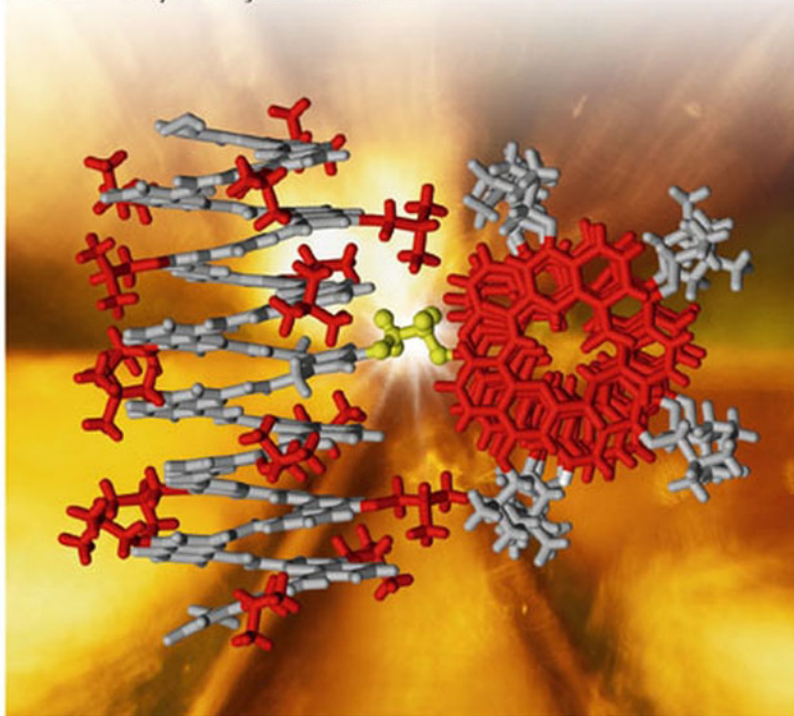Edited by Stefan Hecht and Ivan Huc

WILEY-VCH

Foldamers

Structure, Properties, and Applications

Foreword by François Diederich

Foldamers

Edited by
Stefan Hecht and Ivan Huc

1807–2007 Knowledge for Generations

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce
President and Chief Executive Officer

Peter Booth Wiley
Chairman of the Board

Foldamers

Structure, Properties, and Applications

Edited by
Stefan Hecht and Ivan Huc

Foreword by
François Diederich

The Editors

Prof. Dr. Stefan Hecht
Humboldt University
Institute of Chemistry
Brook-Taylor-Str. 2
12489 Berlin
Germany

Dr. Ivan Huc
Institut Européen de Chimie
et Biologie
2 rue Robert Escarpit
33607 Pessac Cedex
France

Foreword

Biopolymers adopt distinct conformations in order to express functions that are key to life. Examples are the sheet, helix, and turn motifs of proteins, the double and triple helix, quadruplex, or hairpin motifs of nucleic acids, or the helical structures of carbohydrates such as starch. Without these preferred structures, expression and translation, recognition, catalysis, and transport in living systems could not be achieved. While chemists have learned since the middle of last century how to analyze conformational preferences of small molecules and to apply this knowledge to regio- and stereoselective chemical transformations, the control of the three-dimensional structure – and thereby the function – of synthetic oligomers and polymers has only recently become a hot research topic.

Foldamers, i.e. synthetic oligomers with distinct conformational preferences, are at the interface of covalent (molecular) and noncovalent (supramolecular) chemistry. Their investigation will enable chemists to develop geometrically defined oligomers that promise to rival biopolymers in their function and application. Increasingly, foldamers with covalent or supramolecular backbones are switchable under external stimuli between two defined stable states, can be prepared by dynamic combinatorial synthesis, or can assemble to functional foldamer complexes. They will find use as novel biomimetic receptors and catalysts, light and energy capturing and storage devices, delivery and transport systems for synthetic drugs and membrane-impermeable biomolecules, and materials that interface with biological tissues.

The construction of foldamers starts from small, intelligently programmed monomeric modules, which contain the information to generate oligomers with distinct three-dimensional structures. The geometries are controlled by a variety of parameters, including backbone conformational preferences, backbone interchromophoric interactions (such as aromatic–aromatic interactions), side chain interactions, solvophobic interactions, metal ion coordination, and H-bonding molecular recognition. These parameters are logically analyzed in the monograph, resulting in useful design protocols. Functions of synthetic foldamers and their relationships to biopolymers are described for systems spanning from biomimetic oligomers to π-conjugated oligomers. I strongly recommend this monograph to all academic and industrial researchers interested in fascinating

perspectives for future chemical research; it will also take its place in modern graduate student education.

Zurich, September 15, 2006   *François Diederich*

Contents

4   Foldamer Hybrids: Defined Supramolecular Structures from Flexible Molecules   *109*
Carsten Schmuck and Thomas Rehm

5   Control of Polypeptide Chain Folding and Assembly   *147*
Rajkishor Rai and Padmanabhan Balaram

9   Protein Design   *267*
Jean-Luc Jestin and Frédéric Pecorari

Preface

Research in molecular chemistry is essentially devoted to understanding the relationships between chemical structures and their properties and functions. One key parameter of a molecule's structure is its overall shape: its three-dimensional conformation. It is thus no surprise that conformational analysis and strategies to control conformation lie at the heart of many disciplines. Not unexpectedly, Nature has evolved the ultimate realization of function based on controlling and altering conformation of its molecular machinery. Prominent examples include information storage, duplication and translation using DNA and ribosomes and cooperative oxygen transport by hemoglobin. These achievements are based on large and complex yet remarkably defined structures, which are obtained through the folding of long polymeric chains and a subtle balance of noncovalent forces. On the contrary, many synthetic systems with defined conformations rely on covalent restriction of the molecules' flexibility. Pre-organization has long been a cornerstone of molecular design, as exemplified by the fact that most drugs are cyclic or macrocyclic. However, during the past decade, chemists have been inspired by self-organized natural systems and have gained increasing knowledge of how to design molecular strands, so-called foldamers, that are capable of adopting well-defined folded conformations.

Foldamers have been loosely defined by Gellman as ''polymers with a strong tendency to adopt a specific compact conformation'' or more restrictively by Moore as ''oligomers that fold into a conformationally ordered state in solution, the structures of which are stabilized by a collection of noncovalent interactions between nonadjacent monomer units''. Usage of the term foldamer has mostly been targeted to synthetic oligomers (see Chapters 1–4). Artificial folded structures, which in fact are covered by the same definition, were studied extensively long before the term foldamer was coined and include synthetic (non-natural) α-peptide sequences (Chapter 5), artificial proteins (Chapter 9), nucleic acids (Chapter 10), and helical polymers (Chapters 11 and 12), among others.

The aim of this book is to cover the breadth of the rapidly developing field of foldamer research and to unite the different aspects and schools by illustrating the generality of underlying concepts. The central theme is the synthetic construction and functional exploitation of chain molecules with a conformational preference. While the first part of the book is devoted to foldamer design

concepts, the second part covers the use of conformational control to create chemical entities with beneficial functions in biology and materials science.

Synthetic oligomers can be divided into four major families (Chapters 1–4) according to the factors that dominate folding, i.e. local rotational restrictions, interactions between sites remote in the sequence, solvophobic effects, and assembly/hybridization. This division, however, is not exclusive. Folding is often the result of a combination of these factors and, in all cases, requires intrinsic backbone rigidity. Other factors, such as electrostatic and steric repulsions, may play a less visible but no less important role in reducing the accessible (unfolded) conformational space. Experimental studies of synthetic oligomers provide insight into thermodynamics and sometimes kinetics of folding events. In parallel, molecular modeling has advanced to become a useful tool that can aid conformational analysis and ''observe'' missing links, as well as predict preferred folded conformations (Chapter 6). The design of new folding backbones and subsequently, but not necessarily, new functions, may be termed a ''bottom-up approach'' to foldamers (Chapters 1–5). In contrast, ''top-down approaches'' (Chapters 9, 10) start from the well-known folding behavior of proteins and polynucleotides and, through directed evolution techniques or through rational design, target functions while simplifying structures. The dynamic nature and flexibility of foldamers arise from the deliberate utilization of various noncovalent interactions for structure formation. It gives rise to adaptability and responsiveness as key requirements for efficient recognition (''induced fit'') and hence functions (e.g. in sensing). This flexible yet defined shape of foldamer-based chemical systems leads to a large variety of applications ranging from biological, such as inhibitor design and antimicrobial activity (Chapters 7–9), to the materials and nano sciences, such as biomineralization/composite materials, RNA/DNA architectonics, sensors, and functional interfaces (Chapters 7, 10–13).

It is quite surprising to note that only 15 years ago, molecular folding was thought to be associated solely with biopolymers, as if natural building blocks had characteristics unique to themselves. The huge body of recent work on foldamers has clearly demonstrated that multiple 'abiotic' backbone families are able to adopt folded secondary motifs as well. Nowadays, biopolymers can be viewed as one – arguably very important – class of folding molecules among many others. The secondary folding motifs discovered thus far in synthetic backbones do not differ much from those of biopolymers. Turns, helices, linear strands, and multi-stranded systems, such as double helices and sheets, seem to be the most common – perhaps universal – folding motifs. Alternate folding modes, for example knots, are possible but much less common. Furthermore, synthetic systems will undoubtedly benefit from utilizing Nature's hierarchical organization involving control over local conformation, i.e. rotation about bonds, and orientation in larger structures thereby controlling global conformation, i.e. primary → secondary → tertiary → quaternary structure evolution.

Much has been achieved; yet foldamer chemistry is still a young field and a great deal is to be expected. For instance, tertiary abiotic folds with functions remain to be seen and constitute one of the main challenges ahead. The long-term

prospect of building fully synthetic analogs of proteins is not illusionary, though it will require even more powerful design and synthetic strategies than those currently at hand. In this respect, combining bottom-up and top-down approaches, strategies that have thus far evolved independently, may be a promising way to follow. While foldamer-based biomimicry certainly provides deeper insight into Nature's mysteries, it also allows function to be explored in a non-natural context using the increased structural diversity and chemical robustness of foldamers. The potential benefits of this endeavor are enormous. Native folded biopolymers efficiently perform a multitude of functions using sequences based on relatively small alphabets – four nucleobases and roughly 20 amino acids. As shown in artificial proteins and nucleic acids, the same alphabets can be used to achieve numerous non-natural functions. The prospect of extending such alphabets to abiotic folding motifs, either already described in synthetic oligomers or yet to be discovered, thus opens the opportunity for countless applications.

We hope that this book will serve as both inspiration to the non-expert as well as a valuable resource for the specialist and bring together scientists from different disciplines to communicate with each other, engage in a joint effort to unravel one of Nature's mysteries, and create exciting new opportunities for future discoveries.

Last but not least, we want to express our sincere thanks to the authors of the individual chapters for their unique contributions of exceptionally high quality. Furthermore, we are indebted to our students, coworkers, and colleagues, with whom we had the privilege to interact and share the interest and enthusiasm for this exciting field of interdisciplinary research. We also want to thank the Wiley–VCH team, in particular Elke Maase for establishing this fruitful endeavor as well as Manfred Köhl and Steffen Pauly for their professional assistance during the editing and publishing process.

April 2007 *Stefan Hecht and Ivan Huc*
Mülheim an der Ruhr and Bordeaux

List of Contributors

Padmanabhan Balaran
Molecular Biophysics Unit
Indian Institute of Science
Bangalore 560012
India

Jorge Becerril
Yale University
PO Box 208107
New Haven
CT 06520-8107
USA

Richard Cheng
Department of Chemistry
University at Buffalo
The State University of New York
Buffalo
NY 14260-3000
USA

Arkadiusz Chworos
Chemistry & Biochemistry
University of California
Santa Barbara
CA 93106-9510
USA

Jeroen J. L. M. Cornelissen
Radboud University Nijmegen
Department of Organic Chemistry
Institute for Molecules and Materials
Toernooiveld 1
6525 ED Nijmegen
The Netherlands

Louis A. Cuccia
Department of Chemistry and
Biochemistry
Concordia University
1455 DeMaisonneuve Blvd. Ouest
Montréal
Québec, H3G 1M8
Canada

Zrinka Gattin
Laboratorium für Physikalische
Chemie
ETH Hönggerberg, HCI
CH-8093 Zürich
Switzerland

Gilles Guichard
Institut de Biologie Moléculaire et
Cellulaire
15, rue René Descartes
67084 Strasbourg Cedex
France

Andrew D. Hamilton
Yale University
PO Box 208107
New Haven
CT 06520-8107
USA

Ivan Huc
Institut Européen de Chimie et
Biologie
2, rue Robert Escarpit
33607 Pessac Cedex
France

Sebastian Hartwig
Humboldt University
Institute of Chemistry
Brook-Taylor-Str. 2
12489 Berlin
Germany

Stefan Hecht
Humboldt University
Institute of Chemistry
Brook-Taylor-Str. 2
12489 Berlin
Germany

Luc Jaeger
Chemistry & Biochemistry
University of California
Santa Barbara
CA 93106-9510
USA

Jean-Luc Jestin
Unité de Chimie Organique
Institut Pasteur
25-28 rue du Dr. Roux
75015 Paris
France

Phillippe le Grel
Institut de Biologie Moléculaire et
Cellulaire
15, rue René Descartes
67084 Strasbourg Cedex
France

Katsuhiro Maeda
Department of Molecular Design and
Engineering
Graduate School of Engineering
Nagoya University
Furu-cho, Chikusa-ku
Nagoya 464-8603
Japan

Gerald A. Metselaar
Radboud University Nijmegen
Department of Organic Chemistry
Institute for Molecules and Materials
Toernooiveld 1
6525 ED Nijmegen
The Netherlands

Jeffrey S. Moore
The Beckman Institute
University of Illinois
405 N. Mathews Ave
Urbana
IL 61801
USA

Roeland J. M. Nolte
Radboud University Nijmegen
Department of Organic Chemistry
Institute for Molecules and Materials
Toernooiveld 1
6525 ED Nijmegen
The Netherlands

Matthijs B. J. Otten
Radboud University Nijmegen
Department of Organic Chemistry
Institute for Molecules and
Materials
Toernooiveld 1
6525 ED Nijmegen
The Netherlands

Frédéric Pecorari
Unité de Biochimie Structurale
Institut Pasteur
25-28 rue du Dr. Roux
75015 Paris
France

Rajkishor Rai
Molecular Biophysics Unit
Indian Institute of Science
Bangalore 560012
India

Thomas Rehm
Institute of Organic Chemistry
University of Würzburg
Am Hubland
97074 Würzburg
Germany

Johanna M. Rodriguez
Yale University
PO Box 208107
New Haven
CT 06520-8107
USA

Alan E. Rowan
Radboud University Nijmegen
Department of Organic Chemistry
Institute for Molecules and
Materials
Toernooiveld 1
6525 ED Nijmegen
The Netherlands

Ishu Saraogi
Yale University
PO Box 208107
New Haven
CT 06520-8107
USA

Carsten Schmuck
Institute of Organic Chemistry
University of Würzburg
Am Hubland
97074 Würzburg
Germany

Hennie Valkenier
University of Groningen
Chemistry Department
Nijenborgh 4
9747 AG Groningen
The Netherlands

Jan van Esch
University of Groningen
Chemistry Department
Nijenborgh 4
9747 AG Groningen
The Netherlands

Wilfried F. van Gunsteren
Laboratorium für Physikalische
Chemie
ETH Hönggerberg, HCI
CH-8093 Zürich
Switzerland

Eiji Yashima
Department of Molecular Design and
Engineering
Graduate School of Engineering
Nagoya University
Furu-cho, Chikusa-ku
Nagoya 464-8603
Japan

Yan Zhao
Department of Chemistry
Iowa State University
1605 Gilman Hall
Ames
IA 50011-3111
USA

Part 1
Structure: Foldamer Design Concepts

1

Foldamers Based on Local Conformational Preferences

Ivan Huc and Louis Cuccia

1.1
Introduction

Folding, as it occurs in biopolymers, refers to the prevalence of well-defined conformers in solution and, in most cases, to proximity in the folded state between chemical groups that are remote in the molecules' backbones. As illustrated in the first three chapters of this book, a multitude of non-natural folding oligomeric molecules – termed *foldamers* – have been designed, prepared and characterized [1–3]. The factors that promote folding of a linear molecular strand are manifold: specific attractive or repulsive interactions between sites remote in oligomeric sequences, solvophobic effects, local conformational restrictions or any combination thereof. This chapter deals with arguably the most important route to promote well-defined conformations within oligomers. It consists of introducing backbone rigidity through local conformational preferences that stabilize folded structures and also reduce the number, and raise the energy level, of non-folded states. Even when other strong effects are at play as, for example, hydrogen bonding (see Chapter 2) and solvophobic effects (see Chapter 3), their efficiency at promoting folding relies on the premise that the molecular backbone is sufficiently rigid so that the entropic cost of adopting a well-defined conformation is not excessive.

In this chapter, we focus on foldamers whose folding occurs mainly due to backbone rigidity, determined locally at the molecules' rotatable bonds, in the absence of other strong factors. Backbone rigidity can be imparted in many ways and occurs in quite diverse families of folding oligomers. It is not our intention to make an exhaustive presentation, but rather to select representative examples of these families. We may also add that determining which factors dominate a given folding event is often not a clear cut issue: some of the examples presented in this chapter may appear in Chapters 2 or 3 and vice versa. It remains that several general characteristics emerge from the diverse families of "rigid" foldamers presented here: the first is a high level of predictability of the folded conformation, be it by advanced computational means (see Chapter 6) or by a simple paper

sketch, structure prediction in rigid foldamers is, in many cases, reliable. A second general characteristic is a relatively low solvent dependence of the prevailing folded conformations in solution. A third important aspect is that, owing to the narrow distribution of conformations in solution, "rigid" foldamers tend to be much more crystalline than others: a majority of foldamer X-ray crystal structures belong to the families of molecules presented in this chapter.

1.2
Rigidly Locked Molecules

Different levels of rigidity can be imparted to a molecular backbone leading to various degrees of "foldability". Folding is a dynamic process that supposes, *a priori*, an ability to unfold. In extreme cases, the nature of intramolecular connections may result in a single-well energy landscape corresponding to a rigidly locked conformation. Although these molecules may not be considered foldamers *per se* because they show poor capacity to unfold, they do provide a firm starting point for this chapter.

[*n*]Helicenes are π-conjugated helical molecules consisting of *n* all-*ortho* annulated benzene rings (for example, [9]helicene; Fig. 1.1a) [4]. At first glance, one would expect helicenes with more than five fused aromatic rings to be rigidly locked into either a right- or left-handed helical conformation. However, a distribution of molecular deformations over a large number of bonds does allow for racemization of helicenes ranging from hexahelicene to nonahelicene [5, 6]. In geländer helices the building blocks are perpendicular compared with helicenes. In the case of a bridged *para*-terphenylophane geländer (Fig. 1.2b), the molecule
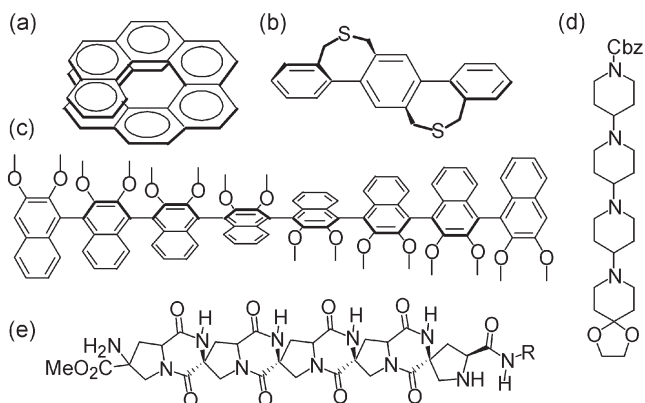
Fig. 1.1 Rigidly locked molecules: (a) helicenes; (b) geländer helices; (c) molecular ribbons; (d)–(e) molecular rods. These molecules are termed oligomeric in the sense that there is a repeating motif.

can be locked into either a right- or left-handed ''spiral staircase'' conformation [7]. An intriguing variation on oligoparaphenylene helices are molecular ribbons reported by Fuji et al. (Fig. 1.2c). The locking mechanism in these configurationally defined oligonaphthalene derivatives relies on restricted rotation about single bonds due to unfavorable steric barriers (that is, atropisomerism) [8]. Molecular rods are rigid ''unfoldable'' oligomers with well-defined molecular dimensions. For example, Semetey et al. have reported the synthesis of a series of water-soluble oligopiperidine molecular rods with as many as ten piperidine rings (Fig. 1.2d). There is NMR evidence for a chair conformation of each piperidine ring with each piperidine unit in an equatorial position with combined nitrogen inversion and chair–chair inversion throughout the well-defined backbone. According to the authors, there is rotation about the C–N bonds, but this results in only small deviations of the linear geometry of the molecule [9]. A more rigid spiro-linked molecular rod was reported by Levins et al. This molecular scaffold was elegantly prepared by rigidifying a flexible oligomer *via* the formation of two diketopiperazine rings (Fig. 1.2e) [10]. With regards to molecular rods, some molecules can have unrestricted rotations along the backbone but do not fold for geometric reasons – that is, if the backbone is linear (180° connectivity) no amount of rotation will cause molecular folding (however molecular ''twisting'' can occur). This is the case for oligo(*para*-phenylene ethylene) molecular wires reported by Schumm et al. [11].

1.3
Predictable Foldamers

Fully predictable foldamers may be defined as oligomeric structures that possess numerous rotatable bonds – in contrast with the oligomers shown in Fig. 1.1 – and that may, in principle, envelop a vast array of conformations, but whose conformational space is narrowed down to a single conformer because a well-defined preference exists at each rotatable bond. There is no need to explore the conformational space accessible to the entire molecule to determine its most stable conformation since it primarily results from local conformational preferences. Taking as an illustration the Ramachandran plots used to map the torsion angles corresponding to stable folded conformations in peptides (see Chapters 2 and 5), the experimentally encountered values for the torsion angles in fully predictable foldamers are reduced to very small areas. A simple example is the secondary amide bond: rotation about this bond is possible but the equilibrium between the possible rotamers is completely shifted in favor of the *transoid* conformation. *Cis*-secondary amides are rarely observed and generally not considered in peptide structures: including a third dimension in Ramachadran plots to describe amide bond rotation is of no use.

Literature pertaining to fully predictable foldamers is already abundant and is steadily growing. Because of their features, they are archetypical structures among the molecules described in this chapter: their structure is very well-

defined, predictable, and these compounds are highly crystalline (thus easily characterized).

1.3.1
Local Conformational Control

Though the definition given above of fully predictable foldamers seems quite general, the families of molecules that fall in this category are, in fact, rather homogeneous and almost all consist of π-conjugated systems – for example, aryls, amides, esters or ureas – connected by single bonds. Evidently, π-conjugation is a very efficient means of restricting rotation about a single bond as it stabilizes conformers where two π-systems are close to being coplanar, allowing the π-orbitals of sp^2-hybridized atoms to overlap. This effect is stronger in true π-conjugated systems where *para-* or *ortho-*connectivity between aryl rings gives rise to resonance, but it remains substantial even for *meta-*connected – cross-conjugated – systems. Upon effecting a 180° rotation about the single bond between two π-systems, two degenerate conformers may be possible, but degeneracy is easily lifted. Fig. 1.2 shows a number of conformational equilibria for which a single stable conformation exists at a single bond connecting two π-conjugated systems. These examples are representative of the families of foldamers described in Sections 1.3.2 and 1.3.3, but it is clear that many alternate schemes could be devised along the same lines.

Thus, aryl–CONH single bonds adopt *syn* conformations when the aryl ring possesses a hydrogen bond donor *ortho* to the amide group (Fig. 1.2a). The hydrogen bond donor may be an exocyclic OH [12] or NH [13–17], or an endocyclic N$^+$H [18]: all three moieties hydrogen bond to the amide carbonyl and repel the amide proton. Conversely, the *anti* conformation of aryl–CONH linkages is stabilized by hydrogen bond acceptors on the aromatic ring (Fig. 1.2b) which attract the amide proton and repel the amide oxygen. The most common groups used as hydrogen bond acceptors are endocyclic nitrogen atoms [13, 15–42] or exocyclic ether oxygen atoms [39, 43–55], but other functional groups have also been shown to be effective, for example, exocyclic fluorine [56], imino nitrogen [57], *N*-oxide oxygen [14] and phenolate oxygen [12].

Conformations about aryl–NHCO linkages are controlled in a very similar way by hydrogen bond donors or acceptors *ortho* to the amide function on the aromatic ring. For instance, a *syn* conformation is favored by a proton [18, 37, 38, 58–60] or a metal ion [40, 61, 62] that can coordinate to an amide carbonyl (Fig. 1.2c). An *anti* conformation is favored when a hydrogen bond acceptor is introduced which binds to the amide proton and/or repels the amide oxygen (Fig. 1.2d). Effective acceptors include exocyclic ether oxygen [12, 27, 28, 40, 44, 45, 47, 49–55] or sulfur [43, 63–65] atoms; exocyclic fluorine [56]; endocyclic nitrogen [18–21, 29, 33, 37, 38, 41]; exocyclic *N*-oxides [66] or phenolates [12]; exocyclic carbonyl oxygen [16, 17, 32, 39, 48, 67–70]; as well as sp^2 nitrogen atoms belonging to connected [71] or fused [24–26, 29, 30, 34–36] aromatic rings.
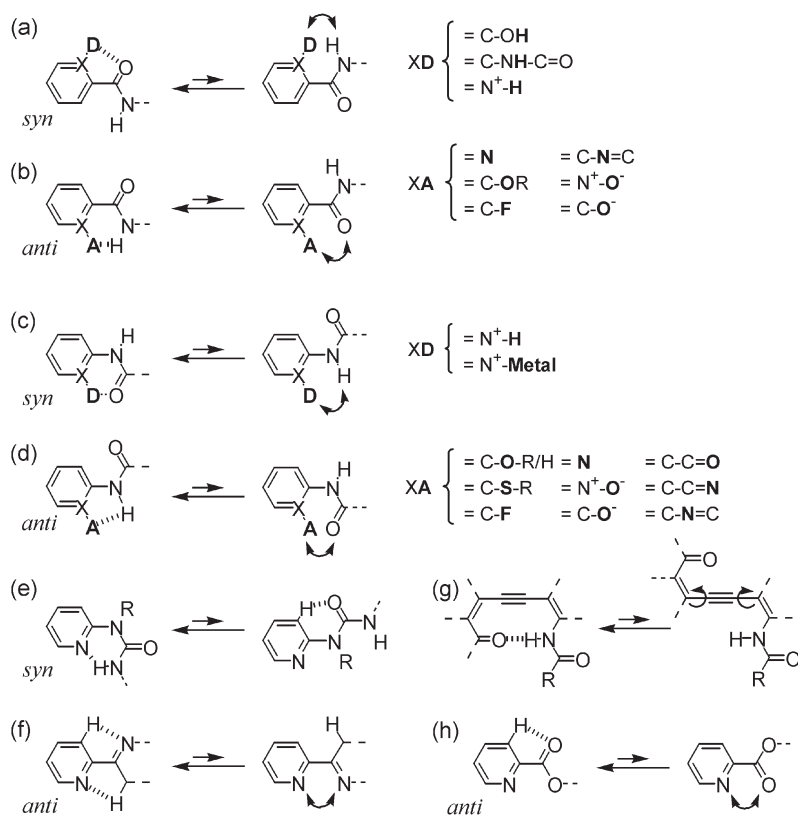
Fig. 1.2 Stabilization of well-defined rotamers through local attractive (hashed lines) and repulsive (double headed arrows) interactions. The stabilized conformer is shown on the left side of each equilibrium. ''X–A'' and ''X–D'' correspond to hydrogen bond donor and hydrogen bond acceptor moieties, respectively. When ''X–D'' does not bear a hydrogen but a metal ion, it is an electron acceptor. (a)–(b) aryl–CONH linkage; (c)–(e) aryl–NHCO linkage; (f) aryl–aryl, aryl–imine and aryl–hydrazone linkages; (g) restricted rotation about an acetylene bond; (h) aryl–carboxyl bond where sp^3 and sp^2 hybridized oxygen atoms are discriminated.

These numerous examples are illustrative of the diversity of aromatic amide foldamers. The attractive interactions that stabilize a given rotamer often consist of bifurcated and not necessarily strong hydrogen bonds. However, the role of repulsive interactions (double headed arrows in Fig. 1.2) should not be underestimated. Although their exact contribution has not been quantified, it is likely that their strength is no less than that of hydrogen bonds.

Related structures involve aromatic sulfonamides [72], hydrazides [73], or diazo groups [42]. In some cases, conformational preference may appear as less obvious, but nevertheless exists. For example, in the 2-pyridyl-carboxyl linkage shown

in Fig. 1.2h, both the sp^2 and sp^3 hybridized oxygen atoms of the carboxyl function are *a priori* involved in repulsive electrostatic interactions with the neighboring endocyclic nitrogen atom. However, a stronger repulsion involving the carbonyl oxygen leads to a stabilization of the *anti* conformer [26]. For the 2-pyridyl ureas shown in Fig. 1.2e, attractive interactions exist both for the *syn* and *anti* conformers [74–78]. While both conformers are at equilibrium when the urea moiety is unsubstituted [76, 79, 80] the *cis* urea conformation and thus the *syn* conformation of the aryl–urea linkage are stabilized by an *N*-alkyl group [81] (see also Section 1.4.1).

The schematic equilibrium shown in Fig. 1.2f covers a number of aryl–aryl connections between pyridine, and other aza-aromatics, as well as aryl–hydrazone linkages. These have been used extensively by the group of Lehn to produce a wide variety of foldamers [82–98].

The restricted rotations presented above are all well documented in the literature. In particular, crystallographic data are available in most cases. The global trend that emerges from these data is a remarkable reliability of the predicted preferred conformations. Only three crystal structures were found showing significant deviations from the predicted more stable conformations. As shown in Fig. 1.3, they all concern the anthranilamide motif and represent interesting snapshots of presumably ill-folded conformations [14, 15].

Finally, rotamers stabilized by local interactions that involve weak π-conjugation or no conjugation at all should also be mentioned. This is the case for the equilibrium shown in Fig. 1.2h where rotation about a diaryl–acetylene linkage is restricted by hydrogen bonding between functions on either side of the triple bond [99–101]. This motif has been introduced within some solvophobically driven foldamers such as oligophenylethynylenes (see Chapter 3) to stabilize the folded conformations in solvents that do not promote solvophobic effects. Other examples include 3,5-linked oligopyrrolin-4-one [102] and some constrained dipeptides [103] (see Section 1.3.3).
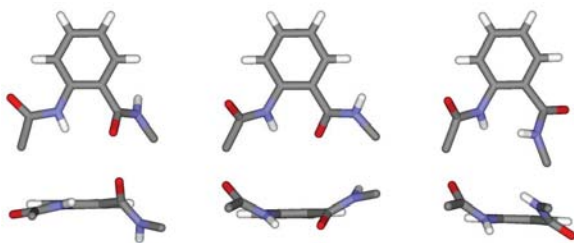
Fig. 1.3 Top view and front view of three anthranilamide motifs observed in the crystalline state showing substantial deviations from the canonical planar six-membered NH···O=C intramolecularly hydrogen bonded ring [14, 15].

1.3.2
Folded Conformations of π-conjugated Systems

The interactions that drive the local conformational preferences in fully predict-able foldamers are largely two-dimensional (2-D); this can be attributed to the pla-nar nature of the repeating unit(s), the linkages, and the relative positioning of consecutive units (Fig. 1.2). The curvature (or lack thereof) resulting from this 2-D folding relies on the nature and geometry of intramolecular hydrogen bond-ing, electrostatic repulsions, and/or steric interactions. The crystal structures of short sequences of monomers actually often reveal rigorously flat arrangements. It is this 2-D nature that makes structure prediction so easy; simple molecular modeling or even paper sketching allows one to determine the relative orientation of the consecutive units according to the conformational preferences shown in Fig. 1.2. When several units are connected within a sequence, the overall confor-mation results from the sum of each preferred rotamer.

1.3.2.1 **Crescents and Helices**

The relative orientation of consecutive units within even a very short oligomer may result in a stable, planar, curved, crescent-like conformation. In this respect, fully predictable foldamers strongly differ from most other foldamer families where strand bending is hardly possible in the absence of contacts between func-tions remote in the sequence as found in helices or hairpin turns. As the length of a crescent-like oligomer is increased, deviation from planarity is imposed by steric repulsions between the extremities of the crescent, and the 2-D structure becomes three-dimensional, giving rise to a helical conformation. As shown in Fig. 1.4, such helices have been designed and characterized using various types of monomeric units and certainly represent the flagship of fully predictable fol-damers. The top views of the structures of Fig. 1.4 are illustrative of the structural differences between the helices highlighting the different number of units per turn and the variable inner diameters. The side views of the structures are re-markably similar despite the completely different nature of backbone composi-tions: in all cases, the helical pitch of these π-conjugated systems equals the thick-ness of one aromatic ring. Most of these oligomers are constructed by alternating two types of symmetrical units and do not possess a helical polarity such as that of peptides. However, the examples shown in Figs. 1.4a and d are built using di-rectional units (3-amino-benzoic acid [45] and 8-amino-2-quinolinecarboxylic acid [97], respectively), but in both cases, these have been connected to a central sym-metrical spacer in order to double the oligomer length in the last synthetic step.

Helices such as those of Fig. 1.4 emerge primarily from the preferred confor-mations of each rotatable bond. Yet, intramolecular π–π stacking between aryl moieties clearly provides additional conformational stabilization. It seems, how-ever, that π–π stacking is not directional to the extent that it has an influence on the actual strand curvature so as to promote specific favorable contacts between aryl groups: curvature in helices is similar to that of crescents where no stacking
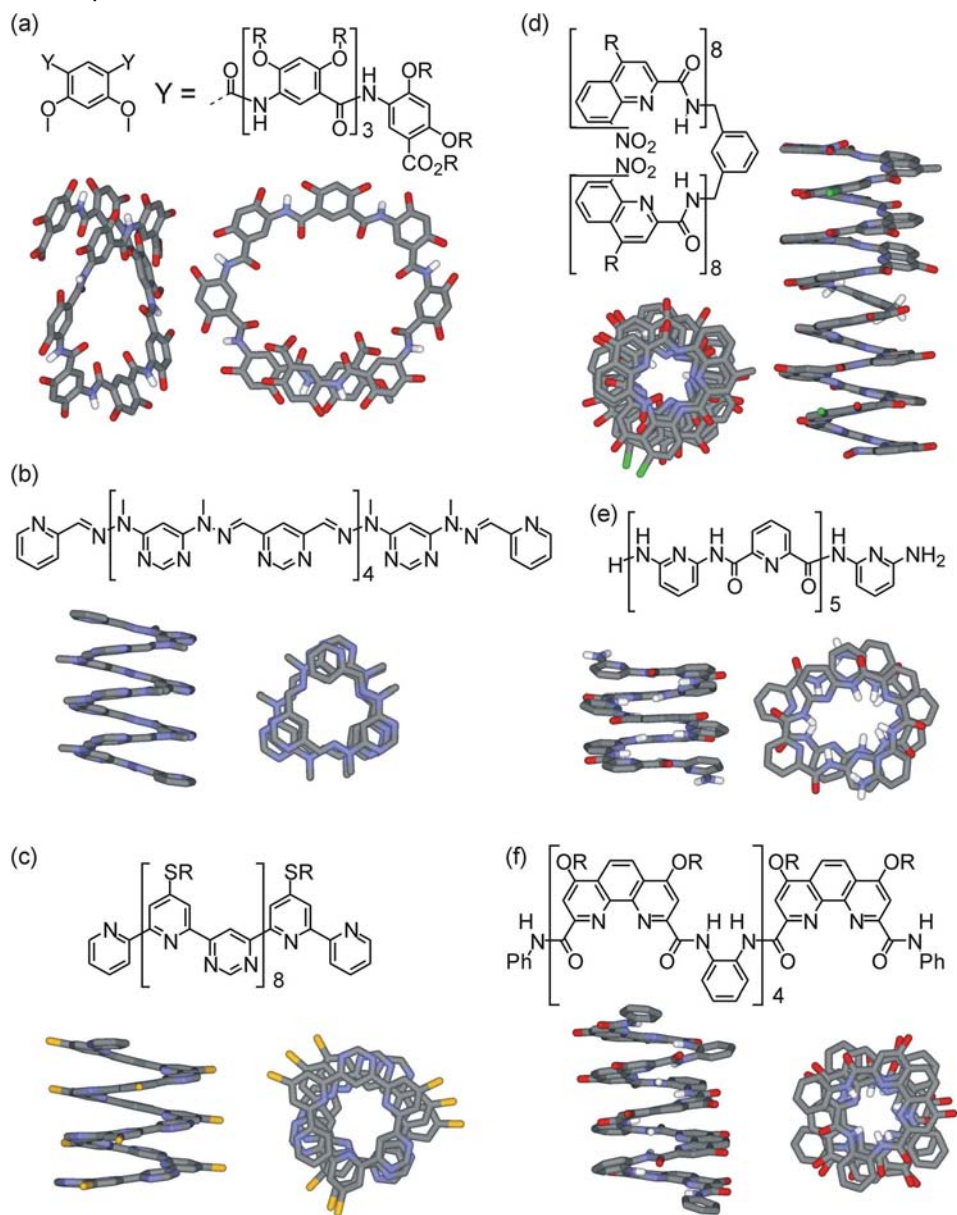
Fig. 1.4 Molecular formulae and X-ray structures – all at the same scale – of helical fully predictable foldamers. References for structures (a), (b), (c), (d), (e), and (f) are found in [45], [97], [93], [26], [20], and [31], respectively. Side chains, including solvent molecules and non-amide hydrogens have been omitted for clarity.

is involved. π–π stacking becomes critical in those cases where local conformational control is not complete [31, 42]. Considering, for example, the oligomer shown in Fig. 1.4f, conformations about aryl–CONH bonds are well defined [31] (see Fig. 1.2b) as opposed to conformations about aryl–NHCO bonds. Helical folding is thus partly driven by π–π stacking and, indeed, is promoted by protic solvents such as methanol [31]. Such foldamers are closely related to those presented in Chapter 3.

The stability of the helical foldamers shown in Fig. 1.4 is illustrated by the fact that their folding apparently occurs in any type of solvent. Though most solution conformations have been characterized in chlorinated, aromatic, or polar non-protic solvents [24, 36, 53], folding has been observed in water as well [30]. Rates of helix handedness inversion are considerably longer than for most other types of oligomers, which also suggest high conformational stability: several hours for the structure shown in Fig. 1.4d [26]. The stability and the compact shape of the helices are likely the origin of their high propensity to crystallize, and crystallography has clearly emerged as a method of choice to structurally characterize these oligomers.

Depending on the geometry of the repeating motifs (that is, the relative orientation between consecutive units in their preferred conformation), the conformation of fully predictable foldamers can display positive, negative, or zero curvature. For example, two 0° units, three 60° units, or six 120° units would be required to complete one turn of a helical foldamer. On the other hand, units that present an angle of 180° or with an overall zig-zag conformation (see Section 1.3.2.2) will not combine to form helical conformations. In essence, one can tune helix diameter by changing the curvature imparted by each monomer in the sequence. Illustrations of this remarkable feature of fully predictable foldamers are given in Fig. 1.5 for aza-aromatic and aromatic amide oligomers. The propensity of oligoheterocyclic pyridine–pyrazine (Fig. 1.5a) [90, 91] pyridine–pyrimidine (Fig. 1.5b) [85, 89, 94] and pyridine–pyridazine (Fig. 1.5c) [86, 87] sequences to form a helical motif stems from structure-inducing "codons" that enforce helical winding due to the strongly favored *transoid* conformation of the α,α'-interheterocyclic bonds (Section 1.3.1; Fig. 1.2f). However, the curvature (and thus diameter) imposed by the position of the nitrogen(s) in the heterocyclic monomers and their connectivity strongly differ in these three oligomers. The geometrically-optimized structure has four heterocycles per helix turn with *ortho–meta* connectivity (Fig. 1.5a); six heterocycles per helix turn with *meta–meta* connectivity (Fig. 1.5b); and twelve heterocycles per helix turn with *para–meta* connectivity (Fig. 1.5c). Another means to increase helix diameter, though less dramatic than going from *ortho*- to *meta*- or *para*-substitution, was to replace the 2,6-substituted pyridine unit (*meta*) with a larger 2,7-substituted 1,8-naphthyridine ring (*pseudo meta*-substitution) [95].

Similarly, curvature and helix diameter have been tuned in amide-linked aromatic foldamers based on the intramolecular hydrogen bonds discussed in Section 1.3.1. Thus, an estimated 20 residues per turn is expected for the motif shown in Fig. 1.5d [45] with *meta–para* connectivity! The expected end-to-end NOE sig-

(a)

(b)

(c)

(d)

(e)

(f)
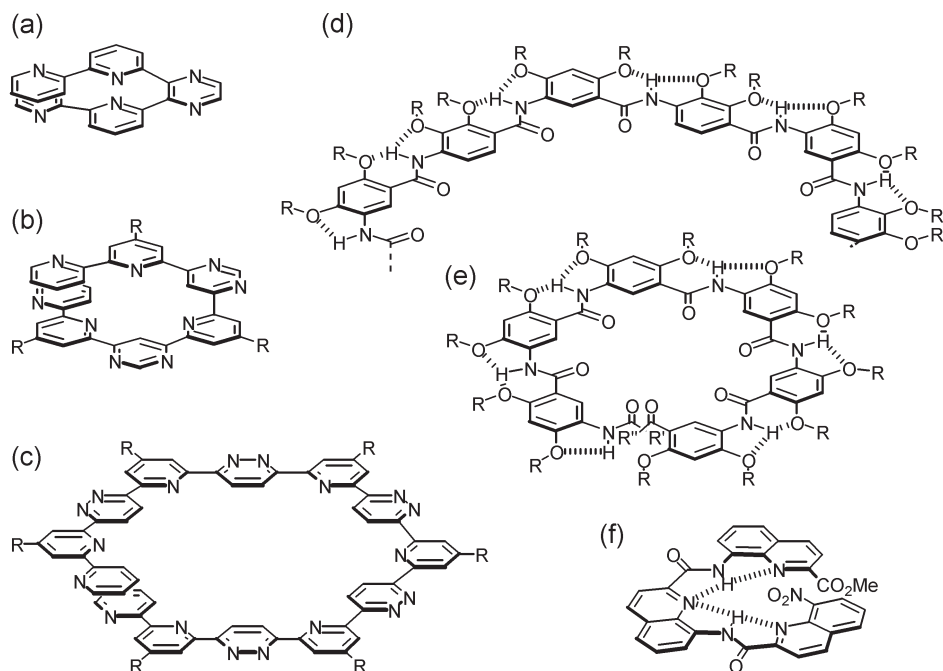
Fig. 1.5 Tuning curvature and helix diameter. The structures shown in (a), (b), (c), (d), (e) and (f) are from references [90, 91], [85, 89, 94], [86, 87], [45], [45], and [35, 36], respectively.

nal was observed in an oligomer possessing 21 units [45]. The foldamer shown in Fig. 1.5e [45] (see also Fig. 1.4a) with *meta–meta* connectivity has approximately 7 residues per helix turn; once again, end-to-end NOE contacts indicate the expected helical conformation. Oligomeric strands based on quinoline amino acid monomers (Figs. 1.5f and 1.4d) possess a high curvature imparted by a *pseudo-ortho* connectivity and comprise 2.5 heterocyclic units per helix turn. It is noteworthy that the number of units observed for these three systems differ from the values expected if *ortho*, *meta* and *para* connectivities were associated exactly with 60°, 120° and 180° angles. In fact, intramolecular hydrogen bonding between consecutive units tends to decrease curvature when it occurs at the helix periphery (Figs. 1.5d and f) [35, 36], and to increase helix curvature when it occurs in the helix interior. It is important to note that decreasing curvature at each unit results in larger and potentially more useful cavities, but also increases the number of units per turn making it more difficult to synthesize multiple turn helices. In contrast, highly curved oligomers give access to helices with high aspect ratios after relatively short syntheses. Such highly tunable systems represent versatile frameworks for developing helical receptors that bind guests into their cavities (see Chapter 7). To this end, the next level of complexity consists of in-

corporating different types of units within the same sequence so as to tune the curvature along the length of the same strand, giving access to conical objects or closed shells [29].

1.3.2.2 Linear Strands

Upon controlling the relative orientation of consecutive units, the diameter of crescent oligomers may be increased indefinitely to produce linear objects such as those represented in Fig. 1.5. When *para–para* connectivity is used (Figs. 1.6c and d) [27, 28, 48], the conformation indeed resembles a linear tape, but crinkle-tapes may also be obtained using *ortho–ortho* (Fig. 1.6e) [14] or *meta–meta* (Figs. 1.6a and b) [18, 37, 55], connectivity where curvature is alternatively positive and negative at each unit. In the strand shown in Fig. 1.6c, the conformation also results from a periodic change of curvature sign, but it occurs at each unit and not between units. When such *para* connections are introduced within a helical sequence, an inversion of helix handedness (and thus a *meso* helix) results [39]. Some of these linear and zig-zag ribbons possess two different edges resulting in different line tensions which eventually gives rise to a slight bending (Figs. 1.6a, b, and d), and may in fact be considered as fragments of very wide helices. Other ribbons have two identical edges and are rigorously straight (Fig. 1.6c and e).

1.3.2.3 Macrocycles

The underlying noncovalent interactions responsible for the conformational driving force of foldamers can be taken advantage of to form what can be considered as ''*self-templated*'' macrocycles. In essence, properly arranged functional groups during a reaction can be used to drive macrocyclization. The macrocycles depicted in Fig. 1.7 contain various combinations of one or two aromatic groups which can hydrogen bond to the linking unit in such a way as to rigidify the linkage in a well-defined conformation. The resulting ''directed conformational preorganization'' is a powerful approach to overcome the unfavorable entropy associated with macrocyclization reactions [104]. Hunter et al. developed a series of macrocycles and catenanes to evaluate the role of intramolecular hydrogen bonding in macrocyclization [22].

It is not surprising that the same starting materials used in crescent or helical foldamers (Section 1.3.2.1), can be used for the synthesis of macrocycles. In fact, by taking advantage of intramolecular interactions during the synthetic process, macrocycles can be prepared from irreversible reactions in one step, without the need for external templates, even at relatively high concentrations (typically from 0.07 to 1 M). As expected, in the case of linear or zig-zag foldamers (Section 1.3.2.2), the geometry of the system does not allow for macrocyclization to occur. Ultimately, it is the precise shapes and conformations that result from intramolecular interactions that allow for macrocyclization.

In this regard, Böhme et al. have prepared heterocyclic formamidine and urea macrocycles in excellent yields from the reaction of 2,6-diaminopyridine with triethyl orthoformate and *N,N*-carbonyldiimidazole, respectively (Fig. 1.7a and b)
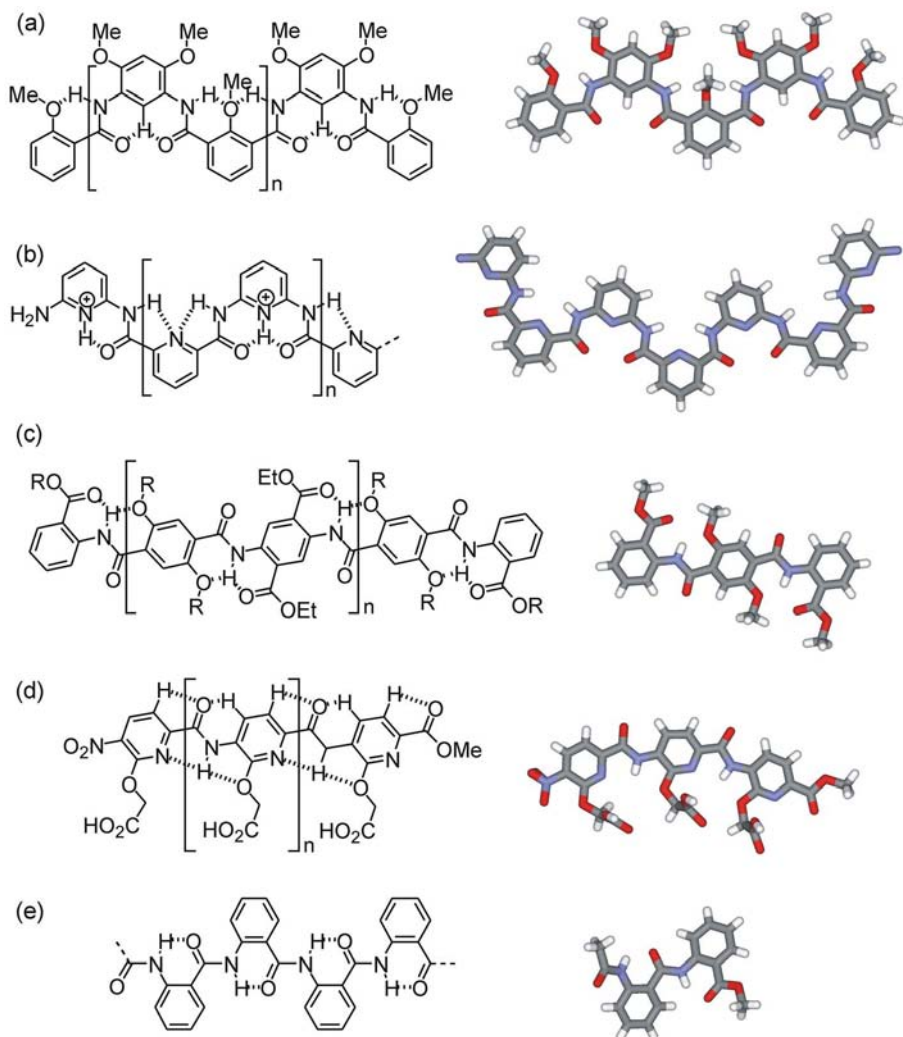
Fig. 1.6 Molecular formulae and representative X-ray structures – all at the same scale – of linear fully predictable foldamers. References for (a), (b), (c), (d), and (e) are from [55], [18, 37], [48], [27, 28] and [14], respectively. Not shown are examples of linear aza-aromatics and aza-aromatic amides [41, 98].

[75, 105]. In both cases, the driving force for macrocycle formation is intramolecular hydrogen bonding between the pyridyl nitrogen atom and the amino hydrogen of the urea or the C–H (or N–H) of the formamidine linkage. The importance of the pyridyl nitrogen for macrocyclization was demonstrated by the fact that when 2,6-diamino pyridine was substituted with 1,3-phenylenediamine, oli-
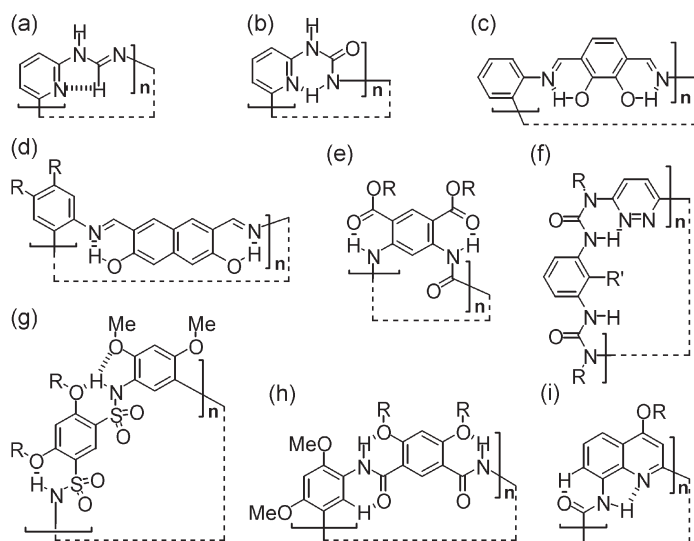
Fig. 1.7 Self-templated macrocycles (% yield). (a) n = 3 (97%) [105];
(b) n = 3 (90%) n = 4 (10%) [75]; (c) n = 3 (91%) [106]; (d) n = 6
(78%) [107]; (e) n = 4 (95%) [108]; (f) n = 2, R′ = H (46%) or CH3
(64%) [78]; (g) n = 2 (38%) [72]; (h) n = 3 (69%) [109]; (i) n = 3
(20%), n = 4 (20%) [110]. R and R′ indicate various types of alkyl
chains.

gomeric polyformamidines or polyureas were exclusively formed. Analogous self-templated macrocyclization reactions were reported by Xing et al. starting from 2,7-diamino[1,8]-naphthyridine for the preparation of the analogous trimeric formamidine macrocycle in 75% yield and the trimeric urea macrocycle in 64% yield [78].

Macrocycles are commonly synthesized by reactions of bifunctional monomers. However the kinetic competition between macrocyclization and polymerization can be a major problem. Chemists often use high-dilution and templating techniques to prevent undesirable polymerization reactions. In some cases dynamic (reversible) covalent chemistry provides an attractive synthetic strategy to yield thermodynamically favored macrocyclic products. The Schiff-base condensation of 2,3-dihydroxybenzene-1,4-dicarbaldehyde with 1,2-phenylenediamine gives a macrocyclic hexaimine in 91% yield (Fig. 1.7c) [106]. The selective formation of this cyclic trimer was rationalized by the stabilization of favorable conformations required for cyclization through intramolecular hydrogen bonding. Furthermore, because this product has low solubility in the reaction solvent, there was a thermodynamic driving force for the formation of this macrocycle. An excellent example of using rigid precursors that are predisposed to a particular geometry was demonstrated in the preparation of a related Schiff-base macrocycle from the 6 + 6 reaction of 3,6-diformyl-2,7-dihydroxynaphthalene with 4,5-diamino-1,2-

dihexyloxybenzene (Fig. 1.7d) [107]. According to Hui et al. the ability to synthesize this 66-membered macrocyle can be attributed to maximizing intramolecular hydrogen-bonding and its poor solubility which again drives the reaction thermodynamically.

Zhang et al. have prepared a urea-linked macrocycle from a $2 + 2$ reaction between two diarylurea units rigidified by intramolecular hydrogen bonding (Fig. 1.7e) [108]. The authors rationalize that the limited conformational freedom of the dimeric starting material and the tetrameric intermediate is expected to facilitate macroyclization. Another variety of urea-linked macrocycles was synthesized from *N-iso*butyl-3,6-diaminopyridazine reacting with either 1,3-phenylene diisocyanate or tolylene-2,6-diisocyanate (Fig. 1.7f) [78]. It is believed that both hydrogen bonding and steric interactions are involved in macrocyclization and that the formation of these intramolecular hydrogen bonds in the transition state directs the irreversible macrocyclic ring closing reaction to occur at such high yields in the absence of a template. The presence of the tolyl groups ($R = CH_3$) versus the phenyl groups ($R = H$) illustrates the importance of steric repulsion to direct macrocyclization. This is supported by the difference in yield of the tolyl-containing macrocycle (64%), compared with the phenyl-containing macrocycle (46%). The tolyl methine group prevents alternative planar conformations due to unfavorable steric interactions with the urea carbonyl oxygens.

The one-step preparation of a macrocyclic polysulfonamide in 38% yield based on intramolecular three-center H-bonding was described by He et al. (Fig. 1.7g) [72]. In related work, Yuan et al. reported highly efficient (69% yield), one-step macrocyclization reactions by treating 4,6-dimethoxy-1,3-phenylenediamine with the appropriate diacid chloride (Fig. 1.7h) [109]. Three-center H-bonds rigidify the backbone and pre-organize the precursor oligomers for macrocyclization. Likewise, Jiang et al. attribute cyclization of oligoamide macrocycles to precursor pre-organization (Fig. 1.7i) [110]. In this case two different macrocycles ($n = 3$ and $n = 4$) were obtained. In all cases, it is believed that the strong conformational preference of the building blocks once the linkage is formed gives rise to a thermodynamic preference for macrocyclization.

1.3.3
Partially π-conjugated Oligomers

Whilst π-conjugated systems represent by far the largest body of fully predictable foldamers, some with aliphatic or partly aliphatic backbones have also been reported (Fig. 1.8). These systems very much resemble those shown in Sections 1.3.2.1 and 1.3.2.2 in that folding may be defined by local conformational preferences directed by interactions between adjacent units, and that such control may in principle operate at every rotatable bond. However, the lack of conjugation increases the number of accessible conformations and hence the fold is entropically destabilized; strong deviations may be observed from the canonical equilibria shown in Fig. 1.2. These oligomers are much related to those described in Chapter 2 (Section 2.3.2).
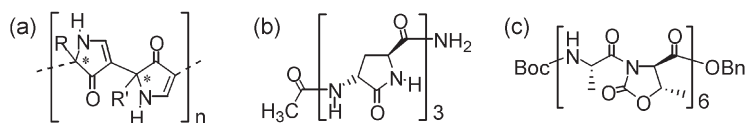
Fig. 1.8 Examples of aliphatic oligomers, the conformations of which are defined by interactions between adjacent units.

The 3,5-linked polypyrrolin-4-one scaffold (Fig. 1.8a) allows for a wide variety of folding motifs. This outcome may be ascribed to the cooperative effect of: (i) the dihedral angles of the backbone that are locked by the pyrrolinone ring, (ii) restricted rotations due to *gauche* steric interactions between the side chains and the neighboring pyrrolinone rings, and (iii) C=O···H–N intramolecular H-bonds between adjacent pyrrolinone rings [102]. The 3,5-linked homochiral polypyrrolinone motif provides for an excellent β-sheet/β-strand peptidomimetic [111]. On the other hand, alternating D,L (heterochiral) polypyrrolinones preferentially adopt a β-turn [112]. Finally, studies of *N*-methylated bispyrrolinones indicate the possibility to obtain β-turn and helical structures [113].

The use of γ-lactams such as (2*S*,4*R*)-4-amino-5-oxopyrrolidine-2-carboxylic acid as conformationally constrained dipeptide building blocks locks every other amide linkage in the *cis* conformation resulting in rigid peptide foldamers with an alternating *cis-trans* amide sequence (Fig. 1.8b). In the solid state, the trimer folds into a crescent shape. Unconstrained backbone ϕ and ψ torsional angles are $100 \pm 15°$ and $138 \pm 13°$, showing that the expected intramolecular hydrogen bonds between the amide NH and the lactam carbonyl and between the lactam NH and the amide carbonyl do not occur. In the solid, intermolecular hydrogen bonds apparently prevail over intramolecular hydrogen bonding [103].

Foldamers that contain the 4-carboxy-5-methyloxazolidin-2-one unit fold into well-defined β-bend ribbon spirals (Fig. 1.8c). Factors that influence folding in these strands include: (i) the rigid –CO–N(CH<)–CO– moiety, which favors a *trans* conformation, (ii) the formation of C=O···H–C$^\alpha$ hydrogen bonds, and (iii) the formation of alternate $1 \leftarrow 4$ intramolecular C=O···H–N hydrogen bonds [114].

1.4
Semi-rigid Backbones

Backbone rigidity, in other words a limited number of unfolded conformations, is the main factor that favors the folding of an oligomeric or polymeric chain. Chapters 2 and 3 present families of foldamers where backbone rigidity is combined with at least a second strong folding force: attractive interactions such as hydrogen bonds (see Chapter 2), or solvent-induced collapse of solphovobic – essentially hydrophobic – moieties (see Chapter 3). When backbone rigidity is

not as overwhelming as in the foldamers presented in Section 1.3 and when no other strong folding force is at play, folding may still arise from less obvious yet efficient factors that determine local conformational preferences, for example steric and stereoelectronic effects, or weak π–π interactions. As shown in the following sections a great diversity of structures may be classified in this category.

1.4.1
Tertiary Aromatic Amides, Imides and Ureas

The oligomers described in this section resemble most of those described in Section 1.3 in that they consist of a string of π-conjugated units separated by single bonds. They thus possess a reduced number of accessible conformers. However, in these cases none of the schemes shown in Fig. 1.2 apply: unambiguous conformational preferences do not exist at the single bonds separating the π-(systems). Rotation about these bonds is nevertheless restricted by various factors, such as intramolecular π–π stacking between aromatic units and steric hindrance associated with bends in the backbone architectures. The bends arise from the preferred *cis* conformation of tertiary amides of aromatic acids and aliphatic–aromatic secondary amines (Fig. 1.9), in which the two aryl groups project to the same side of the amide bond [115]. Even though no clear conformational preferences are defined at the aryl–NCO and aryl–CON linkages, rotation about these bonds is hindered by steric effects and weak attractive electrostactic interactions between aryl groups. When the molecules are soluble in protic solvents, interactions between the aryl groups are reinforced by solvophobic effects (see Chapter 3, Section 3.4.1).
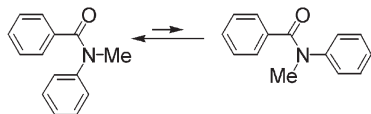
Fig. 1.9 *Cis–trans* equilibrium of aliphatic aromatic tertiary amides. See Itai *et al*. [115].

This pattern has been efficiently exploited in oligomers derived from aliphatic–aromatic tertiary ureas and guanidines [81, 116], imides [117, 118], and amides [119–122]. In urea and imide functional groups, two *cis* conformations are combined, resulting in a strong kink between each aryl group that allows almost perfect face to face stacking of adjacent aryl rings. In the solid state, these oligomers adopt folded ladder-like conformations consisting of a *pseudo*-helical arrangement of imide or urea moieties around a central column of stacked aryl groups (Figs. 1.10a, b and d) [81, 117]. Tertiary amides give rise to less pronounced kinks and thus less pronounced interactions between adjacent aryl groups in the sequence. However, solid state structures show that these oligomers adopt compact helical
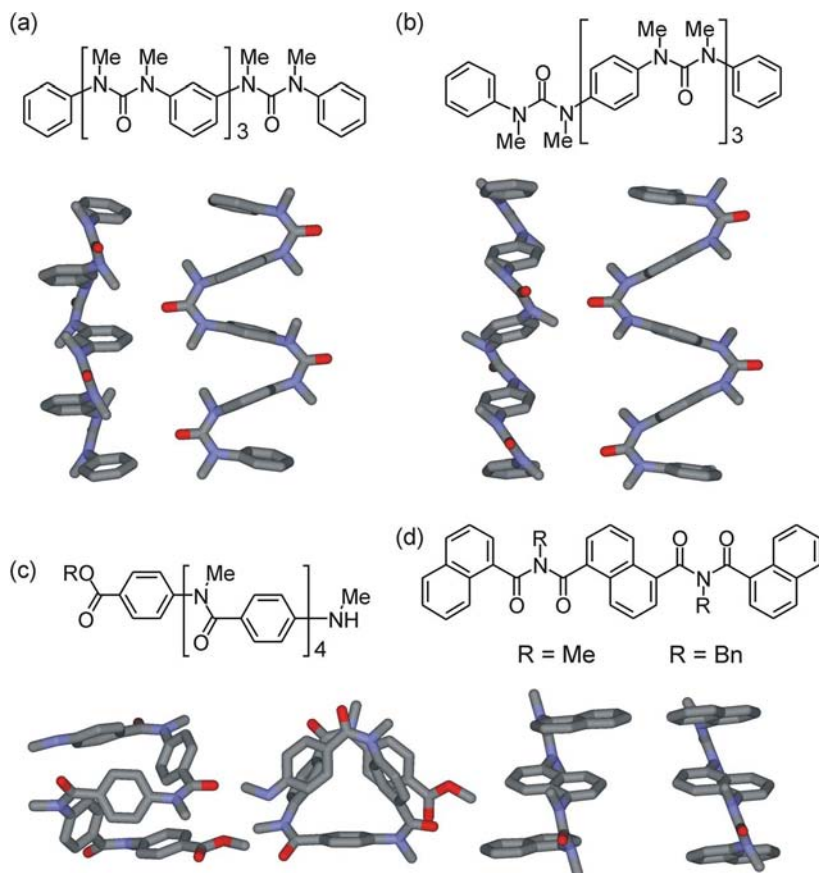
Fig. 1.10 Formulae and crystal structures at the same scale of tertiary ureas, amide and imide oligomers. (a)–(b) front view and side view; (c) top view and side view; (d) two possible conformations. References of examples (a)–(d) are from [81], [81], [120] and [117], respectively.

conformations where aryl–aryl contacts are observed between nonadjacent units (Fig. 1.10c) [120].

Evidence that solution conformations resemble those observed in the solid state come from NMR studies that show NOE and upfield shifts of the proton signals involved in π–π stacking. However, the absence of strong directional interactions in the folded conformation leads one to suspect that the solution structures are not as well-defined as the solid state structures. At each aryl–amide, aryl–imide or aryl–urea linkage, several conformers are compatible with π–π stacking of the aryl moieties. This is well illustrated by the two crystal structures shown in Fig. 1.10d. They correspond to two distinct folded conformations of the same back-bone where the upper and lower naphthyl groups have been flipped upon rota-
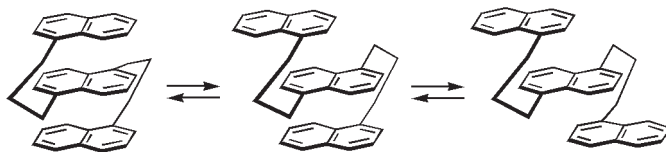
Fig. 1.11 Schematic representation of possible conformers of the oligomers shown in Fig. 1.10d [117].

tion of the naphthyl–imide linkage by 180°. Even though these two structures belong to two molecules which possess methyl and benzyl residues, respectively, there is no reason to exclude an equilibrium between the two conformers for both species, as illustrated in Fig. 1.11. To the extent that the structures shown in Fig. 1.10 may be called helices, such rotations about aryl–imide, aryl–urea and aryl–amide linkages amount to locally inverting the helix handedness. It is remarkable that such inversions may occur many times within an oligomer without disrupting the overall arrangement of the stacked aryl groups. Thus, the solution conformations of aromatic tertiary amide, urea and imide oligomers presumably consist of an ensemble of closely related conformers.

The equilibria involved between these conformers may be shifted upon introducing a chiral bias at each aliphatic substituent of the backbone nitrogen atoms: a chiral side group gives a local preference in favor of the left-handed or the right-handed helical turn. Thus, in oligoimides [118], and polyamides [120, 122] possessing a chiral aliphatic residue at every unit, strong induced circular dichroism bands are observed, suggesting a long range helical order in the backbone. From the studies above however, it might be hypothesized that the effect of chiral residues is local and that "majority rules" and "sergeant and soldiers" principles apply to a limited extent in these systems (see Chapters 11 and 12).

1.4.2
Tertiary Aliphatic Amides: Polyprolines and Peptoids

The helix is ubiquitous in the secondary structure of α-peptides and their β- and γ-homologs (see Chapters 2 and 5). In regular helices formed from secondary α- or β-peptides (Figs. 1.12a and b), intrachain C=O···H–N hydrogen bonding between nonadjacent residues is essential to folding: for example, between C=O$_i$ and N–H$_{i+4}$ for an α-helix (3.6$_{13}$-helix); between C=O$_i$ and N–H$_{i+3}$ for a 3$_{10}$-helix in α-polypeptides; and between N–H$_i$ and C=O$_{i+2}$ and for a 3$_{14}$-helix in β-polypeptides. Intuitively, it might be imagined that upon converting these secondary amides to tertiary aliphatic amides, the resulting loss of hydrogen bond donors may result in the inability of an oligomeric strand to adopt a well-defined folded conformation. This is, however, not the case. Oligomers of the naturally occurring amino acid proline also adopt helical conformations despite the fact that proline is a secondary amine and that its oligomers are tertiary amides unable to form intramolecular hydrogen bonds (Fig. 1.12c). For an α-helix, the ϕ,
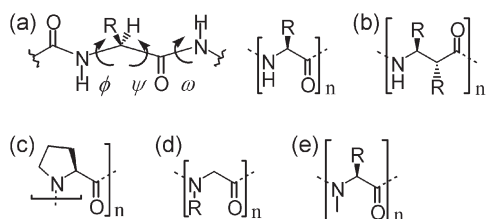
Fig. 1.12 Repeat units of (a) α-polypeptides; (b) β-polypeptides;
(c) polyproline; (d) peptoids (poly(N-substituted glycine)); and
(e) N-methylated polypeptides.

ψ torsion angle distribution (Fig. 1.12a) in the Ramachandran plot lies near
$(-63°, -42°)$ [123]. The Ramachandran plot of polyproline is quite distinct be-
cause the ring structure of the pyrrolidine ring is constrained. The ϕ, ψ distribu-
tion is centered approximately at $-75°$, $145°$ for the left-handed polyproline type
II helix, which is stable in aqueous medium and which contains all-*trans* tertiary
amides ($\omega = 180°$), and at $-70°$, $160°$ for the right-handed polyproline type I he-
lix which contains all-*cis* tertiary amides ($\omega = 0°$) (Fig. 1.13) [123].

The polyproline type II helix is found in both folded and unfolded peptides and
plays important roles in biological signal transduction, transcription, cell motility,
and immune responses [125–127]. The triple helix of collagen consists of three
intertwined polyproline type II helices. It is still unclear why the polyproline
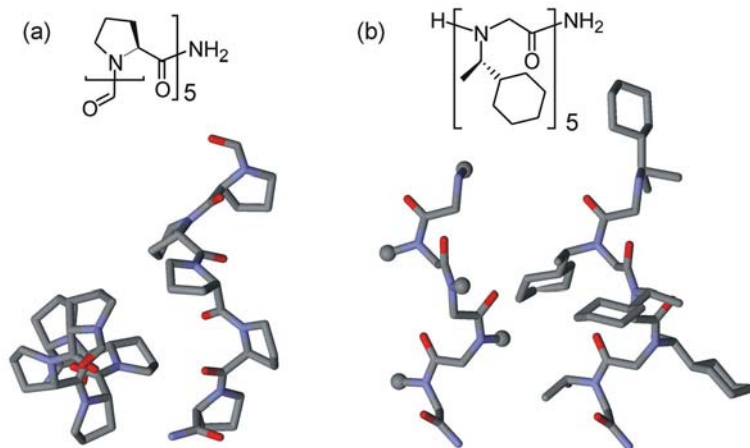
Fig. 1.13 Crystal structures at the same scale of: (a) a left-handed
polyproline type II helix (top view and side view); and (b) a pentameric
peptoid helix [124] (side chains have been replaced by balls in the left-
hand view.

type II helix is intrinsically so stable despite the absence of strong intramolecular interactions. In addition to the rigidity imparted by the pyrrolidine ring which sets the ϕ angle, solvent and stereoelectronic effects apparently determine the preferred ψ values. For example, a (possibly cooperative) stabilizing $n \rightarrow \pi^*$ interaction between O_{i-1} and C'_i has been shown to play a substantial role in stabilizing the polyproline type II helix [125]. Substituents in position 4 of the pyrrolidine ring modulate the ring conformation which in turn enhance or inhibit the $n \rightarrow \pi^*$ interaction and the preference for a *trans* over *cis* amide conformation, thus defining the balance between type II and type I helices [125].

Peptoids (*N*-substituted glycine derivatives; Fig. 1.12d) have many similarities to polyprolines, including their inability to form intrachain hydrogen bonds and the possibility to form *cis* tertiary amide backbone linkages [128]. However, like polyprolines, peptoids do fold into helical structures based on conformational preferences of the backbone chain, side-chain–backbone steric repulsions, and dipole–dipole repulsions between main-chain amide carbonyl electrons [129]. Peptoids with as few as five residues have been shown to form reversible and cooperative stable helical structures in both aqueous and organic solvents [128, 130]. Note that the peptoid backbone is achiral and the chirality of the secondary structure (helix screw sense) is governed by the chirality of the side chains [130]. An X-ray crystal structure of an (*S*)-*N*-(1-cyclohexylethyl)glycine pentamer reported by Wu et al. shows a left-handed helix with similar torsional angles (ϕ, ψ, and ω) as the polyproline type I helix (Fig. 1.13) [124]. Ramachandran plots indicate a greater conformational diversity for peptoids compared with peptides, which is believed to be related to the lack of substitution on the α-carbon (which is achiral) and the absence of an amide N–H capable of hydrogen bonding [129, 131]. Nevertheless, peptoids have a great potential for biological applications (see Chapter 8).

Another family of tertiary aliphatic amides are poly-*N*-methylated peptides (Fig. 1.12e) which have been described as "more congested peptoid-like molecules" [70]. In this case, an extended β-strand conformation rather than a helical conformation is adopted. In fact, X-ray crystallography of (*N*-Me-L-Ala)$_6$ and all-*N*-Me-(Ser(OBz)-Val-Ala-Ser(OBz)-Val-Ala) indicates that the β-strand conformations of poly-*N*-methylated peptides retain an ability to form hydrogen bonds with α-peptide β-strands through their carbonyl groups.

Other classes of foldamers related to peptoids and polyprolines include some β-peptide derivatives containing cyclic tertiary amide linkages that can be broadly classified as cyclic β-peptides (Fig. 1.14). By controlling all the torsion angles of the backbone, Lee et al. [132] have designed a completely non-hydrogen-bonded helical pseudopeptide composed of amide linked oxanipecotic acid units (Fig. 1.14a). In this work, circular dichroism was used to show that an oxanipecotic acid tetramer adopts a non-hydrogen-bonded helical structure more efficiently than the nipecotic tetramer (Fig. 1.14b) [133]. The symmetric chemical environment adjacent to the nitrogen atom in the piperidine ring of nipecotic acid makes the *cis* and *trans* conformation of the amide linkage equivalent whereas this is not the case in oxanipecotic acid.
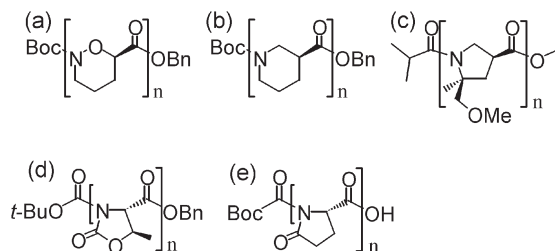
Fig. 1.14 Oxanipecotic acid oligomers (a), nipecotic acid oligomers (b), 2,2-disubstituted pyrrolidine-4-carboxylic acid oligomers (c), benzyl (4*S*,5*R*)-5-methyl-2-oxo-1,3-oxazolidine-4-carboxylate oligomers (d), and pyroglutamic acid oligomers (e).

Using NMR structural analysis and circular dichroism, Huck et al. have studied 2,2-disubstituted pyrrolidine-4-carboxylic acid oligomers (Fig. 1.14c) that fold into helical structures without intramolecular backbone hydrogen bonding [134]. The preferred *cis* conformation of the backbone amide linkage in this foldamer is reminiscent of the *cis* conformation of the backbone in polyproline I helices.

Homo-oligomers of benzyl (4*S*,5*R*)-5-methyl-2-oxo-1,3-oxazolidine-4-carboxylate ($n = 2$–5; Fig. 1.14d) adopt polyproline type II-like helical conformations. Once again, the torsion angle, ϕ, is restricted by the cycle to values between -50 to $-80°$ while the ψ and ω torsion angles are both *trans* (140–160° and 180°, respectively). However, in contrast to the polyproline II helix, this helical conformation is stabilized by intramolecular α-C–H$_i$, 2-oxo-1,3-oxazolidine O=C$_{i+1}$ hydrogen bonds [135]. Similar stabilization by intramolecular C–H···O=C hydrogen bonding was observed for homo-oligomers of pyroglutamic acid (n = 2–4; Fig. 1.14e) [136].

1.4.3
Hindered Polymer and Oligomer Backbones

Attractive interactions play a major role in determining conformation and are easily identified in the structures as they result in a close proximity between the groups involved. Repulsive interactions are no less important. However, since repulsive interactions often result in some distance between the groups that repel each other, they are sometimes overlooked, and definitely less commonly used in design. Steric effects represent one very large class of repulsive interactions. As shown in the following examples, numerous families of oligomeric and polymeric backbones adopt well defined, generally helical, conformations as a result of steric repulsion between bulky peripheral groups. Among the polymer classes that fall into this category are polyisocyanides, isotactic polymethacrylates, polyisocyanates, polyguanidines, and polyacetylenes. Since the behavior of these polymers is presented in great detail in Chapters 11 and 12, they will be only briefly discussed here.
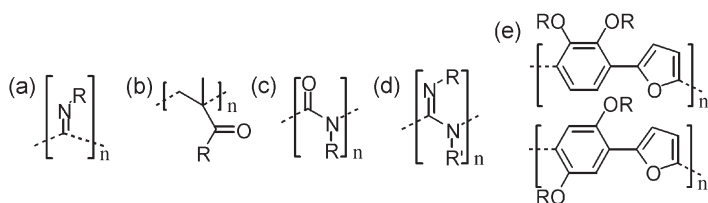
Fig. 1.15 Structures of polyisocyanides (a), polymethacrylates and polymethacrylamides (b), polyisocyanates (c), polyguanidines (d), and poly-(dialkoxy-1,4-phenylene-*alt*-2,5-furan)s (e).

Steric repulsions between the side chains in substituted polyisocyanides (Fig. 1.15a) are large and prevent them from adopting the planar all-*trans* structure. Restricted rotation about the single bonds of the polymer backbone results in right- or left-handed helical polymers with four monomer units per helix turn. The polymers can be resolved using chiral chromatography and the use of chiral isocyanides or chiral catalysts can influence the handedness of the helix [137, 138]. Polymethacrylates (Fig. 1.15b) and polymethacrylamides (Fig. 1.15c) are among vinyl polymers that adopt helical conformations. In order for these helices to be stable, bulky side groups are once again required, and the polymerization process (either anionic or radical) must yield isotactic polymers [139, 140]. The helical conformation of polyisocyanates results from a combination of electronic and steric effects. In essence, the $-C(O)-N(R)-$ linkages are forced to be non-planar because the substituents are bulky [141]. Polyisocyanates have a rich stereochemistry due to the stability of long helical segments and the ease of controlling handedness by incorporating chiral monomers [142]. Polyguanidines (Fig. 1.15d), formed from carbodiimide monomers, have similar helical structures to those of polyisocyanates. The use of bulky (R) and chiral (R′) substituents leads to stable one-handed helical polymers with high barriers for conformational racemization [143]. In alternating poly-(dialkoxy-1,4-phenylene-*alt*-2,5-furan)s (Figs. 1.15e and f), favorable interactions between the oxygen atoms of alkoxy side chains or furan units and the *ortho* hydrogen atoms connected to the aromatic rings lead to restricted rotation about the inter-aromatic bonds and well-defined conformations. By varying the substitution pattern of these conjugated polymers, rigid-rod-like (poly(dialkoxy-2,5-phenylene)-*alt*-2,5-furan)s) or helical (poly-(dialkoxy-1,4-phenylene-*alt*-2,5-furan)s) conformations can be obtained [144].

Several classes of oligomers have also been shown to adopt sterically-induced helical conformations. Their oligomeric nature (as opposed to the polymers shown above) simply reflects that these molecules were synthesized in a stepwise fashion. Similar conformations would be expected had polymerization techniques been used for their preparation. Thus, ortho-oligophenylenes (Fig. 1.16a) adopt a tight helical conformation with three aromatic rings per helix turn that was characterized in the solid state. The geometry and sterics of this hindered structure does not allow free rotation about the interaromatic single bonds [145]. Similarly,
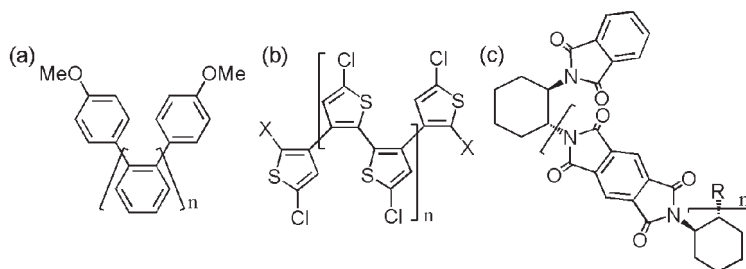
Fig. 1.16 Structure of ortho-oligophenylenes (a), polythiophenes (b), oligoimides of *trans*-1,2-diaminocyclohexane (c).

torsional effects play the major role in biasing a helical conformation of sexithiophenes in the solid state (Fig. 1.16b) [146]. Finally, oligoimides of *trans*-1,2-diaminocyclohexane have been shown to fold into helical conformations (Fig. 1.16c). Although these oligomers have fixed (R,R) chirality, they can fold into either M or P helical conformers. The folding into helical conformations is once again driven by the strict steric constraints of the system which restricts rotation about the single bonds while the preference for M or P helicity has been attributed to intermolecular interactions [147].

1.5
Conformational Transitions

To a large extent, the functions and properties of foldamers rely principally on their structures. This is the case, for example, in molecular recognition using foldamers (see Chapter 7), including molecular recognition of biomolecules to elicit biological responses (see Chapter 8). Foldamers with well-defined structures may serve as scaffolds to display arrays of recognition functions that bind to specific targets. Besides the structures themselves, structural *changes* or "conformational transitions", sometimes of large amplitude, are also useful for a number of applications. For example, sensing and memorizing, as described in Chapter 11, may arise from conformational changes of helical polymers. Large, controlled, conformational changes also provide a framework to elaborate molecular machines [148]. Interest for such phenomena is fueled by their relation to natural molecular factories and by the prospect of elaborating useful chemical devices.

Conformational transitions may be envisaged in any class of folding molecules. The most common transition is that between a folded state and a more or less defined ill-folded state (the so called "random coil"). It may be triggered by a variety of stimuli, such as solvent, temperature, light, guest molecules, or minor chemical transformations. Transitions may also exist between two (or more) distinct well-defined conformations. To achieve control over large conformational transformations, a high level of predictability of the structures is necessary. As
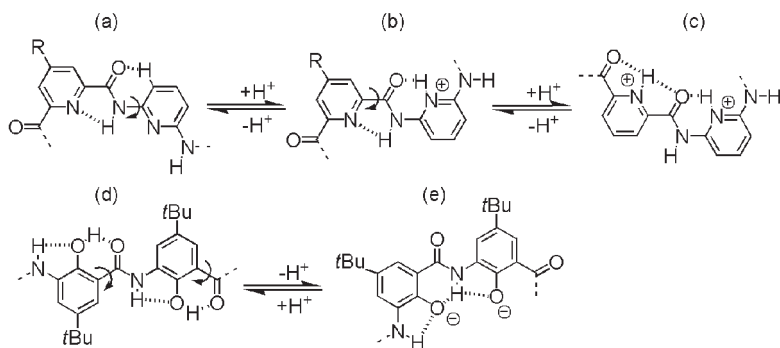
Fig. 1.17 Transitions between helical and extended linear conformations *via* sequential protonation and deprotonation.

shown in Chapter 10, nucleic acids have proven particularly useful in this respect. The foldamers described in this chapter have also been much exploited, particularly the ''fully predictable'' foldamers presented in Section 1.3.

Leaving aside the simplest transitions such as, for example, the dynamic equilibrium between right-handed and left-handed helical conformations, which may be biased in various ways [25, 34, 47], more dramatic transitions between helical and fully extended linear conformations have been implemented in various fully predictable foldamers. For example, 180° rotation about aryl–NHCO and aryl–CONH linkages in aromatic amide foldamers can be effected *via* sequential protonation and deprotonation reactions [18, 37]. Dolain et al. have shown the folding diversity of a single foldamer capable of transforming between two different helical conformations *via* a linear conformation using sequential protonation. TFA was used to protonate diaminopyridine units of a helical foldamer (Fig. 1.17a) giving rise to an extended linear conformation (Fig. 1.17b). Subsequent protonation by a stronger acid of the alternating pyridinedicarboxamide units results in an ''inside-out'' helical conformation (Fig. 1.17c) [18]. An original aspect of this system is that the linear strand shown in Fig. 1.2b may convert into a helix either upon protonation or deprotonation. In an analogous fashion, Kanamori et al. demonstrated linear-to-turn transformations in unsymmetrically linked phenolic oligoamides. In this case, deprotonation was used to switch between the linear conformer (Fig. 1.17d, NH···OH and OH···O=C hydrogen bonding) and a bent conformer (Fig. 1.17e NH···O(oxyanion) hydrogen bonding) [12].

Lehn's group has also used metal coordination to effect conformational changes in oligoheterocyclic molecular strands. In this case, aryl–aryl and aryl–imine linkages undergo 180° rotations upon ion binding that result in linear-to-helical or helical-to-linear conformational transitions of oligomeric strands (Fig. 1.18) [82, 98, 149]. For example, the interaction of Pb(II) with helical oligo–pyridyl–pyrimidine strands generates polymetallic racks – or grid-type supramolecular architectures – by uncoiling of the oligomeric ligand. The reversible inter-
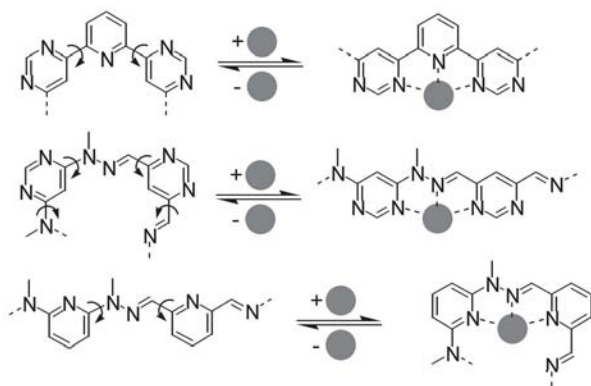
Fig. 1.18 Transitions between helical and extended linear conformations *via* ion binding and unbinding.

conversion between the helical free ligand and the linearly extended ligand was controlled by coupling the ion binding/unbinding process with a competing ligand (tris(2-aminoethyl)amine; tren), whose ion affinity can be modulated with pH. Additional conformational changes can be envisaged in the folded conformations stabilized by metal coordination. For example, electrochemical changes of the oxidation state of the metal may induce a rearrangement of its coordination sphere that, in turn, modifies the conformation of an oligomeric ligand [150].

1.6
Conclusion and Perspectives

Local conformational preferences can be imparted to oligomeric chains in various ways: ring closure, π-conjugation, steric effects, electrostatic attractions and repulsions between neighboring units and combinations thereof. All these factors can contribute to limit (if not completely restrict) bond rotation, and result in different levels of rigidity of molecular backbones, which, in turn, correspond to quite diverse structural families of foldamers. In this chapter, we have grouped these families into three categories. First, oligomeric backbones whose conformational space is essentially reduced to a single conformer; since such molecules cannot unfold, they stand at the border of what the term "foldamer" encompasses. Second, rigid structures that possess numerous rotatable bonds and that undergo dynamic structural changes, but whose structure is perfectly and robustly defined because conformational preferences exist at each and every rotatable bond; to emphasize these characteristics, these oligomers were termed "fully predictable foldamers". Third, semi-rigid structures where strong conformational preferences exist, but not at every rotatable bond; these molecules possess limited conformational spaces and adopt well-defined conformations (though less well-defined and/or less stable than the two categories above) as a result of additional weaker

effects such as electrostatic interactions as involved in π–π stacking and n $\rightarrow$ π^* interactions, at the exclusion of strong hydrogen bonding or hydrophobic effects, which are dominant in the foldamers described in Chapters 2 and 3.

Numerous families of synthetic molecules have been designed and characterized that mimic the secondary structural folded motifs of biopolymers – helices, turns or linear strands. As a result of their folded structures, these molecules exhibit multifaceted properties that are amenable to applications (see Chapters 7, 8, and 10–13). Yet, it remains to be shown whether these same artificial folding molecules can also mimic the structures, and ultimately the functions, of the tertiary and quaternary folds of proteins and nucleic acids. Some of the challenges ahead thus lie in controlling tertiary structures as efficiently as we control secondary structures, thus integrating the structural and functional characteristics of biopolymers with the chemical stability and diversity of synthetic polymers and oligomers.

Fully predictable foldamers are currently the object of intense developments and may prove highly versatile building blocks for this purpose. The very high stability of their folded conformations and the high level of understanding that we have regarding their folded structures should be useful in the context of modular approaches where several secondary motifs would be hierarchically integrated into larger architectures, through covalent or noncovalent assembly. On the other hand, fully predictable foldamers may, in one sense, be considered as too well-defined and do not undergo structural changes as easily as their natural counterparts. Semi-rigid backbones are less well-defined. Their structures are less easy to understand and to predict and, in this respect, have more in common with naturally occurring folding molecules. Semi-rigid backbones also clearly provide a greater variety of chemical structures and it is fair to speculate that only a small part of this variety has been uncovered, and that we are still in the initial stages of this unfolding story.

During the course of the preparation of this chapter, a relevant focus review paper on shape-persistent aromatic amide oligomers was published.

> Z.-T. Li, J.-L. Hou, C. Li, H.-P. Yi,
> *Chem. Asian J.* **2006**, *1*, 766–778.

References

1 D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071–13072.

2 S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.

3 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4011.

4 R. H. Martin, *Angew. Chem. Int. Ed.* **1974**, *13*, 649–660.

5 M. S. Newman, D. Lednicer, *J. Am. Chem. Soc.* **1956**, *78*, 4765–4770.

6 R. H. Martin, M. J. Marchant, *Tetrahedron* **1974**, *30*, 347–349.

7 B. Kiupel, C. Niederalt, M. Nieger, S. Grimme, F. Vögtle, *Angew. Chem. Int. Ed.* **1998**, *37*, 3031–3034.

8 K. Fuji, T. Furuta, K. Tanaka, *Org. Lett.* **2001**, *3*, 169–171.

9 V. Semetey, D. Moustakas, G. M. Whitesides, *Angew. Chem. Int. Ed.* **2006**, *45*, 588–591.

10 C. G. Levins, C. E. Schafmeister, *J. Am. Chem. Soc.* **2003**, *125*, 4702–4703.

11 J. S. Schumm, D. L. Pearson, I. Jones, L. R. Hara, J. M. Tour, *Nanotechnology* **1996**, *7*, 430–433.

12 D. Kanamori, T. Okamura, H. Yamamoto, N. Ueyama, *Angew. Chem. Int. Ed.* **2005**, *43*, 969–972.

13 Y. Hamuro, S. J. Geib, A. D. Hamilton, *Angew. Chem. Int. Ed.* **1994**, *33*, 446–448.

14 Y. Hamuro, S. J. Geib, A. D. Hamilton, *J. Am. Chem. Soc.* **1996**, *118*, 7529–7541.

15 Y. Hamuro, S. J. Geib, A. D. Hamilton, *J. Am. Chem. Soc.* **1997**, *119*, 10587–10593.

16 T. Kawamoto, B. S. Hammes, B. Haggertty, G. P. A. Yap, A. L. Rheingold, A. S. Borovik, *J. Am. Chem. Soc.* **1996**, *118*, 285–286.

17 J. Recker, D. J. Tomcik, J. R. Parquette, *J. Am. Chem. Soc.* **2000**, *122*, 10298–10307.

18 C. Dolain, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2003**, *42*, 2737–2740.

19 V. Berl, I. Huc, R. G. Khoury, M. J. Krische, J.-M. Lehn, *Nature* **2000**, *407*, 720–723.

20 V. Berl, I. Huc, R. G. Khoury, J.-M. Lehn, *Chem. Eur. J.* **2001**, *7*, 2798–2809.

21 V. Berl, I. Huc, R. G. Khoury, J.-M. Lehn, *Chem. Eur. J.* **2001**, *7*, 2810–2820.

22 F. J. Carver, C. A. Hunter, R. J. Shannon, *Chem. Commun.* **1994**, 1277–1280.

23 G. T. Crisp, Y.-J. Jiang, *Tetrahedron* **1999**, *55*, 549–560.

24 C. Dolain, A. Grélard, M. Laguerre, H. Jiang, V. Maurizot, I. Huc, *Chem. Eur. J.* **2005**, *11*, 6135–6144.

25 C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2005**, *127*, 12943–12951.

26 C. Dolain, J.-M. Léger, N. Delsuc, H. Gornitzka, I. Huc, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16146–16151.

27 J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2003**, *42*, 535–539.

28 L. A. Estroff, C. D. Incarvito, A. D. Hamilton, *J. Am. Chem. Soc.* **2004**, *126*, 2–3.

29 J. Garric, J. M. Léger, I. Huc, *Angew. Chem. Int. Ed.* **2005**, *44*, 1954–1958.

30 E. R. Gillies, C. Dolain, J.-M. Leger, I. Huc, *J. Org. Chem.* **2006**, *71*, 7931–7939.

31 Z.-Q. Hu, H.-Y. Hu, C.-F. Chen, *J. Am. Chem. Soc.* **2006**, *71*, 1131–1138.

32 B. Huang, J. R. Parquette, *Org. Lett.* **2000**, *2*, 239–242.

33 I. Huc, V. Maurizot, H. Gornitzka, J.-M. Léger, *Chem. Commun.* **2002**, 578–579.

34 H. Jiang, C. Dolain, J.-M. Léger, H. Gornitzka, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 1034–1035.

35 H. Jiang, J.-M. Léger, C. Dolain, P. Guionneau, I. Huc, *Tetrahedron* **2003**, *59*, 8365–8374.

36 H. Jiang, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2003**, *125*, 3448–3449.

37 E. Kolomiets, V. Berl, I. Odriozola, A.-M. Stadler, N. Kyritsakas, J.-M. Lehn, *Chem. Commun.* **2003**, 2868–2869.

38 V. Maurizot, C. Dolain, I. Huc, *Eur. J. Org. Chem.* **2005**, 1293–1301.

39 V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 10049–10052.

40 V. Maurizot, G. Linti, I. Huc, *Chem. Commun.* **2004**, 924–925.

41 I. Odriozola, N. Kyritsakas, J.-M. Lehn, *Chem. Commun.* **2004**, 62–63.

42 C. Tie, J. C. Gallucci, J. R. Parquette, *J. Am. Chem. Soc.* **2006**, *128*, 1162–1171.

43 S. Choi, D. J. Clements, V. Pophristic, I. Ivanov, S. Vemparala, J. S. Bennett, M. L. Klein, J. D. Winkler, W. F. DeGrado, *Angew. Chem. Int. Ed.* **2005**, *44*, 6685–6689.

44 B. Gong, *Chem. Eur. J.* **2001**, *7*, 4336–4342.

45 B. Gong, H. Zeng, J. Zhu, L. Yua, Y. Han, S. Cheng, M. Furukawa, R. D. Parra, A. Y. Kovalevsky, J. L. Mills, E. Skrzypczac-Jankun, S. Martinovic,

R. D. Smith, C. Zheng, T. Szyperski, X. C. Zeng, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11583–11588.

46 J.-L. Hou, X.-B. Shao, G.-J. Chen, Y.-X. Zhou, X.-K. Jiang, Z.-T. Li, *J. Am. Chem. Soc.* **2004**, *126*, 12386–12394.

47 J.-L. Hou, H.-P. Yi, X.-B. Shao, C. Li, Z.-Q. Wu, X.-K. Jiang, L.-Z. Wu, C.-H. Tung, Z.-T. Li, *Angew. Chem. Int. Ed.* **2006**, *45*, 796–800.

48 Z.-Q. Wu, X.-K. Jiang, S.-Z. Zhu, Z.-T. Li, *Org. Lett.* **2004**, *6*, 229–232.

49 Z.-Q. Wu, X.-B. Shao, C. Li, J.-L. Hou, K. Wang, X.-K. Jiang, Z.-T. Li, *J. Am. Chem. Soc.* **2005**, *127*, 17460–17468.

50 H.-P. Yi, C. Li, J.-L. Hou, X.-K. Jiang, Z.-T. Li, *Tetrahedron* **2005**, *61*, 7974–7980.

51 H.-P. Yi, X.-B. Shao, J.-L. Hou, C. Li, X.-K. Jiang, Z.-T. Li, *New J. Chem.* **2005**, *29*, 1213–1218.

52 L. Yuan, A. R. Sanford, W. Feng, A. Zhang, J. Zhu, H. Zeng, K. Yamato, M. Li, J. S. Ferguson, B. Gong, *J. Org. Chem.* **2005**, *70*, 10660–10669.

53 L. Yuan, H. Zeng, K. Yamato, A. R. Sanford, W. Feng, H. S. Atreya, D. K. Sukumaran, T. Szyperski, B. Gong, *J. Am. Chem. Soc.* **2004**, *126*, 16528–16537.

54 J. Zhu, R. D. Parra, H. Zeng, E. SKrzypczak-Jankun, X. C. Zeng, B. Gong, *J. Am. Chem. Soc.* **2000**, *122*, 4219–4220.

55 J. Zhu, X.-Z. Wang, Y.-Q. Chen, X.-K. Jiang, X.-Z. Chen, Z.-T. Li, *J. Org. Chem.* **2004**, *61*, 6221–6227.

56 C. Li, S.-F. Ren, J.-L. Hou, H.-P. Yi, S.-Z. Zhu, X.-K. Jiang, Z.-T. Li, *Angew. Chem. Int. Ed.* **2005**, *44*, 2–6.

57 B. D. Hudson, R. Kuroda, W. A. Denny, S. Neidle, *Biomol. Struct. Dyn.* **1987**, *5*, 145–158.

58 R. P. Dixon, S. J. Geib, A. D. Hamilton, *J. Am. Chem. Soc.* **1992**, *114*, 365–366.

59 S. J. Geib, S. C. Hirst, C. Vicent, A. D. Hamilton, *Chem. Commun.* **1991**, 1283–1285.

60 R. Ostaszewski, Z. Urbahczyk-Lipkowska, *J. Mol. Struct.* **1999**, *474*, 197–206.

61 W. B. Blanton, S. W. Gordon-Wylie, G. R. Clark, K. D. Jordan, J. T. Wood, U. Geiser, T. J. Collins, *J. Am. Chem. Soc.* **1999**, *121*, 3551–3552.

62 J. D. Epperson, L.-J. Ming, B. D. Woosley, G. R. Baker, G. R. Newkome, *Inorg. Chem.* **1999**, *38*, 4498–4502.

63 H. Tang, R. J. Doerksen, G. N. Tew, *Chem. Commun.* **2005**, 1537–1539.

64 G. N. Tew, D. Liu, B. Chen, R. J. Doerksen, J. Kaplan, P. J. Carroll, M. L. Klein, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5110–5114.

65 H. Tang, R. J. Doerksen, T. V. Jones, M. L. Klein, G. N. Tew, *Chemistry & Biology* **2006**, *13*, 427–435.

66 C. Dolain, C. Zhan, J.-M. Léger, L. Daniels, I. Huc, *J. Am. Chem. Soc.* **2005**, *127*, 2400–2401.

67 R. W. Sinkeldam, M. H. C. J. van Houtem, G. Koeckelberghs, J. A. J. M. Vekemans, E. W. Meijer, *Org. Lett.* **2006**, *8*, 383–385.

68 R. W. Sinkeldam, M. H. C. J. van Houtem, K. Pieterse, J. A. J. M. Vekemans, E. W. Meijer, *Chem. Eur. J.* **2006**, *12*, 6129–6137.

69 J. J. van Gorp, J. A. J. M. Vekemans, E. W. Meijer, *Chem. Commun.* **2004**, 60–61.

70 S. Zhang, S. Prabpai, P. Kongsaereeb, P. I. Arvidsson, *Chem. Commun.* **2006**, 497–499.

71 D. A. P. Delnoye, R. P. Sijbesma, J. A. J. M. Vekemans, E. W. Meijer, *J. Am. Chem. Soc.* **1996**, *118*, 8717–8718.

72 L. He, Y. An, L. Yuan, K. Yamata, W. Feng, O. Gerlitz, C. Zheng, B. Gong, *Chem. Commun.* **2005**, *30*, 3788–3790.

73 J. Garric, J.-M. Léger, A. Grelard, M. Ohkita, I. Huc, *Tetrahedron Lett.* **2003**, *44*, 1421–1424.

74 M. C. Etter, Z. Urbanczyk-Lipkowska, M. Zia-Ebrahimi, T. W. Panunto, *J. Am. Chem. Soc.* **1990**, *112*, 8415–8426.

75 F. Böhme, C. Kunert, H. Komber, D. Voigt, P. Friedel, M. Khodja, H. Wilde, *Macromolecules* **2002**, *35*, 4233–4237.

76 P. S. Corbin, S. C. Zimmerman, P. A. Thiessen, N. A. Hawryluk, T. J. Murray, *J. Am. Chem. Soc.* **2001**, *123*, 10475–10488.

77 M. F. Mayer, S. Nakashima, S. C. Zimmerman, *Org. Lett.* **2005**, *7*, 3005–3008.

78 L. Xing, U. Ziener, T. C. Sutherland, L. A. Cuccia, *Chem. Commun.* **2005**, 5751–5753.

79 P. S. Corbin, S. C. Zimmerman, *J. Am. Chem. Soc.* **2000**, *122*, 3779–3780.

80 C.-H. Chien, M.-k. Leung, J.-K. Su, G.-H. Li, Y.-H. Liu, Y. Wang, *J. Org. Chem.* **2004**, *69*, 1866–1871.

81 A. Tanatani, H. Kagechika, I. Azumaya, R. Fukutomi, Y. Ito, K. Yamaguchi, K. Shudo, *Tetrahedron Lett.* **1997**, *38*, 4425–4428.

82 M. Barboiu, J.-M. Lehn, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5201–5206.

83 M. Barboiu, G. Vaughan, R. Graff, J.-M. Lehn, *J. Am. Chem. Soc.* **2003**, *125*, 10257–10265.

84 M. Barboiu, G. Vaughan, N. Kyritsakas, J.-M. Lehn, *Chem. Eur. J.* **2003**, *9*, 763–769.

85 D. M. Bassani, J.-M. Lehn, G. Baum, D. Fenske, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1845–1847.

86 L. A. Cuccia, J.-M. Lehn, H. J.-C., M. Schmutz, *Angew. Chem. Int. Ed.* **2000**, *39*, 233–238.

87 L. A. Cuccia, E. Ruiz, J.-M. Lehn, J.-C. Homo, M. Schmutz, *Chem. Eur. J.* **2002**, *8*, 3448–3457.

88 K. M. Gardinier, R. G. Khoury, J.-M. Lehn, *Chem. Eur. J.* **2000**, *6*, 4124–4131.

89 G. S. Hanan, J.-M. Lehn, N. Kyritsakas, J. Fischer, *J. Chem. Soc., Chem. Commun.* **1995**, 765–766.

90 F. Heirtzler, M. Neuburger, k. Kulike, *J. Chem. Soc., Perkin Trans. I* **2002**, 809–820.

91 F. R. Heirtzler, M. Neuburger, M. Zehnder, S. J. Bird, K. G. Orrell, V. Sik, *J. Chem. Soc., Dalton Trans.* **1999**, 565–574.

92 R. Krämer, I. O. Fritsky, *Eur. J. Org. Chem.* **2000**, 3505–3510.

93 M. Ohkita, J.-M. Lehn, G. Baum, D. Fenske, *Chem. Eur. J.* **1999**, *5*, 3471–3481.

94 M. Ohkita, J.-M. Lehn, G. Baum, D. Fenske, *Heterocycles* **2000**, *52*, 103–109.

95 A. Petitjean, L. A. Cuccia, J.-M. Lehn, H. Nierengarten, M. Schmutz, *Angew. Chem. Int. Ed.* **2002**, *41*, 1195–1198.

96 J.-L. Schmitt, J.-M. Lehn, *Helv. Chim. Acta* **2003**, *86*, 3417–3426.

97 J.-L. Schmitt, A.-M. Stadler, N. Kyritsakas, J.-M. Lehn, *Helv. Chim. Acta* **2003**, *86*, 1598–1624.

98 A.-M. Stadler, N. Kyritsakas, J.-M. Lehn, *Chem. Commun.* **2004**, 2024–2025.

99 J. M. Cary, J. S. Moore, *Org. Lett.* **2002**, *4*, 4663–4666.

100 X. Yang, A. L. Brown, M. Furukawa, S. Li, W. E. Gardinier, E. J. Bukowski, F. V. Bright, C. Zheng, X. C. Zeng, B. Gong, *Chem. Commun.* **2003**, 56–57.

101 X. Yang, L. Yuan, K. Yamato, A. L. Brown, W. Feng, M. Furukawa, X. C. Zeng, B. Gong, *J. Am. Chem. Soc.* **2004**, *126*, 3148–3162.

102 I. Smith, A. B., T. P. Keenan, R. C. Holcomb, P. A. Sprengeler, M. C. Guzman, J. L. Wood, P. J. Carroll, R. Hirschmann, *J. Am. Chem. Soc.* **1992**, *114*, 10672–10674.

103 M. Crisma, A. Moretto, C. Toniolo, K. Kaczmarek, J. Zabrocki, *Macromolecules* **2001**, *34*, 5048–5052.

104 J. Blankenstein, J. Zhu, *Eur. J. Org. Chem.* **2005**, 1949–1964.

105 F. Böhme, M. Rillich, H. Komer, *Macromol. Chem. Phys.* **1995**, *196*, 3209–3216.

106 S. Akine, T. Taniguchi, T. Nabeshima, *Tetrahedron Letters* **2001**, *42*, 8861–8864.

107 J. K.-H. Hui, M. J. MacLachlan, *Chem. Commun.* **2006**, 2480–2482.

108 A. Zhang, Y. Han, K. Yamato, X. C. Zeng, B. Gong, *Org. Lett.* **2006**, *8*, 803–806.

109 L. Yuan, W. Feng, K. Yamato, A. R. Sanford, D. Xu, H. Guo, B. Gong, *J. Am. Chem. Soc.* **2004**, *126*, 11120–11121.

110 H. Jiang, J.-M. Leger, P. Guionneau, I. Huc, *Org. Lett.* **2004**, *6*, 2985–2988.

111 I. Smith, A. B., M. C. Guzman, P. A. Sprengeler, T. P. Keenan, R. C.

Holcomb, J. L. Wood, P. J. Carroll, R. Hirschmann, *J. Am. Chem. Soc.* **1994**, *116*, 9947–9962.

112 I. Smith, A. B., W. Wang, P. A. Sprengeler, R. Hirschmann, *J. Am. Chem. Soc.* **2000**, *122*, 11037–11038.

113 I. Smith, B. Amos, D. A. Favor, P. A. Sprengeler, M. C. Guzman, P. J. Carroll, G. T. Furst, R. Hirschmann, *Bioorg. Med. Chem.* **1999**, *7*, 9–22.

114 C. Tomasini, G. Luppi, M. Monari, *J. Am. Chem. Soc.* **2006**, *128*, 2410–2420.

115 A. Itai, Y. Toriumi, S. Saito, H. Kagechika, K. Shudo, *J. Am. Chem. Soc.* **1992**, *114*, 10649–10650.

116 A. Tanatani, K. Yamaguchi, I. Azumaya, R. Fukutomi, K. Shudo, H. Kagechika, *J. Am. Chem. Soc.* **1998**, *120*, 6433–6442.

117 F. C. Krebs, M. Jorgensen, *J. Org. Chem.* **2002**, *67*, 7511–7518.

118 H. Masu, M. Sakai, K. Kishikawa, M. Yamamoto, K. Yamaguchi, S. Kohmoto, *J. Org. Chem.* **2005**, *70*, 1423–1431.

119 T. Nishimura, K. Maeda, E. Yashima, *Chirality* **2004**, *16*, S12–S22.

120 A. Tanatani, A. Yokoyama, I. Azumaya, Y. Takakura, C. Mitsui, M. Shiro, M. Uchiyama, A. Muranaka, N. Kobayashi, T. Yokozawa, *J. Am. Chem. Soc.* **2004**, *127*, 8553–8561.

121 K. Yamaguchi, G. Matsumura, *J. Am. Chem. Soc.* **1991**, *113*, 5474–5475.

122 K. Yamazaki, A. Yokoyama, T. Yokozawa, *Macromolecules* **2006**, *39*, 2432–2434.

123 M. Crisma, F. Formaggio, A. Moretto, C. Toniolo, *Biopolymers (Pept Sci)* **2006**, *84*, 3–12.

124 C. W. Wu, K. Kirshenbaum, T. J. Sanborn, J. A. Patch, K. Huang, K. A. Dill, R. N. Zuckermann, A. E. Barron, *J. Am. Chem. Soc.* **2003**, *125*, 13525–13530.

125 J.-C. Horng, R. T. Raines, *Protein Science* **2006**, *15*, 74–85.

126 B. K. Kay, M. P. Williamson, M. Sudol, *FASEB J.* **2000**, *14*, 231–241.

127 Z. Shi, K. Chen, Z. Liu, N. R. Kallenbach, *Chem. Rev.* **2006**, *106*, 1877–1897.

128 P. Armand, K. Kirshenbaum, R. A. Goldsmith, S. Farr-Jones, A. E. Barron, K. T. V. Truong, K. A. Dill, D. F. Mierke, F. Cohen, R. N. Zuckermann, E. K. Bradley, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4309–4314.

129 J. A. Patch, K. Kirshenbaum, S. L. Seurynck, R. N. Zuckermann, A. E. Barron, in *Pseudo-Peptides in Drug Discovery* (Ed.: P. E. Nielsen), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, **2004**, pp. 1–31.

130 K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen, R. N. Zuckermann, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4303–4308.

131 R. Simon, R. Kania, R. Zuckermann, V. Huebner, D. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. Marlowe, D. Spellmeyer, R. Tan, A. Frankel, D. Santi, F. Cohen, P. Bartlett, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9367–9371.

132 M. Lee, K.-Y. Kim, U.-I. Cho, D. W. Boo, I. Shin, *Chem. Commun.* **2003**, 968–969.

133 B. R. Huck, J. M. Langenhan, S. H. Gellman, *Org. Lett.* **1999**, *1*, 1717–1720.

134 B. R. Huck, J. D. Fisk, I. A. Guzei, H. A. Carlson, S. H. Gellman, *J. Am. Chem. Soc.* **2003**, *125*, 9035–9037.

135 C. Tomasini, V. Trigari, S. Lucarini, F. Bernardi, M. Garavelli, C. Peggion, F. Formaggio, C. Toniolo, *Eur. J. Org. Chem.* **2003**, 259–267.

136 F. Bernardi, M. Garavelli, M. Scatizzi, C. Tomasini, V. Trigari, M. Crisma, F. Formaggio, C. Peggion, C. Toniolo, *Chem. Eur. J.* **2002**, *8*, 2516–2525.

137 R. J. M. Nolte, *Chem. Soc. Rev.* **1994**, *23*, 11–19.

138 P. C. J. Kamer, M. C. Cleij, R. J. M. Nolte, T. Harada, A. M. F. Hezemans, W. Drenth, *J. Am. Chem. Soc.* **1988**, *110*, 1581–1587.

139 T. Nakano, Y. Okamoto, *Chem. Rev.* **2001**, *101*, 4013–4038.

140 N. Hoshikawa, Y. Hotta, Y. Okamoto, *J. Am. Chem. Soc.* **2003**, *125*, 12380–12381.

141 J. Roth, D. J. O'Leary, C. G. Wade, D. Miller, K. B. Armstrong, J. D. Thoburn, *Org. Lett.* **2000**, *2*, 3063–3066.

142 R. Mruk, R. Zentel, *Macromolecules* **2002**, *35*, 185–192.

143 H.-Z. Tang, Y. Lu, G. Tian, M. D. Capracotta, B. M. Novak, *J. Am. Chem. Soc.* **2004**, *126*, 3722–3723.

144 S. Dubus, V. Marceau, M. Leclerc, *Macromolecules* **2002**, *35*, 9296–9299.

145 A. J. Blake, P. A. Cooke, K. J. Doyle, S. Gair, N. S. Simpkins, *Tetrahedron Lett.* **1998**, *39*, 9093–9096.

146 M. J. Marsella, K. Yoon, A. Almutairi, S. K. Butt, F. S. Tham, *J. Am. Chem. Soc.* **2003**, *125*, 13928–13929.

147 J. Gawronski, K. Gawronska, J. Grajewski, K. Kacprzak, U. Rychlewska, *Chem. Commun.* **2002**, 582–583.

148 K. Kinbara, T. Aida, *Chem. Rev.* **2005**, *105*, 1377–1400.

149 A.-M. Stadler, N. Kyritsakas, R. Graff, J.-M. Lehn, *Chem. Eur. J.* **2006**, *12*, 4503–4522.

150 F. Zhang, S. Bai, G. P. A. Yap, V. Tarwade, J. M. Fox, *J. Am. Chem. Soc.* **2005**, *127*, 10590–10599.

2
Foldamers Based on Remote Intrastrand Interactions

Philippe Le Grel and Gilles Guichard

2.1
Introduction

In just a decade, the field of foldamers has gained respectability. The early contributions of Gellman [1, 2] and Seebach [3, 4] with foldamers based on remote intrastrand interactions have undoubtedly played a large part in the current dynamism of this rapidly evolving field. In the mid-1990s, both groups demonstrated that designed short-chain homo-oligomers made of unnatural units (β-amino acids) could self-organize in a controlled fashion to form defined secondary structures (e.g. helices, sheets and β-hairpins) reminiscent of those of α-polypeptides. To a large extent, unnatural backbones with folding propensity reported in this chapter are ''proteinomimetics''. In principle, intrastrand interactions in designed oligomers may result from a variety of noncovalent forces including steric repulsion, H-bonds, electrostatic and aromatic–aromatic interactions, coordination to metal ions, as well as solvophobic effects (see also Chapters 3 and 4). H-bonds play an important role in foldamer design, in part because they are robust and directional. H-bonding certainly provides the most versatile way to create intrastrand connections that are useful to stabilize intrinsically flexible oligomeric chains into ordered structures. In this context, the amide linkage has regularly appeared as a motif to elaborate oligomeric strands that self-organize through H-bonding. In recent years, nonpeptide backbones, heterogeneous (hybrid) oligoamide backbones composed of multiple residue types as well as abiotic aromatic oligoamides have emerged as new classes of folded oligomers. These systems, whose propensity for folding is also controlled by noncovalent interactions between non-nearest neighbors, are surveyed in this chapter. The main challenges lie in the ways to control backbone pre-organization, to develop robust predictable secondary motifs and to integrate multiple levels of complexity to create unprecedented folded shapes. Recent studies have highlighted the strength and usefulness of theoretical studies to explore possible secondary structure ensembles in aliphatic oligoamides of various complexities (see also Chapter 6). In a number of systems based on remote intrastrand interactions, the predictability

of folding can now be integrated to develop molecules with function. Successful applications of these foldamers in biology are further discussed in Chapter 8.

2.2
What can be Learned from Strategies used to Control Conformations of α-Polypeptides?

Remarkable progress has been achieved over the last 20 years in understanding the factors that govern folding of linear α-polypeptide strands into well-ordered and compact secondary structures, as well as in the *de novo* design of individual protein secondary structure elements and protein folds from α-polypeptides (see also Chapter 4) [5–10]. The principles that guided the design of well-folded α-polypeptides have been an invaluable source of inspiration to those in the field of foldamers based on remote H-bonding interactions.

Formation of stable and regular secondary structures maintained by intramolecular H-bonds (e.g. helices) requires pre-organization of the main chain so that sequentially remote H-bond donor and acceptors can be positioned in close spatial vicinity, and optimal H-bonding can occur without significant conformational alteration [11]; (in α-polypeptides, the preferred backbone conformation derives in part from minimization of Newman and Pitzer strain, as well as pseudo-allylic a(1,3) strain, which restricts the ϕ and ψ torsion angle values accessible to proteinogenic amino acid residues). A number of approaches have been developed to reinforce the stability of α-peptide folds. Proteinogenic α-amino acids have intrinsic and distinct propensities for helices and sheets and can be selected accordingly to stabilize a given fold [12–14]. For example, in the case of noncharged residues, Ala and Leu are strong helix inducers compared with β-branched Val which instead has an intrinsic propensity to stabilize sheet structures. A higher degree of stabilization can be achieved by restricting further the available conformational space of amino acids in the sequence [5]. This approach has been instrumental in the successful design of foldamers based on remote intrastrand interactions (see following sections). In α-peptides, Thorpe–Ingold effects (C(α)-tetrasubstitution) have been used extensively to impose such a restriction on ϕ and ψ angles [15]. Aib (α-aminoisobutyric acid), the archetype of achiral α,α-dialkylated residue with (ϕ,ψ) values around $(+60°, +30°)$ and $(-60°, -30°)$ is a remarkably strong promoter of helical (3_{10} and α-helices) and β-turn (type III/III′ and type I/I′) structures [5, 10].

The α-helix is polarized from N to C terminus. Hence, its stability can be enhanced by suppressing repulsive electrostatic interactions between the terminal charges and the helix dipole, using appropriate capping groups [12, 16]. Conversely, favorable charge–dipole interactions can be introduced by positioning residues with appropriately charged side chains close to the N and C termini [12]. The creation of ion pairs or salt bridges between oppositely charged side chains of amino acids separated by one helical turn (i and $i + 4$ positions) has been successfully applied to stabilize α-helical secondary structures [17]. Other helix stabi-

lizing $i/i + 4$ side-chain to side-chain interactions include hydrophobic interactions [12] and polar interactions with aromatic side chains [18, 19]. The spatial proximity of residues in a $i/i + 4$ relationship within the α-helical backbone also offers the possibility of constructing various types of helix-stabilizing macrocycles closed by metal [20, 21], disulfide [22], lactam [23] or olefinic [24, 25] bridges.

2.3
Helices from Homogeneous Oligomeric Backbones with Periodicity at the Monomer Level: ω-Peptides and their Analogs

Helices (single-stranded and multi-stranded) represent a major structural motif in biological macromolecules and are interesting scaffolds for displaying appendages in a controlled direction within the 3-D space. Aliphatic oligoamides built from enantiopure β- or γ-amino acid residues, namely β- and γ-peptides are the archetypal helix-forming foldamers. The conformational preference of these homogeneous backbones in solution and in the solid state has been extensively studied and reviewed in detail [2, 4, 26–35]. Five helical shapes differing by the size and orientation of their H-bonded ring have been identified in β-peptides depending upon residue substitution pattern and stereochemistry, thus illustrating the remarkable structural diversity of these synthetic oligomers: namely 14-, 12- (Section 2.3.1), 10- (H-bonding scheme between two consecutive units, mentioned in Section 2.3.1), 8- (Section 2.3.2) and mixed 12/10-helices (see Section 2.4). Noteworthy, the first hint that a polypeptide chain composed of β- or γ-amino acid residues could adopt a regular helical fold structure came from studies of β- and γ-amino acid homopolymers some 40 years ago [36, 37]. It is worth mentioning that a substantial increase in the number of unnatural backbones with propensity to form regular helices came from exploration of the β- and γ-peptide lineages (i.e. β- and γ-peptide mimetics).

2.3.1
Compact Helices with Large (>10 atoms) H-bonded Rings

2.3.1.1 The Homologation Strategy: β- and γ-Peptide Foldamers
Selecting (designing) the right monomeric units to build homo-oligomers that will ultimately show high (helical) folding propensity is obviously a critical and limiting step in the foldamer arena. The choice of β-amino acids and corresponding β-peptides by Seebach [3] and Gellman [1] was not pure serendipity. It was initially guided (for Seebach) by the resemblance of the β-peptide backbone to poly((R)-3-hydroxybutanoic acid) a biopolymer for which a right-handed helical conformation with about three units per turn had been proposed [39, 40], and (for Gellman) by the absence of H-bonding between nearest neighbor amide groups in model β-homoglycine (β-HGly) derivatives [41], a criterion used to evaluate the propensity of the β-peptide backbone to stabilize folds maintained by H-bonds between sequentially remote amide groups.

A

α-peptide

α-helix [(P)-3.6$_{13}$]
or 3$_{10}$-helix

β^3-peptide

14-helix [(M)-3$_{14}$]

β^2-peptide

14-helix [(P)-3$_{14}$]

γ^4-peptide

14-helix [(P)-2.6$_{14}$]

B

NH$_i$ ⟶ CO$_{i+2}$

1

C

CO$_i$ ⟵ NH$_{i+3}$

2
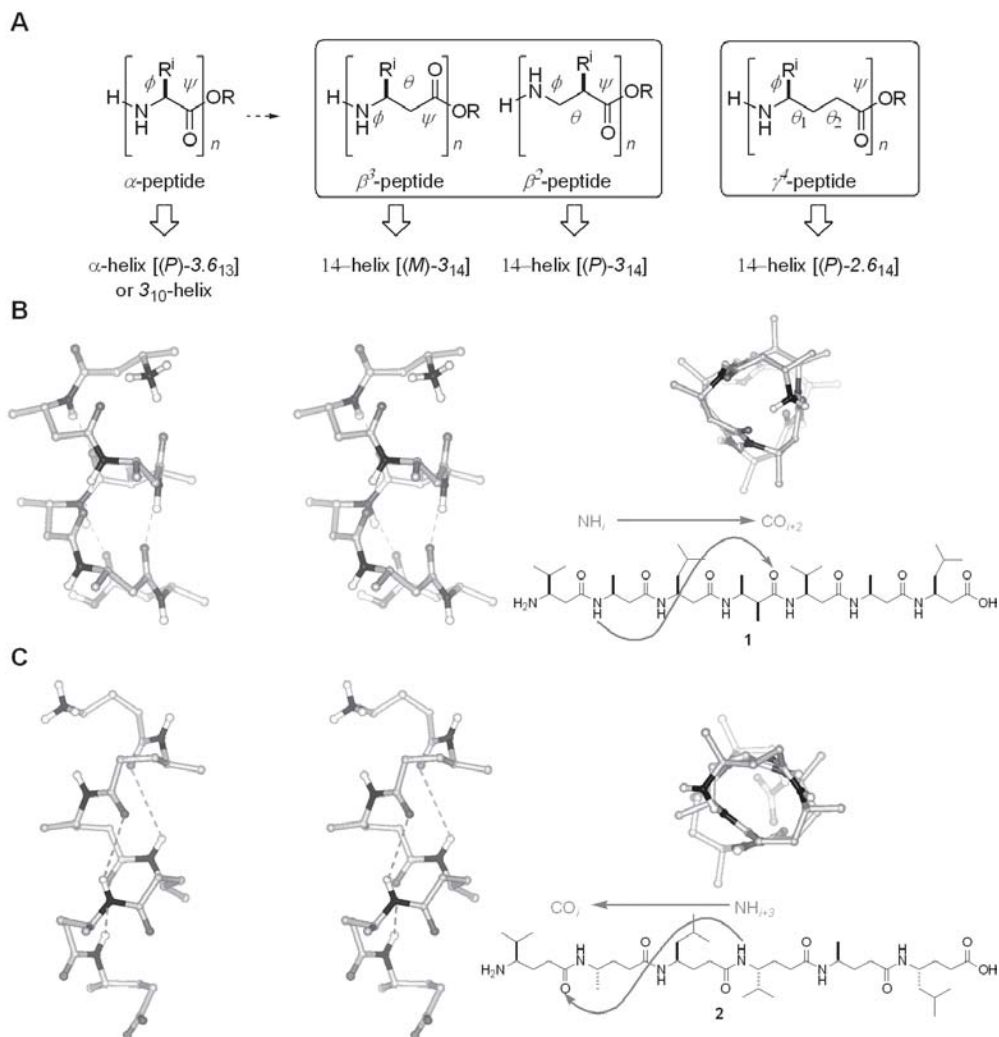
Fig. 2.1 The homologation approach to aliphatic oligoamide foldamers. (A) Peptide foldamers consisting of homologated α-amino acids generated by insertion of one ($\to \beta^2$ and β^3-peptides) or two ($\to \gamma^4$-peptides) CH$_2$ groups, the amino acid side chain remaining unchanged. According to the nomenclature proposed by Balaram [49], the conformational space of β- and γ-peptides can be described by the following sets of backbone torsion angles: $(\omega, \phi, \theta, \psi)$ and $(\omega, \phi, \theta_1, \theta_2, \psi)$, respectively. (B) The β-peptide 3$_{14}$ helical fold. Stereo view along the helix axis and top view of the (M)-3$_{14}$-helix formed by β^3-peptide **1** determined by NMR in CD$_3$OH (adapted from [43, 44]). Average (ϕ, θ, ψ) values are $(-119°, +63°, -147°)$. Side chains have been omitted for clarity. (C) The γ-peptide 2.6$_{14}$ helical fold. Stereo view along the helix axis and top view of a (P)-2.6$_{14}$-helical low energy conformer of **2** (from NMR data in pyridine-d$_5$). Average $(\phi, \theta_1, \theta_2, \psi)$, values calculated on residues 2–5 are $(-127°, +66°, +64°, -140°)$. Side chains have been partially omitted for clarity (adapted from [45]).

Detailed NMR conformational analysis of β-peptides consisting of homologated α-amino acids generated by insertion of one CH_2 group, the amino acid side chain remaining unchanged, revealed a stable ($\beta^3 \rightarrow \beta^2$-peptides) helical fold (3_{14}- or 14-helix) stabilized by 14-membered rings with H-bonds in a forward orientation ($1 \rightarrow 3$ H-bonds between NH_i and $C=O_{i+2}$) (e.g. **1**, [3, 42–44] Fig. 2.1).

Examination of the top view of the helix indicates that the side chains of residues i and $i + 3$ are located nearly on top of each other and suggests that hydrophobic interactions between overlapping aliphatic side chains could play a significant role in stabilizing the overall structure. The distance between facing $C(\alpha)$ atoms at positions i and $i + 3$ is approximately 4.8 Å. The helix is compact with a diameter of ca. 4.8 Å slightly larger than that of the α-helix (4.2 Å). Although less studied, their homologs with one additional methylene group inserted in the backbone of each residue, namely the γ-peptides, have also been found to form stable helical secondary structures in solution [45–48]. γ^4-Peptide chains (e.g. **2**) adopt a 2.6_{14} helical structure stabilized by $1 \leftarrow 4$ H bonds between $C=O_i$ and NH_{i+3} closing 14-membered pseudocycles [45]. While the α-helix of L-α-peptides and the 3_{14} helix of the corresponding β^3-peptides have opposite polarity and helicity, the insertion of two CH_2 groups in the backbone of L-α-amino acids leaves these two helix parameters unchanged, both the α-helix and the 2.6_{14} helix of the resulting γ^4-peptides being right-handed and polarized from N to C terminus. Both (M)-3_{14} and (P)-2.6_{14} helical backbones are characterized by a $(+)$-*synclinal* arrangement (*gauche* conformation) around ethane bonds (e.g. in **1** and **2**, θ, θ_1 and θ_2 values are $\sim +60°$ ($\pm 15°$) (see Fig. 2.1 for definitions of torsion angles [49]).

Besides NMR and X-ray diffraction (see Section 2.3.1.2), circular dichroism (CD) spectroscopy has been instrumental in studying parameters that influence the formation and stability of the β-peptide 14-helix. Typically, 14-helical β-peptides composed of acyclic amino acid residues display a common CD signature with one extremum at 215 nm (negative for (M)-helices and positive for (P)-helices) and the other extremum of opposite sign at 195 nm. In sharp contrast to α- and β-peptides, CD spectra of γ-peptides were not very informative. γ^4-Peptides such as **2** which populate the 2.6_{14} helical fold in solution do not display any characteristic CD signature in MeOH.

2.3.1.2 Imposing Backbone Conformational Restriction/Pre-organization for Optimal Helical Folding

Substantial stabilization of both the β- and γ-peptide 14-helical fold has been achieved by increasing the level of pre-organization of β- and γ-amino acid constituents. However, the rules formulated for α-peptides (see Section 2.2) do not necessarily apply (e.g. $C(\alpha)$- (or $C(\beta)$-) tetrasubstitution of β-amino acid residues is not compatible with β-peptide 14-helix formation [43, 50–54]) and must be transposed. In particular, it has been shown that acyclic $\beta^{2,3}$-amino acids of *like* configuration are more effective than their β^3-counterparts in promoting requisite synclinal arrangement around $C(\alpha)$–$C(\beta)$ bonds in β-peptides (Fig. 2.2A) [43, 55]. With a θ value fixed at approx. $\pm 60°$, *trans*-2-aminocyclohexyl carboxylic acid
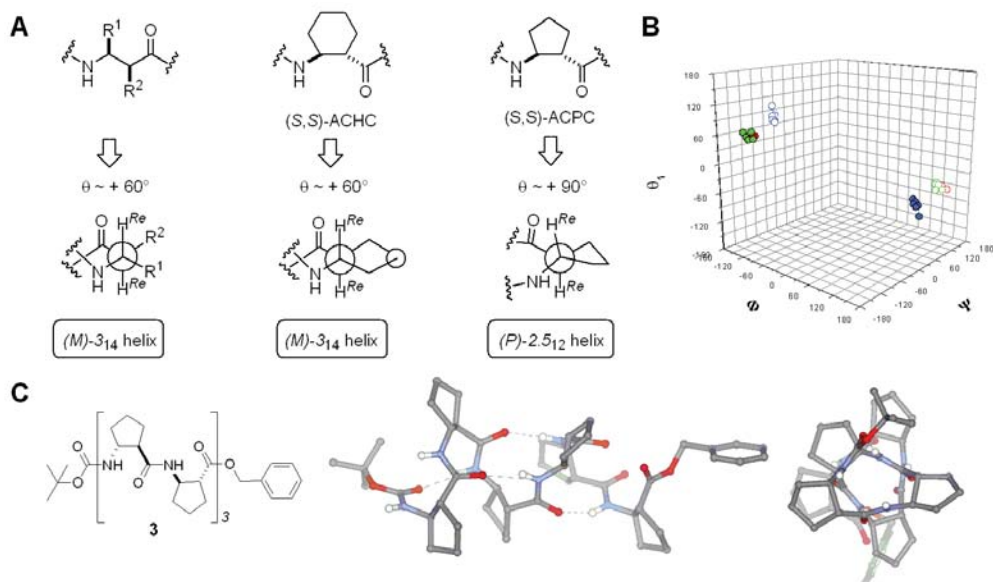
Fig. 2.2 Optimal β-peptide backbone pre-organization for helix formation. (A) Acyclic $\beta^{2,3}$ amino acid residues and (S,S)-ACHC residues which promote gauche conformation arouch the $C(\alpha)-C(\beta)$ bond are 14-helix stabilizers. The related five-membered ring ACPC promotes larger θ values which are not compatible with the 14-helical conformation. (B) Comparison of β-peptide 14- and 12-helices (ϕ, θ, ψ) conformational space. (M)-14-helix (values for **1**, red closed circles, values for **2**, green closed circles), (M)-12-helix (values for **3**, blue closed circles). The enantiomeric positions in the (ϕ, θ, ψ) space are shown as open circles. (C) Side view and top view of the (M)-12-helix formed by (R,R)-ACPC hexamer **3** as determined by X-ray diffraction (adapted from [59]). Average (ϕ, θ, ψ) values for central residues are $(88°, -85°, 98°)$.

(*trans*-ACHC), a cyclic $\beta^{2,3}$-amino acid, in which the $C(\alpha)-C(\beta)$ bond is part of a 6-membered ring, is ideally pre-organized for 3_{14} helix formation (Fig. 2.2A) [1, 56, 57]. In contrast, homo-oligomers (e.g. **3**) consisting of the smaller ring size *trans*-2-aminocyclopentyl carboxylic acid (*trans*-ACPC) for which larger values of θ only are accessible, adopt a stable 12- (2.5_{12})-helix with a $(1 \leftarrow 4)$ H-bonding pattern that differs markedly from the 14-helix (Fig. 2.2) [58–60]. The 14- and 12-helices populated by (S,S)-ACHC and related (S,S)-ACPC oligomers, respectively have opposite polarity and helical screw sense. The (P)-12-helix display a CD pattern distinct from that of the corresponding 14-helix with a maximum at 204 nm, zero crossing at 214 nm and minimum at 221 nm.

β-Amino acids constrained with smaller rings such as *cis*-aminooxetane carboxylic acids have been shown to promote the formation of a 10-helical structure (e.g. **4**) with H-bonds between neighboring amide units $((1 \rightarrow 2)$ H-bonding scheme) [61]. The 10-helix was later identified in short ACHC oligomers (tetramer), suggesting that it could represent a conformational intermediate in the folding pro-

cess toward the thermodynamically stable 14-helix [62]. The incorporation of cyclic ω-amino acid units to fix the peptide backbone in a geometry favorable to helix formation has since been widely utilized (see also following sections) [63–64]. This approach was recently extended to higher oligoamides such as δ-peptides [66–69]. Conformational search using the methods of *ab initio* MO theory identified several H-bonded helical backbones accessible to δ-peptides [70]. Experimental validation came from studies with oligomers composed of carbohydrate-derived tetrahydrofuran amino acids with restricted rotation around $C(\alpha)–C(\beta)$ and $C(\beta)–C(\gamma)$ bonds. In chloroform solution, octamer **5** was found to adopt a well-defined 16-helical fold with $1 \leftarrow 4$ H-bonding pattern, reminiscent of the α-peptide π-helix [69].

R = TBDMS, Bn
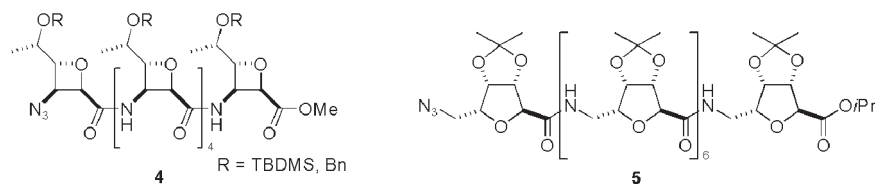
Scheme 2.1

The intrinsic conformational preferences of substituted γ-amino acid constituents of γ-peptides derive in part from avoidance of destabilizing *syn*-pentane interaction [71, 72]. It is recognized that this effect plays a key role in fixing the bioactive conformation of a number of γ-amino-acid-containing natural products such as Bleomycin A_2 [73]. In γ^4-peptides five out of nine conformations generated by rotation around $C(\alpha)–C(\beta)$ and $C(\beta)–C(\gamma)$ bonds are free of *syn*-pentane interaction.

Adding substitutents at the 2-position (*like* configuration, see Fig. 2.3A) or at both 2 and 3-positions (*like,like* relative configuration) reduces to two the number of conformations devoid of *syn*-pentane interaction (conformation II in Fig. 2.3A is almost identical to that found in the γ^4-peptide 14-helical backbone) and thus reinforces optimal pre-organization for 14-helix formation ($\gamma^{2,4}$ and $\gamma^{2,3,4}$-peptides) [45–48]. Other strategies to restrict the conformational space of the γ-amino acid backbone, such as α,β-unsaturation (*cis*-vinologous γ-peptides [74]), tetrasubtitution [oligomers of 1-(aminomethyl)cyclohexaneacetic acid (gabapentin, Gpn) [75]] cyclic γ-amino acids (*cis*-γ-amino-L-proline oligomers [76]) have been reported.

The γ-peptide 14-helical backbone is characterized by large ψ values ranging from 120 to 150° (or −120 to −150°). This observation led to the finding that substituting urea for the $CH_2–CO–NH$ unit in γ-peptide (substitution of nitrogen for $C(\alpha)$ in γ-amino acids) can be used to rigidify the 14-helical fold by fixing the "ψ" angle to a value close to 170–180° (Fig. 2.3B) [77–80]. In methanol or pyridine solution, the resulting enantiopure N,N'-linked oligoureas (e.g. **6**) adopt a well-
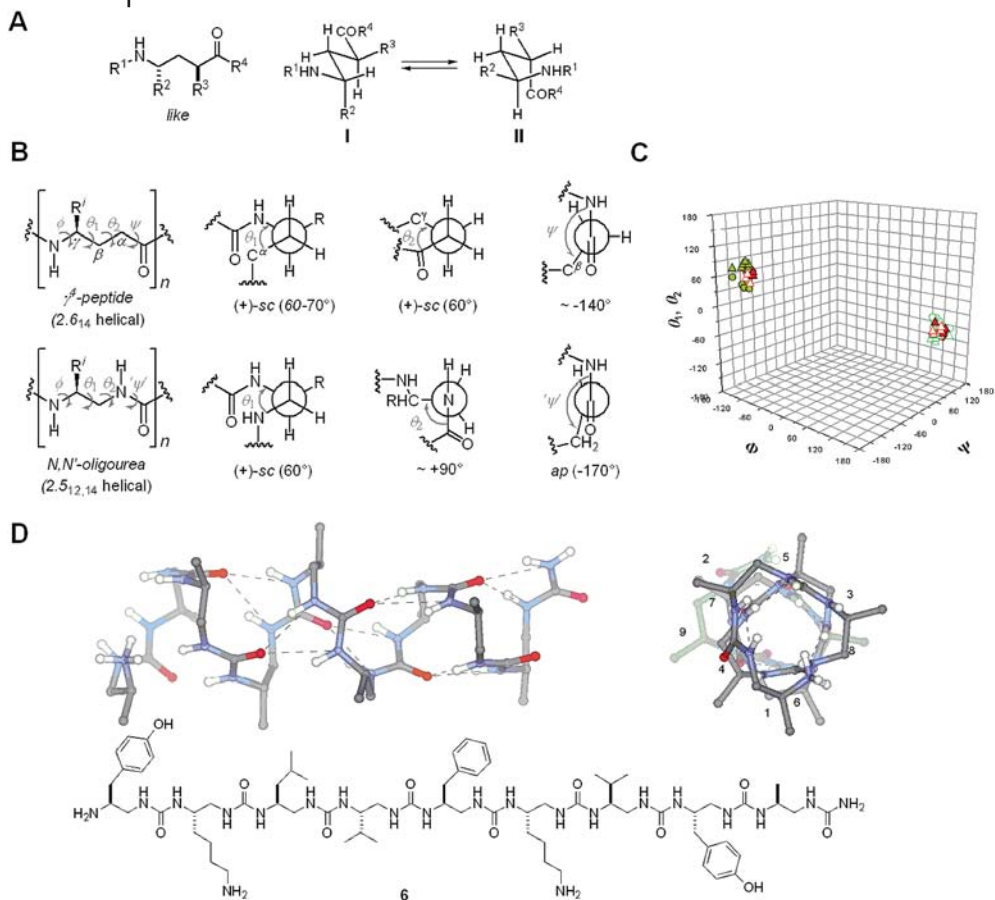
Fig. 2.3 (A) The two conformations free of destabilizing *syn*-pentane interaction [71, 72] in 2,4-disubstituted γ-amino acid derivatives with like configuration. Conformation II is close to that found in the γ^4-peptide 14-helical backbone (see Fig. 2.1) (B) Structural analogy between the γ^4-peptide and *N,N'*-linked oligourea backbones. (C) (ϕ, θ_1, ψ) and (ϕ, θ_2, ψ) maps indicating γ-peptide 14-helix and *N,N'*-linked oligourea 12,14-helix regions. 14-Helix [values for a γ^4- and a $\gamma^{2,3,4}$-peptide [45, 47], red closed circles (ϕ, θ_1, ψ) and closed triangles (ϕ, θ_2, ψ)], 12,14-helix [values for **6**, green closed circles (ϕ, θ_1, ψ) and closed triangles (ϕ, θ_2, ψ)]. The enantiomeric positions in the $(\phi, \theta_1/\theta_2, \psi)$ space are shown as open circles/triangles. (D) The (*P*)-12,14-helical structure of *N,N'*-linked oligoureas. View along the helix axis and top view of a low energy conformer of nonamer **6** as determined by NMR spectroscopy and restrained molecular dynamics calculations in pyridine solution. Average $(\phi, \theta_1, \theta_2, \psi)$ values for central residues are $(-105°, 55°, 88°, -168°)$. With an internal diameter of $\sim$3 Å, the helix is particularly compact and is devoid of empty volume in the interior. Side chains have been partially omitted for clarity (adapted from [78]).

defined 2.5 helical fold, reminiscent of the γ^4-peptide 14-helix (Fig. 2.3C and D). The structure is held by H-bonds closing both 12- and 14-membered rings (12,14-helix) and is characterized by a stable $(+)$-synclinal arrangement around the ethane bond. CD spectra recorded in MeOH display a characteristic signature with an intense maximum near 204 nm [79]. This is in contrast to related helical γ^4-peptides that do not exhibit any characteristic CD signature.

Alternatively, in a manner analogous to α-peptides (see Section 2.2), a helical backbone may be stabilized by creating a covalent linkage (e.g. disulfide bond) between two spatially proximal but sequentially remote side chains (e.g. $i/i+3$ side chains in the 3_{14} helix) [81].

2.3.1.3 Folding in an Aqueous Environment

The nature of the solvent can influence to a large extent the propensity of unnatural oligomers to adopt a given H-bonded fold. Conformational studies aimed at identifying new foldamers are often performed in apolar or moderately polar organic solvents (e.g. chloroform, MeCN, pyridine, trifluoroethanol (TFE), MeOH). However, determination of a folding pattern in an aqueous environment is highly relevant to applications of foldamers in biology.

Considerable efforts have been undertaken to address this issue in the case of 14- and 12-helical β-peptides. To increase water solubility of helical β-peptides composed of ACHC (14-helix promoter) or ACPC (12-helix promoter) oligomers while maintaining the level of backbone pre-organization, Gellman and co-workers developed amino-functionalized versions of *trans*-2-aminocycloalkane carboxylic acids. β-Peptides composed of alternating ACHC/DCHC [82] and ACHC/APiC [83] residues adopt a robust 14-helical conformation in aqueous solution (Fig 2.4A). Similarly, ACPC/APC [84], ACPC/AP [85], ACPC/3-aminomethyl-ACPC [86] repeats promote stable 12-helix formation (Fig 2.4A).

Alternatively, the introduction of a limited number of acyclic β^3- or β^2-Lys residues (1/3) in ACHC- and ACPC-peptides does not preclude the formation of stable 14- and 12-helical structures in aqueous solution [87–89]. Although helicity in water is intrinsically weaker in the absence of strong backbone pre-organization (e.g. β^3-peptides), principles guiding the design of β^3-peptides with high levels of 14-helicity in aqueous solution (e.g. **7**, Fig. 2.4B) have recently been delineated by several groups [90–96]. They parallel those formulated for α-helical α-peptides (see Section 2.2) and include: (i) salt bridge or lactam formation between complementary charged $i/i+3$ side chains (e.g. β^3-HOrn/β^3-HGlu; β^3-HOrn/β^3-HAsp; β^3-HDab/β^3-HAsp; Dab = 2,4-diaminobutyric acid; Orn = ornithine); (ii) maximization of electrostatic interactions with the helix macrodipole (e.g. by free charged termini, appropriate location of charged side chains, appropriate orientation of salt bridges) and (iii) introduction of β^3-amino acids with high intrinsic 14-helix propensitiy. For noncharged β^3-amino acids, 14-helix propensities have been found to differ significantly from α-helix propensities of corresponding α-amino acids. Ala is the most α-helix-stabilizing α-amino acid but the methyl side chain is one of the least 14-helical stabilizing. In contrast, branched side chains of Ile, Thr, and Val which display only moderate to low α-helix propensity are all

A

14-helix

ACHC / DCHC   **ACHC / APiC**

12-helix

ACPC / APC   **ACPC / AP**   **ACPC / 3-aminomethyl-ACPC**

B

7
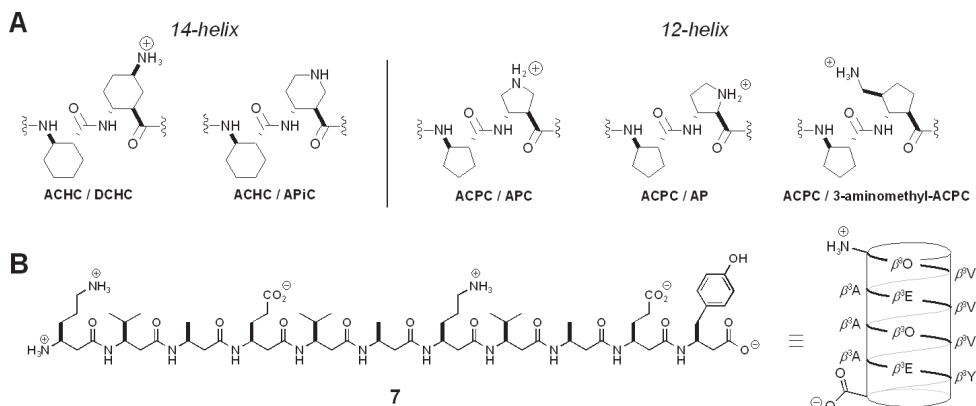
Fig. 2.4 β-Peptides that promote helix formation in water. (A) Using amino-functionalized versions of *trans*-2-aminocycloalcane carboxylic acids [82–86]. (B) *De novo* design of β^3-peptide with high level of 14-helicity in aqueous solution by combining: salt bridges between $i/i + 3$ side chains, favorable electrostatic interactions with the helix macrodipole and side-chain branching (β^3-HVal residues). Mean residue ellipticity of **7** measured by CD at 214 nm: $\Theta_{214} = -13\,320$ deg cm^{-2} dmol^{-1} in 1 mM sodium phosphate/borate/citrate, pH 7.0 at 25 °C [95].

strongly 14-helix stabilizing. Intramolecular interhelical hydrophobic interactions (in a β^3-peptide two-helix bundle) have also been shown to increase 14-helicity in aqueous solution [97].

If β-peptides are indisputably the best characterized helical aliphatic peptide foldamer system in aqueous solution so far, some of the aromatic foldamers described in Chapter 1 fold into water even better than they do in organic solvents without any specific adjustment of their structure. It remains to be seen whether other foldamers based on remote intrastrand H-bonds can be designed to adopt robust helical structures in water.

2.3.1.4 Dynamics of β- and γ-Peptide Helices: Evidence for Noncooperative Folding/Unfolding Processes

The unfolding and folding mechanisms of 14-helical β-, γ-peptides and analogs have been investigated in polar solvents by various approaches. Temperature-dependent-CD and NMR measurements suggest that 14-helical β-peptides [44] and also 14-helical $\gamma^{2,3,4}$-peptides [48] and helical N,N'-linked oligoureas [77–79], undergo noncooperative unfolding upon heating in MeOH. For instance, the intensity of the extremum at 215 nm for β-peptide **1** decreases linearly (by ca 12% per 20 K) and noncooperative break-up of the structure is observed between 295 and 333 K. The scan, which is reversible, suggests that the unfolding and folding route of the helix must be reversible [44]. Insight into the dynamics of β-peptides and evidence for reversible folding were also provided by molecular dynamics (MD) simulations (GROMOS96 force-field [98], reviewed in Chapter 6) in explicit solvents and at different temperatures [99–104]. By simulating on a

time scale that is long compared with the lifetime of any specific conformation (typically $\geq$ 50 ns), it has been possible to determine the population and average lifetimes of the different conformations observed and to explore paths and rates of interconversion between the experimentally observed 14-helix conformation and (partially) unfolded conformations [102, 104]. In the course of a 50-ns simulation and irrespective of its initial conformation, β-peptide **1** folds rapidly (in the order of a few nanoseconds) into the experimentally observed 14-helix conformation (maximum lifetime $\sim$ 10 ns at 340 K) which is populated 50% of the time at 340 K [99, 101].

Recent progress towards the synthesis of the spin-labeled $\beta^{2,3}$-amino acid *trans*-β-TOAC (4-amino-1-oxyl-2,2,6,6-tetramethylpiperidine-3-carboxylic acid) and its incorporation into ACHC-peptides at $i/i + 3$ positions suggest that it might soon be possible to use electron spin resonance (ESR) to investigate further the structures and folding transitions in 14-helical β-peptides [105].

2.3.2
Extended Helices with Small H-bonded Rings Centered at a Single Residue

This section focuses on small (generally $\leq$ 10 atoms) pseudocycles with $1 \leftarrow 3$ H-bond interaction as potent helix building blocks. The pseudocycle is centered at the $(i + 1)$ residue, with an H-bond joining the C=O and NH groups at positions i and $i + 2$ respectively (Fig. 2.5A). This H-bonding scheme recurs in the β- and γ-peptide lineage subject to the presence of strong pre-organization structural elements in monomeric units (e.g. ring, backbone or side chain heteroatoms).

2.3.2.1 α-Peptides: the γ-Helix
In α-peptides, $1 \leftarrow 3$ H-bond interaction corresponds to a seven (C_7) H-bonded ring, i.e. the γ-turn ($\phi \sim +70°$, $\psi \sim -70°$) or the more stable inverse γ-turn ($\phi \sim -70°$, $\psi \sim +70°$) according to the equatorial or axial orientation of the side chain [106]. They are quite common structuring elements of cyclic peptides, and can also be observed in crystallized proteins, although much less frequently than β-turns. A series of consecutive inverse γ-turns generates a theoretical *2.2₇*-helix called γ-helix [107] (Fig. 2.5B), which has, however, not been detected in natural peptides or proteins so far. Recently, a first step was accomplished towards the construction of an artificial γ-helix template. Jiménez et al. showed that the dipeptide Ac-L-Pro-D-c₃Dip-NHMe (**8**) adopts two consecutive γ-turns, induced by the presence of the D-form of the cyclopropane amino acid (c₃Dip) derivative (Fig. 2.5B) [108].

2.3.2.2 ω-Peptides with Specific Conformation-stabilizing Elements

β-Peptides: the 8-helix   The world of β-peptides is mostly associated with the prominent 12- and 14-helical folds. Considering that their topology excludes the presence of axial substituents (see Fig. 2.2), β-peptides consisting of geminally disubstituted amino acids or of $\beta^{2,3}$-amino acids of *unlike* configuration (see
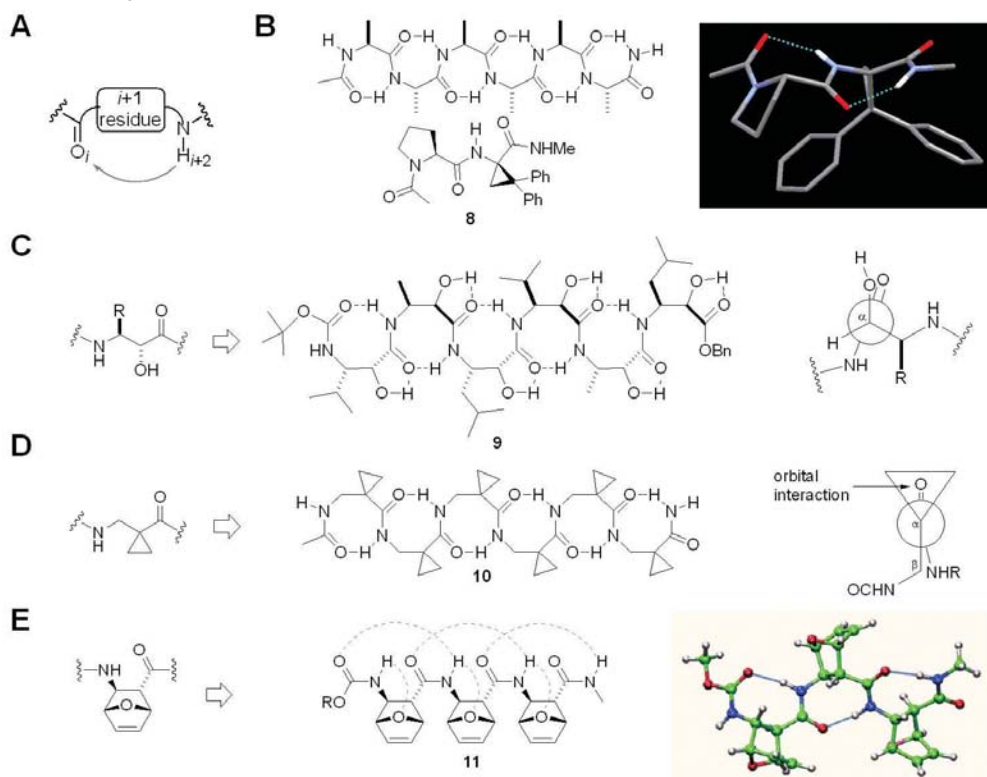
Fig. 2.5 (A) $1 \leftarrow 3$ H-bonding scheme in ω-peptides and analogs. (B) The α-peptide γ-helix and the two consecutive C_7 turns observed experimentally for dipeptide **8** in the solid state [108]. (C–E) Expansion of the α-peptide C_7 structure by the addition of one backbone atom: the C_8 H-bonded conformations of β-peptides **9–11** consisting of specific conformation-stabilizing elements: (C) (2*R*,3*S*)-α-hydroxylated $\beta^{2,3}$-amino acid [109], (D) 1-aminomethylcyclopropane-carboxylic acid [50], and (E) *trans*-oxabornene-β-amino acid residues (B3LYP/6-311++G** minimum energy conformation) [110].

also Section 2.5) cannot fit in any of the two folds. Seebach and collaborators found that β-peptides consisting of (2*R*,3*S*)-α-hydroxylated $\beta^{2,3}$-amino acid residues (e.g. **9**) exist in polar solvent as a helical conformation based on repetitive 8-membered H-bonded rings resulting from $1 \leftarrow 3$ H-bond interactions (C=O$_i$···H–N$_{i+2}$) [109]. The structure is probably further stabilized by an additional interaction between C=O$_i$···H–O$_i$ (Fig. 2.5C). A remarkably similar C_8-based conformation was reported for $\beta^{2,2}$-substituted oligomers consisting of 1-aminomethylcyclopropanecarboxylic acid residues (Fig. 2.5D) [50]. X-ray diffraction studies of short chain oligomers revealed that orbital hyperconjugation between the σ-bonding orbitals (HOMO) of the cyclopropane ring and the π^* nonbonding orbital (LUMO) of the carbonyl constrained ψ values close to 0° (bisected

conformation), thus favoring a C_8 H-bonding pattern. A ribbon-like secondary structure was extrapolated from these data.

Independently, using ^{1}H NMR and density functional theory computations, Klein and co-workers concluded that the β-peptide consisiting of *trans*-oxabornene-β-aminoacid (e.g. **11**, Fig. 2.5E) also adopts a C_8-based helix conformation [110]. These results show that cyclohexyl ring bridging and unsaturation impose angular constraints that translate the robust 3_{14}-helix sustained by *trans*-ACHC-$\beta^{2,3}$ amino acid units (see Section 2.3.1.2) [1, 56, 57], into a new folding pattern. Additional H-bond contacts between amide NHs and the ring oxygens were also hypothesized. Altogether, these results do not question the proposal formulated by Gellman [41] that $1 \leftarrow 3$ H-bonding between nearest-neighbor amide groups in β-peptides is not favored, but rather suggest that extra-interaction or specific angular constraints can overcome this general feature.

γ-Peptides   According to model studies [41], γ-peptides have higher propensities than β-peptides to populate conformations stabilized by H-bonding between nearest neighbor amide groups. Several examples of secondary structures stabilized by $1 \leftarrow 3$ H bonds have been identified in designed γ-peptides incorporating various levels of backbone pre-organization [74–76, 111]. In the case of gabapentin (Gpn), a $\gamma^{3,3}$-geminally disubstituted amino acid, both $C(\alpha)$–$C(\beta)$ and $C(\beta)$–$C(\gamma)$ bonds are locked in a gauche conformation with $\theta_1 \sim \theta_2 \sim \pm 60°$ [75]. In the solid state, Gpn oligomers populate C_9 H-bonded conformations. Because the molecules are achiral, two sets of dihedral angles $(+)$ and $(-)$ of opposite values can be associated with the C_9-pseudocycle. The structure of the dimer shows a $(+, +)$ arrangement, which can be extrapolated to a *2.7*-helix, whereas the structure of the tetramer corresponds to a heterochiral sequence $(+, -, +, -)$ leading to a ribbon structure (Fig. 2.6A). In all cases, the secondary structures rely on consecutive C_9 H-bonded rings [75].

A related C_9-ribbon type structure has been postulated on the basis of NMR spectroscopy data in H_2O for γ-peptide oligomers of *cis*-γ-amino-L-proline (Fig. 2.6B) [76]. In the proposed secondary structure of **12**, the two amide bonds connected to α- and γ-positions of each proline residue are in the same plane, perpendicular to the average plane of the proline rings. The solid phase synthesis approach which authorizes facile incorporation of a variety of acyl and alkyl side chains at the γ-amino position is highly modular.

The carbofuranosyl group was used by Sharma and Kunvar as a conformation-stabilizing side chain to enforce the backbone of β- and γ-peptides into well-defined folded conformations (see also peptides **16** and **22** in Sections 2.4.2.2 and 2.4.3). A left-handed 9-helix with $1 \leftarrow 3$ H-bonding pattern (Fig. 2.6C) was identified from NMR studies in $CDCl_3$ solution and molecular dynamics calculations for a hybrid γ-hexapeptide (**13**) with alternating carbofuranosyl-γ^4-aminoacid and γ-aminobutyric acid (GABA or γ-Gly). It is worth mentioning that the corresponding β-peptides made of alternating $(3S)$-carbofuranosyl-β^3-amino acid and β-HGly residues adopt a mixed helical structure with (M)-helicity. (see Section 2.4.2.2).
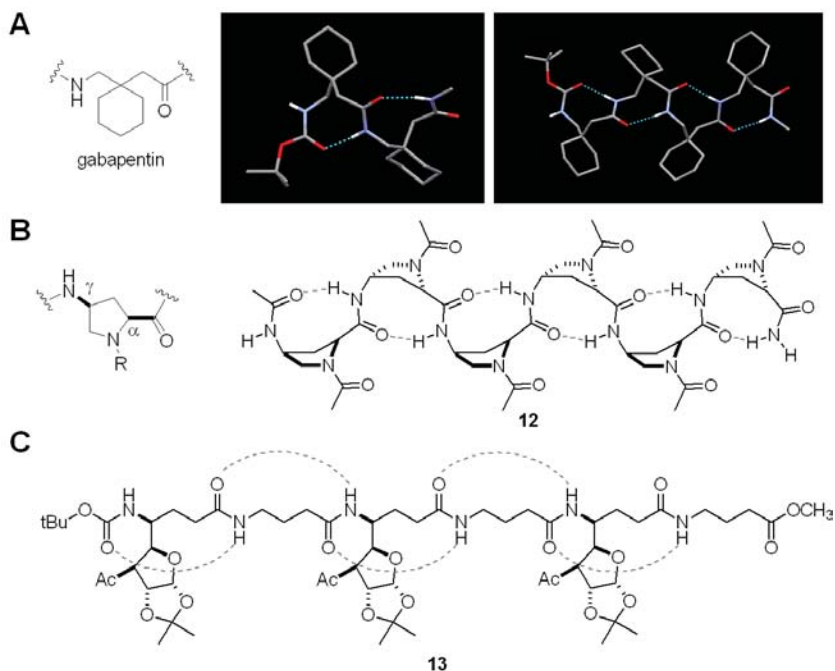
Fig. 2.6 C$_9$ H-bonded conformations of γ-peptides. (A) Gabapentin oligomers. Structure of dimer and tetramer in the solid state [75]. (B) Oligomers of *cis*-γ-amino-L-proline [76]. (C) γ-Hexapeptide **13** with alternating carbofuranosyl-γ^4-aminoacid and γ-Gly residues [111].

2.3.2.3 Stabilizing Local Backbone Conformation by Inverse-Bifurcation Involving an Additional Heteroatom

Oligo α- and β- and γ-aminoxyacids As already mentioned in Section 2.3.1.2 for N,N'-oligoureas (γ-peptide lineage), replacing carbon atoms in an ω-peptide backbone with heteroatoms represents a promising opportunity to design new foldamers. α-Aminoxypeptides formerly obtained by substitution of oxygen for the β-carbon atoms within the β-peptide backbone (Fig. 2.7A) have been investigated in depth by *ab initio* quantum-mechanical calculations and molecular dynamics simulation. This theoretical work predicted that a C$_8$ H-bonded ring (N–O turn, Fig. 2.7A) should stabilize a *1.8$_8$*-helix conformation in homochiral segments [112, 113] (Fig. 2.7B), at least in aprotic solvents [114]. Spectroscopic studies on a series of oligomers supported these predictions and established that such an helical conformation occurs in oligomers as short as a trimer [115]. The high stability of the helix was interpreted as the result of the replacement of the amide bonds by ''amidoxy'' bonds in which the strong withdrawing effect of the oxygen atom considerably enhances the acidity of the NHs and the strength of the H
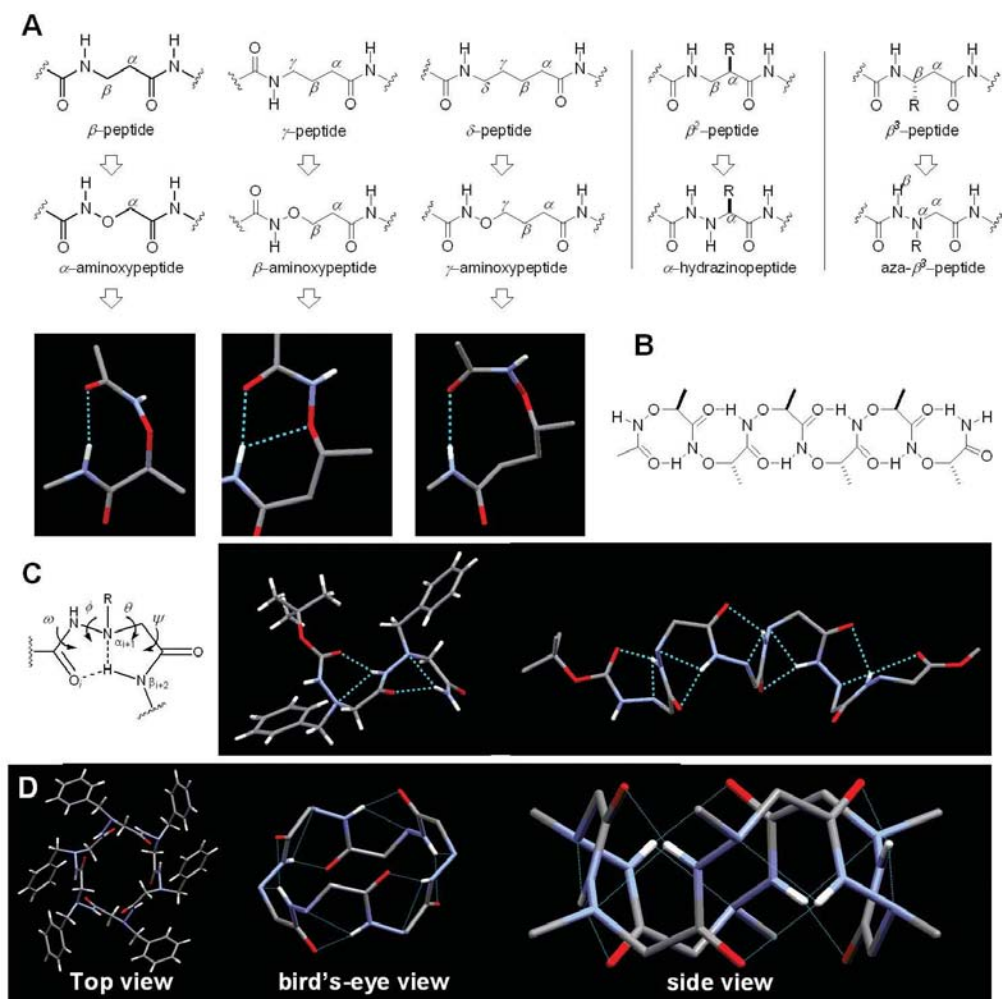
Fig. 2.7 Foldamers with propensities for 1 ← 3 H-bonded conformations by introduction of heteroatoms in the backbone of ω-peptides. (A) General formulae and comparison with related ω-peptide backbones. X-ray structures of N–O turns in α-, β-, γ-aminoxy peptides [112, 113, 117, 120]. (B) The C_8 H-bonded conformation (1.8$_8$-helix) of α-aminoxy peptides [113]. (C) The hydrazino turn and solid state conformation of aza-β^3-peptides (dimer and hexamer) [125]. (D) Solid state conformation of an aza-β^3-cyclohexapeptide showing the uninterrupted framework of bifidic C_8 pseudocycles. For clarity, side chains are omitted on bird's-eye view and side view [128].

bonds. In addition, the lone pair electron repulsion in the N–O segment reduces the flexibility of the backbone, which stabilizes the secondary structure, and promotes H-bonding between adjacent residues by selecting favorable dihedral angles. The crystal structure of a model trimer, in good agreement with the calculated geometry of the N–O turn, comfirm this analysis [116].

Small models of β-aminoxypeptides [117–119] and γ-aminoxypeptides [120] (Fig. 2.7A) were subsequently investigated by the group of Yang. FT-IR and NMR spectroscopy as well as X-ray diffraction studies also indicated a net preference for $1 \leftarrow 3$ H-bond interaction, leading respectively to the formation of C_9 and C_{10} pseudocycles (Fig. 2.7A). The C_9 ring is clearly an inverse bifurcated system where the NH_{i+2} is H-bonded to both O_{i+1} and $C=O_i$ ($d(N_{i+2}-O_{i+1}) = 2.5$ Å). It is not the case for the C_{10} pseudocycle (right) where the distance between O_{i+1} to NH_{i+2} is too long ($d(N_{i+2}-O_{i+1}) = 3.3$ Å).

Hydrazinopeptides Oligomers of hydrazinoacetic acid derivatives are aza^3-analogs of β^2-peptides (Fig. 2.7A). Secondary structure ensembles for these compounds have been examined at various level of *ab initio* MO theory by Günther and Hofmann [121]. The variety of H-bond networks in hydrazinopeptides is increased compared with β-peptides because of the additional $N^\alpha H$ centers which act as potential H-bond donors and H-bond acceptors. A C_8 based helical conformation (calculated *1.75$_8$*-helix), similar to that observed for α-aminoxy peptides, emerged in one mode of calculation. In all other cases the most favorable calculated secondary structure is a new 14-helix that topologically differs from the β-peptide 14-helix because of the participation of the lone pair of the sp^3 N^α atom in the stabilization of the 14-membered H-bonded ring. Although the synthesis of chiral hydrazinopeptides remains challenging, a series of hydrazinopeptides up to the hexamer have been prepared by Seebach and co-workers [122]. In MeOH, the hexamer displays a CD signature that resembles that of the corresponding 14-helical β^2-peptide. However, poor signal dispersion and fast exchange between NHs precluded detailed NMR studies and subsequent molecular modeling. These observations suggest dynamic interconversion between competitive conformations.

Aza-β^3-peptides Aza-β^3-peptides [123], the aza^3-analogs of β^3-peptides, are oligomers of N^α-substituted hydrazinoacetic acid (Fig. 2.7A, right). The main feature of this unnatural backbone is that nitrogen atoms bearing the side chains are sp^3 hybridized. As a result, aza-β^3-peptides are chiral molecules of undefined configuration. In CDCl$_3$ solution as well as in the solid state, their backbone is structured by a continuous set of C_8 pseudocycles (hydrazinoturn [124] or N–N turn) [125]. Examination of X-ray structures of a dimer and a hexamer (Fig. 2.7C) reveals that the former corresponds to a homochiral sequence (both nitrogen centers have the same absolute configuration) that defines an incipient extended *1.75$_8$*-helix whereas the latter is characterized by an heterochiral arrangement (R,R,S,S,R,R chiral sequence) that drives a more folded conformation. In solution, aza-β^3-peptides equilibrate between all possible chiral sequences, all of

which share the same C_8 based H-bond network. The structural resemblance between α-aminoxypeptide and aza-β^3-peptide backbones which both rely on the formation of a C_8 H-bonded network is striking. In both cases, the presence of the two adjacent heteroatoms is the driving force of the folding process. In aza-β^3-peptide C_8 pseudocycles, the ${}^\alpha N_{i+1} \cdots H - {}^\beta N_{i+2}$ distance (~ 2.25 Å) is typical of a H bond but the corresponding angle ($\sim 110°$) is distorted relative to the expected value for a standard H bond. The same holds true for the N–O turn $(d({}^\alpha O_{i+1} \cdots H - {}^\beta N_{i+2}))$. This interaction has been shown to play a crucial role in the stabilization of the secondary structure [126]. The C_8 pseudocycle of aza-β^3-peptide and probably the N–O turn can thus be described as inverse bifurcated systems combining interactions between nearest and non-nearest neighbors [127]. Conformational analyses of aminoxy- and hydrazino-peptides, and also studies of N,N'-linked oligoureas described in Section 2.3.1.2 illustrate that the amide linkage is not the only polar bond that can sustain a robust intramolecular H-bond network for the design of self-organizing oligomers.

It is noteworthy that the intrinsic features and conformational preferences of linear aza-β^3-peptides can be exploited to generate macrocyclized derivatives with remarkable efficiency (for example of self-templated macrocycles, see Chapter 1). This result can be rationalized by invoking a dynamic process, i.e. the ability of linear precursors to populate heterochiral sequences ideally pre-organized for cyclization. The resulting macrocycles retain the alternated chiral sequence (the pyramidal inversion of the chiral nitrogen atoms is now considerably slowed down) as well as the basic C_8 structure (C_3-symmetric hexamer, Fig. 2.7D) [128]. The solid state intramolecular organization of the macrocycles perfectly reflects the major conformation present in solution ($CDCl_3$).

To conclude this section, it is interesting to point out that, in contrast to the natural α-peptides, $1 \leftarrow 3$ H-bond interaction patterns consistently sustain secondary structures in different classes of designed ω-peptides and their related analogs. A survey of the present literature also emphasizes how much folding can be oriented through slight structural modulations, such as stereochemical modifications, side chain tuning, and backbone alteration, which all subtly modulate the set of weak intramolecular interactions and, in the end, the shape of the molecules.

2.4
Oligoamide Mixed Helices

In contrast to common periodic peptide helices in which components along the helix axis of backbone C=O bond vectors point in the same direction, a unique feature of mixed helices is the alternating direction of the C=O bonds along the sequence that results in the formation of a β-sheet-type H-bonding pattern. Mixed helices are thus characterized by a small macrodipole which implies that they are energetically disfavored in a polar environment compared with common periodic helices that have their H bonds oriented in the same direction.

2.4.1
The α-Oligopeptide Precedent: from Antibiotic Gramicidin A to Poly-Gln Aggregates in Huntington's Disease

Mixed helices, also termed β-helices by virtue of this similarity to β-sheet type structures, were proposed in the early 1970s independently by Ramachandran and Chandrasekharan [129], and Urry [130] for α-oligopeptides with a regularly alternating sequence of D and L residues (D,L-peptides) such as the linear penta-decapeptide antibiotic gramicidin A (HCO-Val-Gly-Ala-D-Leu-Ala-D-Val-Val-D-Val-Trp-D-Leu-Trp-D-Leu-Trp-D-Leu-Trp-NH-(CH₂)-OH), i.e. by repeating a pair of (ϕ, ψ) angles with values lying in the allowed β-sheet region of the Ramachandran map for L and D residues, respectively. The H bonds are formed alternatively in forward and backward directions along the sequences thus closing intertwined pseudo rings of different sizes (Fig. 2.8). β-Helices formed by D,L-peptides are characterized by their structural polymorphism [131–135]. They can be either right- or left-handed depending on parameters such as sequence, length and environment. Various periodicity and hence pore size have been predicted and found experimentally. Finally, β-helices can exist either as single helices stabilized by parallel H bonds or as parallel (or antiparallel) double helices. Gramicidin A, for example, depending on its surrounding environment has been found to form either single stranded or double stranded β-helical conformations with residues per turn ranging from 5.6 to 7.2 [136–139].

Proteins also fold into right- and left-handed β-type helical structures with larger periodicities [140–142]. The first parallel β-helix fold (right-handed) was discovered by Yoder and co-workers in the structure of Pectate Lyase C [140]. The repeating unit (i.e. β-helical turn) in such β-helices is formed by three strand segments connected by three loops and generally comprises 15–23 residues. A large "loopless" β-helical nanotube with 20 residues per helical turn has been proposed by Perutz and co-workers to account for the X-ray diffraction patterns of fibers of a poly-L-Gln peptide (exon-1 peptide of huntingtin) [143]. In addition to being relevant to Huntington's disease, this model may also serve to interpret structures of amyloid fibrils associated with various neurodegenerative diseases. It is worth noting that this structure which features repetition of a pair of (ϕ, ψ)
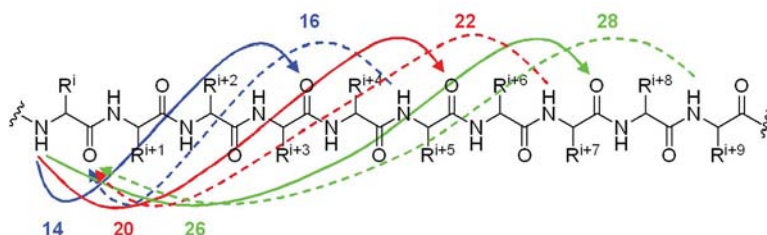
Fig. 2.8 H-bond patterns associated with mixed helical secondary structures in α-polypeptides.

angles with nearly opposite values ($\phi_1 = 160°$, $\psi_1 = -170°$; $\phi_2 = -161°$, $\psi_2 = 178°$); is remarkably close to a D,L-peptide β-helical fold.

2.4.2
Introducing Periodicity at the Level of a Dimer Unit in β-Peptides leads to a Remarkably Stable Mixed Helical Fold

Early studies by Seebach and co-workers established that, by analogy to β-helical α-peptides, β-peptides with an alternating substitution pattern adopt a mixed helical conformation [55, 144, 145]. The resulting structure which consists of 10- and 12-membered H-bonded pseudo rings alternating in forwards and backwards direction, respectively has only a small resulting macrodipole. This particularly robust mixed helical fold was subsequently observed for other β-peptides (Fig. 2.9) exhibiting periodicity at the level of a dimeric unit [146–148].

2.4.2.1 By Mixing β^2- and β^3-Amino Acids
The finding that the conformational preferences of "mixed" β-peptides composed of alternating β^3- and β^2-amino acid residues differed from that of the corresponding homopolymers consisting exclusively of β^3- or β^2-amino acid residues (i.e. the 3_{14} helix) was largely unexpected [144]. In MeOH, β-peptides with (S)-

14 PG = H, R = H
15 PG = Boc, R = Bn

16   **17**   **18**

R = Me

19   **20**   **21**
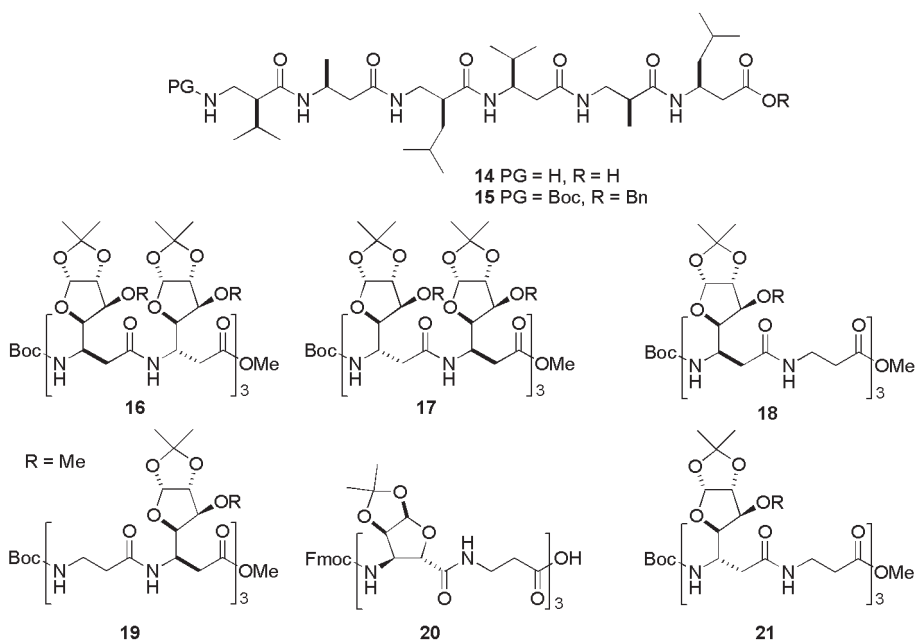
Fig. 2.9 Mixed *12/10*- (*10/12*-) helix forming β-peptides [144–148].

β^2/β^3 (or (S)-β^3/β^2) dipeptide repeats such as **14** and **15** did not display the CD-pattern characteristic of the 3_{14} helix. Their CD spectra showed an intense maximum near 205 nm and no zero-crossing. The dispersion of the chemical shifts as well as the large $^3J(H\text{–}C(\alpha), H\text{–}C(\beta))$ values (>10 Hz) in their NMR spectra recorded in pyridine and CD_3OH indicated that at least one stable secondary structure was populated. ROESY experiments revealed a NOE pattern substantially different from that of the 3_{14} helix with no $i/i+3$ NOE crosspeaks and new $i/i+2$ connectivities not compatible with the 3_{14} helix. Restrained MD calculations based on NOEs and J values yielded a unique mixed helical structure with alternating intertwined 12- and 10-membered H-bonded rings. The helical screw sense was opposite to that of the related 3_{14} helix and the overall macrodipole was strongly reduced because of alternating orientations of backbone carbonyl groups. The *12/10*-structure of a low energy conformer of β^2/β^3-peptide **14** [55, 144] and detailed representations of the 12- and 10-membered pseudocycles are shown in Fig. 2.10. Comparison of the *10/12*-helix turns with the corresponding 14-membered ring of the 3_{14}-helix reveals a common (+)-synclinal arrangement ($\theta_1 \sim 60°$) around the central $C(\alpha)\text{–}C(\beta)$ bond for each amino acid constituents. However, whereas both ϕ and ψ angles are negative for β-amino acids in the regular 3_{14}-helix, β^2-amino acid residues have a positive ϕ value and β^3-amino acid residues a positive ψ value in the *12/10*-helix.

The tendency is that in the absence of any adjacent substituent on the two sides of an amide bond, the 12-membered turn is favored, the 10-membered being formed when the amide bond is flanked by substituted carbons (Fig. 2.10). In fully protected **15**, the NH of residue 1 is engaged in the formation of an addi-
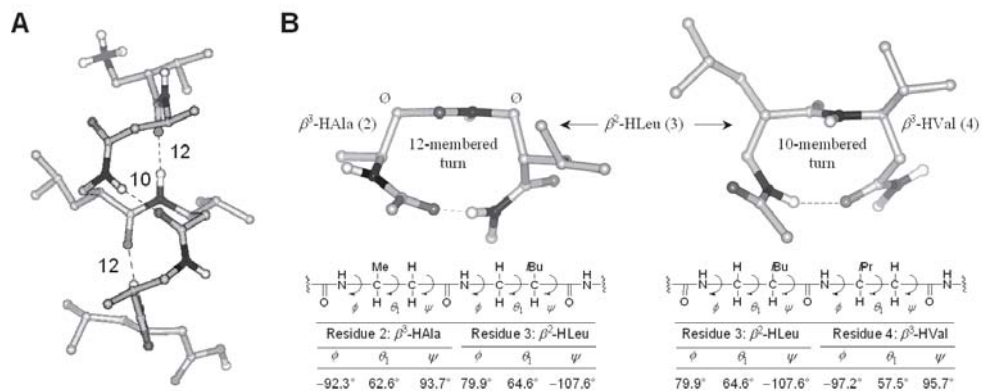
Fig. 2.10 The mixed 12/10 helical structure of β^2/β^3-dipeptide repeats. (A) View along the helix axis of a low energy conformer resulting from NOE-restrained modeling of **14** in pyridine [144, 145]. (B) Comparison of the 12- and 10-membered turns found in the 12/10 helix of **14** together with corresponding backbone dihedral angles.

tional N-terminal 10-membered turn and the pattern of 10- and 12-membered turns is reversed. The strong stabilizing effect associated with N-terminus capping in mixed β-peptides was confirmed by CD studies in MeOH, the ellipticity value at 205 nm for **15** in MeOH being twice that of **14**. In the case of longer β^2/β^3-peptides, the effect of removing the terminal protecting groups was even more pronounced with collapse of the band at 205 nm and restoration of a CD pattern, albeit weak, characteristic of the 3_{14} helical conformation [145]. This observation may suggest the presence of equilibrating conformers and can be explained in term of unfavorable charge–pole interactions in the right-handed *12/ 10*-arrangement of unprotected β^2/β^3-peptides, the positively charged amino terminus being rather a promoter of left-handed 3_{14}-helical structure. Information about the dynamic of (un)folding process in β^2/β^3-(β^3/β^2-) dipeptide repeats (protected or unprotected form) was gained by exploration of conformational ensembles produced by molecular dynamics (MD) simulations in explicit solvent using the GROMOS96 force-field. The results demonstrated reversible folding to the 12/10 helix and were consistent with experimental data (see also Chapter 6). [103, 104, 149–151]. Although alternate conformations such as the pure left-handed 3_{14}-helix and various partially folded conformations were also populated, the right-handed *12/10* helix was the predominant conformation in the simulation of β^2/β^3-(β^3/β^2-) dipeptide repeats in MeOH.

2.4.2.2 Additional Substitution Patterns Stabilizing the Mixed 10/12- (12/10-) Helix

Theoretical studies at various levels of *ab-initio* MO theory (HF/6-31G* and B3LYP/ 6-31G*) provided further insight into the relative preference of β-peptides for the 3_{14} and mixed *10/12*-helices. Quantum mechanics calculations performed on unsubstituted β-peptides (oligo-β-HGly peptides) revealed that the formation of the *10/12*-helix is intrinsically favored over the 3_{14} helix (by 21.4 and 4.8 kcal mol^{-1} in the gas phase and methanol solution for a protected hexapeptide) [152, 153]. Analysis of the influence of substitution patterns [151, 154] confirmed the intrinsic preference of (S)-β^2/β^3 dipeptide repeats for the right-handed 10/12-helical conformation and suggested other patterns of substitutions compatible with the formation of the 10/12-helix such as heterochiral dipeptide repeats (e.g. (R)-$\beta^3/$ (S)-β^3, (R)-$\beta^3/$(S)-β^2, (S)-$\beta^2/$(R)-β^2, (S)-$\beta^2/$(R)-β^3). This prediction was experimentally confirmed by the conformational analysis in solution of carbofuranosyl-β^3-hexapeptides **16** and **17** made of regularly alternating (3R)- and (3S)-building blocks (Fig. 2.9) [147].

In MeOH, both peptides displayed a CD pattern characteristic of a right-handed 10/12- (12/10-) helical structure. More detailed structural analysis by NMR study in CDCl$_3$ revealed that **16** and **17** adopt well-defined (*P*)-12/10 and 10/12 helical structures, respectively (Fig. 2.11).

It is noteworthy that the mixed β-peptide helical backbone can tolerate suppression of side chains every two residues. Thus, β-peptides made of alternating (3R)-carbofuranosyl-β^3-amino acid and β-HGly (3-amino propionic acid) residues (**18** and **19**) and of (2S,3S)-$\beta^{2,3}$-(sugar) amino acid/β-HGly repeats (**20**), respectively
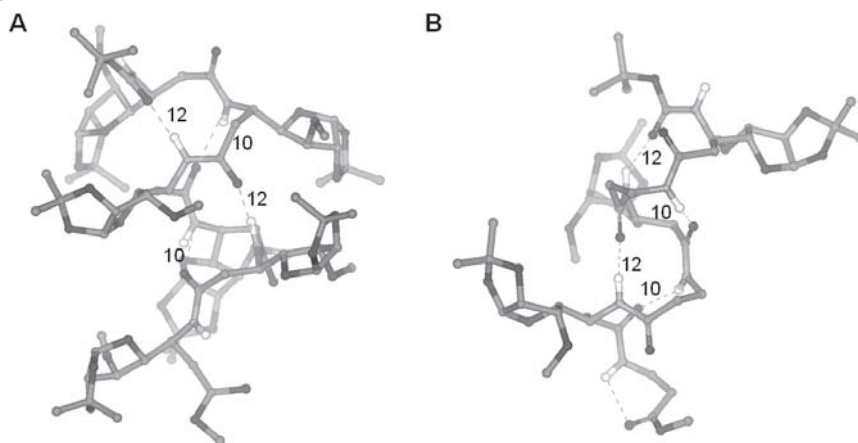
Fig. 2.11 Additional substitution patterns leading to mixed helices in β-peptides. Views along the helix axis of low energy conformers of (A) **16** and (B) **18** obtained by NOE-restrained modeling using NMR data obtained in CDCl$_3$ [147, 148].

have been found to adopt well-defined right-handed (P)-12/10 (**18** and **20**) and (P)-10/12 (**19**) helical structures in nonpolar solvents such as CDCl$_3$ and CD$_3$CN [146, 148]. β-HGly thus behaves like β^2-amino acid residues in β^3/β^2 repeats. This finding is not surprising if one keeps in mind that C(3) substitution is much more effective than C(2) in reducing the flexibility of the β-peptide backbone as already discussed in Section 2.3.1.1. Inverting the configuration of β^3-amino acid residues in β^3/β-HGly repeats (**18** → **21**) caused a switch in helix handedness facilitated by the conformational freedom of β-HGly residues [148].

2.4.3
Extending the Concept of Mixed Helices

The question whether mixed helices might exist in other families of ω-oligoamide backbones has been addressed by the group of Hofmann with methods of *ab initio* theory [153, 154]. Mixed helices were systematically searched within the conformational space of γ- and δ-peptides and their energies compared with those of corresponding periodic helices. The results indicate that (i) while mixed γ-peptide helices (14/12 and 24/22 H-bond pattern) are less stable than the experimentally found periodic 2.6$_{14}$ helix, mixed δ-peptide helices (14/16 and 16/14 H-bond pattern) are significantly more stable according to the density functional theory (DFT) and (ii) in contrast to the situation observed for β-peptides, these mixed helices become more unstable than their periodic alternatives in polar environment such as water (polarizable continuum model, PCM/HF/6-31G*). These
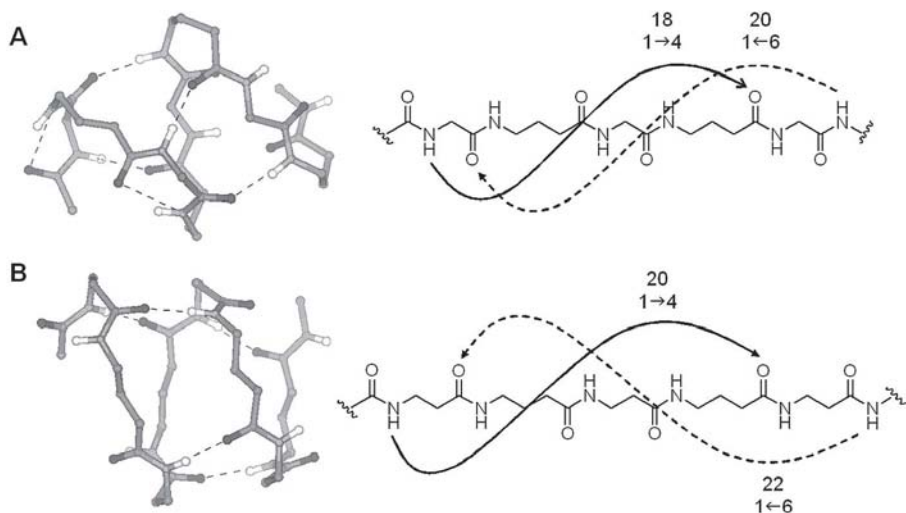
Fig. 2.12 Most stable helices in hybrid (octa)peptides at the HF/6-31G* level of *ab initio* MO theory. (A) The mixed 18/20-helical fold of α,γ-peptides, (B) the mixed 20/22-helical fold of β,γ-peptides [156].

mixed helical structures have not yet been observed experimentally in γ- and δ-peptides synthesized so far.

Another general approach to introduce periodicity at the level of a dimeric unit in oligoamides is to mix two types of ω-amino acid residues in an alternating fashion. Such so-called heterogeneous or hybrid peptide backbones show considerable promise to expand the pool of candidate foldamers, and their experimental folding patterns are discussed with greater details in Section 2.6. Employing the methods of *ab initio* MO theory, Hofmann and co-workers have investigated the ensembles of helical structures attainable with (unsubstituted) α,β-, α,γ- and β,γ-hybrid backbones [155, 156]. Conformational analysis provided three groups of helical structures according to their global H-bonding pattern: helices with all H bonds in one direction (either forward or backwards direction), and mixed helices. For all three heterogeneous backbones, the most stable conformations at the HF and DFT level of *ab initio* theory were found among mixed helices (i.e. 18/16- (and 11/9-), 18/20- and 20/22-helical folds for α,β-, α,γ- and β,γ-peptide hybrids, respectively, see Fig. 2.12). In an aqueous environment (PCM/HF/6-31G* calculations), the stability of helices with unidirectional H-bonding increases at the expense of mixed helices. It is worth mentioning that some mixed helices such as the 11/9-helix of α,β-peptides remain significantly stable. Experimental evidence for the 11/9-helical fold came from NMR studies in CDCl$_3$ of a series of α,β-peptides consisting of L-Ala/(3S)-carbofuranosyl-β^3-amino acid repeat (e.g. **22**, See Fig. 2.13) [157].
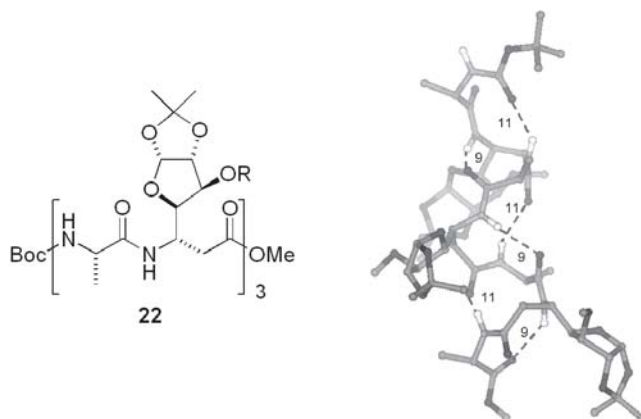
Fig. 2.13 The 9/11-Mixed helical structure of α/β hybrid peptides consisting of L-Ala/ (3S)-carbofuranosyl-β^3-amino acid repeats (e.g. **22**). View along the helix axis of a low energy conformer of **22** generated by NOE-restrained modeling using NMR data obtained in CDCl$_3$ [157].

2.5
Nonperiodic Structures: Open Chain β-Turn-like Motifs and Hairpins in Designed Homo-oligomers

Turn structures and β-hairpins (two H-bonded antiparallel β-strands connected by a turn segment) which cause the peptide chain to reverse its direction are key functional and structural elements of proteins. Significant progress has been made towards our understanding of the principles that govern β-hairpin nucleation (for reviews, see refs [8–10] and Chapter 5). Because these motifs are important for peptide and protein recognition, the design, synthesis and applications of hairpin turn mimetics have attracted considerable attention. In the foldamer area, the *de novo* design of hairpin-type structures from unnatural linear oligomers remains challenging, in part because of the duality of structural elements (turn + sheet) and the absence of periodicity in the motif. So far, homo-oligomers (based on one residue type) programmed to adopt hairpin conformations have been essentially limited to β-peptides. Alternatively, the individual components of hairpin design (i.e. either turn segments or pleated sheet structures) generated from various unnatural oligomeric backbones (e.g. β- and γ-peptides, α-aminoxy acid oligomers) can be assembled with α-peptide strands or β-turn segments to generate hybrid hairpin structures (see Section 2.6.2).

2.5.1
Sheet-forming ω-peptides

Fully extended β-peptide strands can be generated by populating antiperiplanar conformations around the C(α)–C(β) bond (θ_1 values close to 180°). Seebach and

Gellman groups found that this conformational bias can be introduced by using acyclic $\beta^{2,3}$-amino acids of *unlike* configuration bearing alkyl substituents [158, 159]. The resulting $\beta^{2,3}$-peptide chains adopt extended conformation with formation of pleated sheets. This is in contrast to all-*like*-$\beta^{2,3}$-peptides which have be shown to form predominantly 3_{14} helical structures (See Section 2.3.1). Extended conformations for sheet formation are also accessible to γ-peptides subject to sufficient backbone pre-organization (α,β-unsaturation [160], cyclization between C(α) and C(γ) [161]) to restrict rotation around ethylene bond.

2.5.2
Turn Segment for Hairpin Formation

Two types of β-dipeptide turn segments have been used to connect antiparallel β-peptide strands and nucleate β-peptide hairpin conformations.

The first one imagined by Gellman and his group is a 12-membered H-bonded turn generated by a heterochiral dinipecotic acid (Nip, the β^2-homolog of proline: β^2-HPro) sequence [162, 163]. X-ray crystal structure of a designed β-tetrapeptide (**23**, Fig. 2.14A) show the expected antiparallel hairpin conformation, all strand residues displaying antiperiplanar arrangement around the C(α)–C(β) bond [162]. The second approach developed by Seebach and co-workers (**24**, Fig. 2.14B) exploits the propensity of mixed β^2/β^3-dipeptides to populate 10-membered H-bonded turn (see Section 2.4.2.1 and Fig. 2.10) [159, 164]. Detailed NMR spectroscopy and MD simulation analysis of β-hexapeptide **24** in CD$_3$OH revealed significant (an estimation of 20–30% was given by MD) hairpin population. There was however no evidence for secondary structure in water. The structural similarity between the 10-membered turn segment in **24** and a type II′ α-peptide β-turn is illustrated in Fig. 2.14E. Interestingly, the formation of β-peptide 10-membered H-bonded turns was also found to be strongly promoted by $\beta^{2,2}$-geminally substituted amino acids units such as achiral 1-aminomethylcyclohexanecarboxylic acid [51]. The hairpin structures in both **23** and **24** are characterized by the unidirectionality of the C=O and NH bonds within each strand segment. However, as a consequence of their different turn geometry: a 12-membered turn closed by H-bonds between C=O$_i$ and NH$_{i+3}$, and a 10-membered turn closed by H-bonds between NH$_i$ and C=O$_{i+1}$, antiparallel hairpins formed by β-peptides **23** and **24** display opposite sheet polarities. Backbone torsion angle values (X-ray and NMR) for selected β-amino acids residues within extended strand segments of **23** and **24** are close to ideal values for β-peptide pleated sheets: $\phi = -120°$ (or $120°$), $\theta_1 = 180°$, $\psi = 120°$ (or $-120°$).

Optimal pre-organization of the γ-peptide backbone towards the formation of open-chain turn-like motifs was found to be promoted by *unlike*-$\gamma^{2,4}$-amino acid residues [165–166]. This design principle can be rationalized by examination of the two conformers free of *syn*-pentane interaction (I and II, Fig. 2.14C). $\gamma^{2,4}$-peptides built from either homochiral or heterochiral *unlike*-$\gamma^{2,4}$-amino acid units adopt reverse turn-like structure stabilized by 14-membered H bond (e.g. **25**, Fig. 2.14D). γ-Peptide turn conformations compare well with the type II′ β-hairpin
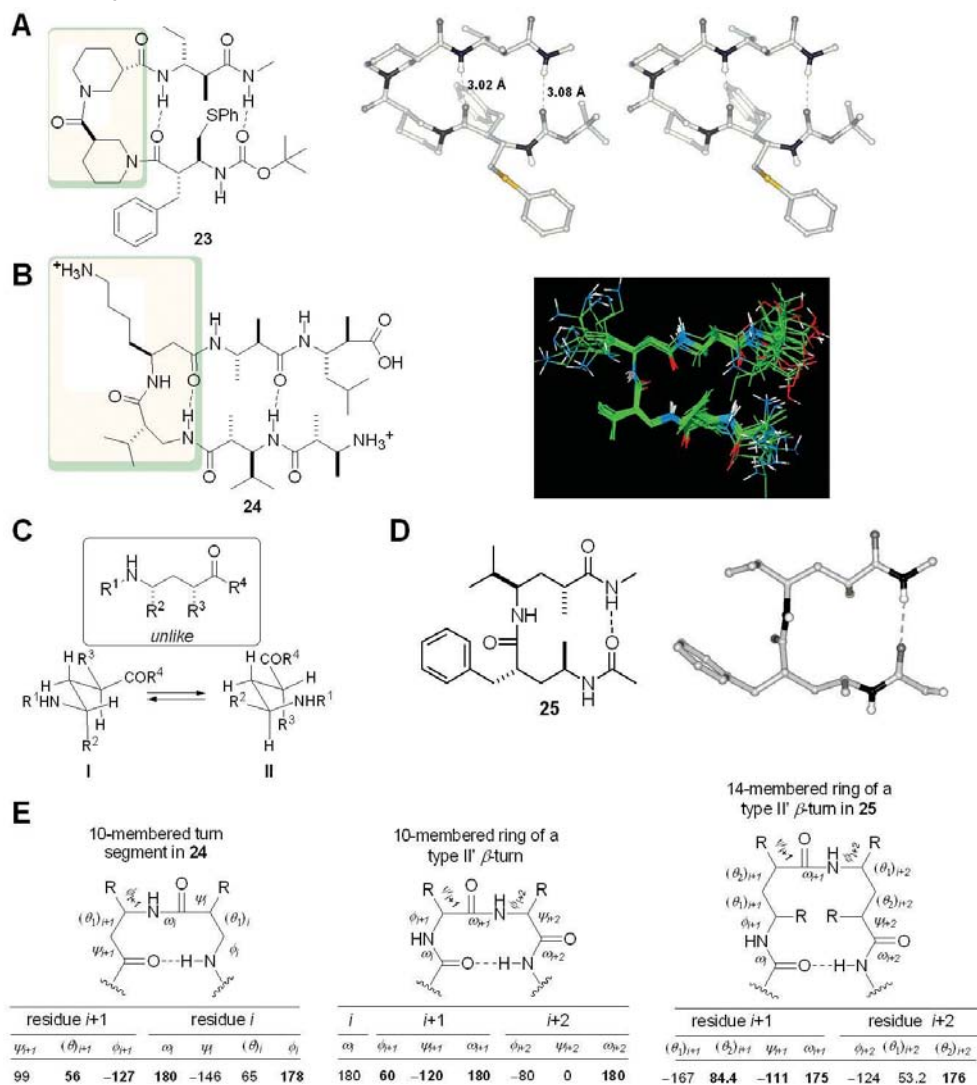
Fig. 2.14 Nonperiodic structures formed by β- and γ-peptides. β-peptide antiparallel hairpin structures with (A) a 12-membered R/S dinipecotic turn segment (e.g. **23**, X-ray structure) [162] and (B) with the 10-membered turn formed by mixed β²/β³-dipeptide sequence (e.g. **24**, structure derived from NMR data in CD₃OH) [164]. (C) The two conformation free of destabilizing *syn*-pentane interaction in γ²,⁴-amino acid with unlike configuration. (D) 14-membered H-bonded turn induced by γ-peptides consisting of unlike-γ²,⁴-amino acid residues (heterochiral sequence) (e.g. **25**, X-ray structure) [166]. (E) Comparison of the turn segment found in β- and γ-peptide **24** and **25** with a naturally occuring type II′ β-turn of α-polypeptides together with backbone dihedral angles in (°). Torsion angles with comparable values are shown in bold [164, 166].

motif of naturally occurring α-peptides, thus suggesting that short-chain γ-peptides with the right substitution pattern could be used as β-turn mimetics in drug design.

2.6
Expanding Structural Diversity with Heterogeneous Backbones

Most structural motifs encountered in the previous sections of this chapter have been generated from oligomeric strands formed by the repeat of monomeric units from a single class. While the design of appropriate building blocks for the construction of homogeneous folding oligomers based on remote intrastrand interactions remains an active area, the possibility to combine these elements and generate new heterogeneous backbones as candidate foldamers has recently emerged. This concept, successfully applied to oligoamides, spectacularly expands the structure space attainable with a relatively small pool of residue types.

2.6.1
From Discrete ω-Amino Acid Guests in α-Helices to Helical α,ω- and β,γ-Peptide Hybrids

Multiple approaches, either experimental [167–172] or theoretical [155, 156], have been considered to delineate the ensemble of helical conformations formed by new hybrid backbones composed of two ω-amino acid residue types. Oligomers with periodicity at the level of dimer (e.g. α,β-, α,γ-, β,γ-peptides) or trimer (e.g. α,α,β- or α,β,β-peptides) units have been studied.

For several years, Balaram and co-workers have been investigated the structural features of β- and γ-amino acid residues (β-HGly, γ-Gly (4-aminobutyric acid), Gpn (gabapentin)) when inserted as guests into host α-peptides [173]. Examination of X-ray crystal structures of peptides **26–28** [172, 174, 175] (see Table 2.1

Table 2.1 Residue patterns (in bold) within peptides **26–28** whose coordinates served to construct models of hybrid helices.

	Sequence	Pattern	Number of atoms in H-bonded ring	Ref
26	Boc-Val-Ala-Phe-Aib-β³-HVal-**β³-HPhe-Aib**-Val-Ala-Phe-Aib-OMe	βα	11	175
	Boc-Val-Ala-Phe-**Aib-β³-HVal-β³-HPhe**-Aib-Val-Ala-Phe-Aib-OMe	αββ	15	
27	Boc-Leu-Phe-**Val-Aib-β³-HPhe**-Leu-Phe-Val-OMe	ααβ	14	172
28	Boc-Leu-Aib-Val-β-HGly-**γ-Gly-Leu**-Aib-Val-Ala-Leu-Aib-OMe	γα	12	174
	Boc-Leu-Aib-Val-**β-HGly-γ-Gly**-Leu-Aib-Val-Ala-Leu-Aib-OMe	βγ	13	

for sequences) revealed that β- and γ-amino acid residues can be substituted for their α-amino acid counterpart at discrete position in α-helical structures without major perturbation of the overall fold.

Experimentally defined backbone torsion angles of expanded $1 \leftarrow 4$ and $1 \leftarrow 5$ H-bonded units formed by hybrid segments in these peptides (marked in bold in Table 2.1) served as starting points to generate a series of energetically favorable models of α,ω-hybrid helices. For example, the 11- and 13-membered H-bonded units encompassing the β^3-HPhe-Aib (β, α) and the β-HGly-γ-Gly (β, γ) segments in X-ray structure of **26** [175] and **28** [174] provided appropriate geometrical parameters to model 11-β,α and 13-β,γ-helices. It is worth noting that a β,γ-dipeptide repeat is isostere to a α-tripeptide and that the 13-helical backbone proposed for β,γ-peptides ($\phi = -106°$, $\theta = 75°$, $\psi = -115°$ (β^3-amino acid) and $\phi = -117°$, $\theta_1 = 66°$, $\theta_2 = 62°$, $\psi = -120°$ (γ-amino acid)), and also identified by quantum mechanics calculations [156] superimposes well to the α-helical backbone (RMSD value of 0.7 Å). This interesting theoretical consideration will need to be verified experimentally.

Parallel to these semi-empirical studies, Reiser, Zerbe and colleagues [167] and Gellman [168–171] independently reported experimental evidence for periodic helix formation in short chain hybrid peptides consisting of alternating α- and β-amino acid residues (mixed helices reported by Sharma and Kunwar [157] have already been discussed in Section 2.4.3). In both cases, backbone pre-organization to enforce folding propensity was introduced at the β-amino acid positions with 2-aminocycloalkanecarboxylic acid residues. α,β-Peptides designed by Gellman and co-workers contained β-amino acid residues constrained with five-membered rings (S,S)-*trans*-ACPC and (S,S)-*trans*-APC (e.g. **29** and **30**, Fig. 2.15).

Detailed NMR analysis of hybrid peptide **29** in CD$_3$OH revealed a complex pattern of $(i, i + 2)$ and $(i, i + 3)$ inter-residue NOE connectivities that could be explained by assuming rapid interconversion between two helical conformations with H bonds in backward direction, namely a 11-helix ($1 \leftarrow 4$ H bonds) and a 14/15-helix ($1 \leftarrow 5$ H bonds) [168]. Although less plausible, a helical conformation with three-center backbone H bonds (both 11- and 14/15 H-bonded rings) was also proposed on the basis of the nonsequential NOEs observed. Factors governing helical folding propensity and helix type repartition in such α,β-peptide hybrids have been delineated. Increase in chain length seems to favor the 14/15 helical shape relative to the 11-helix [168]. The analogy to α-peptides which also show chain length-dependent conformational transition between 3_{10} helix ($1 \leftarrow 4$ H bonds) and α-helix ($1 \leftarrow 5$ H-bonding pattern) is striking (Fig. 2.16).

α,α-Disubstitution of α-amino acids which is well known to promote helical folding among α-peptides [15] was shown to reinforce helix stability and to favor 11-helical folding of α,β-peptide hybrids [170]. Octamer **30** which consists of ACPC/Aib repeats adopts a perfect 11-helical fold in the solid state with all possible H-bonded rings present (Fig. 2.17A). In contrast, β substitution of α-aminoacid residues (e.g. Val, Ile, Thr) or substitution of acyclic β-amino acids for ACPC/APC residues is helix destabilizing [170]. Pre-organization with six-membered ring (S,S)-*trans*-ACHC residues does not support helix formation in

29

30

31

32

33

34
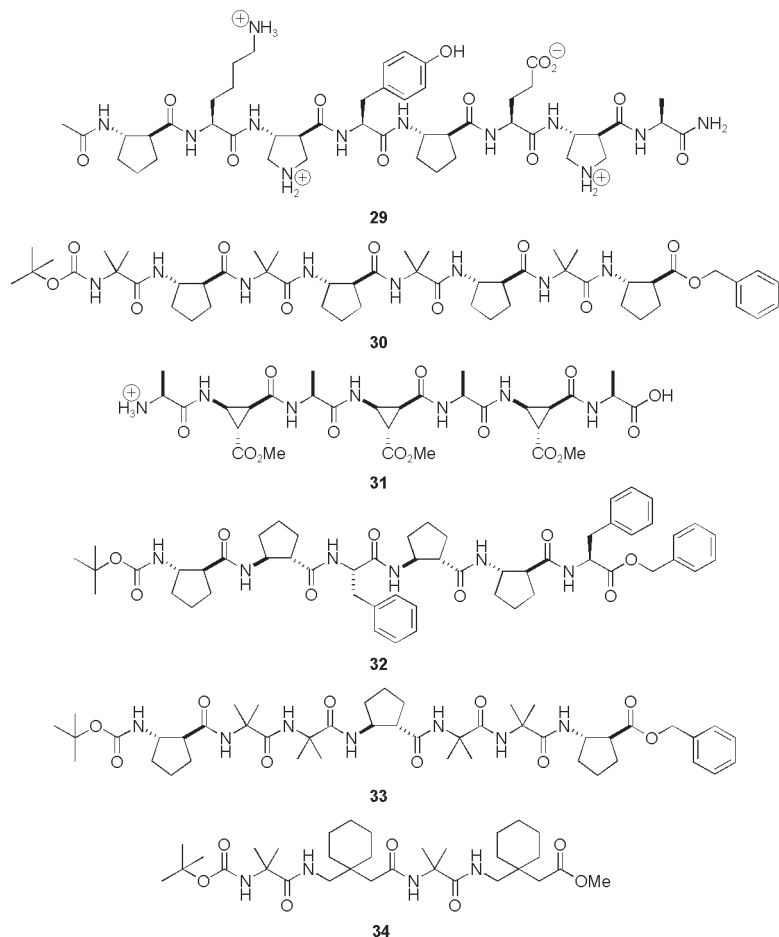
Fig. 2.15 Helix forming α,ω-peptide hybrids [167–172].
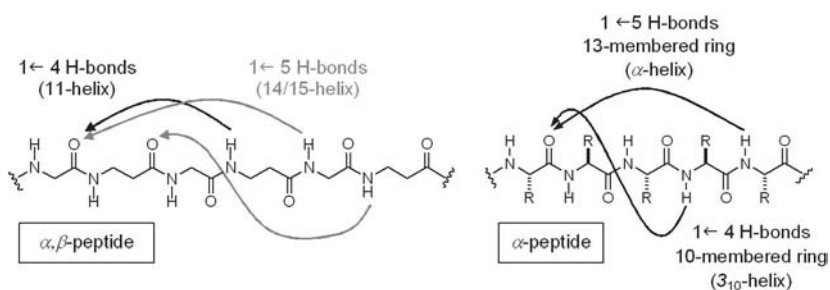
Fig. 2.16 1 ← 4 and 1 ← 5 H-bond patterns of interconverting 11- and 14/15-helices in heterogeneous α,β-peptides such as **29** and analogy with H-bonding scheme of 3_{10} and α-helical secondary structures.
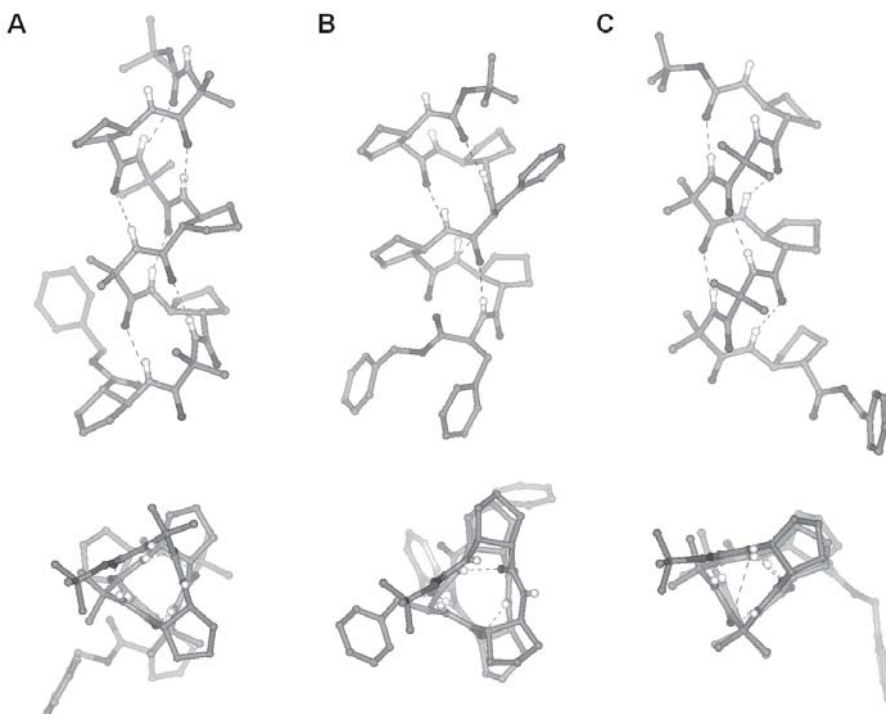
Fig. 2.17 Helical secondary structures of peptide hybrids in the solid state (views along the helix axis and top views). (A) Right handed 11-helix formed by α,β-peptide **30** (adapted from [170]). The backbone torsion angles extracted for central residues 4 and 5 ($\phi = -99.7°$, $\theta = 89.6°$, $\psi = -80.7°$ (β-amino acid) and $\phi = -50.3°$, $\psi = -42.9°$ (α-amino acid)) and those derived from the computer-generated model proposed by Balaram ($\phi = -105°$, $\theta = 80°$, $\psi = -73°$ (β-amino acid) and $\phi = -62°$, $\psi = -44°$ (α-amino acid) [172]) are in good agreement. (B) Right handed helical conformation ($1 \leftarrow 4$ H-bonding pattern) of α,β,β-peptide **32** (adapted from [171]). (C) Right handed helical conformation ($1 \leftarrow 4$ H-bonding pattern) of α,α,β-peptide **33** (adapted from [171]).

α/β-peptide hybrids, probably because homogenous ACHC backbones favor H-bonding in the forward direction (see Section 2.3.1.2) [168].

In a related work [167], Reiser, Zerbe and colleagues used *cis-β*-aminocyclopropanecarboxylic acids (*cis-β*-ACCs) substituted on the 3-position of the ring and investigated oligomers consisting of L-Ala/*cis-β*-ACC repeats (exemplified by heptamer **31**). NMR studies in CD_3OH and molecular modeling calculations led to the identification of a third helical fold (with $1 \leftarrow 3$ H-bonds) for α,β-peptide hybrids.

Introducing periodicity at the level of a trimer unit with β,β,α- and β,α,α-triad repeats (hexamer **32** and heptamer **33**) successfully led to the identification of new helical secondary structures (Fig. 2.17B and C) [171]. While X-ray diffraction

studies led to the characterization of helices with $1 \leftarrow 4$ H bonds, evidence for rapid interconversion between two helical conformations ($1 \leftarrow 4$ and $1 \leftarrow 5$ H-bonding patterns) was gained from NMR studies in CD_3OH.

Experimental evidence that short chain α,γ-peptide hybrids also adopt helical secondary structures came from X-ray diffraction studies of tetramer **34** consisting of Aib-Gpn repeats [172]. The observed helical fold is stabilized by 12 H bonds in the backwards direction. By fixing both ethane bonds in a synclinal conformation (values for both θ_1 and θ_2 are close to 60°), β,β-disusbtitution (in Gpn) ensures conformational space restriction of γ-amino acid residues.

2.6.2
Hairpins from α,ω-Peptide Hybrids

New hairpin type motifs have been generated by combining turn and strand segments made of distinct backbones (Fig. 2.18).

D-Pro-Xaa and Asn-Gly dipeptide sequences are known to support the formation of type I′ or II′ β-turns and to promote nucleation of α-peptide hairpin structures [176–178]. By extension, the D-Pro-Gly turn motif [179] or a related depsipeptide segment L-Pro-glycolate [158] have been used to connect β-peptide
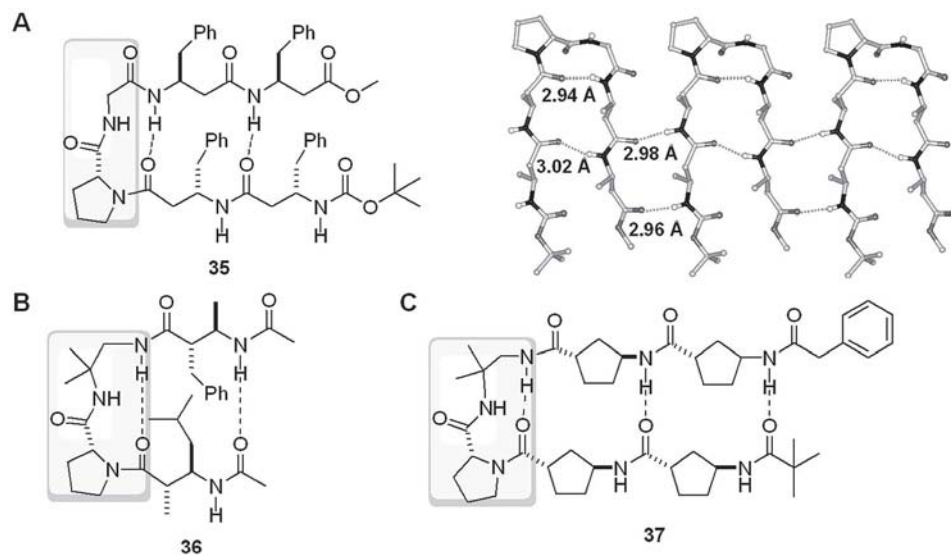
Fig. 2.18 (A) Hybrid β-peptide **35** with a D-Pro-Gly type II′ β-turn segment (grey color): X-ray crystal structure [179]. Intermolecular NH···O=C H bonds (with N–H···O) angles of 160° and 133°) connect the hairpin into an infinitely extended β-sheet. (B, C) Parallel hairpin formation in β- and γ-peptides (e.g. **36**, **37**) by incorporation of a D-Pro-DADME turn segment [158, 183].

strands into new antiparallel hairpin type structures (e.g. **35**). X-ray diffraction studies have shown that β-amino acids residues generally adopt the expected anti-periplanar conformation around the ethylene bond ($\theta \sim 180°$). The same holds true for β-hairpins obtained by incorporation of β- or γ-amino acids at discrete positions into the strand segments (θ_1 and θ_2 values $\sim 180°$) [180–182]. Specific connectors made of several residue types such as D-Pro-(1,1-dimethyl)-1,2-diaminoethyl (D-Pro-DADME) have been reported that allow formation of parallel hairpin type structures from β- (e.g. **36**) and γ-peptides (made of constrained *trans*-3-ACPC residues, e.g. **37**) strand segments [158, 183]. Conversely, studies by the groups of Gellman and Yang have shown that β-peptides (dinipecotic acid heterochiral sequence, see Section 2.5) [184] and α-aminoxy-peptides (D,L-α-aminoxy acid dimer) [116] reverse turn segments can be used to nucleate α-peptide hairpin structures.

2.6.3
Sculpting New Shapes by Integrating H-Bonding, Aromatic Interactions and Multiple Levels of Pre-organization

As outlined in the previous sections, optimal backbone pre-organization for the formation of H bonds and stabilization of well-defined secondary structures (helices and hairpin type structure) has been achieved at an unprecedented level for a whole series of aliphatic oligoamides and related peptidomimetic oligomers. The design principles applied to homogeneous backbones have been extended successfully to heterogeneous backbones. Recent studies with aromatic oligoamides suggest that it is possible to design oligomeric sequences coding for more complex structural information. Hunter and Thomas have shown that **38**, a C_2-symmetric strand built of isophthalic acid, bisaniline and 4-nitro-1*H*-pyrrole-2-carboxylic acid residues adopt in nonpolar solvent (CDCl$_3$) a well defined and unusual fold, stabilized by a combination of H-bonding and aromatic interactions between non-nearest neighbors (Fig. 2.19A) [185].

Experimentally measured folding-induced changes in ^{1}H-NMR chemical shift ($\Delta\delta$) [186] compared with unfolded reference compounds were used to calculate the three dimensional structure of this "tailbiter". Experimental H–D exchange data reflect the difference of free energy in CDCl$_3$ between the H-bonds at work, pyrrole-amide ($\sim$6 kJ mol^{-1}) versus amide-nitro ($\sim$2 kJ mol^{-1}) H-bonds) and support a noncooperative unfolding mechanism. Self-organization properties of related aromatic oligoamides composed of isophthalic acid, and bisaniline units in nonpolar solvents have been exploited for the synthesis of a whole series of supramolecular entities [187, 188], including molecular trefoil knots [189, 190]. Folding of oligoamide **41** – which has been proposed to be formed transiently upon reaction of diamine **39** with 2,6-pyridinedicarboxylic acid dichloride **40** – into a helical loop stabilized by a complex collection of intrastrand H-bonding (both remote and between nearest neighbors) and aromatic interaction is thought to be an early and critical event prior to self-threading and successful closing to amide-knot **42** (Fig. 2.19B).
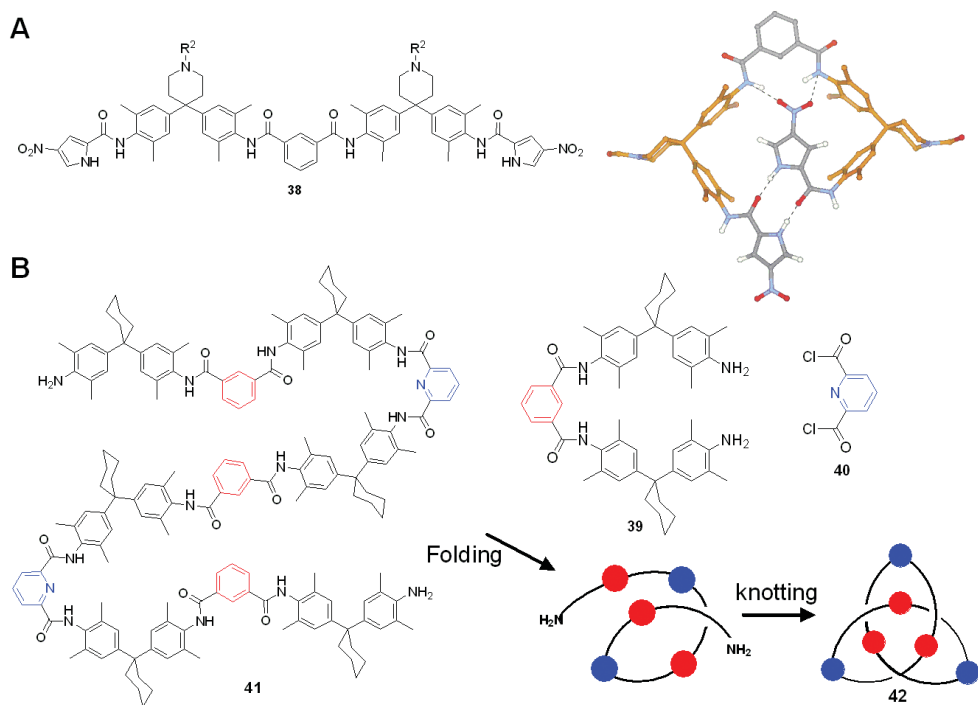
Fig. 2.19 Aromatic oligoamide foldamers (A) ''Tailbiter'' **38** [185]; (B) Proposed folding mechanism for the formation of amide knot **42** [189].

2.7
Conclusion and Outlook

The ''homologation strategy'' which initially led to the discovery of the robust 14-helical β-peptide backbone gave a strong impetus to the search of new foldamers stabilized by remote intrastrand interactions. In 10 years, the number of different secondary structures accessible from β-peptides and higher homologs of α-polypeptides has grown steadily and the basic principles underlying ω-peptide folding have been delineated. Multiple levels of conformational restriction can be applied to ω-amino acid units to pre-organize the oligoamide backbone for folding. For instance, the insertion of heteroatoms into the backbone of ω-peptides to restrict the conformational space results in nonpeptide strands with unique folding patterns. Furthermore, mixing building blocks of more than one type to generate foldamers with hybrid backbones has recently emerged as a promising concept to explore further the structure space attainable with a relatively small set of building units. Although potentially informative, a comparative study of these bioinspired foldamers is not straightforward. Several important issues such as (i) the predictability of folding; (ii) the mechanisms of folding/unfolding; (iii) the ef-

fect of environment on folding; (iv) the chain-length dependence and (v) the conversion of structure into function which have been addressed thoroughly in the case of *β*-peptides, still need to be considered for many oligomers with folding propensity. Despite some success in the formation of supersecondary structures from *β*-peptides (helix bundles and hairpin), a major challenge will be the assembly of foldamer secondary/supersecondary structures into more complex protein-like tertiary structures (tyligomers). Nonetheless, significant progress continues with regards to the design of foldamers with functions (e.g. in biology, see Chapter 8) and one foresees a bright future for folding oligomers based on remote intrastrand interactions.

Acknowledgments

References

1 Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071–13072.

2 S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.

3 D. Seebach, M. Overhand, F. N. M. Kühnle, B. Martinoni, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 913–941.

4 D. Seebach, J. L. Matthews, *Chem. Commun.* **1997**, 2015–2022.

5 J. Venkatraman, S. C. Shankaramma, P. Balaram, *Chem. Rev.* **2001**, *101*, 3131–3152.

6 W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.

7 M. J. I. Andrews, A. B. Tabor, *Tetrahedron*, **1999**, *55*, 11711–11743.

8 E. Lacroix, T. Kortemme, M. Lopez de la Paz, L. Serrano, *Curr. Opin. Struct. Biol.* **1999**, *4*, 487–493.

9 S. H. Gellman, *Curr. Opin. Chem. Biol.* **1998**, *2*, 717–725.

10 G. D. Rose, L. M. Gierasch, J. A. Smith, *Adv. Prot. Chem.* **1985**, *37*, 1–110.

11 G. Quinkert, E. Egert, C. Griesinger, *Macromolecular and Supramolecular Chemistry in Aspects of Organic Chemistry, Structure*, G. Quinkert, E. Egert, C. Griesinger (Eds.), Verlag Helvetica Chimica Acta, Basel, 1996.

12 A. Chakrabartty, R. L. Baldwin, *Adv. Prot. Chem.* **1995**, *46*, 141–176.

13 A. G. Street, S. L. Mayo, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9074–9076.

14 D. L. Jr Minor, P. S. Kim, *Nature* **1994**, *367*, 660–663.

15 C. Toniolo, M. Crisma, F. Formaggio, C. Peggion, *Biopolymers (Pept. Sci.)* **2001**, *60*, 396–419.

16 R. Fairman, K. R. Shoemaker, E. J. York, J. Stewart, R. L. Baldwin, *Proteins: Struct. Funct. Genet.* **1989**, *5*, 1–7.

17 S. Marqusee, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8898–8902.

18 C. A. Olson, Z. Shi, N. R. Kallenbach, *J. Am. Chem. Soc.* **2001**, *123*, 6451–6452.

19 C. D. Andrew, S. Bhattacharjee, N. Kokkoni, J. D. Hirst, G. R. Jones, A. J. Doig, *J. Am. Chem. Soc.* **2002**, *124*, 12706–12714.

20 F. Ruan, Y. Chen, P. B. Hopkins, *J. Am. Chem. Soc.* **1990**, *112*, 9403–9404.

21 M. R. Ghadiri, C. Choi, *J. Am. Chem. Soc.* **1990**, *112*, 1630–1632.

22 D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh, P. G. Schultz, *J. Am. Chem. Soc.* **1991**, *113*, 9391–9392.

23 J. W. Taylor, *Biopolymers* **2002**, *66*, 49–75.

24 H. E. Blackwell, R. H. Grubbs, *Angew. Chem. Int. Ed.* **1998**, *37*, 3281–3284.

25 L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine, S. J. Korsmeyer, *Science* **2004**, *305*, 1466–1470.

26 A. Banerjee, P. Balaram, *Curr. Science* **1997**, *73*, 1067–1077.

27 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4011.

28 R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, *101*, 3219–3232.

29 T. A. Martinek, F. Fülop, *Eur. J. Biochem.* **2003**, *270*, 3657–3666.

30 G. Guichard. In *Pseudopeptides in Drug Development*; Nielsen, P. E., Ed.; Wiley-VCH Verlag: Weinheim, Germany, 2004; 33–120.

31 D. Seebach, A. K. Beck, D. J. Bierbaum, *Chem. Biodiv.* **2004**, *1*, 1111–1239.

32 G. Lelais, D. Seebach, *Biopolymers* **2004**, *76*, 206–243.

33 D. Seebach, T. Kimmerlin, R. Sebesta, M. A. Campo, A. K. Beck, *Tetrahedron* **2004**, *60*, 7455–7466.

34 F. Fülop, T. A. Martinek, G. K. Toth, *Chem. Soc. Rev.* **2006**, *35*, 323–334.

35 D. Seebach, D. F. Hook, A. Glättli, *Biopolymers (Pept. Sci)* **2006**, *84*, 23–37.

36 J. Kovacs, R. Ballina, R. L. Rodin, D. Balasubramanian, J. Applequist, *J. Am. Chem. Soc.* **1965**, *87*, 119–120.

37 J. M. Fernández-santín, J. Ayamí, A. Rodríguez-Galán, S. Muñoz-Guerra, J. A. Subirana, *Nature*, **1984**, *311*, 53–54.

38 H. N. Rydon, *J. Chem. Soc.* **1964**, 1328–1333.

39 H. M. Müller, D. Seebach, *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 477–502.

40 D. Seebach, M. G. Fritz, *Biol. Macromolecules* **1999**, *25*, 217–236.

41 G. P. Dado, S. H. Gellman, *J. Am. Chem. Soc.* **1994**, *116*, 1054–1062.

42 T. Hinterman, D. Seebach, *Synlett* **1997**, 437–438.

43 D. Seebach, P. E. Ciceri, M. Overhand, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 2043–2066.

44 K. Gademann, B. Jaun, D. Seebach, R. Perozzo, L. Scapozza, G. Folkers, *Helv. Chim. Acta* **1999**, 82, 1–11.

45 T. Hintermann, K. Gademann, B. Jaun, D. Seebach, *Helv. Chim. Acta* **1998**, *81*, 983–1002.

46 S. Hanessian, X. Luo, R. Schaum, S. Michnick, *J. Am. Chem. Soc.* **1998**, *120*, 8569–8570.

47 D. Seebach, M. Brenner, M. Rueping, B. Schweizer, B. Jaun, *Chem. Commun.* **2001**, 207–208.

48 D. Seebach, M. Brenner, M. Rueping, B. Jaun, *Chem. Eur. J.* **2002**, *8*, 573–584.

49 A. Banerjee, P. Balaram, *Curr. Science* **1997**, *73*, 1067–1077.

50 S. Abele, P. Seiler, D. Seebach, *Helv. Chim. Acta* **1999**, *82*, 1559–1571.

51 D. Seebach, S. Abele, T. Sifferlen, M. Hänggi, S. Gruner, P. Seiler, *Helv. Chim. Acta* **1998**, *81*, 2218–2219.

52 D. Seebach, T. Sifferlen, P. A. Mathieu, A. M. Häne, C. M. Krell, D. J. Bierbaum, S. Abele, *Helv. Chim. Acta* **2000**, *83*, 2849–2864.

53 D. Seebach, T. Sifferlen, D. J. Bierbaum, M. Rueping, B. Jaun, B. Schweizer, J. Schaefer, A. K. Mehta, R. D. O'Connor, B. H. Meier, M. Ernst, A. Glättli, *Helv. Chim. Acta* **2002**, *85*, 2877–2917.

54 A. Glättli, X. Daura, D. Seebach, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2002**, *124*, 12972–12978.

55 D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, L.

Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1998**, *81*, 932–982.

56 D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1999**, *121*, 6206–6212.

57 J. J. Barchi Jr, X. Huang, D. H. Apella, L. A. Christianson, S. R. Durell, S. Gellman, *J. Am. Chem. Soc.* **2000**, *122*, 2711–2718.

58 D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. Huang, J. J. Barchi, S. H. Gellman, *Nature* **1997**, *387*, 381–384.

59 D. H. Appella, L. A. Christianson, D. A. Klein, M. R. Richards, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1999**, *121*, 7574–7581.

60 L. A. Christianson, M. J. Lucero, D. H. Appella, D. A. Klein, S. H. Gellman, *J. Comput. Chem.* **2000**, *21*, 763–773.

61 T. D. W. Claridge, J. M. Goodman, A. Moreno, D. Angus, S. F. Barker, C. Taillefumier, M. P. Watterson, G. W. J. Fleet, *Tetrahedron Lett.* **2001**, *42*, 4251–4255.

62 A. Hetenyi, I. M. Mandity, T. A. Martinek, G. K. Toth, F. Fülöp, *J. Am. Chem. Soc.* **2005**, *127*, 547–553.

63 J. D. Winkler, E. L. Piatnitski, J. Mehlmann, J. Kasparec, P. H. Axelsen, *Angew. Chem. Int. Ed.* **2001**, *40*, 743–745.

64 S. Izquierdo, M. J. Kogan, T. Parella, A. G. Moglioni, V. Branchadell, E. Giralt, R. M. Ortuno, *J. Org. Chem.* **2004**, *69*, 5093–5099.

65 S. Chandrasekhar, M. S. Reddy, B. N. Babu, B. Jagadeesh, A. Prabhakar, B. Jagannadh, *J. Am. Chem. Soc.* **2005**, *127*, 9664–9665.

66 L. Szabo, B. L. Smith, K. D. McReynolds, A. Parrill, E. R. Morris, J. Gervay, *J. Org. Chem.* **1998**, *63*, 1074–1078.

67 M. D. Smith, D. D. Long, T. D. W. Claridge, G. W. J. Fleet, D. G. Marquess, *Chem. Commun.* **1998**, 2039–2040.

68 T. D. W. Claridge, D. D. Long, N. L. Hungerford, R. T. Aplin, M. D. Smith, D. G. Marquess, G. W. J. Fleet *Tetrahedron Lett.* **1999**, *40*, 2199–2202.

69 T. D. W. Claridge, D. D. Long, C. M. Baker, B. Odell, G. H. Grant, A. A. Edwards, G. E. Tranter, G. W. J. Fleet, M. D. Smith, *J. Org. Chem.* **2005**, *70*, 2082–2090.

70 C. Baldauf, R. Günther, H.-J. Hofmann, *J. Org. Chem.* **2004**, *69*, 6214–6220.

71 R. W. Hoffmann, *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1124–1134.

72 R. W. Hoffman, *Angew. Chem. Int. Ed.* **2000**, *39*, 2054–2070.

73 D. L. Boger, T. M. Ramsey, H. Cai, S. T. Hoehn, J. Stubbe, *J. Am. Chem. Soc.* **1998**, *120*, 9149–9158.

74 C. Grison, P. Coutrot, S. Geneve, C. Didierjean, M. Marraud, *J. Org. Chem.* **2005**, *70*, 10753–10764.

75 P. G. Vasudev, N. Shamala, K. Ananda, P. Balaram, *Angew. Chem. Int. Ed.* **2005**, *44*, 4972–4975.

76 J. Farrera-Sinfreu, L. Zaccaro, D. Vidal, X. Salvatella, E. Giralt, M. Pons, F. Albericio, M. Royo, *J. Am. Chem. Soc.* **2004**, *126*, 6048–6057.

77 V. Semetey, D. Rognan, C. Hemmerlin, R. Graff, J.-P. Briand, M. Marraud, G. Guichard, *Angew. Chem. Int. Ed.* **2002**, *41*, 1893–1895.

78 C. Hemmerlin, M. Marraud, D. Rognan, R. Graff, V. Semetey, J. P. Briand, G. Guichard, *Helv. Chim. Acta* **2002**, *85*, 3692–3711.

79 A. Violette, M. C. Averlant-Petit, V. Semetey, C. Hemmerlin, R. Casimir, R. Graff, M. Marraud, J.-P. Briand, D. Rognan, G. Guichard, *J. Am. Chem. Soc.* **2005**, *127*, 2156–2164.

80 A. Violette, S. Fournel, K. Lamour, O. Chaloin, B. Frisch, J.-P. Briand, H. Monteil, G. Guichard, *Chem. Biol.* **2006**, *13*, 531–538.

81 M. Rueping, B. Jaun, D. Seebach, *Chem. Commun.* **2000**, 2267–2268.

82 D. H. Appella, J. J. Barchi Jr, S. R. Durell, S. H. Gellman, *J. Am. Chem. Soc.* **1999**, *121*, 2309–2310.

83 M. Schinnerl, J. K. Murray, J. M. Langenhan, S. H. Gellman, *Eur. J. Org. Chem.* **2003**, 721–726.

84 X. Wang, J. F. Espinosa, S. H. Gellman, *J. Am. Chem. Soc.* **2000**, *122*, 4821–4822.

85 E. A. Porter, X. Wang, M. A. Schmitt, S. H. Gellman, *Org. Lett.* **2002**, *4*, 3317–3319.

86 M. G. Woll, J. D. Fisk, P. R. LePlae, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 12447–12452.

87 T. L. Raguse, E. A. Porter, B. Weisblum, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 12774–12785.

88 P. R. LePlae, J. D. Fisk, E. A. Porter, B. Weisblum, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 6820–6821.

89 J. S. Park, H.-S. Lee, J. R. Lai, B. M. Kim, S. H. Gellman, *J. Am. Chem. Soc.* **2003**, *125*, 8539–8545.

90 P. I. Arvidsson, M. Rueping, D. Seebach, *Chem. Commun.* **2001**, 649–650.

91 R. P. Cheng, W. F. DeGrado, *J. Am. Chem. Soc.* **2001**, *123*, 5162–5163.

92 T. L. Raguse, J. R. Lai, S. H. Gellman, *J. Am. Chem. Soc.* **2003**, *125*, 5592–5593.

93 T. L. Raguse, J. R. Lai, S. H. Gellman, *Helv. Chim. Acta* **2002**, *85*, 4154–4164.

94 S. A. Hart, A. B. F. Bahadoor, E. E. Matthews, X. J. Qiu, A. Schepartz, *J. Am. Chem. Soc.* **2003**, *125*, 4022–4023.

95 J. A. Kritzer, J. Tirado-Rives, S. A. Hart, J. D. Lear, W. L. Jorgensen, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 167–178.

96 D. A. Guarracino, H. R. Chiang, T. N. Banks, J. D. Lear, M. E. Hodsdon, A. Schepartz, *Org. Lett.* **2006**, *8*, 807–810.

97 R. P. Cheng, W. F. Degrado, *J. Am. Chem. Soc.* **2002**, *124*, 11564–11565.

98 W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland.

99 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 236–240.

100 X. Daura, W. F. van Gunsteren, D. Rigo, B. Jaun, D. Seebach, *Chem. Eur. J.* **1997**, *3*, 1410–1417.

101 X. Daura, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *J. Mol. Biol.* **1998**, *280*, 925–932.

102 X. Daura, W. F. van Gunsteren, A. E. Mark, *Proteins* **1999**, *34*, 269–280.

103 D. Seebach, J. V. Schreiber, S. Abele, X. Daura, W. F. van Gunsteren, *Helv. Chim. Acta* **2000**, *83*, 34–57.

104 W. F. van Gunsteren, R. Burgi, C. Peter, X. Daura, *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 351–355.

105 K. Wright, M. Sarciaux, M. Wakselman, J. P. Mazaleyrat, M. Crisma, F. Formaggio, C. Peggion, A. Toffoletti, C. Corvaja, C. Toniolo, *Proc. 19th American Peptide Symposium*; Blondelle, S. E., Ed.; *American Peptide Society*, **2005**, 557–558.

106 E. Vass, M. Hollosi, F. Besson, R. Buchet, *Chem. Rev.* **2003**, *103*, 1917.

107 M. Crisma, F. Formaggio, A. Moretto, C. Toniolo, *Biopolymers (Pept. Sci.)* **2006**, *84*, 3–12.

108 A. I. Jiménez, G. Ballano, C. Cativiela, *Angew. Chem. Int. Ed.* **2005**, *44*, 396.

109 K. Gademann, A. Häne, M. Rueping, B. Jaun, D. Seebach, *Angew. Chem. Int. Ed.* **2003**, *42*, 1534–1537.

110 R. J. Doerksen, B. Chen, J. Yuan, J. D. Winkler, M. L. Klein, *Chem. Commun.* **2003**, 2534–2535.

111 G. V. M. Sharma, P. Jayaprakash, K. Narsimulu, A. Ravi Sankar, K. Ravinder Reddy, P. Radha Krishna, A. C. Kunwar, *Angew. Chem. Int. Ed.* **2006**, *45*, 2944–2947.

112 D. Yang, F. F. Ng, Z. J. Li, Y. D. Wu, K. W. K. Chan, D. P. Wang, *J. Am. Chem. Soc.* **1996**, *118*, 9794–9795.

113 Y. D. Wu, D. P. Wang, K. W. K. Chan, D. Yang, *J. Am. Chem. Soc.* **1999**, *121*, 11189–11196.

114 C. Peter, X. Daura, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2000**, *122*, 7461–7466.

115 D. Yang, J. Qu, B. Li, F. F. Ng, X. C. Wang, K. K. Cheung, D. P. Wang, Y. D. Wu, *J. Am. Chem. Soc.* **1999**, *121*, 589–590.

116 D. Yang, J. Qu, W. Li, Y. Ren, D.-P. Wang, Y. D. Wu, *J. Am. Chem. Soc.* **2003**, *125*, 14452–14457.

117 D. Yang, Y. H. Zhang, N. Y. Zhu, *J. Am. Chem. Soc.* **2002**, *124*, 9966–9967.

118 D. Yang, Y. H. Zhang, B. Li, D. W. Zhang, J. C. Y. Chan, N. Y. Zhu, S. W. Luo, Y. D. Wu, *J. Am. Chem. Soc.* **2004**, *126*, 6956–6967.

119 D. Yang, D. W. Zhang, Y. Hao, Y. D. Wu, S. W. Luo, N. Y. Zhu, *Angew. Chem. Int. Ed.* **2004**, *43*, 6719–6722.

120 F. Chen, N. Y. Zhu, D. Yang, *J. Am. Chem. Soc.* **2004**, *126*, 15980–15981.

121 R. Günther, H.-J. Hoffmann, *J. Am. Chem. Soc.* **2001**, *123*, 247–255.

122 G. Lelais, D. Seebach, *Helv. Chim. Acta*, **2003**, *86*, 4152–4168.

123 A. Cheguillaume, A. Salaün, S. Sinbandhit, M. Potel, P. Gall, M. Baudy-Floc'h, P. Le Grel, *J. Org. Chem.* **2001**, *66*, 4923–4929.

124 A. Aubry, J. P. Mangeot, J. Vidal, A. Collet, S. Zerkout, M. Maraud, *Int. J. Pept. Protein Res.* **1994**, *43*, 305–311.

125 A. Salaün, M. Potel, T. Roisnel, P. Gall, P. Le Grel, *J. Org. Chem.* **2005**, *70*, 6499–6502.

126 A. Salaün, A. Favre, B. Le Grel, M. Potel, P. Le Grel, *J. Org. Chem.* **2006**, *71*, 150–158.

127 L. Thévenet, R. Vanderesse, M. Marraud, C. Didierjean, A. Aubry, *Tetrahedron Lett.* **2000**, *41*, 2361–2364.

128 P. Le Grel, A. Salaün, M. Potel, B. Le Grel, F. Lassagne, *J. Org. Chem.* **2006** *71*, 5638–5645.

129 G. N. Ramachandran, R. Chandrasekharan, *Indian J. Biochem. Biophys.* **1972**, *9*, 1–11.

130 D. W. Urry, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 672–676.

131 P. De Santis, S. Morosetti, R. Rizzo, *Macromolecules* **1974**, *7*, 52–58.

132 E. Benedetti, B. Di Blasio, C. Pedone, G. P. Lorenzi, L. Tomasic, V. Gramlich, *Nature* **1979**, *282*, 630.

133 G. P. Lorenzi, H. Jaeckle, L. Tomasic, V. Rizzo, C. Pedone, *J. Am. Chem. Soc.* **1982**, *104*, 1728–1733.

134 E. Navarro, R. Tejero, E. Fenude, B. Celda, *Biopolymers* **2001**, *59*, 110–119.

135 E. Navarro, E. Fenude, B. Celda, *Biopolymers* **2004**, *73*, 229–241.

136 D. A. Langs, *Science* **1988**, *241*, 188–191.

137 A. L. Lomize, V. Orekhov, A. S. Arseniev, *Bioorg. Khim.* **1992**, *18*, 182–200.

138 B. M. Burkhart, N. Li, D. A. Langs, W. A. Pangborn, W. L. Duax, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12950–12955.

139 F. Kovacs, J. Quine, T. A. Cross, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 7910–7915.

140 M. D. Yoder, N. T. Keen, F. Jurnak, *Science* **1993**, *260*, 1503–1507.

141 C. R. H. Raetz, S. L. Roderick, *Science* **1995**, *270*, 997–1000.

142 P. Lengar, N. V. Jishi, P. Balaram, *Structure* **2006**, *14*, 529–542.

143 M. F. Perutz, J. L. Finch, J. Berriman, A. Lesk, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5591–5595.

144 D. Seebach, K. Gademann, J. V. Schreiber, J. L. Matthews, T. Hintermann, B. Jaun, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1997**, *80*, 2033–2038.

145 M. Rueping, J. V. Schreiber, G. Lelais, B. Jaun, D. Seebach, *Helv. Chim. Acta* **2002**, *85*, 2577–2593.

146 S. A. Gruner, V. Truffault, G. Voll, E. Locardi, M. Stockle, H. Kessler, *Chem. Eur. J.* **2002**, *8*, 4365–4376.

147 G. V. M. Sharma, K. R. Reddy, P. R. Krishna, A. R. Sankar, K. Narsimulu, S. K. Kumar, P. Jayaprakash, B. Jagannadh, A. C. Kunwar, *J. Am. Chem. Soc.* **2003**, *125*, 13670–13671.

148 G. V. M. Sharma, K. R. Reddy, P. R. Krishna, A. R. Sankar, P. Jayaprakash, B. Jagannadh, A. C. Kunwar, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 3961–3965.

149 R. Baron, D. Bakowies, W. F. van Gunsteren, X. Daura, *Helv. Chim. Acta*, **2002**, *85*, 3872–3882.

150 D. Trzesniak, A. Glattli, B. Jaun, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2005**, *127*, 14320–14329.

151 P. J. Gee, W. F. van Gunsteren, *Proteins* **2006**, *63*, 136–143.

152 Y. D. Wu, D. P. Wang, *J. Am. Chem. Soc.* **1999**, *121*, 9352–9362.

153 C. Baldauf, R. Günther, H.-J. Hofmann, *Angew. Chem. Int. Ed.* **2004**, *43*, 1594–1597.

154 C. Baldauf, R. Günther, H.-J. Hofmann, *Biopolymers* **2005**, *80*, 675–687.

155 C. Baldauf, R. Günther, H.-J. Hofmann, *Biopolymers* **2006**, *84*, 408–413.

156 C. Baldauf, R. Günther, H.-J. Hofmann, *J. Org. Chem.* **2006**, *71*, 1200–1208.

157 G. V. M. Sharma, P. Nagendar, P. Jayaprakash, P. Ragha Krishna, K. V. S. Ramakrishna, A. C. Kunwar, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 5878–5882.

158 S. Krauthauser, L. A. Christianson, D. R. Powell, and S. H. Gellman, *J. Am. Chem. Soc.* **1997**, *119*, 11719–11720.

159 D. Seebach, S. Abele, K. Gademann, B. Jaun, *Angew. Chem. Int.* **1999**, *38*, 1595–1597.

160 M. Hagihara, N. J. Anthony, T. J. Stout, J. Clardy, S. L. Schreiber, *J. Am. Chem. Soc.* **1992**, *114*, 6568–6570.

161 M. G. Woll, J. R. Lai, I. A. Guzei, S. J. C. Taylor, M. E. B. Smith, S. H. Gellman, *J. Am. Chem. Soc.* **2001**, *123*, 11077–11078.

162 Y. J. Chung, B. R. Huck, L. A. Christianson, H. E. Stanger, S. Krauthauser, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **2000**, *122*, 3995–4004.

163 Y. J. Chung, L. A. Christianson, H. E. Stanger, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 10555–10556.

164 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, *J. Am. Chem. Soc.* **2001**, *123*, 2393–2404.

165 S. Hanessian, X. Luo, R. Schaum, *Tetrahedron Lett.* **1999**, *40*, 4925–4929.

166 M. Brenner, D. Seebach, *Helv. Chim. Acta* **2001**, *84*, 2155–2166.

167 S. De Pol, C. Zorn, C. D. Klein, O. Zerbe, O. Reiser, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 511–514.

168 A. Hayen, M. A. Schmitt, F. N. Ngassa, K. A. Thomasson, S. H. Gellman, *Angew. Chem. Int. Ed.* **2004**, *43*, 505–510.

169 M. A. Schmitt, B. Weisblum, S. H. Gellman, *J. Am. Chem. Soc.* **2004**, *126*, 6848–6849.

170 M. A. Schmitt, S. H. Choi, I. A. Guzei, S. H. Gellman, *J. Am. Chem. Soc.* **2005**, *127*, 13130–13131.

171 M. A. Schmitt, S. H. Choi, I. A. Guzei, S. H. Gellman, *J. Am. Chem. Soc.* **2006**, *128*, 4538–4539.

172 K. Ananda, P. G. Vasudev, A. Sengupta, K. M. Poopathi Raja, N. Shamala, P. Balaram, *J. Am. Chem. Soc.* **2005**, *127*, 16668–16674.

173 R. S. Roy, P. Balaram, *J. Pept. Res.* **2004**, *63*, 279–289.

174 I. L. Karle, A. Pramanik, A. Banerjee, S. Bhattacharjya, P. Balaram, *J. Am. Chem. Soc.* **1997**, *119*, 9087–9095.

175 R. S. Roy, I. L. Karle, S. Raghothama, P. Balaram, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16478–16482.

176 T. S. Haque, J. C. Little, S. H. Gellman, *J. Am. Chem. Soc.* **1994**, *116*, 4105–4106.

177 M. Ramirez-Alvarado, F. J. Blanco, L. Serrano, *Nat. Struct. Biol.* **1996**, *3*, 604–612.

178 I. L. Karle, S. K. Awasthi, P. Balaram, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8189–8193.

179 I. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5160–5164.

180 I. L. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3716–3719.

181 H. N. Gopi, R. S. Roy, S. Raghothama, I. Karle, P. Balaram, *Helv. Chim. Acta* **2002**, *85*, 3313–3330.

182 R. S. Roy, H. N. Gopi, S. Raghothama, I. L. Karle, P. Balaram, *Chem. Eur. J.* **2006**, *12*, 3295–3302.

183 J. M. Langenhan, I. A. Guzei, S. H. Gellman, *Angew. Chem. Int. Ed.* **2003**, *42*, 2402–2405.

184 B. R. Huck, J. D. Fisk, S. H. Gellman, *Org. Lett.* **2000**, *2*, 2607–2610.

185 C. A. Hunter, A. Spitaleri, S. Tomas, *Chem Commun.* **2005**, 3691–3693.

186 C. A. Hunter, M. J. Packer, *Chem. Eur. J.* **1999**, *5*, 1891–1897.

187 A. P. Bisson, F. J. Carver, D. S. Eggleston, R. C. Haltiwanger, C. A. Hunter, D. L. Livingstone, J. F.

McCabe, C. Rotger, A. E. Rowan, *J. Am. Chem. Soc.* **2000**, *122*, 8856–8868.

188 C. A. Hunter, *J. Am. Chem. Soc.* **1992**, *114*, 5303–5311.

189 O. Lukin, F. Vögtle, *Angew. Chem. Int. Ed.* **2005**, *44*, 1456–1477.

190 O. Safarowsky, M. Nieger, R. Fröhlich, F. Vögtle, *Angew. Chem. Int. Ed.* **2000**, *39*, 1616–1618.

3
Foldamers Based on Solvophobic Effects

Yan Zhao and Jeffrey S. Moore

Solvophobic effects are responsible for the association of poorly-solvated molecular surfaces and are one of the most important noncovalent forces used by nature in biofoldamers. Unlike specific interactions such as hydrogen bonds, solvophobic effects rely on collective interactions among solvents and solutes instead of specific functional groups for operation. To use solvophobic effects for foldamer design, one must create solvophobic sites or patches along the primary sequence to collapse an otherwise conformationally random chain. A fundamental challenge in this effort is to overcome nonspecificity of solvophobic forces as well as the coexisting and ubiquitous van der Waals interactions. In this chapter, we will start with a general introduction of solvent effects. We will clarify the term ''solvophobic effects'' within the context of our discussion and illustrate how solvophobic effects are employed in inter- as well as intramolecular supramolecules for structural control. We will discuss how solvophobic effects are related to the behavior of both synthetic and natural polymers. Next we will summarize the recent progress in solvophobic foldamers with an emphasis on the strategies for using solvophobic effects, and end the chapter with conclusions and outlooks for some future challenges.

3.1
Introduction

Solvents provide a medium in which reactants encounter one another – this is the major role of solvents played in most covalent (heterolytic and homolytic) reactions. Although inertness is often one of the most important criteria in selecting solvents for a chemical process, the process itself is rarely insensitive to solvents. Chemical equilibria, kinetics, light absorption, and emission are among the numerous processes that can be profoundly influenced by solvents [1]. A classical solvent effect in covalent chemistry, frequently mentioned in undergraduate textbooks, is the contrasting effect of protic and polar aprotic solvents on S_N2 reactions. For a reaction involving anionic nucleophiles such as sodium azide, differ-

ent solvents can easily change the reaction rate over several orders of magnitude. Solvation of the anionic azide by protic solvents (e.g. methanol) occurs through hydrogen bonding, which reduces the availability of electrons on the nucleophile and slows down the reaction. A polar aprotic solvent such as hexamethylphosphoramide (HMPA), on the other hand, preferentially solvates the sodium countercation. Removal of the cation from the nucleophile increases the availability of its electrons, speeding up the reaction as a result.

Solvents are even less innocent bystanders in supramolecular reactions than in covalent reactions because the final products, supramolecules, are formed by noncovalent bonds and both solvent–solute and solvent–solvent interactions are noncovalent in nature. Hence, it is no exaggeration to speak of solvents as active participants or ''reactants'' in supramolecular reactions. Just as a poor choice of solvents can cause slow reactions and low yields of products in covalent reactions, it can completely change ''mechanisms'' and product distribution in noncovalent ones. Two kinds of solvent effects are generally found in supramolecular chemistry, similar to those mentioned in the above S_N2 reactions. The first kind is of a *direct* effect, in which solvents compete for the (supramolecular) reactive sites and impede the reaction. This is the fundamental reason why most hydrogen-bonded supramolecular structures do not survive in water or dimethyl sulfoxide (DMSO). The second kind is *indirect* and is illustrated by the hydrophobic effects. In hydrophobic effects, the driving force for association of nonpolar solutes in an aqueous solution is often not the dispersive interactions among associating solute molecules but rather from the unique properties of water such as strong hydrogen bonds and small molecular size.

Hydrophobic effects are responsible for a variety of processes including membrane formation, protein folding, protein–protein interactions, ligand–receptor binding, laundry cleaning, drug formulation, and countless other processes where nonpolar molecules or groups interact with water [2–20]. Classical hydrophobic effects are thought to be of entropic origin, at least at low temperatures. This is evidenced by the negative (favorable) enthalpies and large positive (unfavorable) entropies of solvation for small nonpolar solutes in water [16]. In the classical model of hydrophobic effects, water forms iceberg-like structures around nonpolar solutes [19–20]. Although frequently seen in water clathrates of small organic molecules, such ordering was not verified by neutron diffraction studies [21]. Instead of focusing on hydrogen bonds, alternative models emphasize other properties of water such as small size and high cohesive energy density (i.e. total intermolecular interactions per unit volume) [8, 9]. In these models, attention is turned to cavities or interfaces created in water for accommodating nonpolar solutes and their influences on hydrophobic aggregation. Importantly, all models agree with the same experimental observation that strength of hydrophobic effects increases with buried hydrophobic surface area upon association. Now, hydrophobic effects are accepted to be a multifaceted phenomenon, with its molecular origin depending not only on temperatures but also on molecular sizes – entropically driven at low temperatures with small solutes and enthalpically driven at high temperatures or with large nonpolar surfaces [2–8]. The energetic

characteristics depend on the nature (e.g. aliphatic or aromatic) as well as the shape (e.g. deep cavity) of the hydrophobic surfaces.

Solvophobic effects are extension of hydrophobic effects and were originally used to describe aggregation of hydrophobic molecules in polar nonaqueous solutions [22, 23]. Although commonly used to refer to association of poorly solvated molecular surfaces, the exact meaning of the term becomes blurred when water, a solvent with extremely small volume, low polarizability, and high cohesive energy density is replaced by organic solvents that have many different and even opposite characteristics. Whereas aggregation of solvent-incompatible solutes in formamide and ethylene glycol may retain many of the energetic profiles of hydrophobic effects, association of solvent-incompatible surfaces in nonpolar solvents such as alkanes, chloroform, or tetrahydrofuran probably deviate significantly. It is clear that many of so-called solvophobically driven supramolecular systems (including the ones reviewed in this chapter) have contributions from both intersolute van der Waals interactions and classical ''solvophobic'' effects – the former may even be the dominant one. Nevertheless, it is convenient to use a single term ''solvophobic effects'' to refer to these solvent-dependent phenomena and the simplification is justified by several reasons. First, both van der Waals interactions and classical solvophobic effects are proportional to the interacting surface area; grouping the two together, therefore, does not change conclusions drawn in most cases. Second, structures of supramolecules and strategies for structural control are often the central points of discussion instead of the exact mechanisms of intermolecular interactions. As separation of van der Waals interactions and ''true'' solvophobic effects is difficult even in well studied systems such as binding of nonpolar molecules by cyclodextrins [24], it may be unfruitful arguing ''how solvophobic really is a particular solvophobic interaction?'' Third, classical hydrophobic effects have contributions from intersolute van der Waals interactions (e.g. dispersive forces) as well. Although possibly not very important in hydrophobic interactions of small gaseous organic molecules, van der Waals interactions probably are quite important in the association of large hydrocarbons. For example, one normally assumes hydrophobic effects to be responsible for insolubility of both methane and hexadecane in water, even though dispersive forces of the two are enormously different (as reflected in the corresponding boiling points of -164 and $287\ ^\circ$C, respectively).

3.2
Learning from Solvophobically Driven Assemblies – Intermolecular Solvophobic Interactions

One way of looking at foldamer chemistry with its efforts in employing noncovalent interactions to stabilize collapsed conformations is classical supramolecular chemistry carried out intramolecularly. If molecular recognition between different molecules is the goal for the latter, recognition between different parts of a flexible molecule is the goal for the former. For specific, directional noncovalent

interactions such as hydrogen bonds, the information for molecular recognition is encoded in the donor–acceptor motif. Thus, a D-D-A hydrogen bonding motif is most suitable to interact with an A-A-D motif. For nonspecific interactions such as solvophobic and van der Waals interactions, complementarity is still needed for selectivity but is required for the entire binding surfaces instead of a localized pair of interacting functional groups. As both forces are proportional to the interacting surface area, geometries (including the size, shape, and distribution) of the binding surfaces logically become the most important parameters one can use to modulate both the stability and the selectivity of interactions. We call this strategy of creating complementary solvophobic surfaces for molecular recognition *geometrical manipulation*. In 1991, Whitesides [25] pointed out that ''van der Waals and hydrophobic interactions ... are ubiquitous in biological systems, but have been difficult to use by design in synthetic systems.'' A long road still lies ahead of chemists in approaching nature's abilities to use these nondirectional interactions for structural control. We are, however, much closer than we were 15 years ago.

Solvophobic effects are most simply demonstrated when an organic solvent such as hexane (''oil'') is poured over an aqueous solution. In such a system, any ''oil'' molecule can freely approach other ''oil'' molecules and no geometrical constraint is present. The requirement of maximum solvophobic contact (or minimal solvophobic/solvophilic contact) naturally calls for complete phase separation of the two components, with the ''structure'' of the ''product'' determined by relative volumes of the two components and shape of the container used.

This is not the case when a head/tail amphiphile is dissolved in water. Contact among the hydrophobic tails is now restricted by the hydrophilic headgroups. Indeed, both the stability and the types of surfactant aggregates are determined by the two distinctively different yet inseparable parts of the molecule. Even with a simple topology, head/tail amphiphiles can form a variety of aggregates ranging from relatively simple structures such as micelles, vesicles, and reversed micelles to lyotropic liquid crystals with complex nanometer-sized phase-separated domains. Despite the large number of possibilities, the preferred aggregates can be understood from simple geometrical considerations. This is the critical packing parameter ($Q = v/a_0 l_c$) proposed by Isrealachvili in which v is the volume of the hydrophobic tails, a_0 the area of the hydrophilic headgroup, and l_c the average critical length of the amphiphile [26]. Spherical micelles are preferred with $Q < 1/3$ and cylindrical micelles with $1/2 < Q < 1/3$. With even higher Q, other structures such as hexagonal, lamellar, bicontinuous cubic, and inverted hexagonal phases are favored. Simple geometrical consideration is able to predict the structure of the aggregates because maximal solvophobic contact allowed by the head and the tail is the major driving force in the system. Similar phase structures are displayed by flexible diblock copolymers and can be reliably predicted by the relative volume fractions of the individual blocks [27–31]. Even though it is not a solvophobic effect in block copolymers (as no solvent is present), minimal contact between chemically different components is the same as in surfactant aggregates and is a manifestation of the simple ''like dissolves like'' principle.

Limited geometrical manipulation gives limited selectivity. Surfactant micelles can solubilize a wide range of hydrophobic molecules with no discrimination. In order to have a higher level of structural control, rigid hydrophobic groups must be employed in the assembling process. When the hydrophobic groups involved have distinctive shape, curvature, and dimensions instead of being a simple flexible chain, only limited packing motifs are allowed in order to maximize solvophobic contact and complex assemblies with highly unique structures may result. This approach has long been used in rod-coil block copolymers in which the "rod" is a stiff, often aromatic-containing block and the "coil" represents a conventional flexible polymer [32–36]. Depending on the exact structures and ratio of the two blocks, various nanoscale objects such as bundles [37, 38], ribbons [39], tubules [40], and vesicles [41] can be formed. One such example was reported by Stupp and co-workers [37], which nicely demonstrated the principle of geometrical manipulation. In triblock copolymer **1**, the rigid rod of biphenyl units tends to self associate into 2D structures [42]. The polystyrene block serves to prevent infinite growth of the 2D aggregates and isolate the crosslinkable middle block of polybutadiene so that crosslinking only happens within a single aggregate. The result is anisotropic (2×8 nm), mushroom-shaped nanoclusters that are completely soluble in organic solvents.

1
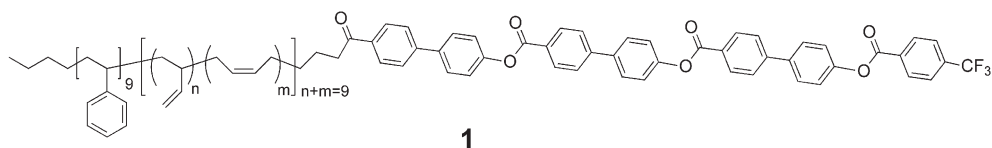
Scheme 3.1

Another field that frequently relies on geometrical manipulation of individual components is liquid crystalline materials [43]. Solvents are involved in lyotropic but not in thermotropic liquid crystals. Nevertheless, the same interplay between enthalpy and entropy expressed in minimal contact between dissimilar components exists in both systems. Therefore, it is not difficult to imagine rod-like molecules such a **2** form *nematic* (Greek for *soap*, layer-like structures) phases and flat aromatic molecules with alkyl side chains (e.g. **3**) form *discotic* phases [44]. Geometrical control is continued to be employed by chemists in search of novel liquid crystalline phases and is especially productive in recent years when unusually shaped molecules such as dendrimers are introduced [45].

Recently, a number of novel nonpolymeric assemblies appeared in the literature based on geometrical manipulation of amphiphilic structures. With two dendritic headgroups and four hydrophobic tails, cone-shaped amphiphile **4** forms a structurally persistent micelle consisting of seven molecules [46]. Unlike conventional micelles that have a distribution of sizes and highly dynamic in nature, micelles from **4** have stable, well defined structures resembling viruses and protein
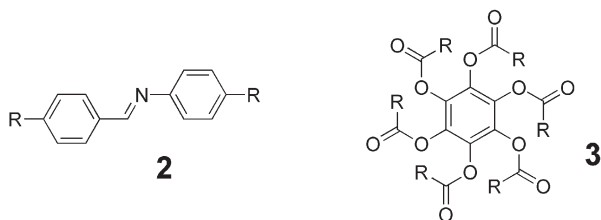
Scheme 3.2

aggregates. Hexabenzocoronene derivative **5** aggregates into 14-nm-diameter tu-bular objects, with the solvophilic tri(ethylene glycol) chains located on the inner and outer surfaces of the tube and solvophobic contacts among the aromatic and aliphatic groups to stabilize the bilayer structure [47]. Rigid–flexible macrocycle such as **6** is particularly interesting with the hydrophilic poly (ethylene glycol) or PEG chains at the two ends of the hexaphenylene joined together by covalent bonds. Solvophobic contact between the aromatic segments is thus confined by the looped hydrophilic side chains. Macrocycles with high proportions of PEO chains assemble into donut-like nanoclusters [48] and those with slightly lower PEO portions give extremely long tubes by coiling of ribbon-like structures [49].
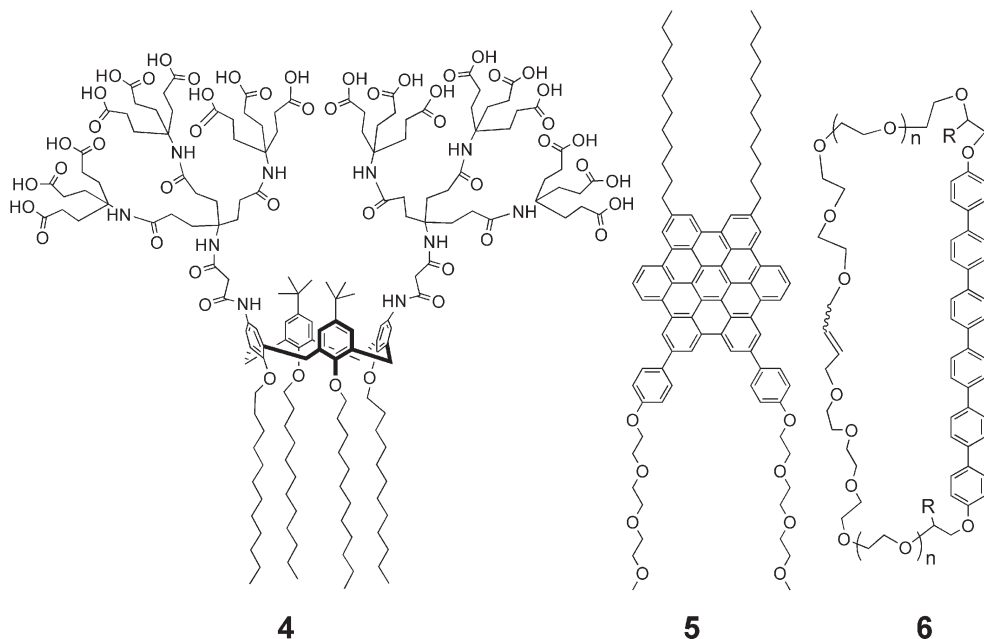
Scheme 3.3

Geometrical manipulation may appear straightforward, but direct prediction of the final assembled structures is often difficult, as the "shapes" of most molecules cannot be represented by simple geometrical objects. Moreover, van der Waals interactions (which often play important roles in the assembling process) in rigid molecules such as fused aromatic rings not only have preferred orientations but also need to be balanced and compromised with other noncovalent forces present in the entire system. Nevertheless, the more closely the solvophobic surfaces can be approximated by simple geometrical shapes, the more predictable the end products will be. This is the case for mesoscale or macroscale self-assembly with components ranging from microns to centimeters in size [50]. At such dimensions, molecular details are no longer important and assembling is driven by capillary forces – the macroscopic manifestation of solvophobic effects in the current case. 2D and even 3D objects with complex structures can be prepared predictably from components with patterned solvophobic and solvophilic surfaces. The assembling process has all the characteristics of the small-molecule counterpart including complementarity in binding surfaces, reversibility, and environmental dependency. Quite interestingly, the concept can be expanded to create mesoscopic "foldamers" with folding motifs resembling those found in natural proteins [51].

3.3
Learning from Synthetic and Biological Polymers

Many biological polymers undergo conformational changes between ordered and disordered states in a highly cooperative manner [52]. Examples include the helix-coil transition in peptides [53] and nucleic acids [54], the beta-sheet to coil transition in peptides [55, 56], and the denaturation of proteins [52] and RNA [57]. Cooperativity can provide a powerful way for supramolecular organization and it is a design principle ideally suited to chain molecules such as foldamers.

Cooperative transitions are characterized by abrupt changes as denaturing conditions such as high temperatures or unfavorable solvents are introduced. The signature of cooperative behavior is a sigmoidal shape in plots of spectroscopic or other conformationally dependent properties as a function of temperature or solvent composition. In contrast, a non-cooperative process only shows gradual changes. Figure 3.1 shows the characteristic behavior of two-state (first order) and one-state (higher order) cooperative transitions. A plot of free energy vs. conformation for a polymer that undergoes a two-state transition shows double minima, separated by a barrier. For such a transition both states are equally populated at the midpoint of the transition. In contrast, a single, broad population distribution and a single minimum in the free energy plot are indicative of one-state behavior. It can be seen from Fig. 3.1 that both types of behavior are capable of producing sigmoidal curves. Thus, it is not possible to distinguish between a one-state or a two-state transition simply on the basis of the appearance of a sig-
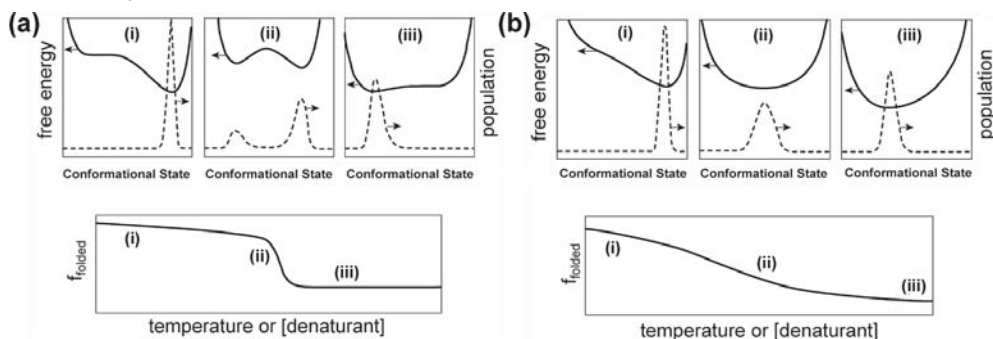
Fig. 3.1 Schematic diagram illustrating (a) two-state and (b) one-state cooperative transitions.

moidal curve. Rather, this distinction must be determined by an analysis of how the population of conformers changes throughout the course of the transition.

Solvents play critical roles in the cooperative behavior of biomolecules. In spite of the fact that the α-helix is the most abundant element of secondary structures in native proteins, many helix-forming sequences are only marginally stable in water [58, 59]. However, solvents such as 2,2,2-trifluoroethanol (TFE) are known to stabilize the α-helical conformation of short peptides [60, 61]. Peptide sequences were reported to undergo cooperative transitions to their α-helical state when TFE was titrated into an aqueous solution [61]. In these cases, helix stability has been shown to follow a linear free energy dependence on TFE composition, analogous to the relationship that is often used to describe the denaturation of proteins by reagents such as urea or guanidinium salts [62].

A change in solvent composition often causes huge conformational changes in synthetic polymers. As the quality of the solvent decreases, flexible homopolymers tend to collapse into globular conformations [63]. Formation of these collapsed conformations is intimately related to the (poor) solvation of the polymer chains. This coil–globule transition is known to be cooperative and is generally described as a second-order process, although simulations have shown that first-order (two-state) transitions are possible for stiff chains [64] and polymers whose segments have long-range attractive potentials [65]. Calorimetric studies on poly(N-isopropylacrylamide) indicate that the coil-globule transition for low molecular weight polymers follows a two-state process [66, 67]. Conformational transitions of this type have thus often been compared to the denaturation of small proteins, a process that also tends to follow a cooperative, two-state transition model between the native and the denatured state [68, 69]. Unlike proteins, however, homopolymers do not generally collapse to a unique conformational state [70]. In fact, it has been shown that the size of the ensemble of compact conformations grows exponentially with chain length [71, 72].

Are there any advantages in using solvophobic and van der Waals interactions to create foldamers with biomolecule-like, cooperative conformational transitions?

An obvious one is related to their strength in aqueous solutions. In fact, water represents the ideal environment for these interactions due to its small size, high cohesive energy, and low polarizability [2–8]. Nevertheless, it seems much better to use strong, selective, and highly directional forces such as hydrogen bonds if one wants to obtain foldamers with highly organized conformations, despite their instability in water.

Lack of directionality in solvophobic and van der Waals interactions may be a huge disadvantage in using them for foldamers if one does not know how to take advantage of this feature. Nondirectionality itself, however, is not a deficiency and can be quite important to the conformational change. A large number of intramolecular interactions need to work together to allow a chain to undergo a cooperative transition to a non-degenerate conformation. Cooperativity asks that these interactions be weak individually but strongly coupled to one another through the conformation and covalent geometry of the foldamer chain. Solvophobic and van der Waals interactions are ideal in this regard, for both are ''collective'' interactions among a large number of solvents or polarizable bonds. In fact, one of the most accepted models for cooperativity in proteins is hydrophobic clustering, which states that hydrophobic collapse, at least in the early stages of folding kinetics, is directly responsible for cooperativity [52].

Because cooperativity demands multiple points of contact, it is easy to imagine highly directionally restricted forces would have little margin for error, whereas weak, non-directional forces tend to be more forgiving. One potential difficulty with highly directional forces (especially strong ones) is that a secondary structural design might quickly become over determined and incapable of yielding the desired conformation. If the ideal geometry of a directional interaction cannot be maintained within a covalently linked backbone, the mismatch would propagate as the chain lengthens. Another issue, probably a most serious one, is related to how well a successfully designed foldamer may serve as a platform to be modified and endowed with functional groups. It is easy to see that conformational motifs are more likely to be retained in foldamers based on ''fuzzy'' interactions such as solvophobic and van der Waals than those based on highly directional forces when different, functional monomers are inserted into the original sequence.

To successfully employ nondirectional interactions in a foldamer, one cannot use backbones with too many degrees of conformational freedom. Flexible homopolymers have rough energy landscapes and highly degenerate native states. Such chains are likely to collapse to a poorly defined compact globule. It seems that a key element for designing a foldamer organized by van der Waals and solvophobic interactions is the molecular contact area per degree of conformational freedom. Maximizing this ratio should help ensure stability of the folded structure as well as reduce degeneracy of the native state. Such a strategy indeed is widely used by nature to simplify the problem of conformational control. For example, amide bonds, which make up one third of the backbone of a natural protein, are rotationally restricted due to conjugation. Polysaccharides are also rigidified by their cyclic monosaccharide units. In the case of solvophobically based foldamers,

rigidity has one additional benefit. When flexible molecules go from the solid phase to a solution phase, flexibility favors the latter by a favorable entropic contribution. Rigid molecules, on the other hand, do not benefit as much in such a transition. Essentially, solvophobicity can be enhanced by the rigidity of a molecule. (Low solubility of rigid molecules is also a result of high crystallinity and/or strong intermolecular forces.)

Another important feature of proteins is amphiphilicity. In order for a solvophobic molecule to be soluble at all, it must be equipped with solvophilic segments. Amphiphilicity, however, is far more important than just providing solubility to biopolymers. Folding of proteins has been studied with the lattice model [70, 73]. In the simplest model, a peptide chain consists only of two types of units, labeled either as hydrophobic (H) or polar (P). These models showed that hydrophobicity was highly restrictive, quite counterintuitively for a nondirectional force. If a ''native'' conformation is defined as one with maximum number of hydrophobic interactions, sequences that can configure only into one or a few native conformations far exceed those that can assume 10 or more native conformations [74]. Although not sufficient to give a single native structure to the peptide chain of a protein, hydrophobicity can constrain the chain into a relatively small number of compact conformations [5]. This distribution of solvophobic and solvophilic segments is part of the connotation we intend for *geometrical manipulation*. In the meso- and macroscale self-assembly, shape, size, and pattern of solvophobic patches are sufficient to yield prescribed final structures [50]. Doing so on the molecular scale is much more challenging, but is a skill chemists have to master in order to approach nature's ability in designing complex structures, especially when secondary folding motifs are integrated to form tertiary and further to quaternary structures.

3.4
Recent Advances in Foldamers Based on Solvophobic Effects

We classify solvophobic foldamers into four categories: (a) foldamers stabilized by adjacent identical aromatic units, (b) foldamers stabilized by adjacent donor–acceptor aromatic units, (c) foldamers stabilized by nonadjacent aromatic units, and (d) foldamers stabilized by aliphatic units. Here, ''adjacent'' and ''nonadjacent'' refer to the relationship of units that contribute significantly to the folding energy (Figs. 3.2a and b). The simplest aromatic foldamer one can envision consists of identical aromatic units tethered by spacers (Fig. 3.2a). In such a structure, the interacting units are separated only by noninteracting spacers. Such a foldamer is considered to be stabilized by adjacent aromatic units even if noninteracting aromatic units may be found in the spacer. The structure in Fig. 3.2b, on the other hand, has the interacting units (**A**–**B** or **C**–**D**) separated by spacers as well as another contributing aromatic unit (**C** or **B**), and is classified as foldamers stabilized by nonadjacent units.
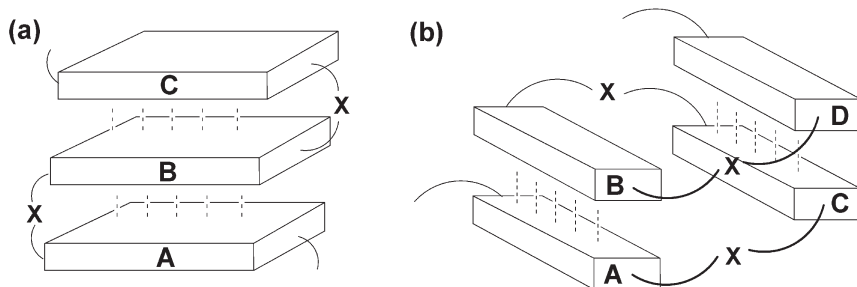
Fig. 3.2 Schematic representation of foldamers stabilized by (a)
adjacent and (b) nonadjacent aromatic units. **A**, **B**, **C**, and **D** represent
interacting aromatic units. **X** is the spacer between the aromatic units.
Dotted lines denote solvophobic, van der Waals, and electrostatic
interactions associated with aromatic interactions.

3.4.1
Foldamers Stabilized by Adjacent, Identical Aromatic Units

From the viewpoint of geometrical manipulation, it is not a good idea to start
designing solvophobic foldamers from flexible aliphatic hydrocarbon chains.
Instead, rigid aromatic rings represent much more reasonable solvophobes with
higher levels of geometrical constraints – effective aromatic interactions are only
obtained in three main geometries: edge–face, offset stacked, and stacked [75–
78]. For aromatic foldamers, geometrical manipulation is achieved through both
the aromatic units and the spacers (Fig. 3.2). Strong van der Waals and solvopho-
bic interactions are obtained with large aromatic units in the stacked or offset
stacked configurations. Length and flexibility of the spacers are also important to
the foldability of the chain.

Some of the earliest aromatic foldamers rely on *cis*-oriented amides, ureas, and
guanidines (made possible by alkyl-substitution on the nitrogen) to form folded
structures [79–84]. Both **7** and **8** adopted folded conformations in the solid state
based on X-ray crystallography. In solution, π–π stacking was confirmed by NOE
and upfield shifts of proton signals in the longer oligomers. The substituent on
the nitrogen is not limited to methyl; bulkier groups such as benzyl and even
(S)-1-(1-naphthyl)ethyl recently were found to promote *cis* conformations in sev-
eral aromatic foldamers (**9**) with naphthalene in the main chain [85]. See Chapter
1 for more details for these and related molecules.

In an effort to synthesize tubular cyclophanes by ring closure of ladder-like
molecules, Vögtle and co-workers prepared oligomers such as **10** [86–92]. Its crys-
tal structure revealed S-shaped conformations with stacked aromatic rings [88].
The authors did not report whether the stacked conformers could unfold under
thermal or solvent-denaturating conditions, but did show these S-shaped confor-
mations were maintained in chloroform.

7 **8** **9a–c**

a, R = benzyl
b, R = (S)-1-phenylethyl
c, R = (S)-1-(1-naphthyl)ethyl

Scheme 3.4

10 **11**
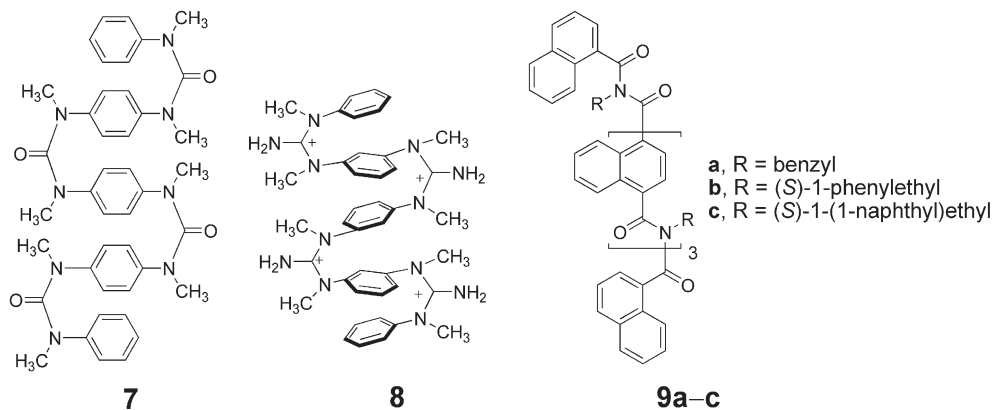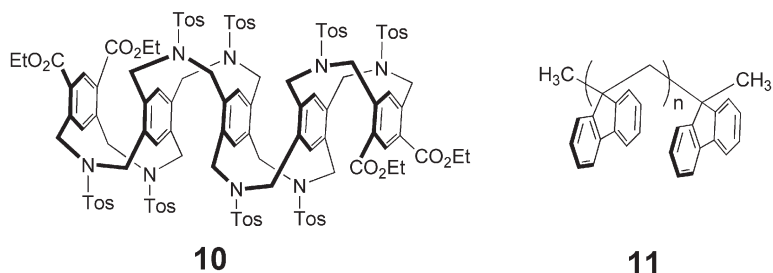
Scheme 3.5

Rathore et al. reported oligofluorenes **11** with simple methylene as the spacer [93]. The difference between this and most other foldamers is that aromatic groups are in the side chains instead of the main chain of the molecule. With the fluorenes closely stacked, oxidation potential decreases from 1.74 V for the monofluorene to 1.14 V for the tetrafluorene derivative. Similar behavior was observed in photoelectron spectroscopy and may be relevant to electron-transport mechanisms in DNAs through π-stacked bases.

Unlike most noncovalent interactions, hydrophobic forces increase in strength with an increase in temperature [2–8]. This feature was exploited by Li et al. in perylene–DNA hybrid foldamer **12** [94–96]. (Because the DNA bases are not involved in base pairing and π–π stacking, the compound is classified as a foldamer stabilized by adjacent aromatic units.) The folding reaction was endothermic ($\Delta H = 2.7, 4.4, 4.8$, and 6.9 kcal mol^{-1} for the folding of two, three, four, and five perylene units) but entropically favored. Consequently, **12** adopted more compact stacking structures at 90 °C than at 20 °C [94]. These foldamers were compared to proteins isolated from thermophilic and hyperthermophilic microorganisms that showed better folding propensity near the boiling point of water. Perylene

units preferred to associate with one another despite the presence of many nucleotide bases present in the structure. The nucleotide bases are less hydrophobic and also do not match perylene in shape and size. Selective association of perylene in the presence of other heterocycles, therefore, is related to the concept of geometrical manipulation mentioned earlier. Perylene bisimide units also were connected with polytetrahydrofuran chains by Janssen and co-workers to form polymers [97]. The polymers were found to form stacked structures in *o*-dichlorobenzene, highlighting the stability of the aromatic interactions (even in aromatic solvents).

12
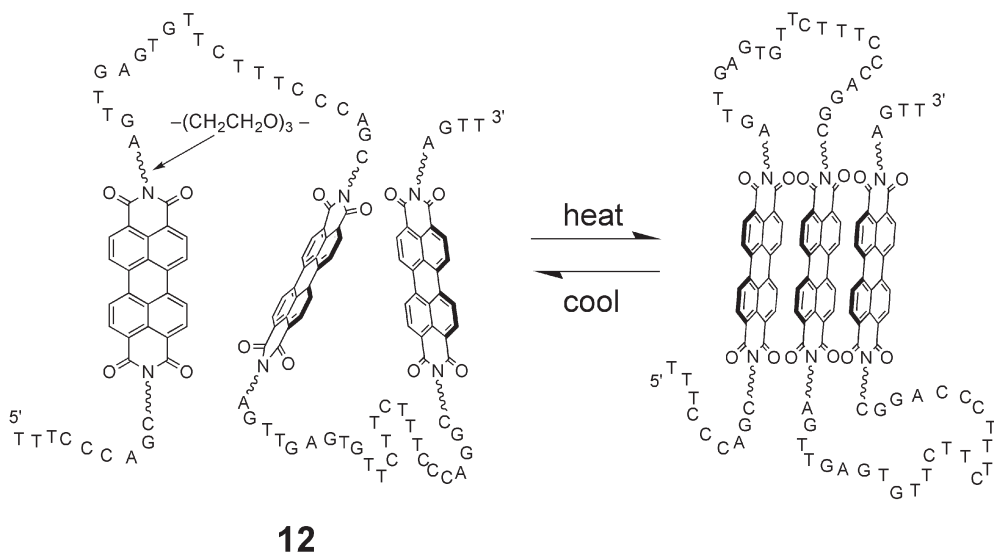
Scheme 3.6

3.4.2
Foldamers Stabilized by Adjacent Donor–acceptor Aromatic Units

One way of enhancing the selectivity in aromatic–aromatic interactions is to polarize the aromatic systems and make one unit rich in electron density and the other deficient. With two different kinds of monomer units in a linear oligomer, different folding motifs may be obtained from the sequence of the monomer units [98]. This highly efficient strategy was pioneered by Iverson and colleagues and the class of foldamers is called *aeda*mers (*a*romatic *e*lectron *d*onor–*a*cceptor oligomers) [99].

1,5-Dialkoxynaphthalene (**13**) and 1,4,5,8-naphthalenetetracarboxylic diimide (**14**) were employed as the donor and the acceptor units in aedamers. For the corresponding monomers, the donor–acceptor association constant (K_a) was nearly

one and two orders of magnitude higher than those for the acceptor–acceptor and donor–donor complexes [100]. Due to different distributions of electron density, the donor and acceptor pair can also have better geometrical overlap, which contribute to stronger solvophobic interactions as well as stronger van der Waals interactions.

13   **14**
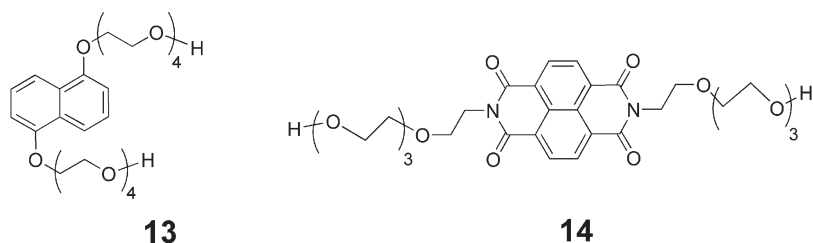
Scheme 3.7

The monomers were derivatized with an amino group on one end and a carboxyl group on the other end. In this way, aedamers not only can be synthesized through standard solid-phase peptide chemistry, but also can incorporate natural amino acids for functionalization and solubility. FMOC chemistry was used to construct oligomers **15a** with different chain lengths using L-aspartic acid as the linker (Fig. 3.3). The folded structure was supported by spectroscopic studies such as UV and ^{1}H NMR spectroscopy. With an increase in the chain length, higher degree of hypochromism was observed, consistent with stacked π systems. Whereas the absorption spectrum of the shorter oligomer ($n = 1$) was similar to that of a 1:1 mixture of the monomeric donor and acceptor, those of the longer oligomers ($n = 2, 3$) differed increasingly with respect to intensity and shape. Additional evidence for the stacked aromatic rings came from NMR studies, which
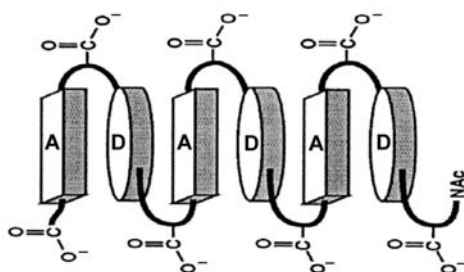
Fig. 3.3 Schematic representation of aedamer-type foldamers. (Reprinted with permission from Ref 102. Copyright 2000, American Chemical Society, Washington, DC.)

15a, R = COO⁻; **15b**, R = *i*-Pr
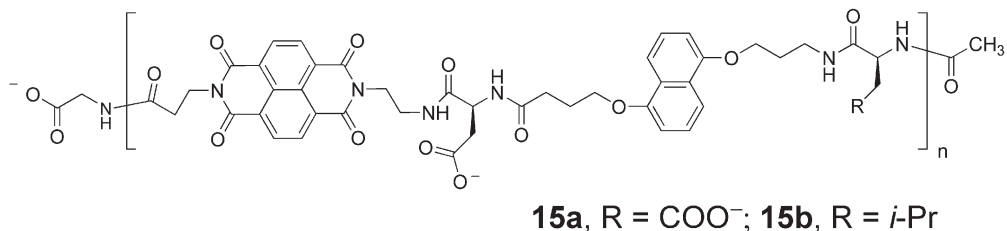
Scheme 3.8

showed significant upfield shifts of aromatic protons, NOE signals between protons on adjacent aromatic units, and diastereotopic methylene hydrogens possibly resulted from restricted rotation due to intramolecular association [99].

Aedamer **15b** was synthesized to mimic leucine zippers found in peptide systems [101]. However, aromatic interactions were complicated by the presence of other hydrophobic units (i.e. leucine). Instead of adopting the desired leucine zipper, **16** folded as typical aedamers at low temperatures. Upon heating to 80 °C, an irreversible conformational change happened as a result of intermolecular association. The red wine colored solution turned into a pale, gel-like solution, apparently due to loss of aromatic stacking. With the visible color change, **15b** could be potentially useful as a thermal sensor indicating onset of a threshold temperature.

In another study, the authors undertook an extensive investigation in the effect of tethering spacers on the conformation of dimeric model systems [102]. Folded structures were found to dominate with both rigid and flexible spacers varying in length, reflecting the robustness of intramolecular aromatic interactions in aqueous solutions. On the other hand, since the spacer could alter the orientation of the donor and acceptor units, it could be used to deliberately put the aromatic rings out of contact of each other. The result was that aromatic units had to aggregate intermolecularly. For example, the charge-transfer (CT) band and hypochromism of **16** decreased dramatically at high temperatures (80 °C), in contrast to all the other folded dimers. Another conclusion from this study was that the stacked conformations of aedamers were highly degenerate, reasonable for molecules with fairly flexible spacers.

16
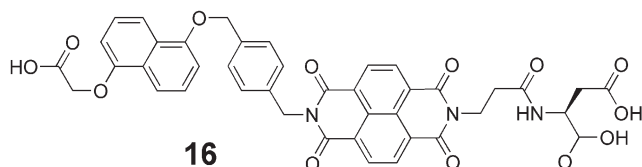
Scheme 3.9

With binding most preferably taking place between donors and acceptors, aromatic associations could be exploited to create double-stranded foldamers from an acceptor strand and a donor strand (see Fig. 3.4.19 in Chapter 4) [103]. (Strictly speaking, double-stranded foldamers are stabilized by nonadjacent units. These examples are included together with other aedamers for the convenience of discussion.) Despite strong repulsions between the two negatively charged chains, K_a increased steadily from 1.3×10^2 M^{-1} for strands containing only a single aromatic unit to 3.5×10^5 M^{-1} for those with four. A high degree of discrimination existed in the binding process, as an excess of either the donor or the acceptor strand ($n = 4$) migrated separately from the complex during polyacrylamide gel electrophoresis.

In addition to forming a duplex with a donor strand, the acceptor strand could do so with a DNA strand. In aedamers **17–19,** several lysine residues were introduced to favor electrostatic interactions between the aedamer and DNA. Compounds **17** and **18** were identified from a library of dimeric derivatives to bind DNA with interesting specificity [104]. The former intercalated DNA with the -Gly$_3$-Lys- linker in the major groove [105], while the latter with the -Ala$_3$-Lys- linker in the minor groove [106]. On the basis of these earlier findings, tetramer **19** was designed to bind sequentially in the minor groove, major groove, minor groove, in a manner similar to how a snake might try to climb a ladder (Fig. 3.4). The binding pattern was confirmed by titration studies and NMR spectroscopy [107]. In the future, these molecules may offer significant opportunities in binding long strands of DNA with sequence specificity.
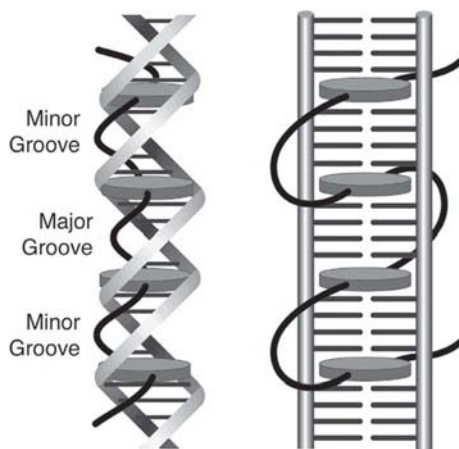
Fig. 3.4 Threading tetraintercalator that binds DNA. (Reprinted with permission from Ref 107. Copyright 2004, American Chemical Society, Washington, DC.)

17

18

19
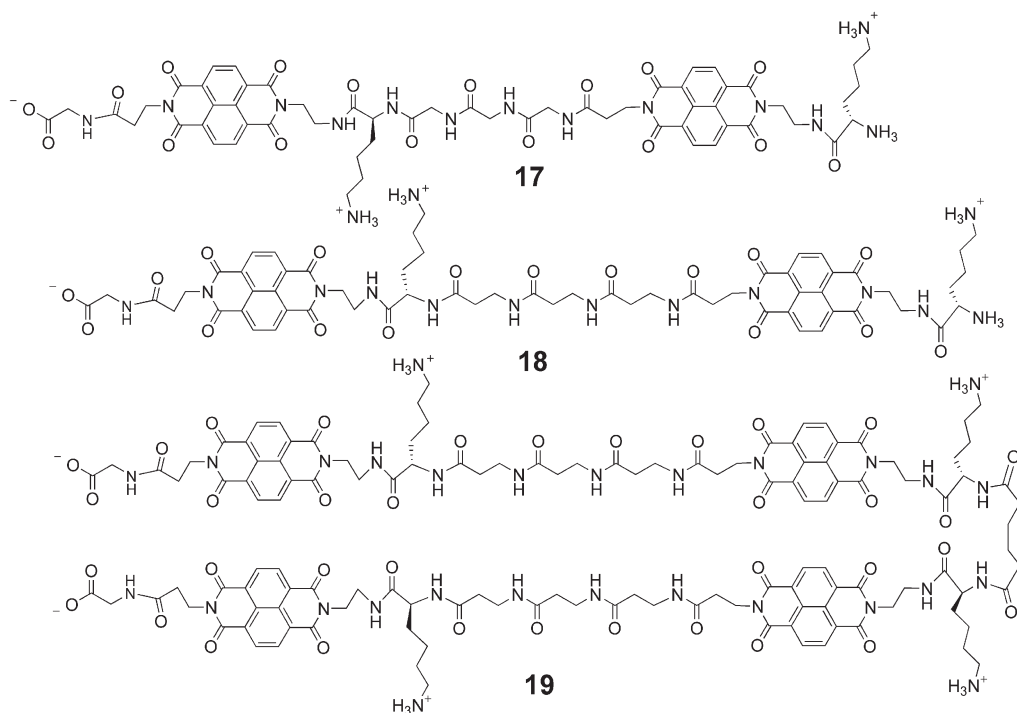
Scheme 3.10

The donor–acceptor motif was also used by Ramakrishnan *et al.* in other systems [108, 109]. Polymer **20** (M_n = 30 000–50 000, PDI ≈ 2) had much longer spacers in between the aromatic units compared to those in most other foldamers discussed so far. Polymers with shorter spacers (x = 3) folded better than those with longer ones (x = 4 and 5). Folding of the latter (**20b** and **20c**), however, could be facilitated by the addition of methanol to enhance solvophobic interactions and van der Waals interactions, or with alkali metal ions to contract the oligo(ethylene glycol) tethers.

20a, x = 3; **20b**, x = 4; **20c**, x = 5;
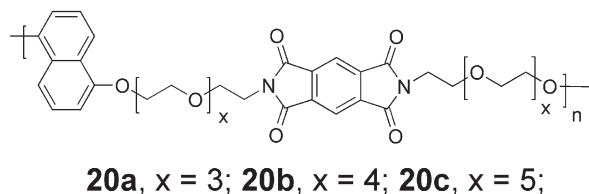
Scheme 3.11

3.4.3
Foldamers Stabilized by Nonadjacent Aromatic Units

In foldamers **7–11**, stacked conformations are either predetermined by the conformations of the linkages or highly favored by the short spacers between the aromatic units. They are unlikely to "unfold" under reasonable conditions. This is in direct contrast to most biofoldamers that are characterized by dynamic conformational behavior. To achieve folding–unfolding reversibility, one must introduce some flexibility in the structure, most likely at the spacers. This was the case in aedamers, which could be "denatured" by cationic surfactant cetyltrimethylammonium bromide (CTAB) [99]. Nevertheless, the folding motif until now is limited to columns of stacked aromatic rings. With flexible spacers, it is difficult to imagine any folding motifs other than the stacked columns illustrated in Fig. 3.2a, whether identical or different aromatic units are involved. In order to obtain foldamers stabilized by nonadjacent units, one must constrain the chain in a way to avoid association of neighboring units.

Such a strategy was successfully employed by Moore and co-workers in the *m*-phenylene ethynylene (*m*PE) foldamers (**22**) [110]. These foldamers were inspired by their discovery that macrocycle **21** self-associated by face-to-face π-stacking interactions [111]. Of the features important to the design, the most critical was the utilization of a semirigid aromatic backbone. As mentioned previously, semirigidity is a strategy universally adopted by nature to simplify the conformational problem. With a 120° angle created by the *meta*-connectivity and linear ethynylene spacers, an *m*PE oligomer was geometrically poised to fold upon itself, forming a conformer resembling macrocycle **21** (Fig. 3.5). Tri(ethylene glycol) (Tg) was
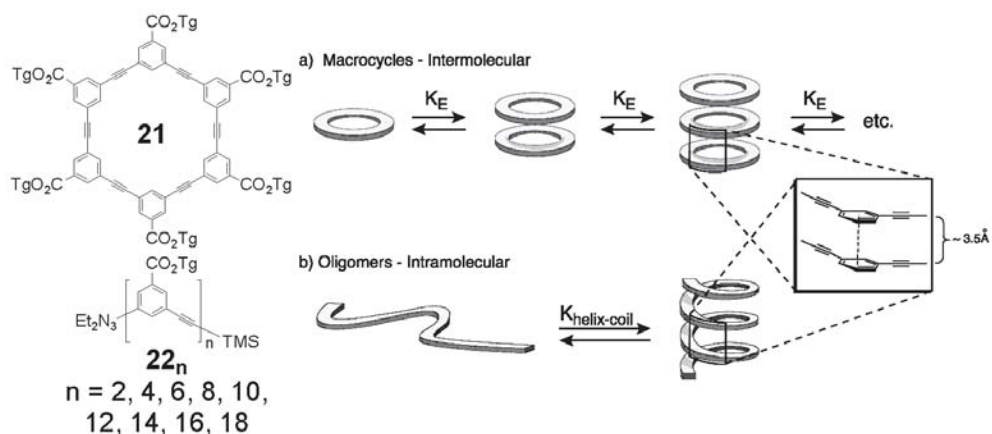
Fig. 3.5 Relationship between intermolecular aggregation of *m*PE macrocyles and the intramolecular folding of linear *m*PE oligomers. (Reprinted with permission from Ref 112. Copyright 2006, American Chemical Society, Washington, DC.)

chosen as the side chain because of its good solubility in polar solvents promoting aromatic interactions. Placement of the side chains at the periphery of the envisaged helix creates a hydrophilic shell around the hydrophobic core, a feature universal in water-soluble proteins.

Unlike α-helices found in proteins, synthetic helices do not have known spectroscopic signatures to allow their quick characterization. Since six units were expected to make one turn in **22**, a natural anticipation was that a critical chain length should exist. For example, in order to benefit from any intramolecular π-stacking interactions, the chain should be longer than the hexamer. Indeed, under dilute concentrations (e.g. 10 μM) at which intermolecular aggregation was minimized, significant upfield shifts of proton signals occurred abruptly in acetonitrile for 22_n with $n > 8$. In chloroform, a solvent that weakened aromatic interactions, the proton signals were essentially independent of chain lengths [110].

This kind of chain-length dependence test (CLDT) turned out extremely useful for the characterization of *m*PE foldamers [112]. When the percentage of acetonitrile was gradually increased in a mixture of acetonitrile and chloroform, absorption for 23_n at 289 nm increased while that at 303 nm decreased (Fig. 3.6a, left). These changes were attributed to a shift of the equilibrium toward the *cisoid* conformers [113]. A plot of the ratio of absorbances (A_{303}/A_{289}) indicated a distinct transition in acetonitrile with n > 8 (Fig. 3.6a, right). Consistent with the folded, stacked conformation, a broad, excimer-like emission band shifted to the red in acetonitrile and replaced the sharp emission at about 350 nm for oligomers longer than the 10-mer (Fig. 3.6b). Such a chain-length-dependency again was absent in chloroform. When the side chains were made chiral, induced circular dichroism (CD) signals were observed in acetonitrile for 24_n with n > 8 (Fig. 3.6c), consistent with helical conformations that were biased in handedness by the asymmetric side chains [114]. During solvent titration experiments, these spectroscopic changes were found to display a sigmoidal relationship with the solvent composition, reminiscent of denaturation curves used to determine thermodynamic stabilities of proteins. The data fit well to a two-state model, in agreement with the proposed helix–coil transition [113]. Spin-labeling experiments later confirmed that six monomers made up one turn in the *m*PE foldamers [115].

Quite surprisingly, the folded state of the *m*PE foldamers was found to be stable in a range of polar and nonpolar solvents including ethyl acetate, DMSO, acetonitrile, methanol, TFE, and even reasonably well in nonpolar, polarizable solvents such as carbon tetrachloride [116]. The only solvents that promoted unfolding of the helices were chlorinated solvents such as chloroform, methylene chloride, and 1,2-dichloroethane. When the Tg side chains were replaced with nonpolar alkyl groups, the resulting foldamers folded well in heptane [117]. Apparently, strong intramolecular interactions were present in the *m*PE foldamers. It is unclear whether van der Waals or solvophobic interactions play the dominant role in the unusual stability of the helix. Solvophobic contributions may be quite substantial, as *m*PE derivatives in general have very poor solubility in most sol-
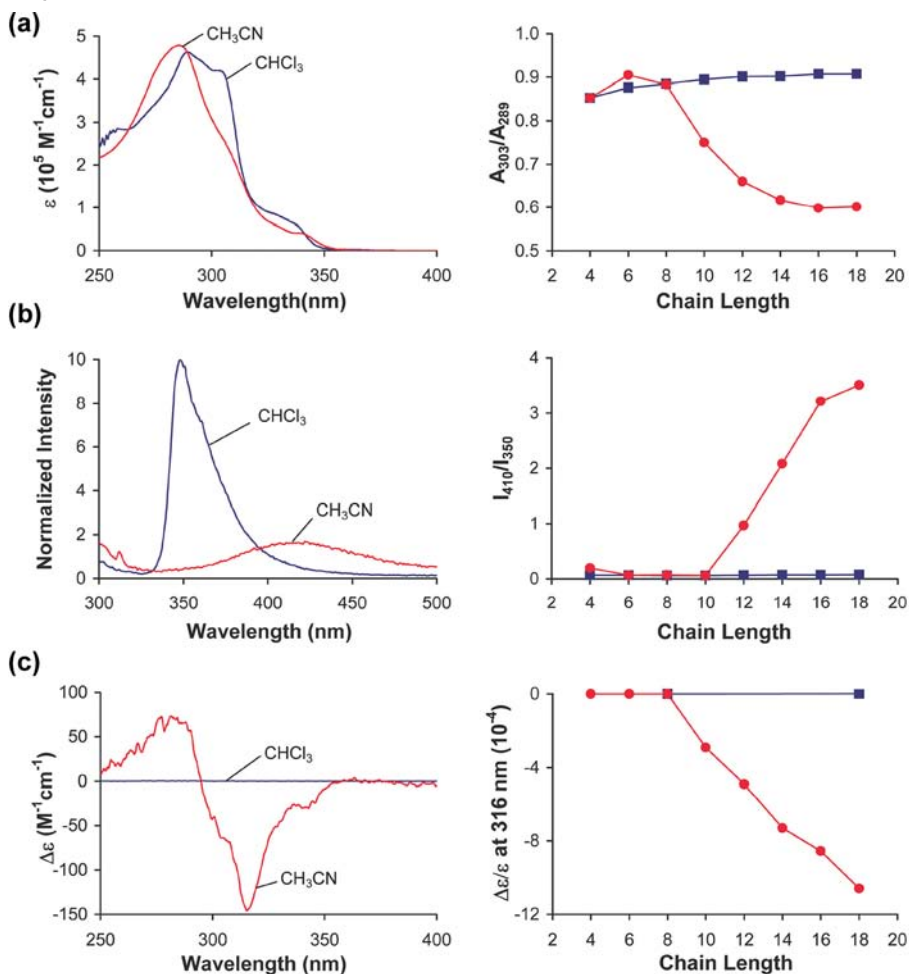
Fig. 3.6 (a) UV spectra of **23₁₈** (left) and the ratio of absorbances at 303 and 289 nm for **23₄–23₁₈** (right). (b) Fluorescence spectra of **23₁₈** (left) and the ratio of fluorescence intensities at 410 and 350 nm for **23₄–23₁₈** (right). (c) CD spectra of **24₁₈** (right) and anisotropy factor (Δε/ε) at 315 nm for **24₄–** **24₁₈** (right). Data collected in CHCl₃ are indicated by blue squares, and those collected in CH₃CN are indicated by red circles. (Reprinted with permission from Ref 112. Copyright 2006, American Chemical Society, Washington, DC.)

vents unless bulky or flexible groups are attached. As mentioned before, although partly due to strong solute–solute interactions, low solubility may also result from poor solvation of a rigid framework that does not gain as much entropy as a flexible one during dissolution.

23$_n$, R = H
24$_n$, R = CH$_3$
n = 2, 4, 6, 8, 10,
12, 14, 16, 18
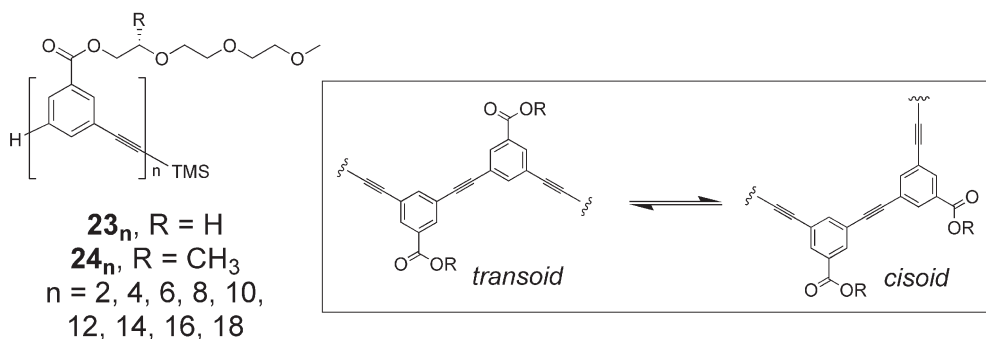
Scheme 3.12

Interestingly, intramolecular photocrosslinking of the helical conformer was demonstrated by Hecht and co-workers as an approach to organic nanotubes [118]. More recently, photoirradiation was employed by the same group to switch the helical conformation of *m*PE foldamers with a central azobenzene unit [119]. This concept has potential applications in smart molecular delivery vehicles.

Upon folding, **23$_n$** forms an internal cavity with a diameter estimated to be 8.7 Å according to molecular modeling. (See Chapter 7 for more detailed discussions on the usage of foldamers for molecular recognition.) In a polar solvent mixture (i.e. 40% water in acetonitrile), the cavity is filled with small solvent molecules that are eagerly waiting to be displaced. Thus, **23$_{12}$** could bind hydrophobic molecules such as monoterpenes with binding energies ($-\Delta G$) in the range of 4–5 kcal/mol [120]. If the chain length is increased, the cavity takes a tubular shape and should prefer a rod-like guest such as **25** [121]. Since burial of poorly solvated surfaces is the driving force for solvophobically driven molecular association, the natural expectation is that tightest binding would occur when the binding cavity of the host molecule and the guest match in terms of size and shape, so that minimum solvophobic surface is exposed. This selectivity in binding derives from geometrical matching of the host and the guest, and is a manifestation of geometrical manipulation mentioned earlier.

25

26
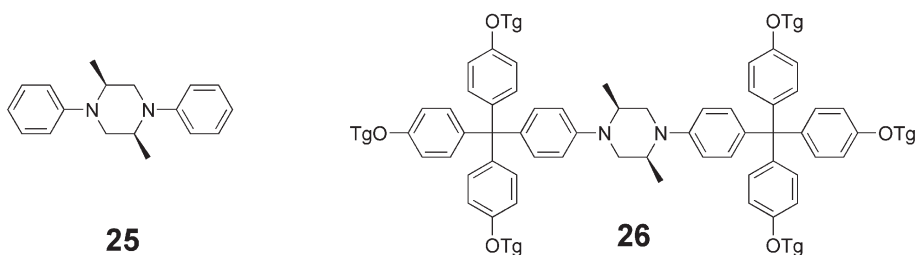
Scheme 3.13
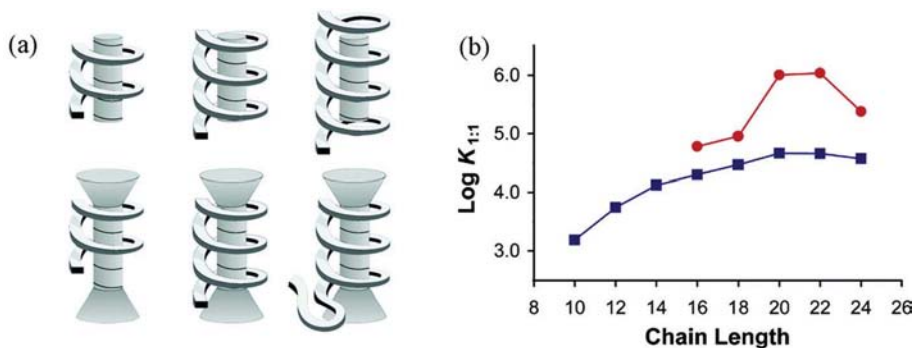
Fig. 3.7 (a) Schematic diagram showing increasing lengths of helical *m*PE oligomer complexed with the rodlike guest **25** (top) and capped guest **26** (bottom). (b) Log of the association constant of guests **25** (blue squares) and **26** (red circles) with series **23** oligomers plotted as a function of the length. Values determined in 40% water in CH_3CN. (Reprinted with permission from Ref 112. Copyright 2006, American Chemical Society, Washington, DC.)

CD titration studies confirmed the hypothesis [121]. As seen in Fig. 3.7b, binding free energies increased gradually from the 10-mer through the 20-mer. The 1:1 binding constants of 20-mer and 22-mer were about 30 times larger than that of the 10-mer. Interestingly, there was a small but experimentally reproducible reduction in binding with the 24-mer, in line with the slight mismatch between the dimensions of the host and the guest. To increase selectivity further, Moore and colleagues synthesized guest **26**, which was capped with two large triarylmethyl groups. Even without the side chain, the diameter (ca. 10.2 Å) of the capping group exceeded the cavity width. Selectivity was indeed much higher. As can be seen in Fig. 3.8b, peaking at $n = 20$ and 22 was much more pronounced for the capped **26**. Also, compared to that of the uncapped guest, there was an overall increase in the binding of the capped guest. This was attributed to favorable aromatic interactions between the triarylmethyl caps and the end of the folded oligomer, which served to bury additional solvophobic surfaces. Another interesting discovery was that binding was essentially complete for $23_{20} \cdot 25$ by the time the first measurement was made at 60 s after mixing, but took more than 1000 s for $23_{20} \cdot 26$ to reach equilibrium. The difference in binding kinetics most likely happened because the folded oligomer had to unfold in order to bind the dumbbell-shaped guest (Fig. 3.7a).

The methylated series 27_n was found to have higher folding stabilities in solution than the parent series 23_n of the same chain length [122]. With methyl groups located in the interior of the helix, fewer solvent molecules could enter the cavity and solvophobic surface area was reduced. Smaller internal cavity of 27_n was also supported by its binding properties: Under identical conditions, 27_{12} bound monoterpene guests with a binding constant two orders of magnitude lower than that of 23_{12} [120]. Solvents apparently were essential to the folded he-

lix in **23**$_n$ because X-ray powder diffraction studies indicated that, in the liquid crystalline state, it packed into lamellar phases with the interlayer spacing linearly related to chain length. Methylated **27**$_n$, on the other hand, had the *d* spacing determined from diffraction independent of chain length, supporting columnar structures obtained from close packing of the helices [123].
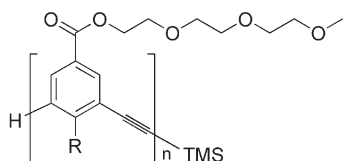
23$_n$, R = H
27$_n$, R = CH$_3$
n = 8, 10, 12, 14, 16, 18

Scheme 3.14

Functionalization of the helix interior can influence binding strength and selectivity. With internal cyano groups, *m*PE foldamers could bind metal ions such as Ag(I) [124]. Oligo(*m*-ethynylpyridine) was shown by Inouye and colleagues to fold in the presence of certain monosaccharides [125]. The state of ionization was found later to strongly influence the propensity of these oligomers to bind saccharides and fold [126]. Details can be found in Chapter 7 of this book.

Folding is not limited to *meta*-connectivity. When three *ortho*-linked phenylene ethynylene (*o*PE) oligomers were studied by Tew and colleagues using ^{1}H NMR spectroscopy, oligomers with as few as four repeat units (**28–30**) were found to adopt folded conformations in acetonitrile [127]. Folding was partly enabled by the 60° internal angle given by the *ortho* linkage, which made it possible to create a turn with only three repeat units. Electronic effects were also important in these structures, as the oligomer (i.e. **30**) consisting of electron-deficient rings and the one (i.e. **29**) with both electron-deficient and electron-rich rings folded better than **28**, which only contained electron-rich aromatic rings. *o*PE polymers were synthesized and studied by Khan and Hecht [128]. Optical spectroscopy indicated transition between extended transoid and helical cisoid conformations induced by solvents, providing evidence for solvophobically driven folding in the *o*PE backbone.

Ethynylhelicene oligomers **31**$_n$ were studied by Yamaguchi et al. using CD and NMR spectroscopy [129]. Helical conformations were identified in oligomers with $n > 7$. Solvents had a tremendous effect on the kinetics of the conformational change, with the rate of unfolding correlating well with the polarizability of aromatic solvents. VPO studies indicated the folded helix to be dimeric and the unfolded coil monomeric. Due to different aggregations in the two conformers, the helical conformation could also be promoted by higher concentration of **31**$_n$.

28 **29**

30

Scheme 3.15

31$_n$, n = 2–9 **32**
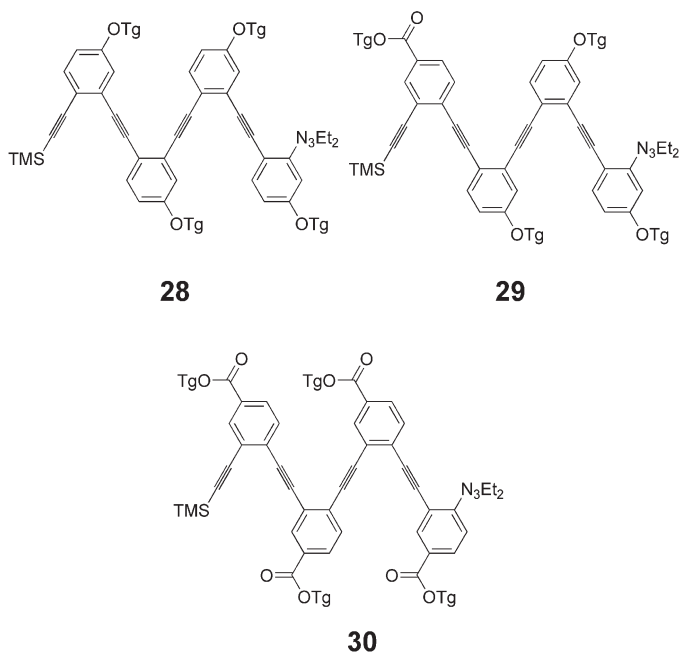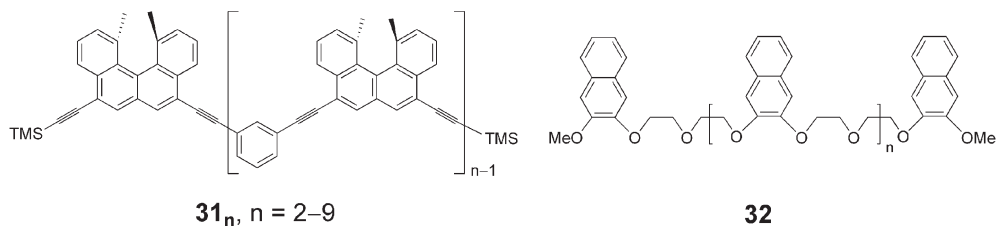
Scheme 3.16

Rigid linkers are not always needed to construct foldamers stabilized by non-adjacent aromatic units. Oligonaphthalene **32** prepared by Li, Chen, and others displayed hypochromism in acetonitrile, a solvent that promoted aromatic association, but not in chloroform, a solvent that disrupted aromatic association [130]. Stacked conformation was also supported by upfield-shifted proton signals and appearance of an excimer-like emission band. When the fraction of acetonitrile was increased in a mixture with chloroform, sigmoidal curves characteristic of cooperativity were observed for these foldamers. Interestingly, upon folding, **32** formed a central cavity resembling crown ethers or kryptands and could be used to bind ethane-1,2-diammonium salt in a 1:1 ratio (see also Chapter 7).

The same group also reported **33**, which had the donor and acceptor units in two blocks instead of in an alternating order [131]. These oligomers folded like a molecular zipper (see also Chapter 4). The folded conformer was quite stable, dominant in both relatively nonpolar solvents (chloroform with 0.5% CF$_3$COOH) and polar aprotic solvents such as DMSO and DMF. For the longer oligomers ($m = n = 2$), the orange color from the intramolecular CT complexation was maintained at up to 150 °C in DMF.

33
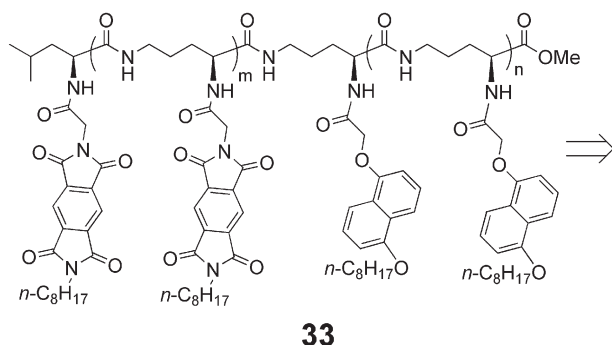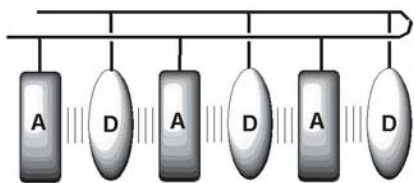
Scheme 3.17

Scheme 3.18

Self-association can occur in any solvophobic foldamers at high enough concentrations and is often a problem encountered during characterization of conformations. In general, ill-defined aggregates are formed during such a process and it is quite difficult to limit self-association to a specific stage such as dimerization. An elegant approach toward duplex foldamers was recently reported by Furusho, Yashima, and colleagues in *meta*-linked oligoresorcinols **34** [132]. When the **5merH** was crystallized from a hydrophobic solvent mixture, chloroform/acetonitrile, it formed a single helix in the solid state (Fig. 3.8). Unlike the *m*PE oligomers, the aromatic units in *meta*-linked oligoresorcinols cannot stack upon one another due to steric congestion. It was clear from the crystal structure that a large surface of hydrocarbon was exposed in the single helix and was in contact with solvent molecules. When the compound was crystallized from
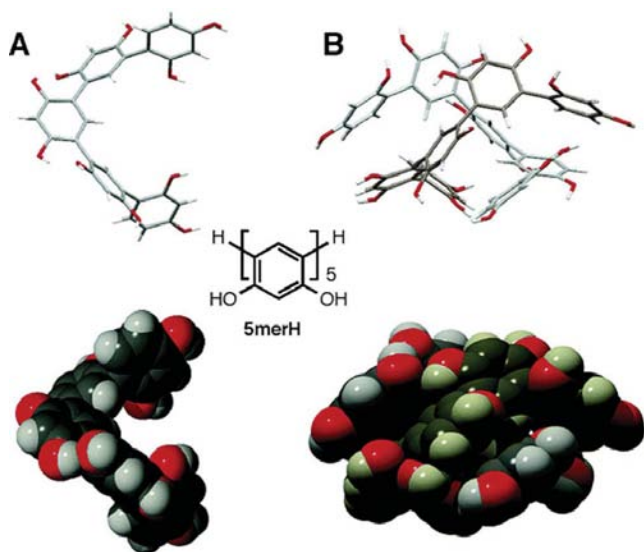
Fig. 3.8 (A) Single helical structure of **5merH**. (B) Double helical structure of **5merH**. (Reprinted with permission from Ref 132. Copyright 2006, American Chemical Society, Washington, DC.)

water, nonpolar solvents were no longer available, and unfavorable exposure of hydrophobic surface was avoided by forming a double helix. Geometrical manipulation through rigid and amphiphilic backbone once again demonstrates its power in limiting self-association to a selective, specific process. The double helical structure was maintained in aqueous solution as shown by UV, NMR, and CD spectroscopy.

34
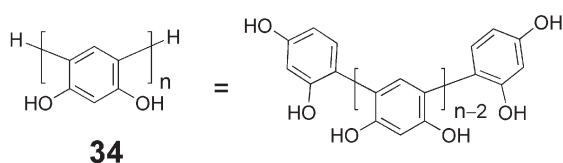
Scheme 3.19

3.4.4
Foldamers Stabilized by Aliphatic Units

It is difficult to use aliphatic solvophobes to construct well defined structures for several reasons. First, their association generally does not have highly preferred orientations as in that of aromatic groups and, hence, puts little geometrical con-

straint on the process of aggregation. The aggregates often vary greatly both in terms of number of molecules and their relative arrangement within in the structure. Second, the most widely used aliphatic solvophobes are flexible hydrocarbon chains. Geometrical manipulation of such objects is not as straightforward as in flat aromatic hydrocarbons. Third, without a well defined shape, precise arrangement of solvophobic and solvophilic groups and creation of amphiphilic pattern (to control aggregation) in aliphatic solvophobes are challenging. Fourth, aliphatic–aliphatic interactions originate from solvophobic interactions and dispersive forces, but aromatic–aromatic interactions have additional contributions such as electrostatic and quadrupolar interactions.

Zhao and co-workers recently described oligomeric cholates derived from cholic acid **35** [133]. Cholic acid has a unique structure. As a metabolite of cholesterol, it is quite rigid with four fused rings. Its rigid backbone is very attractive from the standpoint of geometrical manipulation. It has built-in amphiphilicity with the hydroxylated *α* face and the hydrocarbon-containing *β* face. With a distinctive shape and facial amphiphilicity, its aggregation is much more selective compared to most aliphatic amphiphiles. Unlike conventional head–tail surfactants, sodium cholate tends to form dimers at early stages of aggregation and form large aggregates mostly at relatively high concentrations in water [134]. Another distinguishing feature of cholic acid is its size. The distance between the carboxyl tail and the hydroxyl group at C-3 is about 1.4 nm. A large monomer unit will not only improve the efficiency of synthesis dramatically, but also provide a strong solvophobic driving force, as the strength of solvophobic interaction is directly proportional to the buried solvophobic area. Therefore, most of the problems mentioned above for aliphatic solvophobes are absent in this natural product.

35, Cholic acid
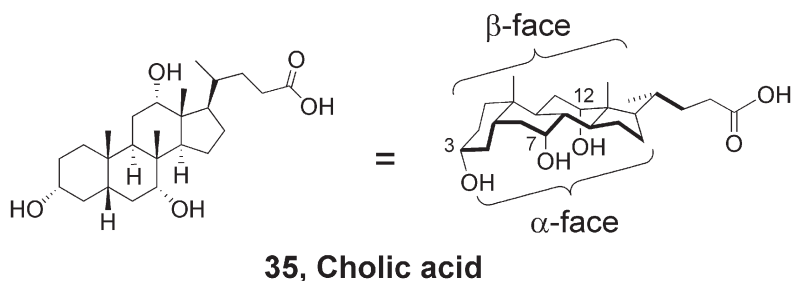
Scheme 3.20

Most of the methods employed in the characterization of aromatic foldamers including hypochromism, excimer formation, and upfield-shifted proton signals could not be used in aliphatic foldamers. NOE techniques were also excluded due to signal overlapping in the ^{1}H NMR spectroscopy. Fluorescence spectroscopy turned out particularly useful. In a mostly nonpolar mixture (e.g. CCl$_4$ or hexane/ethyl acetate) containing a small amount of polar solvent (e.g. methanol

Fig. 3.9 Space-filling molecular models of the unfolded and folded cholate hexamer.

or DMSO), polar solvents were enriched near the fluorophore in 36_n. In addition, quenching of the emission of 36_n by a hydrophilic quencher **37** became more efficient with an increase in the chain length in the above (mostly nonpolar) solvent mixtures, but was independent of chain length in more polar solvents or by the hydrophobic quencher. The results were consistent with folding of the oligocholates to create a helix with a nanometer-sized internal hydrophilic cavity where polar solvents or polar molecules such as **37** could be bound (Fig. 3.9). Folding was confirmed also by fluorescence resonance energy transfer (FRET), a method (essentially a long-ranged version of NOE) widely used in the conformational characterization of proteins and [135–137]. In fact, for foldamers resembling molten-globular proteins, FRET represents a more powerful method than NOE for characterizing conformations. The most interesting result from FRET was that the hexamer (39_4) had a closer end-to-end distance than either the pentamer
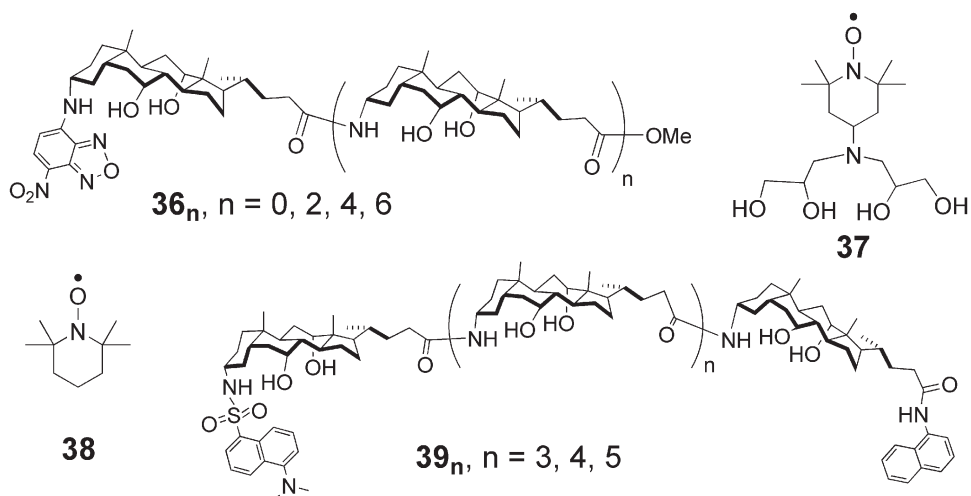
Scheme 3.21

(39_3) or the heptamer (39_5) under folding conditions, as expected from a helix with three monomer units per turn. The data fit well to a two-state transition model, in agreement with the helix–coil transition. The conformational change was extremely sensitive to solvents and could easily detect <0.5% change in solvent composition. High sensitivity toward solvent change was probably due to lack of any other intramolecular interactions besides solvophobic forces. More recently, this highly sensitive, cooperative conformational change was employed to tune the binding affinity between mercury ions and a cholate hexamer with two methionine units incorporated. Simple solvent changes could alter the binding constant over five orders of magnitude [138]. Importantly, the methionine-containing hexamer could fold as well as (actually slightly better than) the parent cholate hexamer. Thus, as expected, functionalized solvophobic foldamers do not deviate (significantly) in foldability from their parent, unfunctionalized versions.

3.5
Conclusions and Outlook

Much progress has been made in the design of foldamers through nondirectional solvophobic and van der Waals interactions. Creating complementary solvophobic surfaces for molecular recognition or the strategy of *geometrical manipulation* seems to be well suited for future endeavors. Nature has perfected this strategy in the folding of peptide chains. Although probably not as rigid as jigsaw puzzles, packing of hydrophobic side chains in the interior of proteins is important to the native conformation and has been proposed to be responsible for the cooperativity observed in conformational transitions of proteins [52, 139]. This may in part explain the difficulty in designing proteins from the primary sequence – it is certainly a lot more challenging to create 3D jigsaw puzzles through forward engineering (the bottom-up approach) than reverse engineering (the top-down approach), especially when the construction material is individual molecules rather than plastics and papers.

Chemists, however, need not be dismayed by this challenge. As illustrated by many examples in the chapter (and in other chapters of this book), advancement in synthetic foldamers does not have to completely parallel the progression of natural foldamers. Whereas most functions are performed at the tertiary and quarternary structural level of proteins, they may be delivered in the secondary structures of synthetic foldamers. Sensing, molecular recognition, and catalysis, although primitive at this point, have already been realized with foldamers (e.g. in aedamers and *m*PE foldamers). Nanometer-sized cavities, typically found at tertiary and quarternary structures of proteins, are obtained in the cholate foldamers prepared in just a few steps from the monomer. Therefore, not limited by a small subset of starting materials such as α amino acids, chemists have the advantage of taking ideas from complex biomolecules in nature and embody them on readily synthesized, water- or organic-compatible molecules, delivering useful functions

at the same time. In the meanwhile, synthetic foldamers may have far greater stability than their natural counterparts, a feature that may be useful in many applications.

The future of foldamer chemistry depends on not only our advancement in the fundamental understanding of conformational control, but also critically on whether superior materials with unique functions can be produced as a result of the fundamental advancement. As mentioned before, solvophobic foldamers have a high tolerance of structural modification and are most likely to retain their conformational properties in the presence of "imperfections" caused by introduction of functional monomers. The most promising approach is probably a marriage between nondirectional (i.e. solvophobic and van der Waals interactions) and directional (e.g. hydrogen bonds and metal–ligand complexation) forces, a relationship benefited by biomolecules. This approach should be important not only to the immediate applications of the currently available solvophobic foldamers, but also to the integration of these structures (and others to be developed) to create higher-order assemblies with sophisticated functions.

References

1 C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, Wiley, Weinheim, **2003**.

2 C. Tanford, *The hydrophobic effect: formation of micelles and biological membranes*, 2nd Ed., John Wiley & Sons, New York, **1980**.

3 A. Ben-Naim, *Hydrophobic interactions*, Plenum Press, New York, London, **1980**.

4 F. H. Stillinger, *Science* **1980**, *209*, 451.

5 K. A. Dill, *Biochemistry* **1990**, *29*, 7133–7155.

6 W. Blokzijl, J. B. F. N. Engberts, *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1545–1579.

7 D. Chandler, *Nature* **2005**, *437*, 640–647.

8 N. T. Southall, K. A. Dill, A. D. J. Haymet. *J. Phys. Chem. B* **2002**, *106*, 521–533.

9 M. Kodaka *J. Phys. Chem. B* **2002**, *105*, 5592–5594, and references therein.

10 A. Marmur, *J. Am. Chem. Soc.* **2000**, *122*, 2120–2121.

11 N. A. M. Besseling, J. Lyklema, *J. Phys. Chem. B* **1997**, *101*, 7604–7611.

12 B. Kronberg, M. Costas, R. J. Silveston, *Dispersion Sci. Technol.* **1994**, *15*, 333–351.

13 N. Muller, *Acc. Chem. Res.* **1990**, *23*, 23–28.

14 P. L. Privalov, S. J. Gill, *Pure Appl. Chem.* **1989**, *61*, 1097–1104.

15 W. Kauzmann, In *Advances in protein chemistry*; C. B. Anfinsen, Jr., M. L. Anson, K. Bailey, J. T. Edsall, Eds., Academic Press, New York, London, **1959**, Vol. 6, pp 1–63.

16 M. H. Abraham, *J. Am. Chem. Soc.* **1982**, *104*, 2085–2094.

17 K. Shinoda, *J. Phys. Chem.* **1977**, *81*, 1300–1302.

18 R. D. Cramer, III. *J. Am. Chem. Soc.* **1977**, *99*, 5408–5412.

19 H. S. Frank, M. W. Evans, *J. Chem. Phys.* **1945**, *13*, 507–532.

20 J. A. V. Butler, W. S. Reid, *J. Chem. Soc.* **1936**, 1171–1173.

21 P. Buchanan, N. Aldiwan, A. K. Soper, J. L. Creek, C. A. Koh, *Chem. Phys. Lett.* **2005**, *415*, 89–93, and references therein.

22 R. Ashoka, *Nature* **1971**, *231*, 313–15.

23 M. Yaacobi, A. Ben-Naim, *J. Phys. Chem.* **1974**, *78*, 175–178.

24 K. A. Connors, *Chem. Rev.* **1997**, *97*, 1325–1357, and references therein.

25 G. M. Whitesides, J. P. Mathias, C. T. Seto, *Science* **1991**, *254*, 1312–1319.

26 J. N. Israelachvili, *Intermolecular and Surface Forces with Applications to Colloidal and Biological Systems*, Academic, London, **1985**, pp 249–257.

27 F. S. Bates, M. F. Schulz, A. K. Khandpur, S. Forster, J. H. Rosedale, K. Almdal, K. Mortensen, *Faraday Discuss* **1994**, *98*, 7–18.

28 M. W. Matsen, M. Schick. *Curr. Opin. Colloid Interface Sci.* **1996**, *1*, 329–336.

29 H. Hasegawa, *Curr. Opin. Colloid Interface Sci.* **1998**, *3*, 264–269.

30 D. J. Lohse, N. Hadjichristidis, *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 171–176.

31 R. J. Spontak, P. Alexandridis, *Curr. Opin. Colloid Interface Sci.* **1999**, *4*, 140–146.

32 S. I. Stupp, *Curr. Opin. Colloid Interface Sci.* **1998**, *3*, 20–26.

33 G. Mao, C. K. Ober in *Handbook of Liquid Crystals, Vol. 3*, Wiley-VCH, Weinheim, **1998**.

34 B. Gallot, *Prog. Polym. Sci.* **1996**, *21*, 1035–1088.

35 K. Loos, S. Munoz-Guerra in *Supramolecular Polymers, Chapter 7*, Marcel Dekker, New York, **2000**.

36 M. Lee, B.-K. Cho, W.-C. Zin, *Chem. Rev.* **2001**, *101*, 3869–3892.

37 S. I. Stupp, V. LeBonheur, K. Walker, L. S. Li, K. E. Huggins, M. Keser, A. Amstutz, *Science* **1997**, *276*, 384–389.

38 G. Das, L. Ouali, M. Adrian, B. Baumeister, K. J. Wilkinson, S. Matile, *Angew. Chem. Int. Ed.* **2001**, *40*, 4657–4661.

39 E. R. Zubarev, M. U. Pralle, E. D. Sone, S. I. Stupp, *J. Am. Chem. Soc.* **2001**, *123*, 4105–4106.

40 S. Park, J.-H. Lim, S.-W. Chung, C. A. Mirkin, *Science* **2004**, *303*, 348–351.

41 D. M. Vriezema, J. Hoogboom, K. Velonia, K. Takazawa, P. C. M. Christianen, J. C. Maan, A. E. Rowan, R. J. M. Nolte, *Angew. Chem. Int. Ed.* **2003**, *42*, 772–776.

42 S. I. Stupp, S. Son, L. S. Li, H. C. Lin, M. Keser, *J. Am. Chem. Soc.* **1995**, *117*, 5212–5227.

43 H. Kelker, R. Hatz, *Handbook of Liquid Crystals*, Verlag Chemie, Weinheim, **1980**.

44 S. Chandrasekhar, B. K. Sadashiva, K. A. Suresh, *Pramana* **1977**, *9*, 471–480.

45 V. Percec, A. E. Dulcey, V. S. K. Balagurusamy, Y. Miura, J. Smidrkal, M. Peterca, S. Nummelin, U. Edlund, S. D. Hudson, P. A. Heiney, H. Duan, S. N. Magonov, S. A. Vinogradov, *Nature* **2004**, *430*, 764–768.

46 M. Kellermann, W. Bauer, A. Hirsch, B. Schade, K. Ludwig, C. Böttcher, *Angew. Chem. Int. Ed.* **2004**, *43*, 2959–2962.

47 J. P. Hill, W. Jin, A. Kosaka, T. Fukushima, H. Ichihara, T. Shimomura, K. Ito, T. Hashizume, N. Ishii, T. Aida, *Science* **2004**, *304*, 1481–1483.

48 W.-Y. Yang, J.-H. Ahn, Y.-S. Yoo, N.-K. Oh, M. Lee, *Nat. Mater.* **2005**, *4*, 399–402.

49 W.-Y. Yang, E. Lee, M. Lee, *J. Am. Chem. Soc.* **2006**, *128*, 3484–3485.

50 N. B. Bowden, M. Weck, I. S. Choi, G. M. Whitesides, *Acc. Chem. Res.* **2001**, *34*, 231–238.

51 T. D. Clark, M. Boncheva, J. M. German, M. Weck, G. M. Whitesides, *J. Am. Chem. Soc.* **2002**, *124*, 18–19.

52 H. S. Chan, S. Bromberg, K. A. Dill, *Phil. Trans. R. Soc. Lond. B.* **1995**, *348*, 61–70.

53 A. Chakrabartty, R. L. Baldwin, *Adv. Protein Chem.* **1995**, *46*, 141–176.

54 W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, **1984**.

55 H. L. Schenck, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 4869–4870.

56 T. Kortemme, M. Ramirez-Alvarado, L. Serrano, *Science* **1998**, *281*, 253–256.

57 D. E. Draper, *Trends Biochem. Sci.* **1996**, *21*, 145–149.

58 S. Marqusee, V. H. Robbins, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5286–5290.

59 K. R. Shoemaker, P. S. Kim, D. N. Brems, S. Marqusee, E. J. York, I. M. Chaiken, J. M. Stewart, R. L. Baldwin,

Proc. Natl. Acad. Sci. USA **1985**, *82*, 2349–2353.

60 F. D. Sönnichsen, J. E. Van Eyk, R. S. Hodges, B. D. Sykes, *Biochemistry* **1992**, *31*, 8790–8798.

61 A. Jasanoff, A. R. Fersht, *Biochemistry* **1994**, *33*, 2129–2135.

62 C. N. Pace, *Methods in Enzymology, Vol. 131* (Eds.: C. H. W. Hirs, S. N. Timasheff), Academic Press, New York, **1986**, pp. 266–280.

63 C. Williams, F. Bochard, H. L. Frisch, *Ann. Rev. Phys. Chem., Vol. 32* (Eds.: B. S. Rabinovitch, J. M. Schurr, H. L. Strauss), Annual Reviews, Palo Alto, California, **1981**, pp. 433–451.

64 H. Noguchi, K. Yosikawa, *Chem. Phys. Lett.* **1997**, *278*, 184–188.

65 Y. Zhou, M. Karplus, J. M. Wichert, C. K. Hall, *J. Chem. Phys.* **1997**, *107*, 10691–10708.

66 E. I. Tiktopulo, V. E. Bychkova, J. Ricka, O. B. Ptitsyn, *Macromolecules* **1994**, *27*, 2879–2882.

67 E. I. Tiktopulo, V. N. Uversky, V. B. Lushchik, S. I. Klenin, V. E. Bychkova, O. B. Ptitsyn, *Macromolecules* **1995**, *28*, 7519–7524.

68 R. Lumry, R. Biltonen, J. F. Brandts, *Biopolymers* **1966**, *4*, 917–944.

69 P. L. Privalov, *Adv. Protein Chem.* **1979**, *33*, 167–241.

70 K. A. Dill, S. Bromberg, K. Yue, K. M. Friebig, D. P. Yee, P. D. Thomas, H. S. Chan, *Protein Science* **1995**, *4*, 561–602.

71 H. S. Chan, K. A. Dill, *Macromolecules* **1989**, *22*, 4559–4573.

72 C. J. Camacho, D. Thirumalai, *Phys. Rev. Lett.* **1993**, *71*, 2505–2508.

73 N. Go, *Annu. Rev. Biophys. Eng.* **1983**, *12*, 183–210.

74 K. F. Lau, K. A. Dill, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 638–642.

75 C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Chem. Soc., Perkin Trans. 2*, **2001**, 651–669.

76 E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250.

77 M. L. Waters, *Curr. Opin. Chem. Biol.* **2002**, *6*, 736–741.

78 C. A. Hunter, J. K. M. Sanders, *J. Am. Chem. Soc.* **1990**, *112*, 5525–5534.

79 K. Yamaguchi, G. Matsumura, H. Kagechika, I. Azumaya, Y. Ito, A. Itai, K. Shudo, *J. Am. Chem. Soc.* **1991**, *113*, 5474–5475.

80 A. Tanatani, H. Kagechika, I. Azumaya, K. Yamaguchi, K. Shudo, *Chem. Pharm. Bull.* **1996**, *44*, 1135–1137.

81 A. Tanatani, H. Kagechika, I. Azumaya, R. Fukutomi, Y. Ito, K. Yamaguchi, K. Shudo, *Tetrahedron Lett.* **1997**, *38*, 4425–4428.

82 A. Tanatani, K. Yamaguchi, I. Azumaya, R. Fukutomi, K. Shudo, H. Kagechika, *J. Am. Chem. Soc.* **1998**, *120*, 6433–6442.

83 R. Fukutomi, A. Tanatani, H. Kakuta, N. Tomioka, A. Itai, Y. Hashimoto, K. Shudo, H. Kagechika, *Tetrahedron Lett.* **1998**, *39*, 6475–6478.

84 H. Kagechika, I. Azumaya, A. Tanatani, K. Yamaguchi, K. Shudo, *Tetrahedron Lett.* **1999**, *40*, 3423–3426.

85 H. Masu, M. Sakai, K. Kishikawa, M. Yamamoto, K. Yamaguchi, S. Kohmoto, *J. Org. Chem.* **2005**, *70*, 1423–1431.

86 W. Josten, D. Karbach, M. Nieger, F. Vögtle, K. Hägele, M. Svoboda, M. Przybylski, *Chem. Ber.* **1994**, *127*, 767–777.

87 W. Josten, S. Neumann, F. Vögtle, M. Nieger, K. Hägele, M. Przybylski, F. Beer, K. Müllen, *Chem. Ber.* **1994**, *127*, 2089–2096.

88 S. Breidenbach, S. Ohren, M. Nieger, F. Vögtle, *J. Chem. Soc., Chem. Commun.* **1995**, 1237–1238.

89 S. Breidenbach, S. Ohren, F. Vögtle, *Chem. Eur. J.* **1996**, *2*, 832–837.

90 S. Breidenbach, S. Ohren, R. Herbst-Irmer, S. Kotila, M. Nieger, F. Vögtle, *Liebigs Ann.* **1996**, 2115–2121.

91 W. Boomgaarden, F. Vögtle, M. Nieger, H. Hupfer, *Chem. Eur. J.* **1999**, *5*, 345–355.

92 H. Schwierz, F. Vögtle, *Synthesis* **1999**, *2*, 295–305.

93 R. Rathore, S. H. Abdelwahed, I. A. Guzei, *J. Am. Chem. Soc.* **2003**, *125*, 8712–8713.

94 W. Wang, L.-S. Li, G. Helms, H.-H. Zhou, A. D. Li, *J. Am. Chem. Soc.* **2003**, *125*, 1120–1121.

95 W. Wang, W. Wan, H.-H. Zhou, S. Niu, A. D. Q. Li, *J. Am. Chem. Soc.* **2003**, *125*, 5248–5249.

96 W. Wang, W. Wan, A. Stachiw, A. D. Q. Li, *Biochemistry* **2005**, *44*, 10751–10756.

97 E. E. Neuteboom, S. C. J. Maskers, E. W. Meijer, R. A. J. Janssen, *Macromol. Chem. Phys.* **2004**, *205*, 217–222.

98 G. J. Gabriel, S. Sorey, B. L. Iverson, *J. Am. Chem. Soc.* **2005**, *127*, 2637–2640.

99 R. S. Lokey, B. L. Iverson, *Nature* **1995**, *375*, 303–305.

100 M. S. Cubberley, B. L. Iverson, *J. Am. Chem. Soc.* **2001**, *123*, 7560–7563.

101 J. Q. Nguyen, B. L. Iverson, *J. Am. Chem. Soc.* **1999**, *121*, 2639–2640.

102 A. J. Zych, B. L. Iverson, *J. Am. Chem. Soc.* **2000**, *122*, 8898–8909.

103 G. J. Gabriel, B. L. Iverson, *J. Am. Chem. Soc.* **2002**, *124*, 15174–15175.

104 V. M. Guelev, M. T. Harting, R. S. Lokey, B. L. Iverson, *Chem. Biol.* **2000**, *7*, 1–8.

105 V. Guelev, J. Lee, J. Ward, S. Sorey, D. W. Hoffman, B. L. Iverson, *Chem. Biol.* **2001**, *8*, 415–425.

106 V. Guelev, S. Sorey, D. W. Hoffman, B. L. Iverson, *J. Am. Chem. Soc.* **2002**, *124*, 2864–2865.

107 J. Lee, V. Guelev, S. Sorey, D. W. Hoffman, B. L. Iverson, *J. Am. Chem. Soc.* **2004**, *126*, 14036–14042.

108 S. Ghosh, S. Ramakrishnan, *Macromolecules* **2005**, *38*, 676–686.

109 S. Ghosh, S. Ramakrishnan, *Angew. Chem., Int. Ed.* **2004**, *43*, 3264–3268.

110 J. C. Nelson, J. G. Saven, J. S. Moore, P. G. Wolynes, *Science* **1997**, *277*, 1793–1796.

111 J. S. Moore, *Acc. Chem. Res.* **1997**, *30*, 402–413.

112 M. T. Stone, J. M. Heemstra, J. S. Moore, *Acc. Chem. Res.* **2006**, *39*, 11–20.

113 R. B. Prince, J. G. Saven, P. G. Wolynes, J. S. Moore, *J. Am. Chem. Soc.* **1999**, *121*, 3114–3121.

114 R. B. Prince, L. Brunsveld, E. W. Meijer, J. S. Moore, *Angew. Chem. Int. Ed.* **2000**, *39*, 228–230.

115 K. Matsuda, M. T. Stone, J. S. Moore, *J. Am. Chem. Soc.* **2002**, *124*, 11836–11837.

116 D. J. Hill, J. S. Moore, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5053–5057.

117 L. Brunsveld, R. B. Prince, E. W. Meijer, J. S. Moore, *Org. Lett.* **2000**, *2*, 1525–1528.

118 S. Hecht, A. Khan, *Angew. Chem. Int. Ed.* **2003**, *42*, 6021–6024.

119 A. Khan, C. Kaiser, S. Hecht, *Angew. Chem. Int. Ed.* **2006**, *45*, 1878–1881.

120 R. B. Prince, S. A. Barnes, J. S. Moore, *J. Am. Chem. Soc.* **2000**, *122*, 2758–2762.

121 T. Tanatani, T. S. Hughes, J. S. Moore, *Angew. Chem. Int. Ed.* **2002**, *41*, 325–328.

122 R. B. Prince, Phenylene Ethynylene Foldamers: Cooperative Conformational Transition, Twist Sense Bias, Molecular Recognition Properties, and Solid-State Organization, Ph.D. Thesis, University of Illinois, Urbana, IL, May 2000.

123 P.-J. Prest, R. B. Prince, J. S. Moore, *J. Am. Chem. Soc.* **1999**, *121*, 5933–5939.

124 R. B. Prince, T. Okada, J. S. Moore, *Angew. Chem. Int. Ed.* **1999**, *38*, 233–236.

125 M. Inouye, M. Waki, H. Abe, *J. Am. Chem. Soc.* **2004**, *126*, 2022–2027.

126 H. Abe, N. Masuda, M. Waki, M. Inouye, *J. Am. Chem. Soc.* **2005**, *127*, 16189–16196.

127 T. V. Jones, M. M. Slutsky, R. Laos, T. F. A. de Greef, G. N. Tew, *J. Am. Chem. Soc.* **2005**, *127*, 17235–17240.

128 A. Khan, S. Hecht, *J. Poly. Sci. Part A: Poly. Chem.* **2006**, *44*, 1619–1627.

129 H. Sugiura, Y. Nigorikawa, Y. Saiki, K. Nakamura, M. Yamaguchi, *J. Am. Chem. Soc.* **2004**, *126*, 14858–14864.

130 J.-L. Hou, M.-X. Jia, X.-K. Jiang, Z.-T. Li, G.-J. Chen, *J. Org. Chem.* **2004**, *69*, 6228–6237.

131 X. Zhao, M.-X. Jia, X.-K. Jiang, L.-Z. Wu, Z.-T. Li, G.-J. Chen, *J. Org. Chem.* **2004**, *69*, 270–279.

132 H. Goto, H. Katagiri, Y. Furusho, E. Yashima, *J. Am. Chem. Soc.* **2006**, *128*, 7176–7178.

133 Y. Zhao, Z. Zhong, *J. Am. Chem. Soc.* **2005**, *127*, 17894–17901.

134 *Sterols and Bile Acids* (Eds: H. Danielsson, J. Sjövall), Elsevier, Amsterdam, **1985**.

135 L. Stryer, *Annu. Rev. Biochem.* **1978**, *47*, 819–846.

136 P. R. Selvin, *Methods Enzymol.* **1995**, *246*, 300–334.

137 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy, 2nd Edn*, Kluwer, New York, **1999**, Chapter 13.

138 Y. Zhao, Z. Zhong, *J. Am. Chem. Soc.* **2006**, *128*, 9988–9989.

139 E. I. Shakhnovich, A. W. Finkelstein, *Biopolymers* **1989**, *28*, 1667–1680.

4
Foldamer Hybrids: Defined Supramolecular Structures from Flexible Molecules

Carsten Schmuck and Thomas Rehm

4.1
Introduction

In Nature, the function of a molecule often requires a specific shape and structure which control or even first enable both its physical properties and also its interactions with other molecules. A most illustrative example are the proteins, linear polymers of amino acids (Fig. 4.1) [1]. Per se such a peptide chain is a rather flexible molecule even though some parts of the conformational space are restricted due to the hindered rotation around the central amide bond. Nevertheless, the resulting three-dimensional (3-D) structure controlled by the covalent framework of the linear peptide strand alone is a random coil at best [2]. However, additional secondary, noncovalent interactions such as H-bonds, electrostatic or hydrophobic contacts between even remote parts of the peptide chain can induce a structural ordering. Parts of the protein first fold into specific secondary structures such as α-helices or β-sheets which then further interact with each other until a fully folded protein with a specific 3-D shape is obtained [3–5]. The correct 3-D structure is vital for the function of the protein. For example, in enzymes only the properly folded state forms the correct active site and hence allows the specific binding and controlled transformation of a substrate [6]. However, even more complex structures can be obtained by the further supramolecular interaction of more than one protein molecule. Again, the resulting overall shape of the aggregate determines its properties. In the case of the protein collagen only a properly folded intertwined triple helical structure formed by the mutual interaction of three peptide strands guarantees the formation of linear fibers with a certain rigidity and stiffness which are needed to build up the extracellular matrix [1].

The folding of a molecule and also its supramolecular interactions with another molecule are controlled by weak and reversible noncovalent interactions. Therefore, structure formation is a dynamic equilibrium process which depends on the number and specific nature of these interactions as well as external parameters (e.g. solvent, temperature). For example, stable α-helices are formed from α-amino acids only with chain lengths of approximately more than 10 amino acids.
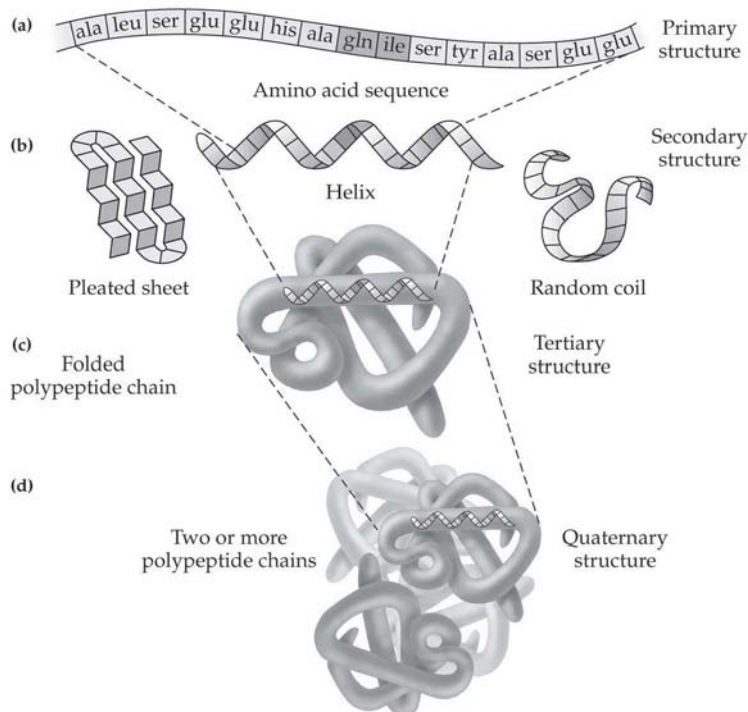
Fig. 4.1 Schematic representation of the different levels of structural order in a protein. The correct function is directly depending on the proper fold of the protein (reprinted with permission; copyright Prentice Hall).

In shorter chains the noncovalent interactions responsible for helix-formation are not strong enough to compensate the unfavorable entropy change associated with the folding of the flexible molecule [3]. Furthermore, the external addition of large concentrations of guanidinium salts or simply heat can reverse the folding of a protein thereby destroying its function. A denatured enzyme does not have any activity any more. Hence, folding is not only vital for the properties of a molecule but the folding (and in consequence everything depending on the fold) can in principle also be externally controlled [1].

Inspired by this overwhelming importance of molecular shape and structure in Nature, chemists have always been interested in designing molecules with specific 3-D structures; some examples are discussed in Chapters 1–3. In principle there are several ways to induce a specific conformation or fold in a molecule (Fig. 4.2): (i) Steric effects in most often rather rigid molecules can be used to induce certain structures and conformations [7]. Examples are the long known aromatic helicenes [8–10] or the shape-persistent aromatic macrocycles introduced in recent years by Moore [11] or Gong [12] for example. (ii) In flexible molecules
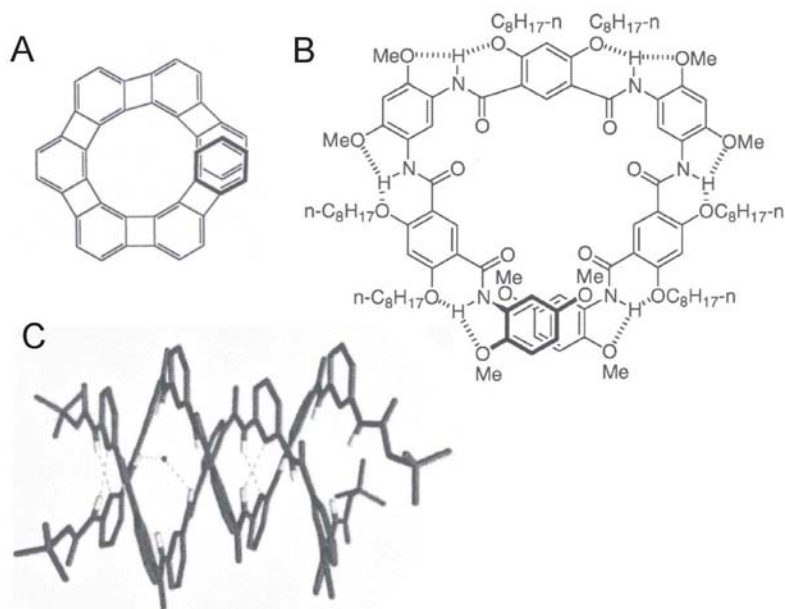
Fig. 4.2 Different possibilities to obtain defined conformations within a molecule: (A) by steric interactions within a rather rigid covalent framework such as in helicenes [8b]; (B) by attractive noncovalent interactions between remote parts of a molecule as in an oligoamide [15a] or (C) by the hybridization induced folding of two or more molecules shown here for a double helix formed from two pyridinedicarboxamide oligomers [33a].

without any built-in biased conformation attractive noncovalent interactions between even remote parts of the molecule can be used to induce a folding as described above for proteins. For such molecules the term "foldamer" has been proposed by Gellman [13]. For example, aromatic polyamides as introduced by Lehn and Huc [33], Gong [14] and Li [15] adopt specific helical shapes due to intramolecular hydrogen bonding and stacking interactions. Flexible oligomers composed of alternating units of electron-rich and electron-poor aromatic as designed by Iverson [55] are another interesting example. (iii) Finally, defined structures can also be achieved not only within one molecule but through a supramolecular interaction of two or more molecules ("hybridization"). The beautiful and fascinating double helical structure of DNA, discovered 1953 by Watson and Crick [16], is one very shining example as well as the structure of collagen already mentioned above. In some cases the individual molecules have a distinct structure even before hybridization more often however at least for artificial systems hybridization induces the 3-D structure of the whole supramolecular assembly while the individual molecules are unstructured.

Whereas examples for the first two approaches can be found in Chapters 1–3, we will describe in this chapter a few instructive examples illustrating the last aspect, the hybridization of foldamers (''foldamer hybrids''). Hence, this chapter deals with *in principle flexible molecules that form supramolecular assemblies with a defined composition and structure.* Of course, we can not cover all work that has been done in this field. Instead we will demonstrate the basic principles and highlight some general aspects using selected recent examples based both on biological (e.g. peptides and nucleic acids) as well as completely artificial foldamers. The choice of examples is subjective and is not intended to question the importance of other contributions not discussed here. We will first concentrate on foldamer hybrids in which the monomers by themselves already have a distinct and well characterized structure (Section 4.2). However, at least in most artificial systems the structure of the underlying monomers is not well defined and the aggregation is the trigger for structure formation. Examples of such hybridization induced folding will be discussed in Section 4.3. The focus in both parts will be on the formation of aggregates with a defined composition such as duplexes or triple helices. The formation of even larger aggregates which unfortunately most often are not really well defined in terms of structure and composition will only be briefly mentioned (Section 4.4). Finally, we will discuss some examples how hybridization can also be exploited to achieve certain functions such as information storage and transfer (Section 4.5).

4.2
Hybridization of Oligomers with Well-defined Structures

4.2.1
Coiled Coils and Helix Bundles

The hybridization of two molecules to form a complex with a defined structure can serve to build a larger suprastructure with additional properties and functions the individual molecules might not possess even if they already do have a defined structure in the absence of their hybridization partner. A classical example for hybridization of structured molecules from Nature is the leucine zipper (Fig. 4.3). Two α-helical peptides of ca. 30 amino acid length with a leucine in every seventh position of the strand dimerize in a parallel orientation [17]. This way the hydrophobic leucine residues can interact with each other, thereby glueing the two helices together via hydrophobic contacts. Even though each helix on its own is structured only the dimerization forms a Y-shaped tweezer which allows the specific interaction with a third molecule in this case DNA. Leucine zippers are a common motif found in DNA transcription factors.

The dimerization, besides forming the binding site needed for the interaction with the DNA, also offers additional means of controlling the interaction via the more pronounced concentration dependence of the dimerization [18]. The leucine zipper is an example of a more general motif found in peptide self-assembly
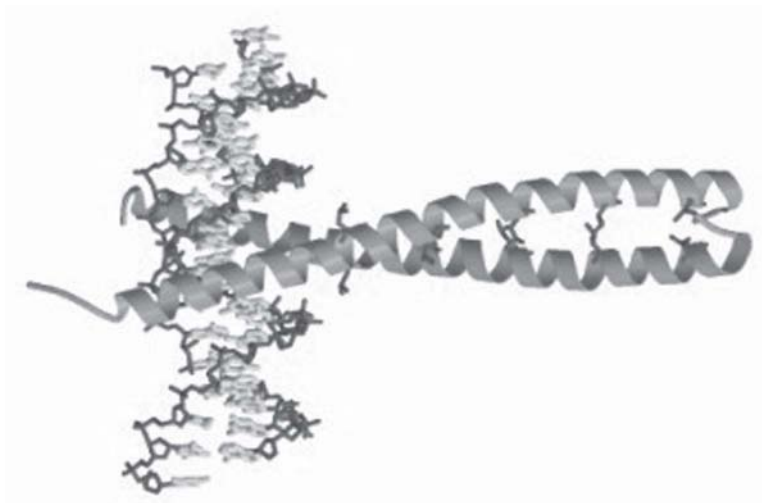
Fig. 4.3 The DNA transcription factor GCN4 as a typical representative
for the leucine-zipper motif. Two α-helices form a Y-shaped dimer, held
together by hydrophobic contacts between opposing leucine residues
(shown to the right).

called coiled-coil motif, in which normally two or three α-helices formed from the
repeat of a heptad with hydrophobic residues in position 1 and 4 interdigitate to
form a stable aggregate [17].

Based on such natural coiled-coil motifs several modifications within artificial
peptides have been introduced to modify the structure or stability of the resulting
aggregates. For example, Tirrell [19] and coworkers incorporated trifluoroleucine
into the leucine zipper protein A1 using *in vivo* expression (Fig. 4.4).

The secondary structures of both the fluorinated and the wild type protein were
identical as determined by CD-spectroscopy (ca. 90% helical) and both proteins
formed stable dimers with $K_{diss} \leq 10$ μM. However, thermal denaturation studies
showed that the fluorination significantly increased the stability of the dimer
($\Delta T_m = +13$ °C, $\Delta\Delta G = 2.4$ kcal mol^{-1}) most likely due to the specific interac-
tions of the fluorinated side chains at the dimer interface within the modified
protein. The same could be shown with a urea titration. The concentration of
urea needed for denaturation of the protein increased within increasing content
of incorporated trifluoroleucine into the protein [20]. The increased stability
due to fluorinated alkyl groups relative to nonfluorinated ones can also nicely be
demonstrated by incorporating semi- or perfluorinated alkyl chains into self-
assembling dendrimers [21, 22]. The higher thermal stability of the correspond-
ing supramolecular system is based on the lower flexibility of a fluorinated chain
and the larger van der Waals radius of fluorine compared to hydrogen. The result-
ing overall increased van der Waals volume of fluorinated chains and their lower
polarizability give rise to an increase in both hydrophobic and lypophobic charac-
ter of a fluorinated molecule relative to a hydrocarbon [23].

H-MRGSHHHHHHGSMASDGL*ENEYA-
-GCL*EREVRSL*EDE AAEL*EQKIEDL*A-
-EIGDL*NNTSGRRPAAKL*N-OH
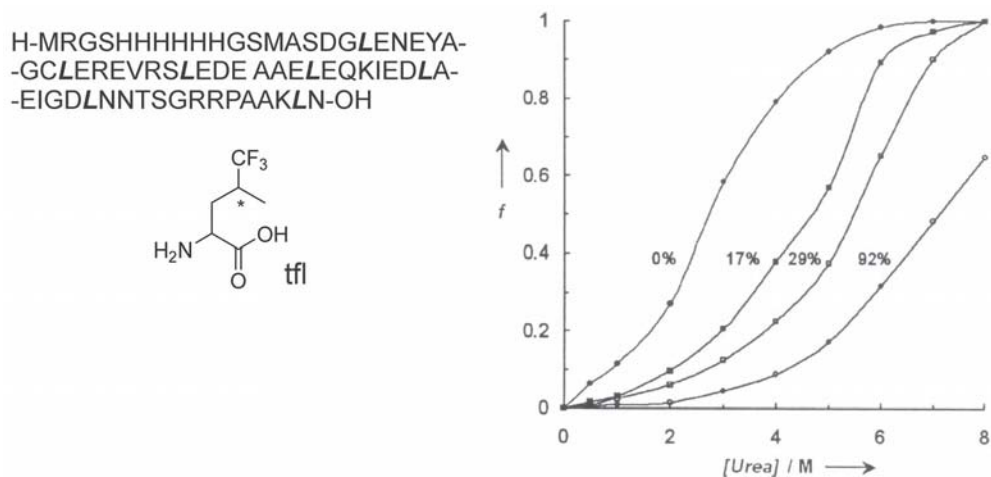
Fig. 4.4 Incorporation of trifluoroleucine (tfl) into the leucine zipper protein A1 at the indicated positions (**L**) increases the thermal stability of the dimer as shown here for a urea titration of wild type A1 (●) against variably fluorinated samples of A1: 17% (■), 29% (□) and 92% (○) incorporated tfl; f represents the ratio of the unfolded state [19].

Similarly, Kennan [24] could increase the stability of a coiled-coil protein trimer derived from GCN4 by replacing a small alanine residue by an unnatural cyclohexylalanine which allows for more extensive hydrophobic contacts at the protein interfaces (Fig. 4.5). The alanine in the natural protein is too small to completely fill the empty space in between the three helices at the trimerization interface. A destabilizing void remains between the three interacting methyl groups of the alanine side chains. Hence, the peptide normally prefers to dimerize instead of forming a trimer, as the dimer interface is more closely packed. However, any increase in hydrophobic contacts at this position should further increase the stability of the trimer thereby shifting the aggregation mode from two to three strands. For example, it had already been shown that hydrophobic ligands such as cyclohexane or benzene present in solution can fulfill that part [25]. In solution these ligands increased the apparent thermal stability of the peptide aggregate and the oligomerization order switched to a trimer. A crystal structure of the peptide shows a single benzene molecule bound directly at the core of the trimer in between the three methyl groups. However, such a ligand with an increased hydrophobic surface can also be directly incorporated into the peptide itself. If in one helix the alanine is replaced by a cyclohexylalanine a 2:1 heterotrimer forms in which the one cyclohexyl residue partially fills the void. This again allows for more efficient and more extensive hydrophobic contacts between the three helices. Hence, the 2:1-heterotrimer is more stable than the homodimer or -trimer of the initial alanine containing peptide. Indeed, CD- and thermal melting
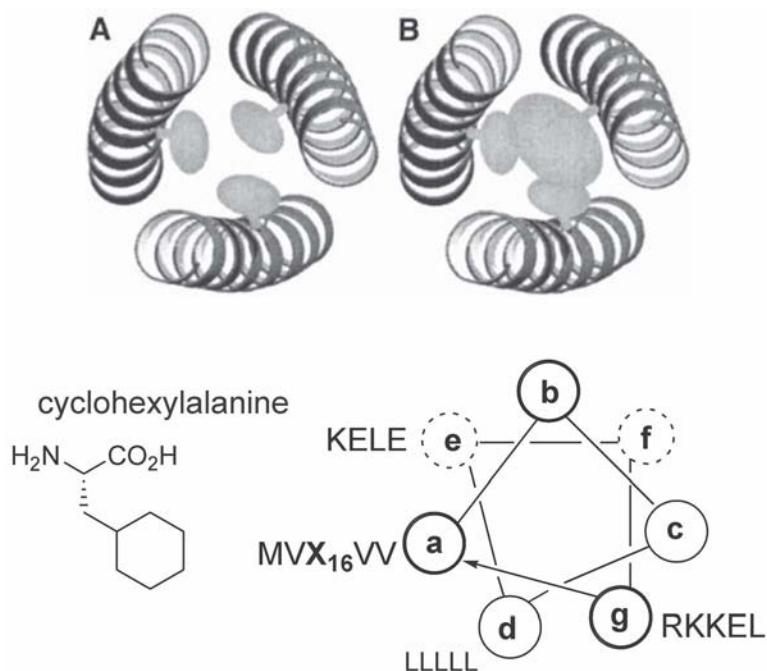
Fig. 4.5 Replacement of an alanine by a cyclohexylalanine (X_{16}) within a helical peptide favors the formation of a coiled-coil 2:1 heterotrimer (B) due to increased hydrophobic contacts at protein interfaces relative to the alanine containing homotrimer (A) [24].

studies confirmed the increased thermal stability of the heterotrimer. The melting temperature of the heterotrimer increased significantly compared to the alanine containing peptide ($\Delta T_m > 30\ °C$). However, thermal melting studies only show that an aggregate forms, but they do not allow its composition to be determined directly . That indeed a heterotrimer is formed in this case could be shown by an affinity tagging experiment. If a His-tag is attached to the cyclohexylalanine peptide, the heterotrimer can selectively be separated from a mixture of the two peptides using affinity chromatography. The retained peptide material had exactly a 2:1-composition even though the initial mixture was enriched in the alanine containing peptide. Therefore, the increased hydrophobic contacts due to cyclohexylalanine residue selectively stabilized the heterotrimer. On the other hand, a naphthalene ring is too large to fit into the void and consequently a peptide with a naphthylalanine residue instead of the cyclohexylalanine does not form stable heterotrimers. This is an instructive case study how specific peptide assemblies can be directed and controlled both in terms of stability as well as composition of the aggregates by fine tuning the hydrophobic contacts which are responsible for the aggregation.

4.2.2
Intertwined Strands

A very common hybridization motif in Nature is the double helix as found in the structure of DNA. Two helical molecules are intertwined to form a helical double strand. In the case of DNA the two strands are held together by H-bonds and stacking interactions [16] though much debate was going on in recent years about the relative importance of these interactions for the molecular recognition of nucleobases [26]. The double helix is crucial for genetic information storage and error-free replication as it ensures the correct reproduction of the information encoded in the primary sequence of the nucleotide strand (see below) [3]. The DNA double helix is also a beautiful supramolecular structure which has intrigued many chemists to devise artificial systems that similarly form helical double strands.

A very interesting class of double helix forming foldamers is based on aromatic oligoamides as introduced by Lehn and Huc (Fig. 4.6) [33a]. These oligomers are formed from alternating 2,6-diaminopyridines and 2,6-pyridinedicarboxylic acids. An intramolecular interaction between the amide NH proton and the pyridine N-atom is responsible for a curved conformation of these molecules. A more detailed description of their conformational behavior can be found in Chapter 1. These helical oligoamides can then also further dimerize forming a stable double helix in solution. Within the double helix the two oligomer strands are held together primarily by arene–arene-interactions between pyridine rings located on top of each other, whereas H-bonds occur intramolecularly within each strand being responsible for the curvature of the helix. The formation of the double helix is accompanied by a spring-like extension of the individual helices, but the inner diameter of the central pore is not significantly affected.

In these oligomers intramolecular H-bonds pre-orientate the monomers and induce the helical structure (Fig. 4.6 A). But hybridization of such oligopyridine–dicarboxamide strands is limited to a certain length of the single strand. Huc
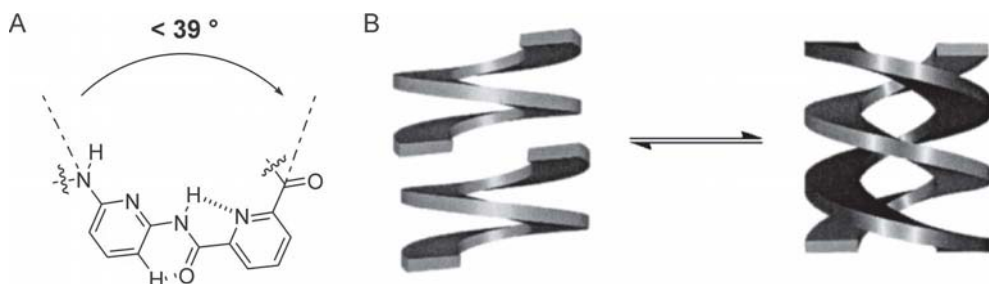
Fig. 4.6 Inducing helical conformations in oligoamides using intramolecular H-bonds (A). Two helical oligoamides can then dimerize to form a supramolecular double helix (B) [33c].

[27, 28], showed that with an increasing length of the single strand the enthalpic price of spring-like extension during the double helix formation is not compensated by intermolecular $\pi–\pi$ interactions. NMR dilution experiments in CDCl$_3$ show that at 25 °C K_{dim} increases from 210 M^{-1} for the smallest strand (5 pyridine units) to 5200 M^{-1} for a medium sized strand (9 pyridine units). In case of the largest strand (15 pyridine units) K_{dim} could not be determined because the single helix was not detected by NMR even at high concentrations. In another experiment the structure of these double helices was investigated by crystallization of pentameric oligoamides from pure DMSO and in the second case from a DMF/Et$_2$O mixture. The comparison of both structures provides evidence that the positions of the single strands in the helix are flexible. Although the crystallographic parameters (space group, unit cell parameters) of both samples indicate high resemblance to each other and even the position of incorporated water molecules differs just slightly, the superposition of both helices shows an offset of the strands of more than 1.5 Å in a plane orthogonal to the helix axis. This leads to the suggestion that the interactions between the single strands are neither directional nor dependent on the distance between the strands. These results are in agreement with the assumption of a certain screw motion based on the freedom of the single strands within the helix, known from investigations in solution.

4.2.3
Stacks of Helical Strands and Macrocycles

Whereas in the case of peptide assemblies based on coiled-coil helices the interaction is mainly controlled by hydrophobic contacts, also hydrogen bonding can be exploited to drive hybridization. In Nature this principle is found e.g. in Gramicidin A, a channel forming linear 15-residue α-peptide composed of alternating D/L-amino acids (more on Gramicidin A can also be found in Section 2.4.1 of Chapter 2). This alternation in the absolute configuration of the amino acids prevents the formation of a normal α-helix; instead a more open β-helix with an inner pore is formed. Two such helices then self-assemble within the cell membrane into a head-to-tail dimer held together by intermolecular H-bonds at the dimerization interface. Through the inner pore of this dimer, ions can then pass through the membrane [3]. The same principle was adopted by Ghadiri [29] and coworker, who designed self-assembling cyclic peptides with alternating D/L-configuration (Fig. 4.7). For example these eight-residue cyclic peptides have flat, ring-shaped conformation and their backbone amide functionalities are perpendicular to the side chains and the plane of the ring structure. This conformation leads to an intermolecular stacking via H-bonds to form nanotubes, which can also serve as pores for ion transport when embedded into a lipid membrane. Tailor-made peptide nanotubes can be achieved by variation of the side chains. Their surface-exposed arrangement leads to a membrane selectivity, which can be applied in antibacterial agents that cause rapid cell death by increased cell wall permeability and collapsed transmembrane ion transport.

A

B

C

D
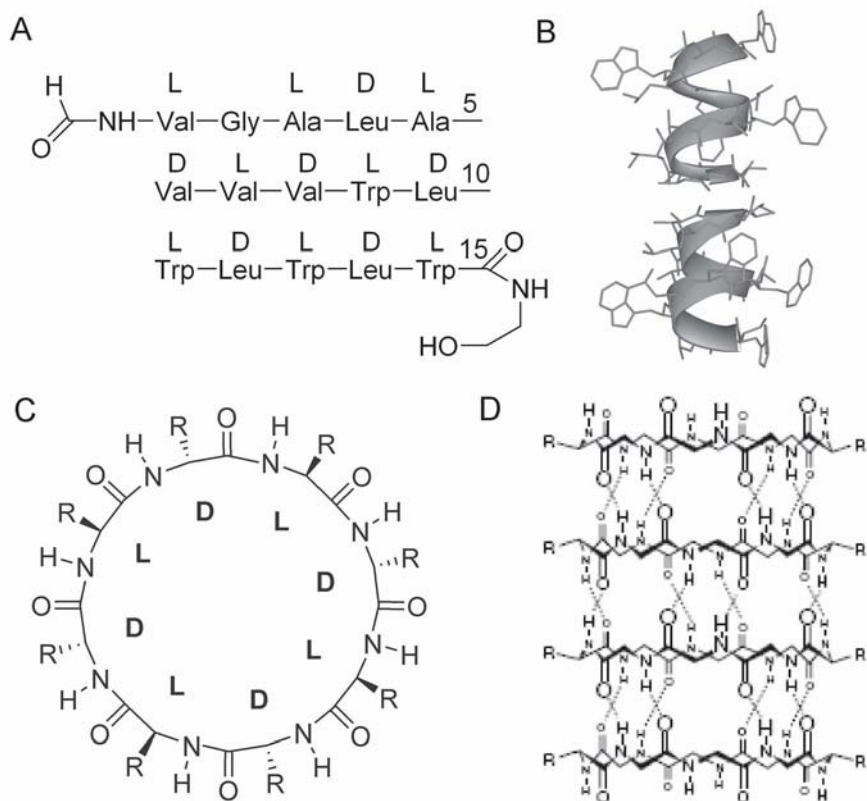
Fig. 4.7 Two helices of Gramicidin A, a 15-residue α-peptide composed of alternating D/L-amino acids (A), form a stacked dimer within a membrane allowing for ion transport though the central pore (B). Cyclic peptides with alternating D/L-configuration (C) can form self-assembled nanotubes via intermolecular hydrogen bonding (D) [29].

Another Gramidicin A analog was introduced by Guichard [30] et al. They synthesized a cyclic oligourea with homochiral residues, which self-assembles in the crystalline state via H-bonding between the urea units. Within the crystal these polar nanotubes are held together by loose van der Waals contacts. A cyclic peptide assembly with hydrophobic cavities was developed by Granja and coworkers [31]. They use 3-aminocyclohexylcarboxylic acid (γ-Acc-OH) alternating with D-α-amino acids to obtain a conformation of a flat peptide backbone with perpendicular arranged amide functionalities (Fig. 4.8). All β-methylene groups of the cyclohexane rings point into the interior of the cyclic peptide, which leads to a cavity with a partial hydrophobic interior. Modifying the C2 of γ-Acc-OH might lead to peptide nanotubes with functionalized inner surfaces.

Also the conformational preference of fused heterocycles can be exploited for helix formation (Fig. 4.9 A), as Lehn [32] et al. could already show a couple
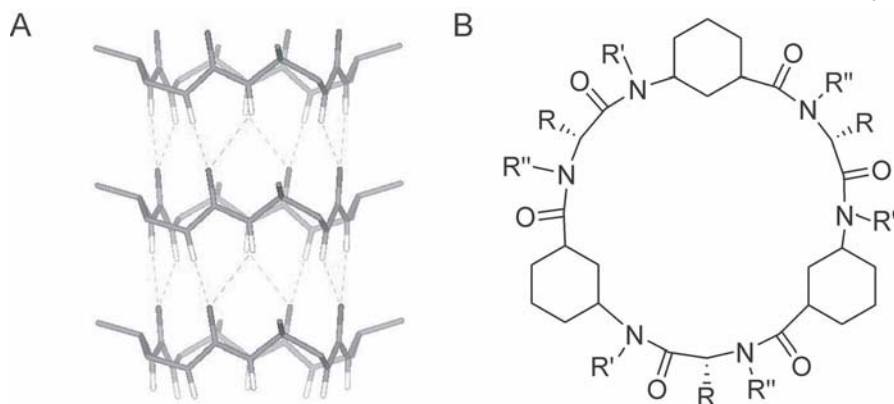
Fig. 4.8 (A) Crystal structure of Guichard's oligourea based cyclic peptide [30]; (B) structural motif of the γ-Acc-OH based cyclic peptide designed by Granja [31].

of years ago. 2,2′-Bipyridines and related heterocycles adopt a transoid conformation due to repulsive dipole-dipole interactions which can be exploited to bias an oligomer towards helix formation. In this context, it was recently shown that by using naphthyridines instead of pyridines an opening of the structure results,
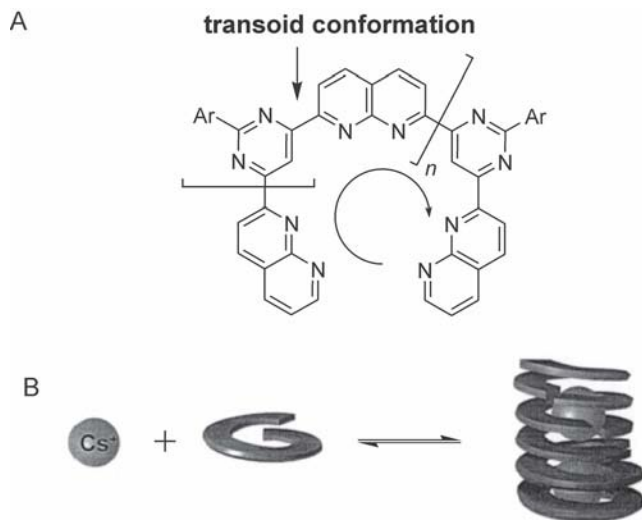
Fig. 4.9 Helical oligoamides (A) can also further interact with other small molecules. Alkali metal cations can induce the formation of large supramolecular helices from small springs which do not self-assemble by themselves (B) [32].

which allows now the inclusion of other molecules within the central pore of the helix: A first step towards helical channels [33]. For example, the naphthyridine pyrimidine oligomer ($n = 2$) in solution ($CDCl_3/CD_3CN$) adopts a conformation representing a single helical turn. In the presence of alkali metal ions such as Cs^+ a supramolecular association of such individual helical springs occurs leading to long hollow tubes, in which, as the authors suggest, cations are incorporated. It is only this mutual interaction of two springs with one cation that stabilizes these supramolecular aggregates, as in the absence of cations only monomeric springs were observed (Fig. 4.9 B). The formation of such cation channels is also supported by electrospray-mass spectrometry and transmission-electron microscopy studies.

4.2.4
Tapes and Hydrogen-bonded Sheets

Besides such helical aggregates also a variety of linear dimers with a tape or ladder-like structure have been designed. These are mostly based on rather rigid aromatic molecules with a distinct alternating H-bond donor and acceptor pattern. Some excellent reviews on this class of H-bonded dimers have been written [15, 34, 46]. One recent example is an ureido-naphthyridine dimer introduced by Zimmermann [35] (Fig. 4.10). This molecule presents a self-complementary AADDAADD H-bond acceptor and donor pattern at the edge of a rigid aromatic
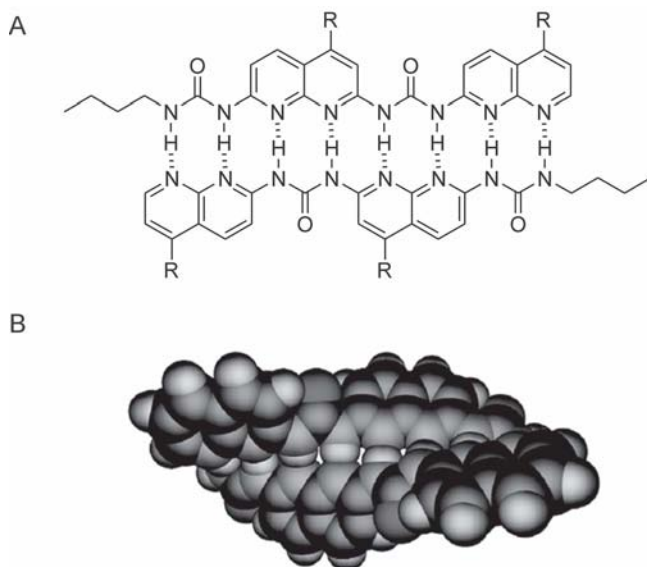
Fig. 4.10 Eight hydrogen-bonds in a self-complementary AADDAADD ureido-naphtyridine monomer lead to a strong dimerization in chloroform (A: schematic representation, B: calculated structure of the dimer) [35].

scaffold. The monomer adopts a helical structure. In chloroform dimerization via eight hydrogen bonds can occur as could be demonstrated by concentration dependent NMR experiments. Due to the fact that the ^{1}H NMR spectra remained unchanged over a concentration range from 423 μM to 13.5 mM in 10% DMSO in CDCl$_3$, a dimerization constant K_{dim} of 4.5×10^5 M^{-1} could be set as a lower limit. Upon increasing the DMSO content to 20%, K_{dim} dropped dramatically to a value of 40 M^{-1}. This clearly demonstrates the significant effect that the polarity of the solvent has on H-bonded complexes (see below).

Therefore, H-bonds alone are not sufficient to achieve strong association in more polar solvents. But in combination with other types of interactions such as ion pairs H-bonds can still be quite important for self-assembly. For example, Schmuck [36] and coworkers designed a self-complementary guanidiniocarbonyl pyrrole–carboxylate zwitterion which dimerizes in DMSO with an association constants of $K > 10^9$ M^{-1} and even in pure water has a $K = 170$ M^{-1}. The ion pair interaction is crucial as could be shown by comparison with a neutral "knock-out" analog which has the same H-bond pattern but lacks the charges (Fig. 4.11). This neutral analog forms isostructural dimers but which are stable only in chloroform ($K > 10^4$ M^{-1}). Already the addition of small amounts of

K > 10^9 M^{-1} in DMSO

K = 170 M^{-1} in H$_2$O

K > 10^4 M^{-1} in CHCl$_3$

K < 10 M^{-1} in 5 % DMSO in CHCl$_3$
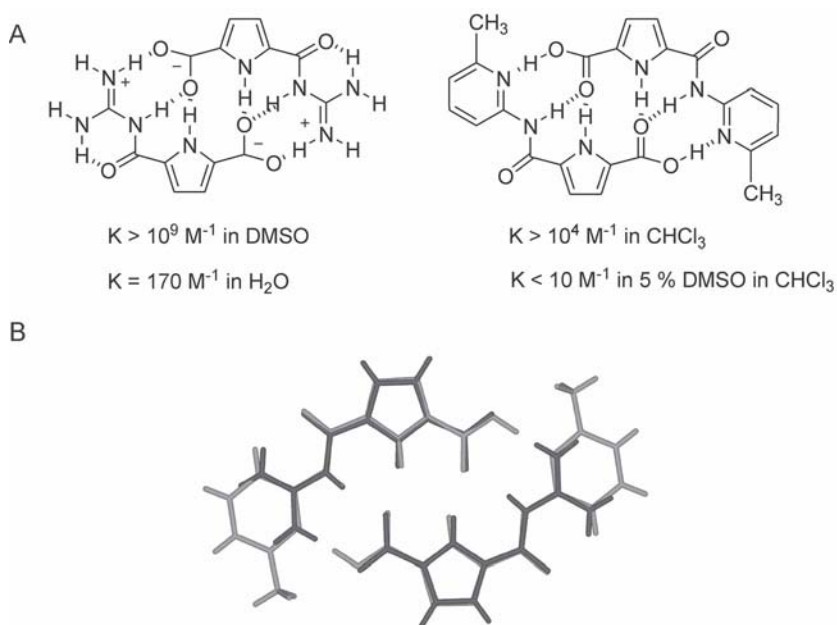
Fig. 4.11 Dimer formation (A) of a zwitterionic guanidiniocarbonyl pyrrole-carboxylate zwitterion (left) and a neutral amidopyridine-carboxylic acid "knock-out" analog (right). Both dimers are isostructural as the solid state structures reveal (B), but differ significantly in their stability [36].

DMSO dramatically reduces the stability of the neutral dimer ($K \approx 10$ M^{-1} in 5% DMSO in CHCl$_3$). However, as a theoretical study of systematically varied ''knock-out'' analogs showed, the charge interaction alone is not sufficient either to explain the large stability of the zwitterionic dimer [37]. Other zwitterions with different H-bond patterns form much less stable dimers. Hence, the stability of the guanidiniocarbonyl pyrrole–carboxylate zwitterions stems from a combination of both the formation of a directed ion pair and this specific H-bond pattern.

4.3
Hybridization-induced Folding of Unstructured Molecules

In contrast to the examples discussed in Part 1, in most artificial systems the structure of the underlying monomers is not well defined and the aggregation is the trigger for structure formation. Hence, this section deals with oligomers which do not possess an intrinsic tendency to adopt a well-defined conformation, but which may do so upon hybridization with another strand.

4.3.1
Hydrogen-bonded Tapes

In continuation with the examples above (Section 4.2.4) similar self-complementary H-bond arrays have been designed even in larger and more flexible molecules. However, most often the monomers do not adopt a defined conformation before hybridization but are rather randomly coiled or exist as a rapidly interconverting mixture of different conformers or even tautomers with similar energy. Hence, structure formation is only induced by the dimerization itself. An interesting example was presented by Gong et al. who designed a six-H-bonded duplex based on an AADADD pattern within an oligoamide consisting of meta-substituted benzene rings linked via glycine residues (Fig. 4.12) [38]. Self-association leads to a linear tape-like duplex which in chloroform is so stable that its dissociation could not be analyzed by concentration dependent NMR studies. A pyrene-labeled derivative allowed to determine a dimerization constant of 7×10^9 M^{-1} via fluorescence studies.

This demonstrates an important point in the problems one often encounters when characterizing supramolecular assemblies. Each physical method such as NMR or UV-Vis spectroscopy has its own specific concentration range where it works best. However, this means that only dissociation constants which are of the same order of magnitude can be accurately measured. For example, NMR requires a sample concentration in the mM range. Hence, only association constants in the range of ca. $K = 100–10\,000$ M^{-1} can be determined reliably. With lower or higher association constants the sample is either fully dissociated or fully complexed at the accessible concentrations. However, for an accurate determination of the association constant, an optimal complexation range of 20–80% is needed [39]. UV or fluorescence studies require much lower sample concentra-
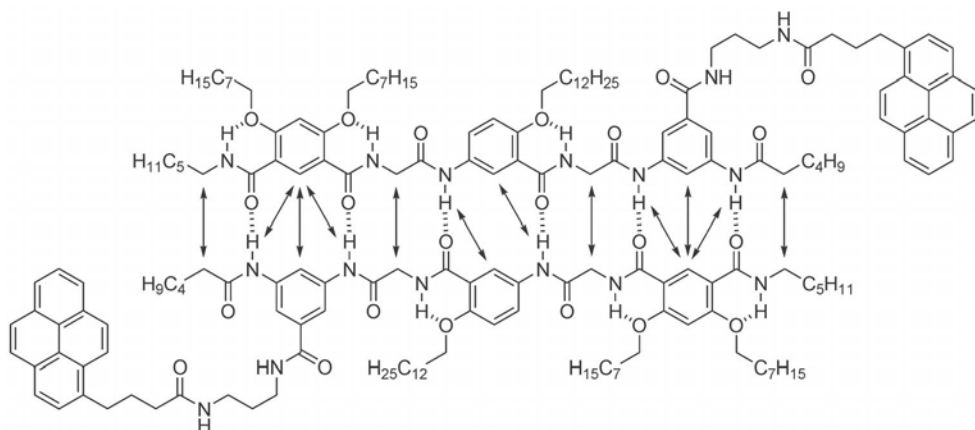
Fig. 4.12 A six-H-bonded duplex based on a self-complementary AADADD pattern. The arrows indicate diagnostic interstrand NOEs. The pyrene units were needed to follow the dimerization using fluorescence studies.

tions, therefore larger association constants can be measured. Hence, not every spectroscopic technique is suitable for every supramolecular system. The use of a more sensitive technique such as fluorescence is sometimes hampered by the absence of appropriate reporter groups. Whereas normally most organic molecules can be analyzed by NMR, fluorescence spectroscopy requires a special chromophore. Therefore, sometimes synthetic modifications have to be introduced before a certain technique can be applied. Here, a pyrene labeled derivative had to be synthesized before the dimerization constant could be measured by fluorescence [38].

Gong [40] also studied several heteroduplexes in which two different molecules with complementary H-bond patterns associate to form ladder-like duplexes (Fig. 4.13). Dimer formation in these cases could be demonstrated by specific interstrand NOEs. Furthermore, as the duplex is overall less polar than the individual monomers, which cannot interact and hence possess free H-bond donors and acceptors. Hence the heterodimer shows a significantly different chromatographic behavior. Whereas the two monomers have a very low mobility on the silica gel TLC ($R_f = 0.0$ and 0.1 using 10% DMSO in $CHCl_3$, respectively), a 1:1-mixture of the two monomers has a R_f value of 0.96 indicating the formation of the overall much less polar heteroduplex! Again the thermodynamic stability was too large for an accurate determination of the association constant by concentration dependent NMR studies. By isothermal titration microcalorimetry (ITC) the association constant could be estimated to be $K \approx 10^9$ M^{-1}. However, as already mentioned in the introduction also the solvent plays an important role in controlling the stability of such aggregates as it dramatically affects the strength of noncovalent interactions which drive the folding or hybridization process. For exam-
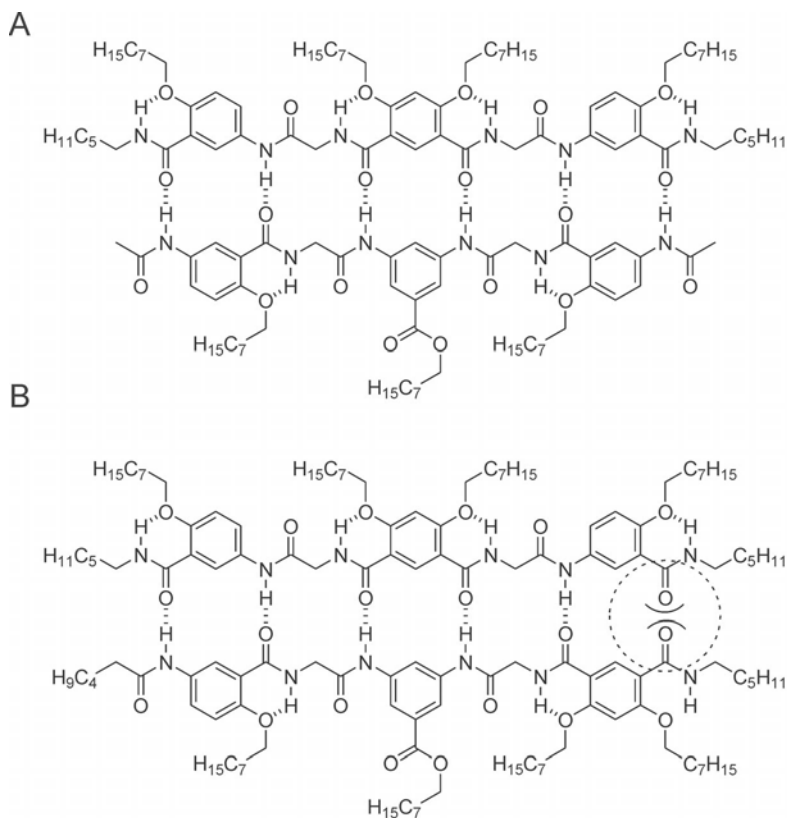
Fig. 4.13 Heterodimerization of two duplex strands held together by six H-bonds which impose a significant sequence specificity (A): the introduction of a mismatch (B: to the right) leads to a decrease in binding affinity by a factor of 40 in this case.

ple, in polar solvents hydrogen bonding is significantly reduced due to the competitive solvation of donors and acceptors by individual solvent molecules. Hence, duplex formation first requires desolvation of both donors and acceptors before their intermolecular interaction can lead to the formation of a dimer. Its stability is of course only the difference of the interactions within the aggregate relative to the interaction of each monomer with the solvent. As the latter can be substantial especially in polar or aqueous solvents, the stability of H-bonded supramolecular aggregates in such solvents is significantly reduced. In this specific case of Gong's six-H-bond heteroduplex already the addition of 5% DMSO led to drop in stability of several orders of magnitude.

Also the introduction of a mismatch caused a significant decrease in the stability of the tapes. In one specific example an attractive H-bond between an amide NH and a carbonyl group was replaced by a repulsive interaction between two car-

bonyl groups. The stability of this mismatched heteroduplex was 40 times lower than of the corresponding matched pair [41]. In Nature sometimes even larger effects are observed for such mismatches. A similar chemical exchange is the reason for the upcoming resistance of bacteria towards the antibiotic Vancomycin [42]. Normally, Vancomycin forms a complex with a bacterial peptide substrate held together by five H-bonds within a hydrophobic environment. In the resistant bacteria the exchange of an amide for an ester within the substrate, replaces an attractive H-bond for a dipole repulsion. The complex stability drops by a factor of 1000. Model studies showed that the loss of the H-bond is responsible for a drop in affinity by a factor 10, whereas the repulsive dipole interaction led to another decrease by a factor of 100 [43].

This specificity of H-bonds in combination with their directionality makes H-bonds so attractive for designing supramolecular structures despite the inherent problem of their weakness in more polar solvents. The individual pattern of H-bond donors and acceptors within one monomer very specifically determines the binding partner required for stable duplex formation. Furthermore, the H-bond pattern also regulates the stability of the duplex due to attractive or repulsive secondary interactions between neighboring binding sites as initially proposed by Jørgensen in 1990 [44]. Therefore, a self-complementary AADD pattern is more stable than an ADAD pattern.
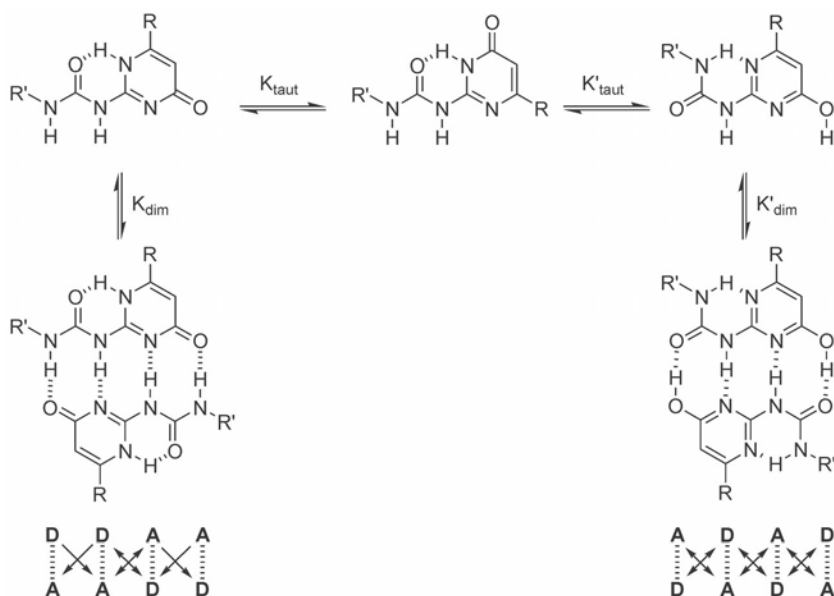
Fig. 4.14 2-ureido-4-pyrimidones exist as a rapidly interconverting mixture of three tautomers in solution with different self-assembling properties depending on their specific sequence of H-bond donor (D) and acceptor (A) sites.

Meijer et al. were one of the first groups who synthesized such quadruple hydrogen bonding motifs. Acylation of diaminotriazines and diaminopyrimidines led to the desired self-complementary ADAD binding motif, which was confirmed by X-ray diffraction. NMR experiments showed that this self-complementary binding motif has association constants of up to 10^5 M^{-1} in chloroform. In respect to Jørgensen a more stable quadruple hydrogen binding motif should result from an AADD array of donor and acceptor sites. Therefore Meijer et al. used 2-ureido-4-pyrimidones as basic molecules for their AADD binding site [45]. The analysis of this bonding motif became quite complicated, because of a complex equilibrium of three tautomers which all coexist in solution (Fig. 4.14). Their composition is both determined by the polarity of the solvent and the concentration of the compound itself. One tautomer has a DDA bonding pattern and can not dimerize. The other two tautomers both present self-complementary quadruple hydrogen binding patterns with either an ADAD or AADD sequence. They can form dimers but with different stabilities. The latter is favored, because of additional stabilizing secondary interactions within the dimer, which aren't possible for the ADAD pattern. Due to the fact that it was impossible to quantify an association constant for the AADD binding motif by NMR experiments, Meijer used excimer fluorescence spectroscopy as an indirect method yielding a value of $6 \cdot 10^7$ M^{-1}. However, this example shows how complicated the situation can be
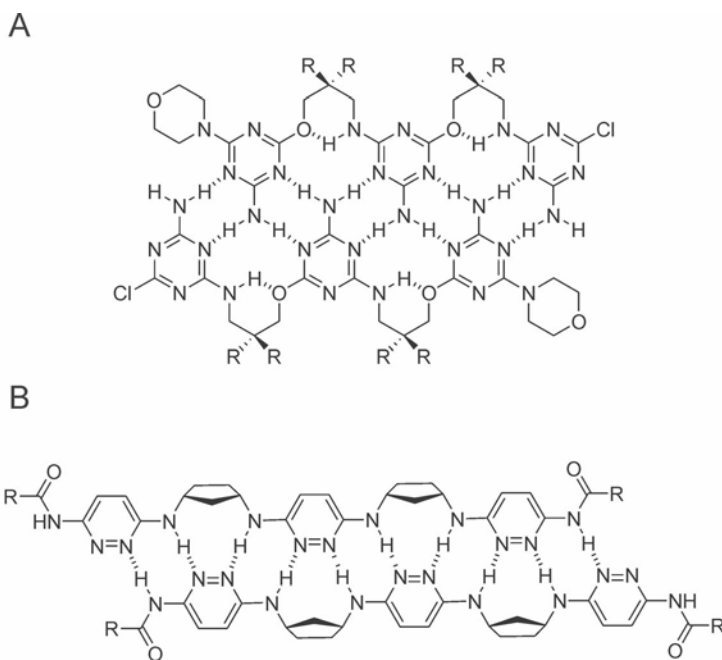
Fig. 4.15 Linear H-bonded tapes based on 2,6-triazines (A) or 2,6-pyridazines (B) as developed by Krische and coworkers.

if the self-assembling monomers do not have a specific well defined structure by themselves.

Another class of linear H-bonded tapes was designed by Krische and coworkers based on aminotriazines, which are covalently linked by flexible aminoalcohols [46]. Several oligomers were prepared, which all present an alternating repeat of an AD-DA H-bond pattern provided by the aminotriazine moiety (Fig. 4.15). This can interact with two other aminotriazine moieties in a second oligomer to form an interdigitated linear tape. ITC experiments in 1,2-dichloroethane at 20 °C showed an increasing association constant K_a from the monomer (4.7 M^{-1}) over the dimer (5×10^3 M^{-1}) to the trimer (6.9×10^8 M^{-1}). The tetramer instead had a lower association constant K_a of 1.1×10^3 M^{-1}, even lower than the dimer. This value resembles the association constants for duplex dimer formation, which suggests an intramolecular folding of the tetramer. Related duplex strands based on 3,6-diaminopyridazines were introduced by Krische later (Fig. 4.15 B) [47]. The association behavior was measured in analogy by ITC under the same conditions mentioned above, resulting in association constants of 5 M^{-1} for the monomer, 870 M^{-1} for the dimer and 8×10^5 M^{-1} for trimer.

4.3.2
Helices Based on Metal-ligand Interactions and Salt Bridges

As mentioned above one of the major disadvantages of H-bonded structures is their limited strength in polar solvents. Hence, most of the folding and structure building described above is limited to chloroform or even less polar organic solvents. The addition of more polar solvents such as DMSO, methanol or water normally immediately destroys the aggregates [48]. Therefore, to achieve stable structures in more polar solvents other interactions are needed. Perhaps the most widely used approach in the context of hybridization induced folding of unstructured molecules is the formation of metal helicates. Monomeric strands are designed with appropriate metal binding sites such as phenanthrolines or bipyridines. The addition of metal ions then can lead to the formation of double, triple or even quadruple-helical complexes. In general, the underlying monomeric ligands are not structured, but the resulting metal helicates are well defined. The exact structure of the helicate is controlled by the complexation geometry of the templating metal ion. Besides the usual solution techniques, such as NMR, also structure determination in the solid state is often possible providing the ultimate proof for the hybridization induced folding; another great advantage compared to H-bonded assemblies, which are often difficult to fully characterize structurally. This vast area of research has been covered in several excellent review articles and will not be discussed here any further [49]. Just one interesting recent example should serve to illustrate this approach. Three bis-pyridylimine organic strands can wrap around two Fe(II)-ion to form a stable triple-helicate. This metal–helicate has the correct size and shape to interact with a DNA three-way junction (Fig. 4.16). As an X-ray structure showed the metal complex sits directly in the middle of the junction. The aromatic phenyl rings of the ligands allow for exten-
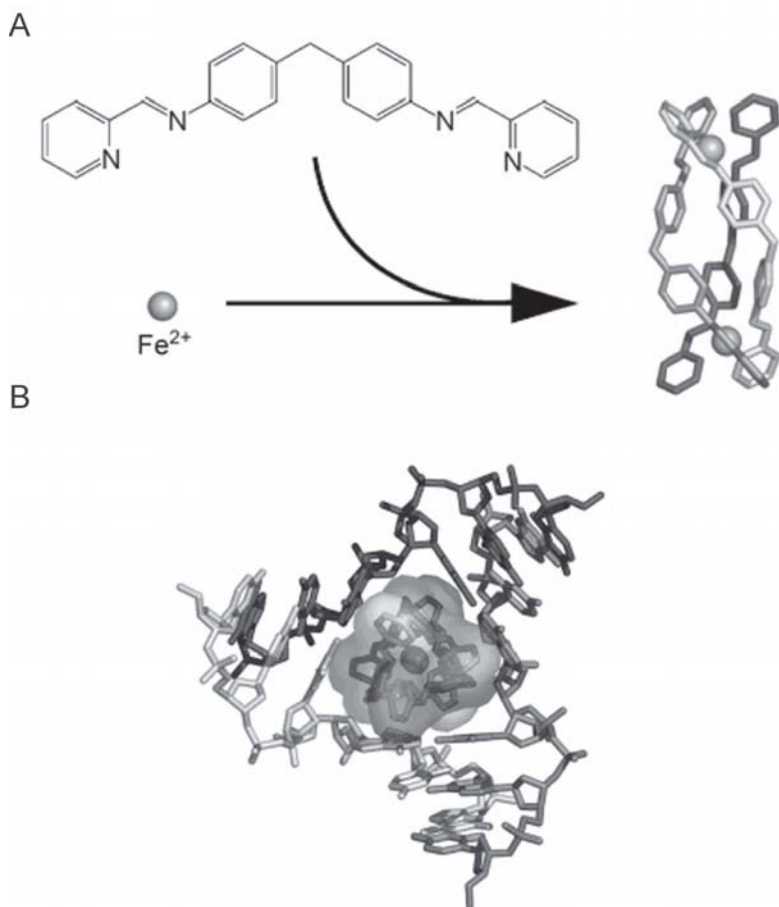
Fig. 4.16 Metal ions can form stable double and triple helices with appropriate ligands (A). Whereas the ligands themselves are normally not structured the metal helicates are. In this case a triple helix is built which can further interact with a DNA three-way junction (B: major groove side view) [50].

sive hydrophobic contacts with the thymine and adenine bases at the junction, whereas the overall cationic charge of the helicate provides additional long range electrostatic interactions with the negatively charged DNA. Furthermore, the threefold symmetry of the DNA junction exactly matches the threefold symmetry of the triple-helicate. The helicate is hence a perfect match in size and complementary interacting sites to fill the void in the core of the three-way junction [50].

Besides metal–ligand interactions, which in general are much stronger than other noncovalent interactions often approaching the strength of covalent bonds, also ion pair formation can be used to increase the stability of duplexes as had been shown above for zwitterionic dimers (Fig. 4.11). This approach was also

used by Yashima [51] who designed meta-terphenyl dimers with either amidinium or carboxylates attached to the central aromatic ring of the terphenyl unit. Chiral groups attached to the amidinium cations were used to induce a defined helical twisting of the monomer. Nevertheless, this monomer exists in solution as a mixture of several conformers as indicated by multiple signal sets in the NMR in chloroform. Addition of the corresponding dicarboxylate resulted in a simplified spectrum of a single species indicating the formation of an ion paired duplex of considerable stability ($K > 10^6$ M^{-1}). As ion pairs between amidinium or guanidinum cations and carboxylates are directional (in contrast to the interaction of spherical ions) [52], the two salt bridges bring the two strands together in a distinct orientation. The chiral twist of the terphenyl units of the diamidinium cation then induces a winding up of both strands in a right-handed double helical structure (Fig. 4.17). Without the chiral inductors a racemic mixture of right and left handed double helices would result. This double helical structure was con-
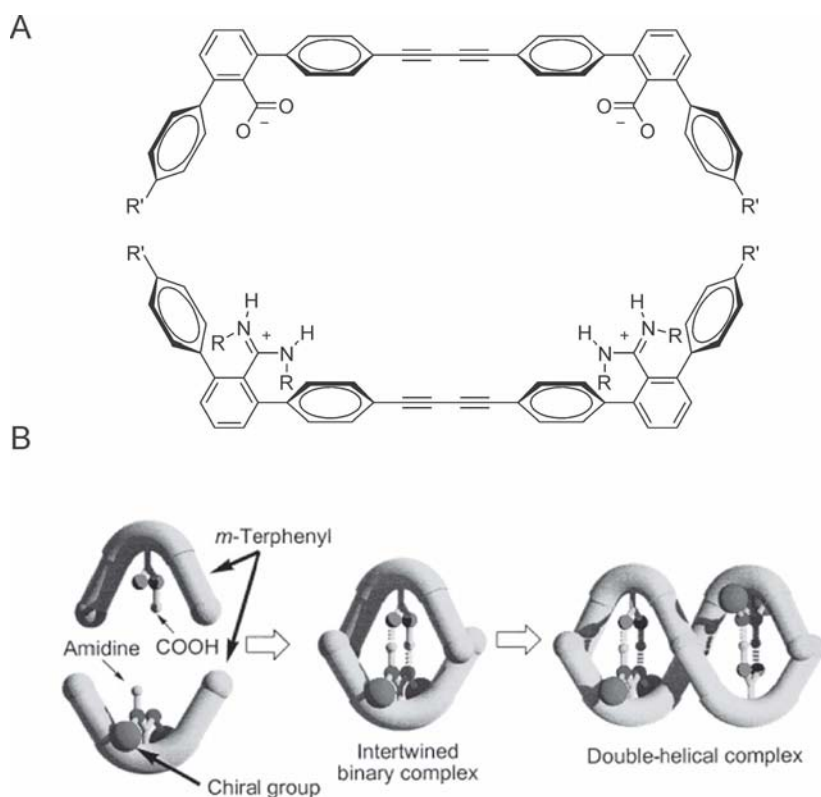
Fig. 4.17 Charge interactions can significantly improve the stability of supramolecular dimers as shown here for a helical meta-terphenyl heterodimer developed by Yashima et al. [51] (A: structure of the two monomers; B: schematic representation of heterodimerization).

firmed at least for the solid state by X-ray analysis. By using the enantiomer of the cation, the helical sense of the double helix can be reversed as could be shown by the mirror-image CD signals of the corresponding duplexes in solution. Due to the stronger electrostatic interactions relative to simple H-bonding, the double helical structure is also retained at least to ca. 70% in the more polar solvent DMSO.

4.3.3
Double-stranded Hybrids Based on Aryl-aryl Interactions and Hydrophobic Contacts

Besides H-bonds and ion pair formation also aromatic and hydrophobic contacts have been used to drive the aggregation of unstructured monomeric strands into folded and structured supramolecular oligomers. Even though these types of interactions even increase in strength in polar and aqueous solvents they are much more difficult to deliberately use. For example, it is much more difficult to control the specificity of a supramolecular interaction based on hydrophobic contacts as any hydrophobic residues will tend to stick together (see also Chapter 3) [48]. Hence, often such unspecific hydrophobic or aromatic interactions are combined with specific H-bonds or charge interactions to control the hybridization specificity. One nice example which combines aromatic edge-to-face stacking interactions with H-bonds was provided by Hunter [53] and coworkers. A series of amide oligomers derived from isophthalic acid and a bisaniline derivative were synthesized. Concentration dependent NMR studies in chloroform revealed that these oligomers form dimers with a zig-zag structure (Fig. 4.18). The stability increased with increasing chain length with some indication of positive cooperativity. The complexation induced changes in chemical shift of both amide and aryl hydrogens were monitored by NMR dilution experiments in chloroform. Upon complexation the amide hydrogens shift downfield, whereas the aromatic hydrogens of the bisaniline shift upfield indicating aromatic edge-to-face interactions which bring them into the shielding cone of the isophthalic acid. The NMR dilution and titration experiments of complementary but different oligomers were
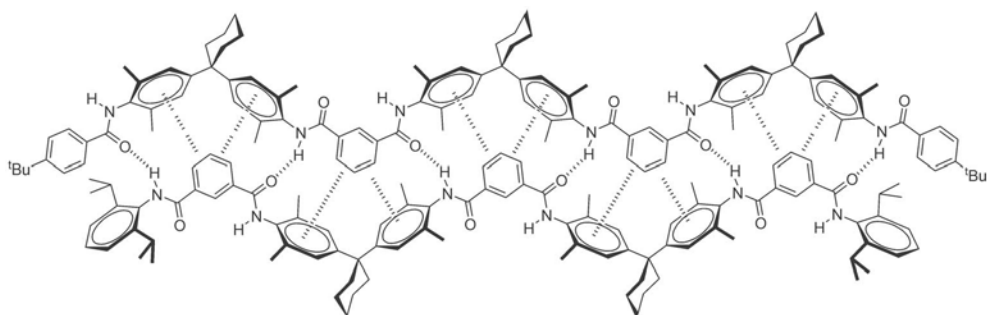
Fig. 4.18 A molecular zipper based on aromatic interactions and H-bonds as developed by Hunter et al.

also used for the determination of association constants and their correlation with the numbers of hydrogen bonds. A two-hydrogen-bonded complex has an association constant K_a of 18 M^{-1}, which increases to a value of 240 M^{-1} for a four-hydrogen-bonded complex and 55 000 M^{-1} for the six-hydrogen-bonded complex. All experiments were measured in a mixture of chloroform and methanol (95:5). This nonlinear increase in complex stability with increasing number of H bonds indicates a positive cooperativity [54]. However, the stability of the dimers is significantly depending on the solvent. Increasing the content of methanol significantly destabilizes the complexes. This suggests that even though aromatic interactions occur the main driving force for association is arising from the H bonds.

A way to increase the specificity of hybridization based on aromatic interactions alone is to use the alternating stacking of electron-rich and electron-poor aromatics as shown by Iverson [55]. Flexible oligomers with alternating electron rich dialkoxynaphthalene units (Dan) and electron poor naphthalenetetracarboxylic diimides (Ndi) adopt stable structures via intramolecular stacking even in aqueous solvents. The dominant driving force for folding in this case is the desolvation of the aromatic surfaces upon their mutual interaction in water. However, the electrostatic complementarity of the aromatic rings determines the extent of stacking (face-centered vs. off-centered or edge-to-face-stacking) leading to the formation of a defined and well structured foldamer. These so called aedamers are described in more detail in Chapter 3. The same stacking interaction can also be used to construct stable heteroduplexes in water. Oligomers of either Dan or Ndi alone do not fold intramolecularly. However, upon mixing these two types of oligomers stable heteroduplexes result (Fig. 4.19). The stability of these
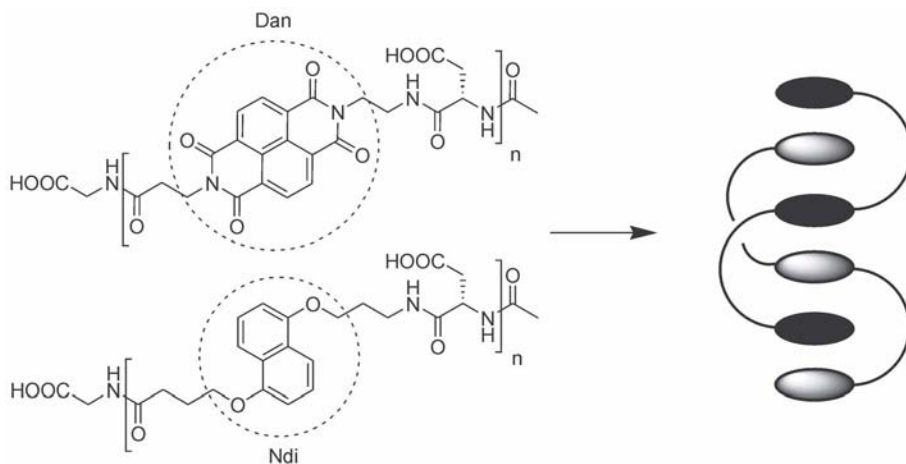
Fig. 4.19 Stacking interactions between electron-poor (Dan) and electron-rich (Ndi) can be used to direct the hetero-association of aromatic foldamers in buffered water. The mutual interaction of these two aromatic units leads to stable helical structure.
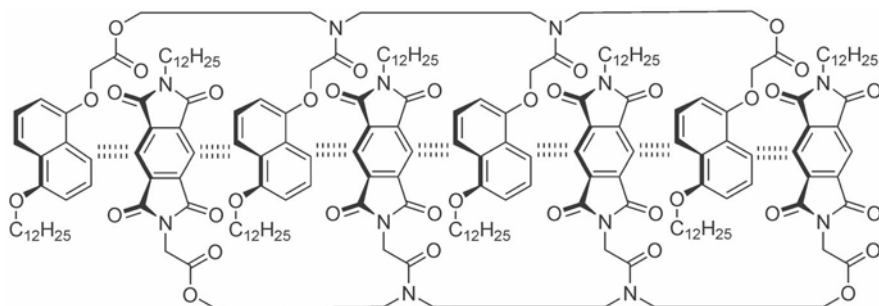
Fig. 4.20 Duplex formation based on the hetero-association of electron-rich and electron-poor aromatic units. In this case, solvophobic interactions play only a minor role for dimerization.

heteroduplexes increases more or less additively with each interacting pair of aromatic units contributing ca. $\Delta G = -1.9$ kcal mol^{-1} in stability. Hence, no cooperativity is observed. Nevertheless, a tetramer has already a surprisingly large stability of $K = 350\,000$ M^{-1} in buffered water. The thermodynamic signature as determined by ITC shows an enthalpy driven association with a negative entropy contribution. The large negative entropy contribution most likely stems from the loss in the flexibility of the linkers. Otherwise simple aromatic stacking interactions are often also entropically favored due to the release of ordered water molecules from the interface (classical hydrophobic effect) [56]. The heteroduplex also seems to be kinetically rather stable as it migrates as one sharp band in a polyacrylamide gel electrophoresis experiment. Unfortunately, the exact structure of the heteroduplex is not known yet [57].

Similar zipper-shaped heteroduplexes were recently presented by Li and coworkers [58]. Again, either electron rich dialkoxynaphthalenes or electron-deficient pyromellitic diimides were linked via simple alkylchains of variable length (Fig. 4.20). Concentration dependent NMR studies with two tetramers in chloroform indicate the formation of heteroduplexes of moderate stability ($K < 3000$ M^{-1}). The chain length or chemical nature of the linker (ester or amide groups) only marginally affected the stability of the duplexes. However, the addition of polar solvents such as methanol or DMSO reduced the stability even so only moderately. This is in contrast to other aromatic stacking systems where solvophobic interactions play an important role [59]. These however increase with increasing polarity of the solvent [60]. Again, the exact structure determination of the duplex remains an open task.

4.3.4
Hybrids Based on DNA-base-pairing Recognition

The archetype of hybridization induced duplex formation of course is found in DNA. The interaction of two complementary linear oligonucleotide strands leads
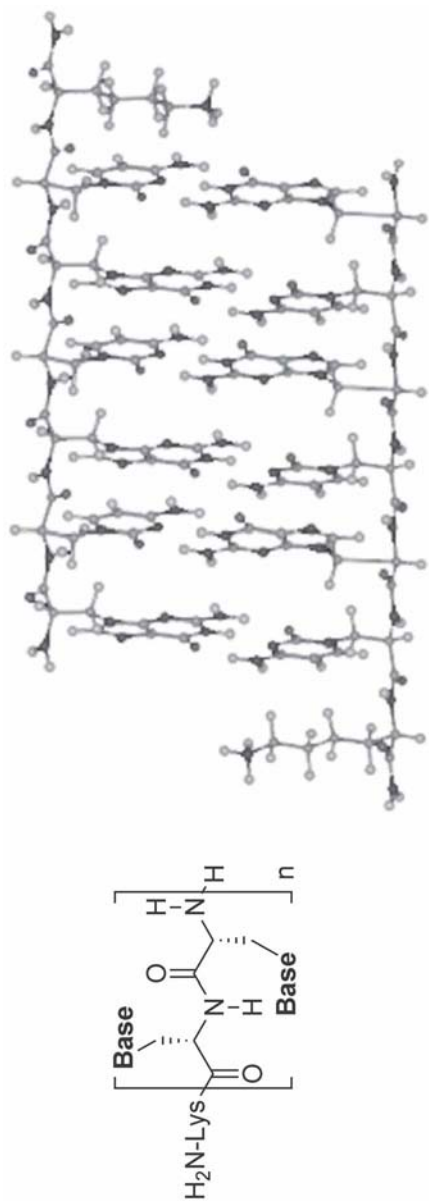
Fig. 4.21 Schematic illustration of a PNA (left) and an antiparallel PNA double strand based on heterochiral alanyl-oligomers with complementary G-C base sequences [64].

to the well known double helical structure. Sometimes, the hybridization is hampered by the intramolecular structuring of the single stranded DNA itself. Long single-stranded DNA and RNA can have considerable secondary structures such as loops or bulbs that can prevent its interaction with the complementary oligonucleotides [61]. Shorter oligonucleotides however are normally unstructured and only obtain their specific fold in the context of the double helix. A variety of DNA analogs with similar pairing and folding properties have been designed based both on natural (nucleotides and peptides) as well as unnatural backbones (hydrocarbon skeletons) [62, 63].

For example, peptide nucleic acids (PNAs) are oligonucleotide mimics in which the sugar phosphate backbone is replaced with an *N*-(1-aminoethyl)glycine polymer carrying A, G, C, and T nucleobases via methylene carbonyl linkages (Fig. 4.21). These strands by themselves are not orderly structured but consist as mixtures of various conformations (e.g. *cis/trans* amide rotamers). However, they possess the same hybridization properties as normal DNA or RNA, forming stable duplexes based on Watson–Crick base pairing either with another PNA (Fig. 4.21) or with both types of nucleotides. In this context PNAs might have some promise for antisense gene therapy or as DNA probes [64].
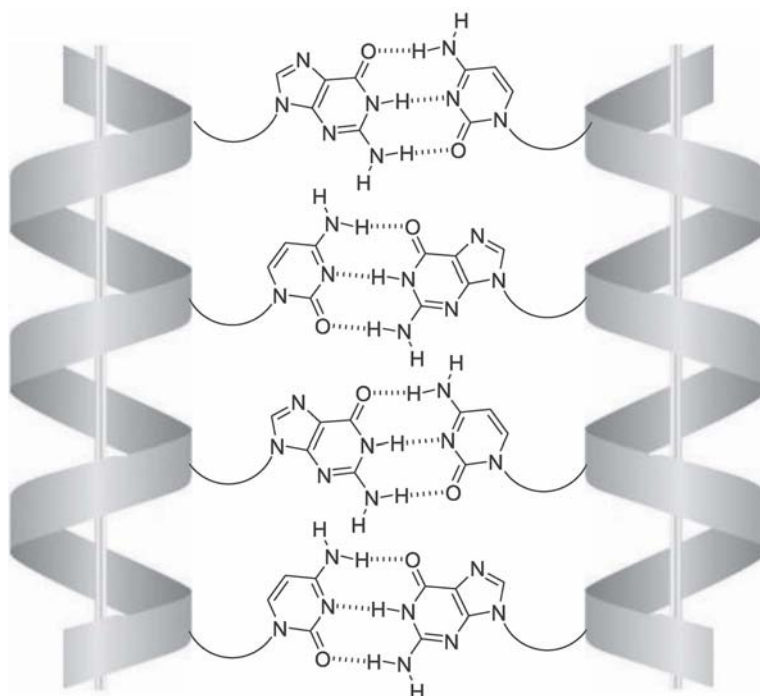
Fig. 4.22 *β*-Peptides with attached nucleobases can form stable duplexes via helix aggregation through antiparallel Watson–Crick base pairing.

Also β-peptides [65] can function as a foldamer scaffold for the formation of self-assembled duplexes. Gellman and Diederichsen [66] designed a β-peptide with nucleobases attached at every third position, which can form a 14-helix. The other positions were assigned with β-homolysine in order to increase solubility in
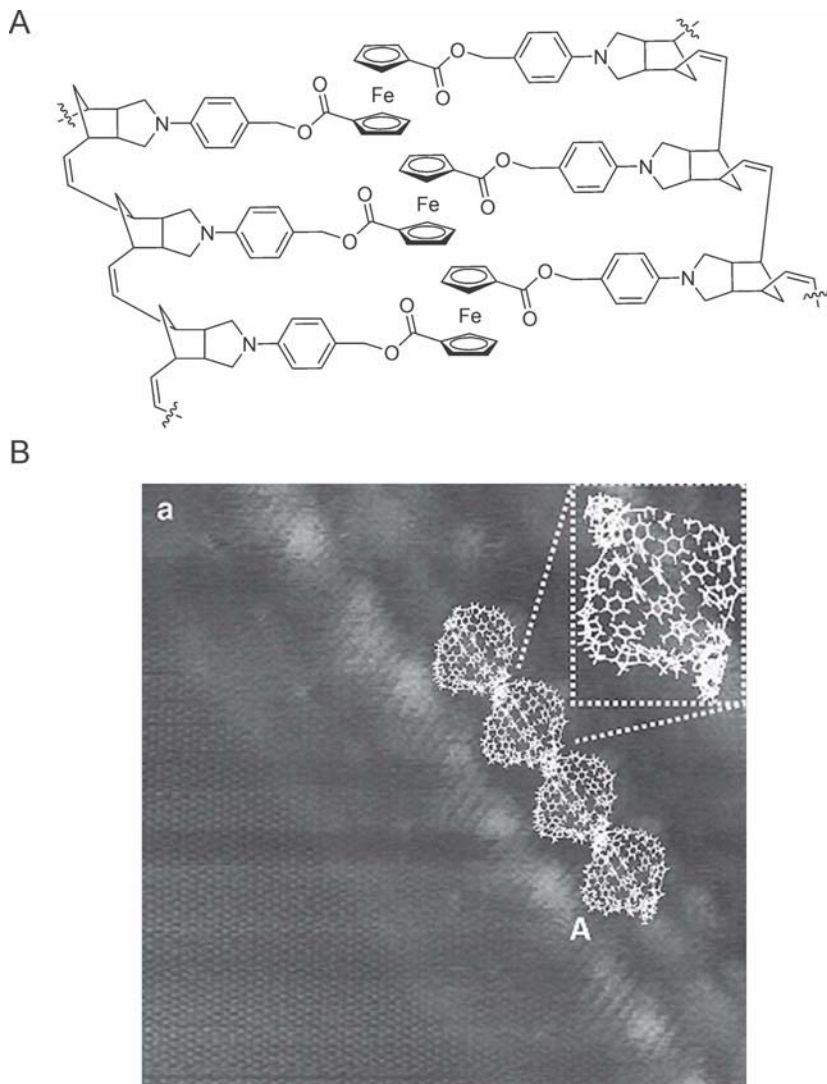
Fig. 4.23 A polymer formed by ring-opening metathesis from a bis-norbornene derivative with a rigid ferrocene bridge, resembles DNA by having a double helical structure with similar geometric parameters as natural DNA (A) [68]. STM experiments on HOPG confirm double helix formation (B) [68].

aqueous media and 2-aminocyclohexane carboxylic acid (ACHC), which should strongly promote the 14-helix formation. The single strands are still very flexible in solution. However, this conformation acts as a scaffold for the presentation of the nucleobases on one side of the helix. Watson–Crick base pairing between two complementary strands then leads to stable duplexes (Fig. 4.22), which were verified by ESI mass spectrometry and temperature dependent UV and CD spectrometry even though the exact conformations and structures of either the single or the double strands have not yet been determined. For example, an equimolar mixture of complementary oligomers, containing the nucleobase sequences ATCA and TGAT, forms a duplex of significant stability ($T_m = 44\ °C$). In order to determine the utility of hydrogen-bond-mediated base pairing for the duplex formation, methylated guanine nucleobases were synthesised and incorporated into the β-peptide helix. These modified nucleobases are no longer able to dimerize over the Watson–Crick site, so that no interaction of the two helices was detected by temperature-dependent UV spectroscopy. Also the CD spectra, recorded at low temperature, gave no indication for base-paired double strands. These measurements confirm the need for free Watson–Crick sites for base pairing. Helical β-peptides by themselvess can self-assemble in water, but in this case rather unspecific aggregates of unknown composition and structure are formed [67].

An interesting DNA analog, albeit remote, based on a complete artificial skeleton was described by Luh [68]. A bis-norbornene derivative with an incorporated ferrocene unit was polymerized using ring opening metathesis leading to a double stranded polymer with an average of 29 repeat units (Fig. 4.23). As scanning tunneling microscopy showed this double strands can adopt several structures one being a classical double helix (besides a supercoiled and a ladder structure, Fig. 4.23 B). This double helix has geometric parameters very similar to the natural DNA as far as for example the number of monomeric units per pitch or the spacing between them is concerned. However, according to molecular mechanics calculations this double helical structure is energetically less favorable compared to the supercoiled or ladder structure.

4.4
Formation of Large Polymeric Aggregates via Self-assembly

Self-assembly is however not limited to the aggregation of a small discrete number of oligomeric strands. Also much larger aggregates can be formed by the hierarchical interaction of suitable monomers. Again, the building blocks of these assemblies normally do not possess a well-defined conformation of their own as monomers but the assemblies can be relatively well defined. For example, Meijer's four H-bonded supramolecular polymers attracted much attention in this context (Fig. 4.24).

The resulting superstructures are not covalent but supramolecular polymers held together by specific noncovalent interactions between bifunctional building blocks. This linear supramolecular polymer then folds into a specific three di-

A

B
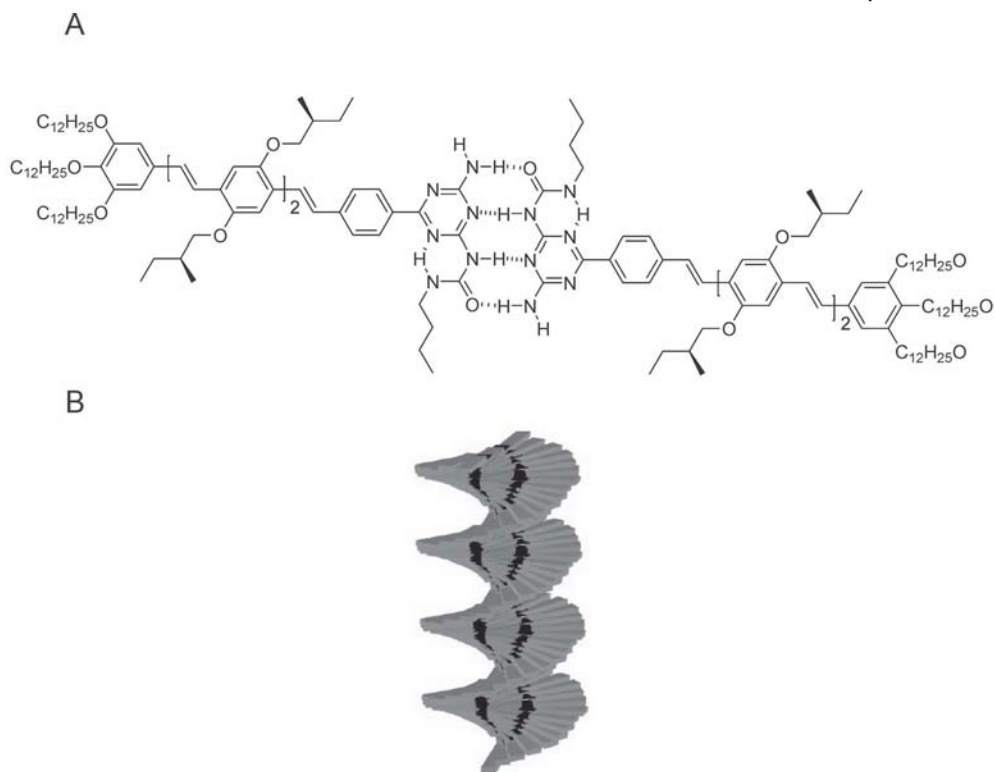
Fig. 4.24 Hierarchical stacking of chiral four H-bonded dimers (A)
leads to large helical super structures (B) [69].

mensional structure [69]. Such a supramolecular foldamer can also further assemble into larger aggregates. For example, a helical supramolecular polymer based on the dipolar stacking of bis-merocyanine dyes forms intertwined linear rods via self-assembly of six such helices. These rods can then even further self-organize into hexagonal stacks (Fig. 4.25) [70]. However, exact structure information on a molecular level is often not available for such large aggregates and hence only reasonable models can be suggested at best.

Peptide amphiphiles (PAs) as introduced by Stupp [71] are another highly interesting class of self-assembling polymers forming large aggregates. They consist of a hydrophilic peptide segment with most often ionizable side chains covalently coupled to a lipid tail (e.g. palmitic acid). The structure of the oligopeptide segment is hence pH sensitive. For example, when ionized the charges prevent structure formation, whereas after protonation at low pH values the peptide adopts a helical secondary structure. This change in protonation state and structure ultimately drives self-assembly through hydrophobic collapse of the lipid tails into cylindrical nanostructures of considerable size (Fig. 4.26). Such aggre-
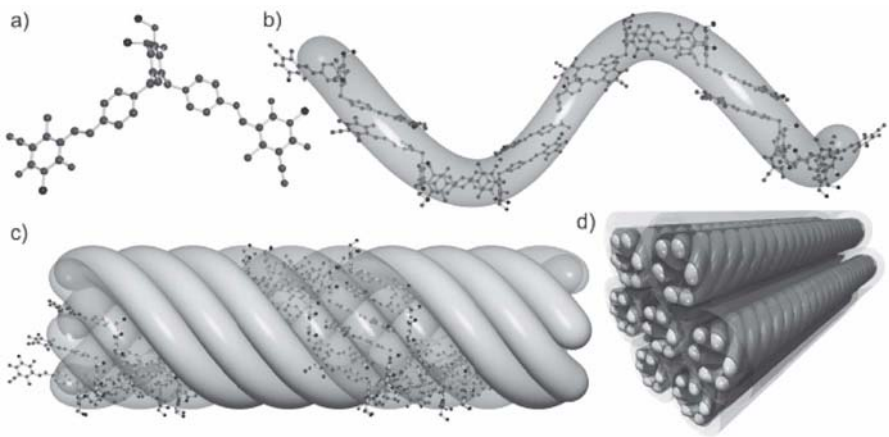
Fig. 4.25 Hierarchical transition of bis-merocyanin dyes (a) from the supramolecular polymer (b) over the rod (c) to a hexagonally ordered pillar (d) [70].

A
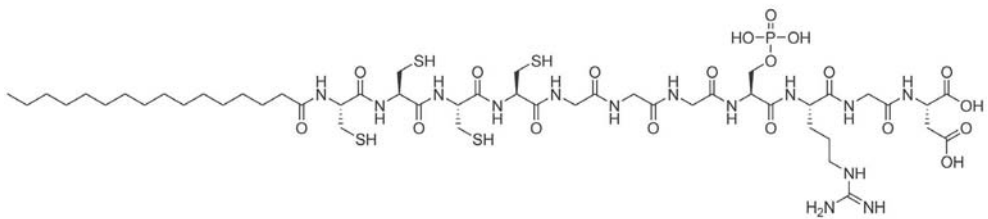
B
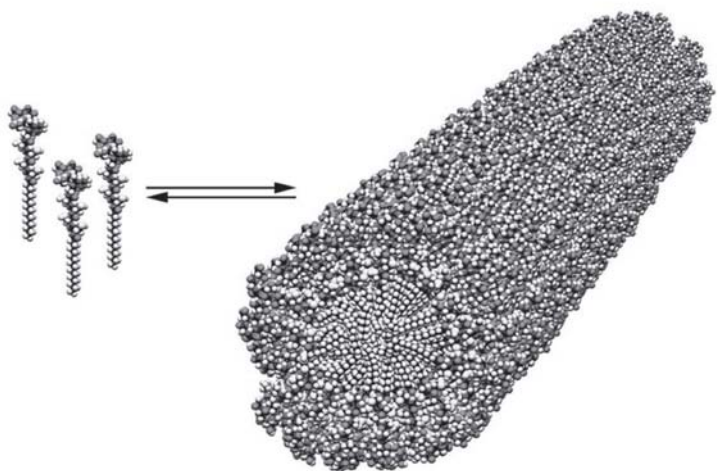
Fig. 4.26 Chemical structure of a peptide amphiphile (A) and the resulting aggregate caused by hydrophobic collapse of the lipid tail (B).

gates have interesting properties. For example, they can be used as scaffolds to direct the mineralization of hydroxyapatite to form composite materials similar to bone growth, where collagen fibrils serve the same function [72]. Also dendritic dipeptides have been shown to form large aggregates via self-assembly. In this case a Tyr-Ala dipeptide is functionalized via the tyrosine OH group with a Frechet-type dendron which long alkyl chains attached to its end. Depending on the stereochemistry of the dipeptide, allosteric self-assembly into helical porous structures occurs [73].

4.5
Applications of Foldamer Hybridization

One reason for self-assembly has already been pointed out in the beginning. Through the noncovalent interaction of smaller building blocks, more complex structures are obtained which display properties not present in the underlying molecules themselves [1, 2, 3, 6]. Hence, supramolecular synthesis provides an alternative approach to more complex systems besides covalent chemistry. The synthesis of smaller building blocks which then self-assemble to produce the desired structure can be easier, less challenging and more economical than to directly synthesize one large molecule with the same properties. An illustrative example is the leucine zipper described in the beginning. Only the synthesis of one peptide is needed which then spontaneously self-assembles to form the Y-shaped dimer with the DNA binding site in the junction. Rather complex structures can thus be obtained from rather simple building blocks [17].

Another even more important aspect of hybridization is information storage and transfer! As had already been pointed out before, hybridization is often very specific if e.g. controlled by H-bonds. This means, that the sequence (e.g. the pattern of H-bond donors and acceptors) of one strand completely determines the sequence of the second strand with which hybridization can occur. This holds both for homodimerization and for heteroduplex formation [14b]. If not the structure but rather the information itself is important, hybridization thus offers a way for safely storing and also reproducing this information. This is what the DNA double helix is for. The linear sequence of the strand contains the genetic information needed to produce proteins. The double helix is the storage form for this information. As two complementary strands are present, both with the same information (otherwise they would not hybridize to form a stable double helix) the risk of information loss due to chemical mutations etc. is greatly diminished [74]. There is always at least one safety copy of the genetic information present. Furthermore, when the information needs to be reproduced (as for cell division or protein production), each DNA strand serves as a templating matrix for either a new DNA strand or a mRNA. Thus the specificity of the hybridization pattern ensures correct information transfer. Nature uses this property of hybridization since the beginning of life most likely [16].

Of course chemists have strived to find artificial systems that also allow a controlled and selective reproduction of information encoded in the sequence of a
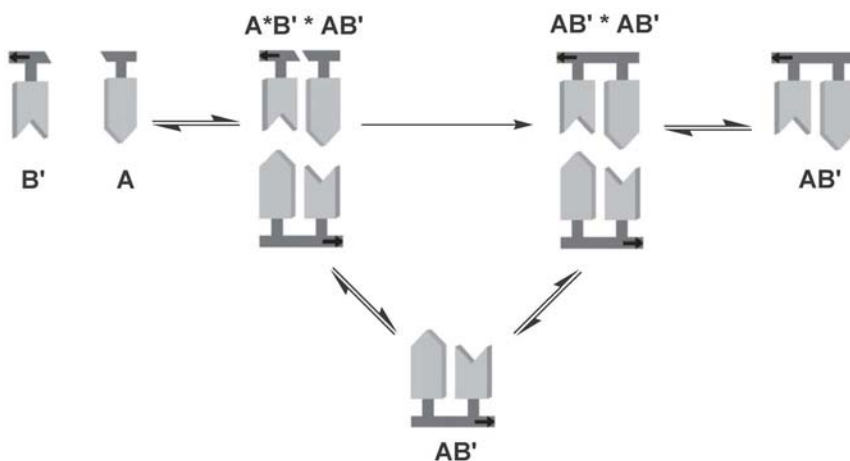
Fig. 4.27 Schematic representation of an autocatalytic self-replication cycle based on a self-complementary template AB′ and two monomers A and B′.

linear polymer. The initial break throughs were obtained with self-replicating nucleotides by Orgel [75] or von Kiedrowski [76]. An oligonucleotide serves as a template to assemble mononucleotides as dictated by the correct Watson–Crick base pairing pattern (Fig. 4.27). An enzyme-free chemical ligation then stitches the mononucleotides together to produce an oligonucleotide with a complementary sequence to the initial coding strand. If the initial oligonucleotide is composed of a self-complementary sequence, the product can also serve as a new template for the next cycle, hence, an autocatalytic system results.

Lateron also self-replicating peptides were designed for example by Ghadiri [77] or Chmielewski [78]. Ghadiri used a leucine zipper motif based on the GCN4 transcription factor. Small heptapeptide segments, which represent the repeating heptad unit of the leucine zipper, were aligned by a larger peptide template and then chemically ligated to produce a copy of the templating peptide strand. Again an autocatalytic self-replicating system resulted (Fig. 4.28). A rather major disadvantage of such systems is of course product inhibition, which reduces the catalytic turnover number. The final product strand forms a more stable complex with the initial template strand than the smaller peptide segments.

A bimolecular complex in general is more stable than a trimolecular one due to entropic reasons. Hence, to increase the autocatalytic efficiency and to approach exponential growth various modifications have been introduced into the initial self-replicating peptide systems such as the use of shorter peptides to destabilize the coiled-coil product and hence to facilitate its dissociation at ambient temperature. Also a proline-kink in the middle of the template strand can help to destabilize the coiled-coil product, whereas the two short peptide segments before ligation can bind to the either side of the kink without much problem. Recently, also
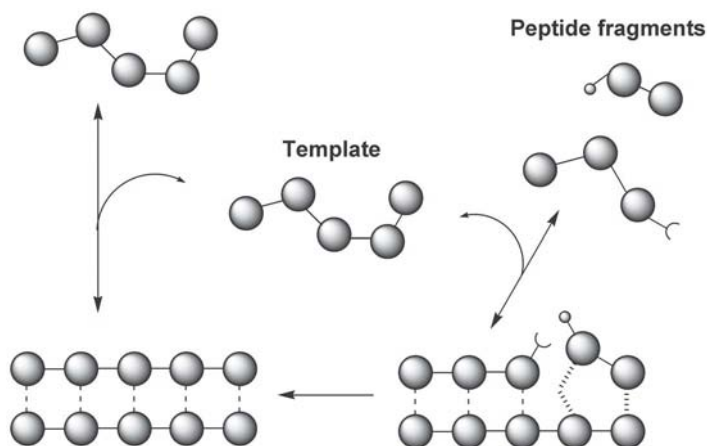
Fig. 4.28 Schematic presentation of a reaction cycle for a self-replicating peptide system. A template strand binds two smaller peptide fragments as dictated by the correct hybridization pattern. Chemical ligation of the two fragments then leads to another copy of the template even though real exponential growth as expected for such an autocatalysis is often hampered by product inhibition.

a self-replicating peptide/RNA system was designed which is based on the cross-hybridization between a RNA aptamer and a linear peptide strand [79]. The RNA aptamer was rationally cut into two halves and modified at each end to allow chemical re-ligation. The peptide strand as a template then binds the two halves of the aptamer, pre-orientates them and hence facilitates their re-ligation.

In a modification of this principle of self-replication, duplex formation between two complementary oligonucleotides can also be used to control a chemical reaction between more or less any two chemical entities A and B which can be attached to two hybridizing oligonucleotides. Duplex formation then brings the two reactants A and B into close proximity thereby increasing their effective molar concentration relative to the situation in free solution and thus enabling a chemical reaction between them. This so called DNA-templated synthesis has been advocated in recent years mainly by Liu and coworkers [80]. In various experiments they investigated the use of different oligonucleotide architectures for amine acylations, Wittig olefinations, 1,3 dipolar cycloadditions and reductive aminations and a variety of other reactions. The advantage of such DNA templated reactions compared to their counterparts in free solution is the control of selectivity. Only those two partners can react that are attached to the correct complementary strands as the reaction only occurs within the hybridized duplex strand (Fig. 4.29 A). For example, as much as 12 different reactants with normally incompatible chemical functionalities selectively form only 6 products if the reactants are attached to different but mutually complementary oligonucleotide sequences.
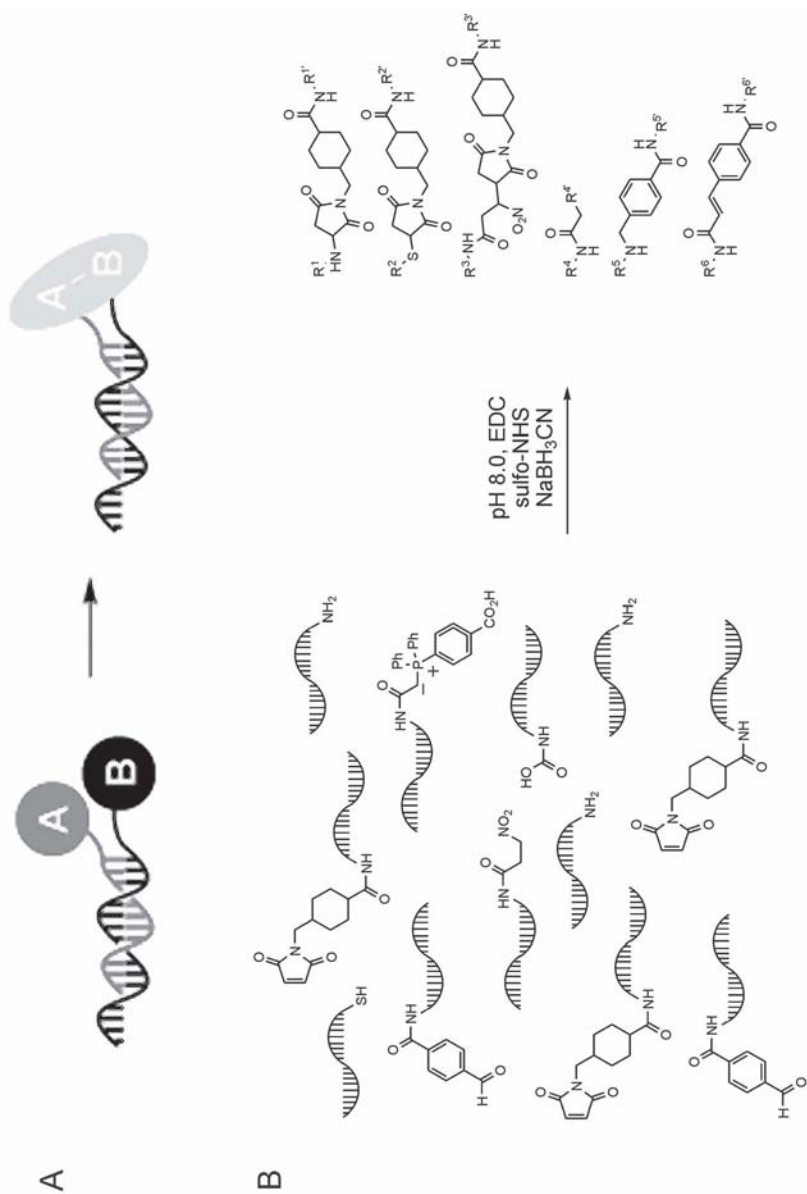
Fig. 4.29 Based on the selectivity imposed by DNA hybridization a mixture of 12 different reactants attached to small pieces of single stranded DNA selectively form only 6 reaction products instead of all possible 28 products that would form in free solution (B). Only those two reactants A and B that form a stable double helix can react (A).

4.6
Conclusion

Hybridization of foldamers not only leads to beautiful structures but actually can serve specific functions, such as information storage and transfer, but only if the hybridization is based on selective interactions. The design of artificial systems with specific and selective folding and hybridization properties remains an interesting and challenging goal for the chemical community. These artificial model systems will help us to better understand Nature and her elegant use of foldamers and their intermolecular interactions in many different applications. The more the underlying molecular principles are understood, the more sophisticated the deliberate use of artificial systems will be. This however requires a sound knowledge of noncovalent interactions, their strength and solvent dependence as well as their specificity and directionality as has been shown above for several examples. There is still a lot to learn despite all the progress that has been achieved in this area in the last few years. We will definitely see even more beautiful and intellectually challenging self-assembling foldamers with surprising properties and interesting applications in the future.

References

1 L. Stryer, *Biochemie*, Spektrum Akademischer Verlag, Heidelberg, **2003**.

2 C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York, **1991**.

3 a) C. Dobson, A. Sali, M. Karplus, *Angew. Chem.* **1998**, 110, 908–935; b) N. Ferguson, A. Fersht, *Curr. Opin. Struc. Biol.* **2003**, 13, 75–81.

4 N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, Wiley-VCH Verlag, Weinheim, **2002**.

5 B. Nölting, *Protein Folding Kinetics*, Springer-Verlag, Berlin, Heidelberg, **2006**.

6 R. Copeland, *Enzymes A Practical Introduction to Structure, Mechanism and Data Analysis*, Wiley-VCH Verlag, **1996**.

7 Even in fully flexible alkanes steric interactions such as the syn-pentane or the 1,3-allylic strain can induce specific conformations: a) R. W. Hoffmann, *Angew. Chem. Int. Ed.* **2000**, 39, 2054–2070; b) R. W. Hoffmann, *Chem. Rev.* **1989**, 89, 1841–1960.

8 a) S. Han, A. Bond, R. Disch, D. Holmes, J. Schulman, S. Teat, P. Vollhardt, G. Whitener, *Angew. Chem.* **2002**, 114, 3357–3361; b) S. Han, R. Anderson, A. Bond, H. Chu, R. Disch, D. Holmes, J. Schulman, S. Teat, P. Vollhardt, G. Withener, *Angew. Chem. Int. Ed.* **2002**, 41, 3227–3230.

9 a) C. Nuckolls, T. Katz, *J. Am. Chem. Soc.* **1998**, 120, 9541–9544; b) C. Nuckolls, T. Katz, G. Katz, P. Collings, L. Castellanos, *J. Am. Chem. Soc.* **1999**, 121, 79–88; c) S. Dreher, D. Weix, T. Katz, *J. Org. Chem.* **1999**, 64, 3671–3678; d) K. Paruch, T. Katz, C. Incarvito, K.-C. Lam, B. Rhatigan, A. Rheingold, *J. Org. Chem.* **2000**, 65, 7602–7608; e) T. Katz, *Angew. Chem.* **2000**, 112, 1997–1999; f) K. Phillips, T. Katz, S. Jockusch, A. Lovinger, N. Turro, *J. Am. Chem. Soc.* **2001**, 123, 11899–11907; g) D. Zhigang, T. Katz, J. Golen, A. Rheingold, *J. Org. Chem.* **2004**, 69, 7769–7771.

10 a) M. Miyasaka, A. Rajca, M. Pink, S. Rajca, *J. Am. Chem. Soc.* **2005**, 127, 13806–1307; b) S. Collins, A.

Grandbois, M. Vachon, J. Côté, *Angew. Chem.* **2006**, 118, 2989–2992; c) D. Harrowven, I. Guy, L. Nanson, *Angew. Chem. Int. Ed.* **2006**, 45, 2242–2245.

11 a) M. Stone, J. Fox, J. Moore, *Org. Lett.* **2004**, 6, 3317–3320; b) J. Moore, C. Ray, *Adv. Poly. Sci.* **2005**, 177, 91–149; c) M. Stone, J. Heemstra, J. Moore, *Acc. Chem. Res.* **2006**, 39, 11–20.

12 X. Yang, L. Yuan, K. Yamato, A. Brown, W. Feng, M. Furukawa, X. Zeng, B. Gong, *J. Am. Chem. Soc.* **2004**, 126, 3148–3162.

13 S. Gellman, *Acc. Chem. Res.* **1998**, 31, 173–180.

14 a) L. Yuan, A. Sanford, W. Feng, A. Zhang, J. Zhu, H. Zeng, K. Yamato, M. Li, J. Ferguson, B. Gong, *J. Org. Chem.* **2005**, 70, 10660–10669; b) A. Sanford, K. Yamato, X. Yang, L. Yuan, Y. Han, B. Gong, *Eur. J. Biochem.* **2004**, 271, 1416–1425; c) L. Yuan, H. Zeng, K. Yamato, A. Sanford, W. Feng, H. Atreya, D. Sukumaran, T. Szyperski, B. Gong, *J. Am. Chem. Soc.* **2004**, 126, 16528–16537.

15 a) H.-P. Yi, X.-B. Shao, J.-L. Hou, C. Li, X.-K. Jiang, Z.-T. Li, *New J. Chem.* **2005**, 29, 1213–1218; b) C. Li, S.-F. Ren, J.-L. Hou, H.-P. Yi, S.-Z. Zhu, X.-K. Jiang, Z.-T. Li, *Angew. Chem. Int. Ed.* **2005**, 44, 5725–5729.

16 W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, **1984**.

17 a) W. Landschulz, P. Johanson, S. McKnight, *Science*, **1988**, 240, 1759–1764: b) E. O'Shea, R. Rutkowski, P. Kim, *Science*, **1989**, 243, 538–542; c) M. Weiss, T. Ellenberger, C. Wobbe, J. Lee, S. Harrison, K. Struhl, *Nature* **1990**, 347, 575–578; d) T. Ellenberger, *Curr. Opin. Struc. Biol.* **1994**, 4, 12–21.

18 a) H. Wendt, L. Leder, H. Härmä, I. Jelesarov, A. Baici, H. Bosshard, *Biochemistry* **1997**, 36, 204–213; b) A. Dragan, A. Privalov, *J. Mol. Biol.* **2002**, 321(5), 891–908.

19 Y. Tang, G. Ghirlanda, W. Petka, T. Nakajima, W. DeGrado, D. Tirell, *Angew. Chem. Int. Ed.* **2001**, 40, 1494–1496.

20 N. C. Yoder, K. Kumar, *Chem. Soc. Rev.* **2002**, 31, 335–341.

21 V. Percec, M. Glodde, G. Johansson, V. Balagurusamy, P. Heiney, *Angew. Chem. Int. Ed.* **2003**, 42, 4338–4342.

22 K. Hoang, S. Mecozzi, *Langmuir* **2004**, 20, 7347–7350.

23 L. Barthel-Rosa, J. A. Gladysz, *Coord. Chem. Rev.* **1999**, 190–192, 587–605.

24 N. Schnarr, A. Kennan, *J. Am. Chem. Soc.* **2001**, 123, 11081–11082.

25 L. Gonzalez, J. J. Plecs, T. Alber, *Nat. Struct. Mol. Biol.* **1996**, 3, 510–515.

26 a) S. A. Benner, *Acc. Chem. Res.* **2004**, 37, 784–797; b) E. T. Kool, *Acc. Chem. Res.* **2002**, 35, 936–943.

27 H. Jiang, V. Maurizot, I. Huc, *Tetrahedron* **2004**, 60, 10029–10038.

28 V. Maurizot, J.-M. Léger, P. Guionneau, I. Huc, *Russ. Chem. Bull. Int. Ed.* **2004**, 53, 1572–1576.

29 S. Fernandez-Lopez, H.-S. Kim, E. Chol, M. Delgado, J. Granja, A. Khasanov, K. Kraehenbuehl, G. Long, D. Weinberger, K. Wilcoxon, R. Ghadiri, *Nature* **2001**, 421, 452–455.

30 V. Semetey, C. Didierjean, J.-P. Briand, A. Aubry, G. Guichard, *Angew. Chem. Int. Ed.* **2002**, 41, 1895–1898.

31 M. Amorín, L. Castedo, J. Granja, *J. Am. Chem. Soc.* **2003**, 125, 2844–2845.

32 A. Petitjean, L. Cuccia, J.-M. Lehn, M. Schmutz, *Angew. Chem. Int. Ed.* **2002**, 41, 1195–1198.

33 a) V. Berl, I. Huc, R. Khoury, M. Krische, J.-M. Lehn, *Nature* **2000**, 407, 720–723; b) V. Berl, I. Huc, R. Khoury, J.-M. Lehn, *Chem. Eur. J.* **2001**, 7, 2798–2809; c) V. Berl, I. Huc, R. Khoury, J.-M. Lehn, *Chem. Eur. J.* **2001**, 7, 2810–2820.

34 T. ten Cate, R. Sijbesma, *Macromol. Rapid Commun.* **2002**, 23, 1094–1112.

35 M. Mayer, S. Nakashima, S. Zimmermann, *Org. Lett.* **2005**, 7, 3005–3008.

36 C. Schmuck, W. Wienand, *J. Am. Chem. Soc.* **2003**, 125, 452–459.

37 S. Schlund, C. Schmuck, B. Engels, *J. Am. Chem. Soc.* **2005**, 127, 11115–11124.

38 H. Zeng, X. Yang, A. Brown, S. Martinovic, R. Smith, B. Gong, *Chem. Commun.* **2003**, 1556–1557.

39 C. S. Wilcox in *Frontiers in Supramolecular Organic Chemistry and Photochemistry* (Eds: H. J. Schneider,

H. Dürr), Wiley-VCH, Weinheim, **1991**, 123–143.

40 H. Zeng, R. Miller, R. Flowers, B. Gong, *J. Am. Chem. Soc.* **2000**, 122, 2635–2644.

41 H. Zeng, H. Ickes, R. Flowers, B. Gong, *J. Org. Chem.* **2001**, 66, 3574–3583.

42 a) D. Williams, B. Bardsley, *Angew. Chem. Int. Ed.* **1999**, 38, 1172–11193; b) C. Walsh, S. Fisher, I.-S. Park, M. Prahalad, Z. Wu, *Chem. Biol.* **1996**, 3, 21–28.

43 a) P. Loll, J. Kaplan, B. Salinsky, P. Axelsen, *J. Med. Chem.* **1999**, 42, 4714–4719; b) C. McComas, B. Crowley, D. Boger, *J. Am. Chem. Soc.* **2003**, 125, 9314–9315.

44 a) W. L. Jørgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, 112, 2008–2010; b) J. Pranata, S. G. Wierschke, W. L. Jørgensen, *J. Am. Chem. Soc.* **1991**, 113, 2810–2819; c) W. L. Jørgensen, D. L. Severance, *J. Am. Chem. Soc.* **1991**, 113, 209–216; d) O. Lukin, J. Leszczynski, J. Phys. Chem. A **2002**, 106, 6775–6782; e) O. Lukin, J. Leszczynski, J. Phys. Chem. A **2003**, 107, 9251–9252.

45 C. Schmuck, W. Wienand, *Angew. Chem. Int. Ed.* **2001**, 40, 4363–4369.

46 E. Archer, M. Krische, *J. Am. Chem. Soc.* **2002**, 124, 5074–5083.

47 H. Gong, M. Krische, *J. Am. Chem. Soc.* **2005**, 127, 1719–1725.

48 a) D. Williams, M. Westwell, *Chem. Soc. Rev.* **1998**, 27, 57–63; b) A. Davis, R. Wareham, *Angew. Chem. Int. Ed.* **1999**, 38, 2978–2996.

49 a) C. Piguet, M. Borovec, J. Hamacek, K. Zeckert, Coor. *Chem. Rev.* **2005**, 249, 705–726; b) M. Albrecht, *Top. Curr. Chem.* **2004**, 248, 105–139; c) M. Albrecht, *Chem. Rev.* **2001**, 101, 3457–3497; d) G. Swiegers, T. Malefetse, *Chem. Rev.* **2000**, 100, 3483–3537.

50 A. Oleksi, A. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. Hannon, M. Coll, *Angew. Chem. Int. Ed.* **2006**, 45, 1227–1231.

51 a) Y. Tanaka, H. Katagiri, Y. Furusho, E. Yashima, *Angew. Chem. Int. Ed.* **2005**, 44, 3867–3870; b) M Ikeda, Y. Tanaka, T. Hasegawa, Y. Furusho, E.

Yashima, *J. Am. Chem. Soc.* **2006**, 128, 6808–6807.

52 In contrast to Coulomb interactions between two point charges which are of course non-directional, charge interactions with organic molecules such as guanidinium cations are directional. This can be due to the formation of H-bond enforced ion pairs and/or the anisotropic solvation of such ions: P. E. Mason, G. W. Neilson, J. E. Enderby, M.-L. Saboungi, C. E. Dempsey, A. D. MacKerell, Jr., J. W. Brady, *J. Am. Chem. Soc.* **2004**, 126, 11462–114710; For a general discussion of the directionality of noncovalent interactions see: J. P. Glusker, *Top. Curr. Chem.* **1998**, 198, 1–56.

53 A. Bisson, F. Carver, D. Eggelston, R. Haltiwanger, C. Hunter, D. Livingston, J. McCabe, C. Rotger, A. Rowan, *J. Am. Chem. Soc.* **2000**, 122, 8856–8868.

54 For reviews on cooperativity, see: a) D. H. Williams, E. Stephens, D. P. O'Brien, M. Zhou, *Angew. Chem. Int. Ed.* **2004**, 43, 6596–6616; b) C. A. Hunter, S. Tomas, *Chem. Biol.* **2003**, 10, 1023–1032; c) G. Ercolani, *J. Am. Chem. Soc.* **2003**, 125, 16097–16103; d) S. L. Tobey, E. V. Anslyn, *J. Am. Chem. Soc.* **2003**, 125, 10963–10970.

55 G. Gabriel, B. Iverson, *J. Am. Chem. Soc.* **2002**, 124, 15174–15175.

56 For reviews on hydrophobic contacts see: a) B. Widom, P. Bhimalapuram, K. Koga, Phys. Chem. Chem. Phys. **2003**, 5, 3085–3093; b) L. R. Pratt, A. Pohorille, *Chem. Rev.* **2002**, 102, 2671–2691; c) N. T. Southall, K. A. Dill, A. D. J. Haymet, *J. Phys. Chem. B* **2002**, 106, 521–533.

57 G. Gabriel, S. Sorey, B. Iverson, *J. Am. Chem. Soc.* **2005**, 127, 2637–2640.

58 Q.-Z. Zhou, X.-K. Jiang, X.-B. Shao, G.-J. Chen, M.-X. Jia, Z.-T. Li, *Org. Lett.* **2003**, 5, 1955–1958.

59 M. Chubberly, B. Iverson, *J. Am. Chem. Soc.* **2001**, 123, 7560–7563.

60 J.-L. Hou, M.-X. Jia, X.-K. Jiang, Z.-T. Li, G.-J. Chen, *J. Org. Chem.* **2004**, 69, 6228–6237.

61 H. B. Gamper, K. Arar, A. Gewirtz, Y. M. Hou, *Biochemistry* **2006**, 45, 6978–6986.

62 a) A. Eschenmoser, *Chimia* **2005**, 59, 836–850; b) A. Eschenmoser, *Chem. Commun.* **2004**, 1247–1252.

63 L. Zhang, A. E. Peritz, P. J. Carrol, E. Meggers, *Synthesis* **2006**, 4, 645–653.

64 a) U. Diederichsen, *Angew. Chem. Int. Ed.* **1996**, 35, 445–448; b) U. Diederichsen, *Angew. Chem. Int. Ed.* **1997**, 36, 1886–1889.

65 For reviews on *β*-peptides see: a) D. Seebach, T. Kimmerlin, R. Sebesta, M. A. Campo, A. K. Beck, *Tetrahedron* **2004**, 60, 7455–7506; b) R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, 101, 3219–3232.

66 a) A. Brückner, P. Chakraborty, S. Gellman, U. Diederichsen, *Angew. Chem. Int. Ed.* **2003**, 42, 4395–4399; b) P. Chakraborty, U. Diederichsen, *Chem. Eur. J.* **2005**, 11, 3207–3216.

67 T. Raguse, J. Lai, P. LePlae, S. Gellman, *Org. Lett.* **2001**, 3, 3963–3966. The structure of this amphiphilic 14-helix depends very much on the influence of the solvent and the used buffer. In 10 mM TRIS with a pH of 8.0 the system self-associates most as a hexamer at a concentration of 1.7 mM. If an acetate buffer is used and the pH is adjusted to a value of 3.8 no clear classification can be done, except for the monomer at a concentration of 0.3 mM.

68 H.-C. Yang, S.-Y. Lin, H.-C. Yang, C.-L. Lin, L. Tsai, S.-L. Huang, I. Chen, C.-H. Chen, B.-Y. Jin, T.-Y. Luh, *Angew. Chem. Int. Ed.* **2006**, 45, 726–730.

69 a) L. Brunsveld, B. Folmer, E. Meijer, R. Sijbesma, *Chem. Rev.* **2001**, 101, 4071–4097; b) F. Hoeben, P. Jonkheijm, E. Meijer, A. Schenning, *Chem. Rev.* **2005**, 105, 1491–1546; c) A. Schenning, P. Jonkeijm, E. Peeters, E. Meijer, *J. Am. Chem. Soc.* **2001**, 123, 409–416.

70 a) F. Würthner, S. Yao, U. Beginn, *Angew. Chem. Int. Ed.* **2003**, 115, 3368–3371; b) A. Lohr, M. Lysetska, F. Würthner, *Angew. Chem. Int. Ed.* **2005**, 44, 5071–5074.

71 J. Hartgerink, E. Beniash, S. Stupp, *Science* **2001**, 294, 1684–1688.

72 a) K. Sugiyasu, S.-I. Tamaru, M. Masayuki, D. Berthier, I. Huc, R. Oda, S. Shinkai, *Chem. Commun.* **2002**, 1212–1213; b) A. Sugawara, S. Yamane, K. Akiyoshi, *Macromol. Rapid. Commun.* **2006**, 27, 441–446; c) W. Zhang, S. Liao, F. Cui, Chem. Mater. **2003**, 15, 3221–3226; d) T. Iwatsubo, K. Sumaru, T. Kanamori, T. Shinbo, T. Yamaguchi, *Biomacromolecules* **2006**, 7, 95–100.

73 a) V. Percec, A. E. Dulcey, M. Peterca, M. Illies, J. Ladislaw, B. M. Rosen, U. Edlund, P. A. Heiney, *Angew. Chem. Int. Ed.* **2005**, 44, 6516–6521; b) M. Peterca, V. Percec, A. E. Dulcey, S. Nummelin, S. Korey, M. Illies, P. A. Heiney, *J. Am. Chem. Soc.* **2006**, 128, 6713–6720.

74 D. M. Noll, T. M. Mason, P. S. Miller, *Chem. Rev.* **2006**, 106, 277–301.

75 a) W. Zielinski, L. Orgel, *Nature* 1987, 327, 346–347; b) I. Kozlov, P. Politis, A. van Aerschot, R. Busson, P. Herdewijn, L. Orgel, *J. Am. Chem. Soc.* **1999**, 121, 2653–2656.

76 a) D. Sievers, G. von Kiedrowski, *Chem. Eur. J.* **1998**, 4, 629–641; b) L. Eckardt, K. Naumann, W. Pankau, M. Rein, M. Schweitzer, N. Windhab, G. von Kiedrowski, *Nature*, **2002**, 420, 286; c) G. von Kiedrowski, *Angew. Chem. Int. Ed.* **1986**, 25, 932–935.

77 a) D. Lee, J. Granja, J. Martinez, K. Severin, R. Ghadiri, *Nature* **1996**, 382, 525–528; b) K. Severin, D. H. Lee, J. A. Martinez, M. R. Ghadiri, *Chem. Eur. J.* **1997**, 3, 1017–1024; c) A. Saghatelian, Y. Yokobayashi, K. Solthani, M. R. Ghadiri, *Nature* **2001**, 409, 797–801.

78 a) I. Gosh, J. Chmielewski, *Curr. Op. Chem. Biol.* **2004**, 8, 640–644; b) N. Paul, G. F. Joyce, *Curr. Opin. Chem. Biol.* **2004**, 8, 634–639; c) X. Li, J. Chmielewski, *J. Am. Chem. Soc.* **2003**, 125, 1182–11821; d) R. Issac, Y.-W. Ham, J. Chmielewski, *Curr. Opin. Chem. Biol.* **2001**, 11, 458–463.

79 M. Levi, A. Ellington, *J. Mol. Evol.* **2003**, 56, 607–615.

**80** X. Li, D. R. Liu, *Angew. Chem. Int. Ed.* **2004**, 116, 4956–4979.

5
Control of Polypeptide Chain Folding and Assembly

Rajkishor Rai and Padmanabhan Balaram

5.1
Introduction

Natural protein chains fold into precise three-dimensional structures by a mechanistically complex process, in which the energetics of the folded states and the rates of their formation are controlled by a delicate balance of intrachain interactions and solvent forces. Approaches to first principles design (*de novo* design), of synthetic polypeptides mimicking structural features found in proteins [1], have relied on strategies to control local chain folding via short and medium range interactions [2, 3]. In some design strategies, patterning of hydrophobic and hydrophilic residues along the sequence helps to drive folding in a predetermined direction by using the hydrophobic effect, in aqueous solution, as the driving force [4–5]. The formation of well-defined local structures (secondary structure elements) has also been achieved using lessons gained from inspection of the large body of available protein three-dimensional structures [6]. The use of Asn-Gly, a β-turn forming segment, to nucleate β-hairpin formation [7–9] and the use of cross-strand Trp–Trp interactions to stabilize antiparallel strands are examples of observations made in proteins that have been effectively exploited in the construction of synthetic polypeptide motifs [10].

This chapter considers an alternative approach in which peptide "foldamer" design is based on the use of non-proteinogenic amino acids to impose local backbone folding constraints, thereby directing the course of chain folding, by biasing the choice of local conformations. It may be parenthetically noted that the term "foldamer" is of relatively recent origin, having been introduced to describe the regular structures of homologated peptide backbones formed by oligomers of the higher homologs of the conventional α-amino acids [11]. The folded structures of polypeptide chains are most conveniently described by using Ramachandran (torsion) angles ϕ and ψ at each residue [12]. In protein structures, the backbone conformations of individual residues are characterized by the regions of ϕ,ψ space in which they are located. The Ramchandran map delineates sterically allowed regions of conformational space. Regular structures like helices and sheets
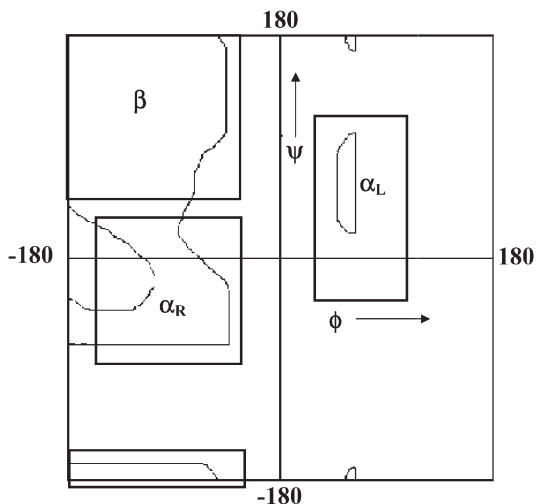
Fig. 5.1 Ramachandran map for the L-alanine residue. The dark boxes enclose the regions of ϕ,ψ space that may be broadly classified as right-handed helical (α_R), left-handed helical (α_L) and extended (β) regions. The extended β-region may be further subdivided into polyproline II (P_{II}) and β-sheet (parallel and antiparallel) secondary structures.

are conveniently identified by a clustering of successive residues in specific regions of ϕ,ψ space (Fig. 5.1).

Figure 5.2 illustrates the distribution of non-Gly and Gly residues in a data set of high resolution protein structures [13]. Glycine is unique amongst the amino acid constituents of proteins, being the smallest in size and also the only achiral residue. These features result in a much larger region of ϕ,ψ space being sterically accessible. The Ramachandran allowed regions are symmetric with respect to origin of the ϕ,ψ map, permitting Gly residue to adopt local conformations which are sterically forbidden for the other 19 L-amino acids. The scatter plot for the non-Gly residues shows a distribution, which is largely concentrated in three broad regions denoted as α_R, β/P_{II} and α_L, which correspond to the right-handed α-helical, extended strand and left-handed helical regions of ϕ,ψ space, respectively. The distribution in the extended region (($\phi \sim -130 \pm 50°$) may be subdivided into two clusters corresponding to β strands ($\beta\phi \sim -120°$) and polyproline II ($P_{II}\phi \sim -60°$). Most individual residues in proteins do not have a strong intrinsic preference for either the helical or extended strand regions, limiting the use of statistically determined propensities in the design of local structures. Two residues, for which conformational biases are significant, are L-Pro and L-Asn. The observed distribution for L-Pro residues in proteins (Fig. 5.2) highlights the narrow distribution of ϕ values ($\phi = -60 \pm 20°$), which results from the constraints imposed by the pyrrolidine ring. Two distinct clusters corresponding to P_{II} ($\phi = -60°$, $\psi = +120°$) and α_R ($\phi = -60°$, $\psi = -30°$) are observed. The gamma turn or C_7 conformation ($\phi = -70°$, $\psi = +70°$) is sparsely populated. Amongst the genetically coded amino acids, the proline residue provides the greatest op-

Non-Gly

Gly
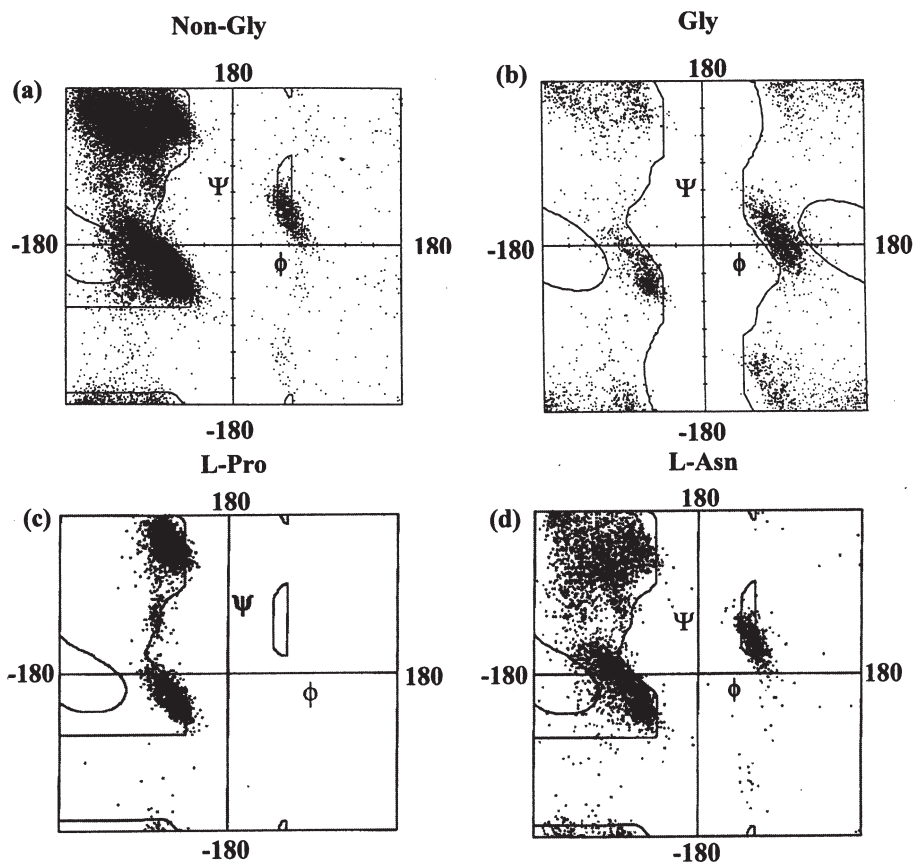
L-Pro

L-Asn

Fig. 5.2 (a) Distribution of ϕ,ψ values of non-Gly residues from 250 protein structures ($\leq$2.0 Å); (b) Distribution of ϕ,ψ values of Gly residues from 250 protein structures ($\leq$2.0 Å). (c and d) Distribution of conformational angles of L-Pro and L-Asn residues from 538 independent protein crystal structures [13]. The protein data set used to generate this figure was derived from the Protein Data Bank using a resolution cutoff of 2.0 Å and a sequence homology cutoff of 40% and contained a total of 47612 non-glycine, 4933 glycine, 4995 proline and 5503 asparagine residues.

portunity for directing local chain folding. The chirality of the L-amino acids in proteins greatly restricts the sterically allowed regions with positive values of ϕ. Asn is the only chiral residue with a high propensity to occur in the left-handed helical (α_L, $\phi = +60°$, $\psi = +30°$) conformations. This local feature is important in generating 'prime' turns (type I′/II′), which in turns act as a nucleus for the formation of β-hairpin structures. A convincing theoretical rationalization of the tendency of Asn residues to favor left-handed helical conformations is still to be achieved. It may be noted that Jane Richardson, in an insightful analysis of protein structures, remarked that by virtue of its CH_2-$CONH_2$ side-chain Asn is the 'least chiral' of the amino acids [14]. An analysis of individual residue conformations in proteins suggests that conformational choices accessible to each amino

acid are wide enough to result in a very large region of sterically allowed structure space, for polypeptides of even limited chain length. Control of chain folding is, thus, a formidable task. One synthetic approach to peptide and protein design involves the use of nonprotein residues to act as local directors of chain folding, restricting available conformations to limited region of Ramachandran space [15].

The use of conformationally constrained residues in the design of folded peptide structures is based on the simple principle that the residue choice must limit conformational excursions to a well-defined region of ϕ,ψ space. The subsequent sections of this chapter illustrate the approaches used in engineering the construction of peptide helices and hairpins, followed by their assembly into mimics of supersecondary structures.

5.2
Helix Promotion by Backbone Substitution

5.2.1
α-Aminoisobutyric Acid (Aib) and Related Dialkyl Amino Acids

Aib is a constituent of several naturally occurring antimicrobial peptides. Notably, Aib is a major constituent of a class of linear peptide antibiotics, produced by soil
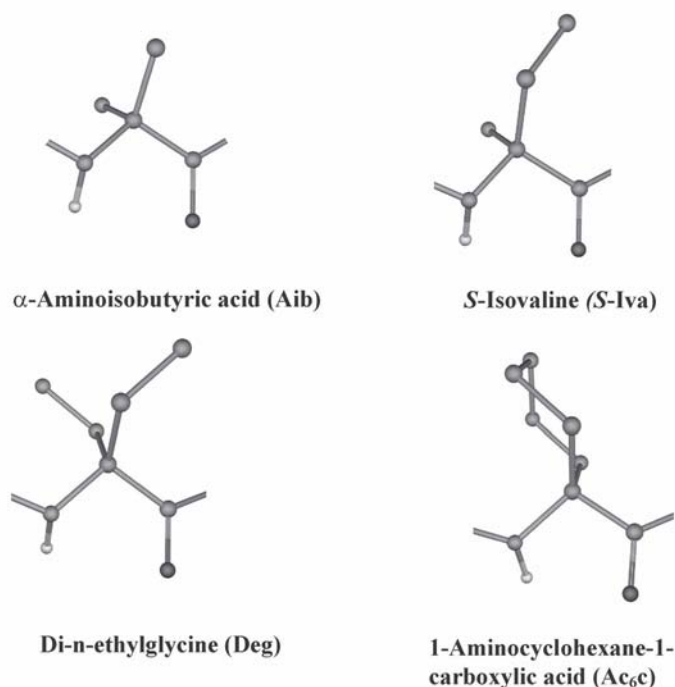

α-Aminoisobutyric acid (Aib)

S-Isovaline *(S*-Iva)

Di-n-ethylglycine (Deg)

1-Aminocyclohexane-1-carboxylic acid (Ac$_6$c)

Fig. 5.3 Chemical structures of some representative α,α-dialkylglycines.

fungi; *Trichoderma* species being the most widely studied. The conformational properties of Aib were uncovered in early structural investigations of synthetic peptide fragments of antibiotics of fungal origin [16], which have later been termed as peptaibols or peptaibiotics [17]. The related chiral residue isovaline (α-methyl-α-ethyl glycine) is also a constituent of these natural products, occurring almost exclusively as *S*-isovaline [18]. A large number of chiral and achiral dialkyl glycines have been synthetically produced (see Fig. 5.3 for representative structures). The presence of a tetrasubstituted C^α-atom in the backbone of polypeptide chains results in a dramatic reduction in 'allowed'conformational space'.

The sterically accessible regions of ϕ,ψ space for Aib may be readily derived by examining the regions of overlap of the Ramachandran map for L-Ala and D-Ala, which are related by inversion about the origin (Fig. 5.4). The two distinct areas of overlap are small and restricted to the right- (α_R) and left- (α_L) handed helical regions, leading to the simple conclusion that Aib and related $C^{\alpha,\alpha}$-dialkyl amino acids may be expected to be strongly helix stabilizing. This expectation has been borne out by a large body of investigations on peptide containing Aib and related residues. Indeed, the largest body of crystal structures of oligopeptides available to date is on sequences containing Aib residues [19–23]. Helical folding in host amino acid sequences containing very few guest Aib residues has been repeatedly demonstrated, suggesting that helix nucleation and stabilization may be readily achieved by strategic incorporation of these conformationally constrained amino acids. Figure 5.5 shows three examples of helices constructed in stable sequences. Both 3_{10} and α-helices may be generated. Mixed helical structures are relatively
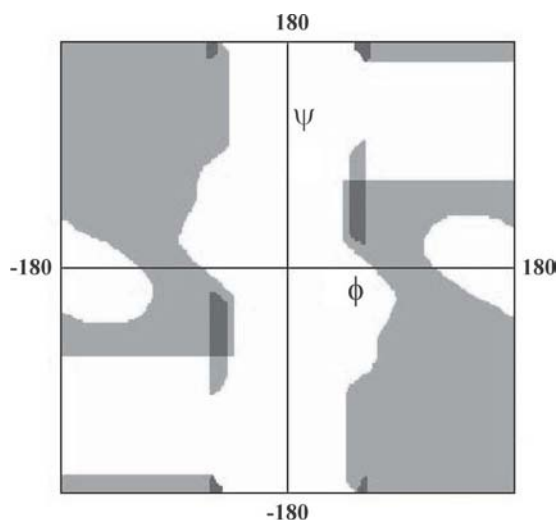
Fig. 5.4 Overlapped Ramachandran maps for *N*-acetyl-L-Ala-*N'*-methylamide and *N*-acetyl-D-Ala-*N'*-methylamide. The shaded regions are stereochemically allowed for both L- and D-Ala and these correspond to the ''allowed'' regions for *N*-acety-Aib-*N'*-methylamide.
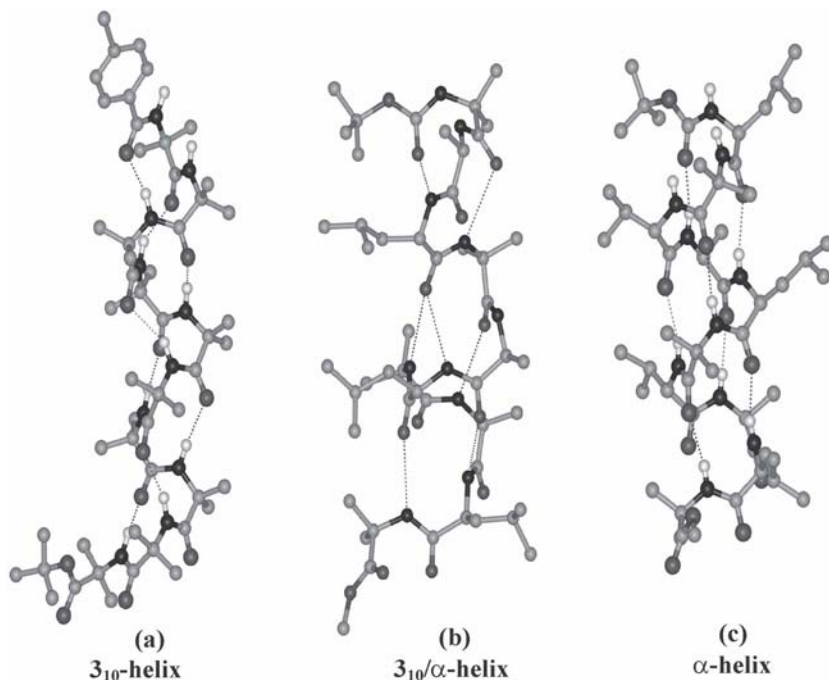
Fig. 5.5 Molecular conformations observed in the crystals of synthetic peptides (a) pBrBz-(Aib)$_{10}$-OtBu [22b]; (b) Boc-Aib-(Ala-Leu-Aib)$_3$-OMe [23]. (c) Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-Ala-Leu-Aib-OMe [21b].

common. In solution, the barriers to helix interconversions, within the limited region of ϕ,ψ space defined for 3_{10} and α-helices, are likely to be small.

The introduction of a few Aib residues is sufficient to stabilize helices in sequences of length upto 20 residues. The structure of a 19-residue peptide containing as many as three internally positioned D-residues provides an example of helix promotion by a guest Aib residue [24]. The crystal structures of normally occurring peptaibols have also established that helical folding is maintained even when several internal proline residues are present. Solid-state conformations of zervamicin and antiamoebin (Fig. 5.6) are examples when an Aib-Pro/Hyp C-terminus segment adopts the β-bend ribbon structure, in which all residues adopt ϕ,ψ values in the helical region of ϕ,ψ space [18b, d].

5.2.2
Diproline Segments

An alternate approach towards nucleating helical folding is the use of diproline templates, positioned at the amino terminus end of synthetic sequences. This approach developed by Kemp and co-workers [25] and subsequently extended by
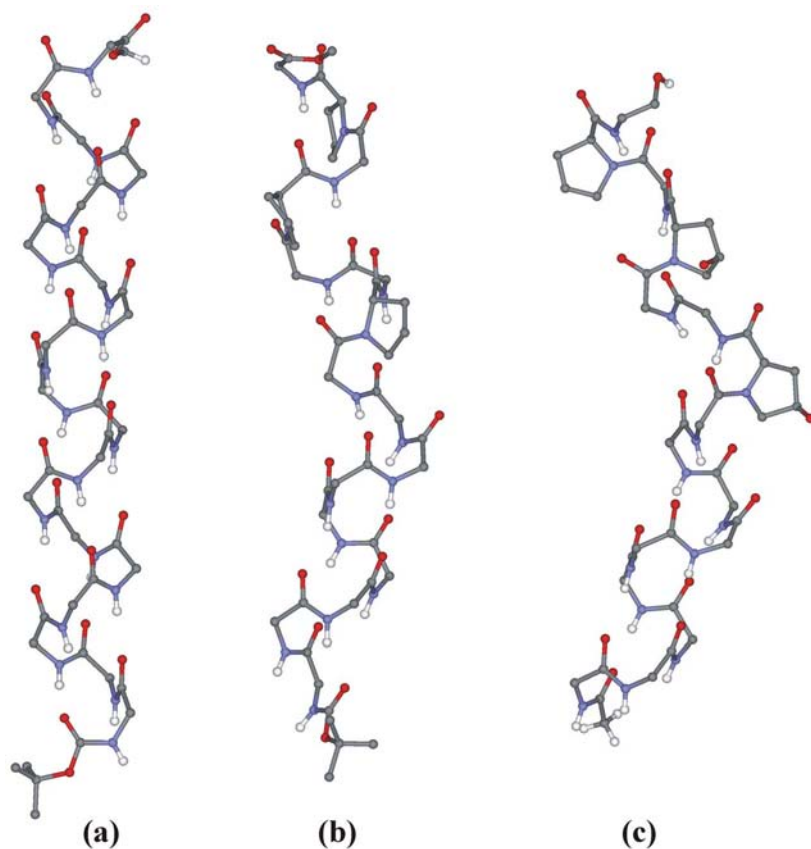
Fig. 5.6 Molecular conformations observed in the crystals of peptides:
(a) Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-^{D}Ala-^{D}Leu-Leu-Val-Phe-Val-Aib-^{D}Val-
Leu-Phe-Val-Val-OMe [24]; (b) Zervamicin analog [18b]; (c) Antiamoebin
[18d].

the group of Hanessian [26], is based on an early observation that the diproline-
containing sequences can form incipient 3_{10}-helical structures in organic solvents,
where consecutive type III β-turn formation is driven by successive hydrogen
bond formation [27]. Figure 5.7a shows the consecutive type III β-turn structure
proposed for the model peptide Pivaloyl-Pro-Pro-^{L}Ala-NHMe. Here, the Pro2 res-
idue occupies the $i + 2$ position of the first β-turn and $i + 1$ position of the second
β-turn. This motif, stabilized by two successive C_{10} hydrogen bonds, constitutes a
single turn of a 3_{10} helical structure in which the torsion angles of all three resi-
due lies in the α_R $(-60°, -30°)$ [27]. Bridging the C^γ atom of Pro(1) and the C^δ
atom of Pro(2) by a thiomethylene group constrains diproline conformations to
local helical structures at both proline residues, thus providing a rigid template
for helix nucleation in attached peptide segments. Extensions of this approach to
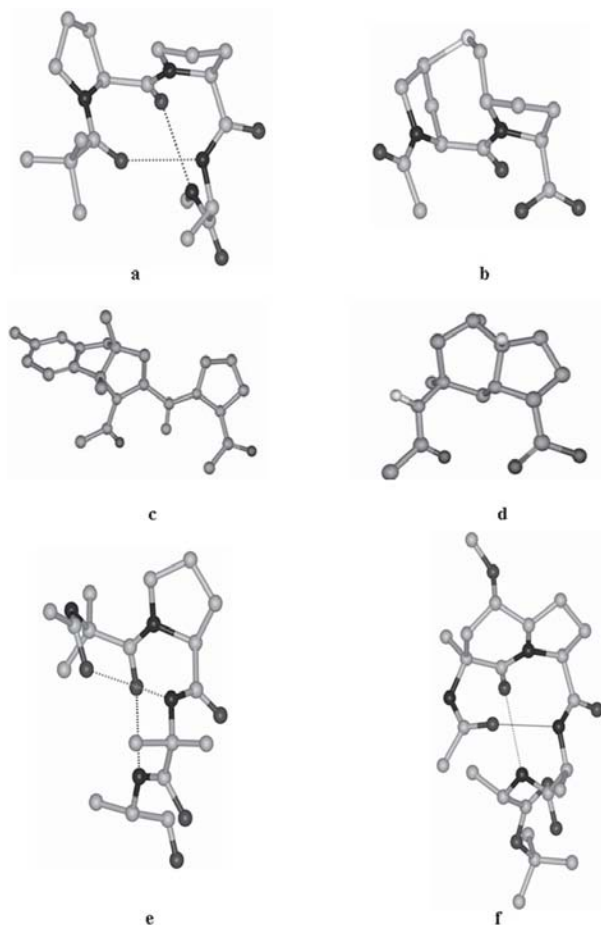
Fig. 5.7 Templates, used to nucleate helical structures (a) NMR model of Piv-[L]Pro-[L]Pro-[L]Ala NHMe [27]; (b) Structure of Kemp's template (2S,5S,8S,-11S)-1-acetyl-1,4-diaza-3-keto-5-carboxy-10-thiatricyclo [2.8.1.0[4, 8]]-tridecane(Ac-Hel1-OH) in crystals [25b]; (c) Ac-L-TcaP-L-Pro-OH (TcaP = tricyclic constrained proline)[26a]; (d) (3S,6S,8S,9S)-6-acetylamino-8-methoxy-6-methyl-5-oxooctahydroindolizine-3-carboxylic acid (LBcaP) [26b]; (e) Crystal structure of the alamethicin segment Ac-Aib-Pro-Aib-Ala [61]; (f) Crystal structure of L-BcaP-L Ala-L Ala-OtBu [26a, 62].

other constrained proline derived structures have been reported [28]. Figure 5.7 summarizes the structures of parent Pro-Pro sequences and related structures.

Application of the diproline mimetic organic templates to synthetic protein design is limited by the complexity of the synthetic protocols used in preparing them. A more readily accessible approach would be to examine the use of unconstrained diproline segments in generating helical structures. A recent analysis focuses on the model hexapeptide Piv-Pro-Pro-Aib-Leu-Aib-Phe-OMe [29]. Solu-

tion NMR studies demonstrate a significant population of helical conformations encompassing the entire length of the peptide, including the N-terminus diproline segment. However, populations of the *cis* Pro-Pro conformer and an apparently unfolded structure, with Pro1 adopting the P_{II} conformation are also described by NMR in solution. In single crystals, a helical fold is observed over the segment, residues 2–5, while Pro1 adopts a P_{II} conformation. This study illustrates the conformational heterogeneity that may be anticipated for diproline segments. The observation of α_R, α_R conformations at Pro1 and Pro2 is encouraging, suggesting that an attempt to bias conformational choices by using local sequence effects, may be worthwhile. An analysis of X-Pro-Pro segments in 1741 protein structures reveals about 25 examples of diproline segments occurring in the α_R, α_R conformations at the N terminus of a helix [29]. The predominant conformation, that is favored for a Pro-Pro unit, is the P_{II}-P_{II} structure, with 256 examples being found in the data set. Directing an unconstrained diproline segment into a helical fold will require a detailed understanding of near neighbor effects on conformational choice.

5.3
Hairpin Design using Obligatory Turn Segments

β-hairpins are ubiquitous elements in protein structures. Ideal hairpins are characterized by hydrogen bond formation between registered antiparallel strands. The connecting element between the two strands is usually a short length of polypeptide, which permits chain reversal. Hairpins that contain two central residue loops are abundant in proteins. The early work of Thornton and coworkers established that type I'/II' β-turns occur most frequently in protein β-hairpins [30]. The recognition that 'prime' turns facilitate antiparallel strand registry, permitted 'first principles design' of β-hairpin structures [6, 21a, 31]. The premise behind this approach is that centrally located prime turns can drive short peptide sequences into β-hairpin conformations. Two different approaches have been adopted in the design of β-hairpins. In one strategy, Asn-Gly sequences have been used to stabilize a β-hairpin fold; the choice of a segment being based on the ability of Asn to adopt α_L values, thus promoting type I' β-turns at the Asn-Gly unit [7–9]. In the second strategy, the prime turn formation is promoted by using the more constrained ^{D}Pro-Gly segment [32], in which the two favored conformations at ^{D}Pro correspond to those required for the $i+1$ residue in type I' (α_L) and type II' (P_{II}') conformations (Fig. 5.8).

5.3.1
^{D}Pro-Xxx Turns

In short oligopeptides, positioning of a central ^{D}Pro-Gly segment has been shown to facilitate registry of N- and C-terminus strand segments, promoting an isolated hairpin [32]. Extensive studies of apolar oligopeptides containing

φ	ψ
-60	-30

φ	ψ
60	30

φ	ψ
- 60	120

φ	ψ
60	-120

L-Pro   D-Pro
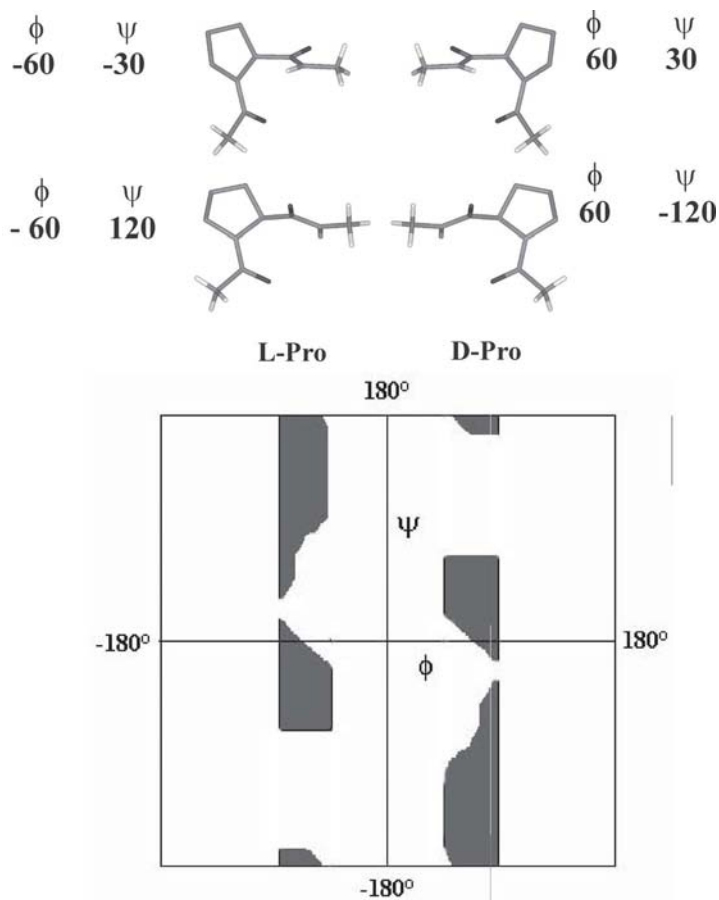
Fig. 5.8 Allowed regions (shaded) of Ramachandran space for L-proline and D-proline. Note that the torsional angle φ is restricted to a relatively narrow range of φ values. L-Pro = −60 ± 20°, D-Pro = +60 ± 20°.

^{D}Pro-Xxx segments have demonstrated hairpin formation in solution by NMR methods and X-ray diffraction in crystal structures [33]. In the vast majority of cases of β-hairpin crystal structures, the ^{D}Pro-Gly sequence adopts a type II′ β-turn conformation. In almost all the cases, three cross-strand hydrogen bonds are observed, with fraying at the N and C termini, resulting in the absence of the fourth hydrogen bond in some structures. Figure 5.9 provides examples of crystallographically determined structures of designed β-hairpins. Tables 5.1 and 5.2 summarize the available experimental evidences for hairpin formation in synthetic sequences soluble in organic and aqueous solvents, respectively. When Xxx = ^{L}Pro, an obligatory type II′ β-turn is formed. Indeed, ^{D}Pro-^{L}Pro sequences have been used to generate stable hairpins in a variety of biologically important
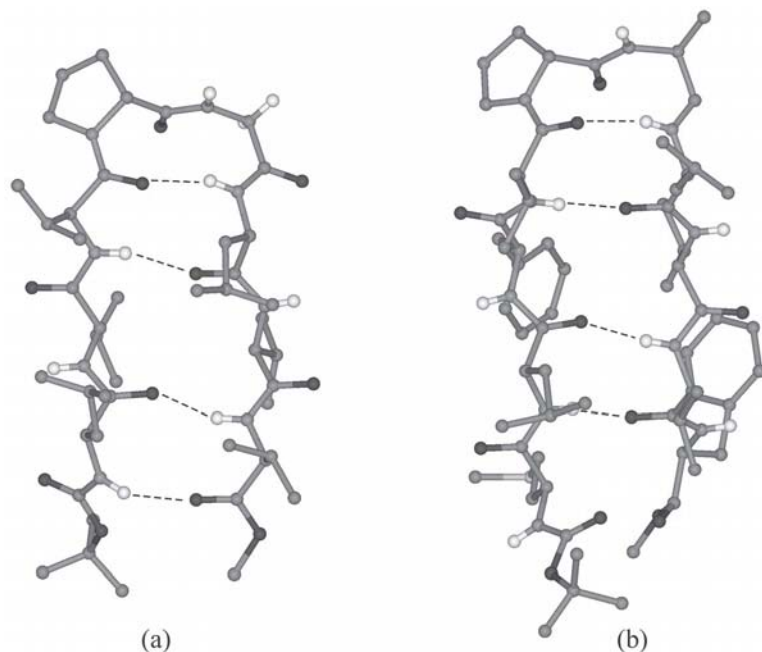
Fig. 5.9 Molecular conformations in crystals observed for the peptides.
(a) Boc-Leu-Val-Val-^{D}Pro-Gly-Leu-Val-Val-OMe [32b]; (b) Boc-Met-Leu-
Phe-Val-^{D}Pro-Ala-Leu-Val-Val-Phe-OMe [33b].

synthetic peptides. The β-hairpin conformation of a model octapeptide Boc-Leu-
Phe-Val-^{D}Pro-^{L}Pro-Leu-Phe-Val-OMe (Fig. 5.10) has been established in solution
by NMR [34] and X-ray diffraction in crystals [unpublished]. It is pertinent to
note that the homochiral diproline segment has been advanced as a potent helix
nucleating template [29], whereas centrally positioned heterochiral diproline seg-
ments strongly favor β-hairpin formation [34, 35]. When the ^{D}Pro-^{L}Pro (hetero-
chiral) sequence is placed at the amino terminus, consecutive β-turn (II$'$-I/III)
formation is observed [36]. In principle, the II$'$-I/III consecutive β-turn structure
at the amino terminus of a peptide sequence can serve to initiate a helical fold.

A variety of residues can be accommodated at the $i + 2$ position in the central
β-turn [33, 37]. In the strands, lengthening of the backbone by incorporation of β
and γ amino acids [32d–g] can occur without disrupting the hairpin structure.
This feature is exemplified in Fig. 5.11.

5.3.2
Aib-^{D}Xxx Turns

The insertion of an Aib-^{D}Ala segment into a host L-amino acids sequence can pro-
mote type I$'$ β-turn formation and consequently lead to generation of a β-hairpin

Table 5.1 Representative examples of water-soluble β-hairpins.

Peptide	Technique	Reference
R-G-I-T-V-N-G-K-T-Y-G-R	NMR	7a
R-G-A-T-A-N-G-A-T-A-G-R	NMR	7a
R-G-A-T-A-N-G-K-T-G-Y-R	NMR	7a
R-T-I-T-V-N-G-A-T-A-G-R	NMR	7a
K-K-Y-T-V-S-I-N-G-K-K-I-T-V-S-I	NMR	8a
R-Y-V-E-V-^{D}P-G-O-K-I-L-Q-NH$_2$	NMR	9a
R-Y-V-E-V-N-G-O-K-I-L-Q-NH$_2$	NMR	9a
R-W-Q-Y-V-^{D}P-G-K-F-T-V-Q-NH$_2$	NMR	9b
R-W-Q-Y-V-N-G-K-F-T-V-Q-NH$_2$	NMR	9b
R-W-Q-Y-V-^{D}P-G-K-S-T-S-Q-NH$_2$	NMR	9b
R-G-W-S-V-Q-M-G-K-Y-T-N-N-G-K-T-T-E-G-R (three-stranded β-sheet)	NMR	51
Ac-V-F-I-T-S-^{D}P-G-K-T-Y-T-E-V-^{D}P-G-O-K-I-L-Q-NH$_2$ (three-stranded β-sheet)	NMR	50
R-G-T-I-K-^{D}P-G-I-T-F-A-^{D}P-A-T-V-L-F-A-V-^{D}P-G-K-T-L-Y-R (four-stranded β-sheet)	NMR	49b
R-G-I-K-V-^{D}P-G-E-T-N-T-^{D}P-S-V-Q-F-**H**-T-I-^{D}P-G-Y-K-T-L-**H**-E-^{D}P-A-R-I-V-L-K (five-stranded β-sheet)	NMR	49c
R-G-E-C(Acm)-K-F-T-V-^{D}P-G-R-T-A-L-N-T-^{D}P-A-V-Q-K-W-H-F-V-L-^{D}P-G-Y-K-C-E-I-L-A (four-stranded β-sheet)	NMR	49d

fold. This feature is illustrated by the structure of an octapeptide Boc-Leu-Phe-Val-Aib-^{D}Ala-Leu-Phe-Val-OMe in crystals (Fig. 5.12a) [38]. Notably, the twist of the antiparallel strands is more pronounced in the case of type I′ β-turn promoted hairpins. Flattened hairpins are obtained with centrally positioned type II′ β-turns. An interesting feature of the apolar β-hairpins is their high solubility in organic solvents like chloroform and methanol and the absence of aggregation in solution, as evidenced by the observation of well-resolved, sharp NMR spectra, even in a poorly solvating medium like chloroform.

The incorporation of a central Aib-^{D}Pro segment strongly favors the formation of type I′/III′ β-turn nucleating hairpins, with ^{D}Pro constraining the torsion angles of the $i + 2$ residue to values that promote type I′/III′ β-turns. (Note: there is only a small variation in torsional angles at the $i + 2$ position in both type I′ and type III′ β-turns. For convenience, it is preferable not to make a distinction between the two closely related turns). Figure 5.12b shows the NMR derived

**Table 5.2** Representative examples of β-hairpins soluble in organic solvents.

Peptide Sequence	Technique	Reference
Boc-L-V-V-^D**P-G**-L-V-V-OMe	X-ray crystallography/NMR	32a–c
Boc-L-V-V-^D**P-A**-L-V-V-OMe	X-ray crystallography	33a
Boc-L-F-V-^D**P-A**-L-F-V-OMe	X-ray crystallography	33b
Boc-L-V-V-^D**P-G**-L-F-V-OMe	X-ray crystallography	33b
Boc-M-L-F-V-^D**P-A**-L-V-V-F-OMe	X-ray crystallography	33b
Boc-L-V-V-^D**P-U**-L-V-V-OMe	X-ray crystallography	33b
Boc-L-F-V-^D**P-Ac₆c**-L-F-V-OMe	X-ray crystallography	33c
Boc-L-F-V-^D**P-Ac₈c**-L-F-V-OMe	X-ray crystallography	33c
Boc-L-F-V-**U**-^D**A**-L-F-V-OMe	X-ray crystallography	38
Boc-L-F-V-^D**P**-^L**P**-L-F-V-OMe	NMR	34
Boc-L-V-V-^DP-^LP-L-V-V-OMe	NMR	Unpublished
Ac-L-Y-V-^D**P-G**-L-Y-V-OMe	NMR	10c
Ac-L-Y-V-^D**P-G**-L-W-V-OMe	NMR	10c
Boc-Y-L-V-^D**P-G**-W-L-V-OMe	NMR	10c
Boc-W-L-V-^D**P-G**-W-L-V-OMe	NMR	10c
Boc-L-L-V-^D**P-G**-Y-L-W-OMe	NMR	10c
Boc-L-L-V-^D**P-G**-Y-W-V-OMe	NMR	10c
Boc-L-Y-V-^D**P-G**-L-L-V-OMe	NMR	10c
Boc-L-V-V-Aib-^D**A**-L-V-V-OMe	NMR	Unpublished
Boc-L-V-V-Aib-^D**V**-L-V-V-OMe	NMR	Unpublished
Boc-L-V-V-Aib-^D**P**-L-V-V-OMe	NMR	Unpublished
Boc L-F-V-^DP-G-L-V-L-A-^DP-G-F-V-L-OMe	NMR	49a
Boc L-F-V-^DP-^LP-L-F-V-A-^DP-^LP-L-F-V-OMe	NMR	Unpublished

structure for the model peptide Boc-Leu-Val-Val-Aib-^DPro-Leu-Val-Val-OMe [unpublished].

5.3.3
Asn-Gly Turns

The high propensity of Asn residue to adopt α_L conformation suggests that Asn-Gly sequences may be used to nucleate type I′ β-turn structures, which in turn can promote β-hairpin formation. Approaches to water-soluble peptide β-hairpins have used the Asn-Gly segment, as the key element in the design strategies [7–9] to promote folding. An analysis by Gellman and coworkers suggests that the ^DPro-Gly segment is a stronger inducer for β-hairpin formation than Asn-Gly segment [9a].
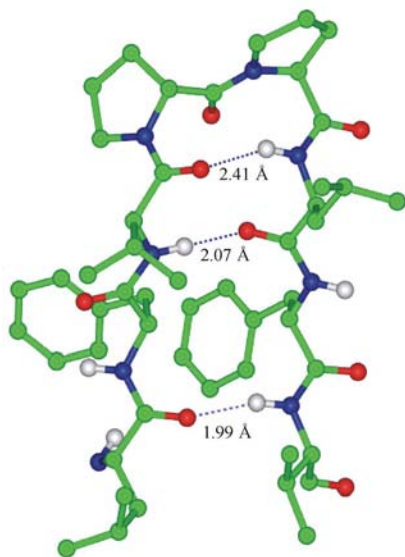
Fig. 5.10 NMR derived structure of peptide Boc-Leu-Phe-Val-^{D}Pro-^{L}Pro-Leu-Phe-Val-OMe in CDCl$_3$ [34].
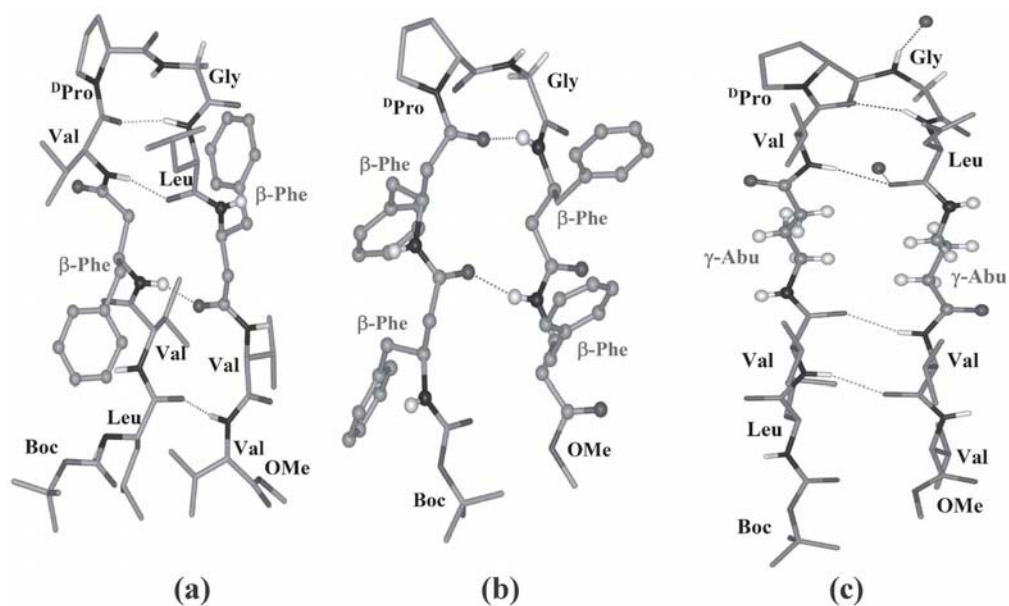
Fig. 5.11 Molecular conformations observed in crystals of peptides containing β- and γ-residues in the strand segments of β-hairpins: (a) Boc-Leu-Val-β^3Phe-Val-^{D}Pro-Gly-Leu-β^3Phe-Val-Val-OMe (^{D}Pro-Gly; type I' β-turn) [32d]; (b) Boc-β^3Phe-β^3Phe-^{D}Pro-Gly-β^3Phe-β^3Phe-OMe (^{D}Pro-Gly; type II' β-turn) [32e]; (c) Boc-Leu-Val-γ-Abu-Val-^{D}Pro-Gly-Leu-γ-Abu-Val-Val-OMe (^{D}Pro-Gly; type I' β-turn) [32f].
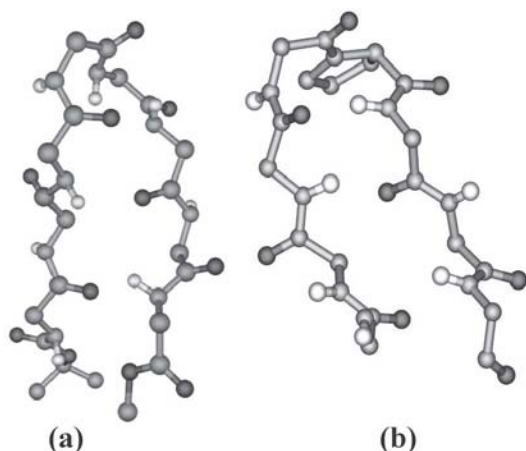
Fig. 5.12 (a) Molecular conformation observed in the crystals for peptide Boc-Leu-Phe-Val-Aib-^{D}Ala-Leu-Phe-Val-OMe [38]; (b) NMR derived structure of the peptide Boc-Leu-Val-Val-Aib-^{D}Pro-Leu-Val-Val-OMe in CDCl$_3$ [unpublished].

5.3.4
Expanded Loop Segments

The construction of β-hairpins with a loop segment greater than two residues provides a significant challenge for design strategies. Considerable information is available on the conformational properties of two-residue turns in proteins. Three-residue loop hairpin structures are much less well characterized. Three-residue turns are relatively common in protein structures [39]. A centrally positioned ^{D}Pro-^{L}Pro-^{D}Ala segment in a model peptide Boc-Leu-Phe-Val-^{D}Pro-^{L}Pro-^{D}Ala-Leu-Phe-Val-OMe has been shown to accommodate the formation of a β-hairpin structure in solution (Figure 5.13), with a central three-residue loop and two strand segments [34]. The observed conformation of the three-residue loop segment is $\alpha_L\alpha_R\alpha_L$, which corresponds to a three-residue α-turn structure, reminiscent of a turn family observed in proteins [40]. The extension of design strategies to four-residue loops and larger connecting elements will undoubtedly be a significant challenge.

5.3.5
Choice of Strand Residues

In designed β-hairpins, a variety of residues can be accommodated in the strand segments. The β-branched residues (Val, Leu, Ile) [33] and bulky nonpolar residues (Phe, Tyr, Trp) [10, 34, 38] have been successfully used in a number of cases. Positioning of aromatic residues in the non-hydrogen-bonding position confers additional stability to the hairpin structures, because the aromatic–aromatic inter-
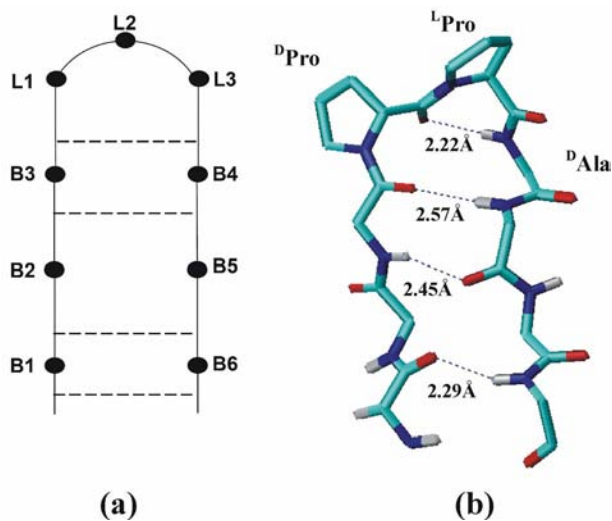
Fig. 5.13 (a) Schematic diagram of β-hairpins with three residue loops; (b) NMR derived structure of peptide Boc-Leu-Phe-Val-^{D}Pro-^{L}Pro-^{D}Ala-Leu-Phe-Val-OMe in CDCl$_3$ [34].

actions are weak [10]. In aqueous solution, Trp-Trp interactions can be used to promote residue clustering and hairpin formation as illustrated by extensive studies on Trp zipper peptides [10a]. Cross-strand interactions also assume additional importance in the case of multi-stranded β-sheet structures, where internal strands interact with the two flanking strands. Nonpolar amino acids like omega amino acids are readily accommodated into the strand sequences of β-hairpins [32d–g]. An interesting residue, which has been used in strand segments, is *m*-aminobenzoic acid. The positioning of amino and carboxyl groups onto a rigid aromatic scaffold has been exploited in incorporating the residue into strand segments of synthetic β-hairpins [41].

5.4
Helix–Helix Motifs

The ability to construct relatively rigid helical peptide modules can be used to advantage in the creation of larger designed structures, in which distinct helical segments are connected by a non-helical linking loop. The orientation of two helical modules will then be determined by the conformational properties of the linking segment. Several attempts have been reported to design antiparallel helix–helix motifs using an Aib-rich helical module. In these cases, Gly, Pro and D-amino acids containing loop segments, generally 2–3 residues in length have been employed. The crystal structures of some representative examples reveal an extended arrangement of the two linked modules. Optimization of side-chain interactions has provided a well characterized example of a helix-loop-helix motif in the pep-
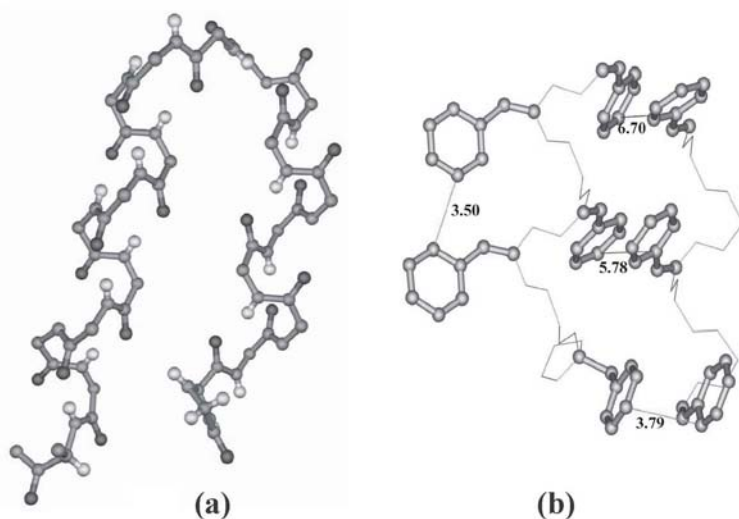
Fig. 5.14 (a) Helix-turn-helix structure of the peptide Ac-Gly-ΔPhe-D-Ala-ΔPhe-ΔPhe-D-Ala-ΔPhe-ΔPhe-L-Ala-(Gly)₄-ΔPhe-L-Ala-L-Leu-ΔPhe-L-Ala-L-Leu-ΔPhe-L-Ala-NHMe [42]; (b) a view of the side chain–side chain interactions showing short distances (Å) [42]. Figures generated using the coordinate set CCDC-153089 [www.ccdc.cam.ac.uk/].

tide   Ac-Gly-ΔPhe-D-Ala-ΔPhe-ΔPhe-D-Ala-ΔPhe-ΔPhe-L-Ala-(Gly)₄-ΔPhe-L-Ala-L-Leu-ΔPhe-L-Ala-L-Leu-ΔPhe-L-Ala-NHMe shown in Fig. 5.14 [42]. In this case, favorable aromatic–aromatic interactions facilitate the formation of a compact structure. In the case of water-soluble peptides, hydrophobic interactions may be used to advantage to drive the *de novo* designed structures into condensed conformations, with solvent forces providing a major impetus for compaction. In such cases, fragment design utilizes the principles of patterning polar and apolar residues in such a manner that there is a clear segregation of residues of one type on distinct faces of the molecule, once secondary structures are formed [5]. The generation of apolar faces permits secondary structure association in apolar tertiary interactions. Table 5.3 lists some representative examples of helix bundles generated by *de novo* design [43–46].

Another approach which has been successfully employed is the use of metal ions as templates with ligating group positioned on distinct secondary structure elements forced to proximity by metal–ligand interactions [47]. Mutter and coworkers have advanced the use of rigid scaffolds to position helical segments in proximity, resulting in systems that have been termed template-assembled synthetic proteins (TASP) [48].

The design of helix–helix motifs in apolar solvents requires substantial control over the conformational properties of the linking segment and an appreciation of the geometrical features of the weakly polar interactions that may be used to control helix–helix orientation. In this approach an understanding of the role of entropically stabilizing side chain–side chain interactions is especially necessary.

Table 5.3 Representative examples of designed water-soluble supersecondary structures.

Molecule	Technique	Reference
α-helix (α-1 at low pH; **1AL1**)[a] (13 residues)	X-ray crystallography	43a
α-helix (α-1 near neutral pH; **1BYZ**)[a] (13 residues)	X-ray crystallography	43b
Four-helix bundle (Peptidergent; single polypeptide chain; **4HB1**) (108 amino acids)	X-ray crystallography	44a
Triple-stranded coiled-coil (Coil-Ser; **1COS**) (31 residues/chain)	X-ray crystallography	43e
Triple-stranded coiled-coil (Coil-Vald; **1COI**) (31 residues/chain)	X-ray crystallography	43c
Trimeric coiled-coil (**1BB1**) (36 residues/chain)	X-ray crystallography	44b
Trimeric coiled-coil (**1GCM**) (34 residues/chain)	X-ray crystallography	44c
Tetrameric coiled-coil (**1GCL**) (34 residues/chain)	X-ray crystallography	44d
Right-handed, tetrameric coiled-coil (**1RH4**) (35 residues/chain)	X-ray crystallography	44e
Four-helix bundle with a diiron-binding center (association of two helix-loop-helix motifs) (Due Ferro 1; **1EC5**) (50 residues/chain)	X-ray crystallography	43d
Helical hairpin (RtR; **1ABZ**) (40 residues)	NMR	44e
Helical hairpin (R-2D; **1QP6**) (35 residues)	NMR	45a
ββα Motif (**1FSD**) (28 residues)	NMR	45b
ββα Motif (**1PSV**) (28 residues)	NMR	45c
Three-helix bundle (single polypeptide chain; **2A3D**) (73 residues)	NMR	46a–b

[a] These are single helices, but the discussion of crystal state aggregation may be relevant to supersecondary structure design.

5.5
Multi-stranded β-Sheets

The design of multi-stranded β-sheets in both polar and apolar structures is conceptually simple, once the principles of β-hairpin design are established. In considering the design of a multi-stranded β-sheet, it is useful to examine the consequences of introducing strong prime turn forming segments at centrally located positions in designed polypeptide chains, as noted in Section 5.3. Incorporation of strong prime turns forming segments like [D]Pro-Xxx [32–33] and Aib-[D]Xxx [38] promotes hairpin formation. The insertion of such nucleating units at multi-
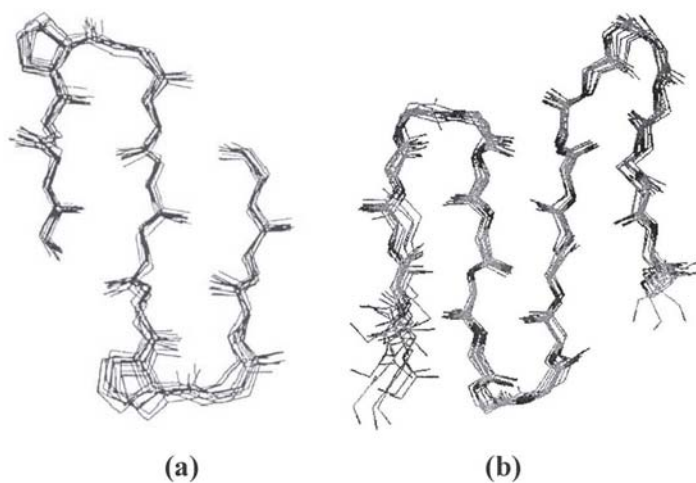
Fig. 5.15 Superposition of 10 NMR derived structures for designed β-sheets: (a) three stranded β-sheet in a 14-residue peptide [49a]; (b) Four stranded β-sheet in a 26-residue peptide [49b], calculated using NOE derived restraints.

ple positions in a polypeptide chain may be an effective strategy for the creation of multi-stranded β-sheets [49–51]. This expectation is clearly borne out by several experimental studies on three-, four- and five-stranded β-sheets. Figure 5.15 illustrates examples of multi-stranded β-sheet structures characterized in organic solvents by NMR. Thus far, no crystallographic analysis of a multi-stranded β-sheet prepared by the *de novo* design approach has been reported. The construction of multi-stranded β-sheet structures stable in organic solvents has proved easier than anticipated because of the driving role of the cross-strand hydrogen bonds in stabilizing β-sheet structures. In β-sheets, the influence of cross-strand side chain–side chain interactions is less dominant, permitting a wide range of strand segments to be chosen. Multi-stranded sheets may be viewed as examples of tertiary structure formation. Mimics of β-sheet motifs in protein structures have been designed, based on naturally occurring toxin models that are stabilized by disulfide bonds [52]. While aggregation is a problem that is frequently encountered in the design of multi-stranded β-sheet structures [53], the positioning of strongly basic residues Arg/Lys at the N and C termini is an effective strategy to limit peptide self association [7a].

5.6
Mixed Helix-Sheet (α/β) Structures

Mixed α,β structures have also been the target of design attempts. The successful construction of the momomeric ββα protein, that contains three elements of protein structures: namely an α-helix, a β-hairpin and a connecting loop, by Imperiali
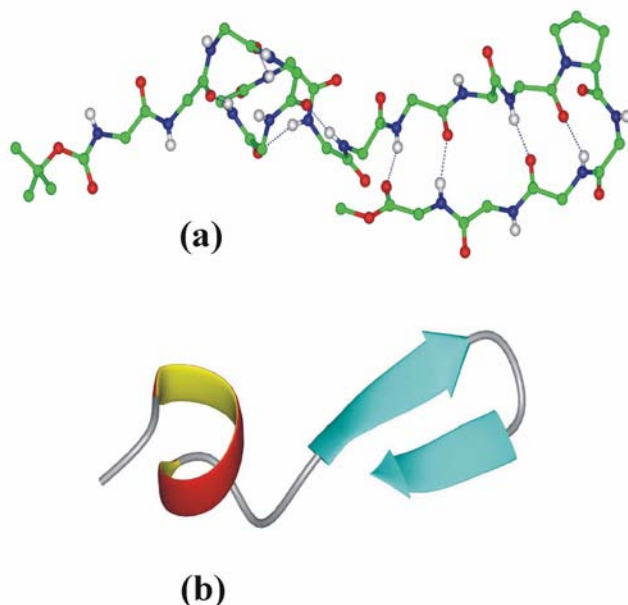
Fig. 5.16 Structure of a designed 17-residue peptide: (a) Conformation in crystals showing intramolecular hydrogen bonds; (b) Ribbon representation of the N-terminal helix linked to a C-terminal β-hairpin [56].

and coworkers is notable [54]. In this design, a single Gly residue is used as a hinge. Subsequent studies revealed that D-amino acids fulfilled the role of a hinge effectively [55]. Mixed α,β structure design has also been explored in completely apolar sequences, by linking conformationally well-defined helical and hairpin modules. A 17-residue sequence Boc-Val-Ala-Val-Aib-Val-Ala-Leu-Gly-Gly-Leu-Phe-Val-^{D}Pro-Gly-Leu-Phe-Val-OMe that has been crystallographically characterized is illustrated in Figure 5.16 [56]. Here, a Gly-Gly linking segment separates the helical and hairpin domains. Gly 8 acts as a helix termination segment permitting the N-terminus helical segment to end in a classical Schellman motif [57]. NMR characterization of an analogous peptide Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-Gly-Gly-Leu-Val-Val-^{D}Pro-Gly-Leu-Val-Val-OMe containing a Gly-Gly linking segment has also been reported (58).

In a further attempt to extend this approach to protein design, a synthetic 19-residue peptide Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-^{D}Ala-^{D}Leu-Leu-Val-Phe-Val-Aib-^{D}Val-Leu-Phe-Val-Val-OMe has been constructed [24]. This sequence is based on the fact that a helical conformation is established for the N-terminus peptide [57c] and a β-hairpin conformation for the C-terminus segment, in isolated peptide fragments [38]. Assembly of these two modules with the linking segment ^{D}Ala-^{D}Leu was anticipated to yield a α/β structure. Surprisingly, the crystal structure of the 19-residue peptide revealed a continuous helix (Fig. 5.6a). A subsequent NMR study in an apolar, non-interacting solvent like CDCl$_3$ revealed

NOEs, which are fully consistent with a continuous helical structure. The incorporation of as many as three D-residues at centrally located positions of a long right-handed helix appears unusual. It must be emphasized that D-residues can be accommodated into right-handed structure (conversely L-amino acids can be incorporated in left-handed helix). The energetic penalty for such accommodation is relatively small and has been estimated to about 1.2 kcal mol^{-1} [59]. This example suggests that the design strategy may be imperfect and that comprehensive unambiguous structural characterization is critical at every stage of *de novo* design projects. X-ray diffraction and NMR spectroscopy are the methods of choice. Peptides in the size range up to 40 residues are often difficult to crystallize. X-ray diffraction and NMR spectroscopy has generally been found to be more facile in the case of apolar peptides, which adopt well-defined conformations and display appreciable solubility in organic solvents. Isolated peptides helices have shown a significantly greater tendency to crystallize than their β-hairpin counterparts. Thus far multi-stranded β-sheet structures have not been characterized in crystals of short and medium size peptides. Quantities of designed sequences, containing unusual amino acids produced by chemical synthesis are often limiting. In the case of NMR spectroscopy, averaging between multiple conformational species is a complicating feature. Observation of nuclear Overhauser effects (nOes), which are incompatible with a single well-defined conformation and the occurrence of selective broadening of specific backbone amide and C$^{\alpha}$H protons resonances is a clear indicator of exchange between multiple conformations.

5.7
Conclusions

Design strategies have evolved over the last two decades to a level where secondary structure modules can be constructed with a fair degree of confidence. Helices and β-hairpins stable in both organic and aqueous solvents are now accessible. In water, hydrophobic effects which promote the clustering of nonpolar side chains can be effectively used in the design of sequences that adopt predictable structures. In nonpolar solvents, local control over backbone conformation is achieved by use of stereochemically constrained residues. In this approach, conformational biasing by the introduction of guest residues into host sequences is used as a device to generate well-defined structures. The creation of tertiary structures by assembling the defined secondary modules is still in an exploratory phase. Limited success has been achieved in the design of helical bundles and multi-stranded β-sheets. The exciting possibilities of designing new folds observed in protein structures has recently been realized [60].

Acknowledgments

References

1 W. F. DeGrado, *Adv Protein Chem.* **1988**, 39, 51–124; (b) W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu Rev Biochem.* **1999**, 68, 779–819; (c) S. F. Betz, J. W. Bryson, W. F. DeGrado, *Curr. Opin. Struct. Biol.* **1995**, 5, 457–463; (d) C. Smith, L. Regan, *Science* **1995**, 270, 980–982.

2 (a) S. Tanaka, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **1975**, 72, 3802–3806; (b) S. S. Zimmerman, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **1977**, 74, 4126–4129; (c) N. Go, H. Taketomi, *Proc. Natl. Acad. Sci. USA* **1978**, 75, 559–563.

3 T. E. Creighton, *Proteins: Structures and molecular properties*; 2 edn; Freeman: New York, 1993.

4 S. Brown, N. J. Fawzi, T. H. Gordon, *Proc. Natl. Acad. Sci. USA,* **2003**, 100, 10712–10717.

5 (a) S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science* **1993**, 262, 1680–1685; (b) K. A. Dill, *Biochemistry*, **1990**, 29, 7133–7154; (c) S. Sun, R. Brem, H. S. Chan, K. A. Dill, *Protein Engineering*, **1995**, 9, 1205–1213; (d) K. A. Dill, H. S. Chan, *Nature Structural Biology*, **1997**, 4, 10–19.

6 J. Venkatraman, S. C. Shankaramma, P. Balaram, *Chem. Rev.* **2001**, 101, 3131–3152.

7 (a) M. Ramirez-Alvarado, F. J. Blanco, L. Serrano, *Nat. Struct. Biol.* **1996**, 7, 604–612; (b) E. de Alba, M. A. Jimenez, M. Rico, *J. Am. Chem. Soc.* **1997**, 119, 175–183.

8 (a) A. J. Maynard, M. S. Searle, *Chem. Commun.* **1997**, 1297–1298; (b) A. J. Maynard, G. J. Sharman, M. S. Searle, *J. Am. Chem. Soc.* **1998**, 120,

1996–2007; (c) S. R. Griffiths-Jones, G. J. Sharman, A. J. Maynard, M. S. Searle, *J. Mol. Biol.* **1998**, 284, 1597–1609.

9 (a) H. E. Stanger, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, 120, 4236–4237; (b) J. F. Espinosa, F. A. Syud, S. H. Gellman, *Protein Sci.* **2002**, 11, 1492–1505.

10 (a) A. G. Cochran, N. J. Skelton, M. A. Starovasnik, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 5578–5583; (b) S. M. Butterfield, M. L. Waters, *J. Am. Chem. Soc.* **2003**, 125, 9580–9581; (c) R. Mahalakshmi, S. Raghothama, P. Balaram, *J. Am. Chem. Soc.* **2006**, 128, 1125–1138.

11 (a) S. H. Gellman, *Acc. Chem. Res.* **1998**, 31, 173–180; (b) D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, 101, 3893–4011; (c) M. A. Schmitt, S. H. Choi, I. A. Guzei, S. H. Gellman, *J. Am. Chem. Soc.* **2006**, 128, 4538–4539.

12 G. N. Ramachandran, V. Sasisekharan, *Adv. Protein. Chem.* **1968**, 23, 283–438.

13 K. Gunasekaran, *Ph.D Thesis, Stereochemical analysis of protein structures – lessons for design, engineering and prediction.* **1997**, Indian Institute of Science, Bangalore, India.

14 J. S. Richardson, D. C. Richardson, *Science* **1988**, 240, 1648–1652.

15 R. Kaul, P. Balaram, *Bioorg. Med. Chem*, **1999**, 7, 105–117.

16 R. Nagaraj, P. Balaram, *Acc. Chem. Res.* **1981**, 14, 356–362.

17 (a) A. Szekeres, B. Leitgeb, L. Kredics, Z. Antal, L. Hatvani, L. Manczinger, Cs. Vágvölgyi, *Acta Microbiologica et*

Immunologica Hungarica **2005**, 52, 137–168; (b) J. K. Chugh, B. A. Wallace, *Biochem. Soc. Trans.* **2001**, 29, 565–570.

18 (a) H. Brückner, G. J. Nicholson, G. Jung, K. Kruse, W. A. Konig, *Chromatographia* **1980**, 13, 209; (b) I. L. Karle, J. L. Flippen-Anderson, M. Sukumar, P. Balaram, *Proc. Natl. Acad. Sci. USA* **1987**, 84, 5087–5091; (c) I. L. Karle, J. L. Flippen-Anderson, S. Agarwalla, P. Balaram, *Proc. Natl. Acad. Sci. USA* **1991**, 88, 5307–5311; (d) I. L. Karle, M. A. Perozzo, V. K. Mishra, P. Balaram, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 5501–5504.

19 (a) N. Shamala, R. Nagaraj, P. Balaram, *Biochem. Biophys. Res. Commun.* **1977**, 79, 996–997; (b) B. V. V. Prasad, P. Balaram, *CRC Crit. Rev. Biochem* **1984**, 16, 307–347.

20 (a) I. L. Karle, P. Balaram, *Biochemistry* **1990**, 29, 6747–6756; (b) P. Balaram, *Current Opinion in Structural Biology* **1992**, 2, 845–851.

21 (a) S. Aravinda, N. Shamala, R. S. Roy, P. Balaram, *Proc. Indian Acad. Sci. (Chem. Sci.)* **2003**, 115, 373–400; (b) S. Datta, *Ph. D. Thesis* **1998**, *Folding of the designed peptides: X-ray crystallographic studies on the structure, conformation, aggregation and interactions of oligopeptides containing conformationally constrained amino acids*, Indian Institute of Science, Bangalore, India.

22 (a) C. Toniolo, E. Benedetti, *Trends Biochem. Sci.* **1991**, 16, 350–353; (b) C. Toniolo, M. Crisma, G. M. Bonora, E. Benedetti, B. DiBlasio, V. Pavone, C. Pedone, A. Santini, *Biopolymers* **1991**, 31, 129–138.

23 I. L. Karle, L. L. Flippen-Anderson, M. Sukumar, K. Uma, *Proc. Natl. Acad. Sci. USA* **1988**, 85, 299–303.

24 I. L. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci. USA* **2003**, 100, 13946–13951.

25 (a) D. S. Kemp, J. G. Boyd, C. C. Muendel, Nature **1991**, 352, 451–454; (b) D. S. Kemp, T. P. Curran, J. G. Boyd, T. J. Allen, *J. Org. Chem.* **1991**, 56, 6683–6697; (c) D. S. Kemp, T. P. Curran, W. M. Davis, J. G. Boyd, C.

Muendel, *J. Org. Chem.* **1991**, 56, 6672–6682; (d) G. E. Job, B. Heitmann, R. J. Kennedy, S. M. Walker, D. S. Kemp, *Angew. Chem. Int. Ed.* **2004**, 43, 5649–5651; (e) B. Heitmann, G. E. Job, R. J. Kennedy, S. M. Walker, D. S. Kemp, *J. Am. Chem. Soc.* **2005**, 127, 1690–1704.

26 (a) S. Hanessian, G. Papeo, K. Fettis, E. Therrien, M. T. Viet, *J. Org. Chem.* **2004**, 69, 4891–4899; (b) S. Hanessian, G. Papeo, M. Angiolini, K. Fettis, M. Beretta, A. Munro, *J. Org. Chem.* **2003**, 68, 7204–7218; (c) S. Hanessian, H. Sailes, A. Munro, E. Therrien, *J. Org. Chem.* **2003**, 68, 7219–7233.

27 Y. V. Venkatachalpathi, P. Balaram, *Nature* **1979**, 281, 83–84.

28 R. E. Austin, R. A. Maplestone, A. M. Sefler, K. Liu, W. N. Hruzewicz, C. W. Liu, H. S. Cho, D. E. Wemmer, P. A. Bartlett, *J. Am. Chem. Soc.* **1997**, 119, 6461–6472.

29 R. Rai, S. Aravinda, K. Kanagarajadurai, S. Raghothama, N. Shamala, P. Balaram, *J. Am. Chem. Soc.* **2006**, 128, 7916–7928.

30 B. L. Sibanda, J. M. Thornton, *Nature* **1985**, 316, 170–174; (b) C. M. Wilmot, J. M. Thornton, *J. Mol. Biol.* **1988**, 203, 211–232; (c) B. L. Sibanda, T. L. Blundell, J. M. Thornton, *J. Mol. Biol.* **1989**, 206, 759–777.

31 P. Balaram, *J. Pept. Res.* **1999**, 54, 195–199.

32 (a) S. K. Awasthi, S. Raghothama, P. Balaram, *Biochem. Biophys. Res. Commun.*, **1995**, 216, 375–381; (b) I. L. Karle, S. K. Awasthi, P. Balaram, *Proc. Natl. Acad. Sci (USA)* **1996**, 93, 8189–8193; (c) S. K. Awasthi, S. Raghothama, P. Balaram, *J. Chem. Soc. Perkin Trans 2*, **1998**, 137–142; (d) I. L. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci (USA)* **2001**, 98, 3716–3719; (e) I. L. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci (USA)* **2002**, 99, 5160–5164; (f) R. S. Roy, H. N. Gopi, S. Raghothama, R. D. Gilardi, I. L. Karle, P. Balaram, *Biopolymers* **2005**, 80, 787–799; (g) R. S. Roy, H. N. Gopi, S. Raghothama, I. L. Karle, P.

Balaram, *Chemistry-A Eur. J.* **2006**, 12, 3295–3302.

33 (a) C. Das, G. A. Naganagowda, I. L. Karle, P. Balaram, *Biopolymers* **2001**, 58, 335–346. (b) S. Aravinda, V. V. Harini, N. Shamala, C. Das, P. Balaram, *Biochemistry* **2004**, *43*, 1832–1846; (c) V. V. Harini, S. Aravinda, R. Rai, N. Shamala, P. Balaram, *Chemistry* **2005**, 11, 3609–3620.

34 R. Rai, S. Raghothama, P. Balaram, *J. Am. Chem. Soc.* **2006**, 128, 2675–2681.

35 J. A. Robinson, *Synlett.* **2000**, 4, 429–441.

36 C. M. K. Nair, M. Vijayan, Y. V. Venkatachalapathi, P. Balaram, *J. Chem. Soc., Chem. Comm.* **1979**, 1183–1184.

37 (a) H. N. Gopi, R. S. Roy, S. Raghothama, I. L. Karle, P. Balaram, *Helv. Chim. Acta* **2003**, 85, 3313–3330; (b) S. C. Shankaramma, S. Kumar Singh, A. Sathyamurthy, P. Balaram, *J. Am. Chem. Soc.* **1999**, 121, 5360–5363.

38 S. Aravinda, N. Shamala, R. Rai, H. N. Gopi, P. Balaram, *Angew. Chem. Intl. Ed.* **2002**, 41, 3863–3865.

39 K. Gunasekaran, C. Ramakrishnan, P. Balaram, *Protein Eng.* **1997**, 10, 1131–1141.

40 (a) C. Ramakrishnan, D. V. Nataraj, *J. Pept. Sci.* **1998**, 4, 239–252; (b) V. Pavone, G. Gaeta, A. Lombardi, F. Nastri, O. Maglio, C. Isernia, M. Saviano, *Biopolymers* **1996**, 38, 705–721.

41 G. Srinivasulu, M. H. V. Ramana Rao, S. Kiran Kumar, A. C. Kunwar, *ARKIVOC* **2004**, 69–86.

42 (a) U. A. Ramagopal, S. Ramakumar, D. Sahal, V. S. Chauhan, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 870–874; (b) Rudresh, S. Ramakumar, U. A. Ramagopal, Y. Inai, S. Goel, D. Sahal, V. S. Chauhan, *Structure* **2004**, 12, 389–396.

43 (a) C. P. Hill, D. H. Anderson, L. Wesson, DeGrado, D. Eisenberg, *Science* **1990**, 249, 543–546; (b) G. G. Prive, D. H. Anderson, L. Wesson, D. Cascio, D. Eisenberg, *Protein Sci.* **1999**, 8, 1400–1409; (c) N. L. Ogihara,

M. S. Weiss, W. F. DeGrado, D. Eisenberg, *Protein Sci.* **1997**, 6, 80–88; (d) A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, 97, 6298–6305 (e) B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado, D. Eisenberg, *Science* **1993**, 259, 1288–1293.

44 (a) C. E. Schafmeister, L. J. Miercke, R. M. Stroud, Science **1993**, 262, 734–738; (b) S. Nautiyal, T. Alber, *Protein Sci.* **1999**, 8, 84–90; (c) P. B. Harbury, P. S. Kim, T. Alber, *Nature* **1994**, 371, 80–83; (d) P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science* **1993**, 262, 1401–1407; (e) P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **1998**, 282, 1462–1467.

45 (a) R. B. Hill, W. F. DeGrado, *J. Am. Chem. Soc.* **1998**, 120, 1138–1145; (b) B. I. Dahiyat, S. L. Mayo, *Science* **1997**, 278, 82–87; (c) B. I. Dahiyat, C. A. Sarisky, S. L. Mayo, *J. Mol. Biol.* **1997**, 273, 789–796.

46 (a) S. T. Walsh, H. Cheng, J. W. Bryson, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 5486–5491; (b) N. L. Ogihara, G. Ghirlanda, J. W. Bryson, M. Gingery, W. F. DeGrado, D. Eisenberg, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, 98, 1404–1409.

47 (a) M. R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.* **1992**, 114, 825–831; (b) M. R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.* **1992**, 114, 4000–4002; (c) K. Suzuki, H. Hiroaki, D. Kohda, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **1998**, 120, 13008–13015.

48 (a) G. Tuchscherer, V. Steiner, K. H. Altmann, M. Mutter, *Methods Mol Biol.* **1994**, 36, 261–285; (b) M. Mutter, S. Vuilleumier, *Angew. Chem.-Int. Edit. Engl.*, **1989**, 28, 535–554; (c) M. Mutter, P. Dumy, P. Garrouste, C. Lehmann, M. Mathieu, C. Peggion, S. Peluso, A. Razaname, G. Tuchscherer, *Angew. Chem. Int. Edit. Engl.*, **1996**, 35, 1482–1485.

49 (a) C. Das, S. Raghothama, P. Balaram, *J. Am. Chem. Soc.* **1998**,

120, 5812–5813; (b) C. Das, S. Raghothama, P. Balaram, *Chem. Commun.* **1999**, 967–968; (c) J. Venkatraman, G. A. Naganagowda, R. Sudha, P. Balaram, *Chem. Commun.* **2001**, 2660–2661; (d) J. Venkatraman, G. A. Naganagowda, R. Sudha, P. Balaram, *J. Am, Chem. Soc.* **2002**, 124, 4987–4994.

50 H. L. Schenck, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, 120, 4869–4870.

51 T. Kortemme, M. Ramirej-Alvarado, L. Serrano, *Science* **1998**, 281, 253–256.

52 J. T. Ottesen, B. Imperiali, *Nature* **2001**, 8, 535–539.

53 J. S. Richardson, D. C. Richardson, N. B. Tweedy, K. M. Gernert, T. P. Quinn, M. H. Hecht, B. W. Erickson, Y. Yan, R. D. McClain, M. E. Donlan, *Biophys J.* **1992**, 63, 1185–1209.

54 (a) M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, 271, 342–345; (b) M. Strutheres, J. J. Ottesen, B. Imperiali, *Folding Des.* **1998**, 3, 95–103.

55 K. A. McDonnel, B. Imperiali, *J. Am. Chem. Soc* **2002**, 124, 428–433.

56 I. L. Karle, C. Das, P. Balaram, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 3034–3037.

57 K. Gunasekaran, H. A. Nagarajaram, C. Ramakrishnan, P. Balaram, *J. Mol. Biol.* **1998**, 215, 917–932; (b) S. Datta, N. Shamala, A. Banerjee, A. Pramanik, S. Bhattacharjya, P. Balaram, *J. Am. Chem. Soc.* **1997**, 119, 9246–9251; (c) A. Banerjee, S. Datta, A. Pramanik, N. Shamala, P. Balaram, *J. Am. Chem. Soc.* **1996**, 118, 9477–9483.

58 C. Das, S. C. Shankaramma, P. Balaram, *Chem. Eur. J.* **2001**, 7, 840–847.

59 R. Fairman, S. J. Antony-Cahill, W. F. DeGrado, *J. Am. Chem. Soc.* **1992**, 114, 5458–5459.

60 B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, *Science* **2003**, 302, 1364–1368.

61 R. O. Fox, F. M. Richards, *Nature (London)* **1982**, 300, 325–330.

62 F. H. Allen, *Acta Crystallogr.* **2002**, B58, 380–388.

6
Simulation of Folding Equilibria

Wilfred F. van Gunsteren and Zrinka Gattin

6.1
Introduction

The prediction of the folding of a protein into its native three-dimensional fold as a function of the external conditions is one of the major computational challenges in structural biology [1]. The folding process of polypeptides in solution is driven by weak, nonbonded interatomic interactions. Such interactions govern the thermodynamic properties of the condensed phase in which the (un)folding occurs. Simulation of folding is therefore most promisingly modeled at the atomic level. Since the temperature (T) range of interest basically lies between room and physiological temperatures and energies involved in (un)folding processes are on the order of 1–10 $k_B T$ (tens of kJ mol^{-1}, k_B is Boltzmann's constant), the folding process is largely determined by the laws of classical statistical mechanics. Although quantum mechanics governs the interaction between the electrons of the atoms and molecules, the nonbonded interactions can be very well described by a classical potential energy function or force field as part of a classical Hamiltonian of the system of interest. The statistical-mechanical nature of the folding equilibrium of a polypeptide complicates its modeling because the entropic contributions to the free energy of (un)folding are sizeable. The state of a polypeptide in solution is generally characterized not by one configuration or structure, but by a Boltzmann ensemble of configurations or structures. Although it is easier to think of and handle single structures than to consider configurational ensembles, a number of (experimental) observations including those concerning folding equilbria can only be understood by an analysis in terms of alternative structures or conformations present in an ensemble and in terms of entropy.

Although the protein folding problem has been extensively studied, both theoretically and experimentally, over many years using proteins as objects [2, 3], the key to unraveling the basic principles of the folding process may lie in the study of other polypeptides or peptoids that also adopt a variety of particular folds, and not only carry different side chains but also vary in composition of the (polypeptide) backbone. A great variety of such foldamers exists [4, 5] (see Chapters 1–5).

In this chapter we consider the theoretical modeling of the folding equilibria of foldamers. To be more precise, we may distinguish three levels of description of fold characteristics:

1. prediction of the most dominant structure or fold,
2. prediction of the folding equilibrium,
3. prediction of the folding pathways and kinetics,

for a variety of foldamers in solution as a function of the thermodynamic conditions (temperature, pressure, pH, ionic strength), and of the composition of the solvent. The first level of description does not consider the folding process, only structure prediction. It does not yield (free) energy differences between conformers, ignores the ensemble character of the condensed phase and folding kinetics. Therefore, methods for structure prediction of foldamers will not be considered here. The second level involves a description of the folding equilibrium in terms of a conformational ensemble and thermodynamics, ignoring kinetics and dynamics, which are considered in addition at the third level of description.

Validation of a simulation of a folding equilibrium is usually done by comparison of simulated and experimentally measured properties of the system. However, this sounds more straightforward than it actually is. First, almost every experiment involves an averaging over time and space or molecules, and therefore, does not yield direct information on all configurations constituting a simulation trajectory. Second, experimental data for foldamers are scarce when compared with the number of degrees of freedom involved, so the problem of deriving the conformational ensemble of the folding equilibrium from experimental data has not been studied extensively. Different ensembles may reproduce the same set of experimental data. Third, the experimental data may be of insufficient accuracy to be used to (in)validate simulation predictions. Data characterizing folding equilibria of foldamers mainly originate from NMR experiments and involve nuclear Overhauser enhancement (NOE) intensities and ^{3}J-coupling constants.

On the theoretical side the situation is not less problematic. Folding equilibria are characterized by small (free) energy differences, on the order of 1–10 $k_\mathrm{B}T$, and a low frequency of (un)folding events compared with the time scale of tens to hundreds of nanoseconds reachable for not too large polypeptides in solution using current computers. This implies that converged equilibrium properties and kinetic data can be obtained only for the smallest, rapidly folding foldamers. A popular way to lengthen the time scale of a polypeptide simulation is to reduce the number of explicitly treated degrees of freedom by omitting the solvent ones and representing them by a mean field, a kind of continuum approximation [6–8]. However, such a mean-field solvent represents only one solvent, generally water, at a particular thermodynamic state point (temperature, pressure, pH, etc). Since one of the goals of foldamer research is to characterize folding equilibria and kinetics as function of variation of the environment (solvent, co-solvents) and thermodynamic state points, we shall leave theoretical work based on mean-field solvation models out of consideration.

In the next sections we consider various aspects of dynamic simulation of folding equilibria and their kinetics. They will be illustrated with examples from our own work. It is not the purpose of the present chapter to review the contributions of various research groups to the field of theoretical modeling of foldamers, but to offer the reader an impression of the current possibilities of simulating dynamical folding equilibria of foldamers.

6.2
## Dynamical Simulation of Folding Equilibria under Different Thermodynamic and Kinetic Conditions

Until nine years ago, computer simulation could only be used to investigate the stability of the folds of proteins or peptides by submitting them in their folded form to strongly denaturing forces, e.g. at non-physiologically high temperatures [9, 10]. Folding into the native structure starting from an arbitrary structure under physiological conditions had not been observed at that time. In 1998 Daura et al. [11, 12] demonstrated the reversible folding of a peptide in solution and showed that the unfolded state was characterized by a limited number of peptide conformations. During the following years other studies of reversible peptide folding appeared [13–21]. It is now possible to investigate the folding equilibrium as function of temperature [12, 22], of pressure [23], of pH [24], of ionic strength [25, 26], and of solvent viscosity [27].

Figures 6.1 to 6.4 illustrate the effects of variation of the mentioned factors upon the folding equilibrium and kinetics for two 7-β-peptides and a 20-β-peptide in solution. The backbone atom-positional root-mean-square deviation (RMSD) from the 3_{14}-helical folded structure is shown as function of time. The 3_{14}-helical model structure (see Chapter 2) had been derived as most populated structure in methanol solution from NMR experiments [28, 29]. The upper panel of Fig. 6.1 shows that the helical fold of the 7-β-peptide is very stable at 298 K, only two major unfolding events are observed within 80 ns and the folded conformation is present for about 97% of the time. At 340 K (second panel) the 7-β-peptide is about 50% folded, in agreement with experimental data. The effect on the folding kinetics at 340 K by a change of the solvent viscosity by $\frac{1}{3}$ or $\frac{1}{10}$ is seen in the third and fourth panels. The folding equilibrium remains the same, but the folding kinetics is much faster [27]. Changing the pressure at 340 K from 1 atm to 1000 atm does shift the folding equilibrium towards the unfolded state, as is illustrated in Fig. 6.2 [23]. Figure 6.3 shows that the population of the helical fold decreases as the terminal groups change from $(NH_3^+, COOH)$ in the upper panel, to (NH_2, COO^-) in the middle panel, to $(NH_2, COOH)$ in the lower panel. The 3_{14}-helical fold is more stable in the absence of protecting groups and is enhanced at acidic conditions [24]. Figure 6.4 shows that the presence of Cl^- counterions stabilizes the 3_{14}-helical fold of the 20-β-peptide carrying all 20 proteinogenic side chains [26] by supporting side-chain salt-bridge formation.
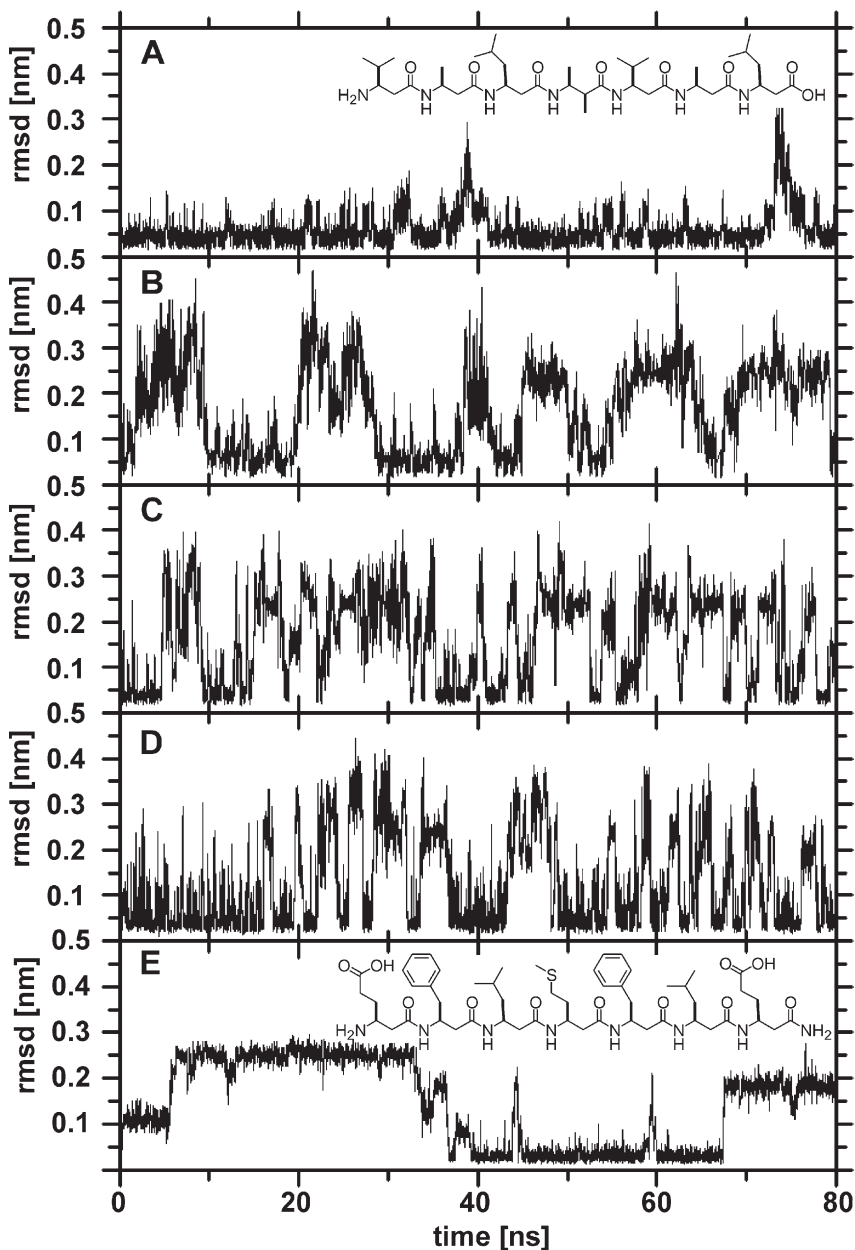
Fig. 6.1 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures with respect to the helical model structures derived from NMR data for β-heptapeptides of identical chain lengths in methanol at 1 atm (the structures are given in panel A and panel E) [57]. The peptide with apolar side chains is simulated at 298 K (A) and at 340 K (B–D). The viscosity of the methanol solvent is reduced by a factor 3 (C) and by factor 10 (D) through mass scaling. The peptide with a few polar side chains is simulated at 340 K in normal methanol (E).
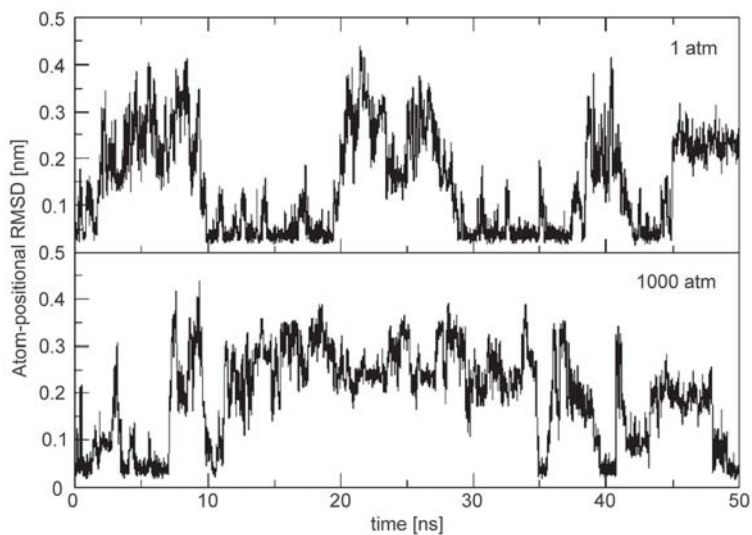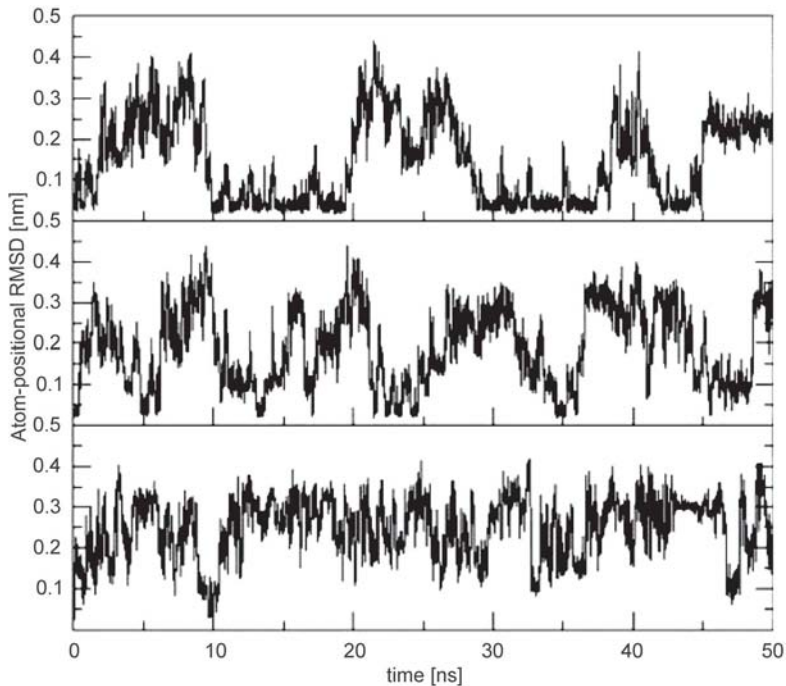
Fig. 6.2 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures (340 K) with respect to the 3_{14}-helical model structure derived from NMR data for a β-heptapeptide (see panel A of Fig. 6.1) in methanol at 1 atm (upper panel) and at 1000 atm (lower panel) [23].

Fig. 6.3 (legend see p. 178)
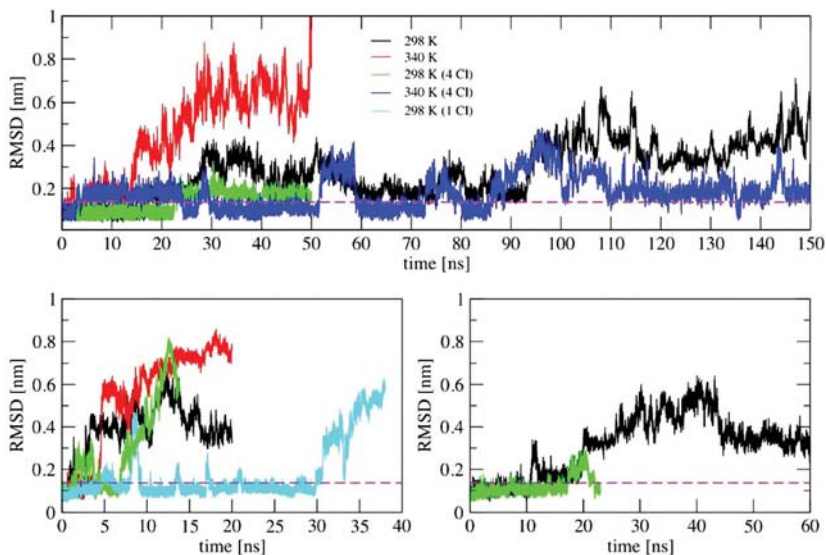
Fig. 6.4 Time series of the backbone atom-positional root-mean-square distance (RMSD) (residues 2–19) of MD trajectory structures with respect to the ideal 3_{14}-helical structure for a 20-β^3-peptide (sequence: Cys, Ala, Ser, His, Asn, Glu, Gly, Trp, Arg, Val, Asp, Gln, Ile, Lys, Thr, Leu, Tyr, Met, Phe, Pro). The top panel shows the results for the simulations in the methanol with the 53A6 force field. The lower left panel shows the results for the simulations in water (53A6 force field) and the lower right panel the results for the simulations in methanol with the 45A3 force field. Colors represent different temperatures and ionic strengths. The magenta horizontal dashed line indicates the minimum RMSD value for which all NMR model structures would belong to the same conformational cluster (0.12 nm) [26].

6.3
Variation of the Composition of the Polypeptide Analogs and the Solvent

Other peptides than α- or β-peptides have been shown, both experimentally and computationally, to fold into stable folds. The folding equilibria for furanose-based carbopeptoids of different chain lengths have been simulated in agreement with NMR experimental data [30, 31]. A simulation study of an α-peptidic equivalent of the 7-β-peptide discussed before predicted that it would not adopt a helical

---

Fig. 6.3 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures with respect to the 3_{14}-helical model structure derived from NMR data for a β-heptapeptide (see Panel A, Fig. 6.1) in methanol for simulations with different charge states of the terminal residues. Upper panel: the N terminus is charged and the C terminus is uncharged (NH_3^+, COOH); middle panel: the N terminus is uncharged and the C terminus is charged (NH_2, COO^-); lower panel: both termini are uncharged (NH_2, COOH) [24].
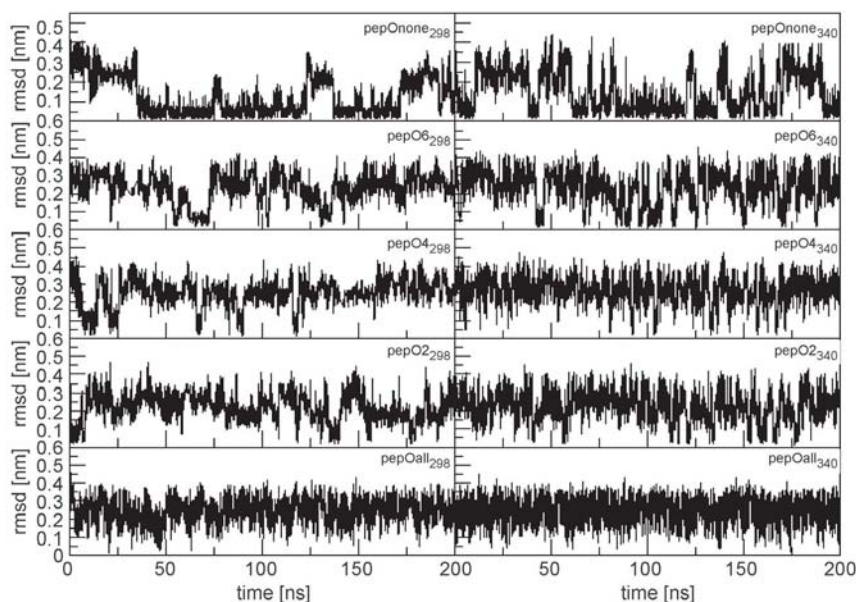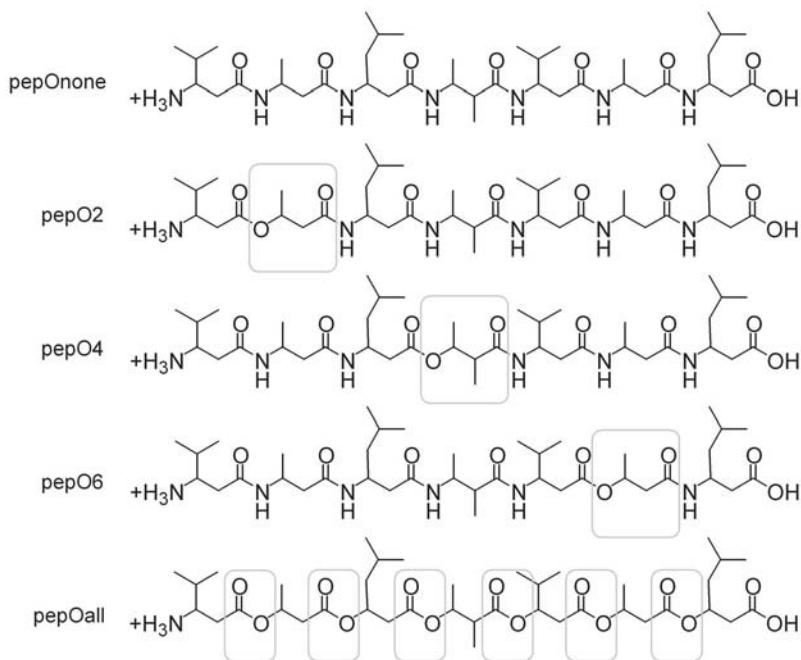
Fig. 6.5 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures with respect to the 3_{14}-helical conformation as a function of time for four β-depsiheptapeptides (see Fig. 6.6) at 298 K and 340 K. For peptide pepOnone the NMR model structure was used as the 3_{14}-helical reference structure, whereas for all other peptides a canonical 3_{14}-helical conformation was taken as reference. The pepOnone notation stands for the 7-β-peptide the structure of which is shown in panel A of Fig. 6.1.

fold, neither in water nor in methanol, in agreement with NMR data [32]. A very short 3-β-aminoxypeptide was shown to adopt a very stable 1.8_8-helix in chloroform and to exhibit no particular fold in water, again in agreement with NMR data [33]. Figure 6.5 shows the effect of replacing an N–H group by an O atom in residue 2 (pep O_2), or residue 4 (pep O_4), or residue 6 (pep O_6), or all residues (pep O_{all}) in 7-β-peptides (Fig. 6.6) that are very similar to the 7-β-peptide (pepO$_{none}$) discussed before. The population of the 3_{14}-helical fold decreases from pepO$_{none}$ to pepO$_6$ to pepO$_2$ to pepO$_4$ to pepO$_{all}$, in agreement with NMR data [34]. Whether the N or C termini of a β-peptide are carrying protective groups or not, also influences the stability of the helical fold in agreement with experiment [35].

Which particular fold (see Chapter 2) a β-peptide will adopt depends on the type of side chain and whether the side chain is located at the α-position (β^2-peptide) or at the β-position (β^3-peptide) in the backbone. The folds observed in NMR experiments were all reproduced in MD simulations based on the GROMOS force field: a left–handed (M)-3_{14} helix [12], a right-handed (P)-10/12 helix [13], a right-handed (P)-2.5_{12} helix [36], or a β-hairpin [18]. Substitution of two methyls at the β-positions in conjunction with standard side chains at the

Fig. 6.6 Sequences of the four β^3-depsi-peptides and the β^3-peptide considered in Fig. 6.5. β-peptide pepOnone: H-β-HVal-β-HAla-β-HLeu-β-HAla(α-Met)-β-HVal-β-HAla-β-HLeu-OH. β-depsipeptide pepO2: H-β-HVal-β-dHAla-β-HLeu-β-HVal-β-HAla-β-HLeu-β-HVal-OH; β-depsipeptide pepO4: H-β-HAla-β-HLeu-β-HVal-β-dHAla-β-HLeu-β-HVal-β-HAla-OH; β-depsipeptide pepO6: H-β-HAla-β-HLeu-β-HVal-β-HVal-β-HLeu-β-dHAla-β-HVal-OH; β-depsipeptide pepOall: H-β-HVal-β-dHAla-β-dHLeu-β-dHVal-β-dHAla-β-dHLeu-β-dHVal-OH; The N-terminal amino- and C-terminal carboxylate groups are both protonated in the simulations as suggested by the experimental data. The depsi-amino acids are denoted with dHAla, dHVal, dHLeu.

α-position prevents helix formation [19], whereas substitution of hydroxyl groups at the α-positions in conjunction with standard side chains at the β-positions leads to the formation of a (P)2.5$_{12}$ helix [36]. The influence of different stereo-centers (SR versus SS) in the backbone of a Val-Phe β-peptide on its conformational preferences was found to be significant, both in simulation and in NMR experiments [37]. Also for carbopeptoids the presence of *cis* versus *trans* linkage across the tetrahydrofuran ring influences the emergence of a particular fold [30, 31]. Whether the presence of side chains with a branching point adjacent to the β-carbon in the backbone (e.g. Ile or Val) [38] or the presence of polar or charged side chains, which would be able to form salt bridges [25, 26], would enhance helix formation in β-peptides was also investigated.

Use of an explicit representation of solvent in the simulations offers the possibility of investigating solvent effects upon fold formation. β-Peptides of different chain lengths that adopt helical folds in methanol, show less to no tendency to do

so when solvated in water [25, 26, 39]. Solvation in chloroform tends to enhance helix formation [33, 40]. For α-peptides, β-hairpin formation in water has been observed [16, 41]. In less polar solvents, such as DMSO, partial helix formation could be observed for a particular 8-α-peptide [42]. Carbopeptoids also showed different folding behavior in DMSO versus chloroform [30, 31]. The observed effects can be rationalized in terms of degree of solvent polarity or dielectric permittivity and competitiveness to form solute–solvent hydrogen bonds. For an example of the complex effects of the addition of co-solvent upon hydrophobic association we refer to [43].

6.4
Convergence of the Simulated Folding Equilibrium

The convergence of a folding equilibrium can be monitored by calculating the number of conformational clusters in a MD trajectory as function of time. A conformational cluster is defined as the set of solute trajectory structures that deviate less than a given limit from each other. Figure 6.7 shows for example trajectory structures of the 7-β-peptide for which the backbone atom-positional root-mean-square deviation (RMSD) for residues 2–6 from the central member structure of the cluster is less than 0.09 nm. The clustering RMSD criteria chosen, 0.09 nm in
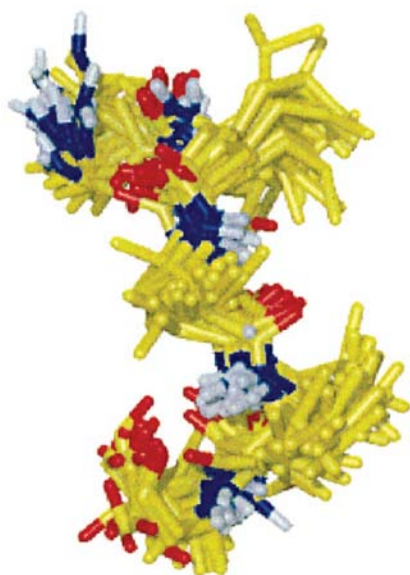
Fig. 6.7 Superposition of the trajectory structures of a β-heptapeptide (see Panel A, Fig. 6.1) at 360 K with RMSD (residues 2–6) from the central structure of 0.09 nm, and a maximum RMSD between any two structures of 0.16 nm [44].
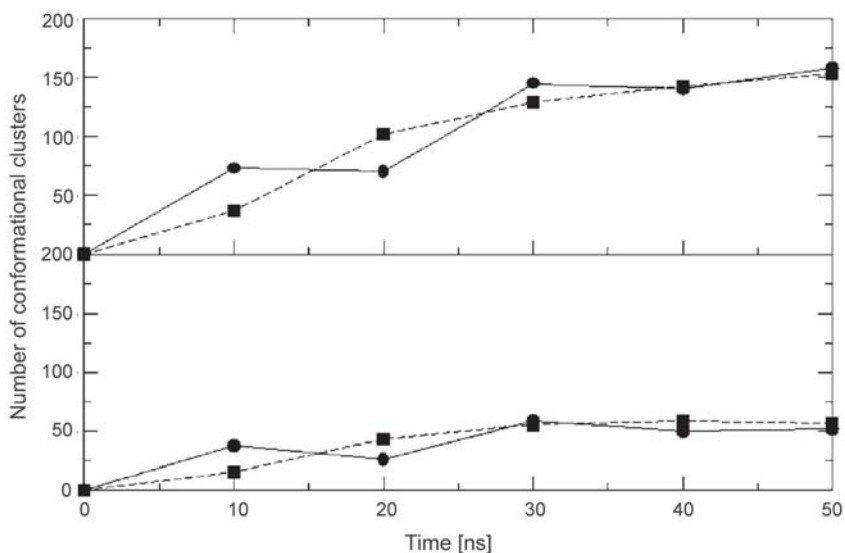
Fig. 6.8 Number of clusters (conformers) of a β-heptapeptide (see Panel A, Fig. 6.1) at 340 K and at a pressure of 1 atm (■) and 1000 atm (●) as a function of time. In the upper panel each point represents the total number of clusters (conformers) at the corresponding time point and in the lower panel the number of clusters (conformers) that make up 95% of the trajectory sampled at the corresponding time point [23].

this case, will determine how precisely a particular cluster is defined. Because the clustering algorithm [44] tends to produce many very sparsely populated clusters after having found the most populated ones, the convergence of the (un)folding equilibrium is better characterized by monitoring not the total number of clusters, but the number of conformational clusters that make up e.g. 95% of the trajectory sampled at the corresponding time point, see Fig. 6.8. This figure shows that the conformational space of the 7-β-peptide is basically completely sampled within about 30 ns. However, to obtain sufficient statistics on (un)folding events much longer simulation times are required as is suggested by Fig. 6.2.

Figure 6.9 shows cases in which the number of sampled conformational clusters does not level off with time, but displays a linear growth with time: the poly-hydroxybutanoate solute continuously accesses new conformations, because there are no hydrogen-bond donor moieties in this chain molecule [45]. Intrasolute hydrogen bonding does not restrict the conformational space accessible to this molecule.

Figure 6.10 demonstrates that for longer chain molecules even 100 ns of sampling at 298 K is not sufficient to find the most dominant P-2.5$_{12}$ helical conformer [36]. Only by simulating at higher temperature, 340 K, it was found, and subsequent simulation starting from this helical structure confirmed its dominance and stability also at 298 K. The lowest panel of Fig. 6.1 illustrates that the presence of polar side chains may slow down the (un)folding process. The conver-
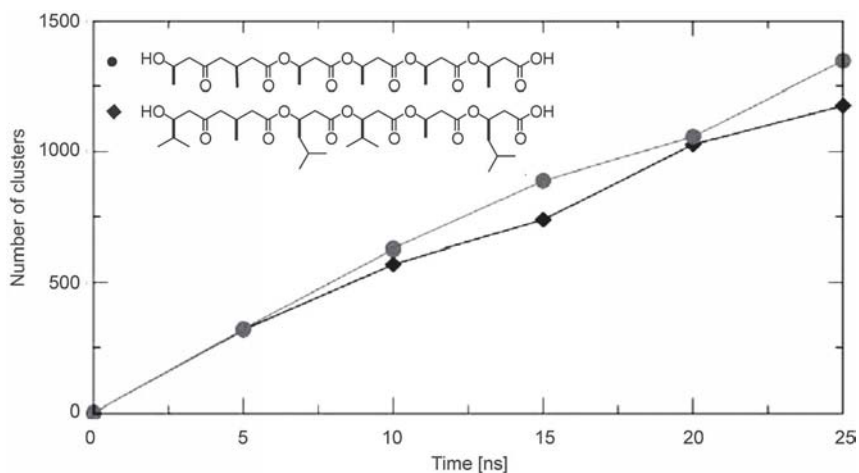
Fig. 6.9 Number of clusters (conformers) of two β-depsihexapeptides which differ in the side chain structure: (●) represents the β-depsihexapeptide with all alanine residues and (◆) represents the β-depsihexapeptide with alanine, valine and leucine side chains at 298 K and 1 atm as a function of time [45].
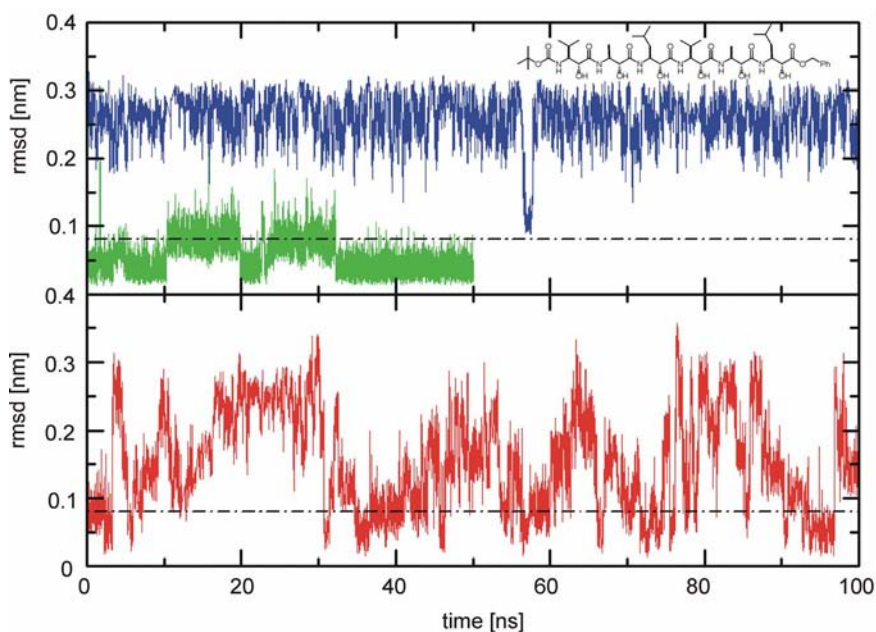
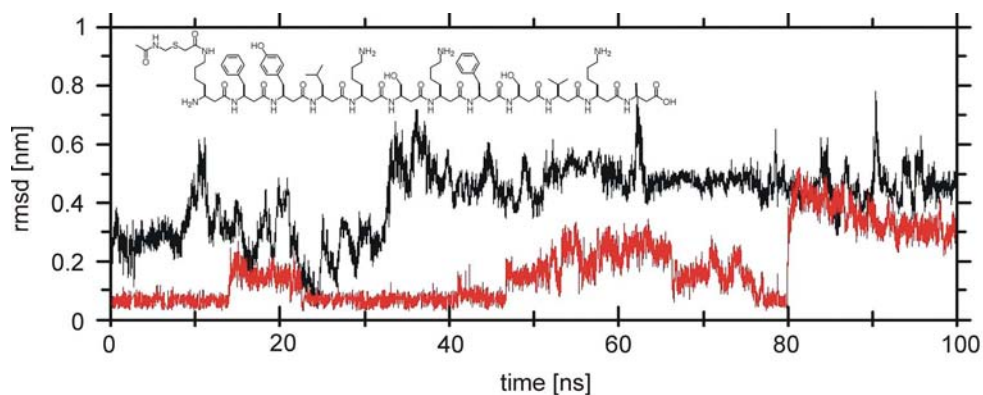**Fig. 6.10** Backbone atom-positional root-mean-square deviation of MD trajectory structures with respect to a 2.5_{12}-helical model structure (residues 2–7) derived from NMR data for a β-octapeptide in methanol at 298 K (upper panel) and 340 K (lower panel) simulated from different starting structures. Blue and red curves: an extended peptide structure; green curve: a 2.5_{12}-helical structure [36].

gence of folding equilibria of α-peptides in water is very much slower than that of β-peptides in methanol [46].

6.5
Sensitivity of the Folding Equilibrium to the Force Field Used

Most of the simulation studies of the folding equilibria of peptoids under various thermodynamic and environmental conditions made use of the GROMOS force field [47, 48], parameter sets 43A1 [49] or 45A3 [50]. These parameter sets contain nonbonded interaction parameters for nonpolar atoms which were optimized to reproduce thermodynamic properties (heat of vaporization, density, free energy of solvation) for liquid hydrocarbons and their aqueous solutions [49, 50]. Because of this thermodynamic basis, the various predominant folds of the different β-peptides with predominantly nonpolar side chains could be found in the MD trajectories in agreement with NMR experimental data [11–13, 16–19]. There seem to be no comparable studies based on other biomolecular force fields that show nearly as good agreement with the experimental data as is obtained using the GROMOS force field [51]. However, the nonbonded interaction parameters for polar atoms in the 43A1 and 45A3 GROMOS force fields had not yet been optimized to reproduce the above-mentioned thermodynamic properties for liquids of polar molecules and their aqueous solutions. Such an optimization led to the GROMOS 53A6 force-field parameter set [48]. It came as no surprise that a simulation using the 45A3 GROMOS force field for a 12-β-peptide with predominantly polar side chains could not reproduce the 3_{14}-helix experimentally observed to be stable in methanol. Only with the 53A6 force-field parameters this helical fold became stable [39], as is illustrated in Fig. 6.11. For the 20-β-peptide

Fig. 6.11 Atom-positional root-mean-square deviation of the backbone atoms of residues 2–11 (the structure of the peptide is given in the figure) with respect to the experimental NMR model structure derived for the peptide in methanol. Parameter sets 45A3 (black) and 53A6 (red) in methanol [39].

in solution (Fig. 6.4), the 53A6 force field seems to preserve the 3_{14}-helix slightly better than the 45A3 force field, but overall the picture is essentially the same for this molecule.

Not surprisingly, a calibration of force-field parameters for small molecules that represent the various moieties present in peptide analogs against thermodynamic data in the condensed phase seems a necessary condition to adequately simulate the folding equilibria of polypeptide analogs in various solvents. Not only the solvation properties in aqueous solution should correspond to experimental data, but also those for other solvents such as chloroform, cyclohexane, methanol, DMSO, acetronitrile and acetone [52], which may be used as solvents for biomolecular studies.

6.6
Comparison of Simulated with Experimentally Measured Observables

The validation of simulated folding equilibria by comparison of simulated properties of the polypeptides with measured ones is not straightforward. First, the experimental data are generally averages over a conformational ensemble. Derivation of an ensemble from average values is impossible. On the other hand, the average of a particular observable, e.g. a NOE or a ^{3}J-coupling constant, may be rather insensitive to the shape of the underlying conformational distribution over which the averaging is performed. For example, the folding equilibrium of the 7-β-peptide discussed before is rather different at 298 K, with 97% 3_{14}-helix present from the ensemble at 360 K with only 25% 3_{14} helical content. Yet the agreement with the 21 measured ^{3}J-values is as good for both quite different ensembles [53], as can be seen in Fig. 6.12. Yet, for another peptide, an 8-α-peptide, in DMSO, the ^{3}J-values are sensitive to differences in the conformational distributions in solution on the one hand and in crystal on the other [42, 54]. In DMSO solution transient M-and P-helical fragments are present, leading to a broad conformational ensemble with $\langle ^3J \rangle = 6.8$ Hz, the experimental value. In the crystal a rather narrow P-helical conformational ensemble is found with $\langle ^3J \rangle = 4.0$ Hz close to the average ^{3}J-value (4.2 Hz) of the X-ray structure.

Second, experimental data on folding equilibria are limited in number and accuracy. They may come from X-ray diffraction on crystals, or CD or NMR measurements in solution. The crystal data may only indicate that the fold that was adopted or preserved upon crystallization from a solution, is likely to be one of the dominant conformers in solution. However, the particular crystalline fold may also be induced by crystal contacts or particular co-solvents required for the crystallization. CD spectra may be very insensitive to the dominant conformers of an ensemble and may actually be determined by a fraction of the ensemble. An example of such a situation was reported in [19] where the CD spectrum was largely due to a conformer that constituted only 18% of the conformational ensemble. Regarding NMR-NOE spectra, it has been shown that different NOE peaks may show a very different sensitivity to the conformational ensemble [37].
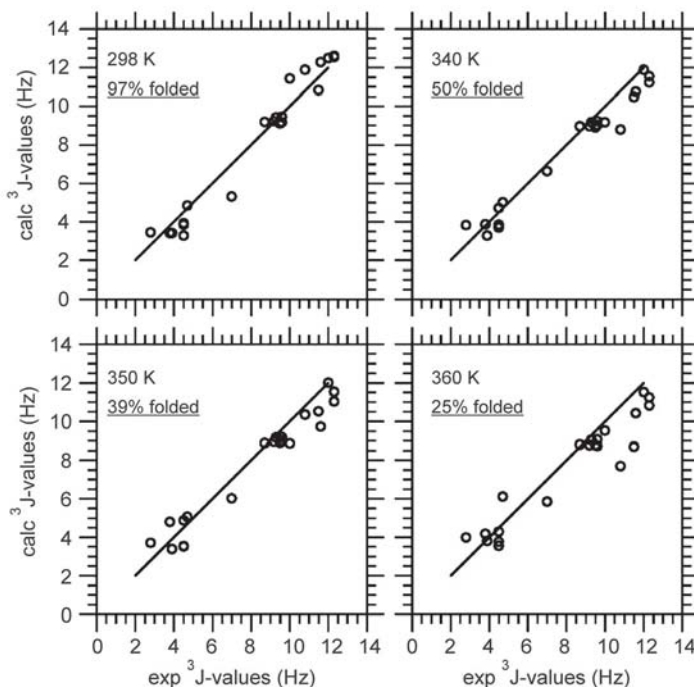
Fig. 6.12 Comparison of the 21 experimental averaged ^{3}J-coupling constants measured at 298 K with the corresponding averaged ^{3}J-coupling constants calculated for the trajectory structures of 50 ns MD simulations of a β-heptapeptide (see Panel A, Fig. 6.1) in methanol at four different temperatures [53].

An example of quite different ensembles reproducing the same experimental NOE and ^{3}J-value data for an 8-β-peptide in methanol can be found in [36]. These data appeared to be insufficient in number to uniquely determine the dominant conformer, a 2.5_{12}-P-helix or a 2_8-8-helix.

6.7
Characterization of the Unfolded State and the Folding Process

The theoretically accessible conformational space of backbone conformations of polypeptide analogs depends on the number of easily rotatable torsional angles along the backbone. Assuming three (trans, gauche$^+$, gauche$^-$) conformations per torsional angle, the theoretical number of conformations of the 7-β-peptide discussed before would be 3^{21} or about 10^9 conformers. Whether all these conformers are accessible under physiological conditions can be investigated by clustering all peptide structures from a MD trajectory of a folding equilibrium into

conformations which are characterized by a maximum atom-positional RMSD between their backbone atoms. This can be done as follows [17, 44]: the number of neighbors (that is the number of structures satisfying a given similarity criterion) is determined for each trajectory structure, with the criteria of similarity between two structures being the positional RMSD value of their main chain atoms. The structure with the highest number of neighbors is then taken as representating the first, most populated, conformation or cluster of structures. After removing the structures belonging to the first cluster from the trajectory, the procedure is repeated to find the second cluster or conformation, and so on. This clustering algorithm can also be applied to two trajectories representing different peptides of the same chain lengths or generated at different thermodynamic conditions for a single peptide. If the two trajectories sample the same part of configuration space and have similar conformational distributions, the resulting clusters will each have a comparable amount of structures from each of the two trajectories. If the two trajectories sample disjunct parts of configurational space, each cluster will only contain members of only one of the two trajectories. The former situation is illustrated in Fig. 6.13, which shows the result of a combined trajectory cluster analysis of a simulation of the 7-β-peptide at 360 K and 1 atm with one at 340 K and 1000 atm [23]. So, the conformations characterizing the unfolded state at higher pressure are the same as those characterizing the unfolded state at higher temperature. This combined trajectory cluster analysis has also been
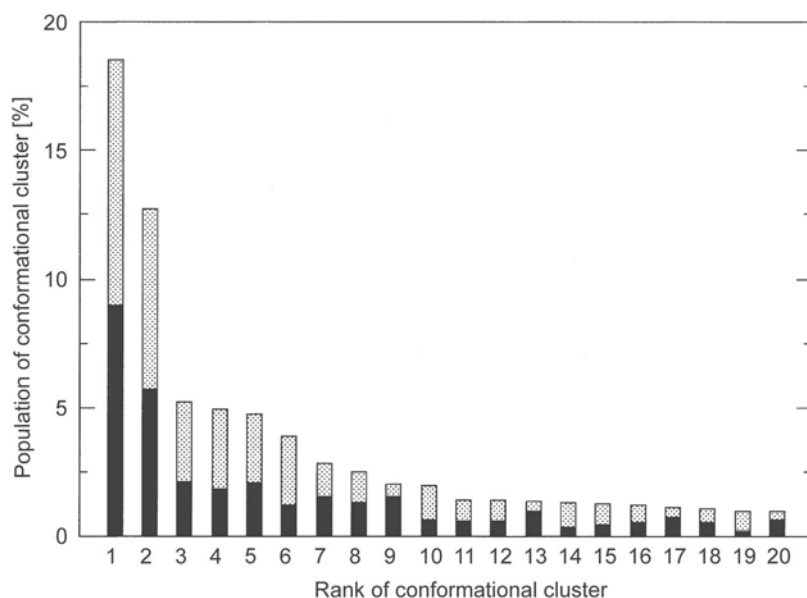

Fig. 6.13 Conformational analysis over the combined 50 ns trajectories of a β-heptapeptide (see Panel A, Fig. 1) at two different conditions: (grey) 360 K and 1 atm, and (black) 340 K and 1000 atm [23].
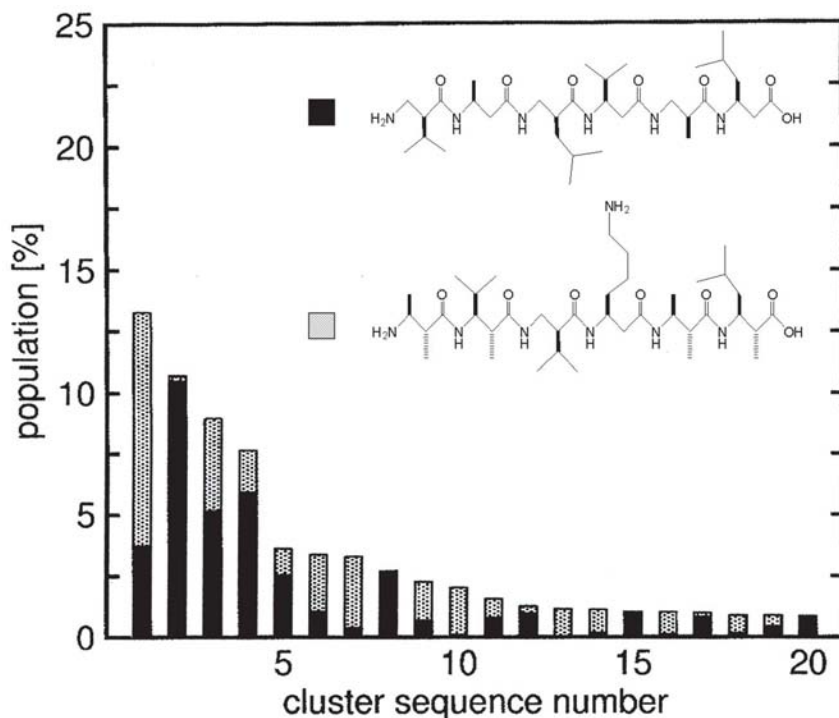
Fig. 6.14 Conformational analysis over the combined 100 ns trajectories of two β-hexapeptides (structures shown in the figure) at 340 K [55].

applied to two 6-β-peptides carrying different side chains, which made them adopt different stable folds: a 3_{14}-helix for one peptide (black bars in Fig. 14) and a hairpin for the other peptide (grey bars in Fig. 6.14) [55]. In Fig. 6.14 the result of the combined cluster analysis of the two trajectories is shown. The conformational distributions are rather different. The most populated cluster is a hairpin and the second most populated cluster a helix, as expected. Figure 6.15 shows an example of two completely disjunct folding equilibrium ensembles of two 6-β-peptides carrying identical side chains, but differing by the presence of two methylgroups at all six α-carbon positions along the main chain. These methyl groups prevent helix formation leading to a completely different conformational ensemble from that of the unmethylated peptide in methanol [19].

Folding pathways can be determined by counting the number of transitions from and to each conformational cluster [41, 44]. Such an analysis has been applied to the dynamic folding equilibrium of the 7-β-peptide [44] which has a 3_{14}-helix as most populated cluster. At 340 K, more than one pathway leads to the helical fold. Figure 6.16 illustrates that these pathways are not necessarily downhill in free energy. We note, however, that in order to obtain statistically
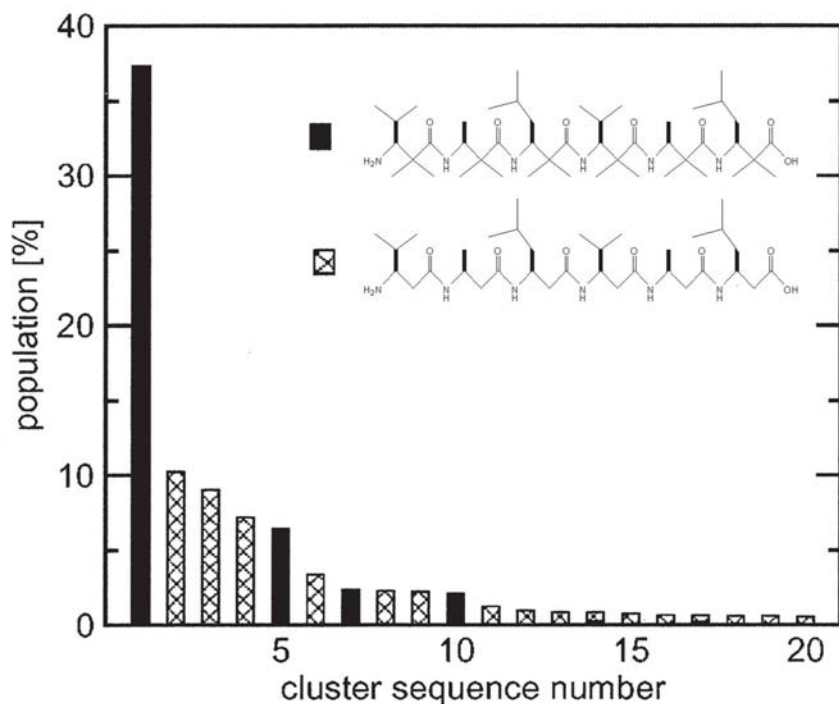
Fig. 6.15 Clustering of the combined 100 ns trajectories of two β-hexapeptides in methanol at 298 K. The plot shows the population in percentage per cluster and the portion of structures per cluster that belongs to the trajectory of each of the peptides [19].
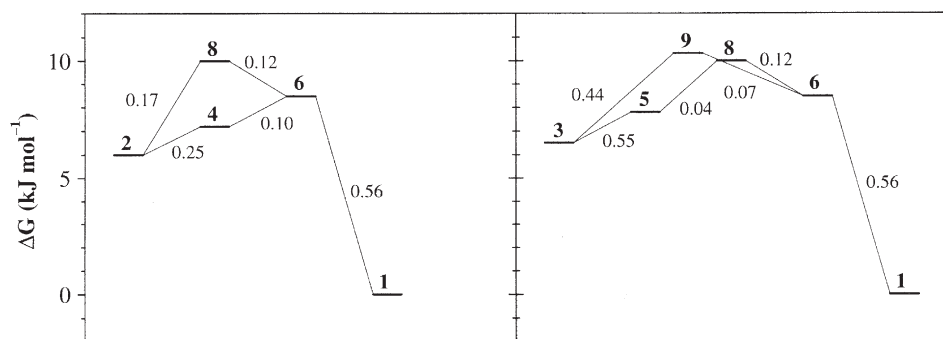
Fig. 6.16 Example of folding pathways at 340 K of a β-heptapeptide (see Panel A, Fig. 6.1), one from conformational cluster 2 (left-hand panel) and one from conformational cluster 3 (right-hand panel). The vertical axis indicates the free energy difference with respect to the helical conformational cluster 1. The transition rate (in ns^{-1}) between consecutive clusters is also indicated. Only the two shortest folding pathways (i.e., those with the minimum number of intermediate clusters) are shown [44].

converged transition rates between conformational clusters very long simulations are required.

Finally, we note that the unfolded state of the β-peptides discussed comprises many fewer conformations than the about 10^9 theoretically accessible ones [17, 56]. The number of different conformers is rather of the order of 10^2–10^3, see Figs. 6.8, 6.13–6.15. This small size of the unfolded state explains why these peptides fold on a nanosecond timescale.

6.8
Conclusion

The dynamical simulation of folding equilibria of a variety of foldamers at the atomic level including explicit treatment of solvent degrees of freedom offers the possibility of analyzing the factors that drive the conformational distribution towards a particular fold. A necessary condition to predict the various stable folds under different thermodynamic and solvation conditions in agreement with experimental data seems to be the use of a biomolecular force field that is calibrated using thermodynamic data of the condensed phase. Currently available computing power only allows adequate sampling of the conformational ensemble and the (un)folding transitions for not too large polypeptide analogs. However, the relatively small size of the denatured or unfolded state of polypeptides and the continuing rapid growth of computing power offer the perspective to simulate the folding equilibrium of a small protein within the next decade.

Acknowledgment

References

1 H.J.C. Berendsen, *Science* **2001**, *294*, 2304–2305.

2 D. Baker, W.A. Eaton, *Curr. Opin. Struct. Biol.* **2004**, *14*, 67–69.

3 C.D. Pande, E.J. Sorin, Y.M. Rhee, V.S. Pande, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43–69.

4 R.P. Cheng, S.H. Gellman, W.F. DeGrado, *Chem. Rev.* **2001**, *101*, 3219–3232.

5 D.J. Hill, M.J. Mio, R.B. Prince, T.S. Hughes, J.S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4012.

6 F. Fraternali, W.F. van Gunsteren, *J. Mol. Biol.* **1996**, *256*, 939–948.

7 M. Schaefer, C. Bartels, M. Karplus, *J. Mol. Biol.* **1998**, *284*, 835–848.

8 P. Ferrara, A. Caflisch, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10780–10785.

9 A.E. Mark, W.F. van Gunsteren, *Biochemistry* **1992**, *31*, 7745–7748.

10 V. Daggett, M. Levitt, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5142–5146.

11 X. Daura, W.F. van Gunsteren, D. Rigo, B. Jaun, D. Seebach, *Chem. Eur. J.* **1997**, *3*, 1410–1417.

12 X. Daura, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, *J. Mol. Biol.* **1998**, *280*, 925–932.

13 X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, *Angew. Chem. Intl. Ed. Engl.* **1999**, *38*, 236–240.

14 M. Takano, T. Yamato, J. Higo, A. Suyama, K. Nagayama, *J. Am. Chem. Soc.* **1999**, *121*, 605–612.

15 D. Roccatano, A. Amadei, A. Di Nola, H.J.C. Berendsen, *Protein Sci.* **1999**, *8*, 2130–2143.

16 A.M.J.J. Bonvin, W.F. van Gunsteren, *J. Mol. Biol.* **2000**, *296*, 255–268.

17 W.F. van Gunsteren, R. Bürgi, C. Peter, X. Daura, *Angew. Chem. Intl. Ed. Engl.* **2001**, *40*, 351–355.

18 X. Daura, K. Gademann, H. Schäfer, B. Jaun, D. Seebach, W.F. van Gunsteren, *J. Am. Chem. Soc.* **2001**, *123*, 2393–2404.

19 A. Glättli, X. Daura, D. Seebach, W.F. van Gunsteren, *J. Am. Chem. Soc.* **2002**, *124*, 12972–12978.

20 G. Colombo, D. Roccatano, A.E. Mark, *Proteins* **2002**, *46*, 380–392.

21 X. Wu, S. Wang, B.R. Brooks, *J. Am. Chem. Soc.* **2002**, *124*, 5282–5283.

22 X. Daura, W.F. van Gunsteren, *J. Chem. Phys.* **2006** in submission.

23 P.J. Gee, W.F. van Gunsteren, *Helv. Chim. Acta* **2006**, *89*, 475–482.

24 P.J. Gee, W.F. van Gunsteren, *Proteins* **2006**, *63*, 136–143.

25 A. Glättli, X. Daura, P. Bindschädler, B. Jaun, Y.R. Mahajan, R.I. Mathad, M. Rueping, D. Seebach, W.F. van Gunsteren, *Chem. Eur. J.* **2005**, *11*, 7276–7293.

26 D. Trzesniak, B. Jaun, R.I. Mathad, W.F. van Gunsteren, *Biopolymers* **2006**, *83*, 636–645.

27 P.J. Gee, W.F. van Gunsteren, *Chem. Eur. J.* **2006**, *12*, 72–75.

28 D. Seebach, P.E. Ciceri, M. Overh, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta*, **1996**, *79*, 2043–2066.

29 D. Seebach, R.I. Mathad, Th. Kimmerlin, Y.R. Mahajan, P. Bindschädler, M. Rueping, B. Jaun, C. Hilty, T. Etezady-Esfarjani, *Helv. Chim. Acta* **2005**, *88*, 1969–1982.

30 R. Baron, D. Bakowies, W.F. van Gunsteren, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 4055–4059.

31 R. Baron, D. Bakowies, W.F. van Gunsteren, *J. Peptide Science* **2005**, *11*, 74–84.

32 T. Soares, M. Christen, K. Hu, W.F. van Gunsteren, *Tetrahedron* **2004**, *60*, 7775–7780.

33 C. Peter, X. Daura, W.F. van Gunsteren, *J. Am. Chem. Soc.* **2000**, *122*, 7461–7466.

34 D. Seebach, Y.R. Mahajan, R. Senthilkumar, M. Rueping, B. Jaun, *Chem. Commun.*, **2002**, *15*, 1598–1599.

35 D. Trzesniak, A. Glättli, B. Jaun, W.F. van Gunsteren, *J. Am. Chem. Soc.* **2005**, *127*, 14320–14329.

36 A. Glättli, W.F. van Gunsteren, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 6312–6316.

37 C. Peter, M. Rüping, H.J. Wörner, B. Jaun, D. Seebach, W.F. van Gunsteren, *Chem. Eur. J.* **2003**, *9*, 5838–5849.

38 A. Glättli, D. Seebach, W.F. van Gunsteren, *Helv. Chim. Acta* **2004**, *87*, 2487–2506.

39 C. Oostenbrink, T.A. Soares, N.F.A. van der Vegt, W.F. van Gunsteren, *Eur. Biophys. J.* **2005**, *34*, 273–284.

40 H. Yu, X. Daura, W.F. van Gunsteren, *Proteins* **2004**, *54*, 116–127.

41 C.M. Santiveri, M.A. Jiménez, M. Rico, W.F. van Gunsteren, X. Daura, *J. Peptide Sci.* **2004**, *10*, 546–565.

42 R. Bürgi, X. Daura, A. Mark, M. Bellanda, S. Mammi, E. Peggion, W.F. van Gunsteren, *J. Peptide Res.* **2001**, *57*, 107–118.

43 N.F.A. van der Vegt, M.-E. Lee, D. Trzesniak, W.F. van Gunsteren, *J. Phys. Chem. B* **2006**, *110*, 12852–12855.

44 X. Daura, W.F. van Gunsteren, A.E. Mark, *Proteins* **1999**, *34*, 269–280.

45 P.J. Gee, F.A. Hamprecht, L.D. Schuler, W.F. van Gunsteren, E. Duchardt, H. Schwalbe, M. Albert, D. Seebach, *Helv. Chim. Acta* **2002**, *85*, 618–632.

46 L.J. Smith, X. Daura, W.F. van Gunsteren, *Proteins* **2002**, *48*, 487–496.

47 W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, **1996**.

48 C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656–1676.

49 X. Daura, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* **1998**, *19*, 535–547.

50 L.D. Schuler, X. Daura, W.F. van Gunsteren, *J. Comput. Chem.* **2001**, *22*, 1205–1218.

51 W. Damm, W.F. van Gunsteren, *J. Comput. Chem.* **2000**, *21*, 774–787.

52 D.P. Geerke, W.F. van Gunsteren, *ChemPhysChem* **2006**, *7*, 671–678.

53 X. Daura, I. Antes, W.F. van Gunsteren, W. Thiel, A.E. Mark, *Proteins* **1999**, *36*, 542–555.

54 H. Yu, M. Ramseier, R. Bürgi, W.F. van Gunsteren, *ChemPhysChem* **2004**, *5*, 633–641.

55 R. Baron, D. Bakowies, W.F. van Gunsteren, X. Daura, *Helv. Chim. Acta* **2002**, *85*, 3872–3882.

56 X. Daura, A. Glättli, P. Gee, C. Peter, W.F. van Gunsteren, *Adv. Prot. Chem.* **2002**, *62*, 341–360.

57 W.F. van Gunsteren, D. Bakowies, R. Baron, I Chandrasekhar, M. Christen, X. Daura, P. Gee, D.P. Geerke, A. Glättli, P.H. Hünenberger, M.A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N.F.A. van der Vegt, H.B. Yu, *Angew. Chem. Int. Ed. Engl.* **2006** *45*, 4064–4092.

Part 2
Function: From Properties to Applications

7
Foldamer-based Molecular Recognition

*Jorge Becerril, Johanna M. Rodriguez, Ishu Saraogi*
and Andrew D. Hamilton

7.1
Introduction

Noncovalent forces such as hydrogen bonding, electrostatics, and hydrophobic effects are ubiquitous in natural systems. Biomolecules combine many of these intrinsically weak interactions to effectively carry out a wide variety of biological processes like protein folding and substrate recognition. Inspired by nature, chemists have been able to use these forces to design new synthetic systems that are able to recognize a variety of targets. Foldamers are one class of molecules that has been developed for this purpose. Their well-characterized structural features make them ideal candidates for the development of new receptors.

Molecular recognition using foldamers can be classified into two main categories. The first one is foldamer recognition of small guests such as ions and small molecules. In general, these types of foldamers are designed to adopt circular or helical conformations that project functionality in a convergent manner to form cavity-like moieties where small molecules can bind. Binding of this type of foldamers to the guest molecules can be described as *endo-recognition* (Fig. 7.1a). Foldamers that form helices and tubular structures have been reviewed elsewhere [1] and we will focus here only on those that have been used as receptors. The second category is foldamer recognition of biomacromolecules such as proteins and DNA. In this type of recognition, the target is usually larger than the receptor molecule. The foldamer is often designed to project functionality in an ordered fashion that will complement the surface of the desired target. This is a case of *exo-recognition* (Fig. 7.1b).

In this chapter, we will focus on foldamers that recognize and bind to small molecules and biomacromolecules. We will also discuss foldamers designed to recognize crystal surfaces in an attempt to mimic biomineralization processes. For the purpose of this chapter, we will include only those foldamers for which there is evidence of a well-organized structure prior to binding.
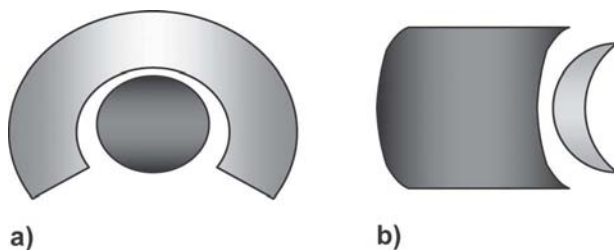
Fig. 7.1 (a) Endo-recognition of a guest molecule (dark gray) by a foldamer host (light gray); (b) Exo-recognition of a host (dark gray) by a foldamer guest (light gray).

7.2
Small Molecule Recognition Using Foldamers

The intricate networks of noncovalent interactions that biological receptors employ to recognize and bind small molecules have been a great source of inspiration for chemists. The formulation of the lock and key principle to explain enzymatic action highlighted the importance of complementarity for successful molecular recognition and played a central role in the later development of the field of supramolecular chemistry. Supramolecular chemists have been very successful in using the principles of pre-organization and complementarity to design receptors with high affinity and selectivity [2–4]. Conformationally constrained molecules such as macrocycles and capsule-like structures have been traditionally used as synthetic receptors for a wide array of small molecules and ions [5–9]. These types of hosts can surround or encapsulate a guest molecule or ion maximizing the number of stabilizing interactions. Even though tremendous progress has been made in this area, the applicability of these receptors has been limited due to intrinsic disadvantages such as synthetically difficult cyclization steps of macrocycles or excessive rigidity of capsule-like systems. There is now compelling evidence that a certain amount of flexibility is in fact beneficial in both biological and synthetic receptors as it allows for some degree of adaptability that leads to better affinity and specificity [10–12].

Oligomers that have a tendency to fold into circular or helical conformations present a great opportunity for the design of receptors that avoid difficult cyclization synthetic steps. Additionally, the inherently weak character of the noncovalent forces that help stabilize the folded structure provides these receptors with some degree of flexibility and adaptability that could be advantageous over the rigid supramolecular receptors.

7.2.1
Receptors for Water Molecules

Even though there are a number of examples of foldamers that bind to molecules of water in the solid state [13–15], there are currently only a few cases in which
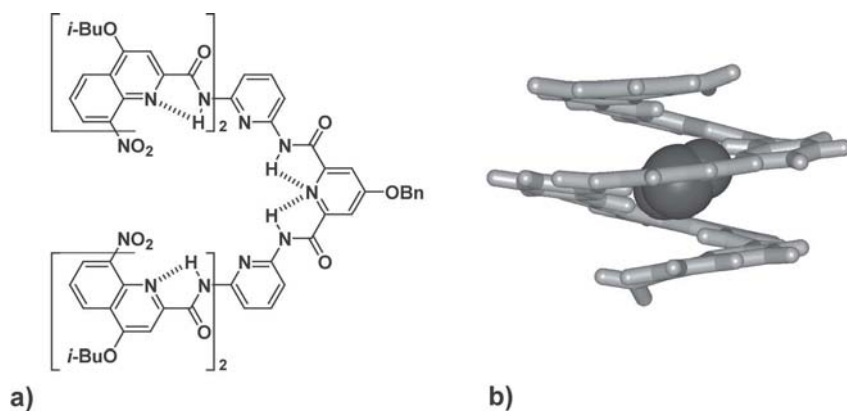
Fig. 7.2 (a) Structure of the quinoline/pyridine oligoamide foldamer. The dashed lines indicate hydrogen bonds that contribute to the stabilization of the helical conformation; (b) X-Ray structure of the quinoline/pyridine oligoamide foldamer (light gray) with a molecule of water trapped in the cavity (dark gray). Isobutyl and benzyl chains have been omitted for clarity (Reproduced with permission from the authors).

the binding to water molecules has been demonstrated in solution [16]. Pyridine oligoamides have been shown to adopt a helical conformation in the solid state and co-crystallize with water molecules in the helical cavity [15]. The restricted rotation about the aryl amide bonds and an extensive network of intramolecular hydrogen bonds are responsible for folding into a rigid structure. It has been shown that water-soluble derivatives of this foldamer maintain the folded structure even in solvents competing for hydrogen bonds such as water. Huc and co-workers transformed this helical scaffold into a capsule by adding two dimeric fragments of quinoline monomers flanking a central trimeric moiety of pyridine monomers as shown in Fig. 7.2a [16].

The X-ray structure obtained from crystals grown in organic solvents showed that the foldamer adopts a capped helical conformation with a water molecule trapped in the inner cavity (Fig. 7.2b). These results validated the design of the foldamer as a receptor for water in the solid state. However, the study of this system in solution proved to be more laborious. Since the molecule of water was completely encapsulated in the crystal structure, the binding event must involve a partial unwinding of the foldamer to allow the entrance of the guest water molecule. To test this hypothesis, Huc et al. performed variable temperature ^{1}H NMR studies in dry and wet organic solvents. When the experiments were carried out at low temperature, two different water peaks were observed which were assigned to free water and encapsulated water. This was supported by the observation that the encapsulated water showed NOEs with some of the amide NH hydrogens in the inner cavity.

7.2.2
Receptors for Ammonium Cations

Over the last few years there have been numerous reports of positively charged molecules that bind in the polar interior of circular or helical foldamers. For example, Li et al. designed and synthesized a series of hydrogen bonded aryl oligoamide foldamers [15] and studied their binding to ammonium salts in chloroform (Fig. 7.3) [17]. The researchers were able to crystallize a fragment of the foldamers' basic repeating unit which revealed a planar rigid conformation locked by a bifurcated hydrogen bond between the amide N–H and the aryl methoxy groups (see Chapter 1). Inspection of the ^{1}H NMR peaks showed significant deshielding of the amide protons indicative of strong hydrogen bonding in solution. NOE experiments showed correlation of the amide NHs with the methoxy hydrogens but not with the aromatic protons. Taken together, the NMR experiments suggest that the scaffolds adopt a planar rigid conformation in solution stabilized by hydrogen bonding. Infrared spectroscopy further corroborated the presence of
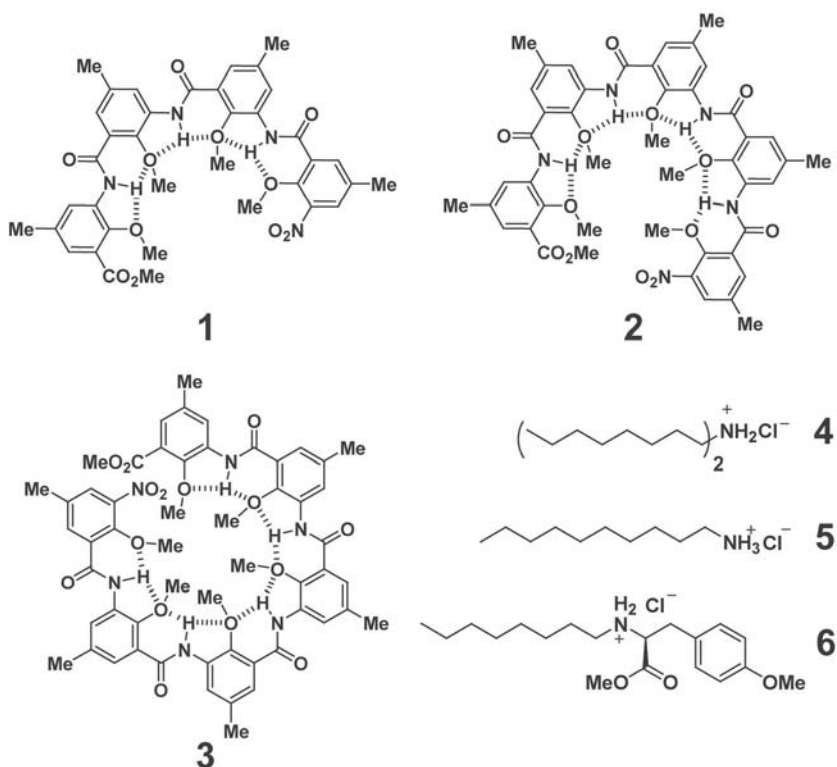

Fig. 7.3 Structures of aryl oligoamide foldamers and the ammonium salts used as guest molecules.

strong hydrogen bonds independent of the concentration which ruled out inter-molecular association.

Addition of ammonium salts **4** and **5** (Fig. 7.3) caused a substantial increase in the fluorescence intensity of the foldamers. Only a small increase in the fluorescence was observed when a quaternary alkyl ammonium salt was added suggesting that both the formal positive charge and the presence of free hydrogen bond donors in the guest molecule are essential for binding. ^{1}H NMR titration experiments were carried out by monitoring the shift of the NH signals induced by the association with the ammonium salt in chloroform. The association constants (K_a) obtained ranged from 14 to 360 M^{-1}, being higher for the less hindered primary amine **5**. The stability of the complex increased with the length of the foldamer, **2** being a better receptor than **1**, presumably because of the presence of more available hydrogen bond acceptors. However, foldamer **3** showed decreased binding affinity which was rationalized as the result of an excess of steric repulsion caused by the methyl substituents within the binding cavity.

More recently the same authors presented a modified scaffold in which the methoxy substituents were replaced by fluorine atoms (Fig. 7.4) [18]. Although there are extensive studies that suggest that covalently bound fluorine is not a good hydrogen bond acceptor, the X-ray structure of the monomer unit proved that the fluorines are involved in hydrogen bonding. The fluorine atoms interact with the amide NHs leading to a rigid well-folded structure. Large down-field shifts in the NMR spectrum were observed for the NH peaks in solution for both foldamers (**7** and **8**) compared to those obtained in control compounds providing evidence of the presence of these F−H−N bonds in chloroform.

Titration experiments with the ammonium salt of dioctylamine **4** (Fig. 7.3) gave stability constants of 4.9×10^6 and 8.1×10^6 M^{-1} for foldamers **7** and **8**, respectively, and Job plots suggested the formation of 1:1 complexes which was further corroborated by ESI-MS. Additionally, a series of NOE cross-peaks were also observed between the ammonium protons and some of the amide hydrogens of the foldamers suggesting inclusion of the ammonium ion into the cavity.
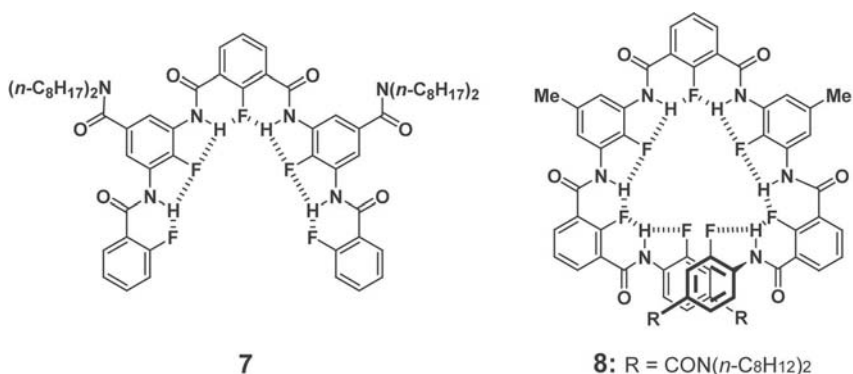

Fig. 7.4 Structure of the fluorine-containing aryl oligoamide foldamers.
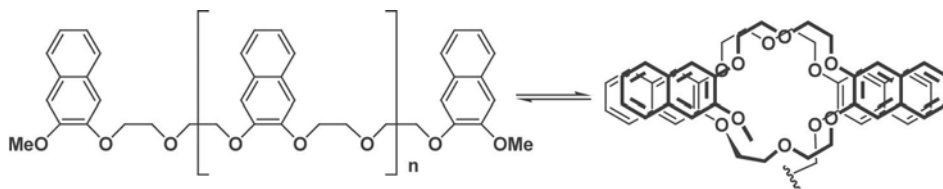
Fig. 7.5 Structure of the naphthalene-based foldamers.

In the same study, the binding of the ammonium salt of a tyrosine derivative **6** (Fig. 7.3) was investigated. The addition of the chiral guest gave rise to a CD signal in the foldamers' absorption wavelength that increased as the concentration of the guest molecule rose. These results, along with the fact that the intensity of the CD signal decreased with the addition of methanol, suggest that the foldamers are indeed binding to the guest through electrostatic or hydrogen-bonding interactions.

Solvophobic effects have been often used as the driving force to obtain well-folded structures (see Chapter 3) [19–22]. Chen and co-workers took advantage of this principle in the design of a new family of foldamers [23]. The scaffolds consisted of a series of naphthalene groups connected by diethylene glycol spacers (Fig. 7.5). In polar solvents the oligomeric chain should adopt a helical conformation in which the hydrophobic naphthalene groups are stacked on top of each other minimizing their exposure to the solvent. As a result, the ethylene glycol spacers should create a cavity that resembles the well-studied crown ethers.

The UV-vis absorption spectra of the foldamers were recorded in different chloroform–acetonitrile mixtures at a concentration low enough to avoid intermolecular association. As the polarity of the mixture was increased with a higher percentage of acetonitrile, the samples underwent a significant hypochromic shift compared to the control. This effect was caused by the π-stacking of the aromatic subunits and was therefore an indicator of the progressive folding of the scaffolds with the increasing polarity of the media. The absorption of the longest oligomers reached a plateau and remained constant for mixtures having more than 50% acetonitrile, indicating that the foldamer was completely folded at that polarity. Moreover, the fact that the hypochromic changes occurred at a lower polarity for longer, more hydrophobic oligomers supports the idea that folding is facilitated by solvophobic forces and takes place in a cooperative manner. ^{1}H NMR and fluorescence experiments also contributed compelling evidence of complete folding in polar solvents.

It is well known that macrocyclic crown ethers have high affinity for ammonium ions [24, 25]. Due to the resemblance of these foldamers to crown ethers, Chen et al. studied their ability to bind to ammonium ions. Indeed, addition of the diammonium salt of ethylenediamine in deuterated acetonitrile induced shifts in the foldamers' ^{1}H NMR resonances confirming the complexation of the

ions with stability constants around 10^4 M^{-1}. Additional Job plots indicated the formation of 1:1 complexes. NOESY experiments of the complex with the longest oligomer (Fig. 7.5, where $n = 5$) revealed intermolecular NOE cross-peaks between the foldamer methylenes and the guest CH_2 protons which further confirmed the formation of the complex. Interestingly, intramolecular NOE cross-peaks were also observed that were not present in the absence of the diammonium salt suggesting that the foldamer becomes even more compact and rigid upon binding.

7.2.3
Receptors for Hydrophobic Small Molecules

Hydrophobic small molecules were the targets of the foldamers developed by Moore et al. [26]. The tendency of aromatic rings to pack together through π-stacking interactions, preventing their exposure to a polar environment, was used in the design of a helical foldamer with a hydrophobic interior (see Chapter 3). This scaffold was based on an oligomer of *m*-phenylene ethynylene units (**9**, Fig. 7.6a) which had been shown to undergo a transition from a random coil in
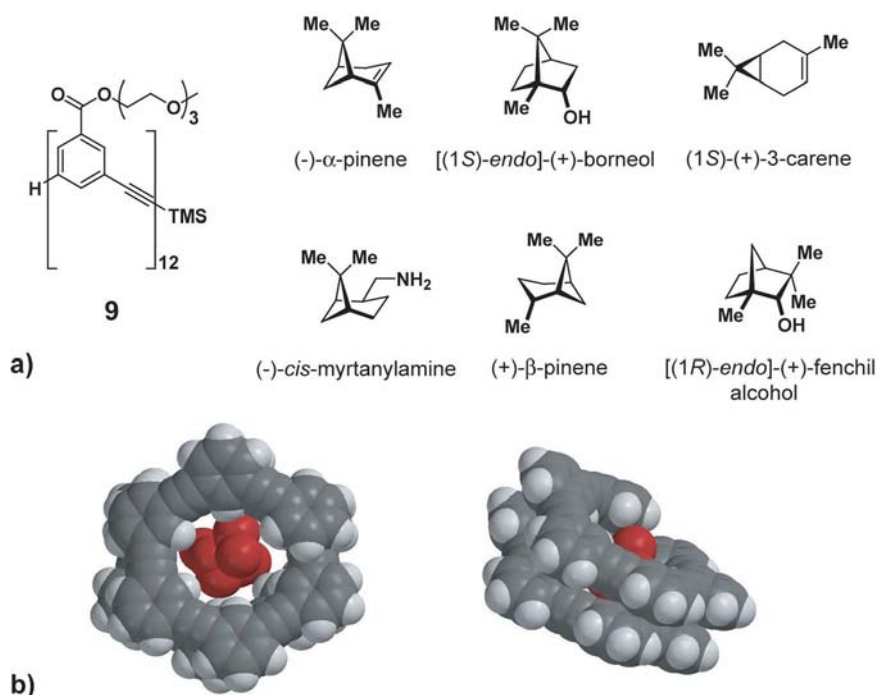
Fig. 7.6 (a) Structure of the *m*-phenylene ethynylene oligomer **9** and hydrophobic small molecules used as guests; (b) Association models of *m*-phenylene ethynylene foldamer (gray) and (−)-α-pinene (red). The polyethyleneglycol chains have been omitted for clarity.

chloroform to a compact helical conformation in more polar solvents such as acetonitrile [21]. Folding created a tubular hydrophobic cavity that could potentially bind apolar molecules. This idea was tested by addition of enantiomerically pure (−)-α-pinene (Fig. 7.6a) to a solution of the foldamer in acetonitrile. The appearance of a strong CD signal was observed in the absorption range of the foldamer which was indicative of its association with the chiral α-pinene. The addition of (+)-α-pinene resulted in the mirror image of the CD trace and further confirmed that the Cotton effect was a direct consequence of the chiral environment present in the foldamer upon binding to the chiral small molecule. Interestingly, addition of the chiral guest did not result in major changes in the UV-vis spectrum of the foldamer. This result was in accordance with the presence of a well-folded structure prior to the addition of the guest molecule since changes in the absorption pattern would have indicated major alterations of the π-stacking interactions.

CD spectroscopy was used to perform a titration and corroborate the formation of a 1:1 complex in mixtures of acetonitrile–water. The value of the binding constant for this interaction increased as the percentage of water rose indicating that binding of the hydrophobic α-pinene to the foldamer was a solvophobically driven process. Even though the foldamer was insoluble in pure water, the complex stability in this solvent was estimated to be 6×10^4 M^{-1} by extrapolating the values of the binding constants with increasing water content. The high value of this binding constant and the perfect fitting of the titration to a 1:1 binding model ruled out the presence of nonspecific interactions. Additional hydrophobic monoterpenes shown in Fig. 7.6a were tested and found to bind to the foldamer at 40% water in acetonitrile. They all formed 1:1 complexes with similar binding affinities. To confirm that the small hydrophobic molecules are binding within the cavity, two modified foldamers were synthesized that included aromatic methyl groups. The methyl substituents were designed to fill the cavity upon folding, disfavoring the binding of small molecules. Previous solvent denaturation studies showed that the addition of these methyl groups did not interfere with the folding process. Indeed, the affinity of these new structures for pure α-pinene was up to 100-fold lower than that of the foldamer with an unfilled cavity and therefore confirmed that binding occurred in the central cavity (Fig. 7.6b).

The insolubility of these receptors in pure water was addressed in a more recent study [27]. Complete water solubility was achieved by incorporating longer polyethylene glycol chains into the design. As before, α-pinene was used as the hydrophobic guest but in this case binding could be seen only when the amount of water was higher than 60%. The authors suggested that one of the polyethylene glycol chains could be folding inwards, interacting with the hydrophobic core and effectively competing with α-pinene at lower water percentages. Varying the water composition from 60 to 90% resulted in a nonlinear increase of the affinity constants ranging from 10^4 to 10^6 M^{-1}. Interestingly, the binding equilibrium was reached more slowly in higher water content mixtures supporting the idea that the formation of a partially unfolded intermediate that exposes the hydrophobic core to the solvent is the rate-limiting step in the binding process.

The cylindrical shape of the cavity led Moore et al. to consider using elongated chiral guest molecules such as the diphenylpiperazine derivative shown in Fig. 3.7 of Chapter 3 [28]. Addition of this small chiral molecule to *m*-phenylene ethynylene oligomers (16, 18, 20, 22 and 24-mer) induced a CD signal allowing the titration of the complex. The best stability constant for a 1:1 complex was 5.6×10^3 M^{-1} in 40% water. Further studies indicated that the piperazine derivative binds through insertion into the cavity rather than by intercalating between the helical loops. In order to obtain some insight into the binding mechanism, the researchers synthesized a guest with a voluminous trisaryl group at each end of a central piperazine core (Fig. 3.7, Chapter 3) [29]. The presence of bulky aryl groups did not prevent complexation although kinetic studies showed a slower rate of complexation suggesting that a partial unfolding was necessary prior to binding.

In an interesting variation of these experiments, the authors used the same hindered piperazine derivative to induce the formation of the optimal foldamer receptor through dynamic templation [30]. In this approach, imine derivatives of a group of oligomeric precursors were allowed to equilibrate with the guest (Fig. 7.7). When the experiment was carried out in an apolar solvent that did not induce folding, the result was a statistical mixture of all the possible products. However, when the experiment was repeated in a polar solvent such as acetonitrile, the composition changed with an increased ratio of the foldamer that had the optimal length that maximized the binding to the piperazine. This elegant study is the proof of concept that dynamic libraries can be used to optimize foldamer receptors.
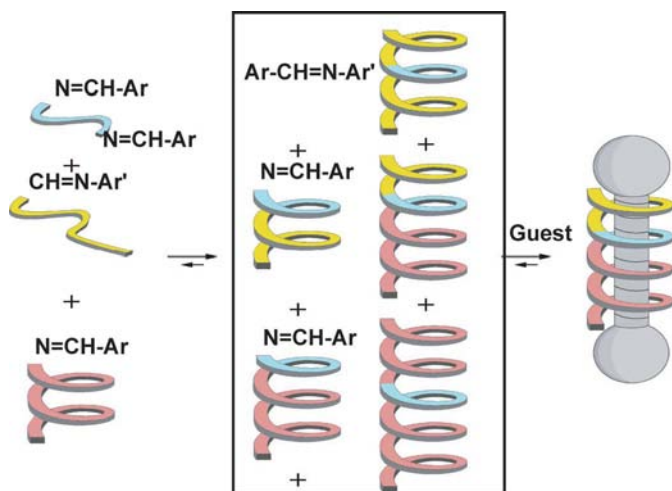

Fig. 7.7 Schematic diagram of the dynamic library approach to find the optimal length of the foldamer (Reproduced with permission from the authors).

The same scaffold has been recently modified by increasing the polarity of the cavity interior [31, 32]. In one of these variations, Moore et al. introduced a series of arylamide monomers into the foldamer sequence. This modification was aimed at the creation of a more hydrophilic cavity and was used to study the recognition of a positively charged guest.

7.2.4
Receptors for Saccharides

Saccharides play key roles in important biological processes such as cell–cell differentiation and recognition and are therefore becoming important therapeutic targets [33]. Synthetic receptors designed to target carbohydrate molecules could be used to prevent viral infectivity, transport saccharides across membranes and facilitate the introduction of drugs into specific cell types. The potential applications of saccharide recognition have prompted the development of numerous synthetic receptors based mostly on macrocycle and calixarene scaffolds [34]. Foldamers that can recognize and bind saccharides could similarly have broad applications.

Prompted by the successful design of intramolecularly hydrogen bonded foldamers, Li and co-workers further developed foldamers **10**, **11** and **12** (Fig. 7.8a) by introducing a series of 1,2-dibenzoyl hydrazide groups that can form three-centered H bonds [35]. The presence of an extensive network of intramolecular hydrogen bonds forces the oligomers to fold into a rigid circular or helical conformation depending on the length of the oligomer. Crystal structures of the hydrazide building blocks confirmed the existence of the predicted hydrogen bonds and ^{1}H NMR and NOESY experiments established the presence of a compact well-folded structure in chloroform which is stabilized by an extensive hydrogen-bonding network. According to computer models, the folded structure should po-
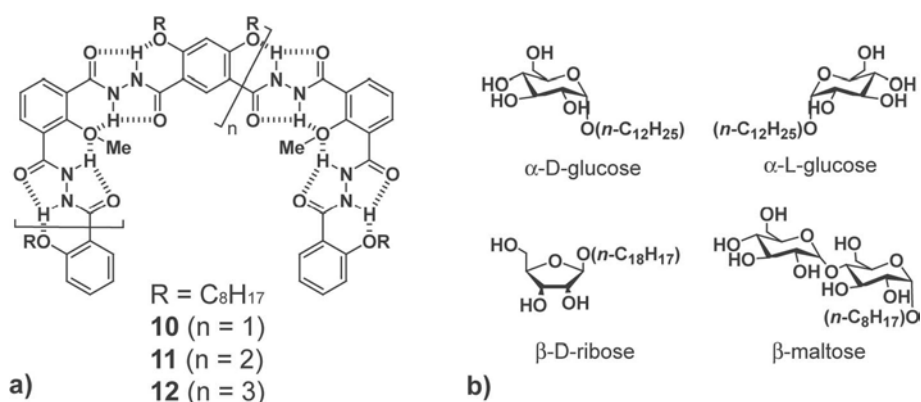
Fig. 7.8 (a) Structures of the aryl 1,2-dibenzoyl hydrazide oligomers; (b) Saccharide derivatives employed for binding experiments.

sition the carbonyl groups in the interior of the cavity creating a circular array of hydrogen bond acceptors.

The addition of mono and disaccharides (Fig. 7.8b) to a solution of the foldamer in chloroform induced a moderate CD signal. When methanol was added the CD absorption disappeared, indicating that the association between the foldamer and the saccharide occurs mainly through hydrogen bonds. The formation of strong complexes with the longer oligomers induced an intense downfield shift of the saccharide OH signals. However, addition of the saccharides to solutions of short oligomers did not cause much change in the ^{1}H NMR peaks. Job plots indicated the formation of 1:1 complexes and the stability constants were obtained by NMR titrations and fluorescence spectroscopy. The highest binding contant was 6.9×10^6 M^{-1} corresponding to the complex of the longest foldamer **12** and the disaccharide β-maltose (Fig. 7.8) indicating that the presence of numerous hydrogen bond donors and acceptors leads to greater stabilization of the complex. Additional NOE experiments showed cross-peaks between the foldamer and the guest molecules suggesting that the saccharides bind in the inner cavity of the foldamer.

Two similar foldamers based on an oligobenzamide scaffold were recently reported (Fig. 7.9) [36]. Binding of these oligomers to triol **15** and some saccharides was investigated by ^{1}H NMR titrations in chloroform. The binding constants ranged from 5.5×10^2 M^{-1} for the shortest foldamer **13** with β-D-ribose, to 7.2×10^3 M^{-1} for the association of **14** with disaccharide β-maltose. The stability constant values indicate that the association is more favorable as more hydrogen bond donors and acceptors are available.
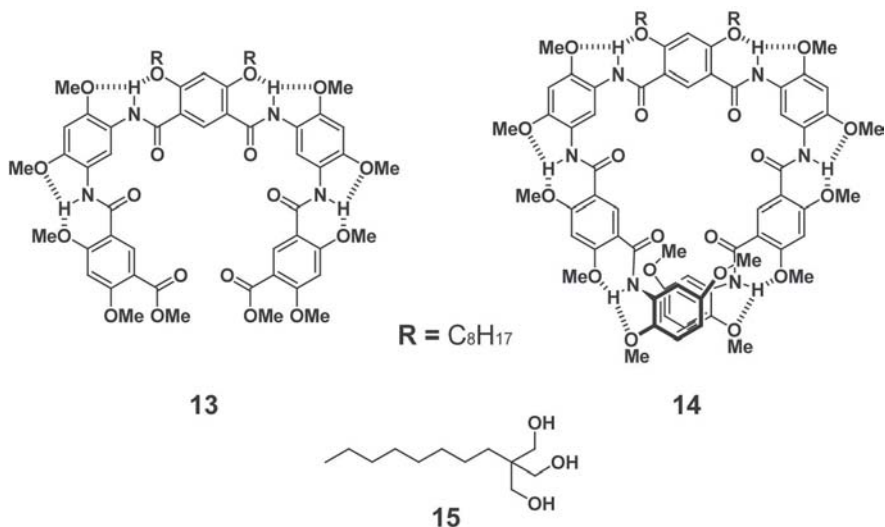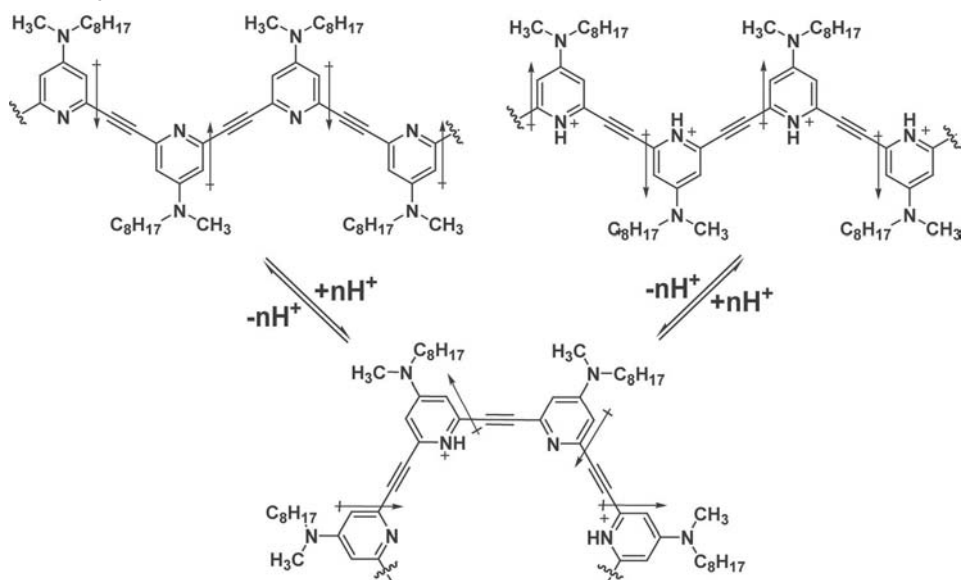

Fig. 7.9 Structure of foldamers **13** and **14**, and triol **15**.

Fig. 7.10 The role of protonation on the local conformation of the foldamer. Partial protonation maximizes dipole-dipole interactions by forcing a folded conformation.

Inouye et al. have recently reported the synthesis of poly(pyridine ethynylene) foldamers and their use as pH-dependent saccharide receptors [37]. In this study, monomers of dialkylaminopyridine derivatives were used to obtain a polymer with an average length of 45 units (Fig. 7.10). Gradual addition of trifluoroacetic acid and subsequent protonation of the pyridines induced a progressive change in the UV-vis spectrum associated with the folding of the polymer into a helical conformation.

It had been previously shown that binding to hexoses can induce the transition of this type of foldamer from an unstructured to a helical conformation in chloroform [38]. The protonation-driven folding of the polymer should yield pre-organized saccharide receptors with higher affinities as entropic penalty upon binding is decreased. In fact, titration of the partially-protonated foldamer with a series of modified hexoses showed an increase in complex stability when compared to binding in the absence of acid. The increase in the binding constants ranged from a 2-fold for β-D-mannose (3300 to 7200 M^{-1}) to a 200-fold increase for β-D-fructose (100 to 20 000 M^{-1}). Interestingly, a great excess of acid seemed to be detrimental as it induced the unfolding of the structure. The folding behavior of the polymer could be rationalized as being the result of the interaction between the monomer dipoles in solution. As shown in Fig. 7.10b, the individual dipoles created by the dialkylaminopyridine groups force the backbone to adopt

an extended conformation to minimize dipole–dipole repulsion. However, protonation of one pyridine group induces an inversion of the local dipole and the rearrangement to a transoid conformation to maximize the dipole–dipole attraction. As more acid is added, the all-transoid conformation is favored leading to complete folding of the foldamer. When an excess of acid is added, however, the electrostatic repulsion between the positively-charged pyridinium ions forces the polymer to unfold. Other effects such as the presence of charge-transfer complexes between pyridinium and pyridine groups in consecutive loops of the helix certainly have an important role as well.

7.2.5
Receptors of Other Organic Molecules

The foldamers presented so far rely on the folding process to create a cavity that can host smaller molecules (endo-recognition). A different alternative is to use the foldamer molecule as a structural scaffold from which to project appropriate functionality that will interact with potential guests in a spatially controlled manner (exo-recognition).

An example of the second approach is the work of Li et al. in which a previously reported polyarylamide foldamer was ingeniously used as the framework from which to project multiple zinc porphyrins in a radial manner [39, 40]. A short oligomer containing two porphyrin units was synthesized and NMR studies demonstrated that the new receptor was able to bind C_{60} and C_{70} derivatives in apolar solvents [39]. As an extension of this work, a longer foldamer was synthesized (Fig. 7.11a) [40]. ^{1}H NMR experiments confirmed the presence of intramolecular hydrogen bonds expected in the folded conformation of the circular core. Fullerenes have been shown to interact with porphyrins through $\pi-\pi$ interactions, whereas pyridine and imidazole functionalities are well-known zinc binders. A series of bidentate guest molecules containing both types of functionalities were therefore synthesized (Fig. 7.11b) and their binding to the foldamer was studied. Both the change in the UV-vis spectra and the quenching of the fluorescence of the porphyrins were used to monitor the titrations and estimate the binding constants and stoichiometry of the complexes. Binding constants of up to 10^6 M^{-1} were obtained in chloroform. Interestingly, the Job plots for these systems revealed that up to six guests bind to one receptor.

Recently, polymers containing an average of 60 *m*-phenylene ethynylene units have been functionalized with alkyl sulfonate and L-alanine derivatives to increase their water solubility (Fig. 7.12a) [41, 42]. Like the previously discussed oligomeric analogs, the polymers have a high propensity to undergo folding in polar solvents. The presence of the sulfonic acid in **16** and the carboxylic acids in **17** ensured that the polymers were completely soluble in water and in other polar solvents such as methanol and DMSO. UV-vis, CD and fluorescence spectra were all consistent with a fully folded structure of **17** at water contents of 60% or more in methanol.
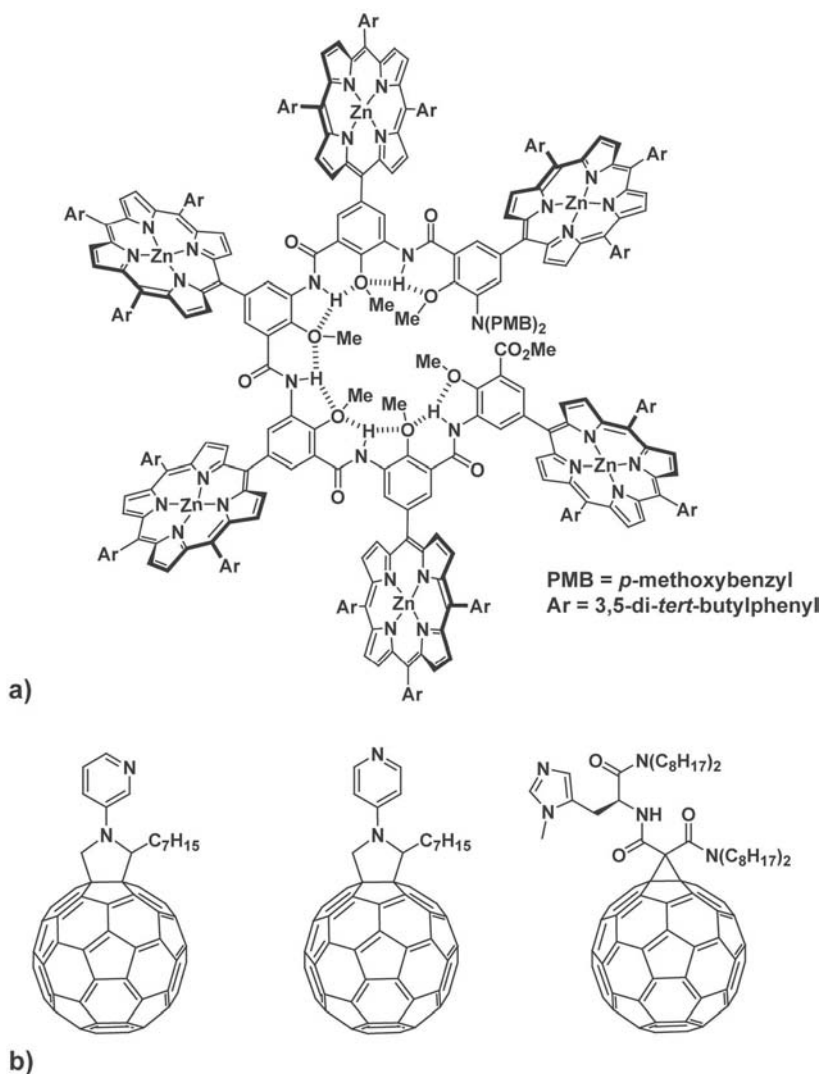
Fig. 7.11 (a) Porphyrin-containing foldamer; (b) Bidentate ligands for the porphyrin-containing foldamer.

The foldamers have a central helical hydrophobic core formed by stacked aromatic units and project the negatively charged groups outwards. This general structure presents interesting analogies with a DNA double helix and the possibility of using well known DNA binders to interact with the foldamers. The complex $[Ru(bpy)_2(dppz)]^{2+}$ is a well-known DNA intercalator that becomes photoluminescent after intercalation into the DNA. When the ruthenium complex was
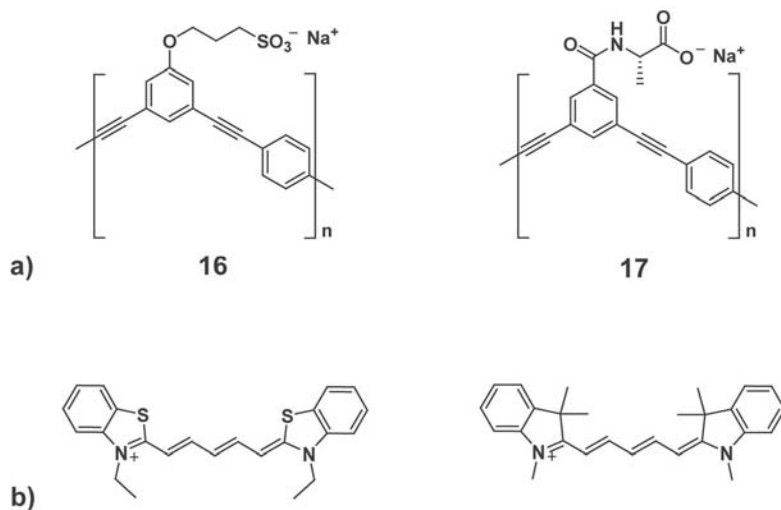
Fig. 7.12 (a) Alkyl sulfonate and L-alanine derivatives of the poly(*m*-phenylene ethynylene) foldamer; (b) Cationic cyanine dyes used in the binding studies.

added to a solution of the foldamers in water the characteristic orange-red photoluminescence of the complex could be observed. The stoichiometry was estimated as being approximately two metal complexes bound per turn of the helix.

Foldamer **17** was used as a receptor for a series of cationic cyanine dyes (Fig. 7.12b). Binding of these species was studied by the changes in the absorption spectra of the dyes. Additionally, the dyes presented a CD signature when bound to the chiral foldamer **17** which led the authors to suggest that the flat dye molecules form an aggregate in which they intercalate between the turns of the helix. This arrangement maintains the positive charges of the guest molecules and the negative charges of the foldamer in close proximity.

In addition to the foldamers discussed so far, there are a growing number of examples of oligomers that acquire a rigid folded conformation upon binding to a guest molecule or ion. For instance, an oligoindole scaffold has been shown to fold into a helical conformation when chloride is added to a chloroform solution [43]. Hydrogen bonding to the chloride ion stabilized the folded conformation even in highly competitive solvents such as 10% water in acetonitrile. In a previous study nitrile groups were added to Moore's phenylene ethynylene oligomer [44]. The modified foldamer underwent folding upon complexation of silver ions in THF. Additionally, many oligomers with coordinating groups such as pyridine and pyrrole have also been studied and their complexes with different metals ions characterized. These complexes receive the general name of helicates and have been extensively reviewed elsewhere [45].

7.3
Protein Recognition

Biomacromolecules such as DNA and proteins use a variety of structural elements to recognize their respective binding partners including α-helices, β-sheets, bulges, and turns. Subtle differences on the interacting surfaces of biomacromolecules are advantageously used to achieve high levels of specificity and affinity. Disruption of these sensitive interactions can affect cellular pathways and ultimately lead to a variety of diseases. The use of molecular recognition principles to design synthetic molecules that modulate the interaction between biomacromolecules, is therefore an attractive strategy for therapeutic intervention.

Within the past decade, peptides and oligonucleotides have been successfully used to recognize macromolecules such as proteins and DNA; unfortunately, their intrinsic susceptibility to enzyme degradation by proteases and peptidases limit their applicability *in vivo*. Foldamers, or well-folded non-natural oligomers, offer an attractive alternative due to their structural similarity to natural biopolymers and stability towards degradation [46]. A commonly used approach to target biomacromolecules of interest is to use their native substrates as templates for the design of new foldamers. The use of this strategy led to the development of a plethora of structural mimetics, such as β-peptides, peptoids, and peptide nucleic acids (PNAs), that mimic naturally occurring biopolymers (see Chapters 1, 2 and 8) [47–49]. A common feature of this family of peptidomimetics is a polyamide backbone similar to the one found in natural systems. Alternatively, totally synthetic foldamers that mimic only the functional epitope of extended regions of proteins but not the polyamide backbone have also been developed [50, 51]. In the following section, we will focus on the design strategies used by various groups to develop foldamers to target biomolecules. As in the previous section, we will discuss only those examples where the presence of a folded structure prior to binding was well characterized.

The development of modulators of protein–protein interactions has remained a challenging task primarily because such large interfacial areas are involved [52]. However, despite these obstacles there is an increasing number of examples where non-natural ligands (small molecules and foldamers) have succeeded in disrupting protein–protein interactions [52–56].

When developing modulators of protein–protein interactions, one of the two interacting partners is often used as a template for the design. Structural analysis of the complex, with methods such as X-ray crystallography, NMR, and alanine scanning, can reveal those structural features that are important for complex formation. With this information in hand, it may be possible to reproduce or graft the functional epitope onto a non-natural scaffold. Furthermore, the affinity of a molecule can be improved by developing stabilizing interactions not present in the natural system through the modification of functional groups.

7.3.1
Abiotic Synthetic Foldamers

Some of the most widely studied protein–protein interactions include those involved in cellular apoptosis. Irregularities associated with the apoptotic pathway can lead to a variety of disorders ranging from cancer to degenerative and auto-immune diseases [57, 58]. Major focus has been placed on the Bcl-2 family of proteins which are key regulators of the apoptotic pathway [59]. In particular, the disruption of the Bcl-x_L/Bak complex has been shown to induce apoptosis in unhealthy cells and is therefore a potential target for pharmacological intervention [60]. Fesik and co-workers found that the Bak peptide forms an amphipathic α-helix upon binding into a hydrophobic cleft formed by the BH1, BH2, and BH3 regions of Bcl-x_L [61]. Bak projects four hydrophobic side chains (Val74, Leu78, Ile81, and Ile85) along one face of its α-helix into the cleft. Furthermore, alanine scanning experiments suggested that these hydrophobic residues corresponding to the i, $i + 4$, $i + 7$, and $i + 11$ positions on the Bak α-helix are key for complex formation. This information makes possible the design of Bak mimetics that can potentially bind to Bcl-x_L.

Hamilton and co-workers have recently reported a novel trispyridylamide foldamer that inhibits the Bcl-x_L/Bak interaction (Fig. 7.13a) [50]. This scaffold was designed to mimic the i, $i + 4$, and $i + 7$ residues of the BH3 domain of Bak. A stabilizing bifurcated hydrogen-bonding network allows **18** to adopt a preferred
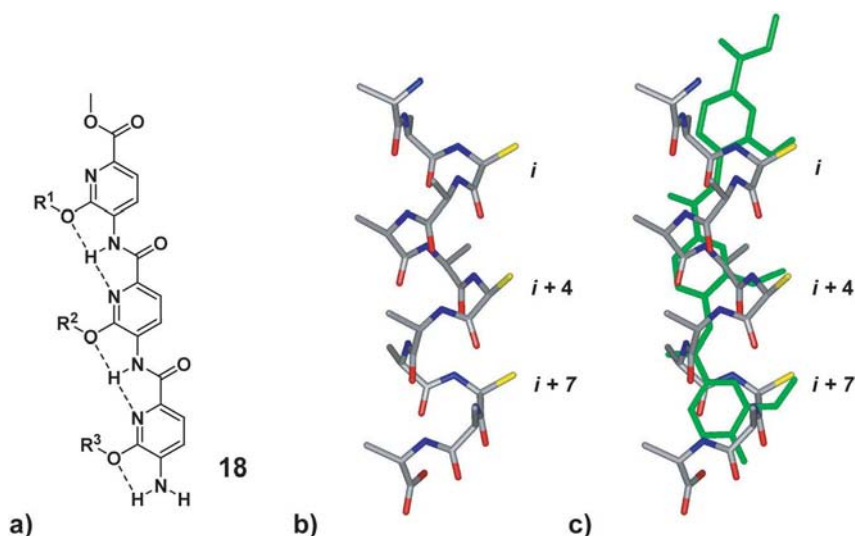

Fig. 7.13 (a) Struture of the trispyridylamide foldamer **18**; (b) Idealized poly-alanine α-helix; (c) Overlay of **18** and poly-alanine α-helix.

conformation in which all three R-functional groups are projected on the same face of the molecule. Overlay of a minimized energy conformation of **18** with the α and β positions of the i, $i+4$, and $i+7$ methyl groups of a poly-alanine α-helix shows close correspondence of the two structures with an RMSD = 0.94 Å (Fig. 7.13 b, c). The overall conformation of the backbone and the position of the side chain functional groups predicted by MM2 minimization calculations were confirmed in the solid state by an X-ray crystal structure of a nitro derivative of **18**. In addition, ^{1}H NMR studies including concentration and temperature dependence experiments were performed to study the stability and hydrogen-bonding pattern in **18**. These experiments suggested that the amide proton is intramolecularly hydrogen bonded in solution both in polar and nonpolar solvents as well as in the solid state.

A competitive binding fluorescence polarization (FP) assay was used to determine the ability of derivatives of **18** to disrupt the Bcl-x_L/Bak complex. Inhibition constants (K_i) in the low micromolar range were found for the trispyridylamide foldamers. The best inhibitor had a side chain substitution pattern of $R^1, R^2 = $ Bn, and $R^3 = i$Pr and a K_i value of 1.6 μM. Differences in the binding affinities were observed when the R-groups were varied. For instance, the derivative with three R = Bn substituents had a K_i value greater than 20 μM, emphasizing the importance of surface complementarity and the selectivity that can be gained by the variation of side chains.

7.3.2
Peptidomimetic Foldamers

Recently, Gellman and co-workers have developed an $(\alpha/\beta + \alpha)$-peptide ligand for the same BH3-recognition cleft of Bcl-x_L [62]. For their design, they studied the different helices adopted by oligomers with a 1:1 alternation of α- and β-amino acid residues along the foldamer backbone (see Chapter 2) [63]. Pure β-peptides, as well as α/β-peptide designs, which formed 11-helices promoted by a five-membered ring constrained β-amino acid, showed no activity in FP assays. On the other hand, those α/β-peptides that adopted a 14/15-helix showed some activity. To gain affinity, Gellman et al. replaced either the N or C terminus of their best α/β-peptide inhibitor with a fragment of the natural α-peptide sequence of Bak. When the N terminus was modified no activity was observed. On the other hand, the C-terminal modified chimeric peptide **19** showed a K_i value of 1.9 nM (Fig. 7.14). In this $(\alpha/\beta + \alpha)$ peptide, a Leu was introduced in the 6th position and a Phe at the 12th position of the 14/15 helix in order to mimic Leu78 and Ile84 of the Bak peptide respectively. In addition, an Arg at position 4 and an Asp at position 11, were included to increase binding via electrostatic interactions with residues on the edge of the Bcl-x_L cleft. Because of the proteolytic susceptibility of the α-segment in **19**, future work will be aimed at replacing it with a nonnatural segment.

The hDM2/p53 complex is another well characterized protein–protein interaction involved in apoptosis [64]. In unstressed normal cells, p53 is present at very
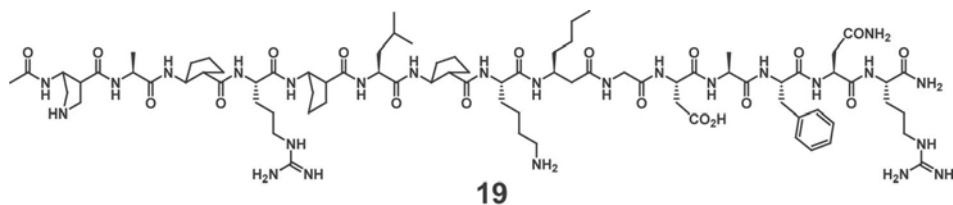
Fig. 7.14 Structure of the $(\alpha/\beta + \alpha)$ chimeric scaffold.

low levels due to rapid degradation via a ubiquitin-dependent proteasome pathway [65]. In cancerous cells, however, hDM2 is overexpressed abrogating the ability of p53 to induce cell cycle arrest and apoptosis by binding p53 and promoting its degradation [66]. For the inhibition of the hDM2/p53 interaction, Schepartz et al. designed a class of 14-helical β-peptide inhibitors [67, 68]. They sought to mimic three residues on the activation domain of p53 (Phe19, Trp23, and Leu26) which had been shown to be important for the heterodimerization of p53 and hDM2 [69]. In general short β-peptides lack a well-defined secondary structure in water. However, Schepartz et al. were able to obtain well-folded 14-helix structures in water by neutralizing the helix macrodipole. To do this, on one of the three faces of the 14-helix β-peptide scaffold, they introduced a positively charged side chain (β^3-homoornithine) near the N terminus and a negatively charged residue (β^3-homoglutamate) at the C terminus while still maintaining a stabilizing salt bridge network (Fig. 7.15). This was the first example where this strategy was used to stabilize short β-peptide 14-helix structures. The second face of the 14-helix was primarily substituted with β^3-homovaline residues, while the third face was used to present the functional epitope necessary for protein recognition. FP assays showed that **20** bound to hDM2 with nanomolar affinity and displaced a p53 peptide derivative from hDM2 with an IC_{50} = 94.5 μM. Scrambling of the β^3-homophenylalanine, β^3-homotryptophan, and β^3-homoleucine resulted in loss of affinity, emphasizing the importance of mimicking the structural features of the side chains in maintaining specificity and affinity.
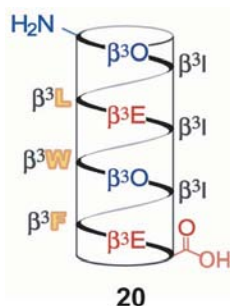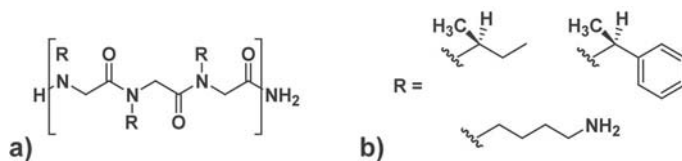

Fig. 7.15 Structure of 14-helix β-peptide **20** (Reproduced with permission from the authors).

Fig. 7.16 (a) General structure of a peptoid; (b) Peptoid monomer side chains used by Barron and co-workers.

Oligomers of *N*-substituted glycines, known as peptoids, have also been developed to mimic proteins [51, 70]. In peptoids, the peptidic side chains are shifted to main chain nitrogen atoms that eliminate stereogenic centers and the possibility of forming hydrogen-bonding networks (Fig. 7.16a). Even though intramolecular hydrogen bonding is not possible, the inclusion of appropriate bulky chiral substituents can induce peptoids to acquire secondary structures (see Chapter 1, Section 1.4.1) [71, 72]. In 2005, Barron et al. reported a class of helical peptoids that mimic the lung surfactant (LS) protein B (SP-B) [73]. During respiratory distress, SP-B interacts with the lipid film by an unknown mechanism to enable breathing. SP-B is an amphipathic 79-amino acid protein with four disulfide bonds. Three of the disulfide linkages are intramolecular, while the fourth is intermolecular forming a homodimer. The presence of multiple disulfide bonds in the SP-B homodimer complicates its synthesis, thus synthetic mimics of SP-B are currently being investigated for use as additives or biomimetic lung surfactant replacements.

In an effort to synthesize a peptoid mimetic of SP-B, Barron and co-workers modeled their design after a small amphipathic N-terminus segment of SP-B (SP-B$_{1-25}$) that had been shown to retain much of the surfactant function. They designed three 17-mer peptoids based on earlier reports describing the structural properties of peptoids whose sequence incorporated certain bulky α-chiral aromatic and aliphatic side chains as shown in Fig. 7.16b. The first peptoid included exclusively α-chiral aromatic side chains with an achiral lysine-like monomer at every third position to create a cationic face and reproduce the amphipathicity of SP-B$_{1-25}$. The second incorporated both α-chiral aromatic and aliphatic residues in addition to the lysine-like monomers. Finally the third mimetic consisted of only α-chiral aliphatic side chains as well as the lysine-like monomers. As expected from earlier reports, the CD spectra of the first two peptoids in water showed features corresponding to a poly-proline type I-like helix. The third was found to be a random coil due to the lack of α-chiral aromatic monomers though it is the most similar to SP-B$_{1-25}$ in that it contains no aromatic residues. All three peptoids, however, exhibited spectra characteristic of helices in a lipid environment.

Four different *in vitro* techniques were used to characterize the surface-active properties of the peptoids. Their activities were compared to two peptides (SP-B$_{1-25}$ and KL$_4$) that had been previously established, by *in vivo* as well as *in vitro* studies, to be good mimics of SP-B. Extensive *in vitro* characterization of the pep-

toids' surface activities showed that their properties were similar to those previously reported for other SP-B mimetics. Barron and co-workers also reported that this similarity appeared to correlate with the overall helicity and hydrophobicity of a given peptoid.

The scope of peptidomimetic compounds is not limited to mimicking α-helices. In fact, β- and γ-peptides have also been used to mimic β-turns (see Chapter 2, Section 2.5). One of the earliest examples where a β-peptide was used to recognize a biomacromolecule was reported by Seebach and co-workers [74]. They developed a cyclic β-tetrapeptide analog of octreotide that would bind to human somatostatin receptors. Octreotide is a cyclic α-octapeptide derived from somatostatin that is currently in use to treat acromegaly and intestinal cancers [75, 76]. SAR studies performed on octreotide revealed that the amino acids in the β-turn (Phe-Trp-Lys-Thr) were required for activity [76]. Though the affinity of the cyclic β-tetrapeptide was an order of magnitude lower than octreotide, this study demonstrated that a β-peptide could be used as a β-turn mimetic.

More recently, Seebach and co-workers designed an unconstrained β-peptide that would be predisposed to fold [77]. They incorporated an α-branched $(\beta^2)/\beta$-branched (β^3) β-dipeptide sequence into their design which had been shown to induce a β-turn structure with a 10-membered intramolecular hydrogen bond [78]. They made a tetrapeptide **21** (Fig. 7.17) as a potential somatostatin (SRIF) mimic by attaching a Lys and a Trp residue to the β^2/β^3 motif, while the other two residues were β^3-HPhe and β^3-HThr. Two other peptides were synthesized for comparison purposes. The first was an all-β^3-amino acid peptide lacking the β^2/β^3 motif **22** and it was expected to fold into an α-helix. The second was an all-α-amino acid peptide expected to be unstructured in solution. One-dimensional NMR spectroscopy and CD were used to obtain information about the secondary structures of the peptides. In the ^{1}H NMR, the C(γ)-H protons of peptide **21** are shifted upfield in agreement with a β-turn conformation where its C(γ)-H protons are in close proximity to the tryptophan indole ring. The C(γ)-H protons of β^3-peptide **22**, showed no such upfield shift indicating the absence of a β-turn. In the CD spectra of **22** and the α-peptide no Cotton effect was observed, while **21** shows a maximum at 203 nm in methanol. This agrees well with the spectrum of a well-characterized β-peptide turn.
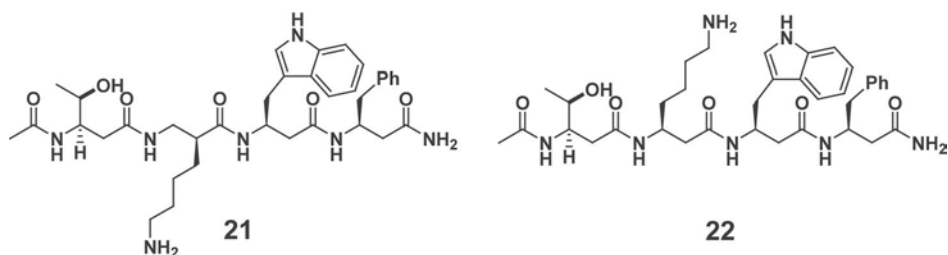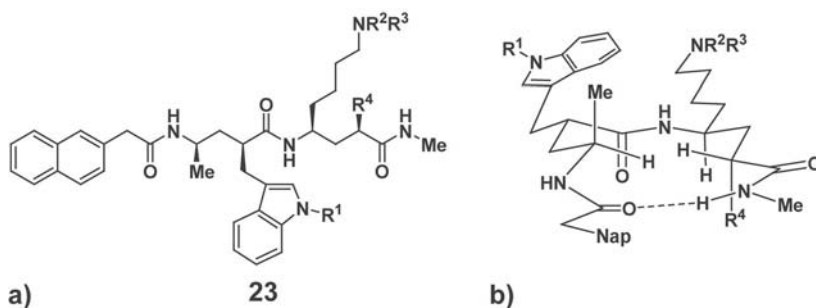

Fig. 7.17 Structure of β-peptides **21** and **22**.

Fig. 7.18 (a) Structure of one of the four γ-peptides **23**; (b) Proposed conformation of **23** showing β-turn structure.

A radioligand-binding assay was performed to evaluate the affinity of the peptides for five human recombinant SRIF receptors. Both the linear β^3-peptide **22** and the α-peptide showed low affinity for the receptors with K_d values higher than 10 μM. On the other hand, **21** showed a significant and selective affinity for the human receptor sst$_4$ ($K_d = 83$ nM). Octreotide was found to have 20 times higher affinity than **21**, whereas somatostatin had an affinity 20 times lower. Interestingly going from peptide **21** to **22**, which only differ in the position of the Lys chain, resulted in nearly a 1000-fold decrease in the affinity for sst$_4$.

Seebach and co-workers have developed a series of γ-dipeptide derivatives that also target human somatostatin receptors [79]. They designed four peptides that included a tryptophan sidechain in the γ^2 position of the first γ-amino acid and a lysine in the γ^4 position of the second (Fig. 7.18). Similar to the β-peptides previously discussed, the NMR of all four γ-dipeptides prepared confirmed the proximity of the lysine chain and tryptophan ring, indicative of a turn conformation as seen for **23** in Fig. 7.18b. In addition, the CD spectra of the peptides exhibited an intensive negative Cotton effect further supporting the idea of the presence of secondary structure. The peptides were tested in a radioligand-binding assay against five human somatostatin receptors. Their best inhibitor exhibited a K_d value of 510 nM.

Synthetic foldamers can also be used to mimic the quaternary structural elements of proteins. In this approach, the recognition occurs between two or more synthetic partners yielding an all-artificial structure that resembles a protein-like assembly. Recently, Schepartz et al. have reported the recognition between structured β-peptide helices that form a complex [80]. Two β-peptide monomers, Acid-1F and Base-1F, were designed to recognize each other (Fig. 7.19). Both contained free N and C termini as well as one face that included alternating positively and negatively charged residues to stabilize the 14-helix conformation. β^3-homoleucine residues were incorporated into positions i and $i + 3$ of the 14-helix and β^3-homophenylalanine residues were incorporated at positions 4 and 7 to favor interhelical interactions. In addition, at positions 1 and 10, a
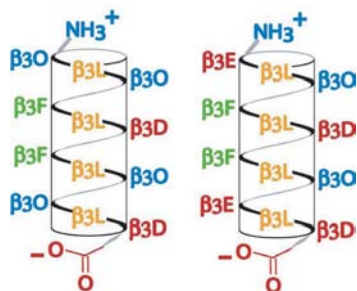
Fig. 7.19 Structure of β-peptides Acid-1F (left) and Base-1F (right) (Reproduced with permission from the authors).

β^3-homoornithine (positively charged) or a β^3-homoglutamic acid (negatively charged) were incorporated into the design to favor heterodimerization.

CD spectroscopy was used to study the helicity of the single 14-helices as well as a 1:1 mixture of the two. The spectra revealed that the monomers had 14-helical structure in solution that was strongly enhanced in a 1:1 mixture indicating higher levels of 14-helix structure. Job plots suggest that the CD signal was due to a 1:1 ratio of heteroligomers. This complex was shown to have a melting temperature of 58 °C similar to that of a well-folded protein. Moreover, sedimentation equilibrium experiments were consistent with either a hetero-hexamer or a hetero-octamer equilibrium complex. Several derivatives of Acid-1F and Base-1F were synthesized to probe the structural requirements necessary for complex formation. Studies done with these variants showed that the aromatic residues at position 4 and 7 are important. In addition the β-homoleucine face was shown to be necessary, probably making multiple interactions within the complex. Disulfide exchange experiments also revealed the presence of homo- and heterodimers. This is a significant step towards the design of artificial protein-like structures.

7.4
Mimicry of Biomineralization: Recognition of Crystal Surfaces Using Foldamers

In the preceding sections, we have discussed how foldamers have been used to recognize small molecules, ions and macromolecules. In this section we will describe the use of foldamers to recognize the ordered arrangement of ions along the faces of a crystal.

7.4.1
### Introduction to Biomineralization

Although the phenomenon of crystallization has been known for a very long time, the roles that factors such as solvent, temperature and impurities play in

this process are still not well understood [81]. A precise control over the crystallization process, which is essentially a type of molecular recognition, is still a formidable challenge [82]. The most remarkable example of such control comes from Nature in the form of biomineralization. In Nature, the growth of inorganic minerals (biominerals) is rigorously controlled by organic biomacromolecules like proteins and polysaccharides. The highly regulated conditions give rise to structures with shape, size, morphology and properties very different from those obtained in the laboratory [83–90]. Interestingly, the biomacromolecules are often incorporated into the inorganic crystal lattice during the growth process, giving rise to organic–inorganic hybrid structures. These composite materials combine the rigidity of inorganic substances and the toughness of organic materials to provide organisms with stronger building units [90, 91]. Biominerals are often superior to man-made materials in terms of their mechanical strength, resistance to fracture and other physical properties [91]. An example of this is calcite, a polymorph of calcium carbonate, which grows in the laboratory as a brittle rhombohedral crystal (Fig. 7.20a). Under biological control, however, it attains a very different shape (Fig. 7.20b) and enough strength to impart protection and mechanical support to the animal, e.g. mollusk shells. One of the most remarkable features of biomineralization is that it occurs under physiological conditions unlike the manufacture of man-made materials, like cement, that requires extremes of temperature and pressure [92]. Due to the desirable properties of biominerals, there is much interest in mimicking biomineralization to obtain superior materials with tailored properties [93].

In Nature, biomineralization is controlled by biomacromolecules that project functionality in an ordered way to selectively recognize the growing faces of a


a)   b)
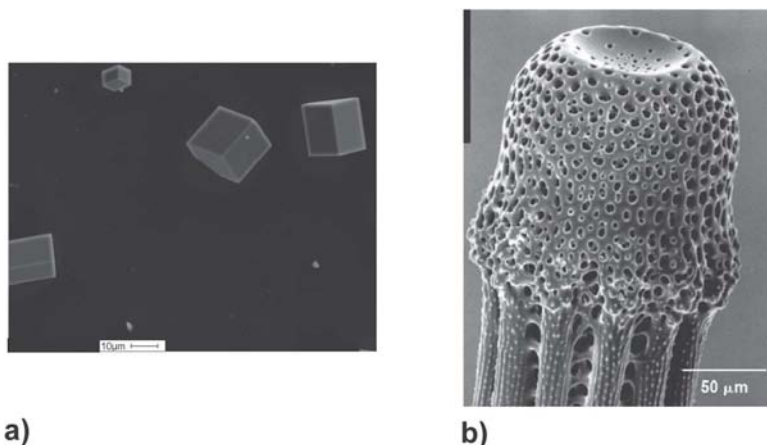
Fig. 7.20 (a) Calcite crystals grown in the laboratory; (b) A mature sea urchin spine composed of calcium carbonate and organic macromolecules (Reprinted with permission from *Science* **2003**, *299*, p. 1192. Copyright 2003 AAAS).
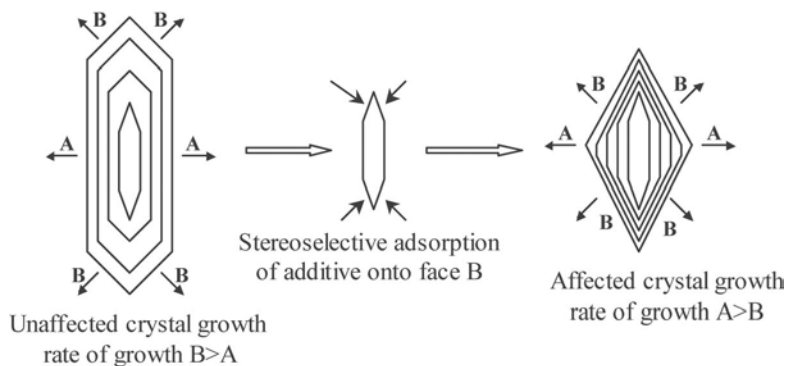
Fig. 7.21 Schematic diagram showing crystal growth in the presence of inhibitor for face B (Reproduced with permission from the authors).

crystal. An effective strategy to mimic this process is to present arrays of functionality designed to recognize and bind to specific crystal faces [90, 94]. This involves molecular recognition between the organic additive and the inorganic ions at the surface of a growing crystal. When an organic additive selectively recognizes a specific face of a growing inorganic crystal, it is adsorbed and further accumulation of inorganic ions on that face is hindered (Fig. 7.21). These faces gradually increase in area while others that are not recognized by the additive, decrease in area and ultimately disappear as the crystal grows. As a result, only the faces that are recognized by the organic molecule are manifested in the equilibrium morphology of the crystal while the rest of the faces are not observed. This implies that examination of the faces that appear in a crystal can provide information about the specific interactions between the additive and the affected face [94]. Conversely, it is possible to design additives that are complementary to a specific face of a crystal that can be used to alter the morphology of a crystal (Fig. 7.22)
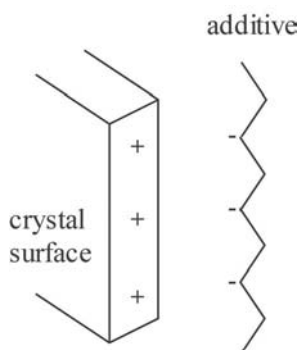
Fig. 7.22 Schematic diagram showing structural complementarity between the interacting groups.

[89]. Charge, size, distance and stereochemical complementarity are important parameters to be considered when designing effective crystal growth modulators [84, 90]. The role of interfacial interactions on the conformational preference of foldamers is discussed in detail in Chapter 13.

During the last two decades, numerous types of templates have been used to obtain various morphologies of common minerals. These include conformationally constrained peptides [95], polymers [96–101], self-assembled monolayers [102–105], other supramolecular assemblies [106] and small molecules [107–111] and have been reviewed elsewhere [85, 112–114]. Since many organic macromolecules involved in biomineralization are known to have acidic moieties, a common design strategy involves projecting negatively charged groups (e.g. carboxylates, phosphates) in an ordered manner [115]. The negatively charged groups presumably recognize the arrangement of cations in the initial stages of crystallization. Foldamers designed to project negatively charged functionality in a rigid manner could interact selectively with specific crystal faces and act as effective crystal growth modifiers.

7.4.2
Biomimetic Synthesis of Calcite Using Foldamers

Calcium carbonate is one of the most widespread biominerals found in nature. Numerous reports of the control of morphology and polymorph selectivity of calcium carbonate have been published. Calcium carbonate has six polymorphs: calcite, aragonite, vaterite, two hydrated forms and one amorphous form [90]. Of these, calcite and aragonite are the most commonly found biominerals while vaterite is rather rare. Calcite, the thermodynamically most stable form of $CaCO_3$, crystallizes in a trigonal unit cell and has rhombohedral morphology. Aragonite crystallizes in an orthorhombic crystal system and forms needle like crystals. Finally, vaterite crystals are spherical with a hexagonal unit cell.

Ueyama et al. have reported the crystallization of $CaCO_3$ in the presence of aromatic polyamide foldamers with regularly repeating carboxylates (Fig. 7.23) [116]. The structures are held in a fixed conformation by intramolecular H bonds as evidenced by IR and NMR. When $CaCO_3$ crystals were grown in the presence of compound **24** the morphology of the calcite obtained was affected. Inspection of the crystals obtained with **24** revealed that the growth of the {401} face had been inhibited. This plane has a Ca^{2+} ion spacing of about 4.96 Å between adjacent calcium ions. Since **24** linearly projects carboxylates at a repeat distance of ~10 Å, the negative charges are optimally placed to interact with alternate Ca^{2+} ions. On the other hand, the trans geometry of the double bonds in **25** precludes the linear arrangement of carboxylates and it is unable to induce a morphological change in calcite. The authors also detected **24** occluded within the inorganic crystals by using cross polarization/magic angle spinning NMR techniques. This further corroborated the fact that **24** was involved in the crystallization of calcite.

In a later paper, the same group reported alternately amidated poly(acrylate)s which form 8-membered rings stabilized by hydrogen bonds between the

a)

24

b)

25
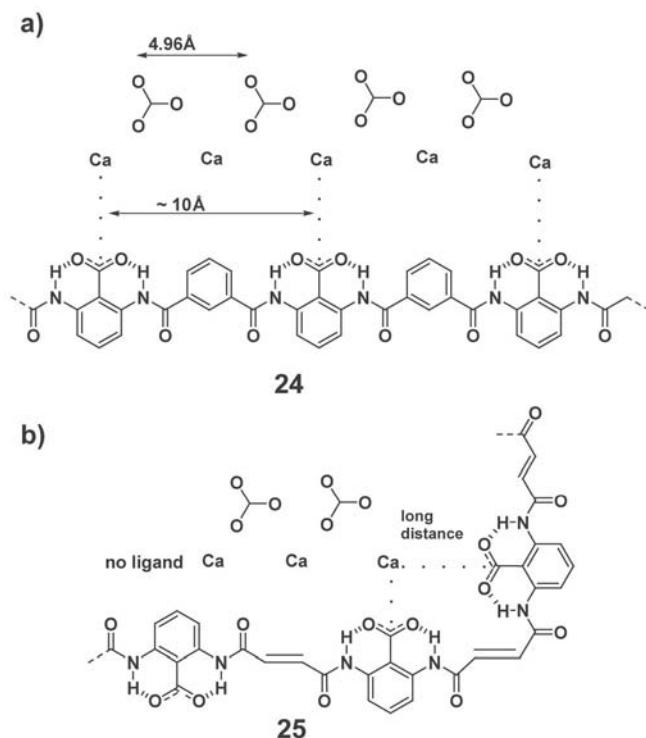
Fig. 7.23 (a) The polyamide **24** with carboxylates regularly repeating at a distance of 10 Å optimal for interaction with the calcite lattice; (b) The polyamide **25** with the trans geometry of the fumaryl spacer not suitable for interaction with the calcite lattice.

carboxylate oxygen and the amide NH (Fig. 7.24) [117]. When the poly(acid)s or their sodium salts were used as additives during the crystallization process, the CaCO₃ crystals obtained were found to be predominantly the vaterite polymorph. In contrast, the crystals obtained with the calcium salts were mainly calcite. The authors proposed that these polymers control the polymorph selectivity at the nucleation stage presumably by stabilizing the nascent nuclei of one polymorph over the others.


R= n-Bu, t-Bu
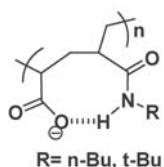
Fig. 7.24 Alternately amidated poly(acrylates) showing 8-membered H-bonded rings used by Ueyama and co-workers.
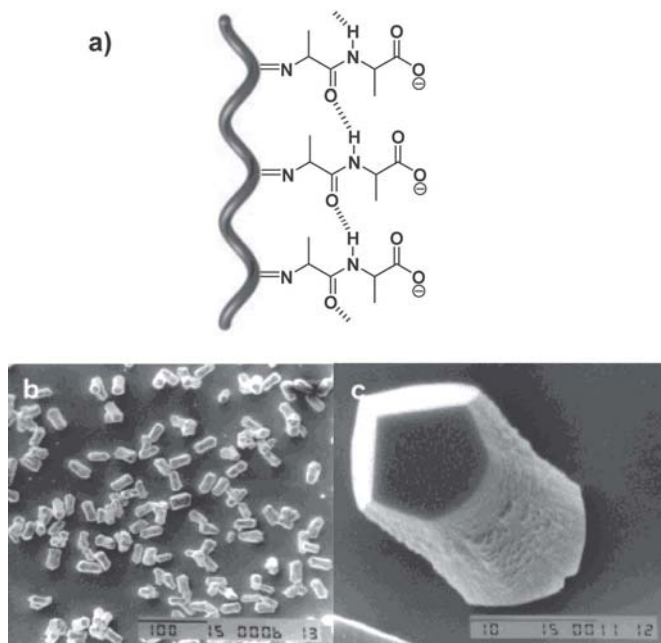
Fig. 7.25 (a) The polyisocyanide-based scaffold, showing the H-bonding arrays between the side chains used by Sommerdijk and co-workers; Calcite crystals obtained in the presence of the polyisocyanide-based scaffold; (b) at low magnification; (c) at high magnification showing the apple core morphology. (Reprinted in part with permission from *J. Am. Chem. Soc.* **2002**, *124*, p. 9700. Copyright 2002 American Chemical Society).

Sommerdijk and co-workers have used peptide based polyisocyanides as organic templates for the crystallization of calcium carbonate [118]. These molecules adopt a helical conformation due to restricted rotation of the polymer backbone and H-bonding interactions between the dipeptide side chains (Fig. 7.25a) [119]. CD data showed that the helical conformation was further stabilized in the presence of calcium ions. Interestingly, the handedness of the helix could be changed by switching the chirality of the amino acids. When the sodium salt of poly(L-isocyanoalanyl-D-alanine) was added to the crystallization solution, the calcite crystals obtained had an "apple core-type morphology" (Fig. 7.25 b, c). The molecular weight and polydispersity of the polymers were shown to have no effect on crystal growth. The faces of the {*hk.0*} family were expressed in the crystals obtained. They identified the nucleating plane as (01.1) and suggested that the orientation of the carboxylates in the polymer mimics the orientation of the carbonates on this face. When the corresponding L,L isomer was used, the crystals obtained were not as uniform, presumably due to a less well-defined structure for the polymer as seen from its CD spectrum. This underlines the importance of well-folded structures and controlled projection of functionality to effectively
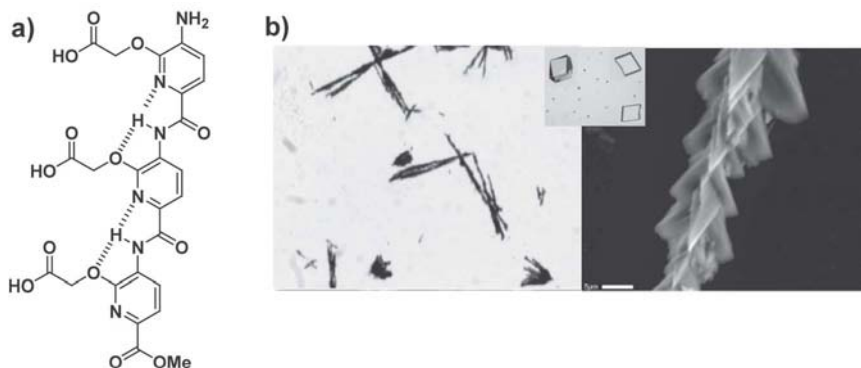
Fig. 7.26 (a) The foldamer used by Hamilton and co-workers; (b) Calcite crystals (90×) obtained in the presence of the foldamer under optical microscope (left), under scanning electron microscope (scale bar: 5 μm) (right) and control calcite crystals (inset).

modify crystal growth. Similar results were obtained with these foldamers in the case of calcium phosphate where rod-like hydroxyapatite crystals were formed [120]. From thermogravimetric analysis, the authors were able to show that the crystals contained 5% by mass of the polymer adsorbed in the crystal.

Hamilton and co-workers reported a synthetic foldamer with a polyamide backbone that is held in a linear conformation due to bifurcated H bonds and projects carboxylates in an ordered manner from one face of the molecule (Fig. 7.26a) [111]. The design was based on a α-helix mimetic that was reported earlier from the Hamilton group [50]. The side chains were designed to include carboxylates that would mimic the aspartates and recognize the arrangement of calcium ions on the surface of a growing calcite crystal. The $CaCO_3$ crystals grown in the presence of these foldamers were shown to be calcite from the characteristic peaks in the IR spectrum at 713 and 876 cm^{-1}. The crystals were elongated and had a "saw-tooth" morphology very different from the usual rhombohedral geometry of control crystals (Fig. 7.26b). Overgrowth experiments were performed in which the elongated calcite crystals were re-introduced into the crystallization solution to study the orientation and identity of the new crystal faces formed. From these results and modeling data, the expressed faces were identified as the $\{-10l\}$ where $0.5 \leq l < 1$. The arrangement of the carbonates on these faces was proposed to be well mimicked by the carboxylates of the foldamer.

Volkmer et al. have used short peptides (4 and 8 residues) containing an alternating Phe-Asp repeat to obtain calcite crystals similar to those obtained from biological samples [121]. They assigned the newly formed faces as {11.0} and {01.2} which are also commonly found in biomineralized calcite. Although the authors have yet to determine the three-dimensional structure adopted by these peptides in solution, they hypothesized that the peptides take up preferred conformations in which the spacing of the carboxylates matches the positions of the calcium ions on the expressed crystal faces.
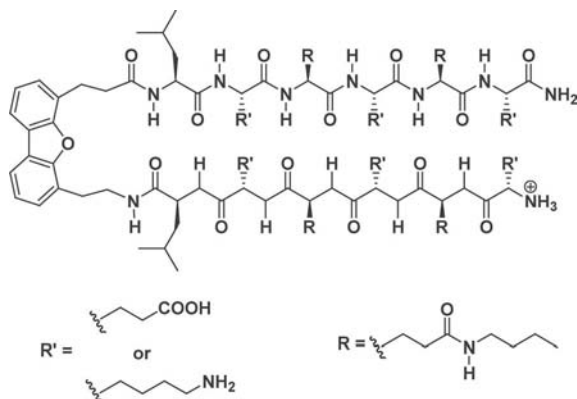
Fig. 7.27 The peptidomimetic used by Kelly and co-workers to obtain controlled growth of CdS crystals.

7.4.3
Biomimetic Synthesis of CdS Using Foldamers

Besides calcite, a variety of other biominerals have been explored including calcium phosphate, cadmium sulfide and silica. Kelly and co-workers have reported amphiphilic peptidomimetics that fold into β-hairpins (Fig. 7.27) which subsequently self-assemble into a β-sheet monolayer at the air–water interface [122]. They have used these peptidomimetics to grow oriented CdS crystals with controlled particle size. Both acidic (–COOH) and basic (–NH$_2$) functionality were used on the side chains to determine the effect of negatively and positively charged groups on CdS growth. Transmission Electron Microscopy (TEM) and Atomic Force Microscopy (AFM) images showed that only the peptidomimetic with –COOH functionality was successful in generating oriented CdS nanocrystals with a narrow size distribution. From electron diffraction studies they concluded that the crystals had the wurtzite lattice structure. The crystals were roughly oriented in the c direction and were nucleated from the {01.0} face. This face has a Cd–Cd distance of 6.7 Å which matches the 6.9 Å distance between the α carbons of the glutamate residues in the peptidomimetic. The fact that the peptidomimetic with –NH$_2$ was not found to give uniform homogeneous CdS crystals, highlights the importance of electrostatic complementarity between the additive and a crystal face.

7.5
Conclusion

During the last decade, there has been considerable progress in the field of foldamer design, particularly in the area of foldamers as useful receptors for the rec-

ognition of molecules. Nature provides the most exquisite examples of well-folded structures that can recognize and bind to desired targets with high affinity and specificity. This is achieved by projecting functionality in a manner that makes optimal use of different noncovalent interactions. In the same way, foldamers designed to recognize a wide variety of molecules and surfaces could find important application in fields such as sensor technology, drug design and delivery, and materials science. The intrinsic structural features of foldamers make them ideal scaffolds for the design of new functional receptors, and we foresee the development of a significant number of new foldamer-based receptors in the years to come.

Acknowledgments

References

1 M.A.B. Block, C. Kaiser, A. Khan, S. Hecht, *Top. Curr. Chem.* **2005**, *245*, 89–150.

2 F. Diederich, *Angew. Chem.* **1988**, *100*, 372–396.

3 M.V. Rekharsky, Y. Inoue, *Chem. Rev.* **1998**, *98*, 1875–1917.

4 T.H. Webb, C.S. Wilcox, *Chem. Soc. Rev.* **1993**, *22*, 383–395.

5 X.X. Zhang, J.S. Bradshaw, R.M. Izatt, *Chem. Rev.* **1997**, *97*, 3313–3361.

6 L. Kovbasyuk, R. Kramer, *Chem. Rev.* **2004**, *104*, 3161–3187.

7 G.W. Gokel, W.M. Leevy, M.E. Weber, *Chem. Rev.* **2004**, *104*, 2723–2750.

8 F. Hof, S.L. Craig, C. Nuckolls, J. Rebek, Jr., *Angew. Chem. Int. Ed.* **2002**, *41*, 1488–1508.

9 R. Warmuth, J. Yoon, *Acc. Chem. Res.* **2001**, *34*, 95–105.

10 A.M. Davis, S.J. Teague, *Angew. Chem. Int. Ed.* **1999**, *38*, 736–749.

11 A.J. Kirby, *Angew. Chem. Int. Ed.* **1996**, *35*, 707–724.

12 J.K.M. Sanders, *Chem. Eur. J.* **1998**, *4*, 1378–1383.

13 V. Berl, I. Huc, R.G. Khoury, J.M. Lehn, *Chem. Eur. J.* **2001**, *7*, 2810–2820.

14 V. Maurizot, J.M. Leger, P. Guionneau, I. Huc, *Russ. Chem. Bull.* **2004**, *53*, 1572–1576.

15 I. Huc, V. Maurizot, H. Gornitzka, J.-M. Leger, *Chem. Commun.* **2002**, 578–579.

16 J. Garric, J.-M. Leger, I. Huc, *Angew. Chem. Int. Ed.* **2005**, *44*, 1954–1958.

17 H.-P. Yi, C. Li, J.-L. Hou, X.-K. Jiang, Z.-T. Li, *Tetrahedron* **2005**, *61*, 7974–7980.

18 C. Li, S.-F. Ren, J.-L. Hou, H.-P. Yi, S.-Z. Zhu, X.-K. Jiang, Z.-T. Li, *Angew. Chem. Int. Ed.* **2005**, *44*, 5725–5729.

19 R.S. Lokey, B.L. Iverson, *Nature* **1995**, *375*, 303–305.

20 S. Sakurai, H. Goto, E. Yashima, *Org. Lett.* **2001**, *3*, 2379–2382.

21 J.C. Nelson, J.G. Saven, J.S. Moore, P.G. Wolynes, *Science* **1997**, *277*, 1793–1796.

22 Y. Zhao, Z.Q. Zhong, *J. Am. Chem. Soc.* **2005**, *127*, 17894–17901.

23 J.-L. Hou, M.-X. Jia, X.-K. Jiang, Z.-T. Li, G.-J. Chen, *J. Org. Chem.* **2004**, *69*, 6228–6237.

24 K.N. Trueblood, C.B. Knobler, D.S. Lawrence, R.V. Stevens, *J. Am. Chem. Soc.* **1982**, *104*, 1355–1362.

25 O. Nagano, A. Kobayashi, Y. Sasaki, *Bull. Chem. Soc. Jpn.* **1978**, *51*, 790–793.

26 R.B. Prince, S.A. Barnes, J.S. Moore, *J. Am. Chem. Soc.* **2000**, *122*, 2758–2762.

27 M.T. Stone, J.S. Moore, *Org. Lett.* **2004**, *6*, 469–472.

28 A. Tanatani, M.J. Mio, J.S. Moore, *J. Am. Chem. Soc.* **2001**, *123*, 1792–1793.

29 A. Tanatani, T.S. Hughes, J.S. Moore, *Angew. Chem. Int. Ed.* **2002**, *41*, 325–328.

30 T. Nishinaga, A. Tanatani, K. Oh, J.S. Moore, *J. Am. Chem. Soc.* **2002**, *124*, 5934–5935.

31 J.M. Heemstra, J.S. Moore, *J. Org. Chem.* **2004**, *69*, 9234–9237.

32 K. Goto, J.S. Moore, *Org. Lett.* **2005**, *7*, 1683–1686.

33 R.A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.

34 A.P. Davis, R.S. Wareham, *Angew. Chem. Int. Ed.* **1999**, *38*, 2979–2996.

35 J.-L. Hou, X.-B. Shao, G.-J. Chen, Y.-X. Zhou, X.-K. Jiang, Z.-T. Li, *J. Am. Chem. Soc.* **2004**, *126*, 12386–12394.

36 H.-P. Yi, X.-B. Shao, J.-L. Hou, C. Li, X.-K. Jiang, Z.-T. Li, *New J. Chem.* **2005**, *29*, 1213–1218.

37 H. Abe, N. Masuda, M. Waki, M. Inouye, *J. Am. Chem. Soc.* **2005**, *127*, 16189–16196.

38 M. Inouye, M. Waki, H. Abe, *J. Am. Chem. Soc.* **2004**, *126*, 2022–2027.

39 Z.Q. Wu, X.B. Shao, C. Li, J.L. Hou, K. Wang, X.K. Jiang, Z.T. Li, *J. Am. Chem. Soc.* **2005**, *127*, 17460–17468.

40 J.-L. Hou, H.-P. Yi, X.-B. Shao, C. Li, Z.-Q. Wu, X.-K. Jiang, L.-Z. Wu, C.-H. Tung, Z.-T. Li, *Angew. Chem. Int. Ed.* **2006**, *45*, 796–800.

41 C.Y. Tan, M.R. Pinto, M.E. Kose, I. Ghiviriga, K.S. Schanze, *Adv. Mater.* **2004**, *16*, 1208–1212.

42 X. Zhao, K.S. Schanze, *Langmuir* **2006**, *22*, 4856–4862.

43 K.J. Chang, B.N. Kang, M.H. Lee, K.S. Jeong, *J. Am. Chem. Soc.* **2005**, *127*, 12214–12215.

44 R.B. Prince, T. Okada, J.S. Moore, *Angew. Chem. Int. Ed.* **1999**, *38*, 233–236.

45 M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457–3497.

46 T. Hintermann, D. Seebach, *Chimia* **1997**, *51*, 244–247.

47 A.R. Sanford, B. Gong, *Current Organic Chemistry* **2003**, *7*, 1649–1659.

48 D.J. Hill, M.J. Mio, R.B. Prince, T.S. Hughes, J.S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4011.

49 K. Kirshenbaum, R.N. Zuckermann, K.A. Dill, *Curr. Opin. Struct. Biol.* **1999**, *9*, 530–535.

50 J.T. Ernst, J. Becerril, H.S. Park, H. Yin, A.D. Hamilton, *Angew. Chem. Int. Ed.* **2003**, *42*, 535–539.

51 B. Liu, P.G. Alluri, P. Yu, T. Kodadek, *J. Am. Chem. Soc.* **2005**, *127*, 8254–8255.

52 H. Yin, A.D. Hamilton, *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163.

53 A. Loregian, G. Palu, *J. Cell. Physiol.* **2005**, *204*, 750–762.

54 A.G. Cochran, *Curr. Opin. Chem. Biol.* **2001**, *5*, 654–659.

55 S. Fletcher, A.D. Hamilton, *Curr. Opin. Chem. Biol.* **2005**, *9*, 632–638.

56 M. Werder, H. Hauser, S. Abele, D. Seebach, *Helv. Chim. Acta* **1999**, *82*, 1774–1783.

57 C.M. Rudin, C.B. Thompson, *Ann. Rev. Med.* **1997**, *48*, 267–281.

58 H. Harada, S. Grant, *Rev Clin Exp Hematol.* **2003**, *7*, 117–138.

59 A. Burlacu, *J. Cell. Mol. Med.* **2003**, *7*, 249–257.

60 T. Chittenden, C. Flemington, A.B. Houghton, R.G. Ebb, G.J. Gallo, B. Elangovan, G. Chinnadurai, R.J. Lutz, *EMBO J.* **1995**, *14*, 5589–5596.

61 M. Sattler, H. Liang, D. Nettesheim, R.P. Meadows, J.E. Harlan, M. Eberstadt, H.S. Yoon, S.B. Shuker, B.S. Chang, A.J. Minn, C.B. Thompson, S.W. Fesik, *Science* **1997**, *275*, 983–986.

62 J.D. Sadowsky, M.A. Schmitt, H.S. Lee, N. Umezawa, S.M. Wang, Y. Tomita, S.H. Gellman, *J. Am. Chem. Soc.* **2005**, *127*, 11966–11968.

63 A. Hayen, M.A. Schmitt, F.N. Ngassa, K.A. Thomasson, S.H. Gellman, *Angew. Chem. Int. Ed.* **2004**, *43*, 505–510.

64 J. Momand, H.H. Wu, G. Dasgupta, *Gene* **2000**, *242*, 15–29.

65 D.A. Vargas, S. Takahashi, Z. Ronai, *Mdm2: A regulator of cell growth and death*, in *Adv. Cancer Res.* 2003. p. 1–34.

66 Y. Haupt, R. Maya, A. Kazaz, M. Oren, *Nature* **1997**, *387*, 296–299.

67 J.A. Kritzer, O.M. Stephens, D.A. Guarracino, S.K. Reznik, A. Schepartz, *Bioorg. Med. Chem.* **2005**, *13*, 11–16.

68 J.A. Kritzer, N.W. Luedtke, E.A. Harker, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 14584–14585.

69 P.H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A.J. Levine, N.P. Pavletich, *Science* **1996**, *274*, 948–953.

70 C.W. Wu, S.L. Seurynck, K.Y.C. Lee, A.E. Barron, *Chem. Biol.* **2003**, *10*, 1057–1063.

71 P. Armand, K. Kirshenbaum, R.A. Goldsmith, S. Farr-Jones, A.E. Barron, K.T.V. Truong, K.A. Dill, D.F. Mierke, F.E. Cohen, R.N. Zuckermann, E.K. Bradley, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4309–4314.

72 C.W. Wu, K. Kirshenbaum, T.J. Sanborn, J.A. Patch, K. Huang, K.A. Dill, R.N. Zuckermann, A.E. Barron, *J. Am. Chem. Soc.* **2003**, *125*, 13525–13530.

73 S.L. Seurynck, J.A. Patch, A.E. Barron, *Chem. Biol.* **2005**, *12*, 77–88.

74 K. Gademann, M. Ernst, D. Hoyer, D. Seebach, *Angew. Chem. Int. Ed.* **1999**, *38*, 1223–1226.

75 P. Brazeau, W. Vale, R. Burgus, N. Ling, M. Butcher, J. Rivier, R. Guillemi, *Science* **1973**, *179*, 77–79.

76 G. Melacini, Q. Zhu, G. Osapay, M. Goodman, *J. Med. Chem.* **1997**, *40*, 2252–2258.

77 K. Gademann, T. Kimmerlin, D. Hoyer, D. Seebach, *J. Med. Chem.* **2001**, *44*, 2460–2468.

78 D. Seebach, S. Abele, K. Gademann, B. Jaun, *Angew. Chem. Int. Ed.* **1999**, *38*, 1595–1597.

79 D. Seebach, L. Schaeffer, M. Brenner, D. Hoyer, *Angew. Chem. Int. Ed.* **2003**, *42*, 776–778.

80 J.X. Qiu, E.G. Petersson, E.E. Matthews, A. Schepartz, *J. Am. Chem. Soc.* **2006**, *128*, 11338–11339.

81 L. Addadi, Z. Berkovitchyellin, I. Weissbuch, J. Vanmil, L.J.W. Shimon, M. Lahav, L. Leiserowitz, *Angew. Chem. Int. Ed.* **1985**, *24*, 466–485.

82 I. Weissbuch, L. Addadi, M. Lahav, L. Leiserowitz, *Science* **1991**, *253*, 637–645.

83 S. Mann, *Nature* **1988**, *332*, 119–124.

84 L. Addadi, S. Weiner, *Stereochemical and structural relations between macromolecules and crystals in biomineralization*, ed. S. Mann. **1989**: VCH.

85 S.I. Stupp, P.V. Braun, *Science* **1997**, *277*, 1242–1248.

86 S. Mann, *Angew. Chem. Int. Ed.* **2000**, *39*, 3393–3406.

87 G.A. Ozin, *Acc. Chem. Res.* **1997**, *30*, 17–27.

88 L. Addadi, D. Joester, F. Nudelman, S. Weiner, *Chem. Eur. J.* **2006**, *12*, 981–987.

89 B.R. Heywood, S. Mann, *Adv. Mater.* **1994**, *6*, 9–20.

90 S. Mann, *Biomineralization: Principles and concepts in bioinorganic materials chemistry.* **2001**, Oxford: Oxford University Press.

91 S. Mann, D.D. Archibald, J.M. Didymus, T. Douglas, B.R. Heywood, F.C. Meldrum, N.J. Reeves, *Science* **1993**, *261*, 1286–1292.

92 M. Lahav, L. Addadi, L. Leiserowitz, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4737–4738.

93 T. Douglas, *Science* **2003**, *299*, 1192–1193.

94 L. Addadi, S. Weiner, *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4110–4114.

95 D.B. DeOliveira, R.A. Laursen, *J. Am. Chem. Soc.* **1997**, *119*, 10627–10631.

96 H. Colfen, L.M. Qi, *Chem. Eur. J.* **2001**, *7*, 106–116.

97 J. Song, E. Saiz, C.R. Bertozzi, *J. Am. Chem. Soc.* **2003**, *125*, 1236–1243.

98 S.G. Deng, J.M. Cao, J. Feng, J. Guo, B.Q. Fang, M.B. Zheng, J. Tao, *J. Phys. Chem. B* **2005**, *109*, 11473–11477.

99 R. Tannenbaum, M. Zubris, E.P. Goldberg, S. Reich, N. Dan, *Macromolecules* **2005**, *38*, 4254–4259.

100 L.A. Gower, D.A. Tirrell, *J. Cryst. Growth* **1998**, *191*, 153–160.

101 J.N. Cha, G.D. Stucky, D.E. Morse, T.J. Deming, *Nature* **2000**, *403*, 289–292.

102 J. Aizenberg, A.J. Black, G.H. Whitesides, *J. Am. Chem. Soc.* **1999**, *121*, 4500–4509.

103 J. Aizenberg, D.A. Muller, J.L. Grazul, D.R. Hamann, *Science* **2003**, *299*, 1205–1208.

104 D. Volkmer, M. Fricke, C. Agena, J. Mattay, *J. Mater. Chem.* **2004**, *14*, 2249–2259.

105 A.M. Travaille, J.J.M. Donners, J.W. Gerritsen, N.A.J.M. Sommerdijk, R.J.M. Nolte, H. van Kempen, *Adv. Mater.* **2002**, *14*, 492–495.

106 E.D. Sone, E.R. Zubarev, S.I. Stupp, *Small* **2005**, *1*, 694–697.

107 R.J. Davey, S.N. Black, L.A. Bromley, D. Cottier, B. Dobbs, J.E. Rout, *Nature* **1991**, *353*, 549–550.

108 P.V. Coveney, R. Davey, J.L.W. Griffin, Y. He, J.D. Hamlin, S. Stackhouse, A. Whiting, *J. Am. Chem. Soc.* **2000**, *122*, 11557–11558.

109 S.B. Mukkamala, A.K. Powell, *Chem. Commun.* **2004**, 918–919.

110 K.M. Roth, Y. Zhou, W.J. Yang, D.E. Morse, *J. Am. Chem. Soc.* **2005**, *127*, 325–330.

111 L.A. Estroff, C.D. Incarvito, A.D. Hamilton, *J. Am. Chem. Soc.* **2004**, *126*, 2–3.

112 L.A. Estroff, A.D. Hamilton, *Chem. Mater.* **2001**, *13*, 3227–3235.

113 K. Naka, Y. Chujo, *Chem. Mater.* **2001**, *13*, 3245–3259.

114 H. Colfen, *Curr. Opin. Colloid Interface Sci.* **2003**, *8*, 23–31.

115 S. Weiner, L. Addadi, *Trends Biochem. Sci.* **1991**, *16*, 252–256.

116 N. Ueyama, T. Hosoi, Y. Yamada, M. Doi, T. Okamura, A. Nakamura, *Macromolecules* **1998**, *31*, 7119–7126.

117 N. Ueyama, H. Kozuki, M. Doi, Y. Yamada, K. Takahashi, A. Onoda, T. Okamura, H. Yamamoto, *Macromolecules* **2001**, *34*, 2607–2614.

118 J.J.M. Donners, R.J.M. Nolte, N.A.J.M. Sommerdijk, *J. Am. Chem. Soc.* **2002**, *124*, 9700–9701.

119 J. Cornelissen, W.S. Graswinckel, A.E. Rowan, N. Sommerdijk, R.J.M. Nolte, *J. Poly. Sci. Part A Poly. Chem.* **2003**, *41*, 1725–1736.

120 J.J.M. Donners, *Macromolecular templates for biomimetic crystallizations.* **2002**, Technische Universiteit Eindhoven: Eindhoven.

121 D. Volkmer, M. Fricke, T. Huber, N. Sewald, *Chem. Commun.* **2004**, 1872–1873.

122 H. Bekele, J.H. Fendler, J.W. Kelly, *J. Am. Chem. Soc.* **1999**, *121*, 7266–7267.

8
Biological Applications of Foldamers

Marc Koyack and Richard Cheng

8.1
Introduction

Biologically active compounds have been traditionally dominated by small molecules [1, 2], oligonucleotides [3, 4], and α-peptides [5]. Small molecules are particularly attractive for developing bioactive compounds, because small molecules can be designed to have both the desirable bioactivity and reasonable *in vivo* half lives suitable for therapeutic drugs [1, 2]. However, the construction of small molecule libraries with high diversity is taxing of both time and resources. Furthermore, small molecules may not be sufficient to bind the large surface areas presented by biomacromolecular systems. Although peptides and oligonucleotides may cover larger surface areas compared to small molecules, both suffer the drawbacks of poor bio-availability and susceptibility to *in vivo* degradation. Nonetheless, libraries of α-peptides [6] and oligonucleotides [7–11] are easy to generate and screen. An ideal scaffold for developing bioactive molecules would combine the desirable *in vivo* half lives of small molecules and the ease of library generation of α-peptides and oligonucleotides.

Foldamers are oligomeric compounds that adopt compact globular structures in solution stabilized through noncovalent interactions between nonadjacent monomer subunits (see Chapters 1–5) [12–14]. Foldamers can be divided into two categories: bio-inspired foldamers and abiotic foldamers [13, 15]. Bio-inspired foldamers have backbones that are structurally similar to naturally occurring biomacromolecules, such as proteins and nucleic acids. In contrast, abiotic foldamers possess backbones that do not resemble any natural biomacromolecule. Both types of foldamer have been shown to fold locally into regularly repeating structures analogous to naturally-occurring secondary structures such as turns, helices, and sheets. Furthermore, many of these foldamers are readily prepared through iterative covalent linkage of monomer building blocks, allowing facile introduction of functional groups at multiple sites. While poor bio-availability may limit the potential of bioactive foldamers as therapeutic agents, certain modifications to foldamers can increase cellular uptake and localization [16–28]. Overall,

foldamers are an attractive platform for designing molecules that can potentially recognize large surface areas presented by biomacromolecular systems. Therefore, a number of bio-inspired and abiotic foldamers have been used as scaffolds to introduce bioactivity. Highlights of the structural aspects and inherent biological properties of some of these foldamers are provided in the following sections.

8.1.1
β-Peptides

β-Peptides are composed of *β*-amino acids [12–14, 29–31], which are naturally occurring *α*-amino acids with a single carbon unit inserted into the backbone (Fig. 8.1). *β*-Amino acids can have side chains present at either the C_α (β^2) or C_β (β^3) positions or both (β^2, β^3). *β*-Peptides can adopt numerous protein-like secondary structures, including sheets, turns, and helices (see Chapter 2) [12, 13, 29, 30]. Independently, Gellman [32] and Seebach [33] revealed the sequence-dependent formation of different *β*-peptide helical structures in organic solvents. In particular biological activity has been introduced into three *β*-peptide helical structures: the 14- [34–45], 12- [46–49], and 10/12-helix [50].

The *β*-peptide helix type can be controlled by incorporating different *β*-amino acids. Homochiral β^2-peptides and homochiral β^3-peptides can promote the 14-helix structure, whereas alternating β^2- and β^3- of residues can promote the 10/12-helix [33]. *β*-Amino acids constrained by cyclohexyl rings can promote 14-helices [32, 51] while cyclopentyl *β*-amino acids can promote 12-helices [32]. Stabilization of the 14-helix scaffold in aqueous solution has been accomplished through various design methods, including favorable electrostatic interaction with the helix macrodipole [52], introduction of residues with branched side chains [53, 54], and intra-helical salt bridges [55, 56].

β-Peptides are readily synthesized by solid-phase techniques using standard amide coupling protocols and protecting groups developed for synthesizing *α*-peptides [57, 58]. However, synthesis of *β*-peptides longer than six residues can be difficult because of impaired deprotection and can be addressed through alter-
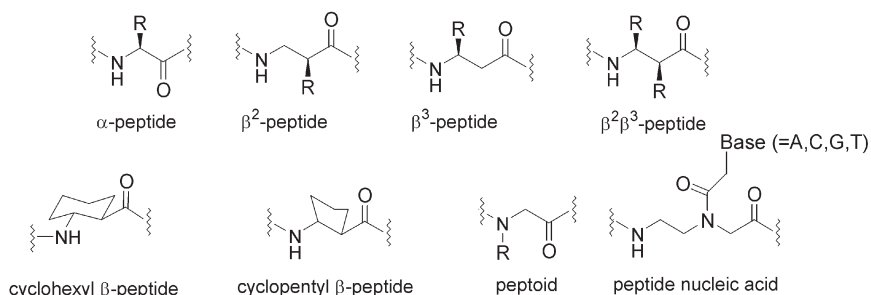
Fig. 8.1 Chemical structures of various residues for α-peptide, β-peptide, peptoid, and peptide nucleic acids (PNA).

native deprotection protocols [59, 60] and increased deprotection temperatures [37]. Nonetheless, *β*-peptide libraries have been synthesized and screened by combinatorial methods [43, 44].

β-Peptides exhibit significant metabolic stability under *in vivo* conditions [61, 62]. *β*-Peptides subjected to various robust peptidases (i.e. proteinase K) remained intact after 48 h, whereas analogous *α*-peptides were completely degraded within 15 min [62]. The pharmacokinetics of *β*-peptides are equally impressive, with elimination half lives of 3–10 h, compared with less than 15 min for analogous *α*-peptides following intravenous injection in rats [62]. *β*-Peptides can also cross cell membranes and localize to the cell nuclei by the appendage of *β*-peptide segments derived from HIV-1 Tat proteins [28]. This permits the design of *β*-peptides that can affect biological targets in the nucleus.

8.1.2
Peptoids

Peptoids are oligomeric peptidomimetics composed of *N*-substituted glycines (see Chapter 1) [12–15]. Compared with natural *α*-peptides, peptoids differ by the appendage of side chains to the amide nitrogen atoms of a peptide backbone, rather than to the *α*-carbons for natural *α*-peptides (Fig. 8.1). Peptoids as short as five residues bearing certain bulky *α*-chiral aromatic or aliphatic side chains can adopt a stable helical structure similar to the type-I polyproline helix, with *cis*-amide bonds, three residues per turn, and 6.0 Å pitch (see Section 1.4.2 in Chapter 1) [12–15, 63–68].

Peptoids are readily synthesized by solid phase techniques [69] and are amenable to combinatorial methods [70, 71]. Unlike other foldamer types, peptoids are prepared by submonomer synthesis [69]. Peptoid synthesis is an iterative two-step process of bromoacetylation of the N-terminal secondary amine followed by S_N2 displacement of the bromide by side chain bearing amines [69]. Synthesis therefore proceeds without the need for producing suitable quantities of a diverse set of protected monomers as in solid phase peptide synthesis. Peptoids exhibit poor oral absorption, as only 3–8% of orally administered peptoids were absorbed in rats [72]. Nevertheless, peptoids have been shown to be stable against multiple proteases *in vitro* [73, 74]. Furthermore, peptoids introduced intravenously into rats were eliminated primarily in the feces without degradation [72], while an *α*-peptide tetramer was completely metabolized in the blood within 2 h.

8.1.3
Peptide Nucleic Acids (PNA)

Peptide nucleic acids (PNA) consist of a pseudopeptide backbone of repeating *N*-(2-aminoethyl)-glycine units linked by peptide bonds, with nucleobase moeities linked to the amide nitrogens of the backbone by methylene carbonyl bonds (Fig. 8.1) [75, 76]. The nucleobases of PNAs can bind those of a complementary

nucleic acid strand, because the number of backbone bonds and the distance between nucleobases are conserved between the two backbone types [77]. In contrast to natural nucleic acids DNA and RNA, the amide backbone of PNAs is non-rigid, achiral, and neutral in charge due to the absence of ribose and phosphate groups. PNA oligomers can bind to complementary strands of DNA [78–80], RNA [80, 81], or PNA [82–84]. The complexes formed are stable duplex and triplex structures that bind through Watson–Crick base pairing (for duplexes) [80, 85] or a combination Watson–Crick–Hoogsteen base-pairing (for triplexes) [77, 79, 86]. Binding occurs through strand invasion, whereby a PNA oligomer displaces a strand of the DNA duplex [78, 87]. Duplex [80–85], triplex [78, 86, 88], and double duplex [89–91] DNA strand invasion have all been described. PNA folding is induced upon hybridization to DNA, RNA, or another PNA oligomer. Experiments with 2-D NOESY and COSY NMR revealed that the DNA in PNA–DNA duplexes adopt a B-form structure with deoxyribose sugars of the DNA strand in the C2′-*endo*-conformation [92]. Crystal structures of PNA–DNA triplexes showed the adoption of a P-form helix, which displays a base-tilt similar to the B-form DNA but with a larger deviation of bases from the helical axis compared to the A-form typically adopted by RNA [86]. Crystal structures of PNA duplexes revealed a helical structure with a wide diameter and large pitch, with nucleobases oriented almost perpendicular to the helical axis [84].

PNAs are readily obtained through solid phase synthesis on polystyrene beads using Boc- or Fmoc-protected monomers and can be synthesized by a combinatorial approach [75–78, 93, 94]. PNA synthesis is an iterative process and employs coupling reagents similarly used for the synthesis of peptides. Synthesis proceeds by deprotection and activation of monomers, followed by amide coupling to the nascent strand and capping. In addition to on-bead synthesis, membranes have also been employed as a support for solid phase PNA synthesis and for the construction of PNA microarrays, although yields have been reported to be lower compared with synthesis on beads [75, 76, 93].

PNAs have a short elimination half life (17 min) and limited bio-availability as revealed by injection of PNAs into rats [75, 95]. While delivery of PNAs into intact cells has generally been limited, the appendage of cationic peptides has increased tissue localization [16–23, 25–27]. Furthermore, increased delivery of PNAs specifically to hepatic cells and rat livers has been achieved by the appendage of GalNAc sugars [22–24].

8.1.4
DNA-Binding Oligoamides

DNA-binding oligoamides are heterocyclic crescent shaped molecules composed of pyrrole and imidazole-based rings [96–98]. Three different heterocyclic rings are commonly used: *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methyl-3-hydroxypyrrole (Hp) (Fig. 8.2). These oligoamides were initially designed based on the natural product Distamycin A, which binds A/T-rich se-
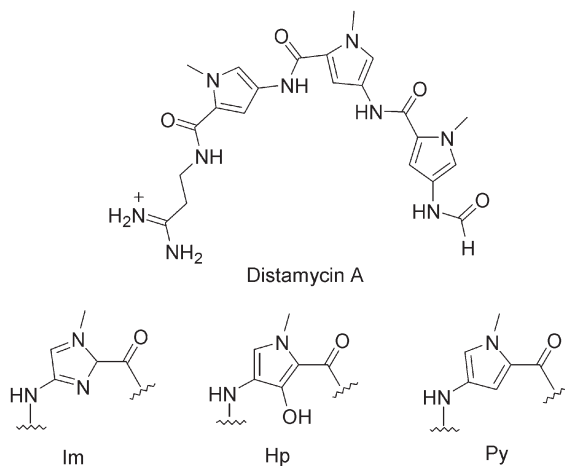
Fig. 8.2 Chemical structures of Distamycin A and pyrole and imidazole based monomers of DNA-binding olioamides (Im = *N*-methylimidazole, Hp = *N*-methyl-3-hydroxypyrrole, Py = *N*-methylpyrrole).

quences in the DNA minor groove as 1:1 or 2:1 Distamycin A-DNA complexes [99]. These oligoamides have been used as site-specific DNA-targeting agents with affinities comparable to DNA-binding proteins [97]. These DNA-binding oligoamides do not adopt specific conformations unless bound to DNA, and therefore are considered foldamers in a liberal sense. Crystal structures revealed that the ring amide groups of these oligoamides recognize the edges of adjacent oligonucleotide strands through a series of hydrogen bonds [100–102]. ''Pairing rules'' for targeting specific sequences of DNA with oligoamide base pairs have been deduced [100–103]. Py/Hp recognizes A-T; Hp/Py recognizes T-A; Py/Im recognizes C-G; Im/Py recognizes G-C.

Structural data of oligoamide-DNA complexes show that the oligoamide rise per residue matches the pitch of the B-form DNA [100–102]. However, the curvature of the oligoamide crescent shape is slightly greater than that of the minor groove of DNA, thus oligoamides longer than 5 units do not completely complement the target DNA, limiting the size of DNA that can be targeted [100]. However, binding sites 10–11 base pairs long have been targeted by oligoamides containing curvature-relaxing moieties such as β-Ala [104] or by multiple oligoamides linked to form hairpin structures [105–108].

Oligoamides can be synthesized through solid phase methods employing Boc- or Fmoc-protected monomers [109–111]. The first residue attached to the resin determines the oligoamide C-terminal residue, which can dictate the DNA-binding specificity. For example, oligoamides with a C-terminal β-Ala specifically target T-A pairs [112] whereas oligoamides without the C-terminal β-Ala target G-C pairs [113].

Localization to nuclei is crucial for DNA targeting with oligoamides. Nuclear localization of oligoamides depends on numerous factors, including the cell type being targeted, oligoamide characteristics (such as charge, shape, sequence), and the identity and placement of dye conjugates [114–117]. Idiosyncratic determinants of nuclear localization have also been observed; the presence of a C-terminal β-Ala residue or the absence of an alkyl amine moiety inhibited nuclear localization, while the presence of an acetylated 2,4-diaminobutyric acid turn promotes nuclear localization [117]. However, nuclear exclusion and localization to other cellular compartments (i.e. lysosomes) often occurs [114, 115].

8.1.5
Aryl Amides and Aryl Ureas

Aryl amides are oligomers that consist of either diaminobenzene and dicarboxylic acid benzene derivatives, or aminobenzoic acid derivatives (Fig. 8.3) [118]. Various functional groups have been introduced to hydrogen bond to the backbone amide to limit the number of low energy conformations, making the aromatic and amide components coplanar. The interactions responsible for structural stabilization are almost exclusively non-cooperative, although cooperative interactions are possible between nonadjacent units (see Chapter 1). Linear [119–121], cyclic [122–124], crescent-shaped [125, 126], and helical [125–129] aryl amides have all been reported. However, bioactivity has only been introduced into the linear extended aryl amides. Aryl amides can be synthesized by Fmoc-based solid phase chemistry but are usually obtained through solution-phase techniques [130]. Carboxylic acids that have been converted to the corresponding acid chlorides are reacted with aromatic amines to produce aryl amide oligomers. Furthermore, a convergent segment doubling strategy can be employed whereby two oligomer segments are efficiently linked in a single step [130]. However, aromatic oligoamides that fold during synthesis can reduce yields because steric hindrance prevents access to the reagents.

Aryl ureas are structurally similar to aryl amides, possessing backbones rigidified by H-bonding groups. Aryl ureas differ from aryl amides by the insertion of an amine moiety between the aromatic rings and carbonyls of the linking amides (Fig. 8.3). Bioactive aryl ureas have been demonstrated to adopt extended conformations. The aryl urea extended conformation is stabilized by H-bonding interac-
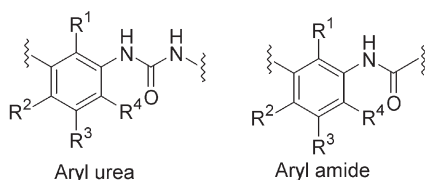

Fig. 8.3 Chemical structures of the monomers for aryl amide and aryl ureas.

tions. The presence of an additional H-bond interaction makes the backbone more rigid compared to aryl amides. Several procedures have been described for the synthesis of urea oligomers, including on solid support via activated carbamate intermediates [131–133] and single-pot reactions using carbonyl diimadazole couplings [134].

8.1.6
meta-Phenylene Ethynylenes (mPE)

meta-Phenylene ethynylenes (*m*PE) are hydrocarbon scaffolds composed of phenyl rings connected by ethynylene linkages at the *meta* position (Fig. 8.4) [13, 135]. *m*PEs can adopt either an extended conformation (*trans*) or a helical conformation (*cis*) (see Chapter 3) [135–138]. The conformation is dependent on solvent [135, 137] and the presence of hydrogen-bonding functionalities [138]. Biological activity has only been introduced into *m*PEs with an extended conformation [139]. *m*PEs composed of alternating cationic primary amines and nonpolar alkyl groups have been designed to adopt an extended all *trans* conformation. *m*PEs are amenable to large scale synthesis and are composed of inexpensive monomers. Various 3-alkoxy-substituted 1,3-diiodobenzenes carrying Boc-protected amines in the 5-position were readily obtained via Mitsunobu etherification, and polycondensation with 1,3-diethynylbenzene under Sonogashira-Hagihara cross-coupling conditions followed by deprotection afforded facially amphiphilic *m*PEs [140, 141].



R= -O-*n*-pentyl, n=20, x=1
-H, n=20, x=2
-H, n=6, x=2
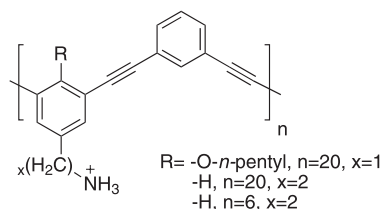
Fig. 8.4 A *meta*-phenylene ethynylene oligomer which adopts an extended (*trans*) conformation and demonstrates antimicrobial activity.

8.1.7
Terphenyls

Terphenyl ring systems are composed of three phenyl rings attached by single carbon–carbon bonds (Fig. 8.5) [142–146]. The rings of the system are oriented to achieve the least strained staggered conformation, resulting in a ''turn'' of the structural axis (see Chapter 7). The twist and pitch between the units are uniform, so substitutions at the *ortho* positions of the rings capture nearly the same
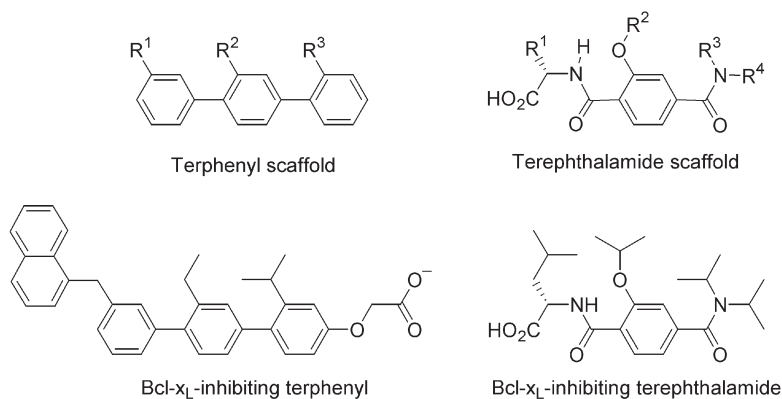
Fig. 8.5 Terphenyl and terephthalamide scaffolds and Bcl-x$_L$ inhibiting derivatives.

spatial arrangement of substituted i, $i+4$ and $i+7$ positions of short α-helical segments [142]. Therefore, the terphenyl scaffold has been utilized to target helix recognizing proteins [142–146]. The terphenyl ring systems have been synthesized using Suzuki [143, 144] or Negishi [142] couplings. Low water solubility of terphenyl foldamers has led to the replacement of the terminal rings with carboxyamides to give terephthalamide derivatives with improved water solubility (see Fig. 8.5) [147, 148].

8.2
Design Strategies

There are four general strategies for developing bioactive foldamers: direct sequence conversion, distribution of physicochemical properties, modular assembly, and grafting bioactive functionality. The biological target and the foldamer employed both dictate the optimal strategy to apply. Foldamers designed using the direct sequence conversion strategy are based on biologically active natural biopolymers such as proteins. The strategy is most likely to be productive if the mere presence of the bioactive functionalities is sufficient to exert bioactivity (which is rare), or if the foldamer scaffold inherently presents the bioactive functionalities with a similar spatial arrangement compared to the natural biopolymer and the backbone is not directly involved in binding. For example, somatostatin receptor binding β-peptides adopt a turn structure similar to the natural α-peptide somatostatin [149–151]. However, biologically active molecules often present functionalities in orientations significantly different from foldamer scaffolds, and therefore more sophisticated design strategies are necessary. Many bioactivities rely on the appropriate distribution of physicochemical properties presented on the surface of bioactive molecules. For instance, the antimicrobial

activity of the α-helical magainin-II peptide relies on amphiphilicity with a positively charged face to interact with the negatively charged outer membranes of bacteria and a hydrophobic face to interact with the interior of the membrane bilayer [152–154]. Therefore, foldamers bearing the same overall distribution of physicochemical properties have been shown to exert antimicrobial activity [34–37, 46–48, 50, 134, 139, 155, 159]. However, when a specific region of a large biopolymer is being targeted (i.e. specific oligonucleotide sequences), relying exclusively on distribution of physicochemical properties might not be sufficient for specific binding. In this case, design strategies that can provide more intricate placement of the bioactive functional groups are necessary. Modular assembly is useful only when appendable monomers that can selectively bind specific components of the targeted specific interaction are available. For instance, peptide nucleic acids that selectively bind a specific oligonucleotide sequence have been constructed by incorporating complementing bases that can pair with the bases on the target oligonucleotide [75, 76]. The most complex strategy is grafting bioactive functionalities onto a scaffold to bind the biological target. This typically requires computational modeling of the designs to evaluate the binding of the foldamer to the target. This strategy is most suitable for targeting protein–protein interactions which involve complicated surface topology with many different functional groups.
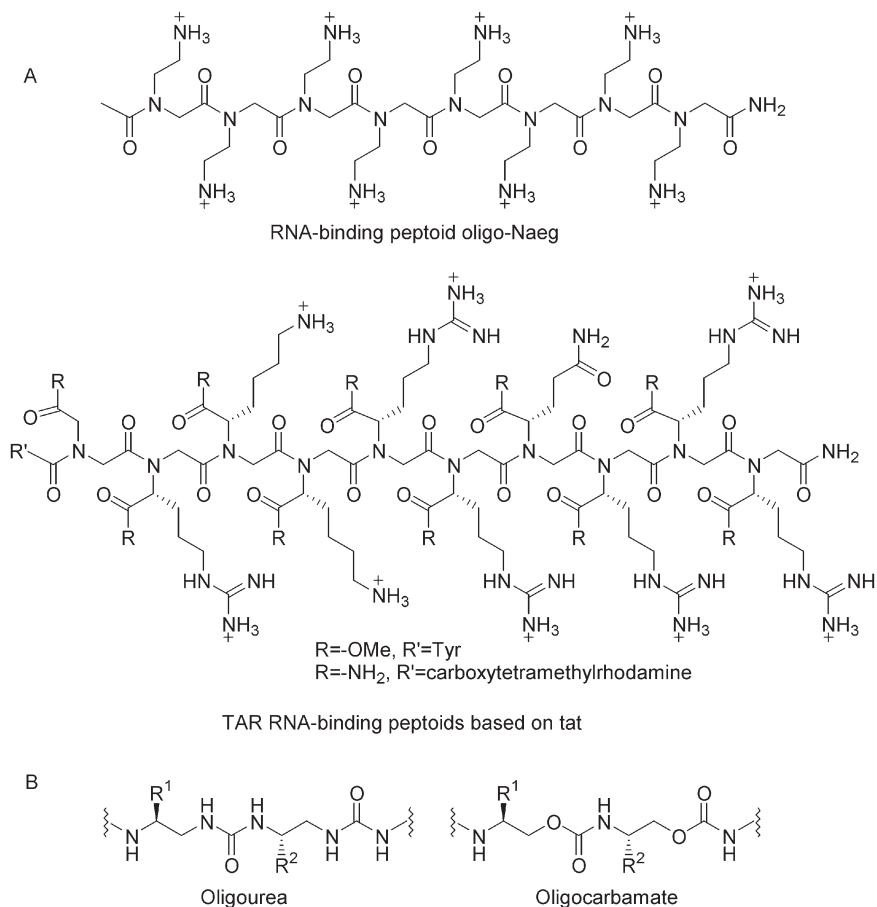
8.2.1
Direct Sequence Conversion

8.2.1.1   RNA-binding Peptoids

Transcription of HIV RNA requires the interaction of the virally encoded Tat protein with the transcriptional activator-responsive (TAR) element [157], a bulged RNA hairpin structure. A single Arg amidst at least three basic residues in the Arg-rich region of the Tat protein is the key determinant for binding the trinucleotide bulge in TAR RNA [158], An α-peptide based on tat with the sequence Ac-Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg-NH$_2$ was found to bind TAR RNA with submicromolar affinity (K_D = 780 nM) based on electrophoretic mobility shift assays [158]. Furthermore, circular dichroism studies with TAR RNA showed a characteristic signal decrease at 265 nm upon addition of Tat peptide, indicating a conformational change of TAR RNA upon peptide binding [159].

Various peptoids have been designed to specifically target TAR RNA (Fig. 8.6) [159, 160]. Peptoids based on the sequence of the Arg rich region of Tat had similar affinities compared with the Tat protein (∼2 μM for peptoid) by gel shift mobility assay [159]. The CD spectrum for wild type TAR RNA exhibited the characteristic change for binding upon adding the peptoid, while no such change was observed for mutant TAR RNA, demonstrating high specificity for targeting TAR RNA. Also, fluorescence resonance energy transfer (FRET) experiments showed a high affinity complex between a peptoid amide analog and TAR RNA with submicromolar affinity (K_D = 155 nM). An ester derivative of the same peptoid demonstrated even higher affinity (K_D = 68 nM) (see Fig. 8.6) [160].

RNA-binding peptoid oligo-Naeg

R=-OMe, R'=Tyr
R=-NH₂, R'=carboxytetramethylrhodamine

TAR RNA-binding peptoids based on tat

Oligourea   Oligocarbamate

Fig. 8.6 TAR RNA binding (A) peptoids, and (B) oligocarbamate and oligourea designs.

8.2.1.2 RNA-binding Oligourea and Carbamate

An oligourea was designed to bind HIV-1 TAR RNA by direct sequence conversion of the Arg-rich region of the Tat protein (Fig. 8.6) [161]. The oligourea bound the TAR RNA with submicromolar affinity ($K_D = 111$ nM), similar to a Tat-derived peptide ($K_D = 780$ nM) as determined by electrophoretic mobility shift assay [161]. Furthermore, mutant TAR RNA could not displace the oligourea from labeled TAR RNA, demonstrating the specificity of the oligourea for binding TAR RNA [161]. Also the oligourea did not bind TAR RNA with mutations immediately surrounding the pyrimidine bulge critical for recognition. Importantly, the oligourea was determined to be protease resistant against proteinase K. Similar results have been obtained for a carbamate derivative (Fig. 8.6) [162]. The oligo-carbamate was found to bind TAR RNA with micromolar affinity ($K_D = 1.13$ µM) as determined by mobility shift assay. Competition experiments on photo

R= $(CH_2)_4NH_3^+$ or $(CH_2)_3NHC(NH)NH_3^+$
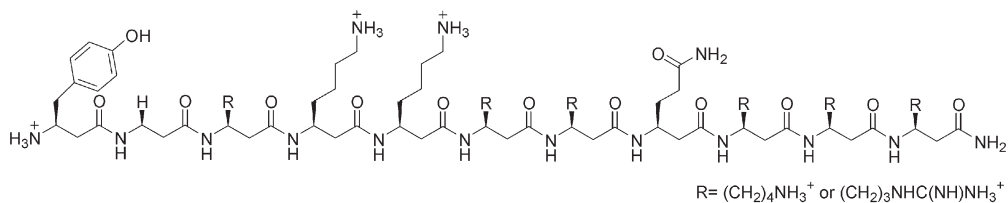
Fig. 8.7 TAR RNA binding β-peptide designs.

crosslinked carbamate-RNA complexes revealed that carbamate binding was specific to the widened major groove of TAR RNA and required the presence of the characteristic trinucleotide bulge [162]. Treatment of the carbamate-RNA complex with proteinase K did not result in a loss of the photo crosslink, indicating much higher proteolytic stability compared with the Tat protein–RNA complex.

8.2.1.3 RNA-binding β-Peptides

TAR RNA-binding β^3-peptides have also been designed as potential HIV therapeutics by direct sequence conversion of a segment of the RNA-binding protein Tat to the β-peptide sequence (Fig. 8.7) [163]. This 11-residue β^3-peptide based on the Arg-rich region of the Tat protein bound TAR RNA with nanomolar affinity ($K_D = 29$ nM) as determined by fluorescence anisotropy [163].

8.2.1.4 Receptor-binding β-Peptides

Somatostatin receptors bind somatostatin, a natural disulfide linked 14-residue cyclic α-peptide, to control the release of various hormones including growth hormone, glucagon, insulin, and gastrin [164]. Various somatostatin receptor subtypes for mediating the different biological activities have been identified [164]. Therefore, developing molecules that selectively target these receptors would be desirable for therapeutic purposes [164]. Somatostatin and most α-peptide somatostatin analogs adopt a β-turn conformation involving the amino acid sequence Phe-Trp-Lys-Thr (Fig. 8.8) [165]. The central Trp-Lys is required for binding somatostatic receptors and thus bioactivity [166, 167]. Octreotide (Sandostatin) is an 8-residue cyclic α-peptide somatostatin analog that may be used for the treatment of acromegaly and certain gastric-entero tumors (Fig. 8.8) [167, 168]. However, octreotide has a relatively short half life of 90 min *in vivo*, making β-peptides with increased bio-availability an attractive alternative.

Design efforts toward somatostatin receptor binding β-peptides have been based on the placement of somatostatin residues onto a β-peptide scaffold that can present the four bioactive side chains in a productive manner (Fig. 8.8). Initially, cyclic β-peptide tetramers were employed to present the side chains of the key residues of somatostatin Phe-Trp-Lys-Thr [149, 150]. Computational modeling showed that the constraints imposed by the cyclic structure matched reasonably well with the type II' β-turn of α-peptides. The cyclic β-peptides bound different types of human somatostatin receptors with micromolar affinities ($K_D = 3.3$–
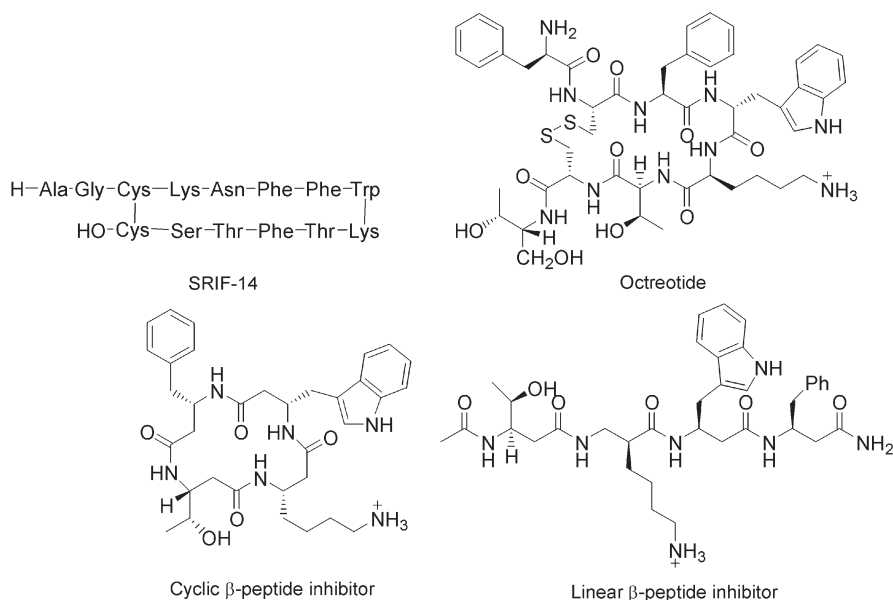
Fig. 8.8 The *β*-turn motif of SRIF-14, the somatostatin receptor inhibitor octreotide, a cyclic *β*-peptide somatostatin analog, and a linear *β*-peptide somatostatin analog.

186 μM). Interestingly, unconstrained linear *β*-peptides, that adopt a turn conformation, bound somatostatin receptors with nanomolar affinities ($K_D = 83$–724 nM) [151, 169], suggesting that the constraints imposed by the cyclic *β*-peptide were less than optimal.

8.2.2
Distribution of Physicochemical Properties

8.2.2.1 Antimicrobial Peptoids
Antimicrobial amphiphilic peptides such as magainin-II and defensins are a widely distributed component of eukaryotic and prokaryotic host immune systems [152–154]. Although magainin-II forms an *α*-helix and defensins form *β*-sheets, both are amphiphilic and display a facial distribution of cationic and hydrophobic groups on the folded structures [152–154]. While the mechanism of action remains to be fully understood, it is generally presumed that the cationic residues direct peptides to the partially negatively charged outer membrane of bacteria, while the hydrophobic residues interact with the lipid bilayer and disrupt membrane integrity, leading to cell death [152–154]. Amphiphilic peptides exert activity by targeting the bacterial phospholipid membranes, therefore there is minimal concern for the development of drug resistance. However, the most common drawback of amphiphilic peptides is nonspecific recognition binding
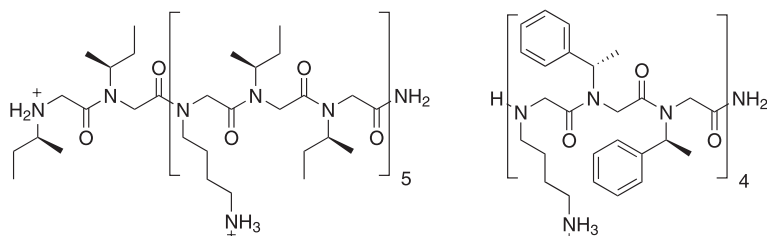
**Fig. 8.9** Helical antimicrobial peptoids.

and lysis of host cells. Since red blood cells are particularly susceptible, hemolysis assays are used to determine the potential for such undesirable cytotoxic effects. The IC_{50} values for magainin-II [152] based α-peptides against *E. coli* have been reported as 1.2 μg mL^{-1}. Varying the helical content of antimicrobial α-peptides affects activity [152, 153]. Increasing helical content by replacing helix-breaking residues with alanine (which is a strong helix former) enhances antimicrobial activity, but at the expense of specificity as hemolysis concomitantly increases [152].

Antimicrobial peptoids have been designed to mimic the distribution of physicochemical properties of the magainin-II amide antibacterial peptide (Fig. 8.9) [170]. The peptoids were designed to be amphiphilic in a helical polyproline type I conformation with a cationic face of lysine-like *N*-(4-aminobutyl)glycine (*N*Lys) side chains and a hydrophobic face of aliphatic or aromatic side chains [170]. Circular dichroism spectroscopy (CD) revealed that some of the peptoids adopted characteristic polyproline type I-like helical structures in aqueous buffer and also in the presence of lipid vesicles [170]. Helical peptoids as short as 12 residues exhibited selective (nonhemolytic) and potent antibacterial activity against both Gram-positive and Gram-negative bacteria with minimal inhibitory concentrations (MICs) in the low micromolar range [170]. This is comparable to previously reported results using a synthetic magainin II analog and other antibacterial α-peptides [152, 153]. Unstructured peptoids were found to be ineffective as antibiotics [170], demonstrating the prerequisite of the helical structure and appropriate distribution of physicochemical properties for antibiotic activity.

Combinatorial libraries of antimicrobial peptoids have also been synthesized and screened [171]. A peptoid library created from a set of multiple statistically unbiased 324 compound library was created. Screening of the library by growth inhibition assays revealed peptoid compounds with antimicrobial activity against a broad range of bacteria, including *E. coli* and *S. aureus*, with MICs in the micromolar range. It was deduced from the screens that antimicrobial activity of peptoids is enhanced (in terms of host range) by the presence of a primary amine and a hydrophobic amine.

8.2.2.2 Antimicrobial β-Peptides

Amphiphilic helix-forming β-peptides with antimicrobial activity comparable to that of a magainin-II analog have been designed based on the 14-helix [34–
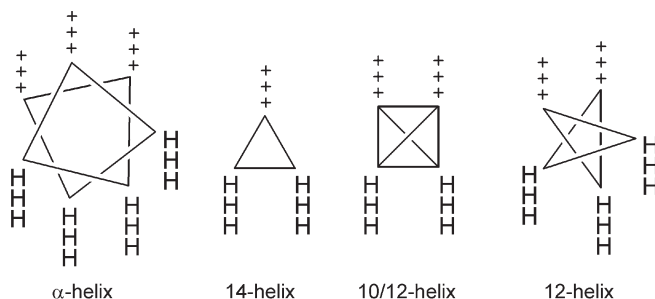
Fig. 8.10 Physicochemical distribution of various helical types for antimicrobial activity (+ = cationic, H = nonpolar).

37], 12-helix [46–48], and 10/12-helix (Fig. 8.10) [50]. 14-Helical designs, consisting of repeating hydrophobic-cationic-hydrophobic triad repeats, display a cationic surface constituting approximately one-third of the helix circumference (Fig. 8.10). Conversely, 12-helical structures consisting of cationic-hydrophobic-cationic-hydrophobic-hydrophobic pentads yield a cationic surface on two-fifths of the helix circumference (Fig. 8.10). Antimicrobial 10/12-helical β-peptides can be designed to have two adjacent cationic faces among the four faces along the helical axis (Fig. 8.10).

It has been shown that varying the helical content of antimicrobial α-peptides affects activity [152, 153]. However, increasing the helical content of 14-helical antimicrobial β-peptides by varying the proportions of rigid *trans*-2-aminocyclohexanecarboxylic acid (ACHC) residues had little effect on antimicrobial activity [37]. Increasing helical propensity of the constituent residues did not alter the MIC values against four bacterial species [37], suggesting that there is no relationship between 14-helical stability and antimicrobial potency. However, the identity of C-terminal groups and the ability of the β-peptide to form an amphiphilic helix were crucial for antimicrobial activity [37]. A series of 9-mer β-peptides were shown to have nearly the same efficacy as (Ala8,13,18)-magainin II amide and melittin to permeabilize the membrane of *B. subtilis* BAU102 [35], as measured by x-gal release. Furthermore, decreasing hydrophobic character by replacing β^3-hLeu with β^3-hAla side chains increases the selectivity for targeting bacterial cells over red blood cells [35].

12-Helical designs have also demonstrated antimicrobial activity comparable to magainin-II peptides against four species of Gram-positive and Gram-negative bacteria [46, 47]. Designs consisting of hydrophobic ACPC and hydrophilic APC residues yielded facially amphiphilic 12-helices that were structurally rigid compared with β^3-peptides because they are composed entirely of cyclic constrained residues [46]. However, 12-helical β-peptides containing noncyclic residues have also demonstrated similar antibacterial activity with comparable specificity to magainin-II peptides [47].

β-Peptides composed of alternating β^2- and β^3-amino acids have been shown to fold into 10/12-helices. A 10/12-helix forming β-peptide composed of hydrophobic and aromatic residues with two β^3-HLys residues presented on one face of the helix was highly active against two species of Gram-positive bacteria and also displayed activity against Gram–negative bacteria [50]. Importantly, these 10/12-helical antimicrobial β-peptides presented low hemolytic activity against human and rat blood cells [50].

8.2.2.3 Antimicrobial Aryl Amides and Aryl Ureas

Aryl amides are smaller than the previously discussed antimicrobial foldamers, and therefore may be advantageous in terms of production costs and bioavailability. Presently, the biological activity of aryl amides is exclusive to those adopting an extended conformation. Aryl amides composed of di-acid and di-amine monomers possess an extended network of hydrogen bonds, which serve to rigidify the backbone and stabilize the extended conformation (Fig. 8.11).

Facially amphiphilic aryl amides have displayed antimicrobial activity comparable to antimicrobial cyclic α-peptides and magainin II derivatives [155, 156]. Chain length studies showed that short aryl amides were the most effective at inhibiting bacterial cell growth, with an optimal length of 8 repeat units [155]. The 8-mer demonstrated bactericidal activity against several species of Gram-positive and Gram-negative bacteria, with MICs below 50 μg ml^{-1} for each strain. Longer oligomers are suggested to be less active due to reduced solubility, lower molar concentration, or the inability to penetrate the proteoglycan layer [155]. The presence of positively charged aminoethyl groups were also found to be requisite for activity, as acetylation of the most highly active 8-mer resulted in loss of antimicrobial activity. Antimicrobial activity was attributed to disruption of phospholipids bilayers, as determined by the ability of aryl amides designs to induce leakage of calcein from unilamellar vesicles composed of mixed phosphatidylserine and phosphatidylcholine lipids [155]. However, these aryl amides were demonstrated to be hemolytic near the MIC. In addition, increasing the hydrophobicity of designed aryl amides has also been demonstrated to increase antimicrobial activity [156]. Aryl amides with increased hydrophobic character display high potency
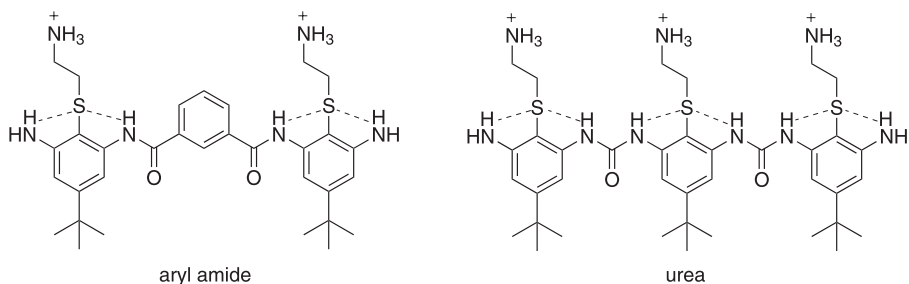

Fig. 8.11 Antimicrobial aryl amide and aryl urea.

against both Gram-negative and Gram-positive bacteria, with MICs of 6–12 μg mL^{-1} against both *E. coli* and *S. aureus* [156]. However, increases in hydrophobicity were found to be directly proportional to hemolytic activity. Introduction of more polar substituents yielded aryl amides that were significantly less toxic towards erythrocytes [156]. One of these compounds displayed potency similar to a magainin II analog and had significantly greater selectivity. Another analog with similar potency was nonhemolytic at concentrations as high as 800 μM [156].

Compared with aryl amides, antimicrobial aryl ureas exhibit greater structural rigidity [134]. In contrast to aryl amides, aryl urea homo-oligomers are composed of only one type of monomer. Importantly, the number of hydrogen bonds per monomer unit is higher than that for aryl amides, therefore aryl urea backbones can be relatively more rigid (Fig. 8.11). The increased rigidity is due to the inherent presence of internal NH-S bonding. Amphiphilic aryl ureas prepared by one-pot synthesis demonstrated potent antimicrobial activity against *E. coli* and *B. subtilis*. A trimer displayed the most potency and specificity, with an MIC = 0.7 mg mL^{-1}.

8.2.2.4 Antimicrobial meta-Phenylene Ethynylenes

meta-Phenylene ethynylenes (*m*PEs) with alternating polar/nonpolar functionalities induce leakage of calcein from large phospholipid vesicles [140, 172], most likely due to the amphiphilic nature of the *m*PE in the extended conformation (Fig. 8.4). The hydrophobicity of *m*PEs is tunable through the appendage of hydrophobic or hydrophilic functionalities onto the hydrocarbon scaffold. This tunability modulates the affinity and selectivity of *m*PEs for bacterial phosphopilid membranes. A 20-mer *m*PE with ethylamine functionalities appended to alternating backbone benzenes exhibited antibacterial activity against both Gram-negative and Gram-positive bacteria, with an MIC = 25 μg mL^{-1} for *E. coli* [139]. A *m*PE composed of six repeat units was less potent (MIC = 50 μg mL^{-1}) but was 20 times more selective than the 20-mer in hemolysis assays, which is comparable to a highly active magainin II analog.

8.2.2.5 DNA-binding Peptoids

DNA has a highly-negatively charged ribophosphate backbone that can be readily targeted by positively-charged compounds. The selective binding of a foldamer to DNA can inhibit the transcription of the targeted gene. To enhance the transcription of certain genes, the foldamer would need to bind the targeted DNA and recruit the appropriate transcription machinery. Also, foldamers that can bind DNA and transport DNA across the cell membrane into the cell (also known as transfection) may serve as gene delivery systems for gene therapy.

Peptoids have been shown to bind plasmid DNA [71], and thus have potential as gene delivery systems. The distribution of basic side chains on peptoid scaffolds can thus serve to bind DNA, while hydrophobic residues render amphiphilicity and facilitate cell entry. A series of peptoids, with varying lengths, frequency of cationic side chains, hydrophobicity, and side-chain shape, were developed to bind DNA bearing the firefly luciferase gene [71]. Many of these peptoids

condensed DNA and protected DNA from nuclease degradation, although only a single type of repeating triplet motif (cationic-hydrophobic-hydrophobic) was able to cross the cell membrane and localize to the nucleus. Direct binding of the peptoid to the target DNA was shown by gel mobility shift assays [71]. A 36-mer peptoid showed the greatest degree of transfection in multiple cell lines in the presence of fetal calf serum [71]. Electron microscopy revealed that the 36-mer peptoid formed highly regular spherical structures (50–100 nm in diameter) when condensed with DNA [71]. Importantly, the efficiency of peptoid-mediated transfection is similar to the standard lipofection method with cationic lipids in serum-free medium.

8.2.2.6 DNA-binding β-Peptides

A 14-helical β-peptide was designed to bind single stranded and duplex DNA [39]. One side of the 14-helix displayed three β^3-hAsn residues to potentially form H-bonds to the DNA bases. Positively charged β^3-hLys residues were incorporated at the termini to potentially form ionic interactions with the negatively charged backbone of DNA and the introduction of β^3-hAla and β^3-hPhe residues rendered the amphiphilic character to the helix and facilitated nuclear entry (Fig. 8.12). Circular dichroism and DNA-melting temperature measurements using UV-Vis revealed a structured interaction occurring between the β-peptide and DNA. CD spectra were characteristic of the β-peptide 14-helix, while signal changes were apparent upon addition of DNA, demonstrating interactions between the β-peptide and DNA. In the presence of β-peptide, melted DNA was unable to reform back to a duplex, suggesting that β-peptide binding interferes with the rewinding process.

8.2.2.7 Cholesterol Uptake-inhibiting β-Peptides

Cholesterol and dietary lipid uptake in the lumen of the small intestine is facilitated by integral proteins in the brush-border membrane (BBM), the so-called scavenger receptors of class B (SRB) type I or II proteins [173]. Uptake occurs
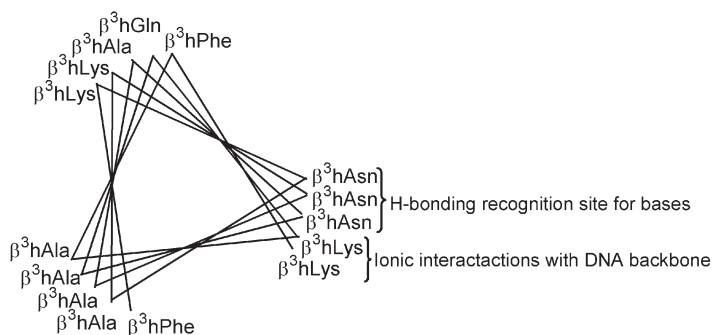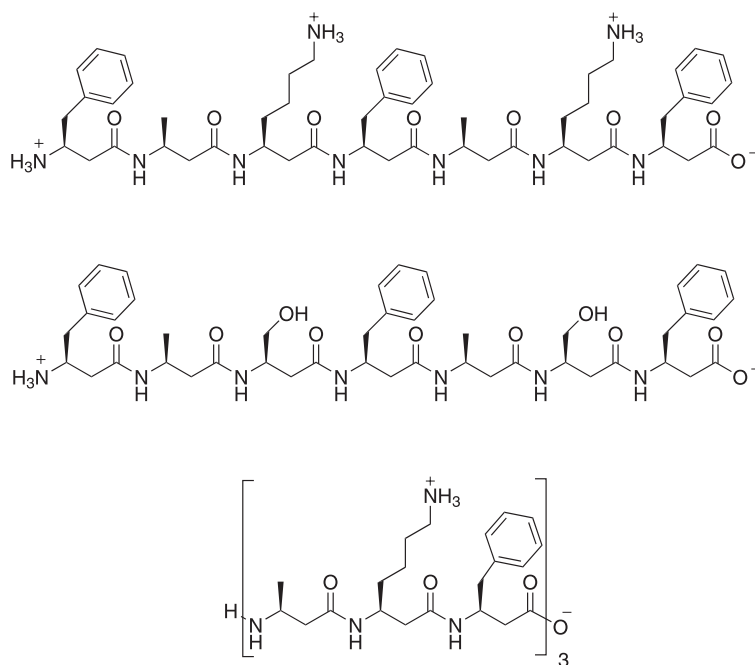

Fig. 8.12 DNA-binding β^3-peptide.

Fig. 8.13 14-Helical *β*-peptide inhibitors of cholesterol absorption.

through the binding of carrier particles such as small unilamellar phospholipid vesicles and mixed bile-salt micelles [173]. In liver and steroidogenic tissues, the SRBI receptor protein binds high density lipoproteins (HDL) which is an important pathway of cholesterol homeostasis and metabolism [174]. More specifically, the SRBI receptor protein binds an amphiphilic *α*-helix motif on HDL [175]. Consequently, amphiphilic helices that bind the SRBI receptor protein have been shown to inhibit cholesterol and lipid uptake at the BBM [173, 175].

Short *β*-peptides (hexamers, heptamers, and nonamers) were designed to mimic the amphiphilic helices of HDL, bind the SRBI receptor protein, and inhibit cholesterol absorption. Amphiphilicity of the *β*-peptides was rendered by the presentation of functionalized β^3-hLys or β^3-hSer side chains along one helical face, and hydrophobic residues (β^3-hVal, β^3-hPhe, and β^3-hAla) on the other helical face two model assays were used to test the bioactivity (Fig. 8.13) [38]. One assay followed the transport of lipids and cholesterol or cholesteroyl esters from unilamellar vesicles into BBM vesicles. The second assay tested uptake into a monolayer of whole CaCO-2 cells. Three of six *β*-peptide designs were capable of inhibiting cholesterol uptake into BBM vesicles, with IC$_{50}$ values as low as 590 µg mL^{-1}, compared with Human apolipoprotein A-I. Interestingly, among the *β*-peptides tested, only those that could form 14-helices in MeOH demonstrated an
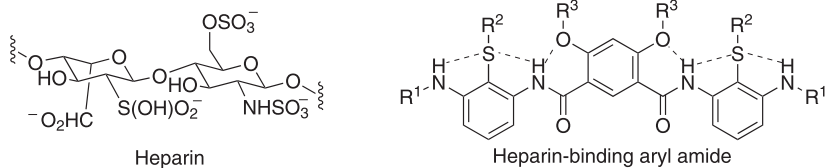
Fig. 8.14 Major repeat of heparin and a heparin binding aryl amide.
Backbone rigidifying H-bonds are shown as dashed lines.

inhibitory effect on cholesterol uptake in either model. On the other hand, while
none of the previously tested α-peptide inhibitors induced effects in whole CaCO-
2 cells, a β-peptide nonamer decreased the absorption rate of radiolabeled choles-
terol to background levels (indicated by an increase in $T_{1/2}$ of cholesterol absorp-
tion rate from 102 min to 2.6 h). This suggested that the β-peptide nonamer was
not susceptible to proteolytic degradation, unlike its α-peptide counterparts.

8.2.2.8 Heparin-inhibiting Aryl Amides

Heparin is a linear, highly sulfated polysaccharide composed of repeating L-
iduronic acid and D-glucosamine units [176]. Heparin demonstrates anticoagu-
lant activity and has therefore been utilized therapeutically as a treatment for
thrombotic diseases [176, 177]. Due to the presence of multiple sulfate function-
alities, heparin carries an overall negative charge, and heparin is therefore recog-
nized by cationic peptides (such as antithrombin) bearing either consensus se-
quence XBBXBX or XBBBXXBX, where B is a basic amino acid residue and X is
any other amino acid residue [178]. Furthermore, it has been suggested that the
basic residues of heparin-binding peptides are localized on one face of either an
α-helix or β-sheet. While several potent heparin-binding peptides have been re-
ported [179–182], these still suffer the drawbacks of proteolytic instability and
limited bioavailability.

Oligomeric 1,3-disubstituted aryl amides have been designed to strongly inter-
act with heparin [183]. Aryl amides were designed to display appropriately spaced
cationic groups to bind the negative charges on heparin (Fig. 8.14). Molecular
dynamic simulations showed that binding was mostly attributed to ionic inter-
actions and that increasing positive charge increased heparin binding [183].
Chromagenic assay studies revealed the ability of aryl amides to compete with an-
tithrombin to bind full-length heparin and low molecular weight heparin analogs
($IC_{50} = 22.5$ μM). Schild plot analysis revealed the designed aryl amide could in-
hibit heparin-antithrombin complex formation with dissociation constants in the
micromolar range (1.8–6.7 μM). Furthermore, the aryl amide designs could in-
hibit heparin induced clotting in activated partial thromboplastin time clotting as-
says with comparable efficacy to the clinically used heparin neutralization agent
protamine. Importantly, the aryl amides were determined to be nonhemolytic at
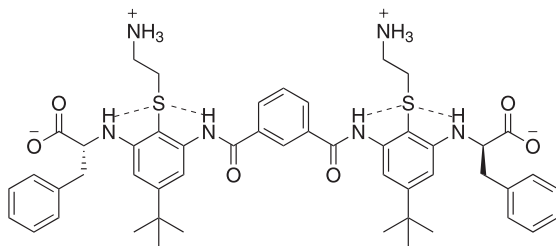concentrations as high as 1 mM.

Fig. 8.15 Aryl amide inhibitor of calmodulin.

8.2.2.9 Calmodulin-inhibiting Aryl Amides

Calmodulin is an attractive therapeutic target due to its implications in multiple biochemical pathways such as metabolism, apoptosis, inflammatory responses, muscle contraction, intracellular movement, short-term and long-term memory, nerve growth and the immune response [184, 185]. Three residues at the i, $i + 3$, and $i + 7$ positions of the smMLCK peptide are critical for CaM-binding [186]. Recapitulation of similar functional groups onto the appropriate scaffold may therefore yield potent and highly selective CaM-binding peptidomimetics.

Molecular modeling studies revealed that the arrangement of *tert*-butyl groups on a thioether substituted aryl amide scaffold closely matched the arrangement of hydrophobic residues presented by the helical CaM-binding peptide smMLCK (Fig. 8.15) [187]. Modeling also showed that D-phenylalanine residues presented on both ends of the aryl amide were adequately accommodated by the hydrophobic binding pockets of CaM globular domains [187]. Inhibitory concentrations (K_i) of the aryl amide inhibitors were determined through fluorescence polarization competition assays by titrating aryl amides against 1:1 complexes of CaM and the high-affinity CaM-binding peptide mastoparin X (MaX). Addition of inhibitors resulted in a shift of the maximum fluorescence emission from 327 to 341 nm, indicating release of MaX from CaM. Aryl amides demonstrated inhibitory concentration in the nanomolar range with the most potent inhibitor having a $K_i = 7.10$ nM [187]. Experiments with 2-D(^{1}H^{15}N)-HSQC NMR also suggested that the aryl amides bound CaM similarly to CaM-binding peptides smMLCK and MARCKS, demonstrating that the designed aryl amides were suitable peptidomimetics though with limited solubility [187].

8.2.3
### Modular Assembly

8.2.3.1 DNA-binding Oligoamides

DNA-binding oligoamides have been designed using the pairing rules discovered through studies on Distamycin A [97, 98]: A-T base pairs are recognized by Py/Hp; T-A base pairs by Hp/Py; C-G base pairs are recognized by Py/Im; and G-C base pairs by Im/Py [97, 98, 100–103]. Single stranded oligoamides bind in an extended fashion in the DNA minor groove [96, 188]. Covalently linking the oli-

goamide strands into a hairpin increases the affinity and specificity of these oligomers [103, 189–195]. Hairpin compounds typically retain the orientation preferences of extended polyamides, aligning N to C with the 5′ to 3′ direction of the closest DNA strand. The classic example of DNA-binding oligoamides is the eight-ring hairpin, in which a γ-aminobutryic acid serves as the linker (Fig. 8.16) [189, 191, 192]. Covalent linkage of complementary oligoamides connecting the carboxyl terminus of one oligoamide to the amino terminus of another to create a hairpin can result in an increase of DNA-binding affinities by 100-fold [189]. Eight-ring hairpins have been demonstrated to bind specific DNA sequences (6bp) with affinities similar to that of DNA-binding proteins ($K_D < 1$ nM) [191]. Undesirable binding events of hairpin oligoamides, such as binding in the opposite intended orientation to target DNA or binding in a 1:1 rather than the intended 2:1 binding mode have been remedied by the introduction of an amino substituent at the α-position of the γ-aminobutryic acid turn [193] or β-Ala [195].

H-pin [111, 194] oligoamides have also been designed, whereby an alkyl chain projecting from the minor groove is used to link oligoamide strands at the central position (Fig. 8.16). The introduction of β-Ala into covalent linkers introduces flexibility to the structure and can relax the curvature of oligoamides, yielding molecules designed after homo-dimeric complexes to bind 11 base pair sequences of DNA with subnanomolar affinities [104]. Tandem hairpin oligoamides [106, 108], in which two hairpins are covalently linked by a 5-aminovaleric acid linker, bound specifically to 10 base pair sites.

Cyclic oligoamides [196, 197], in which the C and N termini of a hairpin have been covalently linked, eliminated the possibility of extended binding (Fig. 8.16). Such oligoamides have demonstrated higher affinities but lower specificities for target DNA sequences compared with analogous hairpin molecules with the same number of cationic groups.

Oligoamides can bind a multitude of specific DNA sites with affinities comparable to DNA-binding proteins [98]. The binding of oligoamides to promoter sites on DNA have inhibited gene transcription by disrupting RNA polymerase activity. In particular, the suppression of 5S RNA transcription RNA polymerase III has been achieved both *in vitro* and in cultured *Xenopus* kidney cells by minor-groove binding oligoamides [198]. Transcription of HIV-1 has also been inhibited through the binding of designed hairpin oligoamides to multiple transcription factor binding sites within the HIV-1 enhancer/promoter site as revealed by cell free assays [199]. The same oligoamides were able to halt viral replication by more than 99% in human blood lymphocytes by directly effecting viral transcription without greatly affecting cell viability. Oligoamides have been designed to target GC rich regions flanking CRE sites to inhibit Tax protein binding and Tax transactivation *in vitro* [200]. Other oligoamides targeting GC rich sequences bound to a 4 base pair site typically cleaved by bacterial gyrase, thus preventing strand cleavage at nanomolar concentrations [201]. The NF-κB heterodimer, a transcription factor which binds in the DNA major groove, has been inhibited by oligoamides that bind the minor groove opposite one of the monomers [202]. Oligoamides conjugated with the tripeptide sequence Arg-Pro-Arg
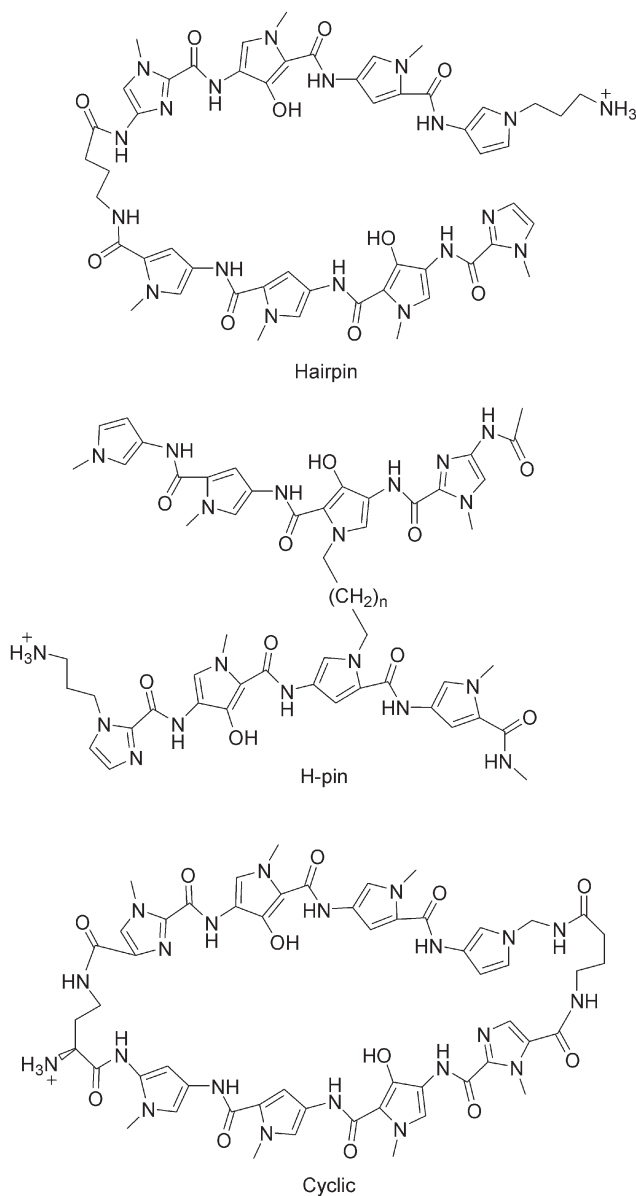
Fig. 8.16 Various types of DNA-binding oligoamides.

can interfere with major groove-binding DNA proteins by distorting DNA through charge neutralization, occupying the major groove to cause steric interference, or binding the backbone phosphate. Such oligoamides were able to inhibit the major groove binding transcription factor GCN4 [203]. Gene expression

has also been affected by oligoamides in *Drosophila*; oligoamides introduced into food sources resulted in the noticeable gain and loss of certain phenotypes, demonstrating the effectiveness of oligoamides in complex organisms [204].

Oligoamides have also been demonstrated to upregulate transcription, either through the inhibition of repressor proteins or recruitment of transcriptional machinery. A hairpin oligoamide was able to block binding of the repressor IE86, upregulating the transcription of human cytomegalovirus MIEP [205]. Upregulation of the HIV-1 promoter has also been achieved by oligoamide binding to the repressor complex sequence in the HIV-1 long terminal repeat, effectively inhibiting the human protein LSF, a protein involved in HIV-1 latency [206]. Oligoamides have also been created which act as artificial transcription factors; a hairpin oligoamide was linked to a 20-residue peptide activation domain by a 36 atom straight-chain linker, resulting in activation of transcription in cell free assays [207]. Appendage of an even smaller yet more potent peptide activation domain was also achieved, resulting in even higher levels of transcription [208].

Hairpin oligoamides have been used to target specific DNA sequences in nucleosomal DNA complexes [209]. Accessibility to certain DNA sequences of the nucleosome is dependent on the positioning and structural implications of the DNA strand wound about 8 histone proteins. Results show that DNA base pairs fully or sometimes partially facing away from the histone octamer are fully accessible [209]. Oligoamides were shown to bind the nucleosome positioning sequence of the sea urchin 5S gene with $K_D \sim 1$ nM, blocking heat-inducible nucleosomal translocation and transcription by the T7 RNA polymerase [209].

8.2.3.2 Nucleotide-binding Peptide Nucleic Acids

Peptide nucleic acids have been primarily used to affect gene transcription, either through enhancement or suppression. In general, transcription enhancement mechanisms are more complex since activity not only relies on DNA binding but also the recruitment of various agents needed by the transcriptional machinery (i.e. transcription factors). The H-bonding pattern between PNA and DNA base pairs is directly consequential to the binding mode. For instance, PNA can pair with single stranded DNA to form hetero-duplexes which are bound together by Watson–Crick base-pairing, while two PNAs can bind a single stranded DNA to form a triple helix via a combination of Watson–Crick and Hoogsteen base pairing. One PNA strand binds DNA via Watson–Crick pairing while the other PNA binds the PNA–DNA heteroduplex via Hoogsteen pairing.

PNAs have been shown to inhibit transcription. A 15-mer homopyrimidine PNA targeted to the IL2-Rα NF-κB binding site inhibited transcription factor binding through strand invasion [210]. Transactivation was inhibited *in vitro* when the PNA was pre-incubated with the target DNA under low salt concentration prior to addition to nuclear extracts. Although direct addition of the PNA to HeLa cells did not inhibit NF-κB mediated transactivation, inhibitory effects were evident up to 24 hours after introduction of preincubated PNA-reporter plasmid complexes. Another strategy to block transcription has been achieved through prevention of transcribed strand elongation. Binding of a

homo-thymidine 10-mer PNA or a mixed sequence 15-mer to the template strand of a G-free transcription cassette have been demonstrated to block site-specific pol II transcription elongation *in vitro* [211]. Binding of the 10-mer to RNA also terminated reverse transcription and *in vitro* translation at the exact site of PNA binding independent of RNase H activity. Furthermore, microinjection of the 15-mer into the nucleus of cells expressing SV 40 T antigen inhibited T antigen expression. A homo-thymidine 10-mer has also been used to target the terminator elements of three yeast class III genes, which are vital to RNA polymerase III (pol III) recycling and thus transcription efficiency [212]. *In vitro* transcription assays revealed nanomolar inhibition of the genes in supercoiled (but not linear) plasmid constructs [212]. Furthermore, PNA concentrations that inhibited multiple rounds of transcription had no effect on the absolute amount of RNA output for a single transcription cycle, suggesting the inhibition of pol III recycling caused by a strand invasion induced "roadblock" to the terminator. Unmodified PNAs are capable of crossing the blood–brain barrier in rats upon intraperitoneal injection [213]. Following intraperitoneal injection, PNAs that bind to the mRNA and DNA of rat neurotensin receptor (NTR1) were demonstrated to inhibit gene transcription and exert behavioral effects in rat specimens [213].

Peptide-linked PNAs have also been designed with antitranscriptional activities. A 17-mer PNA with an appended nuclear localization signal (NLS) peptide successfully localized to the nuclei of Burkitt's lymphomas cells and caused rapid down regulation (35% decrease in 7 h) of c-*myc* oncoprotein expression [21]. Peptide-linked PNAs have also been used to inhibit the expression of human caveolin 1 in both cultured HeLa and primary endothelial cells [26]. A 9-mer PNA conjugated to a nitrogen mustard suppressed HER-2/neu oncogene expression by 80% in intact HeLa cells [27]. Appendage of an alkylating agent presumably facilitates strand invasion and ultimately stabilizes triple helix formation. A trifunctional PNA-peptide-diethylenetriamine conjugate showed sequence-specific RNA cleavage *in vitro* and has potential as an artificial cell-penetrating ribonuclease [25]. A PNA prodrug complementary to the hepatic human microsomal triglyceride transfer protein (huMTP) is rapidly internalized by HepG2 cells due to the appendage of GalNAc sugars [22]. Also, the PNA conjugates accumulate in the parenchymal liver cells of mice to a far greater extent than nonconjugated PNAs following intravenous injection. The MTP mRNA levels in HepG2 cells was consequently down regulated by 35–40% at 100 nM.

PNAs have been demonstrated to act as artificial transcription promoters *in vitro* [214, 215] and *in vivo* [215, 217]. RNA polymerases and transcription factors initiate transcription through sequence-specific recognition of 12 base pair loop structures at the promoter sites of partially melted double stranded DNA (dsDNA). Similar loop structures evolve upon binding of PNAs to dsDNA, as the resulting (PNA)$_2$-DNA complexes display stable D-loop structures on the template strand resembling those of transcription initiation sites [78, 87]. Loops formed by strand invasion of homopyrimidine PNAs are recognized by RNA polymerase which can initiate transcription at PNA-binding sites with efficiencies comparable to that of the robust *E. coli lac UV5* promoter [214]. However, the use

of PNAs as gene promoters *in vivo* has been limited due to impaired cell delivery. Induction of the γ-globin gene, a therapeutic target for sickle cell anemia, has been achieved both *in vivo* and *in vitro* according to studies with reporter gene constructs [215]. Induction of the endogenous gene was achieved in K562 human erythroleukemia cells after introduction into cells by electroporation. Also, PNA length-dependence studies demonstrated that PNAs 14–20 nucleotides long can induce high levels of transcription in a HeLa nuclear extract *in vitro* transcription system [216]. Furthermore, transfection of these same PNAs bound to GFP reporter gene plasmids into human normal fibroblast (NF) cells could induce GFP translation *in vivo*.

The RNA component of the ribonucleoprotein enzyme complex human telomerase is accessible to incoming nucleic acids, and has thus been recognize as a suitable anticancer target. The RNA component of telomerase has previously been targeted by complementary DNA [217], phosphorothioate DNA [218], and 2′-O-methyl RNAs [219], however the inhibitory efficiency of these agents was limited by poor sequence selectivity and bioavailability. Therefore, peptide nucleic acids designed to complement the RNA component of telomerase may provide an alternative with more desirable pharmacokinetics.

Dose-dependent reduction of telomerase activity in cell extracts, tumors, and permeabilized cells has been demonstrated by PNAs [220]. Assays with HME50-5 cell extracts revealed several PNAs of 11 with nanomolar IC_{50} values (0.9–10 nM). The most potent inhibitor also showed an IC_{50} of 50 nM towards permeabilized cells. In comparison to control phosphorothioate DNA oligomers, PNAs demonstrated 10–50-fold increase in binding affinity and increased sequence specificity. The appendage of cationic peptides to 11-mer and 13-mer PNAs were shown to increase cell permeabilization in pretreated melanoma cells, with IC_{50} values in the submicromolar range (360 nM) [221].

8.2.4
Grafting Bioactive Functionalities onto Scaffolds

8.2.4.1 Protein–protein Interaction-inhibiting β-Peptides
Protein p53 is a transcriptional activator critical for stress-induced cell cycle arrest and apoptosis [222]. The overexpression of hDM2 in cancer cells results in unchecked cell growth [223], and is therefore an important target for anticancer therapy. The protein hDM2 inhibits p53 transcriptional activity by binding the p53 transactivation domain [224], exporting p53 from the nucleus [225], resulting in ubiquitination of p53 and subsequent degradation [226, 227]. The binding interface between p53 and hDM2 is well characterized. Three hydrophobic residues of the α-helical activation domain of p53 (F19, W23, L26) are key for hDM2 binding [228, 229]. α-Peptide inhibitors of hDM2 based on the α-helical activation domain of p53 have been reported to induce apoptosis in tumor cells overexpressing hDM2 *in vivo* [230].

The protein-protein interaction hDM2-p53 has been targeted by β-peptides [40–44]. Recapitulation of the three key hydrophobic side chains one turn apart
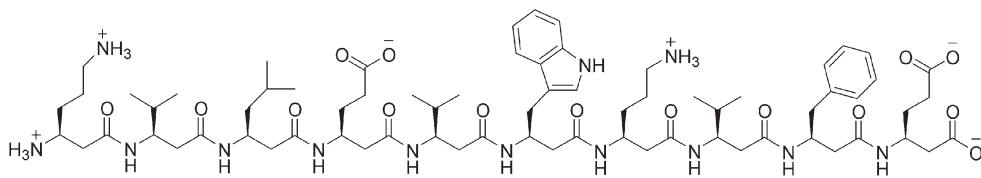
Fig. 8.17 An hDM2-inhibiting β-peptide β53-1.

on a β-peptide 14-helix scaffold as shown to have good overlap with the p53 activation domain bioactive side chains as determined by computational methods (Fig. 8.17) [40]. The designed β-peptide, β53-1, adopted a 14-helix structure as determined by CD and NMR [40]. Fluorescence polarization studies showed that β53-1 bound directly to hDM2 with submicromolar affinity (K_D = 368–583 nM), only 1.6–2.5-fold lower in affinity than an α-peptide analog of the p53 activation domain. β53-1 also competed with a p53-derived peptide for the hDM2 binding site, displacing the peptide with an IC$_{50}$ of 94.5 $\pm$ 4.4 μM [40]. β53-1 variants which displayed altered orientations of β^3-hPhe, β^3-hTrp, and β^3-hLeu side chains demonstrated the importance of these residues for binding affinity and specificity [40].

Combinatorial libraries of β-peptides have also been created to screen antagonists of the p53-hDM2 interaction [43, 44]. An efficient synthesis of a 14-helical β-peptide library on polystyrene macrobeads has been achieved using microwave irradiation, yielding an inhibitor containing cyclically-constrained residues with an IC$_{50}$ of 250 μM [44]. Another library that was created using modified peptide synthesis solid phase protocols resulted in a β-peptide with an inhibitory potency nearly 8-fold greater than β53-1 [43].

14-Helical inhibitors have also been designed to target the N-terminal region of GP-41, a protein implicit to HIV viral fusion to host cell membranes [45]. Three residues comprising the WWI domain of the GP41 C terminus (Trp628, Trp631, and Ile635) were presented on short 14-helical scaffolds consisting of β^3-amino acids (Fig. 8.18). Fluorescence polarization studies demonstrated four inhibitor designs that bound to a GP-41 based model peptide system with micromolar affinity (0.75–2.4 μM), which is comparable to the highest affinity α-peptide of similar size (1.2 μM). The inhibition was also specific for targeting GP41 over carbonic anhydrase II and calmodulin, two proteins that interact with hydrophobic helices. Inhibitor designs were shown to inhibit cell fusion *in vivo*, with EC$_{50}$ values in the micromolar region (5.3–27 μM).

Human cytomegalvirus (HCMV) viral fusion and cell entry is another therapeutic target [231–233] that has been inhibited by β-peptides [49]. HCMV requires interaction between the helical components of two proteoglycans, gB and gH, which presumably bind one another through a coiled-coil [234]. 12-Helical β-peptides were designed to display a series of nonpolar side chains with a spacing similar to coiled coils to mimic the gB heptad and bind gH. Cell-based assays
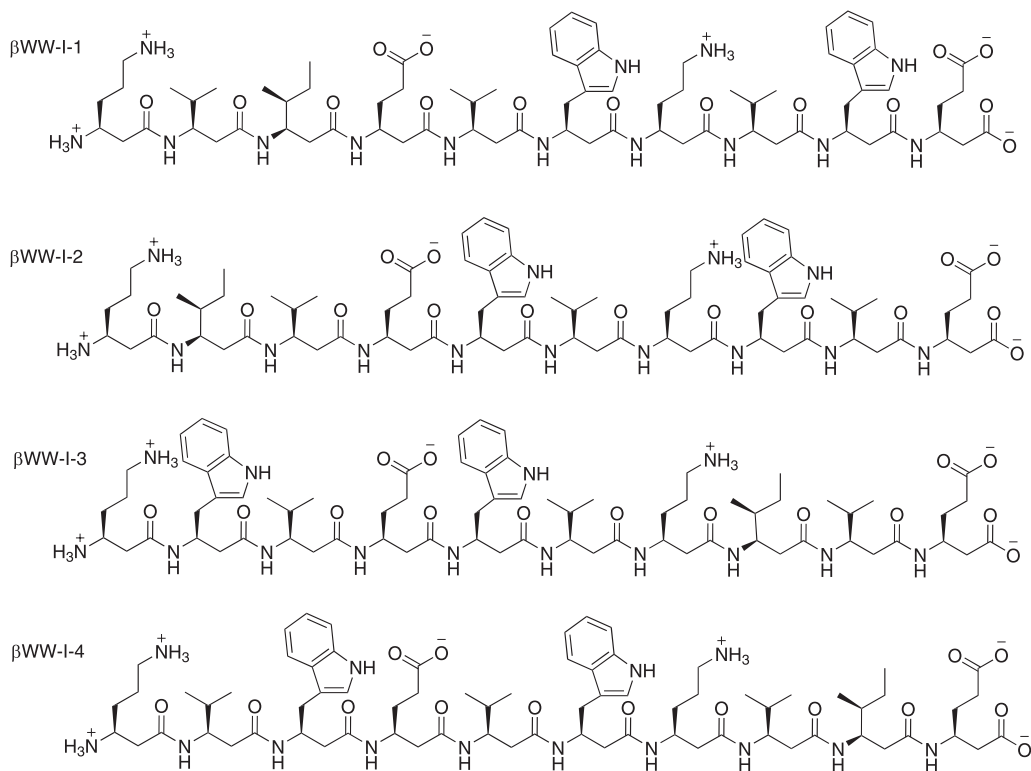
Fig. 8.18 β-Peptide inhibitors of GP-41.

showed that the β-peptide induced inhibition of HCMV with greater potency than the α-peptide model ($IC_{50} = 30 \mu M$) [49].

8.2.4.2 Protein–protein Interaction-inhibiting Peptoids

Using structure-based design, several peptoids were designed to target hDM2 [235]. Three peptoids designed to form a right-handed type-I polyproline helical conformation were found to weakly bind the N-terminal domain fragment of hDM2 (Fig. 8.19) [235]. Synthetic modifications to the side chains to enhance water solubility and improve hydrophobic and nonhydrophobic contacts resulted in a peptoid decamer with improved binding affinity that competed against a peptide analog of the N-terminal region of p53. The IC_{50} for the peptoid was only approximately 2-fold higher than the p53 peptide (6.6 μM versus 3 μM). However, hydrophobic peptoid side chains that mimicked the Phe, Trp, and Leu of p53 may not be limited to binding in the nonpolar cleft of hDM2 based on computer generated models [235]. Interestingly, achiral peptoid designs were found to bind hDM2 with higher affinity than helical peptoid designs, indicating that the helical sense may not be important for binding hDM2.
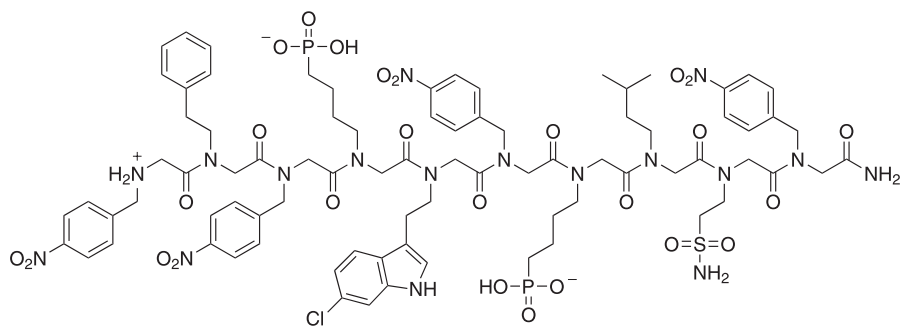
Fig. 8.19 An hDM2-binding achiral peptoid containing a chloroindole residue.

8.2.4.3 Terphenyl Helix Mimetics

The binding epitopes of α-helices are typically found along one face of the helical axis, involving functional side chains at i, $i + 3$, and $i + 7$ positions. The spatial arrangement of functionalities on two α-helical turns can be closely recapitulated onto tris-functionalized $3,2',2''$-terphenyl scaffolds in a staggered conformation [142–146]. Terphenyls displaying hydrophobic functionalities have been designed to mimic the α-helical domain of smooth muscle myosin light chain kinase (smMLCK), a calmodulin (CaM) binding peptide [142]. Calmodulin is a ubiquitous and highly-conserved protein that mediates intracellular Ca^{2+} levels, affecting the activity of regulatory proteins such as kinases and phosphatases and thereby regulating a diverse array of biochemical processes. An enzymatic assay revealed that the terphenyl compound inhibited the phosphodiesterase activity with an IC_{50} of 80 nM; this is only a 10-fold decrease in affinity compared with the 20-residue smMLCK control peptide.

Proteins belonging to the B-cell lymphoma −2 (Bcl-2) family are important determinants of apoptosis, or programmed cell death [236]. Overexpression of the anti-apoptotic Bcl-x_L protein has been demonstrated to decrease the efficacy of anticancer drugs [237]. Bcl-x_L activity is mediated by binding of pro-apoptotic proteins such as Bad- and Bak, which occurs through interactions between α-helices of Bad (or Bak) and α-helices on Bcl-x_L [238]. Helical mimetics of the Bad or Bak would inhibit Bcl-x_L, and may serve as anticancer therapeutics. Initial designs using the terphenyl scaffold displayed aryl or alkyl substituents at the *ortho* positions in addition to carboxylate moieties at either end to electrostatically interact with the positive charge on Bcl-x_L (Fig. 8.5). Fluorescence polarization studies showed that the terphenyl compound bound within the hydrophobic cleft of Bcl-x_L with nanomolar affinity ($K_D = 114$ nM) [143]. Similar derivatives showed increase selectivity for disrupting Bcl-xL-Bak over hDM2-p53 interaction. Treatment of whole human embryonic kidney 293 (HEK293) cells with the terphenyl derivatives demonstrated the inhibition of Bcl-xL binding to Bak *in vivo* [144].

Potent terphenyl-based activators of the pro-apoptotic protein p53 have also been described [145, 146]. Helix mimetics were able to inhibit mDM2-p53 binding in an ELISA assay with IC_{50} values in the micromolar region (10–20 µM). The

terphenyl designs also induced p53 accumulation and activation in cultured HCT116 cells with IC$_{50}$ values in the micromolar region (15–40 µM).

There is an inherent limitation to the solubility of terphenyls due to the hydrophobic nature [147]. To improve the water solubility, the two flanking phenyl groups of terphenyls have been replaced with two functionalized carboxyamide groups to give terephthalamide derivatives (Fig. 8.5) [147]. Terephthalamide analogs designed to inhibit the Bcl-xL/Bak interaction exhibited high *in vitro* activity [147, 148]. Fluorescence polarization studies revealed an analog with submicromolar activity (K_i = 780 nM) [147, 148]. *In vivo* inhibition by terephthalamides was also demonstrated against intact HEK293 cells [148].

8.3
Outlook and Future Directions

The ideal bioactive foldamer needs to be target-specific with high bioavailability, and synthetically readily accessible. Early research efforts on foldamers have mainly focused on synthesizing novel backbones and discovering their three-dimensional structural determinants. Building on these pioneering studies, foldamers are now being developed for applications in material science and pharmaceutics. Early investigations in organic solvents have been important first steps in understanding the structural characteristics of many foldamers. Studies of foldamers in aqueous environments have paved the way for the introduction of bioactivity and for the future development of foldamer pharmaceutics. Accordingly, numerous examples of bioactive foldamers involving several strategies have been discussed in this chapter. Most of these bioactive foldamers have been patterned after known bioactive peptides or oligonucleotides. The foldamer scaffolds have provided more desirable bio-availability over the parent biopolymer for most cases, however the bio-availability of bioactive foldamers remains far from ideal compared to small molecule drugs. Nevertheless, higher bioactivity and specificity for the foldamer has been observed in some cases compared with the parent biopolymer. It seems that the next strategic challenge for developing bioactive foldamers will be the *de novo* discovery of bioactivity either through structure-based design or combinatorial methods.

References

1 G. W. Bemis, M. A. Murko, *J. Med. Chem.* **1996**, *39*, 2887–2893.

2 G. W. Bemis, M. A. Murko, *J. Med. Chem.* **1999**, *42*, 5095–5099.

3 J. H. P. Chan, S. H. Lim, W. S. F. Wong, *Clin. Exp. Pharmacol. Physiol.* **2006**, *33*, 533–540.

4 B. J. Dolnick, *Cancer Invest.* **1991**, *9*, 185–194.

5 W. Williams, D. Weiner, *Biologically Active Peptides: Design, Synthesis, and Utilization, Vol 1*, Technomic, Lancaster, PA, **1993**.

6 C. Falciani, L. Lozzi, A. Pini, L. Bracci, *Chem. Biol.* **2005**, *12*, 417–426.

7 B. Eaton, *Curr. Biol.* **1997**, *1*, 10–16.

8 T. Pan, *Curr. Opin. Chem. Biol.* **1997**, *1*, 17–25.

9 R. R. Breaker, *Curr. Opin. Chem. Biol.* **1997**, *1*, 26–31.

10 M. Famulok, A. Jenne, *Curr. Opin. Chem. Biol.* **1998**, *2*, 320–327.

11 S. Jhaveri, M. Rajendran, A. D. Ellington, *Nat. Biotechnol.* **2000**, *18*, 1293–1297.

12 S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.

13 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4011.

14 R. P. Cheng, *Curr. Opin. Struct. Biol.* **2004**, *14*, 1–9.

15 A. E. Barron, R. N. Zuckermann, *Curr. Opin. Chem. Biol.* **1999**, *3*, 681–687.

16 M. Pooga, U. Soomets, M. Hallbrink, A. Valkna, K. Saar, K. Rezaei, U. Kahl, J. X. Hao, X. J. Xu, Z. Wiesenfeld-Hallin, T. Hokfelt, A. Bartfai, U. Langel, *Nat. Biotechnol.* **1998**, *16*, 857–861.

17 G. Aldrian-Herrada, M. G. Desarmenien, H. Orcel, L. Boissin-Agasse, J. Mery, J. Brugidou, A. Rabie, *Nucleic Acids Res.* **1998**, *26*, 4910–4916.

18 L. J. Branden, A. J. Mohamed, C. I. E. Smith, *Nat. Biotechnol.* **1999**, *17*, 784–787.

19 S. Scarfi, M. Giovine, A. Gasparini, G. Damonte, E. Millo, M. Pozzolini, U. Benatti, *FEBS Lett.* **1999**, *451*, 264–268.

20 T. Ljungstrom, H. Knudsen, P. E. Nielsen, *Bioconjug. Chem.* **1999**, *10*, 965–972.

21 G. Cutrona, E. M.Carpaneto, M. Ulivi, S. Roncella, O. Landt, M. Ferrarini, L. C. Boffa, *Nat. Biotechnol.* **2000**, *18*, 300–303.

22 E. A. L. Biessen, K. Sliedregt-Bol, P. A. C. 'T Hoen, P. Prince, E. Van der Bilt, A. R. P. M. Valentijn, N. J. Meeuwenoord, H. Princen, M. K. Bijsterbosch, G. A. Van der Marel, J. H. van Boom, T. J. C. Van Berkel, *Bioconjug. Chem.* **2002**, *13*, 295–302.

23 P. Sazani, F. Gemignani, S. H. Kang, M. A. Maier, M. Manoharan, M. Persmark, D. Bortner, R. Kole, *Nat. Biotechnol.* **2002**, *20*, 1228–1233.

24 R. Hamzavi, F. Dolle, B. Tavitian, O. Dahl, P. Nielsen, *Bioconjug. Chem.* **2003**, *14*, 941–954.

25 L. Petersen, M. C. de Koning, P. van Kuik-Romeijn, J. Weterings, C. J. Pol, G. Platenburg, M. Overhand, G. A. van der Marel, J. H. van Boom, *Bioconjug. Chem.* **2004**, *15*, 576–582.

26 K. Kaihatsu, K. E. Huffman, D. R. Corey, *Biochemistry*, **2004**, *43*, 14340–14347.

27 Z. V. Zhilina, A. J. Ziemba, P. E. Nielsen, S. W. Ebbinghaus, *Bioconjug. Chem.* **2006**, *17*, 214–222.

28 N. Umezawa, M. A. Gelman, M. C. Haigis, R. T. Raines, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 368–369.

29 D. Seebach, J. L. Matthews, *Chem. Commun.* **1997**, 2015–2022.

30 R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, *101*, 3219–3232.

31 D. L. Steer, R. A. Lew, P. Perlmutter, A. I. Smith, M. Aguilar, *Curr. Med. Chem.* **2002**, *9*, 811–822.

32 D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. L. Huang, J. J. Barchi, S. H. Gellman, *Nature*, **1997**, *387*, 381–384.

33 D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1998**, *81*, 932–981.

34 Y. Hamuro, J. P. Schneider, W. F. DeGrado, *J. Am. Chem. Soc.* **1999**, *121*, 12200–12201.

35 D. H. Liu, W. F. DeGrado, *J. Am. Chem. Soc.* **2001**, *123*, 7553–7559.

36 P. I. Arvidsson, J. Frackenpohl, N. S. Ryder, B. Liechty, F. Petersen, H. Zimmermann, G. P. Camenisch, R. Woessner, D. Seebach, *ChemBiochem* **2001**, *2*, 771–773.

37 T. L. Raguse, E. A. Porter, B. Wiesblum, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 12774–12785.

38 M. Werder, H. Hauser, S. Abele, D. Seebach, *Helv. Chim. Acta* **1999**, *82*, 1774–1783.

39 T. Kimmerlin, D. Seebach, *Helv. Chim. Acta* **2003**, *86*, 2104–2109.

40 J. A. Kritzer, J. D. Lear, M. E. Hodsdon, A. Schepartz, *J. Am. Chem. Soc.* **2004**, *126*, 9468–9469.

41 J. A. Kritzer, O. M. Stephens, D. A. Guarracino, S. K. Reznik, A. Schepartz, *Bioorg. Med. Chem.* **2005**, *13*, 11–16.

42 J. A. Kritzer, M. E. Hodsdon, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 4118–4119.

43 J. A. Kritzer, N.W. Luedtke, E. A. Harker, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 14584–14585.

44 J. K. Murray, B. Farooqi, J. D. Sadowsky, M. Scalf, W. A. Freund, L. M. Smith, J. D. Chen, S. H. Gellman, *J. Am. Chem. Soc.* **2005**, *127*, 13271–13280.

45 O. M. Stephens, S. Kim, B. D. Welch, M. E. Hodsdon, M. S. Kay, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 13126–13127.

46 E. A. Porter, X. Wang, H. S. Lee, B. Weisblum, S. H. Gellman, *Nature* **2000**, *404*, 565.

47 P. R. LePlae, J. D. Fisk, E. A. Porter, B. Weisblum, S. H. Gellman, *J. Am. Chem. Soc. 2002*, *124*, 6820–6821.

48 E. A. Porter, B. Weisblum, S. H. Gellman, *J. Am. Chem. Soc.* **2002**, *124*, 7324–7330.

49 E. P. English, R. S. Chumanov, S. H. Gellman, T. Compton, *J. Biol. Chem.* **2006**, *281*, 2661–2667.

50 P. I. Arvidsson, N. S. Ryder, H. M. Weiss, G. Gross, O. Kretz, R. Woessner, D. Seebach, *ChemBiochem* **2003**, *4*, 1345–1347.

51 D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071–13072.

52 S. A. Hart, A. B. F. Bahadoor, E. E. Matthews, X. Y. J. Qui, A. Schepartz, *J. Am Chem Soc.* **2003**, *125*, 4022–4023.

53 T. L. Raguse, J. R. Lai, S. H. Gellman, *Helv. Chim. Acta* **2002**, *85*, 4154–4164.

54 J. A. Kritzer, J. Tirado-Rives, S. A. Hart, J. D. Lear, W. L. Jorgensen, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 167–178.

55 P. I. Arvidsson, M. Reuping, D. Seebach, *Chem. Commun.* **2001**, 649–650.

56 R. P. Cheng, W. F. DeGrado, *J. Am. Chem. Soc.* **2001**, *123*, 5162–5163.

57 M. Bodanzky (ed.), *Principles of Peptide Synthesis*, Springer-Verlag, Berlin-Heidelberg, **1993**.

58 W.C. Chan, P. D. White (eds.), *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press Inc., New York, **2000**.

59 G. Guichard, S. Abele, D. Seebach, *Helv. Chim Acta* **1998**, *81*, 187–206.

60 J. V. Schreiber, D. Seebach, *Helv. Chim. Acta* **2000**, *83*, 3139–3151.

61 T. Hintermann, D. Seebach, *Chimia* **1997**, *51*, 244–247.

62 D. Seebach, S. Abele, J. V. Schreiber, B. Martinoni, A. K. Nussbaum, H. Schild, H. Schulz, H. Hennecke, R. Woessner, F. Bitsch, *Chimia*, **1998**, *52*, 734–739.

63 P. Armand, K. Kirshenbaum, A. Falicov, R. L. Dunbrack, K. A. Dill, R. N. Zuckermann, F. E. Cohen, *Fold. Des.* **1997**, *2*, 369–375.

64 K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen, R. N. Zuckermann, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4303–4308.

65 P. Armand, K. Kirshenbaum, R. A. Goldsmith, S. Farr-Jones, A. E. Barron, K.T. V. Truong, K. A. Dill, D. F. Mierke, F. E. Cohen, R. N. Zuckermann, E. K. Bradley, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4309–4314.

66 C. W. Wu, T. J. Sanborn, R. N. Zuckermann, A. E. Barron, *J. Am. Chem. Soc.* **2001**, *123*, 2958–2963.

67 T. J. Sanborn, C. W. Wu, R. N. Zuckermann, A. E. Barron, *Biopolymers* **2002**, *63*, 12–20.

68 C. W. Wu, K. Kirshenbaum, T. L. Sanborn, J. A. Patch, K. Huang, K. A. Dill, R. N. Zuckerman, A. E. Barron, *J. Am. Chem. Soc.* **2003**, *125*, 13525–13530.

69 R. N. Zuckermann, J. M. Kerr, S. B. H. Kent, W. H. Moos, *J. Am. Chem. Soc.* **1992**, *114*, 10646–10647.

70 G. M. Figliozzi, R. Goldsmith, S. C. Ng, S. C. Banville, R. N. Zuckermann, *Meth. Enzymol.* **1996**, *267*, 437–447.

71 J. E. Murphy, T. Uno, J. D. Hamer, F. E. Cohen, V. Dwarki, R. N. Zuckermann, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1517–1522.

72 Y. Wang, H. Lin, R. Tullman, C. F. Jewell Jr., M. L. Weetall, F. L. S. Tse, *Biopharm. Drug Dispos.* **1999**, *20*, 69–75.

73 S. M. Miller, R. J. Simon, S. Ng, R. N. Zuckermann, J. M. Kerr, W. H. Moos, *Drug Dev. Res.* **1995**, *35*, 20–32.

74 S. M. Miller, R. J. Simon, S. Ng, R. N. Zuckermann, J. M. Kerr, W. H. Moos, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2657–2662.

75 P. E. Nielsen (ed.), *Peptide Nucleic Acids: Methods and Protocols*, Humana Press, Totowa, NJ, **2002**.

76 P. E. Nielsen (ed.), *Peptide Nucleic Acids: Protocols and Applications*, Horizon Bioscience, Wymondham, U.K., **2004**.

77 M. Egholm, O. Buchardt, P. E. Nielsen, R. H. Berg, *J. Am. Chem. Soc.* **1992**, *114*, 1895–1897.

78 P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1500.

79 M. Egholm, P. E. Nielsen, O. Buchardt, R. H. Berg, *J. Am. Chem. Soc.* **1992**, *114*, 9677–9678.

80 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566–568.

81 L. Good, P. E. Nielsen, *Nature Biotechnol.* **1998**, *16*, 355–358.

82 P. Wittung, P. E. Nielsen, O. Buchardt, M. Egholm, B. Norden, *Nature* **1994**, *368*, 561–563.

83 P. Wittung, M. Eriksson, R. Lyng, P. E. Nielsen, B. Norden, *J. Am. Chem. Soc.* **1995**, *117*, 10167–10173.

84 H. Rasmussen, J. S. Kastrup, J. M. Nielsen, P. E. Nielsen, *Nat. Struct. Biol.* **1997**, *4*, 98–101.

85 P. E. Nielsen, L Christensen, *J. Am. Chem. Soc.* **1996**, *118*, 2287–2288.

86 L. Betts, J. A. Josey, J. M. Veal, S. R. Jordan, *Science* **1995**, *270*, 1838–1841.

87 D. Y. Cherny, B. P. Belotserkovskii, M. D. Frank-Kamentetskii, M. Egholm, O. Buchardt, R. H. Berg, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1667–1670.

88 P. E. Nielsen, M. Egholm, O. Buchardt, *J. Mol. Recognit.* **1994**, *7*, 165–170.

89 J. Lohse, O. Dahl, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11804–11808.

90 V. V. Demidov, E. Protozanova, K. I. Izvolsky, C. Price, P.E. Nielsen, M.D. Frank-Kamenetskii, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5953–5958.

91 E. Protozanova, V. V. Demidov, V. Soldatenkov, S. Chasovskikh, M.D. Frank-Kamenetskii, *EMBO Reports* **2002**, *3*, 956–961.

92 M. Leijon, A. Graslund, P. E. Nielsen, O. Buchardt, B. Norden, S. M. Kristensen, M. Eriksson, *Biochemistry* **1994**, *33*, 9820–9825.

93 J. Weiler, H. Gausepohl, N. Hauser, O. N. Jensen, J. D. Hoheisel, *Nucleic Acids Res.* **1997**, *25*, 2792–2799.

94 S. K. Awasthi, P. E. Nielsen, *Combinatorial Chem. High Throughput Screening* **2002**, *5*, 253–259.

95 B. M. McMahon, D. Mays, J. Lipsky, J. A. Stewart, A. Fauq, E. Richelson, *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 65–70.

96 W. S. Wade, M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794.

97 J. M. Gottesfeld, J. M. Turner, P. B. Dervan, *Gene Expression* **2000**, *9*, 77–91.

98 P. B. Dervan, B. S. Edelson, *Curr. Opin. Struct. Biol.* **2003**, *13*, 284–299.

99 J. G. Pelton, D. E. Wemmer, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723–5727.

100 C. L. Kielkopf, E. E. Baird, P. B. Dervan, D. C. Rees, *Nat. Struct. Biol.* **1998**, *5*, 104–109.

101 C. L. Kielkopf, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, D. C. Rees, *Science* **1998**, *282*, 111–115.

102 C. L. Kielkopf, R. E. Bremer, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, D. C. Rees, *J. Mol. Biol.* **2000**, *295*, 557–567.

103 S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature*, **1998**, *391*, 468–471.

104 S. E. Swalley, E. E. Baird, P. B. Dervan, *Chem. Eur. J.* **1997**, *3*, 1600–1607.

105 J. W. Trauger, E. E. Baird, P. B. Dervan, *Angew. Chem. Intl. Ed. Engl.* **1998**, *37*, 1421–1423.

106 D. M. Herman, E. E. Baird, P. B. Dervan, *Chem. Eur. J.* **1999**, *5*, 975–983.

107 I. Kers, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 3339–3349.

108 P. Weyermann, P. B. Dervan, *J. Am. Chem. Soc.* **2002**, *124*, 6872–6878.

109 E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6141–6146.

110 N. R. Wurtz, J. M. Turner, E. E. Baird, P. B. Dervan, *Org. Lett.* **2001**, *3*, 1201–1203.

111 B. Olenyuk, C. Jitianu, P. B. Dervan, *J. Am. Chem. Soc.* **2003**, *125*, 4741–4751.

112 S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 1113–1120.

113 J. M. Belitsky, D. H. Nguyen, N. R. Wurtz, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 2767–2774.

114 J. M. Belitsky, S. J. Leslie, P. S. Arora, T. A. Beerman, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 3313–3318.

115 K. S. Crowley, D. P. Phillion, S. S. Woodward, B. A. Schweitzer, M. Singh, H. Shabany, B. Burnette, P. Hippenmeyer, M. Heitmeier, J. K. Bashkin, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565–1570.

116 T. P. Best, B. S. Edelson, N. G. Nickols, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12063–12068.

117 B. S. Edelson, T. P. Best, B. Olenyuk, N. G. Nickols, R. M. Doss, S. Foister, A. Heckel, P. B. Dervan, *Nucleic Acids Res.* **2004**, *32*, 2802–2818.

118 I. Huc, Eur. *J. Org. Chem.* **2004**, 17–29.

119 Y. Hamuro, S. J. Geib, A. D. Hamilton, *J. Am. Chem. Soc.* **1996**, *118*, 7529–7541.

120 J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem. Intl. Ed.* **2003**, *42*, 535–539.

121 D. A. P. Delnoye, R. P. Sijbesma, J. A. J. M. Vekemans, E. W. Meijer, *J. Am. Chem. Soc.* **1996**, *118*, 8717–8718.

122 C. A. Hunter, D. H. Purvis, *Angew. Chem. Intl. Ed. Engl.* **1992**, *31*, 792–795.

123 O. Safarowsky, M. Nieger, R. Frohlich, F. Vogtle, *Angew. Chem. Intl. Ed.* **2000**, *39*, 1616–1618.

124 A. R. Sanford, L. Yuan, W. Feng, K. Yamato, R. A. Flowers, B. Gong, *Chem. Commun.* **2005**, 4720–4722.

125 J. Zhu, R. D. Parra, H. Zeng, E. Skrzypczak-Jankun, X. C. Zeng, B. Gong, *J. Am. Chem. Soc.* **2000**, *122*, 4219–4220.

126 B. Gong, *Chem Eur. J.* **2001**, *7*, 4336–4342

127 H. Jiang, J.-M. Leger, I. Huc, *J. Am. Chem. Soc.* **2003**, *125*, 3448–3449.

128 H. Jiang, J.-M. Leger, C. Dolain, P. Guionneau, I. Huc, *Tetrahedron*, **2003**, *59*, 8365–8374.

129 B. Gong, H. Zeng, J. Zhu, L. Yuan, Y. Han, S. Cheng, M. Furukawa, R. D. Parra, A. Y. Kovalevsky, J. L. Mills, E.Skrzypczak-Jankun, S. Martinovic, R. D. Smith, C. Zheng, T. Szyperski, X. C. Zeng, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11583–11588.

130 L. Yuan, A. R. Sanford, W. Feng, A. Zhang, J. Zhu, H. Zeng, K. Yamato, M. Li, J. S. Ferguson, B. Gong, *J. Org. Chem.* **2005**, *70*, 10660–10669.

131 K. Burgess, D. S. Linthicum, H. Shin, *Angew. Chem. Intl. Ed. Engl.* **1995**, *34*, 907–909.

132 K. Burgess, J. Ibarzo, D. S. Linthicum, D. H. Russell, H. Shin, A. Shitangkoon, R. Totani, A. J. Zhang, *J. Am. Chem. Soc.* **1997**, *119*, 1556–1564.

133 J.-M. Kim, Y. Bi, S. J. Paikoff, P. G. Schultz, *Tetrahedron Lett.* **1996**, *37*, 5305–5308.

134 H. Tang, R. J. Doerksen, G. N. Tew, *Chem. Commun.* **2005**, 1537–1539.

135 J. C. Nelson, J. G. Saven, J. S. Moore, P. G. Wolynes, *Science* **1997**, *277*, 1793–1796.

136 R. B. Prince, J. G. Saven, P. G. Wolynes, J. S. Moore, *J. Am. Chem. Soc.* **1999**, *121*, 3114–3121.

137 D. Hill, J. S. Moore, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5053–5057.

138 J. M. Cary, J. S. Moore, *Org. Lett.* **2002**, *4*, 4663–4666.

139 L. Arnt, K. Nusslein, G. N. Tew, *J. Polym. Sci. A1* **2004**, *42*, 3860–3864.

140 L. Arnt, G. N. Tew, *J. Am. Chem. Soc.* **2002**, *124*, 7664–7665.

141 R. B. Brietenkamp, L. Arnt, G. N. Tew, *Polym. Adv. Technol.* **2005**, *16*, 189–194.

142 B. P. Orner, J. T. Ernst, A. D. Hamilton, *J. Am. Chem. Soc.* **2001**, *123*, 5382–5383.

143 O. Kutzki, H. S. Park, J. T. Ernst, B. P. Orner, H. Yin, A. D. Hamilton, *J. Am. Chem. Soc.* **2002**, *124*, 11838–11839.

144 H. Yin, G. Lee, K. A. Sedey, O. Kutzki, H. S. Park, B. P. Orner, J. T. Ernst, H.-G. Wang, S. M. Sebti, A. D. Hamilton, *J. Am. Chem. Soc.* **2005**, *127*, 10191–10196.

145 H. Yin, G. Lee, H. S. Park, G. A. Payne, J. M. Rodriguez, S. M. Sebti, A. D. Hamilton, *Angew. Chem. Intl. Ed. Engl.* **2005**, *44*, 2704–2707.

146 L. Chen, H. Yin, B. Farooqi, S. Sebti, A. D. Hamilton, J. Chen, *Mol. Cancer. Ther.* **2005**, *4*, 1019–1025.

147 H. Yin, A. D. Hamilton, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1375–1379.

148 H. Yin, G. Lee, K. A. Sedey, J. M. Rodriguez, H.-G. Wang, S. M. Sebti, A. D. Hamilton, *J. Am. Chem. Soc.* **2005**, *127*, 5463–5468.

149 K. Gademann, M. Ernst, D. Hoyer, D. Seebach, *Angew. Chem. Intl. Ed. Engl.* **1999**, *38*, 1223–1226.

150 K. Gademann, M. Ernst, D. Seebach, D. Hoyer, *Helv. Chim. Acta*, **2000**, *83*, 16–33.

151 D. Seebach, M. Reuping, P. I. Arvidsson, T. Kimmerlin, P. Micuch, C. Noti, D. Langenegger, D. Hoyer, *Helv. Chim. Acta* **2001**, *84*, 3503–3510.

152 H.-C. Chen, J. H. Brown, J. L. Morell, C. M. Huang, *FEBS Lett.* **1988**, *236*, 462–466.

153 A. Tossi, L. Sandri, A. Giangaspero, *Biopolymers* **2000**, *55*, 4–30.

154 M. Zasloff, *Nature* **2002**, *415*, 389–395.

155 G. N. Tew, D. Liu, B. Chen, R. J. Doerksen, J. Kaplan, P. J. Carroll, M. L. Klein, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 5110–5114.

156 D. Liu, S. Choi, B. Chen, R. J. Doerksen, D. J. Clements, J. D. Winkler, M. L. Klein, W. F. DeGrado, *Angew. Chem. Intl. Ed.* **2004**, *43*, 1158–1162.

157 S.-Y. Kao, A. F. Calman, P. A. Luciw, B. M. Peterlin, *Nature* **1987**, *330*, 489–493.

158 B. J. Calnan, B. Tidor, S. Biancalana, D. Hudson, A. D. Frankel, *Science* **1991**, *252*, 1167–1171.

159 R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, D. C. Spellmeyer, R. Tan, A. D. Frankel, D. V. Santi, F. E. Cohen, P. A. Bartlett, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9367–9371.

160 V. Kesavan, N. Tamilarasu, H. Cao, T. M. Rana, *Bioconjug. Chem.* **2002**, *13*, 1171–1175.

161 N. Tamilarasu, I. Huq, T. M. Rana, *J. Am. Chem. Soc.* **1999**, *121*, 1597–1598.

162 X. Wang, I. Huq, T. M. Rana, *J. Am. Chem. Soc.* **1997**, *119*, 6444–6445.

163 M. A. Gelman, S. Richter, H. Cao, N. Umezawa, S. H. Gellman, T. M. Rana, *Org. Lett.* **2003**, *5*, 3563–3565.

164 Y. C. Patel, *Front. Neuroendocrinol.* **1999**, *20*, 157–198.

165 A. Janecka, M. Zubrzycka, T. Janecki, *J. Pept. Res.* **2001**, *58*, 91–107.

166 R. M. Freidinger, D. S. Perlow, W. C. Randall, R. Saperstein, B. H. Arison, D. F. Veber, *Int. J. Pept. Prot. Res.* **1984**, *23*, 142–150.

167 G. Melacini, Q. Zhu, G. Osapay, M. Goodman, *J. Med. Chem.* **1997**, *40*, 2252–2258.

168 W. Bauer, U. Briner, W. Doepfner, R. Haller, R. Huguenin, P. Marbach, T.

J. Petcher, J. Pless, *Life Sci.* **1982**, *31*, 1133–1140.

169 C. Nunn, M. Reuping, D. Langenegger, E. Schuepbach, T. Kimmerlin, P. Micuch, K. Hurth, D. Seebach, D. Hoyer, *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2003**, *367*, 95–103.

170 J. A. Patch, A. E. Barron, *J. Am. Chem. Soc.* **2003**, *125*, 12092–12093.

171 S. Ng, B. Goodson, A. Ehrhardt, W. H. Moos, M. Siani, J. Winter, *Bioorg. Med. Chem.* **1999**, *7*, 1781–1785.

172 L. Arnt, G. N. Tew, *Langmuir 2003*, *19*, 2404–2408.

173 H. Hauser, J. H. Dyer, A. Nandy, M. A. Vega, M. Werder, E. Bieliauskaite, F. E. Weber, S. Compassi, A. Gemperli, D. Boffelli, E. Wehrli, G. Schulthess, M.C. Phillips, *Biochemistry* **1998**, *37*, 17843–17850.

174 A. Rigotta, B. Trigatti, J. Babitt, M. Penman, S. H. Xu, M. Krieger, *Curr. Opin. Lipidol.* **1997**, *8*, 181–188.

175 D. Boffelli, S. Compassi, M. Werder, F. E. Weber, M. C. Phillips, G. Schulthess, H. Hauser, *FEBS Lett.* **1997**, *411*, 7–11.

176 I. Capila, R. J. Linhardt, *Angew. Chem. Intl. Ed.* **2002**, *41*, 390–412.

177 R. Lever, C. P. Page, *Nat. Rev. Drug Disc.* **2002**, *1*, 140–148.

178 A. D. Cardin, H. J. R. Weintrub, *Arteriosclerosis* **1989**, *9*, 21–32.

179 T. W. Wakefield, P. C. Andrews, S. K. Wroblewski, A. M. Kadell, S. Tejwani, M. S. Hulin, J. C. Stanley, *J. Surg. Res.* **1996**, *63*, 280–286.

180 J. R. Fromm, R. E. Hileman, E. E. O. Caldwell, J. M. Weiler, R. J. Linhardt, *Arch. Biochem. Biophys.* **1997**, *343*, 92–100.

181 L.-C. Chang, J. F. Liang, H.-F. Lee, L. M. Lee, V. C. Yang *AAPS Pharmsci* **2001**, *3*, art. No. 18.

182 S. Onoue, Y. Nemoto, S. Harada, T. Yajima, K. Kashimoto, *Life Sci.* **2003**, *73*, 2793–2806.

183 S. Choi, D. J. Clements, V. Pophristic, I. Ivanov, S. Vemparala, J. S. Bennett, M. L. Klein, J. D. Winkler, W. F. Degrado, *Angew. Chem. Intl. Ed. Engl.* **2005**, *44*, 6685–6689.

184 C. B. Klee, T. C. Vanaman, *Adv. Prot. Chem.* **1982**, *35*, 213–321.

185 S. Kakiuchi, K. Sobue, *Trends Biochem. Sci.* **1983**, *8*, 59–62.

186 W. E. Meador, A. R. Means, F. A. Quiocho, *Science* **1992**, *257*, 1251–1255.

187 H. Yin, K. K. Frederick, D. Liu, A. J. Wand, W. F. Degrado, *Org. Lett.* **2006**, *8*, 223–225.

188 M. Mrksich, W. S. Wade, T. J. Dwyer, B. H. Geierstanger, D. E. Wemmer, P. B. Dervan, *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 7586–7590.

189 M. Mrksich, M. E. Parks, P. B. Dervan, *J. Am. Chem. Soc.* **1994**, *116*, 7983–7988.

190 J. W. Trauger, E. E. Baird, P. B. Dervan, *Chem Biol.* **1996**, *3*, 369–377.

191 J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature,* **1996**, *382*, 559–561.

192 R. P. L. de Clairac, B. H. Geierstanger, M. Mrksich, P. B. Dervan, D. E. Wemmer, *J. Am. Chem. Soc.* **1997**, *119*, 7909–7916.

193 D. M. Herman, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 1382–1391.

194 W. A. Greenberg, E. E. Baird, P. B. Dervan, *Chem. Eur. J.* **1998**, *4*, 796–805.

195 C. R. Woods, T. Ishii, B. Wu, K. W. Bair, D. L. Boger, *J. Am. Chem. Soc.* **2002**, *124*, 2148–2152.

196 D. M. Herman, J. M. Turner, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 1121–1129.

197 C. Melander, D. M. Herman, P. B. Dervan, *Chem. Eur, J.* **2000**, *6*, 4487–4497.

198 J. M. Gottesfeld, L. Neely, J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1997**, *387*, 202–205.

199 L. A. Dickinson, R. J. Gulizia, J. W. Trauger, E. E. Baird, D. E. Mosier, J. M. Gottesfeld, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12890–12895.

200 B. A. Lenzmeier, E. E. Baird, P. B. Dervan, J. K. Nyborg, *J. Mol. Biol.* **1999**, *291*, 731–744.

201 H. Simon, L. Kittler, E. Baird, P. Dervan, C. Zimmer, *FEBS Lett.* **2000**, *471*, 173–176.

202 N. R. Wurtz, J. L. Pomerantz, D. Baltimore, P. B. Dervan, *Biochemistry* **2002**, *41*, 7604–7609.

203 R. E. Bremer, E. E. Baird, P. B. Dervan, *Chem. Biol.* **1998**, *5*, 119–133.

204 S. Janssen, O. Cuvier, M. Muller, U. K. Laemmli, *Mol. Cell* **2000**, *6*, 1013–1024.

205 L. A. Dickinson, J. W. Trauger, E. E. Baird, P. Ghazal, P. B. Dervan, J. M. Gottesfeld, *Biochemistry* **1999**, *38*, 10801–10807.

206 J. J. Coull, G. He, C. Melander, V. C. Rucker, P. B. Dervan, D. M. Margolis, *J. Virol.* **2002**, *76*, 12349–12354.

207 A. K. Mapp, A. Z. Ansari, M. Ptashne, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3930–3935.

208 A. Z. Ansari, A. K. Mapp, D. H. Nguyen, P. B. Dervan, M. Ptashne, *Chem. Biol.* **2001**, *8*, 583–592.

209 J. M. Gottesfeld, J. M. Belitsky, C. Meander, P. B. Dervan, K. Luger, *J. Mol. Biol.* **2002**, *321*, 249–263.

210 T. A. Vickers, M. C. Griffith, K. Ramasamy, L. M. Risen, S. M. Freier, *Nucleic Acids Res.* **1995**, *23*, 3003–3008.

211 J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble, L. E. Babiss, *Science* **1992**, *258*, 1481–1485.

212 G. Dieci, R. Corradini, S. Sforza, R. Marchelli, S. Ottonello, *J. Biol. Chem.* **2001**, *276*, 5720–5725.

213 B. M. Tyler, K. Jansen, D. J. McCormick, C. L. Douglas, M. Boules, J. A. Stewart, L. Zhao, B. Lacy, B. Cusack, A. Fauq, E. Richelson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 7053–7058

214 N. E. Mollegard, O. Buchardt, M. Egholm, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3892–3895.

215 G. Wang, X. Xu, B. Pace, D. A. Dean, P. M. Glazer, P. Chan, S. R. Goodman, I. Shokolenko, *Nucleic Acids Res.* **1999**, *27*, 2806–2813.

216 G. Wang, K. Jing, R. Balczon, X. Xu, *J. Mol. Biol.* 2001, *313*, 933–940.

217 A. I. Glukhov, O.V. Zimnik, S.A. Gordeev, S.E. Severin, *Biochem. Biophys. Res. Commun.* **1998**, *248*, 368–371.

218 J. E. Mata, S. S. Joshi, B. Palen, S. J. Pirruccello, J. D. Jackson, N. Elias, T. J. Page, K. L. Medlin, P. L. Iversen, *Toxicol. Appl. Pharmacol.* **1997**, *144*, 189–197.

219 A. E. Pitts, D. R. Corey, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11549–11554.

220 J. C. Norton, M. A. Piatyszek, W. E. Wright, J. W. Shay, D. R. Corey, *Nat. Biotechnol.* **1996**, *14*, 615–619.

221 R. Villa, M. Folini, S. Lualdi, S. Veronese, M. G. Daidone, N. Zaffaroni, *FEBS Lett.* **2000**, *473*, 241–248.

222 P. Chene, *Nat. Rev. Cancer* **2003**, *3*, 102–109.

223 D. A. Freedman, L. Wu, A. J. Levine, *Cell. Mol. Life Sci.* **1999**, *55*, 96–107.

224 S. M. Picksley, D. P. Lane, *Bioessays* **1993**, *15*, 689–690.

225 J. Roth, M. Dobbelstein, D. A. Freedman, T. Shenk, A. J. Levine, *EMBO J.* **1998**, *17*, 554–564.

226 Y. Haupt, R. Maya, A. Kazaz, M. Oren, *Nature*, **1997**, *387*, 296–299.

227 M. H. G. Kubbatat, S. N. Jones, K. H. Vousden, *Nature*, **1997**, *387*, 299–303.

228 P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine, N. P. Pavletich, *Science*, **1996**, *274*, 948–953.

229 A. Bottger, V. Bottger, C. Garcia-Echeverria, P. Chene, H.-K. Hochkeppel, W. Sampson, K. Ang, S. F. Howard, S. M. Picksley, D. P. Lane, *J. Mol. Biol.* **1997**, *269*, 744–756.

230 P. Chene, J. Fuchs, I. Carena, P. Furet, C. Garcia-Echevaria, *FEBS Lett.* **2002**, *529*, 293–297.

231 J. D. Rich, J. M. Crawford, S. N. Kazanjian, P. H. Kazanjian, *Clin. Infect. Dis.* **1992**, *15*, 609–614.

232 L. Harkins, A. L. Volk, M. Samanta, I. Mikolaenko, W. J. Britt, K. I. Bland, C. S. Cobbs, *Lancet* **2002**, *360*, 1557–1563.

233 C. S. Cobbs, L. Harkins, M. Samanta, G. Y. Gillespie, S. Bharara, P. H.

King, L. B. Nabors, C. G. Cobbs, W. J. Britt, *Cancer Res.* **2002**, *62*, 3347–3350.

234 M. Lopper, T. Compton, J. Virol. 2004, 78, 8333–8341.

235 T. Hara, S. R. Durell, M. C. Myers, D. H. Appella, *J. Am. Chem. Soc.* **2006**, *128*, 1995–2004.

236 J. M. Adams, S. Cory, *Science* **1998**, *281*, 1322–1326.

237 A. Strasser, D. C. S. Huang, D. L. Vaux, *Biochim. Biophys. Acta* **1997**, *1333*, F151–F178.

238 D. T. Chao, S. J. Korsmeyer, *Annu. Rev. Immunol.* **1998**, *16*, 395–419.

9

Protein Design

Jean-Luc Jestin and Frédéric Pecorari

9.1
Introduction

Most natural proteins are not adapted to our needs for therapeutic or even bio-technological uses. The challenge of protein design is to define efficient and comprehensive ways for the identification of new or improved proteins with the required properties. The design of new proteins relies on sequence–structure–function relationships and is limited by our partial knowledge of the protein folding mechanism. But as we will see in this chapter, creating new proteins or proteins with desired properties is clearly possible.

A striking characteristic of natural proteins is that they are able to adopt a folded state despite their very high sequence diversity. About one thousand different folds are known. Natural proteins constitute the richest family of folded molecules. When targeting specific functions and applications, non-natural folded proteins are of obvious interest. They can be created by protein engineering thanks to their plasticity using diverse strategies including mutagenesis (random mutagenesis, recombination or site-directed mutagenesis making use of structures and models), domain fusions and sequence–activity relationships deriving from natural and directed evolution.

This widely applied strategy aims to explore the potential benefit of natural proteins. There is an advantage in knowing the protein's three-dimensional structure, as modifications can be introduced rationally into the protein while trying not to destroy its foldability. This predictive approach can be used to change ligand or reaction specificities, or to increase stability, for example. New proteins can be also designed by combining several natural proteins or their functional domains in one polypeptide chain. Although many successes have been reported with rational approaches to protein design, it remains very difficult in many cases to make drastic changes to a protein. Thus, with the development of powerful molecular biology tools, the combinatorial approach has progressed remarkably during the last 15 years. The principle is to introduce mutations in the gene of interest and create large collections of variants of the encoded protein. The result-

Table 9.1 Comparison of the properties of selection versus screening.

		SELECTION *in vivo* or *in vitro*	SCREENING
1	Analysis	in parallel	in series
2	Number of variants analyzed	10^6 to 10^{13}	10^3 to 10^8
3	Substrate quantities required	µg to mg	g
4	Isolated variants	selected population	one or a few

ing library is then screened or selected to identify clones with the desired property. Several rounds of random mutagenesis and of screening or selection can be conducted in order to obtain variants with better properties (Table 9.1). Such processes represent an accelerated mimicry of the natural evolution of proteins, since there is a generator of variability associated to a selection pressure (Fig. 9.1). These approaches make general use of the link between the genotype (the gene which can be amplified and sequenced) and the phenotype (the protein's function) (Fig. 9.2). The most efficient way to achieve the expected result is to combine the predictive and combinatorial approaches in order to target randomizations in only a part of the protein, for example the active or binding site. The idea is to minimize perturbation of the folding and stability of the protein. It has therefore been possible to obtain new proteins with desired properties and mutations that would have been impossible to predict.

Directed protein evolution

Amplification
Transcription
Translation

Protein library
10^6 to 10^{13}

Screening
or
selection

Selected proteins

Sequencing of corresponding genes
Biochemical characterisation

Characterised proteins
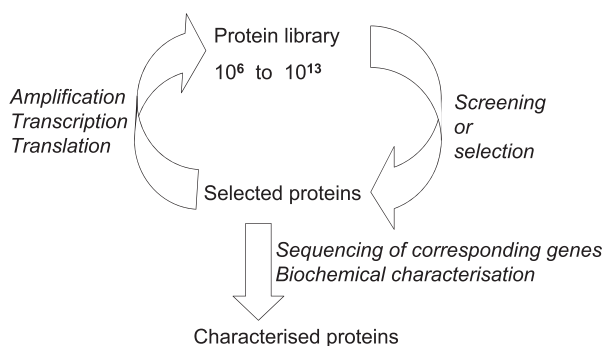
Fig. 9.1 The evolution of protein libraries of more than 10^6 proteins can be directed experimentally by screening or by selection. To enrich the population of variants in proteins of interest, several cycles of selection of proteins and of amplification of their genes are necessary. An evaluation of the number of cycles required can be estimated by measuring enrichment factors [80]. Amplifications may also be mutagenic, thereby generating new libraries from a selected population of proteins of interest.

Phage

Cell

Ribosome

mRNA
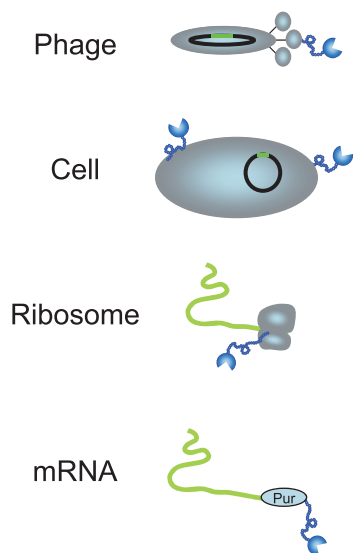
Fig. 9.2 Scheme of several combinatorial tools for selections. A physical entity ensures the link between phenotype (protein = blue) and genotype (gene = green). It can be phages, cells, ribosomes or puromycin molecules for example. These systems allow firstly the isolation of proteins according to their function from a library of proteins and secondly the characterization of selected proteins by amplification and sequencing of the corresponding genes linked to it. Selections are generally done by binding to a ligand immobilized on a solid surface to facilitate washings. After each round of selection, the binders of interest are amplified *via* cell multiplication, phage propagation, or PCR depending on the method of selection used. Finally, after several rounds of selection, the selected variants are isolated, analyzed for the desired properties, and their sequence identified.

After several decades of protein folding studies, it is still a real challenge to design fully *de novo* proteins. Using rules deduced from sequence–structure–function relationships for natural proteins, some proteins with simple folds, such as a four helix bundle fold, have been successfully designed *de novo*.

In this chapter key challenges in protein engineering will be reviewed focusing on the strategies used to isolate a protein with the desired properties. Strategies that make use of natural scaffolds, of domains, of secondary structures, of altered amino acid alphabets and strategies that do not make use of known protein sequences will be addressed here.

9.2
Design of Proteins from Natural Scaffolds

The simplest strategy to design active proteins is to start from proteins with similar functions or from proteins with known folds and to introduce insertions,

deletions or substitutions using amino acids from the standard genetic code. In this case, the idea is to keep the foldability of a protein that has been naturally selected during evolution while introducing targeted substitutions into the protein to change its specificity. This strategy makes the assumption that the foldability and the stability of a protein are robust enough to support quite a high mutational load. The use of screening or selection tools is often powerful enough to eliminate variants that are not folded.

9.2.1
Design of Enzymes

9.2.1.1 Grafting Catalytic Sites in Proteins

A catalytic site with a triose phosphate isomerase activity was grafted onto a ribose binding protein [1]. Computational design was used to create a substrate binding site and to position catalytic residues at the active site. Finally, a powerful *in vivo* selection in *E. coli* allowed the isolation of active variants with rate enhancements of 10^5 to 10^6 when compared to the uncatalyzed reaction.

Similarly, by combining *in silico* modeling of protein insertions and deletions with *in vivo* selection of variants for antibiotic resistance allowed a metallohydrolase to be converted into a cefotaxime hydrolyzing enzyme [2].

9.2.1.2 Endowing Enzymes with Two Catalytic Activities in a Single Domain

Endowing an enzyme with two catalytic activities was achieved using sequence–catalytic activity relationships or using sequence–structure–catalytic activity relationships [3]. Recent experiments tend to show that the simplest strategy to enlarge the spectrum of catalytic activities of a polypeptide may be to direct its evolution towards a new catalytic activity (Fig. 9.2).

Directed evolution of a library of more than 10^7 *Thermus aquaticus* DNA-polymerase I mutants by *in vitro* selection for RNA-dependent DNA-polymerase activity (Fig. 9.3) yielded the isolation of DNA polymerase variants that copy both DNA and RNA with similar catalytic efficiencies [4]. This result is astonishing as the experimental evolution procedure did not include a selection for DNA-dependent DNA-polymerization.

In another study, 2×10^5 *T. aquaticus* DNA polymerase I mutants within the highly conserved motif A were selected *in vivo* for the wild-type activity, DNA-dependent DNA-polymerization. Screening of selected clones for DNA-dependent RNA-polymerization activity allowed the identification of variants with a 10^3-fold greater ribonucleotide incorporation efficiency than that of the wild-type enzyme. Ribonucleotide incorporation was still up to ten-fold less efficient than deoxyribonucleotide incorporation [5]. Direct selection for RNA-polymerization using phage display allowed the isolation of a variant of the Stoffel fragment of *T. aquaticus* DNA polymerase I with similar catalytic efficiencies for RNA- and DNA-polymerization on DNA templates. These catalytic efficiencies were found to be
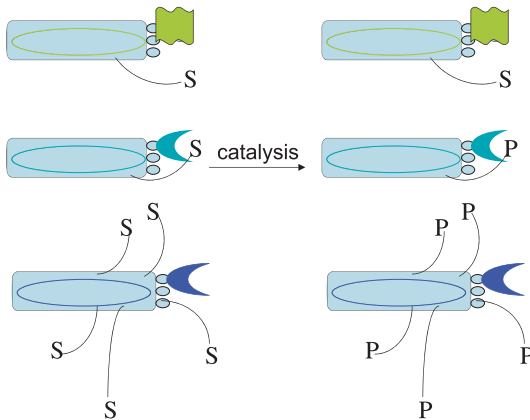
Fig. 9.3 *In vitro* selection for catalytically active proteins using phage display. The *in vitro* selection for enzymes includes a selection for foldability. An unfolded polypeptide (yellow), a folded polypeptide with a low catalytic activity (green) and a folded polypeptide with a high catalytic turn-over (blue). Substrates are coupled to purified phage particles. The *in vitro* selection for catalysis relies on isolation by affinity chromatography for the product bound to phage-enzymes. If the polypeptide is unfolded or inactive, the substrates will not be converted to products; no particles will be purified by affinity chromatography. If the polypeptide has a low catalytic activity, the phage particle will not be isolated efficiently by affinity chromatography for the product. If the polypeptide has a high catalytic turn-over, multiple substrates are converted into products. Because of the chelate effect, this phage particle derivatized by multiple products will be efficiently isolated by affinity chromatography. The genes encoding active enzymes are located within the phage particles. This strategy can be applied to libraries of more than 10^7 variant enzymes.

less than ten-fold lower than that of DNA-dependent DNA-polymerization of the wild-type Stoffel fragment [6].

9.2.1.3 Grafting Allosteric Sites to Regulate Enzyme Activity

Regulation of an enzyme's activity was achieved by introducing mutations that allowed the binding of chosen ligands which could act as regulators.

Libraries of beta-lactamase variants with mutations in loops surrounding the catalytic site were constructed and allowed the isolation of lactamases binding various proteins such as streptavidin, ferritin, beta-galactosidase or antibody fragments [7]. This approach has been used as a method to titrate an antigen (the prostate-specific antigen, PSA) in solution by measuring the lactamase activity. In fact, the enzyme was engineered so as to bind an anti-PSA antibody fragment, forming a complex which was found to inhibit the enzymatic activity by 90% [8].

In another report, beta-lactamase was inserted randomly within a maltose-binding protein. The fusion proteins were selected for ampicillin resistance. Further screening for maltose-dependent activity allowed the isolation of lactamase variants with a 600-fold improved rate in the presence of maltose [9].

9.2.2
Design of Binding Proteins

The key points in designing binding proteins are (i) to choose a starting protein that will be used as a scaffold on the basis of its favorable biophysical properties and (ii) to define a mutagenesis scheme that will allow generation of the new binding property while minimizing negative effects on the foldability of the protein. It has been known for a long time that proteins can tolerate mutations, especially if mutated amino acids are located at their surface [10].

Antibodies, in particular their variable domains, are among the most studied proteins as they exemplify how a protein can show extraordinary adaptability to bind very different kinds of ligands, such as proteins, peptides, nucleic acids, sugars, etc. In antibodies, the scaffold is constituted of β-strands that are connected by six loops. These loops are highly variable in length and in sequence from antibody to antibody and provide the diversity needed for binding the ligand, while the immunoglobulin fold remains constant. Many attempts have been made to recreate this concept *in vitro* with antibodies. A study involving randomization of the most significant loop for antigen-binding and diversity (CDR-H3) [11] opened the door to fully synthetic libraries of antibodies. In this work, the authors showed that the randomization of one loop followed by selections by phage display was sufficient to convert an antibody specific for the human tetanus toxicoid antigen into antibodies specific for fluorescein. This approach was then extended to several loops [12, 13, 14]. With these libraries it was possible to isolate new antibodies with desired specificities for various ligands, often with dissociation constants in the nano- or even picomolar range.

Unfortunately, antibodies or their derivative fragments are often difficult to produce because of their molecular complexity and/or their stability which do not make them ideal molecules for nontherapeutic applications. For these reasons, alternative methods have been developed using other natural proteins engineered *via* a combinatorial mutation/selection approach. The aim was to challenge antibodies for their favorable properties while removing their disadvantages. Many projects have been undertaken during the last decade in this field and we will only focus on those where a three dimensional structure is available for binders generated from a designed scaffold alternative to antibodies.

The staphylococcal protein A is a bacterial receptor that binds the Fc region of immunoglobulin G. It has repetitive subunits of about 58 amino acids that are individually folded into a three-helix bundle (Fig. 9.4A), that is highly soluble and stable. Isolated domains, such as the Z-domain, were shown to tolerate multiple amino acids substitutions at the binding area involved in Fc recognition [15]. Combinatorial libraries corresponding to the randomization of thirteen residues, located into two α-helices of the Z-domain at the binding surface, were used for selections against three different proteins [16]. This work showed that the binding specificity of the Z-domain was changed from Fc to Taq polymerase, insulin or apolipoprotein recognition. Circular dichroism experiments indicated that these engineered binders (called affibodies) had a secondary structure similar to
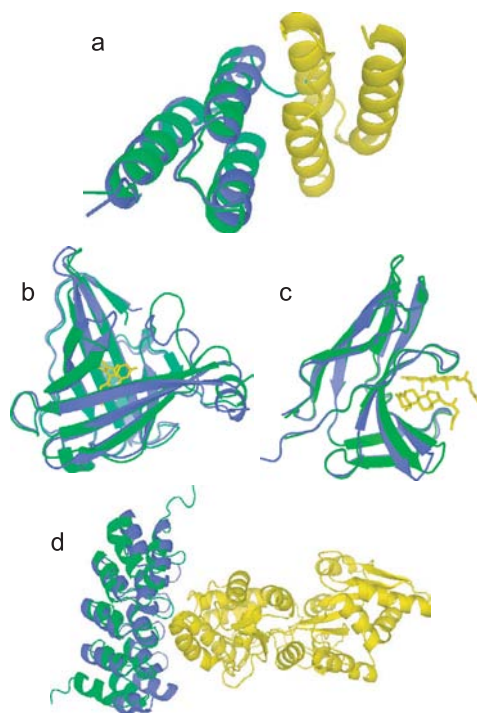
Fig. 9.4 (A) Superimposition of structures of a complex Z$_{SPA-1}$ affibody-ligand (PDB ID: 1LP1) (blue) and of wild-type protein Z domain (PDB ID: 1Q2N) (green); ligand (yellow). (B) Superimposition of structures of a complex FluA anticalin-ligand (PDB ID: 1N0S) (blue) and of wild-type bilin binding protein (PDB ID: 1BBP) (green); ligand (yellow) (fluorescein). (C) Superimposition of structures of a complex neocarzinostatin-ligand (3.24 binder) (PDB ID: 2CBO) (blue) and of wild-type neocarzinostatin (PDB ID: 1NCO) (green); ligand (yellow) (testosterone hemisuccinate). (D) Superimposition of structures of a complex designed ankyrin-ligand (PDB ID: 1SVX) (off7 binder) (blue) and of a wild-type ankyrin (PDB ID: 1AP7) (green); ligand (yellow) (maltose binding protein).

the native Z-domain suggesting that their fold is similar. Initial affinities obtained were in the micromolar range. With an affinity maturation step, *via* random substitution of six residues and a new selection, it was possible to obtain binders with dissociation constants around 40 nM, against the Taq polymerase [17] which is similar in strength to the association of the parent staphylococcal protein A with the Fc fragment. Binders with affinities in the nanomolar range were also obtained against several other protein targets such as the extra cellular domain of the HER2 receptor [18] with dissociation constants as low as 22 pM [19]. The concept of re-directing the specificity of a natural protein was thus clearly demonstrated in this case. However, the foldability of the resulting binders was not demonstrated. To answer these questions structural studies were undertaken for several affibodies, either alone or in complex with their respective partner, by

NMR [20], [21] and by X-ray crystallography [22]. The structures for the binder Z_{SPA-1} revealed that its interaction surface has some characteristics comparable to those observed in antibody-protein complexes (size of the surface area and hydrogen bonding sites) [22] and ten of the thirteen randomized residues were indeed interacting with the target. However, these structural data [20] and biophysical characterizations [23] showed that uncomplexed affibodies behave like a molten globule and that folding occurs upon binding to the target protein. By contrast, it was shown that other affibodies (Z_{Taq} and anti-Z_{Taq}) were well structured, folded even uncomplexed [21] and had higher melting temperature than for Z_{SPA-1} [24]. These studies showed that individual structures of affibodies are superimposable, indicating that the scaffold of the wild-type Z-domain was preserved after mutagenesis, at least in the complex. However, for the same scaffold, the assumption that foldability is retained after surface substitutions have been made, was verified in one case and not in the other. This suggests that selection for a binding property does not mean selection for foldability, at least for the isolated binders.

The bilin-binding protein, a member of the lipocalin protein family has been recruited as a scaffold for the development of new affinity reagents called ''anticalins''. This protein of 174 amino acid residues folds into a β barrel with eight strands connected by loops which are involved in the binding of biliverdin (Fig. 9.4B). These loops indeed form a rather wide and shallow pocket for the hapten. Libraries of lipocalin mutants were created by randomization of sixteen residues distributed across four loops and were used for selection against several low molecular weight compounds. Binders were obtained by phage display with dissociation constants of 35 nM and 295 nM for fluorescein [25] and digoxigenin [26], respectively. The affinity was further improved with a resulting dissociation constant of 12 nM for digoxigenin after an affinity maturation step [27] and 1 nM for fluorescein by rational design [28]. Biophysical characterization of one mutant by circular dichroism showed that some of the residues in a β-strand and α-helical conformation were the same as those deduced from the crystallographic structure of the parent protein. A structural study of the mutant FluA in complex with fluorescein further supported the idea that lipocalin can be mutated in the targeted loops without significantly disturbing its overall conformation [29]. Indeed, the major perturbations were observed in the loops. Hence, the foldability of the resulting variants was similar to the wild-type lipocalin.

Neocarzinostatin, an enediyne-binding chromoprotein of 113 amino acids consists of seven β-strands forming a β-sandwich, an architecture similar to antibodies (Fig. 9.4C). However, the binding of the chromophore, a small organic compound, involves thirteen amino-acids that constitute a binding pocket. To explore the potential of this pocket to bind other haptens, several libraries corresponding to randomization of up to thirteen residues were created and used for selection by phage display [30]. Variants able to bind specifically testosterone have been isolated. These binders showed nanomolar dissociation constants for streptavidin bound testosterone and micromolar K_D values for the free soluble form of testosterone. Thus, it is possible for this scaffold to change its recognition proper-

ties to bind an unrelated ligand. The three dimensional structures of several mutants were obtained to understand the structural basis of the testosterone recognition [31]. These studies revealed that indeed the structure for the binder 3.24 was retained in the free and complexed forms. Uncomplexed binders 1a.15 and 4.1 were found to have the same structure as the wild-type protein, while they formed dimers when complexed with the ligand. Hence, the foldability of the neocarzinostatin was also retained after massive randomization of its amino acids.

Some proteins such as ankyrins are composed of several repeated motifs. A modular approach to library construction has been performed using these kinds of motifs. In ankyrins the repetitive structural units consist of twenty to forty amino acids folded in a β-turn and two α-helices. In contrast to other described scaffolds, for which a strictly natural protein was used as a starting point for library construction, the authors first designed self-compatible repeat modules [32], [33] by consensus analysis of naturally existing ankyrin repeats. Biophysical studies showed that this strategy produced well expressed, soluble, stable and folded proteins [34]. The three dimensional structure of a consensus ankyrin, with three central repeats and two N- and C-terminal caps, showed that the designed ankyrin retained the fold of natural ankyrins [35]. Hence, using the consensus sequence libraries were constructed with two or three randomized ankyrin repeat domains comprising six diversified residues [33]. These libraries were used for selections by ribosome display. Binders with nanomolar dissociation constants were obtained for maltose binding protein, two eukaryotic protein kinases [36] and a bacterial kinase [37]. The three dimensional structures of two binders, in complex either with maltose binding protein [36] (Fig. 9.4D) or aminoglycoside phosphotransferase [38], were obtained. These structures further confirmed that foldability was retained for the binders compared to wild-type ankyrins.

9.3
Design of Proteins from Building Blocks

9.3.1
Design of Proteins from Structural Domains

Foldability is a molecular property that can be selected for and this allows the isolation of stable structural domains. The methods for the isolation of these domains and their applications are highlighted below.

9.3.1.1   Methods for the Identification of Stable Structural Domains
Powerful tools for the identification of stable structural domains are *in vitro* selection for folding or *in vitro* selection for improved stability [39]. These methods allow the isolation of rare structural domains among large repertoires of more than 10^7 protein sequences. They are represented for two types of proteins, those that fold into a stable structural domain and those that do not fold (Fig. 9.5). These strategies make use of the display of proteins on phage or on the ribosome

A.

proteolytic
digestion

B.

proteolytic
digestion
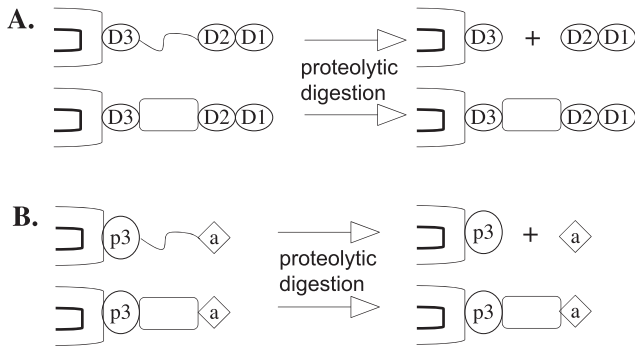
Fig. 9.5 Folded domains can be selected for *in vitro* by phage display, which establishes a link between a protein and the corresponding gene (thick line): the protein is displayed on the surface of bacteriophage particle (*Inovirus*) by fusion with a phage coat protein such as protein p3. The phage tips are represented. (A) A folded domain (rectangle) or an unfolded polypeptide (thin line) is inserted between domain 3 and domains 1 and 2 of the protein p3. Digestion by a proteolytic enzyme cleaves the fusion protein if it is unfolded. Phage particles displaying only domain D3 of p3 are not infectious and are counter-selected. Phage particles with the three domains of p3 are infectious, thereby allowing the corresponding gene to be amplified in *E. coli*. (B) A folded domain (rectangle) or an unfolded polypeptide (thin line) is inserted between protein p3 and an affinity tag 'a'. Digestion by a proteolytic enzyme cleaves the unfolded polypeptide and releases the affinity tag in the solution. Affinity chromatography for the tag 'a' allows the isolation of folded domains and of their genes.

so as to establish a link between the phenotype (the folded or unfolded domain) and the genotype (the corresponding gene which can be amplified) [40, 41]. Proteolytic digestion of unfolded domains is coupled to loss of infectivity of the phage particle, thereby preventing its recovery after infection of *E. coli*. In analogous methods, proteolytic digestion of unfolded domains is coupled to release of an affinity tag, preventing the recovery of the phage particle by affinity chromatography for the tag [42].

9.3.1.2 Identifying New Folds and New Topologies

The *in vitro* selection using phage display described in the previous section was applied to a library of about 3×10^7 random DNA fragments of the *E. coli* genome. 124 unique fragments were found to be resistant to trypsin which cleaves the polypeptide chain at arginine and lysine residues. About two-thirds of the fragments also displayed resistance to the protease thermolysin, which cleaves at aromatic and aliphatic amino acids [42]. This method for identification of structural domains should be useful for structural genomics studies either by NMR which requires small folded domains or by crystallography as crystallization can be improved by removal of the flexible regions in proteins.

A new fold was also identified using the same *in vitro* selection strategy. A library of random fragments of about 40 amino acids in length from *E. coli* were fused to the N-terminal half of the cold shock protein A from *E. coli* and yielded

after selection a fusion with a C-terminal fragment of the ribosomal protein S1 [43]. The protein's architecture was found to be a unique six-stranded beta-barrel that assembles to form a tetramer [44].

Enzyme topology was also altered: in the case of *Methanococcus jannaschii* chorismate mutase, the quaternary structure of the protein was changed from a dimer into a monomer. The catalytic activity as measured by the kinetic parameters was found to be similar for both the dimeric and the monomeric enzymes. This method relies on the introduction of dimer destabilizing mutations and on *in vivo* selection of active enzymes [45].

9.3.1.3 Combining Domains

Given two domains A and B, domain fusions found in nature occur mainly in a single orientation AB or BA. It was estimated that both orientations (AB and BA) domains appear in only about 2% of the cases [46].

Fusion of domains can be used to endow catalysts with new functions. For example, thermostable DNA-dependent DNA-polymerases such as family A *T. aquaticus* DNA polymerase I or family B *Pyrococcus furiosus* DNA polymerase were fused to a thermostable protein binding double-stranded DNA, the *Sulfolobus solfataricus* Sso7D protein. These fusion proteins were found to incorporate more nucleotides per DNA binding event, i.e. to have an improved processivity [47]. This has important applications in polymerase chain reactions (PCR) for efficient amplification of long templates of up to 15 kilobases.

9.3.2
Design of Proteins from Secondary Structures

During the last five years the number of three dimensional structures of proteins available has more than tripled [48] to reach now about 34 000 structures in the RSCB Protein Data Bank. Meanwhile, the number of different protein folds described has less than doubled and currently there are about one thousand folds according to the CATH [49] and SCOP [50] databases. Hence, it has become clear that there are a limited number of protein folds in nature. Proteins are mostly composed of secondary structures such as α-helices and β-strands. The combination of these motifs and their spatial organization define the fold of a protein. Thus, strategies to experimentally explore the occurrence of folded proteins designed from a random combination of secondary structure building modules were undertaken (Fig. 9.6).

By constraining the periodicity of polar and nonpolar residues, sequences with propensity to form defined secondary structures can be generated. The Hecht group used a binary patterning that defines the locations of polar and nonpolar residues but not their precise identity. A combinatorial library was designed with the aim to generate sequences able to fold into four-helix bundles [51]. The encoded proteins had 74 amino acid residues of which only 18 residues were held constant (helix caps and turns). Characterization of the expressed proteins indicated that most of the designed sequences folded into compact α-helical struc-

a

α-sequence:

β-sequence:

b

α-helices   β-strands   β-turns

Sub-libraries of secondary structure modules

Library of randomly assembled secondary structure modules

Screening or selection for
compacity and solubility

Folded proteins
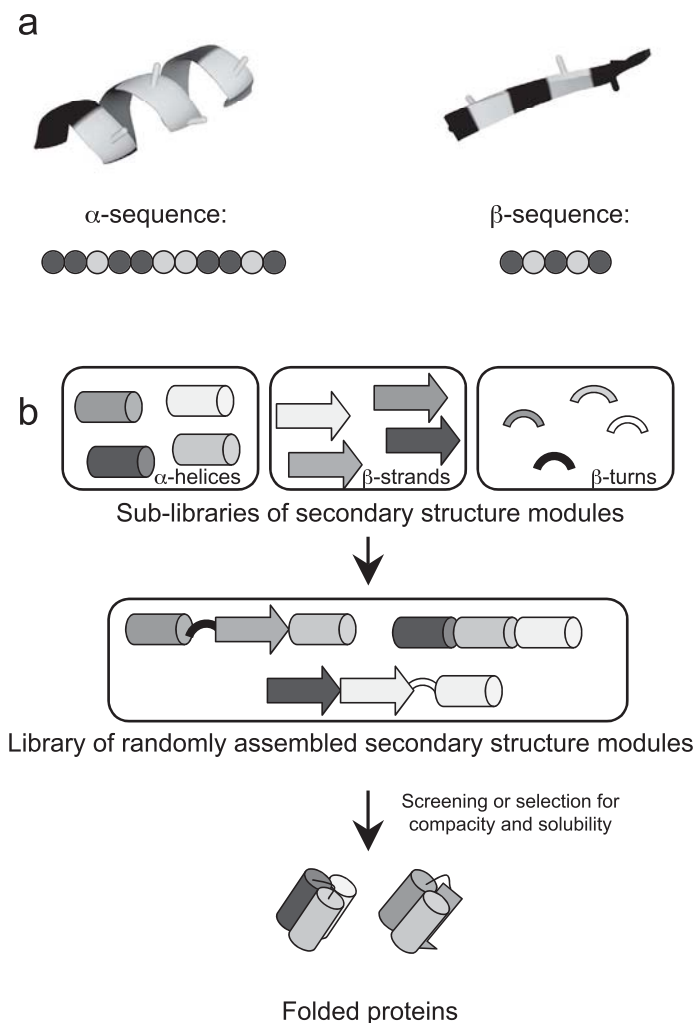
Fig. 9.6 (A) Schematic representation of binary patterning used for amphiphilic α-helix and β-strand. Polar and nonpolar residues are shown in dark grey and light grey; (B) Principle used for the combinatorial design of novel proteins from secondary structure modules.

tures. Thus, a simple binary code of polar and nonpolar residues arranged in the appropriate order can drive polypeptide chains to collapse into globular α-helical folds. Dozens of proteins from this initial library have been purified and characterized. They showed circular dichroism spectra characteristic of α-helical proteins, and some of them had native like properties, such as NMR chemical shift dispersion [52], cooperative chemical and thermal unfolding [53], and protection

Fig. 9.7 Structure of a four helix-bundle protein (clone S-824) isolated from a combinatorial secondary structure motifs library (PDB ID: 1P68 [56]).

of hydrogen exchange [54]. However most of the proteins from this initial library formed fluctuating structures and were probably molten globules. Using a typical molten globule-like protein from the original 74-residue library, a second generation library was constructed to increase length of the four helices (102-residue library) [55]. Biophysical characterization of five proteins by circular dichroism and NMR-measurements showed that stability and native-like properties were improved compared to the initial protein. This was confirmed by a structural study by NMR of a protein isolated from this improved library [56]. The experimentally determined structure was indeed a four-helix bundle as specified by the design (Fig. 9.7). It demonstrated also that the designed protein is not a molten globule and forms a unique structure.

In another study the Plückthun group also used the approach of binary patterning to design novel proteins [57]. The length and amino acid composition of the modules were determined according to rules deduced from secondary structure elements observed in natural proteins. For example, a serine followed by a proline or a glutamic acid were used to design the N-terminal caps of α-helices while the helices themselves were encoded semi-randomly with a binary patterning of polar and nonpolar residues having a propensity to form α-helices. The building blocks were thus generated at the DNA level and randomly assembled until the average DNA fragment length, corresponding to proteins of about 100 amino acids, was reached. Several libraries were built to generate proteins with a mixture of α-helices and β-turns, a combination of β-strands and β-turns, or a combination of α-helices, β-strands and β-turns.

Arbitrarily selected clones from these libraries were tested for expression in *E. coli*. The proportion of clones with detectable expression was found to be between about 8% and 84% depending on the library. Quite a high fraction of proteins in these synthetic libraries were resistant to cytoplasmic degradation. However, the solubility, which is a characteristic of most of natural globular proteins,

was moderate for expressible clones, with 10% to 60% of the proteins found in the soluble fraction. Further characterization by circular dichroism, size exclusion chromatography and sedimentation equilibrium experiments showed that some members from the all α-helices library were indeed helical, possessed a defined oligomerization state and showed cooperative chemical unfolding behavior. However, these proteins also showed properties consistent with a molten-globule state. By contrast, the all β-strand library led mainly proteins prone to aggregation. These results showed that an unexpected proportion of proteins built-up from structural secondary elements without β-strands possessed several of the favorable properties of natural proteins.

In order to avoid fastidious screening, it is an advantage to have a method to select folded proteins from large libraries of random proteins (see Section 9.3.1.1). Such an approach was developed to perform selections by ribosome display [41], based on two observations. Firstly, upon folding, a globular protein hides most of its hydrophobic residues. Secondly, once folded natural proteins often show a certain resistance to proteases due to their compactness. Thus, with a model system, the authors used a combination of hydrophobic interaction chromatography and proteolysis under limiting conditions and were able to select for a natural folded protein (a fragment of protein D from phage lambda) from a mixture including three other proteins previously obtained from the library of secondary structure modules. The potential of this selection approach is evident, and it will probably be applied to the search of folded proteins from combinatorial libraries.

9.4
Design of Proteins using Altered Alphabets

Two types of alphabets will be distinguished here, depending on whether they are larger or smaller than the amino acid alphabet of the standard genetic code. Reduced alphabets provide a tool for the investigation of compact and folded structures. Larger alphabets allow the introduction of unnatural amino acids within proteins.

9.4.1
Design of Proteins using Reduced Alphabets

Is it realistic to obtain folded proteins from libraries of random polypeptides using a less constraining design than for secondary structures but with a reduced alphabet of amino acid residues? Several studies have tried to answer this question, and we will focus on two of them.

In a pioneering study published by the Sauer group in 1994 [58] and in 1995 [59], a library of 80 to 100 residue proteins composed of random combination of glutamine (Q), leucine (L) and arginine (R) was constructed. Glutamine and leucine were chosen as representatives of hydrophilic and hydrophobic residues, respectively. Arginine was arbitrarily chosen to represent charged residues and was

included to increase solubility of the resulting random proteins. A FLAG and $(His)_6$-tag were added at the end of the proteins to facilitate purification and analysis. For this library (named QLR library), the authors claimed that no attempt was made to design secondary structures or tertiary interactions in the encoded proteins. However, it is known that the three residues used have a propensity to form helical structures. The QLR library expression in *E. coli* was analyzed by immunoblotting. It was found that 5% of in-frame proteins were indeed expressed in *E. coli*. More than 60% of these proteins had a size corresponding to the library design. Hence, quite a high proportion of random proteins were well produced in *E. coli* although unfolded proteins are usually degraded in this organism. However, protein solubility was fairly low (two of eleven clones). Three highly expressed proteins were analyzed by circular dichroism showing that their conformation was significantly helical, stable and unfold cooperatively as determined by guanidinium hydrochloride denaturation studies. These proteins were also relatively resistant to proteolysis, a characteristic of folded proteins. As shown by size exclusion chromatography and equilibrium centrifugation experiments, these proteins had a defined oligomeric state, but differed from native proteins in having no slowly exchanging amide hydrogens.

It is probable that the first proteins ever folded on Earth could have been composed with a limited set of amino acid residues. With the hypothesis that the prebiotic residues could be sufficient to obtain folded proteins, the Yanagawa group designed a library of random proteins using only five residues: Valine (V), Alanine (A), Aspartate (D), Glutamate (E), and Glycine (G) [60]. The so-called VADEG proteins were designed to be 100 residues long and were the result of random combinations of these five residues. The set of residues was chosen according to a prebiotic environment deduced from spark-discharge experiments and from the composition of amino acid residues found in meteorites. Eight arbitrarily picked VADEG proteins were found to be highly soluble, with a defined oligomeric state but with a low helical content.

These results support the idea that a certain degree of foldability can be obtained from random libraries using a limited set of amino acid residues. However, in the absence of structural studies for such generated proteins, it is difficult to draw conclusions regarding the extent of folding achieved by these proteins.

9.4.2
Design of Proteins using Extended Alphabets

Extending the alphabet of amino acids that can be incorporated into proteins relies on the engineering of the genetic code [61]. This may be achieved by reassignment of codons: substitutions of amino acids by isosteric amino acids or by amino acid analogues were described. Other strategies rely on the suppression of stop codons using aminoacyl-tRNAs with anticodons complementary to stop codons. Frameshift suppression which makes use of anticodons recognizing more than three-bases long codons, defines another method for introduction of unnatural amino acids within proteins.

9.4.2.1 **By Codon Reassignment Strategies**

Use of auxotrophic strains and supplementation of amino acids analogues in the culture medium provides a means to incorporate unnatural amino acids in the proteins. Histidine analogs such as 2 or 4-fluorohistidine [62] or histidine analogues containing 1,2,3-triazole or 1,2,4-triazole [63] have been introduced into proteins *in vivo* by addition of the amino acid analogues in the culture medium devoid of histidine. In a similar strategy, an unsaturated amino acid, namely an isoleucine analogue with a carbon-carbon double bond between carbon 4 and carbon 5 was incorporated within a protein [64]. In these studies, the unnatural amino acids incorporated in proteins differ from amino acids coded in the genetic code by substitution of a hydrogen atom by a fluorine atom or by introduction of a carbon-carbon double bond within the aliphatic chain of an amino acid: these modifications can be considered as isosteric.

Non-isosteric modifications of amino acids have been also described. They rely on the use of *E. coli* strains with mutations introduced specifically within the editing domains of aminoacyl-tRNA synthetases. Single amino acid substitutions within *E. coli* valyl-tRNA synthetase were found to be sufficient for mischarging ^{Val}tRNA with cysteine and with aminobutyric acid [65]. This observation resulted from the mapping of mutations within strains selected *in vivo* for thymidilate synthase activity in the absence of thymidine. The gene encoding thymidilate synthase contained a valine codon instead of the cysteine codon at a position within the active site known to be highly conserved. Mischarging of cysteine at a valine codon then allowed isolation of *E. coli* strain with reassigned codons within the genetic code.

9.4.2.2 **By Suppression Strategies**

Suppression of stop codons or of frameshifts by aminoacyl-tRNA anticodons that recognize stop codons or four- or five-base long codons respectively, appears to be a powerful tool for selective modification of proteins with unnatural amino acids. More than 30 non-natural amino acids have been incorporated site-specifically into proteins to date [66].

Two main approaches for stop codon suppression have been described depending on whether the aminoacyl-tRNA is synthesized *in vivo* or *in vitro*.

In the first type of approach, an aminoacyl-tRNA synthetase is selected so as to prevent any cross-reaction with the cells tRNAs. The tRNA$_{CUA}$ whose anticodon recognizes the amber nonsense codon TAG was chosen so as to be specifically aminoacylated with the unnatural amino acid by the selected synthetase, and not by endogenous aminoacyl-tRNA synthetases. Extension of the genetic code of *Saccharomyces cerevisiae* with unnatural amino acids was for example described for the introduction of phenylalanine derivatives modified by azide or acetylene groups using a mutant *E. coli* tyrosyl-tRNA synthetase. Proteins having incorporated these residues can undergo further [3+2] azide-alkyne cycloaddition for selective conjugation of the protein [67]. Extension of the genetic code of *E. coli* was achieved using a *Methanococcus jannaschii* tyrosyl-tRNA synthetase variant with four mutations that were selected for charging of *p*-benzoyl-L-phenylalanine on a mutant ^{Tyr}tRNA$_{CUA}$. This unnatural amino acid was incorporated at an amber
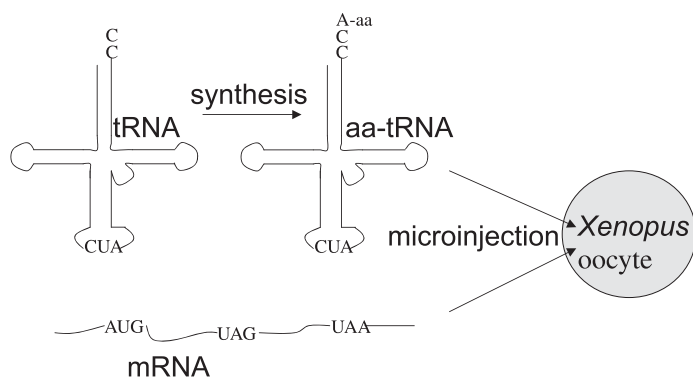
Fig. 9.8 Expression of proteins incorporating unnatural amino acids in *Xenopus* oocytes. An aminoacylated nucleotide (A-aa) that is protected can be ligated to a suppressor tRNA truncated at the last base by T4 RNA ligase. Deprotection yields an aminoacyl-tRNA linking the CUA anticodon to the unnatural amino acid aa. A messenger RNA containing the complementary UAG codon at the site of incorporation of the unnatural amino acid is synthesized *in vitro*. Microinjection of the mRNA and of the aminoacylated tRNA within *Xenopus* oocytes allows the expression of proteins incorporating the unnatural amino acid site-specifically.

codon within a tagged protein produced in *E. coli*. Photocrosslinking experiments with this amino acid may allow the identification of protein-protein interactions [68].

In the second type of approach, the aminoacyl-tRNA was synthesized *in vitro* (Fig. 9.8). A 3′ terminal amino-acylated nucleotide or dinucleotide can be synthesized chemically [69]. These nucleotides can be coupled enzymatically to a tRNA lacking the last nucleotide(s) by T4 RNA ligase. The amino-acylated suppressor tRNA$_{CUA}$ and a mRNA containing a TAG stop codon at the site of incorporation of the unnatural amino acid can then be microinjected into *Xenopus* oocytes for expression [70]. Editing of the aminoacyl-tRNA by endogenous aminoacyl-tRNA synthetases was not reported and may be negligible in this system. Site-specifically modified ion channels such as nicotinic acetylcholine receptors were then characterized by electrophysiological measurements [70]. This defines a method for the introduction of chosen chemical groups at specific sites within proteins.

Frameshift suppression provides an alternative to stop codon suppression methods. A five-base codon allowed the incorporation of nitrophenylalanine into streptavidin at position Y54 using an *E. coli in vitro* translation system. The yield of full-length protein was found to be 3.5 times higher using a five-base codon suppression than for amber codon suppression [71]. Four- and five-base codons are recognized by Watson and Crick base-pairing with extended anticodon loops of tRNAs [72].

Frameshift suppression for two four-base codons combined with amber stop codon suppression allowed the introduction of up to three unnatural amino acids within the heteromultimeric nicotinic acetylcholine receptor [73].

More generally, extended alphabets allow the synthesis of site-specifically modified proteins that could be useful for the study of the proteins' biological function. Extended alphabets should further allow the synthesis of large macromolecular assemblies such as protein conjugates derivatized with a precision approaching the atomic-level. It may also facilitate the design of future vaccines and of therapeutic agents.

9.5
Design of Proteins *de novo*

De novo design of proteins makes no assumptions on the protein sequence and relies on the power of calculations or of selections to define proteins of interest (see also Chapter 6 Section 6.4.1 on computational approaches and Chapter 5 Section 54.2 on the design and chemical synthesis of tertiary structures).

9.5.1
Computational Design of New Folds and Experimental Proofs

The use of empirical rules together with computational approaches allowed the design of a new fold, i.e. alpha-helical bundles with a right-handed superhelical structure [74]. Empirical rules were derived from the analysis of left-handed coiled-coils and from the positioning of amino acids for packing of a hydrophobic core within trimeric and tetrameric bundles. Stable conformations for trimers and tetramers were computed for multiple side chain rotamers and for different backbone conformations. The synthesized trimers and tetramers were stable structured domains as shown by the apparent melting temperatures and had the predicted oligomerization state as evaluated by analytical centrifugation. The tetramer's crystal structure fits the model with atomic level precision, except at the C-terminus which is involved in crystal contacts [74].

Computational design of a new fold was also achieved by iteration of cycles of sequence optimization and of structure prediction [75]. The predicted structure of an α/β protein of 93 amino acids was confirmed by X-ray crystallography. The remarkable stability of the designed proteins was evidenced by circular dichroism spectra which were similar at 25 °C and at 98 °C.

9.5.2
Combinatorial and Experimental Design

Finally, a real challenge for the design of new proteins by a combinatorial approach is to start from fully random proteins without introducing any constraint except their length. The sequence space that corresponds to fully random proteins of 80 amino acids in length for example, contains an astronomical number of sequences (20^{80}). Thus, the fraction of this sequence space that one can explore with the most powerful method of selection is obviously very small (for *in vitro* techniques about 10^{13} clones at best). Given that, a question immediately

arises: is there a substantial chance of finding a folded or even a functional protein in such combinatorial libraries?

To answer this question, Cho and colleagues [76] constructed a random library of about 10^{13} clones in which every position had an equal probability of encoding each amino acid. Arbitrarily, the length of proteins was fixed to 80 amino acid residues and the targeted function was ATP binding. This library was used for selection by mRNA display using immobilized ATP. After several rounds of selection, binders were isolated and showed no sequence homology with known ATP binding motifs. Their dissociation constants for ATP were in the range of 100 nM to 10 µM. Competition with ATP analogues showed that the recognition was specific for some parts of the ATP molecule. This binding was EDTA concentration dependent, and could be restored after dialysis in the presence of Zn^{2+}. One family of binders indeed contained a CXXC motif characteristic of zinc binding domains. Furthermore, a core of only 45 amino acid residues was determined to be sufficient for ATP binding. Finally the monitoring of selection led the authors to conclude that the ATP binding property probably occurs at a frequency of 10^{-11}.

Selected clones were only soluble when fused to the very soluble maltose binding protein. To improve their solubility, a pool of previously selected binders was submitted to rounds of selection in the presence of increasing concentrations of a denaturant over successive rounds. The population was thus selected for the ability to bind ATP in presence of 3 M guanidinium hydrochloride. One clone was shown to have a defined folded structure according to circular dichroism, tryptophan fluorescence and NMR studies [77].

Can this strategy lead to new folds? A structural study by X-ray crystallography of one previously isolated clone was undertaken [78]. The three dimensional structure obtained confirmed that the *in vitro* evolved protein binds a Zn^{2+} ion (confirmed by X-ray fluorescence) *via* a CXXC motif. Although the protein has no apparent sequence homolog, the binding site of ADP/ATP showed features typical of adenine binding proteins and revealed a novel α/β fold (Fig. 9.9). This
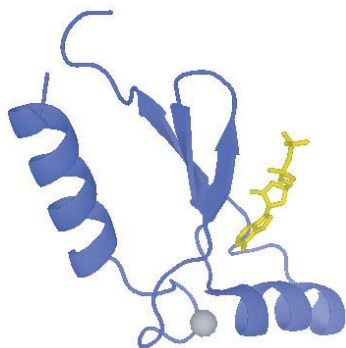
Fig. 9.9 Structure of a *de novo* designed ATP/ADP binding protein selected from a fully random protein library (PDB ID: 1UW1 [78]). Binder = blue; ADP molecule = yellow; Zn^{2+} = grey.

demonstrates that a protein with a tailored function and with the properties of natural proteins can be designed from unconstrained protein libraries.

9.6
Conclusion

We have seen that a large palette of approaches is available to the protein engineer ranging from biological chemistry, molecular modeling and structural biology to microbiology through molecular biology and biochemistry. Success in protein design may well rely on highly interdisciplinary expertise in these various fields.

From all the studies we have presented, we can draw some conclusions. Firstly, it is clear that a highly successful approach is to start from a natural protein in order to have the benefit of its foldability and stability. Practically, to achieve a drastic change in the properties of a protein, the use of predictive and combinatorial approaches in parallel happened to be very efficient. The key points are to choose a starting protein adapted to the targeted properties in terms of architecture, folding and stability, and also to have powerful screening or selection tools to identify clones of interest among the wide diversity of the generated variants. The natural repertoire of twenty amino acids of the genetic code is generally sufficient for the engineering of proteins with new functions (binding, folding, solubility and catalytic activity). The extension of the genetic code towards non-natural amino acids should prove useful for the engineering of protein conjugates defined at the atomic level. Secondly, a lesson from combinatorial approaches is that foldability of a polypeptide chain is a characteristic that can be quite easily determined in a library with minimal rational design. But, it is striking to see that even without any design, a library of fully random polypeptide chains can contain some foldable and functional proteins at a frequency accessible to our tools of selection. Thirdly, recent works using *in silico* approaches to predict and design protein structures have now reached some landmarks and it is realistic to think that in the near future this approach will become more important for the design of new proteins.

The advancements in protein design during the last fifteen years can be realized in a citation by Rainer Jaenicke [79]: ''When and whether the time is approaching when new and even useful proteins will be *de novo* designed, synthesized, and technologically applied is a question of enthusiasm and belief.'' *Ab initio* design of proteins has been mostly limited to small proteins with simple folds, whilst the design of proteins endowed with new or improved functions has generally resulted from screening or selection strategies. Major challenges in protein design continue for the 21st century. The identification of mutations that are associated with new or improved functions and that cannot be interpreted given their sequence–structure–function relationships is certainly an issue of importance. Prediction of protein functions from their primary sequences still remains a challenge and advancements in this area should also facilitate the design of proteins.

Acknowledgments

References

1 M. A. Dwyer, L. L. Looger, H. W. Hellinga, *Science* **2004**, *304*, 1967–1971.

2 H. S. Park, S. H. Nam, J. K. Lee, C. N. Yoon, B. Mannervik, S. J. Benkovic, H. S. Kim, *Science* **2006**, *311*, 535–538.

3 J. L. Jestin, S. Vichier-Guerre, *Res. Microbiology* **2005**, *156*, 961–966.

4 S. Vichier-Guerre, S. Ferris, N. Auberger, K. Mahiddine, J. L. Jestin, *Angew. Chem. Int. Ed.* **2006**, *45*, 6133–6137.

5 P. H. Patel, L. A. Loeb, *J. Biol. Chem.* **2000**, *275*, 40266–40272.

6 G. Xia, L. Chen, T. Sera, M. Fa, P. G. Schultz, F. E. Romesberg, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6597–6602.

7 a) D. Legendre, B. Vucic, V. Hougardy, A. L. Girboux, C. Henrioul, J. Van Haute, P. Soumillion, J. Fastrez, *Protein Science* **2002**, *11*, 1506–1518; b) P. Mathonet, J. Fastrez, *Curr. Opin. Struct. Biol.* **2004**, *14*, 505–511.

8 D. Legendre, P. Soumillion, J. Fastrez, *Nature Biotech.* **1999**, *17*, 67–72.

9 G. Guntas, T. J. Mansell, J. R. Kim, M. Ostermeier, *Proc. Nat. Acad. Sci. USA* **2005**, *102*, 11224–11229.

10 D. Rennell, S. E. Bouvier, L. W. Hardy, A. R. Poteete, *J. Mol. Biol.* **1991**, *222*, 67–88.

11 C. F. Barbas, 3rd, J. D. Bain, D. M. Hoekstra, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 4457–4461.

12 E. Soderlind, M. Vergeles, C. A. Borrebaeck, *Gene.* **1995**, *160*, 269–272.

13 W. P. Yang, K. Green, S. Pinz-Sweeney, A. T. Briones, D. R. Burton, C. F. Barbas, 3rd, *J. Mol. Biol.* **1995**, *254*, 392–403.

14 A. Knappik, L. Ge, A. Honegger, P. Pack, M. Fischer, G. Wellnhofer, A. Hoess, J. Wolle, A. Pluckthun, B. Virnekas, *J. Mol. Biol.* **2000**, *296*, 57–86.

15 K. Nord, J. Nilsson, B. Nilsson, M. Uhlen, P. A. Nygren, *Protein Eng.* **1995**, *8*, 601–608.

16 K. Nord, E. Gunneriusson, J. Ringdahl, S. Stahl, M. Uhlen, P. A. Nygren, *Nat. Biotechnol.* **1997**, *15*, 772–777.

17 E. Gunneriusson, K. Nord, M. Uhlen, P. Nygren, *Protein Eng.* **1999**, *12*, 873–878.

18 M. Wikman, A. C. Steffen, E. Gunneriusson, V. Tolmachev, G. P. Adams, J. Carlsson, S. Stahl, *Protein Eng. Des. Sel.* **2004**, *17*, 455–462.

19 A. Orlova, M. Magnusson, T. L. Eriksson, M. Nilsson, B. Larsson, I. Hoiden-Guthenberg, C. Widstrom, J. Carlsson, V. Tolmachev, S. Stahl, F. Y. Nilsson, *Cancer Res.* **2006**, *66*, 4339–4348.

20 E. Wahlberg, C. Lendel, M. Helgstrand, P. Allard, V. Dincbas-Renqvist, A. Hedqvist, H. Berglund, P. A. Nygren, T. Hard, *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 3185–3190.

21 C. Lendel, J. Dogan, T. Hard, *J. Mol. Biol.* **2006**, *6*, 6.

22 M. Hogbom, M. Eklund, P. A. Nygren, P. Nordlund, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3191–3196.

23 C. Lendel, V. Dincbas-Renqvist, A. Flores, E. Wahlberg, J. Dogan, P. A. Nygren, T. Hard, *Protein Sci.* **2004**, *13*, 2078–2088.

24 J. Dogan, C. Lendel, T. Hard, *J. Mol. Biol.* **2006**, *6*, 6.

25 G. Beste, F. S. Schmidt, T. Stibora, A. Skerra, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1898–1903.

26 S. Schlehuber, G. Beste, A. Skerra, *J. Mol. Biol.* **2000**, *297*, 1105–1120.

27 S. Schlehuber, A. Skerra, *Biophys. Chem.* **2002**, *96*, 213–228.

28 S. Vopel, H. Muhlbach, A. Skerra, *Biol. Chem.* **2005**, *386*, 1097–1104.

29 I. P. Korndorfer, G. Beste, A. Skerra, *Proteins* **2003**, *53*, 121–129.

30 B. Heyd, F. Pecorari, B. Collinet, E. Adjadj, M. Desmadril, P. Minard, *Biochemistry* **2003**, *42*, 5674–5683.

31 A. Drevelle, M. Graille, B. Heyd, I. Sorel, N. Ulryck, F. Pecorari, M. Desmadril, H. van Tilbeurgh, P. Minard, *J. Mol. Biol.* **2006**, *20*, 20.

32 P. Forrer, M. T. Stumpp, H. K. Binz, A. Pluckthun, *FEBS Lett.* **2003**, *539*, 2–6.

33 H. K. Binz, M. T. Stumpp, P. Forrer, P. Amstutz, A. Pluckthun, *J. Mol. Biol.* **2003**, *332*, 489–503.

34 V. S. Devi, H. K. Binz, M. T. Stumpp, A. Pluckthun, H. R. Bosshard, I. Jelesarov, *Protein Sci.* **2004**, *13*, 2864–2870.

35 A. Kohl, H. K. Binz, P. Forrer, M. T. Stumpp, A. Pluckthun, M. G. Grutter, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1700–1705.

36 H. K. Binz, P. Amstutz, A. Kohl, M. T. Stumpp, C. Briand, P. Forrer, M. G. Grutter, A. Pluckthun, *Nat. Biotechnol.* **2004**, *22*, 575–582.

37 P. Amstutz, H. K. Binz, P. Parizek, M. T. Stumpp, A. Kohl, M. G. Grutter, P. Forrer, A. Pluckthun, *J. Biol. Chem.* **2005**, *280*, 24715–24722.

38 A. Kohl, P. Amstutz, P. Parizek, H. K. Binz, C. Briand, G. Capitani, P. Forrer, A. Pluckthun, M. G. Grutter, *Structure (Camb)*. **2005**, *13*, 1131–1141.

39 a) P. Kristensen, G. Winter, *Fold. Design* **1998**, *3*, 321–328; b) V. Sieber, A. Pluckthun, F. X. Schmid, *Nature Biotech.* **1998**, *16*, 955–960.

40 J. D. Marks, H. R. Hoogenboom, A. D. Griffiths, G. Winter, *J. Biol. Chem.* **1992**, *267*, 16007–16010.

41 T. Matsuura, A. Pluckthun, *FEBS Lett.* **2003**, *539*, 24–28.

42 D. Christ, G. Winter, *J. Mol. Biol.* **2006**, *358*, 364–371.

43 L. Riechmann, G. Winter, *Proc. Nat. Acad. Sci. USA* **2000**, *97*, 10068–10073.

44 a) S. de Bono, L. Riechmann, E. Girard, R. L. Williams, G. Winter, *Proc. Nat. Acad. Sci. USA* **2005**, *102*, 1396–1401; b) L. Riechmann, I. Lavenir, S. de Bono, G. Winter, *J. Mol. Biol.* **2005**, *348*, 1261–1272.

45 G. MacBeath, P. Kast, D. Hilvert, *Science* **1998**, *279*, 1958–1961.

46 M. Bashton, C. Chothia, *J. Mol. Biol.* **2002**, *315*, 927–939.

47 Y. Wang, D. E. Prosen, L. Mei, J. C. Sullivan, M. Finney, P. B. Vander Horn, *Nucleic Acids Res.* **2004**, *32*, 1197–1207.

48 a) H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242; b) A. Kouranov, L. Xie, J. de la Cruz, L. Chen, J. Westbrook, P. E. Bourne, H. M. Berman, *Nucleic Acids Res.* **2006**, *34*, D302–305.

49 C. A. Orengo, J. M. Thornton, *Annu. Rev. Biochem.* **2005**, *74*, 867–900.

50 A. Andreeva, D. Howorth, S. E. Brenner, T. J. Hubbard, C. Chothia, A. G. Murzin, *Nucleic Acids Res.* **2004**, *32*, D226–229.

51 S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science*. **1993**, *262*, 1680–1685.

52 S. Roy, K. J. Helmer, M. H. Hecht, *Fold. Des.* **1997**, *2*, 89–92.

53 S. Roy, M. H. Hecht, *Biochemistry*. **2000**, *39*, 4603–4607.

54 D. M. Rosenbaum, S. Roy, M. H. Hecht, *J. Am. Chem. Soc.* **1999**, *121*, 9509–9513.

55 Y. Wei, T. Liu, S. L. Sazinsky, D. A. Moffet, I. Pelczer, M. H. Hecht, *Protein Sci.* **2003**, *12*, 92–102.

56 Y. Wei, S. Kim, D. Fela, J. Baum, M. H. Hecht, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13270–13273.

57 T. Matsuura, A. Ernst, A. Plückthun, *Protein Sci.* **2002**, *11*, 2631–2643.

58 A. R. Davidson, R. T. Sauer, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2146–2150.

59 A. R. Davidson, K. J. Lumb, R. T. Sauer, *Nat. Struct. Biol.* **1995**, *2*, 856–864.

60 N. Doi, K. Kakukawa, Y. Oishi, H. Yanagawa, *Protein Eng. Des. Sel.* **2005**, *18*, 279–284.

61 N. Budisa, *Angew. Chem. Inter. Ed.* **2004**, *43*, 6426–6463.

62 J. F. Eichler, J. C. Cramer, K. L. Kirk, J. G. Bann, *Chembiochem* **2005**, *6*, 2170–2173.

63 Y. Ikeda, S. Kawahara, M. Taki, A. Kuno, T. Hasegawa, K. Taira, *Protein Eng.* **2003**, *16*, 699–706.

64 M. L. Mock, T. Michon, J. C. M. van Hest, D. A. Tirrell, *Chembiochem* **2006**, *7*, 83–87.

65 V. Döring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. De Crécy-Lagard, P. Schimmel, P. Marlière, *Science* **2001**, *292*, 501–504.

66 J. M. Xie, P. G. Schultz, *Curr. Opin. Chem. Biol.* **2005**, *9*, 548–554.

67 A. Deiters, T. A. Cropp, M. Mukherji, J. W. Chin, J. C. Anderson, P. G. Schultz, *J. Amer. Chem. Soc.* **2003**, *125*, 11782–11783.

68 J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *Proc. Nat. Acad. Sci. USA* **2002**, *99*, 11020–11024.

69 a) S. M. Hecht, B. L. Alford, Y. Kuroda, S. Kitano, *J. Biol. Chem.* **1978**, *253*, 4517–4520; b) C. H. Rohrig, O. A. Retz, L. Hareng, T. Hartung, R. R. Schmidt, *Chembiochem* **2005**, *6*, 1805–1816.

70 M. W. Nowak, J. P. Gallivan, S. K. Silverman, C. G. Labarca, D. A. Dougherty, H. A. Lester, in *Ion Channels, Pt B, Vol. 293*, **1998**, pp. 504–529.

71 T. Hohsaka, Y. Ashizuka, H. Murakami, M. Sisido, *Nucleic Acids Res.* **2001**, *29*, 3646–3651.

72 J. C. Anderson, T. J. Magliery, P. G. Schultz, *Chem. Biol.* **2002**, *9*, 237–244.

73 E. A. Rodriguez, H. A. Lester, D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8650–8655.

74 P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **1998**, *282*, 1462–1467.

75 B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, *Science* **2003**, *302*, 1364–1368.

76 G. Cho, A. D. Keefe, R. Liu, D. S. Wilson, J. W. Szostak, *J. Mol. Biol.* **2000**, *297*, 309–319.

77 J. C. Chaput, J. W. Szostak, *Chem. Biol.* **2004**, *11*, 865–874.

78 P. Lo Surdo, M. A. Walsh, M. Sollazzo, *Nat. Struct. Mol. Biol.* **2004**, *11*, 382–383.

79 R. Jaenicke, *Biochemistry* **1991**, *30*, 3147–3161.

80 J. L. Jestin, P. A. Kaminski, *J. Biotech.* **2004**, *113*, 85–103.

10

Nucleic Acid Foldamers: Design, Engineering and Selection of Programmable Biomaterials with Recognition, Catalytic and Self-assembly Properties[1]

Arkadiusz Chworos and Luc Jaeger

10.1
Introduction

The remarkable complexity of living organisms essentially relies on proteins, RNA and DNA, that carry out most of the major cellular functions necessary for life. These biopolymers rely on two basic self-assembly processes: the spontaneous folding of one polymer chain into a stable, well-defined three-dimensional (3-D) structure, and the assembly of multiple subunits into well-defined modular supramolecular architectures. Key characteristics are (i) sequence heterogeneity, (ii) hierarchical organization of conformation (secondary structure versus tertiary structure), (iii) modular components, (iv) stereochemically specific and selective interactions and (v) cooperativity of folding. Proteins and nucleic acids are thus perfect prototypes of functional, biological foldamers.

In the past 15 years, the development of *in vitro* selection and evolution techniques and progress in the rational design of nucleic acids has led to an incredible new world of novel functional nucleic acids capable of exquisite recognition, and having catalytic and responsive properties. Despite the fact that these artificial molecules are typically of greater size than synthetic foldamers, such as those described in Chapters 1–5, and that they have not been described as such in the literature, they can be seen as nucleic acid foldamers.

In the course of this chapter, nucleic acid foldamers are defined as non-natural nucleic acid polymers able to fold into well-defined 3-D shapes with recognition, catalytic and/or assembly properties. By non-natural or artificial nucleic acids, we want to express (i) that they are the product of combinatorial or/and rational synthetic and supramolecular design and (ii) that, besides RNA and DNA, they can be based on analogs that mimic nucleic acid.

This chapter focuses on DNA and RNA foldamers rather than nucleic acid biomimetics and analogs. It aims to provide the reader with a broad outline of the various concepts behind the nanoconstruction of functional nucleic acid based

---

1) A list of abbreviations appears at the end of this chapter.

materials. First we give an overview of the structural and functional properties of RNA and DNA foldamers. Then we emphasize how nucleic acid sequences can be synthesized, engineered and controlled to sculpt new 3-D molecular architectures with catalytic, responsive and supramolecular properties. After a description of the various self-assembly strategies used to direct the assembly of foldamers into nanostructures of increasing complexity, we finally address how RNA and DNA foldamers can potentially lead to the development of novel biomaterials for electronics or biomedical applications.

By exemplifying our present ability to program linear polymer sequences to fold and self-assemble into defined 3-D shapes, nucleic acid foldamers offer a great source of inspiration to the supramolecular chemist [1, 2]. It is our hope that they will pave the way to the creation of useful complex materials based on novel fully synthetic biomimetic foldamers that will be chemically more robust, cheaper and easier to obtain than nucleic acids.

10.2
Principles of Nucleic Acid Foldamers

While proteins are essentially structural components, catalytic or regulatory tools that sustain metabolic pathways, DNA is the stable molecular recipient of the genetic information. RNA molecules have more ambivalent functions, as they can serve as catalytic and regulatory tools as well as the support of the genetic information, more often as transient copies of DNA genes rather than as full genomes, however. These various molecular functions may be considered to be the result of historical evolutionary contingencies, but they are also the direct consequence of the unique biophysico-chemical properties pertaining to each of these classes of polymers.

10.2.1
Structural Principles: Hierarchical Organization and Modularity

RNA and DNA modularity is hierarchically expressed at a chemical, structural and supramolecular level (Fig. 10.1).

10.2.1.1 Chemical Modularity and Stability
From a chemical point of view, RNA and DNA are modular polymers with primary ($1°$) sequences formed of four basic building blocks: A, U, C and G ribonucleotides for RNA and A, T, C and G deoxyribonucleotides for DNA. In cells, deoxyribonucleotides are biologically synthesized by enzymatic modification of RNA ribonucleotides. DNA can be seen as a modified RNA that lacks the $2'$-hydroxyl at the level of the sugar moiety and has an extra methyl at the level of uracil to form thymine. These minor chemical differences increase the resistance of DNA towards spontaneous hydrolysis, and the fidelity of DNA replication, making DNA a better support for the genetic information than RNA. At basic pH values (7–12),
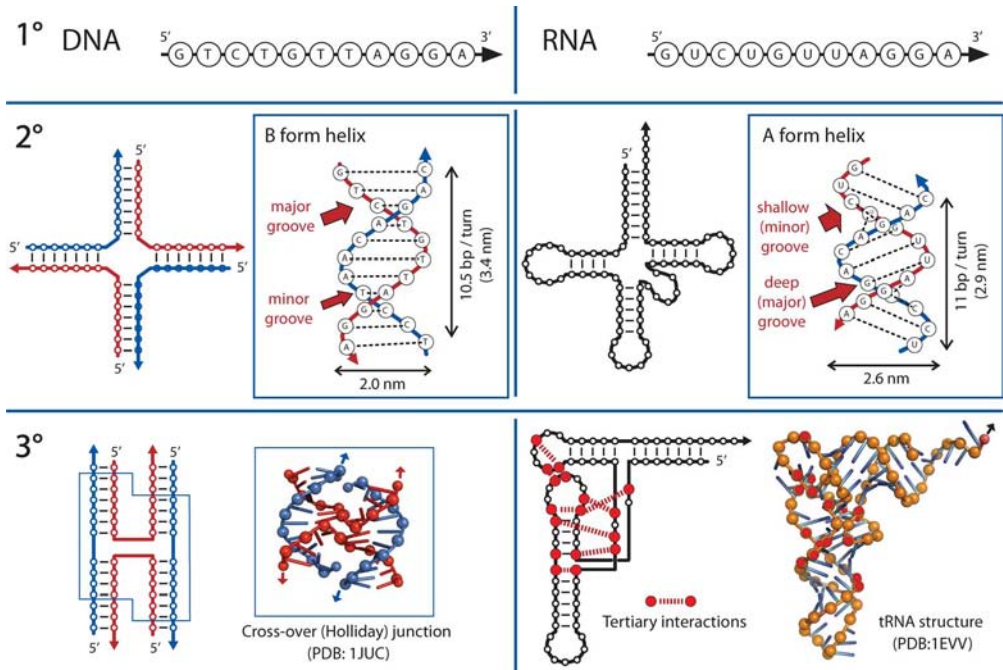
Fig. 10.1 Supramolecular and structural modularity of nucleic acids. DNA and RNA chemical modularity is exemplified by their primary sequence (1°). The structural and conformational modularity of nucleic acids is expressed at the level of their secondary (2°) and tertiary (3°) structures.

RNA backbone hydrolyzes in the presence of divalent ions such as magnesium. Additionally, mutational events resulting from the spontaneous hydrolysis of cytosine into uracil cannot be repaired in RNA as uracils at mutated positions cannot be distinguished from those at non-mutated positions. At acidic pH values (4–6), however, depurination is faster for DNA than RNA [3, 4]. In living organisms, the greater chemical fragility of RNA compared with DNA essentially results from the numerous ribonucleases that biologically target RNA molecules. In a laboratory setting, this can however be easily overcome by taking basic "RNase-free" precautions.

For biotechnological and medical applications, RNA can be a good medium to build transient, biodegradable materials or molecular scaffolds for RNA drugs such as aptamers, ribozymes or siRNAs [5–7] (Section 10.2.2). Alternatively, the combination of RNA moieties with DNA and other nucleic acid analogs can offer limitless possibilities for improving and tuning the chemical and thermodynamic stability of nucleic acid foldamers that would retain the unique structural richness and thermodynamic stability of RNA (Section 10.3).

10.2.1.2 **Secondary Structure Principles**

In contrast to proteins, the secondary (2°) structure of RNA and DNA results from hydrogen bonding between side chains and not between backbone atoms, as for protein alpha-helices and beta-sheets. Therefore, nucleic acid 2° structures are easier to predict than those of proteins. Within the hierarchical framework that characterizes nucleic acid folding and assembly, π-stacking and Watson–Crick base pairings drive the folding and assembly of RNA and DNA 1° sequences into 2° structures through the formation of stable helical elements. Watson–Crick base pairs, with *cis*-glycosyl bonds, form the only set of pairs that are isosteric in anti-parallel helices. Thus, they allow formation of helices with regular sugar–phosphate backbones. RNA and DNA 2° structures are schematically represented in a planar drawing by base-paired segments that specify for various 2° structure motifs such as hairpin loops, bulges, internal loops and multi-helix junctions (Fig. 10.1).

The presence of the 2′-OH in RNA increases the structural rigidity of RNA duplexes that are locked into compact A-form helices with C3′ endo sugar pucker. More polymorphous DNA helices are mostly present in the extended B-form with C2′ endo sugar pucker (Fig. 10.1). As basic modular building blocks, A-form RNA duplexes are thermodynamically more stable than B-form DNA duplexes. According to the base-pair free-energy parameters determined at 1 M NaCl and 37° for RNA and DNA, RNA base pairs are, on average, -0.49 ± 0.35 kcal mol^{-1} more stable than those of DNA [8, 9]. Note, however, that the thermodynamic stability of RNA and DNA duplexes varies as a function of the nucleic acid sequence. The higher thermodynamic stability of RNA duplexes compared with that of DNA duplexes essentially results from a higher enthalpy for duplex formation for RNA, which is consistent with much better hydration of RNA helices than DNA helices [10]. Recently, experimental measurements of the persistence length for RNA and DNA duplexes was performed by single molecule analysis using Fluorescence Resonance Energy Transfer (FRET), Atomic Force Microscopy (AFM) and magnetic tweezers techniques [11, 12]. They corroborate that RNA helices are more compact and stiffer than DNA helices, with persistence length of 55 nm and 63 nm for DNA and RNA, respectively. As the rise per helical turn is 2.9 nm for RNA versus 3.4 nm for DNA, the persistence length calculated in base pairs is 30% greater for RNA than DNA.

Besides the classic Watson–Crick pairs, eleven distinctive "non-canonical" base pairs, which involve at least two hydrogen bonds, can potentially occur between complementary nucleotides [13]. In RNA, they can contribute significantly to the rigidity and thermodynamic stability of RNA structural elements [14]. Design and prediction of RNA and DNA 2° structures can presently be achieved by energy minimization with a reasonable degree of accuracy [14, 15] using software like mfold, RNAfold or RNAsoft [16–19]. Because the formation of mismatches is allowed between strands that are not perfectly complementary, RNA helices have a lower selective informational content than their DNA counterparts. Consequently, for RNA, positive and negative design is particularly critical to maximize the stability of the desired 2° structure while minimizing folding into stable alternatives.

10.2.1.3 Tertiary Structure Principles

Entropic gain through water and ion release is the main driving force for the tertiary (3°) folding of a protein. It leads to a hydrophobic core mostly devoid of water molecules. By contrast, for nucleic acids, both entropy gain through ion and water release and enthalpy gain via the formation of intricate solvent networks lead to the 3° folding mostly devoid of hydrophobic pockets. At the 3° structure level, the 2° structural elements can associate through numerous van der Waals contacts, π-stacking, metal coordination and specific hydrogen bonds via the formation of a small number of additional Watson–Crick and/or non-canonical base pairs that involve single-stranded regions, loops or bulges. For 3° nucleic acid structures, such as large stable RNA molecules, the assembly occurs by metal–ion induced collapse of the 2° structure into compact conformations (reviewed in Refs. [20–22]) (Fig. 10.2). Metal ions essentially screen the negative charges on the phosphate groups. The loosely folded intermediates then undergo further conformational rearrangements before adopting the final 3° structure [22] (Fig. 10.2). This conformational search is essentially dependent on the local folding of recurrent and specific set of nucleotides that specify for modular 3° structure motifs that adopt unique local 3-D shapes and mediate stereochemically precise quaternary (4°) interactions (Fig. 10.3). These structure motifs can be seen as the minimal information for folding a nucleic acid sequence into a specific 2° or 3° structure. As such, they can be seen as basic foldamer modules that can be combined and encoded within a nucleic acid sequence to specify for more complex 3-D shapes. The proper folding of 3° structure motifs requires the formation of 3° interactions between specific nucleotide positions, which are highly dependent on temperature, salts and divalent ion concentration. For instance, without
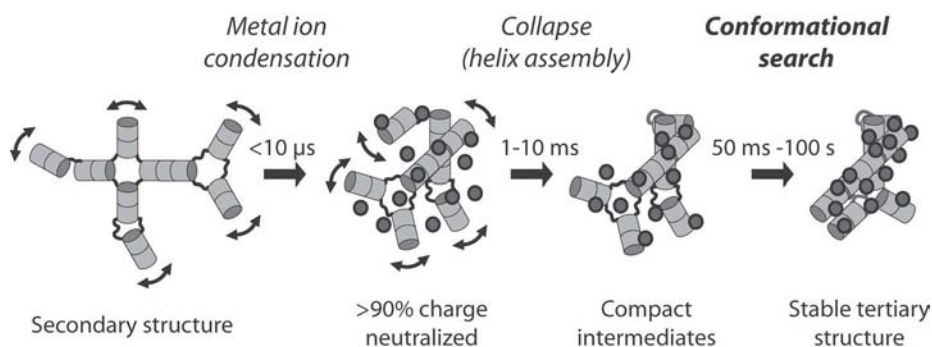
Fig. 10.2 Metal ion-induced folding of RNA. The association of the positive ions (grey spheres) with the unfolded RNA rapidly neutralizes more than 90% of the phosphate charges and induces the collapse of the RNA into more compact conformations. The conformational search, which is the time-limiting step, is dependent on the formation of 3° structure motifs. Cylinders symbolize helical elements. Double arrows indicate helical motions. The diagram is based on Fig. 2 from Ref. [20].
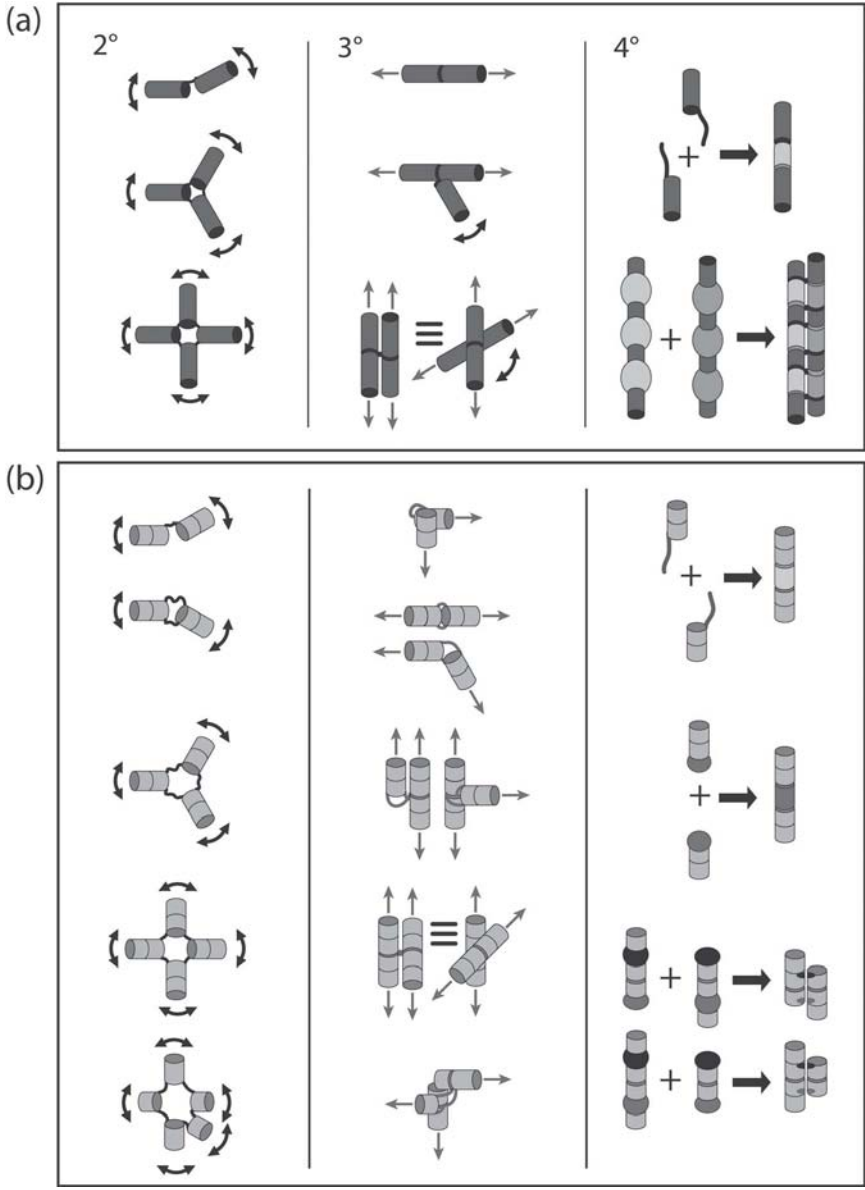
divalent metals such as magnesium, RNA 3° structure motifs often behave as flexible, dynamic 2° structure motifs. However, in the presence of magnesium, they cooperatively fold into stable and unique conformers that can specify a receptor, catalytic site or a precise geometry of helical elements.

In Nature, DNA is essentially found as a long, helical rod and does not seem to be used for purposes other than to carry genetic information. Therefore, our knowledge of 3° DNA is currently very limited and only few natural 3° structure motifs of DNA are known (Fig. 10.3a). As contiguous helices tend to stack on the top of each other, it is possible to constrain the geometry of DNA helical elements to form 3-way or 4-way junctions. In DNA 3-way junctions, one of the helices is generally more flexible than the other two [23, 24]. The quintessential DNA 3° structure motif is the 4-way (Holliday) junction [25] (Fig. 10.1), a motif that forms transiently during recombination and replication cellular processes. Two other classes of 3° DNA motifs are triple helices [26] and guanine quartets or G-tetrads [27]. G-tetrads naturally occur at the level of DNA telomeres [28] and are often found as structural scaffolds in artificial DNA aptamers [29] (Section 10.2.2.1). In the future, novel 3° structure motifs will probably be revealed by NMR or X-ray crystallography of currently available DNA aptamers and enzymes. These structural studies will offer a new wealth of potentially functional foldamer modules for DNA nanoconstruction. Nevertheless, they might not be absolutely required as the remarkable base-pairing selectivity of DNA is particularly well suited to building self-assembling architectures that predominantly rely on the selective formation of 2° structure elements between multiple complementary oligonucleotide strands (Fig. 10.1). This is described further in the later sections of this chapter.

By contrast, the architectural potential of RNA relies more on the ability of an RNA single strand to fold into an exquisite, stable 3° structure [30, 31]. The 3-D

Fig. 10.3 Structural principles of RNA and DNA supramolecular building blocks. Various examples of 2°, 3° and 4° structure motifs commonly used in DNA (a) and RNA (b) nanostructures. Small single arrows specify the geometry adopted by helical elements. (a) Contiguous helices tend to stack on the top of each other. In a DNA 3-way junction, one of the helices is generally more flexible than the other two. DNA 4-way junctions also present some degrees of flexibility [25]. For 4° assembly, DNA units can be joined through noncovalent sticky-tails connectors [159] (top right) or through internal loop–loop interactions that fold into paranemic crossover junctions (PX) that do not interpenetrate topologically [170] (bottom right); (b) RNA 2° structure elements can form flexible hinges at the level of single-stranded regions, internal loops or multi-way junctions (left). Specific set of nucleotides can also direct the formation of rigid 3° structure motifs with distinct helical geometry [32–34]. (Center panel, top to bottom): the ''right angle'' motif [35]; the internal loop E and kink turn motifs; two distinct 3-way junctions that specify for different helical geometries [38]; the 4 way-junction motif from the hairpin ribozyme [104, 176]; the class 2 tRNA 5-helix junction motif. RNA 4° interactions (Right panel) used for generating supramolecular assembling interfaces are (from top to bottom): tails connector [35], loop–loop (''kissing'') interactions [35, 180] and the double GNRA loop–receptor interaction [103].

structures of natural RNA molecules can be seen as mosaics of recurrent and modular 3° structure motifs [30]. Recently, a rich treasure-trove of structural motifs has been identified and compiled by data mining of known NMR and crystallographic atomic structures of RNA [32–34]. They specify a precise geometry of helical elements and can mediate stereochemically precise and readily reversible 3° and 4° interactions (Fig. 10.3b). Among them are single-strand junctions like the U turn, ''hook'' turn and the right-angle motif [35]; terminal loops like thermostable GNRA and UNCG loops, T-loops, internal loops such as loop E and loop C [36, 37]; the kink turn [37]; and different classes of 3 and 4 way junctions motifs [38]; pseudoknots, kissing-loops and loop–receptor motifs (for a complete survey of structure motifs see Refs. [32, 34]). Interestingly, RNA motifs make an extensive use of the 2′-OH group to form specific 3° contacts and numerous long-range RNA–RNA interactions take advantage of the 2′-OH to create compact 3-D RNA structures. Thus, by relying on different structural features, RNA and DNA motifs are structurally different at a 3° structure level. Some 3° motifs, like G-tetrads and triple helices can however form with either RNA or DNA.

Rather than relying solely on 2° structural elements, the structure of an RNA can be engineered at a three-dimensional level by encoding the structural information corresponding to rigid 3° structural motifs within its sequence. Additionally, as thermodynamically stable structural entities, larger RNA domains or full molecules such as the P4-P6 domain of group I ribozyme, the tRNA motif, natural riboswitches and RNA enzymes can themselves be used as scaffolds to engineer new artificial architectures (Sections 2.2.3 and 4). The separation of energy levels between 2° and 3° structures of RNA is distinct for stable natural RNAs, with 2° structure elements being more stable than 3° elements [39, 40]. For a complex RNA object, the dependence of the 3° structure on the presence of the extended and correct 2° structure might therefore be a necessity to avoid kinetically trapped misfolded states.

10.2.1.4 Quaternary Structure Principles

At a quaternary (4°) structure level, RNA and DNA modular units can assemble further into complex and highly modular supramolecular architectures in a predictable manner by using base-pair rules or specific, selective non-Watson–Crick interactions as organizational instructions (Fig. 10.3). The dimensionality of these nanostructures is directly related to the shape, geometry, orientation and number of assembling interfaces present at the level of their constitutive building blocks (Fig. 10.4). Brucale and colleagues proposed an interesting classification of nucleic acid nanostructures according to their topology and dimensionality [41]. Objects of dimensionality zero (**0-D**), which mathematically correspond to a point, are supramolecular architectures of finite size that can best be described as nonreducible modular tiles. These tiles can be modular but formed of distinct nonrepetitive units. Objects of dimensionality one (**1-D**) are made of units with at least two interfaces leading to growth into one direction (Fig. 10.4). Dimensionality two (**2-D**) is based on a rectangular coordinate system commonly defined by two perpendicular axes within a plane. 2-D assemblies require at least three inter-
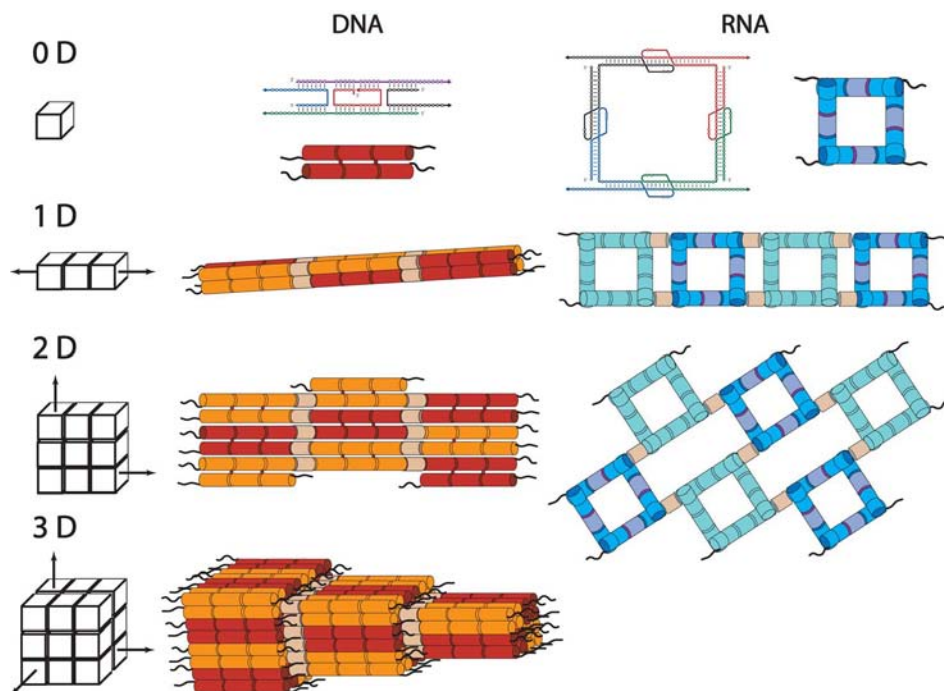
Fig. 10.4 Supramolecular dimensionalities of nucleic acid architectures. Modularity is expressed at the supramolecular level: RNA and DNA units can be engineered to assemble into nanostructures of different dimensionalities [41]. The dimensionality of supramolecular objects can be defined in term of interconnections, spatial arrangement of constitutive units and directional assembly growth.

faces to allow the assembly to grow at least in two directions that are all circumscribed within a plane. Dimensionality three is defined by a **3-D** coordinate system that provides the three physical dimensions of space: height, width, and length. 3-D objects are characterized by units with at least four nonplanar assembling interfaces that allow the assembly growth within and out of the plane in the Cartesian space. Kinetic motion is sometimes described as a fourth dimension (**4-D**).

10.2.2
Functional Principles: Recognition, Switches and Catalysis

Although the range of activities of natural nucleic acids seems limited when compared with that of proteins, a remarkable range of nucleic acid functions (Fig. 10.5) have been unraveled by *in vitro* evolution techniques such as SELEX (Section 10.4) and by the recent developments in molecular and cellular biology [40].
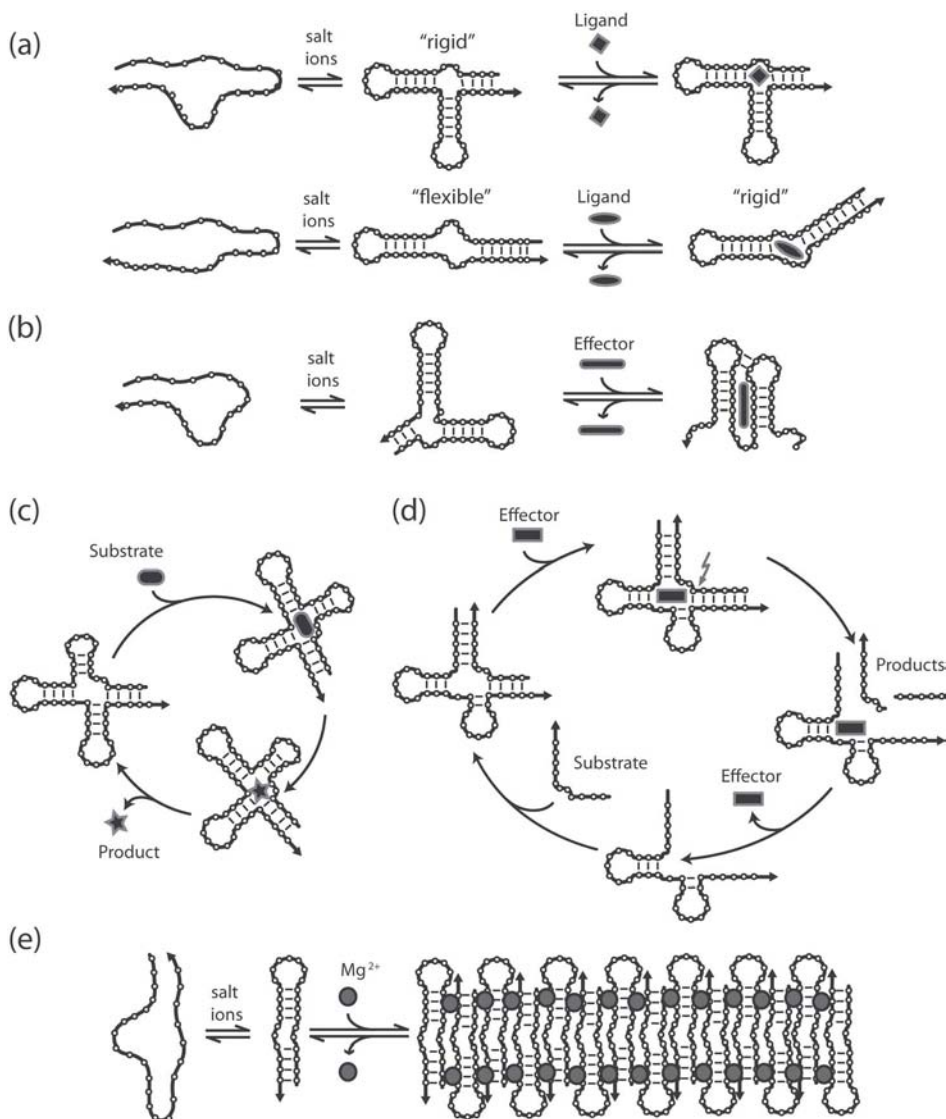
Fig. 10.5 Functional principles of nucleic acid foldamers. All single-stranded RNA and DNA foldamers folds in presence of salt ions. (a) Aptamers: RNA or DNA foldamers able to specifically recognize and bind a ligand. In the absence of ligand, the aptamer can be either flexible or rigid; (b) Riboswitch: RNA foldamer able to switch from one conformation to another upon binding of a ligand effector; (c) Ribozyme or DNAzyme: RNA or DNA foldamer can catalyze chemical reactions with multiple turnovers; (d) Aptazyme (allosteric ribozyme or DNAzyme): responsive catalytic foldamer that is activated by the presence (or absence) of a molecular effector. (e) Self-assembling DNA tiles or tectoRNAs: foldamers that assemble into stable nanostructures in the presence of salts and divalent ions (grey spheres) (Section 10.7).

10.2.2.1 Aptamers and Nucleic Acid Switches

RNA and DNA sequences called aptamers (Fig. 10.5), can act as receptors for a limitless number of ligands such as small bio-organic and synthetic compounds, ions, peptides, proteins, polysaccharides, lipid membranes and mineral surfaces (see reviews [7, 29, 40]). They bind their respective ligands with remarkable specificity and equilibrium constants of dissociation ranging from the micromolar range to the femtomolar range [42, 43]. The specific sequence signature of an aptamer can be surprisingly small and can vary from 10 to 100 nucleotides. In the absence of ligands, an aptamer sequence can readily fold into a rigid, stable 3° structure (Fig. 10.5a). However, it can also adopt a "flexible" metastable conformational state that is cooperatively stabilized into a unique "rigid" structure by induced fit mechanism upon ligand binding (Fig. 10.5a). Several artificial and natural RNA aptamers acting as switches or riboswitches have been found to exploit this property [44–46] (Figures 10.5(b) and 10.6(a)). Natural riboswitches are found to regulate the expression of genes by either activating or inhibiting the transcription or translation upon binding of a small molecular effector [46–49]. These RNA domains are generally able to adopt two distinct alternative conforma-
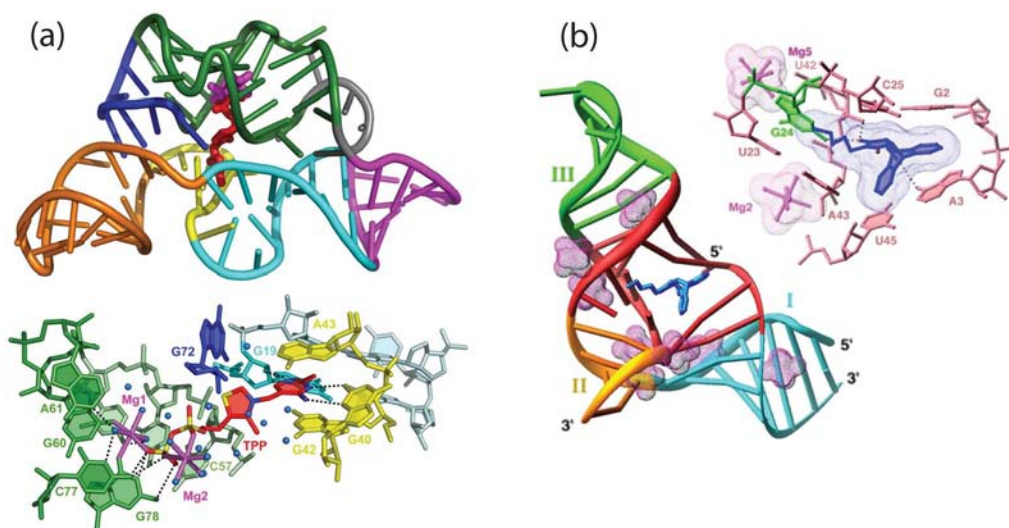
Fig. 10.6 Three-dimensional structures of a natural riboswitch and artificial ribozyme. (a) 3-D crystallographic structure of the thiamine pyrophosphate (TPP) sensing riboswitch: a natural riboswitch involved in gene regulation in bacteria [53]. The RNA bases and bound TPP (red) are shown as cylinders and the backbone is depicted with a ribbon. At the bottom left, detailed view of the TTP binding domain; (b) 3-D crystallographic structure of the Diels–Alder ribozyme: an artificial ribozyme that catalyzes carbon–carbon bond formation between anthracene and *N*-pentyl maleimide. The RNA bases and bound product (blue) are shown as cylinders. The backbone is depicted with a ribbon, whereas the hydrated Mg^{2+} are in a mesh representation [219]. On the right, detailed view of the catalytic site with the bound product. No Mg^{2+} seems to be directly involved in catalysis. Adapted with permission from Refs. [53, 219].

tions that are in equilibrium, one of the conformations being favored upon binding of a small target compound. The conformation that binds the target is generally metastable in its absence, allowing another thermodynamically more favored conformation to occur: the free energy change between the two conformational states is small and depends on few key 3° contacts directly involving ligand binding [50, 51]. The NMR and X-ray structures of several aptamers and riboswitches are presently available and shed light on the molecular recognition characteristic of these molecules [29, 51–56] (Fig. 10.6a). DNA switches have not yet been identified in Nature, but there is no conceptual reason why they should not be engineered.

10.2.2.2 Ribozymes and DNAzymes

Since the discovery of the first catalytic RNA molecules (ribozymes) in the early 1980s, RNA has been shown to exhibit a large repertoire of catalytic functions [57]. This has been extensively reviewed recently [58–60] and will only be described briefly. Ribozymes can efficiently achieve catalysis by bringing the reactive groups close to each other via specific binding, by precise orientation of the reactive groups and by structural complementarity to the substrate transition state (Figures 10.5(c) and 10.6(b)). Ribozymes are often known as metalloenzymes, although they may not directly involve divalent ions in RNA catalysis (Fig. 10.6b). They can also perform general acid/base and covalent catalysis. Besides reactions at phosphoryl centers, RNA is able to catalyze the formation of esters, amides, glycosidic and carbon–carbon bonds as well as alkylation, isomerization, metalation, peroxidation, oxido-reduction and aldol reactions [60, 61]. Recently, pyridyl-modified RNA sequences isolated by *in vitro* selection were found to catalyze the growth of palladium nanocrystals in short reaction times (∼1 min) and with a high degree of shape specificity [62–64], suggesting that RNA can actively take part in the evolution of inorganic materials. The mechanism and exquisite detail of the 3° topology and catalytic site of several ribozymes has recently been revealed by X-ray crystallography [58, 59, 65–68].

The structural diversity of DNA aptamers suggests that DNA can form many of the same secondary structures that are exhibited by ribozymes. Despite the lack of 2′-OH groups, there is now compelling evidence that DNA can also efficiently catalyze various chemical reactions [69–71]. The present scope of reactions catalyzed by DNA enzymes (DNAzymes) is somewhat less impressive than that of ribozymes. Nevertheless, beside reactions of phosphorylation, adenylation, ligation or cleavage occurring on phosphoryl centers, some DNAzymes catalyze more exotic reactions such as peroxidation, porphyrin metalation, DNA depurination and thymine dimer photoreversion [71]. Interestingly, DNA higher order structures do not automatically require the presence of divalent ions, and some deoxyribozymes rely only on potassium ions, which are known to promote the formation of G-tetrads [72].

10.2.2.3 Multifunctional Nucleic Acid Foldamers

Strikingly, most of the new catalytic functions isolated by *in vitro* selection do not require large structural motifs. RNA pools containing fewer than 100 random po-

sitions (N100 random pool) in their sequences are sufficient for finding ribozymes with very different catalytic functions [57]. Longer random pools are generally required to select more complex RNA molecules, however. The degree of functional complexity reached by an RNA molecule apparently correlates with its structural complexity [73]. Thus, the longer the sequence of the random pool, the better is the probability of selecting complex and interesting ribozymes. Moreover, only zeptomoles of nucleic acid might be necessary to find small functional modules within large random pools [74]. Another interesting aspect is that RNA molecules composed of several separate and equally sized modules might have a strong selective advantage compared with larger unique structures [74].

In view of the modular organization of large RNA molecules [30], it is possible to take advantage of a known RNA domain that displays a specific function, such as substrate binding or catalysis, to select multifunctional ribozymes. The assumption is that some molecules selected from an RNA pool which consists of a known RNA module associated with a random domain, will combine the properties of the constant RNA module to a new function. This approach has been successful for the isolation of highly complex template-directed, sequence-independent RNA ligases from a pool of 10^{16} molecules consisting of a pre-existing structural scaffold, appended to random RNA segments [75]. Similar modular approaches were applied to generate a ribozyme that is able to polymerize any RNA sequences up to 14 nucleotides by RNA-template primer extension [76], and a bifunctional ribozyme that can recognize an activated glutaminyl ester and subsequently amino-acylate a tRNA molecule [77]. Interestingly, structurally and functionally complex ribozymes can be isolated from libraries formed of stable structural modules associated with random regions of very limited size (N30 random positions) [78, 79]. Combination of functional motifs was also used to generate complex molecules with dual activities such as RNA cleavage and ligation [80], and allosteric ribozymes [81–84]. The later are also called aptazymes because they result from the combination of an aptamer joined to a catalytic domain by a communication module (Fig. 10.5d). Hammerhead self-cleaving aptazymes have been trained by *in vitro* evolution to switch their activity on or off with remarkable allosteric responses that are orders of magnitude greater than those typically seen for protein enzymes [85]. DNA aptazymes have also been engineered [86]. Recently, bifunctional RNA molecules combining binding and catalytic activities were identified from random pools by a new two-step selection method [87].

In the future, it is likely that novel ligand-responsive RNA self-assemblies will be generated likewise by taking advantages of artificial or natural riboswitches (Fig. 10.5e).

10.3
Synthesis of Nucleic Acid Foldamers and Analogs

RNA and DNA foldamers can be obtained by either chemical or enzymatic synthesis. DNA oligonucleotides of up to 120 nucleotides can presently be synthesized by phosphoramidite technology with reasonable yield [88]. However, the
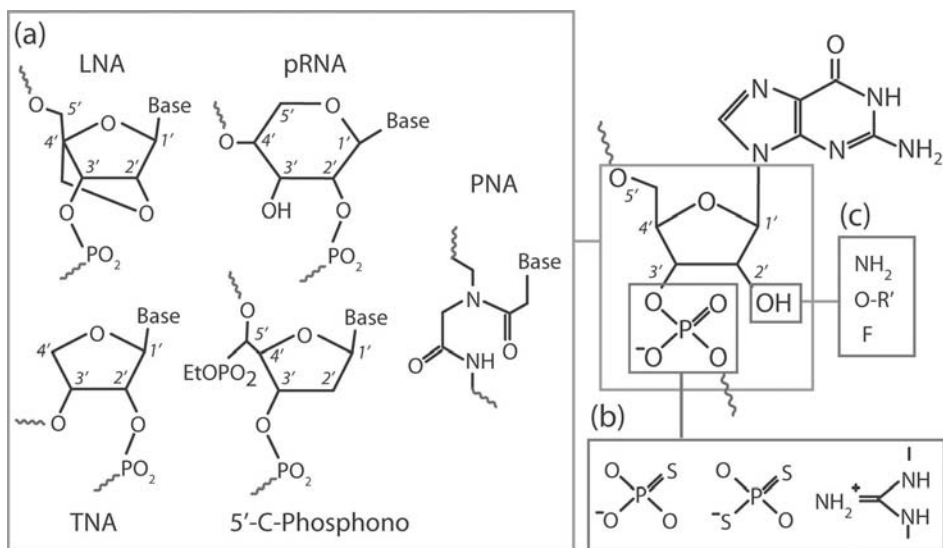
Fig. 10.7 Examples of nucleotidomimetic foldamers. (a) Examples of carbohydrate-phosphate backbone modifications: LNA (Locked-in Nucleic Acid) [90], pRNA (pyranosyl RNA) [91], TNA (alpha-Threofuranosyl Nucleic Acid) [92], 5′-C-phosphononucleotide [93], PNA (Peptide Nucleic Acid) [94, 95]; (b) Examples of nucleotide linkage modifications: phosphorothioate and phosphorodithioate linkages [96] and cationic guanidine backbone modification (deoxynucleic and ribonucleic guanidine) [97, 98]; (c) Examples of modifications at the 2′-OH position: 2′-O-Alkyl [10], 2′-amino [129] and 2′-fluoro [220, 221] modifications. These modifications have been incorporated in RNA and DNA aptamers (e.g. Refs. [221, 222]).

lower coupling efficiency for RNA phosphoramidite synthons limits the length of single-stranded RNA foldamers to 45–50 nucleotides. Similar synthetic approaches have also been used to generate interesting nucleotidomimetic foldamers [89] with modified carbohydrates nucleotides such as 2′-O-Alkylated RNA [10], LNA (Locked-in Nucleic Acid) [90], pRNA (pyranosyl RNA) [91], TNA (alpha-Threofuranosyl Nucleic Acid) [92] and 5′-C-phosphono oligonucleotide [93], or modified backbone linkage like PNA (Peptide Nucleic Acid) [94, 95], phosphorothiates oligonucleotides [96] and cationic DNG and RNG analogs (deoxynucleic and ribonucleic guanidine) [97, 98] (Fig. 10.7). For an exhaustive review of nucleotido-mimetic foldamers before 2001, see the review by Moore and colleagues [89]. The advantage of the synthetic approach is that it can be used to create DNA/RNA analog hybrids or chimeric oligonucleotides and also allows the incorporation of a large variety of modified nucleobase analogs at precise locations within the oligonucleotide sequence [89]. It also offers access to DNA and RNA spiegelmers, single-stranded mirror images of DNA and RNA oligonucleotides that offer the advantage of being extremely resistant to DNA and RNA nucleases [99–101].

Alternatively, RNA and DNA of virtually any size and sequences can be obtained by *in vitro* enzymatic synthesis such as cloning [102], Polymerase Chain Reaction (PCR) or *in vitro* RNA transcription of plasmid and PCR generated templates [35, 103, 104]. Interestingly, a striking number of nucleotide triphosphate analogs are substrates for DNA and RNA polymerases, allowing enzymatic synthesis of DNA or RNA containing modification [7, 105]. In that case, the nucleotide analog is incorporated in the resulting transcripts at all the locations specified by the complementary template nucleotide. This approach can be used to investigate the structure and function of a RNA molecule by nucleotide analog interference mapping NAIM [106–108]. It has also been used to expand the functional scope of RNA or DNA by selection techniques (Section 10.4). The introduction of 2′-O-methyl and other substitutions into RNA and DNA can be facilitated by the selection and evolution of new polymerase variants that can incorporate modified nucleotides [109–114]. Interestingly, a great variety of chemically modified nucleotides with C5 amino-acyl groups can be incorporated during PCR by the KOD Dash DNA polymerase [115]. Until recently, the specific rules of natural DNA and RNA polymerases did not permit incorporation of a nucleotide analog at a specific and unique position within a RNA or DNA. However, the development of novel base pairings has recently been used to incorporate a site-specific fluorescent dye within an RNA, generated by transcription of a modified DNA template with 2-amino-6-(2-thienyl)purine triphosphate (sTP) or 2-amino-6(2-thiazolyl)purine triphosphate (vTP) in the presence of the modified complementary nucleotide triphosphate (2-oxo-(1H)pyridine triphosphate (yTP)) [116]. Alternatively, incorporation of site-specific modification within large RNA and DNA molecules is generally achieved by ligation strategies [117], thereby circumventing the size limitation of synthetic DNA and RNA. The yield of final products reached by these strategies is however still limited.

Particularly interesting are nucleotide analogs that can efficiently mimic the conformation of natural nucleotides helices (Fig. 10.7). For instance, LNA and 2′-O-Me-RNA are both known to adopt A-form helix conformation and as such, are good structural analogs of RNA helices with chemically more stable backbones [10, 118]. For example, DNA/LNA chimeric oligomers have recently been shown to mimic RNA aptamers targeted to the TAR RNA element of HIV [119]. In this specific case, none of the 2′-hydroxyls present at the level of the aptamers is critical for their function [119]. However, within large RNA sequences, the 2′-hydroxyls can be involved at key positions in 3° structure motifs and long-range interactions to promote folding and assembly into complex 3-D shapes. Nevertheless, it is possible to envision the synthesis of chemically more stable RNA foldamer analogs for which most of the ribonucleotide positions are substituted by chemically stable nucleotide analogs as long as these key 2′-OH positions are known and kept unchanged.

It is also worth mentioning that synthetic and natural nucleic acids can be specifically modified by crosslinking [120–125], post-transcriptional modification at their 5′ or 3′ end extremities [126–128] or by specific modification of 2′-hydroxyl positions [129–131] in order to conjugate them with other molecular components.

10.4

Combinatorial Approaches for Isolating Functional Nucleic Acid Foldamers

The great functional diversity of nucleic acids, exemplified in Section 10.2.2, essentially results from combinatorial techniques known as "*in vitro* selection", "*in vitro* evolution" or SELEX (Systematic Evolution of Ligands by EXponential enrichment) techniques (see the reviews [132–134]). It exploits the construction of large libraries of DNA or RNA sequences, with the possibility of amplifying a small subset of selected molecules by PCR or RT-PCR. An example of an *in vitro* selection experiment is presented in Fig. 10.8.

First, a large library of DNA molecules is synthesized from a pool of synthetic semi-randomized oligonucleotides that are amplified by PCR to generate multiple
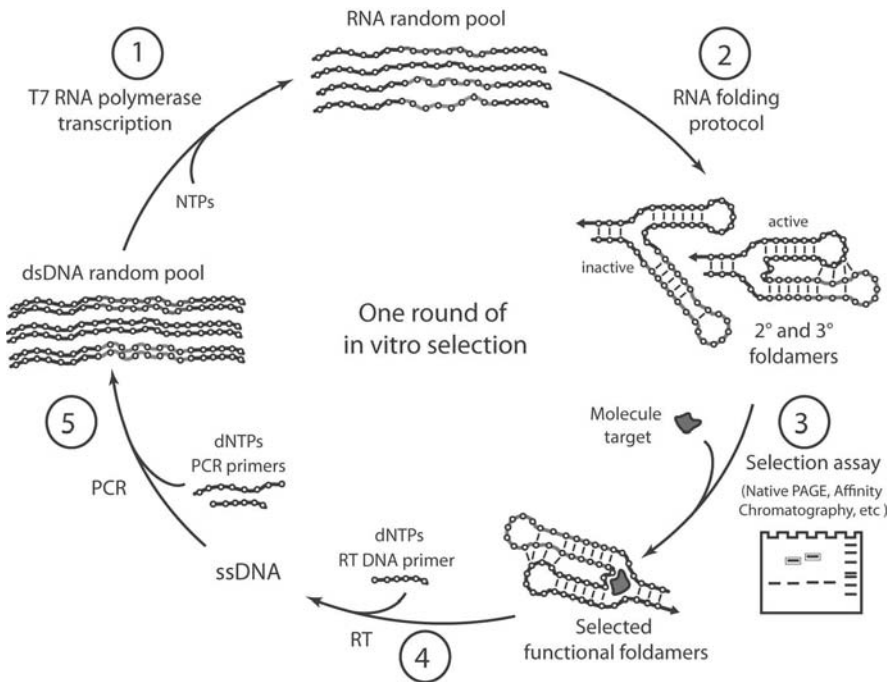
Fig. 10.8 Example of a round of SELEX. (step 1) A random pool of RNA molecules is synthesized by *in vitro* run-off transcription of a random DNA library generated by PCR. After folding in the presence of salts (step 2), the resulting population of 2° and 3° foldamers is submitted to a selection criterion such as binding to a molecule target. The functional molecules can be separated from the inactive ones by a selection assay such as a gel shift or affinity column assay (step 3). The selected molecules are then amplified by reverse transcription (RT) (step 4) and PCR (step 5) and eventually resubmitted to an additional round of selection. Mutations can eventually be introduced during this amplification steps to improve the activity of the pool of functional molecules.

double-strand DNA copies. A library typically contains molecules with two invariable sequence regions at the 5′ and 3′ ends (for amplification purposes) bracketing a random region including up to 200 nucleotides. Initial DNA libraries can contain up to 10^{16} individual molecules [75]. For DNA selection, the double-stranded DNA library is first denatured so that only one DNA strand enters into the selection process. For RNA selection, the DNA library is first transcribed into RNA. The population of single-stranded molecules is then challenged to perform a specific task that can be recognition of a molecular target or catalysis of a specific chemical reaction. This selection step is the most critical of the whole SELEX process and requires efficient physical separation of the functional molecules from the nonfunctional ones. Once separated, the selected molecules are amplified by PCR or RT-PCR and subjected to additional rounds of selection-amplification until the functional activity of the population can be detected through biochemical assay in the pool of selected molecules. The selected molecules can then be cloned, sequenced and tested individually for function and eventually further optimized after introduction of mutations by partial randomization or mutagenic PCR [132–134].

This powerful and very versatile technique has recently been automated so that selection can be performed in a matter of days instead of weeks [135–141]. Another interesting method called continuous *in vitro* evolution offers the possibility of evolving self-modifying ribozymes over several hundreds of generations [132]. The applicability of this system is limited, however.

SELEX strategies can also be applied to nucleic acids analogs synthesized by enzymatic incorporation of modified nucleotides and amplifiable by polymerases variants (see previous section). Alternatively, when nucleic acid analogs might not be amplifiable by PCR or RT-PCR, new methods such as non-SELEX selection can potentially be used to isolate aptamers in only one round of selection [142]. This latter technique developed by analytical chemists is derived from capillary electrophoresis SELEX (CE-SELEX). Because of the higher partition coefficient of CE over more traditional chromatographic techniques, this method allows identification of aptamers with extremely well defined affinity profiles in a limited number of rounds of selection [143, 144]. Another technique of selection, initially developed for protein selection and evolution, has recently been adapted to the isolation of transacting ribozymes through *in vitro* compartmentalization [145–147]. It is worth mentioning that aptamers can also be selected against heterogeneous mixtures of targets such as whole cells offering the possibility of discriminating against different cell types, even when specific biomarkers are not known in advance. This type of approach has been also reviewed recently [148].

10.5
DNA Architectonics

Nucleic acid architectonics is the scientific study of the principles underlying the construction of nucleic acid architectures [31]. This area of research pioneered by

Nadrian Seeman in the mid 1980s, is usually called nucleic acid nanotechnology. For almost 20 years, it has been essentially devoted to the study of DNA nanostructures.

10.5.1
Rational Design of DNA Tiles

Several reviews that describe the rational design and approach used in DNA nanoconstruction have recently been published [41, 149–151]. Readers are particularly encouraged to look at the excellent review by Feldkamp and Niemeyer [151], which presents the best collection of available DNA-based nanoarchitectures generated by rational design and the most comprehensive list of citations published on the subject over the past 20 years.

Because of the lack of stable natural 3° structure motifs, much effort has been devoted to designing robust and rigid DNA self-assembling building blocks, called tiles [25]. The majority of engineered DNA "tiles" are essentially formed
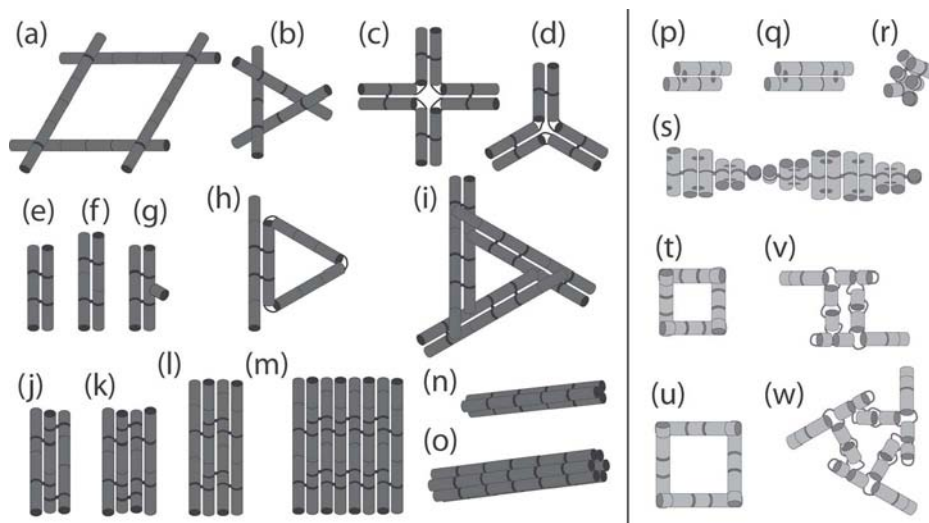
Fig. 10.9 Examples of self-assembling DNA tiles and tectoRNAs. (Left panel) DNA tiles are in dark grey: (a) rhombus (4 strands) [153]; (b) single cross-over triangle (4 strands) [154, 155]; (c) 4 × 4 cross (9 strands) [193]; (d) 3-point-star (7 strands) [156–158]; (e–i) double-crossover (DX) tiles: (e) DAE (5 strands) [159]; (f) DAO (4 strands) [159]; (g) DAE with protruding helix (4 strands) [159]; (h) DX tile with protruding triangle (4 strands) [160]; (i) DX triangle (10 strands) [161]); (j–o) Helix-Bundle (HB) tiles: (j) triple-crossover tile (TDX) (4 strands) [162]; (k) and (l) 4HB (8 or 9 strands) [163, 164]; (m) 8HB (18 strands) [164, 165]; (n) 3HB (9 strands) [166]; (o) 6HB (16 strands) [167]). (Right panel) TectoRNAs are in light grey: (p, q) loop-receptor tectoRNA dimeric particle [103, 176, 179]; (r, s) H-shaped tectoRNA particle and filament [104, 176]; (t, u) small and large "right angle" tectosquares [35]; (v, w) pRNA dimeric and trimeric particles [182, 183].

using a small number of structural rules derived from crossover (Holliday) junction motifs [152]. They are typically assembled from multiple oligonucleotide strands that interact through selective complementary Watson–Crick pairing and intertwine through crossover motifs [153–167] (Fig. 10.1). In particular, the design of robust helix bundles tiles [163–167] offers an attractive frame work for generating 1-D, 2-D or 3-D nanostructures through fine tuning of the positioning of crossover motifs that join parallel helical stacks [168] (Fig. 10.9). As multimolecular assemblies, DNA tiles can readily be considered $4°$ structures but from structural and thermodynamic stand points, no clear distinction can be established between their $2°$, $3°$ and $4°$ structures because their formation is essentially based on $2°$ structure constraints (Figures 10.1 and 10.4). As DNA architectures essentially rely on Watson–Crick base pairing, they are less hierarchical than those of RNA. They assemble by strand invasion processes similar to those that take place during homologous recombination events in cells.

10.5.2
Principle of Tensegrity and Mode of Assembly

A subtle balance of flexibility and stress is required to build good self-assembling tiles [158] but stable rigid $3°$ structural motifs are not an absolute requirement. The vertices of triangulated architectures can be flexible as triangulated structures should be able to resist deformation through tensegrity, a geometrical construction principle that combines stiff helical struts that push outward and flexible junctions that push inward (Fig. 10.10). By taking advantage of this principle, sta-
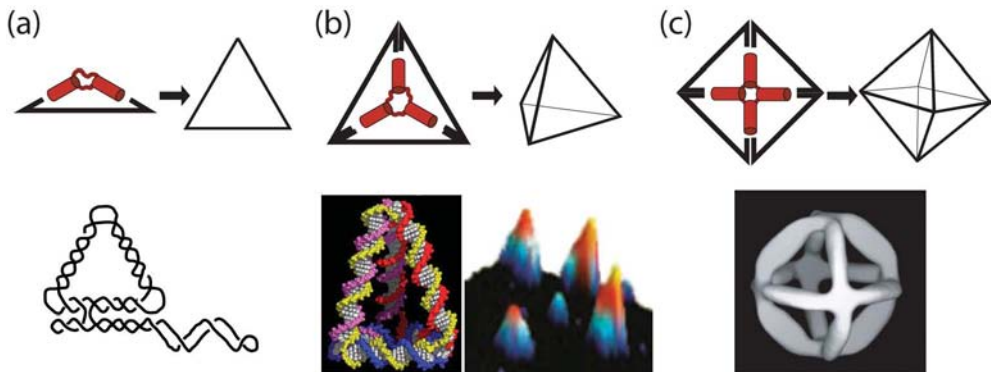
Fig. 10.10 Principle of tensegrity in DNA architectonics. The principle of tensegrity is illustrated in the construction of rigid objects like (a) a DNA triangle [154]; (b) a DNA tetrahedron [169] and (c) a DNA octahedron [102]. (a) (bottom): schematic representation of a DNA triangle that protrudes from a DNA tile [154]; (b) (bottom): 3-D model and AFM image of a DNA tetrahedron (adapted with permission from reference [169]); (c) (bottom): low resolution 3-D structure model of a DNA octahedron obtained by cryo-EM and single image reconstruction (adapted with permission from Ref. [102]).

ble triangular DNA tiles able to assemble into extensive Kagome-like lattices [154, 155], a replicable octahedron cage [102] and rigid tetrahedron building blocks [169] have recently been constructed.

The structure of most DNA tiles imposes strong geometrical constraints over the positioning of their cohesive interfaces (Fig. 10.4). Typically, only a reduced number of different 4° supramolecular architectures can be generated from a particular design of tile. DNA cohesive interfaces are typically formed through complementary Watson–Crick base pairing between collinear tail connectors of adjacent tiles [159]. They can also occur through formation of paranemic crossovers between internal loops that are wrapped around one another and do not interpenetrate topologically [170] (Fig. 10.3). Variation in the number of tail connectors and their thermodynamic stability can be used to modulate the assembly process as a function of temperature, DNA molecules and salt concentration.

In the future, the use of triple helices [26], G-tetrads [27] and non Watson–Crick parallel strands [171] will probably expand the modes of assembly of DNA tiles. In fact, it has already been demonstrated that frayed 2-D and 3-D networks can potentially be generated with guanine-rich DNA oligonucleotides expected to form G quartets [172, 173]. Moreover, a continuous 3-D hexagonal lattice generated from a 13mer DNA oligonucleotide self-assembling through parallel-stranded base pairing was subsequently engineered to produce crystals with enlarged solvent channels [171].

Considering that DNA can fold into stable 3° aptamers and DNAzymes, it is clear that the full potential of DNA 3° structure for nanoconstruction has not been exploited yet. However, the real potential of DNA may lie more in the optimal use of its simple rules of assembly, based on the unique selectivity of Watson–Crick base pairing, rather than its 3° structure diversity, as exemplified by the recent development of scaffolded DNA origami [174] discussed in Section 10.7.1.4.

10.6
RNA Architectonics

The concept of RNA tectonics was initially defined as referring to the modular character of RNA structures that can be decomposed and reassembled to create new modular RNA units, called tectoRNAs, which are able to self-assemble into nanoscale and mesoscale architectures of any desired size and shape [30, 31]. RNA architectonics is the science behind this concept [31].

10.6.1
General Approach

The methodological approach [35, 104, 175] is described in Fig. 10.11. The rational design of artificial 3-D RNA architectures [35, 103, 104, 176] is based on an inverse folding process. Structural fragments corresponding to 3° structure motifs are ''cut and pasted'' from known X-ray or NMR structures (step 1) and inter-
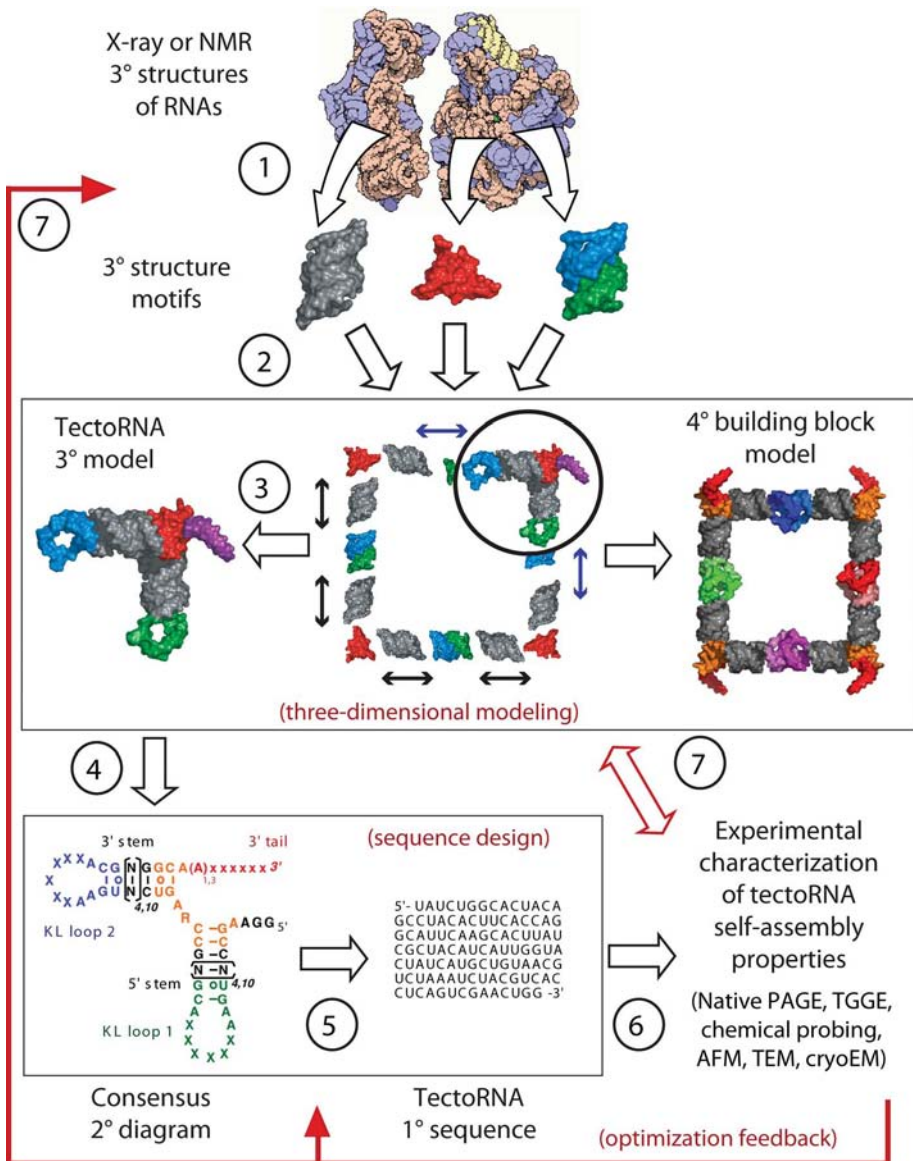
Fig. 10.11 The RNA architectonics methodology. The process of engineering artificial tectoRNA architectures is a multi-step procedure. First, RNA fragments extracted from known crystallographic or NMR data (step 1) are interactively reassembled into artificial RNA molecules by computer 3-D modeling (steps 2 and 3). These 3° models are then used as scaffolds to define consensus 2° diagrams (step 4) that are used as blueprints for designing RNA sequences (step 5) that are optimized by energy minimization [14] to maximize their thermodynamic stability and minimize the occurrence of alternative 2° structure folds [15]. The RNA sequences are then synthesized by chemical or enzymatic methods and characterized for their expected folding and self-assembly properties (step 6). TectoRNAs rational design can be optimized at the sequence or 3-D model levels (step 7).

actively reassembled into novel tectoRNA architectures by computer geometrical modeling with graphic user interfaces (step 2). During this mosaic modeling process [30], 3° interacting motifs can be positioned and oriented precisely by adjusting the lengths of their linking helical elements and the stacking of the helices at multi-helix junctions, thus allowing one to control the supramolecular assembly of RNA units. It is predicted that tectoRNAs will assemble into supramolecular architectures based on the conformation and geometry of their constitutive structural elements (step 3). These 3° models are then used as scaffolds to define consensus 2° diagrams, specifying invariant nucleotide positions to retain 3° structure constraints and positions involved in base pairing (step 4). TectoRNA sequences able to fold into these 2° blueprints are optimized by energy minimization [14] to maximize their thermodynamic stability and minimize the occurrence of alternative 2° structure folds [15] (step 5). The RNA sequences are then synthesized by chemical or enzymatic methods [35, 103] and their expected folding and self-assembly properties characterized by biochemical and biophysical methods like polyacrylamide gel electrophoresis (PAGE), temperature gradient gel electrophoresis (TGGE) and visualization techniques, such as AFM [35, 159, 175] or transmission electron microscopy (TEM) [104, 155] (step 6). The experimental data are then compared with the theoretical models and used to optimize the tectoRNA rational design at the sequence or 3° model level (step 7).

The effect and contribution of specific 3° structure motifs to the overall geometry and stability of the resulting supramolecular architecture can be assessed by introducing sequence mutations at key 3° nucleotide positions within tectoRNA molecules [35, 103, 176–179]. Mutated tectoRNA assemblies are used as negative control for comparison with non-mutated ones. Thus, this approach can also be a powerful way of unraveling the structural properties of 3° and 4° structure motifs for which few experimental data are available.

Although still a new field of investigation, RNA architectonics has already generated a great variety of tectoRNA units able to assemble into highly modular supramolecular architectures of arbitrary shapes (Figures 10.1, 10.4, 10.11 and 10.9). Besides classic cohesive Watson–Crick base pairing, the formation of long-range RNA–RNA interactions, such as loop–receptor or loop–loop interactions, offers a wide range of 4° intermolecular interfaces with various thermodynamic strengths to promote the cooperative assembly in the presence of divalent ions [35, 104, 176, 180] (Fig. 10.9). In the presence of magnesium, kissing loop motifs are more stable than RNA duplexes with identical sequences by two or three orders of magnitude [35, 180]. Moreover, the dynamic equilibrium of assembly through 4° RNA interfaces can be tuned over four to five orders of magnitude by adjusting the magnesium ion concentration and temperature. Thus, the hierarchical self-assembly of tectoRNAs can be monitored in a stepwise fashion to form architectures of increasing complexity [35], as there is a clear distinction between the energies involved in the formation of their 2°, 3° and 4° structures. In contrast to most DNA tiles, the formation of RNA tiles relies on the self-folding of single-stranded tectoRNAs that are characterized by well-defined 2° or/and 3° structures and 4° intermolecular interfaces.

10.6.2
Examples of RNA Nano-architectures

The first tectoRNAs to be generated by RNA architectonics self-assemble through loop–receptor interfaces to form dimeric nanoparticles [103, 176] or micrometer-long RNA filaments [104, 176] (Figs 10.9p–s and 10.12). The atomic structure of a self-dimerizing loop–receptor tectoRNA particle was recently solved by NMR and shown to be in remarkable agreement with the initial 3° structure model [181] (Fig. 10.12a).

Combining rational design of well-defined RNA 3° structures with small-scale combinatorial synthesis holds promise of engineering new functional modules that can accommodate the 3° structural constraints of specific supramolecular architectures [75, 78, 79]. For example, a new class of self-folding RNA molecule similar to domain P4–P6 of the natural *Tetrahymena* group I ribozyme was obtained by RNA architectonics [177] and subsequently used as a scaffold for combinatorial synthesis of new catalytic modules [78].

Several programmable and addressable RNA nanoparticles have been engineered to assemble in a predictable fashion through complementary selective loop–loop interactions [35, 180, 182, 183]. The DNA-packaging motor of bacterial
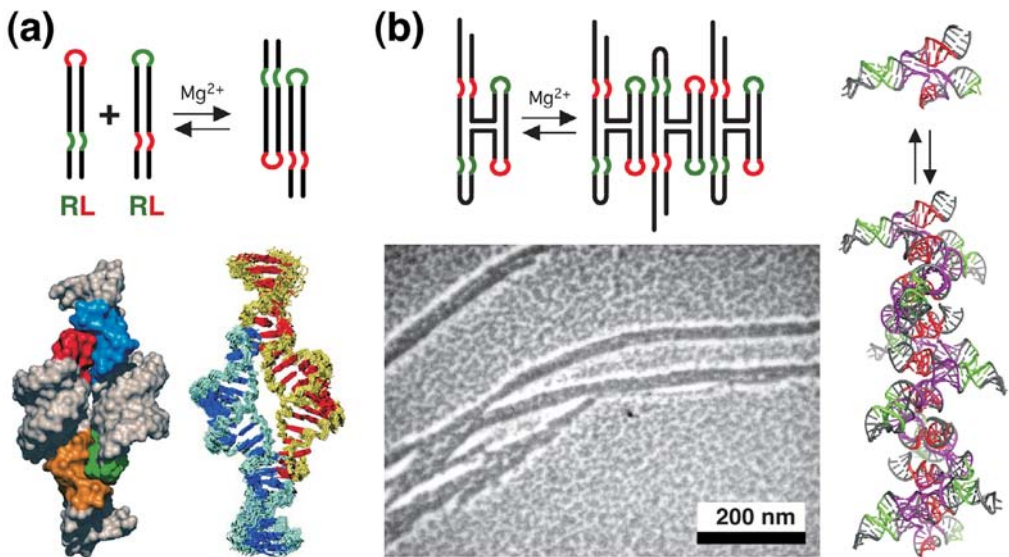
Fig. 10.12 TectoRNA nano-particles and filaments. (a) 0-D: Loop-receptor dimeric tectoRNA particle: the original 3° structure model [103, 176] (left) is in remarkable agreement with the recent NMR structure of the particle [181] (right). (b) 1-D: as predicted by 3° structure modeling (right), "H shaped" tectoRNAs can assemble into programmable, chiral and directional RNA filaments that can be visualized by TEM (adapted from Ref. [104]).
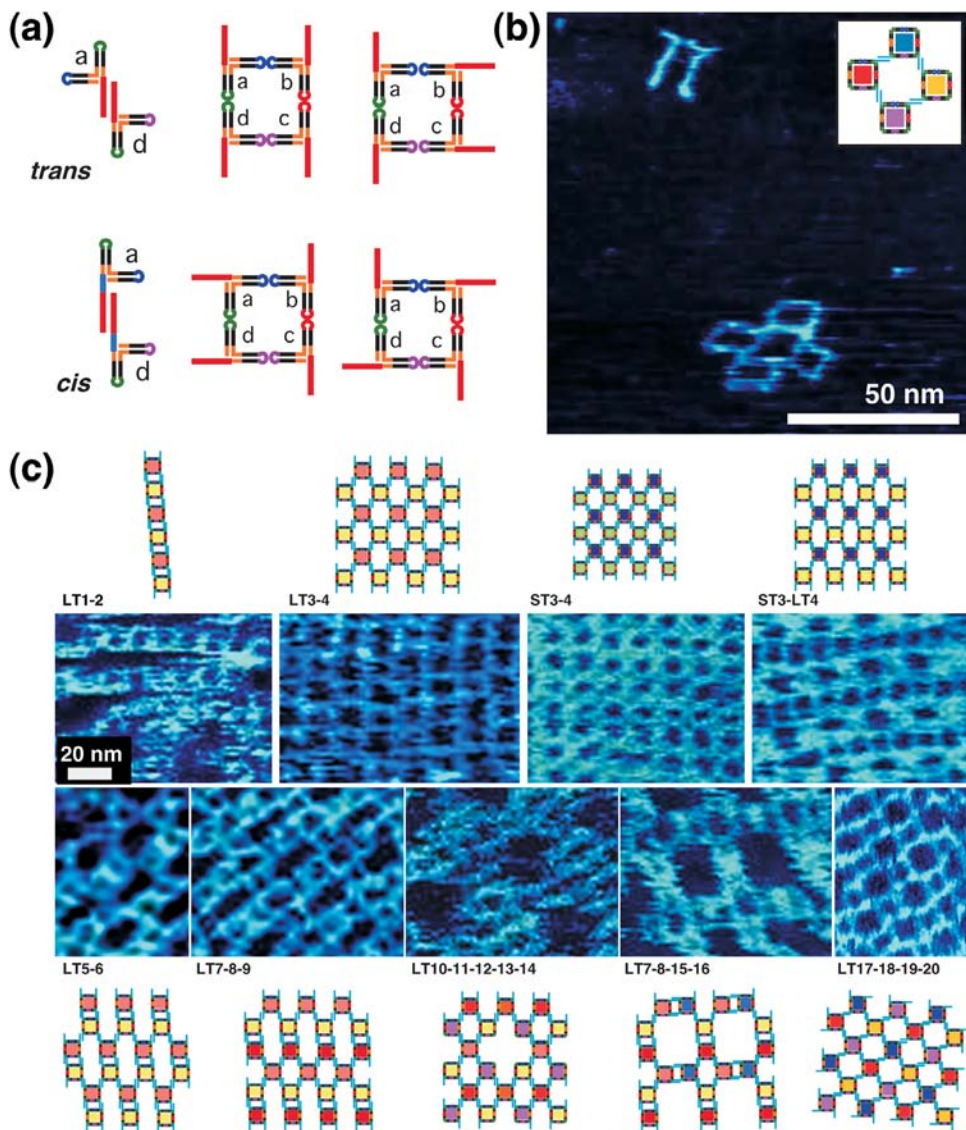
Fig. 10.13 Programmable and addressable 2-D architectures of RNA. (a) RNA tectosquares (TS) are programmable tetrameric nanoparticles. The geometry of TS assembly can be controlled by the orientation and length of their 3′ tail connectors [35]. (b, c) 2-D architectures of tectosquares (adapted from Ref. [35]); (b) The first programmable RNA nano-grid with 16 distinct, addressable positions [35]. This RNA structure is aperiodic with respect of its molecular constituents. (c) Various periodic patterns generated by combination of 22 tectosquares.

virus phi29 contains six-DNA packaging RNAs (pRNAs), which together form a hexameric ring via loop–loop interactions. For example, pRNAs were redesigned to form a variety of predictable structures namely dimers, tetramers, triangles, rods as well as micrometer size bundles of pRNA filaments [182, 183] (Fig. 10.9v–w). Recently, controllable trimeric pRNA particles harboring therapeutic molecules, siRNAs, and a receptor-binding aptamer have been shown to act as a delivery vehicle to cancer cells and induce apoptosis [184]. Collinear kissing loop interactions can generate strong $4°$ interacting interfaces to promote the formation of RNA particles of different sizes [180] (Fig. 10.3). This assembly principle was used in the engineering of a versatile molecular system that takes advantage of a "right angle" $3°$ structure motif to form highly programmable square shaped tetrameric nanoparticles, called tectosquares [35] (Fig. 10.9t–u).

The high modularity and hierarchical supramolecular structure of tectosquares makes it possible to construct a large number of combinatorial variants from a limited set of tectoRNAs that assemble through strong $4°$ interacting loop–loop interfaces [35]. Tectosquares can display an assortment of sticky tail connectors at their corner to control the geometry, directionality and addressability of the self-assembly process (Fig. 10.13). A mixture of them can assemble further into complex 1-D and 2-D architectures with periodic and aperiodic patterns and finite dimensions (Fig. 10.13). Considering that up to 88.5 millions of distinct tectosquares can theoretically be generated from a limited set of 24 tails with two different tails orientations and sizes, an almost infinite number of complex jigsaw puzzle patterns can be designed [35].

10.7
Self-assembly Strategies for Building Complex Nucleic Acid Nanostructures

10.7.1
Programmable Self-assembly

Programmable self-assembly is defined as self-assembly processes whereby the information specified at the molecular sequence level can be controlled with high predictability to fold and assemble into predefined 2-D and 3-D architectures [31].

10.7.1.1 General Principles: "One pot" versus "Step-wise" Assembly
Two main approaches can be distinguished for programmable self-assembly of nucleic acid architectures (Fig. 10.14). The first approach, mostly used with DNA, is a single step assembly strategy in which all the molecules encoding a specific architecture are mixed together and assembled in "one pot" through a slow annealing procedure [153–167, 174, 185] (Fig. 10.14a). According to the energetics of their $2°$ structure pairings, oligonucleotide strands form stable substructures or tiles that assemble through weaker $4°$ interfaces into larger nanoarchitectures when lower temperatures are reached. These structures can eventu-
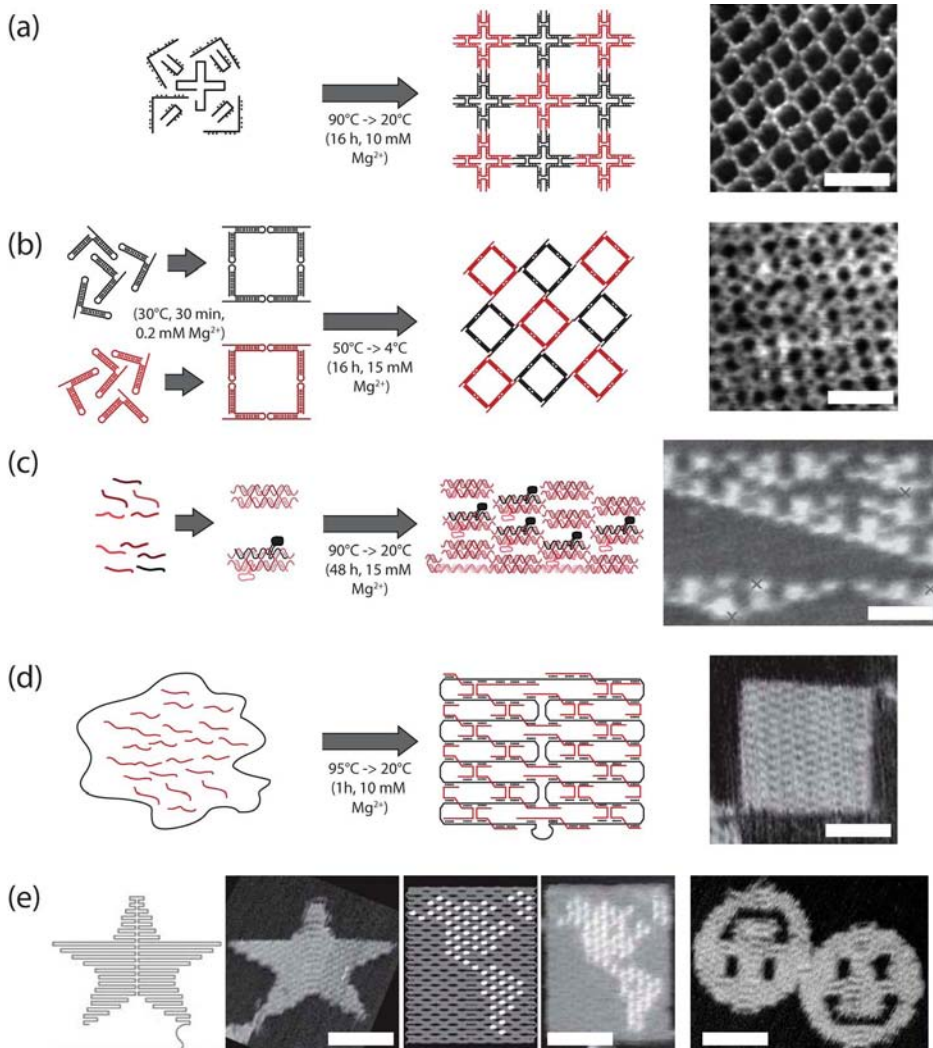
Fig. 10.14 The four main strategies for programmable self-assembly. (a) Single step process of self-assembly whereby all the molecules are mixed together. Most DNA architectures are formed that way (adapted with permission from Ref. [193]); (b) Stepwise hierarchical self-assembly whereby specific sets of molecules are first separately assembled into small supra-molecular entities that are then mixed in a stepwise fashion to form the final architecture [35]; (c) Programmable algorithmic self-assembly: tiles with local pairing implementing the exclusive-or function, are assembled on a template input row to form the Sierpinski triangle pattern (adapted with permission from Ref. [190]); (d) Scaffolded self-assembly where a long single stranded molecule is folded into an arbitrary shape in presence of small oligonucleotides acting as staples [174]; (e) Examples of patterns that can be generated using scaffolded DNA origami (adapted with permission from Ref. [174]). Scale bars are all 50 nm.

ally be ligated together to form robust covalently linked architectures [186] or networks [159].

The second approach, particularly appropriate for RNA assembly, is a stepwise hierarchical self-assembly strategy, in which various small subunits are first separately formed and then mixed together to form the final supramolecular architecture [35, 187] (Fig. 10.14b). This strategy is more time consuming, however. As exemplified by the tectosquare system [35], it can make use of the same 4° interactions and basic molecular units to build a large number of highly modular tiles that can assemble further through weaker 4° interactions. Thus, by separating tile formation from the formation of larger supramolecular assemblies, a reduced number of different connecting interfaces can be used to hierarchically build highly modular architectures [35]. In stepwise assembly, the melting temperature of the tiles and of the resulting supramolecular architecture should be kept well separated. By contrast, this is not absolutely necessary for the ''one pot'' approach, as exemplified by DNA scaffolded origami [174].

These two approaches can make use of additional self-assembly strategies that are not mutually exclusive, such as addressable self-assembly, algorithmic self-assembly, templated (or directed nucleation) self-assembly and scaffolded DNA origami.

10.7.1.2 Addressable Self-assembly

Step-wise assembly can be used to generate addressable architectures of finite size, with the position of each of the constitutive molecules being known without ambiguity within the assembly and therefore addressable within the final architecture. The first demonstration of this approach led to the fabrication of RNA nanogrids of finite size [35, 188] (Fig. 10.13). More recently, the application of this strategy to DNA led to the fabrication of nano-arrays with precisely positioned nanoparticles that form patterns of letters [187] or a pegboard [189].

10.7.1.3 Algorithmic Self-assembly

In algorithmic self-assembly, a set of nucleic acid tiles, defined as Wang tiles, is viewed as the algorithm for a particular computational task leading to the formation of 1-D, 2-D and 3-D patterns [190, 191]. This strategy was used to compute the formation of aperiodic fractal 2-D patterns based on the Sierpinski triangle pattern [190] (Fig. 10.14c). To achieve this task, a minimal set of four DNA tiles with local pairing rules designed to implement the exclusive-or (XOR) function, was assembled on a template input row to facilitate the nucleation of the directional self-assembly growth into a unique pattern [190]. The potential of algorithmic self-assembly is, however, still limited by the presence of various errors, introduced by lattice dislocation, formation of untemplated crystals and mismatched tiles.

10.7.1.4 Templated Self-assembly and Scaffolded DNA Origami

Templated or directed nucleation assembly takes advantage of a nucleic acid template that acts as a scaffold for directing the specific assembly of tiles. This strat-

egy led to the formation of aperiodic 2-D arrays, such as DNA barcodes [192]. The construction of a replicable DNA octahedron [102] was based on a similar scaffolded approach. In this case, a single-stranded DNA molecule that forms helical struts was assembled with the help of four small oligonucleotides into its final shape through the formation of paranemic long-range interactions (Fig. 10.3a). The generalization of these approaches led to the versatile scaffolded self-assembly strategy, also called scaffolded DNA origami [174], which can generate with a remarkable efficiency almost any type of arbitrary shape and pattern (Fig. 10.14d). In this strategy, a long single-stranded DNA scaffold is folded with complementary oligonucleotides that act as staples. The desired shape is designed by raster filling the shape with a 7-kilobase single-stranded scaffold and ∼200 short oligonucleotide staple strands to hold the scaffold in place (Fig. 10.14d–e). Once synthesized and mixed, the staple and scaffold strands self-assemble in one single step. The structure can be programmed into complex patterns, such as words and images (Fig. 10.14e). The success of scaffolded origami stems from several contributing factors, such as efficient strand invasion, excess of staples, cooperative effects and a design that intentionally does not rely on binding between staples [174]. A relatively good yield of defect-free DNA architectures was obtained, despite the fact that the oligonucleotides used were not purified.

10.7.2
Additional Principles of Nucleic Acid Architectonics

10.7.2.1 Principle of Orientational Compensation
The inherent asymmetric nature of RNA and DNA tiles can have a dramatic effect on the larger nanostructures that they form by introducing various degrees of curvature. By using the principle of orientational compensation, whereby two adjacent units are related by a local twofold pseudo-rotational axis of symmetry, one source of asymmetry can be locally eliminated so that asymmetric tiles that are not perfectly flat can still assemble in a plane instead of forming nanotubes [157, 160, 193]. This strategy was also used to favor the assembly of "H-shaped" tectoRNAs into linear filaments instead of rings [104].

10.7.2.2 Applications of Principles of Symmetry
The application of sequence symmetry principles to the design of structurally symmetrical tiles can reduce the sequence size and number of strands necessary for the construction of very complex nanostructures. This approach was extremely powerful for fabricating 2-D DNA arrays up to 1 mm in size that were able to be visualized by fluorescence microscopy [194]. Similarly, the application of symmetry principles to tile assembly can reduce dramatically the number of tiles when constructing nano-arrays of finite sizes [35, 165].

10.7.2.3 Fractal Nano-architectures
As shown previously, fractal patterns can be generated by algorithmic self-assembly [190]. As proposed by Carbone and Seeman, fractal architectures could

potentially be generated by hierarchical stepwise assembly strategies [195]. This remains to be demonstrated, however.

10.8
Ornamentation and Functionalization of Nucleic Acid Architectures

10.8.1
General Principles

Principles of ornamentation of DNA architectures have been reviewed by Gothelf and Niemeyer [149] and Yan and colleagues [196]. Briefly, programmable nucleic acid architectures can direct the spatial organization of other components, like proteins [155, 193, 197], metallic nanoparticles [187, 189, 198–201], small molecules or nanodevices [202] to generate new materials with potential applications in fields as diverse as medicine, molecular biology and device physics [149]. Among the various strategies employed for functionalizing nucleic acid architectures, the use of DNAzymes, ribozymes, therapeutic siRNAs, RNA and DNA aptamers is particularly promising as these molecules can be readily encoded at precise locations within the nucleic acid architecture [184, 197, 203]. The incorporation of functional nucleic acids within self-assembled nanostructures is presently under investigation. For example, DNAzymes have been shown to retain their activity when attached in linear strings within the context of a 2-D DNA array and signaling aptamers immobilized within a self-assembled DNA array have been used for protein detection [204].

10.8.2
Nucleic Acid Foldamers for Sensors, Medicine and Nano-electronics

The availability of many nucleic acid aptamers, riboswitches, ribozymes and DNAzymes has opened a new horizon for a wealth of applications such as novel sensors for protein, RNA, DNA, metabolite and metal–ion detection, drug discovery and nanotechnology. Navani and Li have recently reviewed the fast growing developments in this field [205]. Moreover, several nucleic acid aptamers are presently in clinical trials, illustrating the real potential of these functional nucleic acid foldamers as drugs and therapeutics [7, 206]. Therapeutic agents such as small interfering siRNAs, ribozymes, and antisense RNAs show significant potential in new molecular approaches to down regulate specific gene expression in cancerous or virus-infected cells. The development of safe, efficient, specific and nonpathogenic nanoparticles for the packaging and delivery of multifunctional therapeutic RNA is thus highly desirable. Recent investigations suggest that the use of antigen-free 20–40 nm programmable self-assembled RNA particles presenting multiple therapeutic functionalities [184, 207] might hold promise as delivery and therapeutic systems for the repeated long-term treatment of chronic diseases [6]. Although RNA is chemically more fragile than DNA, such ''instability'' may actually permit a higher degree of flexibility in the design of activatable

structures or triggered assembly or degradation of the engineered target in a timely fashion.

Several strategies have been used to generate nanowires taking advantage of nucleic acids as templates for nano-electronic applications (Fig. 10.15). One of
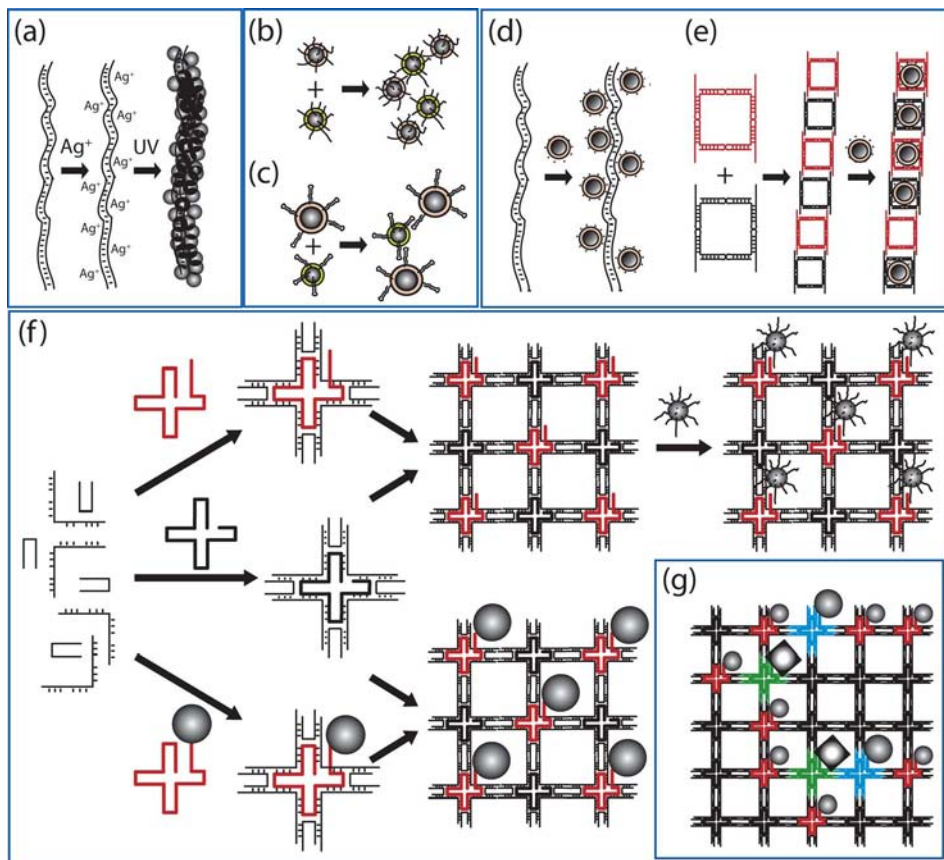
Fig. 10.15 Nucleic acid foldamers for nanoelectronics. (a) Metallic nano-wires can be fabricated from a DNA template by silver deposition and glutaraldehyde reduction [208] or silver photo-induced metallization [209]; (b, c) Metallic particles coated with (b) DNA [211] or (c) RNA [213] can be used to direct the specific bottom-up assembly of nanowires between two electrodes [212, 213]; (d) Positively charged nanoparticles can be aligned on DNA linear structures by electrostatics [215]; (e) Cationic metallic particles can also be positioned with exquisite regularity on RNA 1-D ladder scaffolds by electrostatics, size and shape recognition [214]; (f) Periodic programmable 2-D arrays can be used as template for building regular metallic nanoparticle arrays either by noncovalently binding DNA coated particles or directly incorporating gold particle into the tile design by gold–DNA conjugation [187, 198, 199, 201]; (g) It is possible to use fully programmable and addressable nucleic acid nano-arrays of finite size as templates for exquisite positioning and ordering of different nanoparticles or other components on a surface to create exotic composite materials.

them is to direct metalization of DNA nanostructure by using glutaraldehyde as reducing agent for silver deposition [208], another is to use UV light for silver photo-induced metalization [209]. The former method was used for the fabrication of highly conductive silver nanowires by using DNA self-assembling 1-D architectures as templates [166, 193]. Interestingly, the specific incorporation in nucleic acid of modified triphosphates that bear functions that can be further derivatized with aldehyde groups via the use of click chemistry can offer an alternative route to the selective metalization of DNA or RNA molecules [210]. This method can potentially lead to the development of more complex metallic nucleic acid templated nanowires. Another strategy involves the use of metallic particles coated with DNA to direct their specific assemblies [211] between electrodes [212] (Fig. 10.15). Recently, conductive self-assembling nanowires were constructed by assembling gold-derivatized DNA particles with magnesium dependent loop–receptor tectoRNAs [213]. One RNA-based metallic wire located between lithographically fabricated nano-electrodes was shown to exhibits activated conduction by electron hopping at temperatures in the range 150–300 K [213]. Another interesting potential of well-defined nucleic acid architectures, like tectosquare 1-D ladders [35, 214], is that they can act as scaffolds to control the positioning of cationic nanoparticles not only based on electrostatics [215] but also size and shape recognition [214] (Fig. 10.15). The applicability of such type of assemblies in nano-electronics remains to be demonstrated, however. A very powerful way to precisely organize metallic particles takes advantage of periodic programmable 2 D arrays (Fig. 10.15). By noncovalently binding DNA-coated particles or directly incorporating gold particles into the DNA tile design, it is possible to build very regular metallic particle arrays of different sizes [187, 198–201]. For instance, the use of fully programmable and addressable nanogrids open the way to precise positioning of nanoparticles of different composition and sizes [187] (Fig. 10.15). These nucleic acid based technologies could be particularly promising for the development of several applications including the fabrication of microelectronic architectures, hybrid electronic, optoelectronic devices and sensing.

10.9
Conclusions

For the nanoconstruction of molecular devices with novel properties and functions, there is probably no other polymeric material (i.e. backbone class) that can challenge the versatility of nucleic acid foldamers. Our present understanding of the basic folding and assembly principles pertaining to RNA and DNA allow unprecedented control over the shape and supramolecular assembly of exquisite nucleic acid based nanostructures [31]. Additionally, nucleic acid foldamers offer an almost limitless wealth of molecular properties and functions that are reachable either by *in vitro* combinatorial techniques, by rational design or a combination of both.

The great potential of DNA architectonics is best exemplified by the quasi-digital approach of scaffolded DNA origami. DNA can potentially be shaped into

any arbitrary 1-D, 2-D and 3-D architectures with size ranging from 20 nm to several micrometers and any type of patterns can be drawn with DNA with a pixel definition of 3.6 nm. By contrast, the great potential of RNA architectonics lies in the possibility of sculpting any arbitrary shapes with sizes ranging from 3 to 25 nm, and with moving parts that can be precisely coordinated to generate responsive and directed molecular motion. At the present time, this potential is best exemplified by complex natural RNA nanoparticles such as the ribosome [40].

Several challenges still remain to be overcome, however. The efficiency of formation of nucleic acid nanostructures should be improved by minimizing errors occurring during the folding and supramolecular assembly. The development of computer tools to automate the design and prediction of complex 3-D nucleic acid structures would be particularly helpful in achieving this task, especially in the case of RNA. It will also be important to explore further the principles of nucleic acid architectonics to achieve better control over the movement, dynamic and responsiveness of nucleic-acid-based nanomachines. For instance, DNA-based nanomechanical devices [202] are still far from equalizing the remarkable complexity of efficiency of RNA nanomachine such as the ribosome. Moreover, in order to bring nucleic acid foldamers in the realm of material sciences, new synthetic pathways should be investigated to produce large quantities (grams to kilograms) of nucleic acids of known sequences. For instance, this would offer the unique opportunity to generate responsive, programmable films or plastics based on nucleic acids [216, 217].

Nucleic acid foldamers start to impact areas as diverse as chemistry, biology, medicine and physics. It is anticipated that numerous applications will be derived from this very active and recent field of research. Because of the biodegradability and biological functions of RNA, RNA foldamers might be well suited for bionanotechnology and nanomedicine applications whereas the robustness and chemical stability of DNA might offer greater possibilities for more conventional nanotechnology applications. In a near future, the complementary nature of RNA, DNA and other available nucleic acid analogs will find interesting new developments once combined together. Nevertheless, we can anticipate that nucleic acid foldamers are only the premise of a new generation of polymeric materials that will marry the remarkable programmability of nucleic acid to the chemical stability and robustness of synthetic polymers [218], offering thus a completely new avenue of research in material sciences.

Abbreviations

AFM – Atomic Force Microscopy
CE-SELEX – Capillary Electrophoresis SELEX
Cryo-EM – Cryogenic Electron Microscopy
DX – Double Crossover
FRET – Fluorescence resonance energy transfer
GNRA – tetraloop of sequence GNRA with N for any nucleotide and R for purine
LNA – Locked-in riboNucleic Acid
NMR – Nuclear Magnetic Resonance
PAGE – PolyAcrylamide Gel Electrophoresis
PCR – Polymerase Chain Reaction
PNA – Peptide Nucleic Acid
RT-PCR – Reverse Transcription-Polymerase Chain Reaction
SELEX – Systematic Evolution of Ligands by EXponential enrichment
siRNA –silencing RNA
TEM – Transmission Electron Microscopy
TGGE – Temperature Gradient Gel Electrophoresis
UNCG – tetraloop of sequence UNCG with N for any nucleotide

References

1 G. M. Whitesides, B. Grzybowski, *Science* **2002**, *295*, 2418.

2 J. M. Lehn, *Proc Natl Acad Sci USA* **2002**, *99*, 4763.

3 T. Lindahl, B. Nyberg, *Biochemistry* **1972**, *11*, 3610.

4 J. Pierre, J. Laval, *Biochemistry* **1980**, *19*, 5018.

5 L. Poliseno, A. Mercatanti, L. Citti, G. Rainaldi, *Curr Pharm Biotechnol* **2004**, *5*, 361.

6 P. Guo, *J Nanosci Nanotechnol* **2005**, *5*, 1964.

7 J. F. Lee, G. M. Stovall, A. D. Ellington, *Curr Opin Chem Biol* **2006**, *10*, 282.

8 S. M. Freier, R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, D. H. Turner, *Proc Natl Acad Sci USA* **1986**, *83*, 9373.

9 J. SantaLucia, Jr., *Proc Natl Acad Sci USA* **1998**, *95*, 1460.

10 M. Egli, *Angew. Chemie. Int. Ed.* **1996**, *35*, 1894.

11 J. A. Abels, F. Moreno-Herrero, T. van der Heijden, C. Dekker, N. H. Dekker, *Biophys J* **2005**, *88*, 2737.

12 I. M. Kulic, H. Mohrbach, V. Lobaskin, R. Thaokar, H. Schiessel, *Phys Rev E Stat Nonlin Soft Matter Phys* **2005**, *72*, 041905.

13 N. B. Leontis, J. Stombaugh, E. Westhof, *Nucleic Acids Res* **2002**, *30*, 3497.

14 D. H. Mathews, D. H. Turner, *Curr Opin Struct Biol* **2006**, *16*, 270.

15 R. M. Dirks, M. Lin, E. Winfree, N. A. Pierce, *Nucleic Acids Res* **2004**, *32*, 1392.

16 M. Zuker, *Nucleic Acids Res* **2003**, *31*, 3406.

17 I. L. Hofacker, *Nucleic Acids Res* **2003**, *31*, 3429.

18 M. Andronescu, A. P. Fejes, F. Hutter, H. H. Hoos, A. Condon, *J Mol Biol* **2004**, *336*, 607.

19 M. Andronescu, Z. C. Zhang, A. Condon, *J Mol Biol* **2005**, *345*, 987.

20 S. A. Woodson, *Curr Opin Chem Biol* **2005**, *9*, 104.

21 D. Thirumalai, N. Lee, S. A. Woodson, D. Klimov, *Annu Rev Phys Chem* **2001**, *52*, 751.

22 S. A. Woodson, *Cell Mol Life Sci* **2000**, *57*, 796.

23 V. Thiviyanathan, B. A. Luxon, N. B. Leontis, N. Illangasekare, D. G. Donne, D. G. Gorenstein, *J Biomol NMR* **1999**, *14*, 209.

24 B. Wu, F. Girard, B. van Buuren, J. Schleucher, M. Tessari, S. Wijmenga, *Nucleic Acids Res* **2004**, *32*, 3228.

25 N. C. Seeman, *Methods Mol Biol* **2005**, *303*, 143.

26 S. M. Mirkin, *Curr Opin Struct Biol* **2006**, *16*, 351.

27 A. T. Phan, V. Kuryavyi, D. J. Patel, *Curr Opin Struct Biol* **2006**, *16*, 288.

28 I. G. Panyutin, O. I. Kovalsky, E. I. Budowsky, R. E. Dickerson, M. E. Rikhirev, A. A. Lipanov, *Proc Natl Acad Sci USA* **1990**, *87*, 867.

29 J. F. Lee, J. R. Hesselberth, L. A. Meyers, A. D. Ellington, *Nucleic Acids Res* **2004**, *32*, D95.

30 E. Westhof, B. Masquida, L. Jaeger, *Fold. & Des.* **1996**, *1*, R78.

31 L. Jaeger, A. Chworos, *Curr Opin Struct Biol* **2006**, *16*, 531.

32 S. R. Holbrook, *Curr Opin Struct Biol* **2005**, *15*, 302.

33 N. B. Leontis, A. Lescoute, E. Westhof, *Curr Opin Struct Biol* **2006**, *16*, 279.

34 D. K. Hendrix, S. E. Brenner, S. R. Holbrook, *Q Rev Biophys* **2005**, *38*, 221.

35 A. Chworos, I. Severcan, A. Y. Koyfman, P. Weinkam, E. Oroudjev, H. G. Hansma, L. Jaeger, *Science* **2004**, *306*, 2068.

36 N. B. Leontis, E. Westhof, *Curr Opin Struct Biol* **2003**, *13*, 300.

37 A. Lescoute, N. B. Leontis, C. Massire, E. Westhof, *Nucleic Acids Res* **2005**, *33*, 2395.

38 A. Lescoute, E. Westhof, *RNA* **2006**, *12*, 83.

39 P. Brion, E. Westhof, *Annu Rev Biophys Biomol Struct* **1997**, *26*, 113.

40 R. Gesteland, T. Cech, J. Atkins, *The RNa world*, 3 edn, Cold Spring Harbor Laboratory Press, 2005.

41 M. Brucale, G. Zuccheri, B. Samori, *Trends Biotechnol* **2006**, *24*, 235.

42 H.-A. Ho, M. Leclerc, *J. Am. Chem. Soc.* **2004**, *126*, 1384.

43 T. Hianik, V. Ostatna, M. Sonlajtnerova, I. Grman, *Bioelectrochemistry* **2007**, *70*, 127.

44 M. N. Stojanovic, D. M. Kolpashchikov, *J Am Chem Soc* **2004**, *126*, 9266.

45 B. J. Tucker, R. R. Breaker, *Curr Opin Struct Biol* **2005**, *15*, 342.

46 W. C. Winkler, R. R. Breaker, *Annu Rev Microbiol* **2005**, *59*, 487.

47 M. Mandal, R. R. Breaker, *Nat Rev Mol Cell Biol* **2004**, *5*, 451.

48 J. E. Barrick, K. A. Corbino, W. C. Winkler, A. Nahvi, M. Mandal, J. Collins, M. Lee, A. Roth, N. Sudarsan, I. Jona, J. K. Wickiser, R. R. Breaker, *Proc Natl Acad Sci USA* **2004**, *101*, 6421.

49 W. C. Winkler, *Curr Opin Chem Biol* **2005**, *9*, 594.

50 J. Lim, W. C. Winkler, S. Nakamura, V. Scott, R. R. Breaker, *Angew Chem Int Ed Engl* **2006**, *45*, 964.

51 A. Lescoute, E. Westhof, *Chem Biol* **2005**, *12*, 10.

52 S. Thore, M. Leibundgut, N. Ban, *Science* **2006**, *312*, 1208.

53 A. Serganov, A. Polonskaia, A. T. Phan, R. R. Breaker, D. J. Patel, *Nature* **2006**, *441*, 1167.

54 R. K. Montange, R. T. Batey, *Nature* **2006**, *441*, 1172.

55 A. Serganov, Y. R. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A. T. Phan, C. Hobartner, R. Micura, R. R. Breaker, D. J. Patel, *Chem Biol* **2004**, *11*, 1729.

56 R. T. Batey, S. D. Gilbert, R. K. Montange, *Nature* **2004**, *432*, 411.

57 L. Jaeger, *Curr Opin Struct Biol* **1997**, *7*, 324.

58 J. A. Doudna, J. R. Lorsch, *Nat Struct Mol Biol* **2005**, *12*, 395.

59 M. J. Fedor, J. R. Williamson, *Nat Rev Mol Cell Biol* **2005**, *6*, 399.

60 R. Fiammengo, A. Jaschke, *Curr Opin Biotechnol* **2005**, *16*, 614.

61 S. Fusz, A. Eisenfuhr, S. G. Srivatsan, A. Heckel, M. Famulok, *Chem Biol* **2005**, *12*, 941.

62 D. Liu, L. A. Gugliotti, T. Wu, M. Dolska, A. G. Tkachenko, M. K. Shipton, B. E. Eaton, D. L. Feldheim, *Langmuir* **2006**, *22*, 5862.

63 L. A. Gugliotti, D. L. Feldheim, B. E. Eaton, *J Am Chem Soc* **2005**, *127*, 17814.

64 L. A. Gugliotti, D. L. Feldheim, B. E. Eaton, *Science* **2004**, *304*, 850.

65 A. Torres-Larios, K. K. Swinger, A. S. Krasilnikov, T. Pan, A. Mondragon, *Nature* **2005**, *437*, 584.

66 A. S. Krasilnikov, Y. Xiao, T. Pan, A. Mondragon, *Science* **2004**, *306*, 104.

67 A. S. Krasilnikov, X. Yang, T. Pan, A. Mondragon, *Nature* **2003**, *421*, 760.

68 A. V. Kazantsev, A. A. Krivenko, D. J. Harrington, S. R. Holbrook, P. D. Adams, N. R. Pace, *Proc Natl Acad Sci USA* **2005**, *102*, 13392.

69 S. K. Silverman, *Nucleic Acids Res* **2005**, *33*, 6151.

70 S. K. Silverman, *Org Biomol Chem* **2004**, *2*, 2701.

71 A. Peracchi, *Chembiochem* **2005**, *6*, 1316.

72 Y. Li, R. R. Breaker, *Curr Opin Struct Biol* **1999**, *9*, 315.

73 P. C. Sabeti, P. J. Unrau, D. P. Bartel, *Chem Biol* **1997**, *4*, 767.

74 R. Knight, M. Yarus, *RNA* **2003**, *9*, 218.

75 L. Jaeger, M. C. Wright, G. F. Joyce, *Proc Natl Acad Sci USA* **1999**, *96*, 14712.

76 W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319.

77 N. Lee, Y. Bessho, K. Wei, J. W. Szostak, H. Suga, *Nat Struct Biol* **2000**, *7*, 28.

78 Y. Ikawa, K. Tsuda, S. Matsumura, T. Inoue, *Proc Natl Acad Sci USA* **2004**, *101*, 13750.

79 W. Yoshioka, Y. Ikawa, L. Jaeger, H. Shiraishi, T. Inoue, *RNA* **2004**, *10*, 1900.

80 R. M. Kumar, G. F. Joyce, *Proc Natl Acad Sci USA* **2003**, *100*, 9738.

81 R. R. Breaker, *Curr Opin Biotechnol* **2002**, *13*, 31.

82 J. Tang, R. R. Breaker, *Nucleic Acids Res* **1998**, *26*, 4214.

83 J. Tang, R. R. Breaker, *Chem Biol* **1997**, *4*, 453.

84 M. P. Robertson, A. D. Ellington, *Nucleic Acids Res* **2000**, *28*, 1751.

85 M. Koizumi, R. R. Breaker, *Biochemistry* **2000**, *39*, 8983.

86 M. Levy, A. D. Ellington, *Chem Biol* **2002**, *9*, 417.

87 C. Romero-Lopez, A. Barroso-delJesus, E. Puerta-Fernandez, A. Berzal-Herranz, *Biol Chem* **2005**, *386*, 183.

88 M. H. Caruthers, S. L. Beaucage, C. Becker, J. W. Efcavitch, E. F. Fisher, G. Galluppi, R. Goldman, P. deHaseth, M. Matteucci, L. McBride, et al., *Gene Amplif Anal* **1983**, *3*, 1.

89 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893.

90 H. M. Pfundheller, A. M. Sorensen, C. Lomholt, A. M. Johansen, T. Koch, J. Wengel, *Methods Mol Biol* **2005**, *288*, 127.

91 A. Eschenmoser, *Science* **1999**, *284*, 2118.

92 A. Eschenmoser, *Orig Life Evol Biosph* **2004**, *34*, 277.

93 B. Nawrot, M. Sobczak, M. Wojcik, M. Janicka, M. Nowak, M. Cypryk, W. J. Stec, *Oligonucleotides* **2006**, *16*, 68.

94 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566.

95 P. E. Nielsen, *Mol Biotechnol* **2004**, *26*, 233.

96 W. J. Stec, A. Grajkowski, A. Kobylanska, B. Karwowski, M. Koziolkiewicz, K. Misiura, A. Okruszek, A. Wilk, P. Guga, M. Boczkowska, *J. Am. Chem. Soc.* **1995**, *117*, 12019.

97 R. O. Dempcy, O. Almarsson, T. C. Bruice, *Proc Natl Acad Sci USA* **1994**, *91*, 7864.

98 I. E. Szabo, T. C. Bruice, *Bioorg Med Chem* **2004**, *12*, 4233.

99 A. Vater, S. Klussmann, *Curr Opin Drug Discov Devel* **2003**, *6*, 253.

100 D. Eulberg, S. Klussmann, *Chembiochem* **2003**, *4*, 979.

101 W. G. Purschke, D. Eulberg, K. Buchner, S. Vonhoff, S. Klussmann, *Proc Natl Acad Sci USA* **2006**, *103*, 5173.

102 W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004**, *427*, 618.

103 L. Jaeger, E. Westhof, N. B. Leontis, *Nucleic Acids Res* **2001**, *29*, 455.

104 L. Nasalean, S. Baudrey, N. B. Leontis, L. Jaeger, *Nucleic Acids Res* **2006**, *34*, 1381.

105 S. R. Das, R. Fong, J. A. Piccirilli, *Curr Opin Chem Biol* **2005**, *9*, 585.

106 S. P. Ryder, L. Ortoleva-Donnelly, A. B. Kosek, S. A. Strobel, *Methods Enzymol* **2000**, *317*, 92.

107 J. K. Soukup, N. Minakawa, A. Matsuda, S. A. Strobel, *Biochemistry* **2002**, *41*, 10426.

108 J. C. Cochrane, R. T. Batey, S. A. Strobel, *RNA* **2003**, *9*, 1282.

109 J. Chelliserrykattil, A. D. Ellington, *Nat Biotechnol* **2004**, *22*, 1155.

110 R. Padilla, R. Sousa, *Nucleic Acids Res* **2002**, *30*, e138.

111 P. E. Burmeister, S. D. Lewis, R. F. Silva, J. R. Preiss, L. R. Horwitz, P. S. Pendergrast, T. G. McCauley, J. C. Kurz, D. M. Epstein, C. Wilson, A. D. Keefe, *Chem Biol* **2005**, *12*, 25.

112 G. Xia, L. Chen, T. Sera, M. Fa, P. G. Schultz, F. E. Romesberg, *Proc Natl Acad Sci USA* **2002**, *99*, 6597.

113 M. Fa, A. Radeghieri, A. A. Henry, F. E. Romesberg, *J Am Chem Soc* **2004**, *126*, 1748.

114 F. J. Ghadessy, N. Ramsay, F. Boudsocq, D. Loakes, A. Brown, S. Iwai, A. Vaisman, R. Woodgate, P. Holliger, *Nat Biotechnol* **2004**, *22*, 755.

115 T. Ohbayashi, M. Kuwahara, M. Hasegawa, T. Kasamatsu, T. Tamura, H. Sawai, *Org Biomol Chem* **2005**, *3*, 2463.

116 R. Kawai, M. Kimoto, S. Ikeda, T. Mitsui, M. Endo, S. Yokoyama, I. Hirao, *J. Am. Chem. Soc.* **2005**, *127*, 17286.

117 M. J. Moore, C. C. Query, *Methods Enzymol* **2000**, *317*, 109.

118 K. E. Nielsen, J. Rasmussen, R. Kumar, J. Wengel, J. P. Jacobsen, M. Petersen, *Bioconjug Chem* **2004**, *15*, 449.

119 F. Darfeuille, J. B. Hansen, H. Orum, C. Di Primo, J. J. Toulme, *Nucleic Acids Res* **2004**, *32*, 3101.

120 V. Monjardet-Bas, J. C. Chottard, J. Kozelka, *Chemistry* **2002**, *8*, 1144.

121 K. L. Buchmueller, B. T. Hill, M. S. Platz, K. M. Weeks, *J Am Chem Soc* **2003**, *125*, 10850.

122 Z. Wang, K. Shah, T. M. Rana, *Biochemistry* **2001**, *40*, 6458.

123 T. Bando, A. Narita, I. Saito, H. Sugiyama, *J Am Chem Soc* **2003**, *125*, 3471.

124 S. M. Rink, D. L. Warner, A. Klapars, E. Vedejs, *Biochemistry* **2005**, *44*, 13981.

125 V. Jackson, *Methods* **1999**, *17*, 125.

126 G. Sengle, A. Jenne, P. S. Arora, B. Seelig, J. S. Nowick, A. Jaschke, M. Famulok, *Bioorg Med Chem* **2000**, *8*, 1317.

127 F. Hausch, A. Jaschke, *Bioconjug Chem* **1997**, *8*, 885.

128 B. Seelig, A. Jaschke, *Bioconjug Chem* **1999**, *10*, 371.

129 C. M. Gherghe, J. M. Krahn, K. M. Weeks, *J. Am. Chem. Soc.* **2005**, *127*, 13622.

130 E. J. Merino, K. A. Wilkinson, J. L. Coughlan, K. M. Weeks, *J Am Chem Soc* **2005**, *127*, 4223.

131 E. J. Merino, K. M. Weeks, *J Am Chem Soc* **2005**, *127*, 12766.

132 G. F. Joyce, *Annu Rev Biochem* **2004**, *73*, 791.

133 J. R. Lorsch, J. W. Szostak, *Acc Chem Res* **1996**, *29*, 103.

134 L. Gold, *Harvey Lect* **1995**, *91*, 47.

135 J. C. Cox, P. Rudolph, A. D. Ellington, *Biotechnol Prog* **1998**, *14*, 845.

136 J. C. Cox, M. Rajendran, T. Riedel, E. A. Davidson, L. J. Sooter, T. S. Bayer, M. Schmitz-Brown, A. D. Ellington, *Comb Chem High Throughput Screen* **2002**, *5*, 289.

137 M. Rajendran, A. D. Ellington, *Comb Chem High Throughput Screen* **2002**, *5*, 263.

138 T. S. Bayer, L. N. Booth, S. M. Knudsen, A. D. Ellington, *RNA* **2005**, *11*, 1848.

139 H. Zhang, A. Hamasaki, E. Toshiro, Y. Aoyama, Y. Ito, *Nucleic Acids Symp Ser* **2000**, 219.

140 D. Eulberg, K. Buchner, C. Maasch, S. Klussmann, *Nucleic Acids Res* **2005**, *33*, e45.

141 G. Hybarger, J. Bynum, R. F. Williams, J. J. Valdes, J. P. Chambers, *Anal Bioanal Chem* **2006**, *384*, 191.

142 M. Berezovski, M. Musheev, A. Drabovich, S. N. Krylov, *J Am Chem Soc* **2006**, *128*, 1410.

143 S. D. Mendonsa, M. T. Bowser, *J Am Chem Soc* **2004**, *126*, 20.

144 R. K. Mosing, S. D. Mendonsa, M. T. Bowser, *Anal Chem* **2005**, *77*, 6107.

145 M. Levy, K. E. Griswold, A. D. Ellington, *RNA* **2005**, *11*, 1555.

146 J. J. Agresti, B. T. Kelly, A. Jaschke, A. D. Griffiths, *Proc Natl Acad Sci USA* **2005**, *102*, 16170.

147 A. Aharoni, A. D. Griffiths, D. S. Tawfik, *Curr Opin Chem Biol* **2005**, *9*, 210.

148 A. C. Yan, K. M. Bell, M. M. Breeden, A. D. Ellington, *Front Biosci* **2005**, *10*, 1802.

149 K. V. Gothelf, T. H. LaBean, *Org Biomol Chem* **2005**, *3*, 4023.

150 N. C. Seeman, *Q Rev Biophys* **2005**, *38*, 361.

151 U. Feldkamp, C. M. Niemeyer, *Angew Chem Int Ed Engl* **2006**, *45*, 1856.

152 J. J. Birac, W. B. Sherman, J. Kopatsch, P. E. Constantinou, N. C. Seeman, *J Mol Graph Model* **2006**, *25*, 470.

153 L. Feng, S. H. Park, J. H. Reif, H. Yan, *Angew Chem Int Ed Engl* **2003**, *42*, 4342.

154 D. Liu, M. Wang, Z. Deng, R. Walulu, C. Mao, *J Am Chem Soc* **2004**, *126*, 2324.

155 J. Malo, J. C. Mitchell, C. Venien-Bryan, J. R. Harris, H. Wille, D. J. Sherratt, A. J. Turberfield, *Angew Chem Int Ed Engl* **2005**, *44*, 3057.

156 N. Chelyapov, Y. Brun, M. Gopalkrishnan, D. Reishus, B. Shaw, L. Adleman, *J Am Chem Soc* **2004**, *126*, 13924.

157 Y. He, Y. Chen, H. Liu, A. E. Ribbe, C. Mao, *J Am Chem Soc* **2005**, *127*, 12202.

158 Y. He, C. Mao, *Chem Commun (Camb)* **2006**, *9*, 968.

159 E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **1998**, *394*, 539.

160 X. Yang, L. A. Wenzler, J. Qi, X. Li, N. C. Seeman, *J. Am. Chem. Soc.* **1998**, *120*, 9779.

161 B. Ding, R. Sha, N. C. Seeman, *J. Am. Chem. Soc.* **2004**, *126*, 10230.

162 T. H. LaBean, H. Yan, J. Kopatsch, F. Liu, E. Winfree, J. H. Reif, N. C. Seeman, *J. Am. Chem. Soc.* **2000**, *122*, 1848.

163 D. Reishus, B. Shaw, Y. Brun, N. Chelyapov, L. Adleman, *J Am Chem Soc* **2005**, *127*, 17590.

164 Y. Ke, Y. Liu, J. Zhang, H. Yan, *J Am Chem Soc* **2006**, *128*, 4414.

165 Y. Liu, Y. Ke, H. Yan, *J Am Chem Soc* **2005**, *127*, 17140.

166 S. H. Park, R. Barish, H. Li, J. H. Reif, G. Finkelstein, H. Yan, T. H. Labean, *Nano Lett* **2005**, *5*, 693.

167 F. Mathieu, S. Liao, J. Kopatsch, T. Wang, C. Mao, N. C. Seeman, *Nano Lett.* **2005**, *5*, 661.

168 W. B. Sherman, N. C. Seeman, *Biophys J* **2006**, *90*, 4546.

169 R. P. Goodman, I. A. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, A. J. Turberfield, *Science* **2005**, *310*, 1661.

170 Z. Shen, H. Yan, T. Wang, N. C. Seeman, *J. Am. Chem. Soc.* **2004**, *126*, 1666.

171 P. J. Paukstelis, J. Nowakowski, J. J. Birktoft, N. C. Seeman, *Chem Biol* **2004**, *11*, 1119.

172 M. A. Batalia, E. Protozanova, R. B. Macgregor, D. A. Erie, *Nano Lett* **2002**, *2*, 269.

173 L. T. Costa, M. Kerkmann, G. Hartmann, S. Endres, P. M. Bisch, W. M. Heckl, S. Thalhammer, *Biochem Biophys Res Commun* **2004**, *313*, 1065.

174 P. W. Rothemund, *Nature* **2006**, *440*, 297.

175 H. G. Hansma, E. Oroudjev, S. Baudrey, L. Jaeger, *J Microsc* **2003**, *212*, 273.

176 L. Jaeger, N. B. Leontis, *Angew. Chemie. Int. Ed.* **2000**, *14*, 2521.

177 Y. Ikawa, K. Fukada, S. Watanabe, H. Shiraishi, T. Inoue, *Structure (Camb)* **2002**, *10*, 527.

178 S. Matsumura, Y. Ikawa, T. Inoue, *Nucleic Acids Res* **2003**, *31*, 5544.

179 B. Liu, S. Baudrey, L. Jaeger, G. C. Bazan, *J Am Chem Soc* **2004**, *126*, 4076.

180 S. Horiya, X. Li, G. Kawai, R. Saito, A. Katoh, K. Kobayashi, K. Harada, *Chem Biol* **2003**, *10*, 645.

181 J. H. Davis, M. Tonelli, L. G. Scott, L. Jaeger, J. R. Williamson, S. E. Butcher, *J Mol Biol* **2005**, *351*, 371.

182 D. Shu, L. P. Huang, S. Hoeprich, P. Guo, *J Nanosci Nanotechnol* **2003**, *3*, 295.

183 D. Shu, W. D. Moll, Z. Deng, C. Mao, P. Guo, *Nano Lett* **2004**, *4*, 1717

184 A. Khaled, S. Guo, F. Li, P. Guo, *Nano Lett* **2005**, *5*, 1797.

185 C. Mao, W. Sun, N. C. Seeman, *J. Am. Chem. Soc.* **1999**, *121*, 5437.

186 R. P. Goodman, R. M. Berry, A. J. Turberfield, *Chem Commun (Camb)* **2004**, 1372.

187 S. H. Park, C. Pistol, S. J. Ahn, J. H. Reif, A. R. Lebeck, C. Dwyer, T. H. LaBean, *Angew Chem Int Ed Engl* **2006**, *45*, 735.

188 L. Jaeger, A. Chworos, in *Foundations of Nanoscience 2005: Self-Assembled Architectures and Devices* (Ed.: J. H. Reif), Science Technica, 2005, pp. 157.

189 K. Lund, Y. Liu, S. Lindsay, H. Yan, *J Am Chem Soc* **2005**, *127*, 17606.

190 P. W. Rothemund, N. Papadakis, E. Winfree, *PLoS Biol* **2004**, *2*, e424.

191 R. D. Barish, P. W. Rothemund, E. Winfree, *Nano Lett* **2005**, *5*, 2586.

192 H. Yan, T. H. LaBean, L. Feng, J. H. Reif, *Proc Natl Acad Sci USA* **2003**, *100*, 8103.

193 H. Yan, S. H. Park, G. Finkelstein, J. H. Reif, T. H. LaBean, *Science* **2003**, *301*, 1882.

194 Y. He, Y. Tian, Y. Chen, Z. Deng, A. E. Ribbe, C. Mao, *Angew Chem Int Ed Engl* **2005**, *44*, 6694.

195 A. Carbone, N. C. Seeman, *Natural Computing* **2003**, *2*, 133.

196 C. Lin, Y. Liu, S. Rinker, H. Yan, *Chemphyschem* **2006**, *7*, 1641.

197 Y. Liu, C. Lin, H. Li, H. Yan, *Angew Chem Int Ed Engl* **2005**, *44*, 4333.

198 J. D. Le, Y. Pinto, N. C. Seeman, K. Musier-Forsyth, T. A. Taton, R. A. Kiehl, *Nano Lett.* **2004**, *4*, 2343.

199 J. Sharma, R. Chhabra, Y. Liu, Y. Ke, H. Yan, *Angew Chem Int Ed Engl* **2006**, *45*, 730.

200 J. Zhang, Y. Liu, Y. Ke, H. Yan, *Nano Lett* **2006**, *6*, 248.

201 J. Zheng, P. E. Constantinou, C. Micheel, A. P. Alivisatos, R. A. Kiehl, N. C. Seeman, *Nano Lett* **2006**, *6*, 1502.

202 N. C. Seeman, *Trends Biochem Sci* **2005**, *30*, 119.

203 A. V. Garibotti, S. M. Knudsen, A. D. Ellington, N. C. Seeman, *Nano Lett* **2006**, *6*, 1505.

204 C. Lin, E. Katilius, Y. Liu, J. Zhang, H. Yan, *Angew Chem Int Ed Engl* **2006**, *45*, 5296.

205 N. K. Navani, Y. Li, *Curr Opin Chem Biol* **2006**, *10*, 272.

206 H. Ulrich, *Handb Exp Pharmacol* **2006**, *173*, 305.

207 S. Hoeprich, Q. Zhou, S. Guo, D. Shu, G. Qi, Y. Wang, P. Guo, *Gene Ther* **2003**, *10*, 1258.

208 K. Keren, M. Krueger, R. Gilad, G. Ben-Yoseph, U. Sivan, E. Braun, *Science* **2002**, *297*, 72.

209 L. Berti, A. Alessandrini, P. Facci, *J. Am. Chem. Soc.* **2005**, *127*, 11216.

210 G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen, T. Carell, *J Am Chem Soc* **2006**, *128*, 1398.

211 C. A. Mirkin, *Inorg Chem* **2000**, *39*, 2258.

212 S. W. Chung, D. S. Ginger, M. W. Morales, Z. Zhang, V. Chandrasekhar, M. A. Ratner, C. A. Mirkin, *Small* **2005**, *1*, 64.

213 A. D. Bates, B. P. Callen, J. M. Cooper, R. Cosstick, C. Geary, A. Glidle, L. Jaeger, J. L. Pearson, M. Proupin-Perez, C. Xu, D. R. Cumming, *Nano Lett* **2006**, *6*, 445.

214 A. Y. Koyfman, G. Braun, S. Magonov, A. Chworos, N. O. Reich, L. Jaeger, *J Am Chem Soc* **2005**, *127*, 11886.

215 G. Braun, K. Inagaki, R. A. Estabrook, D. K. Wood, E. Levy, A. N. Cleland, G. F. Strouse, N. O. Reich, *Langmuir* **2005**, *21*, 10699.

216 T. Fukushima, T. Hayakawa, Y. Inoue, K. Miyazaki, Y. Okahata, *Biomaterials* **2004**, *25*, 5491.

217 Y. Hoshino, S. Tajima, H. Nakayama, Y. Okahata, *Macromol. Rapid Commun.* **2002**, *23*, 253.

218 L. Zhu, P. S. Lukeman, J. W. Canary, N. C. Seeman, *J. Am. Chem. Soc.* **2003**, *125*, 10178.

219 A. Serganov, S. Keiper, L. Malinina, V. Tereshko, E. Skripkin, C. Hobartner, A. Polonskaia, A. T. Phan, R. Wombacher, R. Micura, Z. Dauter, A. Jaschke, D. J. Patel, *Nat Struct Mol Biol* **2005**, *12*, 218.

220 R. G. Schultz, S. M. Gryaznov, *Nucl. Acids Res.* **1996**, *24*, 2966.

221 N. C. Pagratis, C. Bell, Y.-F. Chang, S. Jennings, T. Fitzwater, D. Jellinek, C. Dang, **1997**, *15*, 68.

222 A. Rhie, L. Kirby, N. Sayer, R. Wellesley, P. Disterer, I. Sylvester, A. Gill, J. Hope, W. James, A. Tahiri-Alaoui, *J. Biol. Chem.* **2003**, *278*, 39697.

11
Helically Folding Polymers

Eiji Yashima and Katsuhiro Maeda

11.1
Introduction

Macromolecular helicity is one of the most important and basic structural motifs that can often be seen in naturally occurring biological polymers, such as proteins and nucleic acids. They are optically active and fold into a one-handed helical conformation which directs their sophisticated and fundamental functions in living systems. Therefore, in polymer and supramolecular chemistry, the design and synthesis of polymers and oligomers that fold into a one-handed helical conformation has been attracting great attention not only to mimic biological helices, but also for their wide variety of possible applications in materials science, chemical sensing, separation of enantiomers and enantioselective catalysis [1–6].

Starting with the pioneering research by Pino et al. in the 1960s on the structural and chiroptical investigation of isotactic vinyl polymers prepared by the polymerization of α-olefins bearing optically active substituents [1, 7], a helical conformation with an excess helical sense was for the first time revealed in synthetic vinyl polymers in solution. Although the helical polyolefins are dynamic in nature and consist of short helical segments separated by frequently occurring helical reversals among disordered, random coil conformations, this was a significant milestone in the field of synthetic helical polymers, through which a variety of helical polymers and oligomers (foldamers) with a controlled helix-sense have been synthesized. The existing synthetic helical polymers that exhibit optical activity solely due to the helicity [8–12] can be basically classified into two categories with respect to their characteristics in their helical conformations; static (stable) helical polymers and dynamic helical polymers. The former has a rigid helix with a sufficiently high helix inversion barrier. Therefore, optically active helical polymers with an excess of a one helical sense can be prepared by the helix-sense selective polymerization of achiral or prochiral monomers with chiral catalysts or initiators under kinetic control. On the other hand, the dynamic helical polymers consist of interconvertible right- and left-handed helical conformations separated by helical reversals that can readily move along a polymer chain, so that stable

(static) one-handed helical polymers cannot be obtained. However, the helix inversion barriers in dynamic helical polymers are low, and therefore, a predominantly one-handed helical conformation can be induced in the presence of a small amount of chiral residue at the pendant or terminal ends or stimulant. Certain polymers, however, exhibit both features depending on the structures of the monomer units.

In this chapter, we focus on unique features of such static and dynamic helical polymers including the synthesis, structures and functions. The factors that stabilize the helical conformations of essentially all polymers in this chapter are π-conjugation and/or steric hindrance leading to restricted rotations about single bonds. These factors seem to be the same as those that prevail in the folding oligomers described in Chapter 1. Potential applications of synthetic helical polymers in chiral separation and asymmetric synthesis are also briefly described. Noncovalent supramolecular helical assemblies of small chiral molecules, an important area in supramolecular chemistry [13, 14], are beyond the scope of this chapter, but are described in Chapter 4.

11.2
Helical Polymers with High Helix Inversion Barriers (Static Helical Polymers)

When the helix inversion barriers of certain helical polymers are sufficiently high, optically active helical polymers can be synthesized either by the polymerization of optically active monomers or by the asymmetric (helix-sense selective) polymerization of achiral or prochiral monomers with chiral initiators or catalysts [1, 9, 10]. In these synthetic methods, the helical structures of the polymers that include a helix-sense and helical pitch are determined by chiral substituents covalently bonded to the polymer backbone or kinetically during the polymerization. Typical synthetic, static helical polymers exhibiting an optical activity mostly due to their one-handed helical structures are shown in Fig. 11.1. Poly(triphenylmethyl methacrylate) (PTrMA) (**1**), polychloral (**2**), polyisocyanides (**3**) and polyguanidines (**4**) belong to this category. These helical conformations are stable even in solution due to steric repulsion of the bulky side groups. This class of
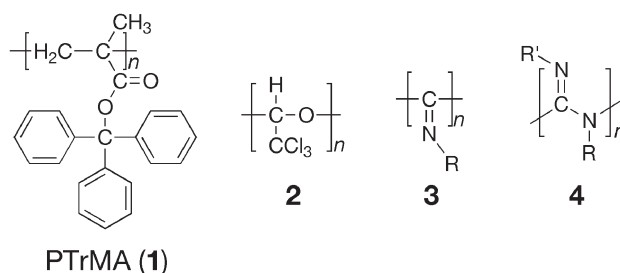
Fig. 11.1 Structures of static helical polymers.

static helical polymers has been extensively explored and thoroughly reviewed [1, 9, 10, 15, 16].

11.2.1
Poly(triarylmethyl methacrylate)s

PTrMA (**1**) is the first vinyl polymer prepared by the helix-sense selective polymerization of an achiral (prochiral) monomer, triphenylmethyl methacrylate (TrMA), using anionic initiators such as 9-fluorenyllithium (FlLi) complexed with chiral ligands such as (−)-sparteine (Sp), resulting in a single-handed, fully isotactic helical polymer with a large optical rotation ($[\alpha]_D$ = ca + 380°) (Fig. 11.2) [9, 17]. The chiral ligand controls the main-chain configuration as well as the helical sense of the polymer. In addition, the bulky triphenylmethyl groups restrict unfolding of the helical conformation produced through the polymerization reaction, so that the optical activity of PTrMA disappears when the triphenylmethyl groups are removed for conversion to methyl esters. The helical conformation of PTrMA was further evidenced by the optical resolution of the optically inactive PTrMA prepared by achiral anionic initiators into fractions showing opposite optical rotations by chiral chromatography. The optically active helical PTrMA shows a remarkable chiral recognition for a variety of racemic compounds, giving a practically useful chiral packing material for high-performance liquid chromatography (HPLC), as described later in Section 11.5.
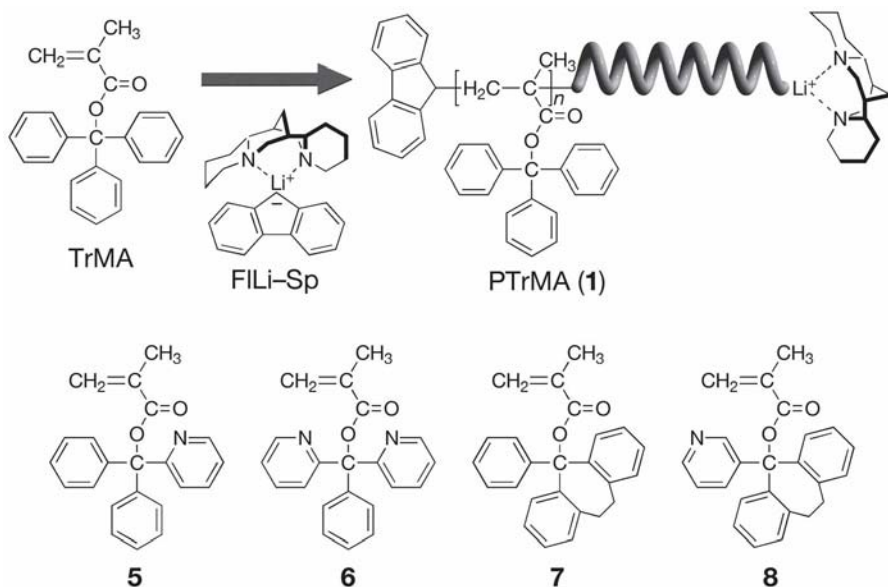
Fig. 11.2 Schematic representation of helix-sense selective anionic polymerization of TrMA and structures of methacrylates bearing a bulky ester group.

Okamoto et al. prepared a series of analogous helical poly(triarylmethyl methacrylate)s by the helix-sense selective polymerization of bulky methacrylates (**5–8**) and investigated their helical structures and chiral recognition abilities as well as the mechanism of the helix-sense selective polymerization. Although anionic polymerization techniques have often been used for the synthesis of the helical polymethacrylates, more versatile, free-radical polymerization has been proved to be an alternative way to produce a helical polymethacrylate from **7** [18]. The free-radical polymerization of **7** with α,α′-azobisisobutyronitrile (AIBN) in the presence of chiral chain transfer agents or cobalt (II) complexes (**9**) interacting with the growing radical [19] or in chiral solvents gave an almost perfect isotactic polymer with an excess of the one-handed helical sense [9]. Radical polymerization of *n*-butyl-substituted *N*-triphenylmethyl methacrylamide (**10**) in menthol produced a highly isotactic, optically active helical polymer soluble in chloroform, although the helix-sense excess of the polymer may not be as high as that of PTrMA [20].

9   **10**: R = C₄H₉
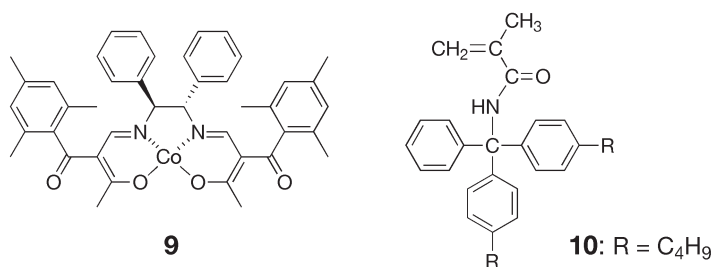
Scheme 11.1

A low molecular weight, one-handed helical poly(diphenyl-2-pyridyl methacrylate) prepared from **5** showed an unprecedented inversion of the macromolecular helicity in solution accompanied by a gradual decrease in the optical rotation with time. This change was ascribed to the helix–helix transition of the main chain, that is, the change from the one-handed helix to a mixture of right- and left-handed helices, as evidenced by the further chromatographic separation into fractions showing opposite optical rotations [21]. The activation energy ($\Delta G^{\ddagger}$) for the helix-sense inversion (racemization) was estimated to be 23 kcal mol^{-1}.

11.2.2
Polychloral

Isotactic polychloral (poly(trichloroacetaldehyde) (**2**), prepared by the helix-sense selective polymerization of chloral with optically active lithium alkoxides and carboxylates as initiators, possesses a 4/1 helical conformation and showed a high optical activity in films ($[\alpha]_D + 4000°$) [15]. Because the polymer is totally insolu-
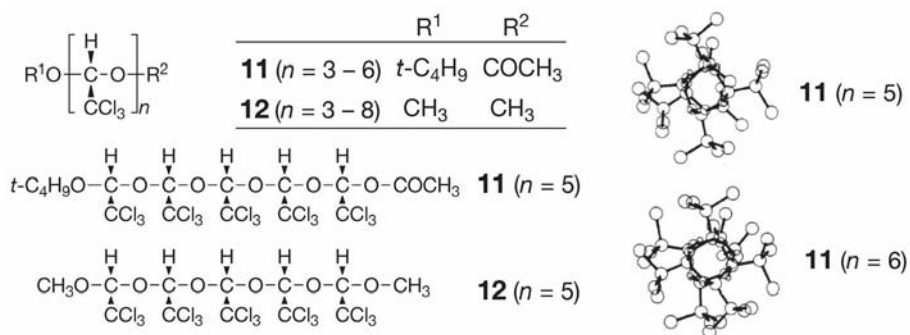
Fig. 11.3 Structures of isotactic chloral oligomers (**11** and **12**) and X-ray structures of 5mer and 6mer of **11**. (Reproduced with permission from Ref. 22. Copyright 1993 The Society of Polymer Science, Japan.)

ble in solvents, further structural explorations in solutions were hampered. However, the 4/1 helical structure in solution as well as in the crystalline state has been proven by NMR and crystallographic analyses of the isotactic uniform oligomers bearing different end groups **11** ($n = 3$–6) isolated by size-exclusion chromatography (SEC) from the oligomerization products, respectively. The individual oligomers were further resolved into enantiomers by chiral HPLC and the enantiomerically pure (−)-pentamer with the (R,R,R,R,R) main-chain configuration was found to adopt a right-handed 4/1 helical conformation by X-ray crystallographic determination (Fig. 11.3) [22]. The difference in the conformational energy (ΔG) of the (−)-pentamer between the diastereomeric right- and left-handed helices in equilibrium was estimated to be 2.7 kcal mol^{-1} based on molecular mechanics calculation, indicating that the right-handed helix is predominant in solution at ambient temperature. The isotactic symmetrical oligomers with the identical end groups **12** ($n = 3$–8) are the "*meso*" isomers and have no chirality in their primary structures. Nevertheless, the pentamer exhibited nonequivalent signals for all proton resonances in the solution NMR spectrum at 30 °C, resulting from an interconvertible enantiomeric helical conformation ($\Delta G = 0$) [23]. Based on variable temperature NMR experiments, the activation energy ($\Delta G^{\ddagger}$) for the helix-sense inversion (racemization) was determined to be 16.4 kcal mol^{-1}; the $\Delta G^{\ddagger}$ value further increased with an increase in the degree of polymerization (n) (19.6 kcal mol^{-1} for $n = 6$). More direct evidence for a stable helical conformation of **12** ($n = 6$–8) was attained by chromatographic separation of the "*meso*" oligomers into enantiomers by chiral HPLC using (+)-PTrMA (**1**) as the chiral stationary phase (CSP) [22]. These model studies using uniform oligomers produced in the polymerization mixture provide a deep insight into the stereochemistry of the helical polymers formed in the further propagation process including the stability of the one-handed helices and the kinetics and thermodynamics of the helix-sense inversion.

11.2.3
Polyisocyanides

Polyisocyanides with a bulky side group adopt a stable 4/1 helical conformation even in solution, as first postulated by Millich [24] and confirmed by Nolte et al. through the direct resolution of poly(*t*-butyl isocyanide) into enantiomeric helices by chiral chromatography [25]. The resolved polymer with a positive rotation was postulated to have a left-handed helical conformation based on a CD spectral analysis [26]. The stable helical conformation of the polyisocyanides was further confirmed by the helix-sense selective polymerization of achiral bulky isocyanides by Nolte [27] and Novak [28]. Nolte and coworkers and other groups have further synthesized wide varieties of helical polyisocyanides with a controlled helicity [10, 29] and these results will be described in detail in Chapter 12.

Although the helical structure of the polyisocyanides has been postulated to be a 4/1 helical conformation on the basis of an X-ray analysis, the absolute configuration of the helical polyisocyanides remains obscure. Advanced microscopy techniques, in particular, atomic force microscopy (AFM) combined with circular dichroism (CD) spectroscopy can reveal the structures of the helical polymers and their helix-senses. Diastereomeric right- and left-handed helical polyisocyanides were prepared from an unprecedented helix-sense controlled polymerization of enantiomerically pure phenyl isocyanides bearing an L- or D-alanine pendant with a long alkyl chain using a nickel catalyst in different solvents (**13**). High-resolution AFM revealed their helical conformations and enabled the determination of the helical sense (Fig. 11.4); poly(phenyl isocyanide)s showing a positive
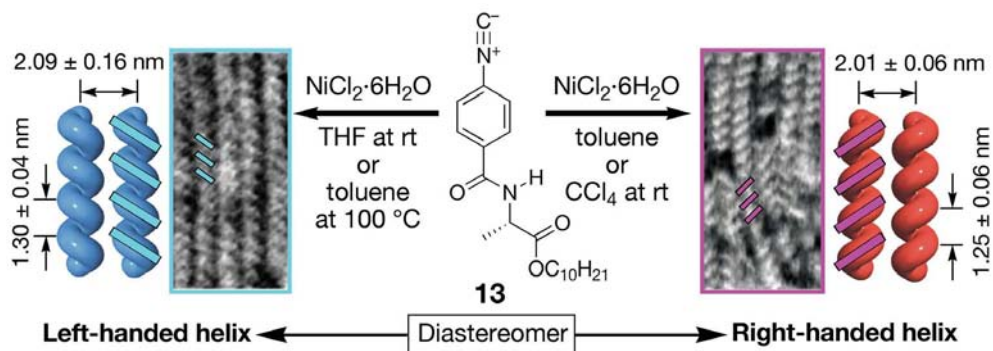
Fig. 11.4 Schematic illustration of diastereomeric helical polyisocyanides produced by the helix-sense controlled polymerization of **13**. The helix-sense can be controlled by the solvent polarity and temperature during the polymerization, resulting in the formation of diastereomeric helical polyisocyanides. The helix-senses of the diastereomeric **13**s were determined by their AFM measurements. AFM phase images of self-assembled **13** on graphite (scale = 10 × 20 nm) with the left-handed (left) and right-handed helical **13** (right) together with their structures determined by X-ray are also shown. (Reproduced with permission from Ref. 30. Copyright 2006 American Chemical Society.)

first Cotton effect sign was assigned to have a right-handed helix [30]. This assignment agrees with that determined by the exciton-coupled CD method [31].

Poly(2,3-quinoxaline)s, which structurally resemble polyisocyanides by condensing two adjacent imine units in each heteroaromatic moiety, have been obtained by the polymerization of 1,2-diisocyanobenzene (**14**) using an organopalladium complex with an optically active imidazoline group (**15**) as the initiator produced a right-handed helical poly(2,3-quinoxaline) with an almost 100% helix-sense selectivity via a living and cyclopolymerization mechanism (Scheme 11.2) [16, 32]. The helical structure and handedness were postulated by X-ray crystallographic analysis of an active pentamer of a diastereomerically pure oligo(2,3-quinoxaline). In sharp contrast to other living polymerization systems, the active growing chains complexed with the palladium can be isolated and subsequent helix-sense selective block polymerization takes place [16].
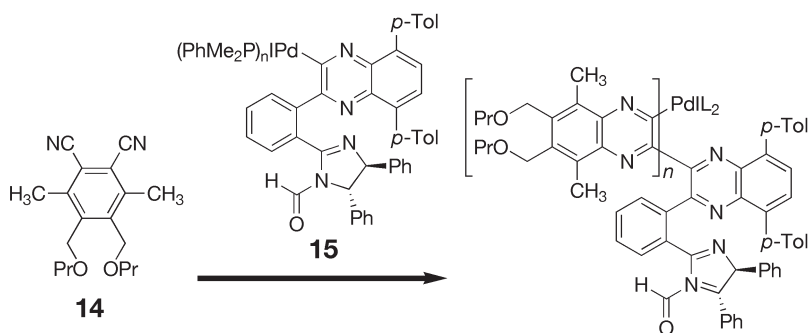
Scheme 11.2

11.2.4
Polyguanidines

Polyguanidines (**4**) prepared by the polymerization of carbodiimides were previously considered to be an interesting class of dynamic helical polymers (Fig. 11.5A) (see Section 11.3.2). Novak et al. found that an optically active polyguanidine (**16**) showed a specific rotation identical to its monomer ($[\alpha]_D = +7.5°$) just after the polymerization. However, the specific rotation significantly increased in solution and reached a plateau value ($[\alpha]_D = -157.5°$) upon annealing at elevated temperatures [33]. This unusual behavior was irreversible and ascribed to the conformational change from a kinetically controlled structure to a thermodynamically controlled helical one with the excess helical sense upon heating. An excess of one helical sense can be induced for an optically inactive polyguanidine (**17**) catalyzed by optically active camphorsulfonic acid [33]. Novak and coworkers further synthesized an optically active helical polyguanidine stable in solution by the
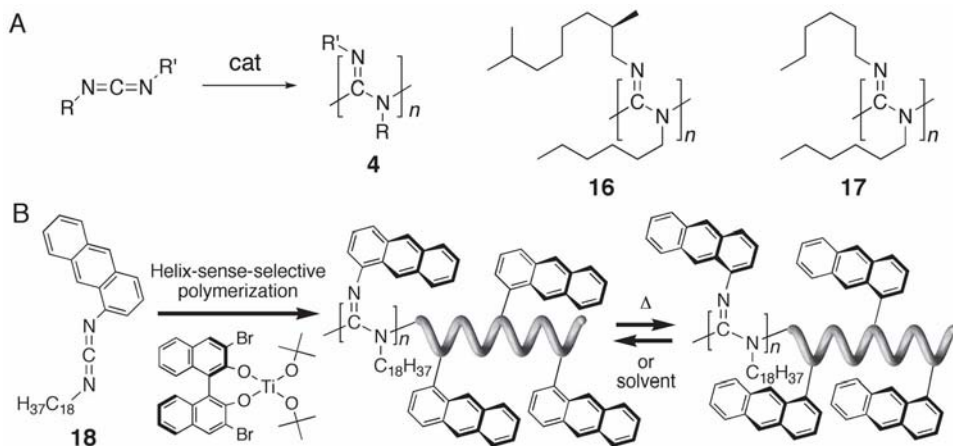
Fig. 11.5 (A) Synthetic scheme of polyguanidines. (B) Schematic illustration of the synthesis of a helical polyguanidine by helix-sense selective polymerization of achiral **18** using a chiral Ti catalyst and reversible switching of the pendant anthracene rings.

helix-sense selective polymerization of an achiral bulky carbodiimide (**18**) with a chiral titanium complex (Fig. 11.5B) [34]. This polymer exhibited a reversible, temperature and solvent-induced chiroptical switch due to a change in the orientation of the pendant anthracene rings as evidenced by vibrational CD measurements, while the main-chain helicity remained unchanged.

11.3
Helical Polymers with Low Helix Inversion Barriers (Dynamic Helical Polymers)

Green and coworkers have thoroughly investigated the structures and chiroptical properties of polyisocyanates, a typical stiff, rigid rod-like polymer with a long persistent length (q), and the substantial nature of the dynamic macromolecular helicity of polyisocyanates has been experimentally and theoretically revealed [8, 35]. The most important feature of dynamic helical polymers, such as polyisocyanates, polysilanes and polyacetylenes, is their high sensitivity to a chiral environment, and therefore, a small chiral bias can be transformed into a main-chain conformational change with a large amplification through covalent or noncovalent bonding with high cooperativity, resulting in the formation of an excess of the preferred helical sense. Such systems provide the basis for the construction of novel chirality-sensing materials. The underlying principle observed in dynamic helical polyisocyanates may be universal and applicable to other polymeric and supramolecular systems [13, 14, 36].

11.3.1
Dynamic Helical Polymers Assisted by Covalent Bonding

11.3.1.1   Polyisocyanates

Polyisocyanates are characterized by an *N*-substituted amide repeat unit (Nylon-1) and possess a helical conformation (8/3 helix) rather than a restricted coplanar conformation. The conjugated partial double-bond characteristic of the backbone amide bonds is responsible for their stiffness. Even the optically inactive poly(*n*-hexyl isocyanate) (**19**, *q* = 20–40 nm) and poly(2-butylhexyl isocyanate) (**20**), which have no stereogenic centers, consist of an equal mixture of interconvertible right- and left-handed helical conformations separated by the rarely occurring helical reversals (Fig. 11.6A). Therefore, helical polyisocyanates in dynamic equilibrium are chiral (or dynamically racemic) macromolecules. However, the helix inversion
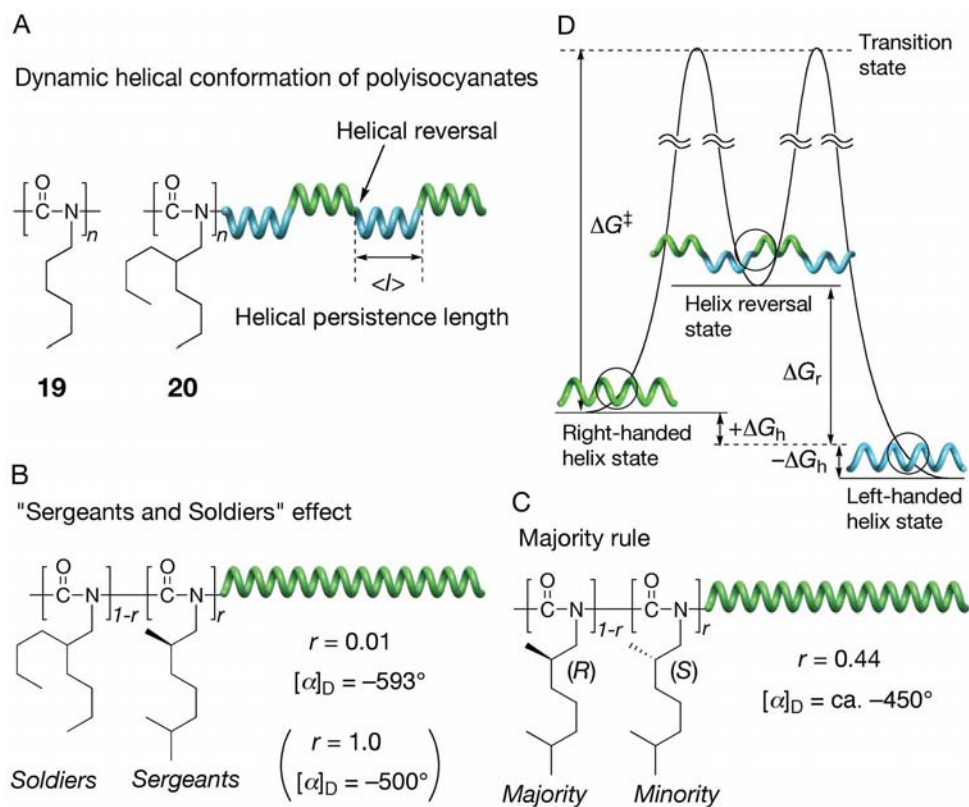
Fig. 11.6 Schematic illustration of dynamic helical conformation of polyisocyanates (A), "sergeants and soldiers" effect (B), majority rule (C), and energy diagram of dynamic helical polymers (D).

barriers are very small, so that optically active polyisocyanates with an excess single-handed helix can be obtained through the copolymerization of achiral monomers using a small amount of optically active monomers (less than 1 mol%) [37, 38] or polymerization of achiral isocyanates with optically active initiators [39]. This can be considered as a typical example of chiral amplification in a polymer. This highly cooperative phenomenon is called the ''sergeants and soldiers effect'' (Fig. 11.6B). The underlying principle for this unique chiral amplification phenomenon was theoretically and quantitatively solved using a statistical theory, where each monomer unit in the helical polymer chains can take either a right-handed helical state, left-handed helical state or helix reversal state [8, 35, 40]. According to Lifson, Green, Teramoto and coworkers, the helix-sense excess of the preferred helical state in helical homopolymers, such as a deuterium-substituted helical polyisocyanate (**21**, Fig. 11.7), can be calculated as a function of the thermodynamic stability parameters, the free energy difference between the right- and left-handed helical states ($2\Delta G_h$), the excess free energy of the helical reversal state (ΔG_r) (per monomer unit), the degree of polymerization (N), and the absolute temperature (Fig. 11.6D) [41]. The key energy parameters ($2\Delta G_h$ and ΔG_r) arising from **21** were estimated to be 0.74 cal mol^{-1} and 3.9 kcal mol^{-1} on a monomer unit basis in hexane at 25 °C, respectively. Importantly, the former value is about three orders of magnitude smaller than the latter. The $2\Delta G_h$ value indicates that **21** favors the right-handed helix over the left-handed helix only by 0.12%, whereas for a longer polymer chain of **21** ($N = 2000$), this minute excess is remarkably amplified by the cooperative mechanism to 67:33, which results in the appearance of intense Cotton effects in the polymer backbone region (Fig. 11.7) as well as a large optical rotation [42]. In addition, the helix reversal costs 3.9 kcal mol^{-1} and appears only once in every 762 monomer
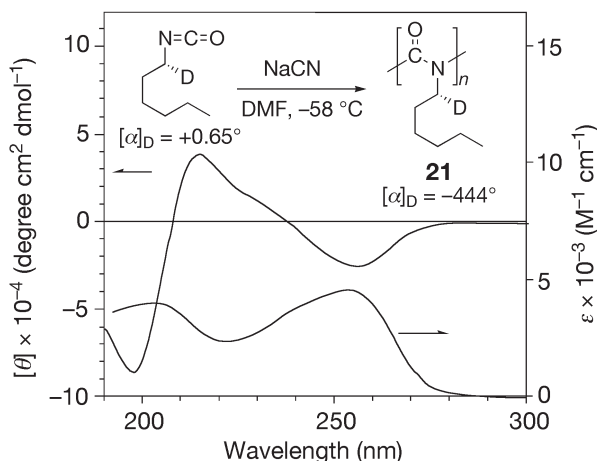
Fig. 11.7 CD and absorption spectra of **21** in hexane. (Reprinted with permission from Ref. 42. Copyright 1988 American Chemical Society.)

units on average. This long helical persistence length (l) arising from the relatively large ΔG_r value directly connects to the observed chiral amplification in polyisocyanates, because the chiral bias $2\Delta G_h$ is multiplied by the number of monomer units between the helical reversals (l), so that many units likely take the same helical sense within the polymer chain. In this way, the small chiral bias of each unit of the polymer chain is significantly amplified [35, 41].

Based on the variable temperature NMR experiments of **20**, the activation energy ($\Delta G^{\neq}$) for the interconversion process between the right- and left-handed helical conformations ($2\Delta G_h = 0$ in this case) was determined to be 19 kcal mol^{-1}, which is significantly greater than the thermodynamic excess energy of the helical reversal ($\Delta G_r = 5$ kcal mol^{-1}) [43]. This result supports the fact that inversion of the helix readily occurs at ambient temperature, but raises a question about the boundary between the static and dynamic helical conformations. Chiral solvation, while its chiral bias ($2\Delta G_h$) seems to be very weak (0.04 cal mol^{-1}), can also be used to induce a helical conformation with a preferential screw-sense in the dynamically racemic poly(n-hexyl isocyanate) (**19**) in nonracemic solvents [44]. Green et al. further demonstrated that the copolymers of the isocyanates composed of a mixture of (R)- and (S)-enantiomers with a small enantiomeric excess (ee) also form an excess single-handed helical conformation (Fig. 11.6C) [45]. The minority units obey the helical sense of the majority units in order to avoid introducing energetic helical reversals. They termed this phenomenon the ''majority rule''.

11.3.1.2 Polysilanes

Polysilanes (**22–26**) are also dynamic helical polymers like polyisocyanates with essentially a 7/3 helix, but different from polyisocyanates with respect to their unique chromophoric and fluorophoric Si σ-conjugated backbones. Wide varieties of optically active polysilanes bearing chiral alkyl or aromatic pendants and co-polymers with achiral monomers have been synthesized by Fujiki and cowork-ers [11, 46]. The structures of the pendant groups, in particular the chain length and position of the branching methyl group at the chiral center, significantly influence the rigidity of the polymer backbones ($q = 6$ (**23**), 70 (**25**), 85 (**24**) and

22

23
(q = 6 nm)

24
(q = 85 nm)

25
(q = 70 nm)

26
(q = 103 nm)
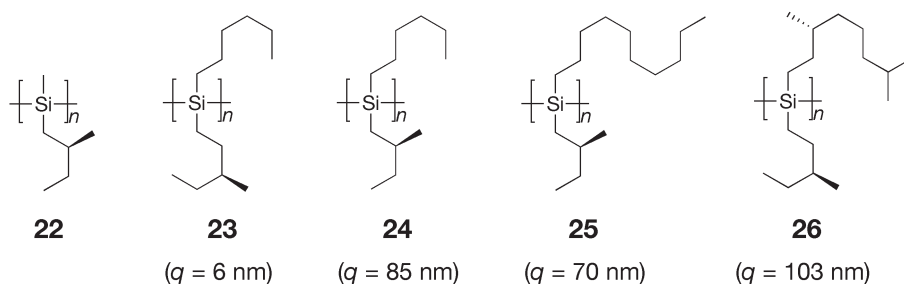
Scheme 11.3

103 nm (**26**)), the thermodynamic stability of the helical conformations and their electric properties including the absorption, CD and fluorescence spectral profiles. The chiroptical properties of helical polysilanes obey the sergeants and soldiers principle and majority rule [46, 47].

11.3.1.3 Polyacetylenes

A large number of π-conjugated, dynamic helical polyacetylenes have also been prepared by the polymerization of phenylacetylenes (**27–32**) [48–51], propiolic esters (**33**) [52] and *N*-propargylamides (**34**) [53, 54] bearing optically active substituents or by copolymerization with achiral acetylenes. Rhodium catalysts, such as [Rh(nbd)Cl]$_2$ (nbd: norbornadiene), are often used to produce stereoregular (*cis-transoidal*) polyacetylenes [55], resulting in the formation of a twisted helical structure. The stereoregularities are essential for the induction of a helical conformation [56, 57]. In contrast to stiff rodlike helical polyisocyanates and polysilanes, helical polyacetylenes are rather flexible, and the reported *q* values are 8.6 and 13.5 nm for a helical poly(4-carboxyphenylacetylene) induced by chiral amines [58] and poly(*N*-propargyl-2-ethylhexylamide) (**34a**) [59], respectively. However,
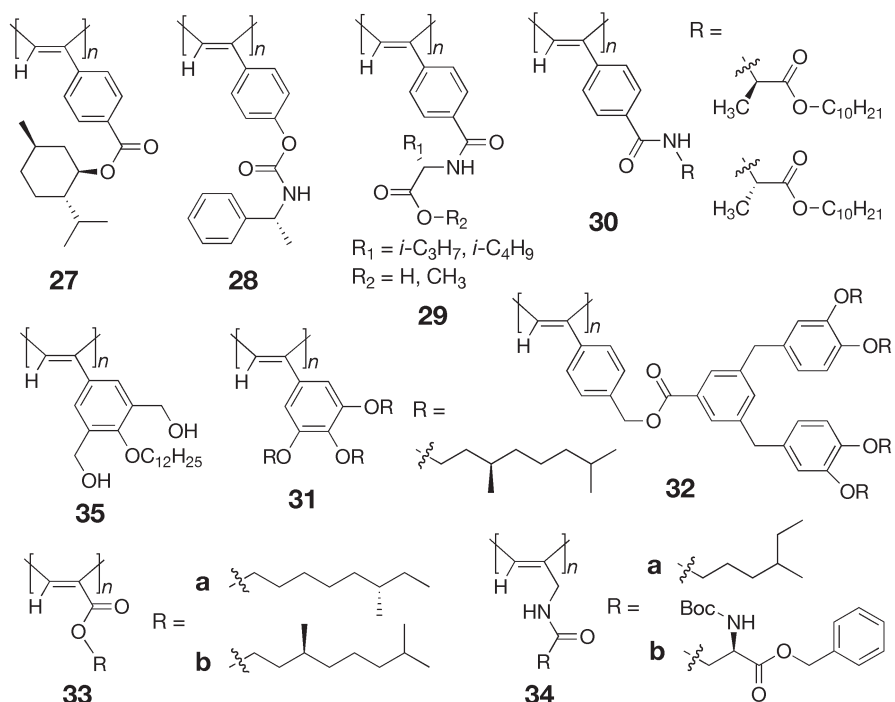
Scheme 11.4

the temperature-dependent changes in the induced circular dichroism (ICD) intensities of a series of homopolymers and copolymers of phenylacetylenes and *N*-propargylamides showed that their ΔG_r values (ca. 3.7 kcal mol^{-1}) are close to or slightly greater than those for polyisocyanates and poly(dialkylsilane)s, indicating that the polyacetylenes consist of long one helical-sense domains (ca. 660 monomer units) separated by rarely occurring helical reversals [60, 61]. Therefore, a similar chiral amplification (sergeants and soldiers effect and majority rule) also takes place in polyacetylenes, although the amplification efficiency was rather low for the covalent systems [52, 61].

Tang et al. and Masuda et al. synthesized helical polyacetylenes bearing various amino acids as the pendants, and their chiroptical properties including their helical conformations and helicity inversion (see Section 11.4) were investigated [49, 53, 62]. The intramolecular hydrogen bonds between the pendant groups appear to induce and stabilize the helical structure, although the hydrogen-bonded poly(*N*-propargylamide) (**34a**) is still semiflexible judging from its short persistence length ($q = 13.5$ nm) in chloroform [59]. An exceptionally stiff helical poly(phenylacetylene) was obtained by the polymerization of phenylacetylenes bearing an L- or D-alanine residue with a long alkyl chain as the pendant (**30**). The resulting *cis-transoidal* poly(phenylacetylene)s form a well-defined lyotropic cholesteric liquid crystalline (LC) phase in concentrated organic solvents based on the main-chain stiffness, which was confirmed by their long persistence lengths of around 40 nm in chloroform [63]. The L- or D-**30** undergoes an inversion of helicity in polar and nonpolar solvents accompanied with a dramatic change in its persistence length (q) from 135 nm in CCl$_4$ to 19 nm in tetrahydrofuran (THF); the former value is the highest among all synthetic helical polymers to the best of our knowledge. The macromolecular helicity inversion process can be directly followed by AFM (see Section 11.4).

Aoki et al. reported that an achiral phenylacetylene bearing two hydroxy groups on the phenyl residue gave an optically active poly(phenylacetylene) (**35**) when polymerized with a rhodium catalyst in the presence of (*S*)- or (*R*)-1-phenylethylamine. The resulting optically active polymer showing an ICD may have an excess of the preferred helical sense stabilized by intramolecular hydrogen bonds, and was stable in chloroform at high temperatures, but the CD disappeared in the presence of dimethyl sulfoxide (DMSO) [64].

Quite recently, the right- and left-handed helical structures of **30** have been directly observed using AFM. Rigid rodlike helical polyacetylenes were found to hierarchically self-assemble on highly oriented pyrolytic graphite (HOPG) upon exposure to benzene vapors (Fig. 11.8). Flat polyacetylene monolayers epitaxially formed on the basal plane of the graphite, on which helical polyacetylenes further self-assembled into chiral two-dimensional (2-D) helix-bundles with controlled helicity [65]. These AFM observations combined with the X-ray structural analysis suggest that the helices of the L-**30** and D-**30** single chains are enantiomers, and both the enantiomeric left- and right-handed helical **30** showing opposite Cotton effect signs (Fig. 11.8A) provide the enantiomorphic 2D structures on graphite.

A

C First layer (top view)

B

Left-handed helix   Right-handed helix
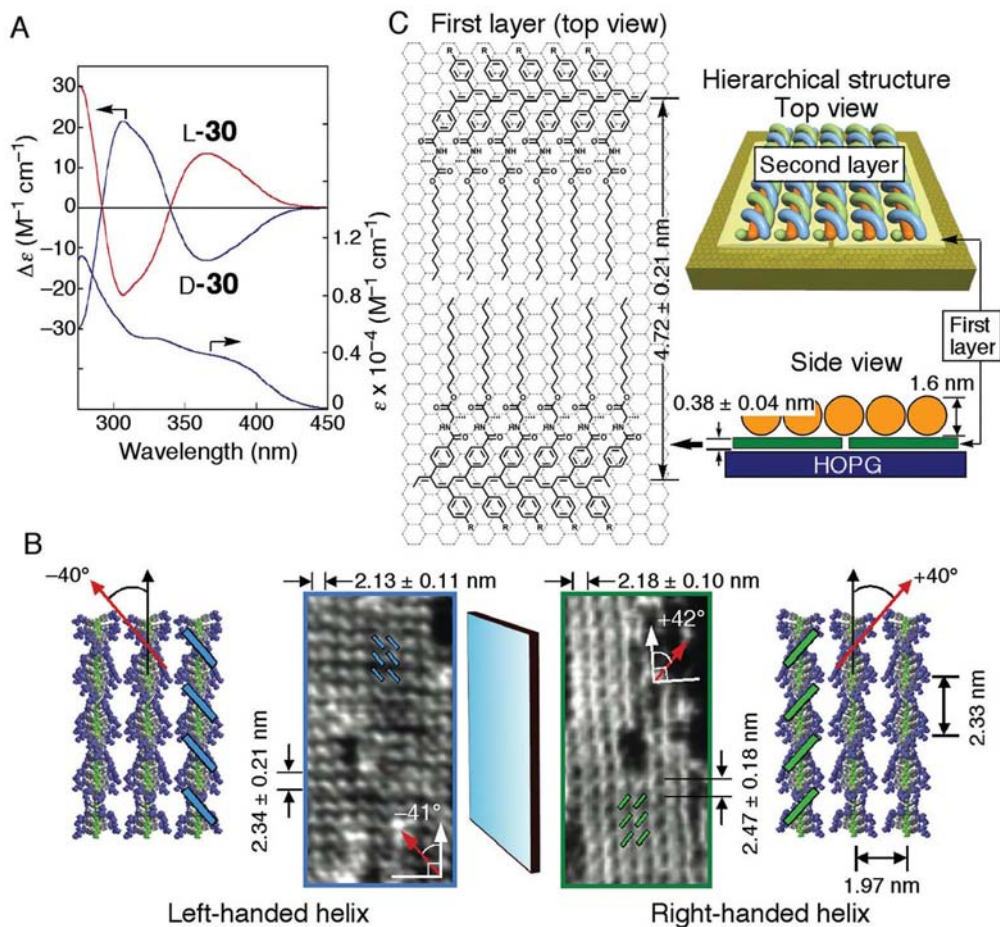
Fig. 11.8 (A) CD and absorption spectra of L- and D-**30** in benzene. (B) AFM phase images of L- and D-**30** (scale: 20 × 40 nm) with schematic drawings of the mirror-image relationship of helical L- and D-**30** 2-D crystals with antipodal oblique pendant arrangements. Possible models (left and right) were constructed on the basis of the X-ray structural analysis. (C) Schematic representation of the hierarchical structure of the self-assembled L-**30** on HOPG. (Reproduced with permission from Ref. 65. Copyright 2006 Wiley-VCH.)

11.3.2
Dynamic Helical Polymers Assisted by Noncovalent Bonding

Macromolecular helicity with an excess helical sense can also be induced in optically inactive, dynamically racemic helical polymers through specific noncovalent bonding interactions. This section mainly describes the helicity induction in poly(phenylacetylene)s through noncovalent chiral interactions.
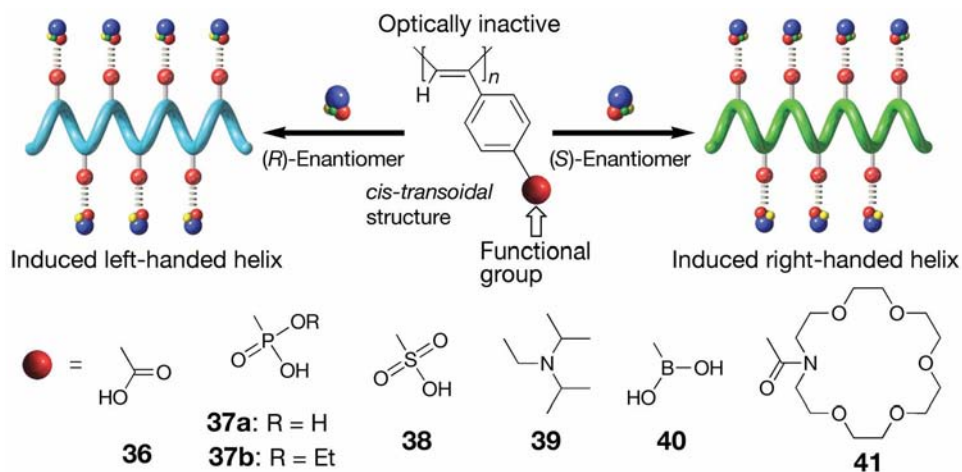
Fig. 11.9 Schematic illustration of helicity induction in poly(phenyl-acetylene)s bearing various functional groups (**36**–**41**) upon complexation with chiral compounds.

11.3.2.1 Induced Helical Poly(phenylacetylene)s

A *cis-transoidal*, stereoregular poly((4-carboxyphenyl)acetylene) (**36** in Fig. 11.9) was the first example of such a one-handed helical polymer induced by noncovalent chiral acid–base interactions [66]. Upon complexation with chiral amines in DMSO, a dynamic one-handed helical conformation is immediately induced in the polymer, resulting in the appearance of a characteristic ICD in the long wave-length region of the polymer backbone (300–500 nm). The typical CD spectra of **36** in the presence of various optically active amines (**42**–**46**) in DMSO are shown in Fig. 11.10. The remarkable CD induction arises from a drastic change in the population of the interconvertible right- and left-handed helices of the polymer. Primary amines and amino alcohols of the same configuration give the same sign of the ICDs, reflecting the helix-sense of **36**, and therefore, the Cotton effect sign of **36** can be used as a probe for sensing the chirality of various primary chiral amines. The magnitude of the ICD tends to increase with the increasing bulkiness of the amines [67]. In sharp contrast to the previously mentioned static and dynamic helical polymers assisted by covalent bonding, the helix-sense can be controlled by the chirality of small chiral molecules after polymerization.

Taking advantage of this helicity induction concept, a variety of chirality-responsive poly(phenylacetylene)s (**37**–**41**) has been synthesized by introducing a functional group as the pendant (Fig. 11.9) [68–72]. The sergeants and soldiers and majority rule effects are also observed for the noncovalent helicity induction in the poly(phenylacetylene)s [12, 67]. Among the poly(phenylacetylene)s prepared so far, a poly(phenylacetylene) (**41**) bearing the bulky aza-18-crown-6-ether, a typical host molecule in host–guest chemistry, as the functional pendant group was the most sensitive to the chirality of chiral molecules, such as amino acids,
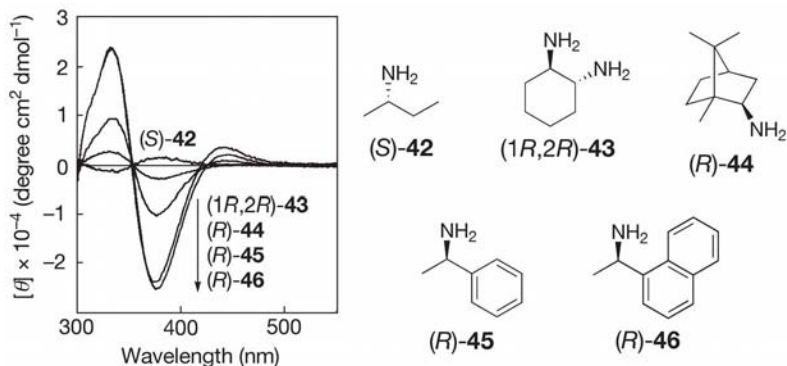
Fig. 11.10 CD spectra of **36** upon complexation with chiral amines in DMSO. (Reproduced with permission from Ref. 67. Copyright 1997 American Chemical Society.)

and an almost one-handed helix was induced in **41** in the presence of 0.1 equiv. L-alanine (L-Ala) in acetonitrile (Fig. 11.11B) [72]. This extremely high sensitivity may be ascribed to the rigidity of the polymer backbone by the bulky pendant group, which may increase the helical segments separated by rarely occurring helix reversals. In addition, **41** showed an apparent ICD even with 0.01 equiv. of L-Ala, indicating a remarkable chiral amplification. Moreover, **41** showed the same Cotton effect signs upon complexation with all the common 19 L-amino acids, indicating that **41** is among the most sensitive and practically useful syn-
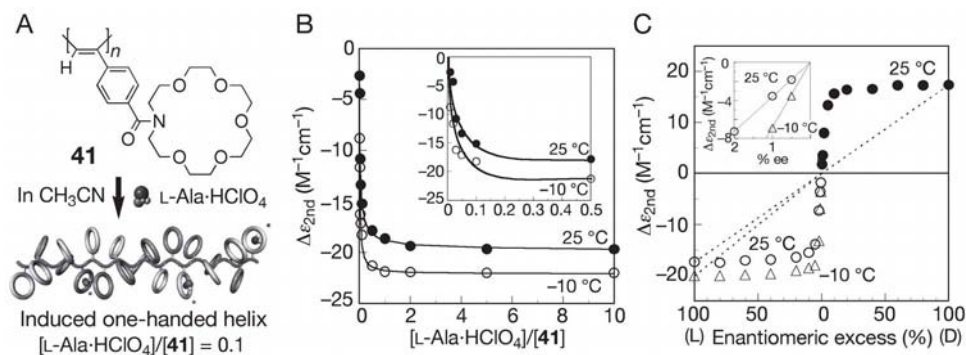
Fig. 11.11 (A) Schematic illustration of helicity induction on **41** with a small amount of L-Ala·HClO$_4$. (B) Titration curves of **41** with L-Ala·HClO$_4$ in acetonitrile at 25 °C and −10 °C. (C) Changes in ICD intensity ($\Delta\varepsilon_{2nd}$) of **41** vs. the % ee of L-Ala·HClO$_4$ during the complexation with **41** in acetonitrile at 25 °C and −10 °C. (Reprinted with permission from Ref. 72. Copyright 2003 American Chemical Society.)

thetic receptors for detecting the amino acid chirality. More interestingly, even a 5% ee of Ala produced the full ICD in **41** as induced by the optically pure Ala (Fig. 11.11C). This majority rule effect of **41** enabled detecting an extremely small enantiomeric imbalance in the amino acids, for instance, Ala of less than 0.005% ee, showing an apparent ICD without derivatization.

The chiral recognition of charged biomolecules with synthetic receptor molecules in water through polar interactions is an attractive challenge, but still remains a very difficult problem, which prompted us to explore the one-handed helicity induction in chromophoric water-soluble **36–39** in water. These are polyelectrolytes and can interact with a variety of charged and noncharged biomolecules involving amino acids, aminosugars, carbohydrates and peptides in water due to an ion condensation effect [73] and hydrophobic interactions, and the complexes exhibited characteristic ICDs in the UV-vis regions [74, 75]. Among the polyelectrolytes, **37b** bearing the bulky phosphonate group as the pendant is the most sensitive induced helical polymer in water; the assay of 19 of the common free L-amino acids produced the ICDs of **37b** with the same Cotton effect signs, which demonstrates that the polyelectrolyte is indeed the first powerful chirality-sensing probe in water.

11.3.2.2 Hierarchical Amplification of Helical-Sense Excess in Liquid Crystals

A positively charged polyelectrolyte, the hydrochloride of **39** (**39**-HCl) bearing an ammonium group, also formed an excess helical sense in the presence of various chiral acids such as **47** by a significant amplification of the chirality in water [76]. The polyelectrolyte function of the **39**-HCl is crucial for such a high chiral amplification in water, because the neutral **39** is not sensitive to the chirality of chiral acids in organic media [70] and requires a large excess amount of chiral acids to exhibit the full ICD.

During the intensive exploration of the chirality amplification mechanism of **39**-HCl, the polymer was found to form a lyotropic, nematic LC in concentrated water (>8 wt%) and the nematic LC phase converted to the cholesteric counterpart by doping with a tiny amount of optically active acids such as (*S*)-**47** or **47** with a low ee [77]. This liquid crystallinity of **39**-HCl is based on the main-chain stiffness in water as evidenced by its long persistence length (*q*) of 26.2 and 28 nm in the nematic and cholesteric LC states, respectively [78]. Interestingly, the helix-sense excess of **39**-HCl induced by (*S*)-**47** in dilute solution was further amplified in the LC state. The addition of increasing amounts of (*S*)-**47** results in a tightening of the cholesteric helical pitch, that reached an almost constant value at 0.05 equiv. of (*S*)-**47** in the LC state, whereas a larger amount of (*S*)-**47** (ca 0.3 equiv.) was required in dilute solution for the full ICD. The **39**-HCl exhibited a clear cholesteric LC phase showing well-defined fingerprint patterns even in the presence of 0.0005 equiv. of (*S*)-**47** (Fig. 11.12B) and 0.1 equiv. of (*S*)-**47** with 5% ee (Fig. 11.12C). In addition, **39**-HCl exhibited a strong majority rule effect for the ee of **47** in the LC phase; the helical pitch decreased with the increasing ee and reached a constant value at about 10% ee. On the other hand, in dilute solution, the ICD value became constant at over 60% ee (Fig. 11.12A). Direct evidence for
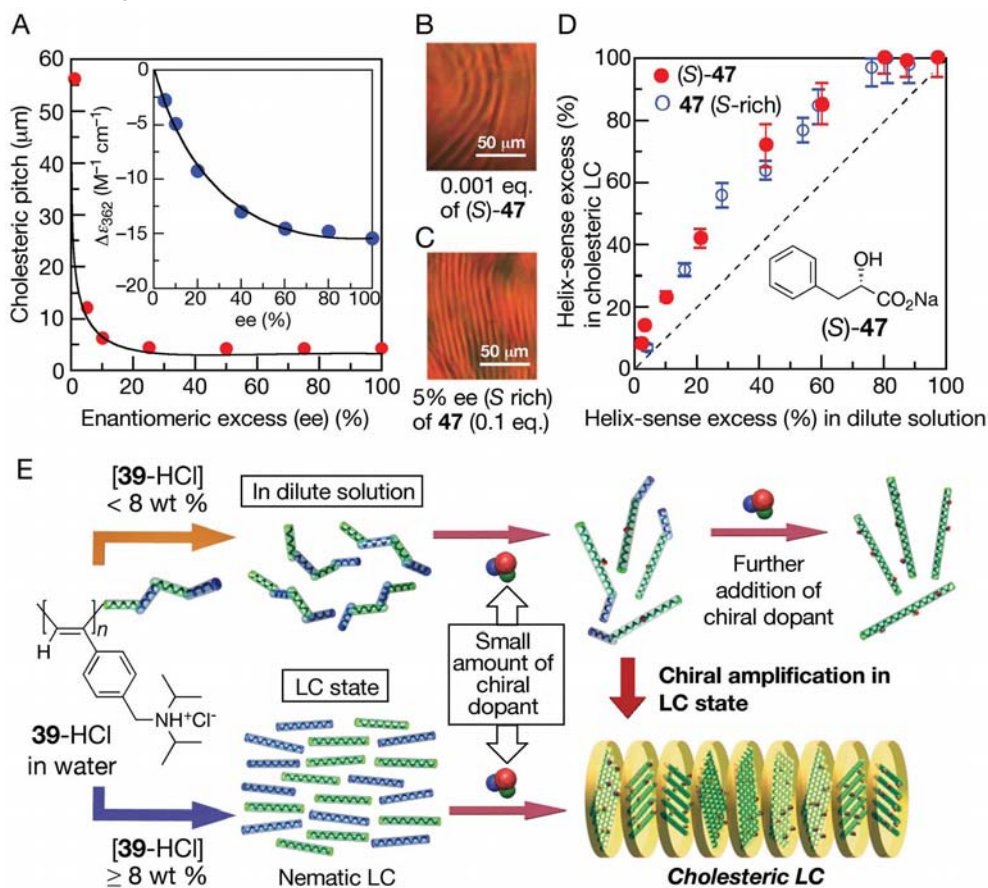
Fig. 11.12 (A) Changes in the cholesteric pitch and ICD intensity of **39**-HCl versus the % ee of **47** (*S* rich) in concentrated (20 wt%) and dilute (inset, 1 mg mL^{-1}) water solutions. (B, C) Polarized optical micrographs of cholesteric liquid crystalline phases of **39**-HCl (20 wt%) in the presence of 0.001 equivalent of (*S*)-**47** and 5% ee (*S* rich) of **47** (0.1 equivalent) in water. (D) Plots of the calculated % ee of helical sense-excess values of **39**-HCl a chiral dopant ((*S*)-**47** (red) and **47** (*S* rich) (blue) in the cholesteric LC state versus those in dilute water. The helical sense-excess values of **39**-HCl in dilute and concentrated LC water solutions were calculated using the maximum $\Delta\varepsilon_{2nd}$ and q_c values of -17.2 and 1.55 as the base values, respectively; $q_c = (2\pi/\text{cholesteric pitch})$. (E) Schematic illustration of hierarchical chiral amplification in macromolecular helicity of **39**-HCl in dilute solution and LC state. (Reproduced with permission from Ref. 78. Copyright 2006 American Chemical Society.)

the hierarchical amplification process of the helical sense excess of **39**-HCl during the cholesteric LC formation was demonstrated by direct comparison of the excess of the one helical sense of the polymer in dilute solution with that in the cholesteric LC state (Fig. 11.12D) [78]. In the LC state, the population of the

helical reversals between the interconverting right- and left-handed helical segments of the polymer may be reduced when compared to that in dilute solution, because the kinked helical polymer chain could likely interfere with the close parallel packing of helical polymer chains in the LC state (Fig. 11.12E) as observed in the LC polyisocyanates by Green et al. [79]. On the basis of the X-ray analyses of the LC samples, the most plausible helical structure of **39**-HCl was proposed to be a 23 unit/10 turn (23/10) helix [78].

11.3.2.3 Other Induced Helical Polymers

Taking advantage of the helicity induction concept, a preferred helical conformation has also been induced in other dynamic racemic, chromophoric polymers or oligomers as a result of noncovalent binding of the nonracemic guests (Fig. 11.13). Aliphatic polyacetylenes (**48–50**) [80–82], polyisocyanates (**19**, **51–53**) [44, 83–85], poly- and oligosilanes (**54–56**) [86–89], poly(phenyl isocyanide) (**57**) [90],
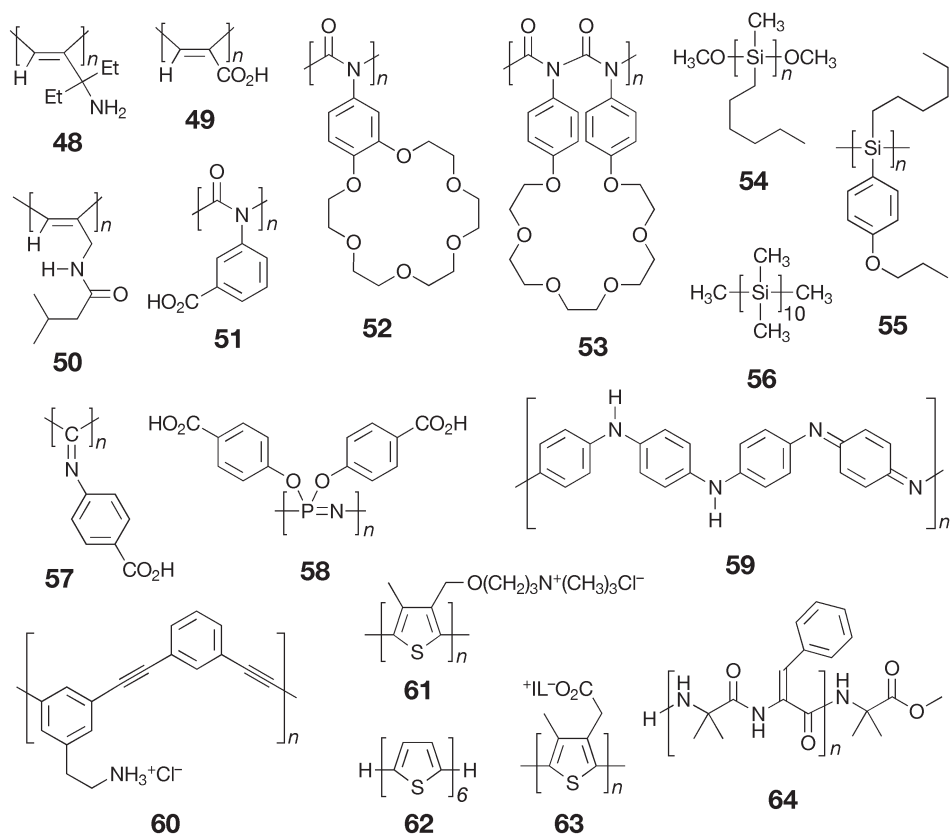
Scheme 11.5
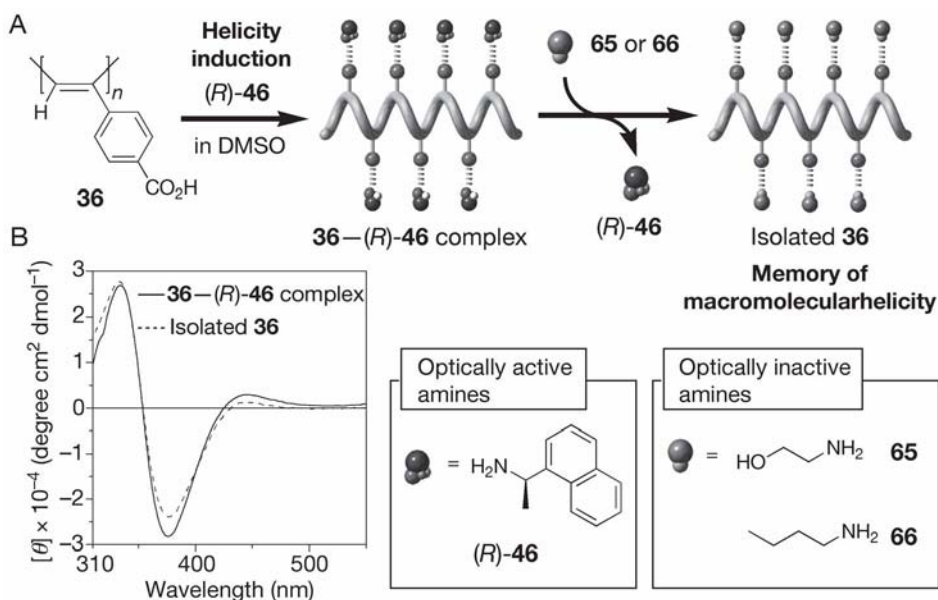
Fig. 11.13 (A) Schematic illustration of helicity induction in **36** upon complexation with (R)-**46** and subsequent memory of the helicity after replacement by achiral amines (**65**, **66**). (B) CD spectra of the **36**–(R)-**46** complex (solid line) and the isolated **36** by SEC fractionation using a DMSO solution containing an achiral amine **36** as the mobile phase (dashed line) in DMSO.

poly(organophosphazene) (**58**) [91], polyguanidine (**17**) [33], polyaniline (**59**) [92] and poly(*m*-phenylene ethynylene)s (**60**) [93] are such examples, in which chiral acid–base, host–guest, ionic or hydrogen bonding interactions play an important role in the helicity induction. A helical conformation with an excess helix-sense was also induced in dynamic helical polysilanes (**54**, **55**) with no functional pendant groups in nonracemic solvents [86, 87].

An oligosilane (**56**) forms an induced helical conformation once entrapped in a hydrophobic chiral cavity created by helical polysaccharides, such as right-handed triple-stranded helical schizophyllan and left-handed helical amylose in water, thus showing ICDs with opposite Cotton effect signs to each other [88, 89]. A similar helicity induction also takes places in a water-soluble polythiophene (**61**) or an oligothiophene (**62**) by wrapping with the schizophyllan interior in water [94, 95].

Inganäs et al. reported that a negatively charged, optically inactive luminescent polythiophene **63** self-assembled into a helix bundle with a positively charged, artificial peptide with a random coil conformation in an aqueous solution. Interestingly, a one-handed helicity and an α-helix were simultaneously induced in both

polymers upon complexation through electrostatic interactions by mixing the two polymers in water [96].

An optically inactive oligopeptide **64** with the N-terminal amino group produced an ICD derived from the one-handed helical conformation of the entire peptide chain upon complexation with chiral carboxylic acids to the N-terminal amino group [97]. This phenomenon was called the ''noncovalent domino effect''.

11.3.3
Memory of Induced Helical Chirality

The macromolecular helicity in poly(phenylacetylene)s **36–38** (Fig. 11.9) induced by chiral amines is not static, but dynamic in nature, so that the ICD due to the helical chirality immediately disappears when exposed to a stronger acid. However, we observed an unusual, but interesting macromolecular helical memory in this dynamic helical polymer system; the helical conformations of **36–38** induced by an optically active amine such as (*R*)-**46** were found to be retained, namely ''memorized'', after the chiral amine was completely removed and replaced by achiral amines, e.g., **65** and **66** for **36** and diamines, such as ethylenediamine, for **37** and **38** in DMSO (Fig. 11.13) [98–100]. The macromolecular helicity memory was not transient, but lasted for an extremely long time (over two years), suggesting that the thermodynamically controlled, dynamic helical conformations transform into kinetically controlled, static ones after the helicity memory is assisted by achiral amines.

The noncovalent helicity induction combined with the helicity memory is a versatile method to produce either a right- or left-handed helix with an excess of the preferred helix-sense. However, the helix-sense is predetermined by the chirality of the enantiomeric amines used. Accordingly, the opposite enantiomeric helicity memory requires the opposite enantiomeric amine, followed by replacement with achiral amines. However, both enantiomeric helices with the mirror image to each other can be produced with a high efficiency from a helical poly(phenylacetylene) (**67**) induced by a single enantiomer (Fig. 11.14) [101]. This ''dual memory'' of enantiomeric helices is based on the inversion of the macromolecular helicity with temperature (see Section 11.4). The poly(phenylacetylene) folds into a one-handed helix induced by (*R*)-**46** at 25 °C in DMSO. The helix-sense further inverts at 65 °C, as evidenced by the Cotton effect inversion. These diastereomeric right- and left-handed helices of **67** obtained at 25 and 65 °C can be further memorized by an achiral diamine such as **68** at these temperatures, resulting in the perfect mirror image Cotton effects and identical absorption spectra [101]. The chiral amplification concept can be applied to this system; a 35% ee of **46** induced as an intense ICD as that with 100% ee of **46** at 25 °C and 65 °C after helicity inversion. Subsequent replacement of the nonracemic **46** yielded the enantiomeric helices of **67** with an excess single-handedness.

The pendant phosphonate complexed with **68** appears to be achiral, but can be converted into its methyl ester with diazomethane, resulting in the generation of
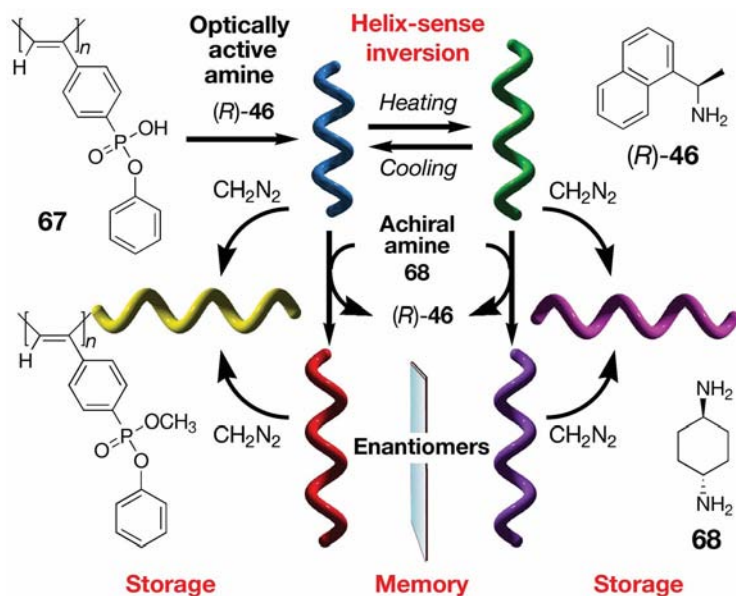
Fig. 11.14 Schematic illustration of an induced one-handed helicity in optical inactive **67**, helix inversion with temperature, subsequent memory of the diastereomeric macromolecular helicity at different temperatures with achiral **68**, and storage of the induced helicity and helicity memory by enantioselective esterification with diazomethane.

a phosphorus stereogenic center with optical activity (Fig. 11.14) [102]. The esterification proceeded enantioselectively when **67** had a helical conformation induced by **46** or macromolecular helicity memory assisted by **68**. Although the enantioselectivity was low, the pendant chirality is significantly amplified in the polymer backbone at low temperatures, resulting in a higher optical activity as an excess single-handed helix than that expected from the ee of the pendant groups; the helix-sense excess of the polymer reached 62% ee at −95 °C.

A macromolecular helicity memory has been achieved in organic solvents, but was unsuccessful in water, because dynamic helical polymers retain their helicity memory only when complexed with achiral molecules, such as achiral amines; therefore, the memory in water is lost. The recently developed layer-by-layer (LbL) assembly technique has made possible the macromolecular helicity memory in water (Fig. 11.15) [103]. A negatively charged helical poly(phenylacetylene) **37b** induced by a chiral amine ((S)-**69**) in water showing a full ICD was deposited on a quartz substrate. Subsequently, an achiral positively charged vinylpolymer such as the hydrochloride of poly(allylamine) (PAH) was LbL assembled. The (S)-**69** molecules used for the helicity induction in **37b** were automatically removed during the LbL assembly process, resulting in optically active multilayer thin films with a macromolecular helicity memory after repeating the alternative deposition
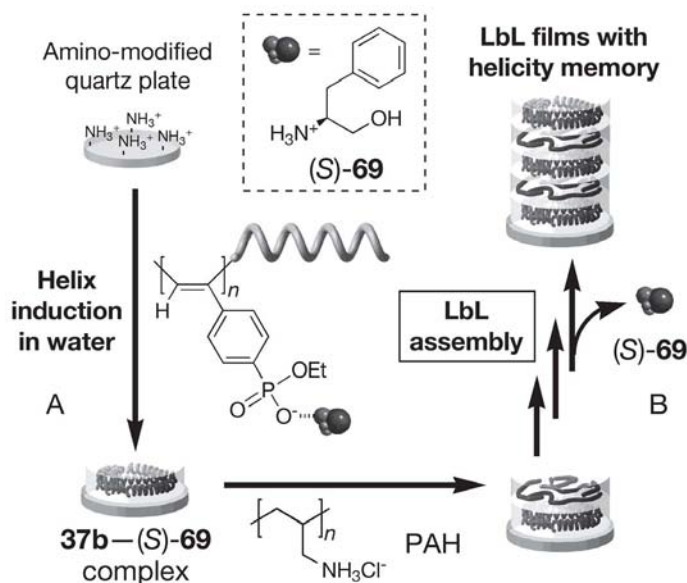
Fig. 11.15 Schematic illustration of the LbL self-assembly of a charged poly(phenylacetylene) with induced macromolecular helicity. (A) An excess of the one-handed helical sense is induced in **37b** with the optically active (*S*)-**69** in water. (B) An induced helical **37b** can be LbL assembled with an achiral polyelectrolyte having opposite charges (PAH), resulting in multilayer thin films with an induced macromolecular helicity memory on a substrate.

cycle. When a positively charged, induced helical **39**-HCl was used instead, the alternative deposition of an achiral vinylpolymer with opposite charges produced a similar thin film with a macromolecular helicity memory [103].

Although the chiral memory effect has also been observed in other dynamic supramolecular systems [104–108], the use of achiral guests is essential for the maintenance of the memory effect. In the absence of the achiral guest, the memory will be instantly lost. However, the sodium salt of helical **57** (**57**-Na) with an excess helical sense induced by (*S*)-**69** was found to be automatically memorized after complete removal of the optically active amine in water (Fig. 11.16A) [109]. In sharp contrast to the conformational memory of the induced helical poly(phenylacetylene)s, the helix formation of **57**-Na may be accompanied by configurational isomerization around the C=N double bonds (*syn–anti* isomerization) (Fig. 11.16B) into one single configuration upon complexation with the chiral amine, which may force the polymer backbone to take an excess helical sense. This is an unprecedented example of the synthesis of a static helical polymer after polymerization through specific noncovalent chiral interactions. The significant advantage of this helicity memory over that of helical poly(phenylacetylene)s is that there is no longer need to use the achiral chaperoning amines to retain the helic-
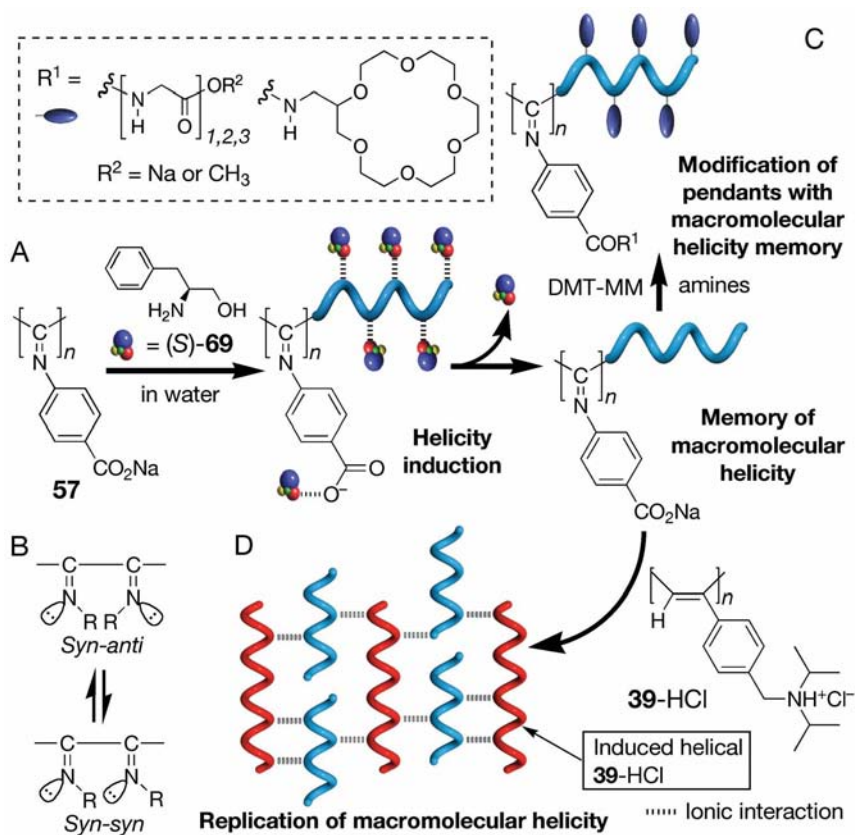
Fig. 11.16 Schematic illustrations of a helicity induction in **57-Na** upon complexation with (*S*)-**69** and memory of the induced macromolecular helicity after complete removal of (*S*)-**69** (A), probably through *syn–anti* isomerization of the C=N bond (B), modification of the pendants with macromolecular helicity memory (C), and the replication of the macromolecular helicity (D).

ity in the polymer, and therefore, further modifications of the side groups such as oligoglycines and crown ethers can be possible along with maintaining the macromolecular helicity memory (Fig. 11.16C).

The negatively charged **57**-Na with a macromolecular helicity memory can serve as the template for further helicity induction in a different, dynamically racemic helical polymer with opposite charges in water ("helicity-replication"), resulting in biomimetic interpolymer helical assemblies with a controlled helicity in water (Fig. 11.16D) [110]. Although the helical **57**-Na no longer has any chiral components and stereogenic centers, the helical chirality of the polymer is efficiently transformed into a dynamically racemic, cationic polyelectrolyte **39**-HCl through electrostatic interaction, resulting in the appearance of an ICD in the **39**-HCl chromophore region due to an excess one-handed helicity.

11.4
Inversion of Macromolecular Helicity

Another interesting and unique feature of dynamic helical polymers is the reversible helix inversion (helix–helix (PM) transition) between right- and left-handed helical conformations regulated by external stimuli, such as a change in temperature, solvent or by irradiation with light. Because the extremely high sensitivity of dynamic helices to a chiral environment arising from a high cooperativity, the formation of an excess of the preferred helical sense can be altered, resulting in the inversion of helicity. Biological polymers such as DNA [111] and polypeptides [112] with specific sequences are known to undergo inversion of the helicity driven by a change in the salt concentration and temperature, respectively. Some static helical polymers (poly-**5**) [21] and chloral oligomers (**11**, **12**) [22, 23] also exhibit a transition in their helicities, but their processes are not reversible, and racemization. Several synthetic, dynamic helical polymers exhibit a reversible PM transition by changing the external conditions, such as temperature (**31**, **70**–**75**), solvent (**29**, **30**, **33a,b**, **34b**, **76**) or by light irradiation (**77**).

Zentel and Mager found that the helix-sense of an optically active polyisocyanate copolymer containing a photosensitive azobenzene side group (**77**) can be switched upon the photochemical *trans–cis* isomerization [113]. The CD spectral pattern of the copolymer was completely inverted upon photoirradiation.
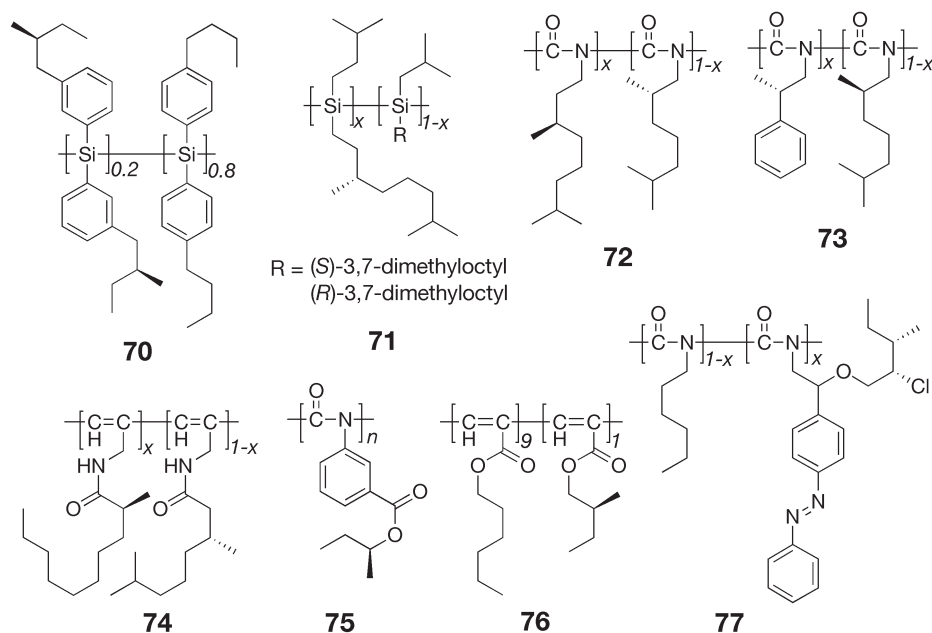
Scheme 11.6
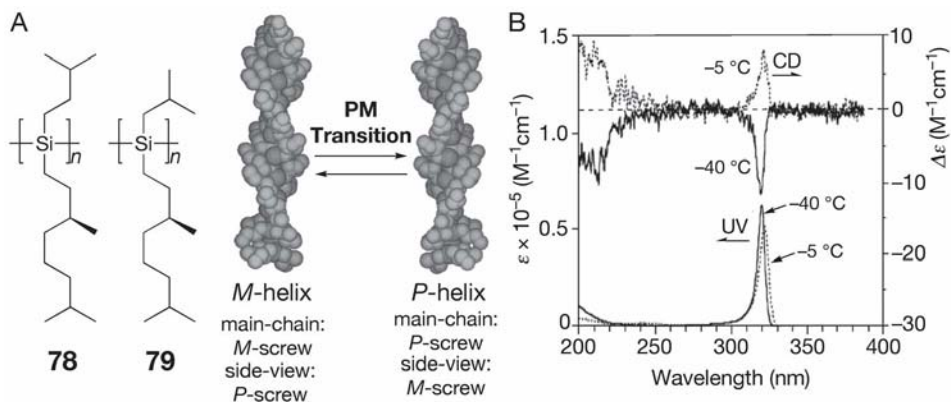
Fig. 11.17 (A) An illustration of the *PM* transition of a rigid polysilylene with *P*- and *M*-7₃ helical motif. (B) CD and UV absorption spectra of **78** at −40 °C (solid line) and −5 °C (dotted line) in isooctane. The right- and left-handed helices of **78** are not enantiomers, but diastereomers because of the presence of enantiopure pendants. Therefore, their CD spectra differ from one another after the helicity inversion. (Reproduced with permission from Ref. 114. Copyright 2000 American Chemical Society.)

Fujiki and coworkers synthesized a series of homopolymers and copolymers of optically active helical polysilanes to develop chiral switchable materials based on the inversion of helicity. They found that poly((*S*)-3,7-dimethyloctyl-3-methylbutylsilane) (**78**), a family of rod-like helical polysilanes, undergoes a thermo-driven PM transition through a transition temperature (T_c) at −20 °C in isooctane; above and below the T_c, the polymer showed opposite Cotton effect signs to each other (Fig. 11.17B) [47, 114]. The inversion of helicity is sensitive to the structure of pendants, and an analogous polysilane (**79**) bearing a slightly different *β*-branched achiral side chain showed no inversion of the CD from −90 to +80 °C. Although the origin of the PM transition remains obscure, Fujiki et al. reported a double-well ("W") shape potential energy curve for **78**, which may be responsible for the thermo-driven PM transition. In contrast, **79** showing no PM transition exhibits an unclear double well potential curve. Fujiki et al. further demonstrated that it is possible to control the PM transition temperature by co-polymerization with appropriate achiral monomers (**70**, **71**) or by changing the molecular shape of the hydrocarbon solvents [47].

Green et al. also reported designer polyisocyanates (**72**, **73**) showing an inversion of the helicity with a desired T_c in dilute solution by the copolymerization of paired structurally different enantiomers, which are in competition with each other in helical sense preferences [115]. They further applied this concept to the lyotropic LC state formed by poly(*n*-hexyl isocyanate) (**19**) using the thermo-driven switchable polyisocyanates as chiral dopants. The addition of **73** showing a T_c near 30 °C to a nematic solution of **19** gave rise to a typical finger texture above and below the T_c due to a cholesteric LC phase. Changing the temperature

Fig. 11.18 Schematic illustration of the macromolecular helicity inversion in dilute solution and 2-D crystal state. (A) CD and absorption spectra of D-**30** measured in THF, chloroform, and benzene. The helix-sense of D-**30** in benzene inverts to the opposite helix-sense in THF and chloroform. (B) Rodlike helical **30** self-assembles into 2-D helix bundles with the controlled helicity upon exposure of each organic solvent vapor. The one-handed 2-D helix bundles of D-**30** further invert to the opposite handedness by exposure to benzene vapor on the substrates. AFM images of 2-D self-assembled D-**30** on HOPG and helical structures of D-**30** proposed by AFM and X-ray analyses are shown. (Reproduced with permission from Ref. 117. Copyright 2006 American Chemical Society.)

thus allows one to control the mesoscopic cholesteric states of opposite twist sense.

Helical polyacetylenes bearing amino acids as the pendants also showed inversion of the helicity (**29**, **30**, **34b**) by changing the temperature or solvent, mainly resulting from the "on and off" fashion of the intramolecular hydrogen bonding between the pendant amide groups in nonpolar and polar solvents, respectively [49, 54, 63, 116]. The direct evidence for the macromolecular helicity inversion of a helical **30** in different solvents has been reported based on AFM observations of the diastereomeric helical structures (Fig. 11.18) induced in polar and nonpolar solvents, followed by deposition on graphite. The diastereomeric helical **30** further self-assembled into ordered, 2D helix-bundles with controlled molecular packing, helical pitch, and handedness on graphite upon exposure to each solvent. The macromolecular helicity of the helical macromolecules deposited on graphite from a polar solvent was further inverted into the opposite handedness by exposure to a specific nonpolar solvent, and these changes in the surface chirality based on the inversion of helicity could be visualized by AFM with molecular resolution [117], and the results were quantified by X-ray diffraction of the oriented liquid crystalline, diastereomeric helical polymer films.

Switching of the macromolecular helicity by chiral stimuli is a current challenging issue, but still remains rare, although such switching materials can be used for sensing the chirality of specific chiral guests. A poly(phenylacetylene) bearing an optically active (1R,2S)-norephedrine residue (**80**) was the first example of helix inversion induced by chiral stimuli [118]. The Cotton effect signs of **80** were inverted in the presence of excess (R)-mandelic acid ((R)-**81**), while the ICD of **80** hardly changed with an excess (S)-**81**, suggesting that **80** undergoes a transition from one helix to another in the presence of (R)-**81**.

80

81

m = 1: **82α**
m = 2: **82β**
m = 3: **82γ**
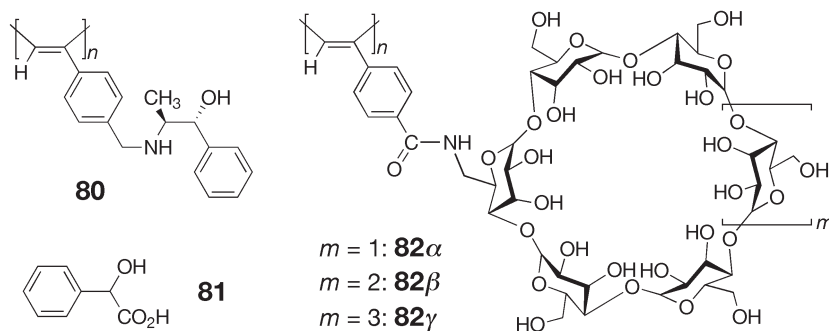
Scheme 11.7

Introducing optically active cyclic host molecules, such as α-, β-, and γ-cyclodextrin (CyD) residues to a dynamic helical polyacetylene backbone as the pendant groups (**82**) provides a conceptually new direct colorimetric detection-system for neutral chemical species including enantiomers as well as solvent and temperature based on the macromolecular helicity inversion (Fig. 11.19A). The helicity inversion was accompanied by a color change due to a change in the twist angle of the conjugated double bonds (tunable helical pitch) that was readily visible by the naked eye and could be quantified by absorption and CD spectroscopies. In particular, **82β** bearing β-CyD residues is sensitive to achiral and chiral stimuli and exhibits an inversion of helicity induced by inclusion complexation with guest molecules into the chiral β-CyD cavity [119]. When 1-adamantanol (**83**) or (−)-borneol (**84**) was added to the **82β** solution, the solution color immediately changed from yellow-orange to red accompanied by the inversion of the Cotton effect signs and a large red-shift of λ_{max} (Fig. 11.19B), whereas, cyclooctanol (**85**) and cyclohexanol (**86**) neither produced such a dramatic color change in the solution nor the Cotton effect inversion. **82γ** also showed a similar CD inversion accompanied by a color change in response to the specific guest molecules capable of interacting with γ-CyD. In addition, the racemic **45** and (R)-rich **45** of 50% ee could not induce a conformational change in **82β**, resulting in almost no change in their absorption and CD spectra, while **82β** is sensitive to the chirality of (S)-**45**, and (S)-rich **45** of 50% ee showed a sig-
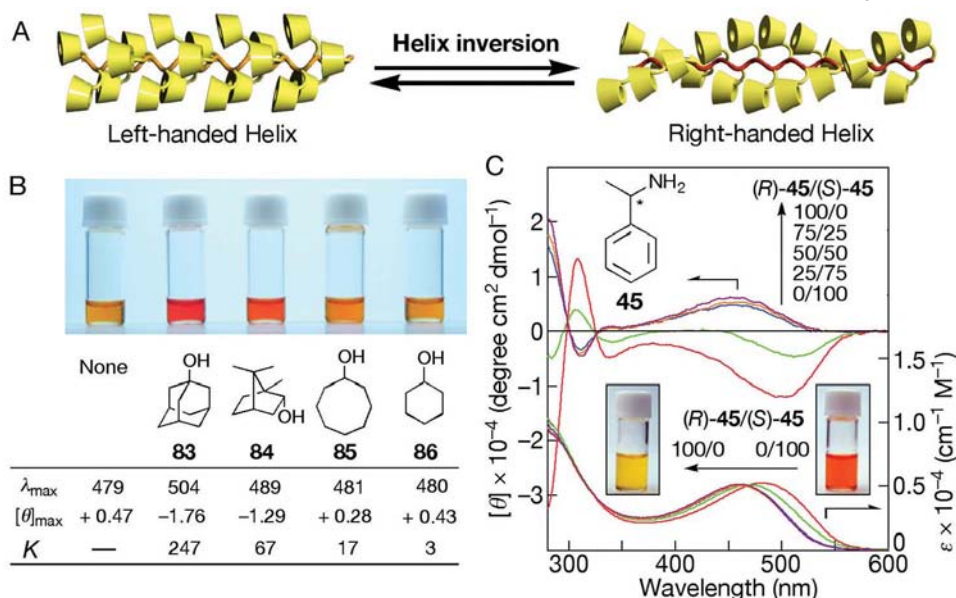
Fig. 11.19 (A) Schematic illustrations of a possible conformational change of **82β**. (B) Visible color changes in **82β** in DMSO-water (8/2, v/v) by the addition of **83–86**. (C) CD and absorption spectral changes of **82β** in alkaline water (pH 11.7)–DMSO (7/3, v/v) in the presence of 0–100% ee of **45** at 25 °C. (Reprinted with permission from Ref. 120. Copyright 2006 American Chemical Society.)

nificant change in the CD and absorption spectra as well as 100% ee of (S)-**45** (Fig. 11.19C) [120]. The dynamic helical conformations of **82β** showing opposite Cotton effect signs in DMSO and alkaline water could be further fixed by intramolecular crosslinking between the hydroxy groups of the neighboring β-CyD units in each solvent. The crosslink between the pendant CyD units suppressed the inversion of the helicity; therefore, the crosslinked **82β**s showed no Cotton effect inversion [120].

11.5
Applications of Helical Polymers

Potential applications of optically active helical polymers that mimic the structures of enzymes, involve enantioselective catalysis and adsorbents [1, 2, 5]. The one-handed helical polymethacrylates prepared by the helix-sense selective polymerization of TrMA and its analog **5** (Section 11.2.1) exhibit excellent chiral recognition abilities when coated on a macroporous silica gel and used as chiral stationary phases (CSPs) in HPLC [121, 122]. These packing materials can resolve a wide range of racemic compounds including chiral drugs and stereochemically
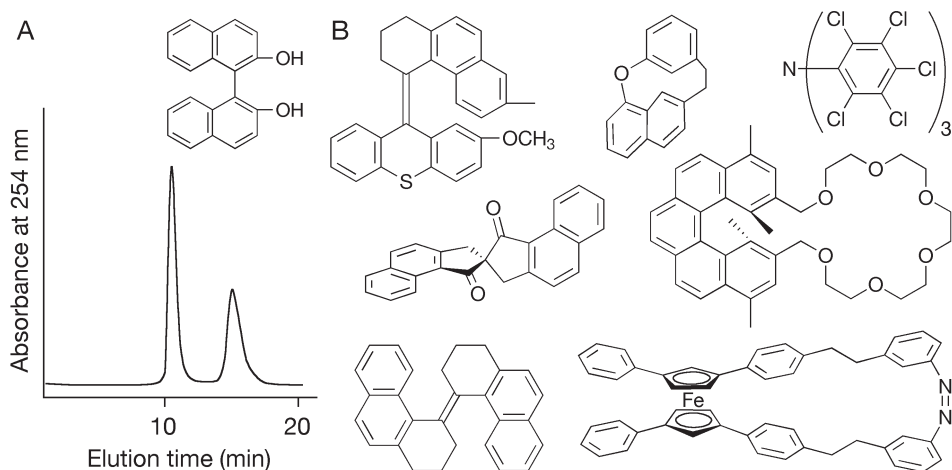
Fig. 11.20 Chromatogram for the resolution of 2,2′-dihydroxy-1,1′-binaphthyl on a PTrMA (**1**) column and the structures of stereochemically interesting compounds resolved on **1**.

interesting molecules, and are commercialized [1]. The typical chromatogram for the separation of 2,2′-dihydroxy-1,1′-binaphthyl and some chiral molecules resolved on PTrMA are shown in Fig. 11.20 [5, 121, 122]. A stereoregular helical poly(phenylacetylene) (**31**) can also be used as a CSP for HPLC, which resolved several enantiomers including Tröger's base and stilbene oxide [57]. However, a stereoirregular poly(phenylacetylene) with an identical chemical structure as **28**, prepared by a different synthetic route showed poor chiral recognition, clearly indicating that the one-handed helical conformation induced by a stereoregular polymer backbone with chiral pendant groups is indispensable for effective chiral recognition. Other helical polyacetylenes such as **27** have been used as enantioselective permeable membranes for separating amino acids and chiral alcohols [123].

Reggelin et al. took advantage of the versatility of helical polymethacrylates developed by Okamoto and reported the first successful catalytic asymmetric C–C bond forming reaction using the helical polymers as a chiral polymeric ligand. The polymethacrylates were prepared by the helix-sense selective anionic polymerization or copolymerization with TrMA [1], producing an isotactic, fully one-handed helical polymer and copolymer with a large optical rotation. Complexed with palladium, the resulting monodentate (**87**) [124] and bidentate (**88**) [125] palladium catalysts promoted the asymmetric allylic alkylation reaction (Scheme 11.8) resulting in the substitution product with ca. 30 and 40–60% ee, respectively. Reggelin et al. further applied this strategy to a dynamic helical polyisocyanate. The copolymer composed of a chiral isocyanate and an achiral isocyanate bearing a phosphine pendant (60:40, mol/mol) (**89**) [125], although its helical sense excess was not perfect, showed a low, but apparent catalytic enantioselective
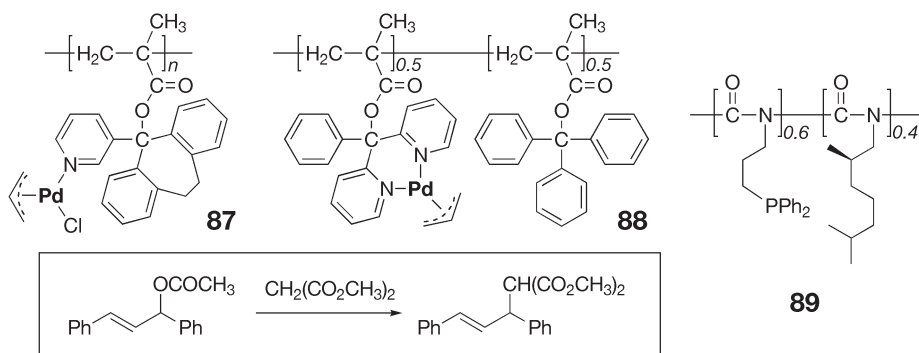
Scheme 11.8

activity in an asymmetric hydrogenation reaction when complexed with a rhodium catalyst, thus producing a hydrogenated product with 14.5% ee. These static and dynamic helical polymers lacking any other elements of chirality except for helicity are a promising new class of ligands for asymmetric catalysis.

A large number of other chiral polymeric ligands have been synthesized from chiral small molecules such as 1,1'-bi-2-naphthol (BINOL) and 2,2'-bis(diphenylphosphino)-1,1'-binaphthyl (BINAP). Some of them may have a helical structure and serve as ligands in various enantioselective transformations [126].

As previously described, the most important and unique feature of dynamic helical polymers is a remarkable amplification of chirality, which may be utilized to construct a novel helical polymer with the desired pendant group in a one-handed helical array along the polymer backbone. In fact, the copolymerization of an achiral phenylacetylene bearing a fullerene pendant with a small amount of an optically active phenylacetylene yielded a helical copolymer with an excess of one helical-sense in which the pendant C_{60} groups adopt a predominant screw-sense along the polymer backbone (Scheme 11.9) [127], because the copolymer exhibited an ICD in the achiral fullerene chromophore region as well as in the polymer backbone region. In a complementary approach, an enantiomerically pure cationic C_{60}-bisadduct (**90**) induced a predominantly one-handed helix in a dynamically racemic poly(phenylacetylene) (**37b**) with the opposite negative charges in DMSO–water mixtures through noncovalent bonding interactions, which further results in a helical array of the C_{60}-bisadducts with a predominant screw-sense along the polymer chain [128].

Optically active helical polymers often show chiral LC phases due to their rigid rodlike backbones. Such liquid crystalline helical polymers combined with a specific property of inversion of the helicity regulated by external stimuli will offer switchable chiral materials suitable for data storage, optical devices and use in other fields involving chiral nanotechnology [129].

A

Helical array of C$_{60}$

B

Chiral C$_{60}$ bisadduct

90

DMSO/water

Helicity induction

37b
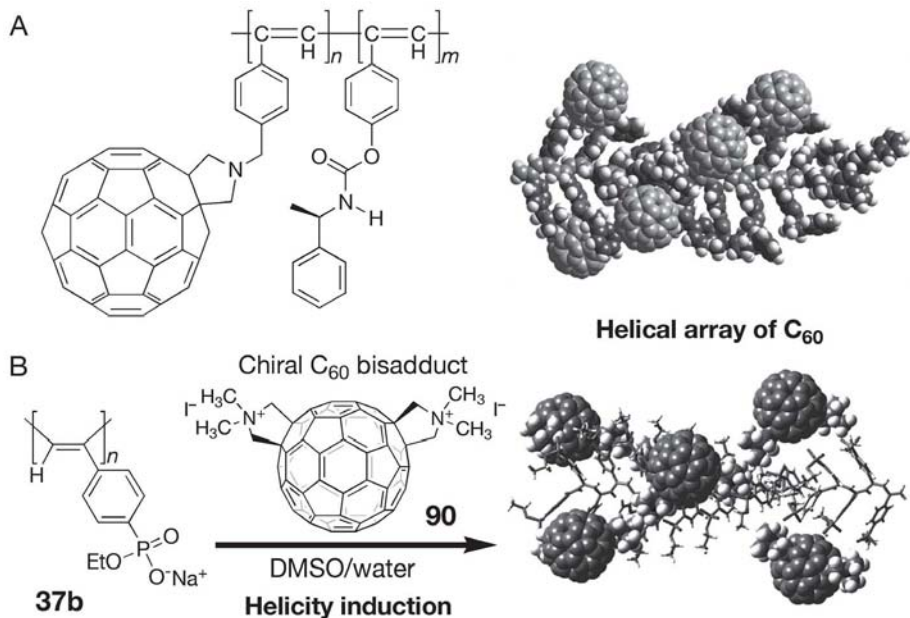
Scheme 11.9

11.6
Conclusion

In this chapter, the synthesis, structures and functions of helical polymers are described. A large number of helical polymers with an excess of a preferred helix-sense have been synthesized by the helix-sense selective polymerization of achiral monomers, thus producing static helical polymers, or by the predominantly one-handed helicity induction in dynamic helical polymers via covalent or noncovalent bonding of chiral pendants. These helical polymers differ through their helix inversion barriers; as a result, the former helical conformations are locked during the polymerization under kinetic control, while the latter helical conformations are under thermodynamic control. However, as described in this chapter, helical conformations of dynamic helical polymers can also be locked, as evidenced by the memory effect of an induced helical poly(phenylacetylene). Either static or dynamic helical polymers with an excess one-handedness have also been prepared by the polymerization of analogous monomers bearing different substituents. The history of synthetic helical polymers extends back to the 1960s; at that time the structural elucidation at a molecular level was a laborious task. However, recent significant developments in spectroscopic and microscopic instruments, coupled with precise polymerization techniques, have made it possible to observe directly the helical structures of certain helical polymers including helical pitch

and handedness. This more detailed information leads to a better understanding of the principles underlying the generation of helical conformations. In addition, biological helical polymers further hierarchically assemble into complicated supramolecular structures, such as the coiled coil (helix bundle) superstructure, which are responsible for their elaborate functions. The next important and attractive challenge, which has implications for biological helices, superstructures and functions, will be not only to mimic biological helices, but also to develop supramolecular helical assemblies with a controlled helix-sense, and this may also provide a clue for the construction of advanced chiral materials [130].

Acknowledgments

References

1 Y. Okamoto, T. Nakano, *Chem. Rev.* **1994**, *94*, 349.

2 Y. Okamoto, E. Yashima, *Angew. Chem. Int. Ed.* **1998**, *37*, 1020.

3 S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173.

4 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893.

5 C. Yamamoto, Y. Okamoto, *Bull. Chem. Soc. Jpn.* **2004**, *77*, 227.

6 I. Huc, *Eur. J. Org. Chem.* **2004**, 17.

7 P. Pino, G. P. Lorenzi, *J. Am. Chem. Soc.* **1960**, *82*, 4745.

8 M. M. Green, N. C. Peterson, T. Sato, A. Teramoto, R. Cook, S. Lifson, *Science* **1995**, *268*, 1860.

9 T. Nakano, Y. Okamoto, *Chem. Rev.* **2001**, *101*, 4013.

10 J. J. L. M. Cornelissen, A. E. Rowan, R. J. M. Nolte, N. J. A. M. Sommerdijk, *Chem. Rev.* **2001**, *101*, 4039.

11 M. Fujiki, *Macromol. Rapid Commun.* **2001**, *22*, 539.

12 E. Yashima, K. Maeda, T. Nishimura, *Chem. Eur. J.* **2004**, *10*, 42.

13 F. J. M. Hoeben, P. Jonkheijm, E. W. Meijer, A. P. H. J. Schenning, *Chem. Rev.* **2005**, *105*, 1491.

14 K. Maeda, E. Yashima, *Top. Curr. Chem.* **2006**, *265*, 47.

15 O. Vogl, G. D. Jaycox, C. Kratky, W. S. Simonsick, Jr., K. Hatada, *Acc. Chem. Res.* **1992**, *25*, 408.

16 M. Suginome, Y. Ito, *Adv. Polym. Sci.* **2004**, *17*, 77.

17 Y. Okamoto, K. Suzuki, K. Ohta, K. Hatada, H. Yuki, *J. Am. Chem. Soc.* **1979**, *101*, 4763.

18 T. Nakano, Y. Shikisai, Y. Okamoto, *Polym. J.* **1996**, *28*, 51.

19 T. Nakano, K. Tsunematsu, Y. Okamoto, *Chem. Lett.* **2002**, 42.

20 N. Hoshikawa, Y. Hotta, Y. Okamoto, *J. Am. Chem. Soc.* **2003**, *125*,

21 Y. Okamoto, H. Mohri, T. Nakano, K. Hatada, *J. Am. Chem. Soc.* **1989**, *111*, 5952.

22 K. Ute, K. Hirose, H. Kashimoto, H. Nakayama, K. Hatada, O. Vogl, *Polym. J.* **1993**, *25*, 1175.

23 K. Ute, K. Hirose, H. Kashimoto, O. Vogl, *J. Am. Chem. Soc.* **1991**, *113*, 6305.

24 F. Millich, *Adv. Polym. Sci.* **1975**, *19*, 117.

25 R. J. M. Nolte, A. J. M. von Beijnen, W. Drenth, *J. Am. Chem. Soc.* **1974**, *96*, 5932.

26 R. J. M. Nolte, _Chem. Soc. Rev._ **1994**, _23_, 11.

27 P. C. J. Kamer, R. J. M. Nolte, W. Drenth, _J. Am. Chem. Soc._ **1988**, _110_, 6818.

28 T. J. Deming, B. M. Novak, _J. Am. Chem. Soc._ **1992**, _114_, 7926.

29 D. B. Amabilino, J.-l. Serrano, T. Sierra, J. Veciana, _J. Polym. Sci. Part A: Polym. Chem._ **2006**, _44_, 3161.

30 T. Kajitani, K. Okoshi, S.-i. Sakurai, J. Kumaki, E. Yashima, _J. Am. Chem. Soc._ **2006**, _128_, 708.

31 F. Takei, H. Hayashi, K. Onitsuka, N. Kobayashi, S. Takahashi, _Angew. Chem. Int. Ed._ **2001**, _40_, 4092.

32 M. Suginome, S. Collet, Y. Ito, _Org. Lett._ **2002**, _4_, 351.

33 D. S. Schlitzer, B. M. Novak, _J. Am. Chem. Soc._ **1998**, _120_, 2196.

34 H.-Z. Tang, B. M. Novak, J. He, P. L. Polavarapu, _Angew. Chem. Int. Ed._ **2005**, _44_, 7298.

35 M. M. Green, J.-W. Park, T. Sato, A. Teramoto, S. Lifson, R. L. B. Selinger, J. V. Selinger, _Angew. Chem. Int. Ed._ **1999**, _38_, 3138.

36 M. A. Mateos-Timoneda, M. Crego-Calama, D. N. Reinhoudt, _Chem. Soc. Rev._ **2004**, _33_, 363.

37 M. M. Green, M. P. Reidy, R. J. Johnson, G. Darling, D. J. O'Leary, G. Willson, _J. Am. Chem. Soc._ **1989**, _111_, 6452.

38 S. K. Jha, K.-S. Cheon, M. M. Green, J. V. Selinger, _J. Am. Chem. Soc._ **1999**, _121_, 1665.

39 K. Maeda, Y. Okamoto, _Polym. J._ **1998**, _30_, 100.

40 A. Teramoto, _Prog. Polym. Sci._ **2001**, _26_, 667.

41 S. Lifson, C. Andreola, N. C. Peterson, M. M. Green, _J. Am. Chem. Soc._ **1989**, _111_, 8850.

42 M. M. Green, C. Andrela, B. Munoz, M. P. Reidy, _J. Am. Chem. Soc._ **1988**, _110_, 4063.

43 K. Ute, Y. Fukunishi, S. K. Jha, K.-S. Cheon, B. Munoz, K. Hatada, M. M. Green, _Macromolecules_, **1999**, _32_, 1304.

44 M. M. Green, C. Khatri, N. C. Peterson, _J. Am. Chem. Soc._ **1993**, _115_, 4941.

45 M. M. Green, B. A. Garetz, B. Munoz, H. Chang, S. Hoke, R. G. Cooks, _J. Am. Chem. Soc._ **1995**, _117_, 4181.

46 M. Fujiki, J. R. Koe, K. Terao, T. Sato, A. Teramoto, J. Watanabe, _Polym. J._ **2003**, _35_, 297.

47 M. Fujiki, _J. Organomet. Chem._ **2003**, _685_, 15.

48 E. Yashima, Y. Okamoto, in _Circular Dichroism: Principles and Applications, 2nd ed._ (Eds: N. Berova, K. Nakanishi, R. W. Woody), John Wiley & Sons, Inc. New York, Chapter 18, p. 521–546.

49 J. W. Y. Lam, B. Z. Tang, _Acc. Chem. Res._ **2005**, _38_, 745.

50 A. P. H. J. Schenning, M. Fransen, E. W. Meijer, _Macromol Rapid Commun._ **2002**, _23_, 265.

51 V. Percec, J. G. Rudick, M. Peterca, M. Wagner, M. Obata, C. M. Mitchell, W.-D. Cho, V. S. K. Balagurusamy, P. A. Heiney, _J. Am. Chem. Soc._ **2005**, _127_, 15257.

52 R. Nomura, H. Nakako, T. Masuda, _J. Mol. Cat. A Chem._ **2002**, _190_, 197.

53 J. Tabei, F. Sanda, T. Masuda, _Kobunshi Ronbunshu_, **2006**, _63_, 286.

54 H. Zhao, F. Sanda, T. Masuda, _J. Polym. Sci., Part A; Polym. Chem._ **2005**, _43_, 5168.

55 T. Masuda, F. Sanda, in _Handbook of Metathesis, Vol. 3_ (Eds: R. H. Grubbs), Wiley-VCH, Weinheim, **2003**, Chapter 3.11, p. 375–406.

56 C. Simionescu, S. Dumitrescu, V. Percec, _J. Polym. Sci., Polym. Symp._ **1978**, _64_, 209.

57 E. Yashima, S. Huang, Y. Okamoto, _J. Chem. Soc., Chem. Commun._ **1994**, 1811.

58 Y. Ashida, T. Sato, K. Morino, K. Maeda, Y. Okamoto, E. Yashima, _Macromolecules_ **2003**, _36_, 3345.

59 R. Nomura, J. Tabei, S. Nishiura, T. Masuda, _Macromolecules_, **2003**, _36_, 561.

60 K. Morino, K. Maeda, Y. Okamoto, E. Yashima, T. Sato, _Chem. Eur. J._ **2002**, _8_, 5112.

61 J. Tabei, M. Shiotsuki, T. Sato, F. Sanda, T. Masuda, _Chem. Eur. J._ **2005**, _11_, 3591.

62 R. Nomura, Y. Fukushima, H. Nakako, T. Masuda, *J. Am. Chem. Soc.* **2000**, *122*, 8830.

63 K. Okoshi, K. Sakajiri, J. Kumaki, E. Yashima, *Macromolecules*, **2005**, *38*, 4061.

64 T. Aoki, T. Kaneko, N. Maruyama, A. Sumi, M. Takahashi, T. Sato, M. Teraguchi, *J. Am. Chem. Soc.* **2003**, *125*, 6346.

65 S.-i. Sakurai, K. Okoshi, J. Kumaki, E. Yashima, *Angew. Chem. Int. Ed.* **2006**, *45*, 1245.

66 E. Yashima, T. Matsushima, Y. Okamoto, *J. Am. Chem. Soc.* **1995**, *117*, 11596.

67 E. Yashima, T. Matsushima, Y. Okamoto, *J. Am. Chem. Soc.* **1997**, *119*, 6345.

68 H. Onouchi, K. Maeda, E. Yashima, *J. Am. Chem. Soc.* **2001**, *123*, 7441.

69 T. Hasegawa, K. Maeda, H. Ishiguro, E. Yashima, *Polym. J.* **2006**, *38*, 912.

70 E. Yashima, Y. Maeda, Y. Okamoto, *Chem. Lett.* **1996**, 955.

71 E. Yashima, T. Nimura, T. Matsushima, Y. Okamoto, *J. Am. Chem. Soc.* **1996**, *118*, 9800.

72 R. Nonokawa, E. Yashima, *J. Am. Chem. Soc.* **2003**, *125*, 1278.

73 G. S. Manning, *Acc. Chem. Res.* **1979**, *12*, 443.

74 M. A. Saito, K. Maeda, H. Onouchi, E. Yashima, *Macromolecules* **2000**, *33*, 4616.

75 H. Onouchi, T. Hasegawa, D. Kashiwagi, H. Ishiguro, K. Maeda, E. Yashima, *Macromolecules*, **2005**, *38*, 8625.

76 K. Nagai, K. Maeda, Y. Takeyama, K. Sakajiri, E. Yashima, *Macromolecules*, **2005**, *38*, 5444.

77 K. Maeda, Y. Takayama, K. Sakajiri, E. Yashima, *J. Am. Chem. Soc.* **2004**, *126*, 16284.

78 K. Nagai, K. Sakajiri, K. Maeda, K. Okoshi, T. Sato, E. Yashima, *Macromolecules*, **2006**, *39*, 5371.

79 M. M. Green, S. Zanella, H. Gu, T. Sato, G. Gottarelli, S. K. Jha, G. P. Spada, A. M. Schoevaars, B. Feringa, A. Teramoto, *J. Am. Chem. Soc.* **1998**, *120*, 9810.

80 E. Yashima, H. Goto, Y. Okamoto, *Polym. J.* **1998**, *30*, 69.

81 K. Maeda, H. Goto, E. Yashima, *Macromolecules* **2001**, *34*, 1160.

82 J. Tabei, R. Nomura, F. Sanda, T. Masuda, *Macromolecules*, **2003**, *36*, 8603.

83 K. Maeda, N. Yamamoto, Y. Okamoto, *Macromolecules* **1998**, *31*, 5924.

84 R. Sakai, T. Satoh, Kakuchi, H. Kaga, T. Kakuchi, *Macromolecules*, **2003**, *36*, 3709.

85 R. Sakai, T. Satoh, R. Kakuchi, H. Kaga, T. Kakuchi, *Macromolecules*, **2004**, *37*, 3996.

86 P. Dellaportas, R. G. Jones, S. J. Holder, *Macromol. Rapid Commun.* **2002**, *23*, 99.

87 H. Nakashima, J. R. Koe, K. Torimitsu, M. Fujiki, *J. Am. Chem. Soc.* **2001**, *123*, 4847.

88 S. Haraguchi, T. Hasegawa, M. Numata, M. Fujiki, K. Uezu, K. Sakurai, S. Shinkai, *Org. Lett.* **2005**, *7*, 5605.

89 T. Sanji, N. Kato, M. Kato, M. Tanaka, *Angew. Chem. Int. Ed.* **2005**, *44*, 7301.

90 M. Ishikawa, K. Maeda, E. Yashima, *J. Am. Chem. Soc.* **2002**, *124*, 7448.

91 E. Yashima, K. Maeda, T. Yamanaka, *J. Am. Chem. Soc.* **2000**, *122*, 7813.

92 M. R. Majidi, L. A. P. Kane-Maguire, G. G. Wallace, *Polymer*, **1995**, *36*, 3597.

93 L. Arnt, G. N. Tew, *Macromolecules*, **2004**, *37*, 1283.

94 C. Li, M. Numata, A.-H. Bae, K. Sakurai, S. Shinkai, *J. Am. Chem. Soc.* **2005**, *127*, 4548.

95 T. Sanji, N. Kato, M. Tanaka, *Org. Lett.* **2006**, *8*, 235.

96 K. P. R. Nilsson, J. Rydberg, L. Baltzer, O. Inganäs, *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 11197.

97 Y. Inai, K. Tagawa, A. Takasu, T. Hirabayashi, T. Oshikawa, M. Yamashita, *J. Am. Chem. Soc.* **2000**, *122*, 11731.

98 E. Yashima, K. Maeda, Y. Okamoto, *Nature* **1999**, *399*, 449.

99 K. Maeda, K. Morino, Y. Okamoto, T. Sato, E. Yashima, *J. Am. Chem. Soc.* **2004**, *126*, 4329.

100 H. Onouchi, D. Kashiwagi, K. Hayashi, K. Maeda, E. Yashima, *Macromolecules*, **2004**, *37*, 5495.

101 T. Miyagawa, A. Furuko, K. Maeda, H. Katagiri, Y. Furusho, E. Yashima, *J. Am. Chem. Soc.* **2005**, *127*, 5018.

102 H. Onouchi, T. Miyagawa, A. Furuko, K. Maeda, E. Yashima, *J. Am. Chem. Soc.* **2005**, *127*, 2960.

103 K. Maeda, Y. Matsushita, M. Ezaka, E. Yashima, *Chem. Commun.* **2005**, 4152.

104 Y. Furusho, T. Kimura, T. Aida, *J. Am. Chem. Soc.* **1997**, *119*, 5267.

105 A. Sugasaki, M. Ikeda, M. Takeuchi, A. Robertson, S. Shinkai, *J. Chem. Soc. Perkin Trans. 1*, **1999**, 3259.

106 L. J. Prins, F. de Jong, P. Timmerman, D. N. Reinhoudt, *Nature*, **2000**, *408*, 181.

107 Y. Kubo, T. Ohno, J. Yamanaka, S. Tokita, T. Iida, Y. Ishimaru, *J. Am. Chem. Soc.* **2001**, *123*, 12700.

108 R. Lauceri, A. Raudino, L. M. Scolaro, N. Micali, R. Purrello, *J. Am. Chem. Soc.* **2002**, *124*, 894.

109 M. Ishikawa, K. Maeda, Y. Mitsutsuji, E. Yashima, *J. Am. Chem. Soc.* **2004**, *126*, 732.

110 K. Maeda, M. Ishikawa, E. Yashima, *J. Am. Chem. Soc.* **2004**, *126*, 15161.

111 F. M. Pohl, T. M. Jovin, *J. Mol. Biol.* **1972**, *67*, 375.

112 H. Toriumi, N. Saso, Y. Yasumoto, S. Sasaki, I. Uematsu, *Polym. J.* **1979**, *11*, 977.

113 G. Maxein, R. Zentel, *Macromolecules* **1995**, *28*, 8438.

114 M. Fujiki, *J. Am. Chem. Soc.* **2000**, *122*, 3336.

115 K. Tang, M. M. Green, K. S. Cheon, J. V. Selinger, B. A. Garetz, *J. Am. Chem. Soc.* **2003**, *125*, 7313.

116 K. K. L. Cheuk, J. W. Y. Lam, J. Chen, L. M. Lai, B. Z. Tang, *Macromolecules*, **2003**, *36*, 5947.

117 S.-i. Sakurai, K. Okoshi, J. Kumaki, E. Yashima, *J. Am. Chem. Soc.* **2006**, *128*, 5650.

118 E. Yashima, Y. Maeda, Y. Okamoto, *J. Am. Chem. Soc.* **1998**, *120*, 8895.

119 E. Yashima, K. Maeda, O. Sato, *J. Am. Chem. Soc.* **2001**, *123*, 8159.

120 K. Maeda, H. Mochizuki, M. Watanabe, E. Yashima, *J. Am. Chem. Soc.* **2006**, *128*, 7639.

121 Y. Okamoto, I. Okamoto, H. Yuki, S. Murata, R. Noyori, H. Takaya, *J. Am. Chem. Soc.* **1981**, *103*, 6971.

122 T. Nakano, *J. Chromatogr. A*, **2001**, *906*, 205.

123 T. Aoki, T. Kaneko, *Polym. J.* **2005**, *37*, 717.

124 M. Reggelin, M. Schultz, M. Hobbach, *Angew. Chem. Int. Ed.* **2002**, *41*, 1614.

125 M. Reggelin, S. Doerr, M. Klussmann, M. Schultz, M. Hobbach, *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 5461.

126 L. Pu, *Chem. Rev.* **1998**, *98*, 2405.

127 T. Nishimura, K. Takatani, S.-i. Sakurai, K. Maeda, E. Yashima, *Angew. Chem. Int. Ed.* **2002**, *41*, 3602.

128 T. Nishimura, K. Tsuchiya, S. Ohsawa, K. Maeda, E. Yashima, Y. Nakamura, J. Nishimura, *J. Am. Chem. Soc.* **2004**, *126*, 11711.

129 J. Zhang, M. T. Albelda, Y. Liu, J. W. Canary, *Chirality*, **2005**, *17*, 404.

130 J. A. A. W. Elemans, A. E. Rowan, R. J. M. Nolte, *J. Mater. Chem.* **2003**, *13*, 2661.

12
Polyisocyanides: Stiffened Foldamers

Matthijs B.J. Otten, Gerald A. Metselaar, Jeroen J.L.M. Cornelissen,
Alan E. Rowan and Roeland J.M. Nolte

12.1
Introduction

In the search for new materials in the field of electronics, biosensing and cataly-sis, materials that not only possess the structural integrity and flexibility of many naturally occurring materials, but also their functionality are considered to be of great potential. Nature teaches us that the creation of well-defined structures is always accompanied by a loss of entropy, which needs to be compensated for by either an increase in favorable enthalpic interactions or a gain in entropy of the environment (e.g. hydrophobic interactions usually lead to increase of entropy due to release of water molecules; see Chapter 3). In well-known robust biological architectures such as the α-helix and the β-sheet, the entropic loss these peptide segments encounter upon folding is repaid by favorable steric, hydrophobic, elec-trostatic and hydrogen-bonding interactions in the secondary structure. Fol-damers, which have been studied as artificial mimics of these biomolecular struc-tures, can adopt well-defined secondary arrangements stabilized by non-covalent interactions. They are dynamic in nature and can be influenced by changing the environment, such as pH, salt concentration, solvent and temperature [1, 2].

Polymers, such as polyacetylenes [3, 4] and polyisocyanates [5, 6] can be consid-ered as polymeric foldamers since they are able to arrange themselves into well-defined dynamic helices under specific circumstances. The formation of well de-fined polymeric foldamer structures is a challenge, since entropy can be expected to play a substantial role here: the longer the polymers becomes the larger the chance of a structural mismatch is [6, 7]. Helically locked polymers constitute a special class of foldamers, that is, they can be considered as kinetically locked fol-damers as a result of, for instance, steric hindrance (atropisomerism) or the pres-ence of strong hydrogen bonding networks. Helical polymers are considered to be stable when their helical inversion barrier exceeds $\sim$85 kJ mol^{-1}. Examples of such polymers are sterically restricted poly(methacrylate ester)s, polychlorals, bi-

naphthyl based polymers, polysilanes with steric bulk and polyisocyanides (see Chapter 11) [8, 9].

This chapter will focus on the latter class of polymers, which were the first to be reported as possessing a stable helical conformation [10]. Their preparation, structure and the use of these polymers in the creation of functional materials will be discussed.

12.2
Preparation

Polyisocyanides, also known as polyisonitriles or polyiminomethylenes, are prepared by the polymerization of isocyanides. The driving force for this polymerization reaction is the transformation of a formally divalent carbon atom in the monomer to a tetravalent carbon atom in the polymer, yielding a heat of polymerization of 81.4 (kJ mol^{-1}) (Fig. 12.1A) [11].

One of the special characteristics of polyisocyanides is the fact that every carbon atom in the polymer backbone bears a substituent. A consequence of this architectural novelty is that the side chains experience a large steric hindrance forcing the polymer to adopt a non-planar conformation (see below). In addition to a
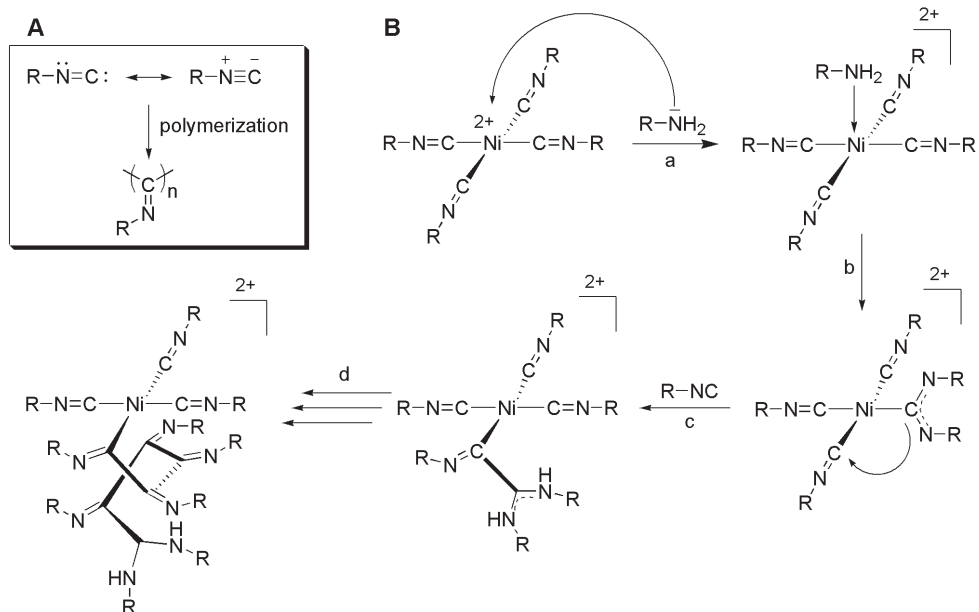
Fig. 12.1 (A) Schematic representation of the resonance structures of isocyanide and polyisocyanide; (B) The ''merry-go-round'' mechanism for the nickel(II)-catalyzed polymerization of isocyanides.

range of polymerization procedures available for isocyanides [12–14], the most successful methods involve the use of a Group 10 metal complex, of which the most widely applied is a Ni(II)-complex, such as Ni(acac) [15], NiCl$_2$ and Ni(ClO$_4^-$)$_2$ [8, 11]. For the nickel-catalyzed polymerization of isocyanides a so-called "merry-go-round" mechanism has been proposed in which the polymerization takes place around the nickel(II) center that pre-organizes the isocyanides for polymerization (Fig. 12.1B) [16]. This mechanism explains many of the features and properties of polyisocyanides, however, detailed mechanistic studies by Deming and Novak on Nickel catalyst **1** (Scheme 12.1) revealed that some aspects of the mechanism are more complex and that the actual catalytic species is probably nickel(I) [17–22]. Nickel catalyst **1** also revealed excellent living polymerization characteristics [17, 18], allowing the formation of block copolymers from two different isocyanides [18]. Block copolymers from polyisocyanide and another type of polymer can be prepared from allyl and amine initiator complexes such as **2** [23, 24], **3** [25–27] and **4** [28], of which the polybutadiene and the poly-
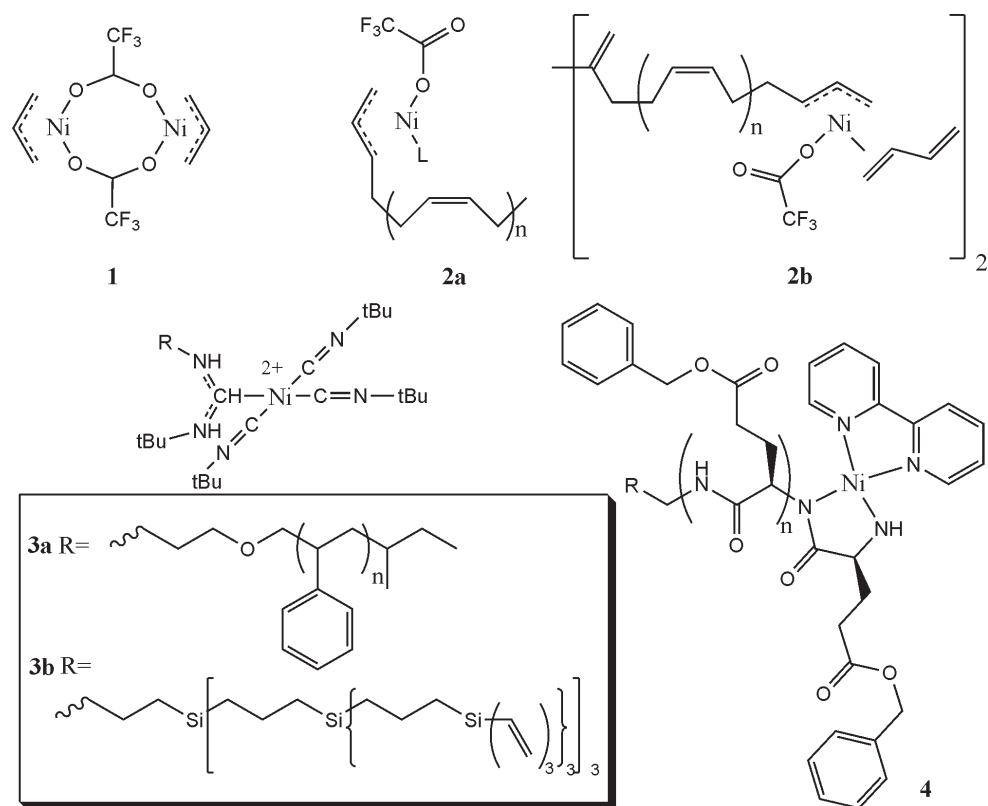
Scheme 12.1

peptide in **2** and **4**, respectively, were polymerized by the same nickel center (Scheme 12.1).

An alternative type of polymerization catalyst, μ-ethynediyl Pd–Pt complex **5a** and μ-ethynediyl Pd–Pd complex **5b** (although the latter catalyst is less efficient), was discovered by Takahashi and coworkers (Scheme 12.2) [29, 30]. They found that **5a** polymerizes aryl isocyanides, but not alkyl isocyanides under reflux conditions in THF. The isocyanides exclusively insert into the Pd–carbon bond, however, the platinum plays an essential role, since only a single insertion of isocyanide was observed for mononuclear complexes **6** in the presence of an excess of isocyanide. The Pd–Pt catalyzed polymerization proved to be living in nature as was illustrated by the low polydispersity of the obtained polymers and the ability of the catalyst to form block copolymers. Even after work-up the Pd end group remains connected to the polymer and polymerization can be continued. Initiators **7** with two and three Pd–Pt μ-ethynediyl units were used to synthesize multi-armed polyisocyanides (Scheme 12.2) [31, 32]. More details on the preparation of polyisocyanides can be found in a recent review by Suginome et al. [14].
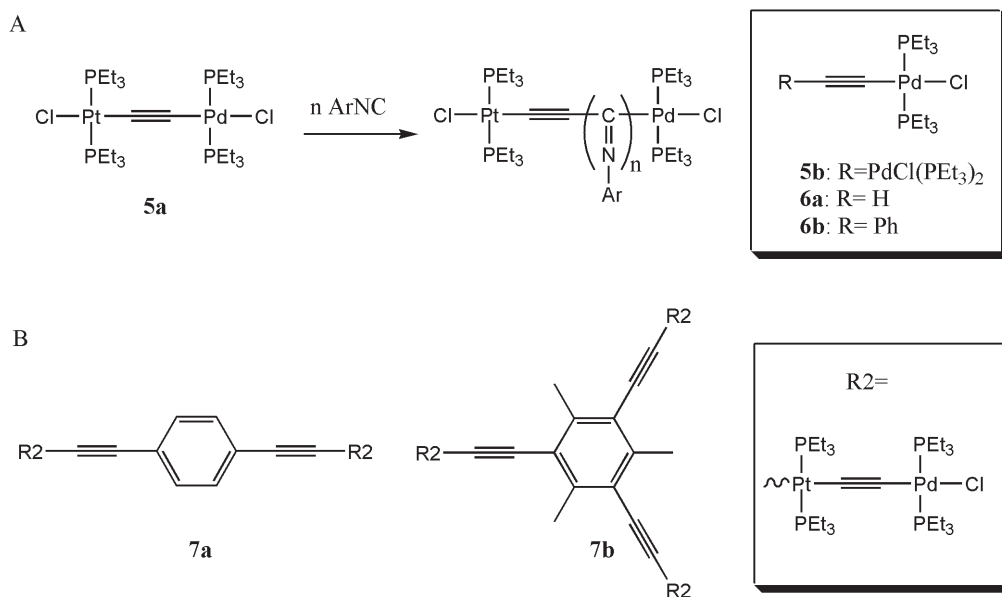
Scheme 12.2

12.3
Conformation

Millich et al. found that the polymerization of the optically active α-phenylethyl isocyanide in the presence of acid treated glass yields polymers with a high

optical rotation per repeat unit [33]. On the basis of this observation, combined with Debye–Scherrer X-ray patterns and space-filling molecular models [34], they proposed that in a polyisocyanide chain the transition dipoles of the absorbing imine chromophores are coupled, leading to a highly organized helical polymer backbone containing 4 repeat units per turn and a pitch of 4.1–4.2 Å [35].

The presumed helical conformation of polyisocyanides was confirmed by Nolte et al. when poly(*t*-butyl isocyanide), which has no chiral centers, was resolved into (+)- and (−)-rotating fractions on the basis of CD spectroscopy [10, 36]. Theoretical studies on the conformation of *t*-butyl isocyanide oligomers using consistent force field conformational calculations indicated that a helical conformation was favored with an increasing number of monomer units. The average dihedral angle N=C–C=N in the hexamer was found to be $\pm 78.6°$, corresponding to 3.75 monomer units per helical turn [37, 38]. The same calculations for a hexadecamer of the *t*-butyl isocyanide resulted in a dihedral angle of 84.3°, corresponding to 3.60 units per helical turn. Substitution of the *t*-butyl group by a methyl-, ethyl-, or isopropyl group was calculated to give a smaller dihedral angle and more units per helical turn. In the case of poly(methylisocyanide), calculations revealed that the methyl group was too small to lead to a fixed dihedral angle and hence no atropisomerism was proposed to be possible.

In the late 1970s, Kollmar and Hoffmann carried out molecular orbital calculations using an extended Hückel approach on a series of polyisocyanides [39], namely RNC, where R = H, CH_3, $C(CH_3)_3$. They concluded that N lone-pair repulsion between the nitrogens that are second nearest neighbor in the polymer chain (Fig. 12.2) plays a dominant role in the structural conformation and as a result the polyisocyanide backbone must adopt a conformation that is not planar.

In the case of isocyanides with bulky R-substituents, electronic repulsion is of minor importance and the non-planar conformation is mainly dictated by the steric interactions between the side groups [39]. According to calculations, the helical angle that is adopted by the polyisocyanide backbone varies from a fairly broad range of helical conformations for the R = H polymer, to a narrow range of configurations around the 4-fold helix as the steric bulk of the substituent in-
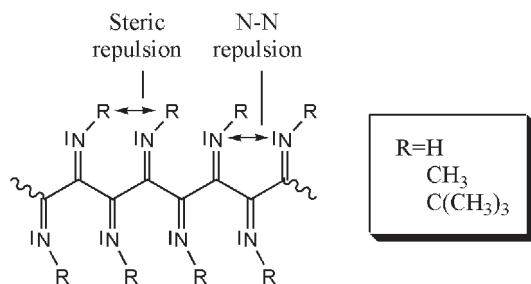
Fig. 12.2 Illustration of the different repulsive interactions driving the all-*anti* imine functions in the polyisocyanide chain out of planarity.
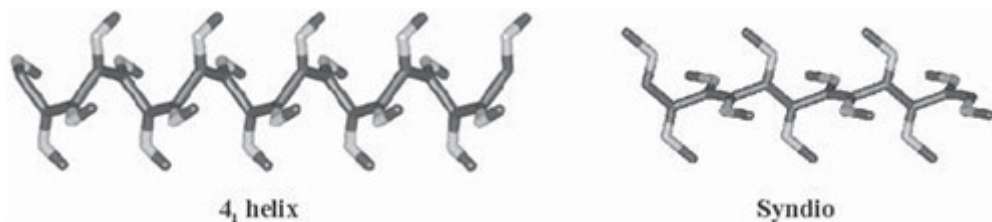
Fig. 12.3 Two proposed most stable conformations for polyisocyanides.

creases to R = C(CH$_3$)$_3$. For intermediate steric bulk (e.g. R = CH$_3$), the authors found two helical minima with different degrees of helicity. The influence of steric bulk of the substituent on the conformation of polyisocyanides was experimentally corroborated by Yamada et al. who found that polyisocyanides derived from phenylalanine readily lose their initial helical conformation in solution when the carboxyl group is protected as an ethyl ester. Increasing the bulk of the ester to a *t*-butyl ester significantly increased the stability of the helical conformation of the polyisocyanide [40].

An alternative conformation for polyisocyanides was postulated on the basis of calculations by Clericuzio et al. [41]. The suggested repulsion between the N lone pairs in the planar all-*anti* conformation of polyisocyanides is absent in the so-called "syndio" conformation in which dimeric sections are alternatively (*E,E*) and (*Z,Z*) (Fig. 12.3). On the basis of both *ab initio* and molecular mechanics (MM) calculations, this conformation was found to be by far the most stable among a number of different possible geometries for polyisocyanides. The syndio conformation is non-helical but highly symmetrical and has a regularly alternating configuration of side chains on C=N double bonds, and an alternating 180° ± 90° conformation of the backbone N=C–C=N angles (Fig. 12.3). The driving force for this conformation seems to lie partly in the large preference for *E,E* trans-planar diiminic units: the rotation around the N=C–C=N central bond in the *E,E* ethane diimine (as calculated by *ab initio*) shows a (s-*cis*)-(s-*trans*) energy difference of ∼34 kJ mol^{-1} in the flexible rotor approximation.

The proposed syndio-conformation was postulated to be very likely for oligo-(phenylisocyanide)s. NMR spectroscopic investigations demonstrated that the 1,3-diphenyl-1,2,3-tri(phenylimino)propane has a planar diimine unit with the third imine located at a 90° dihedral angle from the planar s-*trans* portion [42]. For polymers of this compound, i.e. poly(phenylisocyanide), the authors observed that in halocarbon solvents or tetrahydrofuran solution the conformation slowly changes from a helix to a random coil [43].

In addition to the computational approaches which suggest possible alternative regular conformations for polyisocyanides different from the helical one, Green et al. pointed out that polyisocyanides, especially with small pendant groups, actually adopt a quite irregular conformation. They emphasized the difference in the Mark–Houwink exponent *a* between several polyisocyanides, obtained by viscos-

ity measurements [44]. The Mark–Houwink exponent gives an indication about the rigidity of molecules [45]: if $a < 1$, the polymer has a random-coil character and if $a > 1$, the polymers have a rigid-rod character. The Mark–Houwink exponents for poly(α-phenylethyl isocyanide) (**10**) in toluene, β-phenylethyl isocyanide in tetrahydrofuran and racemic 2-octyl isocyanide in toluene were found to be $a = 1.36$ [46], 0.68 [47] and 1.75 [48, 49], respectively. These results suggest that the structure of the pendant group strongly affects the chain dimensions. Using light scattering experiments Green et al. showed that even the relatively stiff poly(α-phenylethyl isocyanide) **10** has only a limited persistence length of $\sim$3 nm (polyisocyanopeptides by contrast, possess much longer persistence lengths; see Section 12.4). In addition, a large chemical shift dispersion for all carbon atoms, including the backbone carbon, was found in the ^{13}C NMR spectrum of this polymer. This dispersion was even stronger for polyisocyanides lacking an α-substituent, such as **8** and **9** (Fig. 12.4). Since a broad chemical shift dispersion was also observed for polymers from achiral monomer units, Green et al. suggested that this stereo-irregularity is associated with *syn-anti* isomerism about the carbon-nitrogen double bond (Fig. 12.4D).
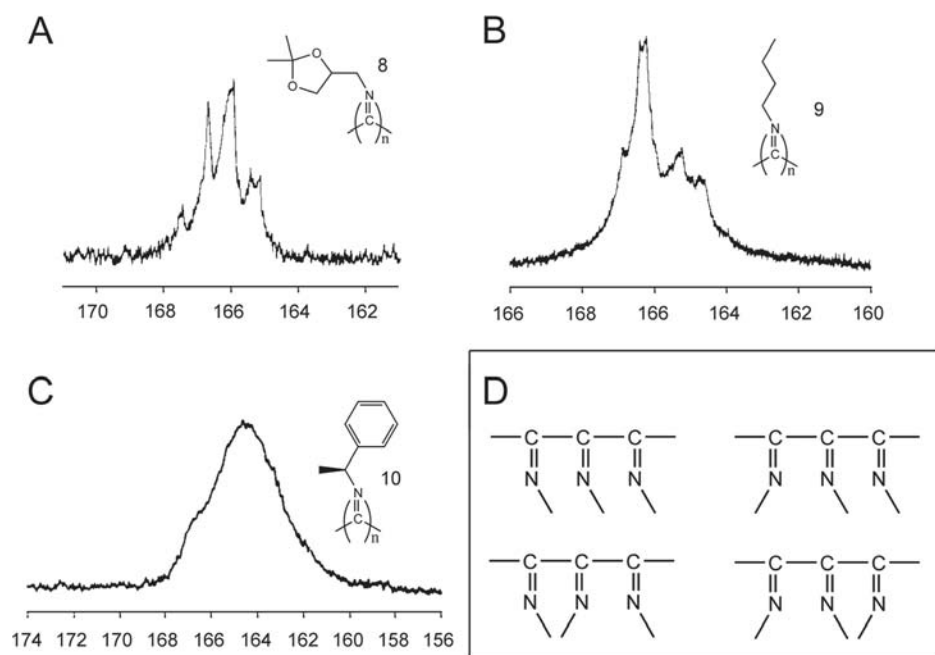
Fig. 12.4 Signal for the imine carbon atom in the polymer backbone in the ^{13}C NMR spectrum for polyisocyanides lacking an α-substituent (A and B) and with an α-substituent (C). (D) Stereoisomeric possibilities for a triad in a polyisocyanide. (Adapted with permission from Ref. 44 Copyright 1988.)
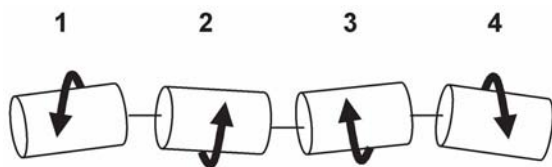
Fig. 12.5 Schematic drawing of the different polymer conformations for polyisocyanides as illustrated by Millich and Baker. Blocks 1–4 correspond to structures with different imine conformations and different helicity (see text).

The carbon nuclei may be sensitive to sequences longer than the triads depicted in Fig. 12.4D. These results indicate that the conformational data for polyisocyanides proposed from interpretation of CD spectroscopy [50–59] are difficult to reconcile with the structural disorder revealed by the high-field ^{13}C NMR spectroscopic data. Millich and Baker [33] already suggested the possibility of blocks with different *syn-anti* isomerism of the imino group together with the possibility of helix reversals as illustrated in Fig. 12.5. Block 1 is obtained by rotation of block 2 around the short axis (similar for 3 and 4), which corresponds to *syn-anti* isomerism of the imino group, while blocks 1 and 3 and blocks 2 and 4 have opposite helix senses. In the case of achiral monomers, 1 and 4, and 2 and 3, are mirror images, whereas for chiral monomer units they are diastereomers.

Takahashi et al. showed that polymerization conditions can have an influence on the stereoregularity of the resulting polymer [60]. Polyisocyanide **12** prepared by polymerizing **11** with NiCl$_2$ in methanol at room temperature revealed a lower specific rotation ($[\alpha]_D^{20} = +354$) and CD ($\Delta\varepsilon_{364} = +3.9$) than polymer **13** ($[\alpha]_D^{20} = +1070$ and $\Delta\varepsilon_{364} = +13.0$), which was prepared by Pd–Pt catalyst **5a** in refluxing THF (Fig. 12.6). By annealing polymer **12** in refluxing THF for 15 h the
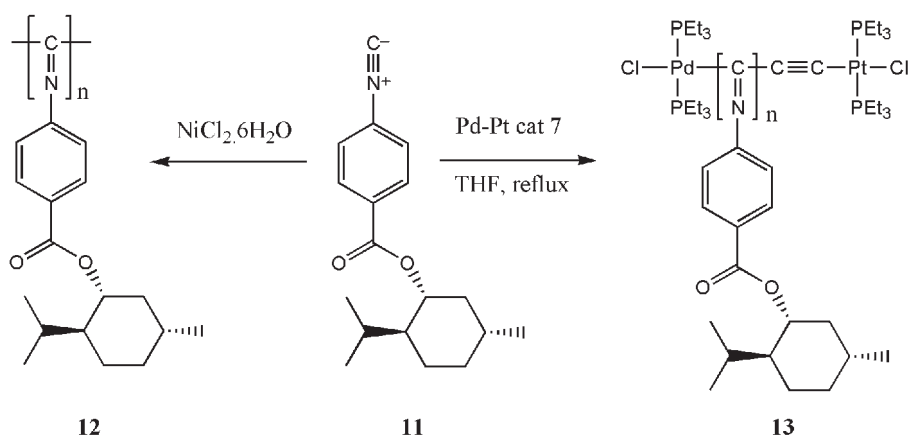
Fig. 12.6 Reaction scheme for the polymerization of **11** by NiCl$_2$·6H$_2$O and Pd–Pt catalyst **5a**.

specific rotation and CD intensity increased to values of $[\alpha]_D^{20} = +1038$ and $\Delta\varepsilon_{364} = +11.6$, whereas no increase was observed for polymer **13**. More information on the conformational changes was obtained by ^{13}C-NMR spectroscopy; a rather broad signal for the imino carbons of the backbone (width at half height: $W_{1/2h} = 208$ Hz) was observed for polymer **12** in contrast to the much sharper signal for polymer **13** ($W_{1/2h} = 133$ Hz). After annealing polymer **12** in deuterated toluene at 80 °C for 15 h, a similar sharp signal as for **13** was found. Comparable results were obtained for other polymers, including achiral polymers, thus the ^{13}C NMR signal distribution for the imino backbone carbon cannot be explained by stereochemical means. The authors therefore proposed, in line with the work of Green et al., that the initial stereo-irregularity in the polymers formed by the nickel-catalyzed polymerization is associated with the existence of both *syn*- and *anti*-isomers of the imino groups in the backbone. The irregular conformation can be transformed into the thermodynamically stable stereoregular form by *syn-anti* isomerization of the imino group at high temperatures. The polymerization at high temperature with Pd–Pt catalyst **5a** immediately leads to the stereoregular conformation.

Yashima et al. further observed that polyisocyanides are not necessarily always present in a stable locked structure, but that some of them can have more dynamic conformations. They provided evidence for a reversible transition between two conformational states of poly(4-carboxyphenylisocyanide) **14** under influence of optically active amines and amino alcohols such as **15** and **16** (Scheme 12.3) [61, 62].

14   **15**   **16**
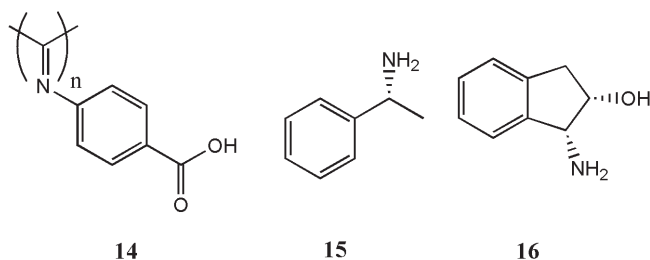
Scheme 12.3

Based on their observations they suggested, that apart from the aforementioned imino *syn-anti* isomerism, these changes are caused by backbone (s-*trans*)-(s-*cis*) isomerism (Fig. 12.7). Initially a 4_1-helix is formed that rapidly looses its regular helix structure to form a irregular structure with s-*trans*, s-*cis* and (s-*cis*)-(s-*trans*) domains (step 1). This structure will slowly transform into the stable all s-*trans* structure at 30 °C resulting in sharper peaks in the ^{1}HNMR spectrum (step 2). It is proposed that the latter structure can give a helical arrangement upon bind-
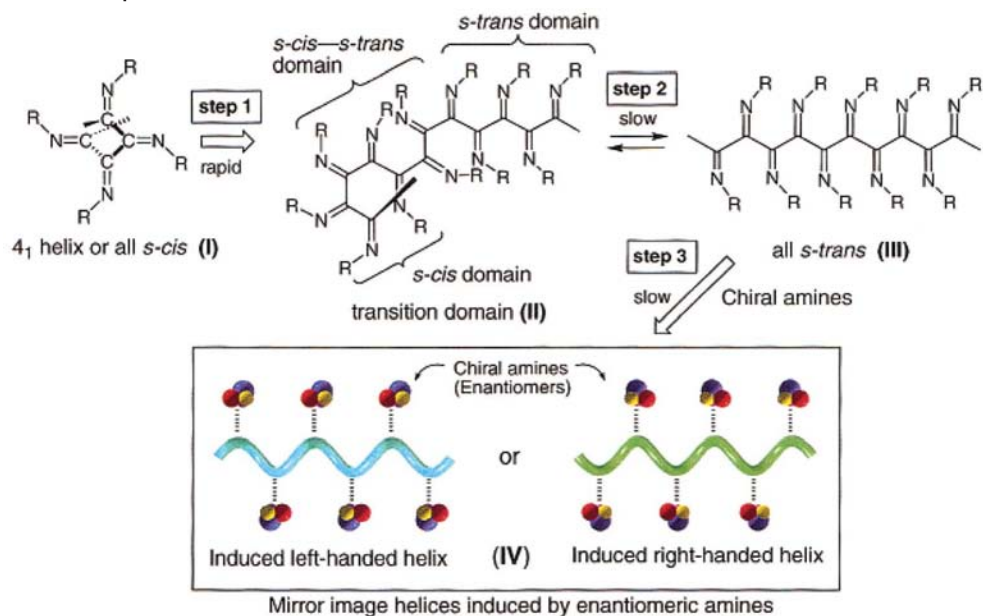
Fig. 12.7 Proposed conformational changes and helix induction in **14**. (Adapted with permission from Ref. 61, Copyright 2002, American Chemical Society.)

ing chiral amines by directing the random twist around the C–C bonds in a single direction (step 3). In water a helix could also be induced, which after removal of the optically active amines maintained its helical conformation at ambient temperatures, but at elevated temperatures the helix unfolded readily. It was postulated that a combination of hydrophobic and chiral ionic interactions in water is responsible for the helix formation and the memory effect because induced helices in DMSO were unable to maintain their helical structure (see also Chapter 11 and Fig. 11.16) [62, 63].

Summarizing, we may conclude that two important structures for polyisocyanides have been proposed: (i) a helix structure that is close to the 4_1-helix initially suggested by Millich et al., which is most likely for polymers with bulky sidearms and (ii) the syndio structure as was calculated by Salvadori et al., which seems to be most favorable for polyisocyanides with small side-groups. In addition to these two distinct structures, Yashima et al. calculated a 12_5-helix for their induced helices [61] and Young et al. both a 4_1-helix and a 3_1-helix [64].

Irregularities that have been observed in polyisocyanides are mainly explained by *syn-anti* isomerism of the imino side-groups and by (s-*cis*)-(s-*trans*) isomerism of the carbon backbone. This latter isomerism leads to structures which are intermediates between a 4_1-helix and the syndio structure. Other possible explanations for observed irregularities are helix inversions as discussed by Millich and

the existence of two types of helical pitches coexisting within one polymer as was calculated by Kollmar and Hoffmann for polyisocyanides with intermediate bulk.

12.4
Stiffening the Helix

In principle, stable, regular polyisocyanides can be formed by polymerization reactions that are under (i) thermodynamic control (driving force = ΔG_{AB}) or (ii) kinetic control (driving force = $\Delta G_B^{\ddagger} - \Delta G_A^{\ddagger}$) (Fig. 12.8). The different possibilities are depicted in Fig. 12.8, for a polymerization of monomer M to a polymer with monomer units in conformations P–M_A and P–M_B, which might correspond to structures that differ in helicity, backbone (s-*cis*)-(s-*trans*) isomerism or imine *syn-anti* isomerism (only two distinct conformations are considered for simplicity). If the barrier between two conformations $\Delta G_{AB}^{\ddagger}$ is sufficiently small (at a certain temperature), the different conformations are kinetically accessible and the polymer structure is under thermodynamic control. The thermodynamically formed conformations are situated on the left side of Fig. 12.8c. Polymerization reactions under thermodynamic control can yield well-defined polymers when the free energy of the most stable conformation is sufficiently lower than that of competing conformations (high ΔG_{AB}); in the case of cooperativity, only small energy differences are sufficient [7].

Only recently has the literature provided some clear examples of regular well-defined polyisocyanides formed under thermodynamic control at elevated temperatures.

Polymerization of aryl isocyanides using the Pd–Pt catalyst as reported by Takahashi et al. is typically done in refluxing THF [29, 30, 65, 66]. Several observations indicate that this type of polymerization is under thermodynamic control. Firstly, as previously mentioned, polymers formed by $NiCl_2$ at room temperature can be converted into a better defined regular conformation at higher temperature, while polymerization of the monomer in refluxing THF with the Pd–Pt catalyst at ∼66 °C already directly leads to this thermodynamic, regular structure [60]. Secondly, from a detailed study on the helix-sense-selective polymerization using chiral oligomer complexes derived from isocyanide **11** [65], it appeared that the rate constants for propagation are virtually identical, independent of whether the monomer that is incorporated has the same or the opposite chirality as the one constituting the initiating oligomer. This observation rules out the possibility of kinetic control. Finally, a nonlinear relation was observed between the amount of chiral monomer excess and the induced helical sense in the polymerization [67]. This is indicative of a thermodynamically driven ''majority rules'' mechanism, of the type observed in the polymerization of polyisocyanates [68, 69]. However, it appeared that only achiral isocyanides with substantial steric bulk could be polymerized with an ongoing helix sense from a chiral oligomer complex, whereas less bulky achiral isocyanides only showed little preference for
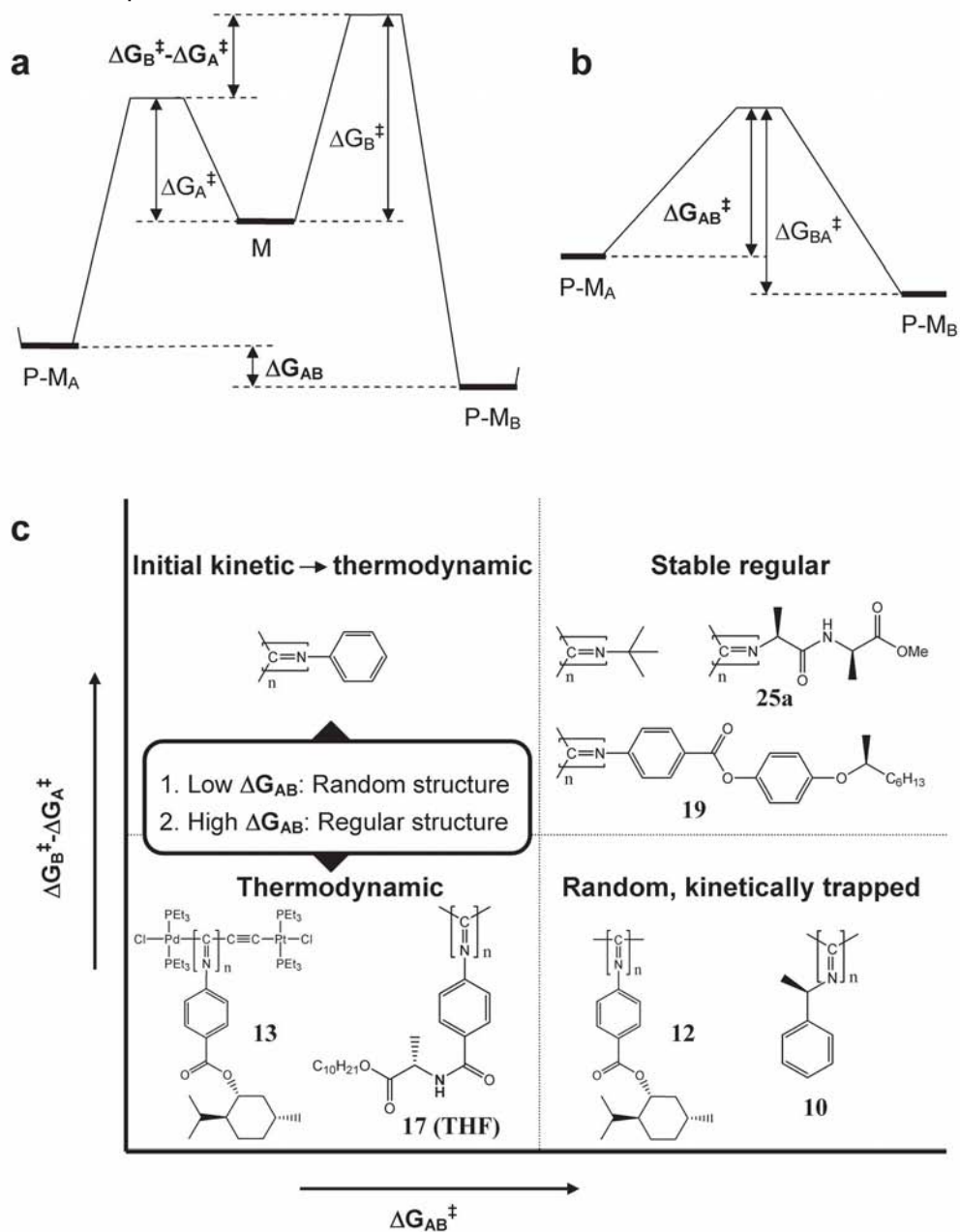
Fig. 12.8 Influence of kinetic and thermodynamic parameters on the polymer structure: (a) Energy diagram for the incorporation of monomer M in conformation A and B of the polymer, with energy levels PM_A and PM_B; (b) Energy diagram for the conversion of a monomer unit in conformation A to conformation B in the polymer; (c) Influence of the degree of kinetic control ($\Delta G_B^{\ddagger} - \Delta G_A^{\ddagger}$) and the kinetic ($\Delta G_{AB}^{\ddagger}$) and thermodynamic stability (ΔG_{AB}) of the monomer units in the polymer on the obtained polymer structures.

a single screw sense [65]. In this case it remains the question whether for the bulky isocyanides, the kinetics rather than the thermodynamics play a role in the helix-sense-selective polymerization.

Yashima showed that the polymerization of phenyl isocyanide **17** (Scheme 12.4) bearing an L-alanine residue can be performed under both kinetic and thermodynamic control [70]. Whereas the polymerization with $NiCl_2$ in toluene and CCl_4 yielded the kinetic product with a positive Cotton effect, the polymerization in THF and even more so in toluene at 100 °C yielded the thermodynamic product, which gave a negative Cotton effect. Under kinetic control, in apolar solvents, hydrogen bonding is thought to play a role in the transition state. In contrast, in polar solvents or at high temperature, hydrogen bonding is suppressed and the thermodynamic product is formed. The role of hydrogen bonding was confirmed by the fact that for the polymer of phenyl isocyanide **18**, which is incapable of hydrogen bonding, a negative cotton effect was observed independent of the polymerization conditions. Remarkably, apart from being manifested by CD spectroscopy, the helix sense of the polymers in self-assembled layers on highly ordered pyrolytic graphite, could be visualized by AFM (See Chapter 11, Fig. 11.4).

17   **18**
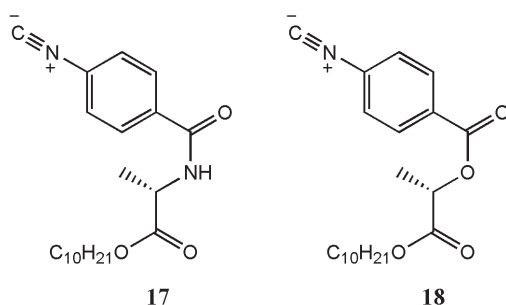
Scheme 12.4

A special case of a thermodynamically formed helix is the previously discussed poly(4-carboxy phenylisocyanide) **14**. The polymer does not form a regular structure by itself, but only upon complexation with an optically active amine, allowing one helix sense to become thermodynamically more favorable than the other. Interestingly, in water the helix structure was retained even without chiral amines present, meaning that the thermodynamically formed structure was kinetically trapped [63].

Polymerization is under kinetic control when the transition state energy $\Delta G_A^{\ddagger}$ for the incorporation of a monomer into the desired configuration P–M_A is sufficiently smaller than that for other configurations P–M_B, that is $(\Delta G_A^{\ddagger} - \Delta G_B^{\ddagger})$ is large. When the formed structure is thermodynamically stable, that is ΔG_{AB} is large, with P–M_A being the lower energy conformation (in this case the polymer could also be formed under thermodynamic control) or when the structure

is kinetically trapped, that is $\Delta G_{AB}^{\ddagger}$ is large, stable well-defined polymers are formed (Fig. 12.8).

Polymerizations catalyzed by nickel at room temperature with bulky monomers are believed to be under kinetic control although in many cases it is not very clear to what extent regular polymers are formed as has been discussed in Section 12.3.

One of the most striking examples of kinetic control is the earlier mentioned polymerization of *t*-butyl isocyanide, which yielded a mixture of M- and P-helices but no mixed M-P-polymers, as was shown by isolating the two polymers by column chromatography using a chiral support [10, 36]. The bulkiness of the *t*-butyl group steers the kinetic control and provides kinetic stability to the formed polymers. Phenyl isocyanide was reported to kinetically give a 4_1-helix during polymerization, however, because of the lack of steric bulk in the side-chain this helix structure is not stable and subsequently unfolds into a random coil polymer [43]. The initial formation of a helix illustrates the important role of the nickel catalyst in providing kinetic control over the reaction, presumably via the merry-go-round mechanism. Deming and Novak also showed that in the polymerization of less bulky isocyanides no complete stereo control is obtained; for a racemic mixture of methylbenzylisocyanide, R and S isomers were mixed in the same helix [71].

For bulky monomers, the occurrence of kinetic control in the nickel(II) mediated polymerization of isocyanides was nicely illustrated by an inhibition experiment [52, 72]. Achiral monomers, which rapidly polymerize, were copolymerized with a slowly polymerizing bulky chiral isocyanide. Instead of imposing its own helix sense (say P) on the achiral monomer, it was found that the chiral isocyanide promoted the formation of a polymer with the opposite screw sense (M). This intriguing result was explained by kinetic inhibition of the formation of one helix type by the bulky monomer. Whereas the achiral isocyanides will normally form both M and P helices, one of the two helices is inhibited in the copolymerization (say P) because of the incorporation of the slowly polymerizing chiral monomer. A variation on this experiment was performed by Amabilino et al. [73], who showed that diastereomers of polymers **19** and **20** could be formed by kinetic inhibition of the growth of the normally occurring helix using the slow polymerizing isocyanide **22** of the same chirality as co-monomer.

Work of the same authors also revealed that apart from steric bulk, other interactions between the monomers can influence the polymerization in a well defined way [74, 75]. Polyisocyanides **19–21** are derived from promesogenic monomers (Scheme 12.5), which are able to induce cholesteric and chiral smectic C phases in nematic and smectic C liquid crystals, respectively [74, 75]. Upon polymerization, in most cases the handedness of the polymers turned out to be the same as that of the monomer induced LC phases. The long range chiral induction by the stereogenic center in the tail was explained by stereoselective interaction of the incoming monomer with the growing polymer in a similar fashion as observed in the LC-phase. It is the rigid nature of the phenyl benzoate group that allows the transfer of chirality from the side chain to the isocyanide func-
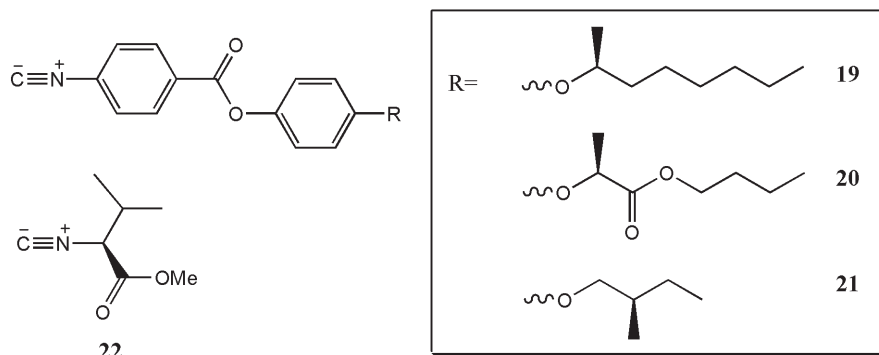
Scheme 12.5

tionality. The CD signal related to the imine backbone was shown to diminish rapidly and invert in sign when the chiral center in the alkyl tail was placed further away from the promesogenic group. More evidence for the role of non-covalent interactions on the stereoselectivity was obtained by performing polymerizations at various concentrations and in different solvents [75]. Variable temperature CD measurements revealed that polymer **19** has a stable conformation up to at least 55 °C. When, however, a nitro group was introduced in the phenyl ring close to the stereogenic center, the polymer was found to adopt a less stable conformation, as was demonstrated by the dramatic and irrecoverable loss of optical activity at 55 °C. This observation confirms that the formed polyisocyanide is a kinetically-determined product [76].

In a more recent publication, the influence of the length and the rigidity of the rod-like spacer located between the isocyanide and the chiral center of **23** were studied (see Scheme 12.6). The chiral induction through spacers **A–G** was investigated in detail and it was concluded that a semi-rigid twisted conformation must be adopted by the spacer in order to be able to effectively transfer the chiral information to a helical polymer backbone. The phenyl benzoate spacer **A** proved to do this most efficiently while no chiral induction is observed for flexible spacer **C**. A 21-Å long teleinduction was observed for spacer **G** [77, 78].

Cornelissen et al. showed the potency of peptide substituents in the formation of stable polyisocyanide helices (Fig. 12.9) [79]. It was found that the helical conformation of a polyisocyanide can be effectively stabilized if a well-defined hydrogen bonding network is present between the peptide side chains at positions n and $(n+4)$, which are stacked above each other at a distance of ~4.6 Å (Fig. 12.9B). Although polyisocyanides derived from peptides had been previously described by the authors [59, 80, 81], they did not recognize at that time the presence of the hydrogen bonds between the peptide side chains and did not utilize this properly to stiffen the helix of polyisocyanides with functional groups (*vide infra*). The stepwise addition of **34a** to a nickel(II) catalyst revealed the kinetic

Spacer=
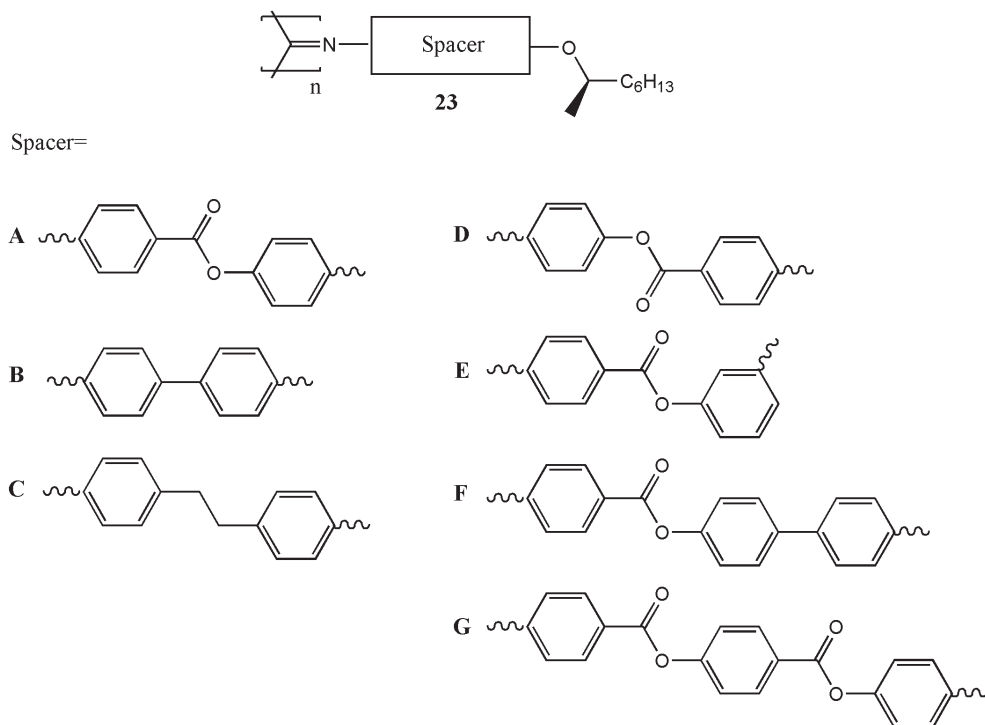
Scheme 12.6

nature of the polymerization, which involves a helical templating effect of the growing polymer on the incoming monomers. After eight monomers were added to the catalyst, a steep increase in the CD signal was observed. At the same time hydrogen bonding interactions between the amide protons started to become visible by ^{1}HNMR spectroscopy [82].

Polyisocyanide **33**, derived from trialanine, contains two amide groups per side chain and is able to fold into a β-sheet-like architecture, mimicking the interactions present in naturally occurring β-helices [83]. Detailed infrared and ^{1}H NMR spectroscopic investigations showed that nearly all amide groups present in polymers **24–33** participated in hydrogen bonding, in a similar way as observed in the crystal structure of **34a** (e.g. see Fig. 12.10A, B). Ordered arrays of hydrogen bonds along the polymeric backbone, however, were not observed for polyisocyanide **27**, which is derived from alanine glycine [79, 84]. It is remarkable that in contrast to **27**, polyisocyanide **26** derived from glycine-alanine did give a well-defined helical structure, suggesting that the steric bulk in the second amino-acid is of great importance, probably because it stabilizes and directs the hydrogen bonding network. Analogous to the denaturation of proteins, the hydrogen bonds in these polymers can be disrupted leading to unfolding of the helix.
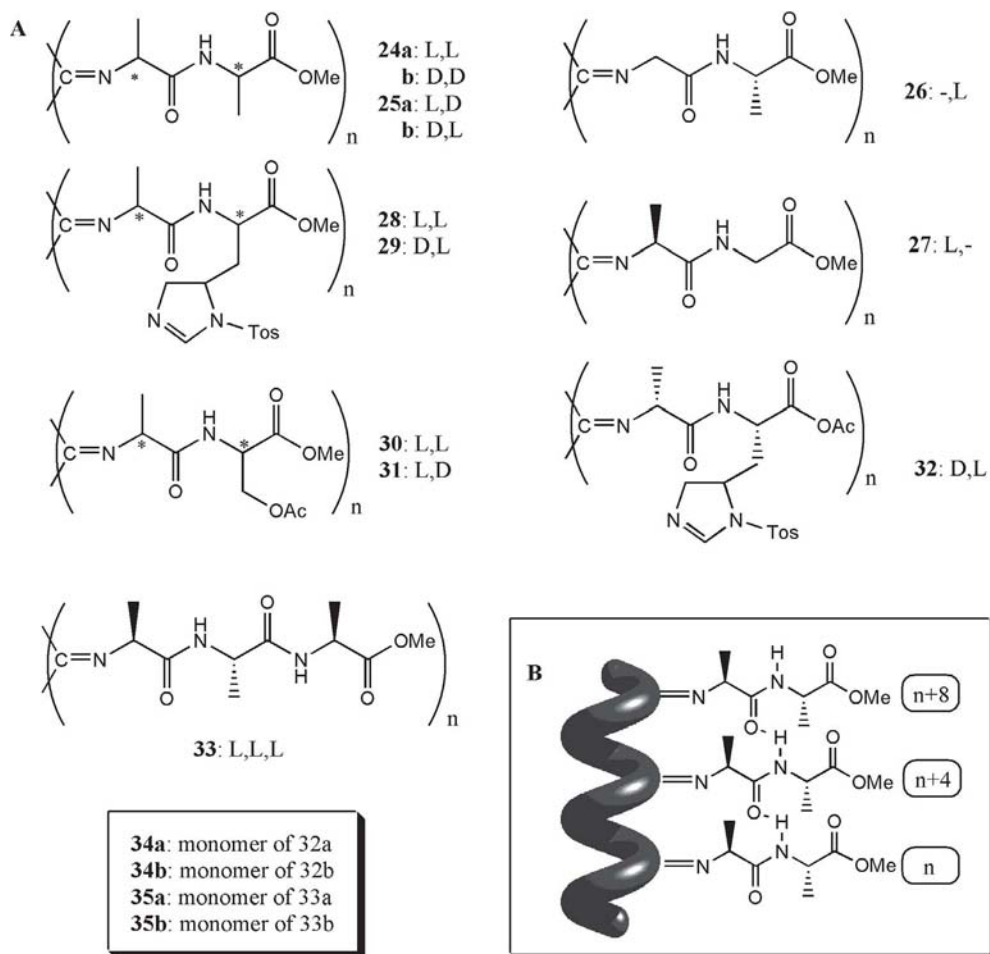
Fig. 12.9 (A) Various polyisocyanides derived from peptides;
(B) Schematic representation of a helical polyisocyanide stabilized
by a hydrogen bonding network between the peptide side-chains.

This unfolding is, however, only possible with strong acids such as trifluoroacetic acid (TFA) and not with hydrogen bonding solvents (e.g. methanol, DMSO), thereby demonstrating the robust character of the hydrogen bonding arrays [79, 81]. Powder X-ray diffraction (PXRD) experiments showed that in the solid state the rigid polyisocyanopeptides are organized in a pseudo-hexagonal arrangement. The acidified samples, which were studied for comparison, in contrast, only gave broad signals pointing to a decreased level of organization in the polymer structure.

The peptide-derived polyisocyanides are stable in solution at room temperature, and as a result of their rigidity, it is possible to visualize the individual macro-
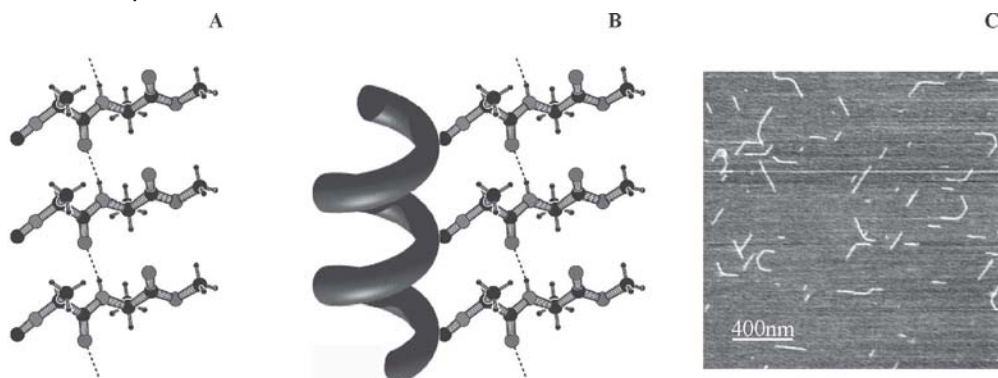
Fig. 12.10 (A) Crystal structure of **34a**; (B) Schematic representation of the proposed orientation of the peptide side chains in **24a**; (C) AFM-micrograph of **24a** prepared with 1/30th equivalent of Ni(ClO$_4$)$_2$·6H$_2$O; (D) AFM-micrograph of **25a** prepared with 1/32th equivalent of TFA.

molecules by atomic force microscopy (AFM) (Fig. 12.10C) [79, 84, 85]. By measuring the contour lengths and by a careful analysis of the curvatures it was possible to determine the molecular weight, the polydispersity, and the persistence length of the polymers. The latter was found to amount to 76 nm, highlighting that these polymers are more rigid than double-stranded DNA [85]. An accurate value of 1.6 nm for the height of the fibers was obtained by AFM measurements under chloroform vapor [86], which corresponds well with the polymer chain as derived from molecular modeling and PXRD measurements [79].

The assignment of the helix sense of peptide-derived polyisocyanides by CD spectroscopy is hampered by the overlap of signals arising from the polymer backbone and the side chains. For an L-alanine based polyisocyanide containing a spectator group (i.e. a diazo chromophore) in the side chains, a right-handed (P) helical geometry was found [87]. Since the helix sense in polyisocyanides is kinetically controlled, this handedness was tentatively assigned to all L-alanine derived polyisocyanides. Selected properties of polyisocyanodipeptides (**24–33**) are presented in Table 12.1 [88]. When hydrogen bonds are present (e.g. **24**, **25** and **30**), a positive optical rotation and a strong positive Cotton effect around $\lambda = 315$ nm indicate the presence of a right-handed (P) helix. When this is not the case (e.g. **27**), the Cotton effect appears at lower wavelength and has an opposite sign and a lower intensity (Table 12.1). From IR and NMR spectroscopic studies it was concluded that, polyisocyanides **24**, **25** and **33** retain their hydrogen bonded helical conformation for significant periods of time [89], even when they are dissolved in water after removal of the methyl ester functions. The thermal denaturation of these water-soluble polymers was also studied in water. It was demonstrated using VT-CD spectroscopy that the denaturation process proceeds in a cooperative fashion [79].

Table 12.1

Compound	[α] (°·dl/g·dm)		Δε(λ) (l/mol·cm)	Screw sense	H-bonds
	Monomer	Polymer			
24a	33	338	5.6 (313)	P	Yes
25a	−5.6	487	5.8 (307)	P	Yes
26	19.2	196	5.1 (321)	P	Yes
27	5.7	−32	−1.5 (290)	P	No
28	164	580	6.5 (304)	P	Yes
			−0.45 (345)		
29	75	−700	1.22 (270)	M	No
30	−58.1	205	8.6 (310)	P	Yes
			0.17 (355)		
31	−12.0	−33	−1.3 (263)	P	No
32	14	−610	−5.5 (300)	M	unclear
33	2.4	188	2.5 (321)	P	Yes

Previously, Millich had found that acid-coated glass can act as a catalyst for the polymerization of isocyanides, although not very efficiently [33, 46]. Metselaar et al. recently reported that the acid (TFA)-initiated polymerization of isocyanopeptides leads to extremely long polymers with lengths up to 14 μm (Fig. 12.10D) [85, 90]. For the polymerization of **35a** at a TFA concentration of 1 mM, kinetic studies revealed a large entropy of activation (-170 J mol^{-1} K^{-1}), which indicates a very high degree of organization in the transition state [82]. At higher acid concentrations instead of polymerization, conversion of the monomer to the corresponding formamide was observed. Based on this result a polymerization mechanism was proposed in which first a helical oligomer is formed acting as a template for the incorporation of subsequent monomer units through a supramolecular complex (Fig. 12.11). In the case of high acid concentrations the template is disrupted and destroyed. The reaction is highly stereospecific since the addition of the enantiomeric monomer **35b** completely blocked propagation of the polymerization of **35a** even when present in only 1%. The diastereomer **34a**, but not **34b**, could be incorporated into the growing polymer, although **34a** itself without **35a** present could not be polymerized with TFA. These subtle differences demonstrate the critical effect of the configuration of the first chiral center of the monomer on the polymerization reaction and the high stereospecificity of the transition state. The fact that **34a** itself cannot be polymerized with TFA was attributed to the inability of **34a** to form a helical template due to larger steric repulsion between the monomeric units in the helix. When a nickel catalyst was used all monomer combinations could be readily polymerized.

Wezenberg et al. showed that polyisocyanides **36** and **37** derived from β-amino acids also form well defined rod-like polymers (Scheme 12.7) [91]. The kinetically
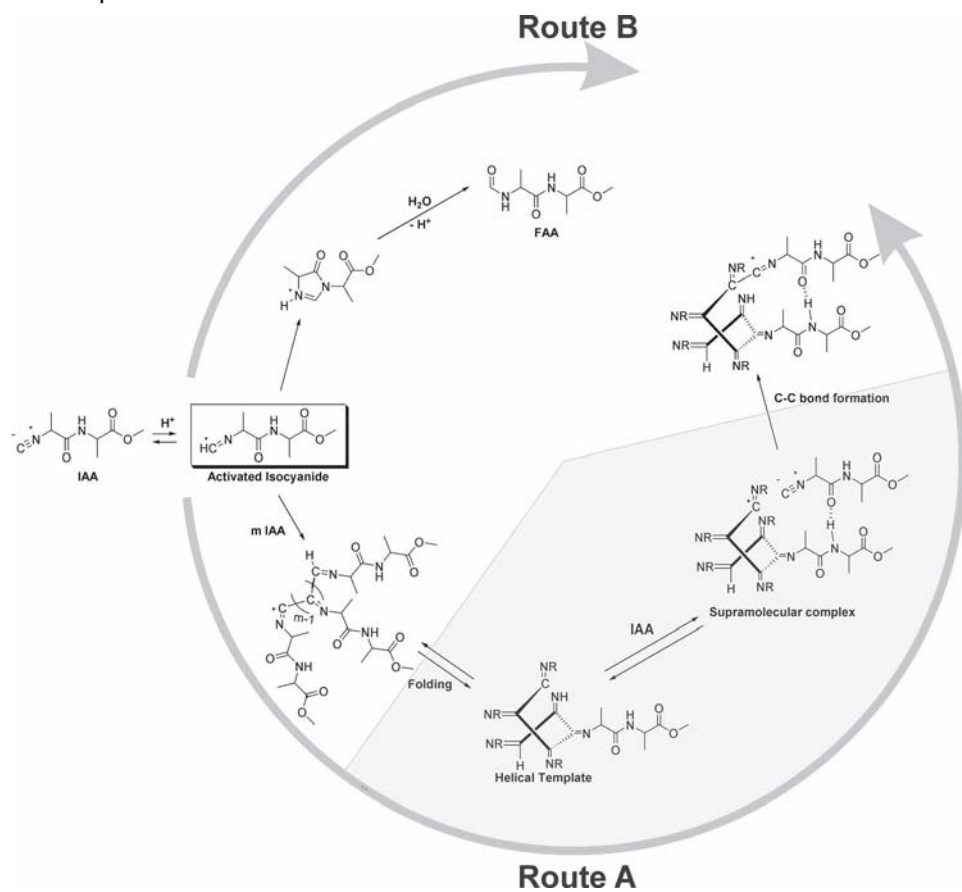
Fig. 12.11 Mechanism of the acid initiated polymerization of **35a**, showing the helical template formation and the subsequent polymerization (Route A). The side reaction to the corresponding formamide, which occurs at high TFA concentrations is shown in Route B.
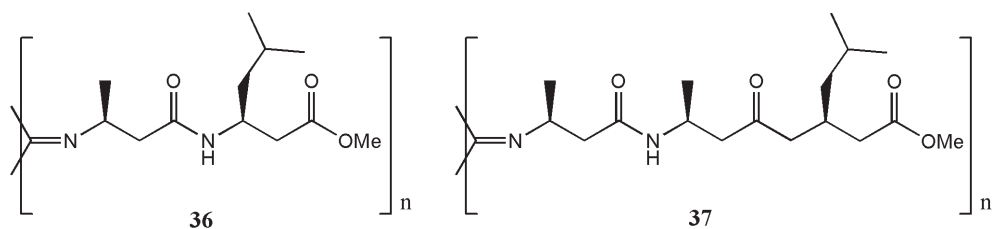
Scheme 12.7

formed polymer, however, turned out to be unstable and transformed into a more stable structure, which possessed a better defined hydrogen bonding pattern as was concluded from IR and temperature-dependent CD studies. The precise structure of the transformed polymer remained unclear.

12.5
Functionalized Polyisocyanides

As was discussed in the previous sections, polyisocyanides possess a unique polymer backbone that can adopt a very stable rigid conformation when bulky side chains or additional stabilizing interactions between the repeat units, like $\pi-\pi$ stacking and hydrogen bonding are present. Polyisocyanides, therefore, are very attractive scaffolds to arrange functional groups in well defined arrays, thus creating new materials with special properties. Many of the functionalized polyisocyanides already have been described in excellent reviews, for example by Cornelissen et al. and very recently by Suginome et al. [8, 9, 11, 14], and will not be discussed here.

Bioinspired sugar (**38–40**) [51, 92, 93] and cholesterol (**41**) [94] derived polyisocyanides and polymers from isocyanides functionalized with ionic side groups (**42–45**) [54, 95–97] have all been reported (Scheme 12.8). Amphiphilic isocyanide **42** forms vesicles in water and the isocyanide functions could be polymerized leading to cross-linking of the bilayer [96]. Iyoda et al. have shown that it is also possible to polymerize isocyanides with extremely bulky groups using Pd–Pt complex **5a** as the catalyst. The dendronized polyisocyanides **46** were synthesized with a polymerization degree exceeding 100 [98]. Using such an approach, polyisocyanides bearing even polystyrene side-groups (**47**) could be synthesized with a polymerization degree of *ca.* 50 [99].

Polyisocyanides also have been used as well-defined templates in catalysis, crystal growth and for the transfer of chirality (Scheme 12.9). Polymeric catalysts have been prepared from polyisocyanide **48**, after coordination of Rh-catalysts to the phosphor ligands in the side chains [100–102]. The stability and catalytic activity of the rhodium complex of **48** towards the hydrogenation of cyclohexene was found to be better than that of the monomeric rhodium complex. The catalytic activity was also better than the activity of analogous polymers with a flexible polystyrene backbone. The latter result was ascribed to the rigidity of the polyisocyanide, which makes the catalytic centers more accessible.

The water-stable peptide-derived polyisocyanide **49a** was used as a supramolecular template for the crystallization of $CaCO_3$. The formation of unusually shaped calcite was found to be controlled by nucleation and adsorption processes involving **49a** [89, 103]. The fact that less control over the crystallization process was obtained when polymer **49b** was used, demonstrated the specificity of the interaction between **49a** and the growing crystal. Single polymer chains of **49a** could be visualized by AFM by complexing the polymer to amino alkanes [104]. By varying

Scheme 12.8

48   **49a**:LD   **50**
            b:LL
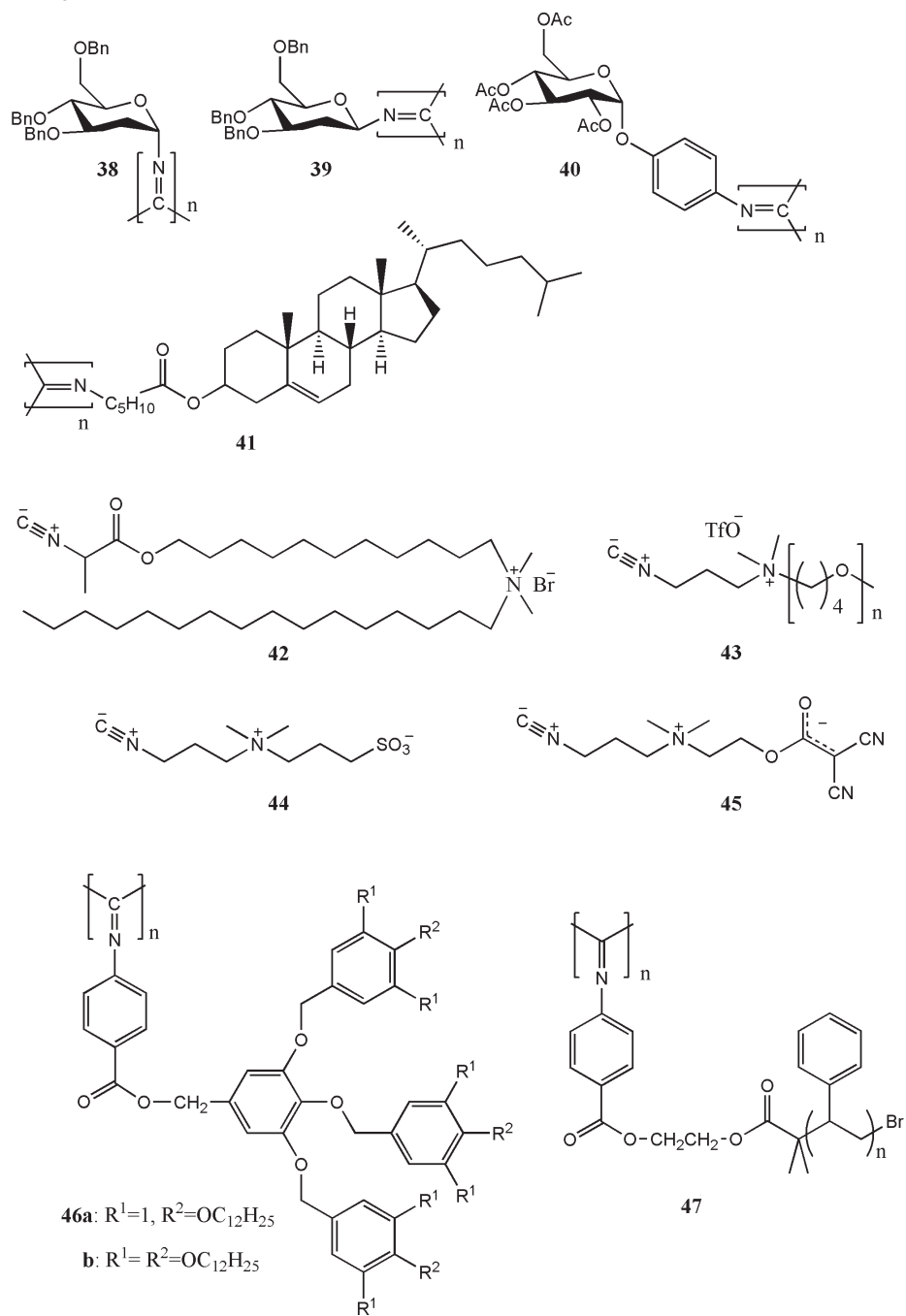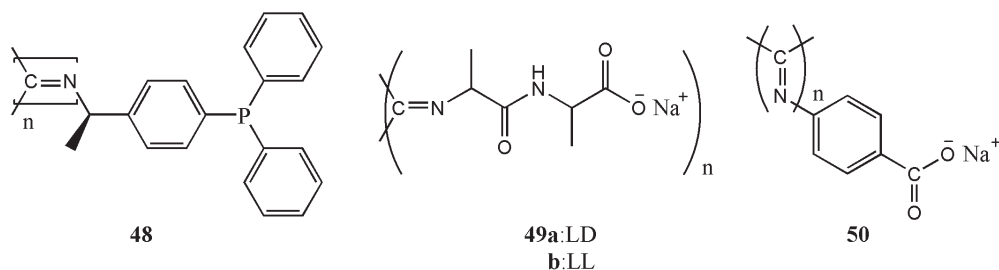
Scheme 12.9

the length of the aminoalkanes, the apparent stiffness of the polymers on the surface could be influenced.

Polyisocyanide **50** with a stable induced helix in water (see Section 12.3, Chapter 11 and Fig. 11.16) [62, 63], was found to be able to induce helicity in another oppositely charged polyelectrolyte based on a polyacetylene [105]. The inter polymer complexation proved to be dependent on both the pH and the salt concentration. This helix replication strategy opens numerous possibilities for the growth of ordered helical assemblies.

Much research has been focused on the use of polyisocyanides as scaffolds for the creation of well defined molecular wires with special magnetic, optical or electronic properties.

Polyisocyanides **51** functionalized with crown ethers of different sizes were reported to yield architectures containing 4 arrays of cofacially stacked rings connected to the polymer backbone (Scheme 12.10) [56, 106]. The polymers appeared to be able to function as ion channels when incorporated within bilayer membranes.

Persoons and coworkers studied polyisocyanide **52** and related compounds, which are functionalized with non-linear optically active (NLO) side groups [107–111]. In solution a first hyperpolarizability exceeding 5000×10^{-30} electrostatic units was measured. Electric field-induced second-harmonic generation measurements revealed a larger non-linear response for the polymer than for the monomer. The second order non-linear response might be further optimized in structures with a smaller angle (60° for this structure) between the polymer backbone and the chromophores [107]. Langmuir–Blodgett films of polymer **52**, showed stable second harmonic generation without the need of poling [110]. This observed NLO effect is thought to originate from the highly defined orientation of the side chains obtained at the air-water interface.

Polyisocyanides **53** and **54** were designed as macromolecular ferromagnets [112], however, no evidence for ferromagnetism or any other short-range ordering among the unpaired spins was observed.

A chiral redox responsive polyisocyanide bearing ferrocenyl groups (**55**) has been reported by Takahashi and coworkers (Scheme 12.10) [113]. Redox cycles of this polymer proved to be completely reversible. CD measurements revealed that
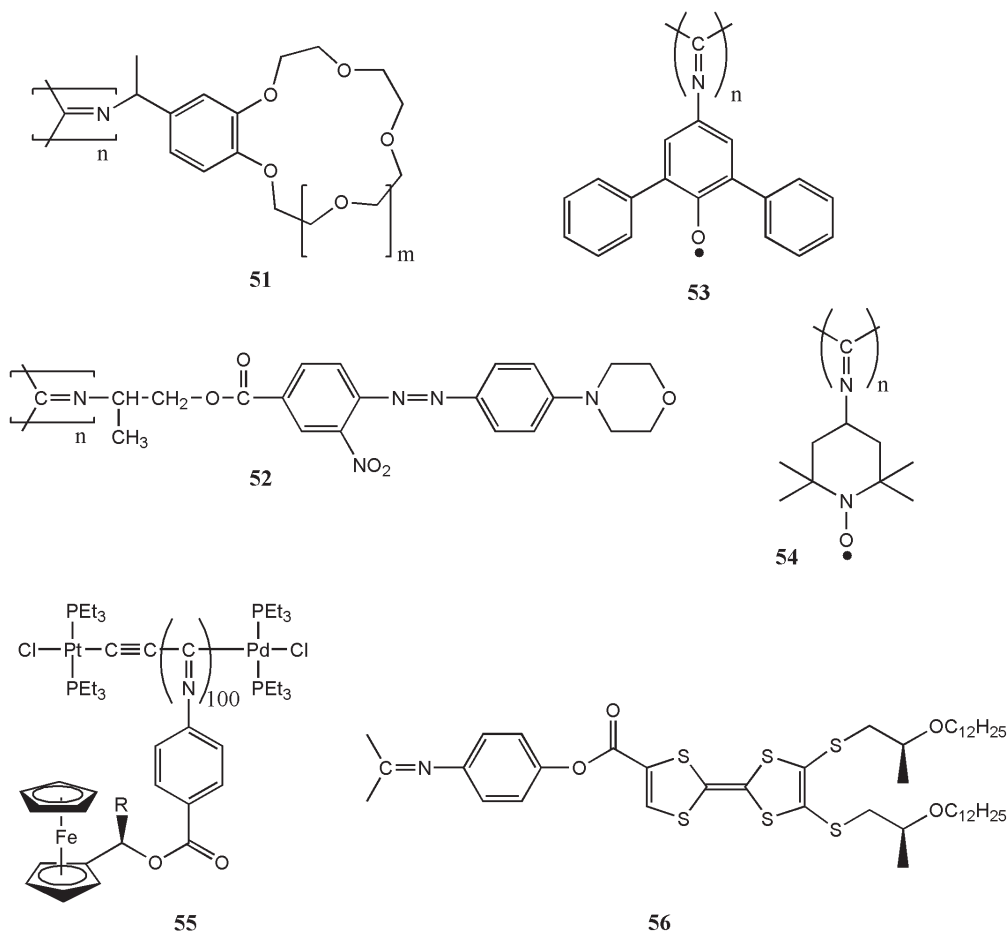
Scheme 12.10

upon oxidation at 1 V, as well as at 1.5 V, the intensity of the positive Cotton effect at 360 nm decreased to 40% of the initial value and that the negative Cotton effect at 250 nm disappeared. The CD signal could be restored by reduction at 0.2 V. The same behavior was observed by chemical oxidation and reduction of the ferrocenyl groups. Upon oxidation, the helical backbone is believed to be transformed into a disordered structure by the electrostatic repulsion of the ferrocenium ions.

In the group of Amabilino, polymer **56** containing tetrathiafulvalene (TTF) units has been designed and synthesized with the objective of constructing a chiral redox polymer [114]. The polymer has three univalent and two wide mixed-valence redox states, which are fully interconvertible (Fig. 12.12). The different redox states of the polymer, in contrast to the monomer, exhibited clearly
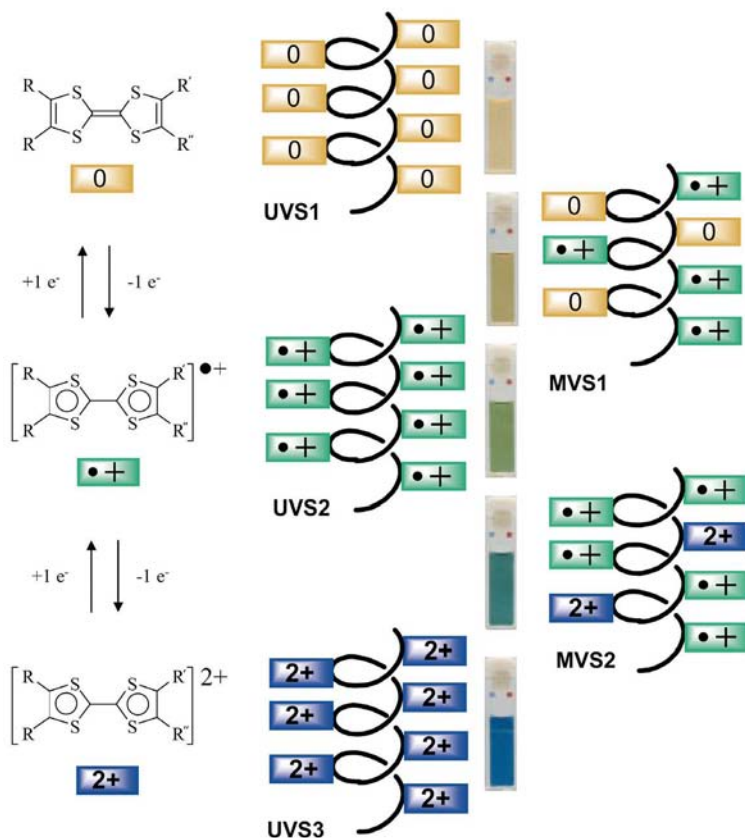
Fig. 12.12 Schematic representation of the three univalent and the two mixed valence redox states of polymer **56** and the colors of the corresponding polymer solutions. (Adapted with permission from Ref. 114 Copyright 2005, Wiley-VCH.)

distinguishable chiroptical properties as was shown by CD spectroscopy measurements and therefore can act as a multistate redox-switchable organic system.

One of the early examples of polyisocyanides bearing electron acceptor and donor groups was reported by Oostveen and Drenth [115, 116]. They synthesized polyisocyanides bearing pyridinium iodide in their side chains. In these polymers the pyridinium functions act as electron acceptors and the iodide ions as electron donors (Fig. 12.13). A large bathochromic shift was observed in the charge transfer spectra of copolymer **57**, which was attributed to intercalation of the iodide ions in the cavities created by the co-monomers leading to strong CT interactions. Similarly, copolymer **58** with short and long side-chains could effectively bind donor molecule **59** leading to electron transfer from the donor to acceptor as was confirmed by EPR measurements. The electron transfer, however, was not ob-
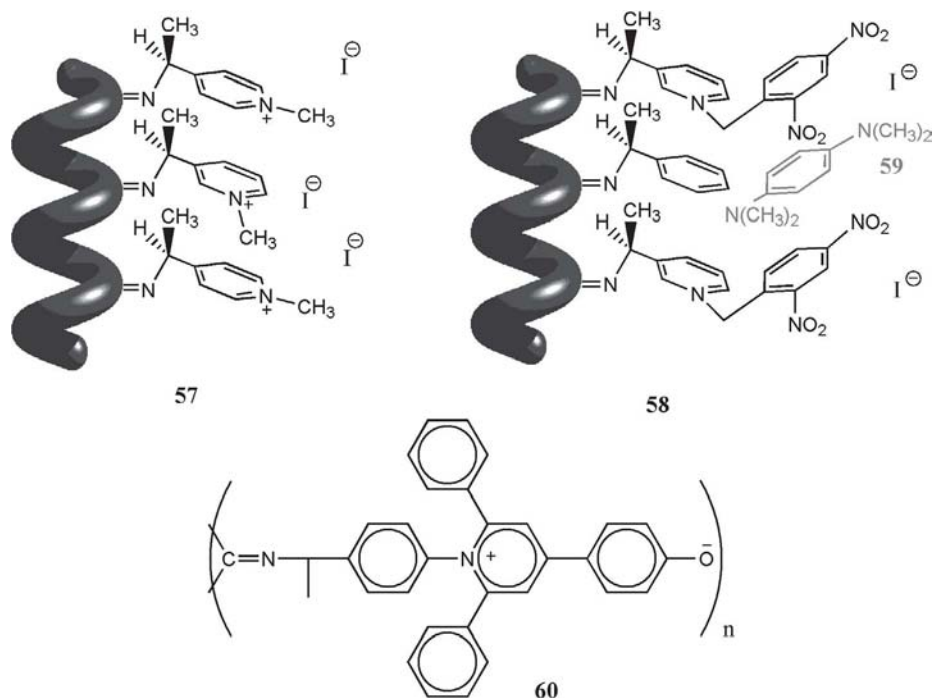
Fig. 12.13 Schematic representation of polyisocyanides bearing acceptor and donor groups, showing the intercalation of the iodide ion and compound **59** in the cavities of the copolymers **57** and **58**. In polymer **60** the acceptor and donor functions are covalently attached to each other.

served between the monomeric, that is, nonpolymerized species. In an attempt to create a polymeric photoconductor, polymer **60** having side groups capable of internal charge transfer, was synthesized [117]. Unfortunately, the polymer was barely soluble and no conductivity measurements could be performed.

A polyisocyanide-based light harvesting system has been constructed by Hong and Fox [118, 119]. Using the nickel catalyst **1**, developed by Deming and Novak, a living polymerization reaction allowed the formation of homo, di- and triblock copolymers incorporating acceptor and donor blocks (Scheme 12.11). From fluorescence spectroscopic studies it was concluded that the stiff polyisocyanide backbone was able to spatially define the chromophores, thereby preventing excimer formation as often observed in more flexible polymers functionalized with chromophores. This is remarkable since the chromophores are connected to the backbone by a relatively flexible ethyl spacer.

Directional singlet energy migration to the acceptor-donor interface was observed in the block copolymers **61** and **62**, resulting in electron transfer at the block interface in the case of **62** and exciplex formation in the case of **61**. Energy

Acceptors   **Donors**

61: R^1=R^2=**A**, R^3=**B**
62: R^1=R^2=**C**, R^3=**B**
63: R^1= **A**, R^2=**D**, R^3=**B**

A

B

C

D

64: R^1=R^2=R^3=H$_2$P
65: R^1=R^2=R^3=ZnP
66: R^1=R^2=H$_2$P, R^3=ZnP
67: R^1=H$_2$P, R^2=ZnP, R^3=H$_2$P

H$_2$P: R^5= H, M=2H
ZnP: R^5= H, M=Zn
H$_2$P-2: R^5= OC$_{12}$H$_{25}$,
                      M=2H
ZnP-2: R^5= OC$_{12}$H$_{25}$,
                      M=Zn

68: R^4= H$_2$P-2
69: R^4= ZnP-2
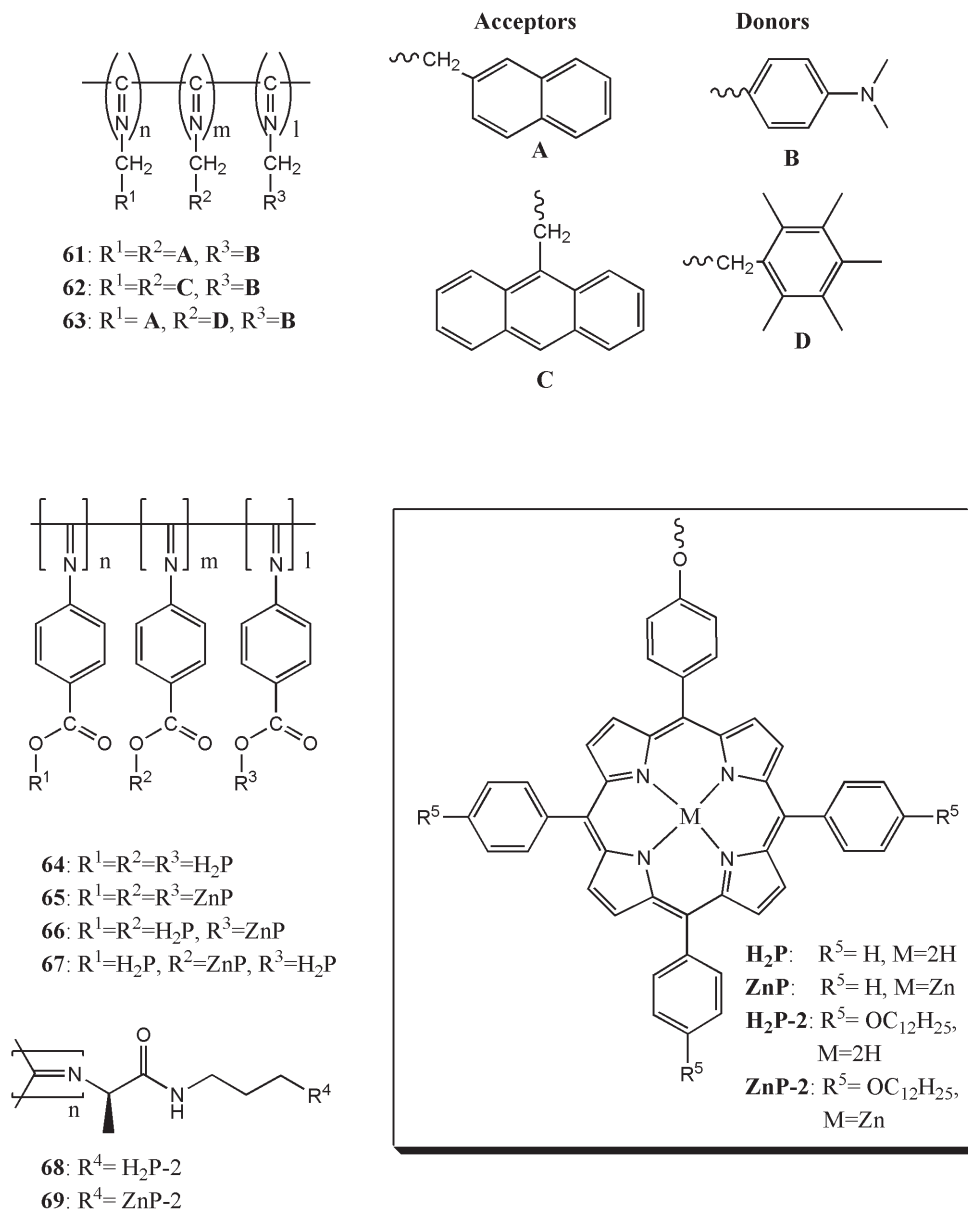
Scheme 12.11

wasting exciplex formation was suppressed in triblock copolymer **63**, which contains an intervening pentamethylphenyl block between the two blocks present in **61**. For block copolymer **62**, transient absorption spectroscopy revealed the formation of a radical ion pair with a lifetime of 1.1 μs.

Extending the field of chromophore functionalized polyisocyanides, the group of Takahashi reported on porphyrin functionalized polyisocyanides **64**–**67** and related compounds, prepared by using Pd–Pt catalyst **5a** [120–123]. The degree of polymerization was varied between 2 and 200 and block copolymers of type **66** with various block length and low polydispersities, mostly below 1.1, were prepared. Photophysical studies revealed that the porphyrins in the stacks are excitonically coupled in a face-to-face manner. It was found that exciton–exciton annihilation rate constants were independent of the length of the polymer indicating a fast exciton migration through the stacks. In the di- and tri-block copolymers **66** and **67**, energy transfer from the zinc to the free-base porphyrins was observed. The rate constants for the excitation energy transfer process appeared to be the same for different block lengths of the free base and zinc porphyrins, again pointing to a fast exciton migration [122].

The achiral porphyrin moieties were also incorporated in different ratios as the middle block in a triblock copolymer, namely between two blocks of chiral isocyanide 11, functionalized with a (L)-menthyl group. The included porphyrins were used as spectator functions to determine the helical sense of the poly(aryl isocyanide)s [124].

Using the well-defined polyisocyanopeptide, as a scaffold, De Witte et al. synthesized helical porphyrin functionalized polyisocyanides **68** and **69** [125]. AFM-measurements revealed that the polymers had an average length of 87 nm, corresponding to an average degree of polymerization of 830. Resonance light scattering measurements demonstrated that at least 25 porphyrins in a stack were interacting with each other and that the slip angle between the porphyrins in a stack was 30°. The presence of a chiral interaction between monomer n and $(n + 4)$ in this slipped conformation and between the neighboring porphyrins in the helix (n and $(n + 1)$) could be observed by CD spectroscopy (Fig. 12.14). Upon addition of the bifunctional ligand dabco, polymer **69**, which has zinc porphyrin side groups, could be switched to a conformation in which the porphyrin stacks
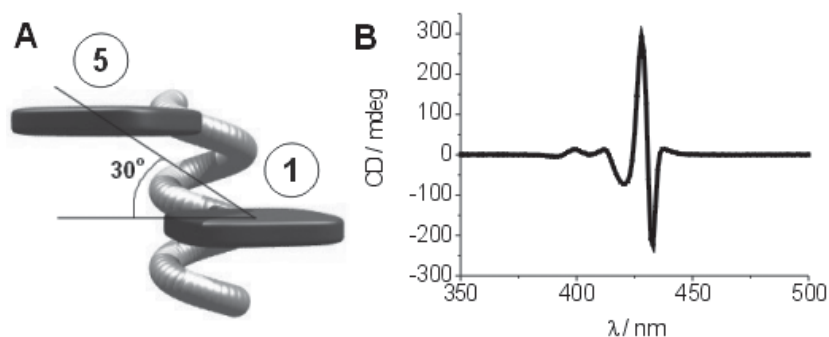
Fig. 12.14 (A) Illustration of the slip angle between porphyrin *n* and $(n + 4)$ in polymer **68** and (B) the CD spectrum of **68** in CHCl$_3$.

possess an opposite helicity, while the helicity of the polymer backbone remains the same [126].

Based on the same polyisocyanopeptide backbone, the groups of Nolte and Rowan prepared thiophene functionalized polyisocyanides **70–72** and perylene functionalized polymers **73** (Scheme 12.12). Thiophene polyisocyanide **70** with only one alanine unit appeared to be less soluble than polymer **71** which contains two alanine units in the backbone [127]. For **70**, only short polymers were observed while longer ones, with lengths up to 300 nm were found for **71**, as revealed by AFM measurements. IR, fluorescence and CD spectroscopic studies showed that the latter polymer contained a better defined backbone structure stabilized by hydrogen bonds, than the former of which the CD spectrum resembled that of the less well-defined polymer **27** derived from alanine-glycine [81]. The prepared polymers might be interesting compounds for the preparation of electron conducting nanowires, for instance, by applying a second topochemical polymerization of the thiophene side groups [128]. In a different approach the monomer from which **70** was prepared, was polymerized using the polystyrene nickel initiator complex **3a** to yield block copolymer **72** [117]. Both in water and organic
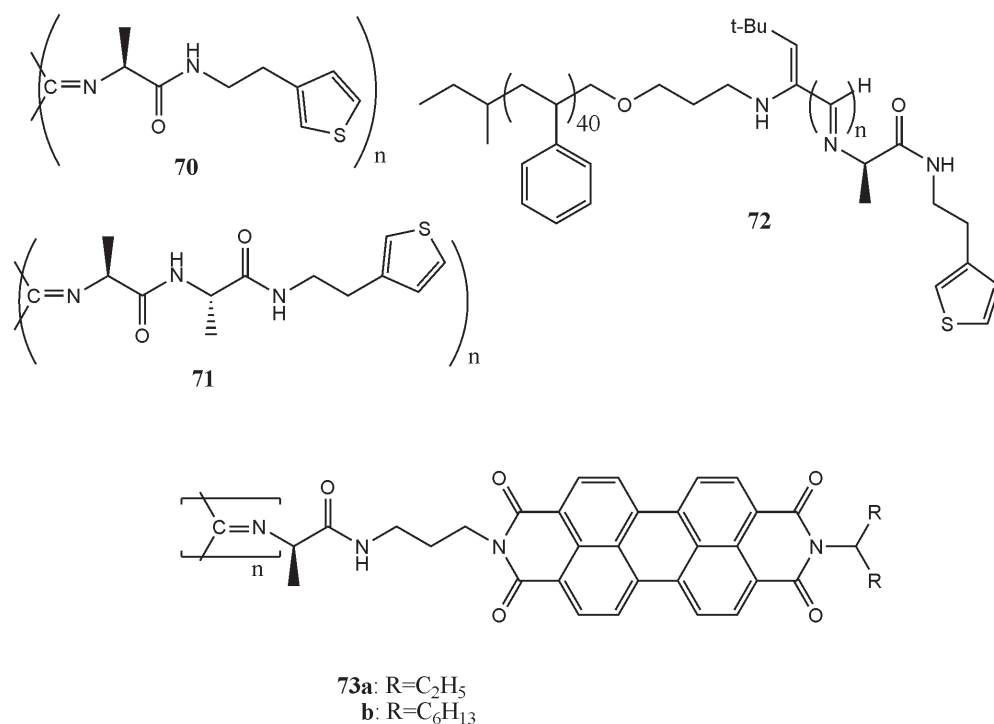
Scheme 12.12

solvents this block copolymer formed polymersomes, which are vesicles derived from amphiphilic polymers. In water, chemical oxidation resulted in cross-linking of the thiophene groups giving rise to electron conducting polymersomes. Catalytically active nanoreactors could be prepared in water by inclusion of enzymes in the interior of the polymersome. The polymersome bilayer proved to be permeable to substrate molecules and the building up of fluorescent product formed by the enzymes in the polymersome interior could be observed by fluorescence microscopy. Giant vesicles with sizes up to 100 μm were prepared from **72**, namely by electroformation [129].

The perylene functionalized polyisocyanides **73** were synthesized as synthetic antenna systems, with possible applications as n-type materials in organic photovoltaics [130, 131]. During polymerization a color change from yellow to red was observed due to the intramolecular stacking of the perylenes, which occurs as a result of the reaction. The polymer fibers were up to 1 μm in length, incorporating *ca.* 10 000 monomer units, as was concluded from AFM measurements (Fig. 12.15). A Cotton effect in the absorption region of the perylene chromophores revealed the helical organization of the perylenes around the helical polymer backbone.

Fluorescence and UV-vis spectroscopic studies on **73a** proved the occurrence of excimer-like species in the close packed perylene arrays. Using a setup combining single-molecule confocal fluorescence and AFM, two species resulting from the polymerization reaction could be distinguished: (i) Ill-defined oligomer species displaying monomer-like fluorescence (green spots Fig. 12.16B) which were too small to be observed by AFM (Fig. 12.16A). The oligomeric character of the species was revealed by their step-wise blinking and bleaching (Fig. 12.16C) and the fact that their fluorescence spectrum was monomer-like (Fig. 12.16E). (ii)
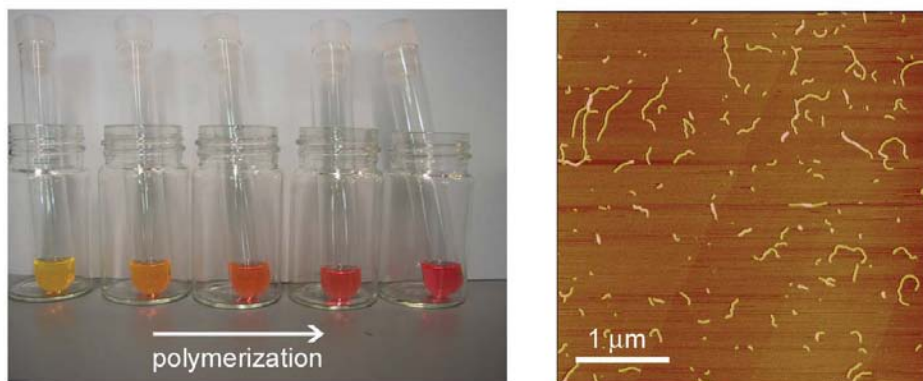
Fig. 12.15 Color change during polymerization of **73b** (left) and AFM picture of the polymer molecules on mica.
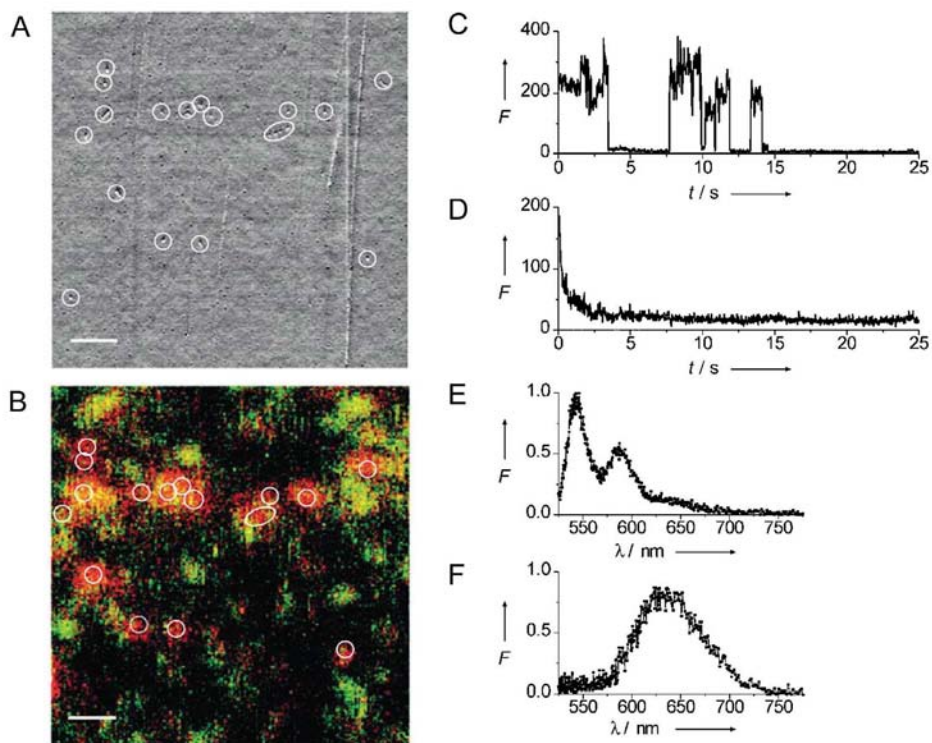
Fig. 12 16 (A) AFM and (B) confocal fluorescence (red: $\lambda > 590$ nm, green $\lambda < 590$ nm) images from the same $3.8 \times 3.8 \ \mu m^2$ area of a diluted solution of polymer **73a** spin-coated on glass (bar = 500 nm; polymers are encircled); (C) Fluorescence intensity trajectory for the green and (D) the red emissions in B; (E) Emission spectra integrated over the whole t = 0–25 s time window for the green emissions; (F) The same for the red emissions. (Adapted with permission from Ref. 130, Copyright 2004, Wiley-VCH.)

Well-defined polymers that could be observed by AFM, showing emission arising from multiple and independent excimer-like sites (red spots Fig. 12.16B). A continuous intensity decrease and an excimer-like emission spectrum, typical for a polymer, were observed for these species (Fig. 12.16D and F). Unfortunately, since polymer **73a** turned out to be only poorly soluble in solvents like toluene and chloroform, application in a device was impossible.

Therefore, polymer **73b** exhibiting similar photophysical properties, but with improved solubility in the aforementioned solvents was synthesized [132]. A photovoltaic cell with an active layer of **73b** as electron acceptor and regioregular polythiophene (**P3HT**) as electron donor was prepared, displaying a 20-fold improved power output as compared to a cell with an active layer of a perylene monomer homolog and **P3HT**.

12.6
Conclusions

It is clear that the field of polyisocyanide chemistry that originated more than 30 years ago is still thriving, primarily due to the unique nature of the polymers, which is only now becoming fully appreciated. It is of interest to note that the initial polymerization catalyst, namely acidic glass discovered by Millich, which was followed by the development of the nickel and then the Pd–Pt catalysts, has now been rediscovered and improved, closing as it were the circle, resulting in polymers with lengths ∼ 15–20 μm long and molecular weights in excess of 20 000 000 Dalton. The helical architecture of the polymers in combination with the functional side-arms, which are precisely positioned along the polymer backbone, results in unique molecules stiffer than DNA, which can act as, for instance, nanowires along which excitons can migrate with ease. It can be foreseen that these polymeric foldamers will continue to raise the interest of chemists and physicists eventually leading to applications as functional materials in optoelectronic and other devices.

References

1 S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.

2 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4012.

3 C. R. Ray, J. S. Moore, *Adv. Polym. Sci.* **2005**, *177*, 91–149.

4 K. Maeda, K. Morino, E. Yashima, *Macromol. Symp.* **2003**, *201*, 135–142.

5 S. Mayer, R. Zentel, *Prog. Polym. Sci.* **2001**, *26*, 1973–2013.

6 M. M. Green, N. C. Peterson, T. Sato, A. Teramoto, R. Cook, S. Lifson, *Science* **1995**, *268*, 1860–1866.

7 M. M. Green, K. S. Cheon, S. Y. Yang, J. W. Park, S. Swansburg, W. H. Liu, *Acc. Chem. Res.* **2001**, *34*, 672–680.

8 J. J. L. M. Cornelissen, A. E. Rowan, R. J. M. Nolte, N. A. J. M. Sommerdijk, *Chem. Rev.* **2001**, *101*, 4039–4070.

9 Y. Okamoto, T. Nakano, *Chem. Rev.* **1994**, *94*, 349–372.

10 R. J. M. Nolte, A. J. M. Van Beijnen, W. Drenth, *J. Am. Chem. Soc.* **1974**, *96*, 5932–5933.

11 R. J. M. Nolte, *Chem. Soc. Rev.* **1994**, *23*, 11–19.

12 R. J. M. Nolte, W. Drenth, *New Methods Polym. Synth.* **1992**, 273–310.

13 M. Yamamoto, K. Onitsuka, S. Takahashi, *Organometallics* **2000**, *19*, 4669–4671.

14 M. Suginome, Y. Ito, *Adv. Polym. Sci.* **2004**, *171*, 77–136.

15 R. J. M. Nolte, R. W. Stephany, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1973**, *92*, 83–91.

16 W. Drenth, R. J. M. Nolte, *Acc. Chem. Res.* **1979**, *12*, 30–35.

17 T. J. Deming, B. M. Novak, *Macromolecules* **1991**, *24*, 326–328.

18 T. J. Deming, B. M. Novak, *Macromolecules* **1991**, *24*, 6043–6045.

19 T. J. Deming, B. M. Novak, *J. Am. Chem. Soc.* **1993**, *115*, 9101–9111.

20 T. J. Deming, B. M. Novak, *Macromolecules* **1993**, *26*, 7092–7094.

21 T. J. Deming, B. M. Novak, *Polym. Prepr.* **1992**, *33*, 1231–1232.

22 T. J. Deming, B. M. Novak, *J. Am. Chem. Soc.* **1992**, *114*, 7926–7927.

23 T. J. Deming, B. M. Novak, *Macromolecules* **1991**, *24*, 5478–5480.

24 T. J. Deming, B. M. Novak, J. W. Ziller, *J. Am. Chem. Soc.* **1994**, *116*, 2366–2374.

25 J. J. L. M. Cornelissen, R. Van Heerbeek, P. C. J. Kamer, J. N. H. Reek, N. A. J. M. Sommerdijk, R. J. M. Nolte, *Adv. Mater.* **2002**, *14*, 489–492.

26 J. Cornelissen, M. Fischer, R. van Waes, R. van Heerbeek, P. C. J. Kamer, J. N. H. Reek, N. Sommerdijk, R. J. M. Nolte, *Polymer* **2004**, *45*, 7417–7430.

27 J. J. L. M. Cornelissen, M. Fischer, N. A. J. M. Sommerdijk, R. J. M. Nolte, *Science* **1998**, *280*, 1427–1430.

28 A. Kros, W. Jesse, G. A. Metselaar, J. Cornelissen, *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 4349–4352.

29 K. Onitsuka, T. Joh, S. Takahashi, *Angew. Chem., Int. Ed. Engl.*, **1992**, *31*, 851–852.

30 K. Onitsuka, K. Yanai, F. Takei, T. Joh, S. Takahashi, *Organometallics* **1994**, *13*, 3862–3867.

31 N. Ohshiro, A. Shimizu, R. Okumura, F. Takei, K. Onitsuka, S. Takahashi, *Chem. Lett.* **2000**, 786–787.

32 K. Onitsuka, K.-I. Yabe, N. Ohshiro, A. Shimizu, R. Okumura, F. Takei, S. Takahashi, *Macromolecules* **2004**, *37*, 8204–8211.

33 F. Millich, G. K. Baker, *Macromolecules* **1969**, *2*, 122–128.

34 F. Millich, R. G. Sinclair, II, *J. Polym. Sci., Pol. Sym.* **1968**, No. 22, 33–43.

35 S. Y. Huang, E. W. Hellmuth, *Polym. Prepr.* **1974**, *15*, 505–508.

36 A. J. M. Van Beijnen, R. J. M. Nolte, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1980**, *99*, 121–123.

37 C. J. M. Huige, A. M. F. Hezemans, R. J. M. Nolte, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1993**, *112*, 33–37.

38 C. J. M. Huige, *Ph.D. Thesis, University of Utrecht, the Netherlands* **1985**.

39 C. Kollmar, R. Hoffmann, *J. Am. Chem. Soc.* **1990**, *112*, 8230–8238.

40 Y. Yamada, T. Kawai, J. Abe, T. Iyoda, *J. Polym. Sci., Part A: Polym. Chem.* **2001**, *40*, 399–408.

41 M. Clericuzio, G. Alagona, C. Ghio, P. Salvadori, *J. Am. Chem. Soc.* **1997**, *119*, 1059–1071.

42 L. Spencer, W. B. Euler, D. D. Traficante, M. Kim, W. Rosen, *Magn. Reson. Chem.* **1998**, *36*, 398–402.

43 J.-T. Huang, J. Sun, W. B. Euler, W. Rosen, *J. Polym. Sci., Part A: Polym. Chem.* **1997**, *35*, 439–446.

44 M. M. Green, R. A. Gross, F. C. Schilling, K. Zero, C. Crosby, III, *Macromolecules* **1988**, *21*, 1839–1846.

45 C. Tanford, *Physical Chemistry of Macromolecules, John Wiley & Sons, New York*, **1961**.

46 F. Millich, *Chem. Rev.* **1972**, *72*, 101–113.

47 F. Millich, E. W. Hellmuth, S. Y. Huang, *J. Polymer Sci., Polymer Chem. Ed.* **1975**, *13*, 2143–2150.

48 A. J. M. Van Beijnen, R. J. M. Nolte, W. Drenth, A. M. F. Hezemans, P. J. F. M. Van de Coolwijk, *Macromolecules* **1980**, *13*, 1386–1391.

49 A. J. M. Van Beijnen, R. J. M. Nolte, A. J. Naaktgeboren, J. W. Zwikker, W. Drenth, A. M. F. Hezemans, *Macromolecules* **1983**, *16*, 1679–1689.

50 A. J. M. Van Beijnen, R. J. M. Nolte, J. W. Zwikker, W. Drenth, *J. Mol. Catal.* **1978**, *4*, 427–432.

51 R. J. M. Nolte, J. A. J. Van Zomeren, J. W. Zwikker, *J. Org. Chem.* **1978**, *43*, 1972–1975.

52 T. Harada, M. C. Cleij, R. J. M. Nolte, A. M. F. Hezemans, W. Drenth, *J. Chem. Soc., Chem. Commun.* **1984**, 726–727.

53 J. M. Van der Eijk, R. J. M. Nolte, V. E. M. Richters, W. Drenth, *Biopolymers* **1980**, *19*, 445–448.

54 M. F. M. Roks, H. G. J. Visser, J. W. Zwikker, A. J. Verkley, R. J. M. Nolte, *J. Am. Chem. Soc.* **1983**, *105*, 4507–4510.

55 H. G. J. Visser, R. J. M. Nolte, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1983**, *102*, 419–420.

56 R. J. M. Nolte, A. J. M. Van Beijnen, J. G. Neevel, J. W. Zwikker, A. J. Verkley, W. Drenth, *Isr. J. Chem.* **1984**, *24*, 297–301.

57 J. G. Neevel, R. J. M. Nolte, *Tetrahedron Lett.* **1984**, *25*, 2263–2266.

58 J. H. Van der Linden, J. Schoonman, R. J. M. Nolte, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1984**, *103*, 260–262.

59 H. G. J. Visser, R. J. M. Nolte, J. W. Zwikker, W. Drenth, *J. Org. Chem.* **1985**, *50*, 3133–3137.

60 F. Takei, K. Onitsuka, S. Takahashi, *Macromolecules* **2005**, *38*, 1513–1516.

61 M. Ishikawa, K. Maeda, E. Yashima, *J. Am. Chem. Soc.* **2002**, *124*, 7448–7458.

62 Y. Hase, M. Ishikawa, R. Muraki, K. Maeda, E. Yashima, *Macromolecules* **2006**, *39*, 6003–6008.

63 M. Ishikawa, K. Maeda, Y. Mitsutsuji, E. Yashima, *J. Am. Chem. Soc.* **2004**, *126*, 732–733.

64 V. Y. Young, E. W. Hellmuth, *Theochem* **2002**, *578*, 1–17.

65 F. Takei, K. Yanai, K. Onitsuka, S. Takahashi, *Chem. Eur. J.* **2000**, *6*, 983–993.

66 F. Takei, K. Yanai, K. Onitsuka, S. Takahashi, *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1554–1556.

67 F. Takei, K. Onitsuka, S. Takahashi, *Polym. J.* **1999**, *31*, 1029–1032.

68 M. M. Green, B. A. Garetz, B. Munoz, H. Chang, S. Hoke, R. G. Cooks, *J. Am. Chem. Soc.* **1995**, *117*, 4181–4182.

69 K. Ute, Y. Fukunishi, S. K. Jha, K. S. Cheon, B. Munoz, K. Hatada, M. M. Green, *Macromolecules* **1999**, *32*, 1304–1307.

70 T. Kajitani, K. Okoshi, S. I. Sakurai, J. Kumaki, E. Yashima, *J. Am. Chem. Soc.* **2006**, *128*, 708–709.

71 T. J. Deming, B. M. Novak, *J. Am. Chem. Soc.* **1992**, *114*, 4400–4402.

72 P. C. J. Kamer, M. C. Cleij, R. J. M. Nolte, T. Harada, A. M. F. Hezemans, W. Drenth, *J. Am. Chem. Soc.* **1988**, *110*, 1581–1587.

73 D. B. Amabilino, E. Ramos, J. L. Serrano, J. Veciana, *Adv. Mater.* **1998**, *10*, 1001–1005.

74 E. Ramos, J. Bosch, J. L. Serrano, T. Sierra, J. Veciana, *J. Am. Chem. Soc.* **1996**, *118*, 4703–4704.

75 D. B. Amabilino, E. Ramos, J.-L. Serrano, T. Sierra, J. Veciana, *J. Am. Chem. Soc.* **1998**, *120*, 9126–9134.

76 D. B. Amabilino, J. L. Serrano, J. Veciana, *Chem. Commun.* **2005**, 322–324.

77 D. B. Amabilino, E. Ramos, J. L. Serrano, T. Sierra, J. Veciana, *Polymer* **2005**, *46*, 1507–1521.

78 A similar approach using a helicene spacer was reported by Chen et al.: J. P. Chen, J. P. Gao, Z. Y. Wang, *Polym. Int.* **1997**, *44*, 83–87.

79 J. J. L. M. Cornelissen, J. J. J. M. Donners, R. de Gelder, W. S. Graswinckel, G. A. Metselaar, A. E. Rowan, N. A. J. M. Sommerdijk, R. J. M. Nolte, *Science* **2001**, *293*, 676–680.

80 J. M. Van der Eijk, R. J. M. Nolte, W. Drenth, A. M. F. Hezemans, *Macromolecules* **1980**, *13*, 1391–1397.

81 J. J. L. M. Cornelissen, W. S. Graswinckel, A. E. Rowan, N. A. J. M. Sommerdijk, R. J. M. Nolte, *J. Polym. Sci., Part A: Polym. Chem.* **2003**, *41*, 1725–1736.

82 G. A. Metselaar, J. Cornelissen, A. E. Rowan, R. J. M. Nolte, *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 1990–1993.

83 C. Branden, J. Tooze, *Introduction to Protein Structure (2nd ed.)*, Garland Publishing, Inc, New York, **1999**, pgs. 84, 288.

84 J. J. L. M. Cornelissen, W. S. Graswinckel, P. J. H. M. Adams, G. H. Nachtegaal, A. P. M. Kentgens, N. A. J. M. Sommerdijk, R. J. M. Nolte, *J. Polym. Sci., Part A: Polym. Chem.* **2001**, *39*, 4255–4264.

85 P. Samori, C. Ecker, I. Goessl, P. A. J. de Witte, J. J. L. M. Cornelissen, G. A. Metselaar, M. B. J. Otten, A. E. Rowan, R. J. M. Nolte, J. P. Rabe, *Macromolecules* **2002**, *35*, 5290–5294.

86 W. Zhuang, C. Ecker, G. A. Metselaar, A. E. Rowan, R. J. M. Nolte, P. Samori, J. P. Rabe, *Macromolecules* **2005**, *38*, 473–480.

87 J. J. L. M. Cornelissen, N. A. J. M. Sommerdijk, R. J. M. Nolte, *Macromol. Chem. Phys.* **2002**, *203*, 1625–1630.

88 J. J. L. M. Cornelissen, *Pure Appl. Chem.* **2002**, *74*, 2021–2030.

89 J. J. J. M. Donners, R. J. M. Nolte, N. A. J. M. Sommerdijk, *J. Am. Chem. Soc.* **2002**, *124*, 9700–9701.

90 M. B. J. Otten, C. Ecker, G. A. Metselaar, A. E. Rowan, R. J. M.

Nolte, P. Samori, J. P. Rabe, *ChemPhysChem* **2004**, *5*, 128–130.

91 S. J. Wezenberg, G. A. Metselaar, A. E. Rowan, J. J. L. M. Cornelissen, D. Seebach, R. J. M. Nolte, *Chem. Eur. J.* **2006**, *12*, 2778–2786.

92 T. Hasegawa, S. Kondoh, K. Matsuura, K. Kobayashi, *Macromolecules* **1999**, *32*, 6595–6603.

93 T. Hasegawa, K. Matsuura, K. Ariga, K. Kobayashi, *Macromolecules* **2000**, *33*, 2772–2775.

94 C. A. Van Walree, J. F. Van der Pol, J. W. Zwikker, *Recl. Trav. Chim. Pays-Bas* **1990**, *109*, 561–565.

95 B. Grassl, S. Rempp, J. C. Galin, *Macromol. Chem. Phys.* **1998**, *199*, 239–246.

96 J. Van Esch, M. F. M. Roks, R. J. M. Nolte, *J. Am. Chem. Soc.* **1986**, *108*, 6093–6094.

97 A. Biegle, A. Mathis, J. C. Galin, *Macromol. Chem. Phys.* **2000**, *201*, 113–125.

98 Y. Q. Tian, Y. Li, T. Iyoda, *J. Polym. Sci., Part A: Polym. Chem.* **2003**, *41*, 1871–1880.

99 Y. Q. Tian, K. Kamata, H. Yoshida, T. Iyoda, *Chem. Eur. J.* **2005**, *12*, 584–591.

100 A. J. Naaktgeboren, R. J. M. Nolte, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1978**, *97*, 112–115.

101 A. J. Naaktgeboren, R. J. M. Nolte, W. Drenth, *J. Am. Chem. Soc.* **1980**, *102*, 3350–3354.

102 A. J. Naaktgeboren, R. J. M. Nolte, W. Drenth, *J. Mol. Catal.* **1981**, *11*, 343–351.

103 J. J. J. M. Donners, E. W. Meijer, R. J. M. Nolte, N. A. J. M. Sommerdijk, *Polym. Mater. Sci. Eng.* **2001**, *84*, 1039–1040.

104 P. Samori, J. J. J. M. Donners, N. Severin, M. B. J. Otten, J. P. Rabe, R. J. M. Nolte, N. A. J. M. Sommerdijk, *Langmuir* **2004**, *20*, 8955–8957.

105 K. Maeda, M. Ishikawa, E. Yashima, *J. Am. Chem. Soc.* **2004**, *126*, 15161–15166.

106 M. F. M. Roks, R. J. M. Nolte, *Macromolecules* **1992**, *25*, 5398–5407.

107 M. Kauranen, T. Verbiest, C. Boutton, M. N. Teerenstra, K. Clays, A. J. Schouten, R. J. M. Nolte, A. Persoons, *Science* **1995**, *270*, 966–969.

108 M. Kauranen, T. Verbiest, E. W. Meijer, E. E. Havinga, M. N. Teerenstra, A. J. Schouten, R. J. M. Nolte, A. Persoons, *Adv. Mater.* **1995**, *7*, 641–644.

109 M. N. Teerenstra, J. G. Hagting, G. T. Oostergetel, A. J. Schouten, M. A. C. Devillers, R. J. M. Nolte, *Thin Solid Films* **1994**, *248*, 110–114.

110 M. N. Teerenstra, J. G. Hagting, A. J. Schouten, R. J. M. Nolte, M. Kauranen, T. Verbiest, A. Persoons, *Macromolecules* **1996**, *29*, 4876–4879.

111 M. N. Teerenstra, R. D. Klap, M. J. Bijl, A. J. Schouten, R. J. M. Nolte, T. Verbiest, A. Persoons, *Macromolecules* **1996**, *29*, 4871–4875.

112 M. Abdelkader, W. Drenth, E. W. Meijer, *Chem. Mater.* **1991**, *3*, 598–602.

113 N. Hida, F. Takei, K. Onitsuka, K. Shiga, S. Asaoka, T. Iyoda, S. Takahashi, *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 4349–4352.

114 E. Gomar-Nadal, J. Veciana, C. Rovira, D. B. Amabilino, *Adv. Mater.* **2005**, *17*, 2095–2098.

115 E. A. Oostveen, W. Drenth, *Proc. IUPAC, I. U. P. A. C., Macromol. Symp., 28th* **1982**, 451.

116 E. A. Oostveen, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1983**, *102*, 35–41.

117 A. J. M. Van Beijnen, R. J. M. Nolte, W. Drenth, *Recl. Trav. Chim. Pays-Bas* **1986**, *105*, 255–261.

118 B. Hong, M. A. Fox, *Can. J. Chem.* **1995**, *73*, 2101–2110.

119 B. Hong, M. A. Fox, *Macromolecules* **1994**, *27*, 5311–5317.

120 F. Takei, K. Onitsuka, N. Kobayashi, S. Takahashi, *Chem. Lett.* **2000**, 914–915.

121 F. Takei, S. Nakamura, K. Onitsuka, A. Ishida, S. Tojo, T. Majima, S. Takahashi, *Chem. Lett.* **2003**, *32*, 506–507.

122 M. Fujitsuka, A. Okada, S. Tojo, F. Takei, K. Onitsuka, S. Takahashi, T. Majima, *J. Phys. Chem. B* **2004**, *108*, 11935–11941.

123 F. Takei, D. Kodama, S. Nakamura, K. Onitsuka, S. Takahashi, *J. Polym. Sci., Part A: Polym. Chem.* **2005**, *44*, 585–595.

124 F. Takei, H. Hayashi, K. Onitsuka, N. Kobayashi, S. Takahashi, *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 4092–4094.

125 P. A. J. de Witte, M. Castriciano, J. J. L. M. Cornelissen, L. M. Scolaro, R. J. M. Nolte, A. E. Rowan, *Chem. Eur. J.* **2003**, *9*, 1775–1781.

126 P. A. J. de Witte, *Ph.D. Thesis, Radboud University Nijmegen, the Netherlands* **2004**.

127 D. M. Vriezema, *Ph.D. Thesis, Radboud University Nijmegen, the Netherlands* **2005**.

128 D. M. Vriezema, J. Hoogboom, K. Velonia, K. Takazawa, P. C. M. Christianen, J. C. Maan, A. E. Rowan, R. J. M. Nolte, *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 772–776.

129 D. M. Vriezema, A. Kros, R. de Gelder, J. Cornelissen, A. E. Rowan, R. J. M. Nolte, *Macromolecules* **2004**, *37*, 4736–4739.

130 J. Hernando, P. A. J. de Witte, E. M. H. van Dijk, J. Korterik, R. J. M. Nolte, A. E. Rowan, M. F. Garcia-Parajo, N. F. van Hulst, *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 4045–4049.

131 P. A. J. De Witte, J. Hernando, E. E. Neuteboom, E. M. H. P. Van Dijk, S. C. J. Meskers, R. A. J. Janssen, N. F. Van Hulst, R. J. M. Nolte, M. F. Garcia-Parajo, A. E. Rowan, *J. Phys. Chem. B* **2006**, *110*, 7803–7812.

132 M. B. J. Otten, E. Schwartz, P. A. J. de Witte, J. J. L. M. Cornelissen, M. M. Wienk, R. A. J. Janssen, R. J. M. Nolte, A. E. Rowan, *PMSE Preprints* **2006**, *94*, 743–744.

13
Foldamers at Interfaces

Jan van Esch, Hennie Valkenier, Sebastian Hartwig,
and Stefan Hecht

13.1
Introduction

A foldamer [1] can be defined as a (macro)molecular strand that is capable of adopting a well-defined, thermodynamically favored conformation in solution [2]. In order to provide a broader description that can be extended to interfaces as well, we view the foldamer's conformational preference as being the result of the action of internal and external boundary conditions, which involve various types of interactions (Fig. 13.1).

In solution in the isolated single foldamer molecule, only internal boundary conditions in the form of intramolecular interactions are involved in stabilizing a specific folded chain conformation. Both covalent internal constraints, such as linkage geometry, rotational preferences etc. (see Chapter 1), and noncovalent intramolecular interactions, such as H-bonding, metal–ligand coordination, π,π-stacking, coulombic, dipolar, and van der Waals interactions, etc. (see Chapter 2), govern the overall conformational preference. In addition, the foldamer's interaction with its surrounding has to be considered giving rise to external boundary conditions. Obviously, solvation that involves solvophilic and solvophobic interactions, such as the hydrophobic effect, as well as intramolecular nanophase separation (see Chapter 3) play an additional important role. Furthermore, *endo*- or *exo*-complexation of guest molecules (see Chapter 7) as well as dimerization and further aggregation (see Chapter 4) can give rise to folding. No further detailed discussion of the conformational preference in solution is needed here and the reader is referred to Part 1 of this book.

At interfaces, however, external boundary conditions in the form of interfacial interactions become crucial and often dictate molecular conformation. Several enthalpic and entropic parameters lead to certain conformational preferences that are often mingled with self-assembly processes, i.e. the external boundary conditions favor certain adsorption conformations and concurrently direct self-assembly (Fig. 13.1). This interplay leads to formation of hierarchically organized
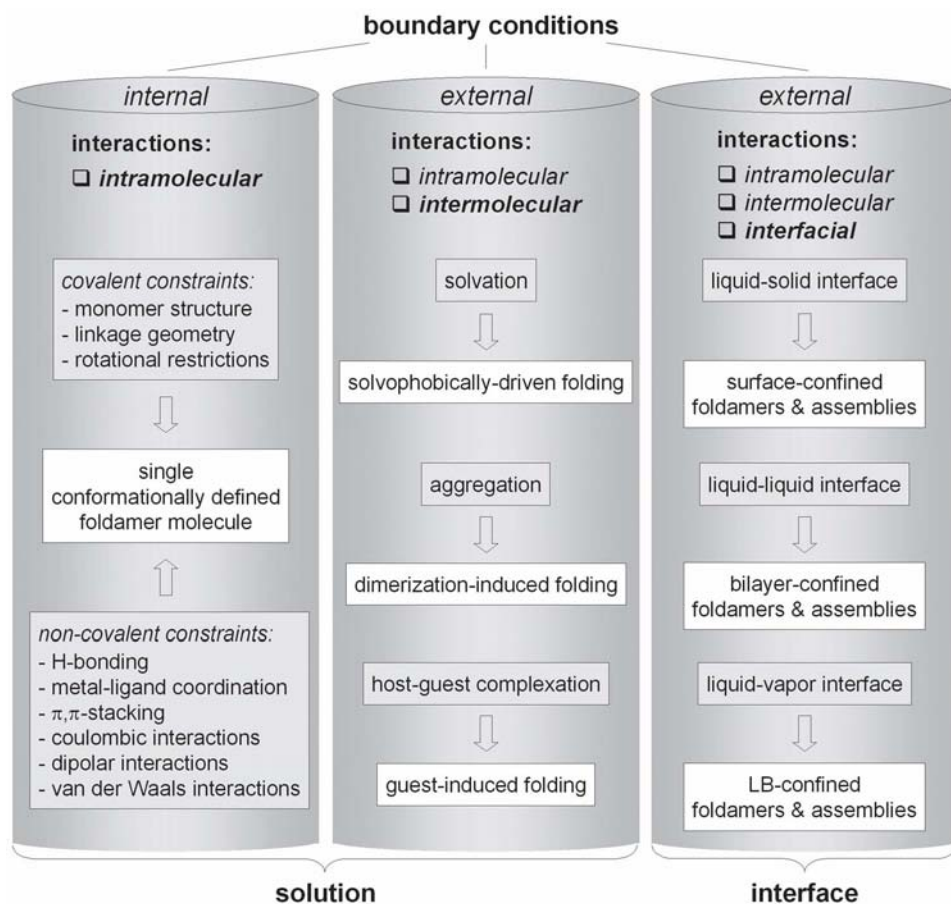
Fig. 13.1 Overview of internal and external boundary conditions influencing the conformational preferences of foldamers and their assemblies.

materials that not only mimic structural evolution in Nature, for example biomineralization, but also hold exciting prospects for various new applications, involving for instance supramolecular electronics.

Here, we want to focus on foldamers at interfaces due to the high scientific interest and technological relevance of this field of research. While the question of how additional interactions at the interface interfere with the folding process and formation of the final secondary structure or higher order assemblies is of fundamental interest, the implication for technology including interfacial processes in recognition, sensors, patterning, catalysis among others is most significant. In addition, a profound knowledge of the underlying principles is key to understand-

ing biological phenomena, such as protein denaturation at interfaces and bio-mineralization. In the first part of this review we will propose a general description of folding at interfaces while in the second part we will highlight and discuss important examples from the literature that have been arranged with respect to the involved secondary structure motifs with particular focus on solid–liquid interfaces.

13.2
Folding in Solution and at Interfaces

In solution, foldamers adopt preferred secondary structures that almost exclusively consist of various helix or sheet types. In biomacromolecules, such as proteins, these secondary structural motifs can further be organized into higher order structures using covalent loops, i.e. tertiary structures, and even further using noncovalent interactions between subunits, i.e. quaternary structures. At interfaces, however, both the conformational preference of the foldamer and the intermolecular interactions between individual foldamer strands might be largely altered (Fig. 13.2). Several examples from the literature (see below) show such behavior.

Solution
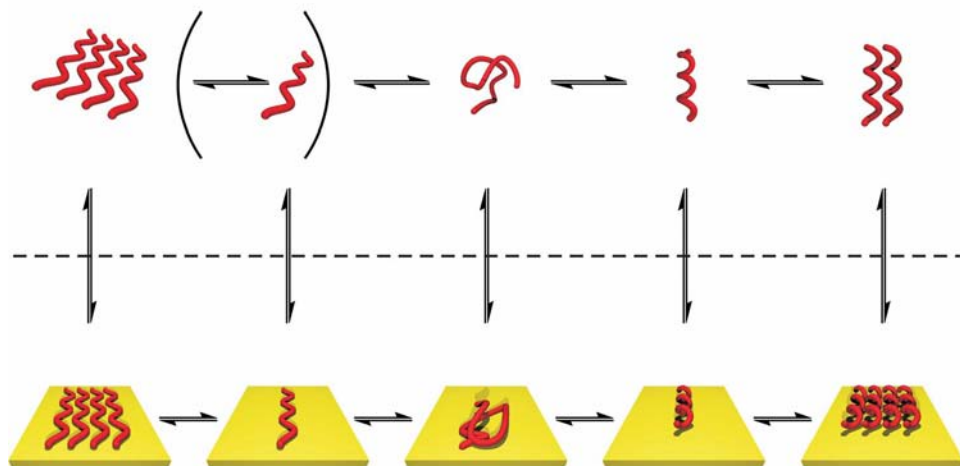
Interface

Fig. 13.2 Overview of the most abundant foldamer conformations in solution and at interfaces and their relationships.

13.2.1
Types of Interactions

In solution, mostly intramolecular interactions as listed in Fig. 13.1 dominate folding. Among the most important conformational equilibria is the coil–helix transition (Fig. 13.2), where intrachain contacts stabilize the helical conformation. Above a critical concentration, intermolecular interactions become important and aggregation can also affect secondary structure formation (see Chapter 4), as illustrated in the formation of β-sheets held together by interchain contacts (Fig. 13.2).

At the interface, we consider that additional molecule–surface interactions, which contribute to the overall conformational preference, become extremely important. The strength of the interaction with the surface determines whether the adsorption conformation is fixed in a local minimum or if it is equilibrating to adopt the global minimum structure, in analogy to the concept of kinetic and thermodynamic control, respectively. If the magnitude of the interaction is rather large and the molecule is strongly bound to the surface, the system is referred to as being chemisorbed, while in the case of a weak interaction, the system is considered as being physisorbed. The interfacial interactions can be specific, involving directional contacts between functional groups, as well as nonspecific, arising from surface free energy minimization. In addition, it is important to note that the intermolecular interactions are likely to become more dominant due to the much higher local concentration within the confinement of the interface.

13.2.2
Thermodynamics

When confining a foldamer strand to a specific conformation, in general the enthalpy gain due to attractive noncovalent interactions has to compensate for the entropy loss associated with the molecules' reduced conformational freedom. In solution, the most important contributions arise from the gain in enthalpy of the intramolecular contacts (ΔH_F) and the loss of conformational entropy ($\Delta S_{conf(F)}$) as shown in eq. 1:

$$\Delta G \sim \Delta H_F - T\Delta S_{conf} \tag{1}$$

We anticipate that the situation at the interface is seemingly more complex involving several enthalpic and entropic terms (eqs. 2 and 3):

$$\Delta H = \Delta H_F + \Delta H_{F@F} + \Delta H_{F@I} + \Delta H_{desolv(F)} + \Delta H_{desolv(I)} \tag{2}$$

$$\Delta S = \Delta S_{conf(F)} + \Delta S_{trans(F)} + \Delta S_{rot(F)} + \Delta S_{desolv(F)} + \Delta S_{desolv(I)} \tag{3}$$

Important additional enthalpic components (eq. 2) are associated with intermolecular association ($\Delta H_{F@F}$), interfacial foldamer–surface interactions ($\Delta H_{F@I}$) as well as desolvation of both the foldamer molecule ($\Delta H_{desolv(F)}$) and the interface ($\Delta H_{desolv(I)}$). Furthermore, additional entropic factors (eq. 3) include the transla-

tional and rotational entropy of the foldamer chain ($\Delta S_{\text{trans(F)}}$ and $\Delta S_{\text{rot(F)}}$) as well as the entropy associated with desolvation of the foldamer ($\Delta S_{\text{desolv(F)}}$) and the interface ($\Delta S_{\text{desolv(I)}}$), respectively. To simplify the thermodynamic treatment, we assume the interfacial foldamer–surface interactions ($\Delta H_{\text{F@I}}$) to be the dominating additional factor. The consequence is illustrated with the aid of a simplified model [3], comparing the folding of a homosequence foldamer with isoenergetic interactions between repeat units into either helix or sheet conformation in solution and at the interface (Fig. 13.3).
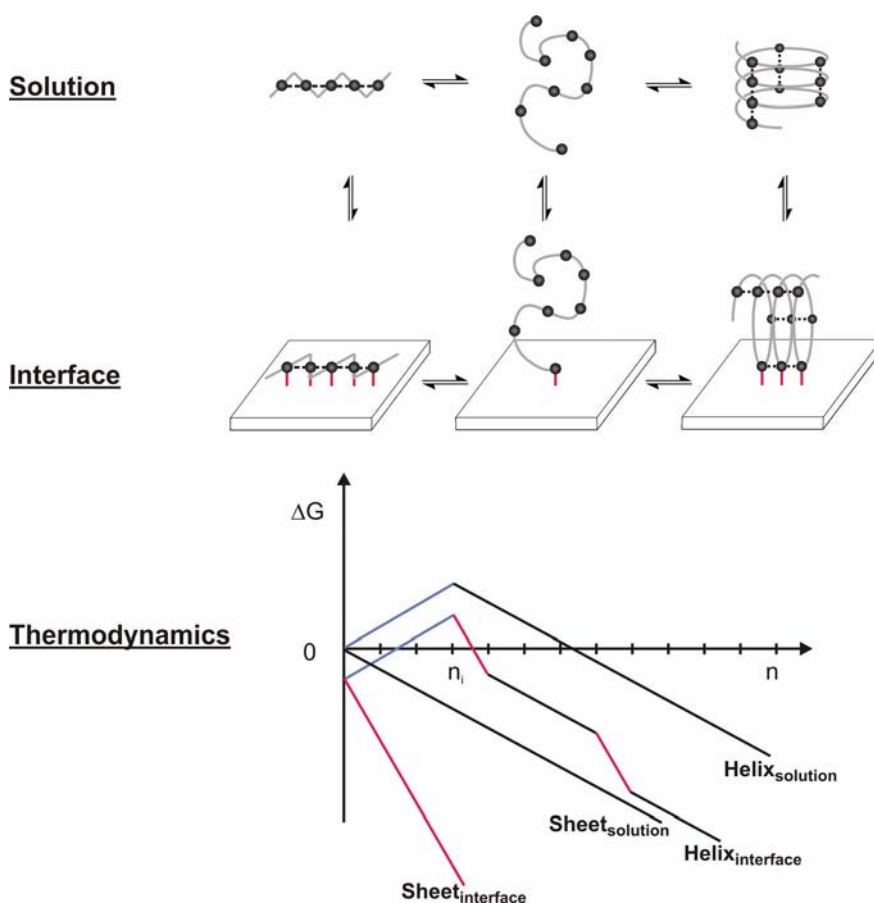
Fig. 13.3 Thermodynamics of folding into helix vs. sheet conformations in solution and at interfaces. Dominant intramolecular (black dotted) and interfacial (red solid) enthalpic interactions give rise to different thermodynamic behavior (bottom), where the most important energetic contributions (bottom) arise from enthalpy gain due to intramolecular (ΔH_F shown in black) and intermolecular ($\Delta H_{\text{F@I}}$ shown in red) interactions and entropy loss due to conformational confinement ($\Delta S_{\text{conf(F)}}$ shown in blue).

In solution, folding into the helix conformation is associated with a nucleation event, which leads to preorganization of the first attractive intrastrand interaction between non-neighboring repeat units and therefore helix folding is cooperative. Folding into a sheet at the single molecule level is however not cooperative since the intrastrand interactions involve neighboring repeat units. It is important to realize that the folding into individual sheets is barely observed due to concurrent aggregation that is mediated by attractive interchain contacts and proceeds in a cooperative fashion due to preorganization.

It is expected that at the interface, the interfacial interactions dominate and as a consequence both helix and sheet conformation are stabilized, however, to a different degree. In our simplified model (Fig. 13.3), the helix can only use every third repeat unit to engage in attractive molecule–surface interactions, while the sheet is able to utilize every single residue. While this particular model is certainly largely simplistic, it illustrates some important differences when comparing folding at the interface to folding in solution, most importantly:

1. The nucleation barrier is lowered.
2. The critical chain length is decreased.
3. A smaller number of helix repeat units is favored.
4. Usually sheet formation is more favored due to the adsorption geometry since usually fewer residues can interact with the surface in the case of the helix as compared to the sheet, i.e. flat vs. curved adsorbate.

Another important aspect is related to the reversibility of the molecule–surface interactions, which is a necessary prerequisite for equilibration and hence defect healing and fidelity of pattern recognition. The "entropic distraction" involved in the recognition of a patterned surface by an oligomer/polymer strand displaying interacting groups complementary to the surface pattern has nicely been illustrated by Muthukumar (Fig. 13.4) [4]. Loop entropy is the reason that the path to the global minimum, i.e. the complex based on correct pattern recognition, does proceed via intermediate stages, which are not the most stable of their kind and hence kinetically less accessible. As a result, the foldamer can readily get "distracted", i.e. enter the wrong reaction funnel, and therefore it is essential that all elementary steps are reversible to adopt the lowest energy adsorption conformation.

13.2.3
Design Considerations

When designing a foldamer sequence to adopt a specific conformation at the interface, "positive" and "negative" design approaches, leading to rational stabilization and destabilization of local geometries by introduction of attractive and repulsive interactions, have to be combined in order to realize the "minimally frustrated" state [5]. One first important aspect concerns the use of hetero-sequences to bias conformational preferences in foldamers. For example, from
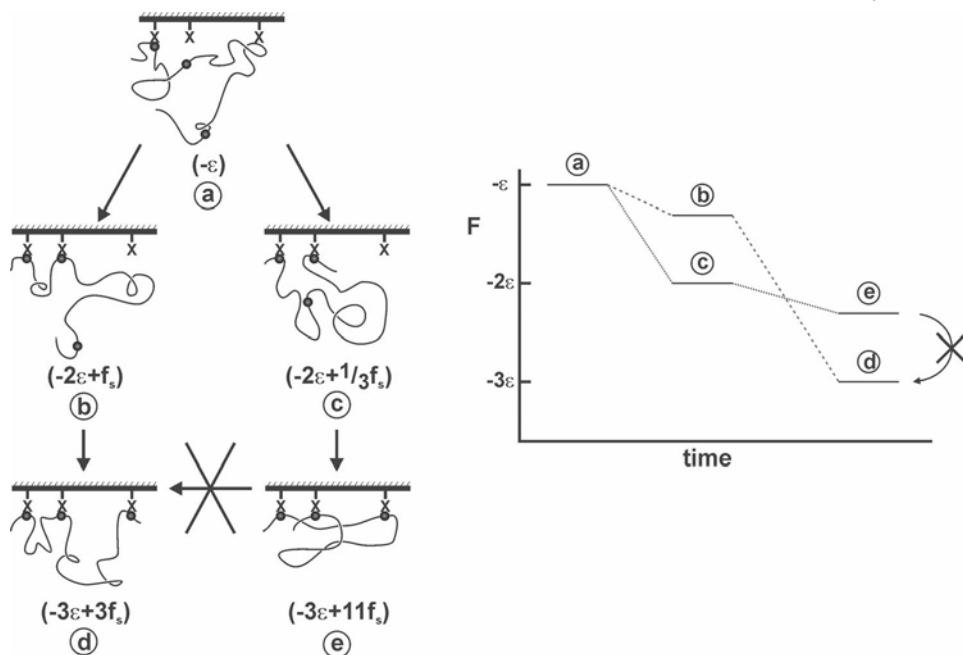
Fig. 13.4 Pathway selection in pattern recognition of foldamers on surfaces. The Muthukumar Model (adapted from ref. [4]) assumes that the foldamer displays three groups, spaced *m* units apart, which recognize the surface functionalities X, spaced by distance *b*. In the case of recognition, the enthalpy of interaction is ε while the loop entropy is $f_s = 3k_B Tb2/2m$. While the intermediate topological state (c) is more stable than (b), it does not lead to the global minimum topological state (d) as shown on the right.

Fig. 13.3 it becomes apparent that in order to favor the helical conformation at the interface, heterosequences are required. A second important strategy avoids complications or local defects due to chain entropy by reducing the number of entropically favored loops. Furthermore, the strength of the interactions, both within and between strands and most importantly between foldamer and surface has to be balanced to assure for defect healing yet conformational stability.

13.2.4
Scope

Secondary structure formation at interfaces is not limited to a specific class of foldamers, any more than their behavior in solution (see Part 1 of this book). The conformational behavior of peptides at interfaces has extensively been studied for several decades, especially in monolayers at the air–water interface, and more recently in self-assembled monolayers (SAMs) of peptides at solid–liquid inter-

faces [6–11]. Remarkably, only a few studies deal with other types of foldamers. In this chapter, we will therefore limit ourselves to representative, relevant examples of peptide-based systems, and focus on secondary structure formation in monolayers of nonpeptide foldamers.

13.3
Helical Structures

13.3.1
Adsorption of Helical Structures at Interfaces

A very abundant secondary structure that can be found in all types of systems is the helical structure. DNA consists of a double-stranded helical structure, many proteins have helical domains, and polymers can form helices too. These specific secondary structures can show different behavior when adsorbed at interfaces, as will be presented in this section. The helix can remain as a helix, but other more or less defined secondary structures can be formed as well in the presence of interfacial forces (see 13.3.2).

An important, early study on the formation of helical structures was performed by De Grado and Lear [12]. They have synthesized peptides consisting of hydrophobic leucine and hydrophilic lysine residues **1**–**3** (Fig. 13.5). Not surprisingly, only the 14-residue peptide **2** with a hydrophobic repeat unit periodicity [13] of 3.5 formed α-helical structures in solution as observed by circular dichroism (CD) spectroscopy. However, the equilibrium between helical and random coil conformations showed a strong dependence on both peptide and NaCl concentration in aqueous solution. The same peptides were studied at interfaces. Surface pressure–area isotherms of the assembly at the air–water interface suggest the predominance of the α-helical conformation. The monolayer was transferred from the air–water interface to quartz slides by the Langmuir–Blodgett (LB) method and the layer was studied by CD and infrared (IR) spectroscopy. The results of these studies are consistent with the formation of a monolayer of α-helical

	Primary Structure	Periodicity	Secondary Structure	
			Solution	Interface
1	Fmoc(Leu-Lys-Lys-Leu-Leu-Lys-Leu)$_1$	3.5	Random coil	Not determined
2	Fmoc(Leu-Lys-Leu-Leu-Leu-Lys-Leu)$_2$	3.5	α-Helix	α-Helix
3	Fmoc(Leu-Lys-Leu-Lys-Leu-Lys-Leu)$_1$	2.0	β-Sheet	β-Sheet

Fig. 13.5 Peptides **1**–**3** with different periodicities, i.e. hydrophobic repeat unit distances, and their conformational preference in solution and at interfaces [12].
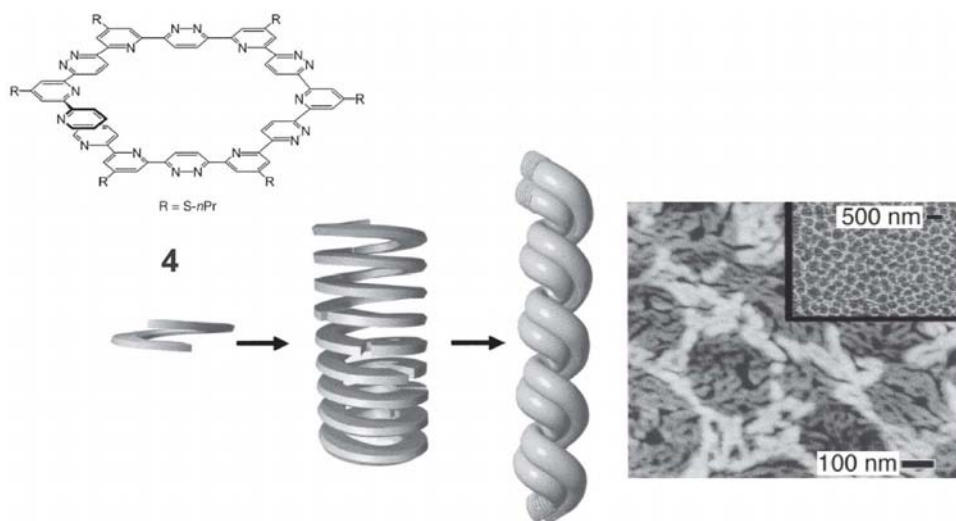
Fig. 13.6 Schematic picture of the self-assembly of the pyridine-pyridazine foldamer **4** into filaments and fibrils (left) suggested by AFM images of drop-cast films of **4** on mica (right) [14]. (Reproduced in part from ref. [14] with permission.)

peptides and support the conclusion that folding into the α-helical conformation at interfaces is supported by the appropriate hydrophobic–hydrophilic periodicity.

The formation of helices at interfaces is not limited to peptides. Helical structures are also formed by the self-assembly of foldamer **4** consisting of 8 pyridine and 5 pyridazine units (Fig. 13.6) [14]. This compound folds into crescent disc-like structures, which stack to form filaments and fibrils. LB films of these compounds have been studied and from the surface pressure–area isotherm it was concluded that the filaments adsorb at the air–water interface in an edge-on arrangement, i.e. with the heteroaromatic repeat units perpendicular to the interface. Atomic force microscopy (AFM) studies on drop-cast films of these compounds on freshly cleaved mica revealed the presence of worm-like structures (Fig. 13.6). These studies show that the self-assembled superhelical structure of these foldamers is well preserved in solution as well as in monolayers at interfaces.

In the examples mentioned above, the respective foldamers form helices in solution as well as at the interface, and there is negligible influence of the interactions between the foldamer and the interface on secondary structure formation. However, this is not always the case. Lu et al. have presented a study on two 15-residue peptides **5** and **6** (Fig. 13.7), where the three tyrosine residues in peptide **5** have been substituted by three tryptophan residues in peptide **6** [15]. In solution both peptides form α-helices. The adsorption of these peptides at the hydrophilic silicon oxide/water interface depends strongly on both pH and concentra-

Primary Structure	
5	**Tyr**-Val-Asn-Ala-Lys-Gln-Tyr-**Tyr**-Arg-Ile-Leu-Lys-Arg-Arg-**Tyr**
6	**Trp**-Val-Asn-Ala-Lys-Gln-Tyr-**Trp**-Arg-Ile-Leu-Lys-Arg-Arg-**Trp**

Fig. 13.7 Peptides **5** and **6** differing in three repeat units (shown in bold) show different adsorption conformations on silicon oxide depending on pH and concentration [15].

tion. In an acidic environment at pH = 5, peptide 5 is highly charged and hydrophilic and does not adsorb on the slightly negatively charged silicon oxide surface. Increasing the pH to pH = 7–9 leads to the formation of stable peptide layers at concentrations as low as 0.01 w/v%. Most interestingly, at these low concentrations the tryptophan-modified peptide **6** forms a loosely packed layer consisting of the peptide in the β-sheet conformation, but at higher concentrations (0.1 w/v%) the peptide adopts the α-helical conformation. Most likely, this structural transition is caused by interchain contacts occurring at higher surface densities.

In addition, the monolayer preparation method and the employed solvent can have a pronounced effect on interfacial conformation. For instance, the backbone of a poly(phenylacetylene) derivative carrying L-valine pendant groups (see Chapter 11 for a more detailed account on related work by the Yashima group, i.e. Figures 11.4, 11.8, and 11.18) was shown by CD and UV-vis spectroscopy to adopt a helical conformation in solution [16]. Films prepared by slow evaporation of a methanol solution resulted in the formation of globular micelles at freshly cleaved mica, whereas slow evaporation of THF solutions resulted in the formation of helical cables. At the air–water interface, on the other hand, the polymer self-assembles into extended fibers as concluded from structural investigations on LB films [17]. This contrasting behavior can be explained by solvent influences. In polar solvents, the polymer prefers to adopt a helical structure since the hydrophobic backbone is directed towards the inside of the helix whereas the more polar valine units are positioned at the outside, thereby increasing the solubility and stability of the helices. However, at the air–water interface, the polymer can adopt a nonhelical conformation in which the valine pendants are located in the water layer and the backbone of the polymer is exposed to the air at the interface.

13.3.2
Loss of Helicity upon Adsorption

The examples provided above show that helical structures can adsorb at interfaces, while preserving (or even increasing) their helical content. However, many foldamers adopt a helical conformation in solution, which is (partially) lost when adsorbed at interfaces. Burkett and Read have synthesized a series of peptides, consisting of anionic aspartate, uncharged alanine, and cationic arginine segments. They investigated the helical conformational behavior in solution and at anionic and cationic colloidal silica substrates by CD and ^{1}H-NMR spectroscopy

[18]. It appeared that the peptide helices showed partial unwinding upon adsorption at both substrates to maximize the number of complementary interactions between charged side chains of the peptide and the nearby opposite surface charges. An in-depth study of one of these peptides provided some remarkable insights [19]. The helices were oriented to the substrate with the complementary charged block. However, it was not necessarily the block oriented to the substrate that showed loss of helicity. Adsorption at both anionic and cationic colloids gave rise to loss of helicity in the arginine segment, while the alanine segment showed only partial helicity loss and the aspartate segment retained its helical conformation completely. In addition, heating of the helix-containing solution gave rise to loss of helicity mainly at the arginine terminus and to a lower extent at the aspartate terminus, whereas no loss of helicity was observed in the alanine segment. It should be noted that in these peptides, the strong dipolar charge distribution of the side chains is complementary to the backbone dipole, thereby stabilizing the helix. Apparently, charge compensation of one of the segments disturbs the intermolecular helix-forming interactions in such a way that the transition to the random coil structure will start and propagate from the less stable terminus. As indicated by the solution phase temperature study, the arginine segment represents the less stable terminus potentially due to a high nucleation barrier, which is further increased upon adsorption.

Another example of helical peptides that have lost helicity upon adsorption at interfaces is reported by Vankann, Höcker, and coworkers [20]. These peptides were built from a leucine-based hydrophobic segment and a hydrophilic head group composed of polar amino acids, e.g. **7** (Fig. 13.8), or oligo(ethylene oxide). The conformational preferences of the peptides were studied by CD spectroscopy both in solution and embedded in liposome lipid bilayers. In the lipid bilayer, the longer peptides retained their helicity to 'fit' in the layer, whereas the shorter peptides unfolded to form β-sheet structures. The peptides were also studied at the air–water interface with a Langmuir balance. The secondary structure of the adsorbed peptides was governed by the size of the head groups. The

7 Fmoc-Leu$_8$-Lys$_5$-Tyr

α-Helix   β-Sheet (extended)
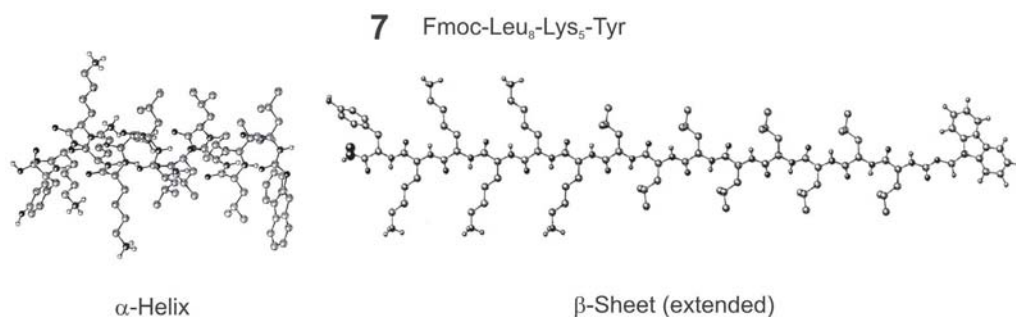
Fig. 13.8 Amphiphilic peptide **7** adopts an α-helical conformation (left) when spread at the air–water interface; however, when the monolayer is compressed the peptide adopts an extended β-sheet conformation (right) [20]. (Reproduced in part from ref. [20] with permission.)

peptides with bulky head groups were assumed to assemble as α-helices perpendicular to the surface, whereas the less dense packing of peptides without bulky head groups suggests a change of conformation into a β-sheet. Compression of the monolayers of **7** led to a conformational change from α-helix to β-sheet (Fig. 13.8), which could be inhibited by the use of the bulky oligo(ethylene oxide) groups. This example shows that the secondary structure of peptides can be governed by the way in which the folded structure 'matches' its surrounding, more than it is determined by intramolecular interactions.

13.3.3
Helical Structures Formed upon Adsorption

Adsorption of unfolded structures at interfaces can also induce the formation of helical conformations. Hydrophobins are among the best biosurfactants and
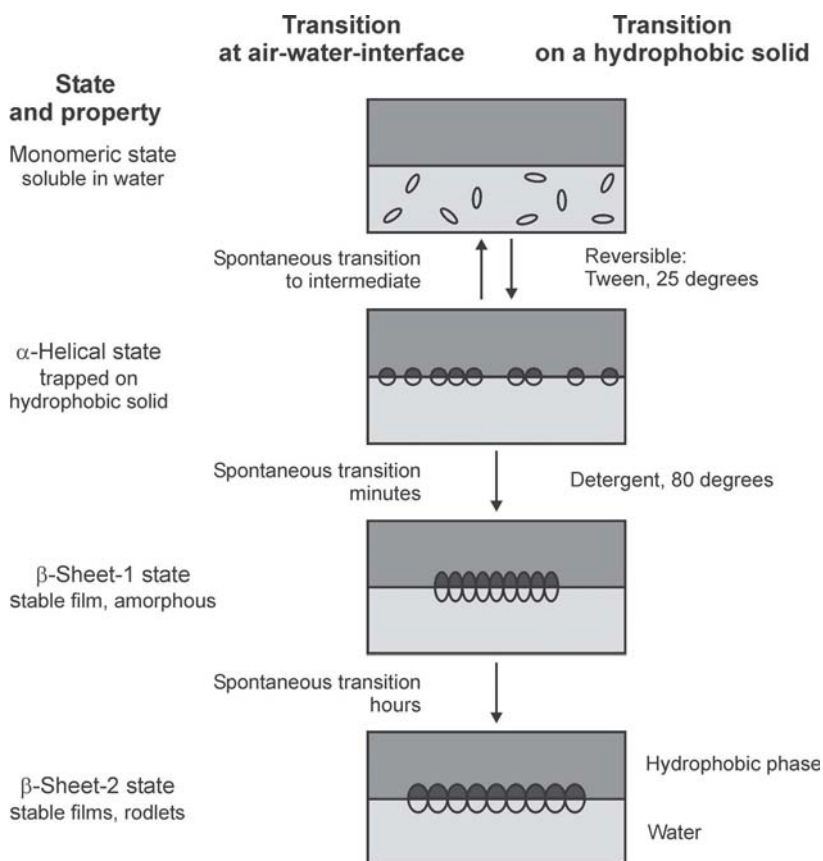
Fig. 13.9 Self-assembly and conformational equilibria of hydrophobin SC3 in different environments and at different interfaces (Tween = detergent poly(ethylene glycol) sorbitan monolaurate with n ∼ 20) [22]. (Adapted from ref. [22].)

show interesting behavior at interfaces. The hydrophobin SC3 is a fungal protein composed of 136 amino acid residues [21]. In solution, these proteins adopt no specific secondary structure. However, when adsorbed at a hydrophobic Teflon surface, an α-helical structure has been observed (Fig. 13.9) [22]. It should be noted that this α-helical structure is not the most stable global minimum state but can be transformed into a β-sheet structure by heating at 80 °C in the presence of a detergent (Tween). These structural transitions of hydrophobins are also observed at the air–water interface. Adsorption of SC3 first leads to an intermediate α-helical state, which is spontaneously converted into an amorphous β-sheet state. This state subsequently reorganizes into a β-sheet state with a rodlet structure, which could be visualized by AFM after transferring the monolayer from the air–water interface onto a carbon film. Although it is clear that these structural transitions of hydrophobins are caused by adsorption at hydrophobic interfaces, the molecular background of these transitions remains to be uncovered.

In a related context, a recent modeling study suggests that DNA naturation, i.e. the formation of the double-stranded helix, can also be stimulated by adsorption at interfaces [23]. This modeling study shows that after adsorption of single stranded DNA at a surface the entropy loss created by dimerization of these two strands will be much smaller than in solution, thereby decreasing the free energy for dimerization. Hence, the dimerization process that leads to helix formation is stimulated by the interface and the nucleation barrier is lowered.

13.4
Sheet Structures

13.4.1
Adsorbed Sheet Structures at Interfaces

Another important class of secondary structures is the sheet motif. β-Sheet structures are abundant peptide and protein secondary structures, and similar secondary sheet structures are also found in a variety of artificial backbones (see Chapter 4). In most sheet structures, the individual foldamer chains are stretched and intermolecular interactions between different strands lead to the formation of two-dimensional ribbons or sheets via dimerization and further aggregation.

A few examples of peptides that adsorb as β-sheets at the air–water interface have been described by Rapaport and coworkers [24–26]. They designed peptides of alternating hydrophobic and hydrophilic amino acids, which are able to assemble at the air–water interface. It was found that proline residues at the N and C termini induce two-dimensional order in the monolayer because prolines are preferentially located at the rim of a ribbon. Other amino acids do not display this preference (Fig. 13.10) [24].

In another study, peptides **8** and **9**, incorporating alternating hydrophobic and hydrophilic amino acids as well as complementary electrostatic interactions between lysine and glutamic acid residues, were designed to favor parallel β-sheets at the air–water interface (Fig. 13.11) [26]. Grazing incidence X-ray diffraction, IR
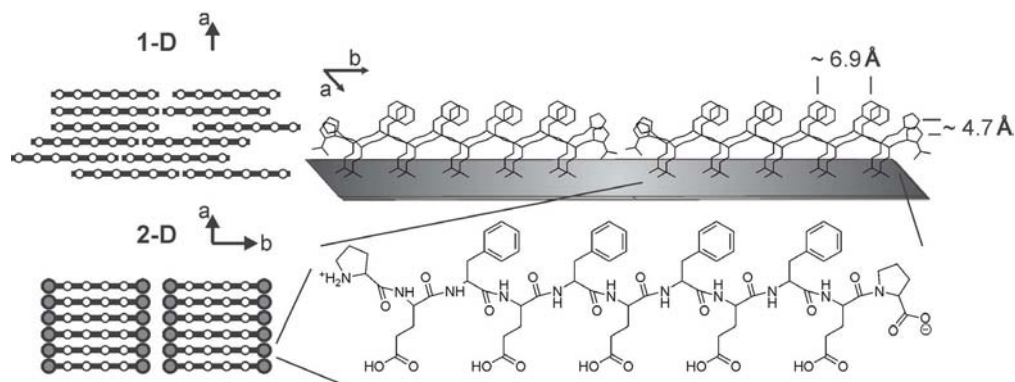
Fig. 13.10 Schematic representation of the self-assembly of β-sheet structures at the air–water interface from amphipathic peptides carrying proline termini (right), resulting in one-dimensional or two-dimensional order (left) [24]. (Adapted from ref. [24].)
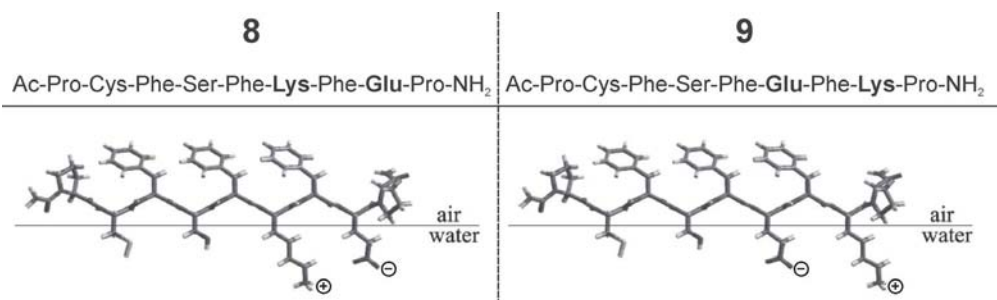
Fig. 13.11 Self-assembly of an equimolar mixture of peptides **8** and **9** having complementary electrostatic interactions (residues shown in bold) to parallel β-sheet structures at the air–water interface [26]. (Reproduced in part from ref. [26] with permission.)

spectroscopy, and mass spectrometry studies on films composed of an equimolar mixture of both peptides confirmed the formation of parallel β-sheets, whereas this is not the case in films solely composed of either one of the peptides alone. This emphasizes the importance of intermolecular electrostatic interactions between hydrophilic amino acid residues in the formation of β-sheets. Unfortunately, in these studies no comparison has been provided between the secondary structures of the peptides at the air–water interface and in solution.

Another example of β-sheet structures at interfaces can be found in the work of Ree, Magonov, and coworkers [27]. They prepared monolayers of short chain poly(benzyloxycarbonyl lysine) **10** on highly ordered pyrolytic graphite (HOPG)
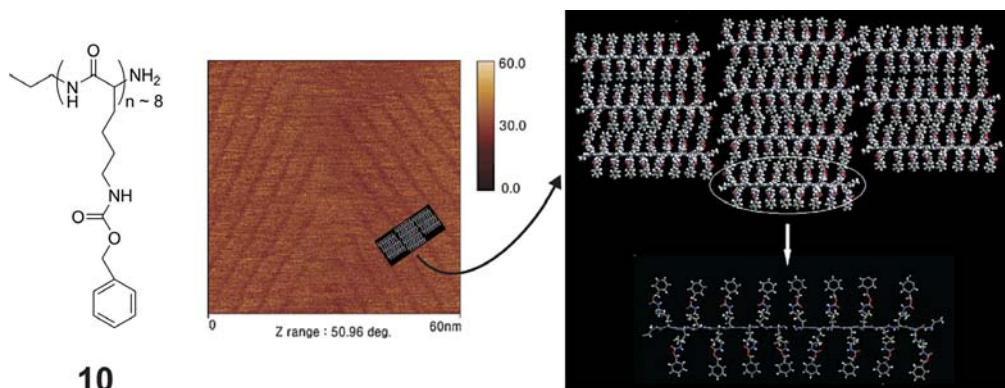
Fig. 13.12 Self-assembled monolayers of short chain polypeptide **10** (left) on HOPG. AFM image of the monolayer formed upon spin-coating (center) and structural analysis (right) [27]. (Reproduced in part from ref. [27] with permission.)

by a solution spin-coating process (Fig. 13.12). The monolayers were characterized by AFM and IR spectroscopy. From inspection of the amide carbonyl stretching vibration in films it was concluded that **10** adopts a β-sheet conformation. However, the analysis of the AFM images suggests that the peptides lie flat on the HOPG surface to maximize favorable contacts between the aromatic side chains and the graphite substrate (Fig. 13.12). Intermolecular π–π interactions between neighboring aromatic side chains rather than hydrogen bonds between amide units lead to maximum surface coverage by subsequent formation of ribbons, which propagate along the direction perpendicular to the long axis of the peptide.

13.4.2
Enhanced Sheet Formation upon Adsorption

Other examples of sheet formation are known in which the comparison between structures in bulk solution and at interfaces is made. In Section 13.3.1 the landmark study of DeGrado and Lear on α-helix formation of leucine- and lysine-consisting peptides **1–2**, both in solution and at hydrophobic–hydrophilic interfaces was discussed [12]. In this paper the authors also describe peptide 3 with a hydrophobic periodicity of 2, i.e. an alternating sequence of leucine and lysine residues. It appeared that this peptide adopts a β-sheet conformation upon adsorption at the air–water interface, whereas α-helices were formed from peptides with a hydrophobic periodicity of 3.5 (Fig. 13.13). As in the case of the α-helices, in solution there is a dynamic equilibrium between β-sheet and random coil conformation. However, at the interface the equilibrium is completely shifted to the β-sheet conformation. This nicely demonstrates the stabilization of the β-sheet structure by interfacial interactions.

| Periodicity ~ 3.5 | vs. | Periodicity ~ 2.0 |

| Hydrophobic Periodicity (Primary Structure) | + | Apolar Surface (Interface) | ⟹ | Induced Secondary Structure | | Hydrophobic Periodicity (Primary Structure) | + | Apolar Surface (Interface) | ⟹ | Induced Secondary Structure |
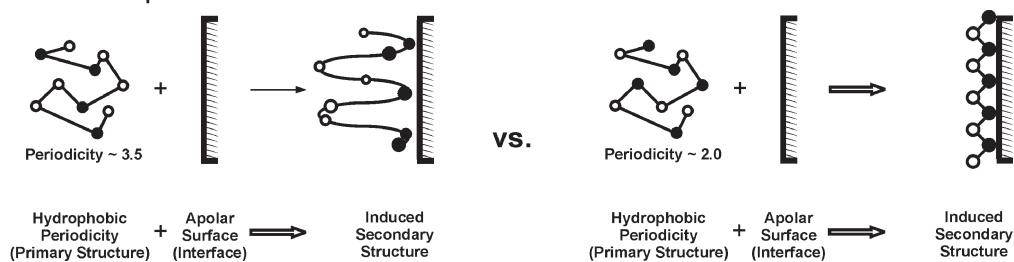
Fig. 13.13 Schematic representation of the adsorption of an amphiphilic peptide with a periodicity of 3.5 into an α-helical structure (left) and of an amphiphilic peptide with a periodicity of 2.0 into a β-sheet structure (right) (Adapted from ref. [12]) [12].

Comparable results were found by Tirrell and coworkers for long peptides consisting of three pairs of alternating alanine and glycine units combined with single glutamic acid residues as turn elements, i.e. [(Ala-Gly)$_3$-Glu-Gly]$_{36}$ [28]. The authors studied the peptides by IR absorption spectroscopy in aqueous solution and by external IR reflectance spectroscopy at the air–water interface. In solution, the peptide adopts the random coil conformation over a wide pH range ($5 \leq pH \leq 14$). However, at the air–water interface the peptide adopts a β-sheet conformation that could further be attenuated when the pH of the solution was changed from basic (pH = 10) to acidic (pH = 5). From this study it can be concluded that sheet formation was stimulated by adsorption.
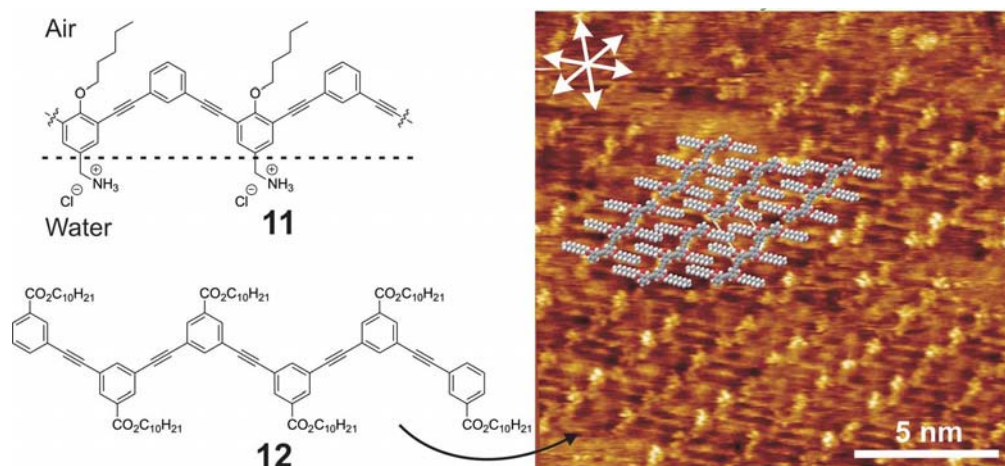
Fig. 13.14 Schematic representation of amphiphilic oligo(*meta*-phenylene ethynylene) **11** adopting a sheet structure at the air-water interface (left) [30, 31] and STM image of *n*-decyl-substituted hexa(*meta*-phenylene ethynylene) **12** on HOPG adopting a zig-zag structure at the solid–liquid interface (right) [32].

In analogy to the aforementioned amphiphilic peptides, other amphiphilic oligomers can form sheet structures at interfaces as well. For instance, amphiphilic oligo(*meta*-phenylene ethynylene) foldamers pioneered by the Moore group (see Chapter 3) adopt a helical conformation in solution in order to maximize solvophilic interactions of the side chains with the solvent and due to π,π-stacking interactions between the aromatic backbone repeat units [29]. Tew and coworkers found that this foldamer family, if appropriately substituted with alternating hydrophobic and hydrophilic groups, i.e. **11**, forms sheet-like structures at the air–water interface when the adsorbed monolayers are compressed (Fig. 13.14 left) [30, 31]. From surface–pressure isotherms and calculations it was concluded that polymer **11** adopted a zig-zag-type *transoid* conformation, with the aromatic rings perpendicular to the interface, i.e. in an edge-on conformation, thereby presumably stabilizing the sheets by π,π-stacking. In related work at the solid–
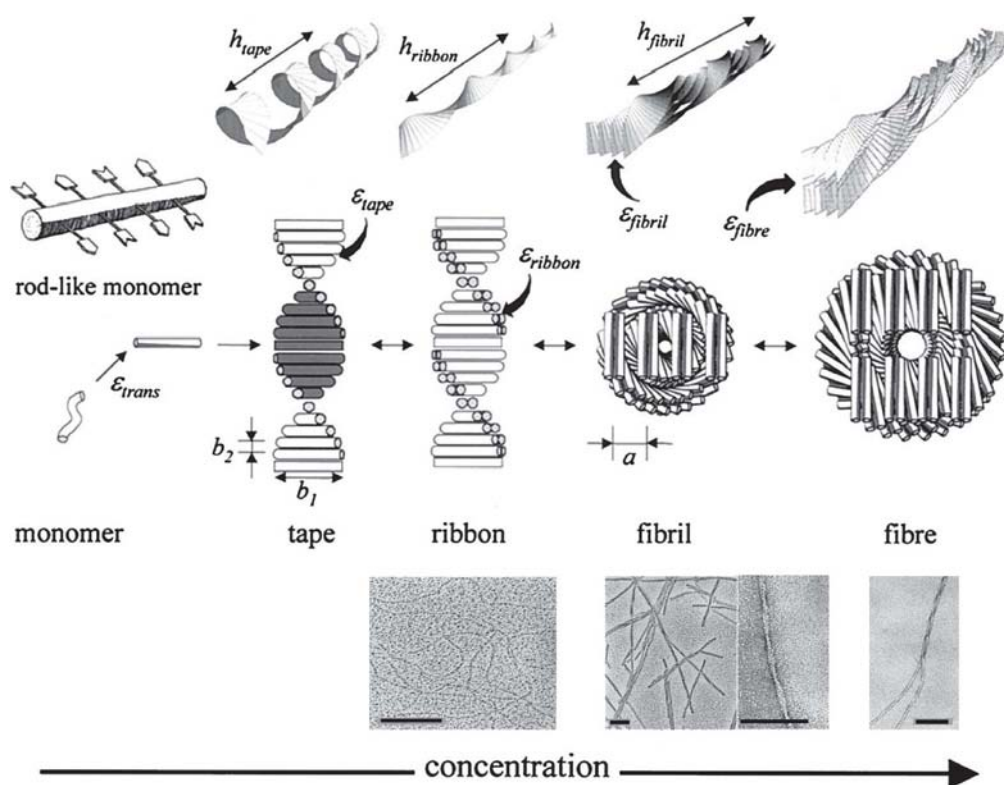
Fig. 13.15 Model for the hierarchical self-assembly of peptide monomer P_{11}-2, i.e. Ac-Gln-Gln-Arg-Phe-Gln-Trp-Gln-Phe-Glu-Gln-Gln-NH_2, into tapes, ribbons, fibrils and fibers as a function of concentration and corresponding electron micrographs (scale bar 100 nm in each image) [33]. (Reproduced in part from ref. [33] with permission.)

liquid interface, Hecht, Rabe, and coworkers studied SAMs of alkyl-substituted hexa(*meta*-phenylene ethynylene) **12** by scanning tunneling microscopy (STM) and could show that the foldamers adopt face-on sheet structures on HOPG (Fig. 13.14 right) [32]. Interestingly, the corresponding *ortho*-linked hexamers show a markedly different behavior by preferentially forming lower dimensional aggregates.

13.4.3
Change in Sheet Structure upon Adsorption

In the group of Boden, the self-assembly of a peptide called P_{11}-2, i.e. Ac-Gln-Gln-Arg-Phe-Gln-Trp-Gln-Phe-Glu-Gln-Gln-NH$_2$, has been studied both in solution and adsorbed on mica substrates [33–35]. In solution, at very low concentrations (<0.05 mM) the peptide was present as random coil, as confirmed by CD spectroscopy. At slightly higher concentrations (∼0.07 mM) β-sheet ribbons start to emerge and at even higher concentrations (>0.6 mM) rigid fibrils and fibers have been observed by electron microscopy (Fig. 13.15) [33].

The observed formation of helical structures from β-sheet assemblies of peptides is generally believed to be based on the chirality of the single peptide building blocks. However, the helical twisting can be overcome by adsorption on mica substrates when the peptide–surface binding energy outweighs the energy gained by twisting [35]. At specific conditions (5 µM P_{11}-2 in 10 vol% water in *n*-propanol and pH = 5.5) tapes resembling antiparallel cross-β-sheet structures were found to grow, as observed by AFM on mica (Fig. 13.16). The formed tapes had a height
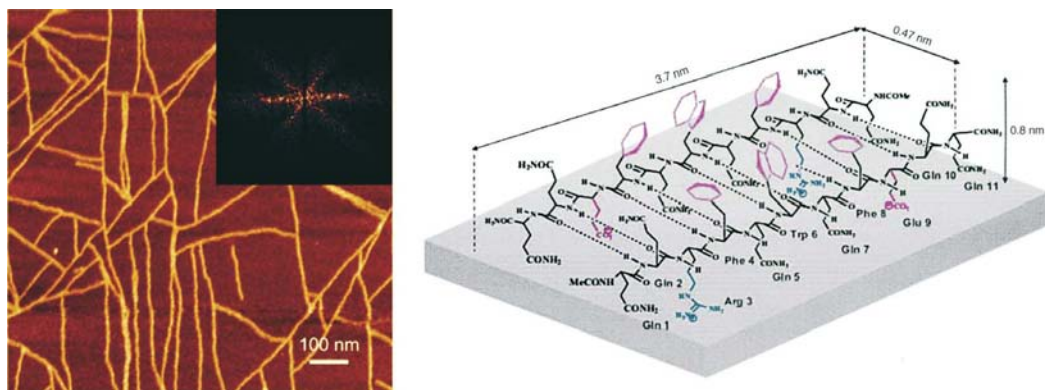
Fig. 13.16 Hierarchical self-assembly of peptide monomer P_{11}-2, i.e. Ac-Gln-Gln-Arg-Phe-Gln-Trp-Gln-Phe-Glu-Gln-Gln-NH$_2$, into tapes as shown by AFM images on mica (left, inset shows FT and hexagonal symmetry) and structural model derived from XPS data and modeling (right) [35]. (Reproduced from ref. [35] with permission.)

corresponding to the height of one molecule and a width corresponding to the length of one molecule. Therefore, the peptides are most likely bound via the replacement of potassium ions of the upper layer of the mica substrate by guanidinium groups of arginine residues while the glutamates' carboxylate groups are expected to bind the potassium ions. Obviously, in this example electrostatic interactions are strong enough to drive the adsorption of flat β-sheets on mica substrates. Furthermore, it is important to note that the critical concentration for the growth of these tapes at a mica surface is more than ten times lower than the critical concentration required for their self-assembly in solution.

13.5
Turn Elements and Hairpins

The last class of secondary structures that will be discussed here are turn elements, of which hairpins will receive the major attention. Van Esch, De Feyter, and coworkers have studied the folding and adsorption of small molecules **13**–**15** that mimic turn elements (Fig. 13.16) [36]. These turn mimics consist of a catechol unit and two alkyl chains, both containing an amide group for additional stabilization of the folded conformation by intramolecular hydrogen bonding. Molecular modeling predicted that folding depends strongly on the length of the spacers between the catechol moiety and the amide groups. Indeed, compound **13** with spacers of equal length did not adsorb in the folded conformation as found by STM analysis since this conformation is rather twisted than flat, rendering adsorption of the folded structure unfavorable. On the other hand, compounds **14** and **15** with spacers differing by one methylene group did adsorb in the folded conformation, and especially **15** yielded highly ordered monolayers at the solid–liquid interface. This study shows that, in addition to the factors discussed above, the 'flatness' of the folded conformation, encoding for maximum interaction with the flat substrate surface, can determine whether the compound adsorbs in the folded or unfolded conformation.

In related peptide work, Kelly, Powers, and coworkers have studied the self-assembly of β-hairpin peptide **16** at interfaces employing various techniques (Fig. 13.18) [37–40]. Peptide **16** consists of two strands with alternating hydrophobic and hydrophilic residues linked by a *D*-Pro-Gly β-turn and labeled with a fluorophore (DMBDY). CD measurements showed that this water-soluble peptide adopts a random coil conformation in deionized water [37]. However, it spontaneously adsorbs at the air–water interface as indicated by fluorescence microscopy. The surface pressure isotherms of the formed monolayers have been studied and LB films have been deposited on mica substrates [38]. The observed area per molecule and the pattern in the monolayer on mica as observed by AFM are in good agreement with adsorption of the peptide in the folded state, i.e. a hairpin conformation (Fig. 13.18). The derived structural model was further supported by a neutron reflection study. A later study revealed that the SAMs at the air–water in-
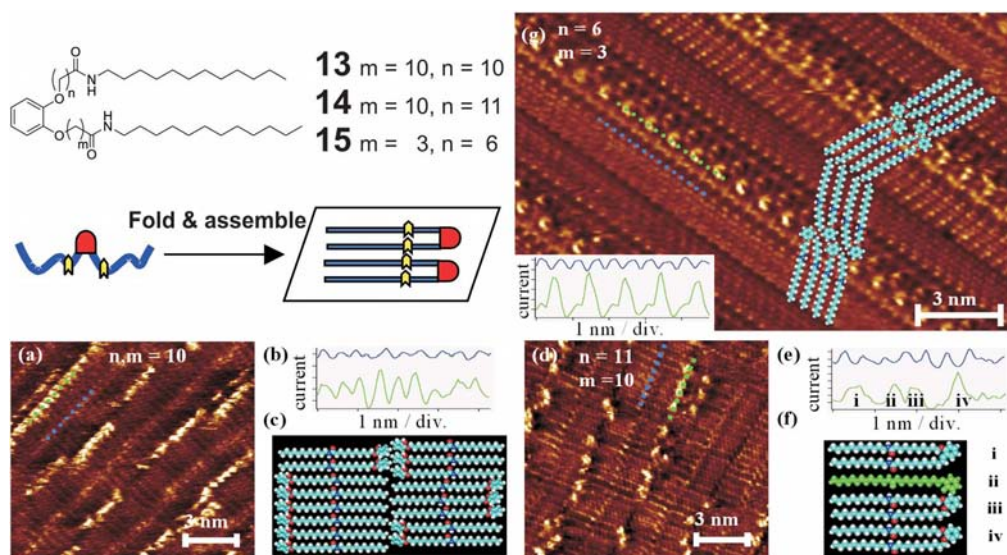
Fig. 13.17 Structure and concept of turn mimics **13**–**15** (top left) and STM images of the adsorbed turn elements on HOPG [36].
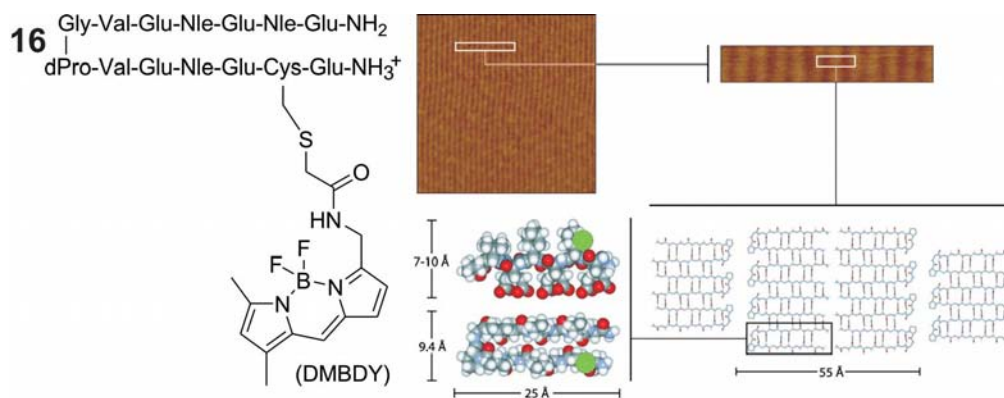
Fig. 13.18 Hierarchical self-assembly of β-hairpin peptide **16** containing a DMBDY fluorophore (left) in tapes as shown by the AFM image on mica and the corresponding structural analysis [38]. (Reproduced from ref. [38] with permission.)

terface are stabilized by the DMBDY fluorophores and the *D*-Pro-Gly β-turn elements [39]. This might be due to van der Waals forces between the DMBDY fluorophores and between the hydrophobic turn elements. Moreover, it was found that at higher pH or higher salt concentration as compared to deionized water,

the glutamic acid residues could get deprotonated resulting in destabilization of the monolayer at the air–water interface. While it becomes apparent that the folding and adsorption of these hairpins is governed by different types of interactions and that specific conditions are required, clearly folding is facilitated by adsorption, probably due to lowering of the nucleation barrier.

13.6
Outlook

From this brief survey of the research, which has been carried out in the area of foldamers at interfaces, it becomes clear that many fundamental aspects of this field still wait to be uncovered and that consequential applications linger to be explored. This certainly is good news – much more can be discovered!

Undoubtedly, a more detailed understanding of the mechanism of folding at the interface including both thermodynamics and kinetics is needed. Important parameters concerning all three intramolecular, intermolecular, and interfacial interactions need to be evaluated and systematically studied in order to comprehend the synergistic interplay of conformation, aggregation, and adsorption. For this purpose, modular synthetic systems, which allow structure–property relationships to be deduced, have to be prepared and studied by advanced techniques for *in situ* characterization at various types of interfaces under a multitude of environmental conditions. These experimental studies have to be supported by high-level theory in the area of molecular dynamic simulations including appropriate molecular and surface models. The majority of studies carried out on natural folding backbones, i.e. peptides and DNA, have to be complemented by investigations involving non-natural foldamers, which might offer beneficial properties with regard to applications in materials science.

In addition to fundamental research, chemists will certainly continue to design increasingly complex systems and implement various functions into foldamers. This will open opportunities for applications ranging from nano- to biotechnology. For example, surface-confined foldamers could be utilized to display various functional groups at defined surface locations to enable high-resolution chemical surface patterning, while other foldamers could control growth of certain inorganic materials analogous to biomineralization and designer peptides could facilitate cell-adhesion to body transplants. Many more possibilities can be anticipated and we are only limited by our imagination!

Last but not least, we would like to seize this opportunity to encourage chemists as well as scientists from neighboring disciplines to embrace *molecular* science. Molecules and their very intimate relationship to one another are characterized to a large extend by their shape and hence their conformation [41]. While the interface certainly adds a new complexity to conformational design and molecular recognition, the essence remains vital since Emil Fischer. There are a lot more new tricks to learn from old dogs!

References

1 The term ''foldamer'' has been coined by Gellman, see: S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.
2 For a comprehensive review consult: D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4012.
3 Folding models are comprehensively reviewed in: K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, H. S. Chan, *Protein Sci.* **1995**, *4*, 561–602.
4 For a commentary see: M. Muthukumar, *Proc. Nat. Acad. Sci. USA* **1999**, *96*, 11690–11692.
5 J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167–195.
6 For a review see: E. T. Kaiser, F. J. Kezdy, *Proc. Nat. Acad. Sci. USA* **1983**, *80*, 1137–1143.
7 For a review see: W. Norde, *Adv. Colloid Interface Sci.* **1986**, *25*, 267–340.
8 For a review see: D. W. P. M. Löwik, J. C. M. van Hest, *Chem. Soc. Rev.* **2004**, *33*, 234–245.
9 For a review see: K. Rajagopal, J. P. Schneider, *Curr. Opin. Struct. Biol.* **2004**, *14*, 480–486.
10 For a review see: R. Fairman, K. S. Åkerfeldt, *Curr. Opin. Struct. Biol.* **2005**, *15*, 453–63.
11 For a review see: H. Rapaport, *Supramolecular Chemistry* **2006**, *18*, 445–454.
12 W. F. DeGrado, J. D. Lear, *J. Am. Chem. Soc.* **1985**, *107*, 7684–7689.
13 The hydrophobic repeat unit periodicity refers to the average number of residues between the start of a sequence of adjacent

hydrophobic residues and the start of the next hydrophobic sequence.
14 L. A. Cuccia, E. Ruiz, J.-M. Lehn, J.-C. Homo, M. Schmutz, *Chem. Eur. J.* **2002**, *8*, 3448–3457.
15 J. R. Lu, S. Perumal, I. Hopkinson, J. R. P. Webster, J. Penfold, W. Hwang, S. G. Zhang, *J. Am. Chem. Soc.* **2004**, *126*, 8940–8947.
16 B. S. Li, K. K. L. Cheuk, L. Ling, J. Chen, X. Xiao, C. Bai, B. Z. Tang, *Macromolecules* **2003**, *36*, 77–85.
17 B. S. Li, S. Z. Kang, K. K. L. Cheuk, L. Wan, L. Ling, C. L. Bai, B. Z. Tang, *Langmuir* **2004**, *20*, 7598–7603.
18 S. L. Burkett, M. J. Read, *Langmuir* **2001**, *17*, 5059–5065.
19 M. J. Read, S. L. Burkett, *J. Colloid Interface Sci.* **2003**, *261*, 255–263.
20 M. Vankann, J. Möllerfled, H. Ringsdorf, H. Höcker, *J. Colloid Interface Sci.* **1996**, *178*, 241–250.
21 M. L. De Vocht, K. Scholtmeijer, E. W. van der Vegte, O. M. H. de Vries, N. Sonveaux, H. A. B. Wöster, J.-M. Ruysschaert, G. Hadziioannou, J. G. H. Wessels, G. T. Robillard, *Biophys. J.* **1998**, *74*, 2059–2068.
22 M. L. De Vocht, I. Reviakine, W.-P. Ulrich, W. Bergsma-Schutter, H. A. B. Wösten, H. Vogel, A. Brisson, J. G. H. Wessels, G. T. Robillard, *Protein Sci.* **2002**, *11*, 1199–1205.
23 A. E. Allahverdyan, Z. S. Gevorkian, C.-K. Hu, T. M. Nieuwenhuizen, *Phys. Rev. Lett.* **2006**, *96*, 098302/1-098302/4.
24 H. Rapaport, K. Kjaer, T. R. Jensen, L. Leiserowitz, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 12523–12529.
25 H. Rapaport, G. Möller, C. M. Knobler, T. R. Jensen, K. Kjaer, L.

Leiserowitz, D. A. Tirrell,
J. Am. Chem. Soc. **2002**, *124*, 9342–
9343.

26 R. Sneer, M. J. Weygand, K. Kjaer,
D. A. Tirrell, H. Rapaport, *Chem.
Phys. Chem.* **2004**, *5*, 747–750.

27 J. Yoon, M. Ree, Y. Hwang, S. W. Lee,
B. Lee, J.-S. Kim, H. Kim, S. N.
Magonov, *Langmuir* **2004**, *20*, 544–
549.

28 H. Chen, S. L. Hsu, D. A. Tirrell,
H. D. Stidham, *Langmuir* **1997**, *13*,
4775–4778.

29 For a review consult: C. R. Ray, J. S.
Moore, *Adv. Polym. Sci.* **2005**, *177*,
91–149.

30 L. Arnt, G. N. Tew, *Langmuir* **2003**,
19, 2404–2408.

31 L. Arnt, G. N. Tew, *J. Am. Chem. Soc.*
2002, *124*, 7664–7665.

32 A. Khan, M. A. Balbo Block, S.
Hecht, N. Severin, J. P. Rabe,
unpublished results.

33 A. Aggeli, I. A. Nyrkova, M. Bell, R.
Harding, L. M. Carrick, T. C. B.
McLeish, A. N. Semenov, N. Boden,
Proc. Nat. Acad. Sci. USA **2001**, *98*,
11857–11862.

34 C. W. G. Fishwick, A. J. Beevers, L. M.
Carrick, C. D. Whitehouse, A. Aggeli,
N. Boden, *Nano Lett.* **2003**, *3*, 1475–
1479.

35 C. Whitehouse, J. Fang, A. Aggeli, M.
Bell, R. Brydson, C. W. G. Fishwick,
J. R. Henderson, C. M. Knobler, R. W.
Owens, N. H. Thomson, D. A. Smith,
N. Boden, *Angew. Chem. Int. Ed.*
2005, *44*, 1965–1968.

36 N. Schuurmans, H. Uji-i, W.
Mamdouh, F. C. De Schryver, B. L.
Feringa, J. van Esch, S. de Feyter,
J. Am. Chem. Soc. **2004**, *126*, 13884–
13885.

37 E. T. Powers, J. W. Kelly, *J. Am. Chem.
Soc.* **2001**, *123*, 775–776.

38 E. T. Powers, S. I. Yang, C. M. Lieber,
J. W. Kelly, *Angew. Chem. Int. Ed.*
2002, *41*, 127–130.

39 S. Colfer, J. W. Kelly, E. T. Powers,
Langmuir **2003**, *19*, 1312–1318.

40 J. R. Lu, S. Perumal, E. T. Powers,
J. W. Kelly, J. R. P. Webster, J.
Penfold, *J. Am. Chem. Soc.* **2003**, *125*,
3751–3757.

41 S. Hecht, *Materials Today* **2005**, *8(3)*,
48–55.

Index

Related Titles

Tramontano, A.

Protein Structure Prediction
Concepts and Applications

approx. 228 pages with approx. 160 figures
2006
Softcover
ISBN-13: 978-3-527-31167-5
ISBN-10: 3-527-31167-X

Buchner, J., Kiefhaber, T. (eds.)

Protein Folding Handbook

2623 pages in 5 volumes with 644 figures and 86 tables
2005
Hardcover
ISBN-13: 978-3-527-30784-5
ISBN-10: 3-527-30784-2

Friesner, R. A. (ed.)

Computational Methods for Protein Folding
A Special Volume of Advances in Chemical Physics

542 pages
2002
Hardcover
ISBN-13: 978-0-471-20955-3
ISBN-10: 0-471-20955-4