaction to scavenge free radicals. Bisulfite-amine solutions are always made up fresh.
4. The transamination reaction is initiated by adding nine volumes of the bisulfite-amine solution to one volume of 100 mM dCTP in 50 mM TEAB of pH 8.0.
5. The sample is incubated with constant vortexing at 42°C for 4 h.
6. The reaction is stopped by adjusting the pH to 8.2 with 5 M KOH.
7. The product is purified by DEAE Sephadex A-25 chromatography as described for the purification of AB-dUTP, **Subheading 3.1.1**, and dae-dCTP eluted from 0.84 to 1.0 M TEAB.
8. Column fractions are analyzed by thin layer chromatography as described in **Subheading 3.1.1**. The R_f of daeCTP is 0.221 *(7)*.

9. Fractions containing product are pooled and concentrated to 10–12 mM. The aryl azido or phenyl diazirine (**Fig. 1**) are coupled to daeCTP as discussed earlier, **Subheading 3.1.1**, for 5-aa-dUTP. Tether length versions of AB-dCTP have also been synthesized with similar lengths of tether to those for the other tether length nucleotides discussed (**Fig. 1**), and the R_f values of these range from 0.067 for ABG-dCTP to 0.078 for ABG_3-dCTP.

3.2. Synthesis of Modified Oligonucleotides

Oligonucleotides are commercially synthesized with a phosphorothioate incorporated near the 3′ or 5′ end and are conjugated to an aryl azide using *p*-azidophenacyl bromide (APB). The phosphorothioate is between the third and fourth nucleotides from either end of the oligonucleotide and is not directly at the end because the modified phosphorothioate destabilizes annealing of the oligonucleotide to the template strand *(14)*. The oligonucleotide is radiolabeled by γ-^{32}P-ATP and T4 polynucleotide kinase (Optikinase from USB) if the modification is near the 5` end; otherwise, the 3` end is radiolabeled by extending the oligonucleotide after annealing to the DNA template strand by Klenow DNA polymerase and α-^{32}P-dNTP.

1. The oligonucleotide is resuspended in 100 mM triethylammonium bicarbonate of pH 8.0 (TEAB) to a concentration of ~200 pmol/μl.
2. The conjugation reaction contains 25 μl of 200 pmol/μl of oligonucleotide and 25 μl of 100 mM APB in dimethylformamide (DMF) or acetonitrile (24 mg/ml).
3. If precipitation does occur upon mixing add small amounts of DMF or acetonitrile.
4. The reaction mixture is incubated at 25°C for 1 h if using DMF or for 4 h if using acetonitrile.
5. The reaction is diluted by the addition of 300 μl of 100 mM TEAB of pH 8.0 and extracted three times with isobutanol that has been saturated with 10 mM Tris–HCl of pH 8.0 and 1 mM EDTA (TE). The oligonucleotide is in the lower aqueous layer and the isobutanol layer is removed and fresh TE saturated isobutanol added for further extraction.
6. The oligonucleotide is extracted three times with TE saturated ethyl ether the same as in **step 5**.
7. After removal of all the ethyl ether, the sample is incubated at 37°C for 10 min with the caps open for evaporation of residual ether in the sample.
8. Oligonucleotide is concentrated and dried by centrifugal evaporation under vacuum using a Labanco system for ~2 h. The oligonucleotide is resuspended in 1 ml of deionized water and concentrated as before.

9. The oligonucleotides are resuspended in 100 μl of 100 mM TEAB to a final concentration of ~50 pmol/μl and stored at −20°C.
10. The extent of modification is determined by radiolabeling the modified and unmodified oligonucleotides and analyzing them side by side on a denaturing 10% polyacrylamide gel containing 8 M urea. Modification causes the oligonucleotide to have a reduced electrophoretic mobility.

3.3. Immobilized DNA Templates

Plasmid DNA is used for immobilizing a single-stranded DNA template that contains either the *SUP4* $tRNA^{Tyr}$ DNA for RNA polymerase III experiments or the 5S rDNA for high-affinity 601 DNA for nucleosome reconstitution *(3, 15–17)*.

1. DNA is biotinylated by initially digesting 200 pmol of plasmid with a restriction enzyme that cleaves at one unique site in the plasmid proximal to the region of interest. An example of this is shown for the pTZ1 plasmid containing the *SUP4* $tRNA^{Tyr}$ DNA that is digested with *Hind*III for nontranscribed strand templates or *Eco*RI for transcribed strand templates (*see* **Note 5**) (**Fig. 2**, step 1).
2. The 5` overhangs are biotinylated by the incorporation of Bio-11-dUTP and Bio-14-dATP (Sigma Chemical Co.) using the exonuclease-free version of the Klenow fragment of DNA Polymerase I (GE Life Sciences). The 200-μl reaction contains 200 pmol of linearized plasmid, 20 μM Bio-14-dATP, dCTP, dGTP, and 25 μM Bio-11-dUTP, and 150 units of Klenow fragment in buffer A (**Fig. 2**, **step 2**).
3. Unincorporated dNTPs are removed by spin column chromatography *(18)* with a 2.5-ml Sephacryl S-200 spin column (GE Life Sciences) equilibrated in buffer B (**Fig. 2**, **step 3**). Aliquots of the samples are removed before and after the spin column to quantify the recovery. Instead of a spin column, Microcon30 can be used to remove unincorporated nucleotides and the sample is washed four times with 500 μl of TE.
4. Biotinylated DNA is precipitated by the addition of 2.5 volumes of ethanol and placing the sample at −20°C overnight. Afterward DNA is subjected to a second restriction enzyme digestion.
5. Sample is resuspended and digested with a second restriction endonuclease that cleaves on the other side of the target region. Generally it is preferred to keep the target DNA in the range of 300–500 bp to facilitate efficient binding to the magnetic beads with streptavidin. In the case of the pTZ1 DNA it is digested with *Eco*RI and *Rsa*I for nontranscribed strand templates or *Hind* III and *Pvu*II for transcribed strand templates to generate a 315-base pair biotinylated DNA fragment containing the *SUP4* $tRNA^{Tyr}$ gene (**Fig. 2**, **step 4**).

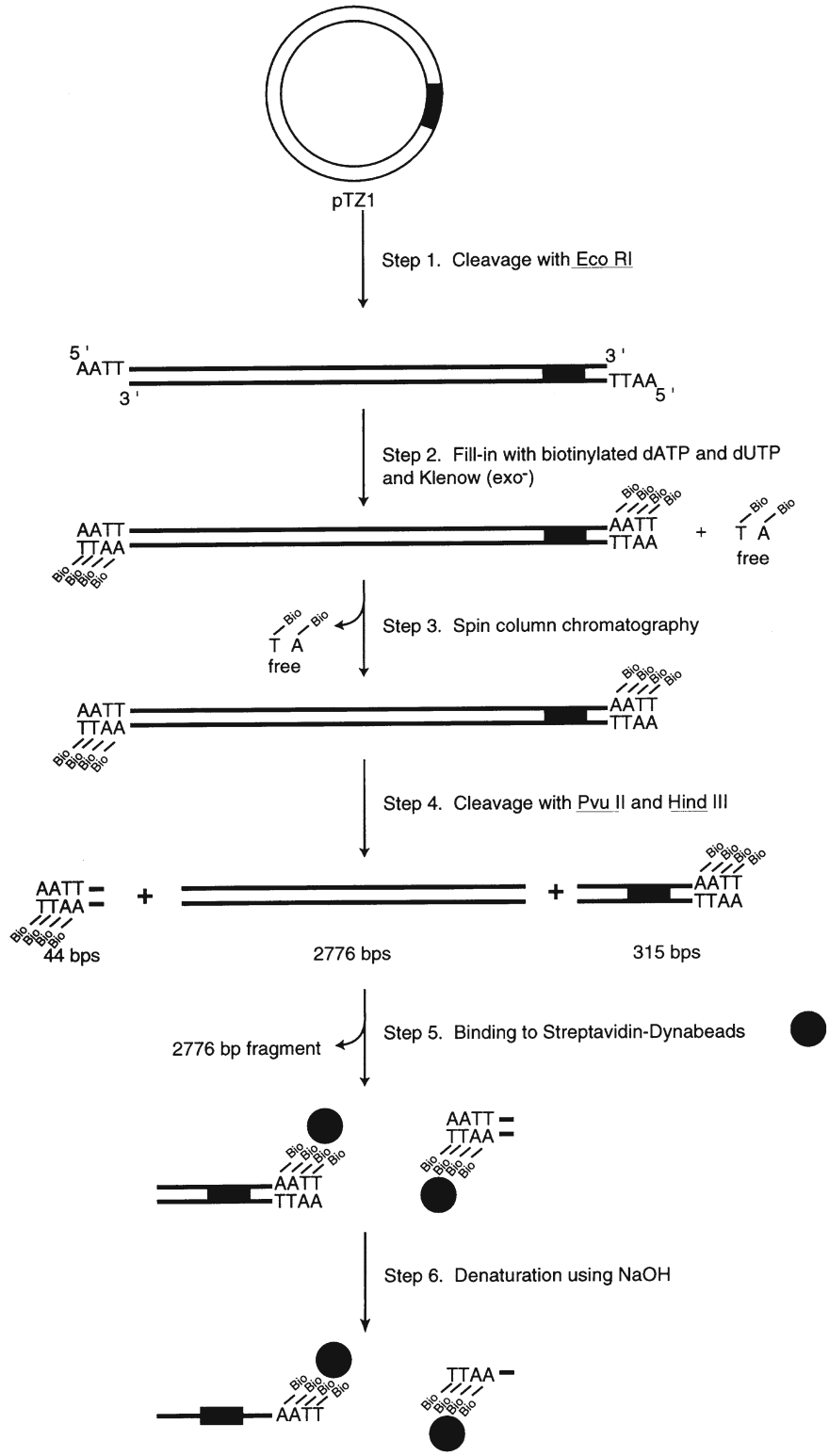


Fig. 2. Diagrammatic representation of DNA template preparation for solid-phase DNA probe synthesis.

Biotinylated DNA (40 pmol) is bound to Dynabeads M-280 Streptavidin (Dynal) in the following procedure.

1. Washing Dynabeads
   (a) The supernatant from 100-µl Streptavidin Dynabeads (10 pmol/µl) is removed using a MagneSphere® Technology magnetic separation stand (Promega), and the beads are resuspended in 150 µl PBS + 0.1 mg/ml BSA.
   (b) The beads are washed one time with 150 µl buffer C, and resuspended in 100 µl 2 × B/W and divided into 200–300 µl aliquots of beads at 5 mg/ml (*see* **Note 7**).
2. Binding reaction
   (a) A reaction is assembled consisting of 100 µl of the 5 mg/ml washed Dynabeads and 100 µl of the 0.4 pmol/µl biotinylated DNA (*see* **Note 6**).
   (b) The reaction is gently vortexed every 1–2 min for 10 s each during which the sample is incubated for 45 min at 37°C (*see* **Note 8**). For this purpose we generally use a Thermomixer (Eppendorf) that can be programmed to perform these repetitive actions.
   (c) Buffer is removed using a magnetic stand and analyzed for binding efficiency by agarose gel electrophoresis (**Fig. 3**, lanes **2** and **5**).
   (d) Beads are washed with 200 µl 0.5× buffer C and the washes saved for further analysis (**Fig. 3**, lanes 3 and 6).
   (e) The beads are resuspended in 100 µl of 0.5× buffer C and stored at 4°C (**Fig. 2**, **step 5**) for extended periods of time.
3. Removal of the nonbiotinylated DNA strand
   (a) Supernatant is removed from the beads and the beads are resuspended in 20 µl of 0.1 M NaOH.
   (b) The sample is incubated with occasional vortexing for 10 min at room temperature.
   (c) The supernatant is removed and beads are washed one time with 50 µl of 0.1 M NaOH and washed four times with sterile deionized water (**Fig. 2**, **step 6**).
4. Dephosphorylation of DNA Beads
   (a) DNA beads are washed three times with 50 µl of Buffer D and resuspended in 100 µl of Buffer D.
   (b) Five units of shrimp alkaline phosphatase (GE Life Sciences) is added to the reaction, and incubated at 37°C for 60 min.
   (c) A second five units of enzyme is added and incubated for an additional 60 min.
   (d) Beads are washed three times with 50 µl of TE (pH 8.0) + 0.1% SDS.

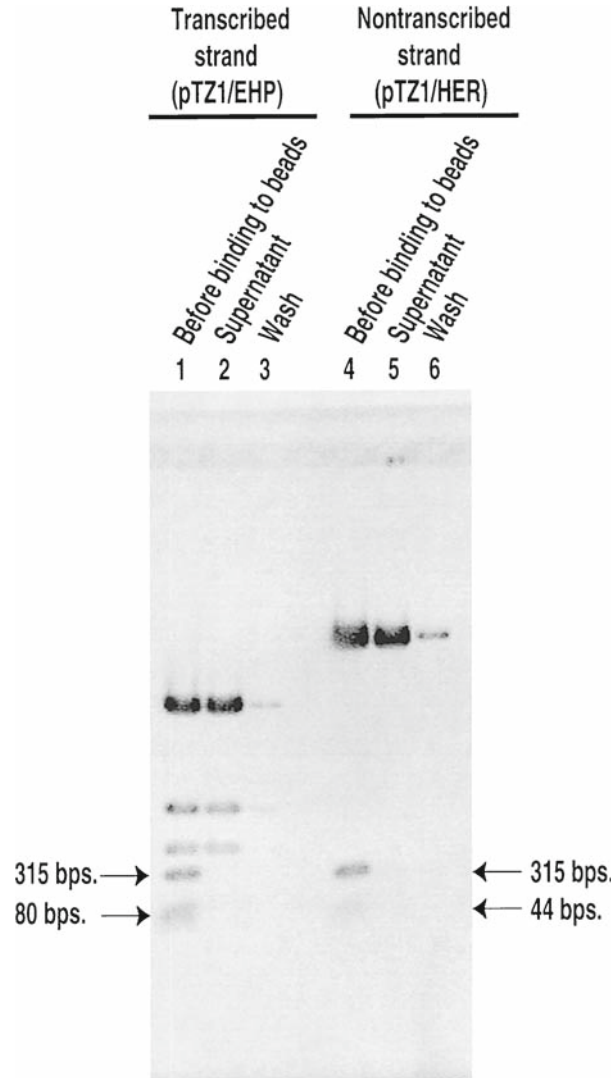


Fig. 3. Agarose gel analysis of DNA bead preparation. *Lanes 1* and *4* are biotinylated pTZ1 DNA cut with *Eco* RI, *Hind* III, and *Pvu* II (pTZ1/EHP) are used for the modification of the transcribed strand, or *Hind* III, *Eco* RI, and *Rsa* I (pTZ/HER) are used for the modification of the nontranscribed strand before binding to Streptavidin Dynabeads. The biotinylated 315-base pair fragments contain the *SUP4*^{Tyr} gene. The 80- and 40-bp fragments are also biotinylated, but are not used as a DNA template. *Lanes 2* and *5* are taken from the supernatant of the binding reaction of DNA with Streptavidin Dynabeads after incubation. The absence of the 315-bp fragments in these samples shows efficient binding of probe DNA to Dynabeads. *Lanes 3* and *6* are samples from a final wash of the beads after binding.

(e) The beads are washed three times with 50 µl of PBS + 0.1 mg/ml BSA, and resuspended in 100 µl of PBS + 0.1 mg/ml BSA (final concentration is 0.2 pmol/µl).

(f) The remaining enzyme is heat inactivated by incubating the reaction at 65°C for 15 min and gently vortexing

every 5 min. The single-stranded DNA beads can be stored for several months up to 1 year at 4°C.

3.4. DNA Probe Synthesis

Immobilized DNA templates made by the previous procedure, **Subheading 3.3**, are used to construct DNA photoaffinity probes on either the transcribed or nontranscribed strand of the gene in the following procedure.

1. First primer extension. Remove 2 pmol of immobilized template and place in a 1.5-ml microcentrifuge tube.
2. Wash the beads three times with Buffer D and resuspend in 10 µl of Buffer D.
3. To each reaction is added 2 µl of sterile deionized water, 1 µl of 2 mg/ml BSA, 2 µl Buffer E, and 5 µl of 6 pmol/µl site-specific oligonucleotide (*see* **Note 9**).
4. Reactions are vortexed and heated at 70°C for 3 min with intermittent vortexing. Samples are cooled with 10°C drop every 10 min with a final incubation at 37°C for 30 min to allow the oligonucleotide primer to anneal (**Fig. 4**,

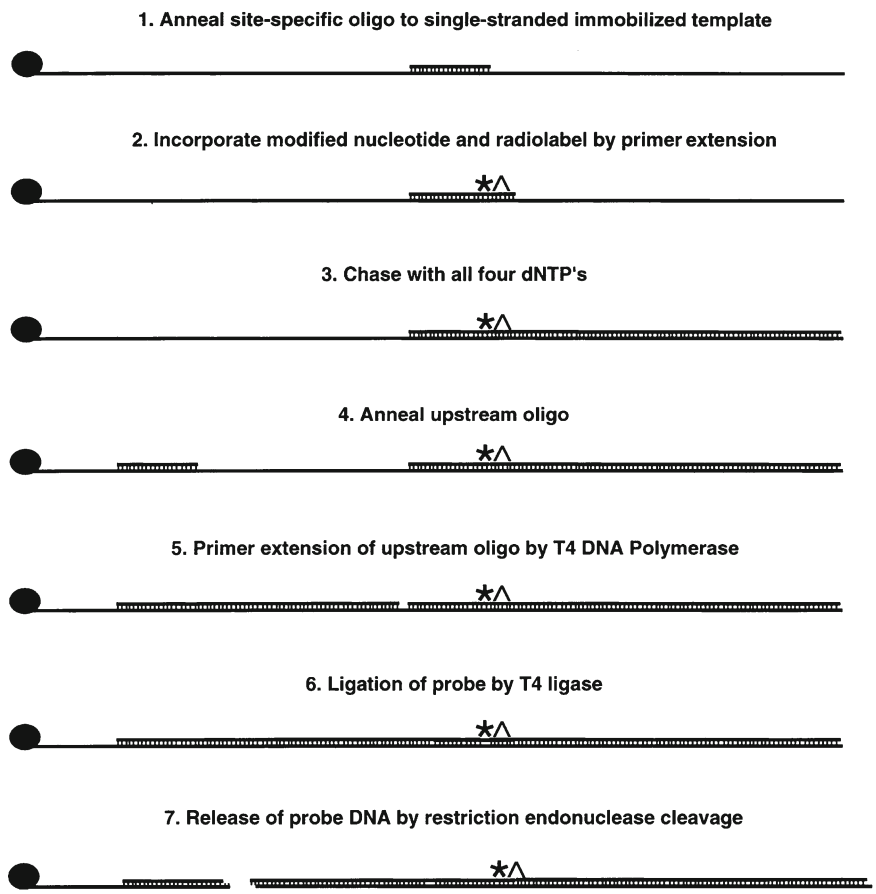


Fig. 4. Schematic representation of solid-phase DNA probe synthesis.

step 1). During all incubations, samples must be vortexed gently for 10 s every 1–2 min to keep the beads suspended.

5. The beads were washed two times each with 50 µl of buffer D and resuspended in 20 µl of buffer D.
6. To the reaction add 1 µl of 100 µM modified nucleotide described in **Subheading 3** and an α-[P^{32}] labeled nucleotide using 0.25 units of the exonuclease-free version of the Klenow fragment of DNA polymerase I (GE Life Sciences) (*see* **Note 11**) and incubation at 37°C for 10 min (*see* **Note 10**) (**step 2** in **Fig. 4** and lanes 1–4 in **Fig. 5**).
7. Next, full-length extension is done by addition of dNTPs to a final concentration of 5 mM and incubation for 15 min at 37°C. (**step 3** in **Fig. 4** and lanes 5–8 in **Fig. 5**).
8. Klenow fragment and dNTPs are removed by washing the beads three times with 50 µl of TE (pH. 8.0) + 0.1% SDS and two times each with 50 µl Buffer D.
9. Beads are resuspended in 20 µl of Buffer D.
10. Three µl of 6 pmol/µl upstream oligonucleotide, 1 µl of 2 mg/ml BSA, and 1 µl of Buffer E are added to a 20-µl reaction containing the DNA beads.
11. The sample is vortexed gently every 5 min and incubated at 37°C for 30 min (**Fig. 4**, **step 4**).
12. The upstream oligonucleotide is extended by the addition of 5 mM dNTPs and 3 units of T4 DNA polymerase (*see* **Note 11**), and incubation for 10 min at 37°C (step 5 in **Fig. 4** and lanes 9–12 in **Fig. 5**).
13. The upstream strand is ligated to the site-specific primer by the addition of 1 µl of 10 mM ATP and 5 units of T4 DNA ligase and incubated at 37°C for 60 min (step 6 in Fig. 4 and lanes 13–16 in **Fig. 5**).
14. Beads are washed two times with 50 µl TE (pH 8.0) + 0.1% SDS, two times with 50 µl of Buffer D, and resuspended in 20 µl of Buffer D.
15. DNA probe is released from the bead by the addition of 12–20 units of specific restriction enzyme (*see* **Notes 11** and **12**) to cut at a site between the *SUP4* $tRNA^{Tyr}$ gene and the attachment site (**step 7** in **Fig. 4** and lanes 17–20 in **Fig. 5**).
16. After restriction enzyme digestion, the probe is washed from the beads in a series of three washes of 50-µl each with Buffer D which are pooled in a fresh microcentrifuge tube.
17. The sample is extracted with phenol:chloroform (1:1) followed by extraction with chloroform.

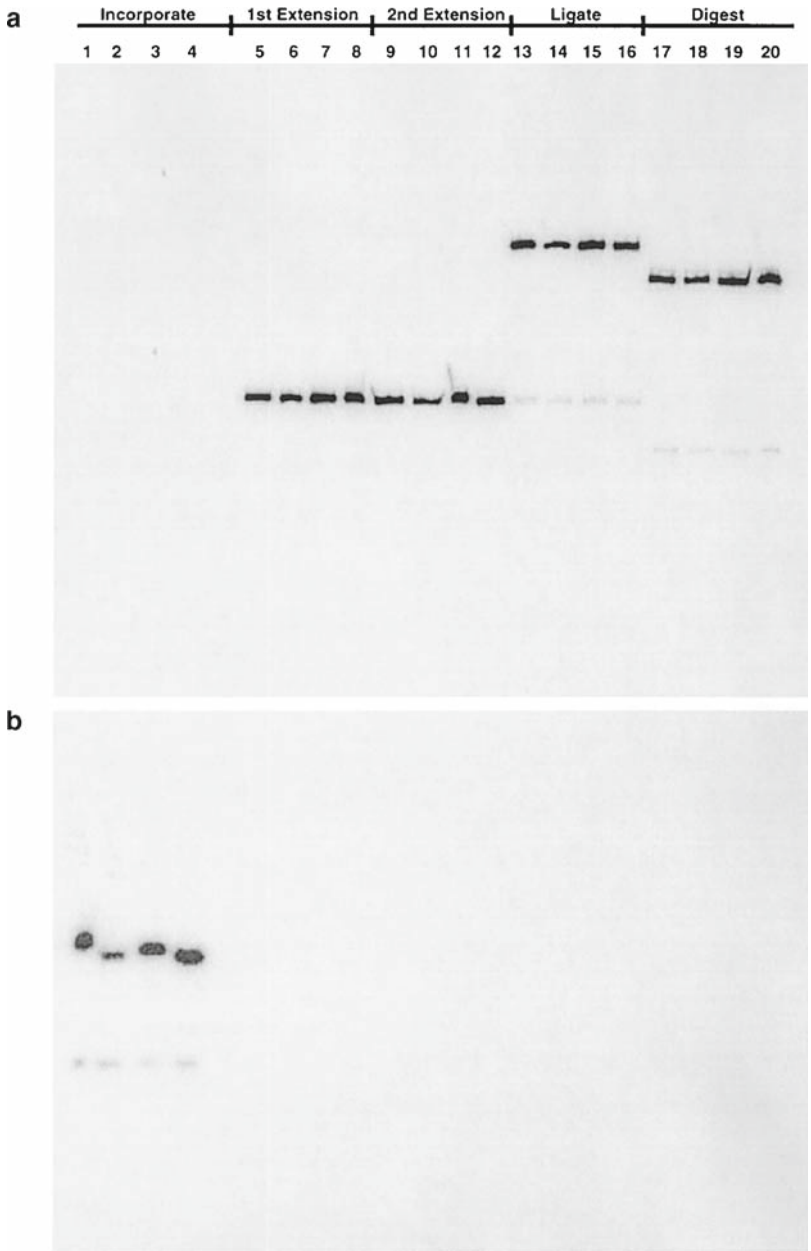


Fig. 5. Analysis of DNA probe synthesis modified at bps + 11 on 10% urea PAGE. Incorporation of BP-dUMP, FAB-dUMP, DB-dUMP, and AB-dUMP and [α-^{32}P] dATP (*Lanes 1–4*). Full-length extension of oligonucleotide primers in the presence of all four dNTPs (*Lanes 5–8*). Upstream oligo annealed to template and extended to site-specific primer by T4 DNA Polymerase in the presence of all four dNTP's (*Lanes 9–12*). Ligation of upstream extension product to site-specific oligonucleotide by T4 DNA Ligase (*Lanes 13–16*). Digestion of Probe DNA to release it from the Streptavidin Dynabeads by *Bam* HI (*Lanes 17–20*).

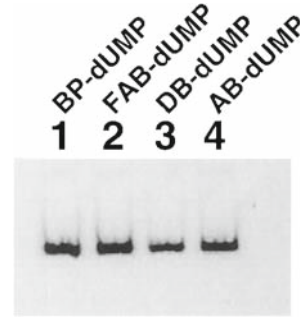


Fig. 6. Analysis of DNA probes modified at bps + 11 on a 4% nondenaturing acrylamide gel.

18. The DNA probe is ethanol precipitated by the addition of 1/10 volume of 10 M lithium chloride and 2.5 volumes of ethanol.
19. Samples are placed at –20°C overnight.
20. Samples are spun down at maximum speed in microfuge at 4°C for 30 min.
21. Supernatant is decanted and pellet is allowed to dry.
22. Samples are resuspended in TE (pH 8.0) + 0.05% Tween 20 at a final concentration of 2–10 fmol/μl and stored at 4°C.
23. After resuspension of the probes 1 μl is removed for analysis on a 4% native acrylamide gel, 20 × 20 cm × 0.8 mm (**Fig. 6**).

3.5. DNA Photoaffinity Labeling

Transcription complexes were formed on probe DNA using the 500-mM KCl fraction from BioRex 70 chromatography of the S-100 extract, made from *Saccharomyces cerevisiae* strain BJ926 (BR500) *(7)*. In addition, transcription complexes were also formed using recombinant TFIIIB and partially purified TFIIIC *(19)*. TFIIIC was obtained from the flow-through fractions of Ni-NTA chromatography of His-tagged RNA Pol III *(7)*.

1. A typical 20-μl photoaffinity labeling reaction contains 4 μl of Buffer H, 9 μl of Buffer I, 1 μl of 500 ng/μl pLNG-56 or pTZ1 (*see* **Note 15**) linearized with EcoRI, 1 μl of 2 fmol/μl DNA probe, and 1–4 μl of BR-500 extract (*see* **Note 14**), and adjusted to 20 μl with deionized water. Optimal protein concentration for transcription activity and photoaffinity labeling is determined by multiple round transcription assays using pTZ1 plasmid DNA and labeled ribonucleotides *(3)*.
2. The photoaffinity labeling reaction is incubated at 25°C for 30 min for assembly of complexes on to probe DNA.
3. The sample is irradiated at this point to crosslink the assembled complex, or a stalled ternary complex can be formed by the addition of a 3 NTP mix. Complexes containing only TFIIIB are formed by addition of heparin (100 μg/ml) to

the assembled transcription complex for release of TFIIIC and Pol III from DNA.

4. After irradiation, the DNA probe is enzymatically digested in two steps to leave a small radioactive tag covalently attached to the crosslinked protein.
5. The first step of digestion is by the addition of 2.3 µl of 0.5 mg/ml DNase I (Invitrogen) to a 21-µl reaction and incubation at 25°C for 10 min (*see* **Note 13**).
6. Immediately add 1 µl of 10% SDS to each sample and incubate at 90°C for 3 min and then place on ice for 5 min and next place at room temperature.
7. Two microliters of the zinc acetate solution and 1 µl of 20 units/µl S1 nuclease (Invitrogen) are added, and samples are incubated at 37°C for 10 min.
8. The reaction is stopped by the addition of a premix of 1 µl of 0.5 M Tris base to adjust the pH to ~7 and 7 µl of 5× DB buffer.
9. The sample is heated at 90°C for 3 min and cooled on ice for 5 min prior to being loaded on to a 4–20% SDS-PAGE that is 20 × 20 cm × 0.8 mm.
10. After electrophoresis, the gel is dried with a slab gel dryer with heating at 80°C and vacuum for 2 h, and exposed to film for autoradiography or to a phosphorimager screen (*see* **Note 16**) (**Fig. 7**).

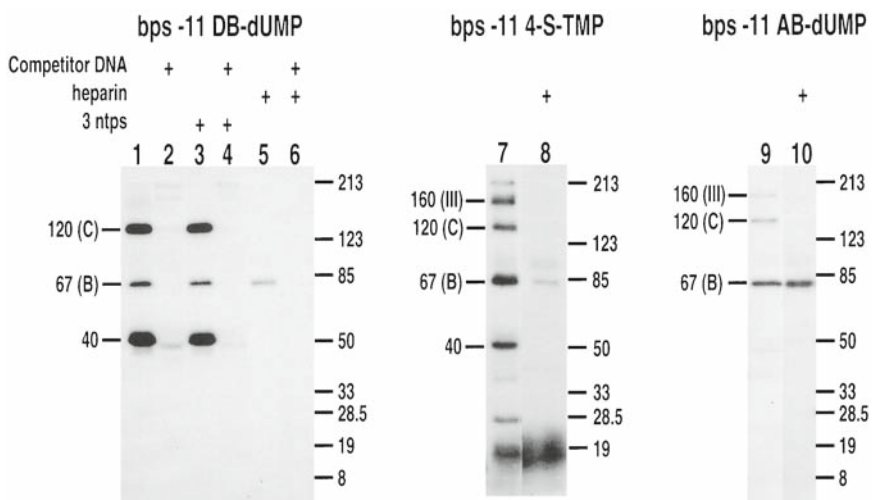


Fig. 7. Comparison of three photoreactive moieties incorporated at bps -11. Subunits labeled in both preinitiation (*Lanes 1*, *7*, and *9*) and heparin-stripped complexes (*Lanes 5*, *8*, and *10*) differ from one photoreactive group to the next. Examples of TFIIIC-specific competition (*Lanes 2*, *4*, and *6*).

3.6. Peptide Mapping

1. Large-scale photoaffinity labeling reactions for peptide-mapping experiments are as described in **Subheading 3.5**, except that everything is scaled up to a final reaction volume of 2 ml (*see* **Note 17**).
2. Samples are irradiated in a multichannel pipettor tray instead of the original sample tube in order to keep the depth of the sample the same as a standard 20-μl labeling reaction.
3. DNase I and S1 nuclease digestion is done as described in **Subheading 3.5** with volumes scaled up to the appropriate amounts for the increased reaction size.
4. Next the samples are concentrated by ultrafiltration using a Centricon 30 (Millipore) to lower the sample volume and allow for loading on a 0.8-mm-thick 8% SDS polyacrylamide gel.
5. Photoaffinity-labeled BRF was excised from the gel and electrophoretically eluted using a Bio-Rad Model 422 Electro-Eluter for 4 h at 10 mA per gel slice into a volatile buffer consisting of 50 mM ammonium bicarbonate and 0.1% SDS.
6. The eluate is dried down by vacuum centrifugation, resuspended in 200 μl sterile deionized water, and dried down again.
7. Gel-purified BRF is treated with 70% formic acid and 1.4% diphenylamine at 70°C for 20 min to further digest DNA and to cleave the protein at Asp-Pro sites (**Fig. 8**).

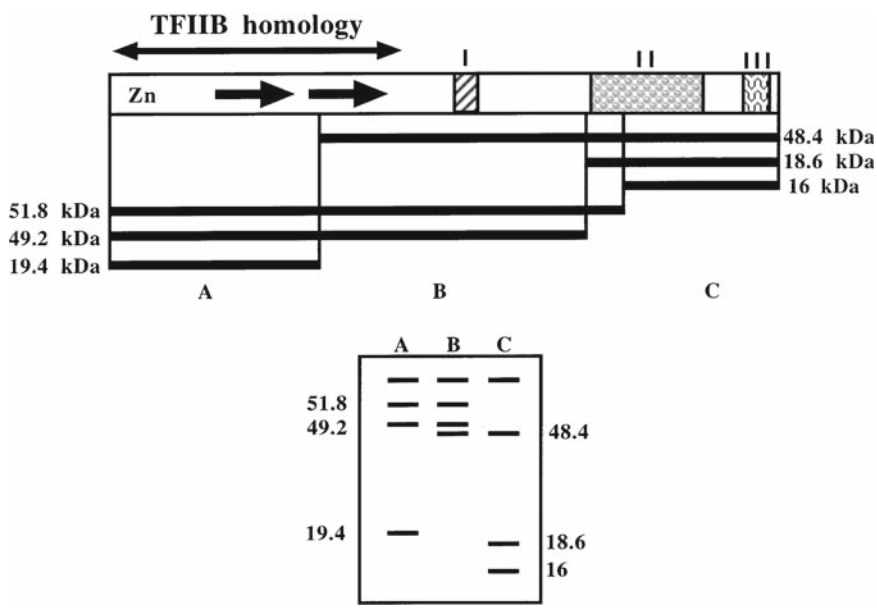


Fig. 8. Representation of possible results of peptide mapping of the BRF subunit of TF IIIB. At the *top* is a linear representation of the protein with Asp-Pro linkages labeled below along with the fragments generated by single-hit digestion. *Below* is a representation of protein gel analysis of labeled proteolytic products from crosslinking occurring in regions A, B, or C.

8. The sample was extracted five times with an equal volume of water-saturated ethyl ether (fresh).
9. Next samples are evaporated to dryness by vacuum centrifugation, resuspended in sterile deionized water, and dried down again.
10. The pellet is resuspended in 40 µl 2% SDS and 0.1 mM 2-mercaptoethanol.
11. Proteins are cleaved with cyanogen bromide by the addition of 1 µl of 1 M hydrochloric acid and 1 µl of 1 M cyanogen bromide in acetronitrile to a 15-µl sample or addition of formic acid to a final concentration of 70% and 1 µl of 1 M cyanogen bromide in acetonitrile for a more complete digestion.
12. Samples are incubated at 25°C for 10 min or 2 h.
13. Samples are resolved on a 10–20% tricine gel *(20)*. After electrophoresis the gel is stained by Coomassie R-250 staining, dried, and visualized by phosphorimaging (*see* **Note 18**) (**Fig. 8**).

4. Notes

1. We store concentrated stocks of modified nucleotides wrapped in foil at –80°C. Only small diluted stocks are stored at –20°C wrapped in foil to help protect the major stock from inadvertent photolysis.
2. We have found that various brands of TLC plates cause products to migrate somewhat differently and may result in different observed Rf values.
3. Photoreactive nucleotides are tested to determine the range of nucleotide concentrations that can be used to incorporate the nucleotides by primer extension and can be compared to the unmodified nucleotide. An aliquot of each sample can be irradiated and the DNA will crosslink the Klenow fragment of DNA polymerase I and cause the labeled oligonucleotide to have a much slower electrophoretic mobility. Some of the oligonucleotide does not get crosslinked to DNA polymerase but is visibly photolyzed as evident by smearing of the free oligonucleotide band.
4. Another question we have addressed is if a modified nucleotide at a given position affects normal DNA–protein interactions in that region. This can be done for transcription complexes by gel shift analysis or performing transcription assays on wild-type DNA and probe DNA to determine if DNA modification affects the level of transcription. Similar experiments are done with modified nucleosomes and chromatin remodeling

complexes through examining the efficiency of remodeling by changes in the electrophoretic mobility of the nucleosomes.

5. It is recommended to have at least 50 bps of DNA between the biotinylation site or attachment site to the bead and the restriction endonuclease cut site, because of potential steric hindrance of the restriction endonuclease if the site is too close to the bead.
6. DNA beads have a preference for binding shorter DNAs with a size restriction of less than ~2,000 bps in length. It is critical to determine the extent of binding the biotinylated DNA by agarose gel electrophoresis.
7. We have found the DNA beads to be sensitive to certain buffers containing high salt in combination with SDS. These conditions lead to the degradation of the magnetic particles, so it is important to avoid washing with buffers containing both SDS and high concentrations of salt.
8. Beads should always be mixed with gentle vortexing action to insure that they remain in solution and avoid vigorous vortexing because it will scatter beads and lead to sample loss.
9. Oligonucleotides for probe synthesis are usually between 18 and 20-mers in length with a GC-rich 5′ end.
10. If the DNA probe is to be used for peptide mapping, the modified nucleotide will have to be incorporated first, followed by dATP or dGTP, to leave a nucleotide tag covalently attached to the crosslinked protein upon chemical degradation of the DNA probe.
11. Enzymes containing DTT need to be avoided since DTT can reduce aryl azides. Some enzymes with DTT can be sufficiently diluted in buffers containing 2-mercaptoethanol (5 mM) in place of DTT.
12. Choosing restriction enzymes with good cutting efficiency to remove probes from the DNA beads can be very beneficial to overall probe yields.
13. Due to the short sequential incubation times it is best to cycle samples into and out of incubation baths to ensure equal incubation times for all samples. This means addition of the enzyme or solution to the first sample and immediately placing it in the temperature bath followed by processing the second sample in the same way and so on. The samples are removed in the same order so that they entered the bath with a 15-s delay between each sample.
14. We have found the crude extract BR-500 to be more efficient for photoaffinity labeling probably due to higher activities of the proteins and factors that may be lost during extensive purification.

15. Proper controls such as DNA competition with specific-pTZ1 (containing an up mutation in the box B region of DNA) vs. nonspecific-pLNG56 DNA (containing a down mutation in the Box B region of the DNA) and heparin striping of samples are necessary to ensure labeling specificity.
16. Crosslinked proteins have a slightly greater electrophoretic mobility than unmodified protein due to the DNA tag left behind and is more noticeable for smaller proteins.
17. Labeling reactions used for peptide mapping are scaled up to insure adequate signal. Sufficient amount of labeled protein is necessary due to losses during the extensive purification and the partial proteolytic conditions creating only a small percentage of labeled proteolytic fragments.
18. Care should be taken in constructing a map of potential proteolytic fragments created by single-hit digests for peptide mapping. Chemical cleavage of protein is preferable to enzymatic cleavage, because chemical cleavage has no apparent specificity or site preference that can make it difficult to interpret the peptide-mapping results.

References

1. Geiduschek, E. P., and Kassavetis, G. A. (2001). The RNA polymerase III transcription apparatus, *J Mol Biol* **310**, 1–26.
2. Kassavetis, G. A., and Geiduschek, E. P. (2006). Transcription factor TFIIIB and transcription by RNA polymerase III, *Biochem Soc Trans* **34**, 1082–1087.
3. Persinger, J., and Bartholomew, B. (2001). Site-directed DNA photoaffinity labeling of RNA polymerase III transcription complexes, *Methods Mol Biol* **148**, 363–381.
4. Gangaraju, V. K., and Bartholomew, B. (2007). Mechanisms of ATP dependent chromatin remodeling, *Mutat Res* **618**, 3–17.
5. Dang, W., and Bartholomew, B. (2007). Domain architecture of the catalytic subunit in the ISW2-nucleosome complex, *Mol Cell Biol* **27**, 8306–8317.
6. Bartholomew, B., Kassavetis, G. A., Braun, B. R., and Geiduschek, E. P. (1990). The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent, *Embo J* **9**, 2197–2205.
7. Lannutti, B. J., Persinger, J., and Bartholomew, B. (1996). Probing the protein–DNA contacts of a yeast RNA polymerase III transcription complex in a crude extract: solid phase synthesis of DNA photoaffinity probes containing a novel photoreactive deoxycytidine analog, *Biochemistry* **35**, 9821–9831.
8. Zofall, M., and Bartholomew, B. (2000). Two novel dATP analogs for DNA photoaffinity labeling, *Nucleic Acids Res* **28**, 4382–4390.
9. Tate, J. J., Persinger, J., and Bartholomew, B. (1998). Survey of four different photoreactive moieties for DNA photoaffinity labeling of yeast RNA polymerase III transcription complexes, *Nucleic Acids Res* **26**, 1421–1426.
10. Bartholomew, B., Braun, B. R., Kassavetis, G. A., and Geiduschek, E. P. (1994). Probing close DNA contacts of RNA polymerase III transcription complexes with the photoactive nucleoside 4-thiodeoxythymidine, *J Biol Chem* **269**, 18090–18095.
11. Persinger, J., and Bartholomew, B. (1996). Mapping the contacts of yeast TFIIIB and RNA polymerase III at various distances from the major groove of DNA by DNA photoaffinity labeling, *J Biol Chem* **271**, 33039–33046.
12. Galardy, R. E., Craig, L. C., Jamieson, J. D., and Printz, M. P. (1974). Photoaffinity labeling of peptide hormone binding sites, *J Biol Chem* **249**, 3510–3518.
13. Burgermeister, W., Nassal, M., Wieland, T., and Helmreich, E. J. (1983). A carbene-generating photoaffinity probe for beta-adrenergic receptors, *Biochim Biophys Acta* **729**, 219–228.

14. Persinger, J., Sengupta, S. M., and Bartholomew, B. (1999). Spatial organization of the core region of yeast TFIIIB-DNA complexes, *Mol Cell Biol* **19**, 5218–5234.
15. Kagalwala, M. N., Glaus, B. J., Dang, W., Zofall, M., and Bartholomew, B. (2004). Topography of the ISW2-nucleosome complex: insights into nucleosome spacing and chromatin remodeling, *Embo J* **23**, 2092–2104.
16. Sengupta, S. M., Persinger, J., Bartholomew, B., and Peterson, C. L. (1999). Use of DNA photoaffinity labeling to study nucleosome remodeling by SWI/SNF, *Methods* **19**, 434–446.
17. Sengupta, S. M., VanKanegan, M., Persinger, J., Logie, C., Cairns, B. R., Peterson, C. L., and Bartholomew, B. (2001). The interactions of yeast SWI/SNF and RSC with the nucleosome before and after chromatin remodeling, *J Biol Chem* **276**, 12636–12644.
18. Bartholomew, B., Tinker, R. L., Kassavetis, G. A., and Geiduschek, E. P. (1995). Photochemical cross-linking assay for DNA tracking by replication proteins, *Methods Enzymol* **262**, 476–494.
19. Bartholomew, B., Durkovich, D., Kassavetis, G. A., and Geiduschek, E. P. (1993). Orientation and topography of RNA polymerase III in transcription complexes, *Mol Cell Biol* **13**, 942–952.
20. Gallagher, S. R. (2007). One-dimensional SDS gel electrophoresis of proteins, *Curr Protoc Cell Biol* **Chapter 6**, Unit 6 1.

Chapter 28

Functional Studies of DNA–Protein Interactions Using FRET Techniques

Simon Blouin, Timothy D. Craggs, Daniel A. Lafontaine, and J. Carlos Penedo

Summary

Protein–DNA interactions underpin life and play key roles in all cellular processes and functions including DNA transcription, packaging, replication, and repair. Identifying and examining the nature of these interactions is therefore a crucial prerequisite understanding the molecular basis of these fundamental processes. The application of fluorescence techniques and in particular fluorescence resonance energy transfer (FRET) to provide structural and kinetic information has experienced a stunning growth during the past decade. This has been mostly promoted by new advances in the preparation of dye-labeled nucleic acids and proteins and in optical sensitivity, where its implementation at the level of individual molecules has opened a new biophysical frontier. Nowadays, the application of FRET-based techniques to the analysis of protein–DNA interactions spans from the classical steady-state and time-resolved methods averaging over large ensembles to the analysis of distances, conformational changes, and enzymatic reactions in individual Protein–DNA complexes. This chapter introduces the practical aspects of applying these methods for the study of Protein–DNA interactions.

Key words: Protein–DNA interactions, Fluorescence spectroscopy, Förster Resonance Energy Transfer, Time-resolved fluorescence, Single-molecule detection.

1. Introduction

Protein–DNA interactions are widespread. Understanding the molecular basis of these crucial biological mechanisms requires therefore a detailed analysis of the nucleoprotein complex structure and the dynamic interactions that govern its assembly and function. Owing to their sensitivity and the recent advances in

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_28

site-specific dye-labeling methods in both nucleic acids and proteins, fluorescence detection-based biophysical assays have proven to be very powerful and versatile techniques to probe the dynamics and function of Protein–DNA complexes *(1–3)*. The appeal of these fluorescence-based approaches relies on the extreme sensitivity of a fluorescence probe to its environment *(4)*, the possibility of monitoring the fluorescence signal continuously in real time to provide accurate kinetic data and, when combined with Förster Resonance Energy Transfer (FRET), the ability to provide insights into the structural basis of DNA or protein-induced conformational rearrangements *(5–7)*.

The basic idea underlying any FRET experiment designed to study a particular Protein–DNA complex relies on the site-specific labeling with a donor and an acceptor dye, located both in the same biomolecule (intramolecular FRET, **Fig. 1a**) or one FRET dye in each interacting partner (intermolecular FRET, **Fig. 1b**). Direct optical excitation of the donor dye results in fast energy transfer to the FRET acceptor, which emits fluorescence at a longer wavelength. The efficiency of this process depends on the sixth power of the average distance between the donor and the acceptor dye *(8–10)* and thus, the changes in fluorescence intensity from donor and acceptor can be used to monitor the interaction between proteins and their DNA substrate with extreme sensitivity and accuracy *(11–13)*.

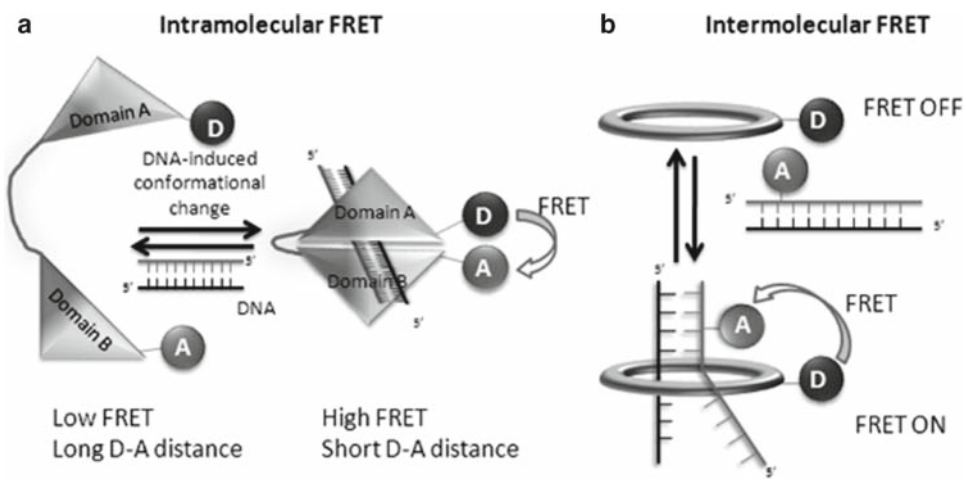


Fig. 1. Protein–DNA interactions can be monitored by intramolecular and intermolecular FRET assays. (**a**) Intramolecular FRET assays are based on the functionalization of the same biomolecule with both donor and acceptor fluorescence labels. Interaction with the partner biomolecule induces a conformational change in the host molecule (i.e., between an open and a closed state) that modifies the donor–acceptor distance and therefore the FRET value. (**b**) Intermolecular FRET assays are engineered with each fluorescence label in different biomolecules. When the two interacting molecules bind together and form a stable complex, FRET occurs. On the other hand, when the complex dissociates, the FRET value diminishes.

Intramolecular FRET assays where both dyes are located in the same biomolecule have been extensively used to monitor protein-induced conformational changes in the DNA substrate and to determine the global structure and assembly dynamics of a variety of nucleoprotein complexes. These studies include analysis of DNA bending upon interaction with a range of DNA-binding proteins such as the TATA protein *(14)*, the high-mobility group box HMG *(5, 15)*, the integration host factor protein IHF *(16, 17)*, the catabolite activator protein CAP *(7, 13)*, and the 5′ Flap Endonuclease FEN1 *(18)*. DNA strand exchange proteins such as RecA *(19, 20)* and its eukaryotic homologs hRad51 and scRad51 *(21)*, single-stranded binding proteins *(22)*, RecBCD-like nucleases *(23)*, and Holliday junction-resolving enzymes such as archaeal Hjc *(24)* have been the subject of intensive research using FRET-related techniques both at ensemble and single-molecule level. FRET has also been used to investigate the relative orientation of single-stranded DNA template primers with respect to the Klenow fragment of *E. coli* DNA Polymerase I *(25)*. More recently, FRET methods have been specifically developed to monitor the movement of RNA polymerase (RNAP) relative to DNA during transcription *(26–28)* and define the three-dimensional structure of transcription complexes in solution *(29)*. Apart from those studies that focused on the understanding of the molecular basis of DNA recognition outlined earlier, the investigation of cleavage reactions of nucleic acids catalyzed by a variety of enzymes is another major area where FRET techniques are already providing a wealth of kinetic information *(30)*. Usually, the efficiency of the enzymatic cleavage process is determined using a DNA substrate doubly labeled with donor and acceptor fluorophores. In the absence of enzymatic reaction, the proximity of both dyes enables efficient energy transfer from the donor to the acceptor, thereby decreasing the fluorescence intensity of the donor moiety *(31)*. Upon incubation with the enzyme, cleavage of the DNA substrate leads to the separation of the donor and acceptor dyes, with the concomitant cease of energy transfer and increase of the donor fluorescence. The main advantage of this fluorimetric assay when compared to more conventional biochemical techniques such as radiolabeling-based electrophoresis or ELISA-based techniques is that it relies on its continuous character, so that the cleavage reaction can be monitored from the initial steps in real time with no need for extensive sample handling. Following this approach, the kinetics of restriction endonucleases such as *PaeR7 (32)*, *Eco*RV *(33)*, S1 nuclease *(34)*, and Endonuclease V *(35)* have all been quantified by FRET techniques. Here, we present methods and protocols for FRET experiments that permit functional studies of Protein–DNA interactions including biomolecular conformational changes and cleavage assays.

1.1. Fluorescence Resonance Energy Transfer

Fluorescence Resonance Energy Transfer (FRET) is a nonradiative process whereby an excited donor fluorophore D transfers energy to a ground-state acceptor A (**Fig. 2a**) as a result of a through-space coupling of their transition dipoles *(8–10)*. According to Förster's theory, the energy transfer efficiency *E* depends on the inverse sixth power of the distance, *R*, separating the D–A pair:

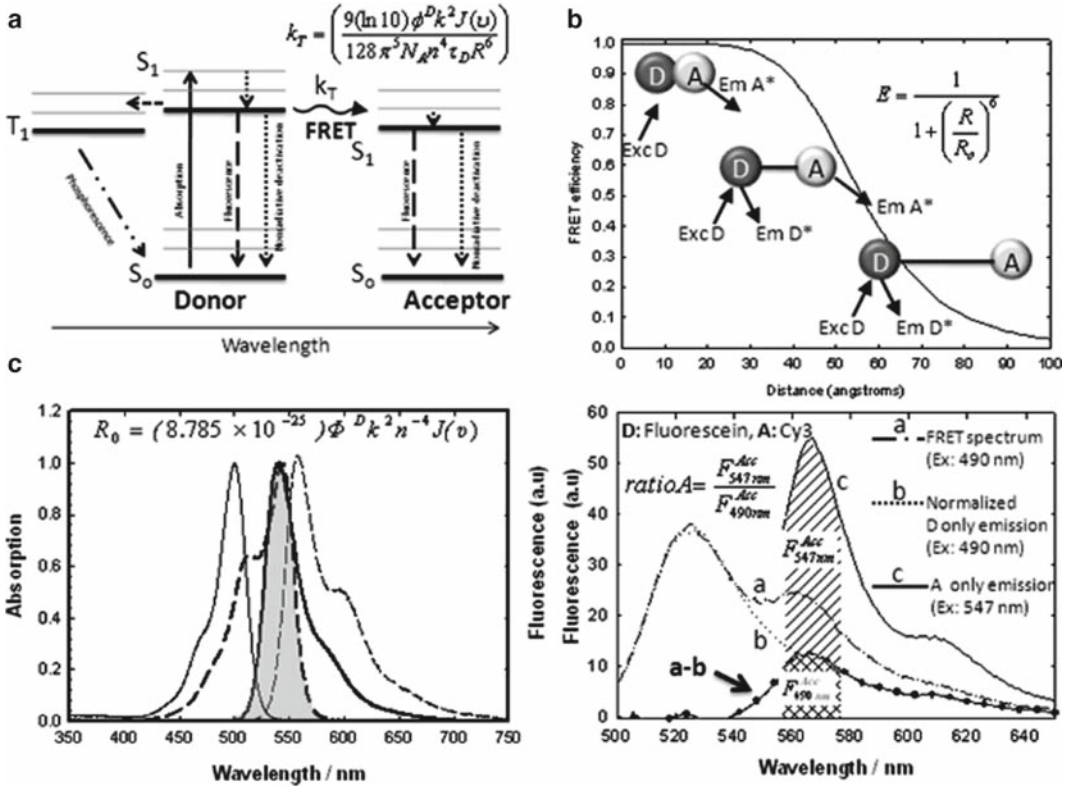


Fig. 2. FRET principle. (**a**) Absorption of a photon by the donor excites an electron to an upper state S_1. The excited donor can spontaneously lose its excess energy and decay to the ground state by a combination of competing processes including fluorescence emission, nonradiative deactivation by interaction with solvent molecules, intersystem crossing to the triplet state T_1 or, in the presence of a nearby acceptor molecule, by resonance energy transfer. The probability of energy transfer, proportional to k_T, decreases rapidly when the distance increases (*inset equation*). (**b**) Distance dependence of the energy transfer efficiency according to Förster theory (**Eq. 1**) for the Fluorescein (donor)-Cy3 (acceptor) FRET pair. The Förster radius R_0, (56 Å) that represents the distance at which 50% of the donor deactivates by energy transfer, is calculated according to (**Eq. 2**). Because of the sixth-power dependence of the energy transfer with the distance, FRET is more sensitive for changes in distance close to the R_0 value. (**c**) Absorption and fluorescence spectra for the FRET pair Fluorescein donor (*continuous line*) and Cy3 (*dashed line*). The overlapping region between donor emission and acceptor absorption is illustrated by the shaded region (inset: Expression for R_0 in terms of wavenumbers). (**d**) Example of steady-state FRET data analysis, for the particular case of fluorescein-Cy3 FRET pair, following the enhanced acceptor emission method. The fluorescence emission of the donor in a donor-only labeled biomolecule (b, *dotted line*) is obtained with excitation at its absorption maximum and normalized in the region 510–530 nm to the experimental FRET spectrum (a, *dash-dotted line*) excited at the same wavelength. The "pure acceptor emission" (**a**–**b**) is obtained by subtracting the normalized donor emission from the experimental FRET spectrum. A fluorescence spectrum of the acceptor only (**c**, *solid line*) is also taken with excitation at its absorption maximum (547 nm), where no donor absorption takes place. The ratio*A* parameter, proportional to the FRET efficiency, is calculated from the ratio $F_{490}^{acc}/F_{547}^{acc}$.

$$E = \frac{R_o^6}{R_o^6 + R^6} = \frac{1}{1 + \left(\frac{R}{R_o}\right)^6}, \quad (1)$$

where R_o is the distance between donor and acceptor at which 50% of the excited D* molecules decay by energy transfer (**Fig. 2b**), while the other half decays through other radiative and non-radiative deactivation channels. This critical transfer distance, the so-called Förster distance, is a characteristic property of the donor–acceptor pair used and can be calculated for a particular FRET pair from the spectral and photophysical properties of the donor and acceptor partners following the expression:

$$R_o^6 = \frac{9{,}000 \ln 10 k^2 Q_D J}{128 \pi^5 n^4 N_A}, \quad (2)$$

where Q_D is the donor's fluorescence quantum yield, n is the refractive index, N_A is Avogadro's number, k^2 is the orientation factor, and J represents the spectral overlap between the donor fluorescence emission spectrum and the acceptor absorption spectrum (**Fig. 2c**) obtained from the expression:

$$J = \int F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda. \quad (3)$$

where $F_D(\lambda)$ is the fluorescence emission of the donor normalized to a total peak area of 1 and $\varepsilon_A(\lambda)$ is the absorption spectrum of the acceptor normalized to the molar extinction coefficient. The quantum yield of the donor chromophore needs to be determined in the context of the biomolecule as this can markedly differ from that in free solution. The refractive index of the medium is assigned a value of 1.4 for biomolecules in aqueous solution. The k^2 factor describing the orientation of the donor–acceptor transition dipole can range from 0 for perpendicular orientation to 4 for parallel. When complete averaging of the relative orientation of the dyes is achieved during the excited-state lifetime of the donor, k^2 adopts a value of two-third, which has proven to be a reasonable approximation in most bimolecular environments. For a more detailed description on the influence of k^2 on the measured distance values, which is beyond the scope of this chapter, we refer to a recent review by van der Meer *(36)*. As an example, the distance dependence of the FRET efficiency for the fluorescein-Cy3 donor acceptor pair is illustrated in **Fig. 2b** (note that for $R = R_o$, the FRET efficiency is 0.5).

Experimentally, the most direct approach to measure FRET efficiencies is based on steady-state fluorescence techniques by measuring the intensity of the donor in the absence I_D and in the presence of acceptor I_{DA} according to **Eq. 4**. From here it is possible to recover the donor–acceptor distance following **Eq. 5**:

$$E = 1 - \frac{I_{DA}}{I_D}, \tag{4}$$

$$R_{DA} = R_o \left(\frac{I_{DA}}{I_D - I_{DA}} \right)^{1/6}. \tag{5}$$

1.2. Time-Resolved Fluorescence Resonance Energy Transfer

It is very often the case that interacting biomolecules are flexible structures showing a broad range of donor–acceptor distances. In these cases, the efficiency of energy transfer calculated from steady-state measurements overestimates the correct distances and represents an average over all possible conformational mixtures present in solution. When the number of different conformations, the relative equilibrium populations of each conformation and their corresponding true distances need to be resolved, time-resolved FRET (tr-FRET) techniques are the best choice *(37)*. In a time-resolved FRET experiment, the measured parameter is the fluorescence lifetime of the donor, which for most organic dyes range in the picosecond to nanosecond time range. In the absence and presence of an acceptor, the fluorescence intensity will decay according to the expressions in **Eqs. 6** and **7**, respectively:

$$I_D(t) = I_o \exp\left(\frac{-t}{\tau_D} \right), \tag{6}$$

$$I_{DA}(t) = I_o \exp\left[\left(\frac{-t}{\tau_D} \right) \left(1 + \left(\frac{R_o}{R} \right)^6 \right) \right], \tag{7}$$

where τ_D represents the lifetime of donor in the absence of acceptor and the second component in the exponential term represents the energy transfer rate. However, it is normally the case that a fluorescence probe attached to a biomolecule can adopt multiple conformations, each of them with an intrinsic fluorescence lifetime. In this case, the fluorescence decay should be modeled as a distribution of lifetimes as represented in **Eq. 8**, where f_i is the relative population of each species and $P_i(R)$ represents the distance distribution.

$$I_{DA}(t) = I_o \sum_i f_i \int P_i(R) \exp\left[-\frac{t}{\tau_D} \left(1 + \frac{R_o}{R} \right)^6 \right] dR \tag{8}$$

The two existing methodologies for measuring time-resolved FRET, phase modulation, and time-correlated single-photon counting (TCSPC), together with the procedures for data analysis, have been recently reviewed by Klostermeier and Millar *(37)*. The expressions outlined earlier are commonly applied to study conformational changes and extract distance information in proteins

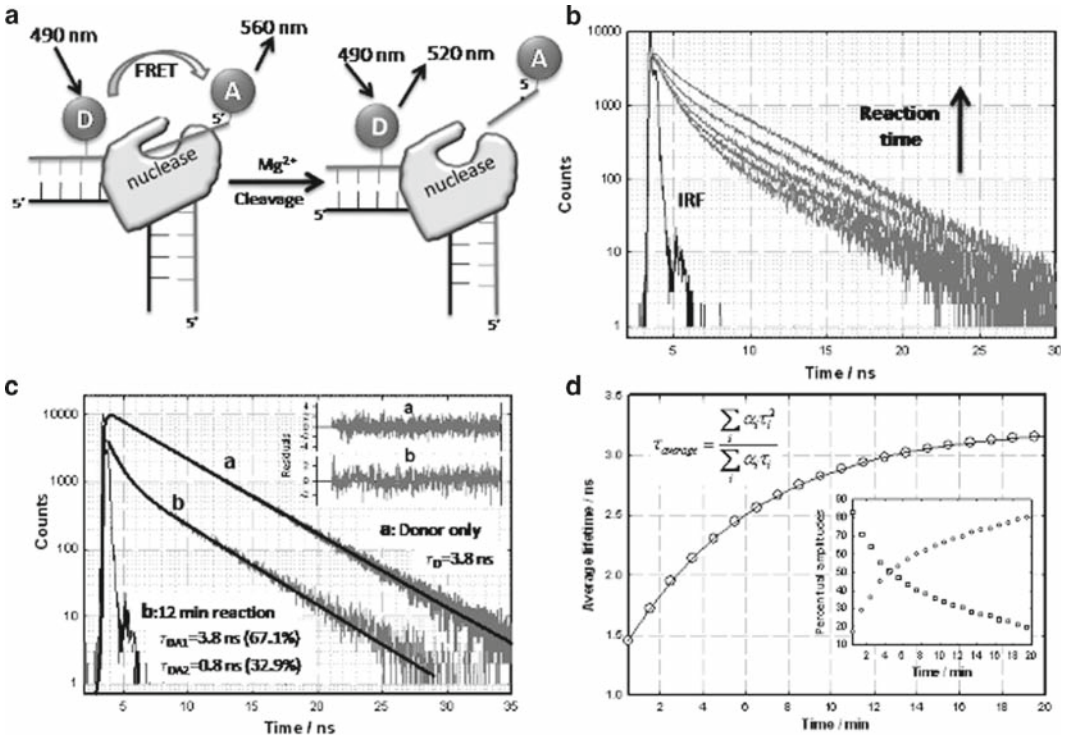


Fig. 3. Example of an enzymatic DNA cleavage assay monitored by time-resolved fluorescence resonance energy transfer. (**a**) A 3′-flap duplex DNA is labeled with donor (fluorescein) and acceptor (Cy3) groups. In the presence of magnesium divalent ions, binding of structure-specific endonucleases such as XPF (Xeroderma pigmentosum, complementation group F) induces cleavage of the 3′-flap with the subsequent release of the FRET acceptor and FRET breakdown (Penedo et al. unpublished data). (**b**) Monitoring the nuclease-induced cleavage of the DNA construct shown in (**a**) by time-resolved fluorescence spectroscopy. As the cleavage progresses, the biexponential decay together with time-dependent contributions from both product (donor only, τ_D = 3.8 ns) and uncleaved substrate (donor + acceptor, τ_{FRET} = 0.8 ns) evolves to a monoexponential donor-only decay. Each fluorescence decay trace was taken at fixed intervals of time (0.5 s) during the progression of the cleavage reaction. The excitation was performed by a 475-nm picosecond pulsed-diode laser (Edinburgh Instruments Ltd, UK) and the emission monochromator was placed at 520 nm, the donor emission maximum. Each fluorescence trace represents the time-correlated single photon-counting histogram accumulated during 0.5 s. The instrument-response function (IRF) used to deconvolute the experimental decays is also shown. (**c**) Comparison between the lifetime decays of the donor-only control and the doubly labeled DNA construct after 12 min progress of the cleavage reaction. The lifetime traces have been fitted to a monoexponential decay in the case of the donor only DNA control and a biexponential decay following the expression: $F(t) = F_0(t) + A_1\exp(-t/t_1) + A_2\exp(-t/\tau_2)$ for the doubly labeled substrate. The pre-exponential factors A_1 and A_2, representing the amplitudes of each exponential at time zero, are proportional to the amounts of product and uncleaved substrate. As an example at 12-min reaction time, the remaining uncut substrate constitutes ~33% of the total intensity. Inset: residuals obtained from the fitting of the experimental decays to a monoexponential function (**a**, donor-only control) and a biexponential function (**b**, doubly labeled substrate). (**d**) Plot showing the average lifetime calculated accordingly (**Eq. 9**) as a function of cleavage time. Inset: percentual amplitudes corresponding to product and uncut substrate as a function of cleavage time. The complete dataset of decay traces obtained every 0.5 s was fitted to biexponential decay following global analysis (Edinburgh Instruments Ltd, UK).

or DNA. However, in enzymatic cleavage assays followed by tr-FRET, it is usually more practical to analyze the changes in the average donor lifetime (**Eq. 9**) as a function of the progress of the cleavage reaction (**Fig. 3d**). In these assays, a DNA substrate labeled with donor and acceptor is engineered (**Fig. 3a**)

so that, due to the enzymatic reaction, the acceptor moiety will be released with the concomitant increase in donor lifetime and breakdown of FRET over time.

$$\overline{\tau} = \frac{\sum_i \alpha_i \tau_i^2}{\sum_i \alpha_i \tau_i}. \quad (9)$$

1.3. Single-Molecule FRET

More recently, FRET experiments at single-molecule level (Sm-FRET) have become possible providing information on protein-nucleic acid dynamics that was previously hidden when using bulk-solution methods *(38–41)*. Single-molecule FRET techniques provide a completely new approach to study the structure-dynamics-function relationship in protein–nucleic acid complexes, allowing the identification of subpopulations in a heterogeneous mixture, the analysis of Protein–DNA complexes lifetimes and the recovery of FRET efficiency distributions, with the advantage compared to time-resolved FRET that it does not require any assumption about the shape of the distance distribution. The efficiency of energy transfer in a single-molecule experiment is usually calculated from the expression:

$$E = \frac{I_A}{I_D + \alpha I_A}, \quad (10)$$

where I_D and I_A represent the fluorescence intensity of the donor and acceptor, respectively. The α parameter corrects for the donor leakage into the acceptor channel and direct excitation of the acceptor at the donor excitation wavelength. All these corrections can be done with donor- and acceptor-only labeled species *(39, 42)*. Single-molecule FRET experiments can be performed either on freely diffusing molecules *(43, 44)*, or on immobilized molecules to record the trajectory of the molecule for extended periods of time *(42, 45)*, but in this case care must be taken to ensure that the immobilization techniques are compatible with the biomolecule(s) under study *(46)*. Single-molecule FRET studies on Protein–DNA interactions have mostly focused on the unwinding of DNA by helicases *(47, 48)*, DNA damage *(49, 50)*, and the analysis of the relative movement of RNAP on the DNA template during the initiation and elongation steps of the transcription processes *(28, 29)*.

2. Materials

2.1. Instrumentation

For steady-state FRET experiments, any scientific-grade commercial fluorimeter equipped with conventional Xe lamps and

with real-time correction for fluctuations in lamp intensity will be appropriate. For time-resolved FRET two different experimental approaches can be used *(37)*, either the frequency-domain method or the time-domain method. In the frequency domain method, the sample is excited with a sinusoidally modulated laser of certain frequency and the fluorescence lifetime is extracted from the phase delay and the demodulation of the emitted light. In the time-domain method, the sample is excited with a short laser pulse, typically of duration ~80 ps or less, and the fluorescence emission is detected with a fast photomultiplier or multichannel plate (MCP) and a time-correlated single-photon counting system (TCSPC). Either in steady-state or in time-resolved conditions, the fluorescence signal must be collected at magic angle conditions, with the emission polarizer placed at 54.7° relative to the vertically polarized excitation, to avoid artifacts due to molecular reorientation.

For single-molecule experiments the two most common approaches are fluorescence correlation spectroscopy (FCS), where molecules are allowed to freely diffuse through the excitation volume, and those techniques requiring the immobilization of the biomolecule under study to allow observation for longer periods of time: total-internal reflection (**Fig. 4**) and scanning confocal microscopy. Single-molecule instrumentation for FCS is already commercially available from a range of manufacturers (Leica, Olympus, Picoquant), and the details of single-molecule FCS instrumentation can be found in several recent reviews *(43, 44)*. For immobilization single-molecule experiments, there is no specific commercial equipment as for FCS, so scientists working in the field need to build their own equipment. It will be assumed in the following that a single-molecule total-internal reflection (Sm-TIR) setup is already available. Single-molecule instrumentation share in common the need for a laser excitation source, an inverted microscope with oil- or water-immersion objectives depending on the particular setup (*see* **Note 1**), and a detection system. Depending whether the aim is to monitor many molecules simultaneously or just one but with higher time resolution, the acquisition system will be based on two-dimensional detectors, such as intensified or electron-multiplying CCD cameras (Andor, UK or Princeton Instruments, USA) or point-detection devices such as avalanche photodiodes APDs (Perkin Elmer, USA). The advantage of FCS is that it is a true solution-based technique and therefore free of perturbations by biomolecule-surface interactions. The main disadvantage is that the observation time is limited to a few milliseconds, the diffusion time of the biomolecule through the confocal volume. To observe single-molecule events for long periods of time, immobilization of the single-molecule on an appropriate surface is the method of choice. Several different immobilization methods have been used depending on the

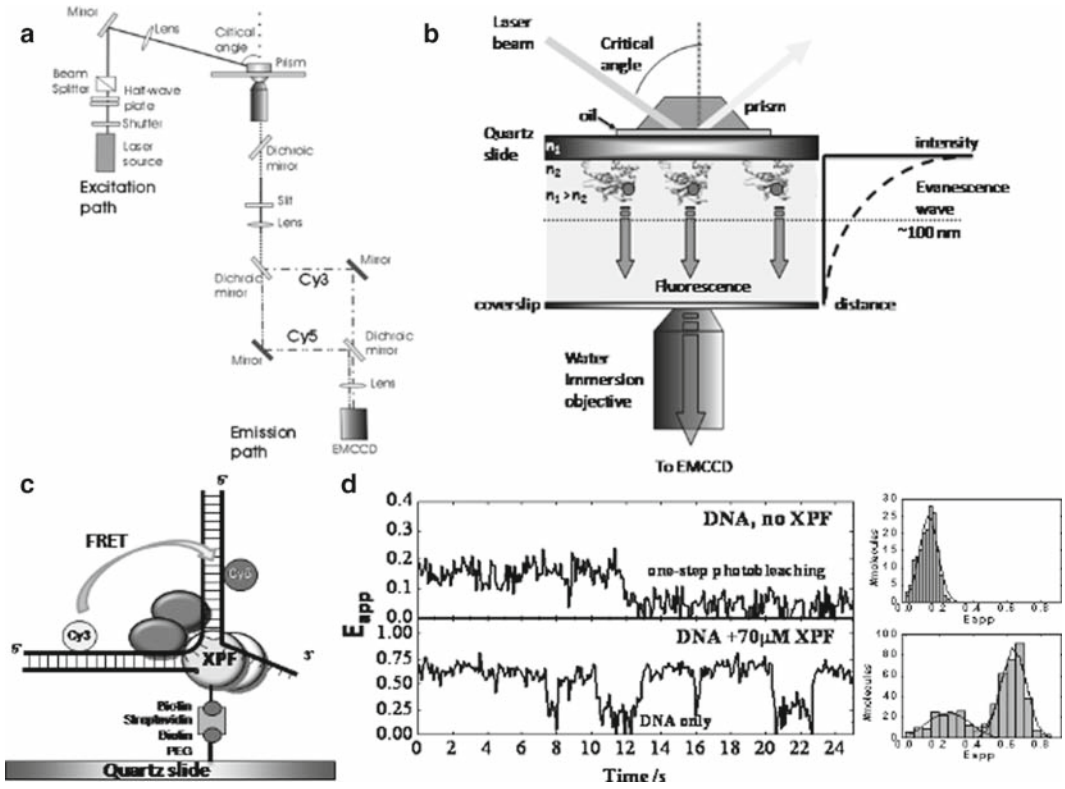


Fig. 4. Single-molecule FRET assays for Protein–DNA interactions by total-internal reflection spectroscopy. (**a**) Schematics of a single-molecule TIR microscope for single-molecule FRET. Briefly, the sample is excited by a continuous wave diode laser (532 nm, Crystalaser) at a critical angle of ~65° by using a prism-type setup. The fluorescence obtained from the sample is collected by a water-immersion objective (60×, NA 1.2, WD: 0.17–0.25) mounted on an inverted microscope (Olympus IX71). A dichroic mirror located underneath the objective is used to block the laser excitation light and allow the fluorescence from donor (Cy3) and acceptor (Cy5) to reach a 3-mm slit (Thorlabs, UK). Using a combination of reflecting and dichroic mirrors, the fluorescence is decomposed in its donor (reflected) and acceptor (transmitted) components and spatially separated so that the donor reaches the left-hand side of the CCD camera chip (Andor Technology, UK) and the acceptor reaches the right-hand side. (**b**) Principle of total-internal reflection illumination. Refractive index differences between the quartz slide (n_1) and water phases (n_2) modulate how much light is refracted or reflected at the interface as a function of angle of incidence of the laser beam. At a specific critical angle, the beam is completely reflected from the quartz/water interface. This reflection generates an evanescence wave (~100 nm) in the aqueous medium, which is used to excite the molecules directly attached to the quartz inner surface. Because the intensity of this evanescence wave decays exponentially with the distance from the surface, very small excitation volumes can be easily achieved, increasing the signal-to-noise ratio to appropriate values for single-molecule detection. (**c**) doubly labeled DNA carrying a 3′-flap engineered for single-molecule binding studies. The construct is immobilized in the quartz surface, which was previously coated with biotinylated-PEG, via biotin–streptavidin interactions. Binding of the structure-specific endonuclease XPF induces a conformational change in the DNA that reduces the distance donor–acceptor, thus increasing the FRET signal as it can be seen in the single-molecule traces represented in (**d**) in the presence and absence of protein. Fluctuations in the FRET signal (bottom plot) are indicative of protein binding-dissociation.

type of biomolecule. These include single-point attachment on a glass or quartz surface via BSA-Streptavidin interactions *(51)*, poly(ethylene glycol) (PEG)-coated surface *(52, 53)* (*see* **Note 2**), His–NTA interactions *(54, 55)*, trapping inside porous gel matrices

of agarose or polyacrylamide *(50)*, and inside biomimetic membranes such as vesicles *(56)* and nanopores *(47)*. Independent of the immobilization method, specific control experiments should be made for each particular system to ensure that the biomolecule is not interacting with the surface, so that any observed heterogeneity is intrinsic and does not arise from molecular interactions with the surface *(46)*.

2.2. Chemicals

Functional studies of Protein–DNA interactions by fluorescence resonance energy transfer require the labeling of the DNA or the protein *(57)*, or both, with a FRET pair suitable for the distance range to be measured, and the subsequent purification steps to get reliable FRET efficiency and distance values. In the particular case of single-molecule applications we should emphasize the need for reagents, buffers, and even water (*see* **Note 3**), absolutely free of fluorescence contaminants. Common buffers such as Tris, phosphates, widely used monovalent and divalent ions solutions and denaturing agents can be obtained with excellent purity from most suppliers. Specific single-molecule additives such as surface-coating molecules (BSA, PEG), oxygen scavenger proteins (glucose oxidase and catalase), and triplet-state quenchers (2-Mercaptoethanol, TROLOX) can be obtained with high purity from more specialized sources (Roche Diagnostics, Pierce Biotechnology, New England Biolabs).

2.2.1. Synthetic DNA

Synthesis of custom DNA oligonucleotides is a procedure now outsourced by most laboratories. Companies such as Operon, DNA Integrated Technologies, and Invitrogen offer custom DNA synthesis at micromole and nanomole scales up to 145–150 nucleotides long. For single-molecule immobilization studies biotin is usually incorporated during chemical synthesis at the 5′ or 3′ end.

2.2.2. Fluorophores

1. A variety of fluorescence molecules can be used as donor and acceptor pairs for FRET studies *(57, 58)*. **Table 1** summarizes the spectroscopic properties of most widely used FRET pairs. Fluorescein-tetramethylrhodamine and fluorescein-Cy3 are very popular FRET pairs for bulk-solution investigations. However, for single-molecule studies, especially under immobilization conditions, Cy3–Cy5 is the most accepted FRET pair because of its wider spectral separation when compared to fluorescein-Cy3 and its higher photostability.
2. The same companies that provide synthetic DNA oligonucleotides can also provide fluorescently labeled-DNA with most common donor and acceptor dyes (Fluorescein, Cyanine and Alexa dyes) at the 5′ and 3′ end terminals and also internal Fluorescein-dT. However, the maximum length available for these fluorescently labeled constructs is ~60–70 nucleotides.

Table 1
Summary of spectroscopic properties, donor lifetime values, and Förster radius R_o for some of the most common FRET pairs used in conventional steady-state and time-resolved FRET, together with those particularly useful in single-molecule FRET applications because of their enhanced photostability

Donor	λ_{exc} (nm)/λ_{em} (nm)	τ_{Donor}/ns	Acceptor	λ_{exc} (nm)/λ_{em} (nm)	R_o
EDANS	336/468	13.0	DABCYL	471/nonfluorescent	33
IAEDANS	336/490	12.0	Fluorescein	494/521	46
GFP	395,475/509	2.6	DsRed	558/583	47
CFP	439/476	3.4(0.6), 1.3(0.4)[a]	YFP	514/527	50
Fluorescein	494/521	4.1[b]	TMR	550/572	55
Fluorescein	494/521	4.1[b]	Cy3	547/565	56
FRET pairs suitable for single-molecule studies					
Cy3	547/560	0.3	Cy5	647/667	53
Cy5	647/667	1	Cy5.5	675/694	--
Alexa 488	495/519	3.9	Alexa 568	578/603	62
Alexa 546	556/573	4.0	Alexa 594	590/617	71
Alexa 594	590/617	3.9	Alexa 647	650/668	85

[a]The numbers in brackets represent the normalized pre-exponential factors or amplitudes at time zero obtained from time-resolved experiments

[b]Fluorescein incorporated in a duplex DNA shows a second component with a lifetime of ~0.8 ns and a small amplitude (<15%) indicating different rearrangements of the fluorophor in the DNA duplex

2.2.3. DNA Postsynthesis Fluorescence Labeling

An alternative to the incorporation of fluorescence labels on the DNA during chemical synthesis relies on the postsynthesis coupling of a succinimidyl ester (NHS) dye derivative (GE Healthcare, Molecular Probes, Invitrogen) to a synthetic DNA carrying a primary amino group at a specific position. This method is widely used to incorporate internal fluorescence labels other than the commercially available fluorescein. For this labeling procedure the following reagents are required:

1. Synthetic single-stranded DNA oligonucleotides carrying a primary amino group modification (Operon, IDT)
2. Cy3-NHS, Cy5-NHS, or other appropriate dye carrying a succinimidyl ester group.
3. Dimethyl sulfoxide (e.g., Fisher Scientific).
4. Labeling buffer: 100 mM sodium tetraborate (Sigma), pH 8.5. Keep this solution aliquoted in the freezer for long-term

storage and avoid long-term exposition to air, which could change the buffer pH.

5. Precipitation solution: 3 M sodium acetate and 100% ethanol. Keep these solutions at room temperature.

2.2.4. Protein Labeling

Currently the most common approach is cysteine labeling with a thiol-reactive maleimide derivative of the dye *(57)*.

1. Protein with a surface-exposed cysteine residue.
2. Labeling buffer: 100 mM phosphate-buffered saline between pH 7.0 and 7.5 Tris or HEPES buffer could also be used.
3. Dimethyl sulfoxide (e.g., Fisher Scientific)
4. Maleimide derivative of the fluorescence dye (Invitrogen, Molecular Probes, GE Healthcare).
5. Dithiothreitol (DTT) (Sigma) or Tris(2-carboxyethyl)phosphine HCl (TCEP) (Sigma) as reducing agents for disulfide bonds. TCEP has the advantages of providing no pungent odor, and it is often not needed to remove TCEP before thiol modification using iodoacetamides or maleimides.
6. Glutathione or mercaptoethanol (Sigma).

2.2.5. Fluorescently Labeled DNA Purification by Polyacrylamide Gel Electrophoresis (PAGE)

1. *5× TBE.* 450 mM Tris, 450 mM borate, and 10 mM ethylenediamine tetraacetic acid disodium salt (EDTA).
2. *Dilution buffer 1×.* 7 M urea dissolved in 1× TBE (store at room temperature).
3. 20% acrylamide/bis solution (19:1) with 7 M urea in 1× TBE. This solution is neurotoxic when not polymerized and so care should be taken not to receive exposure. Keep this solution in the dark at room temperature.
4. Ammonium persulfate (APS): prepare 10% (w/v) solution in water. Keep this solution in the dark at 4°C. (*see* **Note 4**).
5. TEMED (N,N,N′,N′-Tetramethyl-ethylenediamine).
6. Running buffer (1×): 1× TBE.
7. *Molecular weight markers.* Xylene cyanol FF 0.02% and bromophenol blue 0.02% dissolved in pure formamide.

2.2.6. DNA Recovery

1. Electroeluter from Harvard lab shops (http://www.mcb.harvard.edu/bioshop).
2. Ammonium acetate (NH_4OAc) 8 M in water. Keep the solution at room temperature.
3. Electroeluter running buffer: 0.25× TBE.
4. Precipitation solution: 3 M NaOAc and 100% ethanol. Keep these solutions at room temperature.
5. Elution buffer for crush and soak elution: 0.5 M NH_4OAc, 1 mM EDTA, and 0.1% (w/v) sodium dodecyl sulfate (SDS).
6. Microspin columns (GE Healthcare).

2.2.7. Protein Purification

1. Gel filtration column such as Sephadex G-25, BioGel P-30, or similar with the appropriate molecular weight cutoff (GE LifeSciences).
2. 100 mM phosphate-buffered saline between pH 7.0 and 7.5.

2.2.8. Single-Molecule Immobilization

1. Spectroscopic-grade acetone and methanol (Sigma Ltd) and acetic acid.
2. Pegylation buffer: 10 mL ultrapure water + 84 mg sodium bicarbonate
3. Aminopropyl silane (Sigma), Biotinylated-PEG and mPEG (SusTech, Darmstadt, Germany). (*see* **Note 5**)
4. Streptavidin was purchased from Invitrogen and dissolved at a concentration 5 mg/mL in 50 mM Tris–HCl (pH 8.1), 50 mM NaCl. The solution should be stored at 4°C immediately after preparation
5. Imaging buffer: 50 mM Tris–HCl (pH 8.1), 6% (w/w) glucose, 1% 2-mercaptoethanol, 0.1 mg/mL glucose oxidase type II-S from *Aspergillus niger* (Sigma), 0.02 mg/mL glucose catalase (Roche Diagnostics) (*see* **Notes 6** and **7**).
6. Quartz microscope slides 1 in. × 3 in. × 1 mm (Finkenbeiner, Waltham) for prism-type Sm-TIR
7. Fluorescence beads for instrument alignment and to generate the mapping algorithm that allows correlating individual donor spots with individual acceptor spots can be purchased from Molecular Probes (FluoSpheres carboxylated, 0.2 μm). Commercial stock solution should be diluted 1,000-fold in 5 mM HCl before injecting on the quartz slide (*see* **Note 8**).

3. Methods

Independent of whether the FRET assay is going to be performed under steady-state, time-resolved or single-molecule regime, the first steps involve the labeling and purification of the biomolecules under investigation. We assume at this stage that the reader has identified an appropriate FRET pair for the particular question to be addressed. We will start with a description of general methodologies for protein dye-labeling, postsynthetic fluorescent labeling and purification of DNA, and we will continue with a general description of the protocol to follow under each experimental regime.

3.1. Preparation of Fluorescently Labeled DNA

1. Dissolve the amine-modified oligonucleotide in 100 μL of doubly distilled water.

2. Precipitate overnight at –20°C by adding 0.1 volume of NaOAc (3 M) and 2.5 volumes (250 μL) of 100% ethanol (*see* **Note 9**).
3. After centrifugation of the solution at ~16,000 × *g* for 30 min, carefully discard the supernatant and perform a quick spin to remove remaining ethanol. Air dry pellets on bench or under vacuum.
4. Resuspend pellet in water to achieve a final concentration of 25 μg/μL. The solution is stable at –20°C for long periods. Oligonucleotide concentration can be obtained from the absorption spectrum assuming that an optical density of 1.0 corresponds to a DNA concentration of 33 μg/mL in a 1-cm path cell.
5. Dissolve the commercial amine-reactive dye (commonly a succinimidyl ester derivative of the dye) in DMSO to reach a concentration of 250 μg/14 μL. This concentration of dye is optimized to label 100 μg of DNA. Remaining stock of dye in DMSO should be dried under vacuum until next use.
6. Mix 4 μL of the DNA strand to be labeled with the amine reactive dye (250 μg/14 μL) in a total volume of 100 μL of labeling buffer.
7. Incubate the reaction overnight at room temperature with gentle shaking, particularly during the first hour.
8. Precipitate the reaction mixture overnight at –20°C as described in **step 2**. Centrifuge at ~16,000 × *g* for 30 min and carefully remove the supernatant. Redissolve the pellet in 100 μL of 50% formamide and purify as described in **Subheadings 3.2** and **3.3**.

3.2. Purification of Labeled DNA by Polyacrylamide Gel Electrophoresis

1. To purify the fluorescently labeled DNA, pour a 1.5-mm-thick polyacrylamide gel. Acrylamide percentage should be chosen according to the DNA length. In this laboratory we normally use 20% for oligonucleotides less than 25 bases, 15% between 25 and 40 bases, and 10% for longer oligonucleotides. For a 10% gel we mix 30 mL of 20% (w/v) acrylamide (7 M urea) with another 30 mL of 7 M urea/1× TBE. Adding 450 μL of ammonium persulfate (10% w/v) and 45 μL of TEMED will start the polymerization reaction. Leave the gel for at least an hour to ensure complete polymerization.
2. Because urea accumulates in the wells, before prerunning the gel for 1 h at 18 Watts, wash the wells with 1× TBE running buffer using a syringe.
3. Incubate the formamide solution containing the labeled DNA for 2 min at 90°C to disrupt any secondary structure.
4. Just before loading samples, rewash loading wells with a syringe. Use xylene cyanol FF and bromophenol blue in 100%

formamide as size markers. On a 10% gel, xylene cyanol and bromophenol dyes migrate approximately as oligonucleotides of 55 and 10 nucleotides, respectively (*see* **Note 10**).

5. Run the gel until the dyes have migrated a minimum of 2/3 of the plate length. Remove glass plates and carefully put the acrylamide gel on a DarkReader transilluminator (DR45M, VWR International) to visualize DNA bands by fluorescence.

Cut the bands containing the proper length of fluorescent DNA and also the bands corresponding to xylene cyanol or bromophenol blue as control for electroelution. Samples can either be stored in a freezer or used immediately for electroelution.

3.3. DNA Recovery

3.3.1. Recovery by Electroelution

We describe here the protocol to be followed when using a Harvard-type electroeluter. Different steps of this protocol might need to be adapted to each specific electroeluter model.

1. Fill the electroeluter apparatus with 0.25× TBE and run at 200 V for 30 min to clean it. Discard the running buffer, fill the electroluter tank with freshly made 0.25× TBE buffer and add 200 μL of 8 M ammonium acetate in the trapping wells.
2. Carefully place the acrylamide bands in small pieces in the recipient wells and as an indication of the electroelution progress place also a marker band in a separate well. Run the electroluter at 120 V for 1 h or until the maker band is completely free of dye.
3. Block the trapping wells using tips cut at ~5 mm length and remove the buffer from the electroluter tank using a syringe.
4. Collect the fluorescent DNA from each trapping well in separate Eppendorfs and wash each well with running buffer to collect residual DNA sample left.
5. Fill each Eppendorf with 100% ethanol and store overnight at −20°C for precipitation.
6. Resuspend the pellet in appropriate buffer (i.e., 50 mM Tris–HCl, pH 8.0) and calculate the labeling efficiency using an absorption spectrophotometer.

3.3.2. Recovery by Crush and Soak

1. For crush and soak, place the acrylamide band cut in small pieces in a microspin column (GE Healthcare) and fill with buffer to cover all the acrylamide pieces. Seal the column and leave it overnight with gentle shaking.
2. Cut the sealed end of the column, place the column inside an Eppendorf, and centrifuge ~16,000 × *g* for 2 min. Add 1/10 volume of 3 M NaOAc and proceed as in **steps 5** and **6**.

3.4. Preparation of Fluorescently Labeled Protein

1. Dissolve the protein at 50–100 μM in 10 mM Tris or Phosphate buffer at pH 7.0–7.5 at room temperature.

2. Prepare a 1-mM stock solution of the maleimide dye in DMF or DMSO. Protect all dye stock solutions from light as much as possible.
3. Add the maleimide reagent dropwise to the protein solution to give approximately 10 mol of reagent per mole of protein.
4. Allow the reaction to proceed for 2 h at room temperature or overnight at 4°C (*see* **Note 11**).

3.5. Purification of Fluorescently Labeled Protein

1. Sephadex G-25 or similar gel filtration media of the appropriate molecular weight cutoff can be used to separate the protein-dye conjugate from the unreacted labeling reagent. For proteins larger than 5,000 Mr we normally use PD-10 columns (GE Healthcare) with excellent separation.
2. Cut the sealed bottom of the PD-10 column and fill up the column with 25 mL of equilibration buffer (i.e., 50 mM Tris–HCl, pH 7.5) and discard the flow through.
3. Apply the protein sample to the column (maximum: 2.5 mL). If the protein sample to be purified is less than 2.5 mL, add equilibration buffer to adjust the total volume to 2.5 mL. Let the sample pass through and discard the flow through.
4. Elute the protein with 3.5 mL of equilibration buffer and collect each fraction in a test tube or an Eppendorf. The first eluted fluorescent band will correspond to the dye-protein conjugate. Labeled samples should be stored under the same conditions as the unlabeled protein, but protecting the stock as much as possible from light.
5. Calculate the degree of labeling using the expression:

$$\frac{\text{moles dye}}{\text{moles protein}} = \frac{A_{\text{dye}}}{\varepsilon_{\text{dye}}} \times \frac{\text{MW protein}}{\text{mg protein} / \text{mL}}.$$

3.6. Steady-State FRET Experiments

Steady-state FRET experiments can reveal information about Protein–DNA interactions by monitoring (a) the increase in FRET signal as a consequence of formation of the Protein–DNA complex using an intermolecular FRET assay (**Fig. 1a**), (b) the change in intramolecular FRET signal (**Fig. 1b**) as a result of a binding-induced conformational change either in the protein or in the DNA, and (c) the breakdown of the inter- or intramolecular FRET upon protein-induced DNA cleavage (**Fig. 3**). The following protocols describe the basis to perform these assays in a quantitative manner.

3.6.1. Protein–DNA Binding Assay

1. Load 120 μL of a nanomolar substrate solution in a clean quartz microcell and place it in the sample chamber of the fluorimeter, thermostated at the appropriate temperature. In an intermolecular-type assay, the substrate DNA solution will

contain only the donor fluorophor, while in an intramolecular-type assay it will contain a donor–acceptor labeled construct (*see* **Note 12**).

2. Titrate the substrate solution with increasing concentrations of binding partner by adding aliquots of stock solution. For each addition, allow to equilibrate for 5 min and take two emission fluorescence spectra, one exciting at the maximum of the donor and a second one at a wavelength that excites only the acceptor species. We usually collect the fluorescence spectra with 8-nm slit width and 10 nm/s acquisition rate (*see* **Note 13**).
3. If the substrate is labeled with donor and acceptor, an equivalent construct carrying only the donor species should be prepared and used to independently collect the emission fluorescence spectrum of the donor in the absence of acceptor.
4. Normalize the donor-only spectrum obtained in **step 3** to the maximum of the donor fluorescence band in each FRET spectrum in the titration and subtract it from the latter (**Fig. 2d**).
5. The ratio$A$ parameter is proportional to the amount of acceptor fluorescence emission due to energy transfer (**Fig. 2d**) and can be calculated for each data point in the titration by first normalizing the donor-only fluorescence spectrum to the maximum of the donor band in the FRET experimental spectrum, and then by subtracting the normalized donor spectrum from the experimental FRET spectra. The resulting fluorescence curve contains only acceptor signal. The ratio$A$ is then obtained following the expression:

$$\text{ratio}A = \frac{F_{\text{acc}}^{\text{excD}}}{F_{\text{acc}}^{\text{excA}}},$$

where $F_{\text{acc}}^{\text{excD}}$ represents the fluorescence spectrum of the acceptor obtained from the experimental FRET spectrum after subtracting the normalized donor contribution, and $F_{\text{acc}}^{\text{excA}}$ represents the fluorescence spectrum of the acceptor obtained by exciting at its absorption maximum or at a wavelength where only the acceptor absorbs.

6. The ratio A parameter can then be converted in energy transfer efficiency values by applying:

$$E = \frac{\text{Ratio}A - \dfrac{A_{\text{acc}}^{\text{excD}}}{A_{\text{acc}}^{\text{excA}}}}{\dfrac{A_{\text{don}}^{\text{excD}}}{A_{\text{acc}}^{\text{excA}}}},$$

where $A_{\text{acc}}^{\text{excD}}$ represents the absorption of the acceptor at the donor excitation wavelength (i.e., for fluorescein, 490 nm), $A_{\text{acc}}^{\text{excA}}$ represents the absorption of the acceptor at its excitation wave-

length (i.e., for Cy3, 547 nm), and A_{don}^{excD} represents the absorption of the donor at its excitation wavelength (*see* **Note 14**).

3.6.2. Protein–DNA Cleavage Assay

1. Proceed as in **step 1** of **Subheading 3.6.1** *Protein–DNA binding assay*, but instead of collecting the emission of the donor and acceptor as a function of the wavelength, measure a time trace by exciting the donor and collecting simultaneously the emission of the donor and acceptor against time, by continuously switching the monochromator between both wavelengths.
2. Before the enzyme is added to the cell, the relative intensities of both signals and therefore the FRET efficiency should be constant. Then add the enzyme and mix manually with the pipette. Interaction between the substrate and the enzyme will be reported as an increase in the donor signal with a concomitant exponential decrease in FRET efficiency. Monitor the cleavage reaction until both signals reach a plateau.
3. Fit the measured exponential increase of the donor signal or the exponential decrease of the FRET efficiency to an equation of the type: $F(t) = F_o(t) + Ae^{-kt}$, where A represents the amplitude and k is the pseudo-first-order rate constant.

3.7. Time-Resolved FRET Experiments

In tr-FRET, the fluorescence lifetime of the donor is measured in the absence and in the presence of the acceptor fluorophore as a function of increasing concentrations of interacting partners (**Fig. 3b**, **c**). In binding assays, the difference in the lifetimes in the presence and absence of acceptor is used to obtain, for each data point in the titration, the Gaussian distribution of distances that represent the populations present (Eq. 8). In cleavage assays, either the average lifetime (**Fig. 3d**) or the amplitude of substrate and product (**Fig. 3d** inset) is plotted against time to extract the pseudo-first-order cleavage rate constant. Here, we will describe the protocols for these types of experiments using time-resolved FRET based on the TCSPC technique.

3.7.1. Time-Resolved Protein–DNA Binding Assays

1. Place a scattering solution in the time-resolved fluorimeter (i.e., colloidal silica, diluted milk) and measure the instrument response function with approximately 10,000 counts peak intensity and with the excitation and emission wavelengths at the excitation wavelength of the donor. Adjust the detection frequency, so that the maximum number of counts does not exceed the ~5% of the laser repetition rate to avoid counting artifacts.
2. Place a quartz cell of donor-only DNA construct and measure the fluorescence decay up to 10,000 counts peak intensity.
3. Add increasing amounts of acceptor construct (intermolecular assay) or binding partner (intramolecular assay). After

each addition, allow to equilibrate for ~5 min and collect the fluorescence decay at the wavelength of the maximum of the donor fluorescence spectrum.

4. Use the instrument response function previously obtained to deconvolute all fluorescence decays.
5. Fit the deconvoluted donor-only decay to **Eqs. 7** or **8** depending on the number of expected lifetime distributions.
6. Fit the deconvoluted donor–acceptor decay functions to the Gaussian distribution of distances represented by **Eq. 8** and plot their relative percentual contributions as a function of added species (*see* **Note 15**).

3.8. Time-Resolved Protein-Induced DNA Cleavage Assays

1. Proceed as described in **steps 1** and **2** of **Subheading 3.7.1**.
2. Set up the time-resolved fluorimeter to collect consecutive fluorescence decays at the maximum emission wavelength of the donor for the estimated duration of the cleavage reaction. Each donor decay should be collected in time mode in which the fluorescence is collected for a constant period of time. For example, collect continuous decays at ½ min. intervals for a one-hour reaction. Automatically save each of them in a separate file.
3. Add the enzyme and quickly start the acquisition.
4. Use the instrument response function previously obtained to deconvolute all fluorescence decays.
5. Fit the deconvoluted donor-only decay to **Eq. 6** or to a sum of exponential decays.
6. Fit the deconvoluted donor–acceptor decay to a sum of exponentials (i.e., donor-only lifetime that represents the cleavage product plus donor lifetime in the presence of acceptor that represents the substrate).
7. To extract kinetic information it is possible to plot the average lifetime defined as shown in **Eq. 9** or the time evolution of the pre-exponential factors associated to the product and the substrate.

3.8. Single-Molecule FRET

In this section we will describe only those methods related to the measurement of single-molecule FRET between proteins and DNA using the total-internal reflection technique.

3.8.1. Quartz Slides Cleaning Protocol

Independent of whether the quartz slide and the glass cover slip are new or recycled, cleaning the slides and coverslip is an absolutely crucial step for single-molecule applications. In this section we provide the protocol that the authors commonly use in their laboratory.

1. Sonicate quartz slide and coverslip in (a) 20% detergent solution for 15 min, (b) water for 5 min, (c) in acetone for 15 min, (d) water for 5 min, (e) 1 M KOH for 15 min, (f) methanol for 15 min, (g) 1 M KOH for 15 min, (h) water 5 min.
2. Dry with nitrogen or compressed air.
3. Quartz slide and coverslip are passed through a torch flame to remove impurities and moisture.

3.8.2. Quartz Slides Pegylation Protocol

1. Pour 100 mL of methanol in a reaction flask and add 5 mL of concentrated acetic acid with a glass pipette.
2. Add 1 mL of aminopropyl silane using a glass pipette and mix well.
3. Pour the mixture in the slide/coverslip container and incubate for 10 min.
4. Sonicate the reaction container for 2 min and then incubate for another 10 min at room temperature.
5. Rinse slides and/or coverslip with methanol and ultrapure water.
6. Dry with nitrogen or compressed air and put them in clean tip boxes with water at the bottom.
7. For the PEG coating of five slides, take 1–2 mg of biotinylated PEG and 80 mg of mPEG and put them in 1.5-mL Eppendorf tube.
8. Add 320 μL of pegylation buffer and mix gently with a pipette. Spin it for 1 min at 10,000 rpm.
9. Drop 70 μL of the mPEG/Biotin-PEG mixture on each slide.
10. Place very gently a coverslip on the top avoiding the formation of bubbles.
11. Allow them to incubate for 2 h at room temperature in a dark and flat place. After ~10 min of placing the coverslip on the top of the slide, check and restore any misplaced coverslip.
12. Remove the coverslip from the quartz slide, rinse them with ultrapure water, and dry using a flow of nitrogen or air (*see* **Note 16**).
13. Form a channel using double sticky tape sandwiched between the slide and the coverslip with a ~5 mm gap between the two tape strips (*see* **Note 17**).
14. Add water to the channel and look in the microscope to see if the channel is clean before adding the immobilization reagents.
15. Dilute 25-fold the initial Streptavidin stock solution in 50 mM Tris–HCl (pH 8.1), 25 mM NaCl and add 50 μL to the slide channel.

16. Allow another 10 min for Streptavidin binding.
17. Wash unbound Streptavidin with 60 μL of 50 mM Tris–HCl (pH 8.1) and 25 mM NaCl.
18. Add 60μL of a 50–100 pM biotinylated biomolecule in 50 mM Tris–HCl (pH 8.1), 25 mM NaCl and allow 5 min for binding to the Streptavidin-coated surface.
19. Look for the density and quality of labeling in the Sm-TIR microscope and if adequate proceed to add the imaging buffer. If the amount of immobilized molecules is below the desired level repeat **steps 18** and **19** (*see* **Note 18**).

3.8.3. Data Collection and Analysis

1. Prior to collecting single-molecule data, a mapping algorithm that correlates donor spots (left half of the EMCCD camera) with their acceptor counterparts (right half) needs to be created. In the authors' laboratory, this is performed using a solution of 200 nm fluorescence beads and a program written in IDL v. 6 software (ITT Visual Information Systems, USA).
2. Fluorescence data at donor and acceptor wavelengths are acquired from single molecules by using total internal reflection fluorescence microscopy with 532-nm laser excitation.
3. Data are acquired using a laboratory-written Visual C++ v. 6 program with integration times ranging from 16 to 100 ms depending on the sample dynamics. Measurements are performed at room temperature.
4. Single-molecule FRET efficiency after background correction is approximated by $(I_A/[I_A+ I_D])$, where I_A and I_D are the fluorescence intensities of the acceptor and donor, respectively. Because the quantum yields and detection efficiencies of Cy3 and Cy5 are very close, E_{app} (the apparent FRET value observed) closely matches the true efficiency of energy transfer. Data analysis is performed using laboratory-written analysis routines developed in MATLAB 7 (The MathWorks Inc., USA)
5. Single-molecule FRET histograms are obtained by averaging the first ten frames of each FRET trace for every individual molecule after manually filtering photobleaching and blinking effects. States are identified from E_{app} histograms, and dwell times are analyzed only if the time resolution allows the clear observation of transitions (more than five data points per dwell time). Rapidly fluctuating molecules undergo more transitions than slowly fluctuating ones, and thus in order to avoid bias toward fast rates, dwell-time histograms are obtained by using a weighting factor inversely proportional to the number of transitions observed for each molecule. These dwell-time histograms are then fitted to a single-exponential function to obtain the lifetimes of each

state, the inverse of the rate of conversion. For the heterogeneity analysis, the average of the dwell times is calculated for each state for each individual molecule (*see* **Note 19**).

4. Notes

1. Depending on the type of Sm-TIR setup, an oil-immersion or a water-immersion objective would be required. For prism-type sm-TIR, as the one described in **Fig. 4a**, **b**, because the evanescence wave is generated in the interface quartz-slide/water, a water-immersion objective is the appropriate choice due to its higher penetration depth. However, for the objective-type TIR, the evanescence wave generates at the bottom interface (coverslip/water) and thus a higher NA oil-immersion objective that has better performance can be used. In this case it is important to use fused silica cover slides to reduce background from glass luminescence.
2. It has been shown that linear PEG interacts with unfolded proteins preventing refolding *(46)*. Groll et al. *(53)* overcame this problem using branched star-shaped PEG polymers (SusTech, Darmstadt, Germany). Following this alternative, they were able to observe reversible folding and unfolding steps from single protein molecules.
3. As a general rule, all solutions involved in sample preparation should be tested in the single-molecule equipment. A common source of impurities is water, and thus quartz-bidistilled water is recommended but it should be monitored regularly for contaminants. In the authors' laboratory, ultrapure fresh commercial water (Sigma or other general chemical supplier) is always available for testing and sample preparation.
4. This solution is kept at 4°C in the dark. Keep this solution only for 2 or 3 weeks for better efficiency.
5. Once aminopropyl silane has been used, the remaining stock can be reused, but care must be taken to store it properly. It should be dehydrated in vacuum for ~15 min, sealed under nitrogen atmosphere, and kept at –20°C.
6. Trolox, (6-hydroxy-2, 7, 5, 8)-tetramethylchroman-2-carboxylix acid), is a water soluble derivative of vitamin E that has recently been proposed as a good alternative to the triplet-state quencher 2-mercaptoethanol, significantly reducing undesirable Cy5 blinking events.
7. Care should be taken when using higher glucose concentrations as the increase in solution viscosity could artificially

affect the intrinsic dynamics of the biomolecule or complex under investigation.

8. To achieve a homogeneous distribution of fluorescence beads it is convenient to sonicate the solution to be injected on the slides for 5 min. This will substantially decrease bead aggregation. Quartz slides coated with fluorescence beads can be active for long periods of time by sealing them with epoxy resin.
9. In those cases where the DNA is not easily precipitated from solution, add glycogen to a final 0.05–1 μg/μL concentration. Use up to 1 μL of glycogen per 20 μL of the solution; overnight incubation and storage at –20 help to increase the recovery yields.
10. Both xylene cyanol FF and bromophenol blue dyes show UV absorption when visualized on a TLC plate (Merck, Silica gel plate F254 20 × 20 cm). It is thus preferable not to load the dyes in the same wells as the DNA oligonucleotides since it could complicate visualization and contaminate the DNA.
11. After the bioconjugation reaction has reached the desired time, an excess of glutathione or 2-mercaptoethanol may be added to consume the remaining thiol-reactive reagent so that no reactive species are present during the purification step.
12. When using intermolecular FRET to measure Protein–DNA binding affinities with a donor-labeled DNA substrate and an acceptor-labeled protein, it is advisable to choose a FRET pair with an absorption spectral separation as large as possible (i.e., Cy3/Cy5). Thus, direct excitation of the acceptor-labeled protein during the titration is minimized. This is particularly crucial in those systems with poor affinities.
13. Before recording a FRET titration with many data points, it is advisable to check that the fluorescence intensity at any stage of the titration does not saturate the photomultiplier. This can be done using a quick test at the starting and end points of the titration to find the optimum instrumental parameters (slit width and photomultiplier voltage).
14. When calculating the FRET efficiency following the enhanced acceptor emission as described, it is assumed that the acceptor quantum yield is independent of the excitation wavelength, which is normally the case. However, if the shape of an acceptor emission spectra, obtained after subtraction of the normalized donor spectrum from the FRET spectrum, does not correspond to a standard acceptor spectrum, it is clear that both donor and acceptor fluorophores are strongly coupled. In these conditions, energy transfer can indeed take place,

but it is not clear if Eqs. 1 and 2 represent a good estimation of the physical process, suggesting that other mechanisms of energy transfer should be considered.

15. When analyzing the data from tr-FRET, it is usually convenient in order to get highly accurate data to use global fitting procedures instead of fitting each fluorescence decay individually. Thus, lifetimes and amplitudes are the best possible values to reproduce all experimental data
16. Pegylated quartz slides and coverslips can be stored for some time in a dark and dry place (~1 week depending on the laboratory conditions). In our laboratory, we use black-tape-wrapped Corning tubes to avoid light and plastic film to avoid moisture.
17. This defines a volume that is approximately 5 mm wide, 18 mm long, and 100 μm tall. Due to the low volume of solvent required to fill the channel, evaporation is usually quick reducing long-term use of the sample. To avoid this, two procedures can be followed (a) sealing completely the channel with fast-drying time epoxy resin and (b) using a diamond drill (Eternal Tools, UK), make two ~0.75 mm holes on opposite sides of the quartz slide, so that they can be used to inject sample and additives in real time at any stage of the experiment. The sides of the quartz slide not covered by the double-sticky tape are sealed with epoxy resin.
18. To be sure that the fluorescence observed in the CCD camera is arising from biomolecules specifically attached to the quartz slide perform the following test (a) expose a certain region of the slide to high-intensity laser light to promote complete photobleaching and switch off the laser light for 2 min. (b) switch back on the laser light and if molecules can again be observed, it is very likely that the sample contains either free molecules in solution or adsorption/desorption processes from non-specifically attached molecules are taking place.
19. Single-molecule FRET histograms usually show a "zero peak" Gaussian distribution caused mainly by incomplete labeling (lack of acceptor) or fast photobleaching during the first few frames of the experiment. Thus, care should be taken to (a) work at very low laser intensities and (b) to ensure that the lowest value of "true" FRET that can be expected from the biomolecule behavior is well separated from the "zero peak." Usually, with a 10% acceptor leakage in the donor channel, low FRET values higher than ~0.15 can be safely analyzed as arising from an active FRET state of the biomolecule.

Acknowledgments

We thank the Biological and Biotechnology Science Research Council (UK), the Royal Society (UK), and the National Sciences and Engineering Research Council (Canada), and the Universities of Sherbrooke (Canada) and St Andrews (UK) for financial support. We also thank all members of our labs for helpful discussion and critical reading of the manuscript. J. C. P is a Fellow of the Scottish Universities Physics Alliance (SUPA). DAL is a CIHR New Investigator scholar.

References

1. Hillisch, A., Lorenz, M., and Diekmann, S. (2001). Recent advances in FRET: distance determination in Protein–DNA complexes. *Curr. Opinion Struct. Biol.* **11**, 201–207.
2. Holbrook, S.R. (2005). RNA structure: the long and the short of it. *Curr. Opinion Struct. Biol.* **15**, 302–308.
3. Yan, Y., and Marriott, G. (2003). Analysis of protein interactions using fluorescence technologies. *Curr. Opinion Chem. Biol.* 7, 635–640.
4. Michalet, X., Kapanidis, A.N., Laurence, T., Pinaud, F., Doose, S., Pflughoefft, M., and Weiss, S. (2003). The power and prospects of fluorescence microscopies and spectroscopies. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 161–82.
5. Lorenz, M., Hillisch, A., Payet, D., Buttinelli, M., Travers, A., and Diekmann, S. (1999). DNA bending induced by high mobility group proteins studied by fluorescence resonance energy transfer. *Biochemistry* **38**, 12150–12158.
6. Selvin, P.R. (2000). The renaissance of fluorescence resonance energy transfer. *Nat. Struct. Biol.* 7, 730–734.
7. Stuhmeier, F., Hillisch, A., Clegg, R.M., and Diekman, S. (2000). Fluorescence energy transfer analysis of DNA structures containing several bulges and their interaction with CAP. *J. Mol. Biol.* **302**, 1081–1100.
8. Clegg, R.M. (1992). Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.* **211**, 353–388.
9. Stryer, L., and Haugland, R.P. (1967). Energy transfer: a spectroscopic ruler. *Proc. Natl Acad. Sci. USA* **58**, 719–726.
10. Stuhmeier, F., Hillisch, A., Clegg, R.M., and Diekman, S. (2000). Practical aspects of fluorescence resonance energy transfer (FRET) and its applications in nucleic acid biochemistry. *DNA–Protein Interactions.* Edited by Travers A., Buckle, M., Oxford: Oxford University Press, 77–94.
11. Bera, A., Roche, A. C., and Nandi, P. K. (2007). Bending and unwinding of nucleic acid by prion protein. *Biochemistry* **46**, 1320–1328.
12. Lorenz, M., and Diekmann, S. (2006). Distance determination in Protein–DNA complexes using fluorescence resonance energy transfer. *Methods Mol. Biol.* **335**, 243–255.
13. Passner, J. M., and Steitz, T. A. (1997). The structure of a CAP–DNA complex having two cAMP molecules bound to each monomer. *Proc. Natl Acad. Sci. USA* **94**, 2843–2847.
14. Hieb, A. R., Halsey, W. A., Betterton, M. D., Perkins, T. T., Kugel, J. F., and Goodrich, J. A. (2007). TFIIA changes the conformation of the DNA in TBP/TATA complexes and increases their kinetic stability. *J. Mol. Biol.* **372**, 619–632.
15. Dragan, A. I., Klass, J., Read, C., Churchill, M. E., Crane-Robinson, C., and Privalov, P. L. (2003). DNA binding of a non-sequence-specific HMG-D protein is entropy driven with a substantial non-electrostatic contribution. *J. Mol. Biol.* **331**, 795–813.
16. Kuznetsov, S. V., Sugimura, S., Vivas, P., Crothers, D. M., and Ansari, A. (2006). Direct observation of DNA bending/unbending kinetics in complex with DNA-bending protein IHF. *Proc. Natl Acad. Sci. USA* **103**, 18515–18520.
17. Lorenz, M., Hillisch, A., Goodman, S. D., and Diekmann, S. (1999). Global structure similarities of intact and nicked DNA complexed with IHF measured in solution by fluorescence resonance energy transfer. *Nucleic Acids Res.* **27**, 4619–4625.
18. Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B., and Tainer, J. A.

(2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* **116**, 39–50.

19. Xiao, J., and Singleton, S. F. (2002). Elucidating a key intermediate in homologous DNA strand exchange: structural characterization of the RecA-triple-stranded DNA complex using fluorescence resonance energy transfer. *J. Mol. Biol.* **320**, 529–558.
20. McKinney, S. A., Joo, C., and Ha, T. (2006). Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **91**, 1941–1951.
21. Gupta, R. C., Golub, E. I., Wold, M. S., and Radding, C. M. (1998). Polarity of DNA strand exchange promoted by recombination proteins of the RecA family. *Proc. Natl Acad. Sci. USA* **95**, 9843–9848.
22. Kuznetsov, S. V., Kozlov, A. G., Lohman, T. M., and Ansari, A. (2006). Microsecond dynamics of Protein–DNA interactions: direct observation of the wrapping/unwrapping kinetics of single-stranded DNA around the *E. coli* SSB tetramer. *J. Mol. Biol.* **359**, 55–65.
23. Lucius, A. L., Jason Wong, C., and Lohman, T. M. (2004). Fluorescence stopped-flow studies of single turnover kinetics of *E. coli* RecBCD helicase-catalyzed DNA unwinding. *J. Mol. Biol.* **339**, 731–750.
24. Kvaratskhelia, M., Wardleworth, B. N., Bond, C. S., Fogg, J. M., Lilley, D. M., and White, M. F. (2002). Holliday junction resolution is modulated by archaeal chromatin components in vitro. *J. Biol. Chem.* **277**, 2992–2996.
25. Furey, W. S., Joyce, C. M., Osborne, M. A., Klenerman, D., Peliska, J. A., and Balasubramanian, S. (1998). Use of fluorescence resonance energy transfer to investigate the conformation of DNA substrates bound to the Klenow fragment. *Biochemistry* **37**, 2979–2990.
26. Mukhopadhyay, J., Mekler, V., Kortkhonjia, E., Kapanidis, A. N., Ebright, Y. W., and Ebright, R. H. (2003). Fluorescence resonance energy transfer (FRET) in analysis of transcription-complex structure and function. *Methods Enzymol.* **371**, 144–159.
27. Heyduk, T., and Niedziela-Majka, A. (2001). Fluorescence resonance energy transfer analysis of escherichia coli RNA polymerase and polymerase-DNA complexes. *Biopolymers* **61**, 201–213.
28. Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2006). Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophys. J.* **90**, 1419–1431.
29. Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S., and Ebright, R. H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**, 1144–1147.
30. Lee, S. P., and Han, M. K. (1997). Fluorescence assays for DNA cleavage. *Methods Enzymol.* **278**, 343–363.
31. Eggeling, C., Jager, S., Winkler, D., and Kask, P. (2005). Comparison of different fluorescence fluctuation methods for their use in FRET assays: monitoring a protease reaction. *Curr. Pharm. Biotechnol.* **6**, 351–371.
32. Ghosh, S. S., Eis, P. S., Blumeyer, K., Fearon, K., and Millar, D. P. (1994). Real time kinetics of restriction endonuclease cleavage monitored by fluorescence resonance energy transfer. *Nucleic Acids Res.* **22**, 3155–3159.
33. Hiller, D. A., Rodriguez, A. M., and Perona, J. J. (2005). Non-cognate enzyme-DNA complex: structural and kinetic analysis of EcoRV endonuclease bound to the EcoRI recognition site GAATTC. *J. Mol. Biol.* **354**, 121–136.
34. Ray, P. C., Fortner, A., and Darbha, G. K. (2006). Gold nanoparticle based FRET assay for the detection of DNA cleavage. *J Phys Chem B* **110**, 20745–20748.
35. Lin, J., Gao, H., Schallhorn, K. A., Harris, R. M., Cao, W., and Ke, P. C. (2007). Lesion recognition and cleavage by endonuclease V: a single-molecule study. *Biochemistry* **46**, 7132–7137.
36. van der Meer, B. W. (2002). Kappa-squared: from nuisance to new sense. *J. Biotechnol.* **82**, 181–196.
37. Klostermeier, D., and Millar, D. P. (2001). Time-resolved fluorescence resonance energy transfer: a versatile tool for the analysis of nucleic acids. *Biopolymers* **61**, 159–179.
38. Cornish, P. V., and Ha, T. (2007). A survey of single-molecule techniques in chemical biology. *ACS Chem. Biol.* **2**, 53–61.
39. Ha, T. (2001). Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78–86.
40. Ha, T. (2004). Structural dynamics and processing of nucleic acids revealed by single-molecule spectroscopy. *Biochemistry* **43**, 4055–4063.
41. Ritort, F. (2006). Single-molecule experiments in biological physics: methods and applications *J. Phys.: Condens. Matter* **18**, R531–R583.
42. Moerner, W. E., and Fromm, D. P. (2003). Methods of single-molecule fluorescence spectroscopy and microscopy. *Rev. Sci. Instrum.* **74**, 3597.

43. Haustein, E., and Schwille, P. (2007). Fluorescence correlation spectroscopy: novel variations of an established technique. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 151–169.
44. Schwille, P. (2003). TIR-FCS: staying on the surface can sometimes be better. *Biophys. J.* **85**, 2783–2784.
45. Wazawa, T., and Ueda, M. (2005). Total internal reflection fluorescence microscopy in single molecule nanobioscience. *Adv. Biochem. Eng. Biotechnol.* **95**, 77–106.
46. Rasnik, I., McKinney, S. A., and Ha, T. (2005). Surfaces and orientations: much to FRET about? *Acc. Chem. Res.* **38**, 542–548.
47. Cisse, I., Okumus, B., Joo, C., and Ha, T. (2007). Fueling protein DNA interactions inside porous nanocontainers. *Proc. Natl Acad. Sci. USA* **104**, 12646–12650.
48. Myong, S., Bruno, M. M., Pyle, A. M., and Ha, T. (2007). Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase. *Science* **317**, 513–516.
49. Lu, H. P., Iakoucheva, L. M., and Ackerman, E. J. (2001). Single-molecule conformational dynamics of fluctuating noncovalent DNA–protein interactions in DNA damage recognition. *J. Am. Chem. Soc.* **123**, 9184–9185.
50. Segers-Nolten, G. M., Wyman, C., Wijgers, N., Vermeulen, W., Lenferink, A. T., Hoeijmakers, J. H., Greve, J., and Otto, C. (2002). Scanning confocal fluorescence microscopy for single molecule analysis of nucleotide excision repair complexes. *Nucleic Acids Res.* **30**, 4720–4727.
51. Lemay, J. F., Penedo, J. C., Tremblay, R., Lilley, D. M., and Lafontaine, D. A. (2006). Folding of the adenine riboswitch. *Chem. Biol.* **13**, 857–868.
52. Braslavsky, I., Hebert, B., Kartalov, E., and Quake, S. R. (2003). Sequence information can be obtained from single DNA molecules. *Proc. Natl Acad. Sci. USA* **100**, 3960–3964.
53. Groll, J., Amirgoulova, E. V., Ameringer, T., Heyes, C. D., Rocker, C., Nienhaus, G. U., and Moller, M. (2004). Biofunctionalized, ultrathin coatings of cross-linked star-shaped poly(ethylene oxide) allow reversible folding of immobilized proteins. *J. Am. Chem. Soc.* **126**, 4234–4239.
54. Adachi, K., Yasuda, R., Noji, H., Itoh, H., Harada, Y., Yoshida, M., and Kinosita, K., Jr. (2000). Stepping rotation of F1-ATPase visualized through angle-resolved single-fluorophore imaging. *Proc. Natl Acad. Sci. USA* **97**, 7243–7247.
55. Kastner, C. N., Prummer, M., Sick, B., Renn, A., Wild, U. P., and Dimroth, P. (2003). The citrate carrier CitS probed by single-molecule fluorescence spectroscopy. *Biophys. J.* **84**, 1651–1659.
56. Boukobza, E., Sonnenfeld, A., and Haran, G. (2001). Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. *J. Phys. Chem. B* **105**, 12165–12170.
57. Kapanidis, A. N., and Weiss, S. (2002). Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules. *J. Chem. Phys.* **117**, 10953–10964.
58. Sapsford, K. E., Berti, L., and Medintz, I. L. (2006). Materials for fluorescence resonance energy transfer analysis: beyond traditional donor–acceptor combinations. *Angew. Chem. Int. Ed. Engl.* **45**, 4562–4589.

Chapter 29

Single-Molecule FRET Analysis of Protein–DNA Complexes

Mike Heilemann, Ling Chin Hwang, Konstantinos Lymperopoulos, and Achillefs N. Kapanidis

Summary

We present a single-molecule method for studying protein–DNA interactions based on fluorescence resonance energy transfer (FRET) and alternating-laser excitation (ALEX) of single diffusing molecules. An application of this method to the study of a bacterial transcription initiation complex is presented.

Key words: Single-molecule fluorescence spectroscopy, Fluorescence resonance energy transfer (FRET), Alternating-laser excitation (ALEX), Protein–DNA interactions.

1. Introduction

Single-molecule fluorescence methods have advanced our understanding of many important processes that involve nucleic acids. Single-molecule methods have two attractive features. First, they allow observation of subpopulations within a system under equilibrium conditions. Second, they can monitor conformational changes and association/dissociation reactions in real time, something difficult to achieve with ensemble methods since such changes cannot be easily synchronized. Here, we describe protocols for studying protein–DNA interactions using a combination of fluorescence resonance energy transfer (FRET) and alternating-laser excitation (ALEX).

FRET is a non-radiative energy transfer mechanism between two fluorophores: a donor fluorophore is excited by light and transfers its excited state energy to an acceptor fluorophore via a dipole–dipole interaction *(1)*. FRET has been used to study

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_29

distances and distance changes in a range of 2–10 nm and thus can serve as a spectroscopic ruler *(2)*. Single-molecule fluorescence resonance energy transfer (smFRET; *see* **ref. 3***)* is a powerful method for monitoring interactions and conformational changes within single biomolecules. The method has contributed new and important information about how biological machines work *(4, 5)*. In a typical smFRET experiment, a single molecule or complex labelled with a donor–acceptor FRET pair is exposed to light that excites the donor **(Fig. 1a)**; after identification of the signals arising from single molecules **(Fig. 1b)**, FRET-related ratios report on the proximity between the probes; this information is summarized in one-dimensional FRET histograms **(Fig. 1c)**, which report on biomolecular structure.

Since 2004, smFRET has been extended through the use of alternating-laser excitation *(6–10)*, **(Fig. 1d)** which provides additional information on the presence and the state of both donor and acceptor fluorophores in single biomolecules **(Fig. 1e)**; this information is summarized in multidimensional histograms of FRET and relative fluorophore stoichiometry **(Fig. 1f)**, which report on biomolecular structure and stoichiometry, respectively. ALEX helps the virtual sorting of molecules by separating singly labelled species from species that contain both fluorophores (e.g. protein–DNA complexes formed between singly labelled proteins and singly labelled DNA). Initial applications of ALEX focused on studies of gene transcription *(11–13)* and protein folding *(8)*. As with conventional smFRET, ALEX is compatible with studies of both diffusing and immobilized molecules, and with several timescales, allowing monitoring of fast dynamics in single diffusing biomolecules, or simultaneous monitoring of a large number of individual molecules.

This chapter describes how protein–DNA interactions can be studied at the single-molecule level using single-molecule FRET and ALEX. The protocol was developed to study conformational changes in transcription initiation complexes. This includes the formation of a transcription complex after *Escherichia coli* RNA polymerase (RNAP) binds to a doubly labelled DNA fragment and melts the region of DNA around the transcription start site (forming a 'transcription bubble') to generate an RNAP–DNA open complex (RP_o). The formation of the transcription bubble and the DNA bending associated with RP_o formation brings the two fluorophores closer, increasing the FRET efficiency and allowing study of DNA opening as well as subsequent process (e.g. transcription initiation, promoter escape) that change the FRET efficiency. The protocols describe how (a) to prepare and purify fluorescently labelled DNA, (b) to form active transcription complexes, (c) to perform single-molecule experiments using ALEX and smFRET for diffusing biomolecules, and (d) to extract information on distances and stoichiometries.

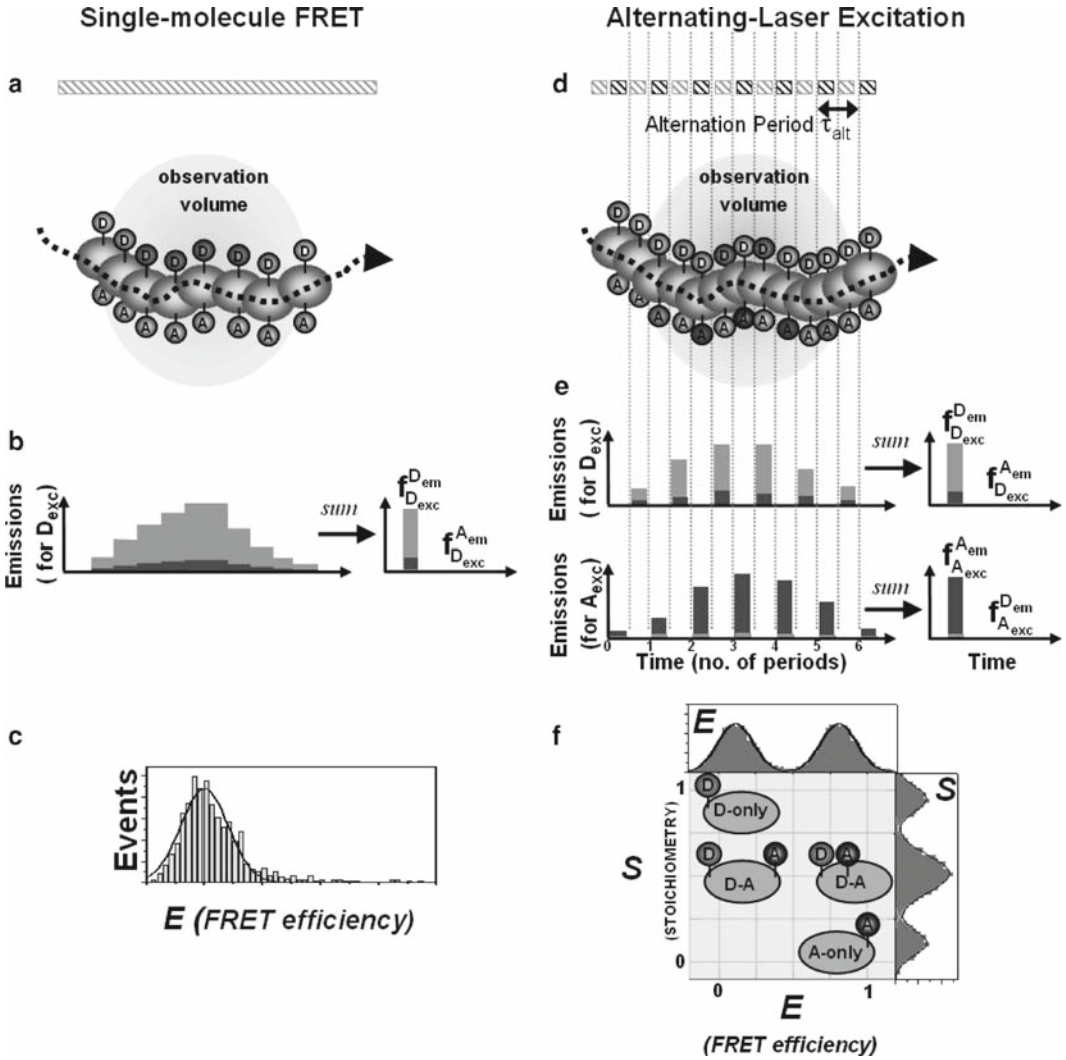


Fig. 1. Concept of ALEX and comparison with single-excitation single-molecule FRET. (**a–c**) Single-molecule FRET using single-laser excitation; diffusing-molecule example. A fluorescent molecule transverses a focused green-laser beam and emits photons in the donor- and acceptor-emission wavelengths. The photon counts at these two wavelengths are used to generate one-dimensional histograms of FRET efficiency *E*. (**d–f**) Single-molecule FRET using alternating-laser excitation. A fluorescent molecule transverses an observation volume illuminated in an alternating fashion using focused green- and red-laser beams. Using the photons emitted in the donor- and acceptor-emission wavelengths for each laser excitation, one can generate two-dimensional histograms of FRET efficiency *E* and relative fluorophore stoichiometry *S*, enabling molecular sorting.

2. Materials

2.1. DNA

DNA was purchased from IBA GmbH (Göttingen, Germany). Alternatively, DNA can be purchased from IDT (Coralville, IA) or MWG (Göttingen, Germany). *See* **Subheading "Design of Synthetic DNA"** for more information about the amount and purity required.

2.2. Reactive Fluorophores

Amine-reactive fluorophores were used for DNA labelling and were purchased as *N*-hydroxy-succinimidyl (NHS) esters: ATTO488 and ATTO647N (ATTO-TEC GmbH, Siegen, Germany); Cy3B and Cy5 (GE Healthcare, Uppsala, Sweden); Alexa488, Alexa647, Rhodamine Green, and Tetramethyl-6-Carboxyrhodamine (TAMRA) (Invitrogen, Carlsbad, CA). For the transcription complexes studied here, fluorophores Cy3B and Alexa647 are used. One milligram of reactive fluorophore was resuspended in anhydrous dimethyl sulphoxide (DMSO, Fisher, Loughborough, UK), aliquots of 100 nmol were dried under vacuum and stored at –80°C where they are stable for at least a year (see also instruction sheet from the manufacturer).

2.3. Chromatography Reagents

1. 1.0 M triethylammonium acetate (TEAA) buffered at pH 7.4 (VWR, West Chester, PA).
2. Acetonitrile (AcCN), HPLC grade (Fisher).

2.4. Buffers

1. *HPLC DNA purification Buffer A*. 0.1 M TEAA, 5% AcCN.
2. *HPLC DNA purification Buffer B*. 0.1 M TEAA, 70% AcCN.
3. *DNA Hybridization buffer*. 20 mM Tris–Cl (pH 8.0), 1 mM EDTA, and 500 mM NaCl.
4. *Gel elution buffer*. 0.5 M ammonium acetate, 10 mM magnesium acetate.

2.5. Polyacrylamide Gel Electrophoresis (PAGE)

1. 40% 29:1 acrylamide:bis-acrylamide solution (Fisher).
2. *N*,*N*,*N*′,*N*′-tetramethylethylene-diamine (TEMED; Fisher).
3. 10% (w/v) ammonium persulphate in water. The solution is stored at –20°C in aliquots of 1 ml and thawed before use.
4. *Running buffer TBE (5×)*. 445 mM of Tris–base (Fisher), 445 mM of boric acid (Fisher), and 10 mM EDTA of pH 8.0 (Fisher).
5. *Molecular weight markers*. 100-bp and 1-kbp DNA ladder (NEB, Beverly, MA).
6. *Loading dye (6×)*. 0.25% bromophenol blue (Sigma), 0.25% xylene cyanol (Sigma) in 30% glycerol in H_2O (*see also* **ref. 14** for details).

2.6. UV-Vis Absorbance Measurements

UV-Vis absorbance measurements were performed using either the NanoDrop® ND-1000 (Labtech, Ringmer, UK) spectrophotometer or the Cary 50 Bio UV-Visible spectrophotometer (Varian, Palo Alto, CA).

2.7. Fluorescence Measurements

Fluorescence emission and excitation spectroscopy was performed using an L-shape fluorimeter (PTI, Birmingham, NJ) equipped with a xenon arc lamp controlled by a lamp power supply (LPS-220B), a motor driver (MD-5020), a shutter (SC-500), and a photomultiplier detection system (814; all parts from PTI).

2.8. Reagents for DNA Labelling

1. 3 M NaCl (Fisher).
2. 70% and 100% ice-cold ethanol (Fisher).
3. Chloroform (Fisher).
4. 0.1 M Sodium borate buffer, pH 9.0.

2.9. Reagents for Transcription Assay

1. *Escherichia coli* RNA polymerase-σ^{70} holoenzyme (Epicentre Biotechnologies, Madison, WI) is supplied in a 50% glycerol solution with 50 mM Tris–HCl (pH 7.5), 250 mM NaCl, 0.1 mM EDTA, and 1 mM dithiothreitol (DTT) and stored at –20°C without a defrost cycle.
2. *lac* promoter DNA variant, *lac*CONS $(-39/+15)^{Cy3B,-15/Alexa647,+15}$ (*see* **ref. 15** for sequence; numbering convention refers to labelling position relative to transcription start site).
3. *Buffer for transcription assays (KG7)*. 40 mM HEPES buffer (pH 7.0, Fisher), 100 mM potassium-L-glutamate (Sigma, St. Louis, MO), 10 mM $MgCl_2$ (Sigma), 1 mM DTT (Sigma), 100 μg/ml bovine serum albumin (BSA, Invitrogen), 5% (w/v) glycerol (Sigma). The buffer is filtered through a 0.1-μm pore inorganic membrane filter (Anotop 25 Plus, Whatman, Maidstone, UK) and stored in aliquots of 1 ml at –20°C (*see* **Note 1** for grade of chemicals used).
4. Heparin or heparin-Sepharose (1 mg/ml, GE Healthcare, Amersham, UK).
5. *Buffer for open complex formation (KG7 + ApA)*. Open complexes were formed in the presence of initiating dinucleotide, 0.5 mM adenylyl (3′–5′)adenosine (ApA, Ribomed, Carlsbad, CA).

2.10. Single-Molecule Fluorescence Spectroscopy Setup

1. A custom-built confocal microscope with single-fluorophore detection capability **(Fig. 2)**.
2. Two laser beams, a 532-nm Nd:YAG laser (Samba; Cobolt AB, Solna, Sweden) and a 635-nm diode laser (Cube; Coherent, Santa Clara, CA), are used for the excitation of two fluorophore labels emitting in the green and red channels, respectively. The red laser is directly modulated via TTL (transistor-transistor-logic; refers to a standardized digital signal) signals and the green laser is modulated using an electro-optical modulator (EOM, Conoptics, Danbury, CT). The alternation period is 50 μs, and the excitation duty cycle is 50% (a constant interval of 3 μs is kept between the two excitation pulses to avoid any temporal overlap of excitation). For full details of laser alternation setup, refer to *(16)*.
3. Place a dichroic mirror (560DRLP, Omega, Brattleboro, VT) at 45° to the green laser beam to combine both laser beams.
4. Focus the laser beams into the aperture of a single-mode optical fibre (Thorlabs, Newton, NJ) using a 10× microscope objective (Comar, Cambridge, UK).

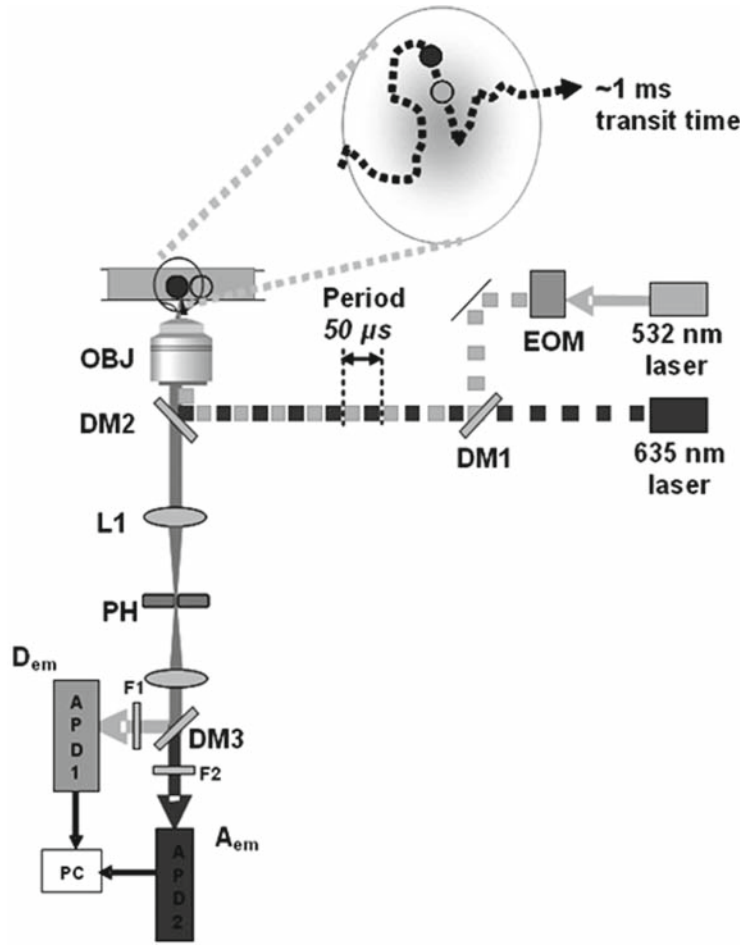


Fig. 2. An alternating-laser excitation (ALEX) setup. A 635-nm laser beam (directly modulated) and 532-nm laser beam modulated with an electro-optical modulator (EOM) are both combined with a dichroic mirror (DM1) into a single-mode fibre and focused by a microscope objective (OBJ) into a fluorescent sample on a coverslip. The fluorescence emission is separated from the excitation light and scattering by a dichroic mirror (DM2) and focused by the tube lens (L1) of the microscope onto a pinhole (PH). The emitted light is split by a dichroic mirror (DM3), spectrally filtered by filters (F1 and F2) into its respective wavelengths, and focused by lenses onto the detectors.

5. At the other end of the optical fibre, use another 10× objective (Comar) to collimate the laser focus into an inverted microscope (IX71, Olympus, Tokyo, Japan).
6. Use a dual-band dichroic beamsplitter (FF545/650-Di01, Semrock, Rochester, NY) to reflect the laser beams into the back aperture of an oil-immersion objective (UPlanSApo 60×, 1.35NA, Olympus). Fluorescence from diffusing molecules is collected through the objective and separated from excitation and scattered light by the dichroic beamsplitter.

7. Align a 100-μm pinhole at the image plane after the tube lens to spatially filter the fluorescence light from background scattering.
8. Insert a lens after the pinhole to collimate the emission light. In our case, we use a lens with a focal length of f = 100 mm.
9. Place a dichroic mirror (630DRLP, Omega) to split the fluorescence emission according to their wavelengths into the green and red emission channels.
10. Insert bandpass and longpass filters (HQ585/F70 and HQ650LP, respectively; both from Chroma, Rockingham, VT) to spectrally filter the emission light for the green and red detection channels, respectively.
11. Align two lenses (f = 20 mm) in front of each detection channel to focus the emission light onto the detection windows of two silicon avalanche photodiodes (SPCM-AQR 14; Perkin Elmer, Waltham, MA). The photon arrival times are recorded with a counting board (PCI-6602, National Instruments, Austin, TX) connected to a personal computer.
12. Measure a standard sample, e.g. TMR or Cy3B (1 nM), to align for the green detector and using the emission crosstalk to align for the red detector.

3. Methods

3.1. Preparation of Doubly Labelled Double-Stranded DNA for Single-Molecule FRET

3.1.1. Preparation Using Strands Prepared with Automated Synthesis

Design of Synthetic DNA

DNA Labelling

Synthetic oligonucleotides were ordered from IBA GmbH and contained amino modification for subsequent labelling with amine-reactive fluorophores. The synthesis scale was typically 0.2 μmol with HPLC-grade purity. For terminal amino modifications, a C6-linker was used (*see* **Note 2** for exceptions concerning guanosine). Internal amino modifications were obtained through insertion of a modified deoxythymidine (dT) nucleobase that contained either an amino-C2 or -C6 linker. Lyophilized DNA was resuspended in distilled deionized H_2O (ddH_2O) at a final concentration of 100 μM and stored at –20°C.

1. Use 5 nmol of DNA dissolved in 100 μl of dH_2O.
2. Extract DNA twice with an equal volume (100 μl) of chloroform; DNA will be in the aqueous phase (top phase).
3. Precipitate DNA by adding 10 μl of 3 M NaCl and 250 μl ice-cold 100% ethanol. Mix well and place at –20°C for 30 min.
4. Centrifuge at 4°C and 16,000 × g for 10 min.

5. Carefully remove the supernatant and rinse pellet with 200 µl ice-cold 100% ethanol. Dry the pellet either at room temperature for 15 min or at 37°C for 5 min.
6. Dissolve the pellet in 95 µl of 0.1 M sodium borate buffer, pH 9. Dissolve 50 nmol of the reactive fluorophore with 5 µl anhydrous DMSO and add the DNA solution to it. Mix solution with a short spin.
7. Protect the sample from light and place it on a table-top orbital shaker (IKA Genius vortex, Sigma), and shake overnight at room temperature.
8. Ethanol precipitate the DNA by adding 5 µl of 3 M NaCl and 125µl ice-cold 100% ethanol to the reaction mix. Mix well and place at –20°C for 30 min.
9. Centrifuge at 4°C and 16,000 × *g* for 10 min.
10. Carefully remove the supernatant, and rinse the pellet with 200 µl ice-cold 100% ethanol. Dry the pellet either at room temperature for 15 min or at 37°C for 5 min.
11. For HPLC purification, dissolve pellet in 500 µl of 0.1 M TEAA, pH 7.4. Otherwise, DNA pellets can be stored at –20°C.

DNA Purification by High-Performance Liquid Chromatography (HPLC)

Labelled oligonucleotides were purified by liquid chromatography using an AKTA™ Basic chromatography system (GE Healthcare) comprising a P-903 pump system fitted with a 2-ml injection loop, a UV-900 monitor, M-925 mixer, a Frac-920 fraction collector, and a reverse-phase column µRPC C2/C18 ST 4.6/100 (all from GE Healthcare). All chromatography buffers (*see* **Subheading 2.4**) were filtered through a 0.2-µm membrane filter and degassed subsequently. A flow rate of 0.5 ml/min was applied throughout.

1. Flush the column with 5 column volumes (CV) of dH_2O until the pressure and the absorption stabilize.
2. Equilibrate column with 5 CV of buffer A.
3. Flush the injection loop with 10 ml of buffer A.
4. Stop the AKTA system and wait until pumps are turned off.
5. Dissolve the labelled DNA pellet in 500 µl of 0.1 M TEAA, pH 7.4.
6. Load half amount of the sample (250 µl) onto the loop.
7. Run the according separation program from the program wizard; typically, the following program was used:
   (a) Equilibration step with 4 CV.
   (b) *Loading step*. The sample is transferred from the loop into the column by flushing 2 CV of buffer A into the loop.
   (c) Two-step gradient for the purification and elution of the DNA:

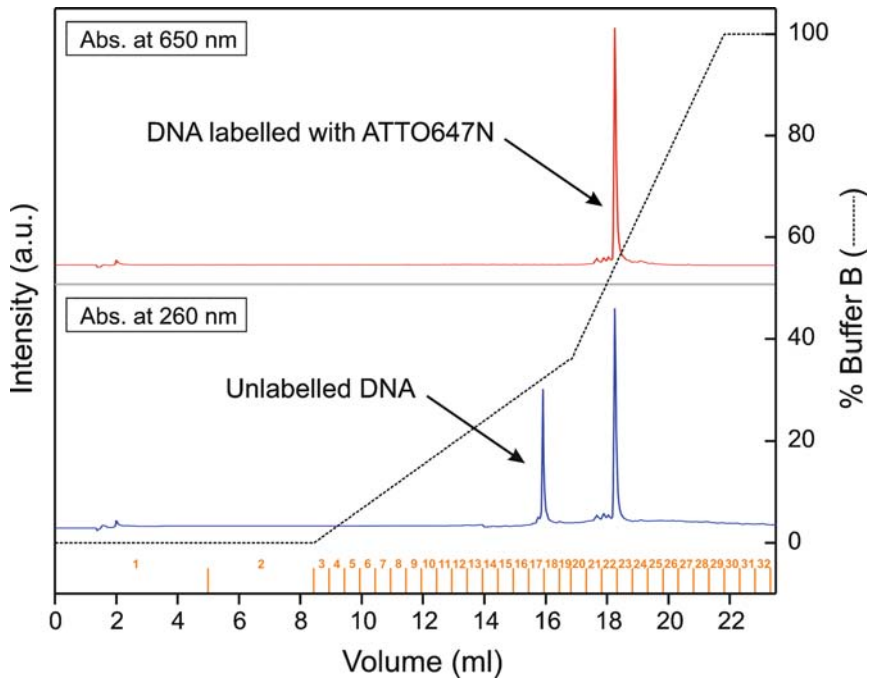


Fig. 3. Typical chromatograms obtained from an HPLC purification of DNA after labelling with ATTO647N. Unlabelled DNA is separated from labelled DNA after applying the two-step gradient. The unreacted fluorophore has been largely removed during the DNA precipitation procedure. The DNA presence is monitored through the absorption at 260 nm (bottom part), while the presence of ATTO647N is monitored through the absorption at 650 nm (upper part). Fractions of 5 ml were collected during the loading and washing steps (fractions 1 and 2), followed by 0.5-ml fractions (3–29) during the two-step gradient (0–100% of buffer B). The column was washed further with 2 ml of Buffer B (fractions 30–32).

(d) After reaching 100% of buffer B, the column is washed with 2 CV of buffer B.

8. *Collect fractions containing labelled DNA*. Simultaneous absorption peaks at both 260 nm (indicates the presence of nucleic acid) and the respective absorption wavelength of the fluorophore indicate elution of labelled DNA. *See* **Fig. 3** for a typical chromatogram.
9. Dry DNA with a centrifugal evaporator (SpeedVac, Savant, Fisher) at ambient temperature overnight and protected from light.
10. DNA was either resuspended with ddH_2O for immediate use or kept as pellet. In both cases, DNA was stored at −20°C.

Measuring the Concentration of Purified DNA and Determine the Degree of Labelling

1. Dissolve DNA pellet in 30 µl of ddH_2O.
2. Use 2 µl of the sample to measure the absorbance using the NanoDrop (*see* **Subheading 2.6**). Note both the absorption maximum of DNA (260 nm) and the fluorophore (in the 400–700 nm region).

3. Calculate concentration of the DNA and the fluorophore according to Lambert–Beer's Law,

$$c = \frac{A}{\varepsilon d},$$

where A is the absorption (in units of OD), ε is the extinction coefficient (in units of $M^{-1}cm^{-1}$), c is the concentration (in units of molar concentration, M), and d is the optical path length of the cuvette (in cm). Note that fluorophores have important contributions to the 260-nm absorption which need to be subtracted from the total 260 nm absorption in order to calculate the correct concentration of the DNA; information on the extent of this contribution is available from suppliers of fluorophores (e.g. Invitrogen; Glen Research, Sterling, VA). This is easily done by first determining the concentration of the fluorophore in the sample,

$$c_{\mathrm{dye}} = \frac{A_{\mathrm{dye},\lambda}}{\varepsilon_{\mathrm{dye},\lambda} d}$$

and then subtracting the contribution of the fluorophore to the absorption at 260 nm,

$$A_{\mathrm{DNA,260\,nm}} = A_{\mathrm{260\,nm}} - \varepsilon_{\mathrm{dye,260\,nm}} c_{\mathrm{dye}} d,$$

$$c_{\mathrm{DNA}} = \frac{A_{\mathrm{DNA,260\,nm}}}{\varepsilon_{\mathrm{DNA,260\,nm}} d}.$$

The labelling efficiency is evaluated by calculating the ratio of the concentration of the fluorophore over the concentration of DNA. Typical ratios are 100 ± 10% (*see* **Note 3** for poor labelling efficiencies).

Hybridization

1. Perform hybridization reactions in 0.6-ml Eppendorf tubes (Eppendorf, Hamburg, Germany) in a final volume of 40 µl.
2. Dissolve the DNA in 40 µl of hybridization buffer at a final concentration of 1 µM for each oligonucleotide.
3. Perform annealing reaction in a thermocycler (PX2 Thermo Fisher, Waltham, MA) with a preheated lid set to 105°C to prevent condensation on top of the tubes. The following annealing protocol was typically used:
   (a) Heat the tubes to 90°C for 4 min.
   (b) Decrease the temperature by 1°C/30 s until it reaches 25°C.
4. In case the hybridization reaction is not efficient, decrease the cooling rate to 1°C/2 min (*see also* **Note 4**).

Purification of Double-Stranded DNA

Double-stranded DNA (dsDNA) was purified using non-denaturing polyacrylamide gel electrophoresis (PAGE). The length of the dsDNA determines the percentage of acrylamide to be used for the gel separation *(14)*.

1. Cast a gel (with 1-mm thickness) in a vertical 16 × 16 cm apparatus (CBS, Del Mar, CA) and pre-run it before loading by applying a low voltage for 15 min. Flush wells with running buffer.
2. Add 1 μl of 50% glycerol to each sample (40 μl) and mix before loading. Load using gel-saver tips (Fisher).
3. Run gel at 150 V for 3 h, or until the loading dye reaches the bottom of the gel.
4. After completion of the run, visualize the gel using either a camera-based gel-imager (e.g. GeneGenius2; Syngene, Cambridge, UK) or a laser-based gel scanner (Molecular Imager Pharos FX, Biorad, Hercules, CA). Individual bands of DNA samples were visualized according to the spectral properties of the fluorophores. For example, fluorophore Cy3B can be visualized using the gel imager, an excitation source at 365 nm (epi-UV) and an emission region centred around 580 nm (defined by bandpass filter 580/60); alternatively, fluorophore Cy3B can be visualized using the gel scanner, an excitation source at 532 nm and an emission region centred around 580 nm (defined again by bandpass filter 580/60).
5. Excise the band containing dsDNA, slice it into small pieces, and place them into a 1.7-ml Eppendorf tube containing 100–200 μl of elution buffer.
6. DNA was eluted overnight at room temperature by leaving the tubes on a mixing platform (Nutator mixer, VWR).
7. Ethanol precipitate DNA by adding 1/10× of the initial volume of 3 M NaCl and 2.5× of the initial volume of ice-cold 100% ethanol.
8. Mix and place at –20°C for 20 min. Centrifuge at 16,000 × *g* using a table-top centrifuge (5415R, Eppendorf) at 4°C for 15 min.
9. Remove the supernatant and wash the pellet with ice-cold 70% ethanol.
10. Centrifuge at 16,000 × *g* at 4°C for 5 min.
11. Remove supernatant and air dry pellet at room temperature for 10 min or at 37°C for 5 min.

3.1.2. Preparation of DsDNA with Polymerase Chain Reaction (PCR)

Primer Design

Primers and full-length DNA were ordered from IBA GmbH; synthesis scale and purification options were chosen as described in **Subheading "Design of Synthetic DNA".** The bioinformatics program NetPrimer (Premier Biosoft International, Palo Alto, CA) was used to ensure that there was no secondary structure

(hairpins) or primer–dimer formation and to calculate the physical parameters of the oligonucleotides (melting temperature T_m).

PCR

PCR was usually performed at 100 µl final volume in 0.6-ml Eppendorf tubes. For larger volumes, increase the number of 100-µl reactions instead of increasing the volume in a single reaction. A typical PCR mix consists of:

1. 1 nM of DNA template
2. 10 µM of each primer
3. 250 µM of dNTPs (Roche, Basel, Switzerland)
4. 5 U of Taq DNA polymerase (NEB)
5. 1× reaction buffer for Taq DNA polymerase (NEB)

The volume was made up to 100 µl with dH_2O and mixed. Before starting the PCR make sure that the lid of the thermo cycler is set to 105°C to prevent condensation.

PCR was performed under the following program:

1. 92°C for 2 min
2. 92°C for 20 s (with ~5 s ramp time)
3. 55°C for 20 s (with ~5 s ramp time)
4. 72°C for 20 s (with ~5 s ramp time)
5. Repeat **steps 2–4** for 39 cycles
6. 72°C for 5 min

(*See* **Note 5** for PCR trouble shooting.)

Purification and Quantitation of Double-Stranded DNA

Purification and quantitation of dsDNA produced by PCR was performed as described in **Subheadings "Measuring the Concentration of Purified DNA and Determine the Degree of Labelling"** and **"Purification of Double-Stranded DNA",** respectively.

3.2. Formation of RNAP–DNA Complexes

1. Prepare 20 µl of a 40-nM RNAP holoenzyme solution in transcription buffer (*see* **Note 6**).
2. Add 0.8 µl of 1 µM labelled DNA fragment (final concentration of 10 nM) and incubate the mixture 15 min at 37°C (*see* **Note 7**).
3. Add 2 µl of 10 mg/ml heparin-Sepharose (1 mg/ml; tap the tube to suspend the beads uniformly in solution). Heparin-Sepharose has a polyanionic structure and serves as a non-specific analogue of DNA and RNA. It is used to disrupt non-specific RNAP-promoter complexes and to remove free RNAP. Mix the solution well with pipette for 30 s and incubate 20 s at 37°C (*see* **Note 8**).
4. Centrifuge the sample for 10 s and immediately transfer 9.5 µl of the supernatant to a tube containing 0.5 µl 10 mM

ApA pre-incubated at 37°C (for formation of open complex, RP_o) at 37°C for 15 min. The initial sequence of *lac*CONS is AATTGTG. Dinucleotide ApA binds to RNAP and increases the stability of the initial transcribing complex by forming two base pairs with the template DNA strand in the transcription bubble. Hence, ApA addition reduces dissociation of open complexes during data collection.

3.3. Single-Molecule Fluorescence Spectroscopy

3.3.1. Sample Preparation

1. On a No. 1 glass coverslip, place a self-adhesive silicone sheet (Invitrogen) with an 8-shaped hole cut out.
2. For DNA measurements, add 50 µl of KG7 onto the gasket and incubate 37°C at 5 min on a flat heating block. (This step ensures coating of the glass surface with BSA molecules present in KG7 buffer and diminishes non-specific binding of complexes to the coverslip.)
3. Add 0.5 µl of the sample in the solution on the coverslip and mix well with a pipette before placing on the microscope objective. The objective is pre-heated to 37°C; a heating stage can also be used instead of heated objectives. The final concentration of complexes is ~100 pM (*see* **Notes 9** and **10**).
4. Place a drop of low-fluorescence immersion oil (Zeiss, Jena, Germany) onto the objective and place the coverslip containing the sample droplet on top.
5. Bring the focus of the objective to the top surface of the coverslip. This will give a strong scattering signal which can be seen at the microscope eyepiece or on a camera that images the focal plane. From there, bring the objective 20 µm into the solution by adjusting the graduated focusing knob.
6. Check that the density of bursts is well spaced. Then, seal the chamber with another coverslip to prevent evaporation.
7. For measurements on RP_o, prepare 50 µl of KG7 + ApA buffer in an Eppendorf tube and place it on a new gasket with an 8-shaped hole. Repeat **steps 3–6**.

3.3.2. Data Acquisition and Analysis

1. Use excitation intensities of 180–360 µW for 532-nm excitation of donor fluorophores (D_{ex}), and 60–120 µW for 635-nm excitation of acceptor fluorophores (A_{ex}); intensities measured when lasers operate in the continuous-wave mode (*see* **Note 11**).
2. Measure the samples for 15–30 min to acquire enough events for data analysis (*see* **Note 12**).
3. Data acquisition and analysis are performed using a custom-built program on LabVIEW software (National Instruments). Assign photons detected at the donor and acceptor emission channels to D_{ex} or A_{ex} and their arrival time, and generate photon streams, $f_{D_{ex}}^{A_{em}}$, $f_{D_{ex}}^{D_{em}}$, $f_{A_{ex}}^{A_{em}}$, $f_{A_{ex}}^{D_{em}}$, where $f_{X_{ex}}^{Y_{em}}$ denotes the

photon count rate of emission of fluorophore *Y*, resulting from excitation of fluorophore *X*.

4. Set thresholds of $f_{A_{ex}}^{A_{em}}$ at 5–10 photons per 50–1,000 μs time bin and 10–25 photons per burst to remove false positives due to background and select molecules containing acceptor fluorophores with photon counts above background. Higher thresholds can be used if separation of closely spaced subpopulations is desired. Analysing only acceptor-containing molecules eliminates complications due to free donor-only molecules present in the reaction mix.
5. Calculate the apparent donor–acceptor stoichiometry parameter,

$$S = \left(F_{D_{ex}}^{A_{em}} + F_{D_{ex}}^{D_{em}}\right) / \left(F_{D_{ex}}^{A_{em}} + F_{D_{ex}}^{D_{em}} + F_{A_{ex}}^{A_{em}}\right)$$

6. Calculate the apparent donor-acceptor energy-transfer-efficiency parameter,

$$E^* = F_{D_{ex}}^{A_{em}} / \left(F_{D_{ex}}^{A_{em}} + F_{D_{ex}}^{D_{em}}\right)$$

7. Plot a 2D E^*-S histogram (**Fig. 4**) to identify the populations of different molecular species as determined by the parameter S, such that when (a) $S = 0.4$–0.9: Molecules of interest contain both donor and acceptor; (b) $S > 0.9$: donor-only molecules, arising from incomplete labelling or acceptor photobleaching; and (c) $S < 0.4$: Molecules contain only an acceptor, arising from incomplete labelling or donor photobleaching. In cases of an interaction of a labelled protein and a labelled DNA, donor-only and acceptor-only species represent free protein and free DNA molecules.
8. To calculate accurate energy transfer efficiencies and corresponding distances between fluorophores, E^* is converted to accurate E by correcting for the (a) background by measuring photon counts from buffer only (typically < 2 kHz), (b) donor leakage by plotting the 2D E^*-S histogram for all species and selecting the D-only species, (c) direct green laser excitation of acceptor by selecting the A-only species, and (d) taking into account the γ correction factor determined by two D-A species with a large difference of E (Note: γ is a correction factor that depends on the quantum yields of the fluorophores and detection efficiencies of their respective emission channels (9)). Accurate E is then used to calculate the distance R, since $E = 1/[1 + (R/R_o)^6]$, where R_o represents the Förster radius (the interprobe distance at which there is 50% energy transfer between donor and acceptor). For more details on data acquisition and analysis *see* **ref. 9**.

(*See* **Notes 13** and **14** for alternative single-molecule fluorescence methods to study DNA–protein interactions.)

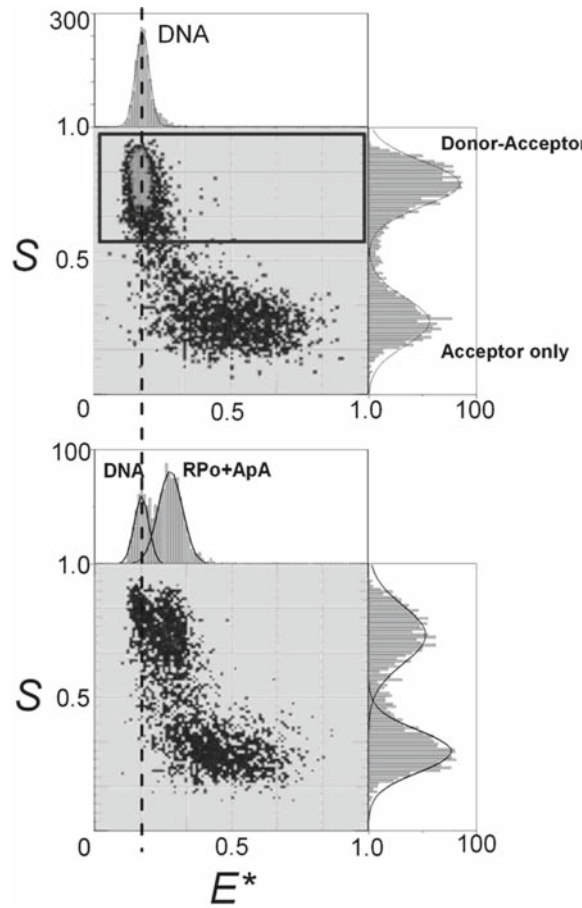


Fig. 4. Two-dimensional *E*-S* histograms of *lac*CONS DNA and transcription complex RP_o + ApA (RP_{itc2}). *Upper Panel*: Doubly labelled DNA (Cy3B as donor at DNA position –15; Alexa647 as acceptor at DNA position +15) represented by the donor–acceptor population (with $0.9 > S > 0.5$, in rectangular box) and collapsed on the 1D *E**-histogram. The acceptor-only population shows $S < 0.5$. In both panels, the donor-only population ($S > 0.9$) is not present, since only acceptor-containing molecules (i.e. molecules with appreciable $f_{\mathrm{Acx}}^{\mathrm{Am}}$ photon count) are displayed. *Lower Panel*: Experiment documenting contraction of DNA between positions –15 and +15. Two subpopulations can be observed in the donor–acceptor species: free doubly-labelled DNA molecules (with mean *E** as in panel A) and transcription complex RP_{itc2} (with higher mean *E**, mainly due to RNAP-induced DNA bending). Free DNA is present due to the dissociation of non-specific RNAP–DNA complexes during heparin challenge.

4. Notes

1. The chemicals should be of high purity (>99%) and of 'low-luminescence' or 'luminescence-spectroscopy' grade when available. All chemicals in buffers used for acquisition should be checked for fluorescence impurities.

2. Some fluorophores (e.g. TAMRA) are known to be quenched when placed on or within one or two nucleotides from a guanosine residue. This should be taken into account when deciding which dyes to use and where to place them on DNA.
3. Suboptimal labelling with fluorescent dyes. We found that use of some fluorophores (such as Alexa488 and ATTO488) results in poor labelling efficiency for internal DNA positions. The efficiency can be increased by repeated additions of reactive fluorophore in the reaction mix, e.g. addition of 10 nmol reactive fluorophore at time 0, an additional 10 nmol after 1 h, and a third addition of 10 nmol, after which the reaction can be left to proceed overnight. If the labelling efficiency is still low, an increase of the temperature of the reaction to 37°C or 55°C should increase the labelling efficiency.
4. Optimizing hybridization reactions. The yield of hybridized DNA can be increased if one doubly labelled oligonucleotide and one unlabelled oligonucleotide are used. In this case, a slight excess (10–50%) of unlabelled oligonucleotide in the annealing reaction can ensure complete hybridization.
5. PCR troubleshooting. In case of difficult sequences (e.g. sequences with palindromic DNA), the PCR may result in low yield of amplified DNA; yields can be improved by using the FailSafe™ Premix Selection Kit (Epicentre), which includes several reactions buffers with different additives that rapidly optimize PCR for different sequences of templates and primers.
6. Preparation of RNAP–DNA complexes takes ~30 min, and the RP_o is stable for 2–3 h for formation of transcription complexes and data collection. Note that this is for the *lac*CONS promoter DNA; for complexes with shorter lifetimes, storage of complexes and data acquisition should be adjusted to ensure that a high fraction of molecules is still in the complexed form.
7. Concentration ratio of RNAP:DNA is ~4:1. This is to ensure that most of the free DNA is bound to RNAP. Higher ratios of RNAP:DNA lead to non-specific binding of more than 1 RNAP molecule to DNA.
8. Excess RNAP present in the reaction mixture is removed by heparin-Sepharose challenge. Heparin-Sepharose is purchased as a powder and is stored as a 100 mg/ml stock suspension in 20 mM Tris–Cl (pH 8) and 20% ethanol at 4°C; the suspension should not be frozen. The suspension is diluted to 10 mg/ml in deionized water. The beads are washed five times by spinning down the suspension, removing the supernatant and adding another round of deionized water. The aliquots in water can be used for the next few weeks when stored at 4°C and should be discarded after a month.

9. Final concentration of sample for single-molecule measurements should be about 50–100 pM to reduce random coincidences between all species that will appear as a smear between the populations on the *E*-S* histogram.
10. Observation of single molecules is usually performed at 50–100 pM. If the equilibrium dissociation constant K_d for the protein–DNA interaction is at the low nM range and one component of the interaction (e.g. dsDNA) is doubly labelled, then unlabelled molecules (e.g. unlabelled protein) can be added to increase complex formation. This is difficult to address in studies of weak interactions between singly labelled components, since both components need to be in the 50–100 pM range, where only a small fraction of molecules is found in complexes when the system is in equilibrium. In this case, it is recommended that the complex is formed at nM–μM concentrations and that data acquisition starts immediately after dilution of the complexes to 100 pM (i.e. in order to observe complexes before they dissociate).
11. The ratio between the donor and acceptor laser excitation power will change the stoichiometry ratio, *S* and shift the populations along the *S* axis in the 2D *E*-S* histogram. This will in turn affect the resolution of separation of the populations along the *S* axis.
12. A data acquisition time period of 30 min at 100 pM typically leads to collection of > 10,000 bursts; the exact number depends on the brightness of the fluorophores and the setup alignment. The maximum photon count rates of Cy3B and Alexa647 are ~400 kHz and ~200 kHz, respectively. At a concentration of <<100 pM, data acquisition time should be adjusted to collect enough events for reliable statistical analysis.
13. Assays with immobilized molecules can be used to monitor dynamics of transcription complexes that occur at timescales slower than the timescale of diffusion. In such experiments, biotinylated DNA is immobilized on a coverslip using a biotinylated PEG-streptavidin method. The surface-immobilized molecules are illuminated by total internal reflection and imaged using an EMCCD camera *(13, 17)*.
14. Gel-based assays can also extend the diffusion time of the molecules in order to monitor dynamics of abortive initiation. Transcription complexes are formed according to Materials and Methods and run in a non-denaturing 5% polyacrylamide gel in 1× Tris–Glycine buffer at 120 V for 1 h for a clear separation of transcription complexes and free DNA. The dual-labelled band of the complexes can be excised and complexes can be monitored using single-molecule fluorescence spectroscopy *(18)*.

Acknowledgments

The help of the rest of the members of the Gene Machines' group at Physics Department, University of Oxford, is greatly acknowledged. We would also like to thank the groups of Richard Ebright and Shimon Weiss for a fruitful collaboration on transcription-related projects. Funding was provided by the UK IRC in Bionanotechnology, the European Union (MIRG-CT-2005-031079), and the National Institutes of Health (NIH R01 GM069709). MH was supported by a fellowship from the German Academic Exchange Service (DAAD).

References

1. Förster, T. (1948). Zwischenmolekulare Energiewanderung und Fluoreszenz, *Ann Phys* **2**, 55–75.
2. Stryer, L. (1978). Fluorescence energy transfer as a spectroscopic ruler, *Annu Rev Biochem* **47**, 819–846.
3. Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor, *Proc Natl Acad Sci U S A* **93**, 6264–6268.
4. Blanchard, S. C., Gonzalez, R. L., Kim, H. D., Chu, S., and Puglisi, J. D. (2004). tRNA selection and kinetic proofreading in translation, *Nat Struct Mol Biol* **11**, 1008–1014.
5. Ha, T. J. (2004). Structural dynamics and processing of nucleic acids revealed by single-molecule spectroscopy, *Biochemistry* **43**, 4055–4063.
6. Kapanidis, A. N., Laurence, T. A., Lee, N. K., Margeat, E., Kong, X. X., and Weiss, S. (2005). Alternating-laser excitation of single molecules, *Acc Chem Res* **38**, 523–533.
7. Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules, *Proc Natl Acad Sci U S A* **101**, 8936–8941.
8. Laurence, T. A., Kong, X., Jager, M., and Weiss, S. (2005). Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins, *Proc Natl Acad Sci U S A* **102**, 17348–17353.
9. Lee, N. K., Kapanidis, A. N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation, *Biophys J* **88**, 2939–2953.
10. Nir, E., Michalet, X., Hamadani, K. M., Laurence, T. A., Neuhauser, D., Kovchegov, Y., and Weiss, S. (2006). Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments, *J Phys Chem B* **110**, 22103–22124.
11. Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S., and Ebright, R. H. (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism, *Science* **314**, 1144–1147.
12. Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonjia, E., Mekler, V., Ebright, R. H., and Weiss, S. (2005). Retention of transcription initiation factor sigma(70) in transcription elongation: Single-molecule analysis, *Mol Cell* **20**, 347–356.
13. Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2006). Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes, *Biophys J* **90**, 1419–1431.
14. Sambrook, D. W., and Russell, J. (2001). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, NY.
15. Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W., and Ebright, R. H. (2001). Translocation of sigma(70) with RNA polymerase during transcription: Fluorescence resonance energy transfer assay for movement relative to DNA, *Cell* **106**, 453–463.
16. Kapanidis, A. N., Heilemann, M., Margeat, E., Kong, X., Nir, E., and Weiss, S. (2007). Alternating-Laser Excitation of Single Molecules,

in Single-Molecule Techniques: *A Laboratory Manual* (Selvin, P. R., and Ha, T., Eds.), Cold Spring Harbor Laboratory Press, New York, NY.

17. Ha, T., Rasnik, I., Cheng, W., Babcock, H. P., Gauss, G. H., Lohman, T. M., and Chu, S. (2002). Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase, *Nature* **419**, 638–641.
18. Hwang, L. C., Santoso, Y., and Kapanidis, A. N. Unpublished data.

Chapter 30

Analysis of DNA Supercoiling Induced by DNA–Protein Interactions

David J. Clark and Benoît Leblanc

Summary

Certain DNA-interacting proteins induce a pronounced bending in the double helix and cause topological stresses that are compensated by the formation of supercoils in DNA. Such supercoils, when forming on a circular plasmid, give rise to a series of topoisomers that run at different speeds during electrophoresis. The number of supercoils introduced in the plasmid can provide information on the protein; it can for example help determine the number of nucleosomes that are assembled on the plasmid or indicate whether the DNA-bending activity of a transcription factor is important enough to cause a topological stress. Because a DNA–protein activity can lead to either an overwinding or an underwinding of the helix, supercoiling can occur in either direction. Determining whether a plasmid contains positively or negatively supercoiled DNA is possible, thanks to an agarose gel containing an intercalating agent known to positively supercoil DNA, such as chloroquine. The speed of migration of the topoisomers varies in a characteristic way in the presence and absence of the agent. Topoisomer standards can furthermore be generated to allow the easy evaluation of the number of supercoils induced in a plasmid by a DNA–protein interaction.

Key words: DNA–protein interaction, Supercoiling, Topology, Chloroquine.

1. Introduction

The two strands of the DNA helix are unable to move around each other freely without the help of a nicking activity. Their overwinding or underwinding therefore results in the accumulation of a topological stress that must be somehow dissipated or compensated, as is easily visualized in an overwound rubber band that seeks to snap back into its original conformation or starts forming supercoils.

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_30

Certain DNA-binding proteins lead to a pronounced bending in the axis of the double helix. This is the case with histone octamers at the core of nucleosomes *(1)*, as well as with nonhistone proteins of the HMGB family *(2)* and with certain transcription factors containing an HMG-box motif such as the ribosomal transcription factor UBF *(3)*, the lymphoid enhancer-binding factor LEF-1 *(4)*, or the yeast mitochondrial factor Abf2 *(5)*. These protein-induced topological changes can be readily visualized as changes in the degree of supercoiling in a closed circular DNA molecule, and the analysis thereof can be performed on a simple agarose gel.

The topology of a circular DNA molecule is defined by three parameters: the linking number *Lk* (which is the number of times one strand of the double helix goes over the other one); the writhe *W* (the path of the DNA duplex through space); and the twist *Tw* (the number of base pairs found in one turn of the DNA helix). The value of *Tw* for a relaxed B-form DNA helix equals roughly 10.5 base pairs per turn *(6)*. Changing this value would normally result in a topological stress in a closed, circular DNA molecule and would have to be compensated by the other parameters. A change in twist can be introduced by the presence of intercalating agents such as ethidium bromide or chloroquine. A change in the linking number, meanwhile, can only occur after the covalent break and repair of at least one of the DNA strands that would allow it to pass over the other strand in one direction or the other; topoisomerases, for example, help relax DNA by performing such an action. The writhe, finally, is obviously influenced by the other two values but can be directly acted upon by a DNA-bending activity. Changes in the three parameters *Lk*, *Wr*, and *Tw* influence each other according to the equation

$$\Delta Lk = \Delta Wr + \Delta Tw.$$

As stated earlier, a topological stress can be compensated in a circular molecule by the generation of supercoils or loops of the double helix upon itself (another way of describing a change in writhe). Therefore, for a constant value of *Tw*, a circular DNA molecule on which a topoisomerase activity changes the linking number by one would compensate by contorting itself into a figure 8, the simplest form of supercoil. As the linking number can be changed in either orientation, the supercoil generated will have a value of −1 if the linking number is reduced or 1 if the linking number is increased.

In the cell, topoisomerases tend to remove supercoils since they relax topological stresses. It is the interaction of DNA with other proteins, such as occurs in the nucleosome, that can stabilize certain levels of supercoiling. In this particular case, it is generally accepted that a typical nucleosome stabilizes one negative supercoil *(7)*.

Circular DNA molecules like plasmids and episomes can be isolated and purified in such a way as to avoid the nicking of their phosphate backbone (which would result in their complete relaxation) and their level of supercoiling can be determined by electrophoresis on an agarose gel alongside topoisomer standards. This can be useful in determining the number of nucleosomes found on an episome, since that number would correspond to that of negative supercoils in the circular molecule *(7)*.

Supercoiling induced by a DNA-bending protein can also be studied *in vitro*. To that end, the protein is incubated with a circular DNA molecule. By forcing the DNA to bend, the protein changes its writhe. Topoisomerase I is then added to the DNA–protein complex and makes it possible for the circular DNA molecule to change its linking number value, helping to dissipate the topological stress caused by the initial bending of the helix. Unlike the writhe, which can be changed with no covalent bond being broken, the linking number remains the same unless a nicking activity is present. When the DNA-binding protein is removed during the next step, along with the topoisomerase, the bending-induced change in writhe will be lost due to the removal of the DNA-binding protein but the modified linking number will remain modified for lack of topoisomerase.

As this remaining change in linking number itself causes a topological stress in a closed circular molecule, it is in turn compensated by a change in writhe and the generation of supercoils. These supercoils, giving rise to a certain number of topoisomers, can be separated and observed on an agarose gel. Because topoisomers on a gel form distinct bands each differing from their immediate neighbors by one turn of helix (or one supercoil), it is possible to count how many turns in the linking number can be added or removed by a particular DNA–protein interaction.

To determine whether a DNA–protein interaction induces positive or negative supercoiling, the topoisomers are run on two gels in parallel, one in the absence and the other in the presence of an intercalating agent such as chloroquine *(8)*. Chloroquine, like ethidium bromide, inserts itself between the strands of the double helix and so artificially unwinds it, reducing the value of *Tw*. This is compensated by a forced change in writhe, and so the value of *Wr* increases. In simple terms, a plasmid run in a chloroquine gel appears more positively supercoiled than it would normally be.

As seen in **Fig. 1**, this can have either of two effects, depending on whether the plasmid was positively or negatively supercoiled to begin with. A supercoiled DNA plasmid runs faster than its relaxed form, irrespective of the orientation of the supercoiling. Topoisomers are therefore distributed on a gel between two positions: a very fast band that corresponds to highly supercoiled DNA, and a very slow band that corresponds to relaxed

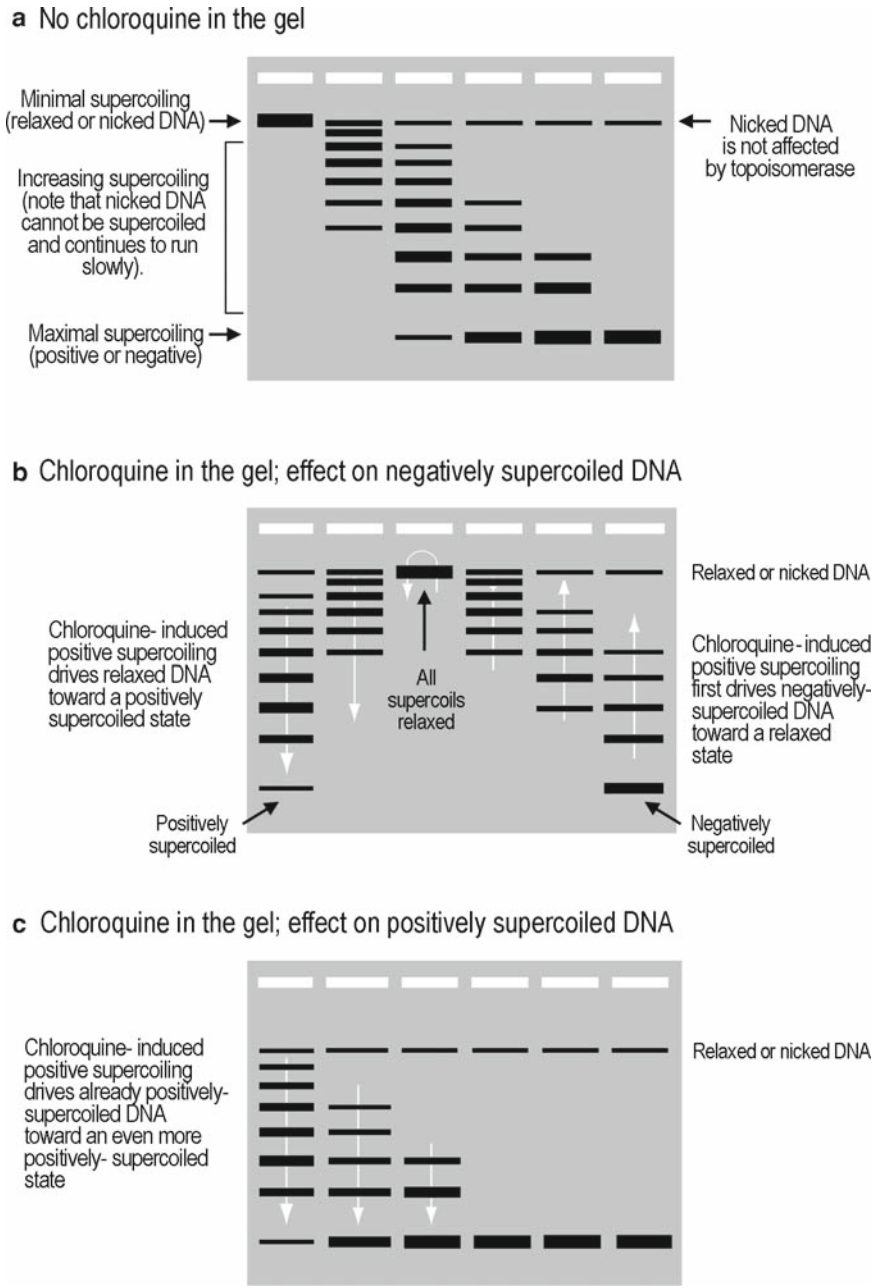


Fig. 1. Distribution of topoisomers on an agarose gel in the absence or presence of chloroquine. **(a)** Expected distribution of topoisomers with increasing levels of supercoiling on an agarose gel. Migration speed will be proportional to the degree of supercoiling, with relaxed plasmids running more slowly and progressively more supercoiled plasmids running faster and faster. Both positively and negatively supercoiled plasmids run faster than their relaxed form. Note that in all preparations of topoisomers, a small fraction of nicked plasmid is usually present and runs as a slow band. **(b)** Expected distribution of the topoisomers seen in **(a)**, if the DNA is initially negatively supercoiled and if the agarose gel contains chloroquine. The effect of chloroquine is to induce positive supercoiling, which here would relax the most negatively supercoiled molecules and cause them to migrate more slowly (on the *right* of the figure) or add positive supercoils to relaxed molecules and make them run faster (on the *left* of the figure). Note that the nicked plasmid running at the same position as relaxed DNA is unaffected by chloroquine. **(c)** Expected distribution of the topoisomers seen in **(a)** if the DNA is initially positively supercoiled and if the agarose gel contains chloroquine. The chloroquine would here increase the positive supercoiling of the DNA and cause it to run faster than in the conditions seen in **(a)**. Comparing the migration of a plasmid in the absence and the presence of chloroquine, it become possible to ascertain the direction of its supercoiling.

(or nicked) DNA. The different topoisomers that contain more than zero supercoils and less than the maximum possible number form a ladder between these two extremes.

If the test plasmid was initially positively supercoiled, it would run as a fast band on a normal agarose gel and even faster (being made even more positively supercoiled) on a chloroquine gel. On the other hand, if a plasmid was initially negatively supercoiled, it would run as a fast band on a normal agarose gel but as a slower band on a chloroquine gel (being made less negatively supercoiled or, in its particular case, less supercoiled, period). Note that with increasing concentrations of chloroquine, it would be possible to induce so much positive supercoiling in the latter plasmid that it would run as a slower and slower band until it reached the level of the relaxed band, after which it would start running faster and faster as an increasingly positively supercoiled molecule.

Topoisomer standards can be prepared from the test plasmid, without the use of a DNA-binding protein, and using only ethidium bromide and topoisomerase I *(7)*. Such standards, run alongside the topoisomers generated during an experiment, can make it easier to determine how many supercoils have been introduced on average in a plasmid by a specific interaction with a protein.

2. Materials

2.1. For the DNA–Protein Interaction Reaction.

1. 2× Reaction buffer (20 m*M* Hepes of pH 7,9, 80 m*M* KCl, 6 m*M* $MgCl_2$, 0.2 m*M* EDTA, 1 m*M* DTT) (*see* **Note 1**)
2. Topoisomerase I (New England Biolabs, M0301S)
3. 3 *M* sodium acetate
4. TE buffer (10 m*M* Tris–HCl, 1 m*M* EDTA, pH 8.0) (1 L)
5. tRNA, 10 mg/mL
6. Phenol:chloroform:isoamyl alcohol (25:24:1)
7. Ethanol
8. Ethanol, 70%
9. Bath at 37°C
10. A suitable plasmid diluted in H_2O to a working concentration of 20 ng/μL (*see* **Note 2**)
11. The protein extract to be tested

2.2. For the Preparation of Topoisomer Standards

1. Ethidium bromide (solid)
2. Topoisomerase I (New England Biolabs, M0301S)

3. 3 *M* sodium acetate
4. 10% Lauryl sulfate (SDS)
5. Phenol:chloroform:isoamyl alcohol (25:24:1)
6. Isopropanol
7. Ethanol, 70%
8. Isoamyl alcohol
9. TE buffer (10 m*M* Tris–HCl, 1 m*M* EDTA, pH 8)
10. Dialysis buffer (10 m*M* Tris–HCl, 1 m*M* EDTA, 3 *M* NaCl, pH 8) (2 L)
11. Dialysis tubing
12. Bath at 37°C
13. Stir plate (in a cold room or a refrigerated unit)

2.3. For Electrophoresis, Transfer, and Hybridization

1. Electrophoresis apparatus
2. TBE buffer (89 m*M* Tris base, 89 m*M* boric acid, 2 m*M* EDTA)
3. Electrophoresis-grade agarose
4. 5× loading buffer (0.1% bromophenol blue, 0.1% xylene cyanol, 15% glycerol, in TE)
5. Chloroquine diphosphate (Sigma #C6628-25G)
6. Ethidium-bromide-staining solution, 5 μg/mL (*see* **Note 3**)
7. Southern blot apparatus
8. Hybond N + transfer membrane (GE Healthcare) or the equivalent
9. Denaturing solution (0.5 *M* NaOH, 1.5 *M* NaCl)
10. Neutralizing solution (1 *M* Tris–HCl of pH 8.0, 1.5 *M* NaCl)
11. Hybridization rotating oven and hybridization bottles
12. 20× SSC (3 *M* NaCl, 0.3 *M* sodium citrate, pH 7.0)
13. 100× Denhardt's solution (2% Ficoll 400, 2% polyvinylpyrrolidone, 2% bovine serum albumin)
14. Prehybridization solution (6× SSC, 10× Denhardt's solution, 0.1% SDS, 0.1% sodium pyrophosphate) (*see* **Note 4**)
15. Single-stranded salmon sperm DNA (ssDNA), 10 mg/mL
16. Labeled probe (*see* **Note 5**)
17. Phosphorimager cassette (*see* **Note 6**)

3. Methods

The experiment should be performed in duplicate. One set of samples will be loaded on an agarose gel containing no chloroquine, the other set on a gel containing it (*see* **Note 7**).

3.1. Analysis of Protein-Induced DNA Bending

1. Clearly label as many Eppendorf tubes as will be required for the experiments. To each, add the appropriate amount of 2× buffer and H_2O with 20 ng of a suitable plasmid roughly 3,000 base pair long. Try to work in as small a volume as possible (*see* **Note 8**).
2. Add the protein sample to the reaction tubes. So as not to miss the most appropriate DNA–protein ratio, protein quantities should be distributed along a reasonable range and in regular increments (for example, 0.3, 9, 27, 91, 273 ng, where each tube contains three times the amount of the previous one). Keep on ice.
3. Allow the DNA–protein interaction to proceed for 20 min (*see* **Note 9**).
4. Add 1 U of topoisomerase I and continue the incubation at 37°C for 10 min (*see* **Note 10**).
5. To each tube, add 200 μL TE, 1 μL tRNA 10 mg/mL as carrier and adjust to 0.3 *M* sodium acetate (*see* **Note 11**).
6. Extract DNA with 200 μL phenol: chloroform: isoamyl alcohol; vortex briefly and centrifuge for 5 min on a tabletop microcentrifuge. Transfer the aqueous (top) phase in a fresh tube.
7. Precipitate DNA with one volume isopropanol. Centrifuge as earlier for 5 min.
8. Wash the DNA pellet with 100 μL ethanol 70%; centrifuge as earlier for 5 min. Air dry the pellet.
9. Add 8 μL TE buffer and 2 μL 5× loading buffer. Load on a 0.7% agarose/1× TBE gel containing either 0 or 10 μg/mL chloroquine. Run the gel at 70 V with buffer recirculation for 5–6 h, or until the xylene cyanol band (light blue) has traveled 7 cm on the gel (*see* **Notes 12** and **13**). An overnight run at a lower voltage is also appropriate (*see* **Note 12**).
10. Denature the DNA by soaking the gel in the denaturing solution with gentle rocking for 30 min.
11. Neutralize the gel by soaking the gel in two changes of neutralizing solution, with gentle rocking for 30 min each time.

12. Transfer the DNA to a Hybond N membrane using Southern blotting. Clearly mark the side on which the DNA is found.
13. Prehybridization: put the membrane in a hybridization bottle, DNA side facing the inside of the bottle. Add as small a volume of prewarmed (65°C) prehybridization solution as will allow complete covering of the membrane (usually 5 mL for small bottles and 10 mL for long bottles). Denature the ssDNA by heating at 95°C for 5 min and add to the prehybridization solution to a final concentration of 100 μg/mL. Allow prehybridization to proceed at 65°C for 1 h, with rotation.
14. Denature the labeled probe by heating at 95°C for 5 min and add to the prehybridization solution (do not pipet the undiluted probe straight on the membrane). Allow the hybridization to proceed for more than 6 h (overnight is fine) at 65°C, with rotation.
15. Wash the membrane with multiple changes of 50 mL SSC of decreasing concentration, 30 min at a time until the desired stringency is reached (3×, 1×, 0.1× SSC, for example) (*see* **Note 14**).
16. Dry the membrane and expose to a phosphorimager cassette or to an X-Ray film.

3.2. Preparation of Topoisomer Standards

It is not necessary to use topoisomer standards to determine if a particular protein induces supercoils in a plasmid, or to determine the direction of the supercoiling. Such standards can however be useful tools to count how many supercoils have been introduced in a molecule, as described in ref. *7*. The standards should be prepared with the same plasmid that will be used in the experiments.

1. Prepare a fresh ethidium bromide solution (2 mg/mL) from the solid stock (*see* **Note 15**). Protect from light.
2. From that solution, prepare a fresh dilution at 0.1 mg/mL. Protect from light.
3. Determine the concentration of the plasmid to be used with a spectrophotometer. Use an intact supercoiled plasmid, with as little nicked material as possible.
4. For each standard, in an Eppendorf tube, add 5 μg plasmid, 2 μL New England Biolabs buffer NEB4 (provided with the topoisomerase), 1 μL (5 U) topoisomerase I, (x) μL ethidium bromide 0.1 mg/mL and (y) μL H_2O (where x and y could be 0 and 7.5, 1.6 and 5.9, 3.1 and 4.6, 6.2 and 1.3, respectively, for a final volume of 20 μL). The higher the concentration of ethidium bromide, the more positively

supercoiled each standard will be. Protect from light (*see* **Note 16**).

5. Incubate in the dark at 37°C for 3 h.
6. Add 264 μL TE, 6 μL SDS 10%, and 30 μL 3 *M* sodium acetate to a final volume of 300 μL (for a final concentration of 0.2% SDS and 0.3 *M* sodium acetate).
7. Extract most of the ethidium bromide by adding one volume isoamyl alcohol to each tube, shaking briefly and discard the top phase. Work in dim light (*see* **Note 17**). Repeat four times.
8. Make sure the volume has remained at roughly 300 μL; if not, top up with H_2O (*see* **Note 18**).
9. Extract the DNA with one volume of phenol: chloroform: isoamyl alcohol. Vortex and centrifuge for 3 min in a table-top microcentrifuge. Recover the top (aqueous) phase.
10. Extract the aqueous phase again with one volume chloroform. Vortex and spin as above. Recover the top phase.
11. Dialyze each standard in an independent piece of dialysis tubing. Dialysis should be performed in the cold room and in the dark. The first dialysis is against 1 L dialysis buffer, overnight. The tubes are then dialyzed against 1 L of fresh dialysis buffer, still in the dark and in the cold room, for the rest of the day (at least 6 h). The tubes are finally dialyzed against 1 L TE buffer, still in the dark and in the cold room, for another overnight period.
12. The standards are recovered and kept in Eppendorf tubes. They should be analyzed by spectrometry or on gel to ascertain their concentration.

3.3. Analysis of the Chloroquine Gels

Topoisomers separated by gel electrophoresis are distributed between two extremes: the fastest running band, which corresponds to highly supercoiled plasmid (whether it be negatively or positively), and the slowest running band, which corresponds to relaxed plasmid (whether nicked or relaxed by topoisomerase). Plasmid DNA extracted from bacteria, as a rule, is negatively supercoiled and runs faster than its linear form.

Protein-induced supercoiling in a plasmid is apparent after the migration pattern of the generated topoisomers has been visualized and compared to the migration of the untreated plasmid.

On a normal agarose gel containing no chloroquine, any activity that introduces *positive* supercoils in a naturally negatively supercoiled plasmid would cause it to run more slowly (higher on the gel), because in effect it relaxes it somewhat. On the other hand, in the case of protein-induced *negative* supercoiling, the band might be expected to actually run faster than the untreated

control since more negative supercoiling is added to the already negative supercoils. That, however, may not be visible since the control DNA might already contain as many negative supercoils as it will tolerate.

To determine if a DNA molecule has been positively or negatively supercoiled, it has to be run on two gels in parallel: one without and one with chloroquine (which introduces positive supercoils in DNA). **Figure 1** shows how chloroquine will alter the course of supercoiled DNA. Panel A shows a normal gel on which a plasmid has been increasingly supercoiled by an increasing quantity of protein. Panel B illustrates what would happen if the samples seen in A were negatively supercoiled and run on a chloroquine gel. Because of the positive-inducing effect of chloroquine, the more heavily negatively supercoiled samples (on the right of the panel) are turned more positive, and so more relaxed, migrating more slowly. At the same time, the samples that were not supercoiled at all (at the left of the panel) acquire positive supercoils, which drive them toward the bottom of the gel. Note that in between these extremes, there is a degree of negative supercoiling that will be exactly compensated by the positive effect of the chloroquine, resulting in a relaxed band (third track from the left). Panel C illustrates what would happen if the topoisomers seen in A were initially positively supercoiled. In such a case, since there is no negative supercoil to be compensated, all bands are driven toward the bottom of the gel by the chloroquine.

4. Notes

1. This buffer has been successfully used with the transcription factor UBF from *Xenopus laevis* on a plasmid containing repeats of the transcriptional enhancers found in the animal's ribosomal gene spacer. It should be optimized for the protein of interest, but is expected to be the same as what would be used for any DNA-binding experiment like electromobility shift assays or DNAse I footprinting. Care should also be given to that buffer's compatibility with the topoisomerase used; certain commercial enzymes have specific salt requirements and incompatibilities. Depending on the degree of purity of the protein, a nonspecific competitor such as poly(dI-dC)·poly(dI-dC) can be included in the reaction.
2. The plasmid should be roughly 3,000 base pair long or shorter, as longer sizes require longer separation runs on agarose. The plasmid should contain one or multiple copies of the protein's binding site, as multiple binding will

amplify any supercoiling effect caused by the DNA–protein interaction.

3. The gels must be run in the absence of any staining agent, so as not to expose the DNA to intercalating agents other than chloroquine when appropriate. They must therefore be stained after migration. Here ethidium bromide is suggested, but other stains are quite acceptable (SYBR green, for example).
4. For 1 L, add 300 mL 20× SSC to 500 mL H_2O, then add 100 mL 100× Denhardt's solution, and only then add 5 mL SDS 20% and 1 g pyrophosphate. Top up with H_2O. This order of addition will prevent precipitation.
5. The probe can be radioactive or fluorescent, depending on availability and preference.
6. The use of a phosphorimager system is not essential; an X-ray film-based system is perfectly suitable.
7. The protocol described uses a small amount of DNA (20 ng) and requires blotting to reveal the supercoils. It would be theoretically possible to use high enough concentrations of DNA, protein, and topoisomerase to make direct visualization on the gel possible (after a postrun staining of the material), but this would require a lot of purified protein and more enzyme than is financially advisable.
8. The DNA–protein interaction will be favored by having all reagents in as concentrated a state as possible. The same holds true for the topoisomerase activity. A reaction volume in the 20–50 μL range is advised, if the protein concentration allows it. The initial reaction volume should allow for the later addition of topoisomerase.
9. Although the reaction described here is performed on ice, some DNA–protein reactions are routinely better observed at 30°C or 37°C. The experimenter must determine the best course for any specific protein analyzed.
10. This amount of topoisomerase I, according to the supplier (New England Biolabs), should relax most of 500 ng of supercoiled plasmid in 15 min. It is however very salt sensitive and its efficiency may vary depending on the exact composition of the reaction buffer used. Since the relaxation of the plasmid is essential to the experiment, it should be determined beforehand how much enzyme must be added to the test plasmid in order to relax it completely after 10 min at 37°C and in the experimental buffer (not the supplier's buffer).
11. The following steps describe the purification and precipitation of the DNA. Alternatively, loading buffer can be directly added to the reaction tubes and the DNA loaded straight

on the agarose gels. Although less clean because it does not guarantee the release of the protein from the DNA, this approach has worked in the past (e.g., **ref. 3**).

12. This represents a fairly long run, but is necessary for the topoisomers to clearly separate from one another. The faster the gel runs, the more it will heat, which could alter the gel and make the bands less sharp. Recirculation of the buffer is in such a case particularly important. Running the gel in the cold room is also an option to limit heating.
13. If a molecular weight marker is used, make sure that it is either diluted enough so as not to give too strong a signal if the probe hybridizes to it nonspecifically, or leave a few blank wells between it and the samples.
14. Stringency will have to be determined empirically.
15. Although ethidium bromide is not as toxic as popular laboratory folklore would have it, it remains a dangerous product and should be handled with care, especially in its solid form. Avoid contact with the skin and avoid generating (and especially breathing) any ethidium bromide dust. Wear protective clothing.
16. Ethidium bromide intercalated in DNA will cause nicking when exposed to light. Although it is not necessary to work in the dark, it is better to dim the lights as much as possible and work in a shaded place (such as a chemical hood).
17. The isoamyl alcohol in the top phase will remove most of the ethidium bromide. DNA will remain in the bottom phase. As isoamyl alcohol is rather noisome and will contain ethidium bromide, it is better to work in a chemical hood (which also helps with the low light).
18. The volume is not absolutely crucial here, but to maintain it around 300 μL will help with the upcoming phenol extractions. As a crude measuring device, we use a home-made graduated Eppendorf tube where we marked volumes at 100 μL increments.

Acknowledgments

The work in the laboratory of B. L. is funded by the Natural Sciences and Engineering Research Council of Canada and by the Canada Foundation for Innovation.

References

1. Simpson, R. T., Thoma, F., and Brubaker, J. M. (1985). Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for study of higher order structure. *Cell* **42**, 799–808.
2. Javaherian, K., Liu, J. F., and Wang, J. C. (1978). Nonhistone proteins HMG1 and HMG2 change the DNA helical structure. *Science* **199**, 1345–1346.
3. Bazett-Jones, D. P., Leblanc, B., Herfort, M., and Moss, T. (1994). Short-range DNA looping by the Xenopus HMG-box transcription factor, xUBF. *Science* **264**, 1134–1137.
4. Giese, K., Pagel, J., and Grosschedl, R. (1997). Functional analysis of DNA bending and unwinding by the high mobility group domain of LEF-1. *Proc. Natl. Acad. Sci. U S A* **94**, 12845–12850.
5. Diffley, J. F. and Stillman, B. (1992). DNA binding properties of an HMG1-related protein from yeast mitochondria. *J. Biol. Chem.* **267**, 3368–3374.
6. Crick, F. H. (1976). Linking numbers and nucleosomes. *Proc. Natl. Acad. Sci. U S A* **73**, 2639–2643.
7. Clark, D. J. (1998). Counting nucleosome cores on circular DNA using topoisomerase I. In Gould. H (ed.) *Chromatin: A Practical Approach*. Oxford University Press, Oxford.
8. Shure, M., Pulleyblank, D. E., and Vinograd, J. (1977). The problems of eukaryotic and prokaryotic DNA packaging and in vivo conformation posed by superhelix density heterogeneity. *Nucleic Acids Res.* **4**, 1183–1205.

Chapter 31

The Cruciform DNA Mobility Shift Assay: A Tool to Study Proteins that Recognize Bent DNA

Victor Y. Stefanovsky and Tom Moss

Summary

So-called architectural DNA binding proteins such as those of the HMGB-box family induce DNA bending and kinking. However, these proteins often display only a weak sequence preference, making the analysis of their DNA binding characteristics difficult if not impossible in a standard electrophoretic mobility assay (EMSA). In contrast, such proteins often bind prebent DNAs with high affinity and specificity. A synthetic cruciform DNA structure will often provide an ideal binding site for such proteins, allowing their affinities for both bent and linear DNAs to be directly and simply determined by a modified form of EMSA.

Key words: Architectural proteins, HMGB-box protein, Bent DNA, Cruciform, EMSA.

1. Introduction

The first reported interaction of an HMGB-box protein (HMGB1) with a bent stable synthetic DNA structure suggested that this family of proteins displayed an intrinsic affinity for bent DNA *(1)*. The assay was based on the electrophoretic mobility shift (EMSA) of an *in vitro* assembled synthetic cruciform when it was bound by an HMGB box. It was found that binding to the prebent cruciform DNA occurred with a much higher affinity and specificity than to linear DNA. Shortly after, the same group demonstrated an identical behavior for the sequence-specific transcription factor SRY, which contains a single HMG box *(2)*, followed by similar reports on the other HMGB-box proteins LEF-1/TCF 1 *(3, 4)*. The RNA Polymerase I transcription

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_31

factor UBF, like HMGB1, is a nonsequence-specific DNA binding protein with multiple HMGB boxes, and as such its binding properties were difficult to analyze by EMSA using linear DNA. However, it was found to bind strongly to cruciform DNA *(5)*. A detailed study showed that the individual HMGB boxes 1 and 2 of UBF, responsible for the rDNA promoter in-phase bending and enhancesome formation *(6, 7)*, each bind with high affinity to cruciform structures *(8)*. Moreover, competition with linear DNA proved to be a valuable tool for detecting changes in the binding affinity of these boxes as a result of mutations or post-translational modifications. Such changes are believed to reflect structural changes in the protein–DNA complex due to an altered bending capacity of the HMGB boxes.

The 14-3-3 family of proteins were found to specifically recognize cruciform structures at origins of DNA replication in a cell-cycle dependent manner and were identified as regulators of DNA replication *(9)*, and the cruciform mobility shift assay has also been used for analyzing these proteins.

The cruciform mobility shift assay has, thus, shown itself to be a useful tool in the study of proteins that display enhanced affinity for bent DNA. Here we present the basic assay using the HMGB boxes as example.

2. Materials

1. Stock solutions of cruciform oligonucleotides at 10 pmoles/μl in ddH_2O;Oligo 1 5′-CCCTATAACCCCTGCATTGAATTCCAGTCTGATAA-3′ Oligo 2 5′-GTAGTCGTGATAGGTGCAGGGGTTATAGGG-3′ Oligo 3 5′-AACAGTAGCTCTTATTCGAGCTCGCGCCCTATCACGACTA-3′ Oligo 4 5′-TTTATCAGACTGGAATTCAAGCGCGAGCTCGAATAAGAGCTACTGT-3′ These oligonucleotides are designed in order to anneal with each other and form a cruciform structure, **Fig. 1**.
2. [γ-^{32}P]-ATP (Perkin-Elmer).
3. T4 Polynucleotide Kinase (NEB).
4. 7.5 M ammonium acetate.
5. 95 and 70% ethanol.
6. TMS annealing buffer: 100 mM NaCl, 10 mM Tris–HCl of pH 7.5, 10 mM $MgCl_2$.
7. 2× Binding buffer: 16% Ficoll, 200 mM NaCl, 20 mM HEPES of pH 7.9, 10 mM KCl, 2 mM EDTA, 2 mM spermidine, 1 mM DTT.
8. TBS:10 mM Tris–HCl of pH 7.5, 100 mM NaCl.

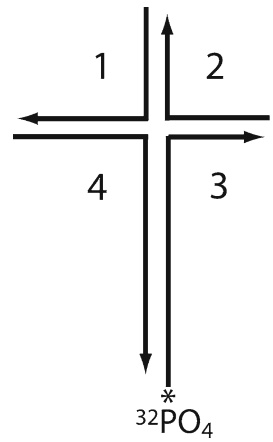


Fig. 1. Schematic formation of a cruciform structure. The arrows indicate the 3′-termini of the annealed oligonucleotides 1, 2, 3, and 4; The asterisk indicates the radioactive $^{32}PO_4$ group on oligonucleotide 3.

9. TBE (10×) Prepare 1 L by mixing 108 g Tris-base, 55 g boric acid, 40 ml 0.5 M EDTA, pH 8.0.
10. 40% acrylamide stock solution (38 parts acrylamide: 2 parts bisacrylamide).
11. 0.1% Xylene cyanol.
12. Gel-loading buffer (10×) 0.25% bromophenol blue, 0.25% xylene cyanol, 25% Ficoll 400.
13. Prepare multiple 0.5-ml Eppendorf tubes with pierced bottom (make several holes with an 18–21-gauge syringe needle) and fill the bottom of the tubes with glass wool (keep stock in 70% ethanol). Autoclave and keep in a sterile jar until use.
14. For competition experiments: double-stranded linear DNA containing the target DNA sequence of interest, stock solution at least 200–500 μg/ml in ddH_2O.
15. For antibody supershift: specific antibodies against the protein of interest in serum dilutions ranging from undiluted to 1:1,000 in TBS.

3. Methods

3.1. 5′ End Labeling of Oligonucleotide 3

1. Take 2 μl (20 pmoles) oligo 3 and add 2 μl 10× Polynucleotide Kinase buffer (NEB) 10 μl [γ-^{32}P]-ATP, 5 μl H_2O and 1 μl T4 Polynucleotide Kinase (~10 units) (NEB)
2. Incubate for 1 h at 37°C.

3. Add 10 ml 7.5 M ammonium acetate (final concentration: 2.5 M).
4. Add 270 µl 95% Ethanol.
5. Leave 5 min, at RT.
6. Spin in an Eppendorf centrifuge 5 min at 13,000 rpm.
7. Carefully discard supernatant into radioactive waste, add 300 µl 70% ethanol.
8. Spin 1 min 13,000 rpm.
9. Carefully discard supernatant into radioactive waste.

3.2. Annealing of the Oligonucleotides and Isolation of the Labeled Cruciform

1. Add 2 µl (20 pmoles) each of oligos 1, 2, and 4 to precipitated, labeled oligo 3.
2. Add 25 µl TMS and mix thoroughly to redissolve oligo 3.
3. Take a small aliquot to determine approximate specific activity, that is total Cerenkov cpm/20 pmole oligo 3.
4. Place in an aluminum heating block at 90°C, switch off heating immediately, and insulate by covering with a Styrofoam box. Leave for 3 h to anneal cruciform.

Alternative method: Leave 3 min at 90°C, then 10 min at 68°C, and 30 min at 37°C. This can be easily performed in a thermal cycler.

5. Add 3 µl 10× gel-loading buffer.
6. Load into a 1-cm wide pocket on a 1-mm-thick, 20-cm-long 6.5% polyacrylamide gel in 1× TBE.
7. Run at 10 V/cm for 2–3 h until bromophenol blue has migrated about 14 cm.
8. Remove the upper plate of the gel, cover with Saran wrap, and autoradiograph for 30 s to 1 min. Ideally, use fluorescent ink markers to allow realignment of film to gel. Alternatively, use one corner of a radiography cassette to align the film with the gel plate during exposure.
9. Realign the developed film under the gel on a transilluminator and excise the cruciform band using a scalpel. The cruciform migrates slightly below the xylene cyanol band on a 6.5% gel (*see* **Fig. 2** and **Note 1**).

3.3. Extraction of Cruciform

1. Place the cut gel fragment containing the cruciform in a 0.5-ml Eppendorf tube. Pierce several holes through the bottom of the tube with a gauge 18–21 syringe needle. Place the tube in a 1.5-ml Eppendorf tube without lid.
2. Centrifuge 20–30 s or until the gel passes through the holes into the bottom tube forming a gel pellet.

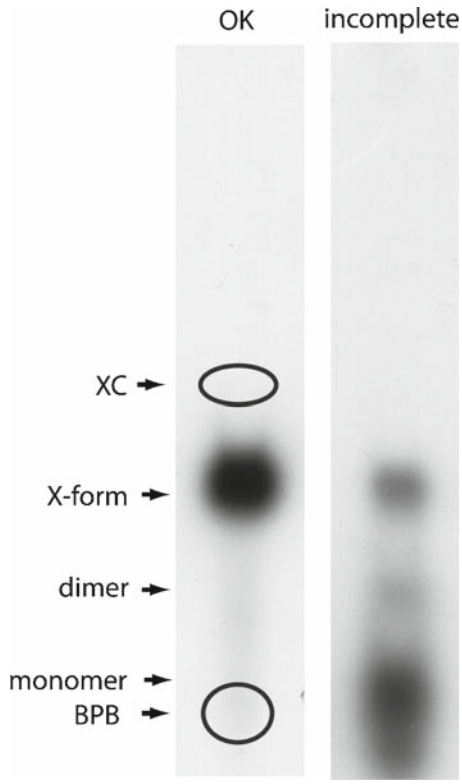


Fig. 2. Migration of the cruciform on the preparative gel. The cruciform structure was resolved on a 6.5% polyacrylamide gel and exposed as described in Methods. *Left panel*: position of the cruciform (X-form) relative to the dyes (*XC* xylene cyanol and *BPB* bromophenol blue) indicated by *black ellipses*. *Right panel*: an example of incomplete annealing. Positions of the monomer and dimer forms are indicated on the *left* side of the figure.

3. Add 300 µl TMS to the pellet to form a slurry.
4. Seal the tube and leave overnight at 4°C.
5. Put a previously prepared glass wool-bottomed, pierced, autoclaved 0.5-ml tube (*see* **Subheading 2**, **item 13**) in an intact 1.5-ml Eppendorf without lid.
6. To recover eluate, transfer all the gel slurry into the 0.5-ml tube using a 1-ml pipette with large tip opening (cut-off) and centrifuge for 20 s.
7. Adjust the eluate volume to 400 µl with TMS.

The eluted cruciform DNA should be at ~50 fmole/µl and nearly all the radioactivity should have been eluted. Cerenkov count a 2-µl aliquot in order to calculate the concentration of cruciform using the specific activity calculated from **Subheading 3.2**, **step 3**. Stored at 4°C the labeled cruciform is stable for at least a week. You will need about 100 fmoles (2 µl) per mobility shift assay.

3.4. Cruciform Mobility Shift Assay

Each mobility shift reaction is performed in 10 µl consisting of 5 µl of 2× Binding Buffer, 2 µl (100 fmole) of cruciform DNA in TMS, and 3 µl of TBS containing varying amounts of the proteins to be assayed.

1. Prepare a 6.5% polyacrylamide gel (acrylamide:bisacrylamide 38:2) at least 15–20 cm long and 1-mm thick in 0.5× TBE and use a 0.5% TBE as running buffer. Prerun the gel for 1 h at 11 V/cm.
2. Prepare a stock mix of 2× Binding Buffer (5 µl × number of reactions plus one) and cruciform (2 µl (100 fmole) × number of reactions plus one).
3. Prepare dilutions of the protein to be analyzed in Eppendorf tubes. Complete the volume of the protein in each tube to 3 µl with TBS. Place 3 µl TBS in a tube to serve as a negative control. Since the K_d of the expected complexes is typically less than µmolar, start with protein amounts ranging from 2 to 30 pmoles.
4. Add 7 µl binding buffer/cruciform mix, as in **step 2**, to each protein dilution.
5. Incubate 10–30 min at room temperature.
6. Add 0.5 µl 0.1% xylene cyanol to each sample and load onto the gel.
7. Electrophorese for 3–4 h at 11 V/cm.
8. Transfer the gel onto a sheet of Whatman 3-MM paper and cover with Saran wrap.
9. Dry the gel for 30 min at 85°C.
10. Expose the gel to radiography film or use a commercial phosphoimaging device to detect cruciform and analyze the results. A typical example of the electrophoretic separation is shown in **Fig. 3** and **Notes 2** and **3**.

3.5. Competition with Linear DNA

The affinity of the protein for linear relative to cruciform DNA can be determined in a simple competition assay (*see* **Fig. 4**). Typically, a 1- to 10-fold molar excess of the linear DNA fragment over the protein is required. Add 1 µl of the double-stranded linear fragment to the cruciform DNA before mixing with the protein sample in **step 3** of **Subheading 3.4**, then proceed as described in that section. A protein concentration that yields less than complete shifting of the cruciform in the absence of linear DNA must be used if the relative affinity of the protein for linear DNA is to be determined.

3.6. Supershifting with a Specific Protein Antibody

In the case of impure protein samples, it may be necessary to determine the identity of the protein that is responsible for the cruciform shift. This can sometimes be achieved by "up-shifting"

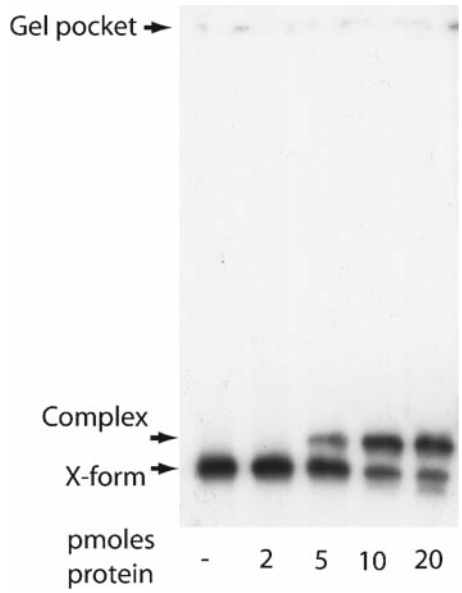


Fig. 3. A typical example of cruciform shift assay. About 100 fmoles of cruciform was incubated with the indicated amounts of HMGB box1 from UBF, *see* **Subheading 2**. The protein–DNA complex (complex) and the naked cruciform (X-form) are indicated.

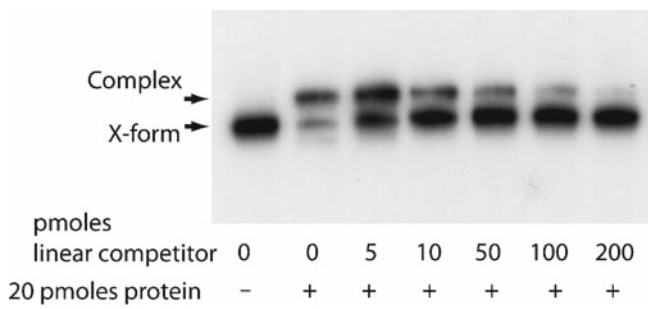


Fig. 4. Competition with linear target sequence. The cruciform structure was incubated with 20 pmoles of UBF HMGB box1(as in **Fig. 3**) and with increasing amounts of linear human rDNA promoter UCE fragment, *see* **ref. 8**. In this case the cruciform shift is efficiently competed with 200 pmoles of the linear promoter fragment. Complex and X-form as in **Fig. 3**.

the cruciform-protein complex by the addition of an antibody before electrophoretic analysis. Add 1 µl of an appropriate range of antibody dilutions to the samples just before applying them to the gel, then proceed as described in that section. An example of such an upshift assay is shown in **Fig. 5**.

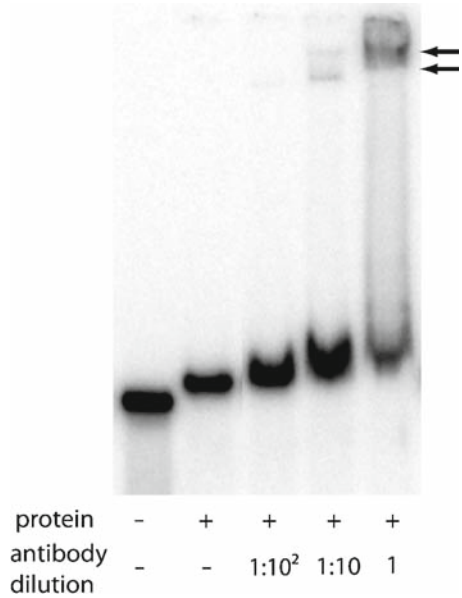


Fig. 5. Specific antibody "supershift." The cruciform structure (100 fmoles) was incubated with 20 pmoles of UBF HMGB box1 and 1 μl diluted polyclonal anti-UBF antibody as indicated. An increasing large "supershift" is observed with increase in antibody concentration, eventually producing antibody-protein–DNA complexes high in the gel indicated by *arrows*.

4. Notes

1. Annealing of the cruciform is not complete (*see* **Fig. 1B**, right panel). Usually, the annealing is so efficient that only the completed cruciform, containing all four oligos, is visible on the gel. In some cases, however, smaller structures containing one or two oligos may be present, running closer to the bromphenol blue. If the cruciform is not the major product:-
   i) Repeat the annealing after verifying the identity and quality of oligonucleotides.
   ii) Check the concentration of your oligonucleotides. Equimolar amounts must be used.
   iii) The gel may have been run too hot. Check the voltage, it should not exceed 11V/cm.
2. More than one shifted band is visible on the gel. There may be a second cruciform binding protein in the protein preparation. A cruciform in some cases may bind more than one protein molecule. *See* also **Note 3**.
3. Certain protein-DNA complexes migrate faster than the naked cruciform (downshift). Dependent on buffer conditions cruciforms may adopt alternative tertiary folds, affecting

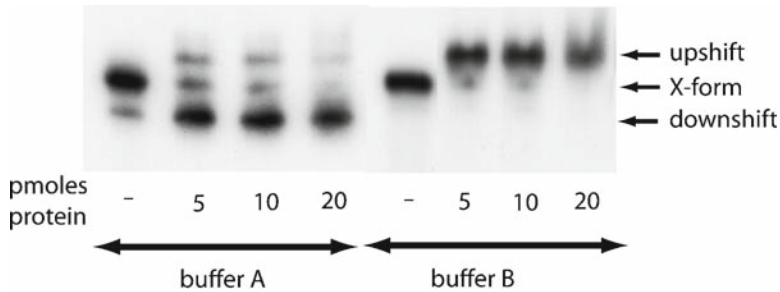


Fig. 6. Differential migration of the protein–cruciform DNA complex depending on buffer conditions. Increasing amounts of UBF HMGB box2 were incubated with the cruciform using two different binding buffers: buffer A, the standard binding reaction (2 mM $MgCl_2$ final concentration) and buffer B, as buffer A but giving final concentrations of 5 mM $MgCl_2$ and 0.3 mM ATP. Positions of the free cruciform (X-form), the "upshift" and the "downshift" are indicated with *arrows*.

their mobility during electrophoresis. For example, HMGB-box 2 of UBF under standard buffer conditions gave a downshift as well as an upshift. However, addition of an extra 3 mM $MgCl_2$ during the binding reaction eliminated the downshift, *see* **Fig. 6**. It is possible that acrylamide concentration or bisacrylamide:acrylamide ratio in the gel may also affect relative migration of cruciform structural isomers.

Acknowledgments

This work was supported by an operating grant from the Canadian Institutes of Health Research.

References

1. Bianchi, M. E., Beltrame, M., and Paonessa, G. (1989). Specific recognition of cruciform DNA by nuclear protein HMG1, *Science* **243**, 1056–1059.
2. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1992). SRY, like HMG1, recognizes sharp angles in DNA, *EMBO J.* **11**, 4497–4506.
3. Lilley, D. M. (1992). DNA–protein interactions. HMG has DNA wrapped up, *Nature* **357**, 282–283.
4. Van de Wetering, M., and Clevers, H. (1992). Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix, *EMBO J.* **11**, 3039–3044.
5. Kuhn, A., Stefanovsky, V., and Grummt, I. (1993). The nucleolar transcription activator UBF relieves Ku antigen-mediated repression of mouse ribosomal gene transcription, *Nucleic Acids Res.* **21**, 2057–2063.
6. Bazett-Jones, D. P., Leblanc, B., Herfort, M., and Moss, T. (1994). Short-range DNA looping

by the Xenopus HMG-box transcription factor, xUBF, *Science* **264**, 1134–1137.

7. Stefanovsky, V. Y., Bazett-Jones, D. P., Pelletier, G., and Moss, T. (1996). The DNA supercoiling architecture induced by the transcription factor xUBF requires three of its five HMG-boxes, *Nucleic Acids Res.* **24**, 3208–3215.
8. Stefanovsky, V. Y., Pelletier, G., Bazett-Jones, D. P., and Moss, T. (2006). ERK modulates DNA bending and Enhancesome Structure by phosphorylating HMG1-boxes 1 and 2 of the RNA polymerase I transcription factor UBF, *Biochemistry* **45**, 3626–3634.
9. Zannis-Hadjopoulos, M., Yahyaoui, W., and Callejo, M. (2008). 14-3-3 cruciform-binding proteins as regulators of eukaryotic DNA replication, *Trends Biochem. Sci.* **33**, 44–50.

Chapter 32

Plasmid Vectors for the Analysis of Protein-Induced DNA Bending

Christian Zwieb and Sankar Adhya

Summary

Bending is not only required to accommodate DNA within the cell but also is a mechanism used by proteins to initiate DNA replication, transcription, and recombination. Determining the angles by which regulatory DNA segments deviate from linearity upon binding of proteins is a necessary step toward a better understanding of a large number of essential biological functions. The pBend plasmids contain duplicate sets of restriction sites and, when combined with "gel shift" experiments, allow the straightforward determination of the bending angle in a DNA molecule. The steps for successfully carrying out a binding/bending experiment are described. They include the cloning of the protein-binding site into the chosen pBend vector, the isolation of a series of DNA fragments with identical in length but variable placing of the protein-binding site, and the gel electrophoretic analysis of the free and protein-bound fragments.

Key words: DNA bending, Protein–DNA complexes, Replication, Recombination, Transcription.

1. Introduction

Bending of DNA by proteins plays an important role in transcription initiation, DNA replication, and recombination. The degree of protein-induced DNA bending can be simply and conveniently determined by combining gel electrophoresis of DNA–protein complexes with the use of special plasmid vectors carrying the bendable DNA sequence *(1, 2)*. The vectors contain duplicate sets of restriction sites in a direct repeat order as well as cloning sites for insertion of protein-binding sequences between the two sets. Restriction enzyme digestion readily generates fragments

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_32

which are identical in size but differ in the location of the binding site (*see* **Fig. 1**). The mobility of a DNA fragment is less when a bend is located at the center of the fragment than when the same bend is located toward one or the other end *(3, 4)*. The bending angle α is defined as the angle by which a segment of the rod-like DNA bends away from linearity. Thus, the bending angle is 0° for a straight DNA fragment. α can be estimated by measuring μ_M (mobility of the complex with the protein bound at the middle of the fragment) and μ_E (mobility of the complex with the protein bound near the end of a DNA fragment) using an empirical relationship $\mu_M/\mu_E = \cos(\alpha/2)$ *(4, 5)*. To carry out a bending experiment, the protein-binding site is inserted into the bending vector. Next, a series of DNA fragments with the binding site located either in the middle or toward the end is generated by digestion of the vector with different restriction enzymes. Finally, DNA–protein complexes formed with the series of DNA fragments are analyzed by gel electrophoresis to determine the bending angle α. As an example, we illustrate the bending at the *lac* promoter by cyclic AMP receptor protein (CRP) *(4, 6)*. However, the method is easily adapted to other protein–DNA complexes if they can be separated from free DNA by gel electrophoresis. General considerations for successful complex formation and for the selection of restriction sites are discussed in **Subheading 4**. pBend3, 4, and 5 (**Figs. 1** and **2)** are different

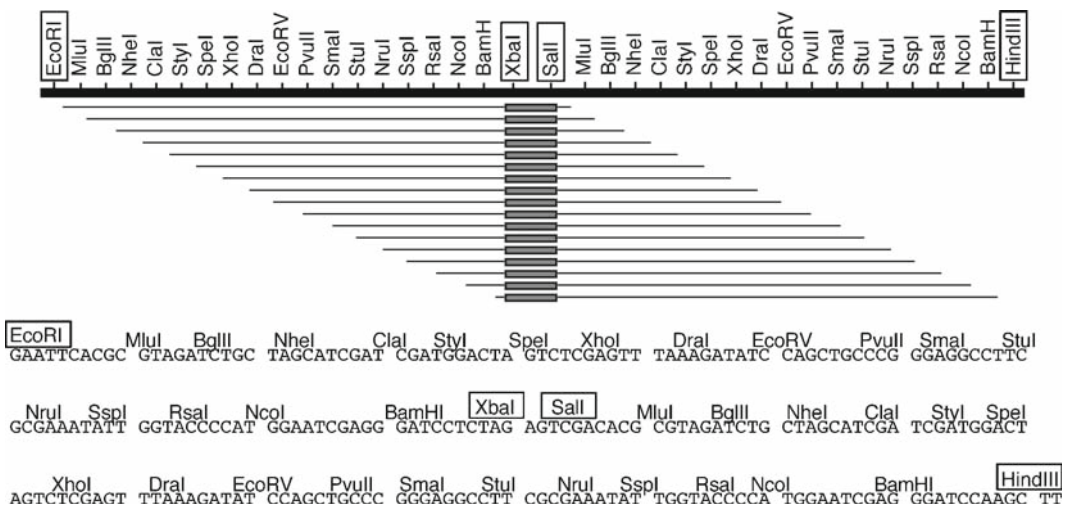


Fig. 1. Schematic representation of the pBend3 insert between the *Eco*RI–*Hin*dIII sites. pBend3 was constructed by cloning of the 236-base pair *Eco*RI–*Hin*dIII fragment of pBend*2 (4)* into pBluescript SK (Stratagene). pBluescript is a high copy number plasmid and generates a large amount of DNA upon plasmid extraction. The *Eco*RI-*Hin*dIII fragment contains 17 duplicated restriction sites to generate DNA fragments of identical length, but in which with the protein-binding sequence (gray rectangle in the upper part of the Figure) is shifted. The sites *Xba*I and *Sal*I (in boxes) are unique and suitable for cloning of the protein-binding sequence. Restriction sites are not drawn to scale. The sequence of the insert is shown in the lower part of the figure.

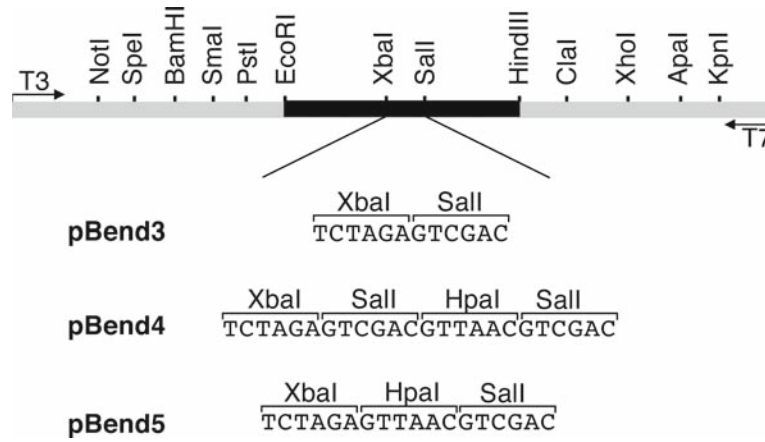


Fig. 2. Restriction and cloning sites of pBend3, 4, and 5. Portions of pBluescript SK (Stratagene) are shown in gray; the *Xba*I and *Sal*I sites of pBluescript are abolished by partial digestion with *Xba*I or *Sal*I, filling in with DNA polymerase (Klenow), and ligation with T_4 DNA ligase. Promoters for T_3 and T_7 RNA polymerase are indicated by the arrows. They can be used to study the bending by proteins which bind to double-stranded RNA or to RNA–DNA hybrids. The region between the *Eco*RI and *Hin*dIII sites (indicated in black) is identical to the one shown in Fig. 1. pBend4 and pBend5 contain additional *Sal*I and *Hpa*I sites as shown.

versions of the bending vector pBend2 *(4)*. pBend2 and pBend3 contain the same 236 base pair *Eco*RI–*Hin*dIII fragment with 17 duplicated restriction sites. Because of the higher copy number of pBend3, preparation of plasmid DNA is more efficient. The remainder of the digested vector DNA is less likely to comigrate with protein–DNA complexes; therefore, high molecular weight DNA-binding proteins that might introduce substantial bending can be analyzed without the tedious purification of individual DNA fragments containing the protein-binding site. pBend4 and pBend5 contain an additional HpaI cloning site to facilitate the insertion of DNAs with blunt ends. The promoters for the T_3- and T_7 RNA polymerases have potential use for analyzing the bending of double-stranded RNA or RNA/DNA hybrids by proteins such as transcription factor TFIIIA *(7)*. Details of the properties of pBend3, 4, and 5 are described in the legends to **Figs. 1–3**.

2. Materials

2.1. Insertion of the Protein-Binding Site into pBend

1. pBend vector (available from the ATCC at http://www.atcc.org/. pBend3 accession number: 87121; pBend4: 87122; pBend5: 87123; pBendBlue *(8)*: 87690).

Fig. 3. Restriction and cloning sites of pBendBlue along with the ORF of the *lacZ*(α) gene passing through the entire permuted sequence of pBend3, including the two cloning sites (*Xba*I and *Sal*I). The 241-base pair permuted sequence of pBend, including the *Xba*I and *Sal*I cloning sites (*shown in gray*), was inserted between the *Sac*I and *Apa*I sites of plasmid pBSK (Stratagene) and is cloned such that the *lacZ* ORF remains uninterrupted. The corresponding amino acid sequence is shown above the DNA sequence. Note that pBSK carries only the α-complementing portion of *LacZ*, necessitating the use of the Δ*lacZ*α host strain (*Epicurian coli*, Stratagene) which is a ω donor. A complementation between α and ω makes a cell β-galactosidase proficient. The insert is bounded by T_3 and T_7 RNA polymerase promoter sequences, both within the ORF.

2. Oligonucleotides or DNA-fragments containing the investigated protein-binding site, compatible with the cloning sites of pBend (*see* **Notes 1** and **2**).
3. *Restriction enzymes. Hpa*I, *Eco*RI, *Hin*dIII, and *Mlu*I at about 10 U/μL with 10 times concentrated digestion buffers as specified or provided by the vendors.
4. *500 mM EDTA of pH 8.* Mix 93 g of disodium ethylene diamine tetraacetate-$2H_2O$ in 400 mL water, adjust pH to 8.0 by adding about 20 g of NaOH pellets (only then will EDTA dissolve completely), adjust volume to 500 mL, autoclave.
5. *TE.* 10 mM Tris–HCl of pH 7.5, 1 mM EDTA, mix 1 mL of 1 M Tris–HCl of pH 7.5 and 200 μL of 500 mM EDTA of pH 8. in a total volume of 100 mL of water. Store at 4°C.
6. *7.5 M ammonium acetate.* Dissolve 57.8 g ammonium acetate in a total volume of 100 mL of water. Store at 4°C.
7. *80% ethanol.* Mix 80 mL ethanol with 20 mL water. Store at 4°C.
8. *Oligonucleotide annealing buffer.* Mix 1 mL of 1 M Tris–HCl of pH 7.5, 200 μL of 500 mM EDTA of pH 8 and 10 mL of 1 M NaCl in a total volume of 100 mL.

9. *Ligation buffer, 5 times concentrated.* 500 mM Tris–HCl of pH 7.6, 100 mM MgCl2, 100 mM DTT. Store at –20°C.
10. *5 mM ATP.* Dissolve 3 mg of ATP (disodium salt) in 1 mL of TE.
11. T_4 polynucleotide kinase: 10 U/μL.
12. T_4 DNA ligase: 1 U/μL.
13. Competent *E. coli* cells: e.g., strain DH5α.
14. *LB-amp plates.* Suspend 10 g of LB powder and 7.5 g of Bacto-agar in 500 mL of water; autoclave and dissolve agar by swirling, cool solution to 55°C in a water bath, add 50 mg of ampicillin (for final concentration of 100 μg/mL), dissolve by swirling and pour plates, store plates in a plastic bag at 4°C.
15. *LB-media.* Suspend 10 g of LB powder in 500 mL water and autoclave. Store at room temperature. For the ampicillin containing LB, add and dissolve the specified amount to the media at room temperature.
16. *Tris–sucrose.* Dissolve 10 g of sucrose and 5 mL of 1 M Tris–HCl of pH 7.8 in a total of 100 mL water (prepare fresh).
17. *Lysozyme solution.* 10 mg/mL in 250 mM Tris–HCl of pH 8, keep frozen aliquots at –20°C, thaw once immediately before use, discard unused portions.
18. *200 mM EDTA, pH 8.* Prepare by dilution of 500 mM EDTA (see earlier).
19. *TLM.* Mix 3 mL of 10% Triton X100, 75 mL of 250 mM EDTA of pH 8, 15 mL of 1 M Tris–HCl of pH 8, and 7 mL water 20. Phenol: add 62.5 mL water to bottle with 250 g phenol, mix, warm as little as possible, add 300 mg 8-hydroxyquinoline, fill 20 mL aliquots in 30-mL Falcon tubes. Upon use, thaw and add 1 mL of 1 M Tris–base to one aliquot. Keep in fridge no longer than one month.
20. *RNaseA.* 250 μg/mL in 10 mM Tris–HCl, pH 7.5.
21. *Tris–acetate, 20 times concentrated.* Dissolve 96.8 g of Tris–base, 22.84 mL glacial acetic acid, 40 mL of EDTA (500 mM, pH 8), add water to a final volume of 1 L.
22. *Agarose gel, 2%, about 5-mm thick, 7-cm long, and 10-cm wide.* Mix 0.6 g agarose (electrophoresis grade), 1.5 mL of 20 times concentrated Tris–acetate, and 30 mL of water in a 100-mL reagent bottle. Be sure that the cap of the bottle is loose before the agarose in the mixture is melted in a microwave oven. Swirl the mixture occasionally and dissolve the agarose completely. Adjust the volume with water to 30 mL and pour the gel. Insert a comb for about 3-mm wide slots and let the agarose solidify. Cover the gel with Tris–acetate

electrophoresis buffer containing 1 μg/mL ethidium bromide (CAUTION, ethidium bromide is mutagenic).

23. *Ethidium bromide solution*. 10 mg/mL; dissolve 1 g of ethidium bromide in 100 mL of water by stirring for several hours. Store in the dark at 4°C.
24. *Agarose-loading buffer*. Mix one volume of 50% glycerol, 1 volume of Tris–acetate electrophoresis buffer, and 1/10 volume of a 2.5% (w/v) bromophenol blue solution in TE.
25. DNA molecular weight markers in the range of 100–1,000 base pairs, e.g., *Hae*III-digest of FX174.
26. UV transilluminator.
27. Horizontal electrophoresis apparatus for agarose gel, approximate dimensions of 7 × 10 cm.

2.2. Detection of Plasmid Clones by Blue/White Color

LB-amp-XG plates: LB-amp plates (see **Subheading 2.1**, **item 14**) containing 0.05 mg/mL of X-gal and 5×10^{-5} M of IPTG. Stocks of 2% X-gal in dimethylsulfoxide and 0.1 M IPTG in water can be made and appropriate amounts can be spread onto LB-amp plates the day before use.

2.3. Purification of Plasmid DNA

1. *E. coli suspension buffer*. 50 mM Tris–HCl of pH 8, 100 mM NaCl, prepare by mixing 50 mL of 1 M Tris–HCl of pH 8, and 100 mL 1 M NaCl in a total volume of 1 L. Store at 4°C.
2. *Tris–glucose–EDTA*. 25 mM Tris–HCl of pH 8, 50 mM glucose, 10 mM EDTA, prepare freshly by mixing and dissolving 2.5 mL 1 M Tris–HCl of pH 8, 1 g glucose and 2 mL 500 mM EDTA of pH 8, in a total volume of 100 mL of water.
3. *NaOH-SDS*. 200 mM NaOH, 1% SDS, prepare by mixing 10 mL of 10% SDS and 2 mL of 10 N NaOH in a total volume of 100 mL of water.
4. *Sodium acetate*. 3 M, pH 4.8 (adjust pH with glacial acetic acid).
5. *2 M Ammonium acetate*. Dissolve 154.2 g Ammonium acetate in a total volume of 1 L of water. Store at 4°C.
6. Isopropanol.
7. Cheese cloth.
8. Quick-Seal polyallomer centrifuge tubes (e.g., Beckman No. 342413).
9. Cesium chloride.
10. *NaCl, 1 M*. Dissolve 58.44 g of NaCl in 1 L of water, autoclave, and store at room temperature.

11. *CsCl-mix*. Dissolve 122.1 g of cesium chloride (DNA grade) in 128 mL TE, and add 4.13 mL of ethidium bromide (10 mg/mL). The final volume is 165 mL.
12. *n-butanol*. Mix *n*-butanol with an equal volume of water and an amount of cesium chloride which leaves some undissolved. Use only the upper (*n*-butanol) phase.

2.4. Analysis of DNA–Protein Complexes

1. Vertical electrophoresis apparatus for polyacrylamide gel; approximate dimensions are 15 × 15 cm.
2. Acrylamide/Bisacrylamide (30/0.8, w/w) 30% aqueous solution.
3. *TBE (ten times concentrated buffer)*. Dissolve 108 g of Tris–base, 55 g of boric acid, and 9.3 g of EDTA (Ethylenediaminetetraacetic acid, disodium salt) in water. Fill up to a total volume of 1 L.
4. *8% polyacrylamide slab gel (about 15 × 15 cm, 1 mm thick)*. Mix 31 mL of water, 13.3 mL of 30% acrylamide/bisacrylamide, 5 mL of ten times concentrated TBE, 600 µL of 10% APS (dissolve 1 g of ammonium persulfate in 10 mL of water), and 60 µL of TEMED (*N*,*N*,*N*′,*N*′-Tetramethylethylenediamide). Pour solution between the glassplates of the assembled vertical electrophoresis apparatus. Insert a comb for about 1-cm wide slots. Let the gel polymerize for several hours, preferably overnight.
5. *Gal repressor protein-binding buffer (five times concentrated)*. 50 mM KCl, 50 mM Tris–HCl of pH 7.6, 50 mM $MgSO_4$, 0.5 mM EDTA, 0.5 mM DTT, 250 µg/mL BSA, 50% glycerol.
6. *Ethidium bromide stain*. Add 100 µL of ethidium bromide solution (10 mg/mL) to 1 L of TE.

2.5. E. coli Host Strains

For pBendBlue *(8)*, *Epicurian coli* SURE2 competent cells (Stratagene) are recommended. For the other pBend plasmids, any transformation proficient *E. coli* strain may be used. The Dam methylase present in most laboratory *E. coli* strains recognizes GATC between the *Nhe*I and *Cla*I sites. If this is of concern, a methylation-defective (dam-) strain should be used.

3. Method

3.1. Insertion of the Protein-Binding Site into pBend

1. Add 25 µL of pBend DNA (1 mg/mL), 10 µL of ten times concentrated *Hpa*I digestion buffer, 55 µL of water, and 10 µL (100 units) of *Hpa*I restriction endonuclease to a 1.5-mL

Eppendorf tube. Mix and incubate at 37°C for 2 h or overnight.

2. To precipitate the restriction enzyme, place the digest on ice, add 3 μL of 500 mM EDTA of pH 8, 100 μL of ice-cold TE, and 100 μL of ice-cold 7.5 M ammonium acetate. Keep the sample at 4°C for 10 min. Centrifuge in a table-top centrifuge for 10 min. Remove the supernatant (containing the DNA) and add it to an Eppendorf tube filled with 600 μL ice-cold ethanol. Incubate at –70°C for 20 min, and centrifuge in a table-top centrifuge for 10 min. Discard the supernatant, add 500 μL of ice-cold 80% ethanol, centrifuge for 5 min, and discard the supernatant. Dry the DNA pellet in a vacuum centrifuge and dissolve the sample in 250 μL of TE. Store the linearized pBend DNA at –20°C.
3. Synthesize two complementary oligonucleotides which, when annealed to each other, form the protein-binding site. The purification of the oligonucleotides is likely to be unnecessary if they are shorter than 30 nucleotides. Dissolve each oligonucleotide in autoclaved distilled water at a concentration of 200 μg/mL.
4. Add 10 μL of each oligonucleotide and 180 μL of oligonucleotide-annealing buffer to a 1.5-mL Eppendorf tube. Incubate the sample for 3 min in a 300-mL beaker with about 150 mL of boiling water. Place the beaker with the tubes in the coldroom at 4°C to allow for annealing of the oligonucleotides over a period of several hours. Store the DNA at –20°C.
5. *Mix in a 1.5-mL Eppendorf tube.* 2 μL of annealed oligonucleotides, 3 μL of five times concentrated ligation buffer, 1 μL of 5 mM ATP, 8.5 μL of water, and 0.5 μL of T_4 polynucleotide kinase. Incubate for 10 min at 37°C. Place the sample on ice and add 3 μL of five times concentrated ligation buffer, 1 μL of linearized vector DNA (from step 2), 1 μL of 5 mM ATP, 9 μL of water, and 1 μL of T_4 DNA ligase. Incubate at 15°C for several hours or overnight. The samples can be stored in refrigerator for several days and aliquots can be used for several transformations.
6. Transform competent *E. coli* cells according to the protocol provided by the vendor and streak on LB-amp plates. Incubate the plates at 37°C overnight or until the colonies appear.
7. For preparation of the plasmid DNA on a small scale, use sterile toothpicks to transfer individual colonies to 15-mL tubes containing 5 mL of LB with 200 μg/mL ampicillin; also, streak cells from each transformant onto an LB-amp

plate. Incubate this master plate at 37°C and shake the liquid cultures at 37°C overnight.

8. Pellet the cells by centrifugation for 15 min at about 700 × *g* at 4°C (e.g., at 3,000 rpm in a Sorvall RT6000B refrigerated centrifuge with an H1000B rotor). Decant the supernatant, add 200 µL of Tris–sucrose, and transfer to 1.5-mL Eppendorf tubes. Add 25µL of lysozyme solution. Mix and add 130 µL of 200 mM EDTA, pH 8, and 130 µL of TLM. Mix and place at 65°C until lysis occurs (which usually takes a few minutes). Vortex briefly and centrifuge for 15 min in a table-top centrifuge. Remove the pellet with a sterile tooth pick, and add half a volume of the prepared phenol and half a volume of chloroform. Vortex for 10 s, centrifuge for 10 min, and carefully remove about 200 µL of the aqueous (upper) phase while staying clear of the interface. Add 400 µL of ice-cold ethanol, mix and centrifuge for 5 min, decant the supernatant, add 1 mL of 80% ethanol, centrifuge for 2 min, carefully decant the supernatant, and dry the pellet in a vacuum centrifuge. Dissolve the pellet in 30 µL of TE with occasional mixing. Store the samples at –20°C.
9. To verify successful insertion of the protein-binding site, digest an aliquot of the DNA with *Eco*RI and *Hin*dIII. To a 5-µL aliquot of the plasmid preparation add 2 µL of water, 1 µL of ten times concentrated *Eco*RI digestion buffer, 1 µL of *Eco*RI, 1 µL of *Hin*dIII, and 1 µL of RNase. As a control, digest 1 µg of pBend DNA. Incubate all samples at 37°C for several hours or overnight. Place digests on ice, add 1 µL of 200 mM EDTA of pH 8, 90 µL of ice-cold TE, and 50 µL of ice-cold 7.5 M ammonium acetate. Keep on ice for 10 min. Centrifuge in table-top centrifuge for 10 min. Collect the supernatant and add it to an Eppendorf tube containing 300 µL ice-cold ethanol. Mix and incubate at –70°C for 20 min. Centrifuge in table-top centrifuge for 10 min. Remove supernatant, add 300 µL ice-cold 80% ethanol to the pellet, centrifuge for 5 min, and discard supernatant. Carefully dry the pelleted DNA in a vacuum centrifuge and dissolve it in 5 µL of TE. Add 5 µL of Tris–acetate-loading buffer and mix briefly.
10. Prepare a 2% agarose gel. Load the samples from **step 9** in parallel with DNA molecular weight markers. Electrophorese at 80 V until the bromophenol blue has migrated about 4 cm. Examine the DNA under a UV transilluminator and take a picture with a Polaroid camera (film type 57 or 55). Successful insertion is indicated by an *Eco*RI–*Hin*dIII fragment of the expected mobility (242 base pairs plus insert). Electrophoresis can be continued to discover minor mobility differences, but then fresh electrophoresis buffer should be used.

Eventually, the nature of the positive clone must be verified by DNA sequencing, which also reveals the orientation of the inserted binding site (*see* **Note 2** for the selection of suitable sequencing primers).

3.2. Detection of the Protein-Binding Site Cloned into pBend Blue by Blue/White Color Screening

After ligation of the DNA segment corresponding to the protein-binding site into pBendBlue DNA *(8)*, as described in **steps 1–5** of **Subheading 3.1**., transform competent *E. coli* cells (*Epicurian* strain) and plate on LB-amp-XG plates. Incubate the plates at 37°C overnight or until the colonies are large enough to distinguish their blue/white color phenotype. By this procedure, usually 1–10% of the transformed colonies on LB-amp-XG plates are white. Verify the white colonies by purifying on LB-amp-XG plates. Sequence verification shows that almost all the white colonies contain the desired insert. Thus, when the cloning efficiency is poor, the color screening allows successful use of pBendBlue in cloning short DNA sequences for studying DNA binding.

3.3. Purification of pBend DNA

1. To obtain pure DNA of the positive pBend derivative, set up a 5-mL culture of the positive clone in LB with 500 μg/mL of ampicillin starting from an individual colony of the master plate (**step 7** of **Subheading 3.1**). Shake at 37°C for several hours until the culture becomes turbid. Transfer the cells to a 2-L sterile Erlenmeyer containing 400 mL of LB with ampicillin. Shake overnight at 37°C.
2. Place the culture on ice and transfer the cells into centrifuge bottles. Pellet the cells by centrifugation at 4°C at about 1,600 × *g* (e.g., 3,500 rpm in a H6000A rotor of a Sorvall RC3C centrifuge). Decant the supernatant, resuspend the pellet in 20 mL *E. coli* suspension buffer, and transfer the cells to a 50-mL centrifuge tube (preferably Nalgene, Cat. No. 3131-0024). Centrifuge at 4°C for 10 min at about 5,000 × *g* (e.g., in Sorvall SS34-rotor at 10,000 rpm.). Freeze the pellet completely by placing the sample on dry ice or in a –80°C freezer.
3. Thaw the pellet and resuspend the cells in 6 mL of Tris–glucose–EDTA. Add 12 mg of lysozyme powder, mix, and keep on ice for 30 min. Bring to room temperature and add 12 mL of NaOH-SDS. Mix and place on ice for 5 min.
4. Add 9 mL of 3 M sodium-acetate of pH 4.8; shake and leave on ice for 30 min. Centrifuge at about 12,000 × *g* at 4°C for 15 min (e.g., in an SS34 rotor at 15,000 rpm).
5. Transfer the supernatant to a new centrifuge tube by filtering through a cheesecloth. Add half a volume of isopropanol to the transferred solution, leave 5 min at room temperature, and centrifuge at 4°C for 10 min at 5,000 × *g* (e.g., in the SS34 rotor at 10,000 rpm).

6. Discard the supernatant and add 6 mL of ice-cold 2 M ammonium-acetate to the pellet. Vortex repeatedly to dissolve the plasmid DNA until only small particles are visible. Centrifuge at 4°C for 10 min at 5,000 × *g*.
7. Transfer the supernatant (containing the plasmid DNA) to a new centrifuge tube, add 4 mL of isopropanol, mix and centrifuge at 4°C for 10 min at 5,000 × *g*.
8. Discard the supernatant and completely dissolve the pellet in 7 mL of TE with occasional shaking. Add 8 g of cesium chloride, 400 µL of 1 M NaCl, 400 µL of 1 M Tris–HCl of pH 8, 160 µL of ethidium bromide (10 mg/mL), and 2.5 mL of CsCl mix.
9. After the CsCl is dissolved, draw the solution into a 20-mL syringe and transfer it into a Quick-Seal centrifuge tube. Fill a second tube with CsClmix and make sure that the two tubes are balanced. Seal the tubes and centrifuge overnight at about 200,000 × *g* at 20°C (e.g., at 50,000 rpm in a Beckman NTV65 rotor).
10. Remove the tubes from the rotor, puncture the top of the tube, then collect the lower of the two visible bands with a syringe by puncturing the side of the tube. Transfer the DNA into a 15-mL Corex glass centrifuge tube.
11. Extract the ethidium bromide by adding 1 mL of *n*-butanol which has been saturated with water and cesium chloride. Vortex and remove the upper phase with a glass pipette. Repeat this process beyond the point where the color becomes invisible (usually about six times).
12. Add 2.5 mL of water and 7 mL of ethanol. Mix and incubate at –70°C for 15 min. (Do not leave too long; otherwise, CsCl will precipitate). Centrifuge for 15 min at 4°C at about 7,000 × *g* preferably in a swinging-bucket rotor (e.g., Sorvall HB4 rotor at 10,000 rpm).
13. Pour off the supernatant, add 5 mL of ice-cold 80% ethanol to the pellet, repeat the centrifugation, discard the supernatant, and evaporate excess ethanol under vacuum.
14. Dissolve the DNA in 500 µL of TE and determine the absorbance at 260 nm. Add the appropriate amount of TE to adjust the concentration of the plasmid DNA to 1 mg/mL (1 A_{260} is equivalent to 50 µg/mL). Store the DNA at 4°C.

3.4. Analysis of DNA–Protein Complexes

1. To generate restriction fragments with the protein-binding site located at the end or in the middle, digest the pBend construct separately with *Mlu*I (end) and *Eco*RV (middle). One digestion contains 100 µL (100 µg) of plasmid DNA from

step 14 of **Subheading 3.3**, 30 μL of ten times concentrated *Mlu*I- (or *Eco*RV) digestion buffer, 160 μL of water, and 10 μL of *Mlu*I or *Eco*RV restriction enzyme (100 U). Incubate the samples at 37°C for several hours or overnight. Place the samples on ice, add 15 μL of 500 mM EDTA of pH 8, 150 μL of ice-cold 7.5 M ammonium acetate and leave at 4°C for 10 min. Centrifuge in a table-top centrifuge for 10 min. Remove and add the supernatant (containing the DNA) to a new Eppendorf tube filled with 900 μL of ice-cold ethanol. Mix and incubate at –70°C for 20 min. Centrifuge in a table-top centrifuge for 10 min. Discard the supernatant, add 500 μL of ice-cold 80% ethanol, centrifuge for 5 min., and decant the supernatant. Carefully dry the pellet in a vacuum centrifuge and dissolve the DNA in 50 μL of water. Verify the success of the digestion by electrophoresis of an aliquot on a 2% agarose gel (described in **item 23** of **Subheading 2.1**).

2. Pour a vertical 8% polyacrylamide slab gel. Assemble the electrophoresis apparatus and pre-electrophorese at room temperature for 1 h at 100 V with TBE buffer in the reservoirs.
3. Isolate the different DNA fragments containing the protein-binding site after gel electrophoresis for end labeling, if necessary (*see* **Note 6**). Label the ends of the DNA fragments by T_4 polynucleotide kinase and [γ-^{32}P]ATP as recommended by the supplier of the enzyme.
4. *Mix at room temperature.* 2 μL of *Mlu*I- or *Eco*RV-digested DNA (from **steps 1** or **3**), 1.6 μL of five times concentrated Gal repressor protein-binding buffer, 4.4 μL of water, 2 μL of diluted protein. Keep the time between diluting the protein and addition to the DNA as short as possible. Do not vortex; mix gently with the tip of the pipette. Prepare a control without added protein. Incubate all samples for 10 min at room temperature.
5. Flush the wells of the polyacrylamide gel with reservoir buffer and load samples without the addition of loading buffer and tracking dyes. The glycerol in the binding buffer gives the sample sufficient density. A long plastic microcapillary tip is helpful to deliver the sample to the bottom of the slot. Glass capillaries should be avoided because proteins tend to stick to glass. In a separate slot, load DNA molecular weight markers with bromophenol blue. Electrophorese at room temperature for 5 h at 200 V. Separate the glass plates and immerse the gel in ethidium bromide stain for visualization of the DNA under UV light. Take a picture and subsequently autoradiograph the gel (consult **Notes 3** and **4** if complexes cannot be detected).

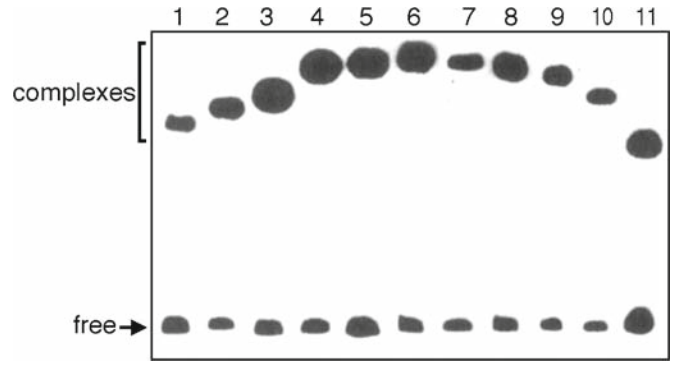


Fig. 4. Gel electrophoresis of permuted fragments of *lac* CRP sites. CRP protein was mixed separately with 11 different 5′-end ^{32}P-labeled DNA fragments. In a volume of 20 μL, a sample of each fragment was mixed with 10 mM Tris–HCl of pH 7.5, 1 mM EDTA, 50 mM KCl, 20 μM cAMP, 50 μg/mL BSA, 10% glycerol, and 1 nM CRP. Polyacrylamide concentration was 10%. The protein and DNA concentrations were such that approximately 50% of the DNA was engaged in complexes with CRP. The DNA fragments used were generated by restriction enzymes, which, from left to right, are *Mlu*I, *Bgl*I, *Nhe*I, *Spe*I, *Xho*I, *Eco*RV, *Pvu*II, *Stu*I, *Nru*I, *Kpn*I, and *BamH*I. As indicated, the fragments near the bottom of the gel are free DNA and those at the upper part are bound to the cAMP–CRP complex.

6. Measure the distances between the slot and the position of the protein–DNA complex of the *Mlu*I- (μ_E) and the *Eco*RV-digest (μ_M). Also, examine the mobilities of the free DNA (*see* **Fig. 4**) to make sure that the DNA fragments contain no intrinsic bending. Calculate the bending angle a using the empirical formula $\mu_M/\mu_E = \cos(\alpha/2)$. *See* **Note 7** for details.

4. Notes

1. Protein-binding sites can be inserted into the pBend vectors using restriction fragments or synthetic oligonucleotides. Restriction fragments should not be considerably larger than the protein-binding site to be tested. Oligonucleotides are normally available with blunt ends; therefore, the HpaI sites of pBend4 or pBend5 (*see* **Fig. 2**) should be used. Newly synthesized oligonucleotides might be designed with "sticky" ends such that they are compatible with the XbaI and the SalI site; they can be cloned efficiently and inserted into the DNA in a single orientation.
2. Insertion of the protein-binding site may not occur if the oligonucleotides are of poor quality. In this case, they should be

purified and checked by polyacrylamide gel electrophoresis. If the transformation efficiency with supercoil control DNA is high, yet very few transformants are obtained with the annealed oligonucleotides, reduce the amount of insert DNA. Multiple insertion of the binding site can occur and is detected by gel electrophoresis and sequencing. For sequencing use primers named T_3, T_7, M13–20, or reverse primer (Stratagene). Do not use the SK and KS primers (Stratagene) because they are not fully complementary to pBend3, 4, and 5.

3. Bending experiments with radioactively labeled DNA are particularly useful if the protein has not been purified or if its availability is limited. Labeling can be accomplished with T_4 polynucleotide kinase and [γ-^{32}P]ATP. Often, the DNA ends generated by the various restriction enzymes are labeled to different degrees. This problem can be overcome by loading the gel with aliquots of the binding reaction adjusted for the efficiency of fragment labeling. It is best to purify and isolate the protein-binding fragments because the radioactively labeled plasmid DNA might obscure the region where the complexes are located. Another potential problem (which is also the case with unlabeled DNA) might be that bands appear which represent minor digestion products. In order to identify those, make sure to include controls without added protein.
4. One of the frustrating aspects of conducting a bending experiment can be the inability to detect a complex on the polyacrylamide gel. Even if the binding and electrophoresis conditions are known one should be careful to avoid solutions and equipment which has been in contact with SDS. If possible, dedicate one electrophoresis setup to "gel-shift"experiments. Many DNA-binding proteins are insoluble in the low salt concentration of the electrophoresis buffer and must be stored at high ionic strength. Limit the time between dilution and addition to the DNA. Avoid vortexing during complex formation and do not add tracking dyes because they interact with the complex and might change its mobility. Larger protein–DNA complexes (e.g., Lac repressor, *(6)*) behave better in low percentage polyacrylamide gels (e.g., 4%) with a high acrylamide/bisacrylamide ratio (80:1). The protein concentration for obtaining about equal amounts of free and complexed DNA should always be determined in a preliminary experiment.
5. If many transformants are obtained, but none contains the protein-binding site, the pBend DNA might not be fully linearized. Alter the DNA-enzyme ratio in favor of the enzyme and confirm complete digestion of an aliquot by electrophoresis on an agarose gel.

6. In the initial bending experiment, it is advisable to restrict only with *Mlu*I or *Bam*HI (to place the binding site close to the ends) and *Eco*RV or *Pvu*II (to place the binding site in the middle). Do not select restriction sites which also occur in the protein-binding sequence. When exploiting the 17 circular permutated restriction sites attention must be paid to the property of some of the restriction enzymes as follows: *Cla*I sites are methylated in most *E. coli* strains; its use is therefore limited to prior growth of the plasmid in a methylation-defective (dam-) host. *Sty*I will also cut at *Nco*I of the repeat; for the purpose of a bending experiment, it can only be used under partial digestion conditions. An additional *Spe*I site is present in the vicinity of the single *Eco*RI site as part of pBluescript SK- (*see* **Fig. 2**). *Spe*I digestion generates an additional small fragment, which contains no protein-binding site and does not interfere with the bending assay. Three additional *Dra*I sites are located in the plasmid corresponding to pBluescript coordinates 1,912; 1,931; and 2,623. Depending on their electrophoretic property, some of the vector-derived DNA fragments might comigrate with certain protein–DNA complexes. Make sure to include a control without added protein. Likewise, two additional *Pvu*II sites correspond to pBluescript coordinates 529 and 977. A *Sma*I site is present close to *Eco*RI (*see* **Fig. 2**). Two additional *Ssp*I sites correspond to pBluescript coordinates 442 and 2,850; and two additional *Rsa*I sites correspond to pBluescript coordinates 665 and 2,526. *Nco*I will also cut at *Sty*I of the repeat and can only be used under partial digestion conditions. An additional *Bam*HI site is present close to the *Eco*RI site (*see* **Fig. 2**).
7. The bending angle α assumes a value of 0° for a straight duplex. Since the mobility of a rigid DNA fragment is related to its end-to-end distance, the latter equals $L\cos(\alpha/2)$, with L being the length of the unbent DNA. The end-to-end distance of a fragment bent at the end will be virtually the same as L. Thus $\mu_M/\mu_E = L\cos(\alpha/2)/L = \cos(\alpha/2)$, where μ_M is the mobility of the complex with the protein bound centrally and μ_E the mobility of the complex with the protein bound at the end of the DNA fragment. The apparent bending angle for two *lac* promoters induced by CRP is 96° (*see* **Fig. 4** of **ref. 2**). We measure the distance between the top of the gel and the front of the band representing the protein–DNA complex. Whatever method is used, one must be consistent. Possible intrinsic bending in the free DNA must be considered in the calculation of the bending angle. It should be noted that the calculated values may be different from absolute bending angles, since factors other than the

end-to-end distance influence the mobility of protein-bound and unbound DNA fragments. The method measures the net bend and cannot distinguish between a single sharp bend at one position and a smooth curving over a larger DNA region. For precise determination of bending angles, at least three independent experiments should be carried out. If possible, control lanes should be added of a similar size complex in which the DNA is bent to a known degree.

References

1. Crothers D.M. and Fried M.G. (1983). Transmission of long-range effects in DNA. *Cold Spring Harbor Symp. Quant. Biol.* **47**, 263–269
2. Zwieb C., Kim J. and Adhya S. (1989). DNA bending by negative regulatory proteins: Gal and Lac repressors. *Genes and Dev.* **3**, 606–611
3. Wu H.-M. and Crothers D.M. (1986). The locus of sequence-directed and protein-induced DNA bending. *Nature* **308**, 509–513
4. Kim J., Zwieb C., Wu C. and Adhya S. (1989). Bending of DNA by gene regulatory proteins: construction and use of a DNA bending vector. *Gene* **85**, 15–23
5. Thompson J.F. and Landy A. (1988). Empirical estimation of protein-induced DNA bending angles: application to site-specific recombination complexes. *Nucleic Acids Res.* **20**, 9687–9705
6. Fried M.G. and Crothers D.M. (1983). CAP and RNA polymerase interaction with the *lac* promoter: Binding stoichiometry and long-range effects. *Nucleic Acids Res.* **11**, 141–185
7. Zwieb C. and Brown R.S. (1990). Absence of Substantial Bending in the *Xenopus laevis* Transcription Factor IIIA-DNA Complex. *Nucleic Acids Res.* **18**, 583–587
8. Sperbeck S.J. and Wistow G.J.(1998). pBend-Blue: modification of the pBend system for color selectability. *BioTechniques.* **24**, 66–68

Chapter 33

Analysis of Distant Communication on Defined Chromatin Templates In Vitro

Yury S. Polikanov and Vasily M. Studitsky

Summary

Regulation of many biological processes in eukaryotes involves distant communication between the regulatory DNA sequences (e.g., enhancers) and their targets over the DNA regions organized in chromatin. However previously developed methods for analysis of communication in chromatin in vitro are artifact-prone and/or do not allow analysis of communication on physiologically relevant, saturated arrays of nucleosomes. Here we describe a method for quantitative analysis of the rate of distant communication in *cis* on saturated arrays of nucleosomes capable of forming the 30-nm chromatin fibers in vitro.

Key words: Chromatin, Nucleosome, Enhancer, Promoter, Transcription.

1. Introduction

Many DNA transactions in eukaryotic nuclei involve protein-mediated interactions between two or more DNA sites widely separated along DNA that is organized into chromatin. Such processes most often are accompanied by direct interaction between proteins bound at the enhancer and the target promoter with accompanying formation of large chromatin loops that includes the intervening chromatin in vivo *(1–6)*. Therefore efficient enhancer action over a distance critically depends on structural and dynamic communication properties of chromatin that are largely unknown.

Several experimental approaches for analysis of distant enhancer–promoter communication have been developed recently.

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_33

The FLP DNA recombination assay was employed for measuring communication over 74-bp to 15-kb distances in chromatin in vivo *(7)*. However applicability of this method for studies of communication in chromatin in vitro has not been evaluated. Furthermore, communication between the DNA sequences required for recombination and transcriptional regulation over a distance may occur by different mechanisms. Thus, recombination can occur between DNA sequences positioned within different, sometimes widely spaced domains of chromatin while enhancer–promoter communication is largely limited by a single chromatin loop *(8)*.

DNA ligation-circularization assay was used as an alternative method for analysis of intramolecular communication over a distance on DNA *(9–11)* and in chromatin in vitro *(12)*. However extensive internucleosomal interactions strongly complicate interpretation of the experiments (Y.S.P., data not shown).

Recently we have developed an experimental assay allowing quantitative analysis of the rate of distant communication between a bacterial transcription regulatory element (enhancer) and its target (promoter) in chromatin *(13–16)*. The use of the bacterial experimental system is dictated by the low efficiency of eukaryotic RNA polymerase II-dependent in vitro transcription systems *(17)* and inconsistency of eukaryotic enhancer action over a distance in vitro *(18, 19)*. At the same time, bacterial transcriptional enhancers can work efficiently over a large distance (up to at least 5 kb) both in vivo and in vitro *(20–22)*. Moreover, pro- and eukaryotic transcriptional enhancers share many key properties, such as the looping mechanism of enhancer–promoter communication *(23)*. The mechanism of action of bacterial transcriptional enhancers has been extensively studied using the *glnAp2* promoter of *Escherichia coli* as a model *(24, 25)*. Transcriptional activity of *glnAp2* promoter is controlled by the NtrC-dependent, σ^{54}-dependent transcriptional enhancer *(22, 26, 27)*. The enhancer is activated by the NtrC protein, which is phosphorylated by the NtrB protein kinase *(28, 29)*. When phosphorylated, enhancer-bound NtrC interacts with the $E\sigma^{54}$ holoenzyme and stimulates conversion of the closed (RP_{closed}) to the open (RP_{open}) initiation complex *(22, 26, 30–33)*. During this direct enhancer–promoter interaction, intervening DNA is transiently looped out *(34, 35)*.

This experimental system is relatively simple, highly efficient, and very well studied. Transcription is strongly (>100-fold) stimulated by the enhancer, and the mechanism of communication can be analyzed both in vitro and in vivo. Activity of the promoter itself in this system does not depend on the level of negative DNA supercoiling (Y.S.P., unpublished data) allowing analysis of communication properties of linear, relaxed or supercoiled DNA and chromatin templates. Using this experimental technique, the mechanisms of distant communication on histone-free DNA *(20, 21, 36)* and in chromatin *(15, 16)* have been studied.

In our previous studies of communication in chromatin subsaturated arrays of randomly positioned nucleosomes were utilized *(15, 16)*. However nucleosomes under physiologically relevant conditions are organized in regularly spaced, saturated arrays of nucleosomes forming higher-order chromatin structure (the 30-nm fibers *(37, 38)*). Here we describe an experimental approach that allows analysis of communication properties of structurally defined, saturated arrays of precisely positioned nucleosomes capable of formation of the 30-nm chromatin fiber.

2. Materials

2.1. Buffers and Reagents

1. *TAE buffer (1×)*. 40 mM Tris–HCl (pH 7.6), 31.2 mM Acetic Acid, 1 mM EDTA.
2. *TBE buffer (0.5×)*. 44.5 mM Tris–HCl (pH 8.3), 44.5 mM Boric Acid, 1 mM EDTA.
3. *Chromatin Reconstitution Buffer A*. 1 M NaCl, 10 mM Tris–HCl (pH 7.5), 0.2 mM EDTA, 0.1% NP40, and 5 mM β-mercaptoethanol.
4. *Chromatin Reconstitution Buffer B*. 10 mM Tris–HCl (pH 7.5), 0.2 mM EDTA, 0.1% NP40, and 5 mM β-mercaptoethanol.
5. *NEBuffer 2 (1×, New England Biolabs)*. 10 mM Tris–HCl (pH 7.9), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT.
6. *ThermoPol Buffer (1×, New Englad Biolabs)*. 20 mM Tris–HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton-X100.
7. *Transcription buffer (1×)*. 50 mM Tris-Ac (pH 8.0), 100 mM KAc, 8 mM $Mg(CH_3COO)_2$, 27 mM $NH_4(CH_3COO)$, 0.7% PEG-8000, and 0.2 mM DTT.

2.2. Native Electrophoresis of Reconstituted Chromatin

1. *Native agarose gel*. 1.2% agarose, 20 mM HEPES-Na (pH 8.0), 0.2 mM EDTA, 5% glycerol.
2. *Running buffer*. 20 mM HEPES-Na (pH 8.0), 0.2 mM EDTA.
3. Electrophoresis was carried out in a vertical apparatus between two glass plates using 1mm spacers at 120 V for approximately 1.5 h until the bromophenol blue front reached the bottom of the gel.

2.3. Denaturing PAGE of Purified DNA and RNA

1. 8% denaturing polyacrylamide gel (19:1) containing 0.5× TBE buffer and 8 M urea was prepared.
2. *Running buffer*. 0.5× TBE buffer.

3. Electrophoresis was carried out in a vertical apparatus between two 20 × 30 cm glass plates using 0.4-mm spacers at 2,000 V and no more than 50 Watts for approximately 1 h until the bromophenol blue front reached the bottom of the gel.

3. Methods

3.1. DNA Template for Analysis of Distant Communication on Defined Arrays of Nucleosomes

Assembly of saturated arrays is impossible using the experimental techniques developed earlier because randomly positioned nucleosomes can assemble on the enhancer and promoter and completely block these DNA elements from binding of the proteins when the level of chromatin assembly approaches saturation *(15, 16)*. To prevent nucleosome assembly on the enhancer and promoter and to form saturated arrays of precisely positioned nucleosomes on the spacer DNA we have employed a high-affinity histone-binding, nucleosome positioning (601) sequence *(13, 14)*. The 601 sequence supports formation of precisely positioned arrays of nucleosomes that can form the 30-nm chromatin fibers in the absence of linker histones H1/H5 *(38, 39)*. Furthermore, if reconstitution of the arrays is conducted in the presence of an excess of competitor DNA and a limited supply of histone octamers, nucleosomes can form on the 601 sequences with a high preference *(40)* leaving the enhancer and promoter histone-free and accessible to corresponding DNA-binding proteins.

1. The pYP05 plasmid (**Fig. 1a**, *see* **Note 1**) used as a template for chromatin reconstitution of an in vitro transcription was purified using the QIAfilter Plasmid Maxi Kit (Qiagen).
2. The overall experimental approach for the transcriptional analysis of distant enhancer–promoter communication in chromatin is outlined in **Fig. 1b**.

3.2. Chromatin Assembly on Supercoiled DNA

1. Reconstitution of chromatin on negatively supercoiled pYP05 plasmid was conducted by a modified transfer method using continuous dialysis from 1 M NaCl *(41, 42)* at 10 nM (50 μg/mL) DNA concentration and 12.5 μg/mL, 25 μg/mL, or 50 μg/mL of donor chromatin from chicken erythrocytes (1:4, 1:2, or 1:1 chromatin/DNA mass ratios, respectively) to achieve subsaturated, saturated, and oversaturated levels of chromatin assembly, respectively.
2. DNA and donor chromatin were mixed in reconstitution buffer A in a total volume of 120 μL, transferred into a small dialysis bag (Spectra/Por, MWCO 8,000), and placed into a glass bottle containing 100 mL of reconstitution buffer A.

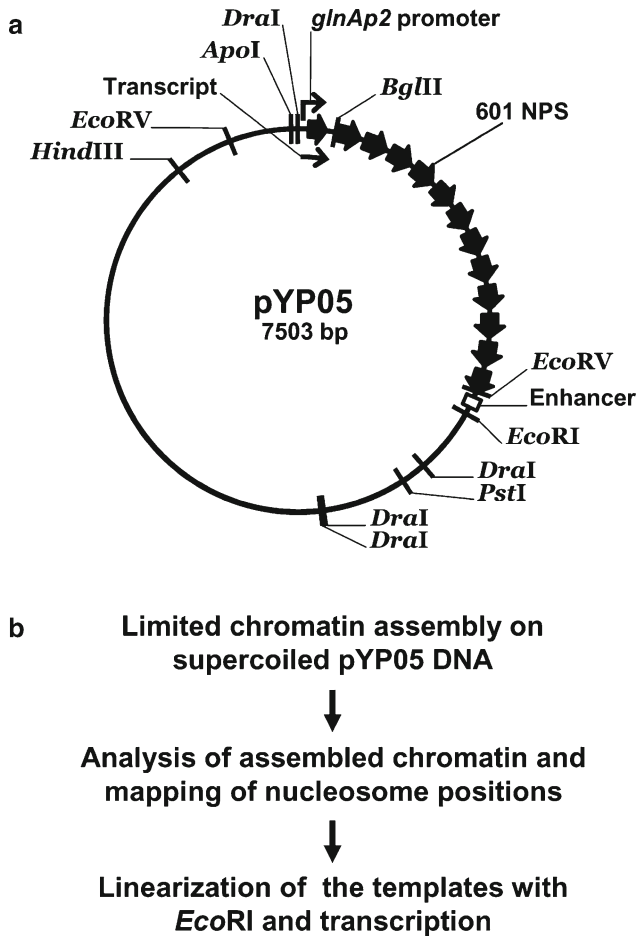


Fig. 1. An experimental approach for the transcriptional analysis of distant enhancer–promoter communication in chromatin. (**a**) Restriction map of the pYP05 template. Positions of the enhancer, promoter, and 13 high-affinity nucleosome positioning 601 sequences (601 NPS) are indicated. 35 bp of the promoter-proximal NPS is truncated from its promoter-distal end and is replaced by bacteriophage T_7 transcription terminator site, limiting the length of the transcript to 176 nucleotides. The length of the internucleosomal spacer DNA between each two NPSs is 29 bp. (**b**) The experimental approach.

3. The bottle was then tightly sealed and connected to the gradient maker; the inner beaker was filled with 500 mL of buffer A and the outer beaker contained 500 mL of buffer B.
4. Dialysis was performed at 4°C in cold room in a gradient maker and on a magnetic stirrer overnight at a flow rate of 1 mL/min.
5. The vast majority (>90%) of the nucleosomes on the saturated and oversaturated reconstituted chromatin samples are positioned on the NPSs (*see* **Note 2**).

3.3. Characterization of the Reconstituted Chromatin in a Native Gel

In a native gel DNA-protein complexes remain intact. The mobility in the gel is determined by the charge, mass, and shape of the complexes.

1. For this experiment 5′-radioactively labeled *Apo*I/*Eco*RI-restriction fragment of pYP05 plasmid (**Fig. 1a**) was used. It comprises the promoter and enhancer separated by 13 NPSs.
2. After nucleosome assembly (*see* **Note 3**) chromatin samples were analyzed by native agarose gel electrophoresis.
3. After the electrophoresis the gel was transferred to Whatman 3-MM paper, covered with polyethylene wrap, and dried for 20 min at 50°C and then for 20 min at 80°C.
4. The dried gel was exposed overnight to PhosphorScreen (Perkin Elmer) and the screen was scanned on Cyclone PhosphorImager (Perkin Elmer, *see* **Fig. 2**).

3.4. Characterization of Reconstituted Chromatin using a Restriction Enzyme Sensitivity Assay

This is a simple and fast assay to evaluate DNA occupancy by nucleosomes. The assay is based on the observation that chromatin assembly results in strong protection of nucleosome-covered DNA from digestion with restriction enzymes *(43, 44)*. The cutting sites should be chosen to generate DNA fragments of different lengths to facilitate their separation in the agarose gel and analysis of the gel. In the case presented later the sites

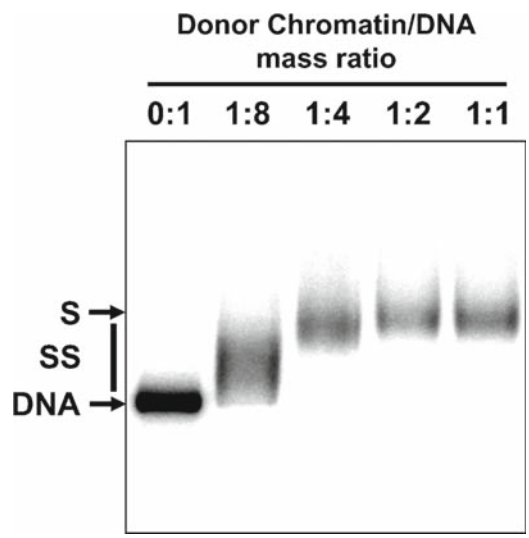


Fig. 2. Analysis of chromatin templates by native PAGE. Chromatin was assembled on the radioactively labeled 2.5 kb *Apo*I/*Eco*RI pYP05 fragment **(Fig. 1a)** at the indicated mass ratios of donor chromatin to template DNA. Chromatin assembly results in a progressive decrease of the mobility of the complexes in the gel. Chromatin reconstituted at 1:2 ratio of the donor-chromatin to DNA is saturated (S), since further increase in the ratio up to 1:1 does not result in further shift of the corresponding band. Saturated chromatin assembled at 1:2 ratio was used in the transcription experiment **(Fig. 5)**.

were chosen in the NPS-free region of the plasmid to evaluate the extent of undesired nucleosome assembly. However a similar technique could be applied to any DNA region on the plasmid *(15, 16)*.

1. 750 ng of DNA or nucleosomal templates (20 μg/mL) was incubated in the presence of an excess of *Dra*I and *Bgl*II restriction endonucleases (ten units each) in the NEBuffer 2 at 37°C for 2 h.
2. DNA was purified by phenol/chloroform extraction followed by ethanol precipitation and analyzed by electrophoresis in 1% agarose/TAE gel. The intensities of the bands in the gel were quantified using the OptiQuant software (Perkin Elmer, *see* **Note 4**).
3. The intensities of the bands corresponding to the final products of digestion are decreased as the efficiency of chromatin assembly is increased (**Fig. 3b**). The extent of chromatin assembly is directly proportional to the decrease in the intensity of the bands. However even on saturated chromatin (1:2, **Fig. 3**) the analyzed restriction sites remain largely accessible suggesting (in combination with the data shown in **Fig. 2**) that nucleosomes are poorly formed on plasmid DNA regions that do not contain NPSs.

3.5. Characterization of Reconstituted Chromatin using a Restriction Enzyme Sensitivity Assay with Primer Extension

This is a more advanced and involved version of the restriction enzyme sensitivity assay (*see* **Subheading 3.4**) allowing more straightforward, quantitative, and high-resolution analysis of nucleosome occupancy and positioning on polynucleosomal templates. Overall experimental approach is illustrated in **Fig. 4a**.

1. Chromatin samples were digested with an excess of restriction enzymes *Alu*I, *Msp*I, or *Sca*I (**Fig. 4b**), followed by phenol/chloroform extraction and ethanol precipitation of DNA.
2. Purified DNA fragments obtained after *Alu*I, *Msp*I, or *Sca*I digestion were further digested with *Eco*RI restriction enzyme (**Fig. 1b**) to set one of the end; the second end is set by the primer (**Fig. 4b**).
3. *Eco*RI-digested DNA was subjected to primer annealing **(Fig. 4b, see Note 5**). The annealed primer was extended with *Taq* DNA polymerase (New England Biolabs) in 1× ThermoPol buffer and conditions recommended by manufacturer.
4. The products of extension were purified by phenol/chloroform extraction, ethanol precipitation, and analyzed by a denaturing PAGE.
5. The distribution of the bands in the gel (**Fig. 4c**) reflects the sensitivity of each restriction site to corresponding restriction enzyme in chromatin and allows quantitative analysis of nucleosome occupancy (*see* **Note 6**).

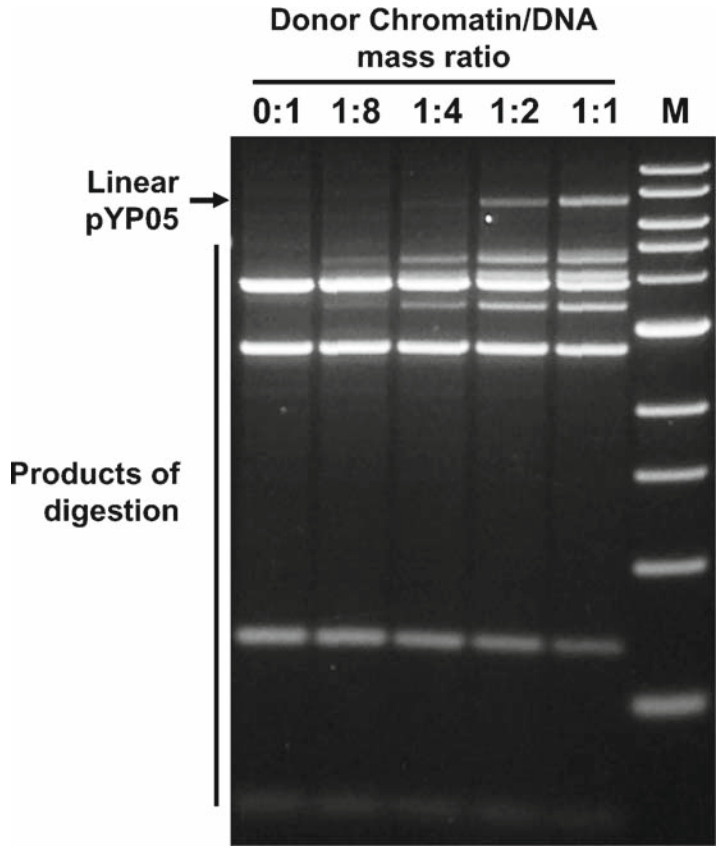


Fig. 3. Characterization of chromatin templates using restriction enzyme sensitivity assay. Chromatin was assembled on supercoiled pYP05 plasmid and digested with an excess of restriction enzymes *Dra*I and *Bgl*II. Then DNA was purified and analyzed in 1% agarose gel. Nucleosomes protect DNA from digestion with restriction enzymes. The restriction sites are localized beyond the NPSs **(Fig. 1b)**; therefore, the increase of donor-chromatin/DNA ratio results in progressively better protection of the template DNA from the enzymes, but this protection is still minimal even at highest ratio of donor chromatin to DNA (1:1), indicating that the majority of nucleosomes were formed predominantly on the desired NPSs. Only a small fraction (<5%) of all nucleosomes formed on the plasmid occupy the NPS-free regions of the pYP05 plasmid. M – 1-kb DNA ladder (New England Biolabs).

3.6. Analysis of Communication in Chromatin Using a Single-Round Transcriptional Assay In Vitro

In vitro transcription assay was optimized for maximal utilization of the chromatin templates and was used previously to accurately measure the rate of enhancer–promoter (E–P) communication (**Fig. 5a** *(16, 24)*). Physical E–P interaction is the rate-limiting step during transcription *(16, 21)*. E–P communication and direct interaction of corresponding DNA-bound proteins are accompanied by looping of the intervening DNA *(34, 35)* and are required for conversion of the closed (RP_{closed}) to the

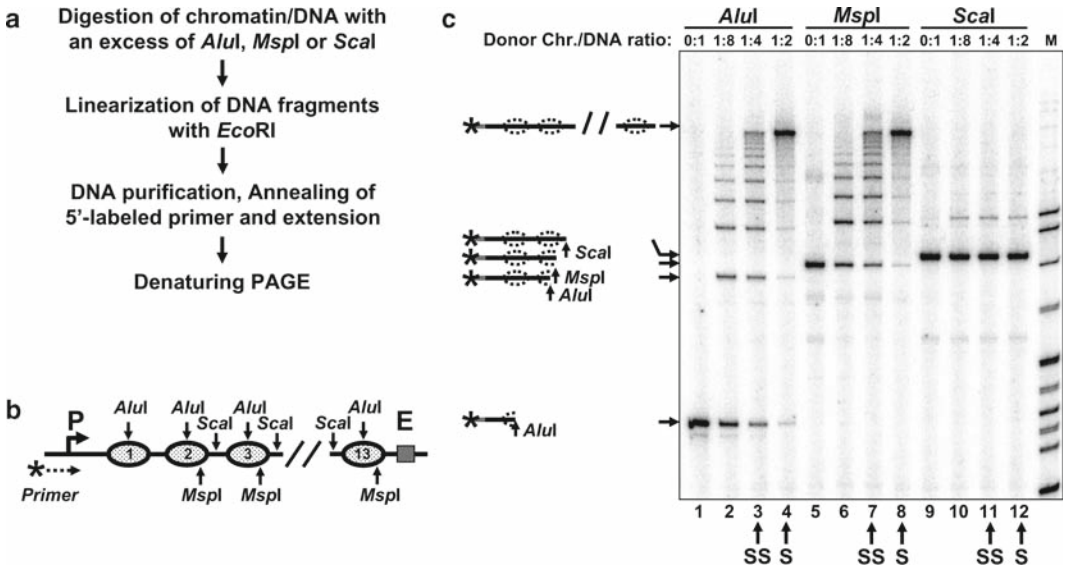


Fig. 4. Analysis of chromatin templates using a restriction digestion sensitivity assay with primer extension. (**a**) The experimental approach. (**b**) Schematic diagram of the fragment of pYP05 plasmid used for in vitro chromatin reconstitution. Nucleosome positioning sequences (NPS) are shown by solid ovals. Restriction enzyme sites within the NPSs are shown by arrows. *Alu*I and *Msp*I sites are localized within NPS, *Sca*I sites – in the middle of NPS-separating spacer DNA. *Msp*I and *Sca*I sites are missing from the first NPS and the spacer DNA, respectively. (**c**) Analysis of the products of primer extension by a denaturing PAGE. The increase in the level of chromatin assembly results in progressively better protection of the templates from *Alu*I and *Msp*I, but not from *Sca*I restriction enzyme. Possible products of digestion are shown on the left. S/SS: Saturated/subsaturated chromatin samples.

open initiation complex (RP_{open}) *(22, 28–32)*. The RP_{open} formed at the promoter is stable in the presence of heparin; however, its formation de novo is strongly inhibited *(45)*. Heparin also disrupts nucleosomes, so they are present only during initiation step, but not during elongation. Therefore when E–P communication is allowed for a limited time (1 min) the overall efficiency of transcription conducted in the presence of heparin can serve as a direct and quantitative measure of the rate of E–P communication (*see* **Note 7** *(24, 36)*).

1. All templates were linearized at the *Eco*RI site **(Fig. 1a)**, and the closed initiation complexes (RP_{closed}) were formed in 50-μL aliquots in 1× transcription buffer (TB) at 1 nM chromatin concentration and 10 nM core RNA polymerase, 300 nM σ^{54}, 120 nM NtrC, and 400 nM NtrB for 15 min at 37°C.
2. 5 μL of 40 mM ATP in 1× TB was added to the reaction volume to 4 mM final ATP concentration, and the reaction was incubated at 37°C for 1 more minute to form the open, elongation-competent initiation complex (RP_{open}).
3. To start elongation and to limit it to a single round (prevent further formation of the RP_{open}) a mixture of all four

ribonucleotide-triphosphates (4 mM each) in 1× TB with 2.5 μCi of [α-^{32}P]-GTP (3,000 Ci/mmol) and 2 mg/mL heparin was added to the reaction.

4. The reaction was continued at 37°C for 15 min and terminated with an equal volume of phenol/chloroform (1:1).
5. The samples were precipitated with ethanol, dissolved in formamide-containing loading solution, denatured at 95°C for 5 min, cooled on ice, and separated by denaturing PAGE.
6. The gel was transferred to Whatman 3-MM paper, covered with polyethylene wrap, and dried for 30 min at 80°C.
7. PhosphorScreen (Perkin Elmer) was placed above the dried gel, exposed (usually overnight), and scanned on Cyclone PhosphorImager (Perkin Elmer). The data were quantified using the OptiQuant software (**Fig. 5b**, *see* **Note 8**).

4. Notes

1. The pYP05 plasmid contains an NtrC-dependent enhancer that strongly activates the *glnAp*2 promoter over 2.5-kb distance *(21, 46)*. It contains thirteen strong 601 nucleosome positioning sequences (NPS) between the enhancer and promoter. The remaining portion of the plasmid does not contain any DNA sequences having high affinity to histones and therefore serves as a "sink" for the excess of histones or histone octamers during reconstitution. Thus chromatin assembly on the plasmid can be conducted either using purified histones or donor chromatin; in both cases the presence of DNA "sink" guarantees complete occupancy of NPSs and minimal promoter/enhancer blockage.
2. After reconstitution of the saturated arrays of 601 nucleosomes only one to two randomly positioned nucleosomes are formed on the DNA region of the plasmid that does not contain NPSs (*see* **Figs. 2 and 3**; Y.S.P., data not shown). Use of regularly spaced nucleosomal arrays allows structural interpretation of the obtained data because the X-ray structures of nucleosomes and the 30-nm fibers formed on the templates have been solved *(39, 47, 48)*.
3. Chromatin assembly on the linear DNA fragment was conducted using the same protocol as in the case of assembly on supercoiled DNA (*see* **Subheading 3.2**). Presence of an excess of donor chromatin in the reaction does not result in binding of an excessive amounts of core histones when chromatin is reconstituted by histone transfer from donor chromatin (**Figs** 2–4).

However reconstitution from purified histones *(42)* has to be conducted much more carefully because even small excess of purified histones can result in complete occupancy of the promoter and enhancer in the absence of DNA "sink" (*see* **Note 1**). In this case, a nonspecific competitor DNA can be added during reconstitution to serve as DNA "sink."

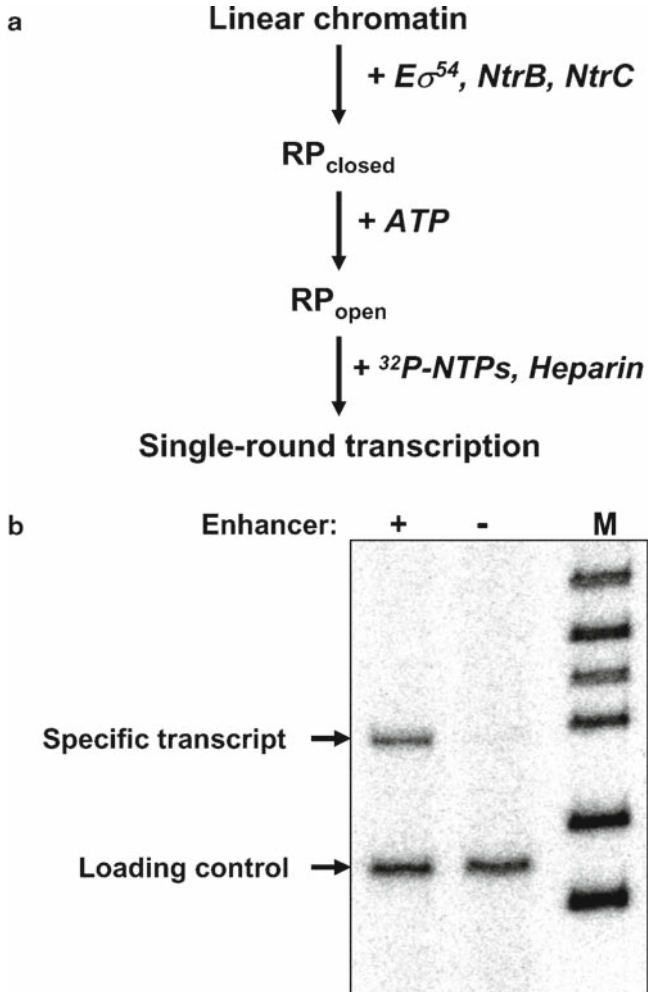


Fig. 5. Analysis of enhancer-dependent *glnAp2* promoter activation in chromatin using the single-round transcription assay. (**a**) The experimental approach. $E\sigma^{54}$: RNA polymerase holoenzyme that recognizes the *glnAp2* promoter. NtrB: protein kinase that phosphorylates NtrC (transcription activator that after the phosphorylation interacts with the enhancer and activates transcription). RP_{closed} and RP_{open} – closed and open initiation complexes, respectively. (**b**) Transcription of enhancer-containing and enhancer-less saturated chromatin templates. The 3.9-kb *Hind*III/*Pst*I fragment of the pYP05 plasmid containing (+) the enhancer or 3.0-kb *Eco*RV fragment missing it (–) were assembled into saturated arrays of nucleosomes and transcribed using the single-round transcription assay **(Fig. 1a)**. Chromatin structure supports efficient transcription that is almost entirely enhancer-dependent. M – end-labeled pBR322-*Msp*I digest. End-labeled DNA fragment was added to the reactions as a loading control.

4. The loading was adjusted to guarantee that intensities of the bands are in linear range of the measuring device. When needed, the standards (DNA fragments with known concentrations) were separated in the same gel.
5. The primer (5′-gaatttcgagggcatgataacgccttttaggg-3′) is localized immediately upstream of the promoter (**Fig. 4b**) and was 5′-end labeled with $\gamma^{[32]}$-ATP (Perkin Elmer) by polynucleotide kinase (NEB) according to the manufacturer's protocol.
6. In the particular example (**Fig. 4c**), histone-free DNA is completely sensitive to each of the restriction enzymes (lanes 1, 5, and 9). In subsaturated chromatin all bands on the gel have similar intensities after *Alu*I or *Msp*I digestion (lanes 3 and 7) indicating that in each chromatin sample on average one of the 13 *Alu*I or 12 *Msp*I sites is accessible to the enzymes cutting within NPSs. Thus on average one of the 13 nucleosomes is missing in subsaturated chromatin. In saturated chromatin all NPSs are >95% protected from *Alu*I and/or *Msp*I digestion (lanes 4 and 8) indicating >95% nucleosome occupancy of each site. At the same time, non-NPS, internucleosomal linker DNA regions are fully sensitive to *Sca*I both in subsaturated and saturated chromatin (lanes 11 and 12) indicating that nucleosomes preferentially occupy NPSs (less than 5% of nucleosomes are formed on the linker DNA regions).
7. Enhancer–promoter communication in chromatin occurs only in *cis (15)* allowing analysis of the same intramolecular interactions that occur in vivo. It depends on the presence of all functional components involved in enhancer action on histone-free DNA *(15, 16)*.
8. The rates of enhancer–promoter communication on DNA, supercoiled or relaxed chromatin templates can be measured quantitatively using the single-round transcription assay. In this case, communication is also initiated by adding ATP after preformation of the RP_{closed} and NtrC–DNA complexes. This approach is similar to the one described in **Subheading 3.6**, but ATP should be added for various time intervals (0, 1, 2, 4, 8, 16, or 32 min) to allow E–P communication for different times. Then $t_{1/2}$ for the rates of communication can be calculated.

Acknowledgments

This work was supported by NSF (0549593) and NIH (GM58650) grants to V.M.S. We would like to thank Dr. T. J. Richmond for providing the12×177-601 plasmid containing an array of twelve 601 NPSs.

References

1. de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* **11**, 447–459.
2. Dean, A. (2004). Chromatin remodelling and the interaction between enhancers and promoters in the beta-globin locus. *Brief Funct. Genomic Proteomic* **2**, 344–354.
3. Tsytsykova, A. V., Rajsbaum, R., Falvo, J. V., Ligeiro, F., Neely, S. R., and Goldfeld, A. E. (2007). Activation-dependent intrachromosomal interactions formed by the TNF gene promoter and two distal enhancers. *Proc. Natl. Acad. Sci. U S A.* **104**, 16850–16855.
4. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. *Nat. Genet.* **32**, 623–626.
5. Kooren, J., Palstra, R. J., Klous, P., Splinter, E., von Lindern, M., Grosveld, F., and de Laat, W. (2007). Beta-globin active chromatin Hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. *J. Biol. Chem.* **282**, 16544–16552.
6. Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* **20**, 2349–2354.
7. Ringrose, L., Chabanis, S., Angrand, P. O., Woodroofe, C., and Stewart, A. F. (1999). Quantitative comparison of DNA looping *in vitro* and *in vivo*: chromatin increases effective DNA flexibility at short distances. *EMBO J.* **18**, 6630–6641.
8. Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.* **7**, 703–713.
9. Shore, D., Langowski, J., and Baldwin, R. L. (1981). DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. U S A.* **78**, 4833–4837.
10. Cloutier, T. E., and Widom, J. (2004). Spontaneous sharp bending of double-stranded DNA. *Mol. Cell.* **14**, 355–362.
11. Crothers, D. M., Drak, J., Kahn, J. D., and Levene, S. D. (1992). DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* **212**, 3–29.
12. Stein, A., Dalal, Y., and Fleury, T. J. (2002). Circle ligation of *in vitro* assembled chromatin indicates a highly flexible structure. *Nucleic Acids Res.* **30**, 5103–5109.
13. Lowary, P. T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42.
14. Thastrom, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M., and Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* **288**, 213–229.
15. Rubtsov, M. A., Polikanov, Y. S., Bondarenko, V. A., Wang, Y. H., and Studitsky, V. M. (2006). Chromatin structure can strongly facilitate enhancer action over a distance. *Proc. Natl. Acad. Sci. U S A.* **103**, 17690–17695.
16. Polikanov, Y. S., Rubtsov, M. A., and Studitsky, V. M. (2007). Biochemical analysis of enhancer–promoter communication in chromatin. *Methods.* **41**, 250–258.
17. Knezetic, J. A., Jacob, G. A., and Luse, D. S. (1988). Assembly of RNA polymerase II preinitiation complexes before assembly of nucleosomes allows efficient initiation of transcription on nucleosomal templates. *Mol. Cell. Biol.* **8**, 3114–3121.
18. Laybourn, P. J., and Kadonaga, J. T. (1992). Threshold phenomena and long-distance activation of transcription by RNA polymerase II. *Science.* **257**, 1682–1685.
19. Ptashne, M., and Gann, A. A. (1990). Activators and targets. *Nature.* **346**, 329–331.
20. Bondarenko, V. A., Jiang, Y. I., and Studitsky, V. M. (2003). Rationally designed insulator-like elements can block enhancer action in vitro. *EMBO J.* **22**, 4728–4737.
21. Liu, Y., Bondarenko, V., Ninfa, A., and Studitsky, V. M. (2001). DNA supercoiling allows enhancer action over a large distance. *Proc. Natl. Acad. Sci. U S A.* **98**, 14883–14888.
22. Popham, D. L., Szeto, D., Keener, J., and Kustu, S. (1989). Function of a bacterial activator protein that binds to transcriptional enhancers. *Science.* **243**, 629–635.
23. Bondarenko, V. A., Liu, Y. V., Jiang, Y. I., and Studitsky, V. M. (2003). Communication over a large distance: enhancers and insulators. *Biochem. Cell. Biol.* **81**, 241–251.
24. Bondarenko, V., Liu, Y. V., Ninfa, A. J., and Studitsky, V. M. (2003). Assay of prokaryotic enhancer activity over a distance in vitro. *Methods Enzymol.* **370**, 324–337.
25. Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y., and Gralla, J. D. (2000). The

bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J. Bacteriol.* **182**, 4129–4136.

26. Sasse-Dwight, S., and Gralla, J. D. (1988). Probing the *Escherichia coli glnALG* upstream activation mechanism *in vivo*. *Proc. Natl. Acad. Sci. U S A.* **85**, 8934–8938.
27. Ninfa, A. J., Reitzer, L. J., and Magasanik, B. (1987). Initiation of transcription at the bacterial *glnAp2* promoter by purified *E. coli* components is facilitated by enhancers. *Cell.* **50**, 1039–1046.
28. Ninfa, A. J., and Magasanik, B. (1986). Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A.* **83**, 5909–5913.
29. Keener, J., and Kustu, S. (1988). Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. *Proc. Natl. Acad. Sci. U S A.* **85**, 4976–4980.
30. Porter, S. C., North, A. K., Wedel, A. B., and Kustu, S. (1993). Oligomerization of NTRC at the *glnA* enhancer is required for transcriptional activation. *Genes Dev.* **7**, 2258–2273.
31. Wedel, A., and Kustu, S. (1995). The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes Dev.* **9**, 2042–2052.
32. Wyman, C., Rombel, I., North, A. K., Bustamante, C., and Kustu, S. (1997). Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science.* **275**, 1658–1661.
33. Buck, M., and Cannon, W. (1992). Activator-independent formation of a closed complex between sigma 54- holoenzyme and *nifH* and *nifU* promoters of *Klebsiella pneumoniae*. *Mol. Microbiol.* **6**, 1625–1630.
34. Su, W., Porter, S., Kustu, S., and Echols, H. (1990). DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. *Proc. Natl. Acad. Sci. U S A.* **87**, 5504–5508.
35. Rippe, K., Guthold, M., von Hippel, P. H., and Bustamante, C. (1997). Transcriptional activation *via* DNA-looping: visualization of intermediates in the activation pathway of *E. coli* RNA polymerase x sigma 54 holoenzyme by scanning force microscopy. *J. Mol. Biol.* **270**, 125–138.
36. Polikanov, Y. S., Bondarenko, V. A., Tchernaenko, V., Jiang, Y. I., Lutter, L. C., Vologodskii, A., and Studitsky, V. M. (2007). Probability of the site juxtaposition determines the rate of protein-mediated DNA looping. *Biophys. J.* **93**, 2726–2731.
37. Thomas, J. O., and Butler, P. J. (1980). Changes in chromatin folding in solution. *J. Mol. Biol.* **144**, 89–93.
38. Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science.* **306**, 1571–1573.
39. Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature.* **436**, 138–141.
40. Huynh, V. A., Robinson, P. J., and Rhodes, D. (2005). A method for the *in vitro* reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone. *J. Mol. Biol.* **345**, 957–968.
41. Walter, W., Kireeva, M. L., Tchernajenko, V., Kashlev, M., and Studitsky, V. M. (2003). Assay of the fate of the nucleosome during transcription by RNA polymerase II. *Methods Enzymol.* **371**, 564–577.
42. Walter, W., and Studitsky, V. M. (2004). Construction, analysis, and transcription of model nucleosomal templates. *Methods.* **33**, 18–24.
43. Studitsky, V. M. (1999). Preparation and analysis of positioned nucleosomes. *Methods Mol. Biol.* **119**, 17–26.
44. Polach, K. J., and Widom, J. (1999). Restriction enzymes as probes of nucleosome stability and dynamics. *Methods Enzymol.* **304**, 278–298.
45. Feng, J., Goss, T. J., Bender, R. A., and Ninfa, A. J. (1995). Activation of transcription initiation from the nac promoter of Klebsiella aerogenes. *J. Bacteriol.* **177**, 5523–5534.
46. Bondarenko, V., Liu, Y., Ninfa, A., and Studitsky, V. M. (2002). Action of prokaryotic enhancer over a distance does not require continued presence of promoter-bound sigma54 subunit. *Nucleic Acids Res.* **30**, 636–642.
47. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature.* **389**, 251–260.
48. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113.

Chapter 34

A Competition Assay for DNA Binding Using the Fluorescent Probe ANS

Ian A. Taylor and G. Geoff Kneale

Summary

Fluorescence spectroscopy is a technique frequently employed to study protein–nucleic acid interactions. Often, the intrinsic fluorescence emission spectrum of tryptophan residues in a nucleic-acid-binding protein is strongly perturbed upon interaction with a target DNA or RNA. These spectral changes can then be exploited in order to construct binding isotherms and the extract equilibrium association constant together with the stoichiometry of an interaction. However, when a protein contains many tryptophan residues that are not located in the proximity of the nucleic-acid-binding site, changes in the fluorescence emission spectrum may not be apparent or the magnitude too small to be useful. Here, we make use of an extrinsic fluorescence probe, the environmentally sensitive fluorophore 1-anilinonaphthalene-8-sulphonic acid (1,8-ANS). Displacement by DNA of 1,8-ANS molecules from the nucleic-acid-binding site of the Type I modification methylase EcoR124I results in red shifting and an intensity decrease of the 1,8-ANS fluorescence emission spectrum. These spectral changes have been used to investigate the interaction of EcoR124I with DNA target recognition sequences.

Key words: Fluorescence spectroscopy, 1,8-ANS, DNA-binding protein, EcoR124I.

1. Introduction

Fluorescence spectroscopy is a useful technique for investigating the interaction of DNA-binding proteins with DNA. Generally, use is made of the intrinsic fluorescence of the protein arising from the aromatic amino acids, which is frequently perturbed in a DNA–protein complex. In some cases, however, changes in the intrinsic fluorescence emission of a protein arising from its interaction with nucleic acid may not be detectable. For example, if tryptophan and/or tyrosine residues are not located in the proximity

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_34

of the DNA-binding site the emission spectrum may not be perturbed by the interaction. Furthermore, in the presence of a large number of tryptophan and tyrosine residues, a relatively small perturbation in the overall emission spectrum brought about by DNA binding may not be observable.

To overcome these problems, an alternative approach is to add an extrinsic fluorescence probe to the system that competes with DNA for the binding site of the protein. One can then measure the change in the fluorescence emission spectrum of the probe as DNA is added. If the fluorescence characteristics of the free and bound probe differ, displacement of the probe by DNA can then be observed.

The fluorescent probe 1-anilinonaphthalene-8-sulphonic acid (1,8-ANS) and its derivatives have long been used to study protein structure *(1, 2)* and more recently to study both protein–nucleic acid and protein–ligand interactions *(3–6)*. ANS has the property that its fluorescence emission spectrum undergoes a 50-nm blue shift along with approx. 100-fold enhancement when transferred from an aqueous environment to a less polar solvent, such as methanol (*see* **Fig. 1**). ANS also binds weakly to hydrophobic patches on protein molecules with an average dissociation constant of 100 μM *(4)*. When molecules of ANS are bound to protein, enhancement and shifting of the fluorescence spectrum similar to that observed in apolar media often occurs (*see* **Fig. 2**). Thus, bound molecules of ANS fluoresce much more strongly and at shorter wavelength than ANS molecules in an aqueous solvent.

The precise reason why ANS molecules bind at DNA-binding sites is not entirely clear, since such sites are not particularly

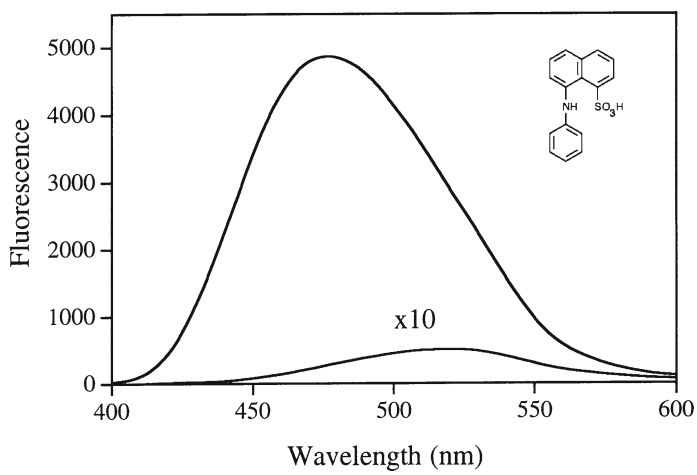


Fig. 1. The effects of solvent polarity on the fluorescence emission spectrum of 1,8-ANS. Fluorescence emission spectra (λ_{ex} = 370 nm) of 50 μM ANS in 100% methanol (*upper*) and in aqueous buffer (*lower*) are shown. *Inset*, the chemical structure of 1,8-ANS.

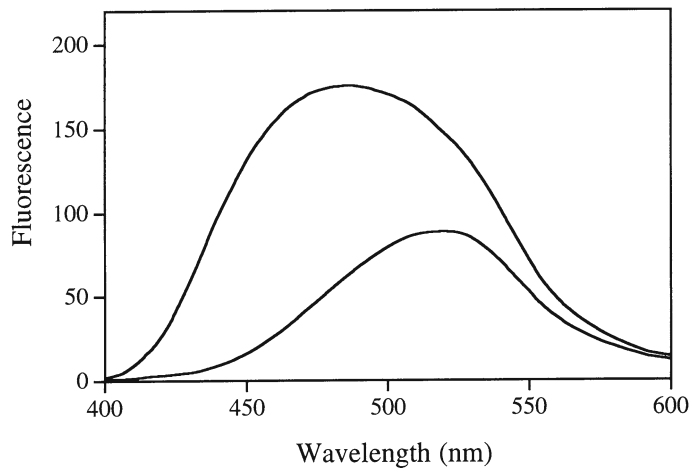


Fig. 2. Enhancement and spectral shift of 1,8-ANS fluorescence induced by the addition of protein. The *lower curve* is the fluorescence emission spectrum (λ_{ex} = 370 nm) of 100 μM ANS. The *upper spectrum* is with the addition of 2.2 μM M.*Eco*R124I.

hydrophobic. Nevertheless, ANS is a planar aromatic molecule (*see* **Fig. 1**) that will have some properties in common with the DNA bases despite the lack of hydrogen-bonding capacity. Furthermore, the negatively charged sulphonate group of ANS may mimic the phosphate group of the DNA backbone.

The protocol described has been successfully applied to an investigation of the DNA-binding properties of a type I modification enzyme (M.*Eco*R124I ***(6)***). It was possible to demonstrate differential binding affinity for an oligonucleotide containing the canonical recognition sequence and one that differs by just one base pair in a non-specific spacer sequence in the enzymes recognition site. It remains to be seen how general the method is since there have been very few instances of its application to DNA-binding proteins. Since it can be established rapidly whether ANS binds to a given protein, and whether there is some release of ANS by the addition of DNA, it is a technique worth investigating. In **Subheadings 3.1** and **3.2** preliminary experiments are described. If these are encouraging, then accurate fluorescence titrations can be undertaken to investigate the DNA-binding characteristics of the protein in more detail (**Subheading 3.3**).

2. Materials

1. A fluorimeter is required that is capable of scanning with both emission and excitation monochromators, and in which the emission and excitation slit widths can be varied. In our laboratory, we routinely use either an ISS PC1 or a Perkin-Elmer

LS50B fluorimeter. Both instruments are controlled by PC, using software written by the manufacturers.

2. It is desirable that the cell holder compartment be thermostatically controlled to ±0.1°C, and some models of fluorimeter have built-in temperature control. Alternatively, this can be achieved by circulating water through the cell holder. In this case, the temperature is controlled by a programmable circulating water bath (e.g. Thermo-NESLAB RTE7).
3. Good-quality quartz cuvettes (preferably stoppered), with all four faces polished are required. A 1 × 0.4 cm (semi-micro) cuvette is preferred, since this minimises the inner filter effect compared to the standard 1× 1 cm (3 ml) cuvettes, see later. The excitation beam should pass through the 0.4-cm path since absorption of the emission beam should be negligible.
4. All buffers should be prepared from the highest-quality reagents and ultrapure water. The buffers should be degassed and filtered to remove any particulate contaminants. In the example provided, the assay was carried out in 10 mM Tris–HCl of pH 8.2, 100 mM NaCl, 5 mM $MgCl_2$ (*see* **Note 1**).
5. High-purity 1-anilino-naphthalene-8-sulphonic acid (1,8-ANS) free of contaminating bis-ANS may be obtained from Molecular Probes, Invitrogen. A 1-mM solution of ANS should be made fresh just prior to use.
6. A 50 μM stock solution of the purified DNA-binding protein of interest.
7. A 100 μM stock of an oligonucleotide duplex containing the DNA-binding site to be investigated (*see* **Note 2**).

3. Methods

3.1. Titration of Protein with ANS

To find optimal conditions for the use of ANS in a DNA-binding experiment it is advisable to first titrate the protein of interest with ANS to check the extent to which the protein binds the fluorescent probe. The precise concentrations of the reagents and composition of the buffer used here work well for M.*Eco*R124I and its subsequent binding to a 30-bp DNA duplex containing a single recognition site. It may be necessary to vary the conditions for other systems.

1. Prepare 250 ml of a degassed and filtered standard buffer that will be used throughout the set of experiments (e.g. 10 mM Tris–HCl of pH 8.2, 100 mM NaCl, 5 mM $MgCl_2$). Dialyse or buffer exchange (*see* **Note 3**) the DNA-binding protein

into this same standard buffer and prepare around 2 ml of a 1 μM solution.

2. Prepare 50 ml of a 1 μM ANS solution again in the same buffer. The concentration of ANS can be determined from its UV absorption spectrum ($\varepsilon_{370\ \text{aqueous}}$ = 5,500 M^{-1} cm^{-1}; *see* **Note 4**).
3. Adjust the excitation and emission slits on the fluorimeter to 2.5 nm (wider slits can be used if the signal is weak) and set the desired temperature. Allow time for the machine to 'warm up' and for the cell holder to temperature equilibrate.
4. Place 1 ml of buffer in the fluorimeter cuvette and using an excitation wavelength (λ_{ex}) of 370 nm record the fluorescence emission spectrum between 400 and 600 nm. Add 2 μl aliquots of the 1 mM ANS solution to the cuvette and mix gently. After each addition, record the fluorescence emission spectrum as before. At the end of the titration measure the OD_{370} of the sample using the same optical path length as used in the fluorescence titration.
5. In a clean cuvette place 1 ml of the solution of 1 μM DNA-binding protein and record the fluorescence emission spectrum between 400 and 600 nm (λ_{ex} = 370 nm). Repeat the ANS titration as in **step 4** of **Subheading 3.1** and again record the OD_{370} of the sample at the end of the titration.
6. To analyse the titration data, choose the wavelength in the emission spectrum that shows the largest difference between the two titrations. This will probably be close to 480 nm, although shorter wavelengths can be used to minimise the background signal from free ANS (*see* **Fig. 2**).
7. To obtain the corrected fluorescence intensity (F_{corr}) firstly if significant correct for any dilution (*see* **Note 5**). Then correct the observed fluorescence intensity (F_{obs}) by application of **Eq. 1** to account for any non-linearity due to inner filter effects (*see* **Note 6**).

$$F_{corr} = F_{obs} \times 10^{(A_{ex}/2)} \quad (1)$$

Apply **Eq. 1** to the data from each point in the titration where *A*ex is the absorbance of the sample at the excitation wavelength (370 nm). Absorbance values may be calculated from the ANS extinction coefficient and the known ANS concentration, taking into account the appropriate path length. However, it should also be checked against the actual OD_{370} value measured at the end of each titration.

8. Plot the corrected fluorescence at the chosen wavelength against the ANS concentration for each titration. A typical case is shown in **Fig. 3**. The titration curve in the absence of

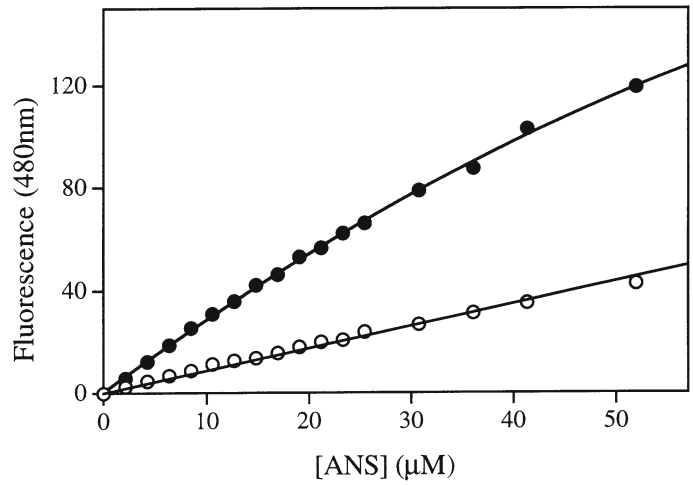


Fig. 3. Titration of M.*Eco*R124I with 1,8-ANS, monitored by fluorescence emission at 480 nm. The *upper curve* shows the fluorescence increase upon successive additions of 1,8-ANS to a 1.5-μM solution of M.*Eco*R124I. The *lower curve* shows the fluorescence increment in the absence of protein. Both titrations have been corrected for dilution and the inner filter effect.

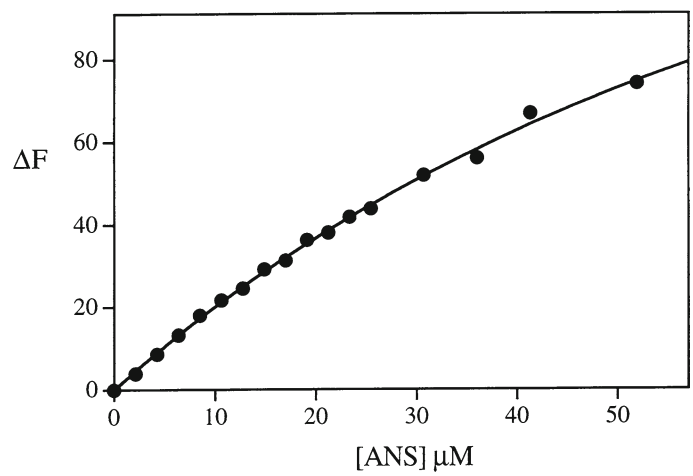


Fig. 4. The binding isotherm for the interaction of 1,8-ANS with M.*Eco* R124I. The data were produced by subtraction of the data points from the *lower curve* from the equivalent points on the *upper curve* in Fig. 3.

protein should be linear, provided the corrections for inner filter and dilution have been correctly applied.

9. Subtraction of the curve of ANS added to buffer from the curve of ANS added to the protein solution yields the binding curve of ANS to the protein as illustrated in **Fig. 4**. The shape of the curve will depend on the ANS-binding properties of the particular protein under study (*see* **Note 7**). In the case shown in **Fig. 4**, this curve is representative of the situation in which several ANS molecules are associated weakly.

3.2. Preliminary Investigation of the Displacement of ANS by DNA

Once satisfied that the protein under investigation binds ANS, one must ascertain if any of the bound ANS molecules are located in the DNA-binding site of the protein. If this is the case, their fluorescence will change upon being displaced by DNA.

1. Make up a solution of 100 μM ANS in buffer (*see* **Note 8**). Measure its fluorescence emission spectrum between 400 and 600 nm (using λ_{ex} = 370 nm).
2. Prepare an identical solution of ANS containing 1 μM DNA. Measure the emission spectrum as in **step 1** of **Subheading 3.2**. These two spectra should be identical as there should be no interaction between the nucleic acid and the ANS.
3. Make up a 100 μM ANS solution containing 1 μM protein, and an identical solution containing in addition, 1 μM DNA. Measure the fluorescence emission spectra of these two samples as earlier.

The presence of protein in the ANS solution should cause a change in the shape and intensity of the emission spectrum. A blue-shifted spectrum accompanied by an increase in quantum yield should be observed. When the nucleic acid is also present and bound to the protein, any ANS (which is weakly bound) in the DNA-binding site of the protein will be displaced and cause the form of the spectrum to move toward that of free ANS (*see* **Fig. 5**).

3.3. Fluorescence Competition Assay

Once it has been established that the DNA fragment of interest shows a measurable displacement of ANS, further investigation of the DNA-binding characteristics of the protein can be carried out. Titrations can be done in a number of dxifferent ways, but we

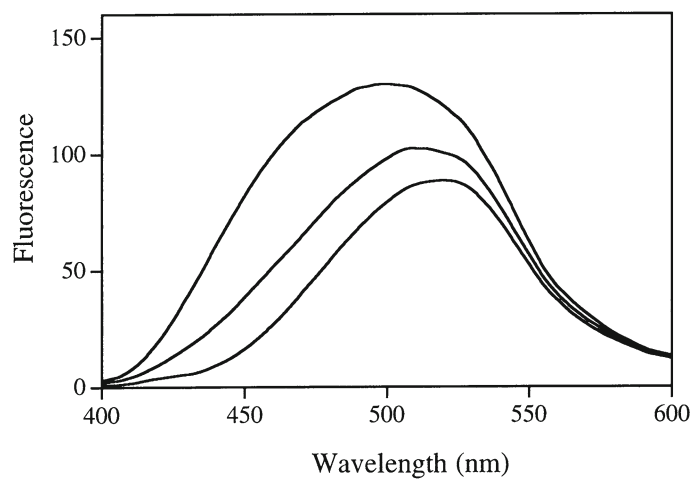


Fig. 5. The fluorescence emission spectra (λ_{ex} = 370 nm) of 100 μM ANS in buffer (*lower*), in a solution of 1 μM M.*Eco*R124 (*upper*), and in a solution of 1 μM M.*Eco*R124I and 1 μM DNA (*middle*).

have found it more reproducible to titrate the protein into a solution of ANS in the presence and absence of DNA. The difference in fluorescence at each point then represents the amount of ANS displaced i.e. the amount of DNA bound. The concentrations used should be those found to be optimal from the earlier experiments. Since the concentration of ANS is constant throughout, the absorption at 370 nm should remain unchanged and no inner filter correction need be applied.

1. Place 1 ml of a 100 μM ANS solution in the cuvette and record the emission spectrum between 400 and 600 nm (λ_{ex} = 370 nm).
2. Add small (2–10 μl) aliquots of the 50 μM stock protein solution to the cell up to a final concentration of 3 μM. After each addition, record the fluorescence emission spectrum.
3. Make up 1 ml of 100 μM 1,8-ANS, this time containing 1 μM DNA, and record the fluorescence emission spectrum as before.
4. Titrate the same small aliquots of the stock 50 μM protein solution into the ANS/DNA mixture and record the emission spectrum after each addition.
5. Select an appropriate wavelength (e.g. 480 nm; *see* **step 6** of **Subheading 3.1**) and plot the fluorescence intensity at each point in the titration against the protein concentration for both experiments (*see* **Fig. 6**).
6. To obtain a binding curve for the protein–nucleic acid interaction, subtract the fluorescence intensities at each point in the two experiments (with and without DNA) and plot the

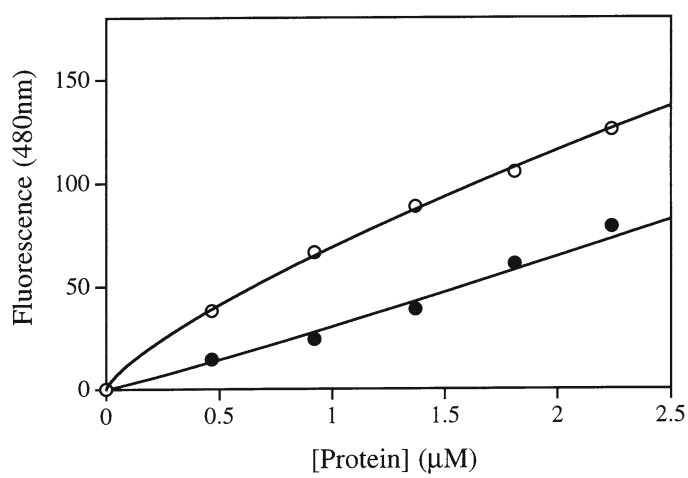


Fig. 6. Data from an ANS competition assay. The upper set of points represent the fluorescence increase resulting from successive additions of M.*Eco*R124I to a 100 μM solution of 1,8-ANS. The lower set of points represent the fluorescence increase when the same titration is carried out in the presence of 1 μM DNA.

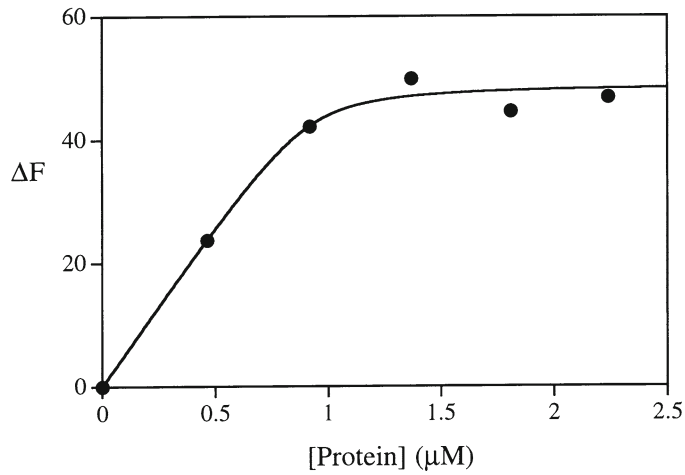


Fig. 7. Binding isotherm for the interaction of M.*Eco*R124I with an oligonucleotide containing its recognition sequence. This curve is generated by the subtraction of the *lower curve* from the *upper curve* in Fig. 6.

resulting DF against protein concentration (*see* **Fig. 7**). This difference represents the amount of bound DNA, since DNA is solely responsible for the decrease in fluorescence through displacement of ANS from the binding site. For a high-affinity DNA protein interaction (with K_d substantially less than the concentration of DNA used in the titration) competition from the ANS will be negligible and a stoichiometric binding curve will be produced; the sharp break in the curve at the stoichiometric point indicates the point at which all the DNA is bound. Curvature of the plot around the stoichiometric point represents a lower-affinity interaction. For a more detailed discussion of DNA-binding curves, *see* Chapter "Fluorescence Spectroscopy and Anisotropy in the Analysis of DNA–Protein Interactions". Variations in the basic method can also be considered (*see* **Notes 9** and **10**)

4. Notes

1. The exact composition of the binding buffer will be dependent on the DNA-binding protein being studied. However, it is advisable to avoid the presence of strongly absorbing compounds and/or quenchers that may interfere with fluorescence measurements. If the interaction has been characterised by another method (for example by gel retardation assay) the fluorescence experiment should initially be carried out in the binding buffer used in these studies.

2. In our laboratory, titrations are carried out using short synthetic DNA duplexes (30mers) containing the protein's recognition sequence. Short oligonucleotides have the advantage that relatively large amounts of highly pure material are readily obtainable. However, the same protocol should be applicable to the use of longer nucleic acids such as restriction fragments or polynucleotides.
3. It is important that all the components in the titration are 'optically matched'. Preparation of a matched protein sample is best achieved by either dialysis or buffer exchange. If the protein sample is limiting, small amounts can be prepared by buffer exchange using NAP5 columns (GE Healthcare) or alternatively by dialysis using Slide-A-lyzer cassettes or Slide-A-lyzer MINI dialysis units (Pierce).
4. The value of $\varepsilon_{370\ \mathrm{aqueous}}$ for ANS in buffer was derived by comparison of the OD_{370} of two equimolar solutions of ANS, one in buffer and the other in 100% methanol for which ε_{370} is known (6,800 M^{-1} cm^{-1}; Molecular Probes). If the buffer used differs significantly from the one used earlier then it is advisable to recalculate a ε_{370}.
5. If the stock solution of ANS is available at high concentration, then the volume of sample in the cuvette during the titration can be assumed to remain constant. However, if a more dilute stock solution is used, there will be significant change in volume (>5%) and it will be necessary to account for this when calculating the ANS concentration at each point in the titration.
6. *See* Chapter "Fluorescence Spectroscopy and Anisotropy in the Analysis of DNA–Protein Interactions" for a more detailed discussion of the inner filter effect. As a guide, for excitation at 370 nm in an aqueous buffer 52 μM ANS has an OD_{370} of 0.11 in a 0.4-cm path length cuvette. This gives rise to an inner filter correction of 1.14 using **Eq. 1**.
7. The binding curve generated for ANS can take many forms. The shape will depend on the number and relative affinity of ANS-binding sites on the protein. If the protein contains high-affinity sites, the curve may be biphasic and may allow the stoichiometry of the strong interaction to be determined. A more likely situation is that there will be numerous ANS-binding sites with differing but weak affinities (K_d > 100 μM). The result of this is a curved plot similar to **Fig. 3**.
8. The concentration of the ANS solution used in the titration must be determined empirically from the previous experiments. It should be high enough to ensure that a good fraction of the ANS-binding sites on the protein are occupied (as determined in **Subheading 3.1**).

9. As well as direct excitation of the fluorescent probe (i.e. with an excitation wavelength of 370 nm for ANS), it may be possible to investigate energy transfer effects between aromatic amino acid residues in the protein and the bound probe. If the excitation is performed at 280 nm to excite both tyrosine and tryptophan, energy transfer to ANS will be apparent from the emission spectrum between 400 and 600 nm. In principle, this effect could also be useful to follow displacement of ANS in a titration with DNA.
10. A further extension of the fluorescent probe approach is to employ the covalent probe 5-((((2-iodoacetyl) amino) ethyl) amino) naphthalene-1-sulphonic acid (1,5-IAEDANS) *(7)*. This reagent reacts with accessible cysteine residues in the protein and has a higher quantum yield than ANS in aqueous solution. One can look at the emission spectrum of the bound probe, or it may be possible to observe energy transfer from aromatic residues in the protein. Any of these fluorescence characteristics could change when the DNA is bound if the probe is located near the DNA-binding site. We have used this technique to study the interaction of M.*Eco*R124I with DNA and found that energy transfer from the protein to the bound probe decreased by over 30% when DNA was bound. As long as the presence of the probe does not inhibit binding, then titrations with DNA can be used to produce DNA-binding curves. However, if the probe does inhibit DNA binding, this can also be informative as the labelled cysteine(s) can be identified by peptide mapping by analogy with the methods reported in chapter "Fluorescence Spectroscopy and Anisotropy in the Analysis of DNA–Protein Interactions".

References

1. Hawe, A., Sutter, M. & Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. *Pharm Res.* **25**, 1487–1499.
2. Brand, L. & Gohlke, J. R. (1972). Fluorescence probes for structure. *Annu Rev Biochem.* **41**, 843–868.
3. Secnik, J., Wang, Q., Chang, C. M. & Jentoft, J. E. (1990). Interactions at the nucleic acid binding site of the avian retroviral nucleocapsid protein: studies utilizing the fluorescent probe 4,4′-bis(phenylamino)(1,1′-binaphthalene)-5,5′-disulfonic acid. *Biochemistry.* **29**, 7991–7997.
4. York, S. S., Lawson, R. C., Jr. & Worah, D. M. (1978). Binding of recrystallized and chromatographically purified 8-anilino-1- naphthalenesulfonate to *Escherichia coli* lac repressor. *Biochemistry.* **17**, 4480–4486.
5. Golynskiy, M. V., Davis, T. C., Helmann, J. D. & Cohen, S. M. (2005). Metal-induced structural organization and stabilization of the metalloregulatory protein MntR. *Biochemistry.* **44**, 3380–3389.
6. Taylor, I., Patel, J., Firman, K. & Kneale, G. (1992). Purification and biochemical characterisation of the EcoR124 type I modification methylase. *Nucleic Acids Res.* **20**, 179–186.
7. Kelsey, D. E., Rounds, T. C. & York, S. S. (1979). lac repressor changes conformation upon binding to poly[dA-T)]. *Proc Natl Acad Sci USA.* 76, 2649–2653.

Chapter 35

Fluorescence Spectroscopy and Anisotropy in the Analysis of DNA–Protein Interactions

Rosy Favicchio, Anatoly I. Dragan, G. Geoff Kneale, and Christopher M. Read

Summary

Fluorescence spectroscopy can be used as a sensitive non-destructive technique for the characterisation of protein–DNA interactions. A comparison of the intrinsic emission spectra obtained for a protein–DNA complex and for free protein can be informative about the environment of tryptophan and tyrosine residues in the two states. Often there is quenching of the fluorescence intensity of an intrinsic emission spectrum and/or a shift in the wavelength maximum on protein binding to DNA. A step-by-step protocol describes the determination of a DNA-binding curve by measurement of the quenching of the intrinsic protein fluorescence.

Fluorescence anisotropy can also be used to obtain a DNA-binding curve if the molecular size of the protein–DNA complex is sufficiently different from the free fluorescing component. Typically an extrinsic fluorophore attached to one or both 5′ ends of single-stranded or duplex DNA is used, for this increases the sensitivity of measurement.

Fitting of the binding curves, assuming a model, can often yield the stoichiometry and association constant of the interaction. The approach is illustrated using the interaction of the DNA-binding domains (HMG boxes) of mouse Sox-5 and mammalian HMGB1 with short DNA duplexes.

Key words: Fluorescence spectroscopy, Anisotropy, Protein–DNA complex, HMG box.

1. Introduction

Changes in the fluorescence emission spectrum of a protein upon binding to DNA can often be used to determine the stoichiometry of binding and equilibrium-binding constants; in some cases the data can also give an indication of the location of particular residues within the protein. The experiments are generally quick and

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_35

easy to perform, requiring only small quantities of material *(1, 2)*. Spectroscopic techniques allow one to measure binding equilibria (unlike, for example, gel retardation assays and other separation techniques which are strictly speaking non-equilibrium methods). Fluorescence is one of the most sensitive of spectroscopic techniques, allowing the low concentrations (typically in the μM range) required for estimation of binding constants for many protein–DNA interactions. Considerable care, however, needs to be exercised in the experiment itself and in the interpretation of results.

The following methods describe the determination of DNA-binding curves by intrinsic fluorescence quenching and by fluorescence anisotropy with an extrinsic fluorophore. Fluorescence Resonance Energy Transfer (FRET) techniques are becoming useful in the area of protein–DNA interactions particularly if one wants to measure the degree of DNA bending, to triangulate or to measure distances, albeit at a low resolution *(2, 3)*. In this method, two or even three extrinsic fluorophores are attached where the emission light of one fluorophore, the donor, is absorbed by another fluorophore, the acceptor, and becomes part of its emission spectrum. Stopped-flow and time-resolved FRET techniques for kinetic and fluorescent lifetime measurements of DNA bending and binding are also advantageous, though they require sophisticated instrumentation *(4, 5)*.

1.1. Fundamental Principles of Fluorescence

A molecule that has been electronically excited with ultraviolet/visible light can lose some of the excess energy gained by a number of processes as it returns to its ground state. In two of these, fluorescence and phosphorescence, this is achieved by emission of light. Phosphorescence is rarely observed from molecules at room temperature and will not be considered further. Whilst electrons can be excited to a number of higher energy states, fluorescence emission in most cases only occurs from the first vibrational level of the first excited state. This has two implications for the measurement of fluorescence emission spectra. Firstly, some of the energy initially absorbed is lost prior to emission, which means that the light emitted will be of longer wavelength (i.e. lower energy) than that absorbed. This is known as Stoke's shift. Secondly, the emission spectrum and therefore the wavelength of maximum fluorescence will be independent of the precise wavelength used to excite the molecule. Thus for tyrosine, the wavelength of the fluorescence maximum is observed around 305 nm, regardless of whether excitation is at the absorption maximum (~278 nm) or elsewhere in the absorption spectrum (e.g. 230 nm). Of course, the fluorescence intensity will change as a consequence of the difference in the amount of light absorbed at these two wavelengths.

The ratio of the number of photons of light emitted as fluorescence to the number of photons initially absorbed, i.e. the

efficiency of the fluorescence process, is known as the fluorescence quantum yield. The value of the quantum yield for a particular fluorophore will depend on a number of environmental factors such as temperature, solvent, and the presence of other molecules which may enhance or diminish the probability of other processes deactivating the excited state. The deactivation or quenching of fluorescence by another molecule, either through collisional encounters or the formation of excited state complexes, forms the basis of many of the fluorescence studies on protein–DNA interactions.

The study of protein–nucleic acid interactions is greatly simplified by the fact that all detectable fluorescence arises from the protein since all of the naturally occurring bases in RNA and DNA are essentially non-fluorescent. Tyrosine and tryptophan residues account for almost all the fluorescence found in proteins. As a general rule, when both residues are present the emission spectrum will be dominated by tryptophan, unless the ratio of tyrosines to tryptophans is very high. The quantum yield of a tyrosine residue in a protein compared to that observed in free solution is generally very low, illustrating the susceptibility of tyrosine to quenching. Tryptophan residues are highly sensitive to the polarity of the surrounding solvent, which affects the energy levels of the first excited state with the result that the emission maximum for tryptophan can range from 330 nm in a hydrophobic environment to 355 nm in water. Thus in proteins containing only one tryptophan, the general environment surrounding the residue can be ascertained.

In protein–DNA interactions, the fluorescence due to tyrosine residues in a protein may be effectively quenched by energy transfer from the tyrosines to the bases of DNA, there being a large overlap in the absorption (DNA) and fluorescence (tyrosine) spectra. Tryptophan fluorescence, like that of tyrosine, can also be quenched (indirectly) by DNA binding. Unlike tyrosine, the emission wavelength maximum can also change if tryptophan is involved in the interaction and this can also be used to monitor DNA binding *(6)*.

The extent to which the fluorescence of a protein is quenched by DNA is proportional to the concentration of quencher. As quenching is due to the formation of a complex between the protein and the DNA, the extent of quenching is proportional to the amount of bound protein. Thus by determining the extent to which the protein fluorescence is quenched when fully bound to DNA (i.e. at saturation), the fraction of bound and free protein at any point in a titration can be determined. From these data, the stoichiometry and binding constant of the interaction can often be obtained. Note that to establish an accurate stoichiometry, a high concentration of DNA is preferred when titrating into protein (i.e. well above the K_d of the complex) to ensure

stoichiometric binding. To establish the binding constant itself, one should be working at much lower concentrations of protein so that at the stoichiometric point, there is a measurable concentration of unbound protein. In the case of protein–DNA interactions having a low K_d, this may not be possible.

If fluorescence quenching is being used to follow DNA binding, it is vital to take account of sample dilution, as well as the increased absorption of the sample as DNA is titrated in. The latter effect is known as the inner filter effect and arises from the absorption of the excitation beam (and generally to a lesser extent, the emission beam) on passing through the sample (*see* **Fig. 1**). One should aim to keep the absorption of the sample (at the excitation wavelength) as low as possible, although absorbances up

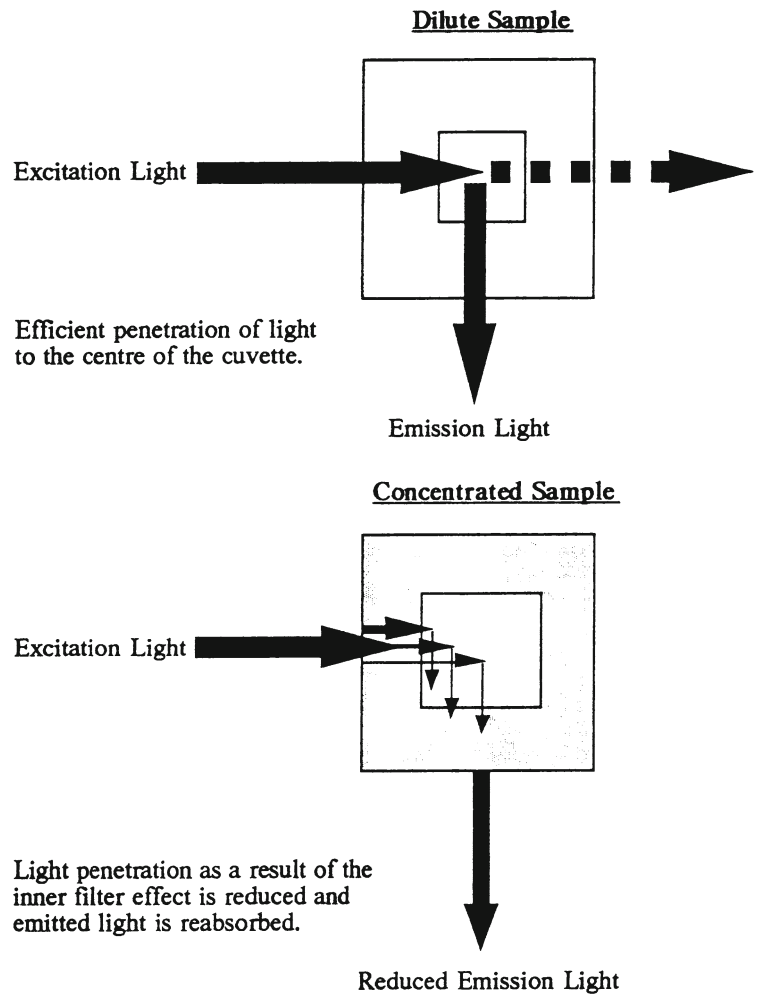


Fig. 1. Schematic representation of the inner filter effect in fluorescence, showing the effect of high concentration on the absorbance of the excitation and emission beams.

to 0.2 can normally be corrected for without too much difficulty. A small pathlength cell will also help. Ideally the absorption of the sample at the excitation and emission wavelengths (A_{ex} and A_{em}) should be measured for each point in the titration (if not, one can calculate these values from the known concentrations of protein and nucleic acid at each point). For normal right-angled geometry of observation, the corrected fluorescence F_{corr} can be obtained from the observed fluorescence F_{obs} by the formula:

$$F_{corr} = F_{obs} \times 10^{(Aex/2 + Aem/2)} \quad (1)$$

Often the value of A_{em} is small enough to ignore (for a detailed treatment of the inner filter correction, *see* **ref.** *7)*. Note that it is equally important to correct for the inner filter effect whether titrating protein into DNA or vice versa.

The use of intrinsic fluorescence, as a method of investigating protein–DNA interactions, is widespread. For example, both binding parameters (K_{obs}) and stoichiometric ratios have been derived for the interaction of the HIV-1 nucleocapsid protein NCp7 with the natural primer $tRNA_3^{Lys}$ and other related RNA molecules *(8)*. Similarly, estimates of binding constants have been determined for the interaction of human replication protein A (hRPA) with single-stranded homopolynucleotides (e.g. poly[dT] and poly[dA]) *(9)*. The interaction of the DNA-binding domain (DBD) of mouse Sox-5, known as the HMG box, with a short DNA duplex *(10)* is used in the text.

1.2. Fundamentals of Fluorescence Anisotropy

Fluorescence anisotropy can be used to measure binding constants if the molecular size of the protein–DNA complex is sufficiently different from the free fluorescing component. Fluorescence emission is not only dependent on the chemical environment of the fluorophore but also on the polarisation of the excitation light used. Fluorescence anisotropy describes the extent of this polarisation of the emission. When excited with plane polarised light a fluorophore often emits light that is also polarised. Rotational diffusion of the fluorophore during the lifetime of the excited state causes depolarisation of the emitted signal. For fluorophores in solution for which the rate of rotational diffusion is much faster than the rate of emission, it follows that the emitted light is fully depolarised and the anisotropy is equal to zero. If the rate of tumbling is much slower than the rate of emission, the emission will have maximal polarisation. If the fluorophore is attached to DNA, the rate of tumbling is such that the emission remains substantially polarised. Binding of a protein to fluorescently labelled DNA will increase the size of the particle and increase the rotational correlation time, i.e. the complex tumbles more slowly in solution than the free DNA. In this case the polarisation of the emitted light will change to give an increase in anisotropy of the

emission. Thus the change in anisotropy can be monitored and subsequently interpreted as a consequence of complex formation. Fluorescence anisotropy, A is defined as:

$$A = \frac{(I_v - I_h)}{(I_v + 2I_h)} \quad (2)$$

where I_v and I_h are the vertically and horizontally polarised emission intensities, respectively, and the value of A is independent of the total intensity of the light emitted. In practice the anisotropy is measured as:

$$A = \frac{I_{vv} - (\mathrm{GF} \times I_{vh})}{I_{vv} + (2 \times \mathrm{GF} \times I_{vh})} \quad (3)$$

and

$$\mathrm{GF} = \frac{I_{hv}}{I_{hh}} \quad (4)$$

where the subscripts h and v in the intensity I indicate vertically and horizontally placed polarisers, the first subscript referring to the excitation polariser position and the second to the emission polariser position. The grating factor, GF, is thus a correction for the polarisation dependence of the instrument transmission, principally the emission monochromator.

Fluorescence anisotropy typically utilizes an extrinsic fluorophore rather than relying on the intrinsic fluorescence emission of the protein. There are several reasons for this. Firstly, to maximise the difference in signal between free and bound states the fluorophore needs to be attached to the smaller of the binding partners: this is usually the DNA. Secondly, extrinsic fluorophores based on fluorescein (known as FAM) or rhodamine derivatives attached to one or both 5′ ends of single-stranded or duplex DNA are available and these have a large extinction coefficient, which increases sensitivity by enabling measurements in the nanomolar concentration range. Thirdly, an extrinsic fluorophore with an emission spectrum in the visible range requires only the use of polarisers constructed of Polaroid rather than the use of quartz polarisers required for emission in the UV from intrinsic Tyr and Trp fluorophores.

Fluorescence anisotropy has been widely used to obtain binding curves for a large number of DNA-binding proteins and domains and include: AT-hooks *(11)*, GCN4-bZIP *(12)*, interferon regulatory factors *(13)*, homeodomains *(14)*, as well as HMG boxes *(15, 16)*. If protein binding results in DNA bending, a double-labelled duplex will display a change in the FRET effect, which can also be used to obtain a binding curve *(15, 16)*. However, unless the bend angle is large, the FRET effect will

be small and this makes the experiment insensitive. However an amplification of small DNA bend angles introduced on protein binding using 'DNA levers' can overcome this to yield binding curves *(17)*.

2. Materials

2.1. Spectrofluorimetry

1. High-quality quartz cuvettes with all four faces polished are required (*see* **Note 1**). A cuvette of square cross-section with light paths of 0.4 and 1.0 cm, a stirring bar, and a close-fitting lid is commonly used (available from Hellma). A sample volume of 1.0–1.5 mL is required.
2. Most commercially available fluorimeters allow scanning by both the excitation and emission monochromators and are suitable for use in these studies. Anisotropic measurements require a fluorimeter with polarisation optics, but since the common fluorophores absorb and emit in the visible range, quartz polarisers are not required and Polaroid is sufficient. Cell sample holders that are thermostatically controlled and contain a magnetic stirrer are essential (*see* **Notes 2** and **3**). We routinely use a Perkin-Elmer LS50B luminescence spectrometer having a Raman S/N ratio of ~1,000. Other more sensitive fluorimeters are available, e.g. the SPEX Fluoromax (from Horiba Jobin Yvon) for work requiring the use of nanomolar concentrations for the determination of very low dissociation constants.
3. To control the spectrofluorimeter and to record fluorescence spectra a computer linked to the fluorimeter is required. The software package FLWinlab running under the WindowsXP operating system is provided for use with the LS50B.

2.2. Reagents and Solutions

1. Reagents used in buffer solutions should be of the highest purity available and the solutions prepared in doubly distilled water. The buffer should have negligible absorbance in the excitation wavelength range (for intrinsic 260–300 nm, for anisotropy 400–500 nm) and should not be used if it shows fluorescence in the emission region (for intrinsic 290–400 nm, for anisotropy 450–600 nm).
2. DNA oligonucleotides were synthesised using phosphoramidite chemistry and purified by reverse-phase HPLC. A 12-bp DNA duplex containing the site-selected AACAAT sequence at its centre *(18)* was used as the Sox-5 HMG box target:

```
5′-GCGAACAATCGG-3′
3′-CGCTTGTTAGCC-5′
```

A 16-bp DNA duplex derived from the Epstein Barr virus BHLF-1 gene promoter and having a TTCAAA core sequence was used as the HMGB1 target *(19)*. For the fluorescence anisotropy experiments, one of the strands was modified at the 5′ end with the fluorescent label 6-FAM:

```
FAM-5′-CGGGTTTCATTAAGGG-3′(FAM-DNAHMG)
     3′-GCCCAAAGTAATTCCC-5′
```

The sequence of the FAM-labelled oligonucleotide is important (*see* **Note 4**). DNA duplexes were constructed by annealing an unlabelled complementary strand at equimolar ratio.

3. The HMG box of mouse Sox-5 (amino acids 182–260 *(18)*) was expressed as a fusion protein in pGEX-2T, using *E. coli* BL21 (DE3) plysS cells. After affinity purification with glutathione-agarose and thrombin cleavage whilst still attached to the column *(20)*, reverse-phase high-performance liquid chromatography (HPLC) was used to purify the protein. Protein was re-dissolved in water and re-folded by extensive dialysis against three changes of 1 L of 100 mM KCl, 10 mM potassium phosphate, 1 mM EDTA (pH 6.0) at 4°C.

   The 100-residue second HMG box of mammalian HMGB1 (amino acids 84–184) and comprising a minimal folded HMG box with a basic C-terminal extension, known as HMGB1-B′ *(21)*, was expressed and purified by anion exchange chromatography and reverse-phase HPLC *(15)*. The minimal second HMG box of HMGB1 (amino acids 92–170), known as HMGB1-B, was expressed as a GST-fusion protein and purified as for Sox-5.

4. Concentrations of all oligonucleotides and DNA duplexes were determined from their UV absorption at 260 nm, after digestion to nucleotides with snake venom phosphodiesterase I (PDE1, from *Crotalus durissus terrificus*, Sigma. (*see* Chapter "Defining the Thermodynamics of Protein/DNA Complexes and Their Components Using Micro-calorimetry"). Account must be made for the contribution to the 260-nm absorption by the fluorescent tag if present (for FAM the extinction coefficient at $\varepsilon_{260\ nm}$ = 28,000 M^{-1} cm^{-1}). Concentrations may also be determined using the fluorophore absorption in the visible region (e.g. for FAM λ_{max} is 496 nm with $\varepsilon_{496\ nm}$ = 78,000 M^{-1} cm^{-1}). Protein concentrations were determined from their UV absorption at 280 nm.

5. Protein and DNA should be dialysed into an identical buffer, in this case 100 mM KCl, 10 mM potassium phosphate, 1 mM EDTA (pH 6.0). Stock solutions of protein and DNA may be prepared in small aliquots and stored at –20°C, assuming this is not detrimental to the protein. Dilution to the working concentration can then be easily made for the fluorescence measurements.

3. Methods

3.1. Intrinsic Fluorescence

The method described in this section assumes that nothing is known concerning the fluorescence properties of the protein or its complex with DNA. Consequently the initial steps described in **Subheading 3.1.1** are concerned with characterising some of the fluorescence properties of the two species such that the optimal conditions for obtaining accurate and reliable data can be obtained. **Subheading 3.1.2** describes a procedure for obtaining data for a protein in which the intrinsic fluorescence is only quenched by DNA on binding, and how these data may be used to obtain information regarding the stoichiometry and dissociation constant of the interaction.

Comparison of the intrinsic emission spectra obtained for protein–DNA complex (at a concentration well above that corresponding to the dissociation constant) and for free protein can be informative. In the case of the Sox-5 HMG box, the complex shows both a quenching of the fluorescence intensity and a shift in the wavelength maxima compared to the free protein (*see* **Fig. 2a**). The difference spectrum having a wavelength maxima at 302 nm is typical of the emission spectrum of tyrosine. This has been interpreted as being due to the fluorescence quenching of three tyrosine residues located close together in the 'minor wing' of the Sox-5 HMG box as they make close approach to the DNA *(10)*. In contrast, fluorescence emission from the two tryptophan residues, located in the hydrophobic protein core and far from the tyrosines in the minor wing, is largely unaffected by DNA binding. A method for obtaining the stoichiometry and dissociation constant of the interaction in which there is fluorescence quenching and a shift in wavelength maxima is outlined in **Subheading 3.1.2.**

3.1.1. Preliminary Experiments

1. Switch on the fluorimeter and allow 10 min for the components to stabilise. Set the excitation and emission slits widths to intermediate values (e.g. 10 nm). Turn on the temperature control.
2. Fill the cuvette with protein solution of ~1 μM concentration (*see* **Note 5**). If the cuvette is rectangular place the smallest pathlength in the excitation beam – this produces the lowest overall absorption. Turn the magnetic stirrer on and allow time for the solution to equilibrate to the temperature of the compartment. To prevent local heating of the solution or possible photodecomposition, the excitation shutter should be kept closed except when taking measurements.
3. If the absorption spectrum of the protein is known set the excitation wavelength to that corresponding to the absorption

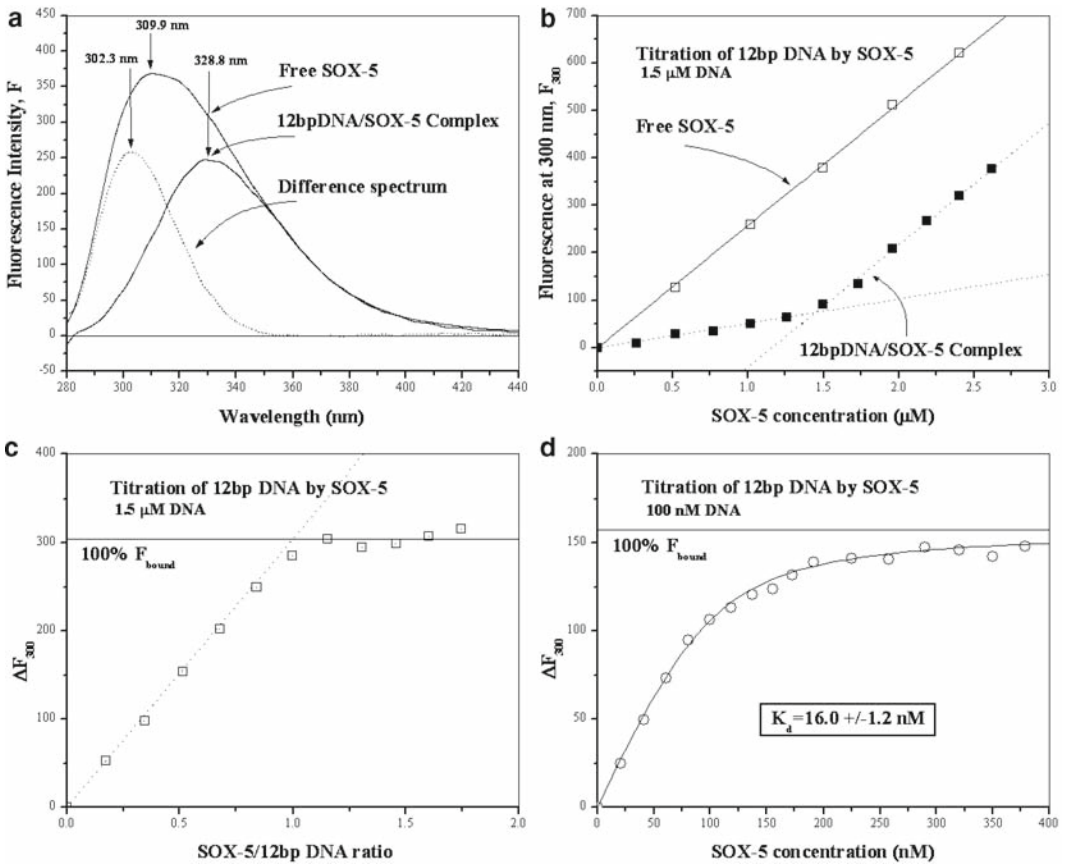


Fig. 2. Titration of the 12-bp DNA duplex with Sox-5 HMG box at 5°C. (**a**) Emission spectra of the free Sox-5, the Sox-5/12-bp DNA complex, and their difference spectrum. The excitation wavelength was 270 nm. The difference spectrum showing a wavelength maxima, λ_{max} of 302 nm, demonstrates that the quenching is of tyrosine fluorescence only. (**b**) Fluorescence intensities at 300 nm for the free Sox-5 and for a titration of 1.50 μM DNA duplex with Sox-5, a concentration well above the dissociation constant K_d. Substantial quenching of the tyrosine fluorescence on binding Sox-5 is evident up to the stoichiometric point (1:1). (**c**) The titration data re-plotted from (**b**) as the difference between the free and bound Sox-5 (ΔF) to show the sharp break at the stoichiometric point. (**d**) The results for a titration of Sox-5 protein into 100 nM DNA duplex, a concentration only 6.25× that of the K_d. The fitted binding curve for a 1:1 interaction is shown. The excitation wavelength was 230 nm to improve sensitivity (though with a larger background).

maximum between 265 and 285 nm; if no such absorption peak exists the protein contains neither tyrosine nor tryptophan residues and it will not fluoresce. If the absorption spectrum is unknown set the excitation wavelength to 280 nm.

4. Open the excitation shutter and quickly scan the emission monochromator between 285 and 400 nm. Identify a wavelength at which there is a maximum value for the intensity. Return the emission monochromator to this wavelength.
5. Find the excitation wavelength maximum between 265 and 285 nm in the same manner, with the emission monochromator

set at the wavelength of maximum fluorescence. Note: The aforementioned FLWinLab software allows for simultaneous scanning of the excitation and emission wavelengths in what is termed a '3D Scan'. This allows both the excitation and the emission wavelength maxima to be determined in one experiment.

6. With both the excitation and the emission wavelengths set at their peak values, adjust the instrument to give a reading corresponding to about 80% of the full scale. Narrow slit widths and a lower amplification (expansion factor, gain) are preferred, and a compromise between the two may have to be found (*see* **Note 6**).
7. Determine the emission spectrum by scanning the emission monochromator over the entire wavelength range over which fluorescence occurs. A scan speed of 100 nm/min is generally suitable.
8. Add a small aliquot of a concentrated DNA solution to the cuvette such that the concentration of DNA is in excess. Mix and immediately check the fluorescence emission at the emission maximum of the protein. Check several times over the next few minutes until a consistent reading is obtained. Allow this time for equilibration in subsequent experiments. Do not adjust the slit widths or the amplification.
9. Obtain an emission spectrum and compare with that obtained for the protein only. If fluorescence quenching is suspected, make sure that allowance for sample dilution has been made. If an inner filter correction is required, measure the absorbance of the sample in a spectrophotometer (in the same cuvette) and correct the observed fluorescence as discussed in the **Subheading 1**.
10. Add aliquots of DNA until there is no further change in fluorescence intensity in the emission spectrum (*see* **Notes 7** and **8**).

3.1.2. Protein–DNA Titrations

1. Examine the emission spectrum of the free protein. If it is characteristic of tyrosine fluorescence check for interference from the Raman band (*see* **Note 9**). If it is characteristic of tryptophan, check for tyrosine contributions which may be masked (*see* **Note 10**).
2. Examine the emission spectrum of the protein bound to DNA. The titration method described in the following passage is particularly applicable when the only change in the spectrum is a change in fluorescence intensity. Several variations of this method are described briefly in **Note 11** including an example where the emission spectrum of the protein shifts on binding DNA.

3. Accurately determine the concentration of protein and DNA solutions, e.g. by UV spectroscopy. As we are titrating DNA into protein try to use a stock concentration of DNA which is at least 20 times the concentration of protein used in the experiment multiplied by the estimated stoichiometric ratio; for example, if the protein concentration used is 10 μM and the stoichiometry estimated to be 1:1 then the DNA concentration should be at least 20 × 10 μM = 200 μM. This would mean that the dilution of the original protein solution will be only 5% at the stoichiometric point.
4. Using the protein solution set up the instrument as described in **steps 1–6** of **Subheading 3.1.1.** If measuring tyrosine fluorescence, use an excitation wavelength near the maximum. This wavelength can also be used for tryptophan excitation if tyrosine fluorescence is insignificant; otherwise, use an excitation wavelength of 295 nm.
5. Run a buffer blank and check that the profile of the emission spectrum is consistent with that previously obtained. Subtract this spectrum from subsequent spectra, if this can be done automatically.
6. Set the emission monochromator to the emission wavelength maximum and ensure that the readout is ~80% of its maximum value. Note down the value.
7. To begin the titration add a small aliquot from the stock DNA solution to the protein in the cuvette. Mix and allow the sample to equilibrate (use the time period determined earlier) before taking a reading. The aliquots should be sufficiently small such that the protein is still greatly in excess and a linear change is observed as more DNA is added.
8. Continue to add the same quantity of DNA for eight to ten points. If changes are still approximately linear at this stage gradually increase the volume of the DNA added, noting the total amount added at each point.
9. When the change in intensity begins to deviate significantly from linearity decrease the size of the aliquot so that more data points are obtained in this region.
10. As quenching approaches the maximum larger aliquots of DNA can be added. Continue until no change in quenching is observed for several points.
11. After the last point check that the emission spectrum of the complex is consistent with that previously obtained for the bound protein.
12. Remove the sample, wash the cuvette thoroughly, and run a blank spectrum consisting of cell plus buffer. This should have negligible or no fluorescence. Subtract any value at the

emission wavelength maximum from the data points, if not already done automatically (*see* **Note 9**).

13. In some cases it may be preferable to titrate protein into DNA, for an example of this method, *see* **ref. *10.*** The procedure is similar to that given here, but in this case the experiment should be repeated by adding protein to the buffer in the absence of DNA as a reference. Subtraction of the two curves should yield a clear binding curve (*see* **Fig. 2**). For a discussion of the merits of whether to titrate protein with DNA or vice versa, *see* **Note 12.**

3.1.3. Data Analysis

1. For each data point calculate the fluorescence quenching as:

$$\Delta F = [F_f - F] \text{ and } Q = 100 \times (\Delta F / F_f) \quad (5)$$

where F is the measured fluorescence intensity and F_f is the fluorescence intensity in the absence of DNA, having made any corrections for inner filter effects and dilution of the sample. Also calculate the DNA and protein concentrations at each point. Then calculate R, the ratio of the concentration of DNA to that of protein.

2. Plot either a graph of Q against R or F against DNA concentration, D. If only one mode of binding is occurring the graph should look like one of the curves shown in **Fig. 3.** If the 'break point' in the titration is sharp, *see* **Fig. 2c**, it indicates a small dissociation constant, K_d (or a large association constant, K_a) compared to the protein concentration. Conversely, weak binding arising from a large dissociation constant (or from too dilute a protein solution) will give a smoothly rising curve with no apparent break point.

3. The stoichiometry of binding is the value of R at which the slope obtained from the initial linear range of the titration crosses the horizontal line defined by Q_{max} (or F_{max}) at which no further change in intensity occurs.

4. Further information can be extracted from the binding curve by fitting to an appropriate model. In the simplest case of a bimolecular interaction ($P + D \Rightarrow PD$), a useful expression to estimate the dissociation constant K_d is:

$$K_d = \frac{(1-\theta) \times \left(D - (\theta \times P_0)\right)}{\theta} \quad (6)$$

where θ is the fraction of bound to total protein at the stoichiometric point and P_0 is the total protein concentration in the cuvette. For a titration in which DNA is added into protein, the fraction of bound protein, θ is given by the analytical function:

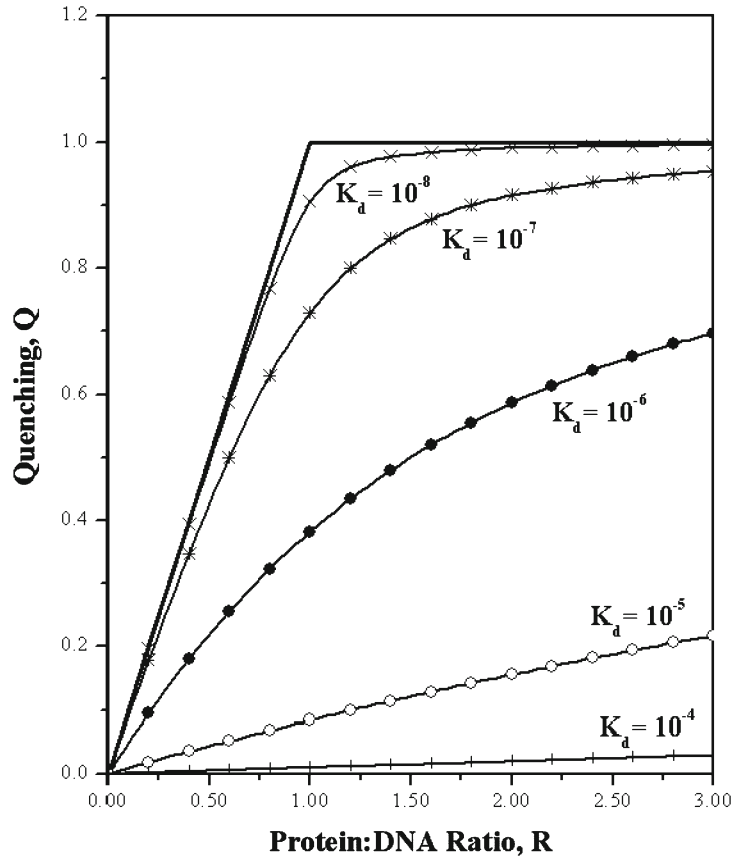


Fig. 3. Graph of fluorescence quenching, Q against protein:DNA ratio, R. The *curves* illustrate the addition of protein to DNA (at 1 μM) for dissociation constants (K_d, M) of 10^{-8} (*times symbol*), 10^{-7} (*asterisk*), 10^{-6} (*filled circle*), 10^{-5} (*open circle*), and 10^{-4} (*plus symbol*), and assuming a bimolecular interaction (1:1 model). The theoretical binding curve for infinitely strong binding (*upper curve*) shows a stoichiometry of one protein bound per DNA.

$$m = 1 + \left(\frac{D}{P_0}\right) + \left(\frac{K_d}{P_0}\right)$$

$$\theta = \frac{\Delta F}{F_f} = \left(0.5m - \sqrt{0.25m^2 - \frac{D}{P_0}}\right) \qquad (7)$$

where D is the DNA concentration at any point in the titration. **Equation** *7* enables fitting of the observed variables by non-linear regression analysis using the Origin fitting/graphical software (*see* **Subheading 3.2.2**).

5. Conversely for a titration in which protein is added into DNA, *see* **Fig. 2**, the fraction of bound DNA, θ is given by the analytical function:

$$m = 1 + \left(\frac{P}{D_0}\right) + \left(\frac{K_d}{D_0}\right)$$

$$\theta = \frac{\Delta F}{\Delta F_b} = \left(0.5m - \sqrt{0.25m^2 - \frac{P}{D_0}}\right) \quad (8)$$

where P is the protein concentration at any point in the titration and D_0 is the total DNA concentration. ΔF_b is the maximum change in fluorescence when all of the DNA is complexed with protein.

6. **Equation 6** also applies to more complex cooperative binding along a linear DNA lattice *(22)*, assuming the cooperativity is sufficiently high, when K_a becomes equal to the apparent binding constant (and approximates to the product of the cooperativity factor and the intrinsic binding constant for one site). For a more extensive discussion of complex DNA-binding equilibria, *see* **ref.** *23.*

3.2. Fluorescence Anisotropy

3.2.1. Anisotropy Experiments

1. With the known working concentration of FAM-labelled-DNA in the cuvette run an emission spectrum using an excitation wavelength of 490 nm (the FAM excitation λ_{max}) over the region 450–600 nm. This serves to check the emission wavelength maxima of ~520 nm and to select slit widths (typically between 2.5 and 12.0 nm) that give a reasonable emission intensity. Slit widths should be varied to increase the signal intensity, depending on the concentration of fluorescently labelled DNA. Use these slit widths in the anisotropy measurements.
2. The anisotropy experiments were performed by titration of protein into DNA. Typically a volume of 5–10 μL of protein is added at each addition into a 1,200-μL volume of FAM-labelled-DNA. Mix the sample by inversion, insert the cuvette into the fluorimeter, and allow the solution to reach equilibrium (~5 min) before measurement of the anisotropy. The sample should be stirred continuously and the sample temperature held constant. DNA concentrations ranging from 30 nM up to 10 μM, depending on the protein construct, were used for the HMGB1-B′ and HMGB1 titrations (*see* **Fig. 4**). Final protein concentrations typically reached 3× that of the DNA. In all cases the dilution factor did not exceed 10%.
3. Fluorescence anisotropy is measured by excitation of the FAM-labelled DNA at 490 nm, setting the excitation path polariser to alternate vertical and horizontal positions. For each position of the excitation polariser, the emission path polariser is also moved alternately to vertical and horizontal positions. At each of the four possible positions the fluorescence intensity is measured at 520 nm to yield I_{vv}, I_{vh}, I_{hh}, and

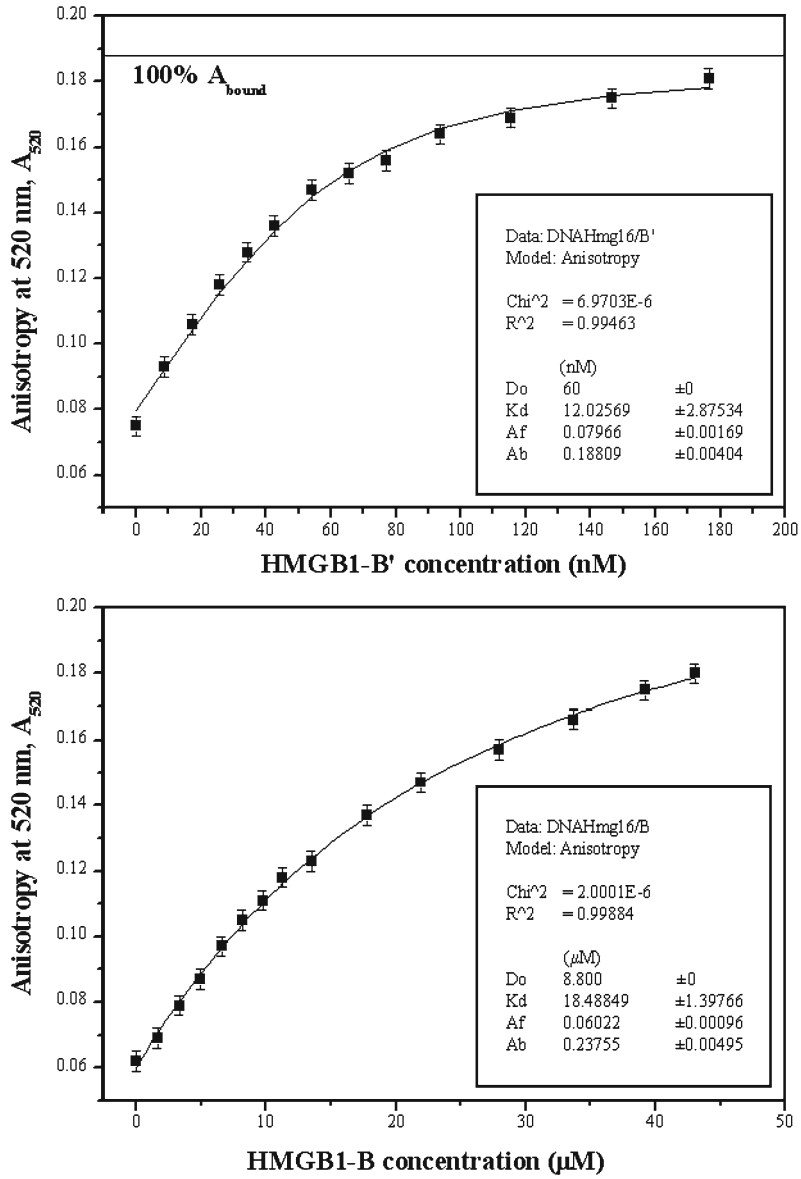


Fig. 4. Plots of the change in fluorescence anisotropy of FAM-DNA^{HMG} upon titration with HMGB1-B′ (**top**) and HMG1-B (**bottom**). Titrations were performed at a number of DNA and protein concentrations; however, the final dataset used for accurate determination of binding constants was recorded with 60 nM FAM-DNA^{HMG} and 1.5 μM HMGB1-B′ (**top**) and 8.8 μM FAM-DNA^{HMG} and 288.0 μM HMGB1-B (**bottom**). The original data for (**a**) are shown in Table 1. The datasets were fitted by non-linear regression to a 1:1 binding model using Origin. The best-fit curves are shown along with the derived K_d values. The anisotropy data fitted to a 1:1 binding model clearly show removal of the very basic C-tail from HMGB1-B′ to generate HMGB1-B raises the K_d from 12 nM to 18.5 μM (i.e. by three orders of magnitude), demonstrating that residues outside the minimal domain help to greatly enhance binding affinity.

I_{hv} (*see* **Subheading 1.2**). Using **Eqs. 3** and **4** the grating factor and anisotropy values can then be calculated for each titration point (*see* **Note 13**).

3.2.2. Data Analysis

1. A tab-delimited text file containing output, the anisotropy values, is read into an Excel spreadsheet (*see* **Table 1**) which can then be directly imported into the 'Data' window of a curve-fitting and graphical software package, such as Origin.
2. A spreadsheet is also useful to calculate the DNA and protein concentrations at each titration point, making any correction for dilution of the sample (*see* **Table 2**). A correction for the inner filter effect could be required at higher FAM-labelled DNA concentrations (at 2.5 μM the absorbance at 496 nm is 0.2). Calculate R, the ratio of the concentration of protein to that of DNA as well. Import into Origin.
3. Plot a graph of anisotropy, A against either protein concentration, P or ratio, R. The graph should be recognisable as a binding curve, similar to that already seen with intrinsic fluorescent quenching data (*see* **Fig. 2** and **Subheading 3.1.3**). The stoichiometry of binding is the value of *R* at which the slope obtained from the initial linear range of the titration crosses the horizontal line defined by the maximum anisotropy, when the DNA is fully bound.

Table 1
Raw data output from a fluorescence anisotropy experiment

Mode:	Anisotropy		Date: ######						
Ex. Slit	Em. Slit	Ex. Wave	Em. Wave	Temp.	I_{vv}	I_{vh}	GF	Int.	Comment
11	11	490	520	20	123.493	104.430	0.951	0.075	60 nM FAM-DNA
11	11	490	520	20	127.373	101.365	0.962	0.093	Plus 7 μL B′
11	11	490	520	20	130.900	99.762	0.967	0.106	Plus 14 μL B′
11	11	490	520	20	133.962	98.694	0.968	0.118	Plus 21 μL B′
11	11	490	520	20	133.221	95.716	0.965	0.128	Plus 28 μL B′
11	11	490	520	20	136.248	95.504	0.970	0.136	Plus 35 μL B′
11	11	490	520	20	136.961	93.429	0.967	0.147	Plus 45 μL B′
11	11	490	520	20	134.135	90.144	0.968	0.152	Plus 55 μL B′
11	11	490	520	20	130.868	86.808	0.971	0.156	Plus 65 μL B′
11	11	490	520	20	123.729	79.592	0.965	0.169	Plus 100 μL B′
11	11	490	520	20	123.142	77.529	0.970	0.175	Plus 130 μL B′
11	11	490	520	20	119.774	74.340	0.969	0.181	Plus 160 μL B′

60 nM FAM-DNA^{HMG} was titrated with successive 7.0 μL aliquots of 1.5 μM HMG1-B′ in 100 mM KCl, 10 mM potassium phosphate buffer (pH 6.0) at 20°C. The excitation wavelength was 490 nm and the emission wavelength was 520 nm, with both slits set at 11 nm. The column labelled 'Int.' is the calculated anisotropy; other column labels are as described in **Subheading 1.2**

Table 2
Spreadsheet example for calculation of protein and DNA concentrations

Titration of 16-bp FAM-HMG DNA with HMGB1-B′					Date: ######	
Ex = 490 nm; Em = 520 nm; Slits = 11.0/11.0 nm; Volume = 1,200 μL						
Buffer:	100 mM KCl, 10 mM Kphosphate (pH 6.0)					
[B′] = 1,500 nM; [F-DNA^{HMG}] = 60 nM						
Add (μL)	Total added (μL)	Total vol. (μL)	Dilution	F-DNA final (nM)	B′ final (nM)	P/F-DNA
0.00	0.00	1,200.00	1.0000	60.0000	0.0000	0.0000
7.00	7.00	1,207.00	1.0058	59.6520	8.6993	0.1458
7.00	14.00	1,214.00	1.0115	59.3081	17.2982	0.2917
7.00	21.00	1,221.00	1.0172	58.9681	25.7985	0.4375
7.00	28.00	1,228.00	1.0228	58.6319	34.2020	0.5833
7.00	35.00	1,235.00	1.0283	58.2996	42.5101	0.7292
10.00	45.00	1,245.00	1.0361	57.8313	54.2169	0.9375
10.00	55.00	1,255.00	1.0438	57.3705	65.7371	1.1458
10.00	65.00	1,265.00	1.0514	56.9170	77.0751	1.3542
15.00	80.00	1,280.00	1.0625	56.2500	93.7500	1.6667
20.00	100.00	1,300.00	1.0769	55.3846	115.3846	2.0833
30.00	130.00	1,330.00	1.0977	54.1353	146.6165	2.7083
30.00	160.00	1,360.00	1.1176	52.9412	176.4706	3.3333

4. A dissociation constant K_d may be obtained by fitting the anisotropy/protein concentration values assuming a binding mode. The equation for a bimolecular (1:1) model is:

$$m = 1 + \left(\frac{P}{D_0}\right) + \left(\frac{K_d}{D_0}\right)$$

$$A = A_f - \left(A_f - A_b\right) \times \left(0.5m - \sqrt{0.25m^2 - \frac{P}{D_0}}\right) \qquad (9)$$

where P is the protein concentration, A is the measured anisotropy upon addition of protein to DNA, A_f is the anisotropy value of the free DNA, A_b is the maximum anisotropy value obtained when the DNA is fully bound to protein, D_0 is the DNA concentration used in the titration, and K_d is the dissociation constant. This function needs to be entered into the fitting software prior to its use in the fitting routine.

5. In the curve-fitting/graphical program the columns of data must be set as follows: the protein concentration P as an independent variable (on the X-axis) and the measured anisotropy A as a dependent variable (on the Y-axis). The DNA concentration (D_0) is a known (fixed) variable. The anisotropy for free and fully bound DNA (A_f and A_b) and the K_d are treated as three unknown variables.
6. The non-linear least squares regression routine is run to best fit to the observed anisotropy/protein concentration data according to the equation for a particular mode of binding (e.g. a 1:1 model, **Eq. 9**). The routine should converge to a solution after a number of iterations, to give values for the unknowns K_d, A_f, and A_b (along with their errors). The curve for the equation representing the best fit of that binding mode may then be plotted, along with the original data (*see* **Fig. 4**). A measure of how well the data are fitted is given by χ^2 and by the regression coefficient r^2. A χ^2 value close to zero and an r^2 value close to one indicate a good fit to the data (*see* **Note 14**).

4. Notes

1. Care should be taken when handling fluorescence cuvettes as both fingerprints and scratches can introduce significant artefacts into the experiment. After use, cuvettes should be thoroughly washed with ethanol or acetone, rinsed with distilled water, and dried thoroughly using a stream of nitrogen or helium gas. If contamination is great, then immerse the cells for several hours in a dilute solution of Hellmanex (from Hellma GmbH), rinse with water, and dry.
2. Changes in temperature affect the viscosity of the solution and hence the number of collisions a fluorophore makes with solvent molecules. Since the viscosity of water is inversely proportional to temperature it follows that there is less collisional quenching and a greater fluorescence at lower temperatures. Both Trp and Tyr fluorescence are highly sensitive to temperature, as is the intrinsic fluorescence of a protein (*see* **ref. *24*** and references therein). For accurate fluorescence measurements temperature control is an essential.
3. A stirred cuvette is similarly required to maintain the thermal equilibrium. If the fluorimeter is not equipped with a magnetic stirrer unit, adequate mixing can usually be achieved by gently drawing the solution in and out through a plastic pipette tip. Avoid the introduction of bubbles into the sample by stirring

or pipetting too fast as this can both denature the protein and cause light scattering.

4. In designing short synthetic oligonucleotides for binding studies, GC base pairs are often added at the end(s) of the DNA duplex to reduce fraying of terminal base pairs. However, the quantum yield of an extrinsic fluorophore attached to the end of a DNA can be affected by the choice of the adjacent (terminal) nucleotide, e.g. in 3′-fluorescein- and 5′-hexachlorofluorescein (HEX)-labelled DNAs an adjacent guanosine nucleotide reduces the quantum yield, i.e. fluorescent intensity by as much as 40% *(25, 26)*. Thus to remove potential quenching effects in fluorescein- (and its derivatives) labelled DNA duplexes but still have a GC terminal base pair, one should design an oligonucleotide with a cytosine nucleotide adjacent to the fluorophore modification, with a base-paired guanosine in the complementary oligonucleotide. Note: 3′-Cy3-labelled DNAs are little affected by the choice of adjacent nucleotide.
5. Fluorescence intensity is only proportional to concentration when the absorbance is no greater than ~0.2 OD units, at the excitation wavelength selected. Remember that the absorption bands for proteins and nucleic acids overlap, and in titrations the contribution of the nucleic acid to the overall absorption must be considered as well as the protein.
6. Whilst it is preferable to have narrow excitation and emission slit widths and a low amplification factor, there may be a need to compromise in order to obtain a stable reading. For proteins displaying tyrosine fluorescence, the small wavelength difference between the excitation and emission maxima suggests that it would be better to maintain narrow slits and increase the signal amplification. For proteins dominated by tryptophan fluorescence the greater the difference between the excitation and emission wavelengths, the greater the feasibility of increasing the slit widths and maintaining a lower amplification. In general, when measuring emission spectra it is better to use a narrow excitation slit width and to widen the emission slit width. For broad banded spectra such as that seen with tryptophan both slits can be widened.
7. If no changes in the emission spectrum of the protein are observed when DNA is added (after inner filter and dilution corrections if necessary), then either the protein is not binding to DNA or binding cannot be detected by this procedure and will need to be assessed by another method such as fluorescence anisotropy or FRET.
8. Note the molar ratio of DNA:protein at which no further changes occur. This will provide a rough guide for future experiments.

9. Tyrosine emission can often be confused with Raman scattering from water molecules which occurs at ~305 nm when an excitation wavelength of 280 nm is used. The presence of the Raman band can be assessed by measuring the emission spectrum using a different excitation wavelength. The fluorescence emission spectrum is independent of excitation wavelength, whereas Raman scattering occurs at a constant wavenumber (=$1/\lambda$ in cm^{-1}) difference from that used for excitation and will shift in the same direction as the change in excitation wavelength. In aqueous solutions this shift is –3,380 cm^{-1}. The contribution of the Raman band to the overall intensity of the signal can be assessed by running an emission spectrum of a buffer blank. Automatically subtract out this spectrum from subsequent spectra where possible. Note: The FLWinLab software includes a 'PreScan Mode' in which the Raman absorption, the Rayleigh scatter, and their second-order peaks are identified automatically.
10. To check the contribution tyrosine may make to a fluorescence emission spectrum dominated by tryptophan, run an emission spectrum using an excitation wavelength of 295 nm. At this wavelength only tryptophan emission will be observed. If the emission spectrum is unchanged then it can be concluded that the contribution from tyrosine residues is negligible. (Of course the intensity will be lower as tryptophan absorption is greater at 280 nm than it is at 295 nm.)
11. In cases where both the emission wavelength maximum shifts and the fluorescence intensity is quenched, the method described can be used provided that an emission wavelength is chosen outside the wavelength region overlapped by the emission spectra of the free and bound protein. Alternatively, the ratio of the intensity of the emission maxima of the free and bound protein can be followed; for an example *see* **ref.** ***6***. The use of a ratio method means that the dilution factor and inner filter correction can usually be ignored, although strictly speaking the ratio is not a linear function of degree of binding.
12. We have found that in some cases different results can be found dependant on the direction of the titration; this can occur when the fluorescence changes observed include contributions from protein–protein interactions accompanying DNA binding in addition to (or instead of) contributions from the interaction with the DNA itself (**ref.** *27)*.
13. In the LS50B the 'Single Read' application of the FLWinLab software automatically controls the change of both polariser positions. Calculation and display of grating factors and anisotropy values is then in real time. All that is required is to enter the excitation and emission wavelengths, with their

respective slit widths and an integration time (typically 1–10 s). The relevant intensities, grating factor, and anisotropy values may be saved into a spreadsheet after each addition of protein (*see* **Table 1**).

14. It can be difficult to decide between different binding modes on the basis of small differences in the fits, e.g. slightly different χ^2 values. One should be careful to not over-interpret, and unless there is good reason the simplest binding model that fits the data should be chosen.

Acknowledgements

We gratefully acknowledge the financial support from a Wellcome Trust grant to the Portsmouth laboratory and an NSF grant (MCB 0519381) to the Baltimore laboratory. R.F. would like to thank IBBS, University of Portsmouth for a bursary.

References

1. Harris, D.A. and Bashford, C.L., eds. (1987). *Spectrophotometry and Spectrofluorimetry:* A Practical Approach. IRL Press, Oxford, England (Chaps. 1 and 4 are particularly relevant).
2. Lakowicz, J.R. (2006). *Principles of Fluorescence Spectroscopy.* 3rd edn. Springer, New York, USA.
3. Clegg, R.M., Murchie, A.I. and Lilley, D.M. (1994). The solution structure of the four-way DNA junction at low-salt conditions: a fluorescence resonance energy transfer analysis. *Biophys. J.* **66**, 99–109.
4. Eis, P.S. and Millar, D.P. (1993). Conformational distributions of a four-way DNA junction revealed by time-resolved fluorescence resonance energy transfer. *Biochemistry.* **32**, 13852–13860.
5. Phillips, N.B., Jancso-Radek, A., Ittah, V., Singh, R., Chan, G., Haas, E. and Weiss, M.A. (2006). SRY and human sex determination: the basic tail of the HMG box functions as a kinetic clamp to augment DNA bending. *J. Mol. Biol.* **358**, 172–192.
6. Kneale, G.G. and Wijnaendts van Resandt, R.W. (1985). Time resolved fluorescence of the Pf1 bacteriophage DNA binding protein: determination of oligo- and polynucleotide binding parameters. *Eur. J. Biochem.* **149**, 85–93.
7. Birdsall, B., King, R.W., Wheeler, M.R., Lewis, C.A. Jr., Goode, S.R., Dunlap, R.B. and Roberts, G.C. (1983). Correction for light absorption in fluorescence studies of protein-ligand interactions. *Anal. Biochem.* **132**, 353–361.
8. Mély, Y., de Rocquigny, H., Sorinas-Jimeno, M., Keith, G., Roques, B.P., Marquet, R. and Gérard, D. (1995). Binding of the HIV-1 nucleocapsid protein to the primer tRNA-3Lys in vitro, is essentially not specific. *J. Biol. Chem.* **270**, 1650–1656.
9. Kim, C. and Wold, M.S. (1995). Recombinant human replication protein A binds to polynucleotides with low cooperativity. *Biochemistry.* **34**, 2058–2064.
10. Privalov, P.L., Jelesarov, I., Read, C.M., Dragan, A.I. and Crane-Robinson, C. (1999). The energetics of HMG box interactions with DNA. Thermodynamics of the DNA binding of the HMG box from mouse Sox-5. *J. Mol. Biol.* **294**, 997–1013.
11. Dragan, A.I., Liggins, J.R., Crane-Robinson, C. and Privalov, P.L. (2003). The energetics of specific binding of AT-hooks from HMGA1 to target DNA. *J. Mol. Biol.* **327**, 393–411.
12. Dragan, A.I., Frank, L., Liu, Y., Makeyeva, E.N., Crane-Robinson, C. and Privalov, P.L. (2004). Thermodynamic signature of GCN4-bZIP binding to DNA indicates the role of water in discriminating between the AP-1 and ATF/CREB sites. *J. Mol. Biol.* **343**, 865–878.
13. Hargreaves, V.V., Makeyeva, E.N., Dragan, A.I. and Privalov, P.L. (2005). Stability and DNA binding ability of the DNA binding domains of interferon regulatory factors 1 and 3. *Biochemistry.* **44**, 14202–14209.

14. Dragan, A.I., Li, Z., Makeyeva, E.N., Milgotina, E.I., Liu, Y., Crane-Robinson, C. and Privalov, P.L. (2006). Forces driving the binding of homeodomains to DNA. *Biochemistry.* **45**, 141–51.
15. Dragan, A.I., Klass, J., Read, C., Churchill, M.E., Crane-Robinson, C. and Privalov, P.L. (2003). DNA binding of a non-sequence-specific HMG-D protein is entropy driven with a substantial non-electrostatic contribution. *J. Mol. Biol.* **331**, 795–813.
16. Dragan, A.I., Read, C.M., Makeyeva, E.N., Milgotina, E.I., Churchill, M.E., Crane-Robinson, C. and Privalov, P.L. (2004). DNA binding and bending by HMG boxes: energetic determinants of specificity. *J. Mol. Biol.* **343**, 371–393.
17. Dragan, A.I., Liu, Y., Makeyeva, E.N. and Privalov, P.L. (2004). DNA-binding domain of GCN4 induces bending of both the ATF/CREB and AP-1 binding sites of DNA. *Nucleic Acids Res.* **32**, 5192–5197.
18. Denny, P., Swift, S., Connor, F. and Ashworth, A. (1992). An Sry-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. *EMBO J.* **11**, 3705–3712.
19. Ellwood, K.B., Yen, Y.M., Johnson, R.C. and Carey, M. (2000). Mechanism for specificity by HMG-1 in enhanceosome assembly. *Mol. Cell. Biol.* **20**, 4359–4370.
20. Read, C.M., Cary, P.D., Preston, N.S., Lnenicek-Allen, M. and Crane-Robinson, C. (1994). The DNA sequence specificity of HMG boxes lies in the minor wing of the structure. *EMBO J.* **13**, 5639–5646.
21. Teo, S.H., Grasser, K.D. and Thomas, J.O. (1995). Differences in the DNA binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *Eur. J. Biochem.* **230**, 943–950.
22. Kelly, R.C., Jensen, D.E. and von Hippel, P.H. (1976). Fluorescence measurements of binding parameters for bacteriophage T4 gene 32 protein to mono-, oligo-, and polynucleotides. *J. Biol. Chem.* **251**, 7240–7250.
23. McGhee, J.D. and von Hippel, P.H. (1974). Theoretical aspects of DNA-protein interactions: cooperative and non-cooperative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **86**, 469–489.
24. Crane-Robinson, C., Read, C.M., Cary, P.D., Driscoll, P.C., Dragan, A.I. and Privalov, P.L. (1998). The energetics of HMG box interactions with DNA. Thermodynamic description of the box from mouse Sox-5. *J. Mol. Biol.* **281**, 705–717.
25. Behlke, M.A., Huang, L., Bogh, L., Rose, S. and Devor, E.J. (2005). Fluorescence quenching by proximal G-bases. *Integrated DNA Technologies Research Report.* http://www.idtdna.com.
26. Bogh, L.D. and Behlke, M.A. (2005). Spectra of DNA-conjugated fluorescent dyes. *Integrated DNA Technologies Research Report.* http://www.idtdna.com.
27. Carpenter, M.L. and Kneale, G.G. (1991). Circular dichroism and fluorescence: analysis of the interaction of Pf1 gene 5 protein with poly(dT). *J. Mol. Biol.* **217**, 681–689.

Chapter 36

Circular Dichroism for the Analysis of Protein–DNA Interactions

P.D. Cary and G. Geoff Kneale

Summary

Circular dichroism (CD) is a well-established technique for the analysis of both protein and DNA structure. The analysis of protein–nucleic acid complexes presents greater challenges, but at wavelengths above 250 nm, the circular dichroism signal from the DNA predominates. Examples are given of the use of CD to examine structural changes to DNA induced by protein binding.

Key words: DNA structure, Protein–nucleic acid complexes, Single-stranded DNA binding proteins, Transcriptional activators and repressors, DNA intercalation.

1. Introduction

The asymmetric carbon atoms present in the sugars of nucleotides and in all the amino acids (with the exception of glycine) result in nucleic acids and proteins displaying optical activity. Further contributions to the optical activity of the polymers result from their ability to form well-defined secondary structures, in particular helices, which themselves possess asymmetry. As a consequence, circular dichroism (CD) has found widespread use in secondary structure prediction of proteins using a number of modelling techniques *(1–3)* and online modelling is available *(4)*. Similar studies, though less widespread, have sought to correlate structural parameters of DNA with their CD spectrum *(5–7)*, with some success particularly in assigning quaternary structures to nucleic acids (e.g. in the case of triplex DNA and G-quartet mediated structures) *(8, 9)*. It follows that the disruption of

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_36

secondary structure by, for example, denaturation or ligand binding can be usefully followed by CD.

Plane polarized light can be resolved into left- and right-handed circularly polarized components. CD measures the difference in the absorption of these two components, i.e.

$$\Delta\varepsilon = \varepsilon_{\mathrm{L}} - \varepsilon_{\mathrm{R}}, \tag{1}$$

where ε is the molar extinction coefficient (M^{-1} cm^{-1}) for the left (L) and right (R) components (*see* **Note 1**). When passing through an optically active sample the plane of polarized light is also rotated; hence, the emerging beam is elliptically polarized. Thus CD is often expressed in terms of ellipticity (θ_λ in degrees) or molar ellipticity ($[\theta]_\lambda$ in degrees cm^2/dm).

$$[\theta]_\lambda = 100\theta_\lambda / c_{\mathrm{m}} l, \tag{2}$$

where c_{m} is the molar concentration and l is the path length in cm. The two quantities are interconvertible with the expression $[\theta]_\lambda = 3{,}298\ \Delta\varepsilon$, or more practically.

$$[\theta] = (\theta_\lambda S M_{\mathrm{mrw}}) / (c_{\mathrm{g}} l 100), \tag{3}$$

S is the instrument scale correction factor, M_{mrw} is the mean residue weight, θ_λ is in degrees, l is in decimetres and c_{g} is in gram per millilitre (*see* **Note 2**).

The overlap of the UV absorption bands of nucleic acids and proteins means that CD studies of protein–DNA interactions can be complicated by the contributions observed from both components. This is particularly true for wavelengths less than 250 nm. In practice CD spectra between 250 and 300 nm are dominated by that of the nucleic acid, the contribution arising from the aromatic chromophores of the protein being weak by comparison and generally few in number in a protein. Changes in conformation can usually be attributed to the polynucleotide as the conformational distribution of aromatic amino acids in a globular protein is such that the majority are normally involved with stabilizing the inner core; this means that a large conformational change throughout the protein would be required to cause a significant change in the CD spectrum.

The low molar ellipticity of polynucleotides means that for accurate CD measurements high concentrations (10^{-4}–10^{-5} M) of nucleotide are required. For this reason CD is not generally used to determine binding constants of protein–DNA interactions. However CD can be used to obtain accurate values for the stoichiometry of protein–nucleic acid interactions *(10)*; in the case of the fd gene 5 protein, CD was used to show the existence of two distinct binding modes *(11)*. CD has also been used to show that conformational changes induced by bound *Lac* repressor are different for operator DNA and for random sequence DNA *(12)*.

Similar studies on *Gal* repressor demonstrated the involvement of the central G–C base pairs of the operator sequence in repressor-induced conformational changes *(13)*. Studies on the interaction of *cro* protein of bacteriophage λ have also revealed different conformational changes for specific and non-specific DNA binding *(14)*. Despite the apparent lack of any direct interaction of the central base pair of the operator sequence with *cro* protein, base substitution at this site was shown to affect the CD spectrum considerably.

Some additional examples in which CD has been used to examine protein–DNA interactions include the SRY-related protein sox5 and the controller protein, C.AhdI protein *(15, 16)*. Both these proteins cause the DNA to bend. Sox-5 is a DNA-sequence-specific protein that has a single globular domain which opens up the DNA by intercalation of the methionine side chain in helix 1.The N- and C-terminal residues stabilize the bending *(15)*, and as a consequence there is a large enhancement of ε and a wavelength shift. On the other hand C.AhdI is a dimer in free solution at CD concentrations; two dimers bind on to the operator sequence 35-bp operator sequence *(16)* causing a large enhancement of $\Delta\varepsilon$ but no wavelength shift (*see* **Fig. 1a**). Hence, although both cause DNA bending, the two proteins must bind in different ways.

It is interesting to note that the 35-bp operator to which C.Ahd binds has two sites, one for each dimer; binding to the left operator, O_L, is much stronger than that to the right operator O_R, so this site is filled first. However, binding of a second dimer to O_R is highly cooperative and under most experimental conditions, only the tetrameric complex is seen. Nevertheless, it was possible to observe binding to the left-hand operator, O_L, when O_R was mutated to a non-specific sequence, which greatly decreases the affinity at this site *(16)*. Titrating the protein into the DNA and observing the change in CD signal (**Fig. 1b**) show that the binding of each dimer causes an identical change in $\Delta\varepsilon$ (i.e. the CD change is additive). This suggests that the structural change is localized to the two sites and that there is no additional perturbation to the DNA structure (at least none observable by CD) arising from the interaction between the two bound C.AhdI dimers.

A large increase in ε was also observed when a type I DNA methyltransferase bound to its DNA recognition sequence, suggesting deformation of the DNA *(17)*. CD has, in addition, been shown to be a useful tool when studying RNA triplex structures *(18)*, binding of peptides TAT and REV *(19, 20)* to RNA and T7 RNA polymerase disordering *(21)*.

It should be emphasized that CD provides complementary data to other spectroscopic techniques such as fluorescence, since such techniques can monitor different components of the interaction.

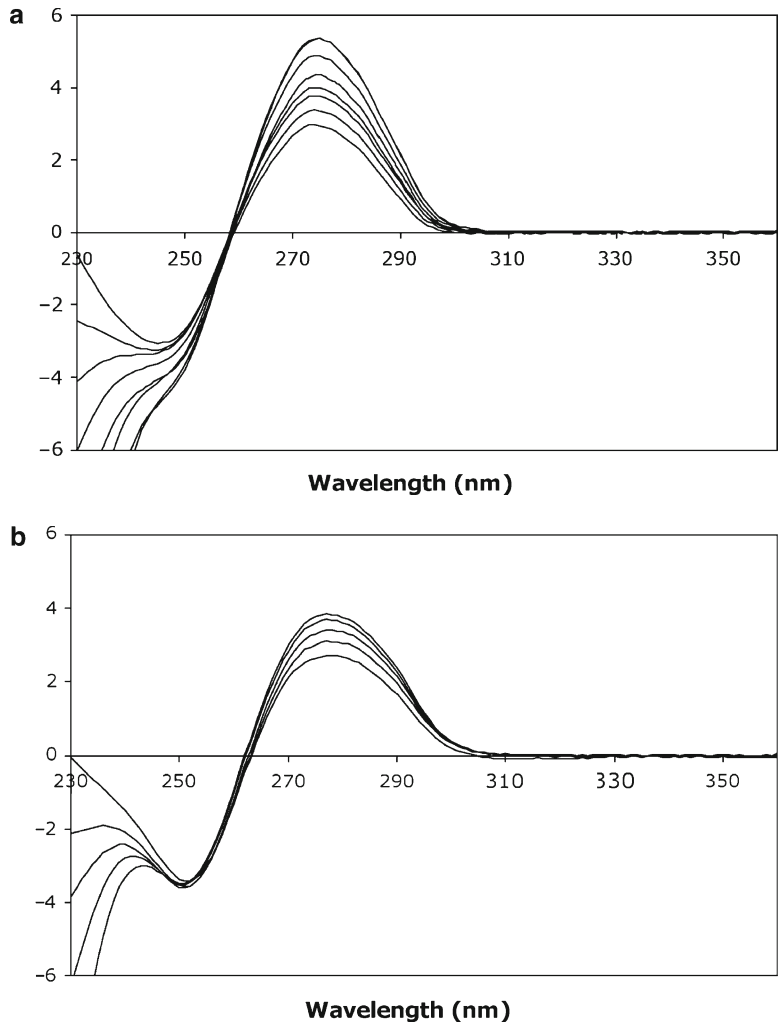


Fig. 1. Circular Dichroism spectra show the result of C.AhdI binding. (**a**) The 35-bp native operator sequence and (**b**) the 35-bp sequence in which the right operator has been mutated. Four scans were taken and averaged from 360 to 230 nm and data collected every 1 nm, following the addition of 50-μM aliquots of C.AhdI. Representative spectra are shown, corresponding to protein: DNA ratios of (**a**) 0, 0.54, 1.1, 1.6, 2.2, 3.0, 4.1, and 4.6 (**b**) 0, 0.49, 1.0, 1.74, and 2.23. All spectra shown were corrected for both baseline differences and dilutions made by titration additions. These experiments were carried out in 40 mM citrate of pH 5.6, 100 mM NaCl, 1 mM EDTA at 20°C in a 4-mm path-length cell (119-004F QS). In (**a**) and (**b**) saturation is close to being reached when the molar ratio of protein: DNA is 4 and 2, respectively. As indicated in Fig. 2, $\Delta\varepsilon$ for the mutated sequence increases at a much lower rate beyond the 2:1 ratio, indicating very weak binding of the second protein dimer to the non-specific DNA sequence that has been introduced in place of the right operator.

Indeed, even the apparent stoichiometry of binding can be significantly different for this reason, even when measured under the same solution conditions *(10)*.

2. Materials

1. A high-quality quartz cell with low strain is required for accurate measurements. The cell path length will depend on the absorption properties and concentration of the sample. Cells with path lengths between 0.01 and 1 cm are often used, depending on the CD signal to be measured and the absorption of the sample (including the buffer). Path lengths of 0.4–1.0 cm is usually recommended for measurement of the nucleic acid signal in the vicinity of 275 nm, where the signal is weak and buffer absorption is negligible. Protein secondary structure analysis requires measurements from 260 nm down to 180 nm (or lower); hence, typically the cell path lengths used will be 0.2–2 mm, dependent on the sample concentration and the solvent absorption properties.
2. Buffers should be prepared using high-quality reagents and water. Use buffers that have low absorbance in the wavelength region of interest. Tris–HCl, citrate, perchlorate, and phosphate are routinely used. Preferably use buffers at low concentrations 1–10 mM and where possible use the salt KF as a replacement for NaCl or KCl to achieve good spectra in the far UV (160–210 nm), having checked by other means that this does not affect binding.
3. Stock solutions of appropriate protein and nucleic acid solutions should be in the same buffer: the protein should be as concentrated as possible avoiding aggregation, to minimize dilution during the titration. If a synthetic DNA fragment containing the recognition sequence is to be used, it should be close to the minimum size required for binding to maximize the change in CD signal and avoid non-specific binding. To avoid denaturation and degradation keep concentrated solutions of protein and DNA frozen in small aliquots, if it has been established that this procedure does not damage the protein. Whilst working at the laboratory bench, keep all solutions on ice to prevent aggregation or degradation and sealed to prevent condensation and evaporation (*see* **Note 4**).
4. A supply of dry nitrogen (oxygen-free) is recommended for any CD measurement as this protects the optics, prevents the production of ozone, reduces the EHT on the photomultiplier, and improves the transmission below 220 nm. The supply may be from a gas bottle or a liquid nitrogen dewar; both systems should have an absorption filter before the optical unit (*see* **Note 3**).
5. Calibration of a CD instrument in millidegree ellipticity is achieved with a solution of (+)10-camphor-sulphonic acid at a concentration of 5.0 mg/ml. The concentration may be

accurately checked by measuring the absorption in a UV spectrophotometer using a molar extinction coefficient of 34.5 at 285 nm. Typically the solution above in a 1-cm path-length cell will have an absorption of 0.743 OD units at 285 nm, and this will give a corresponding ellipticity of 837.5 millidegrees at 290.5 nm in a 0.5-cm path-length cell in the CD instrument. (for calibration scale factor *see* **Note 2**).

6. In the procedure given later, we have made use of an Applied Photophysics pi* -180 spectrometer which is capable of giving reasonable results down to 180 nm. However, more modern machines (for example, the AP Chiroscan or J800 series from Applied Photophysics and Jasco, respectively) offer clear advantages in terms of sensitivity, stability, and extending detection into the far UV down to 163 nm (for setting up, *see* **Note 5**).
7. Temperature control of the sample is essential when comparative measurements are being made in which the instrument may be used for long periods of time, for example a titration or a kinetics experiment. Hence the usage of a Peltier cell with water cooling is an essential part of the spectrophotometer kit and should control the sample temperature ±0.05°C, with a temperature range of –20°C to +110°C.

3. Methods

For most proteins there is no significant CD spectrum between 250 and 300 nm compared to that seen for nucleic acids. Experiments involving the addition of protein can thus be conveniently carried out in this wavelength range, as described below. Below 250 nm both proteins and DNA have optical activity and any experiments here may require resolution of the spectrum into protein and DNA components.

1. To prevent damage to the optics, flush the instrument with nitrogen for 10–15 min before switching the lamp on. Continue to purge the instrument for the duration of the experiment (*see* **Note 3**).
2. Switch on the electronics first and then the lamp and allow the instrument to stabilize for 30 min. After warming up is complete switch on the Peltier temperature controller and set the temperature to 20°C; allow a further wait of 10 min before making any measurements.
3. Whilst waiting for the instrument to warm up, measure the UV spectrum of both the DNA and the protein; calculate the concentration of the stock solutions from their extinction

coefficients. The stock solution of protein should be at as high a concentration as possible to minimize corrections for dilution in subsequent titrations.

4. Measure the UV absorbance of the buffer in the cell to be used in the CD experiment against an empty cell (*see* **Note 4**). This is to check the buffer has little or no absorption in the wavelength region in which the measurements are to be made.
5. Using the data from **steps 3** and **4** determine the concentrations of DNA and protein that can be used in the experiment such that the total absorbance of all components including the cell is less than 1.0 OD unit.
6. Once the CD instrument has warmed up, use a clean 0.4-cm cell filled with water to determine the baseline between 230 and 360 nm (for instrument set-up *see* **Note 5**).
7. Clean the 0.4-cm cell (*see* **Note 6**) and calibrate the instrument using the previously prepared solution of camphor sulphonic acid and measure the CD between 250 and 320 nm. A 5-mg/ml solution in a 0.4-cm path-length cell has an ellipticity of 670 millidegrees at 290.5 nm at 20°C.
8. Take the cell set aside for the experiment, clean it, and fill with buffer. Place the cell in the instrument taking care to note the orientation of its faces in the beam. If using a cylindrical cell place it so that the neck of the cell rests against the side of the cell holder. Run a baseline between 230 and 360 nm, and signal average over nine scans to reduce the noise. This multiple scanning should show reproducible CD spectra and hence a constant condition has been achieved for recording the experiment.
9. Clean the cell and replace the buffer with the DNA solution and run the spectrum under the same conditions. When you remove the cell remember to place the cell back in the holder with the same face towards the light source as before and check the baseline position relative to the buffer in the region 360–320 nm.
10. Accurately pipette a small aliquot of the stock protein solution into the DNA in the cell and mix. Allow time for equilibration and measure the CD spectrum. The volume of the protein solution to be added will be determined by its concentration, the DNA concentration in the cuvette, and a rough idea of the expected stoichiometry. For initial experiments, the molar quantity of protein added at each step should be perhaps 10% of that of the DNA. For every aliquot of protein added the baseline must be checked for any variations and the problem must be resolved before the next addition is made (*see* **Note 5**).

11. Repeat the addition of protein to the DNA until no further changes in the CD are observed.
12. Thoroughly wash the cell and refill with buffer, and re-run the baseline. If the baseline has drifted from its initial value, this will have to be considered when analysing your results. If apparent drifting is greater than 0.2 millidegrees, *see* **Note 5**.
13. Plot the measured CD parameter at a particular wavelength against the concentration of protein added. The stoichiometry can be determined from the point at which a line drawn along the initial slope intersects that of the titration end point, which should be horizontal assuming dilution (if significant) has been corrected for (*see* **Figs. 1** and **2**).
14. Once the spectral changes and stoichiometry have been established, it is often useful to repeat the experiment with rather more titration points, using smaller aliquots of protein. This can be done at a single wavelength, without the need for scanning. Although the maximum CD signal from DNA is normally obtained around 275 nm, one should work at the wavelength that corresponds to the largest difference between free and bound DNA, which may well be different.

The CD spectrum of the DNA at saturation can be used to assess conformational changes that result from protein binding to a given sequence. Although analysis of the spectrum in terms of molecular structure is not straightforward (2), it may be possible

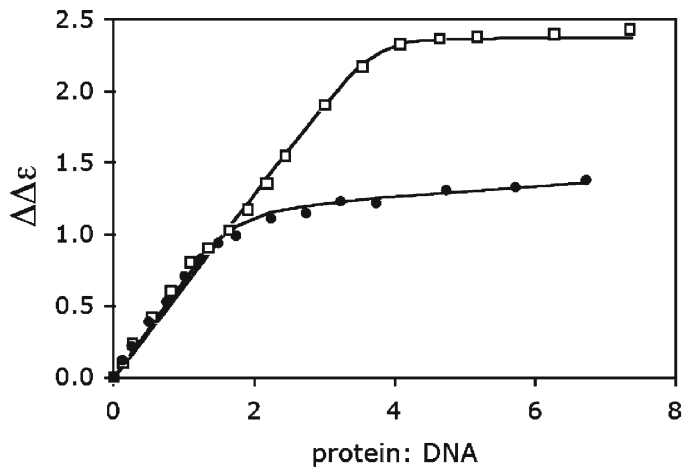


Fig. 2. Circular Dichroism titration curve of the binding of the protein C. AhdI to the native DNA sequence (*open squares*) and the mutated DNA sequence (*filled circles*); values of $\Delta\varepsilon$ were taken at the peak maxima wavelengths of the CD spectra (274.5 and 277.5 nm, respectively) shown in Fig. 1. Data were fitted to a two-step binding model with the binding affinity for the first site, $K_2 = 2 \times 10^7$ M^{-1} in both cases. For binding to the second site, the best fit was obtained with $K_3 = 3 \times 10^7$ M^{-1} for the native operator and $K_3 = 2 \times 10^4$ M^{-1} for the mutant operator, *see* ref. *16* for further details.

to interpret CD spectra of double-stranded DNA in terms of changes in helical twist angle (under-winding or over-winding of the helix).The spectral changes that accompany protein binding to different DNA sequences or with a variety of co-factors can also be informative.

4. Notes

1. For proteins and polynucleotides, molarity is often expressed in terms of moles of amino acid or nucleotide residues (respectively) for such calculations.
2. The instrument scaling factor (S) is the correction scaling factor when calibrating the instrument in millidegrees ellipticity. There is normally an internal scaling correction number that can be adjusted, so the instrument may be read directly in millidegrees ellipticity and hence $S = 1$ or simply use an S value to convert into millidegrees ellipticity. The mean residue weight (M_{mrw}) is simply the Molecular Weight of the protein or nucleic acid divided by the number of amino acids in a protein or the number of nucleotides in the nucleic acid, respectively.
3. High-intensity UV radiation converts oxygen to ozone which damages the optics. Failure to purge will lead to deterioration in instrument performance and be detrimental to the health of those working in the vicinity of the instrument. Flushing the optics with nitrogen gas can be done from a liquid nitrogen dewar (Taylor–Wharton XL-45) with the gas flow at 4 L/min; this gives 3 weeks of continuous purging and the instrument can then be purged for 3 days prior to measurements for the best results. Alternatively you can flush with a nitrogen (oxygen-free) bottle, which will last up to 2 days before a replacement is required.
4. For most experiments at wavelengths 230–300 nm, cells with path lengths of 0.4–1 cm can usually be used. The Hellma low strain 119.004F-QS 10 × 4 mm semi-micro with PTFE stopper is ideal for these measurements; it has a working start volume of 1,200 μl and has two path lengths (0.4 and 1.0 cm) dependent on its orientation to the beam. The stopper prevents evaporation during the experiment; plus it is easy to attach an internal temperature probe through the stopper to measure the actual sample temperature. Longer path-length cells for work with more dilute solutions are available at a price.
5. A number of instrument parameters must be set up before any spectrum run is initiated.These are (1) the entrance and exit slits which normally will be set between 1 and 2 nm for

protein or DNA work, (2) the wavelength start and finish values (say 360 and 230 nm), (3) the number of accumulations required, (4) the data collection settings, (5) temperature set value, and (6) nitrogen purging rate. The wider the slits are set, the more light enters the optical system and hence a better signal-to-noise ratio is achieved, but this results in poorer resolution. For instance, if the wavelength calibration of the instrument is set using the fine structure lines of the xenon lamp (460–490 nm region), the slits should be set to 0.1 nm as these line are extremely sharp and will only be resolved with a narrow setting. The wavelength range for protein, DNA, or RNA structure analysis is typically 360–180 nm, but it will vary within this range dependent on the experiment and the solvent conditions. CD measurements at specific wavelengths are sometimes used; for example, this is often the case with kinetic or temperature melting experiments, etc. Multiple scans or accumulations improve the signal-to-noise ratio – typically a baseline might require nine scans and a sample four scans. Data collection is dependent on how the manufacturer has designed the instrument, with the pi* -180 or Chiroscan from Applied Photophysics, we recommend collection at 0.5- or 1-nm intervals with the number of samplings set to 10,500, operating in conjunction with adaptive sampling, set error to ±0.01 and a maximum sampling of 500,000. For the Jasco 700–800 series, set the step resolution to 0.2 nm (this allows a maximum wavelength range of 400 nm), and a scanning speed of 20 nm/min and a response time of 2 s. If you wish to scan faster or slower, then the response time must be adjusted according to the manufacturer's recommendations. Software packages from both manufacturers allow for smoothing of data, baseline subtraction, zeroing offsets, and simple mathematical options to convert millidegrees into ellipticity per mole residue and compensate for dilution corrections, plus data formatted conversion options to allow the data to be manipulated within Microsoft Excel and other software packages with graphical outputs.

If significant drifting of the baseline has occurred, the experiment may have to be repeated. To minimize the effect of drifting, use two cells: one as a baseline reference and one containing sample. The baseline can then be standardized at each point in the titration. Remember that when swapping cells, it is vitally important to present the same section of the cell to the beam each time. Changes in baseline on a modern instrument with temperature control are in general the order of a fraction of a millidegree, and it is more likely the variations come from the cell not being position correctly each time when the cell has been removed or moved slightly when removing the stopper. Any baseline shifts can be observed in

the region 360–320 nm as there will be no protein or DNA component in this region; hence, it is best practice to scan this region for all measurements as this region should be unchanged.

To re-establish the original baseline, make sure the temperature of the cell solution is constant (20°C), the solution is well mixed, the cell is orientated as before and in the original position, and the cell walls are clean (re-wiping the cell faces is recommended to remove finger marks, etc.).

6. Clean the cell after a buffer or sample measurement, washing the cell continuously with high-quality water and then three times with acetone; wipe the outer surface of the cell before blow drying the cell.

Further Reading

Fasman, G.D. Editor (1996). *Circular Dichroism and the Conformational Analysis of Biomolecules.* Plenum, New York, pp. 1–738. ISBN 0-306-45142-5.

Hammes, G.G. (2005). *Spectroscopy for the Biological Sciences.* John Wiley **4**, 63–84. ISBN 13 978-047171344 9.

Berova, N., Nakarishi, K., and Woody, R.W. Editors (2000). *Circular Dichroism: Principles and Applications*, 2nd Ed. Wiley-VCH, New York, **21**, 601–615, **24–27**, 703–913, ISBN 13 978-047133005.

References

1. Compton, L.A., and Johnson, W.C., Jr. (1986). Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* **155**, 155–167.
2. Johnson, W.C., Jr. (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins: Struct. Funct. Genet.* 7, 205–214.
3. Lees, J.G., Miles, A.J., Wien, F., and Wallace, B.A. (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics.* **22**, 1955–1962.
4. Whitmore, L., and Wallace, B.A. (2004). DICHROWEB: on line server for protein secondary structure analysis from circular dichroism data. *Nucleic Acids Res.* **32**, W668–W673.
5. Johnson, B.B., Dakl, K.S., Tinoco, I. Jr., Ivanov, V.I., and Zhurkin, V.B. (1981). Correlations between deoxyribonucleic acid structural parameters and calculated circular dichroism spectra. *Biochemistry.* **20**, 73–78.
6. Basham, B., Schroth, G.P., and Ho, P.S. (1995). An A DNA triplecode: thermodynamic rules for predicting A and B DNA. *Proc. Natl Acad. Sci. U. S. A.* **92**, 6464–6468.
7. Scarlett, G.P., Elgar, S.J., Cary, P.D., Noble, A.M., Orford, R.L., Kneale, G.G., and Guille, M.J. (2004). Intact RNA-binding domains are necessary for structure-specific DNA binding and transcription control by CBTF122 during Xenopus development. *J. Biol. Chem.* **279**, 52447–52455.
8. Gray, D.M., Hung, S.H., and Johnson, K.H. (1995). Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes. *Methods Enzymol.* **246**, 19–34.
9. Hardin, C.C., Henderson, E., Watson, T., and Prosser, J.K. (1991). Monovalent cation induced structural transition in telomeric DNAs: G-DNA folding intermediates. *Biochemistry.* **30**, 4460–4472.
10. Carpenter, M.L., and Kneale, G.G. (1991). Circular dichroism and fluorescence analysis of the interaction of Pf1 Gene 5 protein with poly(dT). *J. Mol. Biol.* **27**, 681–689.

11. Kansy, J.W., Cluck, B.A., and Gray, D.M. (1986). The binding of fd Gene 5 protein to polydeoxyribonucleotides: evidence from CD measurements for two binding modes. *J. Biomol. Struc. Dynam.* **3**, 1079–1110.
12. Culard, F., and Maurizot, J.C. (1981). Lac repressor-lac operator interaction. Circular dichroism study. *Nucleic Acids Res.* **9**, 5175–5184.
13. Wartell, R.M., and Adhya, S. (1988). DNA conformational change in Gal repressor–operator complex: involvement of central G–C base pair(s) of dyad symmetry. *Nucleic Acids Res.* **16**, 11531–11541.
14. Torigoe, C., Kidokoro, S., Takimoto, M., Kyoyoku, Y., and Wada, A. (1991). Spectroscopic studies on lambda cro protein–DNA interactions. *J. Mol. Biol.* **219**, 733–746.
15. Conner, F., Cary, P.D., Read, C., Preston, N.S., Driscoll, P.C., Denny, P., et al. (1994). DNA binding and bending properties of the post-meiotically expressed Sry-related Protein Sox-5. *Nucleic Acid Res.* **22**, 3339–3346.
16. Papapanagiotou, I., Streeter, S.D., Cary, P.D., and Kneale, G.G. (2007). DNA structural deformations in the interaction of the controller protein C.Ahdl with its operator sequence. *Nucleic Acids Res.* **35**, 2643–2650.
17. Taylor, I.A., Davis, K.G., Watts, D., and Kneale, G.G. (1994). DNA binding induces a major structural transition in a type I methyltransferase. *EMBO J.* **13**, 5772–5778.
18. Pinhero, P., Scarlett, G.P., Rodger, A., Rodger, P.M., Murray, A., Brown, T., Newbury, S., and McClellan, J.A. (2002). Structures of CUG repeats in RNA. *J. Biol. Chem.* **277**, 35183–35190.
19. Calnan, B.J., Biancalana, S., Hudson, D., and Frankel, A.D. (1991). Analysis of the arginine-rich peptites from the HIV TAT protein reveals unusual features of RNA-protein recognition. *Genes Dev.* **51**, 201–210.
20. Tan, R., and Frankel, A.D. (1995). Structural variety of arginine-rich RNA-binding peptides. *Proc. Natl Acad. Sci. U. S. A.* **92**, 5282–5286.
21. Griko, Y., Sreerama, N., Osumi-Davis, P., Woody, R.W., and Woody, A.Y. (2001). Thermal and urea unfolding in T7 RNA polymerase: calorimetry, circular dichroism and fluorescence. *Protein Sci.* **10**, 845–853.

Chapter 37

Defining the Thermodynamics of Protein/DNA Complexes and Their Components Using Micro-calorimetry

Colyn Crane-Robinson, Anatoly I. Dragan, and Christopher M. Read

Summary

Understanding the forces driving formation of protein/DNA complexes requires measurement of the Gibbs energy of association, ΔG, and its component enthalpic, ΔH, and entropic, ΔS, contributions. Isothermal titration calorimetry provides the enthalpy (heat) of the binding reaction and an estimate of the association constant, if not too high. Repeating the ITC experiment at several temperatures yields ΔC_p, the change in heat capacity, an important quantity permitting extrapolation of enthalpies and entropies to temperatures outside the experimental range. Binding constants, i.e. Gibbs energies, are best obtained by optical methods such as fluorescence at temperatures where the components are maximally folded. Since DNA-binding domains are often partially unfolded at physiological temperatures, the ITC-observed enthalpy of binding may need to be corrected for the negative contribution from protein refolding. This correction is obtained by differential scanning calorimetric melting of the free DNA-binding domain. Corrected enthalpies are finally combined with accurate Gibbs energies to yield the entropy factor ($T\Delta S$) at various temperatures. Gibbs energies can be separated into electrostatic and non-electrostatic contributions from the ionic strength dependence of the binding constant.

Key words: Thermodynamics, Calorimetry, Protein/DNA complex, HMG box, AT-hook.

1. Introduction

The many detailed, though static, structures of protein–nucleic acid complexes have revealed the underlying structural determinants of the binding process: a high interface complementarity and a precise orientation of interacting groups. However, another important question is *what drives* the molecules to interact with each other. To answer this question we have to rationalize a structure

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_37

energetically, and this requires a deconvolution of the relatively large negative value of the free energy of association (ΔG) into its enthalpic (ΔH) and entropic (ΔS) contributions. Only then can one appreciate the true nature of the forces that drive the interaction of protein and nucleic acid.

Proteins and nucleic acids typically behave co-operatively and often undergo structural rearrangements on association. These changes range from subtle adjustments of dihedral angles to rearrangements in the orientation of domains and even the refolding of entire binding domains. Many DNA-binding domains are either very flexible, or partly unstructured or even fully unfolded in the free protein and become folded only when bound to the specific DNA target site *(1)*. The DNA itself sometimes undergoes large structural deformation upon complex formation, e.g. bending *(2–6)*. In energetic terms, the view that binding specificity is simply the accumulation of specific favourable interactions between rigid binding partners is frequently not the case and conformational changes in the two components, as well as solvent exclusion and rearrangement in the binding site play an important role. There is thus a complicated energetic profile involving changes in ΔG, ΔH, and ΔS on going from the free components to the final complex.

This chapter focuses on the practical aspects of measuring the equilibrium energetics of associating DNA-binding domains (DBDs) with short DNA duplexes. As examples we will largely consider the HMG boxes from the sequence-specific (SS) protein hLEF-1 and the non-sequence-specific (NSS) HMG-D from Drosophila, in order to illustrate the importance of structural changes that accompany complex formation and the relative roles played by electrostatic and non-electrostatic forces in the binding process *(7, 8)*. HMG boxes contain three α-helices arranged into an L-shaped fold. NMR-derived structures of the HMG box of human LEF-1 complexed to 15-bp DNA *(5)* show that the domain binds into a widened minor DNA groove, generating a large bend with considerable base unstacking, together with the partial intercalation of a single sidechain from the protein. A crystal structure of a minimal domain of HMG-D in complex with DNA *(6)* reveals a very similar folded complex but with *two* partial sidechain intercalations from the protein into the DNA.

Since the structures of the interacting components, in particular the free proteins, depend significantly on temperature *(7, 8)* the temperature dependence of the enthalpy of binding, as measured by isothermal titration calorimetry (ITC), can no longer be simply interpreted as being solely due to a unique heat capacity increment, ΔC_p. The combined use of ITC with differential scanning calorimetry (DSC) is necessary for studying the energetics of protein–nucleic acid complexes, since the DSC data are essential for correcting the data obtained by ITC.

1.1. The Energetics of Biomolecular Interactions

To characterize the thermodynamics of a binding reaction, it is necessary to determine the association Gibbs free energy, ΔG, and its enthalpic and entropic components, ΔH and ΔS, at a given reference temperature. The heat capacity increment on binding, ΔC_p, is a fundamental thermodynamic quantity which reflects the nature of forces involved in the association process (polar, non-polar) and contains information about the conformational changes in the free components on association *(9, 10)*. Furthermore, ΔC_p is required to predict the change of these three quantities with temperature, according to the general thermodynamic relationships, thereby allowing values to be obtained at experimentally inaccessible temperatures.

For a simple 1:1 interaction of protein P with DNA, D: $\mathrm{P} + \mathrm{D} \Leftrightarrow \mathrm{PD}$ the energetics of equilibrium binding can be described by the well-known relationship:

$$\Delta G_{\mathrm{A}} = -RT \ln K_{\mathrm{A}} = \Delta H - T\Delta S, \qquad (1)$$

where ΔG_{A} is the free energy of binding, K_{A} is the association constant, ΔH is the enthalpy change (heat), and ΔS is the entropy change on binding. The Gibbs energy of binding is calculated from the association constant, which is best obtained by optical methods, e.g. fluorescence spectroscopy (*see* later and Chapter "Fluorescence Spectroscopy and Anisotropy in the Analysis of DNA–Protein Interactions"). The enthalpy is obtained from ITC, after correction for refolding of the components on forming the complex. The entropy is then calculated by difference.

Measurement of the ionic strength dependence of the association constant enables the separation of K_{A} (i.e. ΔG_{A}) into its electrostatic and non-electrostatic components. On association of protein with DNA, the formation of ion pairs between cationic residues in the protein and the DNA polyanion results in the release of counterions, the mixing of which with the ions in the bulk solution produces significant entropy increase *(11)*. At relatively low salt concentrations in aqueous solution this entropy effect is simply proportional to the number of released counterions *(12)*. Correspondingly, the logarithm of the association constant of protein with DNA can be presented in just two terms:

$$\log(K_{\mathrm{A}}) = \log(K_{\mathrm{A}})_{\mathrm{ncl}} - N \times \log[\mathrm{KCl}]. \qquad (2)$$

The first term on the right-hand side of **Eq. 2** results from the nonelectrostatic (nel) interactions between the DNA and protein, and the second term results from the electrostatic effects associated with the release of counterions. The slope of this function, N is equal to $(Z\Psi + \beta)$, where Z is the number of DNA phosphates interacting with the protein, Ψ is the number of cations per phosphate released upon protein association, and β is the number of anions displaced from the protein on DNA binding. For short

DNA duplexes Ψ has been empirically determined as 0.64 *(13)* and the product $Z\Psi$ is much greater than β. The number of ionic contacts made on forming a complex can thus be determined from the value of the slope of the linear relationship of $\log(K_A)$ with log[KCl]. It should be noted that counterions are released mostly from the DNA since the low charge density on protein surfaces does not attract a tight coat of counterions *(14)*.

1.2. The ITC Experiment

Modern mixing micro-calorimeters allow precise measurements of the enthalpy of binding, ΔH_{app} between 2 and 80°C. The basic principle is that one binding partner (e.g. the protein, P) is titrated at a constant temperature into a known amount of the other binding partner (DNA, D) placed in the sample cell of the calorimeter. Heat is produced or released when binding occurs. The instrument measures the electrical power (in units of J/s or Watts) required to maintain zero temperature difference between the sample cell and the reference cell, the latter filled with buffer. The contents of the sample cell are stirred to allow rapid mixing and effective heat transfer over the surface of the cell.

If K_A is high and the degree of saturation is still low, electrical power peaks of similar magnitude appear in the thermogram subsequent to each injection, *i*. Integration of a peak with respect to time yields the apparent heat change, $\Delta q_{i,app}$. As the fractional saturation increases, $\Delta q_{i,app}$ gradually decreases and eventually all binding sites become saturated (*see* **Fig. 1**). Small non-specific heat effects, $\Delta q_{i,ns}$ observed after complete saturation may be caused by the heat of protein dilution, by an imperfect match between the buffer composition of the protein solution and the DNA solution, or by other non-specific effects. To eliminate these factors, a separate titration of protein into buffer is carried out and the observed heat, $\Delta q_{i,ns}$ subtracted from the titration with DNA in the cell (*see* **Subheading 3.2**). $\Delta q_{i,app}$ is proportional to the known volume of the calorimetric cell, V_{cell}, to the change in concentration of the bound protein $\Delta[P]_{i,bound} = [P]_{i,bound} - [P]_{i-1,bound}$, and to the apparent molar enthalpy of association ΔH_{app}. Thus ΔH_{app} at temperature T may be calculated from,

$$\Delta q_{i,app} = \Delta q_i + \Delta q_{i,ns} = \Delta[P]_{i,bound} V_{cell} \Delta H_{app}. \qquad (3)$$

As in other binding experiments, the dissociation constant, $K_D(=1/K_A)$, can only be reliably measured in titrations covering a fairly narrow concentration range. If the concentration of binding sites is much larger than K_D, the binding isotherm is of rectangular shape with a sharp step at saturation. In the case of the binding-site concentration being much smaller than K_D, the binding isotherm is very shallow and is of limited use for calculating the dissociation constant and the enthalpy. In practice, values of the ratio of the binding-site concentration to K_D between 10 and 100 represent

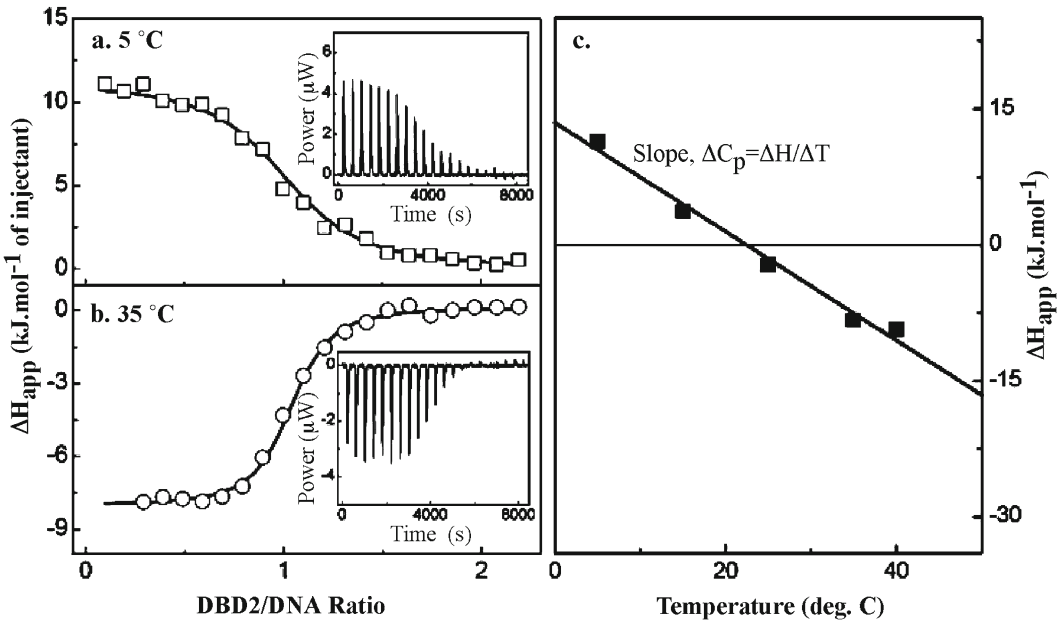


Fig. 1. (**a–b**) ITC titration of the 17-residue AT-hook peptide DBD2 into a 12-bp target DNA (400 μM DNA^{AT}) in 100 mM KCl, 10 mM phosphate buffer (pH 6.0), at two temperatures showing that binding occurs with heat effects of opposite sign. The *insets* give the actual calorimetric traces. The enthalpies of association, ΔH_{app}, were derived by integration of each peak in the titration. The *curves* drawn through the data points correspond to 1:1 stoichiometry fits with binding constants of about 10 μM at both temperatures. (**c**) Plot of the temperature dependence of the binding enthalpy: the *slope* represents the change in heat capacity on binding DBD2 to the DNA.

the optimal window to obtain precise values of the dissociation constant. This is the case for the binding of the AT-hook peptide to DNA^{AT}, as measured at two temperatures (**Fig. 1**): the DNA concentration is 400 μM and the calculated K_D values are ~10 μM, i.e. the ratio is about 40. Unfortunately, this window is not always accessible to ITC. For very low K_D values (tight binding), the optimal concentrations are so low that the released heat is below the sensitivity of the instrument, the specific enthalpy of biomolecular interactions not being very large. This effectively limits the measurement of K_D using ITC to values above about 10^{-7} M (ΔG approx. –40 kJ/mol at 25°C). Alternatively, if K_D is high (weak binding), then very high concentrations are required and the heat of binding may be obscured by aggregation effects.

Two types of ITC measurement are typically made. If the binding is relatively weak, a full titration at a DNA concentration 10–100 times the estimated value for K_D is carried out to obtain n, K_A, and ΔH_{app} with the injection of up to a 2–4 M excess of protein. If the binding is tight, then just a direct measurement of ΔH_{app} is carried out at DNA concentrations much greater than K_D and with the total number of moles of added protein being less than the number of moles of binding sites available, to ensure that virtually all protein becomes bound to the DNA. Each

injection of protein, $\Delta[\mathrm{P}]_{i,\mathrm{total}}$ then produces an approximately equal change in heat, $\Delta q_{i,\mathrm{app}}$. If the overall fractional saturation is still low after completion of a series of injections, then the average may be taken and **Eq. 3** becomes

$$\frac{1}{m}\sum_{i=1}^{m}\Delta q_{i,\mathrm{app}} = \frac{1}{m}\sum_{i=1}^{m}\left(\Delta q_i + \Delta q_{i,\mathrm{ns}}\right)$$
$$= \frac{1}{m}\sum_{i=1}^{m}\left(\Delta[\mathrm{P}]_{i,\mathrm{bound}} V_{\mathrm{cell}} \Delta H_{\mathrm{app}}\right), \qquad (4)$$

where m is the number of injections. If low amounts of protein are injected, aggregation effects and non-specific binding are also minimized.

In general $\Delta C_{p,\mathrm{app}}$ the apparent heat capacity change of association may be calculated from titrations at a number of temperatures. Under ideal circumstances the plot of the apparent enthalpy ΔH_{app} versus temperature will be linear with a slope, $\Delta C_p = \Delta H / \Delta T$, but all too often this is not the case.

1.3. The DSC Experiment

In a DSC experiment, the heat capacity of a macromolecule is measured as a function of temperature. A sample cell containing the macromolecule(s) of interest in buffer and a reference cell containing buffer are electrically heated at a known constant rate. At a temperature-induced transition (typically endothermic), the temperature of the sample cell will lag behind that of the reference cell. An electrical feedback mechanism is used to maintain the reference cell at the same temperature as the sample cell. This amount of compensatory electrical power (in units of J/s) at temperature T divided by the heating rate is the apparent difference in heat capacity between the cell containing the sample and the reference cell, $\Delta C_{p,\mathrm{app}}(T)$ (in units of J/K). Because the sample cell contains a smaller volume fraction of buffer as compared to that in the reference cell, the partial molar heat capacity of the dissolved macromolecule $C_{p,}\phi(T)$ at temperature T (with units of J/mol/K) is given by

$$C_{p,\phi}(T) = \frac{C_{p,\mathrm{buffer}}(T)V_{\phi}^{0}}{V_{\mathrm{buffer}}} - \frac{M\Delta C_{p,\mathrm{app}}(T)}{m}, \qquad (5)$$

where $C_{p,\mathrm{buffer}}(T)$ and V_{buffer} are the partial molar heat capacity and molar volume of buffer, respectively. V_{ϕ}^{0} and M are the partial molar volume and the molar mass of the macromolecule, respectively, and m is the mass of macromolecule in the sample cell.

The excess molar heat capacity function is the heat absorbed in the melting transition above that of the intrinsic heat capacity of the macromolecule. Integration of the excess molar heat capacity function with respect to temperature yields the enthalpy of the melting transition:

$$\Delta H = \int_{T_1}^{T_2} \langle C_p(T) \rangle \mathrm{d}T. \quad (6)$$

The intrinsic heat capacity of the initial (folded) and final (unfolded) states of the macromolecule must be estimated within the melting transition. For monomeric proteins, the heat capacity of the folded state is an approximate linear function of temperature *(15, 16)* whilst the unfolded state is a shallow parabolic function of temperature *(15)*.

The relationship of the partial molar heat capacity to the excess molar heat capacity functions is schematically shown in **Fig. 2**.

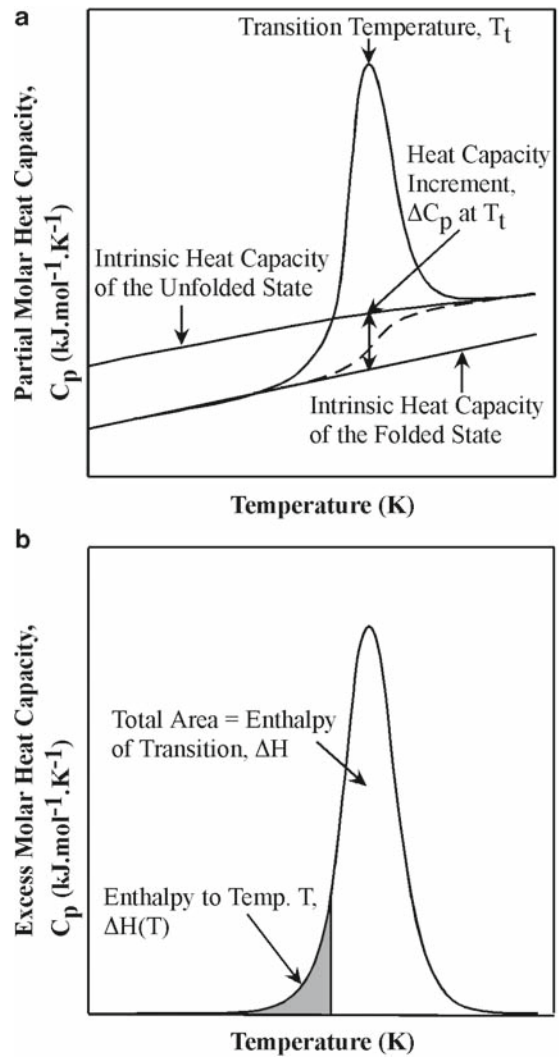


Fig. 2. (**a**) Schematic representation of the partial molar heat capacity function as would be observed for the melting of a single domain monomeric protein. Over the transition region the intrinsic heat capacity function is an interpolation of the folded and unfolded states, weighted in proportion to their relative contributions (*dashed line*). (**b**) replots the C_p/T function above that of the intrinsic heat capacity of the system. This is known as the excess molar heat capacity function.

2. Materials

2.1. Calorimeters

Differential scanning calorimetric experiments were usually carried out on a Nano-DSCIII calorimeter (Calorimetry Sciences Corp./TA Instruments). The instrument's performance and data acquisition are detailed in refs. 17–19. In brief, the instrument operates over the temperature range –20 to 130°C in both heating and cooling regimes. The sample and reference calorimetric cells are of approx. 330 μL volume, with a capillary design for effective heat transfer. Control of the calorimeter and scan acquisition is achieved via the supplied DSC_Acquisition program running on a PC computer connected to the calorimeter.

It is important to note that the scanning calorimeter must not only be precise as regards its absolute determination of C_p in both the heating and the cooling modes but must also have a very stable and reproducible baseline. High stability is achieved only by continuous temperature cycling of the instrument, and this means changing samples at around room temperature without stopping the cycling mode.

Isothermal titration calorimetry was performed on a Nano-ITCIII titration calorimeter (Calorimetry Sciences Corp./TA Instruments) which can operate between 2 and 80°C. The cells are of cylindrical design having approx. 1 mL volume and incorporate a stirrer that can be set to a precise speed. Control of the calorimeter and scan acquisition is achieved via the supplied ITC_Run software. Peak integration, baseline subtraction, and curve fitting were performed with the software, BindWorks and Origin. Technical details on the construction of mixing microcalorimeters, their performance and sensitivity, and on the theory of data analysis have been described elsewhere *(20–23)*.

2.2. Reagents and Solutions

2.2.1. Calorimetry

1. For calorimetric experiments it is important to ensure that the protein and DNA samples are homogeneous, because even a few percent of contaminating species might cause significant heat effects. Homogeneity of the samples is most reliably verified by mass spectrometry.
2. Concentrations must be precisely determined so as to obtain accurate thermodynamic quantities (*see* **Subheading 3.6**). The preparation of equimolar mixtures (of complementary oligonucleotides or HMG box/DNA complexes) can then be made precisely, without further purification. Dilutions are made by weight on a precision balance.
3. The DSC and ITC experiments normally used a standard buffer of 100 mM KCl, 10 mM potassium phosphate (pH 6.0), 1 mM EDTA (*see* **Note 1**). All buffer solutions were pre-

pared from the highest-quality reagents and ultrapure water. For ITC experiments the solutions were filtered through a 0.45-μm membrane and thoroughly degassed prior to use.

2.2.2. Preparation of HMG Box/DNA Complexes and Components

1. The 14–16-bp DNA duplexes used here are specific for the binding of human LEF-1 (DNA^{Lef}) which contains the cognate-binding site TTCAAA, mouse Sox-5 (DNA^{Sox}) containing the binding site AACAAT, human SRY (DNA^{Sry}) containing the binding site CACAAA and a non-specific 16-bp DNA duplex (DNA^{NS}) containing ATAT at its centre *(8)*. For the AT-hook domain 2 from human HMGA1 a 12-bp DNA duplex (DNA^{AT}) containing AAATT at its centre was used.
2. DNA oligonucleotides were synthesised using phosphoramidite chemistry and purified on a Mono Q HR16/10 column fitted to a Pharmacia FPLC system, eluting with a linear 0.1–1.0 M NaCl gradient in 10 mM Tris–HCl, 1 mM EDTA, 20% (v/v) acetonitrile (pH 7.0). Oligonucleotide fractions were precipitated with 3 volumes of ethanol at –20°C overnight, redissolved in water, and then extensively dialysed using Spectrapor tubing (molecular-weight cut-off: 1,000 Da) against three changes of 1 L of standard buffer at 4°C.
3. DNA duplexes were prepared by mixing equimolar amounts of the complementary oligonucleotides in standard buffer and annealed by heating to 95°C in a water bath followed by slow cooling to 4°C over a period of approx. 4 h. DNA duplexes were then extensively dialysed using 3,000 Da molecular-weight cut-off tubing against three changes of 1 L standard buffer at 4°C.
4. The sequence of human LEF-1 encoding amino acids 285–384, with cysteine 321 mutated to serine (LEF86), was expressed using a pET vector, in *E. coli* BL21 (DE3)-RIL cells. The insoluble protein was purified essentially as described in *(24)* and followed by reverse-phase HPLC. A shortened version of LEF-1 encoding amino acids 298–374, with cysteine 321 mutated to serine, LEF79, was expressed as a soluble fusion protein in pGEX-2T, using *E. coli* BL21 (DE3) plysS cells. After affinity purification on glutathione-agarose and thrombin cleavage whilst still attached to the column *(7)*, it was purified by reverse-phase HPLC.

The DNA sequence encoding the entire drosophila HMG-D protein, HMG-D100, and a 25-residue C-terminal shortened version, HMG-D74, was expressed using a pET vector, in *E. coli* BL21 (DE3) cells. Briefly, the soluble proteins were purified by ammonium sulphate fractionation, DEAE- and SP-ion exchange chromatography, and finally by reverse-phase HPLC *(7)*.

Dried purified proteins were dissolved in water and refolded by extensive dialysis against three changes of 1 L of standard buffer at 4°C. Electrospray mass spectrometry indicated that all the expressed proteins lacked an initiator methionine residue.

5. The HMG box/DNA complexes were prepared by mixing equal volumes of the components at equimolar concentration in standard buffer at 4°C. Protein was added in 10% aliquots, at 5-min intervals, to the DNA. The complex was then extensively dialysed using 3,000-Da molecular-weight cut-off tubing against three changes of 1 L standard buffer at 4°C.

2.2.3. DNA Concentrations

1. 100 mM Tris–HCl (pH 8.0).
2. Snake venom phosphodiesterase I (PDE1, from *Crotalus durissus terrificus*, Sigma).

3. Methods

3.1. Differential Scanning Calorimetry

1. Prior to starting experiments ensure that the Nano-DSC calorimeter sample and reference cells are at room temperature, thoroughly cleaned, and thermally balanced.
2. Fill both sample and reference cells with standard buffer (100 mM KCl, 10 mM potassium phosphate, 1 mM EDTA (pH 6.0)) at room temperature (*see* **Note 2**). It is most important that no air bubbles are present in the calorimeter cells. The presence of air bubbles will yield incorrect values for the partial molar heat capacity of the macromolecule, because the volume of solution in the sample and reference cells will not be identical (assumed in **Eq. 5**). Furthermore, as heating proceeds the expansion of even the smallest of air bubbles will result in significant heat output (and vice versa on cooling).
3. Seal the top of the chamber with the screw-threaded piston and let the calorimeter settle to thermal equilibrium. Apply approx. 3 atm. of over-pressure to the cells by screwing the piston down (*see* **Note 3**).
4. Scan to obtain a buffer–buffer baseline. Initially, a number of scans are recorded in order to produce a number of reproducible baselines (to within 0.1 μW, in the linear region). A complete cycle of heating and cooling at a rate of 1 K/min actually takes about 4 h. because at the end of each heating or cooling scan, there is a period of thermal equilibration.
5. Remove the buffer from the sample cell. This is most easily accomplished with a vacuum line attached to a water pump.

6. Slowly and carefully apply a 1.0 mL volume of dialysed sample to the cell, using the method outlined in **Note 2**. The excess sample is put to good use: firstly to determine sample concentration and secondly, it is diluted (using the final dialysate) for subsequent calorimetric runs at lower concentrations. Again apply an overpressure of approx. 3 atm. to the cells.
7. Perform one heating and cooling scan to obtain the denaturation and renaturation curves of the sample buffer. Usually samples were heated from 0 to 80°C and then cooled to –8°C, both at a rate of 1 K/min. The heat effects of unfolding and refolding should appear as mirror images, slightly shifted in temperature due to the slower kinetics of refolding. Reproducibility demonstrates that the heat-induced unfolding/refolding is reversible. However, this is not always the case (*see* **Note 4**).
8. Remove the sample from the cell. Thoroughly wash out the cell with buffer and then perform another scan of buffer. A heat-absorption peak may be observable if the previous sample is not completely removed (*see* **Note 5**).

3.1.1. DSC Data Analysis

The subtraction of buffer scans to obtain an accurate baseline, the conversion to partial molar heat capacities, and the deconvolution of the excess molar heat capacity function into separate transitions were all performed using the CPCalc program. CPCalc provides a simple interactive mechanism, based on Data Exchange Ports, for the parsing of data from one step to another.

1. The functions of differential power and temperature versus time are extracted from the DSC data acquisition files. Each file may contain more than one scan, but further analysis is on a one-scan basis.
2. A matching buffer–buffer scan is subtracted from a sample-buffer scan. Both must be from a heating (or cooling) scan. Since the absolute values of the molar heat capacities depend on good baseline subtraction and the buffer–buffer scans vary to a small extent, it is important that a number of subtractions are tried using different scans from the acquisition file.
3. The subtracted compensatory power curve is converted to the partial molar heat capacity function by use of **Eq. 5**. Accurate values for the concentration (**Subheading 3.6**), molecular mass, and partial specific volume of the macromolecule are required (*see* **Note 6**).
4. Appropriate functions for the intrinsic heat capacity of the folded and unfolded states (*see* **refs. 15, 16, 19**) may then be introduced onto the observed partial molar heat capacity function to generate the 'background' heat capacity function (**Fig. 2a**). The excess heat capacity function above this 'background' then yields the calorimetric ΔH, the melting temperature, and ΔS at this temperature (**Fig. 2b**).

5. If the melting profile of the sample is complex and contains several overlapping transitions, their individual energetic parameters can be obtained by deconvolution of the molar heat capacity function as shown in **Fig. 3a**.

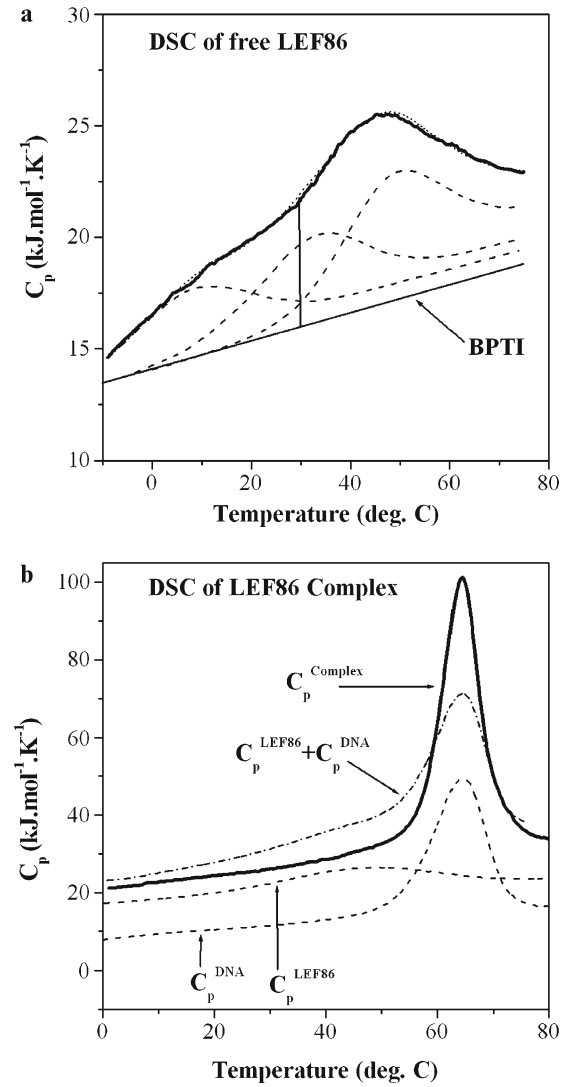


Fig. 3. (**a**) Deconvolution of the partial molar heat capacity function for free LEF86, determined over the temperature range from −10°C (i.e. a super-cooled solution) to 75°C (*solid line*). Three transitions are observed (*dashed lines*) and their sum is shown as a *dotted line*. The baseline represents the partial molar heat capacity of the fully folded protein BPTI, corrected to the molecular weight of LEF86, together with an added contribution from the seven residues of the unfolded C-tail of LEF86. The vertical line at 30°C defines the wedge-shaped area that corresponds to the total heat absorbed from the start of melting (at −10°C) up to 30°C. This heat is the correction to be added to the ITC-observed heat of binding at 30°C (*see* **Fig. 4**). (**b**) Partial molar heat capacity functions of free LEF86 (C_p^{LEF86}), the free DNA duplex (C_p^{DNA}), and the LEF86/DNA complex ($C_p^{complex}$). The *dotted-dashed line* shows the summed heat capacities of free DNA and protein ($C_p^{LEF86} + C_p^{DNA}$).

Figure 3a shows the partial molar heat capacity function of the free LEF86 DBD, together with a baseline representing the C_p/T function of a 'standard fully folded protein', BPTI, the function is corrected for the difference in molecular weight between BPTI and LEF86. It can be seen that even at –8°C there is still a small residual unfolded component in LEF86: this largely represents the unstructured C-terminal tail that constitutes 9% of the total residues. The complete melting profile of LEF86 can be deconvoluted into the cooperative melting of three components and whilst their interpretation in terms of sub-domain structure is still unclear, a high degree of structural disorder in the LEF-1 HMG box at room temperature has been amply demonstrated by NMR *(24)*.

3.2. Isothermal Titration Calorimetry

Before a series of experiments, the calorimeter was calibrated either by applying electrically generated heat pulses or by standardised chemical reactions (e.g. the protonation of tris[hydroxymethyl] aminomethane or the binding of Ba^{2+} to 18-crown-6 ether *(25)*). It is recommended to install the calorimeter in an air-conditioned room and to equilibrate it at a set temperature overnight since this substantially improves baseline stability. The speed of the stirring paddle both equilibration and experiment was usually 200 rpm (*see* **Note 7**).

1. Sample and reference cells are first rinsed with dialysis buffer. The reference cell is then filled with buffer and the sample cell filled with the DNA solution. The system is heated to the working temperature and equilibrated until the differential power signal levels off. The injection syringe containing protein is inserted into the sample cell, stirring is initiated, and the baseline is established over a period of 1–2 h. Typically, a baseline drift (differential power signal drift) of ~40 nW/h and an rms noise of ~4 nW indicates complete thermal equilibration of the system under stirring (*see* **Note 8**).
2. The experiment is started with a small injection of 1–2 μL. The reason for this is that during the long equilibration period, diffusion through the injection port of the syringe occurs, thus causing a change in protein concentration near the syringe needle tip. The proper injection schedule is then executed. To measure the enthalpy of LEF86 binding to DNA, a further nineteen injections each of 10 μL and of ~10 s duration were performed, with a 2.5 min interval between injections (*see* **Note 9**). A typical thermogram is depicted in **Fig. 4a**.
3. After completion of the experiment the cells are thoroughly cleaned. Cleaning of the calorimetric cells and filling syringes follows a standard laboratory protocol (*see* **Note 10**).

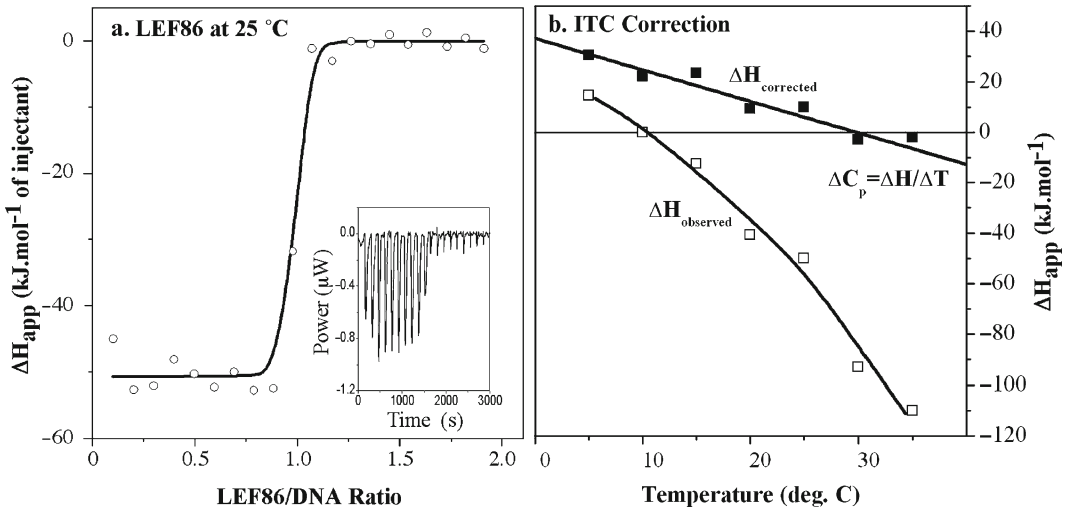


Fig. 4. (**a**) ITC titration of LEF86 into a 15-bp target duplex DNA[Lef] at 10 μM in 100 mM KCl, 10 mM phosphate buffer (pH 6.0) at 25°C. The *inset* gives the actual calorimetric trace. The enthalpy of association, ΔH_{app}, was derived by integration of each peak in the titration. The *curve* drawn through the data points corresponds to a 1:1 stoichiometry fit with a binding constant of ~0.1 nM. (**b**) The ITC-measured enthalpies of binding LEF86 to DNA[Lef] plotted as a function of temperature. The observed binding enthalpies ($\Delta H_{observed}$) were then corrected for protein refolding ($\Delta H_{corrected}$) using fractional proportions of the C_p/T function (*see* **Fig. 3**). The slope of the $\Delta H_{corrected}/T$ function represents the heat capacity change, ΔC_p, on binding fully folded protein to the DNA.

4. The cells may then be filled with buffer and **step 2** repeated. Injections of protein into buffer yields the heat associated with protein dilution and other non-specific effects. Because the heats obtained in **steps 2** and **4** are directly subtracted in the data analysis, it is crucial that *(1)* the blank titration is performed at exactly the same temperature as the main experiment, *(2)* the same protein solution and dialysis buffer are used, and *(3)* an identical injection scheme is executed.
5. The titrations are performed at a number of temperatures to collect data on the temperature dependence of the binding enthalpy, *see* **Fig. 4b**.

3.2.1. ITC Data Analysis

If a complete binding isotherm has been recorded over an optimal concentration range (as in the AT-hook case, *see* **Fig. 1**), the data can be subjected to a nonlinear least-squares analysis to obtain a full set of parameters (ΔH_{app}, K_A and n) according to **Eq. 3**. In the case of LEF86 binding to DNA (**Fig. 4a**), the experiments were designed to measure only the enthalpy and thereby the heat capacity change. Thus integration of the differential power peaks (*see* **Note 11**) collected as in **step 2** of **Subheading 3.2** yielded $\Delta q_{i,app}$ and the non-specific heats $\Delta q_{i,ns}$ are obtained by integration of the peaks collected as in **step 4** of **Subheading 3.2**. Under

conditions of total association, $\Delta[P]_{i,bound}$ equals $\Delta[P]_{i,total}$ and is simply calculated from the known concentration of protein in the injection syringe, the volume of each injection and the volume of the cell, V_{cell}. The enthalpy of association, ΔH_{app} at temperature T may then be calculated from **Eq. 4.** The procedure is repeated at several temperatures.

The experimentally observed enthalpies (ΔH_{app}) can then be plotted against temperature, as illustrated for LEF86 binding to DNA (**Fig. 4b**). It can be seen that the temperature dependence of the binding enthalpy is very non-linear. Although its negative slope is a sign that substantial dehydration of apolar surface occurs on forming the complex, the strong curvature suggests that ΔC_p is very temperature dependent over a narrow range. The curvature is in fact due to the components, in particular the protein, refolding on forming the complex, with consequent heat release: this is an extra negative contribution to the observed binding enthalpy. Since the degree to which the protein is unfolded increases as the temperature of the ITC experiment rises, this negative contribution to ΔH increases with temperature and the $\Delta H/T$ plot is consequently non-linear.

3.3. Correction of ITC-Derived Enthalpies for Partially Unfolded Components

Correction of the observed enthalpy of binding for the negative contribution of protein unfolding can be achieved if an accurate determination is made of the thermal unfolding of the DBD using DSC (**Subheading 3.1**). For an ITC experiment carried out at 30°C, the heat of refolding of LEF86 is given by the wedge-shaped area under the C_p/T function, as obtained by DSC, from 30°C down to approx. –10°C, the temperature in which LEF86 is fully folded (*see* **Fig. 3a**). At approx. –10°C the C_p/T function of LEF86 coincides with the C_p/T function of BPTI, the 'standard fully folded protein'. For each of the temperatures used in the ITC measurements the heats of refolding are obtained by calculating the area under the C_p/T function (from DSC), and this is then a positive correction to the ITC-measured enthalpies of association to yield corrected heats (*see* **Fig. 4b**). At 30°C the heat of refolding is +90 kJ/mol. Since the total enthalpy of LEF86 unfolding is 265 kJ/mol (from DSC), the domain could be said to be 34% unfolded at this temperature.

Is there any unfolding of the DNA or even of the complex at the temperatures used for ITC, for which correction also needs to be made? NMR studies of short duplexes show that the two terminal base pairs are very mobile, but this flexibility does not extend to the penultimate base pairs. The duplexes used for the formation of complexes with DBDs are usually slightly longer than the contact length of the protein (to ensure that no contacts are omitted), so the terminal base pairs are not constrained in the complex, i.e. there is no change in their state on complex formation. No positive correction for DNA refolding is normally

needed therefore. In regard to the complex, for ITC measurements made at temperatures as high as 35–40°C there could be some temperature-induced unfolding of the complex: a negative correction would then need to be made. This correction can be determined by calculating the area under the C_p/T function of the complex, obtained from DSC, as described earlier for the free protein.

A simpler approach in which all three possible corrections are made simultaneously is illustrated for LEF86, *see* **Fig. 3b**. At the lowest temperatures, where all three elements (free protein, free DNA, and complex) are virtually fully folded the observed C_p/T function of the complex is slightly less than the summed heat capacities of free protein and DNA, meaning that formation of fully folded complex from fully folded components is accompanied by a small reduction in heat capacity (i.e. a negative ΔC_p – as measured with much greater precision from the corrected ITC data). At temperatures of ~35°C and below (a typical ITC temperature range), there is an increasing interval between the observed and the 'summed' C_p/T functions for the complex: this wedge-shaped area (from the lowest temperature up to the temperature of an ITC measurement) represents the net enthalpy of changes in all three elements on forming the complex and is the correction to be applied to the ITC-observed enthalpy. In practice, this correction comes almost entirely from protein refolding and can usually be taken from the C_p/T function of the protein alone (e.g. **Fig. 3a** for LEF86). When this (substantial) correction is made, the $\Delta H/T$ function is no longer curved (*see* **Fig. 4b** for LEF86), implying that ΔC_p is largely temperature independent over this narrow temperature interval. For LEF86, a value of ΔC_p = –1.2 kJ/mol/K is obtained and not ~–4 kJ/mol/K as appears from the uncorrected data.

3.4. Obtaining Electrostatic and Non-electrostatic Free Energies

The Gibbs free energy of binding is calculated from the association constant, which is best obtained by optical methods, e.g. fluorescence spectroscopy. The change in fluorescence anisotropy on titration of LEF86 into a 16-bp FAM-labelled DNA^{Lef} duplex yielded a binding isotherm fitting a 1:1 interaction having K_D ~0.1 nM (**Fig. 5a**). This DBD consists of a folded domain and a C-terminal extension of eight residues, seven of which are basic (K or R). The binding of a truncated form lacking the C-tail (LEF79) shows that removal of the very basic C-tail raises K_D to ~200 nM.

The ionic strength dependence of the association constant is instructive in this example (*see* **Fig. 5b**). The experiment is most simply carried out by gradually increasing the salt concentration of a single solution: knowing the value of the anisotropy at saturation and for free DNA, measured at the same salt concentration, interpolation gives the fraction bound and hence the K_D value at

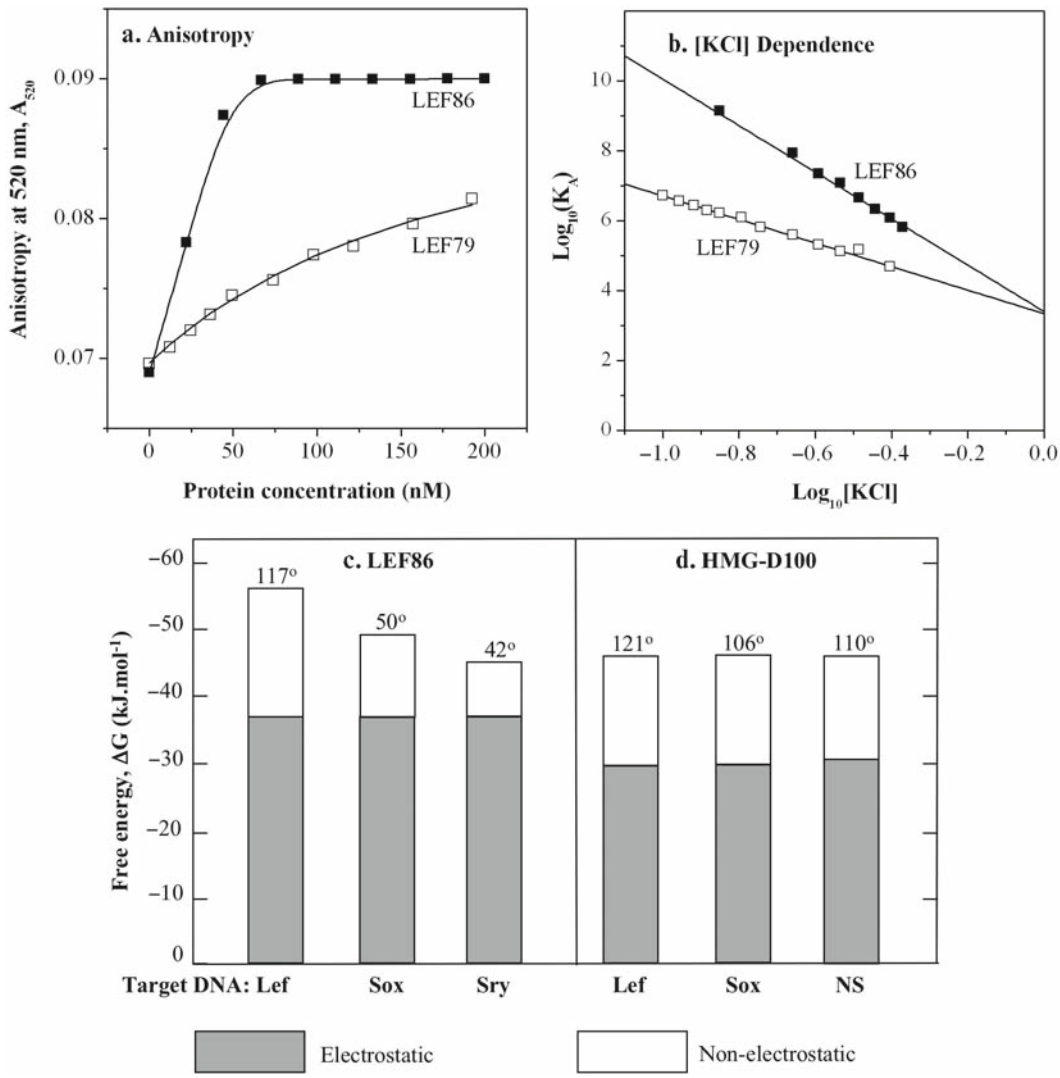


Fig. 5. (**a**) Change in fluorescence anisotropy upon titration of LEF86 into a 16-bp FAM-labelled duplex DNA^{Lef} at 20°C in 100 mM KCl, 10 mM phosphate buffer (pH 6.0). Since the DNA concentration was 50 nM and K_D is ~0.1 nM for LEF86, this binding curve cannot be used to derive K_D, but it does demonstrate the 1:1 stoichiometry. Removal of the very basic C-tail lowers K_D to ~200 nM, so the binding isotherm for LEF79 under the same conditions exhibits substantial curvature and fitting to a 1:1 binding model allows an accurate determination of its K_D. (**b**) Dependence of the log values of the association constants of LEF86 and LEF79 with DNA^{Lef} on the log value of the KCl concentration (K_D values obtained by measuring the FRET effect from doubly labelled DNA). (**c**) and (**d**) The separated electrostatic and non-electrostatic components of the total Gibbs free energy of binding LEF86 and HMG-D100 with DNAs of various sequences in the standard buffer at 25°C, derived by extrapolating the binding constant to 1 M KCl so as to obtain ΔG_{nel}. The *numbers* above the bars indicate the induced DNA bend angles, as measured by FRET, also in the standard buffer at 25°C.

each ionic strength. For maximum accuracy however, it is best to obtain complete binding isotherms at several ionic strengths to check the reliability of the 'single solution' approach. In 1 M KCl both the LEF86 and LEF79 domains have the same association constant ($\log(K_A)_{ncl} = 3.3$), which represents the non-electrostatic

contribution to binding of the folded domain and is the basis of its sequence specificity. It follows that interaction of the C-tail with the target duplex is entirely electrostatic. In the standard buffer (containing 100 mM KCl, i.e. log[KCl] = –1), log(K_A) for the intact LEF86 is 10.2 and for the truncated construct lacking the basic C-tail it is 6.7, i.e. the tail contributes more than three orders of magnitude to K_A under these conditions. However, in the absence of *all* ionic contacts log(K_A) has dropped to 3.3, and this makes it clear that the globular domain (of 78 residues) also makes a substantial number of ionic contacts with the DNA. In fact, the change in log(K_A) on removing the tail under standard conditions (10.2–6.7 = 3.5) is almost the same as the change on losing all the ionic contacts from the globular domain (6.7–3.3 = 3.4), so the tail and the globular domain make the same number of ionic contacts with the DNA. The slopes N in **Fig. 5b** (i.e. $Z\Psi$, assuming $\beta = 0$) lead to the conclusion that both domains of LEF86 make about five ionic contacts with the DNA, a conclusion well corroborated by the NMR structure of the globular domain, although for the C-tail this value cannot be verified due to its extreme mobility.

Knowledge of $(K_A)_{nel}$ yields ΔG_{nel}, so that if we obtain ΔG_{total} from the binding curve determined under the standard conditions, we derive the electrostatic contribution to the Gibbs energy of binding as $\Delta G_{el} = \Delta G_{total} - \Delta G_{nel}$. The electrostatic contribution is in fact entirely entropic (as demonstrated by the observation that binding enthalpies – as determined by ITC – are *independent* of the salt concentration, see later), and the entropy increase comes from the release of bound cations from the DNA phosphates into the bulk solution. Since the protein component of the complex is also highly charged, one might ask whether release of counterions (of either sign) from its surface also gives an entropy contribution. Generally, it appears that protein counterions are not so tightly bound as to result in a significant mixing entropy when released. However, tightly coordinated anions could play a role with basic DBDs, as seen in a recent example of homeodomain binding to DNA *(26)*. When KCl was substituted by NaCl, there was no change in the affinity but substitution by KF led to a rise in the affinity, i.e. the nature of the anion was important, though change of the monovalent cation – which exchanges with the DNA phosphates – was not. This suggests that chloride ion(s) are tightly bound to some of the basic residues on the homeodomain surface that participates in interactions with DNA in the complex.

The insights that can be obtained from separating the electrostatic and non-electrostatic contributions to the free energy of binding are well illustrated by comparing the SS LEF86 with the NSS HMG box from Drosophila HMG-D100. **Figure 5c, d** gives a breakdown of their Gibbs energies of binding to several

DNA sequences under the standard buffer conditions. For both DBDs the electrostatic forces dominate and when the basic tails are removed, in both cases the electrostatic contribution drops to about one-half that for the full-length DBD *(8)*, i.e. the tail and the globular domain of each DBD make equal numbers of DNA contacts. Comparing the two DBDs bound to DNA^{Lef}, it is seen that the mix of electrostatic and non-electrostatic contributions is almost the same, as is the bend angle generated, so where does the difference between SS and NSS binding lie? When the target sequence is altered to DNA^{Sox} and DNA^{Sry}, somewhat less favourable for binding of LEF86, the electrostatic contribution remains unchanged (i.e. all the phosphate contacts are maintained, as expected), but the non-electrostatic contribution drops, as does the bend angle (measured by FRET). Since the non-electrostatic contribution comes from van der Waals contacts and H-bonds between the protein and DNA, reduced complementarity at the interface would be expected to lower this contribution and lead to less bending: this is the manifestation of sequence specificity. For HMG-D100, not only is the electrostatic contribution independent of the target DNA sequence but so also is the non-electrostatic contribution and, by and large, the bend angle. The unchanged non-electrostatic contribution must mean that the van der Waals contacts and H-bonds at the interface are also unaffected by the DNA sequence: this is the manifestation of non-sequence specificity *(6)*.

3.5. A Full Energetic Profile for a Protein/DNA Association

This can now be derived as follows:

1. When the salt concentration approaches 1 M (**Fig. 5b**), the electrostatic term in **Eq. 2** becomes zero (since log[KCl] = 0) and the Gibbs energy of association then represents the non-electrostatic component, ΔG_{nel}. By subtracting this term from the overall Gibbs energy in the standard buffer, one can calculate the electrostatic term of the Gibbs energy: $\Delta G_{el} = \Delta G_{total} - \Delta G_{nel}$.
2. The enthalpy ΔH_{total} is obtained from the corrected ITC enthalpies. Since the enthalpy of HMG box binding to DNA has been experimentally demonstrated to be independent of the salt concentration (e.g. SRY ***(8)***), it can be entirely assigned to the non-electrostatic component of the Gibbs energy.
3. The non-electrostatic entropy factor can then be determined as $T\Delta S_{nel} = \Delta H_{nel} - \Delta G_{nel}$. Since the electrostatic component of the Gibbs energy of association is entirely entropic in nature, the electrostatic component of the entropy factor, $T\Delta S_{el}$ equals $-\Delta G_{el}$.
4. We now have the total Gibbs free energy (ΔG_{total}), the enthalpy ΔH, and the entropy factor $T\Delta S$, the values being under the standard buffer and temperature conditions. The enthalpy and

entropy can be extrapolated to other temperatures using the corrected value of ΔC_p (e.g. from **Fig. 4b**), thereby yielding the Gibbs energies at other temperatures. It is typically found that the total Gibbs free energy (ΔG_{total}) has only a small temperature dependence.

Obtaining the full energetic profile of a DBD/DNA interaction can allow a detailed interpretation to be made of the forces driving the interaction. This is illustrated by comparing the LEF86 DBD which binds to DNA^{Lef} sequence specifically, with that of the HMG-D100 DBD and its truncated form HMG-D74, both of which bind to DNA^{Lef} non-sequence specifically. **Figure 6** shows that sequence-specific binding of the LEF86 DBD is actually opposed by a small positive enthalpy (+13 kJ/mol). LEF86 binding is therefore driven by a substantial negative entropy factor, $T\Delta S$ (i.e. a large entropy increase). The Gibbs energy of binding HMG-D100 to DNA^{Lef} is almost as negative as that for LEF86 (*see* **Fig. 6**) but in this case binding is strongly opposed by a large positive enthalpy: correspondingly, binding is driven by an even larger negative entropy factor, principally non-electrostatic.

Since the structures of the two complexes are rather similar *(5, 6)*, the question arises as to why they have such disparate energetic signatures. In both cases, structures show that the highly ordered water in the AT-rich minor groove is totally excluded on binding protein, and this must result in a large positive enthalpy and a large negative non-electrostatic entropy factor, as seen for HMG-D100. However, for the sequence-specific LEF86/DNA^{Lef} complex there is an *additional* negative enthalpy from the formation of multiple van der Waal's contacts and H-bonds at the complementary interface, a contribution absent from the D-100 complex. The result is that the net enthalpy is much less positive in the case of the LEF86 complex than for HMG-D100. As regards the entropy factors, immobilization of groups at the interface of the LEF86 complex results in a reduction in the non-electrostatic contribution to $T\Delta S$, in contrast to the HMG-D100 complex where there is little immobilization of groups at the interface due to lack of complementarity. When the C-tail is removed from HMG-D100 (to give HMG-D74) the Gibbs energy is only about 60% of the value for HMG-D100. **Figure 6** shows that this is entirely due to a reduction in the electrostatic entropy factor. This is due to the loss of multiple ionic contacts in the C-tail.

3.6. Determination of Concentrations

The protein and DNA samples must be highly homogeneous and their concentrations accurately determined by UV-absorption spectroscopy. Firstly, this enables an accurate preparation of a DNA duplex from its complementary oligonucleotides and of the HMG box-DNA complex from protein and DNA duplex. In DSC, the accuracy with which the sample concentration can be determined

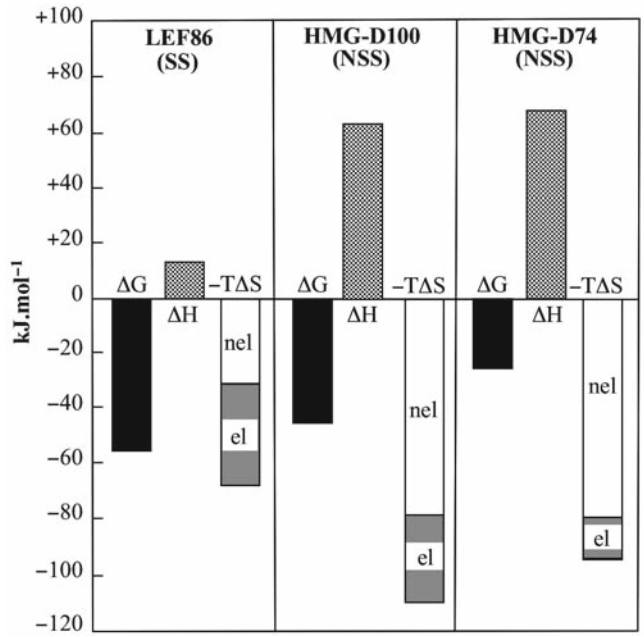


Fig. 6. Thermodynamic profiles of LEF86 (a sequence-specific binding HMG box), HMG-D100 and HMG-D74 (non-sequence-specific binding HMG boxes) binding to DNA^{Lef}, showing the Gibbs free energy, ΔG, and its contributions from the enthalpy, ΔH, and entropy factor, $-T\Delta S$. The entropy factor is shown separated into its electrostatic (el) and non-electrostatic (nel) contributions.

directly affects the molar heat capacity values obtained. In ITC, unlike other binding assays, errors in the concentration of protein are directly reflected in the values of ΔH_{app} (**Eqs. 3 and 4**). Thus, in both ITC and DSC, concentrations are determined after dialysis. In the DSC experiments it is also possible that a small amount of buffer from a previous scan remains in the gold capillary cell. To eliminate this potential source of error, concentrations were determined using the excess sample that remains after filling the calorimeter cell.

3.6.1. DNA

For oligonucleotides and DNA duplexes, concentrations were determined from their UV absorption at 260 nm, after digestion to nucleotides with snake venom phosphodiesterase I (PDE1).

1. Accurately dilute the solution of DNA to about 1.0 mL with 100 mM Tris–HCl (pH 8.0) so as to give an absorbance at 260 nm of about 0.5 in a 1-cm path-length cell. To overcome any reliance on the presumed accuracy of the pipettes being used, the dilutions are performed on a precision balance.
2. Record the UV-absorption spectrum of the diluted DNA solution, using as a reference a blank cell containing buffer. Note that for small oligonucleotides and DNA duplexes with

a biased nucleotide composition, the absorption maximum is not necessarily at 260 nm. This is in fact an optional step, but the recording of UV spectra both before and after the addition of PDE1 indicates that digestion has occurred and enables a determination of the hypochromicity of the DNA solution.

3. Remove the diluted DNA solution from the cell and place into a screw-capped 1.5-mL tube. Add either 0.008 units (for oligonucleotides) or 0.08 units (for DNA duplexes) of PDE1 to the solution, mix well, and incubate at 37°C overnight. If necessary, a time course of the digestion may be followed (at room temperature) in the UV spectrophotometer by direct addition of PDE1 to the DNA solution in the quartz cell.
4. Accurately record the UV-absorption spectrum of the DNA solution after PDE1 digestion, using as reference a cell containing buffer. The contribution to the absorbance at 260 nm arising from overlap of the PDE1 280-nm peak may be neglected.
5. Optionally, calculate the percent hypochromicity, %H from the equation

$$\%H = \frac{(A_a - A_b)}{A_a} \times 100,$$

where A_b and A_a are the absorbances at 260 nm before and after PDE1 digestion respectively, with dilution due to the addition of PDE1 taken into account. Values of %H are approx. 20% for oligonucleotides and 35–40% for duplex DNA. Lower values indicate that digestion may not be complete, that the original DNA is degraded, or that there is a failure to form duplex DNA.

6. Calculate the molar nucleotide concentration and thus the molar DNA concentration from the following equations:

$$[\text{Nucleotide}] = \frac{A_a(\text{dilution})N}{(12,010G) + (15,200A) + (8,400T) + (7,050C)}$$

and

$$[\text{DNA}] = \frac{[\text{Nucleotide}]}{N}$$

where N is the total number of nucleotides in the DNA; and G, A, T, and C are the number of dG, dA, dT, and dC nucleotides, respectively. The extinction coefficients at 260 nm for the four nucleotides are from **ref. 27**.

3.6.2. Proteins

LEF79 and LEF86 concentrations (in standard buffer) were determined from their UV absorption at 280 nm using an extinction coefficient of 17,780 M^{-1} cm^{-1}. This extinction coefficient is based on the addition of the extinction coefficients of tryptophan and tyrosine (of which there are 2 and 5 in LEF79 and LEF86, respectively). The concentration of HMG-D74 and HMG-D100

was similarly determined using an extinction coefficient at 280 nm of 19,252 $M^{-1}cm^{-1}$. The accuracy of the concentration determination is within 5%.

3.6.3. HMG Box/DNA Complexes

The concentration of complex was determined from its UV absorbance at 260 nm. At this wavelength the absorption is mainly due to the DNA, but there is a significant absorption from the protein that must be taken into account. There may also be additional hyperchomic effects since the DNA of the HMG box–DNA complex is highly bent with considerable base unstacking. The approach taken was therefore to measure the UV absorbance at 260 nm for a solution of DNA duplex and then again after the addition of an equimolar amount of protein to form the complex.

4. Notes

1. The choice of buffer depends on the system under study and on the requirements of other assays performed in the context of a particular investigation. If a titration is conducted at near neutral pH, it is possible that titratable groups in the protein, e.g. the imidazole ring of histidine, change protonation level on complex formation. This is of course a change in its p*K*a and is a large enthalpic process. In the case of the HMG boxes HMG-D74 and HMG-D100 there are no histidine residues, whilst in LEF79 and LEF86 there are three, so it is possible that the observed differences in the enthalpies on binding DNA in the standard buffer at pH 6.0 result from protonation/deprotonation of histidine(s). The heat of DNA binding by LEF86 was therefore measured at three different pH values: pH 5.0, 6.0, and 7.0 at 20°C and the binding enthalpies found to differ by less than 2 kJ/mol, i.e. within experimental error.
2. Each cell of the Nano-DSCIII calorimeter should be filled by the following method using a minimum of 1.0 mL of solution. Attach a 2.5-mL pipette tip with a short length of silicon tubing to one port of the cell. Draw up the solution using a 2.5-mL pipettor also fitted with a 2.5-mL tip and silicon tubing. Connect this to the other port of the calorimeter cell and slowly introduce the solution. Since the cell is only ~330 μL in volume, excess solution rises up into the other 2.5-mL pipette tip. The solution is then pipetted up and down, at first slowly and then more rapidly, for a sufficient time to completely expel all air bubbles from the cell. Eventually, excess solution may be removed by care-

fully withdrawing it into the tip connected to the pipettor and then disconnecting by simultaneously pulling off both 2.5-mL tips with their connecting tubing. This excess solution is used for concentration determinations.

3. As pressure is applied note the compensatory power reading – it should not change by more than ~10 μW. If greater, then air bubbles are still present in one or both of the cells, which must then be emptied and refilled.
4. One reason may be temperature-induced aggregation, manifesting itself as a large peak of heat evolution that typically commences close to the point of maximum heat absorption in the co-operative denaturation transition (approximately at T_m). A second possible reason for irreversibility is the presence of disulphide bonds in the protein, which may disrupt on heating but on cooling form mixed disulphide bonds with incorrect Cys residues. This may be overcome by using proteins (or domains) with no Cys residues (as in HMG-D74 and HMG-D100) or by using engineered versions in which the Cys residues have been mutated to Ser (as in LEF79 and LEF86). This strategy will, however, depend on the contribution of the disulphide bonds in stabilising the overall protein fold. A third possibility is that heating the protein to high temperatures can cause some chemical degradation, with the result that the amount of protein capable of proper refolding has dropped.
5. The calorimetric cells may be thoroughly cleaned by filling them with 50% (v/v) formic acid, followed by heating from 25 to 75°C (at 1 K/min) and back to 25°C. The cells are then thoroughly rinsed with water and buffer. Multiple baseline scans of the buffer are recommended before proceeding with the next sample.
6. The partial specific volume used for DNA duplexes was 0.540 mL/g. The partial specific volume of LEF86 was calculated from its composition as 0.728 mL/g, using the known partial specific volumes of the amino acids. The LEF86/15-bp DNA complex partial specific volume of 0.640 mL/g was calculated as the weight average of the partial specific volumes of LEF86 and the DNA duplex.
7. In setting the stirring speed in the titration calorimeter, the particular system under study must be considered. Some proteins do not resist the forces arising from rapid stirring of the solution placed in the narrow calorimetric cell and aggregate or unfold in consequence. On the other hand, very slow stirring may result in low rates of heat transfer thus causing broadening of the peaks observed in the thermogram and thus a decrease in sensitivity.

8. When working below room temperature, it is recommended to fill the cells with cold solutions; otherwise the equilibration time may be very long. Any particles in the stirred sample cell as well as the formation of bubbles during filling may cause problems in establishing the baseline.
9. There is no general rule about the number, volume, and duration of injections. To obtain an entire binding isotherm, a 2–4 molar excess of ligand over the concentration of receptor-binding sites should be injected by at least 10–15 additions. Five to ten injections suffice to reliably measure the enthalpy of reaction. In this case, the degree of saturation should be <0.5 at the end of experiment. The volume and duration of injections must be chosen in such a way that the released heat is well above the threshold of sensitivity and sharp peaks appear in the thermogram. Typically, 5-min intervals between injections are sufficiently long for the signal to return to the baseline. However, for very slow reactions this interval should be prolonged. Because the heat of reaction is a temperature-dependent quantity, the injection scheme might have to be changed for experiments carried out at different temperatures.
10. Alternating cycles of washing with 0.5 M NaOH and isopropanol are routinely used. To remove heavily precipitated proteins, the cell can be rinsed with a hot solution of 20% sodium dodecyl sulphate (SDS). The protocol for cleaning will depend on the particular system under study and the material from which the cell is made.
11. Integration requires construction of a proper baseline. Some software products support automatic procedures for baseline determination. However, it is often found that manual adjustment of the baseline is a better practice, particularly in cases when the signal-to-noise ratio is low or if the instrument baseline has drifted. In constructing the baseline manually, it is important to use the same integration window for all the peaks observed, both in the specific binding and in the blank titration experiments. This will avoid the introduction of non-random bias in the data.

Acknowledgements

We would like to thank Peter Privalov (Johns Hopkins University, Baltimore), in whose laboratory the calorimetry was performed. Financial support from an NIH grant to the Baltimore laboratory (GM48036-06) and a Wellcome Trust grant to the Portsmouth laboratory are gratefully acknowledged.

References

1. Patikoglou, G. and Burley, S.K. (1997). Eukaryotic transcription factor–DNA complexes. *Ann. Rev. Biophys. Biomol. Struct.* **26**, 289–325.
2. Schultz, S.C., Shields, G.C. and Steitz, T.A. (1991). Crystal structure of a CAP–DNA complex: the DNA is bent by 90°. *Science* **253**, 1001–1007.
3. Kim, J.L., Nikolov, D.B. and Burley, S.K. (1993). Co-crystal structure of TBP recognising the minor groove of a TATA element. *Nature* **365**, 520–527.
4. Werner, M.H., Huth, J.R., Gronenborn, A.M. and Clore, G.M. (1995). Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY–DNA complex. *Cell* **81**, 705–714.
5. Love, J.J., Li, X., Case, D.A., Giese, K., Grosschedl, R. and Wright, P.E. (1995). Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **376**, 791–795.
6. Murphy, F.V., Sweet, R.M. and Churchill, M.E.A. (1999). The structure of a chromosomal high mobility group protein–DNA complex reveals sequence neutral mechanisms important for non-sequence specific DNA recognition. *EMBO J.* **18**, 6610–6618.
7. Dragan, A.I., Klass, J., Read, C., Churchill, M.E., Crane-Robinson, C. and Privalov, P.L. (2003). DNA binding of a non-sequence-specific HMG-D protein is entropy driven with a substantial non-electrostatic contribution. *J. Mol. Biol.* **331**, 795–813.
8. Dragan, A.I., Read, C.M., Makeyeva, E.N., Milgotina, E.I., Churchill, M.E., Crane-Robinson, C. and Privalov, P.L. (2004). DNA binding and bending by HMG boxes: energetic determinants of specificity. *J. Mol. Biol.* **343**, 371–393.
9. Murphy, K.P. and Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behaviour in proteins. *Adv. Protein Chem.* **43**, 313–361.
10. Spolar, R.S., Livingstone, J.R. and Record, M.T. Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry* **31**, 3947–3955.
11. Manning, G.S. (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11**, 179–246.
12. Ha, J.H., Capp, M.W., Hohenwalter, M.D., Baskerville, M. and Record, M.T. Jr. (1992). Thermodynamic stoichiometries of participation of water, cations and anions in specific and non-specific binding of lac repressor to DNA. Possible thermodynamic origins of the "glutamate effect" on protein–DNA interactions. *J. Mol. Biol.* **228**, 252–264.
13. Olmsted, M.C., Bond, J.P., Anderson, C.F. and Record, M.T. Jr. (1995). Grand canonical Monte Carlo molecular and thermodynamic predictions of ion effects on binding of an oligocation (L8+) to the center of DNA oligomers. *Biophys. J.* **68**, 634–647.
14. Manning, G.S. (2003). Is a small number of charge neutralizations sufficient to bend nucleosome core DNA onto its superhelical ramp? *J. Am. Chem. Soc.* **125**, 15087–15092.
15. Privalov, P.L. and Makhatadze, G.I. (1990). Heat capacity of proteins II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *J. Mol. Biol.* **213**, 385–391.
16. Makhatadze, G.I. and Privalov, P.L. (1995). Energetics of protein structure. *Adv. Protein Chem.* **47**, 307–425.
17. Privalov, G., Kavina, V., Freire, E. and Privalov, P.L. (1995). Precise scanning calorimeter for studying thermal properties of biological macromolecules in dilute solution. *Anal. Biochem.* **232**, 79–85.
18. Privalov, G.P. and Privalov, P.L. (2000). Problems and prospects in the microcalorimetry of biological macromolecules. *Methods Enzymol.* **323**, 31–62.
19. Privalov, P.L. and Dragan, A.I. (2007). Microcalorimetry of biological macromolecules. *Biophys. Chem.* **126**, 16–24.
20. McKinnon, I.R., Fall, L., Parody-Morreale, A. and Gill, S.J. (1984). A twin titration microcalorimeter for the study of biochemical reactions. *Anal. Biochem.* **139**, 134–139.
21. Wiseman, T., Williston, S., Brandts, J.F. and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–37.
22. Freire, E., Mayorga, O.L., and Straume, M. (1990). Isothermal titration. *Anal. Chem.* **62**, 950A–959A.
23. Breslauer, K.J., Freire, E. and Straume, M. (1992). Calorimetry: a tool for DNA and ligand–DNA studies. *Methods Enzymol.* **211**, 533–567.

24. Love, J.J., Li, X., Chung, J., Dyson, H.J. and Wright, P.E. (2004). The LEF-1 high-mobility group domain undergoes a disorder-to-order transition upon formation of a complex with cognate DNA. *Biochemistry* **43**, 8725–8734.
25. Briggner, L.E. and Wadso, I. (1991). Test and calibration processes for microcalorimeters, with special reference to heat conduction instruments used with aqueous systems. *J. Biochem. Biophys. Methods* **22**, 101–118.
26. Dragan, A.I., Li, Z., Makeyeva, E.N., Milgotina, E.I., Liu, Y., Crane-Robinson, C. and Privalov, P.L. (2006). Forces driving the binding of homeodomains to DNA. *Biochemistry* **45**, 141–151.
27. Wallace, R.B. and Miyada, C.G. (1987) Oligonucleotide probes for the screening of recombinant DNA libraries. *Methods Enzymol.* 152, 432–442.

Chapter 38

Surface Plasmon Resonance Assays of DNA–Protein Interactions

Peter G. Stockley and Björn Persson

Summary

Assaying sequence-specific DNA–protein complex formation in vitro often involves the use of specific labelling or modification of the components of the complex to provide unique signals that can be used to assess the affinity of the interaction. Surface plasmon resonance (SPR) spectroscopy is an optical technique that can be used without radio- or other labelling of the components of a complex provided that one of the partners can be immobilised to a solid support. For DNA oligonucleotides this can easily be achieved by the incorporation of a biotin end label, but proteins can also be immobilised if they carry conventional tags for affinity purification, such as GST or polyhistidine extensions. The SPR effect relies on changes in the refractive index of solutions adjacent to the immobilised surface and is extremely sensitive. The continuous flow systems developed by BIAcore AB (now GE Healthcare Biosciences AB) permit real-time recording of the binding and dissociation of analyte species to the immobilised ligand, resulting in both rapid stoichiometric kinetic, affinity, and thermodynamic measurements. These assays can be carried out with complex mixtures of analytes, providing a powerful addition to the techniques available to probe such interactions. We illustrate the use of such assays here using the example of the *E. coli* methionine repressor, MetJ, which is also described in Chapters "Filter-Binding Assays" and "Ethylation Interference Footprinting of DNA–Protein Complexes."

Key words: Surface plasmon resonance, Spectroscopy, DNA–protein interactions, Affinity and kinetics, MetJ, Methionine repressor.

1. Introduction

The procedures for biomolecular interaction analysis described in this chapter are based on biosensor technology, which offers an opportunity to study the interaction between an immobilised partner, the ligand, and an analyte molecule in solution

Tom Moss and Benoît Leblanc (eds.), *Methods in Molecular Biology, DNA-Protein Interactions, vol. 543*

DOI: 10.1007/978-1-60327-015-1_38

(*see* **Note 1**). The interaction is monitored continuously and the binding events are studied in real time, with no need for labelling of any component. The technology is generally applicable to the study of any molecular structures which express affinity for one another, including proteins, nucleic acids, carbohydrates, lipids, low molecular weight substances, but also large particles like vesicles, bacteria, and cells. The wealth of information provided by this technology includes qualitative data such as binding specificity, but quantitative aspects of binding like stoichiometry, affinity, concentration, and kinetics are also amenable to detailed analysis *(1, 2)*.

1.1. Principle of Detection

The principle of detection relies on the optical phenomenon of surface plasmon resonance (SPR), which is sensitive to the optical properties of the medium close to a metal surface. Various configurations of instrumentation may be used to exploit the SPR effect and the following descriptions are all based on the commercial instruments available from GE Healthcare Biosciences AB, Uppsala (formerly BIAcore AB). The specific details of instruments from other manufacturers will vary, but many of the protocols outlined should be adaptable in most cases *(3)*. **Figure 1** shows the detection unit of the instrumental set up in the Biacore system. The detection system consists of a monochromatic, plane-polarised light source and a photodetector that are connected optically through a glass prism. The thin metal film positioned on the prism is in contact with the sample solution. The fan-shaped

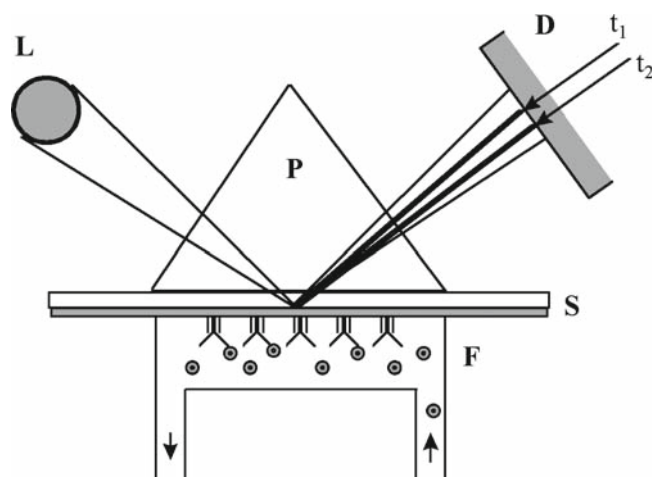


Fig. 1. Surface plasmon resonance detection unit. *L* plane-polarised light source, *D* photodiode array detector, *P* prism, *S* sensor surface, *F* flow cell. The two *dark lines* in the reflected beam projected on to the detector symbolise the light intensity drop following the resonance phenomenon at time = t_1 and t_2. The *line* projected at t_1 corresponds to the situation before binding of analyte species to the immobilised target molecules on the surface and t_2 is the position of resonance after binding.

light beam incident on the back side of the metal film is totally internally reflected on to the diode-array detector.

SPR occurs when light incident on the metal film couples to oscillations of the conducting electrons, plasmons, at the metal surface. These oscillations create an electromagnetic field commonly referred to as the evanescent wave. This wave extends from the metal surface into the sample solution on the back side of the surface relative to the incident light, i.e. the non-illuminated face. When resonance occurs, the intensity of the reflected light decreases at a sharply defined angle of incidence, the SPR angle, which is dependent on the refractive index penetrable by the evanescent wave close to the metal surface. The SPR angle is dependent on several instrumental parameters, e.g. the wavelength of the light source and the metal of the film. When these parameters are kept constant, the SPR angle shifts are dependent only on changes in refractive index of a thin layer adjacent to the metal surface. If the SPR angle shift is monitored over time, a gradual increase of material at the surface will cause a successive increase of the SPR angle, which is detected as a shift of the position of the light intensity minimum on the diode array.

The SPR angle shifts obtained from different proteins in solution have been correlated to surface concentration determined from radiolabelling techniques and found to be linear within a wide range of surface concentration. The instrument output, the resonance signal, is given in resonance units (RU); 1,000 RU corresponds to a 0.1° shift in the SPR angle, and for an average protein this is equivalent to a surface concentration change of about 1ng/mm^2. Note, the relationship between RU change and surface concentration may be different for non-protein species. It has been reported that for DNA 780 RU = 0.1° shift *(1, 4, 5)*, but this has been disputed recently with a claim that nucleic acids and protein have similar dependencies *(6)*.

1.2. Surface Chemistry

The sensor chips are available commercially and consist of a glass substrate on to which a 50-nm-thick gold film has been deposited. The gold film is covered with a long-chain hydroxyalkanethiol, which forms a monolayer at the surface. This layer serves as a barrier to prevent analyte coming into contact with the gold, but also as the attachment point for carboxymethylated dextran chains that create a hydrophilic surface to which proteins or other molecules can be covalently coupled. A range of surface-derivatised sensor chips with differing dextran layers and/or chemistries are available *(2, 7)* (*see* **Note 4**). In a typical kinetic analysis, one of the two interacting partners, commonly referred to as the ligand, is immobilised on the dextran. For example, substances containing primary amines can be immobilised after activation of the dextran matrix with carbodiimide/*N*-hydroxysuccinimide. Several other procedures are also available for the immobilisation of

ligands, e.g. DNA is commonly synthesised with a 5′-biotin label and captured on streptavidin that is immobilised on the sensor chip either by amine coupling or directly onto pre-coupled commercial streptavidin sensor chips (*see* **Subheading 2.2**).

An important feature of the sensor chip construction is that the dextran layer extends typically ~100 nm out from the surface and thereby a small volume, penetrated by the evanescent wave, is created where the analyte/ligand interaction can be studied. The specificity of the analysis is determined by the identity and biological activity of the ligand.

1.3. Liquid-Handling System

Sample solutions containing analytes are injected into a running buffer that flows continuously over the sensor surface. A typical kinetic analysis consists of (1) establishment of a baseline response in running buffer, (2) an association phase, i.e. injection of sample in running buffer, and (3) the dissociation phase, i.e. wash out with running buffer (*see* **Subheading 2.2**). For kinetic analysis of fast interactions, it is important that the flow system can perform virtually dispersion-free injections into the detector flow cell; otherwise, small deviations in analyte concentration will perturb the results.

The mass transfer of analyte molecules from the bulk of the sample solution in the flow cell to the ligand on the sensor surface is another crucial parameter in kinetic analysis. Fast mass transfer is obtained by having: (1) a high linear velocity of the sample solution, thereby diminishing the thickness of the unstirred layer near the surface; and (2) a thin flow cell, facilitating an unperturbed flow regime. It is also important to keep the ligand concentration low; otherwise, depletion of analyte near the surface will perturb the effective analyte concentration and thus the kinetic constants will be underestimated. A good baseline stability is required for kinetic analysis, e.g. when the dissociation phase of a stable complex is measured, but also in equilibrium analysis to obtain a reliable steady-state value during sample injection.

Interacting partners in solution are delivered to the surface via a system of flow channels, formed when a microfluidic cartridge is brought into contact with the sensor surface. The number of flow cells is two (Biacore X100) or four (Biacore 2000 and 3000, Biacore T100) and enables the corresponding number of sensor surfaces to be addressed. Each flow cell allows separate ligand immobilisation. In multi-channel detection mode these surfaces can be addressed in series which makes it possible to run the same analyte solution over two or four surfaces, respectively, offering an opportunity to use one surface as a reference whose signal may be subtracted from that of the active surface.

Biacore A100 has four parallel flow cells, each containing five detection spots. Here hydrodynamic addressing allows injected samples to be directed to selected spots or groups of spots within

the same flow cell. This technique enables multiple interactants to be immobilised on detection spots in a single flow cell, allowing simultaneous analysis of interactions with a reference spot in the same flow cell as the ligands. There is no lag time between samples reaching the active and reference spots, so that very rapid kinetics can be measured.

In recent versions of the instrumentation (Biacore T100, Biacore A100, Biacore X100) thorough method and evaluation software support is available. Application wizards provide guidance in setting up experiments for assay development and execution. Separate wizards are offered for different purposes such as immobilisation, concentration determination, or measurement of kinetic constants. Each wizard consists of an ordered series of dialog boxes, ensuring that the essential features of the application setup are correctly defined. If more flexibility (and conversely less guidance) is required the graphical interface called *Method Builder* provides full flexibility in method definition while still retaining a simple interface for running assays based on established methods. The evaluation software offers general functions for presentation of results (sensorgrams, report point plots) and application-specific evaluation, e.g. kinetics, affinity, and thermodynamic measurements.

For example, in the kinetics wizard it is recommended to run a concentration series of at least five analyte concentrations and one zero concentration. At least one analyte concentration should be measured in duplicate. A further option for kinetic analysis is single-cycle analysis. This assay type includes a series of sample injections (from low to high concentrations) in one cycle, with a dissociation period after the last injection. The surface is not regenerated between injections. The method is useful for determining kinetics for interaction systems that are difficult to regenerate. Also, when ligand capture is used, consumption of ligand is minimised since one capture injection is used for all sample concentrations.

2. Materials

2.1. Coupling Streptavidin to Carboxymethyl-Dextran Sensorchips

1. This protocol assumes access to a Biacore instrument. For Biacore 2000 and 3000 the instrument would be set up to run sample buffer at 20°C with rack D in the first position and rack A in the second position.
2. Streptavidin (e.g. from Pierce) resuspended in 0.22-μm filtered, distilled water to a final concentration of 5 mg/mL. This preparation may be stored at 4°C for up to 3 months.

3. HBS buffer (10 mM HEPES of pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.005% (v/v) BIAcore surfactant, such as Tween 20).
4. *N*-ethyl-*N*′-(diethylaminopropyl) carbodiimide (EDC) and *N*-hydroxy succinimide (NHS) purchased from Biacore as lyophilised powders, resuspended in 0.22-μm filtered, distilled water to final concentration of 100 mM each.
5. 1 M Ethanolamine hydrochloride (pH 8.5), purchased from Biacore, stored at 4°C.
6. Sensorchip surface CM5 research grade installed in the BIAcore apparatus according to the manufacturer's instructions and pre-primed with HBS buffer (*see* **Notes 2** and **5**).
7. Reaction vials for the Biacore (small, plastic = 7 mm; medium, glass = 16 mm; and large, glass = 2 mL) were purchased from Biacore.

2.2. Immobilisation of the DNA

1. Access to a Biacore instrument. For Biacore 2000 and 3000 the instrument would be set up to run sample buffer at 20°C with rack D in the first position and rack A in the second position.
2. End-biotinylated DNA suspended in HBS buffer, or equivalent, to 10 μg/mL.
3. HBS buffer.
4. Streptavidin-activated sensorchip CM5 research grade, or a pre-derivatised streptavidin sensorchip (SA) installed in the Biacore apparatus and pre-primed with HBS buffer.

2.3. Kinetic Analysis of MetJ-Operator Interaction

1. A stock solution of MetJ dialysed against the running buffer.
2. A stock solution of AdoMet (100 mM) in running buffer.

3. Methods

3.1. Experimental Design

Traditional approaches to the assay of protein–DNA interactions rely on alterations of the properties of the partners of the complex as the interaction proceeds. For instance, DNA fragments become retarded relative to unbound fragments in gel electrophoresis, or are retained on nitrocellulose filters by virtue of their interaction with proteins in filter-binding assays (*see* **Chapter** "Filter-Binding Assays") or the accessibility of functional groups in the DNA becomes reduced in footprinting assays (*see* **Chapters** "DNase I Footprinting," "Hydroxyl Radical Footprinting of Protein–DNA Complexes," "Ethylation Interference Footprinting of DNA–Protein Complexes," etc.). These are necessarily

techniques based on the separation of equilibrium mixtures or competition assays, neither of which is ideal for accurate quantitation. In some cases, the molecules concerned, usually the proteins, have unique spectral properties which permit analysis of the kinetics and equilibrium distribution. However, such situations are the exception rather than the rule.

The SPR biosensors provide an excellent tool to overcome the various limitations of these traditional techniques. In many cases it is possible to obtain stoichiometric, kinetic, thermodynamic, and affinity data very rapidly with relatively small amounts of material. The first task in setting up such SPR experiments is to design the assay for the interaction of interest. For practical purposes, it is often easiest to immobilise the target DNA fragments, via biotin–streptavidin complexation, and address these ligands with solutions of the test protein analyte. There are many advantages to this approach. DNA molecules are conformationally flexible and are naturally part of very large polymers which approximate to the immobilised state. There should therefore be little effect from the immobilisation, although this is not true when the DNA is used as the template for polymerases *(4)*. Regeneration of the sensorchip is also straightforward, requiring only removal of the protein analyte which is easily achieved without denaturing the DNA ligands.

The following sections describe the results obtained for the interaction of the *E. coli* methionine repressor, MetJ, and a series of target DNA molecules encompassing operator sites. MetJ is a well-characterised protein structurally, biochemically and genetically *(7–10)*, and details of its interaction with DNA can be seen in **Fig. 1** of "Filter-Binding Assays."

3.2. DNA Immobilisation

Sequence-specific DNA–protein complexes characteristically have very rapid forward rate constants, reflecting the functional need to identify the DNA target against the large background of competing non-specific sites. In order to avoid problems caused by mass transport limitations, the level of immobilised DNA ligand should be relatively small (~100–300 RU). This can conveniently be achieved by sequential injections of the appropriate low concentrations of biotinylated DNA ligands for differing contact times. Immobilisation is usually linear during this period (*see* **Fig. 2**).

3.2.1. Coupling Streptavidin to Carboxymethyl-Dextran Sensorchips

1. Prime the apparatus with HBS buffer.
2. The thawed EDC solution in an Eppendorf tube with the top removed is placed in rack 1 position a1 (r1a1).
3. The thawed NHS solution in an Eppendorf tube with the top removed is placed in r1a2.
4. Streptavidin (5 mg/mL, 50 μL) in an Eppendorf tube with the top removed is placed in r1a3.

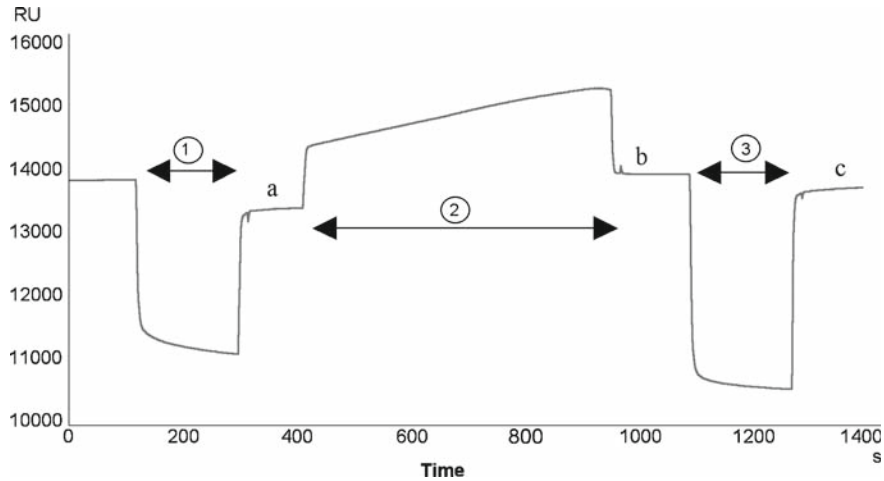


Fig. 2. Sensorgram showing the different steps of the immobilisation of DNA on an SA-5 sensor chip. *Phase 1*: drop in signal as NaOH is passed over the sensor chip surface. A new baseline signal (a) is reached at the end of this injection due to the loss of loosely bound streptavidin from the surface. *Phase 2*: injection and binding of biotinylated DNA to the surface. At the end of this injection an increased baseline signal is seen (b), due to the bound DNA. *Phase 3*: injection of SDS to remove loosely bound DNA from the surface of the sensor chip. At the end of this injection a slightly lowered baseline signal is established (c). If the baseline signal, a, is subtracted from this new reading, c, the amount of DNA immobilised onto the surface can be calculated *(1)*.

5. 2 mL of filtered (0.2 μm) distilled water is placed in a large glass vial in r2f7.
6. 1 M Sodium acetate buffer (1 mL, pH 4.5) is placed in a large glass vial in r2f3.
7. 1 M Ethanolamine (200 μL) is placed in a large tube in r2f4.
8. Two small clean plastic vials are placed in positions r2a1 and r2a2.
9. An empty large glass vial is placed in r2f5.
10. The following method, or equivalent, is programmed into the Biacore instrument, checked for errors and run. Note, modern instrument software contains simple routines (wizards) that offers prompts for most common protocols. The EDC/NHS reagents are mixed according to the manufacturer's instructions and the activation of the surface allowed to proceed for 7 min. Streptavidin (150–500 μg/ml) is passed across the surface in 10 μm acetate buffer, again for 7 min, before the reaction is terminated by passing across the ethanolamine, again for 7 min.

3.2.2. Immobilisation of the DNA

1. It is advisable to run this sensorgram (**Fig. 2**) manually so as to monitor the degree of immobilisation.
2. Select a surface pre-treated with streptavidin. For Sensor Chip SA inject three 60-s aliquots of 50 mM NaOH to remove loosely retained streptavidin.

3. Flow HBS buffer at 20 μL/min across the surface until a steady baseline signal is achieved. This can take several hours.
4. Inject the DNA solution across the surface, set the baseline to the point of injection, and monitor the change in RU during the injection phase. Ideally, between 20 and 50 RU of DNA should be immobilised (see later).
5. Wash the surface with a 50-μL injection of 1 M NaCl in filtered, distilled water.
6. Wash the surface with a 15-μL injection of 0.05% (w/v) sodium dodecyl sulphate (SDS).
7. Allow the surface to re-equilibrate in HBS buffer to a stable baseline; the difference in RU between the beginning of the injection phase and the end of the wash period reflects the amount of DNA bound.

3.2.3. Kinetic Analysis of MetJ-Operator Interaction

1. Establish a steady baseline with running buffer flowing over the sensor chip derivatised with the target, two met box DNA duplex, at the highest possible flow rate, e.g. 30 μL/min for the example shown in **Fig. 3**. For studies with MetJ the buffer was TKE (20 mM Tris–HCl of pH 7.4, 200 mM KCl, 0.5 mM EDTA, 0.0005% (v/v) Tween 20).

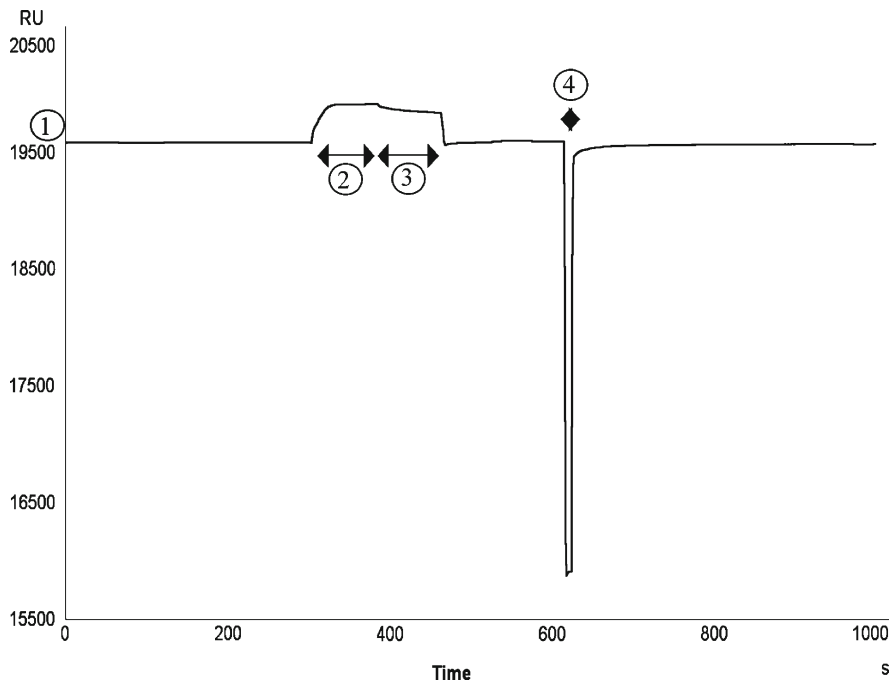


Fig. 3. (**a**). A typical sensorgram showing MetJ binding to and dissociating from the operator-derivatised sensor chip surface and the regeneration of the sensor chip. *Phase 1*: initial baseline signal when running buffer passes over the sensor chip. *Phase 2*: first injection containing MetJ and AdoMet. The change in signal is due to the binding of the holo-repressor to the chip surface. *Phase 3*: second injection of running buffer containing AdoMet alone, allowing the dissociation in the presence of effector to be measured. *Phase 4*: regeneration of the sensor chip surface by the injection of SDS *(1)*.

2. Make a serial dilution of the MetJ stock protein into running buffer plus saturating (1 mM) AdoMet. On completely automated instruments it is sometimes best to do this dilution via the instrument software.
3. Program triplicate injections of 30 μL of each concentration of MetJ + AdoMet followed by co-injection of 30 μL of AdoMet in buffer across the flow cell. Each analyte injection should be separated by a wash phase of 15 μL 0.05% (w/v) SDS to ensure complete dissociation of repressor before a second sample is injected. We program the triplicates to occur in series allowing the activity and concentration changes (e.g. due to evaporation) to be monitored over many hours. In our hands each derivatised flow cell can be used up to 100× before it begins to show deterioration.

3.3. Binding Curve Analysis

The mathematical treatment of the data obtained is usually carried out using the software supplied by the manufacturer. The basic equations describing the simplest analyses are shown as follows.

3.3.1. Stoichiometry and Equilibrium Analysis

For an immobilised DNA fragment (D), the interaction with a mobile protein (P) can be written as:

$$\mathrm{D} + \mathrm{P} \underset{k_{\mathrm{diss}}}{\overset{k_{\mathrm{ass}}}{\rightleftarrows}} \mathrm{DP}. \qquad (1)$$

A classical Langmuir adsorption isotherm requires that the fraction of available sites on the DNA occupied by the protein (θ_{D}) is given by:

$$\theta_{\mathrm{D}} = \frac{[\mathrm{DP}]}{[\mathrm{D}_{\mathrm{T}}]} \ . \qquad (2)$$

Furthermore, in such a simple case the equilibrium association constant K_{a} is given by the expression:

$$\theta_{\mathrm{D}} = \frac{K_{\mathrm{a}}[\mathrm{P}]}{1 + K_{\mathrm{a}}[\mathrm{P}]} \ . \qquad (3)$$

Thus, assuming that in the continuous flow system, $[\mathrm{P}]_{\mathrm{T}} = [\mathrm{P}]$, a plot of ($\theta_{\mathrm{D}}$) against $[\mathrm{P}]_{\mathrm{T}}$ should allow a direct fit by **Eq. 3** to give an estimation of K_{a}.

3.3.2. Kinetic Analysis

The protein that is injected across the surface should after an infinite time arrive at an association equilibrium giving a signal R_{eq}, and the resonance signal R at time t during this process following injection at $t = 0$ when $R = R_0$, should in simple instances, obey the expression:

$$R(t) = R_0 + (R_{\mathrm{eq}} - R_0)[1 - \mathrm{e}^{(-k_{\mathrm{obs}}t)}]. \qquad (4)$$

Similarly, for the dissociation of the bound protein

$$R(t) = R_0 + (R_{eq} - R_0)e^{(-k_{off}t)} \quad (5)$$

assuming that the bound molecule completely dissociates from the immobilised ligand.

Consequently, the observed reaction rate k_{obs} for the interaction is given by

$$k_{obs} = k_{ass}[P] + k_{diss}. \quad (6)$$

There is thus a linear relationship between the value for k_{obs} and the total concentration of protein [P]. The value for k_{obs} can be obtained from a direct fit of the association phase using Eq. 4, or by linear regression of a semi-log plot. It follows that linear regression analysis of the dependence of k_{obs} on [P] allows calculation of k_{ass} and k_{diss} using Eq. 6. If we assume that the reaction is in fact activation controlled (were it otherwise then the association rate would be of the order of 10^9 M^{-1} s^{-1} which is well beyond the range of current SPR devices), then:

$$k_{diss} = k_{ass}K_d \quad . \quad (7)$$

Thus the equilibrium dissociation constant (K_d) can be obtained from the ratio of the off and on rates.

3.3.3. Practical Considerations

The interaction between molecules in BIAcore is monitored over time and presented as a sensorgram, i.e. a plot of response units (RU) versus time (s). To be able to extract kinetic parameters from the sensorgram, the curve must be analysed in terms of a defined mathematical model. The Langmuir isotherm is one example of a model that describes the interaction between a monovalent analyte and a ligand with *n* independent binding sites. The concentration of analyte is assumed to be constant during the sample injection phase and zero during the dissociation phase.

There are several things that must be considered for the design of a proper kinetic analysis of binding.

Multivalent analytes. It is important to avoid multivalent analytes in kinetic analysis of binding. Antibody–antigen interactions are one such example. In studies of such interactions it is preferable to immobilise the bivalent antibody and use the antigen as analyte. If the antibody is used as analyte some molecules may bridge two antigens and give rise to an avidity effect, which is not taken into account in the simple Langmuir model.

Mass transport. Mass-transport-limited binding of analyte occurs in instances where the association rate is particularly elevated and the diffusion rate of the non-immobilised molecules

is not especially fast. In this situation the interaction of the free molecules with the immobilised ligand may deplete a layer of solvent immediately surrounding the immobilised ligand such that the rate-limiting step for association now becomes the rate of repletion of this layer from the bulk solvent. Fortunately, this type of effect can fairly easily be identified by injecting the analyte at different flow rates. Generally, a change of the flow from 15 to 75 μL/min should not affect rate constants more than 5–10% if there is no significant mass transfer effect. Two practical solutions to avoid mass-transfer-limited binding are, firstly, to use a low immobilisation density, and secondly to use relatively elevated flow rates (>20 μL/min). It is generally recommended to perform kinetic measurements with R_{max} values of 100–1,000 RU, although lower values may be used if high-quality signals can be attained (*see also* **Note 3**).

Recapture of analyte may occur when an analyte dissociates from the immobilised surface. The analyte may then subsequently be recruited to an adjacent molecule. The effect of this will be to decrease the numerical values ascribed to derived dissociation constants. This effect is important in systems with fast kinetics of binding and is pronounced when the immobilisation level is high. Thus in practice the density of immobilisation must be adjusted so that the dissociation rate is independent of immobilised ligand density (**Fig. 4**). If this proves difficult then it is possible, with considerable error, to extrapolate to infinite dilution. Finally, free ligand may be included during the dissociation phase in order to

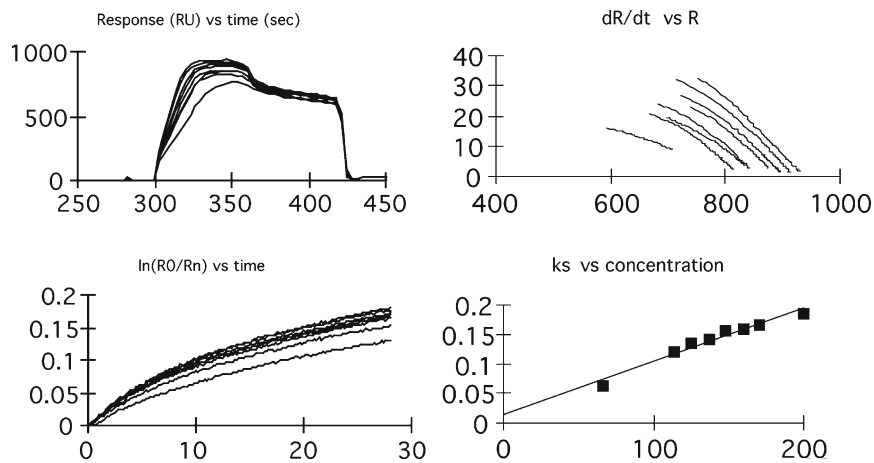


Fig. 4. Calculation of kinetic constants using the linear kinetic evaluation software. Plots used for calculating the apparent kinetic constants (see text) gave the following parameters: $k_{ass} = 9.06(\pm 0.77) \times 10^5\ M^{-1}\ s^{-1}$; $k_{diss} = 11.3(\pm 1.5) \times 10^{-3}\ s^{-1}$, yielding an apparent equilibrium dissociation constant, $k'_d = 12.0(\pm .1.9) \times 10^{-9}$ M MetJ dimer. This value compares with values ranging from 4 to 14 nM MetJ dimer obtained from other techniques (*see* **Chapters** "Filter-Binding Assays" and "Ethylation Interference Footprinting of DNA–Protein Complexes").

calculate an affinity for the competitor and thus allow an estimation of a true dissociation constant.

Association rate constants for proteins are normally in the range 1×10^4 to 1×10^7 M^{-1} s^{-1} and values $\approx 1 \times 10^6$ can generally be determined in BIAcore. Faster rate constants can be determined by the use of non-linear regression analysis procedures provided with the BiaEvaluation software 3.0. In this procedure the parameter k_t, the mass transfer constant, is fitted to the experimental curve along with the other parameters. It is related to the generally accepted mass transfer coefficient k_m by the expression:

$$k_t \approx 10^9 \times \mathrm{MW} \times k_m,$$

where MW is the molecular weight of the analyte. Values for k_t are typically of the order of 10^8 for proteins of molecular weight 50–100 kDa.

Heterogeneity. It is important to point out that heterogeneity is highly prevalent in preparations of biological molecules. This may be induced by purification schemes or can be inherent in the population of native molecules. The immobilisation procedure used to couple the ligand to the surface may also introduce heterogeneity. If the time of injection of the analyte is varied, this should not affect the shape of the dissociation curve provided that the interacting species are homogeneous. When heterogeneity of either ligand or analyte is at hand, different complexes having different properties will form at the sensor surface and their relative abundance will vary with time. This will affect the shape of the dissociation curve. The first remedy in this case is to consider an alternate immobilisation strategy, either chemical procedures for covalent attachment or capturing, e.g. by antibodies or by other means.

Scatchard analysis and kinetic binding. The equilibrium dissociation constant, K_d, can be derived from a plot of the steady-state response R_{eq} versus C (concentration of analyte). The value thus obtained should be identical with k_{diss}/k_{ass} according to Eq. 7. However, to obtain a reliable K_d from R_{eq} responses, several concentrations of analyte in the range 0.1–10 times K_d must be determined. It may take many hours or even days to determine steady-state responses, while kinetic analysis is performed within minutes at concentrations of 10–1,000 times K_d.

3.4. Results and Discussion

3.4.1. MetJ Sensorgrams

A representative SPR sensorgram for the interaction of MetJ with an immobilised, uniquely labelled 5′ biotin DNA duplex encompassing a minimal, consensus operator site (16 bp, two tandem met boxes, *see* **Fig. 1** of Chapter "Filter-Binding Assays") is shown in **Fig. 3**. Holo-repressor, i.e. MetJ in the presence of saturating amounts of the co-repressor, *S*-adenosylmethionine (AdoMet), was passed over the surface and the binding curve,

sensorgram, recorded. Dissociation of the repressor–operator complexes formed by return to running buffer in the absence of co-repressor was too rapid to permit kinetic analysis. Therefore dissociation was monitored in experiments in which consecutive injections, the first containing MetJ and AdoMet and the second only AdoMet, were made. Binding was clearly both repressor concentration and AdoMet-dependent, as expected. Furthermore, binding saturation occurred at RU levels consistent with the expected stoichiometry of one MetJ dimer per met box, which is based on both crystallographic and biochemical data.

Binding data from triplicate injections of a range of MetJ protein concentrations (60–200 nM) were evaluated using the linear kinetic evaluation package in the BIAevaluation software, as described previously *(10)* (*see* **Note 6**). Briefly, in order to calculate a value for k_{ass}, data from sensorgrams were used to plot the binding rate (dR/dT) as a function of the response (R) for each repressor concentration (**Fig. 4**). The slopes (k_s) of these resulting straight lines were then calculated. By plotting these values of k_s versus concentration, k_{ass} was obtained as the slope of the resultant straight line. The value for k_{diss} was obtained from the gradient of the plot of $\ln(R_0/Rn)$ versus $(tn–t_0)$, where R_0 and Rn are the response values obtained along the dissociation curve at times (t) 0 and n.

It was found that, although the minimal system with the two met box DNA target involves the interaction of a dimer of repressor dimers on the DNA, the association phase fits well to a simple 1:1 binding model. The dissociation phase in the presence of AdoMet, however, is clearly biphasic. We have ascribed the biphasic dissociation behaviour to the increasing effect of operator rebinding (see earlier) as the concentration of free holorepressor increased. This seemed likely because of the very high forward rate constant (~10^6 M^{-1} s^{-1}). Apparent rate constants for the dissociation reaction could be calculated for two roughly linear regions of the sensorgram: the initial pseudo-first-order dissociation of the repression complex and at later times the slower phase consisting of both dissociation and rebinding. The calculated apparent equilibrium constant for the interaction (the ratio of the apparent kinetic constants) was in the nM range, close to those determined by more traditional binding assays such as nitrocelluose filter binding and gel mobility shifts *(9)* (*see* **Chapters** "Filter-Binding Assays" and "Electrophoretic Mobility Shift Assays for the Analysis of DNA–Protein Interactions").

Detailed descriptions of the physical basis of the SPR effect and the analysis of binding data are beyond the scope of this chapter, and readers should consult appropriate reference in the literature *(1, 11, 12)*, as well as the detailed help routines in the Biacore instrument software and on the company's web site *(2)*.

4. Notes

1. It is important to remember that SPR assays using Biacore are not free solution measurements. There are certain to be effects on the interaction because of the immobilisation of one of the ligands. Nucleic acid immobilisation in our hands is the preferred option because it is easy to tether from a defined end, e.g. via a terminal biotin residue, and they are inherently flexible molecules that, in the case of DNA, show little long-range order. As a result it would be difficult to imagine that the base pairs within a short recognition sequence embedded in a tethered duplex would behave differently to the same sequence that is part of a chromosomal DNA molecule. The situation is different in the case of polymerases because polymerase action leads to torsional stress (*see* ref. 4). It is also different for proteins where long-range order within the globular structure, and molecular dynamics linked to function, are the norm. Immobilisation of proteins by chemical cross-linking produces a heterogeneous mixture of ligands for analyte binding. Use of specific protein tethers, e.g. a His-tag with an NTA-sensor chip, solves this problem but could cause unexpected effects on affinity if the tethering hinders internal motion. For these reasons it is always best to refer to the parameters measured via Biacore as 'apparent', e.g. apparent association rate constant, k_{ass}, explicitly acknowledging the fact that the values are for the interaction at the sensor chip surface rather than for the reaction in solution. We also try to study the effects of making small changes to the interacting ligands so that the principal focus is on the effect(s) of this difference, all other factors being equal.
2. We routinely carry out desorb and sanitise steps with the sensor chip docked before starting immobilisations.
3. Problems can also be caused by rapid, high-affinity interactions. The very high forward apparent rate constant observed for the binding of MetJ to its operator is typical for sequence-specific DNA–protein interactions. Care must be exercised in such cases to ensure that sensible values are obtained for the apparent rate constants. Two phenomena are relevant to this discussion. The first is the rate of mass transfer of the analyte into the region of the immobilised target. This can be increased by increasing flow rates and analyte concentrations, but the principal approach to avoid problems caused by mass transport effects is to work at very low immobilised target concentrations, avoiding depletion of the analyte concentration in the layer adjacent to the surface. Low target densities mean that the rate of analyte binding is not determined by simple mass transfer kinetics.

4. Another problem which has been reported is that the presence of a dextran support matrix results in hindered diffusion of the analyte within the layer, dramatically slowing the rate of target binding *(13, 14)*. This proposal has been tested directly for MetJ and other interactions using sensor surfaces having different lengths of dextran support or where the streptavidin is attached directly to the hydroxyalkane chain on the gold surface. There are no discernible effects from the presence of the dextran surface *(7)*, suggesting that hindered diffusion is not a significant problem under these conditions.
5. A possible problem is that the DNA may not couple onto a CM-derivatised sensor chip. Oligonucleotides are repelled by the negatively charged carboxymethyl groups on some surfaces. Such simple problems can be overcome by immobilisation in high salt concentrations. There are now a wide range of sensor chips available from the manufacturer with different surface chemistries, and customised surfaces can be designed in consultation.
6. Referees will not accept the interpretation of SPR assays. It is important to realise that the interpretation of sensorgrams relies on accurate insights into the binding events taking place on the sensor surface. The software associated with the instruments is very powerful but only does what the user tells it to do, and it mostly will always return values for the parameters associated with binding events. Therefore if the sensorgrams do not make biochemical sense, e.g. when the maximum RU signal exceeds that which is theoretically possible for an expected interaction misleading results can be obtained. For MetJ the sensorgrams recorded over concentrations where specific operator recognition is known from other techniques to occur, and which are reasonable in vivo, yield Rmax values corresponding to formation of complexes with roughly one MetJ dimer/8 bp (met box) of DNA, as expected. We might have expected that this binding reaction would need to be analysed to reflect this stoichiometry, but the data fit very well to a simple 1:1 binding model, probably because under these conditions recruitment of the second MetJ dimer to the initial complex is not rate limiting.

Acknowledgements

We thank various colleagues for their help with the SPR experiments reported here, especially Dr Isobel D. Parsons. We also acknowledge previous discussions and interactions with Dr Malcolm Buckle, and Dr Francis Markey for valuable comments on the presentation of Biacore technology.

References

1. Persson, B., Buckle, M. M. and Stockley, P. G. (2000). Kinetics of DNA interactions studied by surface plasmon resonance spectroscopy. In *DNA–Protein Interactions: A Practical Approach.* (Editors, Travers, A. and Buckle, M. M.), Chapter 19, pp. 257–279, Oxford University Press, Oxford, UK.
2. Biacore web page (General Electric Company); http://www.biacore.com.
3. Rich, R. L. and Myzska, D. G. (2006). Survey of the year 2006 commercial optical biosensor literature. *J. Mol. Recognit.* **20**, 300–366.
4. Buckle, M., Williams, R. M., Negroni, M., and Buc, H. (1996). Real time measurements of elongation by a reverse transcriptase using surface plasmon resonance. *Proc. Natl Acad. Sci. U. S. A.* **93**, 889–894.
5. Fisher, R. J. and Fivash, M. (1994). Surface plasmon resonance based methods for measuring the kinetics and binding affinities of biomolecular interactions. *Curr. Opin. Biotechnol.* **5**, 389–395.
6. Di Primo, C. and Lebars, I. (2007). Determination of refractive index increment ratios for protein–nucleic acid complexes by surface plasmon resonance. *Anal. Biochem.* **368**, 148–155.
7. Parsons, I. D. and Stockley, P. G. (1997). Quantitation of the E. coli methionine repressor–operator interaction by surface plasmon resonance is not affected by the presence of a dextran matrix. *Anal. Biochem.* **254**, 82–87.
8. Somers, W. S. and Phillips, S. E. V. (1992). Crystal structure of the met repressor operator complex at 2.8 Å resolution: DNA recognition by β strands. *Nature* **346**, 586–590.
9. Phillips, S. E. V., Manfield, I., Parsons, I. D., Davidson, B., Rafferty, J. B, Somers, W. S., Cohen, G., Saint-Girons, I. and Stockley, P. G. (1989). Cooperative tandem binding of *met* repressor of *Escherichia coli. Nature* **341**, 711–715.
10. Parsons, I. D., Persson, B., Mekhalfia, A., Blackburn, G. M., and Stockley, P. G. (1995). Probing the molecular mechanism of action of co-repressor in the *Escherichia coli* methionine repressor–operator complex using surface-plasmon resonance (SPR). *Nucleic Acids Res.* **23**, 211–216.
11. Schuck, P. (1997). Reliable determination of binding affinity and kinetics using surface plasmon resonance biosensors. *Curr. Opin. Biotechnol.* **8**, 498–502.
12. Schuck, P. and Minton, A. P. (1996). Kinetic analysis of biosensor data: elementary tests for self-consistency. *Trends Biochem. Sci.* **21**, 458–460.
13. Schuck, P. and Minton, A. P. (1996). Analysis of mass transport-limited binding kinetics in evanescent wave biosensors. *Anal. Biochem.* **240**, 262–272.
14. Schuck, P. (1996) Kinetics of ligand binding to receptor immobilized in a polymer matrix, as detected with an evanescent wave biosensor. I. A computer simulation of the influence of mass transport. *Biophys. J.* **70**, 1230–1249.

INDEX

D

E

F

G

H

I

L

M

N

O

P

R

S

T

U

W