The Biochemistry of Cell Signalling

Ernst J. M. Helmreich

Cover Page p. 329 | **Edition Info** p. 332 | **Preface** p. 334 | **Acknowledgements** p. 337 | **Abbreviations** p. 339

Plate Section p. 348 | **INDEX** p. 323

Back Page p. 364

Contents

Part 1 **The machinery of signal transduction** 1

1 Molecular basis of signal transduction 3

2 Activation of receptors by oligomerization 24

3 Components of signalling networks: linkers and regulators 31

4 Signal transduction pathways through small monomeric G proteins 57

5 Signal transduction pathways through heterotrimeric G proteins: transmission of hormonal and sensory signals 76

6 Signal transduction pathways controlling morphogenesis and haematopoiesis 102

7 Control of signalling by phosphorylation and dephosphorylation 122

8 Regulation by a hormone: the insulin response 137

Part 2 **Cell signalling and gene transcription** 155

9 Machinery of gene transcription 157

10 Regulation of gene transcription by growth factors and cytokines 172

11 Regulation of gene transcription by hormones 190

Part 3 **Global cell regulatory programmes** 211

12 Regulation of the cell cycle 213

13 Regulation of cell death 234

14 Regulation of the immune response 250

Part 4 **Loss of the regulatory control and its consequences** 267

15 Transformation of normal cells to tumour cells 269

16 Loss of developmental controls in cancer 284

17 The causes of cancer 294

Glossary 303

Index 323

1

The machinery of signal transduction

1

Molecular basis of signal transduction

In this chapter, the signals, the growth factors, and the receptors receiving the signals are introduced. Receptor–ligand interactions at the cell surface are the first step in cellular signalling. This introduction will acquaint the reader with some of the major players in cellular signalling.

Properties of receptors

Cellular signal transduction is a two-step process: First, a signalling molecule is sensed by a receptor at a target cell and then the receptor is activated (Fig. 1.1). Membrane-bound receptors respond to a large spectrum of extracellular signals. External signals range from light and odours to hormones, growth factors, and cytokines. When the receptor sensing the signal is a catalyst, a kinase, the response is amplified.[1] As diversified as the signals, are the proteins which respond to them. In each case, binding of a signalling molecule converts the dormant receptor to an active state. A mechanism involved in the transition of a receptor from its inactive to its active state is receptor oligomerization. This is discussed in Chapter 2.

All receptors that transmit signals from the surface of the cell to the interior, and finally to the nucleus and the genes, have two features in common:

(1) the signalling molecule binds to the extracellular domain of the membrane-inserted receptor; and

(2) ligand binding triggers, in a cooperative manner, a change in the domain inside the cell.

Although the molecular details of this conformational transition are not yet elucidated in all cases, it is to be expected that they differ as much as the structure of the receptor. For all membrane-bound receptors, the effects of the environment, the lipid bilayer, should be taken into account, although little is known about such influences.

Upon binding the ligand and activation, ligand–receptor complexes are eventually internalized. Internalized ligand–receptor complexes are dissociated in acidic

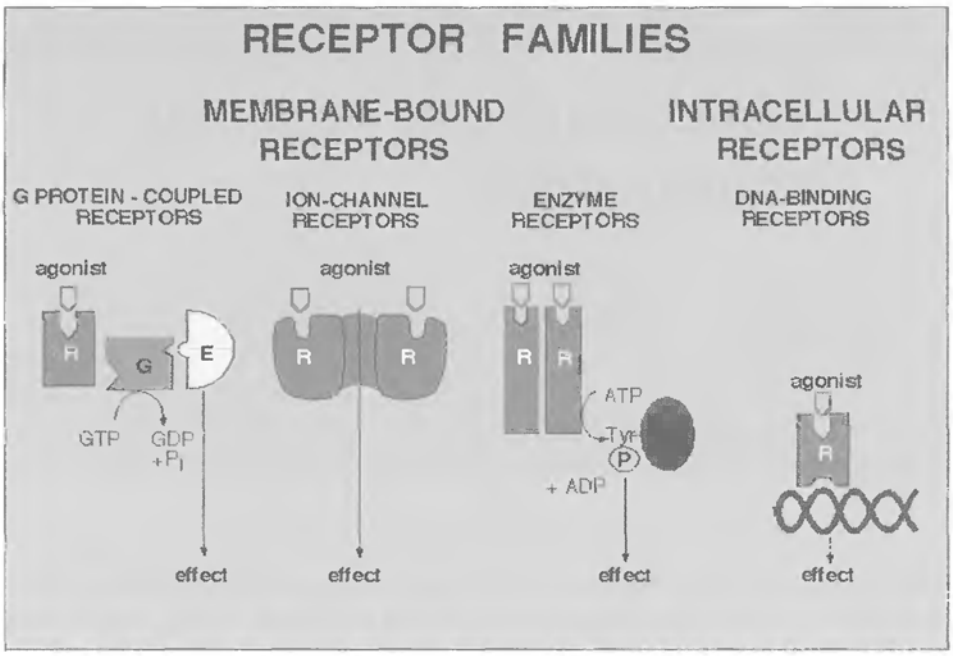


Fig. 1.1 Four major families of receptors are shown: From left to right: G-protein-coupled receptors. A representative example of this receptor family is rhodopsin, which is responsible for visual perception in the eye. R is the receptor; G is the GTP-binding protein; E is the effector. Effectors are often enzymes which form second messengers. An example is adenylyl cyclase, which catalyses the synthesis of cAMP from ATP. In the case of the visual response, the effector is a cGMP phosphohydrolase, which converts cGMP to GMP (see Chapter 5). A representative of ion-channel receptors is the nicotinic acetylcholine receptor. This type of receptor will not be discussed in this book. Enzyme receptors are transmembrane receptors with intrinsic enzymatic activity. Examples are the receptor tyrosine kinases (RTKs), which autophosphorylate their own tyrosine residues, such as the growth factor receptors and the insulin receptor. And, finally, there are the intracellular DNA-binding receptors. They bind lipophilic ligands that pass through the membrane. They address genes directly. Examples are the steroid hormone receptors (see Chapter 11). (This figure was donated by Professor Martin Lohse, University of Würzburg.)

endocytosolic vesicles and the ligand is degraded in lysosomes, whereas the receptor may be degraded or recycled back to the cell surface. Receptor–ligand complexes may be internalized together with proteins which regulate their endocytosis and degradation.[2] Receptor desensitization by removal from the membrane and endocytosis is a feature shared by single-pass tyrosine kinase receptors and serpentine, heptahelical G-protein-coupled receptors (see Chapter 5).

It is not clear whether membrane-anchored growth factors (not bound to receptors), are also endocytosed, like their receptors, and if so how. The domain structure of some membrane-anchored growth factors and the high degree of conservation of their cytosolic domain have raised the question whether they may also be receptors, but there is no evidence to support such a function.

A process which, in most cases, modulates receptor signalling is phosphorylation. In the case of G-protein-coupled heptahelical receptors, the interaction with specific kinases is the first step in shutting off their action. In other cases, binding of a growth factor to a receptor triggers the intrinsic receptor kinase activity and leads to autophosphorylation. The important point is that the phosphates introduced in the receptor are essential for recognition and binding of other proteins, adaptors and transducers, which are often cytosolic protein kinases and phosphatases. Signalling triggered by growth factor–receptor interactions leads to a response which is often of global nature, such as growth, proliferation, and differentiation of cells. Growth factors affect the cell cycle and the cell death programmes, which determine the fate of the cell. Although, many processes vital for the cell are affected, the main target is the genome. The essence of cellular signalling is the transmission of signals from the surface of the cell to the nucleus and the subsequent expression of genes. Dysfunction of the regulatory mechanisms controlling these processes can cause malignant transformation of cells and other diseases.

I begin with signals that regulate cell growth and proliferation and describe the receptors that respond to these signals. This survey is by no means complete; only those growth factors and cytokines that we shall encounter again later are described.

Properties of signals

Although, principal features of receptor activation are preserved, different ligands and receptors account for the diversity of the biological effects.

Growth factors and hormones can engage in autocrine, juxtacrine, paracrine, and endocrine stimulation. If the receptor resides on the same cell where the ligand is expressed, the resulting cellular stimulation is autocrine; when the growth factor diffuses from the cell to neighbouring cells in the same organ, stimulation is paracrine, and juxtacrine stimulation is a mode of signalling reserved for those membrane-anchored growth factors which interact with receptors located on neighbouring cells. When a factor is transported through the bloodstream from the place of synthesis to other tissues equipped with receptors that recognize the factor, stimulation is endocrine, as in the case of hormones such as insulin.

Growth factors, cytokines, and hormones are, in large part, proteins. They are synthesized by the usual mechanisms of protein synthesis and transported through common routes of cellular protein traffic. Often inactive precursors are converted by limited proteolysis into active factors and shed by the cells.[3]

Processing of growth factors

Proteolytic processing is an ancient pathway, conserved from worms to humans. Processing of membrane-anchored growth factors is complex (cf. ref. 4). It usually involves cleavage of specific peptide bonds at the N- and C-termini of a large precursor. Cleavage and shedding of the ectodomains of plasma membrane proteins is carried out by transmembranous and soluble metalloproteinases, such as the matrix metalloproteinases (MMPs) and their relatives.[5] Serine proteinases have also been implicated.[6]

The first such 'sheddase' characterized was the tumour necrosis factor-α (TNF-α) converting enzyme (TACE).[10] TACE is a membrane-anchored proteinase and a member of the ADAM (*A* *d*isintegrin *a*nd *m*etalloproteinase) family of proteins that combine cell-surface adhesion and proteinase activity.[7–10]

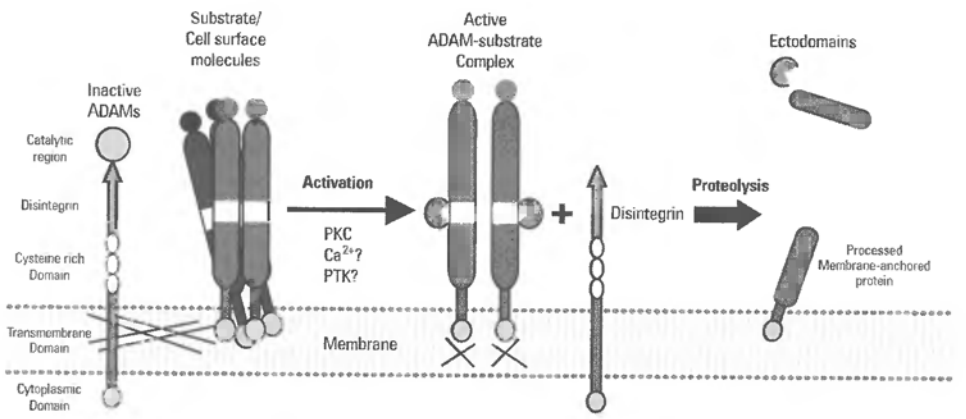


Fig. 1.2 A scheme of the activation of ADAMs and of the shedding process. The dimeric ADAM proteases and their substrates are anchored in the membrane, but are separated from each other. Upon activation (via protein kinases?), the protease is disengaged from disintegrin and associates with the substrate. Proteolysis takes place and the free, soluble ectodomains of membrane-anchored cell-surface proteins are shed from the membrane-anchored substrates and released. (Reproduced from ref. 3, with permission of the authors and *Science*.)

ADAMs, such as TACE, have a common domain organization (Fig. 1.2); they are zinc-dependent metalloproteinases, closely related to the MMP family. The action of the shedding enzymes is controlled by endogenous inhibitors, the TIMPs (tissue inhibitors of metalloproteinases).[11] The three-dimensional structure of TACE has been solved.[8] It is widely expressed and has several functions, among them the production of ligands for the epidermal growth factor (EGF) receptor. TACE is essential for epithelial development.[12] However, TACE and related metalloproteinases process not only growth factor ligands, they also cleave the ectodomains of receptors[13] and other membrane proteins.

Processing of membrane proteins by the ADAMs and other sheddases requires that both the membrane-anchored enzyme and its substrate are located on the same cell. The regulation of cleavage and the removal of the split products from the cell surface has been studied with angiotensin-converting enzyme (ACE), a typical sheddase.[14] These studies raised intriguing questions, of how the proteinases are kept apart from their substrates until shedding is triggered and, conversely, what brings the proteinase and its substrate together. Another question is why only ectodomains of certain transmembrane proteins, out of many such proteins on the cell surface, are targetted.

As for the first question: cytoskeletal interactions are apparently involved in keeping the proteinases and their transmembrane substrates apart, keeping them in distinct domains of the plasma membrane. See also: ref. 15. Upon activation, the cytoskeletal attachments seem to be loosened so that the proteinases and their substrates can approach each other and interact. How spatial restriction and compartmentalization of enzymes and substrates controls their action is of more general importance. We shall come back to that (Chapter 7).

Proteolytic processing of membrane-anchored proteins is an important event. TNF-α, TGF-α, EGF, heparin-binding EGF-like growth factor (HB-EGF), the Kit ligand, and other growth factors must all be processed to become functional. Processing has other

important consequences: cleavage of adhesive molecules profoundly affects cell–cell adhesions and cell–cell interactions in mammary epithelial cells; dysfunction of processing may cause cancer.[16] Processing of the Notch receptor plays a central role, not only in cell adhesion but also in neurogenesis in vertebrates and Drosophila.[17] In Drosophila Notch controls cell–cell interactions that are involved in the diversification and differentiation of cells during development. The Notch receptor participates in cell development and in the selection of cell lineages. It is activated by protein ligands, localized on the surface of neighbouring cells, rather than by soluble ligands. This kind of receptor, interacting with adhesive ligands, may also participate in pathfinding in the nervous system. Finally, shedding can also control cell death by processing ligands for death receptors (see Chapter 13).[18]

Many growth factors that become soluble after having been processed are not shed and remain anchored to the membrane. Therefore, the distinction between membrane-anchored[19] and diffusible growth factors is somewhat ambiguous. In malignant transformed cells, diffusible factors may be retained in the membrane of the same cell where they are synthesized and processed. This can stimulate the cell in an autocrine fashion with deleterious effects.

A few specific details on the processing of growth factors and hormones follow.

Processing of pro-TGF-α

TGF-α is a growth factor.[20] Its processing is of interest. Pro-TGF-α is a 160-amino-acid protein which is heterogeneously N- and O-glycosylated.[21] The whole extracellular ectodomain of pro- TGF-α, with the sugars, is cleaved off and secreted.[22]

Mutations have identified the C-terminal valine, in the highly conserved cytoplasmic region of pro-TGF-α, as the critical determinant for cleavage. Substitutions elsewhere throughout the cytoplasmic region have no effect on cleavage, as long as valine remains the C-terminal amino acid. Replacement of this residue abolishes cleavage, and even conservative substitutions of this valine markedly impair processing, suggesting a highly specific recognition event. The simplicity of a single residue, a C-terminal valine, as the pro-TGF-α cleavage signal is striking. How this C-terminal valine determines cleavage is a matter of speculation.[23]

The colony-stimulating factors

The granulocyte–macrophage colony-stimulating factors (GM-CSFs) are cytokines (Chapter 6). The gene for CSF-1 encodes a 554-amino-acid precursor protein. A huge portion is removed and only a small region of 36 amino acids is retained. The final macrophage colony-stimulating factor is a soluble, heterogeneous proteoglycan.[24,25]

TNF-α and TNF-β

Two kinds of tumour necrosis factors exist: TNF-α and -β. They are either membrane bound or soluble. TNF-α is identical with cachectin,[26]. It is produced in monocytes and macrophages; TNF-β is synthesized mainly in lymphocytes. Only TNF-α is processed and cleaved from a membrane-bound glycoprotein.[27] The processed, mature human TNF-α is a 145–157-amino-acid peptide chain, depending on the cleavage site. The chains associate, forming a soluble trimer.

Vascular endothelial growth factors

Vascular endothelial growth factors (VEGFs) are cytokines which increase vascular permeability and promote angiogenesis and blood vessel growth during embryogenesis[28]

(a second family of vascular growth factors are the angiopoietins). VEGFs also promote the formation of lymphatic vessels.[29] One of the VEGFs, VEGF-C, is synthesized as a precursor and must be processed for activation. Apart from VEGF, growth factors such as fibroblast growth factor (FGF) and TGF-α also promote angiogenesis. Angiogenic factors have an important role in the growth of cancers and in metastasis.[30] Therefore there is great interest in the development of blockers of VEGFs to slow down the growth of cancer.

Processing of pro-insulin

Inclusion of insulin in the group of growth factors may be disputed. Processing as exemplified above is a post-synthetic process, whereas processing of pro-insulin is part of the synthesis of insulin, taking place at the site of synthesis, the β-cells of the pancreas. It is the final step in the formation of an active hormone. But since insulin processing is known in great detail, mainly thanks to the work of Donald Steiner and his laboratory,[31] and since insulin signalling is discussed later (Chapter 8), insulin processing will be included here.

The gene for insulin is located on the short arm of chromosome 11. Expression of the insulin gene yields a precursor protein (preproinsulin), of 104–109 amino acids, depending on the animal species, with a 24-amino-acid signal peptide. After removal of the signal peptide in the endoplasmic reticulum, proinsulin is formed. Proinsulin resembles the insulin-like growth factors, IGF-1 and IGF-2 (7 kDa proteins), and the female hormone, relaxin, which are larger than the finally processed insulin. The next step involves generation of inter- and intrachain disulphide bridges, facilitated by the removal of the C-peptide. The separate A and B chains of mature insulin are formed and held together by interchain disulphide bridges (Fig. 1.3).

Processing is carried out by a family of well-characterized proteases. Conversion takes place in secretory granules and the excised C peptide is secreted together with mature insulin into the bloodstream. Peptide C is thought to link the A and B chains in the precursor in a way that facilitates folding and efficient formation of the interchain disulphide bonds. No receptors for the C peptide have been found, and all the available evidence suggests that the C peptide does not interact with a receptor, although the C peptide was recently claimed to have biological activity.[33]

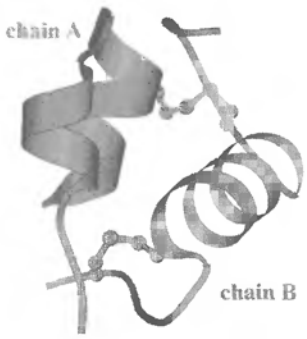


Fig. 1.3 Parts of the disulphide-linked A and B chains of bovine insulin. (Reconstructed from data[32] deposited in the Brookhaven protein databank, with permission of Professor D. L. D. Caspar and the Biophysical Journal.)

Receptors for growth factors and their ligands

Growth-factor signalling is mainly mediated through receptors with tyrosine kinase activity (RTKs) (Table 1.1). Activation of these receptors involves dimerization and autophosphorylation on tyrosyl residues. The phosphotyrosyls then become docking sites for intracellular regulatory proteins, linkers, which are instrumental in transmitting the signal (Fig. 1.4).[34]

Table 1.1 The properties of some growth factors and their cognate receptors

Growth factor	Properties	Receptor	Properties
Epidermal growth factor (EGF)	EGF-unit with 6 conserved cysteines	EGF receptor	Single pass Tyrosine kinase receptor
Transforming growth factor-α, (TGF-α)	40% homologous with EGF	EGF-like receptor	Dimerization Autophosphorylation
Neuregulin	Neuregulin signalling is involved in the development of the heart and the peripheral nervous system[48]	ErB2, ErB3, ErB4	Truncated EGF receptors
Platelet-derived growth factor (PDGF)	Homo- and heterodimeric AA, BB, AB forms	PDGF-receptor-α PDGF-receptor-β	Homo- and heterodimeric Tyrosine kinases
VEGF	Vascular endothelial growth factor expressed in endothelial cells	VEGFR	PDGF-like receptors
MCSF-1	Macrophage colony factor	MCSFR	" "
Acidic and basic fibroblast growth factors (FGFs)	Tissue-specific functions FGFs bind to heparan-sulphate glycosaminoglycans	FGFRs	Homo and heterodimeric tyrosine kinases
Keratinocyte growth factors (KGFs)	Wound healing	FGF-like receptor	" "
Kit/KL: mast cell growth factor, stem cell or steel factor	Two heterogeneously glycosylated forms. KL-1, KL-2. Role in haemalopoiesis and in the development of cell lineages	c-Kit receptor	Receptor tyrosine kinase
Insulin	Circulating hormone, formed and secreted from the β-cells of the pancreas Blood sugar regulation	Tyrosine kinase receptor with constitutive oligomeric structure	Heterodimer consisting of α- and β-disulphide-linked subunits Autophosphorylation Forms a tetramer on ligand binding
Insulin-like factor	Regulates metabolism and longevity in *Caenorhabditis elagans*	Insulin-like receptors, DAF-2 in *Caenorhabditis elegans* and *Drosophila*	Tyrosine kinase domains
HGF, SF hepatocyte growth factor, scatter factor	Dissociates sheets of epithelial cells and promotes cell mobility, stimulates cell proliferation and induces cell polarity. Mutations cause cancer	Insulin-like receptors	Tyrosine kinases
HGF-like factor	Factors implicated in axonal pathfinding	Axl receptors	Tyrosine kinases, related to the insulin receptor
Ephrins	Ephrins may be involved in cell–cell signalling	Eph-like receptors	Tyrosine kinases

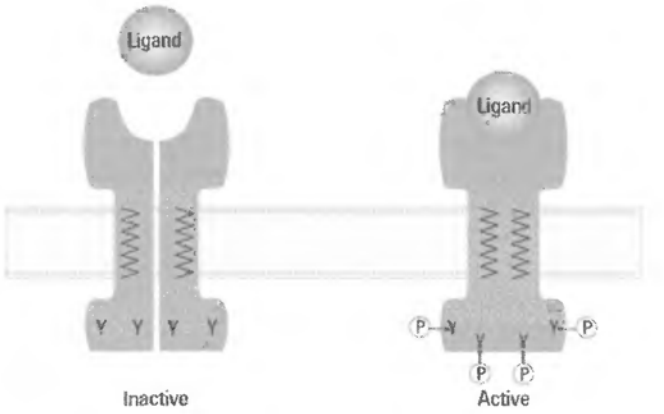


Fig. 1.4 Receptor dimerization and autophosphorylation of tyrosines (Y) on ligand binding. (The molecular basis of this mode of receptor activation is discussed in Chapter 2).

The EGF family of growth factors and the cognate receptors

EGF was the first growth factor to be identified. This was achieved by Stanley Cohen. see ref. 35. It is the prototype of a heterogeneous family of growth-promoting proteins, which all share one or more EGF-like structural units.

The EGF-unit is defined by six conserved cysteines in a stretch of 35–40 amino acids: $CX_7CX_{3-5}CX_{10-12}CXCX_5GXRC$, where C is cysteine; G, glycine; R, arginine; and X any other amino acid. The six cysteines form three disulphide bonds, C1–C3, C2–C4, and C5–C6. This motif is conserved in all EGF-like growth factors. Some of them contain calcium-binding motifs (for a review see ref. 36).

A variety of quite different factors belong to the family of EGF-like growth factors, among them TGF-α. TGF-α shares 44% and 33% homology with human and mouse EGF, respectively.[37] Whereas TGF-α binds to a typical receptor tyrosine kinase, TGF-β binds to receptors which are serine–threonine-specific kinases (Chapter 6). Other growth factors with EGF structural units are the heparin-binding EGF-like growth factor (HB-EGF)[38] and the vaccinia virus growth factor (VGF). They all bind to EGF receptors. The EGF unit of HB-EGF contains a N-terminal extension, rich in basic amino acids, that is responsible for binding heparin. Many different growth factors bind to one and the same EGF receptor, although the amino-acid sequence identity between these factors and EGF is less than 35%. This points to plasticity of the receptor interaction surface. Furthermore, each of these different factors elicits a distinct biological response, depending on the cell type (e.g. epithelial or mesenchymal cells). The cell-specific biological response implicates the existence of unique signal-transducing pathways in each differentiated cell. In some cases, the cell provides a different adaptor which directs the signal to a different target.

As EGF is a prototypical growth factor, EGF receptors (EGFRs) are prototypes of receptor tyrosine kinases (RTKs). Homologues of the EGFR exist in Drosophila (DER)[39] and in Caenorhabditis elegans (LET-23).[40]

The non-activated EGF receptor is a single-pass transmembrane protein of 170 kDa (Fig. 1.5).

All receptors of the EGF family have two cysteine-rich clusters in the extracellular region and a tyrosine kinase domain in the cytoplasmic part. Autophosphorylation sites are in the C-terminal regions of the cytoplasmic domain. On activation, five tyrosines are

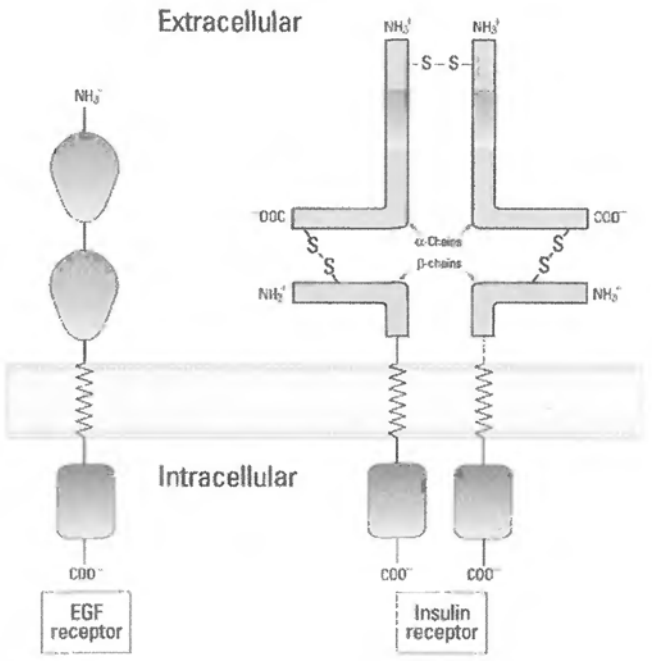


Fig. 1.5 The EGF receptor is compared with the insulin receptor, which is a constitutive heterodimer, consisting of α- and β-chains.

autophosphorylated. With the help of these specific phosphotyrosyls, coupling partners are recruited with amazing precision.[41]

There exists a family of truncated, EGF-like receptors, (ErB2, ErB3, ErB4), which regulate the neu gene (neu is an oncogene derived from the DNA of a rodent neuroblastoma, it is also known as *ERBB2* or *HER2*).

Platelet-derived growth factors and their receptors

Platelet-derived growth factors (PDGFs) exist in three different dimeric compositions, containing disulphide-linked A-, or B-, or A- and B-chains (AA, BB, or heterodimeric AB). The composition of the ligand dimers matches the composition of the receptor dimers to which they bind (Fig. 1.6). Therefore, a specific response of cells to PDGF depends on the composition of the ligand dimer which is recognized by a fitting receptor dimer. One of the aims of research in this active field is to identify different signalling pathways for each receptor–ligand combination.[42] Platelet-derived growth factor receptors, PDGFR-α and PDGFR-β, have quite similar primary sequences and domain structures, but are expressed in different cells. PDGFRs have extracellular immunoglobulin-like domains and a cytoplasmic domain with a large tyrosine kinase insert which presents the phosphotyrosines for coupling to signalling molecules. Upon binding a matching PDGF, the receptors form non-covalent homo- and heterodimers.

PDGF-like receptors are VEGFR, the receptor for the vascular endothelial cell growth factor, which is expressed only in endothelial cells, and the receptor for the MCSF-1, the macrophage colony-stimulating factor-1. PDGFRs have, like EGFRs, multiple autophosphorylation sites, each of which is, again, specific for a particular recognition domain of a coupling partner. For example, phosphorylation of tyrosine 1021 in the receptor tail of PDGFR-β is responsible for binding to phospholipase C (PLC-γ1). A mutated receptor,

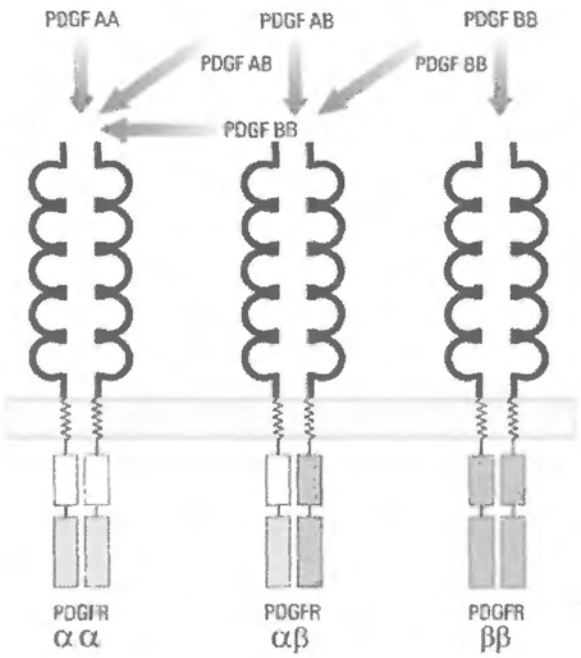


Fig. 1.6 Homo- and hetero-dimerization of the PDGFR-α and -β chains is matched by the assembly of the homo- and heterodimeric ligands, PDGF A and B. (From Lemmon and Schlessinger,[43] with permission of the authors and Trends Biochem. Sci.)

where Tyr1021 was replaced by Phe, could not activate the PLC pathway, but could still activate other signalling paths.[44,45] PLC-γ is a common receptor target. It couples with high affinity to the EGF receptor. This interaction is also mediated by a phosphotyrosyl residue (Y992 of the EGFR). The consequences in both cases are the same, phosphorylation and activation of PLC-γ[46,47] and formation of the second messenger, IP_3 inositoltrisphosphate. Other enzymes which are downstream mediators of growth factor receptor signals are the phosphatidylinositol-3-OH kinase[48] which, together with PLC-γ, propagates the mitogenic signal of the PDGF receptors.

Among proteins addressed by growth factor receptors are also Raf kinase (and other cytosolic kinases) and also monomeric GTPases, such as Ras. They play a central role in the regulation of growth factor signalling. (These pathways are discussed in Chapter 4). Each of the different partners recruited by the activated receptor initiates a different signalling pathway, making possible a great variety of cellular responses.[49] I shall come back to this when explaining how these proteins recognize each other and build cellular signal transmission chains over which the signal travels from the receptor on the membrane to the genes in the nucleus. At this point, I only wish to emphasize the coupling versatility of growth-factor RTKs.

Fibroblast growth factors and their receptors

At present, seven members of the fibroblast growth factor (FGF) family are known. Among them are the acidic and basic FGFs, aFGF and bFGF,[50] the keratinocyte growth factor (KGF),[51] which is involved in wound healing in the skin, and the angiogenic factors.[52,53]

In common with PDGFRs, the FGF receptors are single-pass polypeptide chains with a membrane-spanning domain, an extracellular region with three immunoglobulin-like

domains, and a cytoplasmic region with a catalytic tyrosine kinase domain. A short insert splits the kinase domain into two nearly equal halves. Differential splicing in the exons for the extracellular region of fibroblast growth factor receptor 1 generates receptor variants with different ligand-binding specificities.[54] FGFR variants with different ligand-binding specificities also arise from different exons in the FGFR1 and FGFR2 genes, expressed in a tissue-specific manner (exons are those parts of the genome that are expressed and transcribed into proteins).[55]

FGF receptor isoforms, FGFR1, FGFR2, and FGFR3, form both homodimeric and heterodimeric receptor species. The multiplicity of FGF receptors[56,57] explains the selective, individual responsiveness of cells and tissues to either acid or basic FGFs.

Binding of FGFs to their receptors has the familiar consequences: activation of the receptor tyrosine kinase and autophosphorylation (probably intermolecular) of the dimerized receptor.[58] Binding of FGFs to their cognate receptors is facilitated by heparan sulphate-type glycosaminoglycans.[59] These molecules belong to the ground substance, the glycocalix surrounding cells. They are made up of disaccharide units, containing a derivative of an amino sugar, either glucosamine or galactosamine with N- and O-linked, negatively charged sulphate and N-acetyl groups.[60] Crystal structures of complexes of bFGF with tetra- and hexasaccharides, derived from heparin, showed that no significant conformational change in bFGF occurs on association with saccharides, supporting the notion that glycosaminoglycans in the extracellular matrix only help to bring FGF to the membrane where the receptor is, but have no function in signalling (the same has been assumed in the case of HB-EGF). Thus, the idea was, until recently, that heparan sulphate proteoglycans, HSPGs, help to bind and protect ligands and concentrate them at the cell surface, but have no specific functions. However, there are proteoglycans, functioning as co-receptors, which participate in growth-factor signalling. One of them is a membrane-bound core protein, syndecan. Syndecans are transmembrane, heparan sulphate proteoglycans which interact with extracellular ligands and may facilitate the formation of receptor–ligand complexes.[61] Proteolysis of syndecan converts the syndecan co-receptor to a potent inhibitor of FGF-2.[62] New information from genetic studies in Drosophila and mice are beginning to reveal unique functions of HSPGs in specific signalling pathways involved in cell differentiation and morphogenesis.[63] The study of mutant phenotypes of either the enzymes synthesizing the heparan sulphate side-chains of the core proteins or synthesizing the core proteins themselves have indicated that the HSPGs are critical for signalling, specifically for the interactions of extracellular ligands with their signal-transducing receptors in the cell membrane.

Like other growth factors, FGF has distinct and different effects in different cell types. For example, in PC12 cells FGF promotes neurite outgrowth, in NIH3T3 cells it regulates cell proliferation and mitogenesis, and in Xenopus embryos it controls the formation of the mesoderm.

Interactions between the activated, autophosphorylated FGFRs and their coupling partners are as specific as EGFR and PDGFR interactions. For example, when Y766 is replaced by F (phenylalanine) the mutant receptor can neither associate with nor phosphorylate PLC-γ, and consequently some of the characteristic FGF responses are missing, such as an increase in intracellular calcium due to formation of second messengers (such as IP3), which is promoted by PLC-γ. But cells expressing this FGFR mutant, Y766F, can still respond to fibroblast growth factors with stimulation of cellular proliferation, suggesting that PLC-γ is not involved in this signalling pathway.[64] From this and

similar evidence one can conclude that, by engaging different transmitters, different signalling routes can be triggered by the same receptor, giving rise to different cellular responses.

The 'KIT' receptor and its ligand

Since later chapters deal with the development of differentiated cell lineages (Chapters 6 and 14), the receptor c-KIT, and its ligand (KL), are introduced here as an example of RTKs participating in the control of haematopoiesis. The c-KIT receptor is encoded by the proto-oncogene of a feline sarcoma retrovirus. It is involved in the development of haematopoietic stem cells and of melanocytes. Autophosphorylation upon ligand binding initiates signal transduction and recruitment of cytosolic signal transducers.[65] Signal transmission is regulated by the receptor tyrosine kinase activity of the c-KIT receptor.[66,67]

KL or 'KIT' is also known as the 'mast cell growth factor', or 'stem cell' or 'steel' factor.[68] KL and the cognate receptors are transcribed from genes in loci essential for mouse haematopoiesis.[69] Alternative splicing yields two KL forms, KL-1 (273 amino acids) and KL-2 (245 amino acids).[70,71] Both forms are heterogeneously glycosylated, giving rise to various cell-bound and soluble forms. On splicing, the shorter KL-2 form loses its main proteolytic processing site and remains bound to membranes.

Receptors for neurotrophins

The neurotrophins and their receptors are included in this introductory chapter on ligands and receptors participating in cellular signalling. References are provided so that the interested reader has access to the literature, to give him or her a more comprehensive view of this interesting field, since signalling in the neuronal system is not treated explicitly in this book.

Currently, there are six members of the neurotrophin family, including nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), glial-cell-line-derived neurotrophic factor, (GDNF),[72] and 'neurturin', which is structurally related to GDNF,[73] as well as several neurotrophins, NT-3, -4, and -5.

Three genes for vertebrate neurotrophin RTKs have been cloned. The three RTKs, A, B, and C, are highly selective with respect to the ligands they accept: the ligand for RTKA is NGF and BDNF, and the ligands for RTKB are the neurotrophins, NT-4/5, whereas NT-3 is bound exclusively to RTKC (Table 1.2).[74–76]

NGF was discovered by Rita Levi Montalcini in Victor Hamburger's laboratory at the Washington University in St Louis. It was also the first neurotrophin of which the structure (Fig. 1.7) was solved (reviewed in ref. 77,78). Originally it was thought that the 'cystine-knot' motif[79] in the ligand is important for recognition and for discrimination between different receptor tyrosine kinases. But this can no longer be upheld: ligands with a cystine-knot motif bind to structurally and functionally dissimilar receptors. The cystine-knot motif is a core structure, found also in transforming growth factor-β (TGF-β), PDGF, and human chorionic gonadotrophin (hCG). Whatever the role of the cystine-knot motif, it is probably not a structural motif, determining receptor–ligand interactions because, if that were the case, one would expect a matching motif in the diverse group of receptors which all accept ligands with a 'cystine-knot motif'. But such a motif has not been found.

Table 1.2 Summary of the neurotrophins and their receptors

Growth factor	Properties	Receptor	Properties
Nerve growth factor (NGF)	Cystine knot motif Promotes cell death in the nervous system, when bound to $p75^{NGFR}$	Two types of receptor, co-receptor? NGFR-α,β	$p75^{NGFR}$, have no tyrosine kinase activity. They are cytokine-type receptors like the TNF-receptor
NT-3, -4, and -5	Several neurotrophic factors, including NGF Support survival of neurons	NTFR	Tyrosine kinase receptors, 140 kDa, RTKs A, B, C
Brain derived neurotrophic factors (BDNF)	Support survival of neurons	BDNFR	Tyrosine kinase with large cytoplasmic region. Several subtypes
Glial-cell-line derived sine neurotrophic factor (GDNF)	Several neutrophic factors including NGF, ciliar neurotrophic factor, and neurturin, related to GDNF. Supports survival of glial cell lines	GDNFR-α and RET receptor	RET receptor is a tyro- kinase. GDNFR-α is a co-receptor which is required for activation of the RET receptor

Two different classes of transmembrane receptors are involved in signal transduction by the neurotrophins; one is an RTK, whereas the other, $p75^{NGFR}$, has no tyrosine kinase activity. $p75^{NGFR}$, has been classified as a member of the TNF-receptor family,[80] although the intracellular part of $p75^{NGFR}$ shows no homology to any known protein. The p75 NGF receptor is widely expressed in neuronal and also in non-neuronal cells and tissues, but its function is still controversial. The original assumption that NGF is a trophic factor for sympathetic neurons must be qualified, because how NGF functions depends on the kind of receptor to which it binds. NGF bound to $p75^{NGFR}$ receptors actually promotes cell death in the nervous system,[81] because $p75^{NGFR}$ is, like the TNF receptor, a 'death receptor' (see Chapter 13). On the other hand, binding of NGF to RTK receptors supports survival of neurons. This class of neurotrophin receptors comprises typical 140 kDa tyrosine kinases, with large cytoplasmic regions. Binding of neurotrophins leads to receptor dimerization, autophosphorylation, and activation.

Interaction of neurotrophin RTKs with $p75^{NGFR}$

The fact that the NGF-binding RTK and the $p75^{NGF}$ receptor are co-precipitated by antibodies in cell extracts raised the possibility that both types of receptors may interact with each other.[82–84] Some kind of a co-receptor function of $p75^{NGFR}$ for activation of the neurotrophin p140 RTKs, would be reminiscent of the co-receptor function of GDNFR-α in the activation of RET, the RTK which mediates signalling of GDNF, the glial-cell-line-derived neurotrophic factor. The co-receptor, GDNFR-α, has no cytosolic domain and therefore is incapable of signalling. For signalling, it needs the RET eceptor.[85,86] The co-receptor GDNFR-α is found in all cells and tissues that are GDNF responsive. Expression of GDNFR-α seems to be controlled developmentally. Several isoforms of GDNFR-α are expressed in different parts of the brain. A homologue of GDNFR-α facilitates binding of another neurotrophic factor, ciliar NTF, to its receptor.

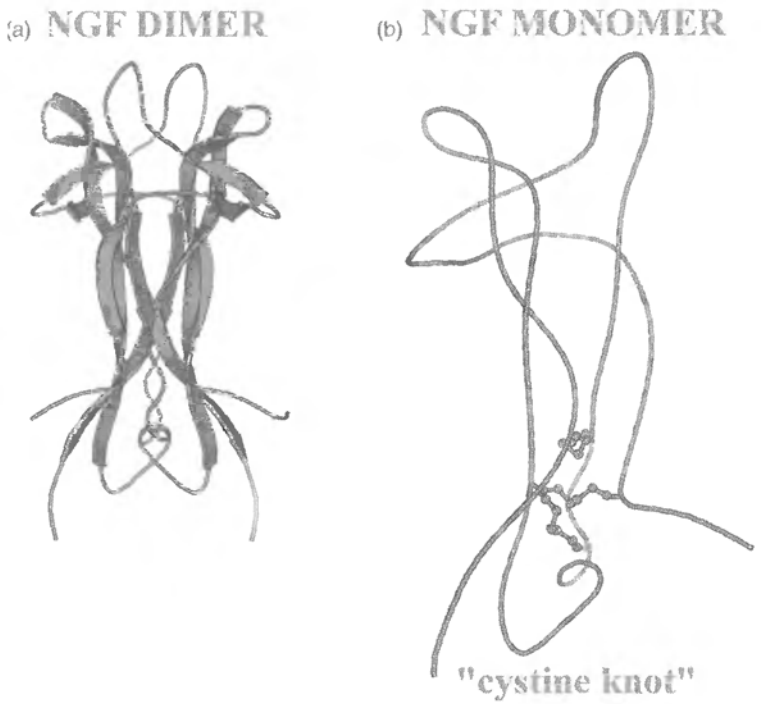


Fig. 1.7 (a) The crystallographic structure of 7S dimeric mouse NGF. (Reproduced with permission of Professor Thomas Blundell and Structure from ref. 78, using data from protein data banks). The dimeric NGF molecule is rather asymmetric and has a flat, elongated shape, a feature for which the presence of two pairs of antiparallel β-strands in each NGF subunit are mainly responsible. These β-strands are linked at the top to three short, highly flexible loops. The flexibility of the loops accounts for the fact that NGF assumes different structures in different crystal preparations. At the other end of the NGF molecule are three neighbouring disulphide bridges with a characteristic topology. Two of these disulphide bridges (Cys58–Cys108 and Cys68–Cys110) form a ring-like structure through which the third disulphide bridge (Cys15–Cys80) passes and forms a tightly packed 'cystine knot' motif. NGF was the first protein shown to have a cystine knot structure. The 'cystine knot' motif of the NGF monomer is reconstructed and shown in (b). (b) The 'cystine knot' arrangement allows the two pairs of β-strands from each subunit of NGF to pack against each other, generating an extensive interface in the dimer. The fact that mainly aromatic amino acids, capable of hydrophobic interactions, are at the interface explains the tight association of the subunits.[86] The preservation of these structural features in neurotrophins explains why they can readily form mixed heterodimers with each other.

It was suggested that GDNF associates first with the co-receptor, GDNFR-α, and the GDNF–GDNFR-α ligand–co-receptor complex is then presented to the RET receptor. RET receptors are tyrosine kinases. The RET receptor is the product of the c-ret proto-oncogene. (The name was given to a transforming human oncogene, ret, that originated from a T cell lymphoma DNA). A single point mutation in the tyrosine kinase domain is a gain-of-function mutation and results in deregulated activation of the RET receptor

tyrosine kinase, often found in humans with a predisposition to endocrine cancers, such as thyroid cancers. Loss-of-function mutations were found in Hirschsprung's disease, a developmental disorder with defective intestinal innervation. Moreover, somatic rearrangements of the *ret* proto-oncogene have been found in about 60% of papillary thyroid cancers in children exposed to radiation in the Chernobyl catastrophe (Chapter 15).

Hormone receptors and orphan receptors

The insulin receptor will be discussed here as a prototype of a receptor with tyrosine kinase activity which transmits a hormonal signal. The other major classes of receptors which transmit hormonal signals are the heptahelical, G-protein-coupled receptors and the nuclear receptors for steroid and non-steroid hormones (see Fig. 1.1); they are discussed later. The receptor for the peptide hormone insulin is one of the most extensively studied RTKs (receptors with tyrosine kinase activity).

RTKs that interact with peptide hormones

The insulin receptor

The insulin receptor is the prototype of a tyrosine kinase receptor with a constitutive oligomeric structure. The signalling form is a $\alpha 2/\beta 2$ tetramer, but the unit structure is a heterodimer, consisting of α- and β-subunits. But, in contrast to the homo-and heterodimers of other RTK growth factor receptors, the formation of the insulin receptor oligomer is constitutive and not dependent on ligand binding (see Fig. 1.5).

The α-chain and the β-chain of the receptor are linked by disulphide bridges. The α-chains are exclusively extracellular, whereas the β-chains have extracellular-, transmembrane-, and cytosolic domains. The extracellular domain of the α-subunit contains the ligand-binding site. Binding of insulin leads to autophosphorylation of the β-chains. Three groups of phosphotyrosyls are crucial for signalling. They are in the cytosolic part of the β-chains, in their carboxy-terminal tail, and in a region close to the membrane. When Y960, located close to the membrane, was replaced, the mutant receptor was inactive, although insulin could still stimulate the receptor tyrosine kinase activity. This is another impressive example of the specific role of a single tyrosyl phosphate for receptor signalling (see refs 89–92).

A central question is, of course, the relationship of the tyrosine kinase activity of the receptor to the action of the ligand, the hormone, insulin.[93] This will be discussed in Chapter 8.

Receptors that share similarities with the insulin receptor

Homologues of the mammalian insulin receptor, IR, are the DAF-2 receptor in Caenorhabditis elegans[94] and the IR-like receptor in Drosophila.[95] The DAF-2 receptor shares 35% of its amino-acid sequence with the human insulin receptor and 34% with the insulin-like growth factor receptor-1. The tyrosine kinase domain of the DAF-2 receptor is 70% similar and 50% identical to the tyrosine kinase domain of the human insulin receptor. A ligand for DAF-2 has not yet been identified, but an insulin-like peptide is anticipated. Since a typical insulin receptor substrate, like IRS-1 or IRS-2, has not been found in C. elegans, it is assumed that a COOH-terminal extension of the DAF-2 receptor serves as a built-in receptor substrate, which when phosphorylated, helps to recruit signalling proteins.

The close relationship of the human insulin receptor with the IR-like receptors in C. elegans and Drosophila indicates a common origin of these receptors, going back about 700 million years, to a time before invertebrates and mammals diverged. The fascinating similarities of signalling through the DAF receptor in nematodes and the insulin receptor in humans are discussed in Chapter 8.

The hepatocyte growth factor (HGF) receptor also belongs to the group of insulin-like receptors. The ligand for this receptor, HGF, is identical with the scatter factor, SF. This factor has been named 'scatter factor' because it dissociates sheets of epithelial cells and stimulates their motility. Binding of the factor activates the receptor tyrosine kinase and leads to autophosphorylation and association with a variety of signalling proteins, opening the door to different signalling pathways.

HGF is essential for the development of epithelia.[96,97] It stimulates cell proliferation, motility, and induction of cell polarity. Genetic knock-out of either the ligand or the HGF-receptor is lethal for the embryo. It is not surprising, therefore, that deregulation and overexpression of the *met* gene, which encodes the HGF receptor, can cause invasive growth of epithelial cells, contributing to the metastatic properties of cancer cells. Germ-line and somatic mutations of the *met* gene were found in patients with kidney carcinoma. The met gene was first identified in cells treated with the mutagen N-methyl-N-nitrosoguanidine.

A receptor more distantly related to this group of receptors is the bacterial (Escherichia coli) aspartate receptor, which promotes movement of the bacteria and chemotaxis in response to the attractant, aspartate. The aspartate receptor is not a tyrosine kinase (see also Chapter 2).

RTKs in search of a ligand

Finally, one should keep in mind that there exists a large family of RTKs, the so called Eph-like receptor tyrosine kinases, for which the ligands (the 'Ephs') are not well defined (*eph* is a gene, named from an erythropoietin-producing hepatocellular carcinoma cell line). At least seven genes have been identified which code for these receptors, and more can be expected (cf. ref. 98). Eph-like receptor tyrosine kinases interact with unusual membrane-associated and membrane-spanning molecules, the 'ephrins', which appear to be involved in cell–cell signalling.[99] An Eph-like receptor, the Axl receptor tyrosine kinase, is related to the insulin receptor and has been implicated in axonal pathfinding.[100]

The G-protein-coupled receptors

The G-protein-coupled heptahelical receptors are the largest transmembrane receptor class. They may take up 2% of the genome. Altogether, there may be thousands of different receptor molecules of this type that all transmit their signals through a heterotrimeric GTP-binding protein. These receptors are dealt with in Chapter 5.

The receptors for TGF-β and the cytokine receptors

The receptors for TGF-β and cytokines either are Ser/Thr kinases or have no kinase activities. They are considered in Chapter 6.

Conclusion

In this chapter growth factors and receptors have been surveyed and processing of ligands and receptors has been discussed. The first step in cellular signalling, the recognition of extracellular signals by membrane-bound sensors and the autophosphorylation of the

tyrosine kinase receptor upon ligand binding and activation, has been introduced, and it has been shown that the introduction of a single phosphate group, esterified with the OH-group of a tyrosine on a strategic position on the recognition surface of the receptor, determines which partner is chosen for coupling. The recruitment of the partner is specific. This event may determine the route the signal travels. The structural basis for protein–protein recognition will be discussed in Chapter 3.

In Chapter 2, I shall explain a general principle of activation of membrane single-pass tyrosine kinase receptors—receptor dimerization.

References

1. A. Ullrich and J. Schlessinger. Signal transduction by receptors with tyrosine kinase activity. *Cell*, **61** (2), 203–212, 1990.
2. P. De Camilli, K. Takei, and P. S. McPherson. The function of dynamin in endocytosis. *Curr Opin Neurobiol*, **5** (5), 559–565, 1995.
3. Z. Werb and Y. Yan. A cellular striptease act [comment]. *Science* **282**, 1279–1280, 1998.
4. J. Arribas, F. Lopez Casillas, and J. Massagué. Role of the juxtamembrane domains of the transforming growth factor-alpha precursor and the beta-amyloid precursor protein in regulated ectodomain shedding. *J Biol Chem*, **272**, 17160–17165, 1997.
5. M. Suzuki, G. Raab, M. A. Moses, C. A. Fernandez, and M. Klagsbrun. Matrix metalloproteinase-3 releases active heparin-binding EGF-like growth factor by cleavage at a specific juxtamembrane site. *J Biol Chem*, **272**, 31730–31737, 1997.
6. L. Lum and C. P. Blobel. Evidence for distinct serine protease activities with a potential role in processing the sperm protein fertilin. *Dev Biol*, **191**, 131–145, 1997.
7. R. A. Black, C. T. Rauch, C. J. Kozlosky, J. J. Peschon, J. L. Slack, M. F. Wolfson, *et al.* A metalloproteinase disintegrin that releases tumour-necrosis factor-alpha from cells. *Nature*, **385** (6618), 729–733, 1997.
8. K. Maskos, C. Fernandez Catalan, R. Huber, G. P. Bourenkov, H. Bartunik, G. A. Ellestad, *et al.* Crystal structure of the catalytic domain of human tumor necrosis factor-alpha-converting enzyme. *Proc Natl Acad Sci, USA*, **95**, 3408–3412, 1998.
9. C. P. Blobel. Metalloprotease-disintegrins: links to cell adhesion and cleavage of TNF alpha and Notch. *Cell*, **90** (4), 589–592, 1997.
10. Z. Werb. ECM and cell surface proteolysis: regulating cellular ecology. *Cell*, **91** (4), 439–442, 1997.
11. A. Amour, P. M. Slocombe, A. Webster, M. Butler, C. G. Knight, B. J. Smith, *et al.* TNF-alpha converting enzyme (TACE) is inhibited by TIMP-3. *FEBS Lett*, **435** (1), 39–44, 1998.
12. J. J. Peschon, J. L. Slack, P. Reddy, K. L. Stocking, S. W. Sunnarborg, D. C. Lee, *et al.* An essential role for ectodomain shedding in mammalian development [see comments]. *Science*, **282**, 1281–1284, 1998.
13. P. J. Miettinen, J. E. Berger, J. Meneses, Y. Phung, R. A. Pedersen, Z. Werb, and R. Derynck. Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. *Nature*, **376** (6538), 337–341, 1995.
14. R. Sadhukhan, G. C. Sen, R. Ramchandran, and I. Sen. The distal ectodomain of angiotensin-converting enzyme regulates its cleavage-secretion from the cell surface. *Proc Natl Acad Sci, USA*, **95** (1), 138–143, 1998.
15. E.J.M. Helmreich and E. L.Elson. Mobility of proteins and lipids in membranes. In *Advances in cyclic nucleotide and protein phosphorylation research*, Vol.18, (ed. P. Greengard and G.A. Robison). Raven Press, New York, pp. 1–61, 1984.
16. A. Lochter, S. Galosy, J. Muschler, N. Freedman, Z. Werb, and M. J. Bissell. Matrix metalloproteinase stromelysin-1 triggers a cascade of molecular alterations that leads to stable epithelial-to-mesenchymal conversion and a premalignant phenotype in mammary epithelial cells. *J Cell Biol*, **139** (7), 1861–1872, 1997.
17. D. Pan and G. M. Rubin. '*Kuzbanian*' controls proteolytic processing of Notch and mediates lateral inhibition during *Drosophila* and vertebrate neurogenesis. *Cell*, **90** (2), 271–280, 1997.
18. M. Tanaka, T. Itai, M. Adachi, and S. Nagata. Downregulation of Fas ligand by shedding [see comments]. *Nature Med*, **41**, 31–36, 1998.
19. J. Massagué and A. Pandiella. Membrane-anchored growth factors. *Annu Rev Biochem*, **62**, 515–541, 1993.

20. J. Massagué. Transforming growth factor-alpha. A model for membrane-anchored growth factors. *J Biol Chem*, **265** (35), 21393–21396, 1990.

21. R. Derynck, A. B. Roberts, M. E. Winkler, E. Y. Chen, and D. V. Goeddel. Human transforming growth factor-alpha: precursor structure and expression in *E. coli*. *Cell*, **38** (1), 287–297, 1984.

22. J. Teixido and J. Massagué. Structural properties of a soluble bioactive precursor for transforming growth factor-alpha. *J Biol Chem*, **263** (8), 3924–3929, 1988.

23. M. W. Bosenberg, A. Pandiella, and J. Massagué. The cytoplasmic carboxy-terminal amino acid specifies cleavage of membrane TGF alpha into soluble growth factor. *Cell*, **71** (7), 1157–1165, 1992.

24. S. Suzu, T. Ohtsuki, N. Yanai, Z. Takatsu, T. Kawashima, F. Takaku, *et al.* Identification of a high molecular weight macrophage colony-stimulating factor as a glycosaminoglycan-containing species. *J Biol Chem*, **267** (7), 4345–4348, 1992.

25. L. K. Price, H. U. Choi, L. Rosenberg, and E. R. Stanley. The predominant form of secreted colony stimulating factor-1 is a proteoglycan. *J Biol Chem*, **267** (4), 2190–2199, 1992.

26. B. Beutler and A. Cerami. The biology of cachectin/TNF—a primary mediator of the host response. *Annu Rev Immunol*, **7**, 625–655, 1989.

27. D. M. Jue, B. Sherry, C. Luedke, K. R. Manogue, and A. Cerami. Processing of newly synthesized cachectin/tumor necrosis factor in endotoxin-stimulated macrophages. *Biochemistry*, **29** (36), 8371–8377, 1990.

28. K. G. Peters, C. De Vries, and L. T. Williams.Vascular endothelial growth factor receptor expression during embryogenesis and tissue repair suggests a role in endothelial differentiation and blood vessel growth. *Proc Natl Acad Sci, USA*, **90** (19), 8915–8919, 1993.

29. M. Jeltsch, A. Kaipainen, V. Joukov, X. Meng, M. Lakso, H. Rauvala, *et al.* Hyperplasia of lymphatic vessels in VEGF-C transgenic mice. *Science*, **276** (5317), 1423–1425, 1997. [Published erratum *Science*, **277** (5325), 463, 1997.]

30. N. Weidner, J. P. Semple, W. R. Welch, and J. Folkman. Tumour angiogenesis and metastasis; correlation in invasive breast carcinoma. *N Engl J Med*, **324**, 1–8, 1991.

31. D. F. Steiner, W. Kemmler, J. L. Clark, P. E. Oyer, and A. H. Rubenstein. The biosynthesis of insulin. In *Endocrine pancreas*, Am. Physiol. Soc. Handbook of Physiology, Sect. 7, Vol. 1, (ed. D. F. Steiner and N. Freinkel). American Physiological Society, Washington, DC, 1972.

32. O. Gursky, J. Badger, Y. Li, and D. L. D. Caspar. Conformational changes in cubic insulin crystals. *Biophys J*, **63**, 1210–1220, 1992.

33. F.D. Steiner and A. H. Rubenstein. Proinsulin C-peptide-biological activity? *Science*, **277**, 531–532, 1997.

34. W. J. Fantl, D. E. Johnson, and L. T. Williams. Signalling by receptor tyrosine kinases. *Annu Rev Biochem*, **62**, 453–481, 1993.

35. G. Carpenter and S. Cohen. Epidermal growth factor. *J Biol Chem*, **265** (14), 7709–7712, 1990.

36. C. G. Davis. The many faces of epidermal growth factor repeats. *New Biol*, **2** (5), 410–419, 1990.

37. H. Marquardt, M. W. Hunkapiller, L. E. Hood, and G. J. Todaro. Rat transforming growth factor type 1: structure and relation to epidermal growth factor. *Science*, **223** (4640), 1079–1082, 1984.

38. S. Higashiyama, K. Lau, G. E. Besner, J. A. Abraham, and M. Klagsbrun. Structure of heparin-binding EGF-like growth factor. Multiple forms, primary structure, and glycosylation of the mature protein. *J Biol Chem*, **267** (9), 6205–6212, 1992.

39. B. J. Rutledge, K. Zhang, E. Bier, Y. N. Jan, and N. Perrimon.The *Drosophila* spitz gene encodes a putative EGF-like growth factor involved in dorsal–ventral axis formation and neurogenesis. *Genes Dev*, **6** (8), 1503–1517, 1992.

40. R. J. Hill and P. W. Sternberg. The gene *lin-3* encodes an inductive signal for vulval development in *C. elegans* [see comments]. *Nature*, **358** (6386), 470–476, 1992.

41. G. M. Walton, W. S. Chen, M. G. Rosenfeld, and G. N. Gill. Analysis of deletions of the carboxyl terminus of the epidermal growth factor receptor reveals self-phosphorylation at tyrosine 992 and enhanced *in vivo* tyrosine phosphorylation of cell substrates. *J Biol Chem*, **265** (3), 1750–1754, 1990.

42. C. H. Heldin and B. Westermark. Platelet-derived growth factor: mechanism of action and possible *in vivo* function. *Cell Regul*, **1** (8), 555–566, 1990.

43. A. Lemmon and J. Schlessinger. Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem Sci*, **19**, 151–155, 1994.

44. L. Ronnstrand, S. Mori, A. K. Arridsson, A. Eriksson, C. Wernstedt, U. Hellman, *et al.* Identification of two C-terminal autophosphorylation sites in the PDGF beta-receptor: involvement in the interaction with phospholipase C-gamma. *EMBO J*, **11** (11), 3911–3919, 1992.

45. A. Kashishian and J. A. Cooper. Phosphorylation sites at the C-terminus of the platelet-derived growth factor receptor bind phospholipase C gamma 1. *Mol Biol Cell*, **41**, 49–57, 1993.

46. Q. C. Vega, C. Cochet, O. Filhol, C. P. Chang, S. G. Rhee, and G. N. Gill. A site of tyrosine phosphorylation in the C terminus of the epidermal growth factor receptor is required to activate phospholipase C. *Mol Cell Biol*, **12** (1), 128–135, 1992.

47. H. K. Kim, J. W. Kim, A. Zilberstein, B. Margolis, J. G. Kim, J. Schlessinger, and S. G. Rhee. PDGF stimulation of inositol phospholipid hydrolysis requires PLC-gamma 1 phosphorylation on tyrosine residues 783 and 1254. *Cell*, **65** (3), 435–441, 1991.

48. M. Valius and A. Kazlauskas. Phospholipase C-gamma 1 and phosphatidylinositol 3 kinase are the downstream mediators of the PDGF receptor's mitogenic signal. *Cell*, **73** (2), 321–334, 1993.

49. A. Eriksson, A. Siegbahn, B. Westermark, C. H. Heldin, and L. Claesson Welsh. PDGF alpha- and beta-receptors activate unique and common signal transduction pathways. *EMBO J*, **11** (2), 543–550, 1992.

50. E. Johnson, P. L. Lee, J. Lu, and L. T. Williams. Diverse forms of a receptor for acidic and basic fibroblast growth factors. *Mol Cell Biol*, **10** (9), 4728–4736, 1990.

51. P. W. Finch, J. S. Rubin, T. Miki, D. Ron, and S. A. Aaronson. Human KGF is FGF-related with properties of a paracrine effector of epithelial cell growth. *Science*, **245** (4919), 752–755, 1989.

52. D. Gospodarowicz, G. Neufeld, and L. Schweigerer. Molecular and biological characterization of fibroblast growth factor, an angiogenic factor which also controls the proliferation and differentiation of mesoderm and neuroectoderm derived cells. *Cell Differ*, **19** (1), 1–17, 1986.

53. J. Folkman and M. Klagsbrun. Angiogenic factors. *Science*, **235** (4787), 442–447, 1987.

54. S. Werner, D. S. Duan, C. de Vries, K. G. Peters, D. E. Johnson, and L. T. Williams. Differential splicing in the extracellular region of fibroblast growth factor receptor 1 generates receptor variants with different ligand-binding specificities. *Mol Cell Biol*, **12** (1), 82–88, 1992.

55. D. E. Johnson and L. T. Williams. Structural and functional diversity in the FGF receptor multigene family. *Adv Cancer Res*, **60**, 1–41, 1993.

56. H. H. Reid, A. F. Wilks, and O. Bernard. Two forms of the basic fibroblast growth factor receptor-like mRNA are expressed in the developing mouse brain. *Proc Natl Acad Sci, USA*, **87** (4), 1596–1600, 1990.

57. C. A. Dionne, G. Crumley, F. Bellot, J. M. Kaplow, G. Searfoss, M. Ruta, *et al.* Cloning and expression of two distinct high-affinity receptors cross-reacting with acidic and basic fibroblast growth factors. *EMBO J*, **9** (9), 2685–2692, 1990.

58. F. Bellot, G. Crumley, J. M. Kaplow, J. Schlessinger, M. Jaye, and C. A. Dionne. Ligand-induced transphosphorylation between different FGF receptors. *EMBO J*, **10** (10), 2849–2854, 1991.

59. W. H. Burgess and T. Maciag. The heparin-binding (fibroblast) growth factor family of proteins. *Annu Rev Biochem*, **58**, 575–606, 1989.

60. S. Faham, R. E. Hileman, J. R. Fromm, R. J. Linhardt, and D. C. Rees. Heparin structure and interactions with basic fibroblast growth factor. *Science*, **271** (5252), 1116–1120, 1996.

61. P. Zimmermann and G. David. The syndecans, tuners of transmembrane signalling. *FASEB J*, **13** (Suppl.), 91–100, 1999.

62. M. Kato, H. Wang, V. Kainulainen, M. L. Fitzgerald, S. Ledbetter, D. M. Ornitz, and M. Bernfield. Physiological degradation converts the soluble syndecan-1 ectodomain from an inhibitor to a potent activator of FGF-2. *Nature Med*, **4** (6), 691–697, 1998.

63. N. Perrimon and M. Bernfield. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature*, **404**, 725–728, 2000.

64. M. Mohammadi, C. A. Dionne, W. Li, N. Li, T. Spivak, A. M. Honegger, *et al.* Point mutation in FGF receptor eliminates phosphatidylinositol hydrolysis without affecting mitogenesis. *Nature*, **358** (6388), 681–684, 1992.

65. R. Rottapel, M. Reedijk, D. E. Williams, S. D. Lyman, D. M. Anderson, T. Pawson, and A. Bernstein.The Steel/W transduction pathway: kit autophosphorylation and its association with a unique subset of cytoplasmic signaling proteins is induced by the Steel factor. *Mol Cell Biol*, **11** (6), 3043–3051, 1991.

66. R. Herbst, M. S. Shearman, A. Obermeier, J. Schlessinger, and A. Ullrich. Differential effects of W mutations on p145c-kit tyrosine kinase activity and substrate interaction. *J Biol Chem*, **267** (19), 13210–13216, 1992.

67. A. D. Reith, C. Ellis, S. D. Lyman, D. M. Anderson, D. E. Williams, A. Bernstein, and T. Pawson. Signal transduction by normal isoforms and W mutant variants of the Kit receptor tyrosine kinase. *EMBO J*, **10** (9), 2451–2459, 1991.

68. K. Nocka, J. Buck, E. Levi, and P. Besmer. Candidate ligand for the c-kit transmembrane kinase receptor: KL, a fibroblast derived growth factor stimulates mast cells and erythroid progenitors. *EMBO J.* **9** (10), 3287–3294, 1990.

69. E. S. Russell. Hereditary anemias of the mouse: A review for geneticists. *Adv Genet.* **20**, 357–459, 1990.

70. J. G. Flanagan, D. C. Chan, and P. Leder. Transmembrane form of the kit ligand growth factor is determined by alternative splicing and is missing in the Sld mutant. *Cell*, **64** (5), 1025–1035, 1991.

71. E. J. Huang, K. H. Nocka, J. Buck, and P. Besmer. Differential expression and processing of two cell associated forms of the kit-ligand: KL-1 and KL-2. *Mol Biol Cell*, **3** (3), 349–362, 1992.

72. L. F. Lin, D. H. Doherty, J. D. Lile, S. Bektesh, and F. Collins. GDNF: a glial cell line-derived neurotrophic factor for midbrain dopaminergic neurons [see comments]. *Science*, **260** (5), 1130–1132, 1993.

73. P.T. Kotzbauer, P. A. Lampe, R. O. Heuckeroth, J. P. Golden, D. J. Creedon, E. M. Johnson Jr, and J. Milbrandt. Neurturin, a relative of glial-cell-line-derived neurotrophic factor. *Nature*, **384**, 467–470, 1996.

74. D. S. Middlemas, R. A. Lindberg, and T. Hunter. trkB, a neural receptor protein-tyrosine kinase: evidence for a full-length and two truncated receptors. *Mol Cell Biol*, **11** (1), 143–153, 1991.

75. R. Klein, D. Conway, L. F. Parada, and M. Barbacid. The trkB tyrosine protein kinase gene codes for a second neurogenic receptor that lacks the catalytic kinase domain. *Cell*, **61** (4), 647–656, 1990.

76. F. Lamballe, R. Klein, and M. Barbacid. trkC, a new member of the trk family of tyrosine protein kinases, is a receptor for neurotrophin-3. *Cell*, **66** (5), 967–979, 1991.

77. Q. N. McDonald and M. V. Chao Structural determinants of Neurotrophin action. *J Biol Chem*, **270**, 19669–19672, 1995.

78. B.D.V. Bax, T. L. Blundell, J. Murray-Rust, and N. Q. McDonald. Structure of mouse NGF: A complex of nerve growth factor with four binding proteins. *Structure* (London), **5**, 1275, 1997.

79. P. D. Sun and D. R. Davies. The cystine-knot growth-factor superfamily. *Annu Rev Biophys Biomol Struct*, **24**, 269–291, 1995.

80. L. A. Tartaglia and D. V. Goeddel. Two TNF receptors. *Immunol Today*, **13** (5), 151–153, 1992.

81. S. Rabizadeh and D. E. Bredesen. Is p75NGFR involved in developmental neural cell death? *Dev Neurosci*, **16** (3–4), 207–211, 1994.

82. B. L. Hempstead, D. Martin Zanca, D. R. Kaplan, L. F. Parada, and M. V. Chao. High-affinity NGF binding requires coexpression of the trk proto-oncogene and the low-affinity NGF receptor [see comments]. *Nature*, **350** (6320), 678–683, 1991.

83. J. M. Verdi, S. J. Birren, C. F. Ibanez, H. Persson, D. R. Kaplan, M. Benedetti, *et al.* p75NGFR regulates Trk signal transduction and NGF-induced neuronal differentiation in Mah cells. *Neuron*, **12** (4), 733–745, 1994.

84. J. Masssagué. Neurotrophic factors: Crossing receptor boundaries, News and views. *Nature*, **382**, 29–30, 1996.

85. J. J. S. Treanor, L. Goodman, F. de Sauvage, D. M. Stone, K. T. Poulsen, C. D. Beck, *et al.* Characterization of a multicomponent receptor for GDNF. *Nature*, **382**, 80–83, 1996.

86. S. Jing, D. Wen, Y. Yu, P. L. Holst, Y. Luo, M. Fang, *et al.* GDNF-induced activation of the ret protein tyrosine kinase is mediated by GDNFR-alpha, a novel receptor for GDNF. *Cell*, **85** (7), 1113–1124, 1996.

87. P. Durbec, C. V. Marcos-Guitierrez, C. Kilkenny, M. Grigoriou, K. Wartiowaara, P. Suvanto, *et al.* GDNF signalling through the Ret receptor tyrosine kinase. *Nature*, **381**, 789–793, 1996.

88. M. Trupp, E. Arenas, M. Fainzilber, A.-S. Nisson, B.-A. Sieber, M. Grigoriou, *et al.* Functional receptor for GDNF encoded by the c-ret proto-oncogene. *Nature*, **381**, 785–789, 1996.

89. M. F. White, J. N. Livingston, J. M. Backer, V. Lauris, T. J. Dull, A. Ullrich, and C. R. Kahn. Mutation of the insulin receptor at tyrosine 960 inhibits signal transmission but does not affect its tyrosine kinase activity. *Cell*, **54** (5), 641–649, 1988.

90. M. F. White and C. R. Kahn. Cascade of autophosphorylation in the beta-subunit of the insulin receptor. *J Cell Biochem*, **39** (4), 429–441, 1989.

91. P. A. Wilden, J. M. Backer, C. R. Kahn, D. A. Cahill, G. J. Schroeder, and M. F. White. The insulin receptor with phenylalanine replacing tyrosine-1146 provides evidence for separate signals regulating cellular metabolism and growth. *Proc Natl Acad Sci, USA*, **87** (9), 3358–3362, 1990.

92. M. S. Murakami and O. M. Rosen. The role of insulin receptor autophosphorylation in signal transduction. *J. Biol. Chem.* **266** (33), 22653–22660, 1991.

93. C. R. Kahn and M. F. White. The insulin receptor and the molecular mechanism of insulin action. *J. Clin. Invest.*, **82** (4), 1151–1156, 1988.

94. K. D. Kimura, H. A. Tissenbaum, Y. Liu, and G. Ruvkun. *daf-2,* an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science*, **277**, 942–946, 1997.

95. R. Fernandez, D. Tabarini, N. Azpiazu, M. Frasch, and J. Schlessinger. The *Drosophila* insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential. *EMBO J.* **14** (14), 3373–3384, 1995.

96. C. Boccaccio, M. Ando, L. Tamagnone, A. Bardelli, P. Michieli, C. Battistini, and P. M. Comoglio. Induction of epithelial tubules by growth factor HGF depends on the STAT pathway. *Nature*, **391** (6664), 285–288, 1998.

97. P. Comoglio. In *Hepatocyte growth factor-scatter factor (HGF-SF) and the c-MET receptor*, (ed. I. D. Goldberg and E. M. Rosen). Birkhäuser Verlag, Basel, 1992.

98. H. Hirai, Y. Maru, K. Hagiwara, J. Nishida, and F. Takaku. A novel putative tyrosine kinase receptor encoded by the eph gene. *Science*, **238** (4834), 1717–1720, 1987.

99. K. Brückner, E. B. Pasquale, and R. Klein. Tyrosine phosphorylation of transmembrane ligands for Eph receptors. *Science*, **275**, 1640–1643, 1997.

100. J. O'Bryan, R. A. Frye, P. C. Cogswell, A. Neubauer, B. Kitch, C. Prokop, *et al.* Axl, a transforming gene isolated from primary human myeloid leukemia cells, encodes a novel receptor tyrosine kinase. *Mol Cell Biol*, **11** (10), 5016–5031, 1991.

2

Activation of receptors by oligomerization

Binding of the ligand leads to the activation of a membrane-bound receptor and initiates the cellular response to hormones, growth factors, and cytokines. The structural basis of ligand binding and receptor activation is therefore a central problem of cellular signalling. The question is how a single polypeptide chain which passes through the membrane only once, is activated on binding the ligand. The answer is by dimerization.

Dimerization is not restricted to receptors. It is a property of many regulatory proteins and enzymes. Homo- and heterodimerization occur in many signal transmitters and transcriptional activators, as we shall see later. Out of many examples of receptor dimerization, I have selected the growth hormone receptor and closely related receptors, because X-ray crystallographers have provided detailed information on the structural basis of ligand-induced oligomerization.

Growth hormone (GH) is a cytokine, produced in the pituitary gland. It controls the growth of the long bones in the body and determines its size. Human growth hormone (hGH) is used for the treatment of children with stunted growth. Since the gene for hGH has been cloned and expressed, hGH can now be produced by gene technology. This has provided safe and abundant material for therapy. Before this breakthrough, hGH had to be taken from the deceased; it was rare and sometimes contaminated with infectious agents.

The role of receptor dimerization in receptor activation

Studies on receptor mutants, incapable of dimerization, have confirmed that receptor dimerization of single-chain receptors is essential for signalling and the biological response *in vivo*.

Complexes of the hGH receptor and related receptors with their ligands

hGH is a polypeptide that can bind to either the growth hormone or the prolactin (PRL) receptor. The hGH and hPRL receptors both have an extracellular hormone binding

domain of 250 residues. Following the extracellular domain is a single transmembrane spanning region of only 25 residues and an intracellular domain of varying length, with little similarity with other single-chain receptors.

Dimerization of the GH receptor is strictly ligand dependent. The unliganded receptor does not dimerize, even at very high (>100 mM) concentrations. If the ligand is a single-chain monomer, it binds two copies of the same receptor chain, forming an active 1 : 2 ligand–receptor complex; and when the ligand is a dimer, each monomer likewise binds two receptor chains, forming a 2 : 4 ligand–receptor complex.

Kossiakoff's group[1,2] has determined the structures of five ligand–receptor complexes, using a separated extracellular hormone-binding domain, ECD. The separate extracellular domain of the hGH receptor is named hGHbp, human growth hormone binding protein. The ligand, hGH, has two distinct sites, each binding one copy of the ECD of the receptor (see Plate 1).

Three-dimensional structure

hGHbp contains two immunoglobulin-like domains with β-sheet structure, whereas the ligand, the hGH, is a bundle of four helices, structurally similar to other cytokines. hGH binds at the interface of the two β-sheets of the receptor. The surface of site 1 of the first receptor molecule, and to a lesser degree the surface of site 2 on the second receptor molecule, is concave with respect to the face of the binding partner. The site 1 area of the receptor, which is buried on binding of hGH, is considerably larger than the site 2 area: ~1300 Å^2 compared with 800 Å^2 .This difference is compensated by an area of 500 Å^2 that is buried between the two receptor hGHbps, on binding of hGH to the β-sheet domain.

It is remarkable that the amino-acid side-chains of the ligand, hGH, that bind to the first and those that bind to the second receptor molecule are entirely different, although the corresponding binding sites at each of the receptor chains are virtually identical. Thus, the same amino-acid side-chains of each receptor molecule bind to very different side-chains of the ligand molecule. Accordingly, the conformations of the two receptor-binding sites which bind to the two different hormone-binding sites are also only slightly different and the changes in the receptor structure on binding the hormone involve movements of less than 4 Å. It is also remarkable that the structures of the free and bound hormone are also nearly identical, indicating that large conformational adjustments of the ligand do not occur on binding to the receptor.

hGH also binds to the prolactin (hPRL) receptor, but in this case additional binding interactions are required, which are contributed by a Zn^{2+}. Both hGH and the prolactin receptor contribute to the zinc-binding site. Structural studies of a complex of hGH and the extracellular domain of the hPRL receptor have shown that the structure of the first site of the hPRL receptor to which the hormone binds is virtually the same as that in the hGHbp. This is an example of two different receptors sharing the same contact surface.

In Plate 1, ref: 2, 3, the crystal structures of hGH complexed with the extracellular domain of its receptor, referred to as hGH-binding protein, and of hGH complexed with the prolactin receptor are compared. Moreover, the effects of mutations of the hormone and the receptor on ligand binding are shown.

Another receptor of great medical interest is the erythropoietin (EPO) receptor. EBP is the binding domain of the receptor for erythropoietin. EPO is a cytokine with an important function in the formation of red blood cells.[3] (see Chapter 6). EPO has been cloned and expressed and is now produced by gene technology. The structure of the EPO-

binding protein is similar to that of the GH- and the PRL-receptor binding domains. However, the two EBPs in the dimeric receptor complex have little contact with each other; the contact surface between them is only 75 Å^2, as compared with 500 Å^2 in the case of the dimeric hGHbp. Thus, the dimer assembly differs from that of the hGH-receptor dimer. Apparently, there is more than one way to form a dimer, although the ligand binds to the same kind of hydrophobic amino acids in both cases.

The medical importance of EPO stimulated a search for a smaller peptide, still capable of binding and activating the EPO receptor. EPO has 165 amino acids. A EPO peptide, only 20 amino acids long, was found to bind and activate the EPO receptor, although with a K_d of 200 nM, 1000-fold larger than that of the full-length EPO. The peptide did, however, activate EPO receptors on cells *in vivo,* apparently engaging the same signalling routes as the full-length wild-type EPO. The structure of the complex of this peptide with the EPO-binding domain, EBP, was solved. The EBP dimerizes, on binding two molecules of the EPO peptide, just like on binding wild-type EPO (Fig. 2.1).

The fact that the EPO-mimetic EMP1 is a peptide, unrelated in sequence and structure to the wild-type EPO, but still capable of inducing dimerization of the receptor and eliciting a biological effect is remarkable. This gives hope for the design of small peptides and even unrelated, non-peptidic molecules that mimic the natural hormone.[5] To find such

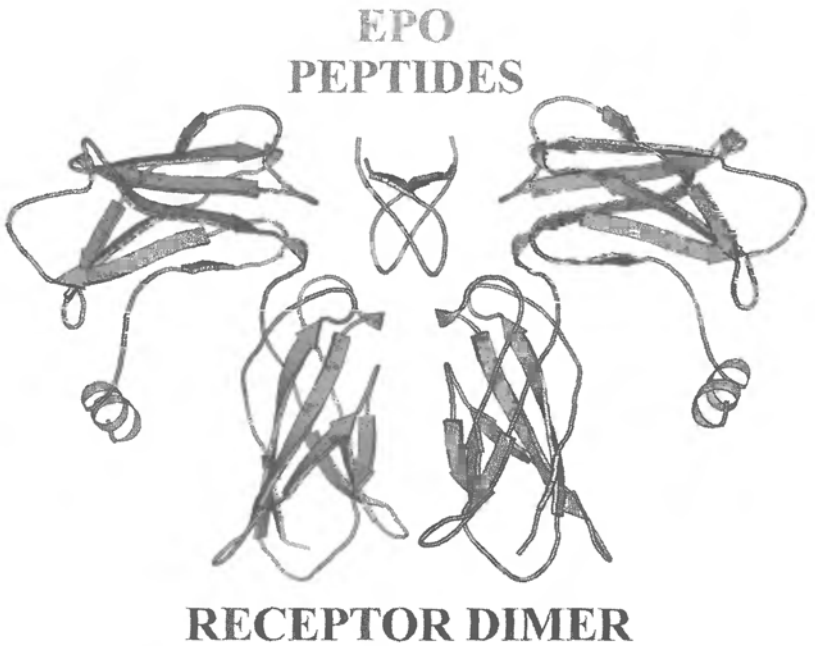


Fig. 2.1 Crystal structure of a complex of the extracellular domain of the human erythropoietin receptor (EBP) and an agonist peptide, the EPO-mimetic peptide 1 (EMP1). The EBP dimer is shown on the left- and right-hand sides of the diagram. The EMP is in the centre at the top. Only that part of the structure of the EMP is shown that is crucial for the interaction with the loops in each EBP molecule. (Reproduced with permission of the authors and Science from data in ref. 4 deposited in databanks.)

structures is a central issue in designing drugs today. Understanding the molecular details of the structural and chemical requirements for competent ligand–receptor interactions is the basis of success.

The core of the actual binding region is remarkably small

The contact interface of EMP1 with the receptor is considerably smaller than that of the wild-type EPO, leading to the conclusion that the functionally relevant binding region might be small. Screening of alanine mutants indicated that only about half of the 31 side-chains of hGH that become buried on binding to the receptor are directly involved in the binding interactions. The actual core of the binding region of the hormone is a small hydrophobic region with two tryptophans, surrounded by a few hydrophilic residues.[7] Consequently, only 8 of the 31 side-chains of the hormone determine the affinity of the hormone for site 1 of the first receptor molecule, but account for more than about 85% of the total binding energy. In the case of the receptor, the situation is similar and only 9 out of 33 side-chains participate in the binding reaction and are buried on binding the ligand, but these residues accounted for virtually all the binding affinity.

Clackson and Wells[6] have pointed to the remarkable fact that many side-chains in the binding domains of the hormone and the receptor do not contribute to binding, although they appear perfectly well capable of such interactions, since they could form good hydrogen bonds and salt bridges. Thus, it seems that the role of many side-chains at the interface of the hormone–receptor complex is to maintain the complex, rather than to form it in the first place. The conclusion is, that the ligand is guided to the receptor by diffusion and long-range electrostatic interactions. Once the hormone has found the receptor, a number of weak collisional complexes are formed, eventually leading to a single, defined hormone–receptor complex.

Thus, the 'functional epitope' is much smaller than the 'structural epitope', meaning that only a few residues determine specificity and affinity of binding.

Order of binding

A further aspect of the binding interaction is the order of binding. Binding of hGH through its site 1 to the first molecule of hGHbp is the first step, followed by binding of the hormone through site 2 to the second molecule of hGHbp.

James A. Wells[7] has rationalized why a step-by-step formation of a ligand receptor complex is more economical. He has reasoned that a sequential addition of the receptor requires only one encounter of the ligand with the receptor, involving three-dimensional diffusion of soluble hGH to a membrane-bound receptor. But, once the first ligand molecule is bound, the 1 : 1 hGH : receptor complex diffuses over a short distance laterally in the two-dimensional membrane to encounter a second, empty receptor, forming the final 1 : 2 ligand : receptor complex. The fact that the receptor is anchored in the membrane further limits the degrees of rotational freedom of the receptor and facilitates the ligand–receptor interaction and dimerization. The point is that formation of a 1 : 2 ligand : receptor complex, involving three-dimensional diffusion of the ligand to reach the receptor on two occasions, would require much higher concentrations of ligand.

TNF-receptor–ligand complexes

The interaction of hGH with its receptor is a paradigm of ligand binding, resulting in receptor dimerization. Other cytokines and growth factors form, likewise, dimeric and

even trimeric receptor complexes. The stepwise formation of receptor–ligand complexe is similar, except that in some cases one hormone molecule binds to three receptors, on after another. Examples are the complexes of TNF-β with the extracellular bindin; domains of the corresponding receptor, of which several structures have been solved. The structural evidence indicated that the receptor exists initially as a monomer in th membrane. It dimerizes without ligand, but it trimerizes apparently only upon ligan binding. The active ligand–receptor complex has three receptor molecules bound symmetrically to one TNF-β. The extracellular part of the receptor is an elongated arrangement of four similar folding domains, forming the characteristic fold of the TNF receptor which fits into a groove between two adjacent TNF-β subunits. This fold is likely to be representative of the nerve growth factor NGF/TNF receptor family as a whole.[9,10]

The structure of the ligand–receptor complex defines the orientation of the ligand with respect to the cell membrane. The total surface area buried in the TNF-β–$(TNFR)_3$ complex (1.120 Å^2) is about the same size as that in the hGH–$(hGHbp)_2$ complex. Although the quaternary structures of free and liganded TNF-β receptor are different, the structure and conformation of free and bound ligand, TNF-β, are, like free and bound hGH, virtually identical, as shown by Naismith *et al.*[9]

Because, most biological functions require the assembly of proteins in multicomponent complexes, a better understanding of the quantitative aspects of protein–protein interactions is important.

The lesson to be learned from the binding interaction of hGH with its receptor (hGHbp) has been summarized[11] by James A. Wells, as follows: binding is quite fast, the K_{on} is $3 \times 10^5\ s^{-1}\ M^{-1}$, but the ligand–receptor complex is fairly long lived. To rationalize the kinetics one must assume longer-lived collisions. Thus, when proteins collide they may not diffuse away immediately, but may stay together for a while. This may give them a chance to probe the surface and make contacts with a much larger surface area than would be the case if they were to bump into each other only once for a single collision. This could also explain why the 'structural epitope' participating in the formation of the initial complex is much larger than the 'functional epitope' selected in the final complex.

Summary

A striking property of receptors is their plasticity.[11] A receptor can recognize and accept structurally quite different ligands, and a given ligand may interact with different receptors.

James A. Wells has summarized the important events that are responsible for specific and high-affinity interactions between ligand and receptor. These events help to explain receptor plasticity. He concludes that:

1. Association of the hormone hGH and its receptor is controlled by diffusion.
2. The hormone finds the receptor by random, rapid collisions. The location of the receptor in the two-dimensional space of the plasma membrane facilitates collisional encounters.
3. The initial encounter may not involve the same protein surface or the same amino-acid side-chains which represent the final binding interface.

4. In the final hormone–receptor complex only a few side-chains determine the affinity of the hormone for the receptor. The most important contacts are made by hydrophobic side-chains, whereas polar and charged interactions are of less importance.
5. The major kinetic effect of these contacts is to slow the off-rate and keep the hormone bound, once it has reached its final binding site (in the case of the hGH–receptor complex, K_{off} is 2.7×10^{-4} s^{-1}).
6. Monomeric proteins which do not form oligomers have fewer hydrophobic interfaces than those forming oligomeric proteins. Thus, hydrophobic side-chains seem to favour oligomerization.
7. In protein oligomers, the subunits are more closely packed, with fewer planar interfaces and fewer intersubunit hydrogen bonds than would be the case if the same components were more fortuitously arranged in random aggregates.

Finally, to address the question: 'How obligatory is ligand-dependent dimerization for activation of single-chain receptors (passing the membrane only once)?'. The advantage of dimerization is presumably the presentation of more intracellular signalling domains, creating larger and more diverse interfaces for recognition of cytosolic transducer and linker molecules. Moreover, a ligand-induced monomer–dimer transition of a receptor is the simplest way to bring sensing and signalling domains together and have them communicate with each other in a cooperative manner.

While dimerization is certainly an attractive way to activate a single-chain receptor upon binding a ligand, it may not be the only way. For example, the receptor in *E. coli* for aspartate (TAR) is a typical, polarized chemotaxis receptor,[12] which does not dimerize on ligand (aspartate) binding. TAR is present either as a constitutive dimer, like the insulin receptor, or it signals as monomer. TAR has, like the hGH receptor, a soluble extracellular binding or sensing domain and an intracellular signalling domain. However, the X-ray structure of the sensing domain of the bacterial receptor is not at all like that of hGHR. Whereas the ligand-binding region of hGHR is composed of two 7-stranded β-sandwiches (see Plate 1), the sensor domain of TSR (the serine chemotaxis receptor, which is closely related to TAR) is a bundle of parallel α-helices.[13] The results of paramagnetic resonance spectroscopy of the liganded and unliganded receptor were in line with a small piston-like movement of the two transmembrane helices which, on binding the ligand, slide downward relative to each other by only 1 Å. Thus, TAR signalling results from conformational transitions within a signalling unit and does not require ligand-dependent receptor oligomerization. The 'piston model' proposed by Daniel E. Koshland Jr and colleagues[14] may apply to other types of receptors, for example G-protein-coupled, heptahelical receptors (see Chapter 5). Although certain types of G-protein-coupled, heptahelical receptors in the brain, such as the GABA (γ–aminobutyric acid) receptor and the opioid receptors, form heterodimers,[15] whether dimerization of heptahelical receptors is obligatory rather than permissive is not clear.

Finally, some mammalian receptors (certain NGF and interleukin receptors) are apparently capable of signalling with only one chain. However, in these cases, signalling seems to require the formation of heterodimers with accessory proteins, e.g. co-receptors, or interactions with other molecules, such as proteoglycans.[16] Furthermore, there are the insulin receptors, which are constitutive dimers with the monomers covalently linked by

intersubunit disulphide bonds. Although this arrangement excludes a ligand-induced monomer–dimer transition, a conformational change in the relative orientation of the monomers in a pre-existing dimer is possible and may be critical for activation. This kind of conformational transition may be the reason why receptors for insulin and insulin-like growth factors exhibit negative cooperativity in ligand binding, a property they share with TAR.

References

1. A. M. De Vos, M. Ultsch, and A. A. Kossiakoff. Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science*, **255**, 306, 1992.
2. W. Somers, M. Ultsch, A. M. De Vos, and A. A., Kossiakoff. The x-ray structure of a growth hormone–prolactin receptor complex. *Nature*, **372**, 478, 1994.
3. K. Maruyama, K. Miyata, and A. Yoshimura. Proliferation and erythroid differentiation through the cytoplasmic domain of the erythropoietin receptor. *J Biol Chem*, **269** (8), 5976–5980, 1994.
4. O. Livnah, E. A. Stura, D. L. Johnson, S. A. Middleton, L. S. Mulcahy, N. C. Wrighton, *et al.* Functional mimicry of a protein hormone by a peptide agonist: The EPO receptor complex at 2.8 Å. *Science*, **273**, 464, 1996.
5. N. C. Wrighton, F. X. Farrell, R. Chang, A. K. Kashyap, F. P. Barbone, L. S. Mulcahy, *et al.* Small peptides as potent mimetics of the protein hormone erythropoietin. *Science*, **273**, 458–463, 1996.
6. T. Clackson and J. A. Wells. A hot spot of binding energy in a hormone–receptor interface. *Science*, **267** (5196), 383–386, 1995.
7. J. A. Wells. Hormone mimicry. *Science*, **273**, 449–450, 1995.
8. D. W. Banner, A. Arcy, W. Janes, R. Gentz, H. Schoenfeld, C. Broger, *et al.* Crystal structure of the soluble human 55 kd TNF receptor–human TNF beta complex: Implications for TNF receptor activation. *Cell*, **73** (3), 431–435, 1993.
9. J. H. Naismith, T. Q. Devine, B. Brandhuber, and S. R. Sprang. Crystallographic evidence for dimerization of unliganded tumor necrosis factor receptor. *J Biol Chem*, **270**, 13303–13307, 1995.
10. L. E. Rodseth, B. Brandhuber, T.Q. Devine, M. J. Eck, K. Hale, J. H. Naissmith, and S. R. Sprang. Two crystal forms of the extracellular domain of type I tumour necrosis factor receptor. *J Mol Biol*, **239**, 332, 1994.
11. S. Atwell, M. Ultsch, A. M. De Vos, and J. A. Wells. Structural plasticity in a remodeled protein–protein interface. *Science*, **278**, 1125, 1997.
12. J. R. Maddock and L. Shapiro. Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science*, **259** (5102), 1717–1723, 1993.
13. K. K. Kim, H. Yokota, and S.-H. Kim. Four helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature*, **400**, 787–792, 1999.
14. K. M. Ottemann, W. Xiao, Y.-K. Shin, and D. E. Koshland Jr. A piston model for transmembrane signaling of the aspartate receptor. *Science*, **285**, 1751–1754, 1999.

    See also: Comments by M. Gerstein and C. Chotia. Proteins in motion. *Science*, **285**, 1682–1683, 1999.
15. B. A. Jordan and L. A. Devi. G-protein-coupled receptor heterodimerization modulates receptor function. *Nature*, **399**, 697–700, 1999.
16. J. Schlessinger, I. Lax, and M. Lemmon. Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell*, **83** (3), 357–360, 1995.

3

Components of signalling networks: linkers and regulators

Signal transduction cascades are built from many different components which are glued together by structural motifs, which are the 'putty', the 'cement', that holds the 'bricks' together. The phosphotyrosyls[1] and other markers on the surface of RTKs recruit linker or adaptor proteins which contain structural motifs, such as Src-(Rous sarcoma virus)-homology 2 (SH2), Src-homology 3 (SH3),[2] pleckstrin-homology (PH), phosphotyrosine-binding (PTB), and other domains.[3] With the help of these motifs, the adaptors link the receptor to regulatory proteins in the cell, building signalling networks. Adaptors and linkers are evolutionary conserved. Homologues are expressed in *Caenorhabditis elegans* and in *Drosophila*. Many linkers are without enzymatic activity, but some are cytosolic kinases and phosphatases or monomeric GTP-binding proteins and their regulators, GDP-exchange factor (GEFs), and GTPase-activating proteins (GAPs). These enzymatic linkers play a special role, because they not only transmit and control the receptor signals, they also amplify them.

Linkers without enzymatic activities

Figure 3.1 pictures a hypothetical adaptor protein, containing common motifs, PH, SH2, SH3, and PTB domains.

Each domain recognizes a different, structurally complementary region in the coupling partner. SH2 and PTB domains recognize phosphotyrosyls, SH3 domains polyprolyl sequences and PH domains bind to phospholipids and to the plasma membrane. Many adaptor proteins have more than one distinct structural module, each with a different binding preference.[4,5] This enables them to connect a variety of structurally and functionally different proteins, building a versatile signal transduction network, serving the needs of the cell. Common linkers such as Crk, Vav, and Shc were first identified as retroviral oncogenes and are named according to their retroviral source. Crk is the *CT*10 virus *r*egulator of *k*inase. The sequence of Crk is similar to that of phospholipase C, which generates lipid messengers and plays an important role in signal transduction. Vav was found in *va*ccinia *v*irus. It contains a whole collection of distinct structural motifs: cysteine rich,

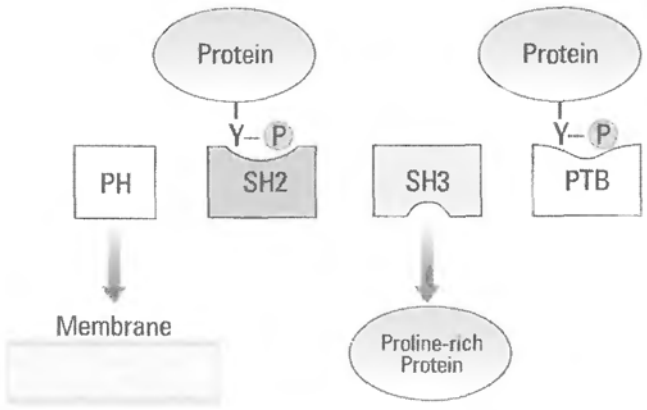


Fig. 3.1 Common structural motifs, instrumental in forming interconnectivity of components in signal transduction chains.

Zn-finger-HLH, helix–loop–helix DNA-binding motifs, PH- and SH2- and SH3-homology domains. The rich endowment of different structural modules allows Vav to function as a linker, a signal transmitter, and a regulator of gene transcription in lymphocytes and other myeloid cells. Deletion of the *vav* gene is accompanied by defects in antigen receptor signalling and impairment of lymphocyte proliferation and lymphopenia.[6] Shc is a linker that binds to activated, autophosphorylated EGF-, PDGF-, and FGF-receptors.[7] It has a PTB domain. PTB is a structural module which is an alternative to SH2 domains.[8] It binds with high affinity to sequences with the motif NPXpY (N is asparagine, P is proline, X is any amino acid, and pY is a phosphotyrosyl) in RTKs. But, in contrast to SH2 domains, Shc also accepts unphosphorylated peptides as partner. Shc also helps to link the monomeric G protein, Ras, to RTKs[9] (see p. 46).

Crk, Shc, Grb2 (growth-factor-receptor-binding protein), and IRS1 (insulin receptor substrate1) have related functions.

When Shc and related proteins are attached to activated RTKs, they become phosphorylated and, when phosphorylated, the pTyr (Y-P) on the adaptor serves as a marker, just like the pTyr on the receptor. In that way Shc is linked, for example to the SH2 domain of the adaptor Grb2. Grb2 has two SH3 and one SH2 domains. Each structural module of Grb2 seems to function independent of the other.[10] Thus, binding of the receptor phosphotyrosyls to the SH2 domain of Grb2 does not affect interactions of either one of the SH3 domains with prolyl-rich peptides, and conversely the binding of a prolyl-rich peptide to one of the SH3 domains of Grb2 does not affect the interaction with the other. This enables Grb2 to form heterologous complexes with several different proteins. For example, Grb2 recruits the prolyl-rich Sos (the product of the 'son of sevenless' gene in *Drosophila*). Sos is a guanine nucleotide exchanger, a GEF, and a regulator of the activity of the monomeric GTP-binding protein, Ras. Moreover, Grb2 can accept a proline-rich sequence of a GAP. These connections bring G proteins, such as Ras, via GAP and the Sos-GEF, under the control of growth factors, and make Ras a central checkpoint, controlling transmission of growth factor signals to the Ras/MAP kinase (mitogen-activated protein kinase) pathway, a universal cellular signalling route. (Fig. 3.2). (see Chapter 4).

One way to reduce complexity of cellular regulation is to look for conserved structural modules, present in many different regulatory proteins. These recurrent motifs help

to find the likely coupling partners. They serve as a road map, that shows the way through the maze of signalling routes.

Protein–protein recognition in signal transduction pathways

In the past two decades, a large amount of genetic, biochemical, and structural data have become available on conserved structural modules in regulatory proteins. The number of discrete structural motifs in regulatory proteins is still growing and it is safe to conclude that more structural domains will be found as more signalling proteins are discovered. This survey considers only some of the most common motifs. But this information is important for understanding cellular signalling, because protein–protein interactions through these recognition surfaces are often modulated by reversible post-translational modifications, such as phosphorylation dephosphorylation. An equally important aspect is cellular compartmentalization that regulates the accessibility of interacting partners (see Chapter 7).

A principal function of the phosphorylated tyrosine sites on receptors is to recruit other proteins with complementary recognition sites. These can be enzymes, such as phospholipase C, or regulatory proteins and linkers of all kinds.[16]

The recruited proteins may have quite different functions, but their structures often have in common that they contain consensus sequences, 50–100 amino acids long, although the adjacent regions may be quite different. Among sequences recognizing phosphotyrosyls are the Src homology domain, SH2, and the PTBs.

The SH2 domains

SH2 domains recognize peptide sequences bearing phosphotyrosine (pTyr) residues. They bind to compatible phosphopeptides with an affinity in the 10–100 nM range and bind to phosphopeptides of random sequence with a 1000-fold lower affinity. They have practically no affinity for unphosphorylated peptides.[17–20] The topography of a tyrosyl-binding SH2 domain is well known, thanks to structural studies (Plate 2).[21,22]

The phosphotyrosyl binding site of a SH2 domain has two parts. A conserved pocket binds the pTyr. This pocket contains only one crucial invariant residue, an arginine, which forms hydrogen bonds with two oxygens of the

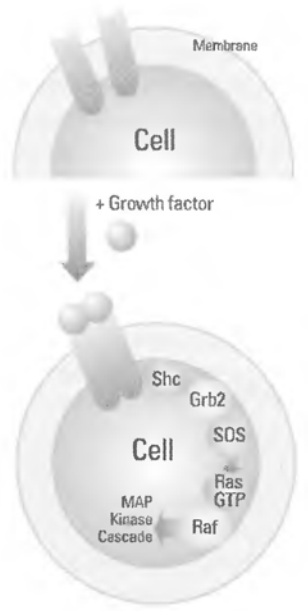


Fig. 3.2 Ras is associated with the membrane and receives signals from RTKs through the intervention of adaptors, such as Shc, which bring Ras under the control of GDP-exchange factors and GAPs. The consequence is the formation of the active GTP-bound form of Ras. Ras in the GTP-bound form recruits Raf kinase, a serine/threonine kinase, to the membrane. The activated Raf kinase triggers the MAP kinase, (mitogen-activated protein kinase), phosphorylation cascade. But Ras can also control other signalling routes. For example, when Grb2 couples to dynamin, a large, 120 kDa GTPase, clathrin-mediated endocytosis of proteins becomes responsive to growth factor signals.[11–13] Still another Grb2-docking protein is Gab-1 (for Grb2-associated binder-1).[14] Gab-1 shares homology and several structural features with the insulin receptor substrate-1 (IRS-1). Therefore, Grb2 associated with Gab-1 can respond to the insulin-receptor signal. The IRS-1 has a phosphotyrosine-binding domain, through which it binds to the tyrosine phosphorylated insulin receptor.[15]

pTyr. The second part of the binding site recognizes the amino acids immediately C-terminal to the pTyr. This recognition site can distinguish between different sequences adjacent to a pTyr. It is responsible for the great selectivity with which different pTyr-containing proteins are recognized.

The biological consequences of the interaction of pTyr-containing proteins with SH2-domain proteins depend on the function of the protein recruited. Thus, a phosphorylation–dephosphorylation cascade may be activated by recruitment of a kinase[23] or a phosphatase,[24] or a target enzyme may be stimulated which produces a second messenger in response to a receptor signal,[25] or signalling proteins may be relocated and transcription factors activated.[26] Eventually, most of these events lead to the activation of gene transcription.

The SH3 domains

SH3 (Src-homology 3) domains are, like SH2 domains, small, non-enzymic motifs, which are widely distributed in proteins, notably in those that participate in tyrosine kinase receptor (RTK) signalling cascades. The binding partners for SH3 domains are proline-rich peptides.[27,28] In contrast to the binding of pTyr-containing peptides to SH2 domains, binding of proline-rich peptides to SH3 domains is remarkably weak, with dissociation constants in the range of 5–50 mM. Thus, the weak interactions must be strengthened synergistically by simultaneous interaction with other SH3 domains and, indeed, the coordinated binding of several domains enhances the overall affinity up to 1000-fold. This may explain why typical linkers and adaptors, such as Grb2, have more than one SH3 domain (Fig. 3.3).

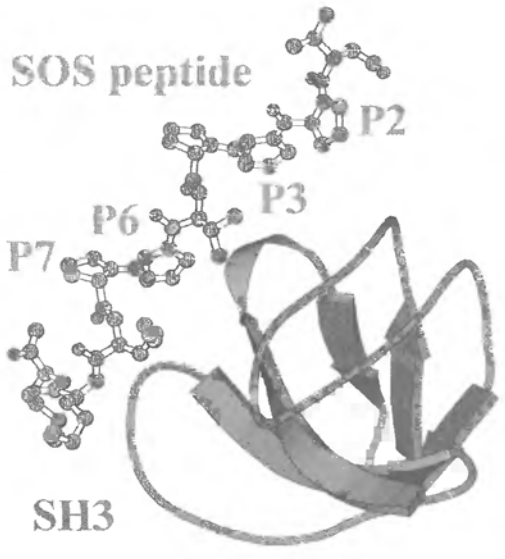


Fig. 3.3 The structure of the N-terminal SH3 domain of Grb2 bound to a proline-rich Sos peptide has been determined by NMR.[29,30] The structure of the Gbr2 N-terminal SH3 domain, complexed with a 10-residue peptide, comprising residues 1134–1144 (VPPPVPPRRR-NH_2) of Sos, is shown. The prolyl residues, P2, P3, P6, and P7, which interact with the SH3 domain of Grb2 are marked. (The ribbon model was reproduced with permission of the authors and J. Mol. Biol. from data in ref. 30, available in databanks.) A variation of this scheme is the recognition of a proline-rich sequence (APTMPPPLPP) in the GAP protein for Rho by the SH3-domain of the cytosolic c-Abl tyrosine kinase.[31] This interaction couples the Rho/GAP tightly to this cytosolic tyrosine kinase and brings the momomeric G protein, Rho, under the control of phosphorylation by the kinase.

Grb2 is built almost entirely from SH2 and SH3 domains. With its SH2 domain, Grb2 couples to activated, autophosphorylated growth-factor receptors, such as the EGFR, by recognizing a pTyr–X–Asn–X consensus sequence around Tyr1068, and with its SH3 domain Grb2 can couple to other and different signalling proteins. Each coupling partner may open another signalling path. A primary coupling partner of Grb2 is Sos.[32] The proline-rich sequence of Sos bridges the N- and C-terminal SH3 domains of Grb2, connecting them. Both SH3 domains of Grb2 are necessary for stable binding *in vivo*. Sos activates the Ras GTPase by enhancing GDP release and promoting formation of the active GTP-bound form. Association of Sos, the Ras nucleotide-exchange protein, with Grb2 is important for RTK signal transduction (see p. 48, Fig. 3.14).[33,34] An SH3 domain bound to a proline-rich peptide is shown in Plate 3.[35,36]

SH3-binding prolylpeptides have pseudosymmetry. They can bind in either direction, e.g. amino- to carboxy-terminal or carboxy- to amino-terminal.[37] The orientation of the prolylpeptide associated with the SH3 domain will determine the spatial orientation of the complex. This, in turn, determines the specificity of the recognition surface for other coupling partners. Thus, the possibility of accepting prolylpeptide ligands in either direction expands the range of potential binding partners and consequently the spectrum of the biological response. The core region of prolylpeptides which bind to SH3 domains is conserved. It is usually seven residues long and is centred around five residues with the consensus sequence X–P–(P)–X–P or XPPXPPX, where X is an aliphatic residue and P are the conserved prolines. The intervening residue, (P), also tends to be a proline. Each X–P pair fits into a hydrophobic pocket of the SH3 domain, formed by conserved aromatic residues. This mode of interaction provides the principal binding energy and stabilizes a left-handed polyproline type II helix. Prolylpeptides adopt this structure when they bind to SH3 domains.[38] Another site of interaction of SH3 proteins with prolylpeptides is more variable, although it frequently contains an arginine.

To sum up—mainly based on structural information, the SH3 prolylpeptide interactions can be divided into three groups (cf. ref. 39):

(1) interactions directly associated with the XP(P)XP motif;

(2) interactions involving the side-chains of residues preceding the conserved prolines; and

(3) interactions associated with residues at the termini of the prolylpeptides.

The PH and PTB domains

The pleckstrin-homology domain is named after pleckstrin, where it was found first (pleckstrin is a major platelet protein which is a substrate of protein kinase C).[40,41] Pleckstrin and brain spectrin contain two domains,[42] about 120 residues each, with similar sequences in their amino- and carboxy-terminal parts. To date, at least 90 such PH-domain sequences have been described in proteins with very diverse functions, such as spectrin,[43] cytosolic protein kinases, phospholipase C isoforms (PLC-β, -γ, -δ) and the GTPase dynamin,[44,45] guanine nucleotide exchange factors for monomeric GTP-binding proteins, the GTPase-activating protein for Ras, the Ras-GAP, and cytoskeletal and several other proteins.

A characteristic feature of PH proteins is their lack of sequence identity. The sequence identities of some of these proteins are as low as 15% and lower. Only a tryptophan near the carboxy-terminal end is conserved in all PH domains. Quite often, PH domains are

associated with other structural motifs. For example, the PH domain of PLC-γ has an insertion with two SH2 and one SH3 domains.

A role as a recognition surface in connecting proteins, similar to the role of SH2 and SH3 domains, has been postulated for PH domains. However, protein partners specifically recognizing PH domains are rare. Therefore, a role of PH domains in anchoring soluble proteins to membranes succeeded the original assumption that PH domains determine protein–protein interactions. In that case, PH domains would be like covalently attached fatty-acid residues, which also help to attach proteins to the membrane. The structural properties of PH domains support a role in bringing cytosolic proteins to the membrane. The structure of a PH domain is shown in Plate 4a. Structures of complexes of inositol phosphates with a PH domain were solved, both by X-ray crystallography and NMR.[46–48] The structural features are shown in Plate 4b.

Interactions of proteins with membranes are quite unspecific. Accordingly, typical PH-domain proteins, such as pleckstrin and spectrin, bind with no apparent stereospecificity and with quite low affinity to phosphotidylinositol bisphosphate ($PtdInsP_2$) and IP_3 in lipid vesicles ($K_d \geq 30\ \mu M$ and 40 μM, respectively). Moreover, whether all PH-domain proteins associate with negatively charged membrane phospholipids is not certain. However, there is the case of PLC-δ1, where the PH domain is instrumental in bringing the soluble, cytosolic enzyme to its membrane-bound substrate, $PtdInsP_2$. This function is supported by deletion mutants of PLC, lacking the PH domain, which cannot find the membrane-bound substrate. Therefore, while the role of the PH-domain motif in protein–protein interactions and interconnectivity is questionable, its role in membrane attachment of soluble, cytosolic proteins is quite persuasive.

Perhaps the most convincing evidence for a biologically important function of a PH domain comes from the consequences of a single missense mutation in the X-linked *btk* gene, the gene for a cytosolic tyrosine kinase (Bruton's agammaglobulinaemia tyrosine kinase, BTK; this tyrosine kinase is a member of the Src family of cytosolic tyrosine kinases).This mutation alters a single amino acid, a surface-exposed arginine in the PH domain of BTK to a cysteine. (In human pleckstrin this arginine is next to one of the conserved lysines (K22).) The mutation in BTK abolishes binding to the lipid messengers, phosphatidylinositol-3,4,5-trisphosphate and IP_3, and consequently severely impairs murine B-cell development, causing agammaglobulinaemia.[50,51]

Since PTB domains adopt the same fold as PH domains, J. Schlessinger has suggested that the PTB domain may be another class of PH domain, a relationship that was overlooked because of the absence of sequence homologies.[52] This classification is supported by differences between PTB domains and classical SH2 domains. The crystal structure of the phosphotyrosine-binding domain (PTB) of a human, neuron-specific peptide (X11) that contains a C-terminal PTB domain, provides relevant information.[53] With its PTB domain, this peptide binds to the cytoplasmic domain of the β-amyloid precursor protein (β-APP), found in the brain of patients with Alzheimer's disease. The domain with which the peptide interacts is an internalization motif of β-APP (Fig. 3.4).

To sum up: while the SH2 domains recognize C-terminal residues adjacent to the phosphotyrosine, the PTB domains recognize the N-terminal residues adjacent to the phosphotyrosyl residue. Moreover, in PTB domains a β-turn with the motif NPXY* (N is asparagine, P is proline, X is any amino acid, and Y* is a phosphotyrosine) is critical for recognition. However, the main difference between PTB domains and SH2 domains is that the former also bind to unphosphorylated peptides and proteins, whereas the latter only recognize phosphorylated proteins.

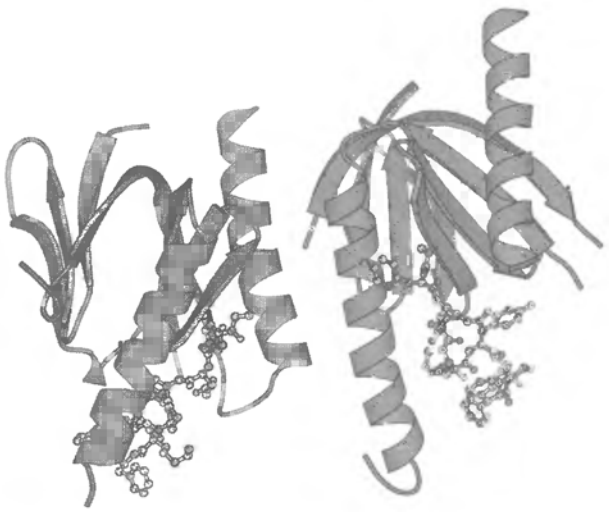


Fig. 3.4 The structure of a PTB domain that recognizes the internalization motif of Alzheimer's amyloid precursor protein, β-APP. A human PTB-domain fragment of X11 is shown that binds to a 14mer peptide of the internalization motif of the Alzheimer's amyloid precursor protein (β-APP). Please note that the peptides are unphosphorylated. Eight residues make specific contacts with the PTB domain. Collectively they achieve high affinity. Two ribbon presentations at different orientations of the same PTB domain in complex with the β-APP peptides are shown. The peptides are shown as ball-and-stick models and the X11 PTB-domain structures are shown as ribbon models on the left and the right, respectively. (These ribbon models were reproduced with permission of the authors and the EMBO J. from data in ref. 53. available in protein databanks.)

The notch/ankyrin repeats are evolutionarily conserved structural modules

Finally, a structural motif is examined that is evolutionarily conserved and participates in cell–cell interactions in embryonic development. These are the 'notch/ankyrin' repeats, which are found in vertebrates, invertebrates, and the metazoan *Hydra vulgaris*.[54] Notch is a transmembranous signal receptor, which regulates neurogenesis.[55,56] Ankyrin repeats are also found in tyrosine kinases in immune cells[57] and in other human proteins. They may interact with tumour suppressors and interfere with their function, thus contributing to carcinogenesis (see Part 4). An example is the helical ankyrin repeats in a p53 tumour-suppressor-binding protein, p53BP2 (p53-*b*inding *p*rotein 2) (Fig. 3.5).[58]

The list of linkers and domain modules discussed here is far from being complete. In Chapter 1 I mentioned the Eph-receptor tyrosine kinases, which play a role in axonal pathfinding of neurons and cell–cell interactions. Therefore, I conclude this part by mentioning a structural motif of Eph receptors, SAM, the '*s*terile *α*-*m*otif'. The X-ray structure of a Eph receptor–SAM domain has been solved.[59] SAM domains are involved in oligomerization of Eph receptors.[60] Figure 3.6 shows the NMR structure of the SAM domain of the Eph B2 receptor.[61]

Linkers with enzymatic activities

Cytosolic tyrosine kinases and tyrosine phosphatases are representatives of components of signalling pathways with enzymatic functions. They phosphorylate and dephosphorylate regulatory proteins and form phosphorylation cascades, enhancing or dampening signal transmission from receptor tyrosine kinases (RTKs). Cytosolic tyrosine kinases and tyrosine phosphatases are substrates of RTKs and are controlled by them by

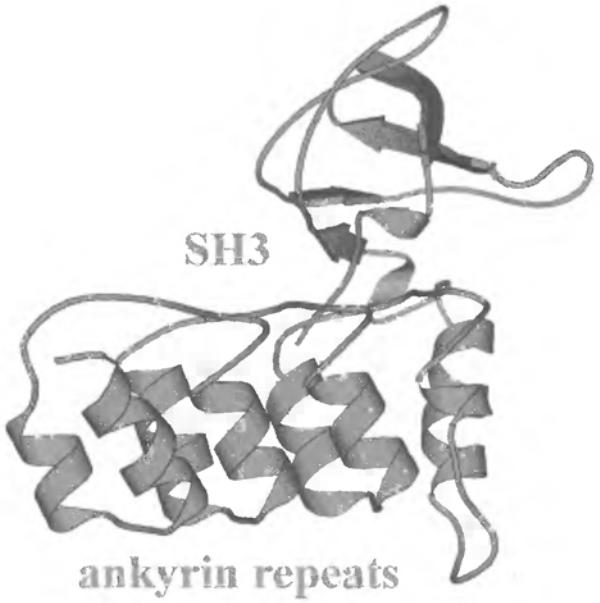


Fig. 3.5 Part of the helical ankyrin repeat structures and part of the SH3 domain of human p53BP2 with which it binds to the tumour suppressor p53. The region of p53BP2 that binds to the tumour suppressor p53 comprises about 200 amino acid residues, and the complementary region of the tumour suppressor, p53, comprising residues 97–287, is about the same size. (It should be noted that the ankyrin repeats are helical repeats, like those found in the structure of GAPs; see Chapter 4.) (This ribbon model was reproduced with permission of the authors from data in ref. 58 and Science deposited in the Brookhaven protein databank.)

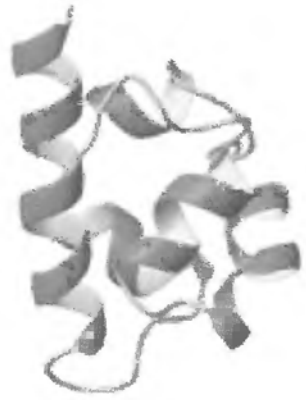


Fig. 3.6 This is a stereo view of the SAM structure determined by NMR. The characteristic five helices are shown. (The ribbon structure is reproduced from ref. 61, with permission of the authors and Protein Science.)

phosphorylation. But they have their own set of substrates, which they modify by phosphorylation and dephosphorylation. The specificities of cytoplasmic and receptor tyrosine kinases match closely, reflecting the preferences of their recognition motifs. This suggests that membrane receptor tyrosine kinases and cytosolic tyrosine kinases may have co-evolved.

Tyrosine kinases: the Src and Abl kinases

Typical cytoplasmic tyrosine kinases are the c-Src and the c-Abl kinases (c-Src is the product of the cellular proto-oncogene of the Rous sarcoma virus-oncogen, v-Src, and c-Abl is the cellular representative of the v-Abelson leukaemia virus tyrosine kinase). The family of c-Src kinases comprises, at present, nine members. All Src tyrosine kinases have SH3 and SH2 domains, a catalytic kinase domain, and a short carboxy-terminal tail.[31,62] Some cytosolic tyrosine kinases have additional domains.

The role of the phosphorylation activity of cytosolic Src kinases in growth-factor signalling was demonstrated convincingly with a catalytically inactive Src kinase fragment, overexpressed in *src*$^{-/-}$ mouse fibroblasts. Although, this protein still contained the SH2 recognition domain for the receptor and the myristoyl membrane anchor, PDGF- and EGF-induced mitogenesis and cell proliferation were abolished in these cells. Thus, the Src-family kinases are not only linkers, what matters is their capability for protein phosphorylation. The constitutively high tyrosine kinase activity of the viral oncogenic variant, v-Src is responsible for

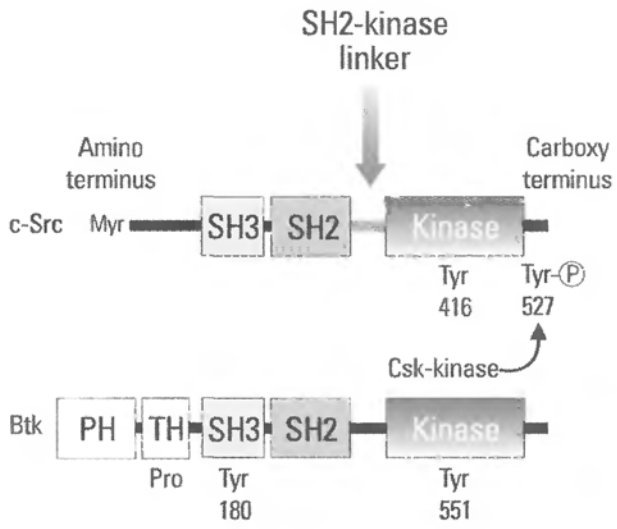


Fig. 3.7 The domain organization of c-Src- and Bt-kinases. The presence of several structural modules explains the coupling versatility of cytosolic tyrosine kinases. They can accept not only signals from membrane-anchored RTKs, but also from other cytosolic tyrosine kinases, such as Csk (C-terminal Src kinase).[63] The amino-terminus of the c-Src kinase is myristoylated. The fatty acid is attached irreversibly as acylamide (Myr), whereas in the Btk a PH domain fulfils the role of the fatty acid and brings the enzyme to the membrane. The phosphorylation state of two tyrosines determines the activity of the kinase: phosphorylation of a carboxy-terminal Tyr (Tyr527 in c-Src) inactivates, and dephosphorylation by a tyrosine phosphatase activates, the kinase. In the active state Tyr527 is unphosphorylated, and Tyr416 in the kinase domain is phosphorylated. When Tyr527 in c-Src is replaced by a phenylalanine, the tyrosine kinase activity is enhanced and c-Src becomes oncogenic. (In the oncogenic v-Src, Tyr527 is deleted by a mutation.).When Tyr527 is phosphorylated by another cytosolic tyrosine kinase (Csk), the Src activity is repressed, because the phosphotyrosyl 527 locks the kinase in an inactive conformation.[64] This has been verified by structure determination (Plate 5).

malignant transformation of cells, expressing the oncogenic v-Src-kinase. Thus, the activity of the cytosolic kinases must be restrained and regulated tightly, in order to keep the cells healthy.

Thanks to crystal structures of human c-Src and its close relative, the hemopoietic cell kinase, Hck, the control of the activity of these cytosolic tyrosine kinases by phosphorylation dephosphorylation is now understood (Fig. 3.7 and Plate 5).

In conclusion, there are two ways to activate Src kinases:[66]

(1) either dephosphorylation of Tyr527 by a phosphotyrosine phosphatase; or

(2) binding of proteins that recognize the SH3 domain that blocks the catalytic domain and remove the barrier, preventing access to the catalytic site.

Both modes of activation are realized: an example of activation by dephosphorylation is the activation of the cytoplasmic, lymphocyte tyrosine kinase (Lck) by the membrane-bound CD45-phosphatase in hemopoietic cells, whereas a release of the inhibitory

domain has been implicated in the activation of Src kinase (and related kinases, such as Fyn and Yes), on binding to activated RTKs, such as PDGFR-β. (Fyn and Yes are products of the cellular protooncogenes, *c-fyn and c-yes*, respectively. The names refer to the viral oncogenes. For example, *v-yes* is the oncogene of the Yamaguchi sarcoma virus).

The activation of the Bt-kinase is a variation of the same theme: phosphorylation of a Tyr, (Tyr180), in the SH3 domain of the Bt-kinase relieves the blockade by interrupting the intramolecular interaction of the catalytic domain with the proline-rich (TH) domain, leading to activation. Moreover, several substrates of Src kinases have proline-rich domains which recognize SH3 domains and could de-inhibit the kinase.

One of the kinases that are critical for the regulation of the activity of Src kinases is Csk (C-terminal Src kinase). This kinase binds in lymphocytes (like the *l*ymphocyte *k*inase, Lck) to a phosphatase, PTP, which dephosphorylates the activating phosphorylation site in the c-Src kinase. The Csk–PTP complex inhibits the Src kinase in two ways: it adds an inhibitory tyrosine phosphate to the Src kinase and it removes an activating phosphate. This assigns a central role to Csk and related kinases in the regulation of the activity of the c-Src tyrosine kinases. see: Fig. 3.7. The question then is how Csk is regulated. Csk itself is not tyrosine phosphorylated, neither on an inhibitory or an activating site. It now seems that intracellular compartmentalization of this kinase is an important point of regulatory control (see also Chapter 7). Csk is in the cytosol, but can be relocated to the membrane, where it meets the Src kinases. Csk has a SH2 domain with which it binds to a tyrosine-phosphorylated *C*sk-*b*inding *p*rotein (CBP).[71] The binding protein is phosphorylated, probably by an Src kinase. CBP is a typical membrane receptor-like protein with extracellular, transmembranous, and intracellular regions.

Cytosolic tyrosine kinases are often found attached to the plasma membrane. Some are associated with subcellular structures and the cytoskeleton and others are in the nucleus. To coordinate the activities of the many protein kinases (and phosphatases) in the membrane, membrane 'rafts' would be an attractive location for c-Src kinases and their regulators, the Csk–CBP complex. 'Rafts' might be visualized as a kind of railway station where extracellular signals enter the trains which bring them to their destinations in the cell. We do have an idea how the Src-family tyrosine kinases are activated and deactivated, but we have not yet identified all the trains bringing the signals to their destinations in the cell. (Membrane 'rafts' are membrane domains with high cholesterol and glycolipid content; see Chapter 8; ref. 27).

Considering the multitude of signals to which a cell responds, it is likely that several different upstream kinases may play a role. Indeed, more than one signal pathway seems to be required for the full activation of the Jun/Fos-controlled early response genes, expression of which is essential for cell growth and proliferation. *c-Jun* is a family of proto oncogenes encoding transcription factors. The related *v-jun* is an oncogene from the avian sarcoma virus, ASV17. The product of the cellular proto oncogene is Jun. Fos was originally found in murine osteosarcoma viruses. The product of the *c-fos* gene, Fos, is a growth-promoting transcription factor, found in the nucleus of cells. Jun forms heterodimers with Fos which like the homodimers of Jun bind to the AP1 consensus site of early genes which promote cell growth and proliferation. Jun is a typical DNA-binding leucine zipper protein. DNA-binding and dimerization of Jun is regulated by phosphorylation, de-phosphorylation. (See: chapter 10). Apparently, both the phosphatidylinositol 3-kinase pathway and the Ras/Raf/MAP kinase cascade are needed for the promotion of cell proliferation. (The name PtdIns 3-kinase is imprecise, because the kinase accepts not only phosphatidylinositols but also proteins as substrates).

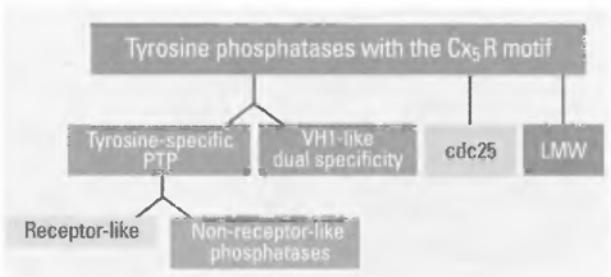


Fig. 3.8 The protein tyrosine phosphatases (PTPs) with the $C(X)_5R$ motif are divided into tyrosine-specific PTPs; the VH1-like dual specificity, serine/threonine- and tyrosine-specific phosphatases; the CDC25 phosphatase; and the low molecular weight (LMW) phosphatases. The low molecular weight phosphatases are acid phosphatases without distinct regulatory or targeting domains. Their function is not known. The tyosine-specific phosphatases are further subdivided into receptor-like and non-receptor-like phosphatases. (This scheme is reproduced with permission of the authors and Trends Biochem. Sci. from ref. 76.)

Phosphotyrosine phosphatases

The regulatory role of protein phosphotyrosine phosphatases received attention rather late, in the 1980, mainly thanks to the work of E. H. Fischer and his laboratory in Seattle.[73]

Phosphotyrosine phosphatases[74] are integrated just like tyrosine kinases into signalling pathways. They interact with receptors and have recognition motifs that direct them to their targets.[75] Protein phosphotyrosine phosphatases downregulate tyrosine phosphorylation and play a role in cellular regulation as important as that of protein tyrosine kinases.

Phosphotyrosine phosphatases are only one of several classes of phosphatases. It has been estimated that about a thousand phosphatases of different specificities exist. Here, we consider only cytosolic and receptor tyrosine phosphatases (Fig. 3.8) (phosphoserine/threonine phosphatases are discussed in Chapter 7).

The catalytic site of protein tyrosine phosphatases has a conserved sequence motif, H/V–C–$(X)_5$–R–S/T–G/A/P, where X is any amino-acid residue; H is histidine; V, valine; C, cysteine; R, arginine; S, serine; T, threonine; G, glycine; A, alanine; and P, proline. The cysteine in this site is involved in catalysis, in binding and removing the phosphate group of the substrate.

VH1 (encoded by the vaccinia virus late gene 1) was the first dual-specificity phosphatase to be characterized. Although these phosphatases hydrolyse phosphoseryls and phosphothreonyls as well as phosphotyrosyl, they have a preference for tyrosyl phosphates. Among other dual-specificity phosphatases, the cdc25 (cell division control) phosphatase is of special interest.[77] cdc25 regulates cyclin-dependent kinases (cdks). This is a crucial point of control of the cell cycle. Mammals have three cdc25 phosphatases, each directed to a specific cyclin-dependent kinase; (this is discussed in Chapter 12). Although, no structure of cdc25 is available as yet, it is expected that the hydrolysis reaction proceeds like in other phosphatases with a $C(X)_5$ R motif.

The best-studied protein tyrosine phosphatases are the high molecular weight cytoplasmic enzymes of the PTP family. X-ray structures of the human PTP1B cytosolic tyrosine phosphatase, have been solved by David Barford *et al.*,[78,79] and that of a *Yersinia* tyrosine phosphatase by Fauman *et al.*[80,81] In Fig. 3.9a and b the structures of the

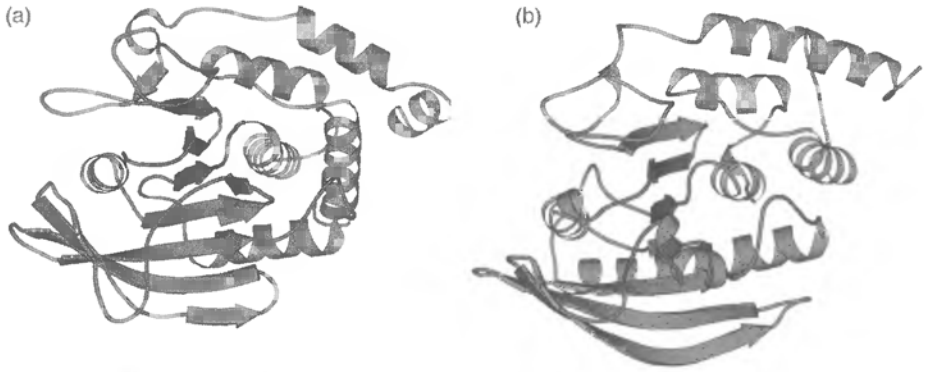


Fig. 3.9 (a) A ribbon presentation of the structure of the human tyrosine-specific phosphatase 1B, (PTP1B). It is a single-domain protein with a seven-stranded, mixed b-sheet, flanked by a-helices. The catalytic site is in the central region, in a shallow cleft. The phosphate-recognition site is characteristic for PTPs. It is formed by a loop containing the $C(X_5)R$ motif, which contains the catalytically essential cysteine and arginine residues. (Reproduced with permission of the authors, D.Barford, A. J. Flint, and N. K.Tonks, and Science from ref. 78.) (b) For comparison is shown a ribbon presentation of the *Yersinia* tyrosine phosphatase (reproduced with permission of E. B.Fauman, Ch. Yuvaniyama, H. L. Schubert, J. A. Stuckey, and M. A. Saper and the J. Biol. Chem. from ref. 81).

tyrosine-specific phosphatase PTP1B and of the *Yersinia* tyrosine phosphatase are juxtaposed to emphasize the great similarity of the active site of all tyrosine phosphatases (this also includes the serine/threonine phosphatases, see Chapter 7).

The crystal structure of a complex of a Cys215Ser mutant of PTP1B with a peptide containing a pTyr-autophosphorylation site of the epidermal growth factor receptor has been solved (not shown). The phosphotyrosine side-chain anchors the phosphopeptide substrate to its binding site. Hydrogen bonds between peptide main-chain atoms and the phosphatase contribute to high binding affinity, whereas interactions of acidic residues of the peptide with basic residues on the surface of the phosphatase confer specificity. The structure revealed a conformational change on binding the substrate. This transition creates a phosphotyrosine recognition pocket and induces a catalytically competent form of the phosphatase.

Some of the non-receptor phosphotyrosine phosphatases have structural features that target them to other proteins or to special cellular locations. For example PTP1B has a membrane-binding domain and PTP2s have SH2 domains with which they interact with phosphotyrosyl sites on activated RTKs and with a class of transmembranous receptor-like glycoproteins, the SIRPs (signal-regulatory proteins). Although the function of SIRPs is not yet clear, it is likely that they dampen signalling through tyrosine kinase receptors.[82] A representative member is SIRP-α1 which is phosphorylated by RTKs. The tyrosine-phosphorylated form of SIRP-α1 recruits the cytosolic tyrosine phosphatase SH-PTP2 through SH2 domains and brings it to the receptor on the membrane, where it downregulates and dampens signalling by dephosphorylation of RTKs.

Probably, most growth factor RTKs are subject to regulation by dephosphorylation. Linkers could bring the phosphatases to the receptor, where they modulate its state of phos-

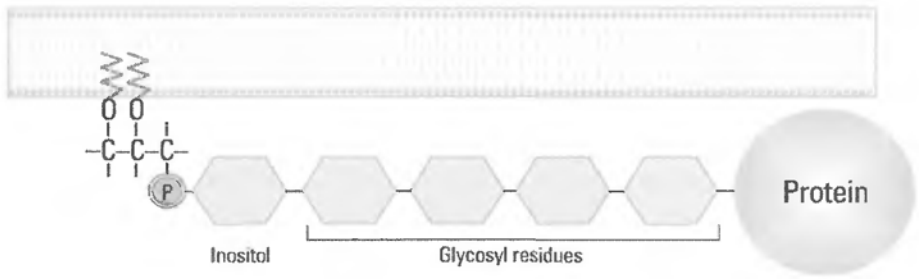


Fig. 3.10 A protein anchored to the membrane with a glycosylphosphatidylinositol anchor.

phorylation. Since the phosphatase activity is itself controlled by phosphorylation, a steady state of receptor phosphorylation/dephosphorylation could be established, controlling the receptor activity. Moreover, phosphorylation could also create recognition sites on the phosphatase for the recruitment of linkers, just as in the case of the receptor kinases.

Membrane-anchored receptor tyrosine phosphatases exist, just as there are RTKs. A typical receptor tyrosine phosphatase (RPTP) is CD45. It plays a pivotal role in B- and T- cell activation (see Chapter 14). Although RPTPs are mostly orphan receptors with unknown ligands, their cysteine-rich, highly glycosylated extracellular region suggests that they may have a role in cell–cell interactions and in interactions with the extracellular matrix. For example, the phosphatase PTPξ has a extracellular carbonic anhydrase (CAH) domain, anchoring the enzyme to the surfaces of neurons. Other RPTPs have fibronectin domains. These receptor-like, membrane-anchored tyrosine phosphatases could regulate cell–cell interactions by removing phosphates attached to tyrosine residues on cell adhesion molecules. An example is contactin, which is associated with a transmembrane protein, the contactin-associated protein, Caspr. The extracellular domain of the RPTP-β contains a CAH domain, through which it binds contactin.[83] RPTP-β is located on the surface of glial cells and contactin on the surface of neuronal cells, where it is anchored through a glycosylphosphatidylinositol (GPI) moiety (Fig. 3.10). This is how RPTP-β enables cross-talk between glial cells and neuronal cells.

To sum up: pathway interconnectivity is established through protein–protein interactions which are often modulated by phosphorylation–dephosphorylation reactions.

Ras, a G protein as regulator

The active form of all monomeric and heterotrimeric G proteins is the GTP-bound form. All G proteins have in common that the GTP-bound form is slowly converted to the inactive GDP-bound form by the protein's intrinsic capacity to hydrolyse GTP to GDP and inorganic phosphate (Pi). However, monomeric G proteins, such as Ras have practically no measurable intrinsic GTPase activity. Therefore, the 'switch-on' of activity on replacing GDP by GTP and the 'switch-off' by the GTPase are all-or-nothing phenomena. The 'switch-on' is controlled by GTP/GDP-exchange or GDP-release factors and GDIs (GDP-dissociation inhibitors). These factors have positive or negative effects, whereas the GTPase-activating proteins (GAPs) switch off the active state by promoting the hydrolysis of GTP to GDP and Pi (Fig. 3.11).

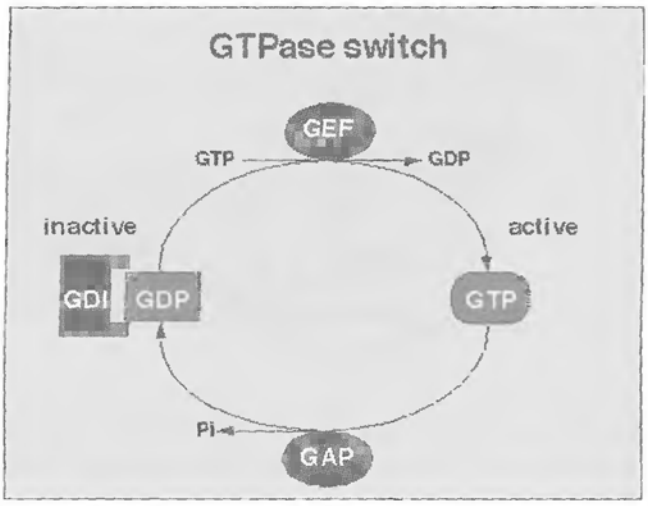


Fig. 3.11 The switch function of GTPases. The points of regulatory control are indicated: positive effects are due to interaction of GTPases with GDP-exchange factors (GEFs), and negative effects to interactions with factors inhibiting GDP release, GDP-dissociation inhibitors (GDIs). The GTPase-activating proteins (GAPs), shorten the lifetime of the active GTP-bound state. (This scheme was made available through the generosity of Professor Martin Lohse, Department of Pharmacology, the University of Würzburg, Medical School.)

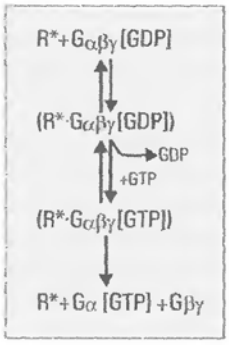


Fig. 3.12 Replacement of GDP by GTP leads to the dissociation of the αβγ-complex into the active, GTP-bound α-subunit and a complex of βγ-subunits, each of which controls independently the activities of target enzymes. In the case of the heterotrimeric G proteins, the receptor, activated by a ligand, is the most effective GDP/GTP exchanger (GEF) (see Chapter 5).

The GDP/GTP cycle

The GDP/GTP cycle of heterotrimeric G proteins is shown in Fig. 3.12. Heterotrimeric G proteins hydrolyse GTP at rates 1000 times (or more) faster than Ras.

Plate 6 depicts the concerted movements of three switch regions, which is the crucial step in the transition from the inactive to the active state of a typical G-protein α-subunit on binding GTP. The structural change in an α-subunit on binding GTP takes place in a relatively small region. In G_{tr}-α, the three switch regions contain only 14% of all the amino acids of an α-subunit. G_{tr}-α is the α-subunit of the heterotrimeric G protein transducin. According to Noel *et al.*,[84] the primary structural change, triggered by the γ-phosphate of GTP (or GTPγS), is propagated through a set of polar and non-polar linkages to the reactive contact surfaces of the α-subunit for interaction with other proteins. This loosens the contacts of the active, GTP-bound form of G-α with the receptor and the βγ-subunits, and leads to the dissociation of the heterotrimeric complex in separate α- and βγ-subunits, each of which can now establish contacts with effector enzymes and targets.

The structure of the αβγ-holocomplex of transducin from bovine rod outer segments of the retina, which was solved in Paul Sigler's laboratory,[85,86] is presented in Plate 7.

Looking at the structure of the α-subunit in Plate 7 reveals features that are unique. For example, the large helical domain on the left is present only in the α-subunits of the larger heterotrimeric G proteins and is missing in Ras. As was pointed out by Henri Bourne, this

domain functions like a GAP. It is responsible for the higher intrinsic GTPase activity of the α-subunits of the heterotrimeric G-proteins.[87,88] The higher rates of GTP hydrolysis of heterotrimeric G α-subunits are a consequence of a better structured α-2-helix, which stabilizes the transition state, where the γ-phosphate of GTP interacts with the side-chain of a conserved arginine. This becomes clearer when we look at the catalytic mechanism. For that purpose, we rely on the high-resolution structure of the catalytic site of the α-subunit of G_i-α1 determined by Steven Sprang and Alfred G. Gilman.[89] Despite the fact that G_i-α1 and G_{tr}-α are expressed in different tissues, interact with different receptors and different βγ-subunits, and regulate different effectors, the structures of G_i-α and G_{tr}-α are remarkably similar. Plate 8 describes the catalytic domain of G_i-α. G_i-α is the α-subunit of the heterotrimeric G protein G_i. (i = inhibitory). G_i regulates adenylylcyclase.

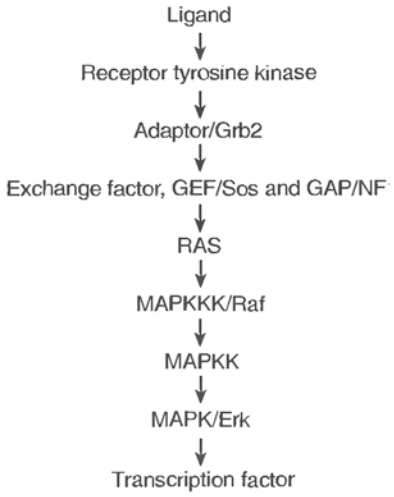
Hydrolysis of GTP proceeds by an (S_{N2}) in-line attack of a water molecule on the γ-phosphate of GTP. A Gly, a Gln, and an Arg residue are involved in catalysis. Mutations at any of these sites largely abolish GTP hydrolysis and stabilize the active GTP-bound state.[90] A constitutive, unregulated, permanent activity of the GTP-bound state can be harmful, as indicated by the consequences of mutations at these sites in G_s-α (G_s-α is the common stimulatory G protein, see Chapter 5). Mutated G_s-α is found in human endocrine tumours[91,92] and comparable mutations in Ras are oncogenic.[93]

The central role of the Arg in GTP hydrolysis is also apparent from the fact that its guanidino group is the target of ADP ribosylation by cholera toxin, which stabilizes the active GTP-bound state of G α-subunits of heterotrimeric G proteins by blocking hydrolysis of GTP.[94] The uncontrolled, constitutive activity of G-α in the intestine is the cause of the deleterious loss of fluid and salt in cholera.[95], Plate 8.

The three-dimensional structures of GTPγS–G_i-α1 of Coleman *et al.*,[89] of the transducin α-GDP–AlF_4^- complex by Sondek *et al.*,[96] and G_{tr}-α from Paul Sigler's laboratory, ref. 84, show how the crucial Arg-residue contacts the γ-phosphate of GTP directly. Thus, the presence or absence of the γ-phosphate in the guanine nucleotide defines the active or inactive state of a G α-subunit.

Comparing the structures in Plates 6 and 8 with the structure shown in Plate 9, shows that the catalytic domains of the Ras–GAP complex[97] and of the α-subunit of heterotrimeric G proteins are remarkably similar. This is not surprising because the catalytic mechanism is the same in the α-subunits of heterotrimeric and Ras-like monomeric G proteins.[98] According to Wittinghofer and his colleagues (Plate 9), a water molecule is the nucleophile which accepts the γ-phosphate, transferred in catalysis. The hydrogen-bonding distance of the water to the γ-phosphate, and its position in the catalytic centre, makes proton transfer possible. The γ-phosphate of GTP acts as a generalized base and Gln61 and Gly12 assist in stabilizing the geometry of the active site and the transition state.

To sum up: The crystallographers have explained the catalytic mechanism of the GTPase reaction of G proteins. Moreover, Wittinghofer *et al.* have provided a plausible explanation for the lack of the GTPase activity of Ras and have explained how the Ras GTPase activity is triggered by binding a GAP, the GTPase-activating protein. The important point is that during phosphoryl transfer, the γ-phosphate acquires a partial negative charge which must be neutralized. This is the function of the guanidino group of the Arg. This Arg is missing in Ras. Its absence explains the very low, hardly measurable GTPase activity of Ras. A GAP contributes the missing Arg residue and restores GTPase activity.



(NF1 is the neurofibromin GAP)

Ras

Ras is the master controller of a central cellular signalling pathway, the MAP kinase–phosphorylation cascade. Ras accomplishes its regulatory role in a way that is different from the function of the other regulatory components of signalling networks discussed up to now. Ras is a GTPase. In the GTP-bound state its activity is switched on and in the GDP-bound state, switched off.

c-Ras is a cellular proto-oncogene, related to a *rat sarcoma* retroviral oncogene. It is a monomeric 21 kDa protein and the prototype of monomeric G proteins. Ras is anchored to the membrane, where it interacts with linkers, which recruit guanine nucleotide exchange factors (GEFs) and GTPase activators (GAPs) which regulate the activity of Ras. Linkers connect Ras with membrane-bound RTKs, from which it receives signals from growth-promoting factors. Ras activates the Raf kinase, the first MAPKKK in the MAP kinase-regulated phosphorylation cascade, and turns on the whole downstream cascade, as illustrated below. (We shall learn more about the design of this phosphorylation cascade in Chapter 4.)

Control of the activity of Ras

There is ample evidence affirming the biological importance of Ras activity and its control by GAPs.

The GAPs

There are many examples emphasizing the importance of the control of Ras activity and explaining the great interest in finding smaller, GAP-like molecules that are capable of re-activating the GTPase. One of these examples is the role of Ras activity in T-cell activation.[99] Stimulation of antigen receptors in T cells turns down GAP activity (in conjunction with activation of protein kinase C) and consequently active Ras promotes cell proliferation, necessary for the immune response (see Chapter 14).

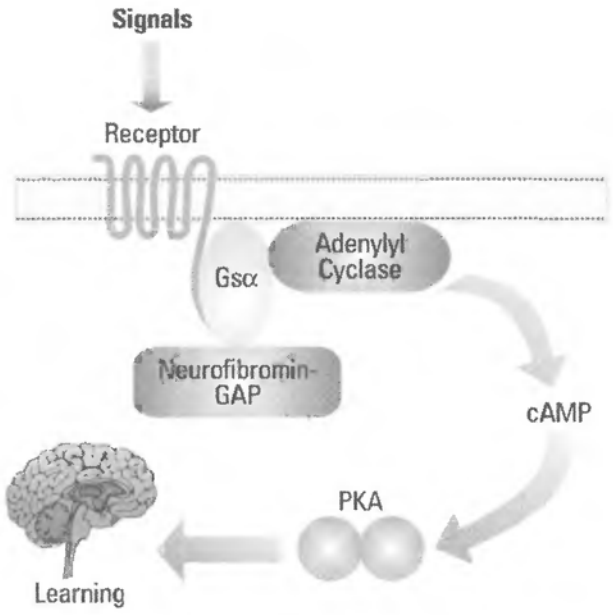


Fig. 3.13 Neurofibromin signalling in learning. G-protein-coupled receptors are linked to adenylyl cyclase through the α-subunit of the stimulatory G protein, Gs. Activated receptors stimulate adenylyl cyclase to synthesize cyclic AMP (cAMP), which activates the downstream effector, protein kinase A (PKA). Guo *et al.*[103] show that the neurofibromin protein is required in this pathway for normal odour learning in *Drosophila*, perhaps by direct interaction with a heterotrimeric G protein or with the adenylyl cyclase encoded by the *rutabaga* gene (see ref 104 for further information). (This scheme is based on information published in Fig. 1 of ref. 104 and reproduced with permission of the author and Nature.)

The role of neurofibromin 1

In humans, defect mutants of the gene coding for the neurofibromin GAP, NF1, cause Recklinghausen's neurofibromatosis. One in every 3500 individuals suffers from the genetic disease neurofibromatosis type 1 (NF1). The many symptoms of the disease, including tumours of the nervous system,[100] are accompanied by a deficiency in the protein neurofibromin. About half of NF1 patients suffer from learning disabilities, which also have been attributed to aberrant neurofibromin control. In the fruitfly, *Drosophila*, a loss of a functional neurofibromin 1 causes a learning disability (inability to discriminate odours).[101] However, in *Drosophila*, a defect in the neurofibromin control of adenylyl cyclase activity, rather than of Ras activity, is responsible for the learning defect. Accordingly, application of cyclic AMP did restore the learning defect,[102] and, as expected, expression in transgenic flies of an active form of the cAMP-dependent protein kinase A, a downstream component of the cyclic AMP pathway, corrected the learning deficiency of NF1 mutants. The defect in adenylyl cyclase activity in NF1 mutants resembles the consequence of a mutation in the *rutabaga* gene. The *rutabaga* gene encodes the neurofibromin-regulated adenylyl cyclase in *Drosophila* (Fig. 3.13).

This novel function of neurofibromin in *Drosophila* has not yet been found in mammals and, conversely, attempts to show that neurofibromin regulates Ras in *Drosophila* have failed. But, since adenylyl cyclase activity in humans is controlled by heterotrimeric G-protein-coupled receptors for hormones, one should check whether neurofibromin affects heterotrimeric G-protein activation in general, and in the nervous system in particular. Since some alternatively spliced variants of mammalian neurofibromin are expressed in neurons, it is also possible that modulation of cyclase activity in neurons is carried out by a particular isoform of neurofibromin.

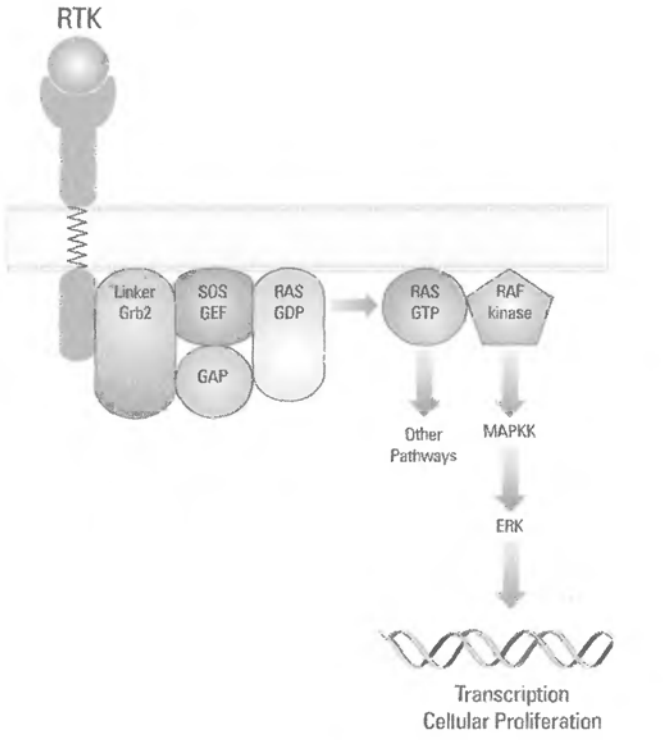


Fig. 3.14 The chain of events by which a RTK activates Ras. A complex is formed, containing the activated, phosphorylated RTK, and the linker, Grb2 (or Shc), which recruits the exchange factor, Sos. Contacts are through SH2 and SH3 domains of the coupling partners.[106] It is assumed that activation of Sos is due to its translocation from the cytosol to the membrane where the receptor and Ras are located.The activated, GTP-bound form of Ras is formed, which in turn activates the Raf kinase and the MAP kinase cascade.

A function of neurofibromin in human brain similar to that in *Drosophila* is quite likely, because mammalian and *Drosophila* genes are highly conserved, including cAMP-responsive genes involved in learning and behaviour, such as *NF1*, *dunce*, *rutabaga*, *CREB* (cAMP response element-binding protein), and others.[104]

The GNRPs

The first step in the activation of G proteins is the replacement of GDP by GTP. Whereas in the case of heterotrimeric G proteins, GDP–GTP exchange is catalysed by G-protein-coupled heptahelical receptors, monomeric G proteins, such as Ras, recruit GDP exchange factors (GEFs) and guanine nucleotide release proteins (GNRPs).[105] These factors promote formation of the active, GTP-bound form of Ras and, because they are linkers, connect Ras with the RTK.

When cells are treated with a growth factor, Ras is activated by Sos, the GEF for Ras. The lifetime of Ras-GTP is controlled by GAP (Fig. 3.14).

Ras is a distributor of many signals

Its versatility to interact with different adaptors makes Ras a distributor for a variety of different signals. To rearrange a cellular response, Ras can couple to other receptors and linkers and can transmit other signals. For example, signalling in neuronal cells is normally controlled by Ras, as in other cells. But sometimes Ras must be silenced, in order to change the cellular response. This is the case when neuronal cells are stimulated by NGF. In these cells the Ras-controlled Raf/MAP kinase is bypassed and, instead, Src kinase is activated. The consequence is a different cellular response, characterized by neurite outgrowth.[107] Ras also

responds to Ca^{2+}-calmodulin and other second-messenger-mediated signals. Moreover, Ras can recruit GNRPs, which interact with heterotrimeric G proteins (I shall come back to this, when discussing signalling by heterotrimeric G proteins in Chapter 5). Nitrous oxide (NO) and regulation of the redox state of a cell has also been implicated in the regulation of Ras.[108] *S*-nitrosylation of a cysteine, Cys118, enhanced the guanine nucleotide exchange activity of isolated p21 Ras, but whether a reaction of Ras with NO occurs in cells and is physiologically relevant remains to be seen.

The long and growing list of adaptors channelling signals supports the central role of Ras as a relay station for accepting, controlling, and distributing a multitude of signals.

Considering the role of Ras in signalling, it is to be expected that loss of Ras control is often associated with malignant transformation. It is estimated that 25–30 % of all human tumours are associated with a mutated Ras, having abnormal properties.

The coupling versatility of Ras and its engagement with different linkers could explain, at least in part, why each differentiated cell responds differently to growth factors and cytokines, and why the flow of information is dependent on the cellular environment. Thus, in order to correct deregulation and protect cells from becoming malignant, we need to know all possible signalling routes and their specific manifestations in differentiated cells. A sobering outlook, underscoring how much we still need to learn. But, to fulfil this role, Ras must be at the right place in the cell. Therefore, translocation to the site of action at the membrane is another important point of control.

Lipid modifications and membrane attachment of Ras

Membrane attachment of Ras requires covalent addition of lipids. Virtually all members of the Ras superfamily are prenylated.[109] Farnesylation is essential. Defects in lipid modification make Ras non-functional. A 15-carbon farnesyl or 20-carbon geranylgeranyl chain is attached to a cysteine in the COOH-terminus of Ras. The linkage is by a thioether bond and irreversible, and therefore not a point of regulatory control.[110]

The enzymes catalysing these reactions have been characterized.[111] Geranylgeranyl transferases type I and II attach one or two C_{10} prenyl groups to proteins, whereas farnesyl transferase (FTase) transfers C_{15} farnesyl groups much better than C_{20} geranylgeranyl groups. The X-ray structure of a mammalian FTase explains why this enzyme prefers farnesyl diphosphate (FPP) as the prenyl donor.[112] The reason is that the binding pocket of the FTase is rigid and has a limited length which still can accomodate C_{15} FPP, but not C_{20} geranylgeranyl diphosphate (GGPP). Interest in inhibitors of FTase is great.[113] Figure 3.15 shows that the primary modification by farnesyl transferase is followed by other modifications.[114]

The final step in the acylation of Ras is the formation of a palmitoylester bond with a cysteine. Palmitoylation is reversible and could be subject to regulation. In yeast cells, where the protease that removes the COOH-terminal AAX sequence has been deleted by mutation, Ras can not be palmitoylated. In these cells Ras has lost its orientation and ends up in the interior of the cell, rather than at the plasma membrane. Therefore it is thought that farnesylation may target Ras quite indiscriminately to all kinds of cellular membranes, including endoplasmic membranes, and that only palmitoylation traps Ras in the plasma membrane, where the palmitoyl transferase is localized.[115]

However, not every Ras is palmitoylated: c-K-(Kirsten)-Ras does not have a palmitoylation site, instead it has a cluster of eight basic residues that substitute for the missing palmitoyl residue and attach c-K-Ras to the inner leaflet of the plasma membrane by means of electrostatic interactions with acidic phospholipids.

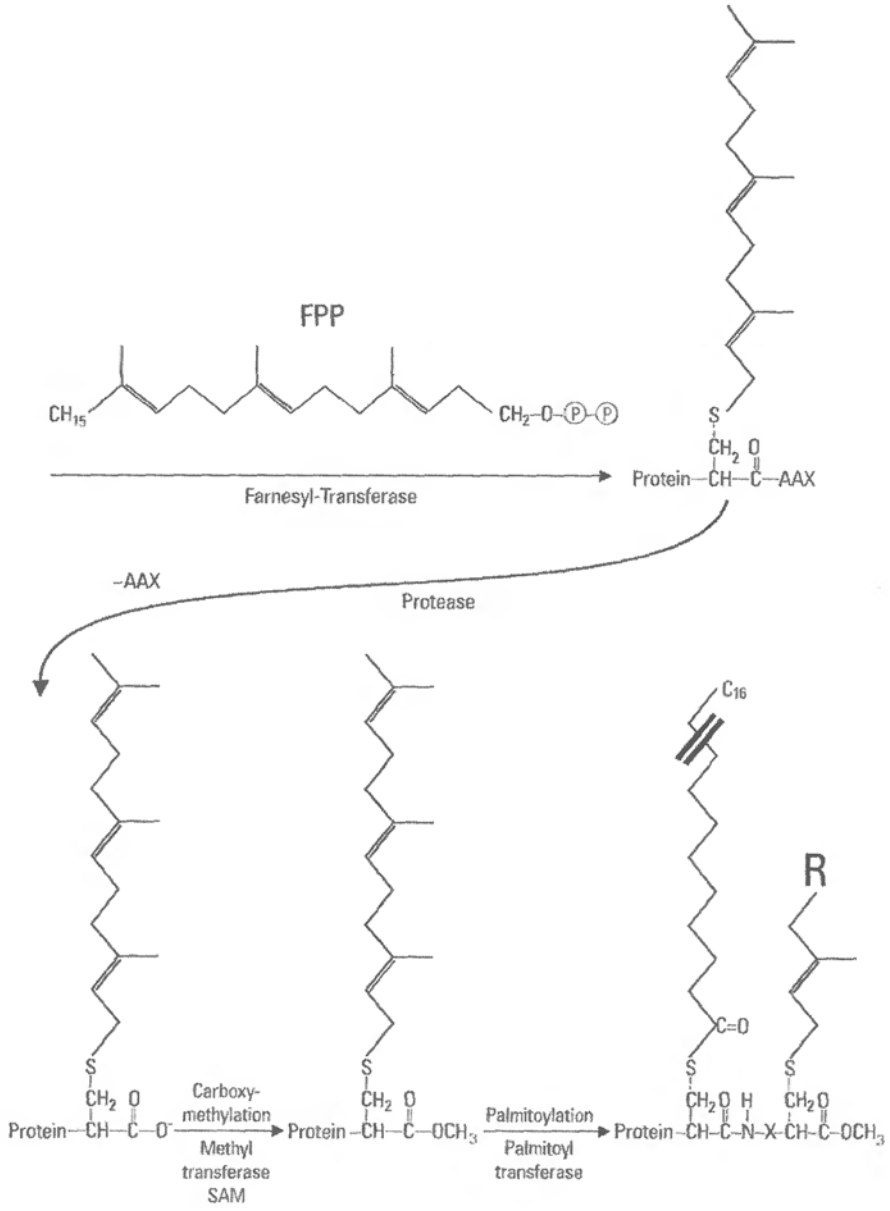


Fig. 3.15 FTase transfers the farnesyl group from farnesyl diphosphate (FPP) to the SH group of a cysteine near the COOH-terminus of the protein. The COOH-terminal tripeptide, AAX (A is an aliphatic amino acid and X is Val–Leu–Ser, or any other amino-acid residue), is removed by a protease. Subsequently, a methyl transferase donates the methyl (CH_3) group from S-adenosylmethionine (SAM) to the COOH-terminal *S*-farnesylated cysteine. The final step is the attachment of palmitoyl groups to the cysteines near the farnesylated, carboxymethylated-terminus by a specific palmitoyl transferase.

To sum up

Most regulatory proteins are modified covalently and post-translationally (after translation from mRNA). The groups attached are phosphates, sugars, and lipids. Lipids mainly serve as localization signals, targeting proteins to membranes. In many cases, relocation of cytosolic regulator proteins to the membrane is essential for their function, although some regulatory proteins have structural properties that make acylation superfluous. Co-localization of proteins on to a two-dimensional space such as the plasma membrane, may build up high local concentrations and consequently enhance the rate of encounter. Therefore, the interest of pharmaceutical chemists in manipulating protein acylation is great. However, the individual role of each different lipid modification is far from being understood. Why some proteins are modified by myristoylation, palmitoylation, or farnesylation, others by geranylgeranylation, and still others by geranylgeranylation and palmitoylation, needs to be explained.

Although lipid modifications are believed to be necessary for bringing soluble, cytoplasmic proteins to their membrane-bound partners, it is not obvious why transmembrane receptors with structural features favouring their anchorage in the lipid bilayer must be acylated (palmitoylated). Therefore, lipid modifications of proteins may have other roles than attachment to membranes. In some cases lipid modification may change the conformation of proteins, just like phosphorylation; an example is the Raf kinase, where the membranous environment activates the kinase conformationally (see Chapter 4).

References

1. W. J. Fantl, J. A. Escobedo, G. A. Martin, C. W. Turck, M. del Rosario, F. McCormick, and L. T. Williams. Distinct phosphotyrosines on a growth factor receptor bind to specific molecules that mediate different signaling pathways. *Cell*, **69** (3), 413–423, 1992.
2. T. Pawson and G. D. Gish. SH2 and SH3 domains: from structure to function. *Cell*, **71** (3), 359–362, 1992.
3. T. Pawson. Protein modules and signalling networks. *Nature*, **373** (6515), 573–580, 1995.
4. M. K. Rosen, T. Yamazaki, G. D. Gish, C. M. Kay, T. Pawson, and L. E. Kay. Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. *Nature*, **374** (6521), 477–479, 1995.
5. R. B. Birge, J. E. Fajardo, B. J. Mayer, and H. Hanafusa. Tyrosine-phosphorylated epidermal growth factor receptor and cellular p130 provide high affinity binding substrates to analyze Crk-phosphotyrosine-dependent interactions *in vitro*. *J Biol Chem*, **267** (15), 10588–10595, 1992.
6. R. Zhang, F. W. Alt, L. Davidson, S. H. Orkin, and W. Swat. Defective signalling through the T- and B-cell antigen receptors in lymphoid cells lacking the vav proto-oncogene. *Nature*, **374** (6521), 470–473, 1995.

   Also: A. Tarakhovsky, M. Turner, S. Schaal, P. J. Mee, L. P. Duddy, K. Rajewsky, and V. L. Tybulewicz. Defective antigen receptor-mediated proliferation of B and T cells in the absence of Vav. *Nature*, **374** (6521), 467–470, 1995.
7. G. Pelicci, L. Lanfrancone, F. Grignani, J. McGlade, F. Cavallo, G. Forni, *et al.* A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. *Cell*, **70** (1), 93–104, 1992.
8. Kavanaugh, W. M. and Williams, L. T. An alternative to SH2 domains for binding tyrosine-phosphorylated proteins. *Science*, **266**,1862–1865, 1994.
9. M. Rozakis Adcock, J. McGlade, G. Mbamalu, G. Pelicci, R. Daly, W. Li, *et al.* Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases. *Nature*, **360** (6405), 689–692, 1992.

10. D. Cussac, M. Frech, and P. Chardin. Binding of the Grb2 SH2 domain to phosphotyrosine motifs does not change the affinity of its SH3 domains for Sos proline-rich motifs. *EMBO J*, **13** (17), 4011–4021, 1994.

11. H. Miki, K. Miura, K. Matuoka, T. Nakata, N. Hirokawa, S. Orita, *et al.* Association of Ash/Grb-2 with dynamin through the Src homology 3 domain. *J Biol Chem*, **269** (8), 5489–5492, 1994.

12. R. Scaife, I. Gout, M. D. Waterfield, and R. L. Margolis. Growth factor-induced binding of dynamin to signal transduction proteins involves sorting to distinct and separate proline-rich dynamin sequences. *EMBO J*, **13** (11), 2574–2582, 1994.

13. K. Seedorf, G. Kostka, R. Lammers, P. Bashkin, R. Daly, W. H. Burgess, *et al.* Dynamin binds to SH3 domains of phospholipase C gamma and GRB-2. *J Biol Chem*, **269** (23), 16009–16014, 1994.

14. M. Holgado Madruga, D. R. Emlet, D. K. Moscatello, A. K. Godwin, and A. J. Wong. A Grb2-associated docking protein in EGF- and insulin-receptor signalling. *Nature*, **379** (6565), 560–564, 1996.

15. M. J. Eck, S. Dhe Paganon, T. Trub, R. T. Nolte, and S. E. Shoelson. Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell*, **85** (5), 695–705, 1996.

16. D. Anderson, C. A. Koch, L. Grey, C. Ellis, M. F. Moran, and T. Pawson. Binding of SH2 domains of phospholipase C gamma 1, GAP, and Src to activated growth factor receptors. *Science*, **250** (4983), 979–982, 1990.

17. E. Piccione, R. D. Case, S. M. Domchek, P. Hu, M. Chaudhuri, J. M. Backer, *et al.* Phosphatidylinositol 3-kinase p85 SH2 domain specificity defined by direct phosphopeptide/SH2 domain binding. *Biochemistry*, **32** (13), 3197–3202, 1993.

18. S. Felder, M. Zhou, P. Hu, J. Urena, A. Ullrich, M. Chaudhuri, *et al.* SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. *Mol Cell Biol*, **13** (3), 1449–1455, 1993.

19. G. Panayotou, G. Gish, P. End, O. Truong, I. Gout, R. Dhand, *et al.* Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor beta-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. *Mol Cell Biol*, **13** (6), 3567–3576, 1993.

20. Z. Songyang, S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell*, **72** (5), 767–778, 1993.

21. C. H. Lee, D. Kominos, S. Jacques, B. Margolis, J. Schlessinger, S. E. Shoelson, and J. Kuriyan. Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure*, **2** (5), 423–438, 1994.

22. P. S. Charifson, L. M. Shewchuk, W. Rocque, C. W. Hummel, S. R. Jordan, C. Mohr, *et al.* Peptide ligands of pp60 (c-src) SH2 domains: A thermodynamic and structural study. *Biochemistry* **36**, 6283, 1997.

23. M. J. Eck, S. E. Shoelson, and S. C. Harrison. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. *Nature*, **362** (6415), 87–91, 1993.

24. S. Sugimoto, T. J. Wandless, S. E. Shoelson, B. G. Neel, and C. T. Walsh. Activation of the SH2-containing protein tyrosine phosphatase, SH-PTP2, by phosphotyrosine-containing peptides derived from insulin receptor substrate-1. *J Biol Chem*, **269** (18), 13614–13622, 1994.

25. D. Rotin, A. M. Honegger, B. L. Margolis, A. Ullrich, and J. Schlessinger. Presence of SH2 domains of phospholipase C gamma 1 enhances substrate phosphorylation by increasing the affinity toward the epidermal growth factor receptor. *J Biol Chem*, **267** (14), 9678–9683, 1992.

26. X. Y. Fu. A transcription factor with SH2 and SH3 domains is directly activated by an interferon alpha-induced cytoplasmic protein tyrosine kinase(s). *Cell*, **70** (2), 323–335, 1992.

27. R. Ren, B. J. Mayer, P. Cicchetti, and D. Baltimore. Identification of a ten-amino acid proline-rich SH3 binding site. *Science*, **259** (5098), 1157–1161, 1993.

28. H. Yu, J. K. Chen, S. Feng, D. C. Dalgarno, A. W. Brauer, and S. L. Schreiber. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell*, **76** (5), 933–945, 1994.

29. H. Terasawa, D. Kohda, H. Hatanaka, S. Tsuchiya, K. Ogura, K. Nagata, *et al.* Structure of the N-terminal SH3 domain of GRB2 complexed with a peptide from the guanine nucleotide releasing factor SOS. *Nature Struct Biol*, **1**, 891–897, 1994.

30. M. Wittekind, C. Mapelli, V. Lee, V. Goldfarb, M. S. Friedrichs, C. A. Meyers, and L. Mueller. Solution structure of the Gbr2 N-terminal SH3 domain complexed with a ten-residue peptide derived from Sos: Direct refinement against NOES, J-coupling 1H and 13C chemical shifts. *J Mol Biol*, **267**, 933, 1997.

31. B. J. Mayer and D. Baltimore. Mutagenic analysis of the roles of SH2 and SH3 domains in regulation of the Abl tyrosine kinase. *Mol Cell Biol*, **14** (5), 2883–2894, 1994.

32. J. Downward. The GRB2/Sem-5 adaptor protein. *FEBS Lett*, **338** (2), 113–117, 1994.

33. S. E. Egan, B. W. Giddings, M. W. Brooks, L. Buday, A. M. Sizeland, and R. A. Weinberg. Association of Sos Ras exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation [see comments]. *Nature*, **363** (6424), 45–51, 1993.

Also: N. Li, A. Batzer, R. Daly, V. Yajnik, E. Skolnik, P. Chardin, *et al.* Guanine-nucleotide-releasing factor hSos binds to Grb2 and links receptor tyrosine kinases to Ras signalling. *Nature*, **363** (6424), 85–88, 1993.

Also: N. W. Gale, S. Kaplan, E. J. Lowenstein, J. Schlessinger, and D. Bar Sagi. Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on Ras. *Nature*, **363** (6424), 88–92, 1993.

34. L. Buday and J. Downward. Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell*, **73** (3), 611–620, 1999.

35. M. E. Noble, A. Musacchia, M. Saraste, S. A. Courtneidge, and R. K. Wirenga. Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J*, **12**, 2617–2624, 1993.

36. A. Musacchio, M. Saraste, and M. Wilmanns. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides [see comments]. *Nature Struct Biol*, **1** (8), 546–551, 1994.

37. S. Feng, J. K. Chen, H. Yu, J. A. Simon, and S. L. Schreiber. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science*, **266** (5188), 1241–1247, 1994.

38. M. Wittekind, C. Mapelli, B. T. N. Farmer, K. L. Suen, V. Goldfarb, J. Tsao, *et al.* Orientation of peptide fragments from Sos proteins bound to the N-terminal SH3 domain of Grb2 determined by NMR spectroscopy. *Biochemistry* **33** (46), 13531–13539, 1994.

39. Cowburn, D. Helical encounter: Insight into how Src Homology domains interact with proline-rich ligands is provided by structures of Abl-SH3 with and without peptide and a Fyn-SH3/ligand complex. *Nature Struct Biol*, **1**,489–491, 1994.

40. A. Musacchio, T. Gibson, P. Rice, J. Thompson, and M. Saraste. The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends Biochem Sci*, **18** (9), 1993.

41. G. Riddihough. More meanders and sandwiches [news]. *Nature Struct Biol*, **1** (11), 755–757, 1994.

42. L. H. Davis and V. Bennett. Identification of two regions of beta G spectrin that bind to distinct sites in brain membranes. *J Biol Chem*, **269** (6), 4409–4416, 1994.

43. M. J. Macias, A. Musacchio, H. Ponstingl, M. Nilges, M. Saraste, and H. Oschkinat. Structure of the pleckstrin homology domain from beta-spectrin. *Nature*, **369** (6482), 675–677, 1994.

44. K. M. Ferguson, M. A. Lemmon, J. Schlessinger, and P. B. Sigler. Crystal structure at 2.2 Å resolution of the pleckstrin homology domain from human dynamin. *Cell*, **79** (2), 199–209, 1994.

45. D. Timm, K. Salim, L. Gout, L. Guruprasad, M. Waterfield, and T. Blundell. Crystal structure of the pleckstrin homology domain from dynamin. *Nature Struct Biol*, **1** (11), 782–788, 1994.

46. K. M. Ferguson, M. A. Lemmon, J. Schlessinger, and P. B. Sigler. Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. *Cell*, **83** (6), 1037–1046, 1995.

47. M. Hyvoenen, M. J. Macias, M. Nilges, H. Oschkinat, M. Saraste, and M. Wilmans. Structure of the binding site for inositol phosphates in a PH domain. *EMBO J*, **14**, 4676, 1995.

48. H. S. Yoon, P. J. Hajduk, A. M. Petros, E. T. Olejniczak, R. P. Meadows, and S. W. Fesik. Solution structure of a pleckstrin-homology domain. *Nature*, **369** (6482), 672–675, 1994.

49. J. E. Harlan, P. J. Hajduk, H. S. Yoon, and S. W. Fesik. Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature*, **371** (6493), 168–170, 1994.

50. J. D. Thomas, P. Sideras, C. I. Smith, I. Vorechovsky, V. Chapman, and W. E. Paul. Colocalization of X-linked agammaglobulinemia and X-linked immunodeficiency genes. *Science*, **261** (5119), 355–358, 1993.

51. D. J. Rawlings, D. C. Saffran, S. Tsukada, D. A. Largaespada, J. C. Grimaldi, L. Cohen, *et al.* Mutation of unique region of Bruton's tyrosine kinase in immunodeficient XID mice. *Science*, **261** (5119), 358–561, 1993.

52. M. A. Lemmon, K. M. Ferguson, and J. Schlessinger. PH domains: Diverse sequences with a common fold recruit signaling molecules to the cell surface. *Cell*, **85**, 621–624, 1996.

53. Zh. Zhang, Ch-H. Lee, V. Mandiyan, J.-P. Borg, B. Margolis, J. Schlessinger, and J. Kuryan. Sequence-specific recognition of the internalization motif of the Alzheimer's amyloid precursor protein by the X11 PTB domain. *EMBO J*, **16**, 6141–6150, 1997.

54. M. Peifer, S. Berg, and A. B. Reynolds. A repeating amino acid motif shared by proteins with diverse cellular roles. [letter] *Cell*, **76** (5), 789–791, 1994.

55. M. E. Fortini and S. Artavanis Tsakonas. Notch: neurogenesis is only part of the picture. *Cell*, **75** (7), 1245–1247, 1993.

56. J. Lewis. Neurogenic genes and vertebrate neurogenesis. *Curr Opin Neurobiol*, **6** (1), 3–10, 1996.

57. T. A. Chan, C. A. Chu, K. A. Rauen, M. Kroiher, S. M. Tatarewicz, and R. E. Steele. Identification of a gene encoding a novel protein-tyrosine kinase containing SH2 domains and ankyrin-like repeats. *Oncogene*, **9** (4), 1253–1259, 1994.

58. S. Gorina and N. P. Pavletich. Structure of the p53 tumour suppressor bound to the ankyrin and SH3 domains of p53BP2. *Science*, **274**, 1001, 1996.

59. D. Stapleton, I. Balan, T. Pawson, and F. Sicheri. The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nature Struct Biol*, **6**, 44–49, 1999.

60. C. D. Thanos, K. E. Goodwill, and J. U. Bowie. Oligomeric structure of the human Eph B2 receptor SAM domain. *Science*, **283**, 833–836, 1999.

61. M. Smalla, P. Schmieder, M. Kelly, A. Ter Laak, G. Krause, L. Ball, *et al.* Solution structure of the receptor tyrosine kinase EphB2 SAM domain and identification of two distinct homotypic interaction sites. *Protein Science*, **8**, 1–8, 1999.

62. R. J. Rickles, M. C. Botfield, Z. Weng, J. A. Taylor, O. M. Green, J. S. Brugge, and M. J. Zoller. Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. EMBO J, **13** (23), 5598–5604, 1994.

63. H. Sabe, A. Hata, M. Okada, H. Nakagawa, and H. Hanafusa. Analysis of the binding of the Src homology 2 domain of Csk to tyrosine-phosphorylated proteins in the suppression and mitotic activation of c-Src. *Proc Natl Acad Sci, USA*, **91** (9), 3984–3988, 1994.

64. G. Superti Furga, S. Fumagalli, M. Koegl, S. A. Courtneidge, and G. Draetta. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *EMBO J*, **12** (7), 2625–2634, 1993.

65. F. Sicheri, I. Moarefi, and J. Kuriyan. Crystal structure of the Src family tyrosine kinase Hck. *Nature*, **385**, 602–609, 1997.

66. T. Pawson. Protein tyrosine kinases: New impressions of Src and Hck. *Nature*, **385**, 582–585, 1997.

67. W. Xu, S. C. Harrison, and M. J. Eck. Three-dimensional structure of the tyrosine kinase c-Src. *Nature*, **385**, 595–602, 1997.

68. J. Zheng, *et al.* 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with Mn ATP and a peptide inhibitor. *Acta Crystallogr*, **D49**, 362–365, 1993.

69. H. Yamaguchi and W. A. Hendrickson. Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature*, **384** (6608), 484–489, 1996.

70. S. Mori, L. Ronnstrand, K. Yokote, A. Engstrom, S. A. Courtneidge, L. Claesson Welsh, and C. H. Heldin. Identification of two juxtamembrane autophosphorylation sites in the PDGF beta-receptor; involvement in the interaction with Src family tyrosine kinases. *EMBO J*, **12** (6), 2257–2264, 1993.

71. M. Kawabuchi, Y. Satomi, T. Takao, Y. Shimonishi, S. Nada, K. Nagai, *et al.* Transmembrane phosphoprotein Cbp regulates the activities of Src-family tyrosine kinases. *Nature*, **404**, 999–1003, 2000.

72. L. A. Cary and J. A. Cooper. Molecular switches in lipid rafts. *Nature*, **404**, 945–947, 2000.

73. E. H. Fischer, H. Charbonneau, and N. K. Tonks. Protein tyrosine phosphatases: a diverse family of intracellular and transmembrane enzymes. *Science*, **253**, 401–406, 1989.
Also: N. K. Tonks, C. D. Diltz and E. H. Fischer, Purification of the major protein-tyrosine-phosphatases of human placenta. *J. Biol. Chem.* **263**, 6722–6730. 1988

74. T. Yi, A. L. Mui, G. Krystal, and J. N. Ihle. Hematopoietic cell phosphatase associates with the interleukin-3 (IL-3) receptor b chain and down-regulates IL-3-induced tyrosine phosphorylation and mitogenesis. *Mol Cell Biol*, **13** (12), 7577–7586, 1993.

75. L. J. Mauro and J. E. Dixon. 'Zip codes' direct intracellular protein tyrosine phosphatases to the correct cellular 'address'. *Trends Biochem Sci*, **19** (4), 151–155, 1994.

76. E. B. Fauman and M. A. Saper. Structure and function of the protein tyrosine phosphatases. *Trends Biochem Sci*, **21** (11), 413–417, 1996.

77. B. Sebastian, A. Kakizuka, and T. Hunter. Cdc25M2 activation of cyclin-dependent kinases by dephosphorylation of threonine-14 and tyrosine-15. *Proc Natl Acad Sci, USA*, **90** (8), 3521–3524, 1993.

78. D. Barford, A. J. Flint, and N. K. Tonks. Crystal structure of human protein tyrosine phosphatase 1B [see comments]. *Science*, **263** (5152), 1397–1404, 1994.

79. Z. Jia, D. Barford, A. J. Flint, and N. K. Tonks. Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science*, **268** (5218), 1754–1758, 1995.

80. J. A. Stuckey, H.L. Schubert, E. B.Fauman, Z. Y. Zhang, J. E. Dixon, and M. A. Saper. Crystal structures of *Yersinia* protein tyrosine phosphatase at 2.5 Å and the complex with tungstate. *Nature*, **370**, 571, 1994.

81. E. B. Fauman, Ch. Yuvaniyama, H. L. Schubert, J. A. Stuckey, and M. A. Saper. The X-ray crystal structures of *Yersinia* tyrosine phosphatase with bound tungstate and nitrate. Mechanistic implications. *J Biol Chem*, **271**, 18780–18788, 1996.

82. A. Kharitonenkov, Z. Chen, I. Sures, H. Wang, J. Schilling, and A. Ullrich. A family of proteins that inhibit signalling through tyrosine kinase receptors. *Nature*, **386**, 181–186, 1997.

83. E. Peles, M. Nativ, P. L. Campbell, T. Sakurai, R. Martinez, S. Lev, *et al.* The carbonic acid anhydrase domain of receptor tyrosine phosphatase b is a functional ligand for the axonal cell recognition molecule contactin. *Cell*, **82**, 251–260, 1995.

84. J. P. Noel, H. E. Hamm, and P. B. Sigler. The 2.2 Å crystal structure of transducin-alpha complexed with GTP gamma S. *Nature*, **366** (6456), 654–663, 1993 [Comment in *Nature*, **366** (6456), 628–629, 1993].

85. D. G. Lambright, J. P. Noel, H. E. Hamm, and P. B. Sigler. Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. *Nature*, **369** (6482), 621–628, 1994 [Comment in *Nature*, **369** (6482), 611–612].

86. D. G. Lambright, J. Sondek, A. Bohm, N. P. Skiba, H. E. Hamm, and P. B. Sigler. The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature*, **379**, 311–319, 1996.

87. D. W. Markby, R. Onrust, and H. R. Bourne. Separate GTP binding and GTPase activating domains of a G alpha subunit. *Science*, **262** (5141), 1895–1901, 1993.

88. H. R. Bourne. G proteins. The importance of being GTP [comment]. *Nature*, **369** (6482), 611–612, 1994.

89. D. E. Coleman, A. M. Berghuis, E. Lee, M. E. Linder, A. G. Gilman, and S. R. Sprang. Structure of active conformation of Gia1 and the mechanism of GTP hydrolysis. *Science*, **265**, 1405, 1994.

    Also: M. A. Wall, D. E. Coleman, E. Lee, J. A. Inignez Lhui, B. A. Posner, A. G. Gilman, and S. R. Sprang, The structure of the G protein heterotrimer G_i alpha 1 beta 1 gamma 2. Cell, 83m (6), 1047–1058, 1995.

90. M. P. Graziano and A. G. Gilman. Synthesis in *Escherichia coli* of GTPase-deficient mutants of Gs alpha. *J Biol Chem*, **264** (26), 15475–15482, 1989.

    Also: M. Freissmuth and A. G. Gilman. Mutations of Gs alpha designed to alter the reactivity of the protein with bacterial toxins. Substitutions at Arg187 result in loss of GTPase activity. *J Biol Chem*, **264** (36), 21907–21914, 1989.

91. C. A. Landis, S. B. Masters, A. Spada, A. M. Pace, H. R. Bourne, and L. Vallar. GTPase inhibiting mutations activate the alpha chain of Gs and stimulate adenylyl cyclase in human pituitary tumours. *Nature*, **340** (6236), 692–696, 1989.

92. J. Lyons, C. A. Landis, G. Harsh, L. Vallar, K. Grunewald, H. Feichtinger, *et al.* Two G protein oncogenes in human endocrine tumors. *Science*, **249** (4969), 655–659, 1990.

93. M. Barbacid. ras genes. *Annu Rev Biochem*, **56**, 779–827, 1987.

94. D. Cassel and T. Pfeuffer. Mechanism of choleratoxin action: Covalent modification of the guanyl nucleotide-binding protein of the adenylate cyclase system. *Proc Natl Acad Sci, USA*, **75**, 2669, 1978.

95. G. W. Sharp and S. Hynie. Stimulation of intestinal adenylcyclase by choleratoxin. *Nature*, **229**, 266–269, 1971.

96. J. Sondek, D. G., Lambright, J. P. Noel, H.E. Hamm, and P. B. Sigler. GTPase mechanism of G proteins from the 1.7 Å crystal structure of transducin α GDP AlF_4^-. *Nature*, **372**, 276–279, 1994.

97. K. Scheffzek, R. M. Ahmadian, W. Kabsch, L. Wiesmüller, A. Lautwein, F. Schmitz, and A. Wittinghofer. The Ras–Ras GAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*, **277**, 333–338, 1997.

98. E. P. Pai, U. Krengel, G. A. Petsko, R. S. Goody, W. Kabsch, and A. Wittinghofer. Refined crystal structure of the triphosphate conformation of H-*ras* p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J*, **9**, 2351–2359, 1990.

99. J. Downward, J. D. Graves, P. H. Warne, S. Rayter, and D. A. Cantrell. Stimulation of p21ras upon T-cell activation [see comments]. *Nature*, **346** (6286), 719–723, 1990.

100. S. Ozonoff. Cognitive impairment in neurofibromatosis type 1. *Am J Med Genet*, **89**, 45–52, 1999.

101. M. S. Grotewiel, C. D. Beck, K. H. Wu, X. R. Zhu, and R. L. Davis. Integrin-mediated short-term memory in Drosophila. *Nature*, **391** (6666): 455–460, 1998.

102. H.-F. Guo, T. Inge, F. Hannan, A. Bernards, and Y. Zhong. Requirement of Drosophila NF1 for activation of adenylyl cyclase by PACAP38-like neuropeptides. *Science*, **276** (5313): 795–798, 1997.

103. H.-F. Guo, J. Tong, F. Hannan, L. Luo, and Y. Zhong. A neurofribromatosis-1-regulated pathway is required for learning in *Drosophila*. *Nature*, **403**, 895–898, 2000.

104. R. L. Davis. Neurofibromin progress on the fly. *Nature*, **403**, 846–847, 2000.
Also: R. L. Davis. Physiology and biochemistry of Drosophila learning mutants. *Physiol Rev*, **6** (2) 299–317, 1996.

105. S. Tanaka, T. Morishita, Y. Hashimoto, S. Hattori, S. Nakamura, M. Shibuya, *et al.* C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. *Proc Natl Acad Sci, USA*, **91** (8), 3443–3447, 1994.

106. J. Downward. Signal transduction. New exchange, new target. *Nature*, **396**, 416–417, 1998.

107. M. Matsuda, Y. Hashimoto, K. Muroya, H. Hasegawa, T. Kurata, S. Tanaka, S. Nakamura, and S. Hattori. CRK protein binds to two guanine nucleotide-releasing proteins for the Ras family and modulates nerve growth factor-induced activation of Ras in PC12 cells. *Mol Cell Biol*, **14** (8), 5495–5500, 1994.

108. H. M. Lander, A. J. Milbank, J. M. Tauras, D. P. Hajjar, B. L. Hempstead, G. D. Schwartz, *et al.* Redox regulation of cell signalling [Letter]. *Nature*, **381** (6581), 380–381, 1996.

109. J. F. Hancock, A. I. Magee, J. E. Childs, and C. J. Marshall. All ras proteins are polyisoprenylated but only some are palmitoylated. *Cell*, **57** (7), 1167–1177, 1989.

110. F. L. Zhang and P. J. Casey. Protein prenylation: molecular mechanisms and functional consequences. *Annu Rev Biochem*, **65**, 241–269, 1996.

111. K. Yokoyama and M. H. Gelb. Purification of a mammalian protein geranylgeranyltransferase. Formation and catalytic properties of an enzyme-geranylgeranyl pyrophosphate complex. *J Biol Chem*, **268** (6), 4055–4060, 1993.

112. H. W. Park, S. R. Boduluri, J. F. Moomaw, P. J. Casey, and L. S. Beese. Crystal structure of protein farnesyltransferase at 2.25 angstrom resolution [see comments]. *Science*, **275** (5307), 1800–1804, 1997. [Published erratum appears in *Science*, **276**, (5309), 21, 1997.]

113. K. S. Koblan, N. E. Kohl, C. A. Omer, N. J. Anthony, M. W. Conner, S. J. deSolms, *et al.* Farnesyltransferase inhibitors: A new class of cancer chemotherapeutics. *Biochem Soc Trans*, **24** (3), 688–692, 1996.

114. V. L. Boyartchuk, M. N. Ashby, and J. Rine. Modulation of Ras and a-factor function by carboxyl-terminal proteolysis [see comments]. *Science*, **275** (5307), 1796–1800, 1997.

115. J. A. Glomset and C. C. Farnsworth. Role of protein modification reactions in programming interactions between ras-related GTPases and cell membranes. *Annu Rev Cell Biol*, **10**, 181–205, 1994.

4

Signal transduction pathways through small monomeric G proteins

Ras-controlled signal transduction: the MAP kinase and the PtdIns 3-kinase pathways

Having introduced Ras, the master switch of the MAP kinase (mitogen-activated protein kinase) signalling cascades, we are now ready to put the pieces of the puzzle together and sum up the essential features of a typical cellular signalling pathway, the Ras/MAP kinase phosphorylation cascade.

The Ras/MAP kinase cascade is a main signalling route for growth factors. It drives cell-cycle progression, proliferation, and differentiation, and may also transmit signals that order cells to die. Rather than describe all possible signal inputs and cellular responses, I shall concentrate on two aspects: (1) how the switch function of Ras controls signal input; and (2) how the signal is propagated by the MAP kinases (Fig. 4.1).

The phosphorylation cascade consists of a sequence of kinases, starting with MAPKKKs, (MAP kinase kinase kinases) followed by MAPKKs (MAP kinase kinases), and ending with MAPKs (MAP kinases).

Representative of upstream kinases which accept receptor signals are the Raf kinases. They are cytosolic Ser/Thr kinases. Raf kinase is recruited to the membrane and its activity is controlled by Ras. Raf phosphorylates and activates dual specificity, e.g. Ser/Thr- and Tyr-MAPKKs. The MAPKKs have a conserved serine/threonine phosphorylation site which is recognized and phosphorylated by Raf.

Signal input

The role of the Raf kinase

Whereas, activation of MAPKKs and MAPKs (also known as extracellular signal-responsive kinases, ERKs), through *trans*-phosphorylation reactions is fairly well understood, activation of the upstream MAPKKKs, such as the Raf family of protein kinases, is more complicated. There are at least three human Raf kinases: Raf-1, Raf-A, and Raf-B. The activation of the Raf kinase is controlled by Ras. This is a central point of control of the whole MAP kinase cascade.

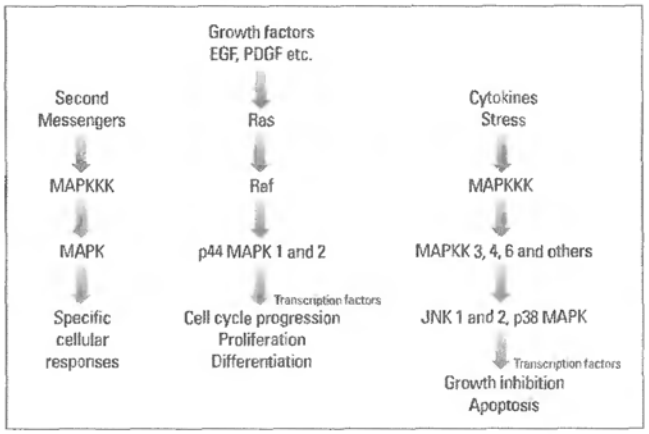


Fig. 4.1 Growth-factor signals are channelled in the Raf/Ras pathway, resulting in cellular proliferation and differentiation, whereas stress and cytokine signals are directed via MAPKKKs and MAPKKs to JNKs 1 and 2 and the p38 MAP kinase. They cause growth inhibition and apoptosis. Second messengers, such as lipid messengers, may also adress the MAP kinase cascade and elicit specific cellular responses. Included in the scheme are upstream kinases, such as Raf, that in turn activate the MAPKKs (not shown), and eventually the p44 MAPKs 1 and 2 (or ERKs). (This scheme is simple and not complete, because some kinases participating in these signalling pathways are not yet defined.)

Ras is attached to the inner surface of the cell membrane, through its prenyl residues; it brings the Raf kinase to the membrane. This translocation to the membrane is essential for the activation of the Raf kinases. Raf is not a substrate of receptor tyrosine kinases. Raf-1 mutants that lack tyrosine phosphorylation sites still have biological activity and Raf-B does not even have tyrosine phosphorylation sites. Thus, Raf can only respond to receptor signals through the intervention of Ras. Besides Ras, other proteins contribute to the activation of Raf. Candidates for such a role are linkers, such as the so-called 14-3-3 proteins. The name comes from the position of the bands in gel electrophoresis. This group of proteins has been implicated in the regulation of protein kinase C and the Raf kinase. Binding of linker proteins to Raf can bring other regulatory proteins in contact with Raf , increasing the signalling repertoire. The details of the activation of Raf still need to be worked out. Although, Raf-1 does form dimers in cells treated with growth factors, it is not clear whether the interaction of Ras in its activated GTP-bound form with Raf-1 is responsible for dimerization, nor is it clear whether oligomerization of Raf is associated with autophosphorylation of serine/threonine sites.[1]

Other Ras-controlled kinases

It is expected that upstream kinases other than Raf kinases, and G proteins other than Ras, may be engaged in signal input because more than one signal or signal pathway is required for the full activation of the early response genes for cell proliferation. Other cell-type-specific upstream kinases exist. An example is the role of a novel family of protein tyrosine kinases, named Pyk, identified by J. Schlessinger's laboratory.[2] Pyks are highly expressed in neuronal cells. They participate in a cell-type-specific upstream phosphorylation cascade, where Pyk1 phosphorylates and activates Pyk2. The Pyks respond to a bewildering variety of different signals, including lipid messengers and other signals, transmitted through G-protein-linked receptors. They are responsive to TNF-α and stress factors, such as UV-irradiation and changes in osmolarity. Because some of these factors also

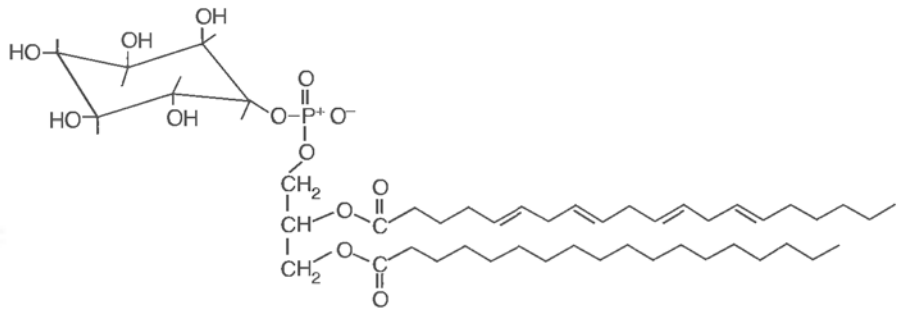
increase intracellular calcium, it was reasoned that Pyks may help to bring calcium channels under the control of the JNK (Jun N-terminal kinase)/MAPK pathways.

Another common signalling route is the phosphatidylinositol 3-kinase pathway. Class I PtdIns 3-kinases, PI 3-kinases, are heterodimers, consisting of a catalytic and a regulatory subunit. Regulatory subunits of the Class I_A PtdIns 3-kinases (the p85 subunits) serve as adaptors and bind cytosolic tyrosine kinases through their SH2 and SH3 domains (Table 4.1).

Phosphatidylinositols and their phosphorylated derivatives are substrates for several PtdIns kinases. For example, PtdIns is converted by PtdIns 3-kinase to PtdIns-3-P, by PtdIns 4-kinase to PtdIns-4-P, and by PtdIns 5-kinase to PtdIns-5-P. Moreover, PtdIns-3-P is converted by phosphatidylinositol phosphate 4-kinase, (PIP 4-kinase), to PtdIns-3,4-P_2, and PtdIns-4-P is converted by PIP 3-kinase to PtdIns-3,4-P_2 and to PtdIns-4,5-P_2 by PIP 5-kinase. Finally, PtdIns-4,5-P_2 is hydrolysed by phospholipase C to diacylglycerol (DAG) and IP_3, and the PtdIns bisphosphates (PtdIns-3,4-P_2 and

Table 4.1 Summary of the structural properties of the three classes of phosphoinositide kinases (PI 3-kinases) and their substrates

Phosphatidylinositol



Phosphoinositide kinases

PI 3-kinases	Catalytic subunit	Regulatory subunit		Substrates
Class I_A	3 isoforms	3α	isoforms	PtdIns-4,5-P_2
		1β		PtdIns-4-P
		1γ		PtdIns
				PtdIns-5-P
Class I_B	1 isoform	1 subunit		PtdIns-4,5-P_2
				PtdIns-4-P
				PtdIns
Class II	2 isoforms	—		PtdIns
				PtdIns-4-P
Class III	1 subunit, no isoforms	1 subunit		PtdIns

It should be noted that phosphatidylinositol, PtdIns, itself is not considered to be a proper phosphoinositide. Therefore, enzymes which phosphorylate only PtdIns, such as the Class III PI 3-kinases, are not strictly phosphoinositide kinases, although they have sequence homologies with other PI 3-kinases and are therefore included in the superfamily of phosphoinositide kinases. This information is taken from ref. 3, with permission of the authors and Annu. Rev. Biochem.

PtdIns-4,5-P_2) are converted to the trisphosphate, PtdIns-3,4,5-P_3, by PIP 5-kinase and PIP 3-kinase, respectively.

PtdIns 3-kinases and the lipid messengers they produce play an important role in growth regulation and in several other cellular responses, including the insulin response (Chapter 8). The Class 1_A PtdIns 3-kinases are regulated by Ras. Ras–GTP binds to the regulatory subunit of the kinase. Cells expressing Ras mutants that can not interact with PtdIns 3-kinases have lost certain responses.

MAPKKs (MEKs)

The upstream Raf kinases transmit signals to the MAPKKs. MAPKKs phosphorylate and activate the MAPKs, which seem to be their only substrates. They are remarkably specific. The structural properties that determine the specific recognition of a particular downstream MAP or JNK kinase remain to be clarified. The specificity and selectivity of substrate recognition by the MAPKKs is in striking contrast to the great variety of signals to which Raf and other kinases, upstream of the MAPKKs, respond. Thus, the MAPKKs, filter and sort the signals. Here is one of the controls, preventing cross-reactions. This is important, because Ras/Raf/MAP kinase cascades are addressed by many signals, each with a different message and each capable of eliciting a different biological response, even in the same cell.

MAP kinases

The last step in the phosphorylation cascade involves the MAP kinases (mitogen-activated protein kinases) or ERKs (extracellular signal-responsive kinases). There are two major kinds of MAP kinases: the MAPKs and the JNKs (Jun N-terminal kinases; Jun is a universal, growth-regulating transcription factor) or SAPKs (stress-activated protein kinases). There are six MAP kinases in yeast, and a number of mammalian MAP kinases have been cloned and more are to be expected. The MAP kinases are serine-threonine kinases and are activated by phosphorylation of a tyrosine and a threonine residue.

MAP kinase pathways and cellular responses[4,5]

Three major mammalian Ras/MAPK routes have been identified, each with a different terminal MAP kinase, and each associated with a different cellular response:

(1) A pathway with p42–p44 MAPKs;

(2) a cascade with p38 MAPK; and

(3) the pathways using p46–p54 JNKs or SAPKs.

Each signalling pathway is Ras-controlled.

MAPKs are highly specific in their selection of substrates. Each member of the two MAPK families phosphorylates different substrates. The JNKs/SAPKs, and also the p38 MAPK, transmit signals mainly in response to cytokines and environmental stress.[6] Growth factors turn on the activation of the p42/p44 class of MAPKs which regulate cell proliferation and drive cell-cycle progression.

Structure of an MAP kinase

The three-dimensional structure of the unphosphorylated form of MAP kinase p38 is similar to that of other protein kinases (Fig. 4.2).

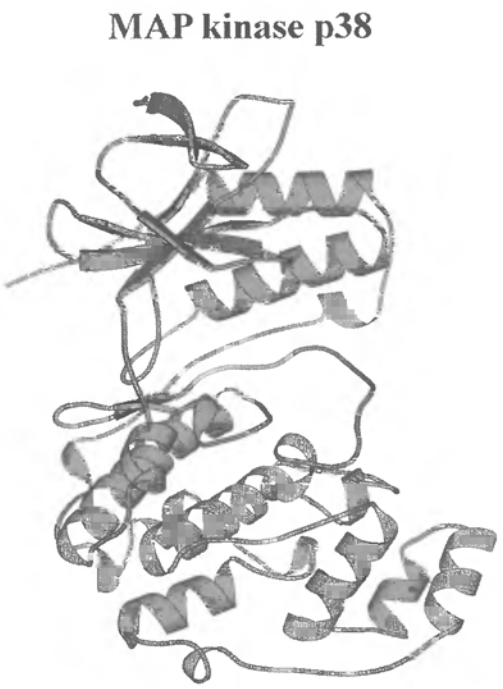


Fig. 4.2 The monomeric enzyme consists of a smaller N-terminal and a larger C-terminal domain, connected by a linker peptide. The catalytic site to which ATP binds is deeply embedded between the interface of the two domains. (The information is taken from PDB files and reproduced with permission of Professor E. J. Goldsmith and the Proc. Natl. Acad. Sci. USA from ref. 7.)

Activation of MAP kinase p38

Phosphorylation of the kinase leads to activation and triggers rather substantial conformational changes, which result in a rotation of the N- and C-terminal domains. MAP kinases are phosphorylated on threonines and on tyrosines. The rotation realigns the catalytic residues in the active site. In analogy with the structure of the cAMP-dependent kinase, it was suggested that the domain rotation brings a basic residue in the N-terminal domain (e.g. Arg65 in MAPK 2) into a position that allows it to bind to the phosphothreonyl (e.g. phosphothreonyl 183). In a similar fashion, the side-chain of the phosphorylated Tyr185 is relocated by a local conformational change, allowing its interaction with Arg189 and Arg192. These changes form a structure of the active site which is suited for high catalytic activity.

Regulation of MAP kinase p38

A discrete region of 40 residues in the amino-terminal part of p38 MAPK is mainly responsible for screening of extracellular signals, transmitted by the upstream kinases, whereas the carboxy-terminal half of the MAPK binds to downstream substrates, such as transcription factors. The amino-terminal recognition region of the MAPK contains an exposed α-helix in the proximity of the catalytic cleft, which is the phosphorylation site recognized by an upstream kinase. Phosphorylation opens the catalytic cleft and activates the kinase. This seems to be a common structural feature of kinases participating in sequential reactions in phosphorylation cascades.

Activation of MAP kinase by nuclear translocation

Whereas the upstream kinases of the MAP kinase phosphorylation cascade are cytoplasmic, the MAP kinase is translocated to the nucleus. Nuclear translocation of the MAP

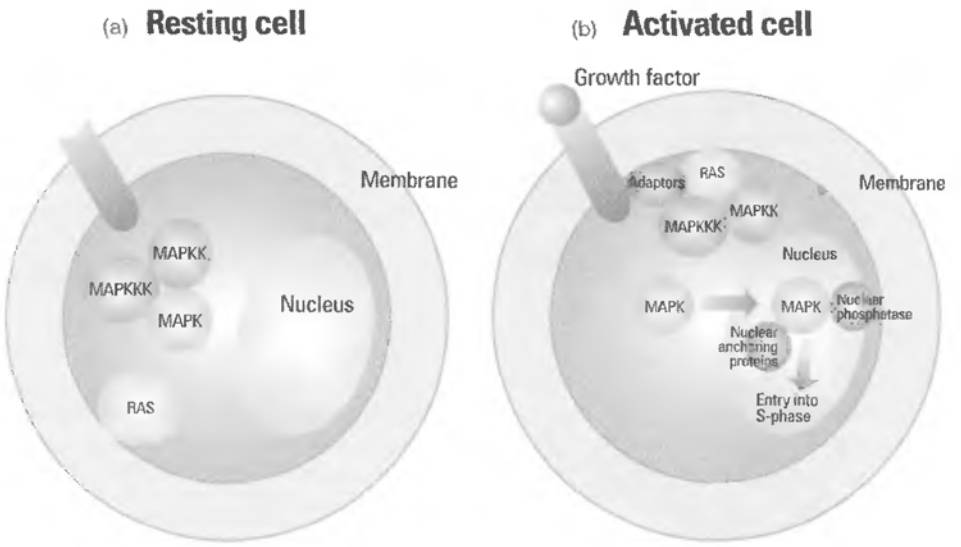


Fig. 4.3 Pouysségur and colleagues[8] found that in the resting dormant state of a cell (a) the MAP kinases are kept in the cytoplasm by interaction with the upstream cytosolic kinases of the MAP kinase cascade. (b) Activation of the MAP kinase cascade by a growth factor uncouples the MAPK from upstream MAPKKs and initiates translocation of MAPK to the nucleus, where it signals entry into the S phase of the cell cycle. In the nucleus, MAPK is retained by short-lived nuclear anchoring proteins. When their proteolytic removal is blocked, the residence time of MAPK in the nucleus is prolonged.

kinase is necessary for a positive response to growth-promoting factors. When the p44 MAP kinase is retained in the cytoplasm, entry into the S phase of the cell cycle and DNA synthesis do not occur (Fig. 4.3).

There is a delicate relationship between the duration of MAPK signalling and growth promotion. Two factors are involved:

1. The activity of MAP kinases is poised and kept in check by several MAP kinase phosphatases. This is a kind of feedback regulation, such that when the early genes involved in cell growth and proliferation are turned on by growth-factor signals through the MAP kinase cascade, the same genes are called up to express MAP kinase phosphatases.
2. Signalling by different growth factors leads to longer or shorter residence times of the MAP kinases in the nucleus. Eventually as the activity of growth-promoting factors declines, the activity of the MAP kinases also abates. Thus, the temporal activation pattern determines the biological response.

The functional repertoire of MAP kinases

We must admit that we do not yet know all the main routes of the Ras/MAPK pathway for the control of cell growth. A recent report makes that clear. Graves *et al.*[9] have discovered a new target for MAP kinase: an enzyme that provides the building bricks for the synthesis of RNA and DNA—a reaction of central importance. As we shall discuss in more detail

in Chapter 12, cell proliferation involves the replication of the cell's chromosomes. Chromosome replication requires *de novo* formation of DNA, whereas the synthesis of proteins for the newly formed cell requires formation of messenger RNAs and other forms of RNA. In both cases an adequate supply of nucleotides is needed. Graves *et al.*[9] found a link between an enzyme involved in the synthesis of pyrimidine nucleotides and the Ras/MAPK pathway. Since both DNA and RNA contain pyrimidine nucleotides (cytosine, thymine, and uracil), the regulation of carbamoyl phosphate synthase II (CPS II), which catalyses the initial, rate-limiting step in the *de novo* synthesis of pyrimidine nucleotides by MAP kinase 1, is obviously very important. In cells treated with EGF, MAP kinase activates CPS II by phosphorylation of a threonine. CPS is a prototypical, allosterically regulated enzyme. Phosphorylation makes CPS II more responsive to the allosteric activator, phosphoribosyl pyrophosphate (PRPP) and less responsive to allosteric inhibition by uridine triphosphate (UTP), an end product of pyrimidine nucleotide synthesis. Since the threonine phosphorylation site is conserved in CPS I, a closely related enzyme which participates in the synthesis of urea, it seems quite likely that MAP kinase also regulates this enzyme and the synthesis of urea.

But, this is not the only way by which MAP kinase promotes cell growth and proliferation, it also turns on a group of transcription factors that activate the so-called rapid or early response genes, encoding growth-promoting proteins.[10] As we shall discuss in Part 2, transcription factors regulate the transcription of genes into pre-messenger RNAs, the precursors of messenger RNAs, from which the information is then translated into proteins. Gene transcription is activated either directly by phosphorylation of transcription factors by MAP kinase or by opening access of transcription factors to the genes in the chromosomes.[11,12] Finally, MAP kinase appears to play a role not only in the transcription of DNA into mRNAs, but also in the regulation of protein synthesis. MAP kinase helps to recruit factors required for protein synthesis and activates, by phosphorylation, the eukaryotic translation initiation factor-4E (eIF-4E).[13,14]

In principle, all these processes are ancillary to the cell cycle, the most important point of control of cell growth and proliferation. Therefore it is not surprising that the MAP kinase pathway also regulates the cell cycle. When the MAP kinase signalling pathway is turned on in response to growth-factor signals, transcription of the cyclin D1 gene is increased.[15] Formation of cyclin D1 then promotes the formation of an active cyclin D1–Cdk4 complex,[16] which is rate limiting for cell-cycle progression and growth and proliferation. This will be discussed in more detail in Chapter 12.

What guarantees the specificity of the biological response?

The nearly unlimited functional repertoire of MAP kinase signalling requires strict control. The Ras/Raf/MAPK pathway is constructed in a way to guarantee control. This explains why constitutively active MAPK mutants have never been found. The presence of a constitutively active, unregulated MAPK would likely be deleterious for the cell. Accordingly, when ERK/MAPK is switched on all the time, the result is unregulated growth of fibroblast cells and tumour formation in mice,[17,18] and when constitutively active mutants of the MAPKKs that phosphorylate and activate the MAPKs, are expressed in mice, the cells are transformed and tumours arise. Conversely, thymocytes (immature T cells) isolated from mice that lack one member of the ERK/MAPK family, ERK-1, grow more slowly in response to growth-promoting stimuli than do thymocytes isolated from normal mice.[19]

The remarkable specificity of recognition among kinases engaged in the *trans*-phosphorylation reactions of the MAP kinase cascade guarantees the specificity of the biological response. Each MAP kinase is activated and phosphorylated by a different class of upstream kinases. For example, MAPKK 1 and 2 phosphorylate only p44 MAPKs, whereas MAPKK 3 phosphorylates only p38 MAPK, and MAPKK 4 and 6 phosphorylate only the JNKs. There is no cross-talk. The specificity of kinase interactions is matched by the selectivity of the signal response: in the same cell, p44 MAPK is preferentially activated by growth factors such as EGF and PDGF and serum, whereas p38 MAPK and JNK are activated by stress-inducing agents and cytokines. But, whereas the involvement of the ERK/MAPK family of growth-factor-activated kinases in the regulation of the cell is beyond doubt, although perhaps not yet known in all details, the role of the stress-activated MAPKs and of the JNKs is not so clear. However, it is known that the p44 MAPKs, the p38 MAPKs, and the JNKs each address different transcription factors and different genes.

Summary

1. The Ras/Raf/MAPK pathway is the prototype of a phosphorylation cascade. Understanding the construction principle of such a signalling cascade should make it easier to deal with other signalling pathways. The complexity of cellular signalling networks is, in most cases, not a consequence of redundancy. The design has a purpose. Each new sensor, linker, and receiver increases the range of signals and the scope of cellular responses. The amazing versatility of cellular signalling is the result of 'evolutionary tinkering',[20] ever improving function by adapting the signalling network to new challenges the cell has to face and to needs that have to be met, so that the cell and the organism can survive and grow in a changing environment.

2. I have presented information on how activation of MAP kinase by growth-factor stimulation of cells leads to the expression of genes promoting growth and proliferation, and I have referred to targets, controlled by MAP kinases which are central for the regulation of cell growth and proliferation. The pivotal importance of the control of DNA, RNA, and protein synthesis is demonstrated by the fact that the tools required, such as CPS II and its allosteric regulator (PRPP), eIF-4e, and others, are all highly active in many types of rapidly proliferating cancer cells (see Part 4).

3. Although the role of the MAP kinase phosphorylation cascade in the control of cell growth and cellular proliferation is undisputed, other pathways and pathway ramifications contribute to that control. This requires communication between signal transduction pathways. MAP kinase also seems to play a part in pathway interconnections, for example in connecting signals from G-protein-coupled receptors with signals from growth-factor-activated receptors (discussed in Chapter 5).

Control of signalling by the Rho/Rac/Cdc42 GTPases

The Ras superfamily of small monomeric GTP-binding proteins consists of at least 13 small GTPases. We have already mentioned the enormous functional diversity of the monomeric G proteins, of which many aspects are still unclear (see Chapter 3). For example, the Rab GTPases regulate vesicular protein traffic and direct the transport vesicles to their 'targets in eukaryotic cells'.[21] Another monomeric G protein, Ran, regulates transport in and out of the nucleus. Moreover, Ran in the active, GTP-bound form

induces microtubule organization and formation of the mitotic spindle in *Xenopus* eggs (see Chapter 12).[22] There is also new information from R. A. Cerioni's laboratory at Cornell University on the role of Cdc42.[23] By expressing mutants of Cdc42 in cells they showed that this monomeric G protein in the active, GTP-bound form directly interacts with part of the machinery (the cytoplasmic coatomer complex) that coats membrane regions and promotes budding and formation of vesicles. The assembly of the coatomer complex on the membrane is controlled by Arf. (Arf is a small monomeric G protein. Its name, ADP-ribosylation factor comes from its role in assisting choleratoxin-mediated ADP-riborylation of the heterotrimeric G-protein G_s). In that way Cdc42 regulates vesicle transport between the endoplasmic reticulum and the Golgi apparatus. But the interaction of Cdc42 with a subunit of the coatomer complex also gives instructions to cells leading to malignant transformation. Although the Cdc42 signal differs from Ras signals, the route the Cdc42 signal travels, via interaction with the coatomer complex, to rearrange the programme controlling cell growth, proliferation, and cell-cycle progression remains to be clarified.

There is no doubt that each of these multiple actions is important and of great interest, but I had to make a quite arbitrary choice in selecting for discussion only one of the many functions of monomeric G proteins in the control of fundamental cellular processes. I have chosen the role of the Rho/Rac/Cdc42 GTPases in the regulation of cell proliferation, mainly because it gave me an opportunity to introduce cell–cell interactions and the role of the cytoskeleton.

Rho/Rac/Cdc42 GTPases belong to a subgroup of the Ras superfamily of small monomeric GTP-binding proteins, comprising Rho's A, B, C, and G, Rac 1 and Rac 2, and Cdc42. The activity of these GTPases is regulated, like the activity of the Ras GTPases, by GAPs and GEFs.[24,25] The Rho/Rac/Cdc42 GTPases control and activate the PAK family of serine/threonine kinases,[26] just like Ras controls Raf. (PAK is the abbreviation of p21-activated kinase. p21 are monomeric G proteins with a Mrw. of 21 kDa) The PAKs have binding sites for Cdc42 and Rac and a catalytic domain, homologous to that of protein kinase C. (PKC is the membrane associated protein kinase activated by DAG, diacylglycerol). In muscle dystrophy, a mutated form of a human cytosolic serine/ threonine protein kinase, closely related to PAK, has been implicated in the pathology.

Good candidates for downstream kinases in a Rho/Rac/Cdc42-controlled phosphorylation cascade are the JNKs, rather than the MAP kinases. The Rho/Rac 1/Cdc42 G proteins transmit signals to the stress-responsive JNKs much more effectively than Ras. The Rho/Rac/Cdc42-GTPases regulate the cytoskeleton,[27] and respond to lysophospholipid, and growth factor, and hormonal signals. They also participate in cell-cycle regulation.[28]

The assembly of the cytoskeleton is controlled by Rho, Rac, and Cdc42

Each of the three small GTPases has a somewhat different role. Rho controls the assembly of actin filaments and stress fibres and promotes formation of focal adhesion plaques and integrin complexes. Rac promotes actin polymerization and controls the assembly of integrin complexes at the plasma membrane and the formation of membrane ruffles and lamellopodia. Cdc42 forms integrin complexes, associated with actin-rich filopodial extensions. In growing yeast cells, Cdc42 controls polarization, and in the immune system it assists in directing helper T cells to antigen-presenting cells. Of special interest is the control of cell invasion in tumour metastasis, which is intimately linked to cell adhesion and cytoskeletal rearrangements.

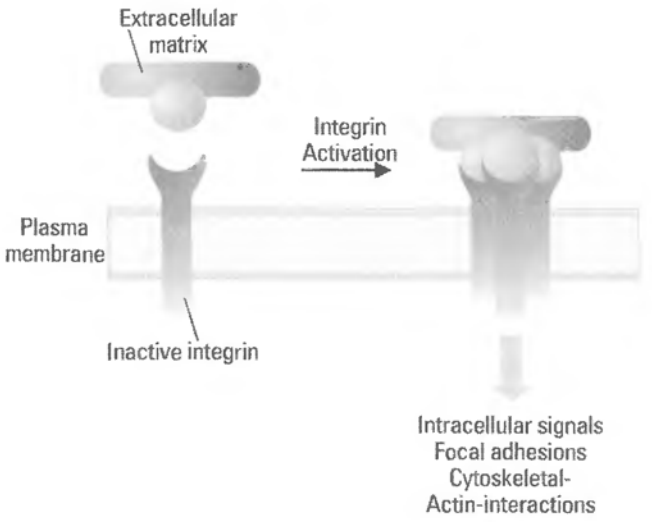
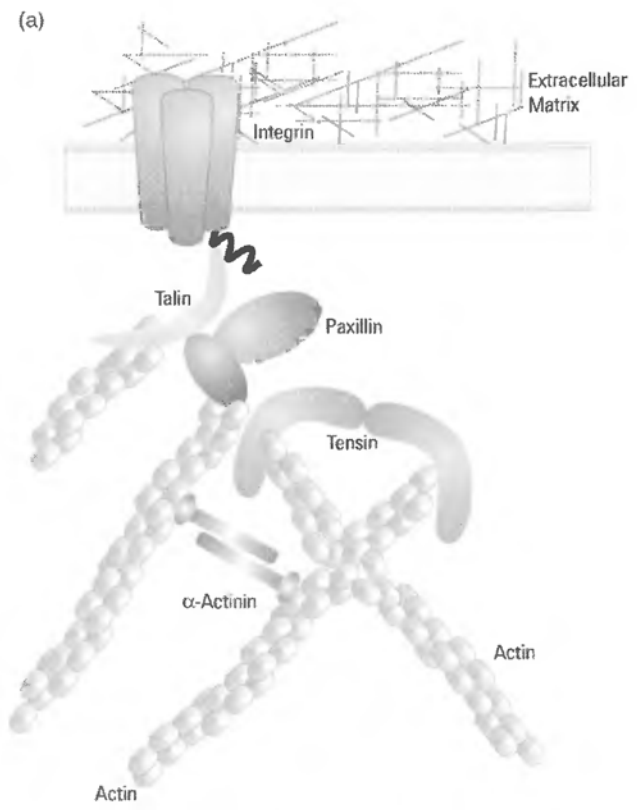


Fig. 4.4 Integrins are transmembrane receptors. They are activated by components of the extracellular matrix, such as fibronectin and collagen. These interactions lead to clustering of integrin receptors and the formation of focal adhesions, intracellular cytoskeletal complexes, and formation of actin bundles.[30] Activation of integrins is part of the cell–cell adhesion process. Integrins also bind to adhesion molecules on adjacent cells.

Platelet-derived lipid messengers, such as lysophosphatidic acid (LPA) and sphingosine-1-phosphate (S1P), can cause cytoskeletal contraction, leading to a loss of cell–cell communication. LPA binds to a G-protein (G_{13})-coupled receptor and signals to Rho,[29] inducing rapid remodelling of the actin cytoskeleton. Rho/Rac also respond to signals from the peptides bradykinin, thrombin, and bombesin. These peptides also signal through heptahelical receptors, coupled to heterotrimeric G proteins of the G_{12} and G_{13} variety.

Adhesive interactions play critical roles in the migration of cells. These interactions are mediated by cell-surface receptors, the integrins that bind to components of the extracelluar matrix and/or to ligands presented by adjacent cells. Consequently, these interactions lead to the formation of intracellular cytoskeletal assemblies and actin bundles (Fig. 4.4).

Fig. 4.5 (a) Integrins are diverse, heterodimeric, αβ-structures which bind to components of the extracellular matrix, such as the matrix proteoglycans, fibronectin, and collagen, through a RGD (Arg, Gly, Asp) binding motif. Integrins form clusters, the focal adhesion plaques. They also bind to soluble ligands such as fibrinogen, and to receptors on adjacent cells, leading to cell aggregation. In both cases, when integrins bind ligands or engage in clustering of cells, focal adhesions are formed. The focal adhesion plaques contain intracellular cytoskeletal components, such as talin, paxillin, and tensin, which tether them to actin and, with the help of α-actinin promote the formation of bundles of actin filaments. (b) Signalling by integrins is controlled by phosphorylation. On the left is shown that cell-surface integrin clusters recruit FAK, the focal adhesion kinase. FAK serves as scaffold. Autophosphorylation of FAK creates a phosphotyrosyl site (Y397), to which cytosolic Src tyrosine protein kinases bind through their SH2 domain.[32] Attention is directed to the two phosphotyrosyls of the integrin–FAK complex which are crucial for the regulation of focal adhesions. The SH2 domain of the Src kinase recognizes the phosphotyrosyl,Y397, of FAK, leading to the activation of the Src kinase. The activated Src kinase then binds an actin-fibre-associated protein (AFAP) which targets the kinase to actin fibres.[33] Next, the Src kinase phosphorylates Y925 in the carboxy-terminal FAT domain of FAK and creates docking sites for Grb2 and Sos and for paxillin. This is shown in (b). The FAT domain is a focal adhesion targeting domain, to which the cytoskeletal protein, paxillin is bound. The Src-family kinases phosphorylate the cytoskeletal protein, paxillin (and tensin), which in turn recruits the guanine nucleotide exchange factor, C3G, with the help of the adaptor Crk, and the docking protein CAS ($p130^{CAS}$). This is the port of entry for monomeric G proteins, such as Rho. Signalling is controlled by PTPs (phosphotyrosine phosphatases) and by the cytosolic tyrosine kinase, Csk, which suppresses the activity of the Src kinase and dampens the integrin/FAK signals. (Reproduced by permission of the authors and Science from Fig.1 in ref. 30.)


(a)
Extracellular Matrix
Integrin
Talin
Paxillin
Tensin
α-Actinin
Actin
Actin

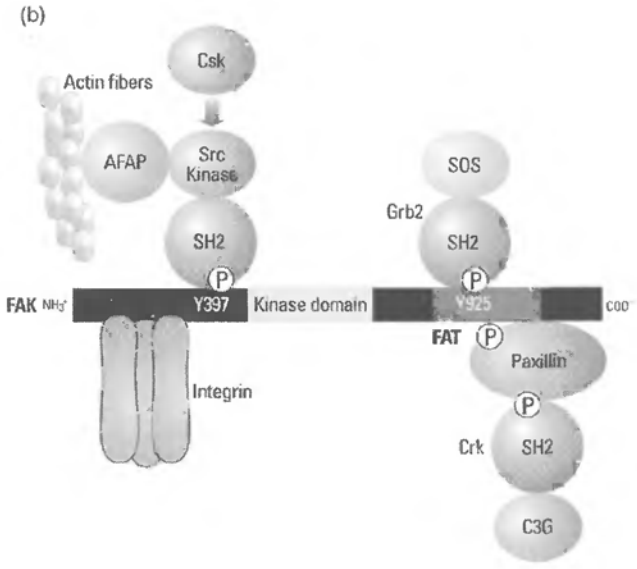

(b)
Csk
Actin fibers
AFAP
Src Kinase
SH2
P
FAK NH3+
Y397
Kinase domain
Y925
COO
SOS
Grb2
SH2
P
FAT
P
Paxillin
P
Crk
SH2
C3G
Integrin

Table 4.2 Components of the integrin-FAK signalling system. (Reproduced with permission of the authors and Science from Table 2, ref. 30)

Components of integrin signalling	Properties
Kinases	
FAK	Focal adhesion kinase; associates with paxillin and integrins
Src	Tyrosine kinase, associates with FAK, paxillin, and integrin–cytoskeletal complex
Other cytosolic tyrosine kinases	*Activated by integrins, associate with FAK and paxillin*
PKC *PAK*	Serine–threonine kinases, localized to focal adhesions, activated by Ca^{2+} and monomeric G proteins
MAP kinase	Regulated by growth factors. Is a serine–threorine kinase activated by integrin cross-linking
Non-enzymic linkers	
Crk	Associates with paxillin and C3G
Grb2	Associates with FAK and Sos
Enzymic linkers	
PI-3K	Phosphorylates phosphatidylinositol at 3-OH position; regulated by integrin, associates with FAK and Src and the cytoskeleton; activated by Rho
PLC	Phospholipase, regulated by integrins and growth factors though tyrosine phosphorylation
GTPases	
Ras	Activated by integrins
Rho	Essential for formation of focal adhesions and stress fibres
GTPase regulators	
Sos	Guanine nucleotide exchange factor; binds to ras and Grb2
C3G	Guanine nucleotide exchange factor; binds to Rho
GAP	GTPase-activating protein; binds to Rho
Phospholipid enzymes	
PtdIns-4-P-kinase	Kinase regulated by Rho.
PLA_2	Phospholipase, hydrolyses glycerophospholipids and forms lysophospholipids and arachidonic acid. Regulated by MAP kinase and Ca^{2+}
5-lipoxygenase	Transforms arachidonic acid into leukotrienes by oxygenation; Essential for spreading of cells on collagen.

Although all these processes are connected in one way or another with the cytoskeleton, how the rearrangement of the cytoskeleton is actually related to the response of cells to signals controlling cell polarity, cell movement, cell adhesion, and cell division is not known. These complex processes are not yet fully understood, but it seems that the point of control of Rho/Rac/Cdc42 is a phosphorylation cascade, where integrin and the focal adhesion kinase (FAK), and cytosolic Src kinases, play a central role.

The role of the integrins in cell adhesion

Integrins are cell-surface proteins that mediate cell–cell adhesion.[31] Integrin-mediated interactions are involved in embryonic development, in tumour metastasis, control of the cell cycle and cell death, blood clotting, and in the response of cells to mechanical stress. Integrin signals modulate the activity of growth-factor receptors, cytosolic kinases, and ion channels, but, most importantly, they regulate the organization of the intracellular actin cytoskeleton. The properties of integrins and the cytoskeletal components with which they interact are summarized in Fig. 4.5. Some of the components of the integrin-FAK signalling system are summarized in Table 4.2.

Summary

1. Integrin complexes control cell adhesion, cell shape, locomotion, and spreading of cells. As the name implies, integrins are integrators, linking the ECM and the cytoskeleton. The fact that tumour cells have lost their anchorage to ECM (see Part 4) makes modification of integrins by suitable modifiers an especially attractive goal for drug design. Drugs targeted to integrins have been developed.

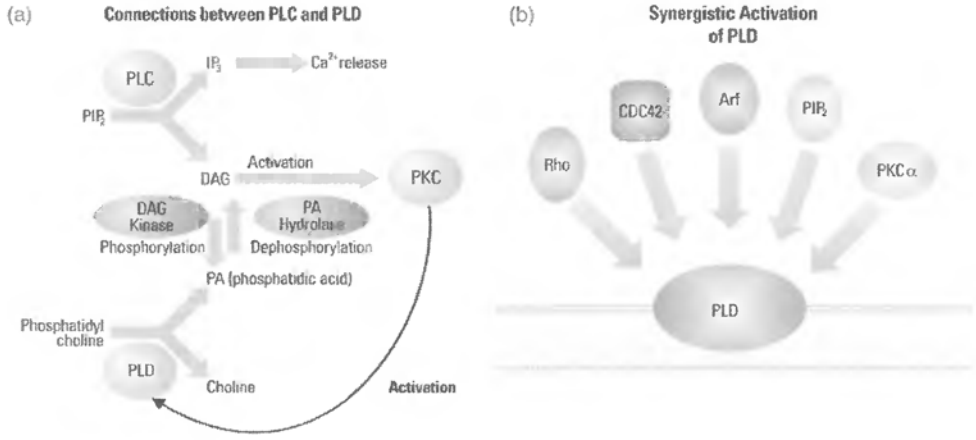


Fig. 4.6 (a) Phospholipase C hydrolyses PIP_2, forming diacylglycerol (DAG) and IP_3 (the crystal structure of a mammalian phosphoinositide-specific phospholipase Cδ has been solved[38]). DAG in turn activates protein kinase C and IP_3 triggers the release of calcium from intracellular stores. Phospholipases of the D family (PLDs), hydrolyse phosphatidylcholine to phosphatidic acid, (PA) and choline. PA is hydrolysed and dephosphorylated to diacylglycerol, DAG, by PAH, phosphatidic acid hydrolase. Conversely, DAG is phosphorylated and converted to PA by DAG kinase. PLC and PLD are connected through diacylglycerol, which is formed directly by PLC and indirectly by PLD via phosphatidic acid and the PAH reaction. Another connection between PLC and PLD is through protein kinase C. DAG activates protein kinase C, which in turn activates PLD. (Drawn with permission of Professor Paul Sternweis and Annu. Rev. Biochem. according to Fig. 1 in ref. 39). (b) Phospholipase D is regulated by the GTPases Rho, Cdc42, and Arf, by the phospholipid messenger PIP_2, and by protein kinase Cα. Rho and Arf, in concert with PIP_2, activate phospholipase D synergistically. The formation of the active GTP-bound state of Rho is controlled by a GDP-dissociation inhibitor, a GDI.)

2. Integrin complexes are regulated by a phosphorylation cascade. The master regulator is a Src kinase, which activates FAK. Activation is modulated by an other cytosolic *k*inase, Csk. Moreover, the guanine nucleotide exchangers Sos and C3G link the integrin–FAK complex to monomeric G proteins and bring them under the control of these G proteins. There is a bewildering variety of connections; for example, to the JNK kinase pathway (to the Ras/MAP kinase pathway?), to PKC and other kinases, to phospholipases, to the PtdIns 3-kinase route and to other PtdIns kinases and to lipoxygenase.

Integrins associate with the membrane protein caveolin.[34] Although little is yet known about this particular interaction, it could assist in forming integrin clusters at the plasma membrane which might serve as a scaffold for the assembly of multi-component signalling complexes. This possibility is attractive, because caveolin binds cholesterol and glycosphingolipids and forms membrane 'rafts', enriched in myristoylated and palmitoylated Src-family kinases (Chapter 8).[35]

Control of phospholipases

It is fitting to introduce the phospholipase pathways here because they are controlled by small G proteins. Phospholipases C and D (PLC, PLD) are controlled by Rho/Arf and Cdc42 and Src tyrosine kinases participate in the control of PLD.[36] Phospholipases form potent second messengers, such as inositol trisphosphate (IP_3) and diacylglycerol (DAG). In Fig. 4.6a the reactions catalysed by phospholipases C and D and the connections between phospholipases C and D are summarized, and in Fig. 4.6b the regulatory pathways that activate phospholipase D synergistically are summarized.[37]

PLC, PLD, and protein kinase C (PKC) respond to many signals and participate in many cellular processes.. The activity of PLD and members of the PLC family, such as PLC-β2, is regulated by the small GTPases Cdc42, Rho, Ral, Rac 1, and Arf, as well by heterotrimeric G proteins. In blood platelets, PtdIns-P_2 (PIP_2) activates PLD. Increased levels of PtdInsP_2 have been related to actin polymerization in platelets and in blood clotting.

Chemotaxis: The role of phosphoinositides for the orientation of the cell towards stimuli

As an example demonstrating the role of phosphoinositide lipid messengers, the regulation of the directed movement of cells along a concentration gradient of chemical attractants will be discussed here. Chemotaxis is essential for the function of phagocytic cells, which help to defend us against the invasion of microbes, by attacking and eventually engulfing and removing them. Movement of cells along a concentration gradient of molecules also plays a critical role in the immune response and in wound healing, as well as in embryogenesis, neurogenesis, and angiogenesis. But it is also found in amoebae. A classical example is the response of *Dictyostelium discoideum* to cAMP.[40]

A gradient of chemotactic stimuli (or of morphogens) polarizes cells. The reorientation of cells is accompanied by morphological changes, including changes of the cytoskeleton, and involves redistribution of many intracellular components, bringing them to the leading edge of the moving cell. For the orientation of cells towards stimuli, phosphatidylinositol-3,4,5-phosphate (PIP_3) and the enzyme that forms it, a PIP_2 kinase isoenzyme play a central role (Fig. 4.7).

A wide range of cells is affected by a loss of the isoenzyme γ-subunit of PIP_2 kinase γ. Loss of this particular PIP_2 kinase impairs the ability of neutrophils and peritoneal macrophages to move towards chemoattractants. Consequently, the deficient cells amass

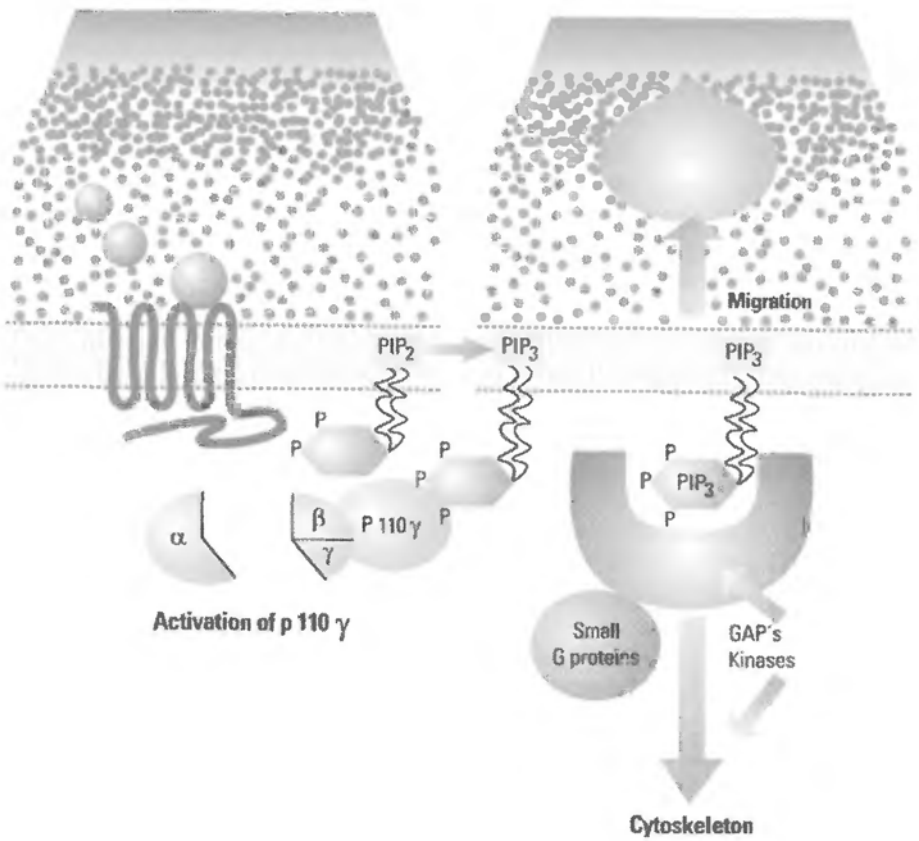


Fig. 4.7 The role of the catalytic subunit p110γ of the PIP_2 kinase γ in asymmetric signal transduction in chemotaxis. A G-protein-coupled heptahelical chemoattractant receptor (left) senses and binds a ligand. The signal is transmitted by a heterotrimeric G protein and the p110γ subunit of the PIP_2 kinase γ is activated through an interaction with the G protein βγ-subunits. The activated p110γ complex then phosphorylates PIP_2, resulting in the formation of PIP_3 (in the middle). PIP_3 associates with PH-domain-containing proteins, which in turn recruit cytosolic kinases, small G proteins and G-protein regulators (GAPs) (on the right-hand side). All these components are involved in the translation of the initial signal into a migratory response of the cell (symbolized above on the right-hand side) and in the activation of the cytoskeleton (symbolized below on the right-hand side). This becomes apparent from the similar phenotypes of knock-out mice lacking either the p110γ catalytic subunit of the PIP_2-kinase γ or the small GTPase Rac 2.[42] This makes it attractive to speculate that Rac 2 (or a Rac 2 GDP-exchange factor, or a GAP) is a downstream component of the PIP_2 kinase signal transduction pathway in chemotaxis. Rac and other small G proteins could mediate the cytoskeletal rearrangements accompanying cell migration and phagocytosis (for a review see ref. 43). Multiple PIP_2 kinase isoforms with different catalytic subunits (p110α, β, δ, γ) catalyse the formation of PtdIns-3,4,5-trisphosphate from PtdIns-4,5-bisphosphate, but only the PIP_2-kinase isoenzyme γ with the catalytic subunit p110γ is activated by the βγ complex of heterotrimeric G proteins and G-protein-coupled receptors (Chapter 5). The other isoenzymes are activated by tyrosine kinase receptors. As expected all signal transduction pathways transmitted by G-protein-coupled receptors and regulated by PIP_3 were impaired in p110γ-deficient cells, whereas the tyrosine kinase controlled PIP_3 pathways were functional. (For further information, see ref. 44). The role of the G-protein-controlled PIP_2 kinase in the cellular response to tissue damage in inflammation—examples are acute inflammation of the respiratory tract and inflammatory bowl disease—will make the p110 catalytic subunit γ of the PIP_2 kinase γ isoenzyme an attractive target for anti-inflammatory drugs.[45] (Reproduced with permission of the authors and Science from a figure in L. V. Dekker and A. W.Segal. Signal transduction: signals to move cells.[41])

in the circulation, since they are unable to move into tissues. This resembles the phenotype of a human disease, the leucocyte adhesion deficiency syndrome, where leucocytes lack the adhesion molecule β_2-integrin. Although this particular PIP_2 kinase γ isozyme is certainly an important mediator of chemotactic responses, other PIP_2 kinase isozymes may also play a role in cell migration. But since all of these PIP_2 kinases catalyse the formation of PIP_3, it is clear that this phospholipid molecule is the main second messenger and regulator of cell migration.

Moreover, other PtdIns-regulated kinases, such as the phosphoinositide-dependent protein kinase PKB,[46] also participate in the cell polarity change induced by the chemoattractant. PKB contains PH domains[47] to which phosphoinositides can bind (see Chapter 3). When the gene encoding the PH domains of PKB was fused with the gene for the green fluorescent protein and the fused gene was expressed in mammalian HL-60 cells by Bourne and colleagues,[48] or in *Dictyostelium discoideum* (slime-mould cells) by Devreotis and colleagues,[49] and the distribution of the fusion protein in the cell was followed by immunofluorescence microscopy, it became apparent that PKB was distributed differently, depending on how the cells were stimulated. For example, when the chemoattractant was applied uniformly to the cells, the fusion protein was also distributed uniformly. However, when the cells were exposed to a shallow, unequal gradient of chemoattractant, the localization of PKB was restricted to the area of the cell membrane which was directly in contact with the stimulus. What ever the reason, an asymmetric generation of intracellular messenger molecules, such as PIP_3 and their effector kinases, could help to guide the cell towards the chemoattractant gradient. This is where the G proteins come into play: Rac, Rho, and Cdc42 could control the localization of PIP_3-responsive PH-domain-containing effector proteins, and thus assist in establishing an asymmetric distribution. Although the way in which asymmetry is established is still disputed, there is no doubt that asymmetric signalling can increase the sensitivity of cells to chemoattractants at that pole of the cell which is exposed to the signal. This might be of crucial importance for the response of cells to morphogens in embryonic development.

Connections and branching of signalling routes

Finally, I address the question, whether Rho/Rac/Cdc42 and Ras signal through separate or interconnected pathways. There are examples that indicate that Rho/Rac/Cdc42-controlled and Ras-controlled signalling pathways communicate with each other. When Ras is overexpressed in response to cytokines and stress signals, a cell-cycle-inhibiting protein (a cyclin-dependent kinase inhibitor, CDKI) is activated and the cell cycle and cell proliferation are stopped (this CDKI is the p21Waf-1 protein which inhibits cyclin-dependent kinases; see Chapter 12).[50] But when Rho is expressed together with Ras, the CDKI is turned off and the stress-responsive JNK/ SAPK cascade is activated, which turns on genes, promoting cell proliferation. Thus, Ras needs the assistance of Rho to turn on the cell cycle and proliferation because Rho relieves the blockade of the cell cycle in stressed cells.[51] Because of its role in cell-cycle control, the cross-talk between the Ras- and the Rho/Rac/Cdc42 pathways also plays a role in several phases of oncogenesis, especially in tumour invasion (see Part 4). [52]

Although, the Rho/Rac/Cdc42-controlled and Ras-controlled signalling pathways remain responsive to each other's signals, the two pathways must have separated at some point, each sending signals along its own way. This implicates bifurcations of a common signalling cascade into several branches, each targeted to different effectors. But where

Rho and Ras signals depart from each other is not so clear. The situation is further complicated, because even in the same kind of cells, the path the signal travels may vary, depending on the type of growth factors and cytokines that the cell is exposed to. Nevertheless, regulated interconnectivity of signalling routes is certainly an attractive feature, because it increases the density of the signal network, bringing an ever greater area under control. An example is the pathway interconnectivity between heterotrimeric G proteins and monomeric G proteins, which will be discussed in Chapter 5.

Outlook

Goals for future research are to find where the pathways that control cytoskeletal organization separate from the routes that control cell division. How reorganization of cell-surface integrin complexes and of the actin cytoskeleton promote cell division, and vice versa, is a central point. Answers to this question are of great importance for an understanding of malignant transformation and tumour invasion. But it is quite likely that the cell is using the tools that we have already encountered above to control these processes, namely monomeric G proteins. An attractive candidate is Ran GTPase.[53] Furthermore, in the link between control of the assembly of actin filaments and control of growth and proliferation, proteins such as 'WASP' (Wiskott–Aldrich syndrome protein) may play a role, the activity of which is controlled by Cdc42.

WASP is implicated in the Wiskott–Aldrich immunodeficiency syndrome, a hereditary disease accompanied by severe defects of haematopoietic cell function, thrombocytopenia, and immunodeficiency. What makes WASP especially attractive as a linker is its conformational flexibility. It has a GTPase-binding domain (GBD), which in the inhibited state hides the carboxy-terminal region of WASP which contains a Cdc42/Rac interactive binding domain (CRIB) and prevents its interaction. But on binding Cdc42/Rac, WASP undergoes a dramatic conformational change which frees its C-terminus, which now can interact with the actin regulatory machinery. WASP is thought to link Cdc42 directly to the rearrangement of the actin cytoskeleton.[54] A NMR-based structure,[55] of a complex of Cdc42 with a WASP peptide reveals that the GBD domain of WASP, which contains the CRIB motif, contacts the regulatory switch regions of Cdc42. These contacts with the switch 1 region of Cdc42 sense the GDP/GTP switch and the conformational change in Cdc42 on binding GTP. Other motifs outside the CRIB region in the GBD domain may allow WASP to interact with other targets.

References

1. C. . Marshall. Raf gets it together, News and Views. *Nature*, **383**, 127–128, 1996.
   Also: M. A. Farrar, J. Alberola-Ila, and M. Perlmutter. Activation of the Raf-1 kinase cascade by coumermycin-induced-dimerization. *Nature*, **383**, 178–181, 1996.
   Also: Z. Luo, G. Tzivion, P. J. Belshaw, D. Vavvas, M. Marshall, and J. Avruch. Oligomerization activates c-RAF-1 through a Ras-dependent mechanism. *Nature*, **383**, 181–185, 1996.
2. S. Lev, H. Moreno, R. Martinez, P. Canoll, E. Peles, J. M. Musacchio, *et al.* Protein tyrosine kinase PYK2 involved in Ca^{2+}-induced regulation of ion channel and MAP kinase functions. *Nature*, **376**, 737–745, 1995.
3. D. A. Fruman, R. E. Meyers, and L. C. Cantley. Phosphoinositidekinases. *Annu Rev Biochem*, **67**, 481–507, 1998.
4. M. H. Cobb and E. J. Goldsmith. How MAP kinases are regulated. *J Biol Chem*, **270**, 14843–14846, 1995.

5. P. Van der Geer, T. Hunter, and R. A. Lindberg. Receptor protein tyrosine kinases and their signal transduction pathways. *Annu Rev Cell Biol*, **10**, 251–337, 1994.
6. A. Brunet and Pousségur, J. Identification of MAP kinase domains by redirecting stress signals into growth factor responses. *Science*, **272**, 1652–1655, 1996.
7. Z. Wang, P. C. Harkins, R. J. Ulevitch, J. Han, M.H. Cobb, and E. J. Goldsmith. The structure of mitogen-activated proteinkinase p38 at 2.1 Å resolution. *Proc Natl Acad Sci, USA*, **94**, 2327, 1997.
8. A. Brunet, P. Lenormand, J.-M. Brondello, F. McKenzie, and J. Pouysségur. Growth factor activated MAP kinases: Mechanism of nuclear translocation and role in growth control. Abstracts XIIIth International Congress of Pharmacology, 26–31 July 1998, Munich, Germany. *Arch Pharmacol*, **358** (Suppl. 2, Number 1), Abstr. SE 3.3.p: R. 383, 1998.
9. Also: L. M. Graves, *et al.* Regulation of carbamoyl phosphate synthetase by MAP kinase. *Nature*, **403**, 328–332, 2000.
10. A. J. Whitmarsh and R. J. Davis. Transcription factor AP-1 regulation by mitogen-activated protein kinase signal transduction pathways. *J Mol Med*, **74**, 589–607, 1996.
11. P. Sassone-Corsi, *et al.* Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. *Science*, **285**, 886–891, 1999.
12. S. Thomson, *et al.* The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. *EMBO J*, **18**, 4779–4793, 1999.
13. A. J. Waskiewicz, *et al.* Phosphorylation of the cap-binding protein eukaryotic translation initiation factor 4E by protein kinase Mnk1 in vivo. *Mol Cell Biol*, **19**, 1871–1880, 1999.
14. S. Pyronnet, *et al.* Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate e1F4E. *EMBO J*, **18**, 270–278, 1999.
15. M. H. Ladha, *et al.* Regulation of exit from quiescence by p27 and cyclin D1-CDK4. *Mol Cell Biol*, **18**, 6605–6615, 1998.
16. M. Cheng, V. Sexl, C. J. Sherr, and M. F. Roussel. Assembly of cyclin D-dependent kinase and titration of p27Kip1 regulated by mitogen-activated protein kinase (MEK1) *Proc Natl Acad Sci, USA*, **95**, 1091–1096, 1998.
17. S. Cowley, H. Paterson, P. Kemp, and C. J. Marshall. Activation of MAP kinase is necessary and sufficient for PC12 differentiation and for transformation of NIH3T3 cells. *Cell*, 77, 841–852, 1994.
18. S. J. Mansour, *et al.* Transformation of mammalian cells by constitutively active MAP kinase kinase. *Science*, **265**, 966–970, 1994.
19. G. Pagès, *et al.* Defective thymocyte maturation in p44 MAP kinase (Erk 1) knockout mice. *Science*, **286**, 1374–1377, 1999.
20. F. Jacob. Evolution and tinkering. *Science*, **196**, 1161–1166, 1997.
21. T. Lazar, M. Götte, and D. Gallwitz. Vesicular transport: how many Ypt/Rab-GTPases make a eukaryotic cell? *Trends Biochem Sci*, **22**, 468–472, 1997.
22. T. Ohba, M. Nakamura, H. Nishitani, and T. Nishimoto. Self-organization of microtubule asters induced by *Xenopus* egg extracts by GTP-bound Ran. *Science*, **284**, 1356–1358, 1999.
    Also: A. Wilde and Y. Zheng. Stimulation of microtubule aster formation and spindle assembly by the small GTPase Ran. *Science*, **284**, 1359–1362, 1999.
23. W. J. Wu, J. W. Erickson, R. Lin, and R. A. Cerione. The γ–subunit of the coatomer complex binds Cdc 42 to mediate transformation. *Nature*, **405**, 800–804, 2000.
    Also: C. J. Der and W. E. Balch. GTPase trafic control, Comments. *Nature*, **405**, 749–752, 2000.
24. C. Nobes and A. Hall. Regulation and function of the Rho subfamily of small GTPases. *Curr Opin Genet Dev*, **41**, 77–81, 1994.
25. J. Settleman, C. F. Albright, L. C. Foster, and R. A. Weinberg. Association between GTPase activators for Rho and Ras families. *Nature*, **359**, (6391), 153–154, 1992.
26. T. Leeuw, C. Wu, J. D. Schrag, M. Whiteway, D. Y. Thomas, E. Leberer. Interaction of a G-protein beta-subunit with a conserved sequence in Ste20/PAK family protein kinases. *Nature*, **391**, 191–195, 1998.
27. M. Symons. Rho family GTPases: the cytoskeleton and beyond. *Trends Biochem Sci*, **21**, (5), 178–181, 1996.
28. M. F. Olson, A. Ashworth, and A. Hall. An essential role for Rho, Rac and Cdc 42 GTPases in cell cycle progression through G1. *Science*, **269**, 1270–1272, 1995.
29. A. Gohla, R. Harhammer, and G. Schultz. The G protein G13, but not G12 mediates signaling from lysophosphatidic acid receptor via epidermal growth factor receptor to Rho. *J Biol Chem*. 273, 4653–4659, 1998.

30. E. A. Clark and J. S. Brugge. Integrins and signal transduction pathways: The road taken. *Science*, **268**, 233–239, 1995.

31. F.G. Giancotti and E. Ruoslahti. Integrin signalling. *Science*, **285**, 1028–1032, 1999.

32. B. S. Cobb, M. D. Schaller, T. H. Leu, and J. T. Parsons. Stable association of pp60src and pp59fyn with the focal adhesion-associated protein tyrosine kinase, pp125FAK. *Mol Cell Biol*, **14**, (1), 147–155, 1994.

33. D. C. Flynn, T. H. Leu, A. B. Reynolds, and J. T. Parsons. Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated pp60src substrate. *Mol Cell Biol*, **13**, (12), 7892–7900, 1993.

34. Y. Wei, X.Yang, Q. Liu, J. A Wilkins, and H. A. Chapman. A role for caveolin and the urokinase receptor in integrin-mediated adhesion and signaling. *J Cell Biol*, **144**, 1285–1294, 1999.

35. T. Harder and K. Simons. Caveolae, DIGs, and the dynamics of sphingolipid-cholesterol microdomains. *Curr Opin Cell Biol*, **9**, 534–542, 1997.

36. J. G. Song, L. M. Pfeffer, and D. A. Foster. v-Src increases diacylglycerol levels via a type D phospholipase-mediated hydrolysis of phosphatidylcholine. *Mol Cell Biol*, **11**, (10), 4903–4908, 1991.

37. H. Exton. Phospholipase D: enzymology, mechanisms of regulation, and function. *Physiol Rev*, **77**, (2), 303–320, 1997.

38. L.-O. Essen, O. Peridic, R. Cheung, M. Katan, and R. L. Williams. Crystal structure of a mammalian phosphoinositide-specific phospholipase C. *Nature*, **380**, 595–602, 1996.

39. W. D. Singer, H. A. Brown, and P. C. Sternweis. Regulation of eukaryotic phosphatidylinositol-specific phospholipase C and phospholipase D. Annu Rev Biochem, **66**, 475–509, 1997.

40. C. A. Parent and P. N. Devreotes. A cell's sense of direction. *Science*, **284**, 765–770, 1999.

41. L. V. Dekker and A. W. Segal. Signal transduction: signals to move cells. *Science*, **287**, 982—985, 2000.

42. A. W. Roberts *et al.* Deficiency of the hematopoietic cell-specific Rho family GTPase Rac2 is characterized by abnormalities in neutrophil function and host defense. *Immunity*, **10**, 183–196, 1999.

43. B. Vanhaesebroeck *et al.* Phosphoinositide 3-kinase: a conserved family of signal transducers. *Trends Biochem Sci*, **22**, 267–272, 1997.

44. Tian Jin, Ning Zhang, Yu Long, C. A. Parent, and P. N. Devreotes. Localization of the G protein complex in living cells during chemotaxis. *Science*, **287**, 1034–1036, 2000.

    Also: Zhong Li, Huiping Jiang, Wei Xie, Zuchuan Zhang, A. V. Smrcka, and Dianqing Wu. Roles of PLC-2 and -3 and PI3K in chemoattractant-mediated signal transduction. *Science*, **287**, 1046–1049, 2000.

    Also: . Takehiko Sasaki, J. Irie-Sasaki, R. G. Jones, A. J. Oliveira-dos-Santos, W. L. Stanford, B. Bolon, *et al.* Function of PI3K in thymocyte development, T cell activation, and neutrophil migration. *Science*, **287**, 1040–1046, 2000.

    Also: E. Hirsch, V. L. Katanaev, C. Garlanda, O. Azzolino, L. Pirola, L. Silengo, *et al.* Central role for G protein-coupled phosphoinositide 3-kinase in inflammation. *Science*, **287**, 1049–1053, 2000.

45. A. Hall. Rho GTPases and the actin cytoskeleton. *Science*, **279**, 509–514, 1998.

46. D. R. Alessi *et al.* Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase B alpha. *Curr. Biol.* 7, 261–269, 1997.

47. M. J. Bottomley, K. Salim, and G. Panayotou. Phospholipid-binding protein domains. *Biochim Biophys Acta*, **1436**, 165–183, 1998.

48. G. Servant, O. D. Weiner, P. Herzmark, T. Balla, J. W. Sedat, and H. R. Bourne. Polarization of chemoattractant receptor signaling during neutrophil chemotaxis. *Science*, **287**, 1037–1040, 2000.

49. Tian Jin, Ning Zhang, Yu Long, C. A. Parent, and P. N. Devreotes. Localization of the G protein complex in living cells during chemotaxis. *Science*, **287**, 1034–1036, 2000.

50. M. F. Olson, H. F. Paterson, and C. J. Marshall. Signals from Ras and Rho GTPases interact to regulate expression of p21 $^{Waf1/Cip1}$. *Nature*, **394**, 295–299, 1998.

51. C. J. Marshall. Ras effectors. *Curr Opin Cell Biol*, **8**, (2), 197–204, 1996.

52. S. Mittnacht, H. Paterson, M. F. Olson, and C. J. Marshall. Ras signalling is required for inactivation of the tumour suppressor pRb cell-cycle control protein. *Curr Biol*, **7**, (3), 219–221, 1997.

53. E. Pennisi. Nuclear transport protein does double duty in mitosis. *Science*, **284**, 1260–1261, 1999.

54. A. S. Kim, L. T. Kakalis, N. Abdul-Manan, G. A. Liu, and M. K. Rosen. Autoinhibition and activation mechanisms of the Wiskott–Aldrich syndrome protein. *Nature*, **404**, 151–158, 2000.

55. N. Abdul-Manan, B. Aghazadeh, G. A. Liu, A. Mjumdar, O. Ourfelli, K. A. Siminovitch, and M. K. Rosen. Structure of Cdc 42 in complex with the GTPase-binding domain of the 'Wiskott–Aldrich Syndrome' protein. *Nature*, **399**, 379–383, 1999.

5

Signal transduction pathways through heterotrimeric G proteins: transmission of hormonal and sensory signals

Signalling by G-protein-coupled receptors

G-protein-coupled receptors (GPCRs) are expressed in eukaryotes, from yeast to humans. Many hormones bind to GPCRs. Targets of hormonal signals, transmitted by GPCRs, are often enzymes, synthesizing second messengers such as cAMP. Through second messengers, G-protein-coupled receptors regulate many metabolic and cellular functions, including ion channels. GPCRs also transmit signals to cellular phosphorylation cascades, such as the MAP kinase pathway. Sensory signals, light, smell, and taste signals are also received by receptors with seven transmembrane helices and transmitted and amplified by heterotrimeric G proteins.

GPCRs are heptahelical, serpentine receptors, which pass the membrane seven times. They have no kinase activity. Prototype G-protein-coupled receptors are the β-adrenergic receptors, which transmit the signals of the hormone adrenaline (epinephrine) and rhodopsin (which transmits the light signal in the rod outer segment membrane and initiates the visual response in the eye). More than 40 sequences alone for the β-adrenergic receptors have been derived from cDNAs.[1]

The G-protein-coupled receptors

Only recently was the first high-resolution atomic structure of a G-protein-coupled receptor solved, namely that of rhodopsin,[2] although lower-resolution spatial structural information based on two-dimensional crystals and electron diffraction and NMR structures was available.[3,4] This information makes it certain that all heptahelical receptors have the same topological arrangement of the polypeptide chains. The amino- and carboxy-termini are oriented in the same way, with the amino-terminus outside and carboxy-terminus on the cytoplasmic side. Valuable structural relationships between different G-protein-coupled receptors for hormones have also come to light, mainly thanks to comparisons of cDNA-derived sequences.[5]

Hormonal signalling by G-protein-coupled receptors

Many G-protein-linked hormone and neurotransmitter receptors have been modelled.[6] The docking sites on the 5-hydroxytryptamine (5-HT, or serotonin) receptor for G protein were modified by 'RNA editing'.[7] Variants of the gene for the 5-HT 2C-subtype of the serotonin receptor were obtained by conversion of the adenosines in the pre-messenger RNA to inosines by adenosine deaminase. The variants of the 5-HT gene were expressed in cells. The changes in the gene resulted in amino-acid changes at three sites in the receptor and pointed to the second intracellular loop, i_2, as one of the docking sites for G proteins. Modification at this site caused a 10- to 15-fold reduction in G-protein coupling.

The 5-HT 2C receptors are expressed throughout the human central nervous system and play a major role in the control of higher-order cerebral functions. 5-HT 2C receptor mutants might be related to cerebral dysfunctions, causing depression and epilepsy. The serotonine receptor is only one of many G-protein-coupled hormone and neurotransmitter receptors. There are the dopaminergic, α- and β-adrenergic, opioid, and muscarinic receptors, and so forth. They all have important, medically relevant functions. This is the reason why the study of GPCRs and their ligands is one of the most active fields in pharmacological research.

Figure 5.1 shows schematically the structure of the β_2-adrenergic receptor (β_2-AR).

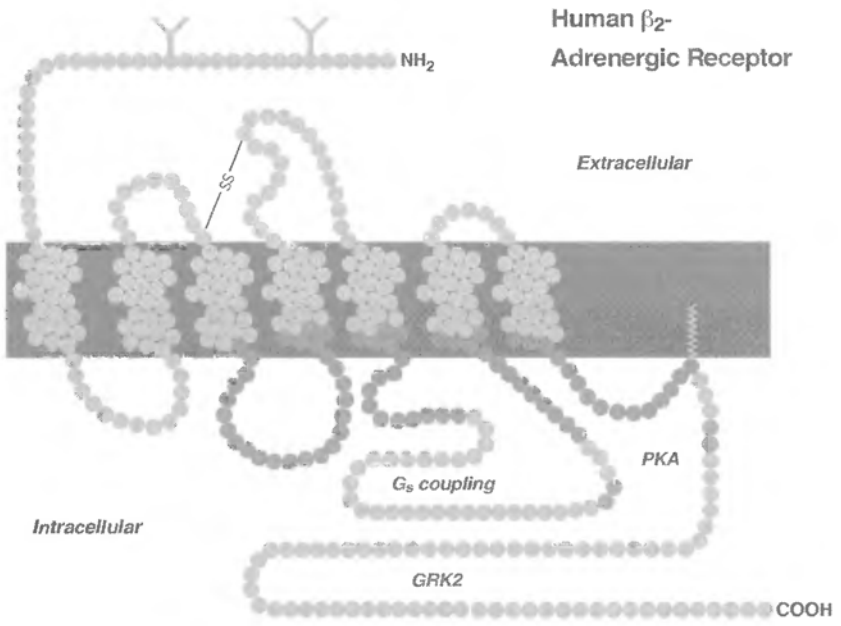


Fig. 5.1 Parts of the intracellular loops of the receptor to which the heterotrimeric, αβγ-G-protein binds and the sites of phosphorylation by the G-protein-coupled receptor kinase, GRK2, and by the cAMP-dependent kinase, PKA. The glycosylation sites on the amino-terminal region are also indicated. (Reproduction of this scheme was made possible through the generosity of Professor Martin Lohse, Department of Pharmacology, the University of Würzburg, Medical School.)

A fundamental problem is how these receptors can form different active states, each presenting a different interaction surface to G proteins. Whether G-protein-coupled heptahelical receptors form dimers or higher-order signalling assemblies, called 'transducisomes' is a matter of discussion.[8] This would be the simplest way to form new interaction surfaces (Chapter 2). Also there is information that dimerization of GPCRs is important for receptor function in some cases (such as in the case of $GABA_B$ (γ-aminobutyric acid) and opioid receptors),[9] but we have to see how general and functionally significant this phenomenon is. But when high-resolution structures of heptahelical receptors in complex with a G-protein become available we shall get information on the signalling-competent state of GPCRs. Only once this information is available will it be possible to understand how the cytoplasmic loop regions, which are instrumental in the recognition of the different G proteins (and other partners) are rearranged on activation of the receptor by the ligand.

Signalling by the β-adrenergic receptor

A well-studied example of a signal response, mediated by a heterotrimeric G protein, is the activation of the enzyme adenylyl cyclase by the hormone adrenaline. Adenylyl cyclase catalyses formation of cAMP from ATP (Fig. 5.2).

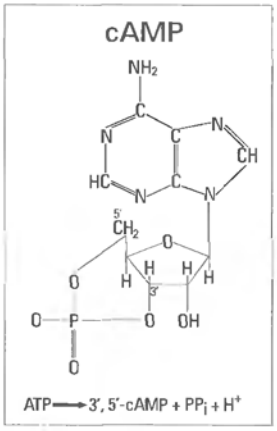


Fig. 5.2 cAMP is formed enzymatically through the reaction: ATP → cAMP + PP_i + H^+.

Earl Sutherland,[10] who discovered cAMP, which is formed in the liver in response to adrenaline, named the hormone the first messenger and cAMP the second messenger. On activation by binding a ligand or receiving a sensory input signal, such as a quantum of light, the heterotrimeric G protein, G-αβγ, bound to the receptor is activated. The receptor catalyses GDP/GTP exchange, and when GTP is bound, the αβγ-complex dissociates into α-GTP and βγ-subunits, both of which transmit signals to targets. G proteins are encoded by a large gene family.[11,12] Heterotrimeric G proteins are present in all eukaryotic cells.

Figure 5.3 shows how the second messenger, cAMP, formed from ATP by adenylyl cyclase, transmits the signal by activating a second-messenger-dependent serine/threonine protein kinase, the cyclic AMP-dependent protein kinase A (PKA).

The solution of the crystal structure of a soluble, catalytically active fragment of adenylyl cyclase in a complex with the α-subunit of the heterotrimeric G protein G_s (stimulatory G protein), by Stephen Sprang and Alfred Gilman,[13] is a great achievement, because adenylyl cyclase is membrane-bound and a notoriously unstable and difficult enzyme to work with. The efforts have paid off. Four different structures of the catalytic domain complexed with different ATP analogues and various ions (Mg^{2+}, Mn^{2+}, Zn^{2+}) could be determined. The structures of the enzyme–ATP complex showed that two metal ions bind to the active site. The similarity of the active site of adenylyl cyclase with that of DNA polymerase indicated that adenylyl cyclase also catalyses a phosphoryl transfer reaction with the help of two metal ions, and that the two enzymes, DNA polymerase and adenylyl cyclase, have probably evolved from a common ancestor (see Plate 10).[14,15]

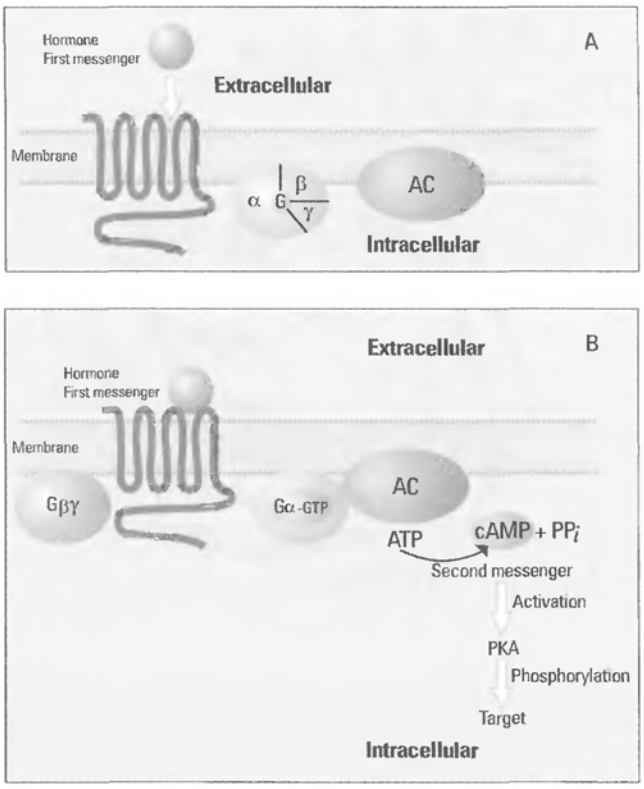


Fig. 5.3 The situation (A) before and (B) after the receptor, activated by the hormone, had made contact with the transducer, a heterotrimeric αβγ-G protein. After coupling to the receptor and activation, the αβγ-G protein dissociates into the βγ-subunits and the α-subunit, which is now in the active, GTP-bound, 'on' state. The active G-α transducer finds the adenylyl cyclase in the two-dimensional space of the membrane, by collision coupling. The activated adenylyl cyclase forms the second messenger cAMP from ATP, which in turn activates the cAMP-dependent kinase, PKA, and downstream targets. A simple ribbon diagram illustrating the relative position of a heterotrimeric G protein in relation to the receptor embedded in the membrane is shown in plate 11.

The role of the βγ-subunits

Although dissociation of heterotrimeric G proteins has been difficult,[16] to demonstrate in the living cell, it is clear that each G-α GTP, and also the G-βγ complex, is a cellular signal transducer.[17] A variety of targets have been identified that are regulated by G-βγ alone or synergistically with G-α , including ion channels, enzymes such as adenylyl cyclase, phospholipase C-β, and β-adrenergic receptor kinases, and regulatory proteins, such as phosducin and others (see below).

The structure of the contact region of the subunits changes drastically when α- and βγ-subunits combine, forming the inactive αβγ-complex. Crystal structures of the αβγ-complex and of separated α- and βγ-subunits of heterotrimeric G proteins point to a small region which is crucial for interactions with the β-subunits.

Of the 13 most variable sites in five different G β-subunits, 11 are within this contact region. These sites form a highly variable solvent-exposed surface, accounting for the fact that the same β-subunit pairs with different γ-subunits.[18,19] Fifteen genes for the α-subunit, 11 for β-, and 8 for γ-subunits have been identified so far,[20] but there are probably more. There exist about 20 splice variants and isoforms of β-subunits alone.[21] Each βγ combination might possibly interact with a different partner. For the assembly of a signal-responsive heterotrimeric, αβγ-G-protein complex and for proper interaction with the α-subunit, farnesylation of the βγ-complex is required.[22]

Fig. 5.4 A ribbon model of the G β-subunit, which has WD (Trp–Asp) repeats and forms β-propellers. This structure is like a scaffold, presenting a surface to which other proteins can bind. (The structure was solved in Paul Sigler's laboratory[25] and is reproduced with permission of the authors and Nature.)

For the remarkable coupling versatility of G β-subunits, structural features—the WD (Trp–Asp) repeats—are thought to be responsible.[23] (For a review see ref. 24 and Fig. 5.4.)

Control of hormonal signalling

Signalling by G-protein-coupled receptors is terminated by binding of the receptor kinase to the receptor, followed by phosphorylation. Receptor kinases are serine/threonine kinases, recognizing the active conformation of the receptor. A carrier function is attributed to βγ-subunits in bringing the kinase to the receptor in the membrane.[25] For that function, isoprenylation of the βγ-complex is required.[26] This function is comparable to that of Ras, bringing the cytosolic Raf kinase to the membrane for activation. Both the γ-subunits of the βγ-complex and Ras are modified by prenylation (see Chapter 3).

Control of G proteins by RGS proteins

Not only can receptor activity be turned off, but also activity and signalling of the α-subunit of heterotrimeric G proteins. The latter activity is controlled by the RGS proteins (regulators of G protein signalling). In yeast, RGS proteins regulate the response to pheromones.[28] Pheromone activates a heterotrimeric G-protein-linked heptahelical receptor and the RGS proteins control mating of haploid cells and the formation of diploid cells by turning down the response to the pheromone.[29] RGS proteins are GAPs for the α-subunit of heterotrimeric G proteins.[30] They shorten and downregulate G-protein-dependent signalling in *C. elegans*, yeast, and in mammalian cells. A mammalian *rgs* gene family encodes the structural and functional homologues of the EGL-10 proteins, encoded by the *egl-10* gene in *C. elegans*.[31] The RGS proteins in yeast are the SST2p proteins.[32] The human *rgs* gene can actually replace the *sst2* gene in yeast,[33] demonstrating how little these genes have changed from worm and yeast to humans. RGS proteins,[34] apparently interact with all α-subunits of heterotrimeric G proteins (perhaps with the

exception of G_s-α. Steven Sprang has recently solved the structure of a complex of an active G α-subunit, G_i-α-AlF_4^-, with an RGS protein.[35] Although the RGS-GAPs have no sequence homology with Ras-GAPs, both are helical proteins and both bind to the switch regions of the G protein, stabilizing the catalytic domain and triggering the GTPase activity of the α-subunit (see Chapter 3).

Deactivation and receptor desensitization

Sustained stimulation leads to a decreased responsiveness of G-protein-coupled receptors. Desensitization guards against aggressive effects of second messengers and destructive overexposure to light. Receptor desensitization has been studied in depth with the β-adrenergic receptor–G_s adenylyl cyclase system, primarily in R. J. Lefkowitz's laboratory (see ref. 36).

Changes in responsiveness of GPCRs range from seconds to days. Long-lasting down-regulation involves changes in receptor gene expression.[37] Homologous and heterologous desensitization are usually distinguished: homologous desensitization involves only those receptors activated by the same agonist, whereas heterologous desensitization is a more generalized process, affecting a variety of receptors responding to different agonists.

Here, only homologous, rapid receptor desensitisation will be dealt with (see ref. 38 and Fig. 5.5).

A critical analysis of the on and off rates of ligand–receptor interactions made it clear that deactivation of a heptahelical receptor can not be determined by the decay of the

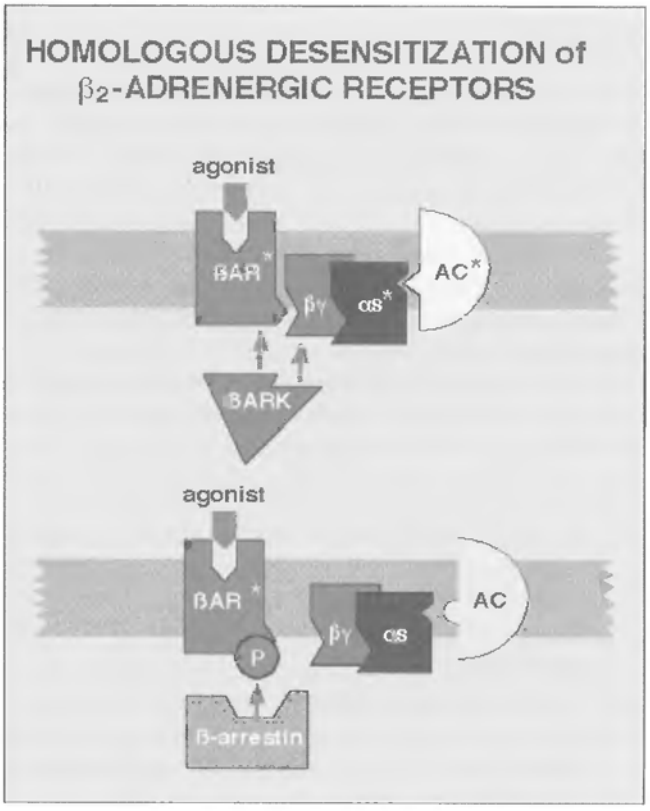


Fig. 5.5 Desensitization of the activated β-adrenergic receptor, βAR*. βARK is the β-adrenergic receptor kinase. αs* and βγ are the activated α-subunit and the βγ-subunits of the heterotrimeric G protein, G_s. AC* is the activated effector, the enzyme adenylyl cyclase. Phosphorylation of activated GPCRs enhances binding of a family of proteins, the arrestins, which prevent access of G proteins and uncouple the receptor from G proteins. The arrestin involved in the desensitization of the β-adrenergic receptor is β-arrestin. (Reproduction of this scheme has been made possible through the generosity of Professor Martin Lohse, Department of Pharmacology, University of Würzburg, Medical School.)

receptor–ligand complex and the off rate of the activating ligand.[39] A plausible explanation for receptor shut off had to wait until the discovery of specific receptor kinases in R. J. Lefkowitz's laboratory, which recognize and phosphorylate only the active conformation of the receptor, formed on binding agonists, and the discovery of the arrestins. The protein arrestin is bound to the phosphorylated receptor. Arrestin prevents access of the G protein and interrupts G-protein coupling to the receptor and signalling.

Arrestins are a class of ubiquitous cytosolic proteins of which several variants exist. The first arrestin was discovered by Hermann Kühn in rod outer segments of the retina, where it participates in the deactivation of the visual response.[40]

All signal transmission systems employing G-protein-coupled receptors have a similar programme to turn off activated receptors, and employ the same tools. However, the resensitization process is different. Whereas desensitized hormone receptors are removed from the membrane, separated from the ligand, and transferred into vesicles for dephosphorylation before they are returned to the membrane, in the case of rhodopsin receptor resensitization does not start with translocation of the receptor from the membrane, instead retinal (the light-absorbing photochemical group) is first inactivated by separation from the receptor protein, opsin. Retinal is oxidized to retinol,[41] and free opsin is dephosphorylated (see below).[42] Eventually opsin is recombined with retinal, forming rhodopsin.[43]

Fig. 5.6 Arrestin is an elongated molecule. The fold of the polypeptide chains is shown. The crystal structure of bovine arrestin comprises two domains of antiparallel β-sheets connected through a hinge region and one short α-helix on the back of the amino-terminal fold. The region where arrestin binds to the phosphorylated light-activated rhodopsin is located at the N-terminal domain, on top. This is also supported by binding studies with N-terminally truncated arrestin. (This ribbon model of arrestin is reproduced with permission of the authors and Cell

Arrestin

A phosphorylated serine in the receptor serves as a structural marker for the recognition of arrestin. Arrestin has practically no affinity for the unphosphorylated receptor, but the phosphorylated receptor, rhodopsin, binds with a K_D in the order of about 50 nM. Three-dimensional structures of arrestin from bovine rod outer segments have been reported at 3.3 and 2.8 Å resolution.[44] We show the more recent, higher-resolution structure from Paul Sigler's laboratory (Fig. 5.6).[45]

R. J. Lefkowitz's laboratory presented data suggesting that arrestin may have more than one function.[46] The possible functions of arrestin are:

(1) it participates in the deactivation of GPCRs.;

(2) it may also be a linker, involved in the internalization of the receptor–arrestin complex—arrestin is assumed to combine with clathrin, helping to remove the receptor from the membrane through clathrin-coated pits; and, finally,

(3) it may recruit and activate an Src tyrosine kinase, promoting activation of the MAP kinase pathway and linking shut-off of the G-protein-linked receptor to activation of the MAP kinase pathway (Fig. 5.7).

Links of G-protein-coupled receptor signalling to the MAP kinase phosphorylation cascade

It has been known for quite some time that signalling pathways controlled by growth factors and cytokines accept signals from heptahelical receptors, coupled to heterotrimeric G proteins. However, the individual steps that funnel hormonal signals via G-protein-coupled receptors and heterotrimeric G proteins into the Ras/MAP kinase pathway were not well defined.[47] This state of affairs has changed recently.[48] We now have an idea, although not yet complete, how hormones may contribute to the regulation of cell proliferation via G-protein-coupled receptors.

The ports of entry of signals

One link is through arrestin, activating a c-Src cytosolic tyrosine kinase; the other link is through the momeric G protein, Rho, which communicates with the heteromeric G proteins, $G\alpha_{12}$ and $G\alpha_{13}$ (see Chapter 4); and still another link is the monomeric G protein, Rap 1, which responds to cAMP.

As shown in the laboratory of R. J. Lefkowitz by Luttrell *et al.*[46] and Lin *et al.*,[49] activation of c-Src and the MAP kinase by arrestin is coupled to β_2-AR and its removal from the membrane via clathrin-coated pits. Arrestin mutants that can not interact with clathrin are unable to internalize the receptor and to activate c-Src and MAP kinases. On the other hand, arrestin mutants that interact poorly with c-Src kinase, but interact with clathrin, are still functional in receptor sequestration, but can no longer activate MAP kinases in a β_2-AR-dependent manner.

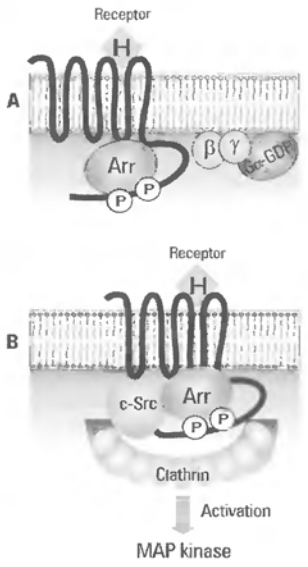


Fig. 5.7 (A) Arrestin binds to the phosphorylated receptor and deactivates the receptor. (B) Arrestin recruits clathrin to the receptor–arrestin complex and promotes receptor internalization through clathrin-coated pits. Clathrin is like the deck of an aircraft carrier, forming a platform where the receptor–arrestin complex can 'land'. After 'landing', the receptor–arrestin complex is transported by the carrier to its destination in the cell. Moreover, arrestin recruits the cytosolic c-Src tyrosine kinase to the arrestin–receptorcomplex and links receptor desensitization to the activation of the MAP kinase signalling pathway. (According to Zuker and Ranganthan.[48] Reproduced with permission of the authors and Science.)

This raises the question of how the linker function of arrestin is regulated. A likely possibility is phosphorylation dephosphorylation, because the free, unbound cytosolic form of arrestin is phosphorylated and is dephosphorylated on binding to the receptor. Moreover, only dephosphorylated arrestin can interact with c-Src kinase and activate MAP kinase. Luttrell *et al.* showed that arrestin interacts with the SH3 domain of c-Src, suggesting that arrestin binds to the SH3 domain which blocks activation of the kinase (see Chapter 3). What is not yet clear is how internalization of the receptor complex signals activation of the cytosolic tyrosine kinase and the MAP kinase. But this may turn out to be quite complicated. De Fea et al.[50] found that MAPK activation by the receptor PAR2 requires endocytosis of this G protein-coupled receptor. (PAR2 is a protease-activated receptor coupled to the heterotrimeric G protein, Gq). They proposed that activation of MAPkinase involves an endosomal multiprotein complex, which contains besides PAR2, arrestin, Raf and MAPK. So far so good, where there not the puzzling observation that endosomal activation blocks nuclear import of the activated MAPK. (When the MAPK

was activated by a non-endosomally-localized PAR-mutant, it ended up in the nucleus). How the cellular localization of the activation of MAPK by of a G protein-coupled receptor controls nuclear import of MAPK is not known, but it was speculated that there might be another, yet unknown role of the endosomal pathway. Another open question is how representative the PAR-pathway is for GPCR's in general.[51] While these question remain open, an endosomal signalling scaffold, a„ signalosome" bringing receptors together is an attractive model.[52]

The other opening to the MAP kinase phosphorylation cascade is an Rho-GEF, a guanine nucleotide exchange factor of Rho. As shown in Alfred G. Gilman's laboratory,[53] Rho-GEF contains an RGS-like domain through which it connects with the α-subunits of the G proteins, $G\alpha_{12}$ and $G\alpha_{13}$. Interaction of the RGS domain of the p115 Rho-GEF with the activated GTP-bound form of $G\alpha_{13}$ enhances the nucleotide-exchange activity of the Rho-GEF, whereas binding of $G\alpha_{12}$ to Rho-GEF interferes with activation by $G\alpha_{13}$. Thus, $G\alpha_{13}$ bound to Rho-GEF promotes GDP/GTP exchange and forms the active GTP-bound form of Rho. This is the pivotal step. In that way G-protein-coupled receptors which signal through $G\alpha_{13}$ activate Rho and Rho-controlled cellular processes. Important are also the signals to which receptors coupled to $G\alpha_{13}$ respond. As shown by Günter Schultz's laboratory these receptors are activated by lipid messengers,[54] by biologically active proteins such as bradykinin, bombesin, and thrombin, by cytokines such as interleukin-8 (IL-8), and by other factors (discussed in Chapter 4).

The p115 Rho-GEF is multifaceted. The structure contains, in addition to a RGS domain, a GAP domain and SH2-, SH3-, and PH-homology domains, proline-rich regions, and Ser/Thre phosphorylation sites. These different structural domains may mediate multiple protein–protein interactions, making Rho-GEFs versatile linkers. Moreover, additional regulatory controls may have to be considered, such as phosphorylation of Rho-GEFs.

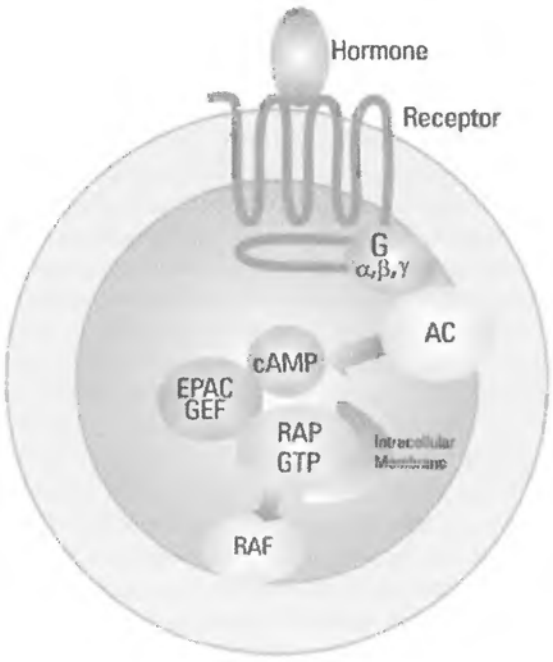


Fig. 5.8 A scenario where adenylyl cyclase is activated by a hormone-activated GPCR. The elevated cellular level of cAMP activates EPAC, the exchange protein directly activated by cAMP. The GDP/GTP exchange activity of EPAC promotes the formation of the active, GTP-bound form of the monomeric G protein, Rap1. Rap1 recruits and activates Raf-B, the upstream kinase of the MAP kinase pathway. (This scheme is shown with permission of the authors and Nature. ref. 56.)

Still another point of entry of signals is the monomeric G protein, Rap1. Rap1 is activated by cAMP. cAMP activates a GDP/GTP exchanger for Rap1, named EPAC (exchange protein directly activated by cAMP). EPAC shares homologies with other GEFs, but it also has sequences related to the regulatory subunit of the cAMP-dependent protein kinase, PKA. These sequences enable EPAC to bind cAMP directly. When the cAMP level in cells is high, EPAC is activated and its exchange activity turns on Rap1 by forming the GTP-bound form of Rap1. Rap1 then activates a homologue of the Raf kinase (Raf-B). The crystal structure of the binding domain of the Raf-1 serine/threonine kinase in a complex with Rap1A and a bound GTP analogue (GppNHp) has been solved.[55] It shown that Rap1 is a close relative of Ras. Although it is not yet clear how the Rap1 pathway affects cell function, such effects are expected, because Rap1 transmits signals of the second messenger, cAMP, which is known to change profoundly cell function and metabolism. Rap1 is the first monomeric G protein shown to respond directly to cAMP (Fig. 5.8).

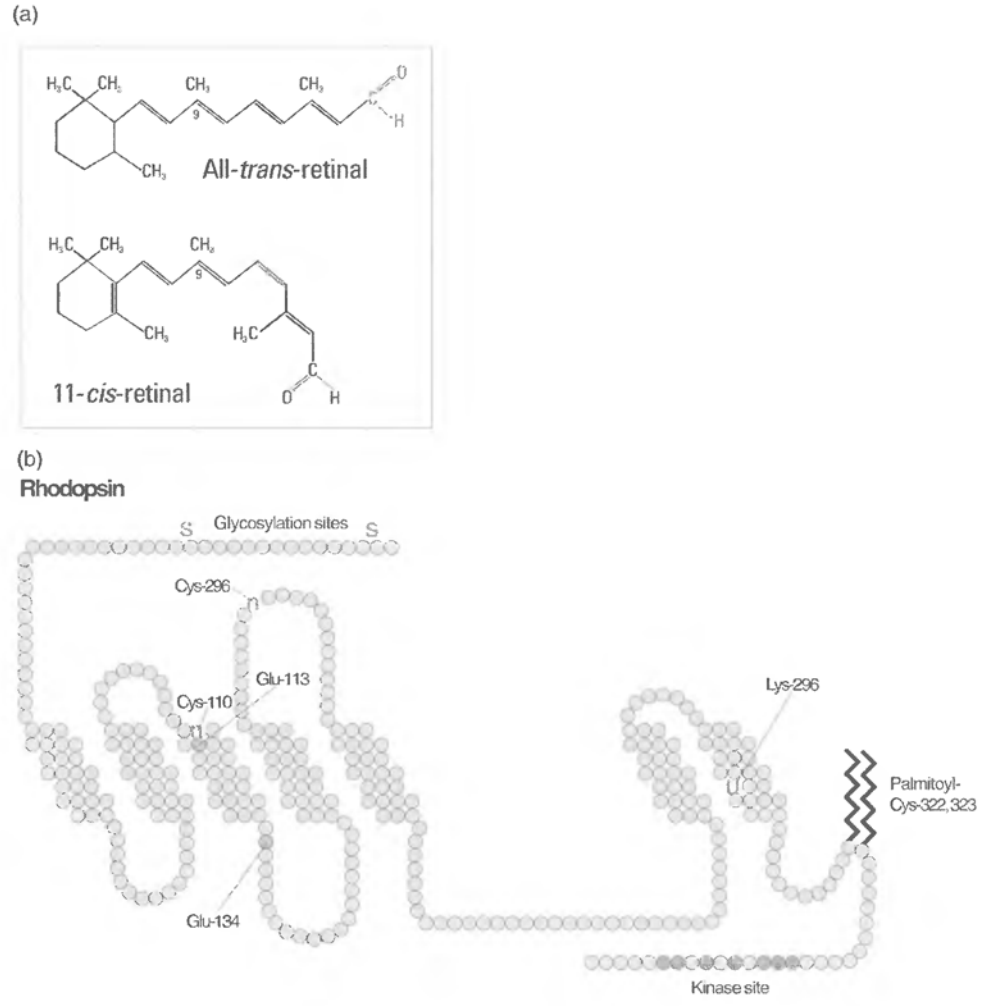


Fig. 5.9 (a) The structures of 11-*cis*- and all-*trans*-retinal. (b) The structure of rhodopsin.

Activation of the light sensor, rhodopsin

The light sensor rhodopsin is a heptahelical receptor with a built-in chromophore. Thanks to spectral changes in rhodopsin on illumination, discrete, sequential steps in the activation process could be distinguished. Since the relation of different activity states to their structural and molecular conformers, and finally to their functional states, is an issue that applies to all receptors, the information obtained with rhodopsin is considered in some detail (see ref. 57).

In the inactive, ground state, the chromophore, 11-*cis*-retinal, is bound via a protonated Schiff base to a lysine (Lys296) in the centre of the seventh transmembrane helical stretch. The inactive conformation is stabilized by a salt bridge between the protonated Schiff base and the carboxyl side-chain of Glu113 (Fig. 5.9).

Both in the case of sensory rhodopsin in humans and of bacteriorhodopsin (a heptahelical membrane protein in halobacteria which is not coupled to a G protein) translocation of a Schiff-base proton is the essential step in making the protein functional (reviewed in ref 58). In rhodopsin the conversion of the inactive MI state to the MII state that binds to the G protein is coupled to proton transfer from the Schiff base to the counterion, Glu113, and proton uptake from the medium to the highly conserved Glu134 , which serves as proton acceptor. Based on that similarity, one could consider sensory rhodopsin as an incomplete proton pump. Furthermore, a property shared by all G-protein-coupled receptors is a triplet, formed by residues 134–136 in rhodopsin, consisting of Glu–Arg–Tyr. The consequences of mutational replacement of Glu134 supports the notion that the state of protonation of this amino acid is crucial for activity, and that its protonation triggers the conformational transition of the receptor from the inactive to the active state.

Absorption of a photon leads to the efficient transition, with a high quantum yield (0.7), of a more twisted 11-*cis*- to a more stretched all-*trans*-conformation (Fig. 5.10).

Two-thirds of the energy acquired from the absorption of the photon remains in the opsin protein. From this energy-rich state rhodopsin passes through several intermediate states on the way to the signalling state which is able to bind and activate the G protein, transducin. The 9-methyl group of retinal is the allosteric trigger that opens the cross-talk

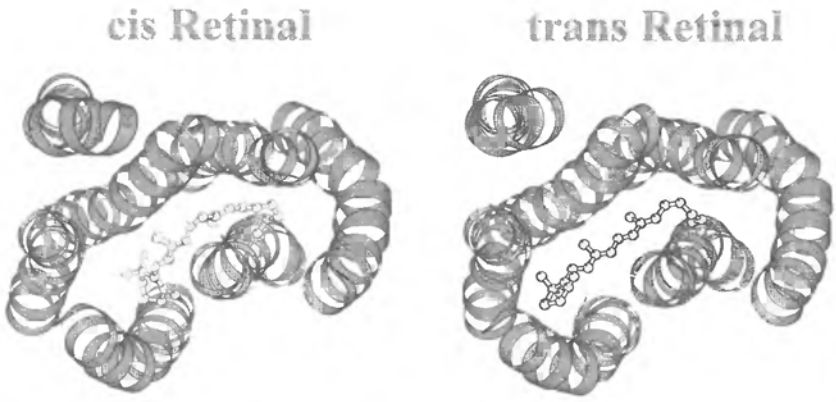


Fig. 5.10 Representation of the conformation of all-*trans*-retinal, and 11-*cis*-retinal as bound to rhodopsin.[58,59] Reproduced from ref. 58 with permission of Elsevier Science.

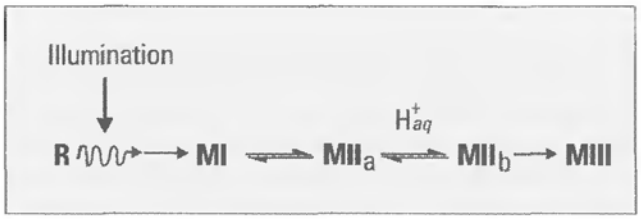


Fig. 5.11 The MII forms of rhodopsin (R) are in equilibrium with their parent compound MI and the progeny MIII. The equilibrium between the meta-states, MI, MIIa, and MIIb, are attained within a few milliseconds after illumination. The MI/MIIa transition is accompanied by the intramolecular translocation of the Schiff-base proton to Glu113, located in a region close to the lipid membrane. The subsequent MIIa/MIIb transition is dependent on proton uptake from the aqueous phase and occurs in a water-exposed area. Fast binding of the G protein, transducin, occurs only with MIIb.(Reproduced from ref. 58 with permission of Elsevier Science.)

between the retinal pocket and the intracellular interaction surface for the G protein. Correlated with the signalling states and the presentation of the interaction surfaces to the G protein is the formation of the so-called 'meta-states', MI, MIIa, and MIIb. There are at least five intermediate states through which rhodopsin passes on illumination, starting from an inactive state, with the apo-protein opsin liganded to 11-*cis*-retinal, and ending with the signalling all-*trans*-retinal state of rhodopsin (Fig. 5.11).

The most general statement that can be made is that a typical heptahelical receptor exists in several conformational states. Each of these states presents a topologically different domain to the G protein (or other regulatory proteins, interacting with the receptor). In rhodopsin there are three cytoplasmic loops that represent the actual interface with transducin. Each of these loops may contribute to a greater or lesser extent to the binding of transducin. If one compares the structure of rhodopsin with the structure of the β_2-adrenergic receptor, it becomes apparent that the β_2-adrenergic receptor has much larger and more elaborate intracellular loops than rhodopsin, suggesting that rhodopsin is an older, less elaborate and more specialized heptahelical receptor.

Studies on rhodopsin using a variety of physical methods by Gobind Khorana and colleagues[60] indicated a surprisingly small change in the protein on activation, involving residues near the N-terminal part of the second cytoplasmic loop. A primary goal is to characterize the intermediate states through which rhodopsin passes on the way to activation. The recent advances in high-resolution X-ray structures of bacteriorhodopsin,[61] will open the way for the identification of the intermediate states in the photocycle of the archaebacterium.[62] The present state of the art is reviewed by Dieter Oesterhelt and colleagues.[63] From these developments and the recently solved crystal structure of rhodopsin by Palczewski and colleagues,[2] the elucidation of the molecular photocycle events in illuminated rhodopsin will certainly profit.

Nucleotide exchange catalysis

Binding of the nucleotide, GTP, to the G protein, transducin, and binding of the G protein to activated rhodopsin, R*, are mutually exclusive.[64] It follows that the nucleotide-binding site on the α-subunit of the G protein must be empty when the

G protein should bind to R*.[65] Since other G proteins allow substantial nucleotide exchange without the receptor, an intermediate state with an empty nucleotide-binding site on the α-subunit in the R*–G complex may not be applicable to all receptor-catalysed nucleotide exchange reactions with G proteins. Once GTP is bound to the receptor, it triggers a conformational transition that leads to the dissociation of the R*–G complex and the separation of α-GTP and βγ-subunits, which transmit the signal to the target.

Conclusions

Activation of a typical heptahelical receptor and binding and activation of a G protein proceed in several steps: the formation of multiple signalling states with the sequential display of recognition surfaces for coupling partners applies to all G-protein-coupled receptors and, in a more general way, to all receptors. Multiple signalling states demonstrate the diversity and selectivity of receptor interactions. In the case of heptahelical receptors, the intracellular loop structures, exposed in a time-ordered manner, are instrumental in sequentially calling up different coupling partners. This is a kind of editing function and requires structural flexibility in the interplay of the ligand-binding domain with the regions involved in G-protein coupling. The behaviour of receptor mutants where G-protein activation was uncoupled from ligand binding to the receptor, confirms this kind of structural flexibility.

Signalling by heterotrimeric and monomeric G-proteins: analogies and differences

There follows a summary and comparison of the cogent features of signalling by heterotrimeric and monomeric G proteins.

Heterotrimeric G proteins

1. Heterotrimeric G proteins transmit and amplify signals received from heptahelical receptors in response to hormones and sensory signals.

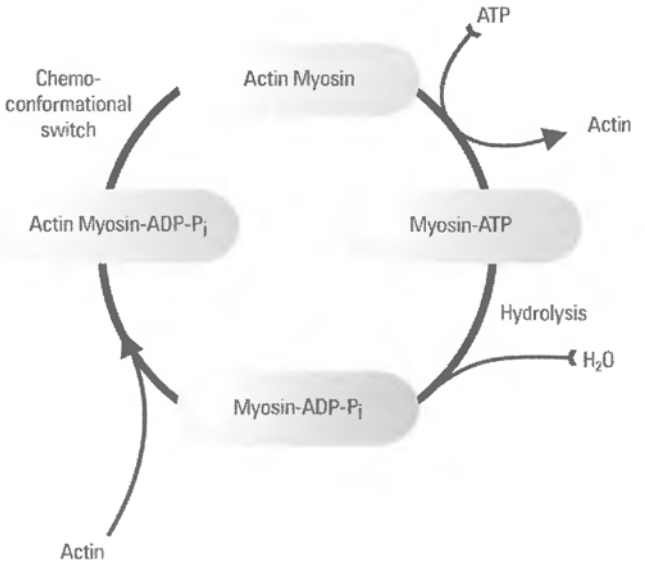


Fig. 5.12 ATP/ADP-regulated actin–myosin interaction. This scheme is based on the information presented by Drs Geeves and Conibear in ref. 66.

2. Nucleotide exchange, the first step of activation, is catalysed in the case of heterotrimeric G proteins by the receptor itself. The nucleotide-exchange reaction always proceeds from the GDP-bound to the GTP-bound form of the G protein, since only GTP is transformed catalytically by the receptor to the product GDP. This makes the binding of GTP a quasi-unidirectional step. It may be compared with a chemo-conformational transition, analogous to the ATP/ADP-regulated actin–myosin interaction, which also proceeds through a sequence of collisional coupling steps and a nucleotide-dependent interaction (Fig. 5.12).

3. Although the lifetime of the active state of monomeric and heterotrimeric G proteins is determined by the GTPase activity in both cases, the consequences of activation are different. Upon activation by an agonist or a sensory signal, heptahelical receptors and heteromeric G proteins interact directly with each other and produce many copies of two signalling entities, GTP-bound α-subunits and $\beta\gamma$-subunits. Signalling is terminated by receptor phosphorylation and disruption of receptor–G protein coupling. The signal response to heteromeric G proteins is highly amplified and large, mainly due to the activation of target enzymes synthesizing potent second messengers such as cAMP, or of enzymes catalysing their rapid decomposition, such as the removal of cGMP by a cGMP-specific phosphodiesterase in the retina. Second messengers regulate cellular metabolism and vital cellular functions, often through activation of messenger-specific kinases, such as PKA.

Monomeric G proteins

On the other hand, the activity of monomeric G proteins is regulated indirectly by receptors, through the intervention of separate linkers bringing GEFs and GAPs to the G protein. The role of monomeric G proteins is not so much to amplify the signal, but rather to modulate the connectivity between components of a signalling cascade and, if necessary, to translocate them to the membrane or to other locations in the cell. In that way, monomeric G proteins control the activity of phosphorylation cascades and signalling from the surface of the cell to the nucleus. Signals transmitted in that way lead to subtle, spatiotemporally ordered changes in gene expression.

Connections

Connections between heterotrimeric G-protein-coupled receptors, transmitting hormonal signals and phosphorylation cascades controlled by growth factors, and monomeric G proteins are generally made through linkers. These linkers need not always be GEFs or GAPs. An example is arrestin. But all linkers must have structural modules that facilitate recognition, making connections possible between the components of separate signal-transducing pathways, thus greatly increasing the range of signals to which a cell can respond.

G proteins and chemotaxis

A particularly important connection between signalling pathways controlled by heterotrimeric and monomeric G proteins is chemotaxis. Chemotactic signals are received by G-protein-coupled receptors and transmitted by $\beta\gamma$-subunits of heterotrimeric G proteins, notably to phosphoinositide 3-kinase (PtdIns 3-kinase) and phospholipase C. Finally, monomeric G proteins direct the signal to the cytoskeleton, promoting crawling and rolling of cells (see discussion of the role of phosphoinositides for the orientation of the cell towards stimuli in Chapter 4).

The chemotactic response to cAMP of the slime mould *Dictyostelium discoideum* is presented as an example. When these cells are starving, they sense cAMP signals, and in response to the hunger signal, cAMP, the cells differentiate (reviewed in ref. 67). In *Dictyostelium discoideum* the response to cAMP is mediated by G-protein-coupled heptahelical receptors and is transmitted by the βγ-subunits of a heterotrimeric G protein. In response to the chemoattractant, cAMP, a homologue of PKB (protein kinase B) is rapidly activated by phosphorylation through a PtdIns kinase specific for the 3-OH position (see also Chapter 4).

For human cells typical chemoattractants are cytokines, such as interleukin-8 (IL-8), and a formylated peptide (*N*-formyl-Met-Leu-Phe, fMLP).

The visual response

Vertebrates have two types of light-responsive photoreceptor cells: the rods and the cones. The photoreceptor cells convert light into nerve impulses. Cones are responsible for colour vision and function only in bright light, whereas rods also function in dim light. The retina in the human eye contains about 3 million cones and 100 million rods.

The target of the G protein transducin, G_{tr}-α, activated by rhodopsin and light in the retina, is an enzyme that decomposes the second messenger cGMP. The phosphodiesterase, PDE-αβγ2, is activated by G_{tr}-α GTP. Activation is due to removal of the inhibitory γ-subunit of the PDE. Consequently, the large decrease in cGMP in the membrane of the rod outer segment, (ROS), closes a cGMP-regulated cation channel. The resulting hyperpolarization triggers neuronal stimulation in the retina. Since both the transmitter (the G protein, which is a GTPase) and the effector are enzymes, signal amplification in the visual system is large. In response to a single quantum of light, 100 000 molecules of cGMP are hydrolysed to GMP.

'Shut-off' of the response

For termination of the visual response, the same tools are employed as in the shut-off of other hormonal and sensory responses. While phosphorylation of rhodopsin is an essential step in the shut-off reaction, phosphorylation alone is not sufficient for shut-off.[68] Even with three phosphates incorporated, the activity is only decreased tenfold. Therefore phosphorylation does not so much decrease the activity of rhodopsin, as introduce a phosphate marker for arrestin.

Electrophysiologists have pointed out that in daylight all rods are fully activated. Therefore in order to distinguish contrasts the response must be dampened. Damping of the response is part of the control of visual perception. It also protects the visual system against damaging overexposure to light.

Damping of the response

Several proteins downregulate signalling by G proteins. The RGS proteins (regulators of G-protein signalling) have already been discussed in this chapter. Phosducin and phosducin-like proteins have a comparable role in the visual system, and perhaps elsewhere. Other factors, such as recoverin, intercept Ca^{2+} signals and regulate dephosphorylation and resensitization. Both phosducin and recoverin are regulated by phosphorylation. Only the role of phosducin will be considered here.

Phosducin and phosducin-like proteins

Phosducin is a soluble, 33 kDa protein. It was originally discovered in the retina and in the pineal gland, a tissue related developmentally to the retina. In the rod outer segment membrane of the retina, as much phosducin as transducin is present. It is now established that phosducin and phosducin-like proteins (PhlPs) are evolutionarily conserved proteins, expressed in many species and tissues, where they may control G-protein signalling in other tissues than the eye or where they may have other, not yet known, functions. Whereas RGS proteins interact with the α-subunits of G proteins, phosducins bind to both α- and βγ-subunits, but with pronounced preference for βγ-subunits (Plate 12).

Binding of α- and βγ-subunits to phosducin in light-adapted rods blocks the re-formation of the αβγ-complex and reduces the turnover of the G-protein cycle. Consequently, phosducin dampens the response to light. In the dark-adapted eye, phosducin is inactive, because it is phosphorylated by the cAMP-dependent protein kinase A,[71] and perhaps by other kinases. In the light phospho-phosducin is reactivated by dephosphorylation.

The role of the glutamic-acid-rich proteins in rod photoreceptors

Another type of protein that may dampen unnecessary signalling activity in daylight, when all rods are fully saturated, is the GARP-like protein. This reduces cGMP turnover in the light. Rod photoreceptors, but not cone receptors, contain glutamic-acid-rich proteins (GARPs).[72] They are multivalent proteins that interact with the major components of the signalling assemblage in rod photoreceptor cells (Fig. 5.13). A soluble splice form, GARP2, associates more strongly with the light-activated than with the non-activated phosphodiesterase, inhibiting its activity.

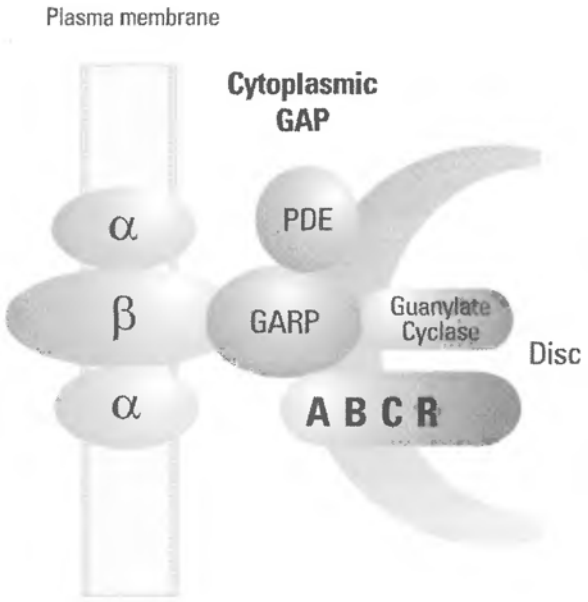


Fig. 5.13 GARPs are located in a cytoplasmic GAP, in close proximity to the guanylate cyclase in the disc, which synthesizes the second messenger cGMP from GTP. GARPs also have access to the cGMP phosphodiesterase, PDE. Moreover, the β-subunit of the cGMP-gated channel contains a large cytoplasmic domain which interacts with GARP. Finally, GARPs interact with a retina-specific, ATP-binding transporter, a member of the ABC family of transporters. ABC transporters belong to a family of proteins that carry many different molecules through the membrane. Transport is driven by ATP. ABC is an abbreviation of *ATP-binding cassette*. The ABC protein is located at the outer rim of the discs. (It is also called the 'rim' protein, ABCR).[73] (Reproduced with permission of the authors

Resensitization by dephosphorylation

Dephosphorylation is a step in the resensitization of all G-protein-coupled receptors. Dephosphorylation of rhodopsin in the human eye is regulated by Ca^{2+} and recoverin, a Ca^{2+}-regulated protein. Rhodopsin phosphatases are present in humans and mice,[75] and mutations of rhodopsin phosphatase have been implicated in degenerative retinal diseases, including retinitis pigmentosa, in humans. The phosphatase dephosphorylating rhodopsin is apparently a typical serine/threonine phosphatase. The laboratories of Zvi Selinger and B. Minke at the Hebrew University in Jerusalem have studied phototransduction in the eye of *Drosophila* and have shown that downregulation of the photoresponse in the fly involves arrestin and a Ca^{2+}-regulated phosphatase which dephosphorylates *Drosophila* rhodopsin.[76] The *Drosophila* eye phosphatase is a GPCR-coupled receptor-regulated serine/threonine phosphatase. This phosphatase has an N-terminal domain that has high sequence similarity with the catalytic domain of other Ser/Thr protein phosphatases (Chapter 7). The C-terminal domain has multiple EF-hand motifs and binds Ca^{2+}. The EF-hand is a helix-turn helix motif, (see Chapter 9), that binds Ca^{2+}. It was called EF-hand because it was first recognized in parvalbumin in which helices E and F form the motif. Mutations in *Drosophila*, resulting in loss of the phosphatase, are associated with severe defects in the light response and retinal degeneration. The phenotypes of these *Drosophila* mutants have hyperphosphorylated rhodopsin.

Summary

Considering that the visual (and the olfactory) systems operate with 10- to 100 000-fold amplification rates, it is plausible that the turning-off switch must work as reliably and as fast as the turning-on switch. If the off switch does not work properly, the retina becomes damaged, as in some cases of retinitis pigmentosa. In the active state, rhodopsin can catalyse GDP–GTP exchange at a higher rate than 1000 molecules/s. In the inactive state, in the dark, this rate is unmeasurably low. This implies that the interaction sites of rhodopsin with the G protein are exposed effectively in the active state and concealed perfectly in the inactive state. Deactivation is initiated by phosphorylation, catalysed by receptor kinases, and resensitization is initiated by serine/threonine-specific receptor phosphatases, preventing hyperphosphorylation of rhodopsin, which is detrimental to the integrity of the photoreceptor cell.

Olfaction

Olfaction, like visual and taste perception, is an ancient process. Olfaction plays a role in sexual arousal. The olfactory system in mammals is remarkable with respect to the number of receptors engaged in monitoring odours. There are several thousand heptahelical G-protein-coupled receptors in the olfactory epithelium and the nasal organ of a dog, and still about 1000 receptors in the corresponding human organs. It has been estimated that nearly 1% of all genes code for olfactory receptors alone.

The efficacy of odour recognition is very variable, not only in different species but also in individuals. The condition, (anosmia), where an individual cannot perceive an odour at concentrations that are detected readily by others is frequent. Desensitization on long-time exposure to odorants and resensitization are probably similar to the corresponding processes occurring in other sensory systems, although these processes are not as well-defined as in the hormonal and the visual systems.

Olfactory neuronal connections: one neuron, one receptor

One thousand or more different olfactory receptors transmit their signals through the axons of olfactory neurons. Olfactory neurons are clustered in the olfactory epithelium,[77] and project their axons to a small number of topographically fixed loci, the glomeruli, in the olfactory bulb of the brain, where the information is processed. Each individual glomerulus in the olfactory bulb apparently corresponds to a distinct type of receptor. Olfactory neurons have specific marker proteins (OMP, olfactory marker protein) which have facilitated the localization of olfactory neurons.

Local neuronal circuits among the glomerular loci in the olfactory bulb fine tune the odour response. The signal is then transmitted to output neurons. Oscillatory discharges among output neurons may then help to integrate the odorant signals in the olfactory regions of the brain cortex.[78] (Reviewed in ref. 79.)

Each olfactory neuronal cell has only one receptor, although this is disputed. Visualization of individual axons from sensory neurons, each having only one receptor, has suggested that activation of the olfactory receptor guides the axon to its site in the brain.

Many fascinating problems remain to be solved. For example, one would like to know to what extent the odour perception map in the brain is engraved by usage, and/or to what degree it is genetically preformed. Another question is how the information is processed and modified in the neocortex. After all, odour perception can cause revulsion but also can give pleasure, a fact exploited by a huge industry producing perfumes and toiletries.

Odour perception has been compared with the immune system. But receptor activation in the olfactory system seems to have an instructive rather than a selective role. An instructive role would be in agreement with the large genetic repertoire coding for a receptor for each type of odorant molecule.

Odorants

Small proteins in the mucus surrounding the sensory epithelial cells bind the odorants and bring them across the mucus to the receptor. These odorant-binding proteins (OBPs) have been cloned and identified (the binding proteins for retinol (vitamin A) and cholesterol belong to the same family). The OBPs function as carriers of small lipophilic compounds. Terrestrial animals, in contrast to aquatic animals, can only smell volatile, lipophilic odorants.

Receptors recognizing hydrophobic, lipophilic, and hydrophilic odorants seem to belong to different classes. Odorant receptors are also expressed in mammalian spermatocytes, where they bind odorants that guide them to the Fallopian tubes.

Whereas there are thousands of odorant molecules and odorant receptors, there are only some tens of OBPs in any one species. The three-dimensional structure of a bovine odorant-binding protein has been solved.[80] Structural information has provided a clue to how a single dimeric OBP can accomodate so many different odorant molecules. But it is not clear whether OBPs are the only proteins carrying odorants to their receptors, nor is it clear how they find the right receptor. Moreover, at present it remains questionable whether OBPs have other functions, for example as scavengers, removing odorants.

The inactivation and removal of odorants is essential, because the preceding wave of odorant must be extinguished to make place for a fresh wave of new odorants. Odorants are rapidly decomposed by enzymes catalysing detoxification reactions. The biochemical transformation makes odorants water soluble and insoluble in lipids, so that they can no longer go to the odorant receptor and are excreted instead.

Odorant receptors

The discriminatory power of the olfactory system in mammals is extraordinary. It recognizes a multitude of structurally different odorous molecules. Specific recognition of odorants occurs at the receptor level. Linda Buck and Richard Axel have cloned the first odorant receptor.[81] Subsequent work has proved the existence of a multigene family of G-protein-coupled heptahelical receptors which are expressed exclusively in the olfactory epithelium, confined to a small subset of neurons. Odorant receptors have been cloned from different vertebrate species, including human, rat, mice, and fish, from nematodes, and from *Drosophila*.[82] The high degree of sequence variability in the central helical transmembrane region of olfactory receptors, containing the odorant-binding pocket, corresponds to the diversity of odorants.

The odorant receptors are closely related to rhodopsin. Like the light-responsive receptor, they have relatively short intra- and extracellular loops and short N- and C-termini. This simple structural design has been thought to indicate that the sensory receptors are evolutionarily the oldest members of the superfamily of G-protein-linked heptahelical receptors. The human olfactory receptor genes have no introns and are clustered. Coding regions for 16 olfactory receptor genes were localized in a 0.35 Mb stretch (where 1 Mb = 10^6 base pairs) on human chromosome 17. The clustered location of olfactory receptor genes in the chromosome corresponds to defined localized zones of receptors in the olfactory epithelium.

From a comparison of protein sequences (Fig. 5.14) it became apparent that the *Drosophila* olfactory receptors have no homologies with the corresponding receptors of other species. Thus, olfaction is, unlike other fundamental signalling pathways, where family relationships can be traced back over hundreds of millions of years of evolution,

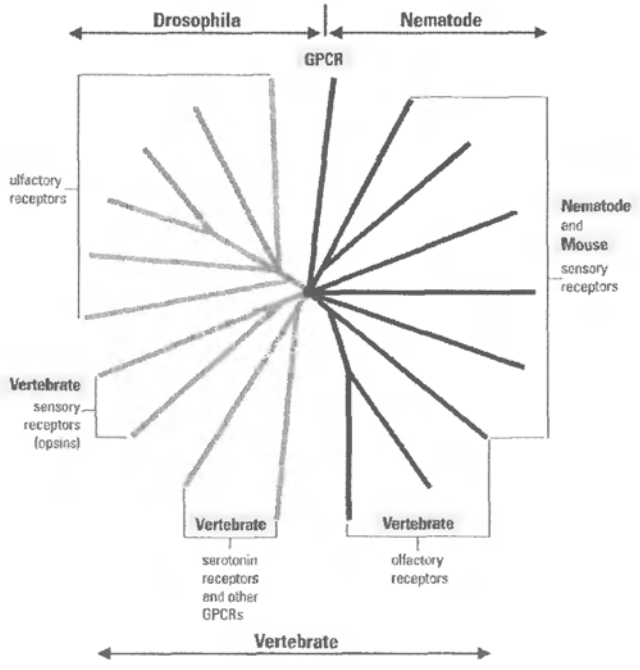


Fig. 5.14 Dendrogram of the protein sequences of chemosensory receptors and G-protein-coupled receptors, GPCRs. Although rhodopsin and serotonin receptors are clearly related between humans and *Drosophila*, the *Drosophila* olfactory receptors are unrelated to the olfactory receptors in vertebrates. (Reproduced from ref. 83 with permission of the authors and Nature.)

from mammals and vertebrates to *Drosophila* and *C. elegans* (83 see also ref. 84). Olfactory receptors are an example of convergent evolution, reminiscent of the relationship between rhodopsin and bacteriorhodopsin, two proteins that have no sequence similarities but a similar structure. Yitzhak Pilpel and Doron Lancet, have pointed out that olfactory receptor genes of *Drosophila* and *C. elegans* have introns (non-coding sequences), whereas GPCRs, including the olfactory receptor genes in vertebrates, are intronless. They have related the loss of introns in vertebrates to the extensive gene duplication required for the development of the large repertoire of olfactory receptors in vertebrates.[83,85]

When each individual olfactory sensory cell should expresse only one specific receptor it would require extraordinarily strict control of gene expression. For the study of gene expression, R.R. Reed and his laboratory have used mutants in *Drosophila melanogaster* and *C. elegans* that have defects in olfactory function.[86–89] These studies led to the characterization of olfactory transcription factors, which control the transcription of olfactory receptors. The transcription factor, Olf-1, identified by R. R. Reed, activates a number of olfactory receptor genes.[89] Olf-1 is expressed only in olfactory epithelia. It is one of several related, highly conserved transcriptional activators that control the differentiation of olfactory neurons.

In the immune system, each immune-competent cell of a certain clone also expresses only one kind of receptor of a given specificity (Chapter 14), but in the immune system this is achieved by a selection process in the course of development of lymphocytes and a combinatorial arrangement of constant immunoglobulin chains with a number of variable chains, generated through somatic mutations and recombination. This allows the expression of a nearly unlimited number of antigen receptors, one for each committed cell clone, and at the same time avoids an unrealistic expansion of the genetic repertoire. In contrast to the immune system, a huge repertoire of genes has apparently been set aside for the olfactory response. In the immune system, a ligand (the epitope of an antigen) no longer needs to make choices between different cells and receptors, because for each ligand a competent immune cell has already been selected. The selected cell, when stimulated by the matching epitope, divides, forming a clone of the same cell. On the contrary, each odorant molecule must screen and discriminate between many different olfactory cells, each expressing an olfactory receptor, until it finds the receptor with the best fit. Another difference is that once the receptor is found the bound odorant cannot order the receptor-containing neuronal cell to divide, forming a clone of cells, each recognizing the same ligand.

There exists a family of olfactory receptor kinases, similar to the β-adrenergic and the rhodopsin kinases. In the olfactory system, a cytosolic G-protein-coupled receptor kinase, GCRK 3, binds prenylated γ-subunits of G proteins which assist in the translocation of the kinase to the receptor in the membrane. The receptor is phosphorylated and binds a homologue of arrestin, expressed in the olfactory system.[90] Thus, olfactory receptor deactivation is, in principle, like deactivation of hormonal and visual receptors, and so apparently is reactivation, which is assumed also to involve dephosphorylation.

Signal transduction in olfactory neurons

The principal target of odorant-receptor G-protein-mediated signals is adenylyl cyclase. cAMP formed by the cyclase activates a cAMP-dependent Na^+/K^+ channel. This is like

the activation of a cation channel by cGMP in the retina. But, whereas in the visual response the channel is closed on receptor activation, because cGMP is rapidly degraded to inactive GMP, in the olfactory system the channel is opened as cAMP rapidly accumulates. The response is highly amplified, as in the visual system. The olfactory neurones are stimulated and an excitatory, electrical signal is transmitted to the brain. The olfactory channel has been cloned. It belongs to the family of CNGCs (cyclic nucleotide-gated channels) that are structurally similar to voltage-gated K^+ and Ca^{2+} channels. The same kinds of channels are found in photoreceptor-, olfactory-, and taste-receptor cells, and also in non-sensory cells.

A potent activator of the cAMP, cyclic nucleotide-gated (CNG) Na^+/K^+ channel in olfactory cells is NO (nitric oxide), which is many times more potent than cAMP. However, NO is present only in developing olfactory neurons, where it seems to react directly with cysteine SH-groups in the β-subunit of the channel. It was speculated, therefore, that NO might be a channel activator early in development or in regenerating olfactory neurons (see: ref. 91).

But not all odorants activate adenylyl cyclase. This has motivated to look for additional signal transduction pathways. Inositol trisphosphate (IP_3) may be an alternative second messenger in response to odorant stimulation. This would be in agreement with the existence of a Ca^{2+}- responsive channel in olfactory cells, because intracellular Ca^{2+} is released in response to IP_3. Ca^{2+} may also play a role in damping the response to odorants by modulating the activity of the CNGC and/or the activity of other Ca^{2+}-responsive proteins, regulating signalling. The latter effects of Ca^{2+} would be similar to the action of recoverin in the visual response.

Taste transduction

Taste transduction is initiated when taste stimuli interact with exposed cells in the apical microvilli of the tongue. The receptor–ligand interaction leads to membrane depolarization and to activation of afferent gustatory neurons. The perception of the common human taste qualities—sweet, bitter, salty, and sour—has been assigned to groups of taste neurons. But a characteristic property of taste transduction is that the taste cells can also be stimulated directly without the intervention of receptors.

In contrast to visual perception, where the sole stimulus is a photon, and even in contrast to olfaction with its structurally much more diversified stimuli, taste perception is exceptional, because the taste stimuli differ even more than odorants in size and chemical complexity, ranging from H^+ ions to carbohydrates, amino acids, and proteins. Consequently, taste transduction might involve different mechanisms for different stimuli. In the rhesus monkey,[92] taste reception has been located anatomically to defined loci at either the anterior or the posterior part of the tongue.

Most ionic taste stimuli depolarize the taste cells by interacting directly with ion channels, without receptor interactions. For example, the salt taste involves first translocation of Na^+ ions into the cell through Na^+ channels, or by passage through tight junctions. In the taste cell, Na^+ activates a chloride channel. Sour taste is transduced through a variety of mechanisms, including entry of protons through K^+ and Na^+ channels.[93] In contrast to ionic stimuli, amino acids (such as glutamate, an ingredient in Chinese food), synthetic sweeteners, and most bitter-tasting stimuli bind to G-protein-coupled heptahelical membrane receptors. The first taste receptors in vertebrates have recently been cloned.[94]

These receptors are coupled to a variety of taste-cell-specific heterotrimeric G proteins, such as gustducin, which has been localized, together with transducin, in taste buds. Transgenic mice lacking gustducin have lost their aversion for bitter stimuli, but also have less liking for sucrose than normal mice. Thus, gustducin may participate in transduction of sweet and bitter stimuli. Gustducin has sequence homology with transducin. Therefore, a function like that of transducin in visual signal transmission would be plausible. However, in contrast to visual signalling, a cyclic nucleotide-gated cation channel is opened, rather than closed, in taste transduction, and cAMP rather than cGMP regulates the cation channel. This is more like the situation in olfaction, where an increase in cAMP, rather than a decrease of the cyclic nucleotide, controls channel activity. Because cAMP mimics the effects of sweeteners, it is assumed that sugars and sweeteners activate adenylyl cyclase via G-protein, gustducin-coupled receptors, and cAMP activates a cAMP-regulated cation channel in taste buds, producing an excitatory electrical signal, which eventually activates afferent gustatory neurons. cAMP may not be the only messenger in taste transduction. A participation of Ca^{2+} in taste reception is likely: bitter stimuli and some common synthetic sweeteners, such as aspartame (a dipeptide composed of phenylalanine and aspartate), have been shown to stimulate phospholipase C, causing formation of IP_3 and release of Ca^{2+} from intracellular stores.

Summary

1. Receptors participating in sensory perception are G-protein-coupled heptahelical receptors, expressed in cells specialized for sensory responses. The mechanisms involved in sensory signal transduction are basically the same as in hormonal signal transmission, although the signals are very different.

2. A characteristic property of olfactory neurons is that each neuronal cell has apparently only one receptor. In the immune system, a competent cell also expresses only one receptor and recognizes also only one kind of ligand (stimulus), but whereas in the immune system the gene repertoire required for recognition of all conceivable antigenic stimuli has been restricted in the course of development of immune cells, in the olfactory system a large gene repertoire has been set aside to endow each individual neuronal cell with a specific receptor. This burdens each odorant with the task to screen all the receptors that it encounters, until it finds the one with the best fit.

References

1. C. D. Strader, T. M. Fong, M. P. Graziano, and M. R. Tota. The family of G-protein-coupled receptors. *FASEB J*, **9** (9), 745–754, 1995.
2. K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, **289**, 739–745, 2000.
3. V. M. Unger and G. F. Schertler. Low resolution structure of bovine rhodopsin determined by electron cryo-microscopy. *Biophys J*, **68** (5), 1776–1786, 1995.
4. P. L. Yeagle, J. L. Alderfer, and A. D. Albert. Structure of the carboxy-terminal domain of bovine rhodopsin. [letter]. *Nature Struct Biol*, **2** (10), 832–834, 1995.
5. R. A. Dixon, B. K. Kobilka, D. J. Strader, J. L. Benovic, H. G. Dohlman, T. Frielle, *et al.* Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. *Nature*, **321** (6065), 75–79, 1986.

6. S. Trumpp Kallmeyer, J. Hoflack, A. Bruinvels, and M. Hibert. Modeling of G-protein-coupled receptors: application to dopamine, adrenaline, serotonin, acetylcholine, and mammalian opsin receptors. *J Med Chem*, **35** (19), 3448–3462, 1992.
7. C. M. Burns, H. Chu, S. M. Rueter, L. K. Hutchinson, H. Canton, E. Sanders-Bush, and R. B. Emeson. Regulation of serotonine-2C receptor G-protein coupling by RNA editing. *Nature*, **387**, 303–308, 1997.
8. S. Tsunoda, *et al.* A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade. *Nature*, **388**, 243–249, 1997.
9. J.H. White, *et al.* Heterodimerization is required for the formation of a functional $GABA_B$ receptor. *Nature*, **396**, 679–682, 1998.

   Also: B. A. Jordan and L. A. Devi. G-protein-coupled receptor heterodimerization modulates receptor function. *Nature*, **399**, 697–700, 1999.

   Also: R. Kuner, *et al.* Role of heteromer formation in $GABA_B$ receptor fiunction. *Science*, **283**, 74–77, 1999.
10. E. W. Sutherland. Studies on the mechanism of hormone action. *Science*, **177**, 401–408, 1972.
11. A. G. Gilman. G proteins: transducers of receptor-generated signals. *Annu Rev Biochem*, **56**, 615–649, 1987.
12. M. I. Simon, M. P. Strathmann, and N. Gautam. Diversity of G proteins in signal transduction. *Science*, **252** (5007), 802–808, 1991.
13. J. J. G. Tesmer, R. K. Sunahara, A. G. Gilman, and S. R. Sprang. Crystal structure of the catalytic domains of adenylylcyclase in a complex with Gsa GTPyS. *Science*, **278**, 1907–1919, 1997.
14. J.G. Tesmer, R. K. Sunahara, R. A. Johnson, G. Gosselin, A. G. Gilman, and S. R. Sprang. Two-metal-ion catalysis in adenylyl cyclase. *Science*, **285**, 756–760, 1999.
15. P. J. Artymiuk, A. R. Poirrette, D. W. Rice, and P. Willett. A polymerase I palm in adenylyl cyclase? *Nature*, **388**, 33–34, 1997.
16. H. Heithier, M., Fröhlich C. Dees, M. Baumann, M. Haring, P. Gierschik, E. Schiltz, W. L. Vaz, M. Hekman, E. J. M. Helmreich. Subunit interactions of GTP-binding proteins. *Eur J Biochem*, **204** (3), 1169–1181, 1992.
17. L. Birnbaumer. Receptor-to-effector signaling through G proteins: roles for beta gamma dimers as well as alpha subunits. *Cell*, **71** (7), 1069–1072, 1992.
18. U. Mende, C. J. Schmidt, F. Yi, D. J. Spring, and E. J. Neer. The G protein gamma subunit. Requirements for dimerization with beta subunits. *J Biol Chem*, **270** (26), 15892–15898, 1995.
19. A. Katz and M. I. Simon. A segment of the C-terminal half of the G-protein beta 1 subunit specifies its interaction with the gamma 1 subunit. *Proc Natl Acad Sci, USA*, **926** 1998–2002, 1995.
20. K. Ray, C. Kunsch, L. M. Bonner, and J. D. Robishaw. Isolation of cDNA clones encoding eight different human G protein gamma subunits, including three novel forms designated the gamma 4, gamma 10, and gamma 11 subunits. *J Biol Chem*, **270** (37), 21765–21771, 1995.
21. N. Ueda, J. A. Iniguez Lluhi, E. Lee, A. V. Smrcka, J. D. Robishaw, and A. G. Gilman. G protein beta gamma subunits. Simplified purification and properties of novel isoforms. *J Biol Chem*, **269** (6), 4388–4395, 1994.
22. J. B. Higgins and P. J. Casey. *In vitro* processing of recombinant G protein gamma subunits. Requirements for assembly of an active beta gamma complex. *J Biol Chem*, **269** (12), 9067–9073, 1994.
23. E. J. Neer. G proteins: critical control points for transmembrane signals. *Protein Sci*, **3** (1), 3–14, 1994.
24. E. J. M. Helmreich and K.-P. Hofmann. Structure and function of proteins in G protein-coupled signal transfer. *Biochim Biophys Acta*, **1286**, 285–322, 1996.
25. J. Sondek, A. Bohm, D. G. Lambright, H. E. Hamm, and P. Sigler. Crystal structure of a G protein $\beta\gamma$ dimer at 2.1 Å resolution. *Nature*, **379**, 369–374, 1996.
26. J. A. Pitcher, J. Inglese, J. B. Higgins, J. L. Arriza, P. J. Casey, C. Kim, *et al.* Role of beta gamma subunits of G proteins in targeting the beta-adrenergic receptor kinase to membrane-bound receptors. *Science*, **257** (5074), 1264–1267, 1992.
27. J. Inglese, W. J. Koch, M. G. Caron, and R. J. Lefkowitz. Isoprenylation in regulation of signal transduction by G-protein-coupled receptor kinases. *Nature*, **359** (6391), 147–150, 1992.
28. J. Kurjan. The pheromone response pathway in *Saccharomyces cerevisiae*. *Annu Rev Genet*, **27**, 147–179, 1993.
29. G. F. Sprague Jr and J. Thorner. In Pheromone response and signal transduction during the mating process of Saccharomyces cerevisiae, (ed. J. R. Broach, J. R. Pringle, and E. W. Jones), pp. 657–744. Cold Spring Harbor Laboratory Press, New York, 1992.
30. N. Watson, M. E. Linder, K. M. Druey, J. H. Kehrl, and K. J. Blumer. RGS family members: GTPase activating proteins for heterotrimeric G-protein α-subunits. *Nature*, **383**, 172–175, 1996.

31. M. R. Koelle and H. R. Horvitz. EGL-10 regulates G protein signalling in the *C. elegans* nervous system and shares a conserved domain with many mammalian proteins. *Cell*, **84**, 115, 1996.
32. H. G. Dohlman, *et al.* Inhibition of G-protein signalling by dominant Gain-of-function mutations in Sst2p, a pheromone desensitization factor in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **15**, 3635–3643, 1995.
33. C. Dietzel and J. Kurjan. Pheromonal regulation and sequence of the *Saccharomyces cerevisiae* SST2 gene: a model for desensitization to pheromone. *Mol Cell Biol*, 7 (12), 4169–4177, 1987.
34. T. W. Hunt, T. A. Fields, P. J. Casey, and E. G. Peralta. RGS10 is a selective activator of Gai GTPase activity. *Nature*, **383**, 175–177, 1996.
35. J. J. Tesmer, D. M. Berman, A. G. Gilman, and S. R. Sprang. Structure of RGS4 bound to AlF_4^- activated Gαi ; stabilisation of the transition state for GTP hydrolysis. *Cell*, **89**, 251–261, 1997.
36. W. P. Hausdorff, M. G. Caron, and R. J. Lefkowitz. Turning off the signal: Desensitization of β-adrenergic receptor function. *FASEB J*, **4**, 2881–2889, 1990.
37. S. Collins, M. G. Caron, and R. J. Lefkowitz. Regulation of adrenergic receptor responsiveness through modulation of receptor gene expression. *Annu Rev Physiol*, **53**, 497–508, 1991.
38. M. J. Lohse. Molecular mechanisms of membrane receptor desensitization. *Biochim Biophys Acta*, **1179**, 171–188, 1993.
39. E. J. M. Helmreich. Hormone–receptor interactions. *FEBS Lett*, **61**, 1–5, 1976.
40. U. Wilden, S. W. Hall, and H. Kuhn. Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-kDa protein of rod outer segments. *Proc Natl Acad Sci USA*, **83** (5), 1174–1178, 1986.
41. S. Ishiguro, Y. Suzuki, M. Tamai, and K. Mizuno. Purification of retinol dehydrogenase from bovine retinal rod outer segments. *J Biol Chem*, **266**, 15520–15524, 1991.
42. K. Palczewski, J. H. McDowell, S. Jakes, T. S. Ingebritsen, and P. A. Hargrave. Regulation of rhodopsin dephosphorylation by arrestin. *J Biol Chem*, **264**, 15770–15773, 1989.
43. K. Palczewski, S. Jäger, J. Buczylko, R. Crouch, D. L. Bredberg, K. P. Hofmann, M. A. Asson-Batres, and J. C. Saari. Rod outer segment retinol dehydrogenase: substrate specificity and role in phototransduction. *Biochemistry*, **33**, 13741–13750, 1994.
44. J. Granzin, U. Wilden, Hui-Woog Choe, J. Labahn, B. Krafft, and G. Büldt. X-ray crystal structure of arrestin from bovine rod outer segments. *Nature*, **391**, 918–921, 1998.
45. J. A. Hirsch, C. Schubert, V. V. Gurevitch, and P. B. Sigler. The 2.8 Å crystal structure of visual arrestin. *Cell*, **97**, 257, 1999.
46. L. M. Luttrell, S. S. G. Ferguson, Y. Daaka, W. E. Miller, S. Maudsley, G. J. Della Rocca, *et al.* β-arrestin-dependent formation of β2 adrenergic receptor –Src protein kinase complexes. *Science*, **283**, 655–661, 1999.
47. A. Hall. G proteins and small GTPases: Distant relatives keep in touch. *Science*, **280**, 2074–2075, 1998.
48. C. S. Zuker and R. Ranganathan. The path to specificity. *Science*, **283**, 650–651, 1999.
49. F. T. Lin, K. M. Krueger, H. E. Kendall, Y. Daaka, Z. L. Fredericks, J. A. Pitcher, and R. J. Lefkowitz. Clathrin-mediated endocytosis of the beta-adrenergic receptor is regulated by phosphorylation/dephosphorylation of beta-arrestin1. *J Biol Chem*, **272** (49), 31051–31057, 1997.
50. K. A. De Fea, J., Zalevsky, M. S., Thoma, O. Dery, R. D. Mullins, and N. W. Bunnett, β-Arrestin-dependent endocytosis of proteinase-activated receptor 2 is required for intracellular targetting of activted ERK1/2. *J. Cell Biol*, **148**, 1267–1281, 2000.
51. J. S. Gutkind, Regulation of mitogen-activated protein kinase signaling networks by G protein-coupled receptors. *Science STKE:* wwwstke.org/cgi/content/full/OC-sigtrans; 2000/47/pe 1.

    See also: J. Whistler, E. Beattie, M. von Zastrow, Tales from the Crypt: Evidence for Heptahelical Receptor Signaling in the Endocytic Pathway. *Science's STKE*-http://www.stke.org/cgi content full/OC-sigtrans; 2000/47/pe 1.
52. W. R. Burack, and A. S. Shaw, Signal transduction: hanging on a scaffold. *Curr. Opin. Cell Biol*, **12**, 211–216, 2000.
53. M. J. Hart, X. Jiang, T. Kozasa, W. Roscoe, W. D. Singer, A. G. Gilman, *et al.* Direct stimulation of the guanine nucleotide exchange activity of p115 RhoGEF by Ga13. *Science*, **280**, 2112–2114, 1998.
54. A.Gohla, R. Harhammer, and G. Schultz. The G-protein G_{13} but not G_{12} mediates signalling from lysophosphatidic acid receptor via epidermal growth factor receptor to Rho. *J Biol Chem*, **273**, 4653, 1998.
55. N. Nassar, G. Horn, C. Herrmann, A. Scherer, F. McCormick, and A. Wittinghofer. The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf 1in complex with Rap 1A and a GTP analogue. *Nature*, **375**, 554–560, 1995.

56. J. de Rooij, F. J. T. Zwartkruis, M. H. G. Verheijen, R. H. Cool, S. M. B. Nijman, A. Wittinghofer, and J. L. Bos. Epac is a Rap-1 guanine-nucleotide exchange factor directly activated by cyclic AMP. *Nature*, **396**, 474–477, 1998.
57. L. Stryer. Cyclic GMP cascade of vision. *Annu Rev. Neurosci*: 9, 87–119, 1986.
58. E. J. M. Helmreich and K. -P. Hofmann. Structure and formation of proteins in G protein-coupled signal transfer. *Biochim. Biophys. Acta*, **1286**, 285–322, 1996.
59. J. M. Baldwin. *Curr Opin Cell Biol*, **6**, 180–190, 1994.
60. Z. T. Farahbakhsh, K. D. Ridge, H. G. Khorana, and W. L. Hubbell. Mapping light-dependent structural changes in the cytoplasmic loop connecting helices C and D in rhodopsin: a site-directed spin labeling study. *Biochemistry*, **34**, 8812–8819, 1995.

    See also: H. J. Steinhoff, R. Mollaaghababa, C. Altenbach, K. Hideg, M. Krebs, H. G. Khorana, and W. L. Hubbell. Time-resolved detection of structural changes during the photocycle of spin-labeled bacteriorhodopsin. *Science*, **266**, 105–107, 1994.
61. J. L. Spudich. A chloride pump at atomic resolution. Perspectives: Structural Biology. *Science*, **288**, 1358–1359, 2000.

    Also: M. Kolbe, H. Besir, L.-O. Essen, D. Oesterhelt. Structure of the light-driven chloride pump halorhodopsin at 1.8 Å resolution. *Science*, **288**, 1390–1396, 2000.

    Also: H. Luecke, B. Schobert, H.T., Richter, J.P., Cartailler, K. Lanyi. Structure of bacteriorhodopsin at 1.5 Å resolution. *J Mol Biol*, **291**, 899, 1999.
62. K. Edman, P. Nollert, A. Royant, H. Beirhali, Pebay-Peyroula, J. Haydu, *et al.* High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature*, **401**, 822–826, 1999.
63. U. Haupt, J. Tittor, and D. Oesterhelt. Closing in on bacteriorhodopsin: Progress in understanding the molecule. *Annu Rev Biophys Biomolec Structure*, **28**, 367–399, 1999.

    See also: D. Oesterhelt. The structure and mechanism of the family of retinal proteins from halophilic archaea. *Curr. Opin. Struct. Biol*, **8**, 489, 1998.
64. N. Bennett and Y. Dupont. The G-protein of retinal rod outer segments (transducin). Mechanism of interaction with rhodopsin and nucleotides. *J Biol Chem*, **260** (7), 4156–4168, 1985.
65. F. Bornancin, C. Pfister, and M. Chabre. The transitory complex between photoexcited rhodopsin and transducin. Reciprocal interaction between the retinal site in rhodopsin and the nucleotide site in transducin. *Eur J Biochem*. **184**, (3), 687–698, 1989.
66. M. A. Geeves and P. B. Conibear. The role of three-state docking of myosin S1 with actin in force generation. *Biophys J*, **68** (4), 194S–199S, and 199S–201S, 1995.
67. C. A. Parent and P. N. Devreotis. A cell's sense of direction. *Science*, **284**, 765–769, 1999.
68. K. Palczewski, G. Rispoli, and P. B. Detwiler. The influence of arrestin (48 K protein) and rhodopsin kinase on visual transduction. *Neuron*, **8** (1), 117–126, 1992.
69. R. Gaudet, A. Bohm, and P. B. Sigler. Crytal structure at 2.4 Å resolution of the complex of transducin β,γ and its regulator, phosducin. *Cell*, **87**, 577–588, 1996.
70. K. Blüml. W. Schnepp, S. Schröder, M. Beyermann, M. Macias, H. Oschkinat, and M.J. Lohse. A small region in phosducin inhibits G protein function. *EMBO J*, **16** (16), 4908–4915, 1997.
71. P. H. Bauer, S. Müller, M. Puzicha, S. Pippig, B. Obermaier, E. J. M. Helmreich, and M. J. Lohse. Phosducin is a protein kinase A-regulated G-protein regulator. *Nature*, **358**, 73–76, 1992.
72. Y. Sugimoto, K. Yatsunami, M. Tsujimoto, H. G. Khorana, and A. Ichikawa. The aminoacid sequence of a glutamic acid-rich protein from bovine retina as deduced from the cDNA sequence. *Proc Natl Acad Sci, USA*, **88**, 3116–3119, 1991.
73. M. Illing, L. L. Molday, and R. S. Molday. The 220-kD rim protein of retinal rod outer segment is a member of the ABC transporter superfamily. *J Biol Chem*, **272**, 10303–10310, 1997.
74. H. G. Körschen, M. Beyermann, F. Müller, M. Heck, M. Vantler, K.-W. Koch, *et al.* Interaction of glutamic-acid-rich proteins with the cGMP signalling pathway in rod photoreceptors. *Nature*, **400**, 761–766, 1999.
75. J. Vinós, K. Jalink, R. W. Hardy, S. G. Britt, and C. S. Zuker. A G protein-coupled receptor phosphatase required for rhodopsin function. *Science*, **277**, 687–690, 1997.
76. T. Byk, M. Bar Yaacov, Y. N. Doza, B. Minke, and Z. Selinger. Regulatory arrestin cycle secures the fidelity and maintenance of the fly photoreceptor cell. *Proc Natl Acad Sci, USA*, **90** (5), 1907–1911, 1993.
77. J. Strotmann, I., Wanner, T. Helferich, A. Beck, and H. Breer. Rostro-caudal patterning of receptor-expressing olfactory neurones in the rat nasal cavity. *Cell Tissue Res*, **278**, 11–20, 1994.

78. K. Mori, H. Nagao, and Y. Yoshihara. The olfactory bulb: Coding and processing of odor molecule information. *Science*, **286**, 711–715, 1999.

79. *Science*, **286**, 22 October 1999 (entire volume).

80. M. A. Bianchet, G. Bains, P. Pelosi, J. Pevsner, S. H. Snyder, H. L. Monaco, and L. M. Amzel. The three-dimensional structure of bovine odorant binding protein and its mechanism of odor recognition [see comments]. *Nature Struct Biol*, **3** (11), 934–939, 1996.

See also: M. Tegoni, R. Ramoni, E. Bignetti, S. Spinelli, and C. Cambillau. Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nat Struct Biol*, **3**(10), 863–867, 1996.

Also: L. M. Amzel, *et al. Nature Struct Biol*, November issue, 1996.

81. L. Buck and R. Axel. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*, **65** (1), 175–187, 1991.

82. G. Glusman, S. Clifton, B. Roe, D. Lancet. Sequence analysis in the olfactory receptor gene cluster on human chromosome 17: recombinatorial events affecting receptor diversity. *Genomics*, **37** (2), 147–160, 1996.

See also: D. Lancet, N. Ben Arie, S. Cohen, U. Gat, R. Gross Isseroff, S. Horn Saban, *et al.* Olfactory receptors: transduction, diversity, human psychophysics and genome analysis. *Ciba Found Symp*, **179**, 131–141, 1993, discussion pp. 141–146.

83. Y. Pilpel and D. Lancet. Olfaction. Good reception in fruitfly antennae. *Nature*, **398**, 285–287, 1999.

84. P. Mombaerts. Seven-transmembrane proteins as odorant and chemosensory receptors. *Science*, **286**, 707–711, 1999.

85. A. J. Gentles and S. Karlin. Why are human G-protein-coupled receptors predominantly intronless? *Trends Genet*, **15**, 47–48, 1999.

86. R. R. Reed. Genetic approaches to mammalian olfaction. *Cold-Spring-Harb Symp Quant Biol*, **61**, 165–172, 1996.

87. I. C. Griff and R. R. Reed. The genetics of olfaction. *Curr Opin Genet Dev*, **5** (5), 657–661, 1995.

88. R. R. Reed. The molecular basis of sensitivity and specificity in olfaction. *Semin Cell Biol*, **5** (1), 33–38, 1994.

89. R. R. Reed. (The information reported here is a transcript of a lecture given by R. R. Reed on the mechanisms of sensitivity and specificity in olfaction. See abstracts of a symposium at the occasion of the the 40th anniversary of the Unilever research laboratory in Vlaardingen, The Netherlands, 1996.)

See also: R. R. Reed. Opening the window to odor space. *Science*, **279**, 193, 1998.

90. I. Boekhoff, J. Inglese, S. Schleicher, W. J. Koch, R. J. Lefkowitz, and H. Breer. Olfactory desensitization requires membrane targeting of receptor kinase mediated by beta gamma-subunits of heterotrimeric G proteins. *J Biol Chem*, **269** (1), 37–40, 1994.

91. H. Zhao, L. Ivic, J. M. Otaki, M. Hashimoto, K. Mikoshiba, and S. Firestein. Functional expression of a mammalian odorant receptor [see comments]. *Science*, **279** (5348), 237–242, 1998.

(The information reported here is a transcript of a lecture given by S. Firestein on cyclic nucleotide gated channels in olfactory transduction. See abstracts of a symposium at the occasion of the the 40th anniversary of the Unilever research laboratory in Vlaardingen, The Netherlands, 1996.)

92. G. Hellekant, V. Danilova, Y. Ninomiya. Primate sense of taste: behavioral and single chorda tympani and glossopharyngeal nerve fiber recordings in the rhesus monkey, *Macaca mulatta*. *J Neurophysiol*, 77 (2), 978–993, 1997.

(The information reported here is a transcript of a lecture given by G. Hellekant on the use of electrophysiological and behavorial techniques in animals to solve organoleptic questions in humans. See abstracts of a symposium at the occasion of the the 40th anniversary of the Unilever research laboratory in Vlaardingen, The Netherlands, 1996.)

93. The information reported here is a transcript of a lecture given by S. C. Chinaman on taste transduction: A diversity of mechanisms. See abstracts of a symposium on the occasion of the 40th anniversary of the Unilever research laboratory in Vlaardingen, The Netherlands, 1996.

94. M. A. Hoon, E. Adler, J. Lindemeier, J. F. Battey, N. J. Ryba, and C. S. Zuker. Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity. *Cell*, **96**, 541–545, 1999.

6

Signal transduction pathways controlling morphogenesis and haematopoiesis

Signals controlling morphogenesis: transforming growth factor β

The prominent role of the transforming growth factor β (TGF-β)/activin/ bone morphogenetic factor (BMP) family of growth factors in the control of development and morphogenesis of multicellular organisms, from *Drosophila* to humans, sets the superfamily of TGF-β factors apart from other growth factors. In humans, BMPs control bone formation and bone repair, and in *Drosophila* dorsal patterning and morphogenesis of eye and wings. In humans, activins regulate pituitary function and the secretion of follicle-stimulating hormone (FSH), and in frogs they induce the mesoderm. Thus the spectrum of cellular responses is large and diversified and so is the repertoire of genes encoding members of the transforming growth factor-β superfamily (Table 6.1)

In Fig. 6.1 a prominent feature of the three-dimensional structure of the bone morphogenetic protein-7 (also known as osteogenic protein-1, OP-1) is shown. The structure displays a cystine-knot motif, characteristic of proteins of the TGF-β superfamily (and other growth regulators).

BMP-2 has in common with BMP-7 and other members of the TGF-β family, a structural scaffold, consisting of a cystine-knot motif and two finger-like double-stranded β-sheets which determine the mode of dimerization. Secondary-structure differences between BMP-2 and BMP-7 account for the recognition and specific interaction with different binding partners. The crystal structure of human bone morphogenetic protein-2 is shown in Fig. 6.2.

The receptors for TGF-β-like growth factors

TGF-β proteins bind first to a type II receptor. The receptor–ligand complex then combines with a type I receptor, forming an oligomeric signalling receptor complex. Both type I and type II receptors have cytoplasmic serine/threonine kinase domains. TGF-β signalling is transmitted by the SMADs, a family of signal transducers and transcriptional activators (Fig. 6.3).

Table 6.1 Members of the TGF-β superfamily (reproduced with permission of Professor J. Massagué and Annu Rev. Biochem. from Table 1 of ref. 1)

Factors	Cellular responses
TGF-β	Cell-cycle arrest in epithelial and haematopoietic cells. Control of growth and differentiation of the mesenchyme. Wound healing. Formation of extracellular matrix. Immunosuppression
Activins	Stimulation of production of pituitary follicle-stimulating hormone (FSH). Differentiation of haematopoietic cells
Bone morphogenetic factors (DPP in *Drosophila*)	Regulation of development Gastrulation Neurogenesis Chondrogenesis Development of eyes and wings and dorsalization in *Drosophila*
BMP 3/Osteogenin	Bone differentiation and bone formation
Growth and differentiation factors	
GDFs	Chondrogenesis
Nodal	Mesoderm induction in frog and fish
Dorsalin	Cell differentiation in the neural system
Distantly related members of the TGFβ/Activin/BMP family	
MIS (Müllerian inhibiting substance) or AMH (anti-Müllerian hormone)	Regression of the Müllerian duct
Inhibin α	Inhibition of FSH production
GDNF, glial-cell-line-derived neurotrophic factor	Survival of neurons; kidney development
CDMP, cartilage-derived morphogenetic protein	Cartilage development

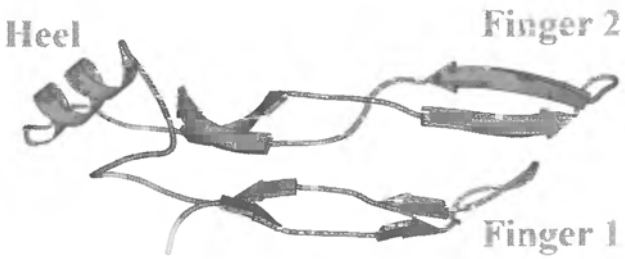


Fig. 6.1 Ribbon model of bone morphogenetic protein-7. From the cystine knot emerge four antiparallel β-sheets, which form the finger 1 and 2 projections. An α-helix on the left-hand side of the knot lies perpendicular to the axis of the two fingers and forms the heel of the hand. (The cystine knot forming the core of the monomer consists of three disulphide bonds. Two disulphides form a ring through which the third disulphide passes.) (Reproduced with permission of the authors of ref. 2, and the Proc. Natl. Acad. Sci: USA, from data filed with the Brookhaven protein databanks.)

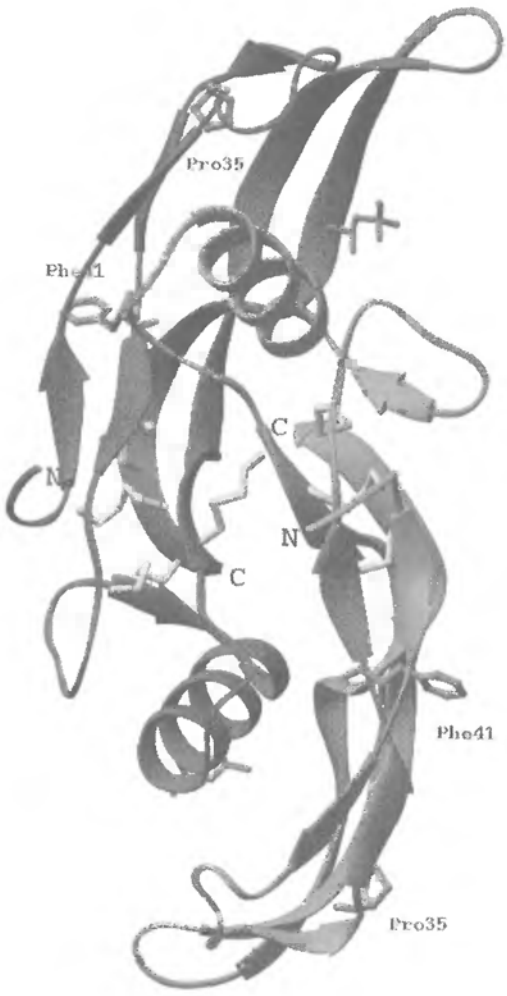


Fig. 6.2 A stereoview of the topology of the native BMP-2 dimer. α-Helices are indicated as spirals, disulphide bonds as sticks. The interactions which are responsible for dimerization occur between the helix α-3 and the β-strands on the other subunit. A unique β-strand, β-5a is located near the helix α-3. Differences in the surface charge of the region adjacent to the heel helix, which interacts with the receptor, are thought to determine the fit between the factor and the matching receptor, and are responsible for functional differences between these factors (Phe41 and Pro35 are involved in the recognition of the receptor by the ligand). (Reproduced by courtesy of Professor Walter Sebald, Department of Physiological Chemistry, University of Würzburg and by permission of the Journal Mol. Biol. (ref: 3). Coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank (PDB code 2BMP). (ref. 3).

TGF-β receptor activation is a good example of the mutual dependency of dimerization and transphosphorylation. Type I receptors have a highly conserved, 30 amino-acid-long region which precedes the kinase domain. This region is the GS (glycine–serine) domain, because it contains the sequence, GSGS. On interaction with the liganded type II receptor, the GSGS sequence in receptor I is *trans*-phosphorylated by the type II receptor. Phosphorylation of Ser165 in the GSGS sequence is crucial and determines the intensity of signalling.

Although both type I and II receptors have a kinase domain, each has a different function. Whereas the type I receptor kinase phosphorylates the substrates, the SMADs, the type II receptor kinase autophosphorylates itself and the type I receptor. Only the type I receptor in the I/II heterodimer interacts with R-SMADs. No TGF-β responses have been found in cells lacking type I receptors, although the type II receptors should be able to signal, because they have a functional kinase, leaving open the possibility that type II

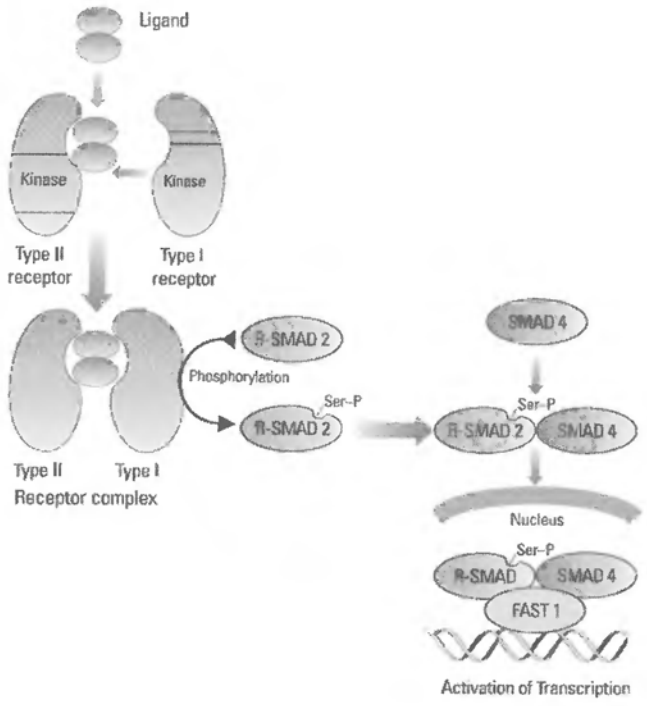


Fig. 6.3 Binding of a member of the TGF-β family to receptor type II recruits receptor type I. The signalling receptor complex phosphorylates a receptor-regulated SMAD (R-SMAD 2). The activated R-SMAD forms a complex with the common Co-SMAD (SMAD 4), and the SMAD 2–SMAD 4 complex moves to the nucleus, where it combines with the DNA-binding protein Fast-1,[4] and activates transcription.

receptors may have as yet unknown substrates. The actual signalling complex is a heterotetramer composed of two molecules of TGF-β receptor I and two molecules of receptor II (for further information, see refs 1,5,6).

TGF-β signalling

The SMADs[7]

The SMADs are like the STATs in cytokine signalling signal transducers and transcriptional activators (see later in this chapter). The first SMAD homologue identified was the product of the *Drosophila* gene, *Mad* (*mothers against dpp*; Dpp is the *Drosophila* homologue of BMP-2). Subsequently, *Mad* homologues were identified in *Caenorhabditis elegans* and named '*sma*s', *small-sized,* because mutations of the *sma* genes lead to small-sized worms. That is why the vertebrate homologues were named SMAD (*SMA-MAD*-related). The identification of the SMAD family of signal transducers has provided a clue to the role of the TGF-β/activin/BMP family in regulating the formation of organs and differentiated tissues in the body.[8]

SMADs are cytoplasmic proteins with molecular weights ranging from 42 to 60 kDa. Pathway-restricted SMADs and common or Co-SMADs are distinguished. The pathway-restricted, or receptor-regulated SMADs (R-SMADs), are phosphorylated on serines in the MH2 domain by TGF-β receptor I. The phosphorylated, receptor-regulated SMADs associate with a constitutive, common Co-SMAD (SMAD 4), and the heterodimeric SMAD complex is translocated to the nucleus where it binds either directly or with the help of other transcriptional factors to enhancers and promoters of specific gene elements (Fig. 6.4) (see also Chapter 9).

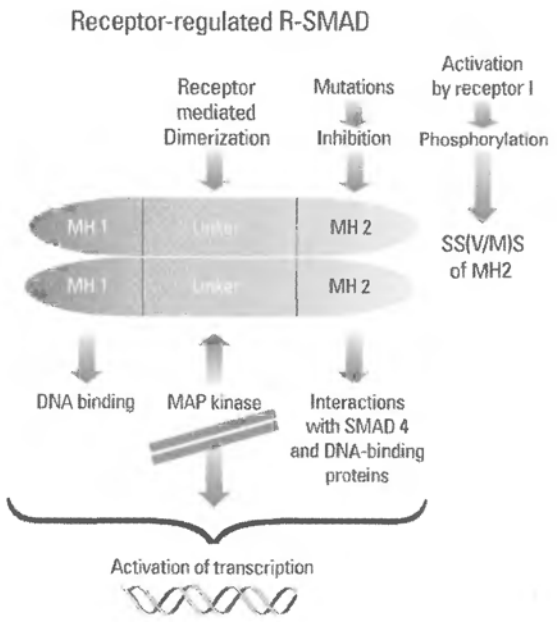


Fig. 6.4 Receptor-regulated SMADs (R-SMADs) have two homology regions, one at the amino- and one at the carboxy-termini, MH1 and MH2. They are linked by a proline-rich linker region. The linker region is highly variable in size and sequence. It participates in the formation of homo-oligomeric structures and contains phosphorylation sites for the MAP kinase. Phosphorylation by the MAP kinase has a negative effect, because it prevents nuclear translocation of the SMADs. R-SMADs interact through the MH2 domain with the activated receptor I and are phosphorylated at the C-terminal SS(V/M)S motif, where S is serine; V, valine; and M, methionine. Both, MH1 and MH2 domains are essential for DNA binding and the recruitment of DNA-binding proteins.

Regulation of growth

Growth inhibition by TGF-β is a consequence of cell-cycle arrest. TGF-β/SMAD upregulate an inhibitor of a cyclin-dependent kinase, p15 INK,[9] and block expression of the Cdc25A phosphatase (Chapter 12).[10] SMADs have this function in common with tumour suppressors (see Part 4).[11] Both SMADs and tumour suppressors control growth and proliferation of cells. But, TGF-β and related factors, such as BMP-4, also send signals promoting cell-cycle progression. This signal pathway involves activation of a member of the MAPKK family of kinases, TAK-1 (TGF-activated kinase-1).[12,13]

Summary

Figure 6.5 summarizes signalling by TGF-β and related factors.

There are several points of control. The first is at the level of the ligand–receptor interaction: TGF-βs 1–3 bind to membrane-attached proteo-β-glycans ('betaglycans').[14] These proteoglycans have no signalling function, but assist in binding TGF-β to the receptor with high affinity. Another membrane-bound glycoprotein which is engaged in TGF-β binding is 'endoglin'. It is expressed in high concentrations in endothelial cells. Mutations of the endoglin gene are associated with a hereditary human disease, haemorrhagic telangiectasia.

Several proteins have been identified which interfere with binding of TGF-β to the receptor. For example, LAP (latency-associated peptide) is a negative regulator which keeps TGF-β in a latent, inactive form and prevents binding to the receptor. Attached to LAP by disulphide bonds is a large glycoprotein, LTBP (latent TGF-β binding protein). LTBP is believed to play a role in anchoring TGF-β to the extracellular matrix. Mutations of LTBP1 are found in arachnodactyly, Marfan's disease. Then there are the inhibins, distantly related to the TGF-β family of proteins. They form heterodimers with activin and interfere with activin receptor binding and activin signalling. Another protein that

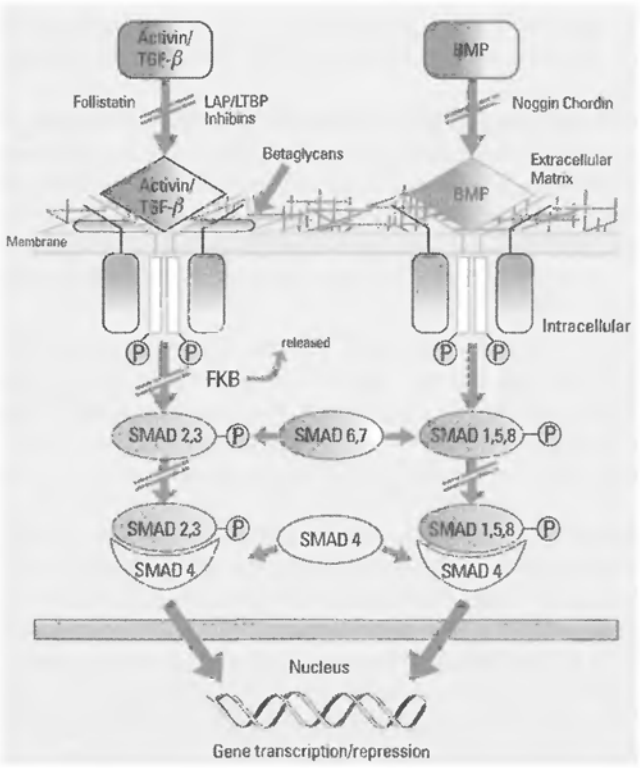


Fig. 6.5 Signalling through SMAD proteins. Receptor complexes, formed with transforming growth factor-β (TGF-β) and related factors, activins, and bone morphogenetic proteins (BMPs), consist of two serine/threonine receptor kinases, the type II and type I receptors. Binding of the ligand results in autophosphorylation of type I receptors .The activated type I receptors in turn phosphorylate and activate R-SMADs: SMADs 1, 5, and 8 are activated by BMP receptors and SMADs 2 and 3 are phosphorylated by activin and TGF-β receptors. The SMADs then dissociate from the receptors and form complexes with SMAD 4. The SMAD complexes move into the nucleus, where they may activate or repress transcription of genes. SMADs 6 and 7 inhibit signalling, probably by competing with R-SMADs for receptor binding and/or by interfering with association of R-SMADs with the common SMAD 4.[7] Phosphorylation is represented by a circled P. (This model is based on Fig. 6 in ref. 1 and is reproduced with the permission of the author and Annu Rev. Biochem.)

interferes with TGF-β–receptor interactions is follistatin, which inhibits the production of FSH (follicle-stimulating hormone) in the pituitary. Finally, the peptides noggin[15] and chordin,[16] which belong to a family of polypeptides that bind specifically to the bone morphogenetic protein, BMP-4, preventing its interaction with the receptor. These peptides control BMP-4, which determines whether adjacent cells form neural rather than epidermal, or dorsal rather than ventral mesoderm.

The TGF-β receptor itself is also a point of control. Receptor I has a GSGSGLP (Gly/Ser/Gly/Ser/Gly/Leu/Pro) motif to which the immunosuppressor-binding protein, FKBP12 binds. When the receptor II/ receptor I complex is activated by ligand and phosphorylated, the binding protein is released. This is probably a way to prevent spontaneous autoactivation of the receptor in the absence of ligand.[17] The next checkpoint is at the level of the receptor-regulated R-SMADs. They are controlled by inhibitory SMADs 6 and 7, which inhibit both TGF-β/activin and BMP signalling. TGF-β can regulate the intensity of its signal by turning on expression of genes coding for inhibitory SMADs.

Thus, each step from the ligand–receptor interaction to SMAD activation, including nuclear transport of SMADs and finally gene activation, is a point of regulatory control (see Chapter 10).

TGF-β controls development and morphogenesis of multicellular, metazoan organisms

Our understanding of the role of TGF-β in the regulation of morphogenesis in multicellular organisms has greatly improved through studies of the morphogenesis of *Drosophila* which is presently the best-suited object for the study of embryonic development.

A multicellular, metazoan organism, consisting of billions of differentiated cells which are arranged in a precise pattern, emerges from a single, fertilized egg in a time span ranging from a few days to weeks and months. Cells in an early embryo are initially alike and totipotent and then the stem cells differentiate and develop specific functions. In the human body, there are only about 200 different cells. When differentiation is completed, the cells may grow but the organs they form preserve their characteristic organizational pattern.

A key morphogen for the development of *Drosophila* is DPP, encoded by the *d*eca*p*enta-*p*legic (*dpp*) gene. In *Drosophila,* DPP interacts with a TGF-β receptor homologue and elicits various morphogenetic responses. The signal transducer of DPP is the Mad protein, a homologue of SMAD. Thus, TGF-β signalling pathways are evolutionary conserved in vertebrates and invertebrates, in all metazoan, multicellular organisms.

Control of morphogenesis is dependent on the concentration of the signal. Accordingly, expression of DPP/MAD is gene-dosage dependent. In *Drosophila,* an individual cell is able to sense different concentrations of DPP. The selective recognition of concentration gradients of regulatory molecules, morphogens, is a fundamental aspect of morphogenesis, because pattern formation in the early embryo requires distinct, spatially differently distributed positional signals. Such gradients are already set up in the early embryo. The spatiotemporal pattern of expression of these genes leads to a localized accumulation of mRNAs from which transcriptional gene regulator proteins, such as the SMADs, are translated. Cells in the early 'anlagen' of a vertebrate embryo already have positional information that determines their fate, long before the final differentiation programme is called up. This information determines, for example, the asymmetric position of visceral organs such as the heart. TGF-β-related factors, such as nodal[18,19] and lefty,[20,21] are expressed and distributed asymmetrically and determine the left–right asymmetric positions of organs. Malfunction of the R-SMAD 2 regulator of gene transcription is correlated with abnormalities in the left–right asymmetric patterns of organs. Nomura and Li,[22] Waldrip *et al.*,[23] and Sirard *et al.*[24] have inactivated the genes for SMAD 2 and SMAD 4. The phenotype of SMAD 2-deficient mice was more severe than the phenotype of mice where one or the other TGF-β-related factor was missing. This means that several TGF-β-related factors apparently signal through the same SMAD 2. And indeed, SMAD 2 is activated by two activins and three different TGF-βs. Thus, SMAD 2 has a central position, where several TGF-β/activin-related signals come together. This explains the severity of the SMAD 2-deficient phenotype. Equally important is the common SMAD 4. SMAD 4-deficient embryos have a defect in early development, e.g. in gastrulation. (a very early developmental event in which the three primary layers of the animal body—ectoderm or outer layer, endoderm or inner layer, and mesoderm—are selected and positioned for further development)[25] (see the discussion of Rik Derynk[25]).

The phenotypes of mutations of components of the TGF-β/SMAD signalling pathway are summarized in Table 6.2.

Table 6.2 Consequences of mutations of genes encoding components of the TGF-β/SMAD signalling pathway (reproduced with permission of the author and Annu Rev. Biochem. from Fig. 7 in ref. 1)

Mutations	Phenotype
Ligands	
MIS/AMH	Persistent Müllerian duct syndrome (PMDS)
CDMP (cartilage-derived morphogenetic protein)	Hereditary chondrodysplasia (HC)
Receptors	
TGF-β receptor II	Gastrointestinal cancer
Receptor for MIS/AMH	PMDS
ACK 1 (activin-receptor-like kinase)	Hereditary haemorrhagic telangiectasia (HHT)
Co-receptors	
Endoglin	Hereditary haemorrhagic telangiectasia (HHT)
SMADs	
SMAD 2	Colon cancer
SMAD 4	Colon, pancreatic, and other cancers

The same repertoire of growth factors and cytokines that regulate development and morphogenesis in the embryo, also regulate growth and proliferation in the adult organism. Growth stops when mammals and humans reach a certain size. But even when growth of the body and of the organs stops, new cells are formed continuously in most of our tissues (probably with the exception of the nervous system and muscles). TGF-β and related factors play a central role also in the control of proliferation of cells in the adult organism. Loss of TGF-β-mediated growth inhibition is often associated with cancer,[27] and deregulation of TGF-β signalling causes diseases that manifest themselves in abnormal inflammatory responses,[28] and accumulation of interstitial matrix material in the lung, kidney, liver, and other organs.[29]

Earlier work tried to distinguish between factors that stimulate and factors that inhibit growth, on the basis of the biochemical properties of their receptors. Stimulation has been attributed to growth factors that signal through receptors with tyrosine kinase activity, and inhibition to the TGF-β family of factors which signal through receptors with Ser/Thr kinase activity. But this classification lost its meaning as the many and different functions of TGF-β became better understood. As we have seen, growth inhibition is actually only one of the many manifestations of TGF-β as regulator of the development of the embryo and its role in the preservation of the differentiated state and the maintenance of morphology in the adult.

The much too restricted discussion of the morphogenetic role of TGF-β reveals the limited scope of this book. Because this book deals primarily with the instruments, the hardware, of cellular signalling, the biology of embryonic development and the progress made in this fascinating field is outside its scope. Fortunately, detailed reviews[29] and excellent textbooks, such as the *Molecular Biology of the Cell*,[30] are available and should be consulted.

Summary

Each class of growth factors and cytokines interacts with a special class of receptors and recruits its own set of transducers and transcription factors. Moreover, TGF-β and

related factors employ a number of co-receptors, co-factors, and inhibitors. This partly accounts for the complexity of cellular signalling. However, such complexity is realistic, considering the intricate role of TGF-β in the implementation and maintenance of the morphogenetic programme in the embryo and in the adult organism.

The diversity of the effects of TGF-β is only matched by the effects of cytokines. Actually, the cytosolic signal transducers and activators of transcription—the SMADs, employed by TGF-β, and the STATs, recruited by cytokines—have much in common, as we shall see below.

Signals governing the differentiation of haematopoietic cell lines: the cytokines

Cytokines control the development of haematopoietic cell lines. They regulate the immune response and are involved in the development of immune-competent cells. They play a role in embryogenesis and act as hormones, like growth hormone, and a cytokine subgroup, the interferons, has antiviral activities. Cytokines bind to transmembrane proteins that lack kinase activity.[31] Thus, in order to signal they must recruit cytosolic tyrosine kinases, mostly the JAKs (Janus kinases). For transmission of signals from the cell surface to the gene, cytokines use the signal-transducing transcriptional activators (STATs). The JAK/STAT pathway was discovered in studies of interferon signalling, but it soon became clear that it is used by nearly 40 other cytokines, giving rise to a variety of cellular responses.

Cytokine receptors

Ligand-induced receptor dimerization is also the first step in cytokine signalling, followed by tyrosine phosphorylation of the receptor by a recruited cytosolic kinase. Principles and structural details of receptor dimerization were discussed in the case of a type I cytokine receptor, the growth hormone (GH) receptor, in Chapter 2.

Structure of cytokine receptors

Based on structural properties, cytokine receptors are divided into four groups (Fig. 6.6).[32,33] The largest group I comprises receptors for interleukins and all the colony-stimulating factors (CSFs) which do not signal through RTKs. The receptors for the ciliary neurotrophic factor (CNTF), the leukaemia inhibitory factor (LIF), and the receptors for erythropoietin (EPO), thrombopoietin, prolactin (PRL), and growth hormone (GH) also belong to group I. All group I cytokine receptors share with the hGH receptor a similar structure. Group II includes receptors for interferons (IFN-α/β and -γ). Group III is functionally heterogeneous. Included in this group are all receptors without tyrosine kinase activity, such as the receptors I and II for TNF-α and those TNF-related NGF receptors which lack tyrosine kinase activity (see Chapter 1). We include the integrins in this group, because they are (like the cytokine receptors) membrane-spanning proteins that lack kinase activity (see Chapter 1). Included in group III are also the apoptosis receptors, CD40R and CD95R/FasR (see Chapter 13). Finally, class IV is reserved exclusively for the type I and II IL-1 receptors.

The crystal structure of human interleukin-4 (IL-4) bound to the ectodomain of the α-chain of the IL-4 receptor has been solved.[34] The IL-4 receptor is a type I receptor. The ligand, IL-4, binds first with high affinity to the receptor α-chain and then the inter-

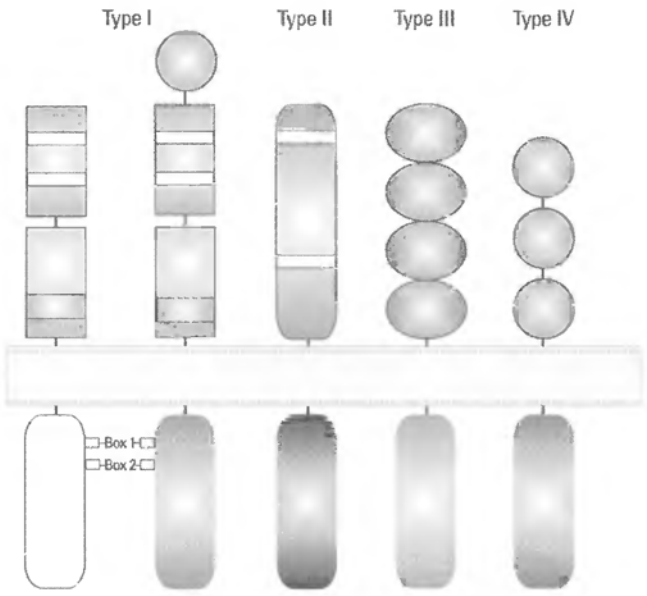


Fig. 6.6 Cytokine receptors have on their extracellular side a variety of different structural domains. Type I receptors have fibronectin-like domains, type III receptors have cysteine-rich domains, and type IV receptors have immunoglobulin-like regions. Type I receptors have intracellular, membrane-proximal Box 1 and Box 2 regions, which are docking sites for the JAK2 tyrosine kinase with which type I receptors primarily associate. (Reproduced from Fig. 1A. p. 252 of ref. 33, with permission of Professor Taniguchi and Science.)

mediate complex recruits the common γ_c-chain. The structure of IL-4 bound to its receptor is a paradigmatic example of the binding interaction of a family of short-chain, four helix-bundle cytokines (IL-2, IL-7, IL-9, and IL-15), which all recruit on binding a common γ-receptor chain. The structure of the complex is of interest, because the spatial orientation of IL-4 and the corresponding receptor-BP (binding protein) to which it binds, and the structural properties of the contact region, are different from the binding site of the human growth hormone (hGH) for its binding protein (see Chapter 2). Thus, several structural possibilities exist for cytokine–receptor interactions.

IL-4 plays a key role in the response of a T-helper cell (TH-2), which activates IgE-producing B cells, eosinophils, and mast cells. This IL-4 response is of great medical interest, because it is associated with allergic, inflammatory reactions, with asthma, rhinitis, conjunctivitis, and dermatitis. Therefore, the structure of the IL-4–receptor complex may help in the development of a more effective therapy of allergies (Fig. 6.7).

The group II receptor complexes of interferon-γ and IFN-α/β receptors will be presented as examples of cytokine–receptor complexes. They are expressed in nearly all cells. A three-dimensional structure of an interferon-γ complex has been solved,[36] it has two subunits, IFNγR1, (IFNGR1), and IFNγR2, (IFNGR2) (Fig. 6.8).

There are at least 12 IFN-αs, but only one IFN-β in humans. The interferon α/β receptors have, like the IFNγRs, two subunits: IFNAR1 and IFNAR2 (Fig. 6.9).

Signalling through receptor JAKs and STATs[37]

Cytokine receptors sometimes need auxiliary factors and co-receptors. An example is the LIF-receptor-α (LIFR-α), a type I cytokine receptor which requires LIFR-β as co-receptor for signalling.

Although more factors modulating cytokine signalling will certainly be found, in principle, the mode of cytokine signalling, like TGF-β signalling, is quite simple, since

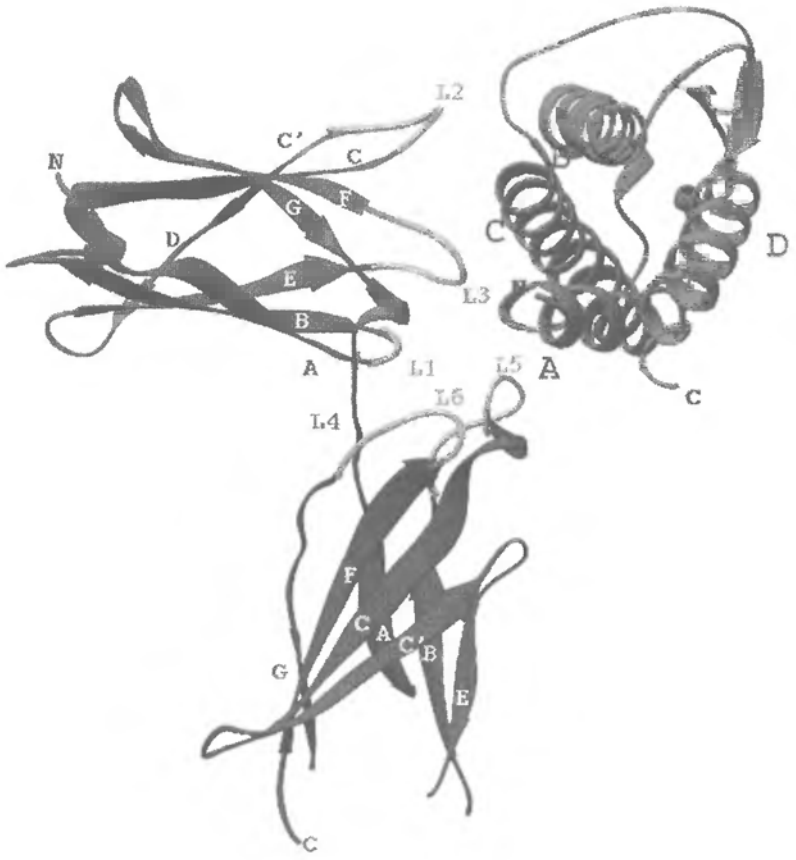


Fig. 6.7 The crystal structure of human interleukin-4 bound to the ectodomain of the α-chain of the IL-4 receptor (IL-4 BP), at 2.3 Å resolution. This is a stereo-ribbon view. The helical IL-4 is on the right and the IL-4-binding protein is on the left. The amino- and carboxy-termini are labelled N and C, respectively. The four loops of the IL-4 BP that interact with IL-4 are labelled L1, L2, L3, and L5, according to de Vos *et al.*[35] The loops are ordered in a stack-like fashion, comprising four levels, from top to bottom: L2, L3, L1, and L5 and L6. L4 connects the two domains of the receptor. It does not interact with IL-4. This kind of receptor–ligand interaction may be characteristic for all ligands of this type, which first bind with high affinity to an IL-binding protein and then recruit the common receptor γ-chain (γ_c). To this family of ligands belong IL-2, IL-4, IL-7, IL-9, and IL-15. (Reproduced by courtesy of Professor Walter Sebald, Department of Physiological Chemistry, The University of Würzburg and Cell, from ref. 34.)

the signal transmission chain consists of only three major components: the ligand–receptor complex (with associated protein kinases and protein phosphatases), and the signal transducing transcriptional activators, the STATs or SMADs (see Chapter 10).

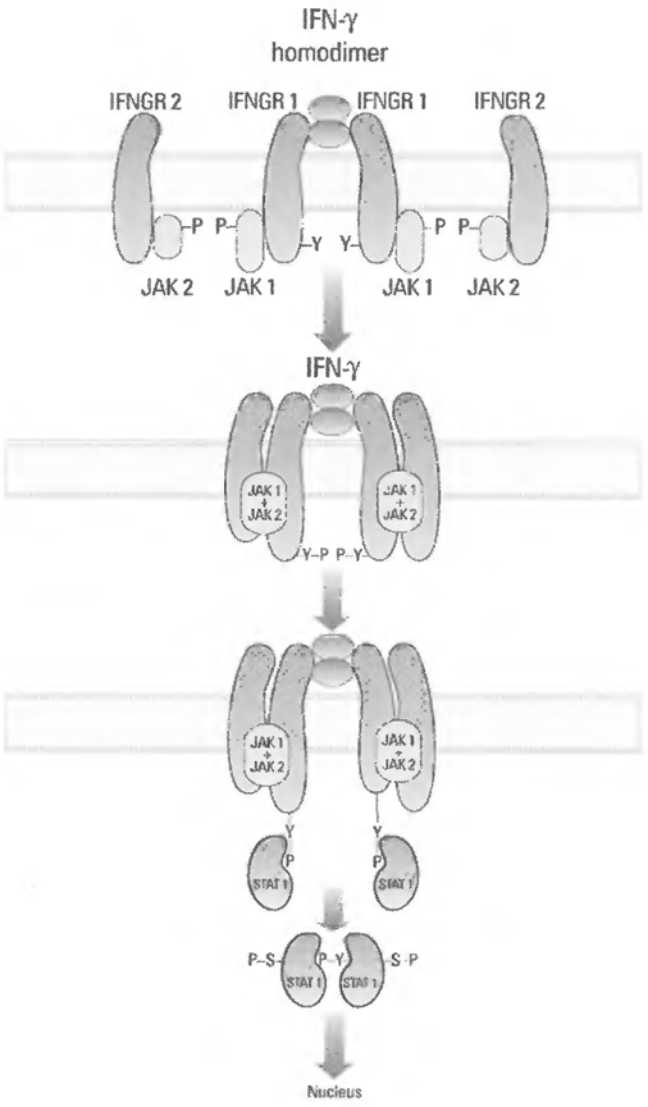


Fig. 6.8 The IFNGR1binds the tyrosine kinase JAK1. JAK1 is then phosphorylated and in turn phosphorylates a tyrosine in the cytoplasmic domain of IFNGR1, creating a docking site for STAT 1. IFNGR2 then recruits JAK2. JAK2 is also phosphorylated. The next step is the assembly of the signalling receptor complex from IFNGR1 and IFNGR2. The binding site for interferon-γ is on the IFNGR1. Two IFNGR1 receptor subunits associate with an interferon-γ homodimer. Two IFNGR2 subunits are then added, forming a symmetrical complex, consisting of two IFNGR1 and two IFNGR2 subunits. The receptor/JAKs activate the STATs by phosphorylation, forming STAT 1 homodimers, which are released from the receptor complex and move to the nucleus.[37]

Mutant cells lacking JAKs are unresponsive to cytokines, and restitution with JAKs restores the response. That JAKs are essential was also apparent from the phenotype of homozygous knock-out mice, in which the *jak3* gene was disrupted by gene targeting. These mice had drastically reduced numbers of lymphocytes and functionally defective B and T cells.[39] X-chromosome-linked, Severe Combined ImmunoDeficiency in humans (SCID) is caused by mutations of the IL-2 receptor, which prevent association with JAK3.[40] This demonstrates the importance of cytokine signalling for the differentiation of immune-competent cell lines.

JAKs have molecular weights ranging from 120 to 130 kDa. The sequence identity is 35–45%. They have several homology regions, JH1–JH7, including a carboxy-terminal kinase domain, JH1, but they have no SH2 and SH3 domains, instead they use their

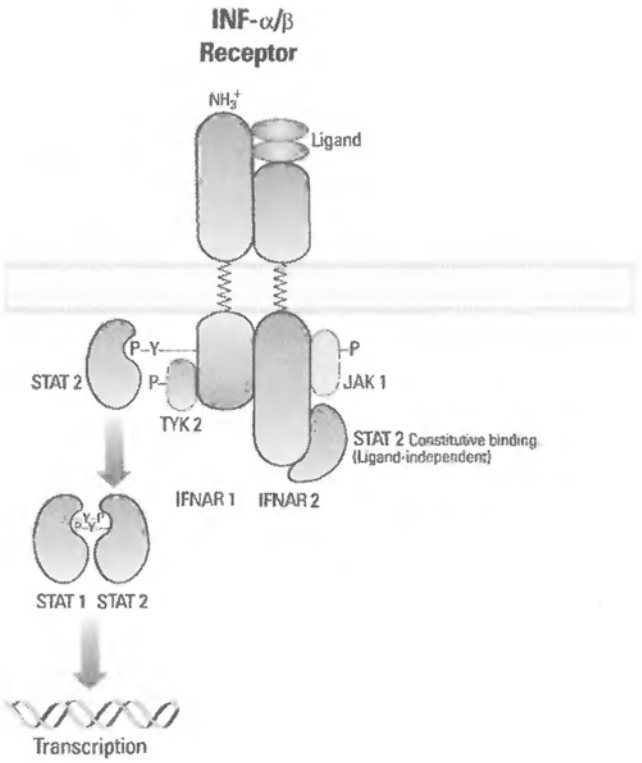


Fig. 6.9 The unliganded IFNAR1 binds constitutively the cytosolic tyrosine kinase TYK2, whereas the IFNAR2 binds constitutively STAT 2. On binding the ligand on IFNAR2 and activation, IFNAR1 and IFNAR2 associate. IFNAR2 recruits and activates JAK1 and IFNAR1 activates TYK 2, creating a docking site for another STAT 2. STAT 2, attached to IFNAR1 is now phosphorylated, creating a docking site for a STAT 1 molecule, which in turn is also phosphorylated and forms a STAT 1/STAT 2 heterodimer. Binding of the STATs is sequential. Only when STAT 2 is phosphorylated, can STAT 1 be bound. Finally, The STAT 1–STAT 2 heterodimer is released and goes to the nucleus, where it functions as transcription factor.[38]

homology regions as recognition sites for coupling proteins. Both constitutive and ligand-dependent association of JAK and receptor is possible. Receptor-bound JAKs are autophosphorylated in the KEYY (lysine, glutamic acid, tyrosine site) of their kinase activation loop. Phosphorylation at this site activates the JAK, which then phosphorylates the receptor, creating docking sites for the STATs. The presence of two SH2 domains in each STAT molecule provides the necessary binding energy for the dimerization of the phosphorylated STATs, and the cooperative interaction between the two STATs overcomes their strong association with the receptor complex, eventually leading to the release of the active, phosphorylated STAT dimer. How the STATs are directed to the nuclei in IFN-treated cells is not clear. In the nucleus, STATs, together with other DNA-binding proteins, regulate gene transcription (see Chapter 10).

STATs are phosphorylated in response to IFNs not only on tyrosines but also on serine/threonines. Serine phosphorylation is carried out by a dual-specificity kinase (recall that SMADs, the transcriptional activators that transmit TGF-β signals are serine phosphorylated; see above). Since tyrosine and serine/threonine phosphorylation sites are different, phosphorylation of STATs on tyrosines and serines would be expected to diversify and enlarge the repertoire of STAT/DNA interactions. Although JAKs are engaged primarily with cytokine signalling, in some cases JAK/STATs accept signals from other receptors, such as growth factor receptor tyrosine kinases and G-protein-linked heptahelical receptors.[41] For example, the hormone angiotensin II, which signals through a G-protein-coupled heptahelical receptor, has been shown to activate transcription through

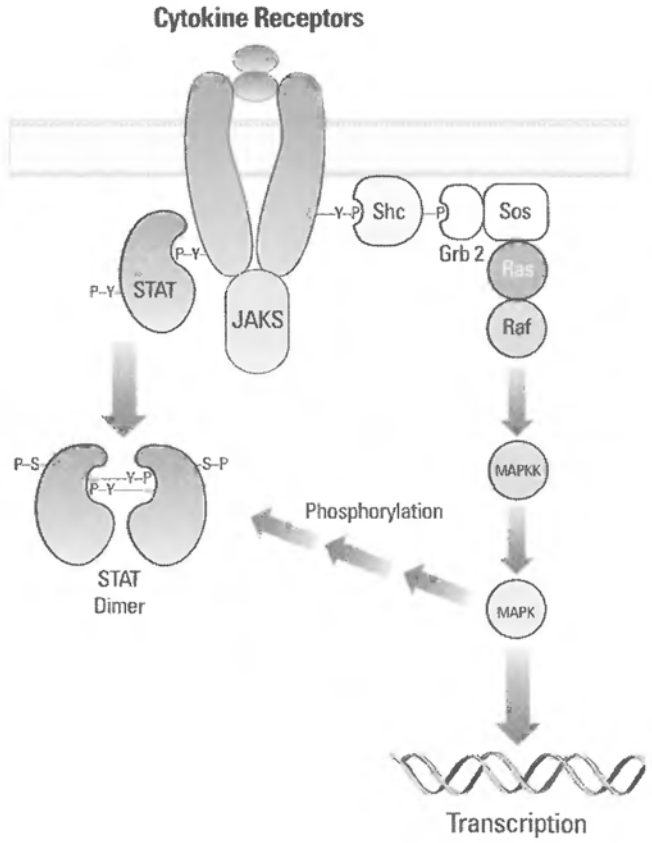


Fig. 6.10 A possible route of communication between cytokine receptors and the RAS/MAP kinase pathway. The role of the MAP kinase in the phosphorylation of STAT is

the JAK/STAT pathway,[42] and IFN-α/β and other cytokines activate the insulin receptor substrate (IRS-1) through the JNK cascade.[43] Figure 6.10 outlines how recruitment of linkers by cytokine receptors activates the Ras/MAP kinase cascade and how the MAP kinase, through phosphorylation of STATs, can open communication between growth factor/RTKs and cytokine signalling.

Control of cytokine signalling

Activation of STATs is transient, like that of other transcriptional regulators. When cytokine signalling is terminated, the concentrations of active STATs in the nucleus decline, because STATs are either inactivated by dephosphorylation or removed by proteolysis, or both. Accordingly, the half-life of STAT 1 can be prolonged by preventing dephosphorylation, through inhibition of phosphotyrosine phosphatases. Finally, there are inhibitors of JAK/STAT signalling that are induced by STATs.[44]

Phosphotyrosine phosphatases dephosphorylate JAKs and silence receptor–JAK complexes. They also inactivate JAKs which have been released from the receptor complex and are out of control. But dephosphorylation of tyrosines may also have positive effects and enhance cytokine signalling. A related phosphotyrosine phosphatase, actually stimulates IL-6 signalling, suggesting that the phosphotyrosines in the IL-6 receptor–JAK complex, which are attacked by the phosphatase, are inhibitory. The phosphatase activi-

ty itself is regulated by phosphorylation on tyrosines, but it is not clear whether a JAK or another cytosolic tyrosine kinase does that.

The biological relevance of negative control of cytokine signalling by dephosphorylation becomes apparent in the phenotype of the mouse mutant, 'motheaten', which lacks a functional phosphotyrosine phosphatase. This mouse suffers from numerous haematopoietic defects and uncontrolled cell proliferation.

However, dephosphorylation is not the only way to downregulate cytokine signalling. IL-6 signalling is also downregulated by regulators of cytokine signalling, the SOCS, suppressor of cytokine signalling, and CIS, cytokine-inducible inhibitors of signalling.[45] These proteins dampen cytokine signalling by binding through a SH2 domain to the tyrosine phosphorylated receptor–JAK complexes. One of these inhibitor proteins is JAB, a *JA*K-binding protein that recognizes JAK1, 2, and 3.[46] A related inhibitory protein (SS-1) is induced in cells stimulated by several cytokines, such as IL-4 and IL-6, the leukaemia-inhibitory factor (LIF), and the granulocyte-colony-stimulating factor (G-CSF). Overexpression of SS-1 inhibits JAK2 and TYK2 (another JAK-like tyrosine kinase) and interrupts cytokine signalling.

How selective is cellular control by cytokines?

When any particular cytokine-receptor–JAK complex would activate only one STAT, specificity would be guaranteed. However, some STATs, like the SMADs in the case of TGF-β signalling, are activated by more than one cytokine, also each cytokine evokes a specific and unique biological response. This raises the question of how the apparently promiscuous activation of STATs can be reconciled with a specific biological response to each cytokine. Several possibilities may be considered. We can forget JAKs, because they do not determine specificity of cytokine signalling. This leaves us with four possibilities:

1. Not all receptors may be expressed in the same cell. A case in question is the restricted, cell-specific expression of receptors for IL-2, erythropoietin, thrombopoietin, prolactin, and growth hormone, all of which activate the same STAT.
2. But even when different cytokines activate the same STAT in the same cell, the activation may be quantitatively different and not all at the same time.
3. Also, post-transcriptional and post-translational processing may give rise to STAT variants, transcribed from the same *stat* genes. And indeed, at least a dozen different STATs are transcribed from only seven *stat* genes. Each of these STAT variants could be addressed to a different gene locus. And, last but not least, more STATs can be expected to be discovered.
4. There may be a host of secondary transcriptional regulators, such as the IRF's, the interferon regulatory factors that could specify interferon signalling.

Biological role of cytokines

Embryonic development

STATs play a role in embryonic development, like SMADs. In *Drosophila* embryos, STATs seem to have a role in the expression of early genes, like STAT 2 and STAT 3 in mammalian embryos. Defects in a STAT homologue are lethal in the larvae of the fly, and the same is true for mammalian embryos: embryos of *stat3* knock-out mice die before they ever form mesoderm, and *stat2* knock-outs do not reach term.

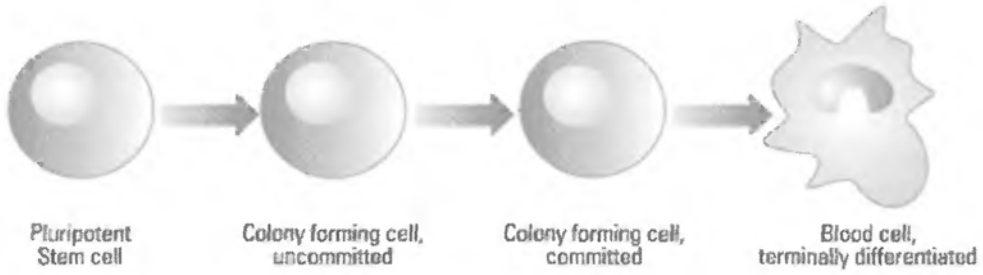


Fig. 6.11 The way is shown from a pluripotent stem cell to a multipotent, uncommitted colony-forming cell, to a committed colony-forming cell (CFC) with an already restricted repertoire, and finally to a terminally differentiated blood cell. (For more information, see also ref. 30.)

In the slime mould, *Dictyostelium discoideum,* a STAT-like DNA-binding protein, DIF (the differentiation-inducing factor), is expressed. DIF is about 700 amino-acid residues long, and has substantial homology with mammalian STATs. DIF is, like STAT, activated by phosphorylation of tyrosines. It is essential for differentiation of the stalk cells of *D. discoideum*.[47] The long evolutionary path that the STATs have travelled from *Dictyostelium* to mammals underlines their conserved and essential role as signal transmitters and transcriptional activators in early development.

Differentiation of cell lineages

Cytokines have a central role in the differentiation of haematopoietic cell lines (the role of interleukins in the differentiation of lymphocytes will be discussed in more detail in Chapter 14). Haematopoiesis is the process by which terminally differentiated cells—red, white, and lymphoid blood cells—are continually formed from undifferentiated, pluripotent stem cells (Fig. 6.11).

In the adult, most of the blood cells develop in the bone marrow. The pluripotent stem cells form committed colony-forming cells (CFCs). The CFCs divide when stimulated by a colony-stimulating factor (CSF) and differentiate into various, mature blood cells, which usually die after a few days or weeks. Colony-stimulating factors are produced in one type of cells and act on other cells. CSFs signal through cytokine receptors, but some also signal through RTKs. A list of CSFs is given in Table 6.3. CSFs were isolated and characterized first by Don Metcalf and his laboratory at the Walter and Eliza Hall Institute in Melbourne.[48]

A central question is how the expression of lineage-specific genes in haematopoiesis is timed. Selection and activation of lineage-specific genes requires a mechanism that regulates activation of a particular set of genes at a critical moment, while at the same time other genes are silenced. Once the uncommitted pluripotential stage has been passed, factors come into play which address only genes that participate in the development of a certain lineage, while genes leading to the expression of other lineages remain inactive and/or are silenced. For that purpose, cell-restricted transcriptional regulators are used. We shall learn more about them in Chapter 11.

Timing of a signal is important. The same signal acting on the same cell at different stages of maturation, early or late in the life of a cell, may evoke different responses. How important the timing of a signal is was shown by Harvey Lodish and his group. Different

Table 6.3 Colony-stimulating factors (CSFs) (based on information published in Table 22.2 (p. 1171) of ref. 30

Factor	Target cells	Producing cells	Receptors
EPO	Blood cells	kidney cells	Cytokine Receptors
interleukins	pluripotent stem cells, and many terminally differentated cells	T lymphocytes,	
GM-CSF	Progenitor cells of granulocytes and macrophages (GM)	T lymphocytes, endothelial cells, fibroblasts	
G-CSF	GM progenitor cells, neutrophils	macrophages, fibroblasts	
M-CSF	GM progenitor cells, macrophages	fibroblasts, macrophages, endothelial cells	Growth factor Tyrosine kinase Receptors
Steel factor	haematopoietic stem cell	bone-marrow cells, and many other cells	

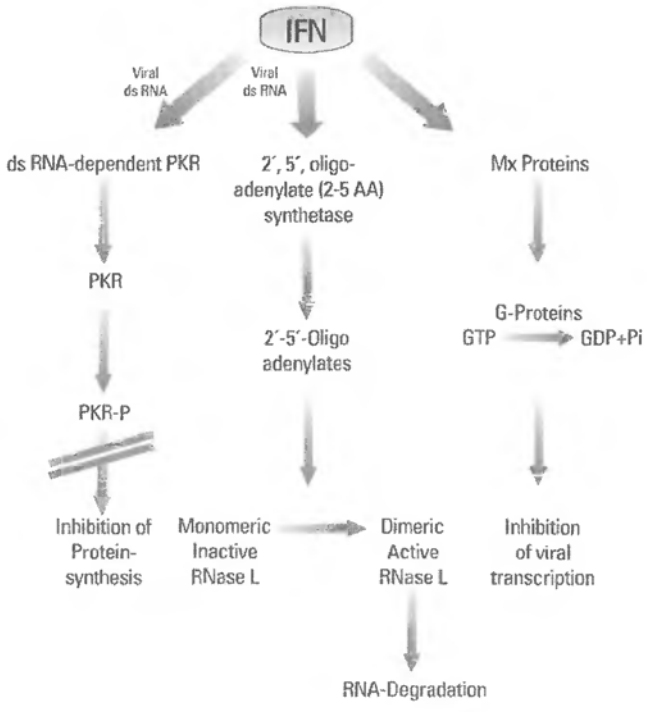


Fig. 6.12 IFN-inducible pathways. On the left is shown the pathway of the serine/threonine protein kinase (PKR), which is activated by viral double-stranded RNA (dsRNA). PKR is activated and phosphorylated, and phosphorylates and inactivates the elongation factor eIF2, resulting in inhibition of protein synthesis and viral replication. On the right is shown the Mx pathway. The Mx proteins are IFN-inducible, helical GAPs, belonging to the dynamin superfamily of GAPs. The activation of the Mx pathway by IFNs inhibits transcription and viral replication. In the middle is shown an antiviral pathway in which IFN-inducible oligoadenylate synthetases are stimulated by viral double-stranded RNA (dsRNA). These synthetases produce short 2′,5′-oligoadenylates (pppA′2p5′A2p5′A), which stimulate a 2′,5′-oligoadenylate-dependent RNase L, by forming the active, dimeric state of the RNase L. Activation of this pathway leads to extensive cleavage of single-stranded RNA and inhibition of viral replication. (This figure uses the information published by G. R. Stark *et al.*, in Fig. 5 of ref. 37. It is reproduced here by the generosity of the authors with permission of the Annu Rev. Biochem.)

signalling pathways are turned on at different stages of the differentiation programme. For example, at some stage in the differentiation of erythroid cell lines, the KIT receptor is activated by the stem-cell factor and the EPO receptor is switched off (see Chapter 1). This results in a switch from one signalling pathway to another, because the KIT receptor turns on another signalling route, different from that triggered by the EPO receptor (see 49).

Antiviral effects

IFNs were originally discovered more than 40 years ago by Isaacs and Lindemann because of their antiviral effects.[50] Today interferons are used mainly in the treatment of infectious viral hepatitis B and C. IFNs control genes involved in antiviral responses. Mice lacking the IFN-responsive transcriptional activator and transmitter, STAT 1, are defenceless against viral infections. Three antiviral pathways are induced by IFNs (Fig. 6.12).

References

1. J. Massagué. TGF-β signal transduction. *Annu Rev Biochem*, **67**, 753–791, 1998.
2. D. L. Griffith, P. C. Keck, T. K. Sampath, D. C. Rueger, and W. D. Carlson. Three dimensional structure of recombinant human osteogenic protein 1: Structural paradigm for the transforming growth factor β superfamily. *Proc Natl Acad Sci, USA*, **93**, 878, 1996.
3. C. Scheufler, W. Sebald, and M. Hülsmeyer. Crystal structure of human bone morphogenetic protein-2 at 2.7 Å resolution. *J Mol Biol*, **287**, 103–115, 1999.
4. X. Chen, M. J. Rubock, and M. Whitman. A transcriptional partner for SMAD proteins in TGF-β signalling. *Nature*, **383**, 691–696, 1996.
5. R. Derynck. TGF-beta-receptor-mediated signaling. *Trends Biochem Sci*, **19** (12), 548–553, 1994.
6. J.L. Wrana, L. Attisano, R. Wieser, F. Ventura, and J. Massagué. Mechanism of activation of the TGF-β receptor. *Nature*, **370**, 341–347, 1994.
7. C.-H. Heldin, K. Miyazono, and P. ten Dijke. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*, **390**, 465, 1997.
8. C. Niehrs. Growth factors. Mad connection to the nucleus [news; comment]. *Nature*, **381** (6583), 561–562, 1996.

   Also: F. Liu, A. Hata, J. C. Baker, J. Doody, J. Carcamo, R. M. Harland, and J. Massague. A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature*, **381** (6583), 620–623, 1996.
9. G. J. Hannon and D. Beach. p15INK4B is a potential effector of TGF-β-induced cell cycle arrest. *Nature*, **371**, 257–261, 1994.
10. A. Ivarone and J. Massagué. Repression of the CDK activator Cdc25A and cell cycle arrest by cytokine TGF-β in cells lacking the CDK inhibitor p15. *Nature*, **387**, 417–422, 1997.
11. A. Hata, R. S. Lo, D. Wotton, G. Lagna, and J. Massagué. Mutations increasing auto inhibition inactivate tumour suppressors Smad 2 and Smad 4. *Nature*, **388**, 82–87, 1997.

    Also: Y. Shi, A. Hata, R. S. Lo, J. Massagué, and N. Pavletich. A structural basis for mutational inactivation of the tumour suppressor Smad 4. *Nature*, **388**, 87–93, 1997.
12. H. Shibuya, K. Yamaguchi, K. Shiakabe, A. Tonegawa, Y. Gotoh, N. Ueno, *et al.* TAB 1: An activator of the TAK1 MAPKKK in TGF-β signal tramsduction. *Science*, **272**, 1179–1182, 1996.
13. J. Collignon, I. Varlet, E. J. Robertson, H. Shibuya, K. Yamaguchi, K. Shirakabe, *et al.* TAB1: an activator of the TAK1 MAPKKK in TGF-β signal transduction. *Nature*, **381** (6578), 155–158, 1996.
14. J. Schlessinger, I. Lax, and M. Lemmon. Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell*, **83** (3), 357–360, 1995.
15. L. B. Zimmerman, J. M. De Jesus Escobar, and R. M. Harland. The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell*, **86**, 4599–4606, 1996.
16. S. Piccolo, Y. Sasai, B. Lu, and E. M. De Robertis. Dorsoventral patterning in *Xenopus*: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell*, **86**, 4589–4598, 1996.

17. Y. G. Chen, F. Liu, and J. Massagué. Mechanism of TGF-β receptor inhibition by FKBP 12. *EMBO J*, **16**, 3866–3876, 1997.

18. J. Collignon, I. Varlet, and E. J. Robertson. Relationship between asymmetric nodal expression and the direction of embryonic turning [see comments]. *Nature*, **381**, 155–158, 1996.

    Also: L. A. Lowe, D. M. Supp, K. Sampath, T. Yokoyama, C. V. Wright, S. S. Potter, *et al.* Conserved left–right asymmetry of nodal expression and alterations in murine situs inversus. *Nature*, **381**, 158–161, 1996.

19. L. A. Lowe, D. M. Supp, K. Sampath, T. Yokoyama, C. V. Wright, S. S. Potter, *et al.* Conserved left–right asymmetry of nodal expression and alterations in murine situs inversus [see comments]. *Nature*, **381**, 158–161, 1996.

20. C. Meno, Y. Ito, Y. Saijoh, Y. Matsuda, K. Tashiro, S. Kuhara, and H. Hamada. Two closely-related left-right asymmetrically expressed genes, lefty-1 and lefty-2: their distinct expression domains, chromosomal linkage and direct neuralizing activity in *Xenopus* embryos. *Cell*, **82** (5), 803–814, 1995.

21. M. Levin, R. L. Johnson, C. D. Stern, M. Kuehn, and C. Tabin. A molecular pathway determining left–right asymmetry in chick embryogenesis. *Nature*, **371** (6494), 257–261, 1994.

22. M. Nomura and E. Li. Smad2 role in mesoderm formation, left-right patterning and craniofacial development [see comments]. *Nature*, **393** (6687), 786–790, 1998.

23. W. R. Waldrip, E. K. Bikoff, P. A. Hoodless, J. L. Wrana, and E. J. Robertson. Smad2 signaling in extraembryonic tissues determines anterior–posterior polarity of the early mouse embryo. *Cell*, **92**, 797–808, 1998.

24. C. Sirard, J. L. de la Pompa, A. Elia, A. Itie, C. Mirtsos, A. Cheung, *et al.* The tumor suppressor gene Smad4/Dpc4 is required for gastrulation and later for anterior development of the mouse embryo. *Genes Dev*, **2** (1), 107–119.1998.

25. G. Winnier, M. Blessing, P. A. Labosky, and B. L. Hogan. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev*, **9** (17), 2105–2116, 1995.

    Also: R. Derynck. Developmental biology: SMAD proteins and mammalian anatomy. *Nature*, **393**, 738–739, 1998.

    Also: J. Wrana and T. Pawson. Mad about SMADs. News and views. *Nature*, **388**, 28–29, 1997.

    Also: A. Hata, R. S. Lo, D. Wotton, G. Lagna, and J. Massagué. Mutations increasing auto inhibition inactivate tumour suppressors Smad 2 and Smad 4. *Nature*, **388**, 82–87, 1997.

    Also: X. Chen, M. J. Rubock, and M. Whitman. A transcriptional partner for MAD proteins in TGF-β signalling. *Nature*, **383**, 691–696, 1996.

26. T. M. Fynan and M. Reiss. Resistance to inhibition of cell growth by transforming growth factor-beta and its role in oncogenesis. *Crit Rev Oncog*, **45**, 493–540, 1993.

27. M. M. Shull, I. Ormsby, A. B. Kier, S. Pawlowski, R. J. Diebold, M. Yin, *et al.* Targeted disruption of the mouse transforming growth factor-beta 1 gene results in multifocal inflammatory disease. *Nature*, **359**, (6397), 693–699, 1992.

28. W. A. Border and E. Ruoslahti. Transforming growth factor-beta in disease: the dark side of tissue repair. *J Clin Invest*, **90** (1), 1–7, 1992.

29. J. Massagué. How cells read TGF-B signals. *Nature Reviews/Molecular Cell Biology*, **1**, (3), 169–178, 2000.

30. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell*. Garland Publishing, New York, 3rd edition, 1994.

31. J. N. Ihle and I. M. Kerr. Jaks and Stats in signalling by the cytokine super family *Trends Genet*, **11**, 69–74, 1995.

    Also: J. N. Ihle. Cytokine receptor signalling. *Nature*, **377** (6550), 591–594, 1995.

32. C. Schindler and J. E. Darnell Jr. Transcriptional responses to polypeptide ligands: the JAK-STAT pathway. *Annu Rev Biochem*, **64**, 621–651, 1995.

33. T. Taniguchi. Cytokine signaling through nonreceptor protein tyrosine kinases. *Science*, **268** (5208), 251–255, 1995.

34. T. Hage, W. Sebald, and P. Reinemer. Crystal structure of the interleukin-4 receptor α chain complex reveals a mosaic binding interface. *Cell*, **97**, 271–281, 1999.

35. A. M. de Vos, M. Ultsch, and A. A. Kossiakoff. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, **255**, 306–312, 1992.

36. M. R. Walter, W. T. Windsor, T. L. Nagabhushan, D. J. Lundell, C. A. Lunn, P. J. Zauodny and S. K. Narula. Crystal structure of a complex between interferon-gamma and its soluble high-affinity receptor [see comments]. *Nature*, **376** (6537), 230–235, 1995.

37. G. R. Stark, I. Kerr, M. Williams, R. H. Silverman, and R. D. Schreiber. How cells respond to interferons. *Annu Rev Biochem*, **67**, 227–264, 1998.

38. J. E. Darnell Jr, I. M. Kerr, and G. R. Stark. Jak-STAT pathways and transcriptional activation in response to IFN's and other extracellular signalling proteins. *Science*, **264**, 1415–1421, 1994.

39. T. Nosaka, J. M. A. van Deursen, R. A. Tripp, W. E. Thierfelder, B. A. Witthuhn, A. P. McMickle, *et al.* Defective lymphoid development in mice lacking JAK3. *Science*, **270**, 800–802, 1995.

40. S. M. Russell, N. Tayebi, H. Nakajima, M. C. Riedy, J. L. Roberts, M. J. Aman, *et al.* Mutation of Jak3 in a patient with SCID: Essential role of Jak3 in lymphoid development. *Science*, **270**, 797–800, 1995.

41. J. N. Ihle, B. A. Witthuhn, F. W. Quelle, K. Yamamoto, W. E. Thierfelder, B. Kreider, and O. Silvennoinen. Signalling by the cytokine superfamily: Jak's and Stat's. *Trends Biochem Sci*, **19**, 222–227, 1994.

42. G. J. Bhat, T. J. Thekkumkara, W.G. Thomas, K. M. Conrad, and K. M. Baker. Angiotensin II stimulates *cis*-inducing-factor-like DNA binding activity. *J Biol Chem*, **269**, 31443–31449, 1994.

43. M. S. Burfoot, N. C. Rogers, D. Watling, J. M. Smith, S. Pons, G. Paonessaw, *et al.* Janus kinase-dependent activation of insulin receptor substrate 1 in response to interleukin-4, oncostatin M, and the interferons. *J Biol Chem*, **272** (39), 24183–24190, 1997.

44. T. Naka, M. Narazaki, M. Hirata, T. Matsumoto, S. Minamoto, A. Aono, *et al.* Structure and function of a new STAT-induced STAT inhibitor. *Nature*, **387**, 924–929, 1997.

45. R. Starr, T. A. Willson, E. M. Viney, L. J. L. Murray, J. R. Rayner, B. J. Jenkins, *et al.* A family of cytokine-inducible inhibitors of signalling. *Nature*, **387**, 917–920, 1997.

46. T. A. Endo, M. Masuhara, M. Yokouchi, R. Suzuki, H. Sakamoto, K. Mitsui, *et al.* A new protein containing an SH2 domain that inhibits JAK kinases. *Nature*, **387**, 921–924, 1997.

47. R. A. Firtel. Integration of signaling information in controlling cell-fate decision in Dictyostelium. *Genes Dev*, **9**, 1427–1444, 1995.

    Also: J. Williams and A. Morrison. Prestalk cell-differentiation and movement during the morphogenesis of Dictyostelium discoideum. *Prog Nucleic Acid Res Mol Biol*, **47**, 1–27, 1994.

48. D. Metcalf. The hemopoietic regulators; an embarassment of riches. Bioassays. **14**, 799–805, 1992

49. H. Wu, U. Klingmüller A. Acurio, J. G. Hsiao, and H. Lodish. Functional interaction of erythropoietin and stem cell factor receptors is essential for erythroid colony formation. *Proc. Natl. Acad. Sci. USA*, **94**, 1806–1810, 1996.

    Also: H. Wu, U. Klingmüller, P. Besmer and H. Lodish. Interaction of the erythropoietin and stem-cell factor receptors. *Nature*, **377**, 242–246, 1995.

50. A. Isaacs and J. Lindemann. Virus interference. I. The interferon *Proc R Soc London, Ser B*, **147**, 258–267, 1957.

7

Control of signalling by phosphorylation dephosphorylation

Chapter 7 will sum up the facts concerning phosphorylation dephosphorylation, a mechanism involved in nearly all cellular regulatory processes. As example of the role of phosphorylation dephosphorylation the response of the body to insulin will be examined in Chapter 8.

We have learned:

1. that the control of interconnectivity of the components of signalling pathways regulates the flow of information from the receptor on the cell surface to the gene; and
2. that the major mechanism controlling interconnectivity and the flow of signals is phosphorylation dephosphorylation.

Protein phosphorylation dephosphorylation regulates the activity of enzymic and nonenzymic proteins in eukaryotic cells. Ten per cent or more of all proteins in a cell are modified in that way. The phosphates are transferred from ATP and esterified with hydroxyl groups of serine, threonine, or tyrosine residues. They are removed and transferred to water by phosphatases. There are at least about 2000 kinases and about 1000 phosphatases to carry out these reactions (Fig. 7.1).[1]

Phosphorylation dephosphorylation modifies protein structure and function and protein–protein interactions. The consequences are many. Phosphorylation of a protein may signal its destruction or may prevent it. It may control the activity of an enzyme or may determine its location in the cell. Phosphorylation dephosphorylation can modulate the activity of transcription factors and control gene expression, and phosphorylation dephosphorylation may regulate cell–cell interactions.

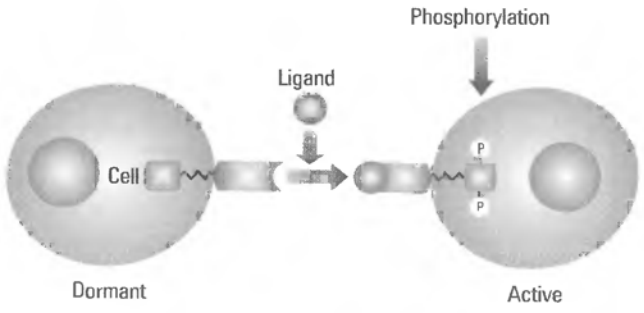


Fig. 7.1 Often, the first step in the response of a cell to a signal is receptor activation and phosphorylation. The signal is turned off by dephosphorylation

In principle, all signalling mechanisms involve conformational switches. This applies to receptors, linkers, transducers, and transcription factors. Conformational transitions of the receptor are induced by ligand binding and propagated by receptor oligomerization and phosphorylation reactions. The consequences are either changes in interconnectivity or enzymic activity, or both. G proteins and receptors turn signalling on or off. Specific recognition among signalling proteins selects the route where a signal travels and phosphorylation dephosphorylation regulates the traffic.

The organization of phosphorylation cascades

Receptors on the cell surface accept the signal and send it on its way through the cell to the nucleus.

Receptor kinases

Receptors with intrinsic or acquired kinase activity generate the input signal. The response of receptor kinases to activation by the ligand is slow and so is the cellular response. Phosphorylation of receptors initiates phosphorylation cascades and cytosolic kinases amplify and diversify the signal.

Cytosolic kinases

The response of cytosolic kinases to the input signal can be regulated by monomeric G proteins. An example is the control of the Raf-1 kinase in the MAP kinase cascade by the GTP/GDP on/off switch function of Ras. Signals passing through the Ras control point can be routed in different directions. The right of way is controlled by kinases. The specificity of the kinases overrules the coupling promiscuity of Ras, and other linkers.

Cells can choose between several phosphorylation cascades, to cope with a multitude of extra- and intracellular signals. Cells have kinases which respond to second messengers, such as PKA, PKC and Ca^{2+}-calmodulin-dependent kinases. Each of these kinases can establish a separate phosphorylation cascade.

Second-messenger-controlled kinases

The classical 'second messenger' is cAMP, which is formed by the enzyme adenylyl cyclase in response to hormones such as adrenaline. The cAMP signal is transmitted through a cAMP-responsive protein kinase, PKA. Kinases, such as the phosphatidylinositol 3-kinase also form potent second messengers. An example is the formation of Ptdlns-3,4,5-P_3 (PIP_3) (discussed in Chapter 5). The Pdtlns 3-kinase phosphorylation cascade plays a role in many essential cellular and developmental events, such as dorsoventral patterning in *Xenopus* and *Drosophila*. In neurons, NGF controls the Ptdlns 3-kinase phosphorylation cascade and decides over survival or death of these cells. Finally, the Ptdlns-3,4,5-P_3 messenger-controlled phosphorylation cascade has been implicated in the insulin response (Chapter 8).

Phosphoprotein phosphatases: localization and targeting

An essential requirement of all regulatory processes is reversibility. Prime examples are the on/off switch of a monomeric G protein, Ras, which controls signalling through the MAP kinase cascade and the reversal of phosphorylation by phosphatases, catalysing

phosphoprotein dephosphorylation. This allows fine tuning of signalling by continously checking and downregulating the amplitude of signalling.

A mode of reciprocal regulation such as phosphorylation dephosphorylation is like a valve, releasing the pressure of excess activity.[2] Phosphorylating and dephosphorylating enzymes are often linked functionally; thus the phosphatase may be activated by the kinase and the kinase inactivated by the phosphatase. In some cases, protein kinases and protein phosphatases are even associated with each other, forming a complex. This may facilitate reciprocal control. But not every individual phosphorylation reaction may be shut off by the activation of a corresponding phosphatase. Instead, one should consider that a level of phosphatase activity may be maintained constitutively, continously damping and controlling the state of phosphorylation and establishing a state of activity, tailored to the needs of the cell. Damping of signalling by dephosphorylation may have a different time scale than phosphorylation, and may be controlled differently and by different signals.

For dephosphorylation reactions, many enzymes with different specificities are available: There are phosphatases, specific for phosphotyrosyl residues, for Ser/Thr-bound phosphates, and dual-specificity phosphatases, recognizing both phosphotyrosyls and phosphoserines Tyrosine phosphatases and dual-specificity phosphatases have already been introduced (Chapter 3). Here, the properties of serine/threonine phosphatases will be described[3] and their regulation by cellular relocation introduced. Much of what we know about the regulation of this class of phosphatases we owe to the work of P. Cohen and his colleagues. Table 7.1 lists common phosphoserine/phosphothreonine phosphatases of eukaryotes.

There are three modes of regulation of the activity of protein phosphatases:

1. Regulation by phosphorylation, just like in the case of some protein kinases.
2. Regulation by relocation. Mechanisms to overcome cellular compartmentalization are pivotal points of regulatory control. Relocation may involve targeting proteins. Targeting subunits guide both protein kinases and phosphatases to their substrates, located in different subcellular compartments.
3. Some phosphatases are activated, just like some kinases, by second messengers.

Table 7.1 The properties of phosphoprotein (Ser/Thr) phosphatases (reproduced from ref. 4 with permission of Dr D. Barford and Trends in Biochemical Sciences)

Catalytic subunit	Regulatory subunit
PP1	Numerous, targeting, G_M, and regulatory subunits
PP2A	Regulatory A- and B-subunits
PP2B	Regulatory B-subunit, AKAP.
	Calmodulin makes the PP2B phosphatase in T cells responsive to Ca^{2+} signals, regulating nuclear transfer of the transcription factor NF-AT
PP2C	A large family of protein phosphatases. Reverses stress-induced, SAP kinase phosphorylation cascades

PP, protein phosphatase; G_M is a targeting subunit guiding PP1 in skeletal muscle to glycogen and myosin. See Fig. 7.4.; AKAP, A-kinase-associated protein; NF-AT, nuclear factor of activated T cells; SAP, stress-activated protein. PP2B, calcineurin is a target of the immunosuppressant FK 506 bound to the FK binding protein FKB. (FK is thought to be an abbreviation of this class of compounds which were originally developed by the pharmaceutical industry because of their strong fungicidal action).

X-ray crystallography of the protein phosphatase-1 (PP1) catalytic subunits,[5] and of a holoenzyme (the human Ca^{2+}-calmodulin-activated PP2B-type phosphatase, calcineurin), and of a complex of calcineurin with the immunosuppressant, FK506 bound to its binding protein FKB,[6] has helped in understanding the mechanism of catalysis and regulation of this important class of regulatory enzymes. The serine/threonine phosphatases are unrelated in sequence and structure to the protein tyrosine phosphatases (Chapter 3) (compare with ref. 7; for more information, see ref. 7 and ref. 8).

Ser/Thr phosphatases are metalloenzymes. They have been classified into four groups: 1, 2A, 2B, and 2C. In the case of PP1 and PP2B, the catalytic metal ions are Mn^{2+} and Fe^{2+} and Zn^{2+} and Fe^{2+}, respectively. Thanks to structural evidence, provided mainly by David Barford and his colleagues,[8] we have a good idea of the catalytic mechanism. The two metal ions are coordinated in the centre of the catalytic site by aspartate and asparagine residues, assisted by histidine residues. The side-chain of an aspartate and a water molecule bridge the two metal ions and form a metal centre. The metal-bound water acts as a nucleophile and attacks the phosphorus atom of the phosphate of the substrate (by an S_{N2} mechanism). The side-chain of a histidine near the Ser or Thr side-chains donates a proton to the leaving

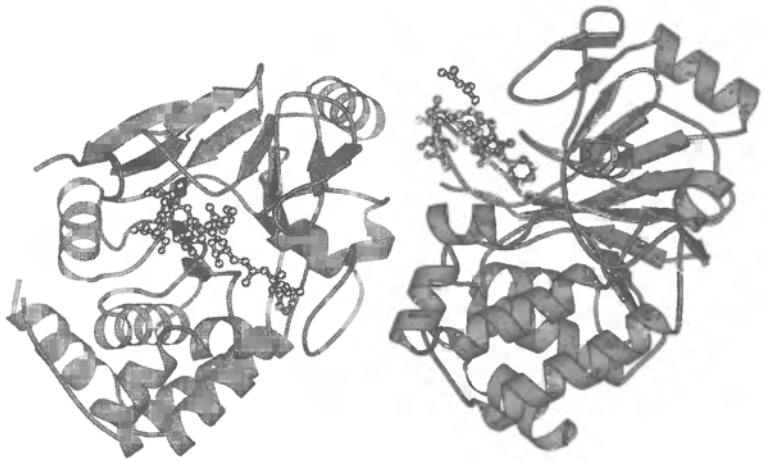


Fig. 7.2 The crystal structure of mammalian Ser/Thr protein phosphatase-1, complexed with the toxin mycrocystin was determined at 2.1 Å resolution. PP1 has a single domain with a fold, distinct from that of the protein tyrosine phosphatases. The Ser/Thr protein phosphatase-1, is a metalloenzyme with two metal ions positioned at the active site with the help of a β–α–β–α–β scaffold. A dinuclear ion centre consisting of Mn^{2+} And Fe^{2+} Is situated at the catalytic site that binds the phosphate moiety of the substrate. Ser/Thr phosphatases, PP1 and PP2A, are inhibited by the membrane-permeable ocadaic acid and by cyclic hexapeptides, known as microcystins. The toxin molecule is depicted as a 'ball-and-stick' structure. On the left and on the right, two different views of the same molecule are shown. Microcystin binds to three distinct regions of the phosphatase: to the metal-binding site, to a hydrophobic groove, and to the edge of a C-terminal groove in the vicinity of the active site. At the surface are binding sites for substrates and inhibitors. These ribbon models are reproduced with permission of the authors and Nature from ref. 9.

group. The structural data indicate that dephosphorylation is catalysed in a single step by a metal-activated water molecule. This single-step reaction is consistent with the effect of mutations of the catalytic residues and with the fact that a phosphoryl enzyme intermediate is not formed during the reaction. (This contrasts with protein tyrosine phosphatases, acid and alkaline phosphatases, which form phosphoryl–enzyme intermediates.)

The three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1 is shown in Fig. 7.2.[9] Plate 13 shows the X-ray structure of the ternary complex of a fragment of the serine/threonine phosphatase, calcineurin A (PP2A), with the FK-binding protein, FKBP12, and the immunosuppressant drug FK506.[10] This structure is shown because it gives an explanation why and how the phosphatase calcineurin is inhibited by a complex of the immunosuppressor FK506 with its binding protein, FKBP12. This interaction is central to the immunosuppressive actions of the FKB–FK506 complex.

Comparing the structures in Fig. 7.2 and Plate 13 reveals how highly conserved the catalytic domains of Ser/Thr phosphatases are. Therefore their functional diversity must be due to different regulatory and targeting subunits that direct the phosphatase to discrete subcellular locations (Fig. 7.3).

The specificity of targeting is remarkable: for example, the smooth muscle PP1 phosphatase combined with its targeting subunit dephosphorylates exclusively smooth muscle myosin, whereas the targeting subunit of skeletal muscle PP1 guides the phosphatase to skeletal muscle myosin, but not to smooth muscle myosin. In skeletal muscle, a substantial portion of PP1 is associated with glycogen particles through its targeting subunit, G_M. The G_M subunit can direct the PP1 phosphatase either to glycogen particles or to the membranes of the sarcoplasmic reticulum, which are devoid of glycogen.

The control of targeting and the regulation of the protein phosphatase $PP1_G$, which is targeted to glycogen, is summarized in Fig. 7.4. The information on which these drawings are based can be found in ref. 11.

Another important location of the phosphatase is the nucleus. In yeast, a targeting subunit directs the PP1 phosphatase to the nucleus. Ca^{2+} signals are probably transmitted in a similar way. The phosphatase PP2B is dependent on Ca^{2+} for activity (PP2B is identical with calcineurin). The enzyme has a A-subunit with an amino-terminal catalytic domain and a carboxy-terminal regulatory B-subunit, which binds calmodulin (calmodulin is a common, Ca^{2+}-binding protein). In order to activate the phosphatase maximally, Ca^{2+} must bind to both the B-subunit and calmodulin, and both moieties must associate with the A-subunit of PP2B. This opens the way for the active A-subunit of PP2B to address cellular targets and dephosphorylate, in a Ca^{2+}-dependent manner, reponsive phosphoenzymes.

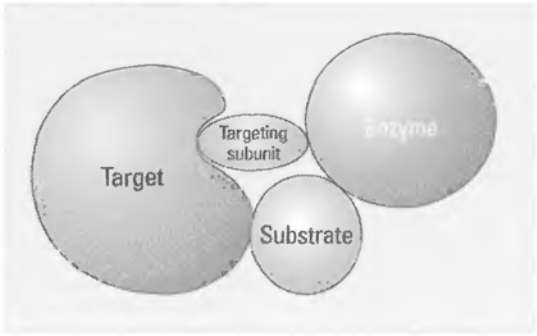


Fig. 7.3 The targeting subunit is a discrete entity, different from the substrate.

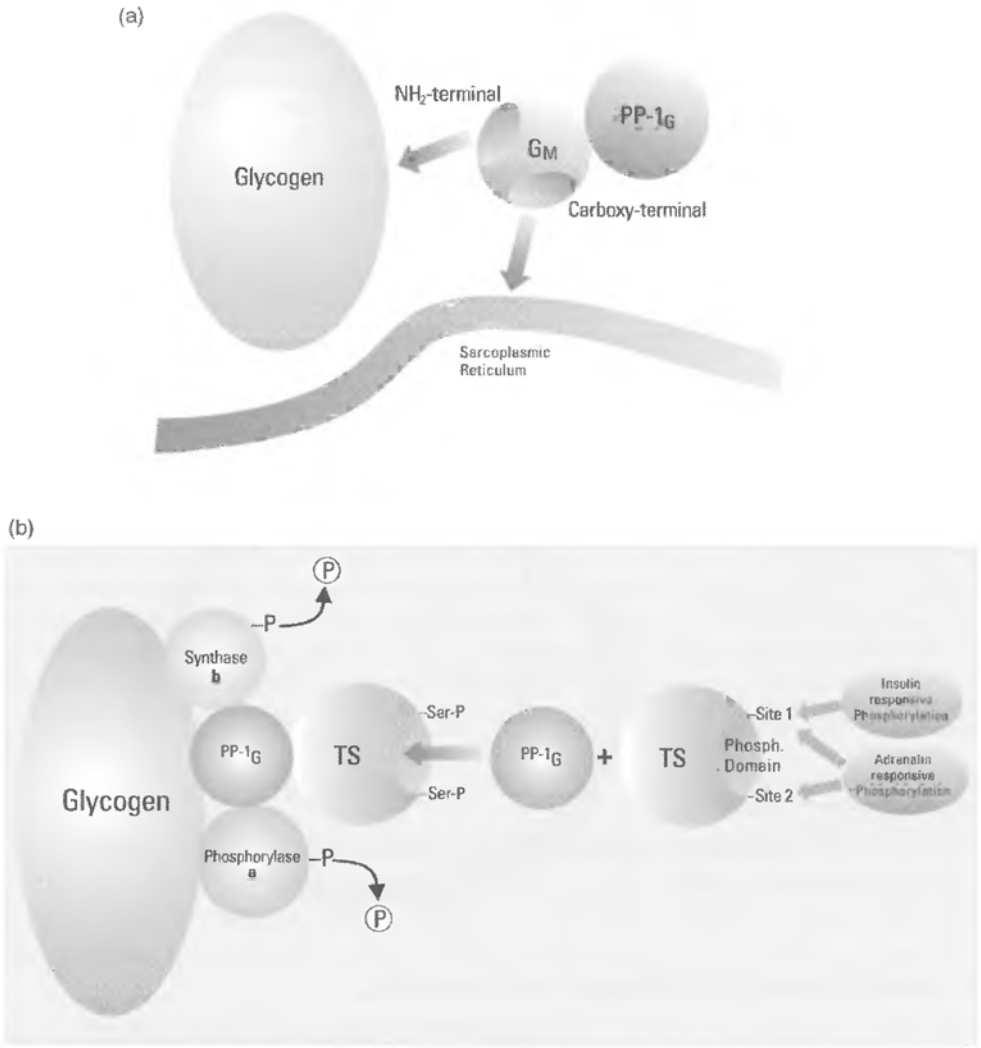


Fig. 7.4 (a) Binding of the phosphatase, PP-1_G with the help of the targeting subunit, G_M, either to the sarcoplasmic reticulum or to glycogen. Binding sites for glycogen are in the amino-terminal regions, whereas the recognition site for membranes is in the hydrophobic part of the carboxy-terminal region of the targeting subunit, G_M. What actually determines the choice between the two targets is not clear, but the targeting subunit G_M is a point of control in the hormonal regulation of glycogen metabolism. (b) PP-1_G in striated muscle is controlled by insulin and adrenaline. Each of these hormonal signals is transmitted through specific kinases and phosphatases; and each of these hormones can change the phosphorylation state of the site 1 and site 2 serines in the amino-terminal part of the targeting subunit, TS, of PP-1_G in a different manner. For example, insulin-mediated phosphorylation of TS modulates the interaction of the PP-1_G phosphatase with the glycogen-bound phosphoenzymes, the inactive glycogen synthase *b* and the active glycogen phosphorylase a. Dephosphorylation activates glycogen synthase by forming the active, dephosphorylated form a, and inactivates glycogen phosphorylase *a*, forming the inactive *b*-form of the enzyme. In that way, insulin can promote glycogen synthesis.

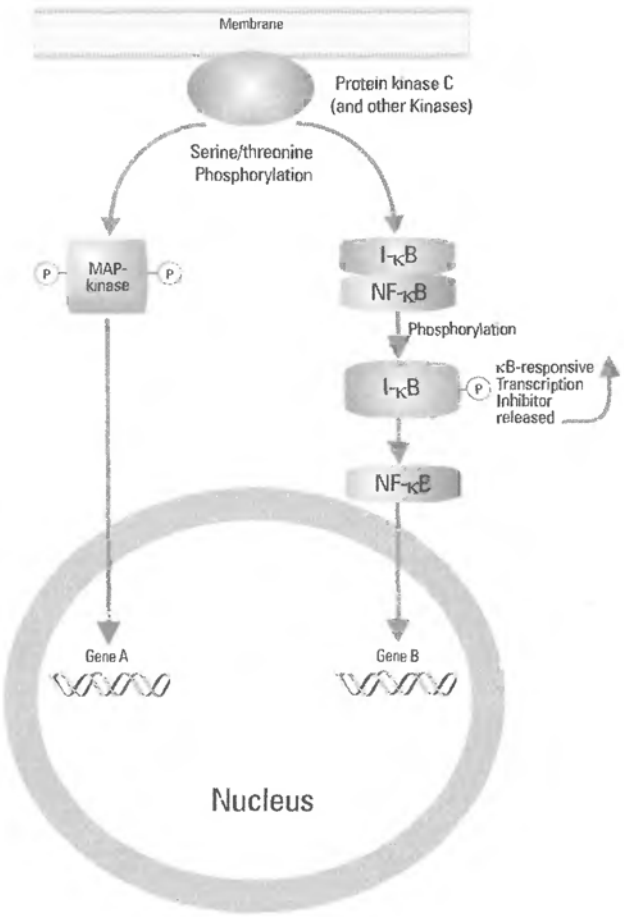


Fig. 7.5 There are two protein kinase C-controlled signalling routes. In one, a membrane-bound, activated C kinase phosphorylates a MAP kinase and transmits signals to the MAP kinase. The other route circumvents the MAP kinase. PKC phosphorylates instead the inhibitory Iκ-B protein which releases a transcription factor, NF-κB, which travels to the nucleus and activates gene transcription. The point to remember is that one and the same kinase, PKC, transmits signals to different transcription factors and to different genes, resulting in different biological responses. (Based on the information in Fig. 15–32 in Chapter 15 of *Molecular biology of the cell*.[13]

Protein kinases: localization and targeting

The role of targeting subunits was first recognized in the case of protein kinase C, as it became clear that isoforms of protein kinase C and their substrates have different subcellular locations, raising the question of how the enzymes find their substrates. PKCs are membrane-attached Ser/Thr protein kinases. They were discovered by Y. Nishizuka.[12] They are activated by diacylglycerol (DAG) and Ca^{2+}. Diacylglycerol is formed from phosphatidylinositol-4′, 5′-P_2 by phospholipase C. The other product of this reaction is IP_3, a lipid messenger, releasing Ca^{2+} from intracellular stores. The protein kinase C family has at least 11 members; the largest amounts of protein kinase C are present in the brain. Some routes of PKC signalling in the brain are indicated in Fig. 7.5; there are others.

Most cells express one or more isoforms of PKC, of which the substrates are localized in different cellular compartments. This means that either the PKC must be brought to the same compartment where the substrate is, or the substrate must be brought to the enzyme. It turned out that the PKC is directed to the substrate. The PKCs find their subcellular locations with the help of targeting proteins. There exist two groups of targeting proteins for PKCs. To one group belong components of the cell structure, the cytoskeleton and the

membrane. The other targeting proteins belong to a class of regulatory proteins. An example is PICK-1, a perinuclear protein that binds to the catalytic core of PKCα and inhibits the enzyme (PICK stands for perinuclear inhibitor of C kinase). There may be a third class of proteins that interact specifically with PKC. These are cytosolic kinases, which may recognize PKC with the help of their PH-domains. Finally, there are also 'targeting' proteins for PKCs, differing from other targeting subunits, because they are substrates and are phosphorylated by PKCs. In this case, binding of the 'targeting protein' is like an enzyme–substrate interaction. There is still another mode of interactions of PKCs. PKCs attached to cell membranes interact with receptors that do not have a kinase activity, and form RACKs (receptor-associated C-kinases). These receptors are like cytokine receptors, and just as cytokine receptors recruit JAKs, these receptors bind PKC. Since PKC binds to the receptor at a site distinct from the substrate-binding pocket of the kinase, binding does not interfere with phosphorylation of substrates of PKC. The part of the RACK kinase interacting with the receptor has some similarity to the β-subunit of heterotrimeric G proteins. It contains seven WD motifs and probably has a β-propeller structure, like the β-subunit of G proteins. This scaffold-like structure of RACK may help to accommodate the receptor and the substrates of the kinase.

Another second-messenger-dependent kinase which interacts with targeting proteins, is the cAMP-dependent protein kinase A. PKA interacts with AKAPs (A-kinase-associated proteins). AKAPs bind to the dimeric form of the regulatory subunit of PKA. They are multivalent linkers, which bind not only to PKA, but also to other kinases, such as PKC, and the Ca^{2+}/ calmodulin-dependent kinase II and the phosphatase PP2B (Fig. 7.6).

AKAPs have been found in nearly every subcellular structure and organelle; in the ER, the Golgi, in mitochondria, secretory granules, membranes, nuclear matrix, microtubules, and the centrosomes.

Are phosphorylation cascades multi-enzyme organelles?

From subcellular compartmentalization it is not far to the organization of kinases and phosphatases as supramolecular multi-enzyme organelles. What advantages could the arrangement of a protein kinase cascade such as, for example, the MAP kinase cascade,

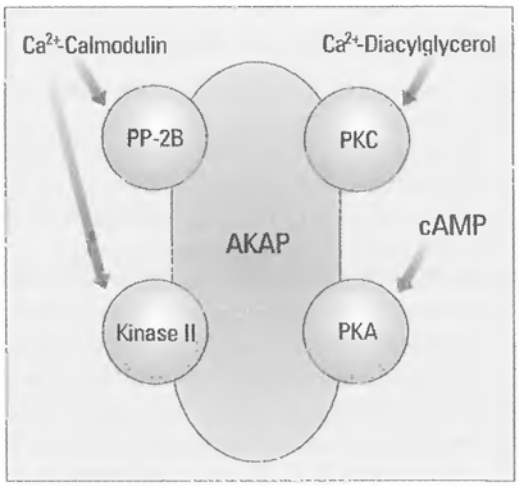


Fig. 7.6 The complex of PP2B, PKA, Kinase II, and PKC with AKAP. AKAP targets the kinases and the phosphatase to subsynaptic sites in neurons. Potential substrates are glutamate- and NMDA (*N*-methyl *D*-alanine) receptors and synaptic Ca^{2+} channels. (The information on which this scheme is based can be found in Fig. 3 of ref. 14. It is reproduced with permission of the authors and Trends in Biochem. Sci.)

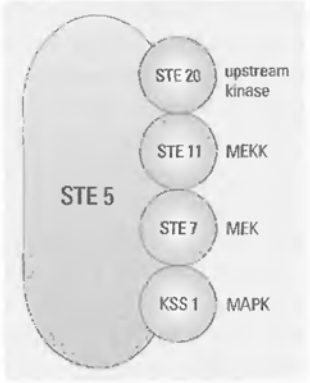


Fig. 7.7 The assembly of the MAP kinase phosphorylation cascade in yeast with the scaffolding protein, STE 5 (The original drawing is in Fig. 3 of ref. 14. It is reproduced with permission of the authors and Trends in Biochem. Sci.)

have as an interconnected multi-enzyme assembly? An answer is easily found when one looks at a classical multi-enzyme complex, such as the fatty acid synthase complex in yeast and humans, where one polypeptide chain carries several enzyme activities, allowing one enzyme to hand its product to the next enzyme in line, which then can use it as substrate. Intermediates are not lost in solution and do not need to be recovered. The assembly of a phosphorylation cascade as a 'functional organelle' might have similar advantages. The pheromone mating response in yeast is arranged in such a way. In the mating response of haploid yeast cells to pheromone, a kinase cascade is turned on by activation of an upstream kinase (STE 20), a MAPKKK (MEKK) homologue (STE 11), and a MAPKK (MEK) homologue (STE 7), which in turn phosphorylates and activates the MAP kinase homologues (FUS3 or KSS1) (Fig. 7.7). (MEK is an abbreviation of mitogen and extracellular signal-responsive kinase. It is used here to simplify the term MAPKK which is now MEK and MAPKKK is MEKK). The orderly, stepwise progression of these phosphorylation reactions is facilitated by a multivalent protein, STE 5, which is like a scaffold that couples to each kinase and regulates the access of substrate in a sequential order.

Is the mammalian Ras/Raf/MAP kinase cascade organized in a similar way with the help of an organizer protein? We do not know. But, one might speculate that Ras and Ras-like proteins and linkers, which are multivalent binding proteins and are integrated in the cell structure through attachment to the membrane, might function as organizers, like STE 5. Although this remains open for the time being, one can expect that more scaffolding proteins will be found to help in the organization of kinase cascades.

Regulation of enzymes and signalling pathways by phosphorylation dephosphorylation: a comparison

There are two major ways of control. One mechanism involves reversible covalent modifications, such as phosphorylation dephosphorylation, the other requires conformational transitions by binding an allosteric ligand or regulator protein. It follows an example of regulation of an enzyme, of which the activity is subject to control by both mechanisms, then we compare the regulation of an enzyme with regulation of components of cellular signalling pathways, of which many have no enzymic activity.

There are many examples of the regulation of enzymes by phosphorylation dephosphorylation and by allosteric ligands. A classic example is the regulation of the enzymes of glycogen metabolism. The first example of the activation of an enzyme by phosphorylation was muscle glycogen phosphorylase, an enzyme that degrades, with phosphate rather than water, 1,4-linked glucosyl residues in glycogen ($glucose_n$) to glucose 1-P and limit dextrin (glucose n–1). The enzyme was discovered and characterized by Carl F. Cori and G. T. Cori in the late 1930s.[15] E. H. Fischer and E. G. Krebs, then found, in 1958,[16]

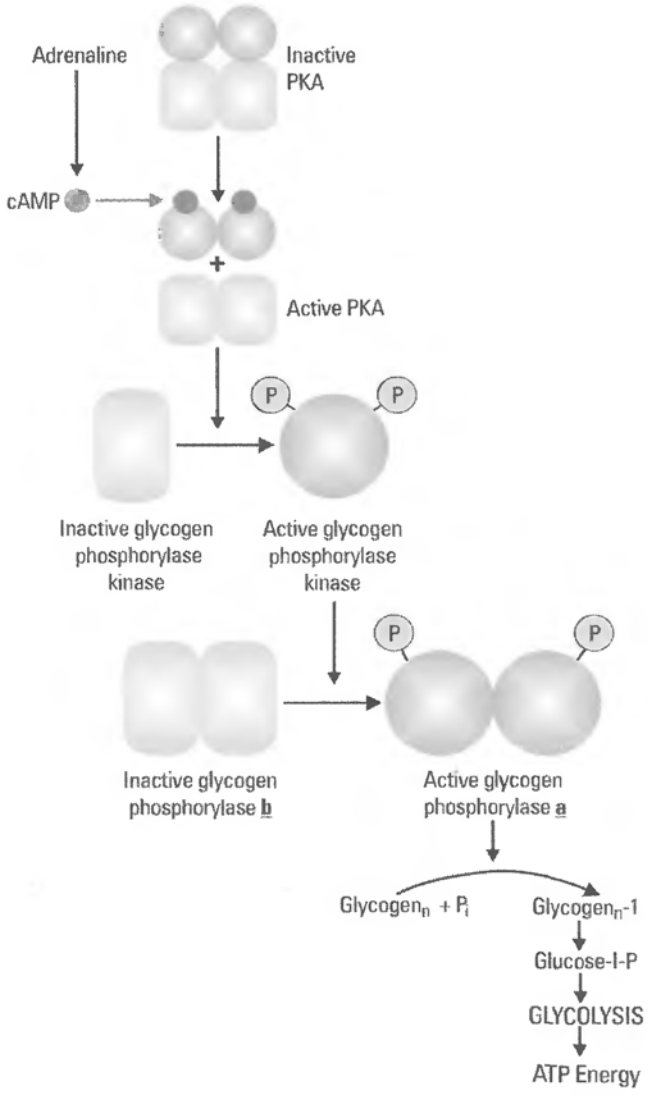


Fig. 7.8 Glycogen phosphorylase is activated by the hormone adrenaline (and by Ca^{2+} In the working muscle). Consequently, glycolysis is turned on in muscle, producing readily available energy (ATP). Adrenaline binds to β-AR, the G-protein-coupled *β*-adrenergic receptor, and activates a G protein, G_S. The target of G_S is the enzyme adenylyl cyclase, which produces the second messenger cAMP from ATP. cAMP initiates a phosphorylation cascade by activating the cAMP-dependent protein kinase A which, in turn, activates a glycogen phosphorylase-specific kinase. The phosphorylase kinase converts an inactive, non- phosphorylated form, glycogen phosphorylase *b* to an active, phosphorylated form, phosphorylase a. (Phosphorylation is reversed by a separate cascade of phosphatases, not shown.)

an enzyme activity that catalysed the covalent attachment of a phosphate group, derived from ATP, to a serine of glycogen phosphorylase *b*, converting it to the active phosphorylated form, phosphorylase *a* (Fig. 7.8).

Regulation of glycogen phosphorylase and glycogen synthase

The phosphorylated and unphosphorylated forms of glycogen phosphorylase have markedly different conformations, as shown by high-resolution X-ray crystallography, carried out by Louise Johnson and David Barford,[17] in Oxford and by Robert Fletterick and Steven Sprang in San Francisco.[18] The structures provide an explanation for the low activity of the unphosphorylated enzyme *b* and the high activity of the phosphorylated enzyme *a* (Plate 14).

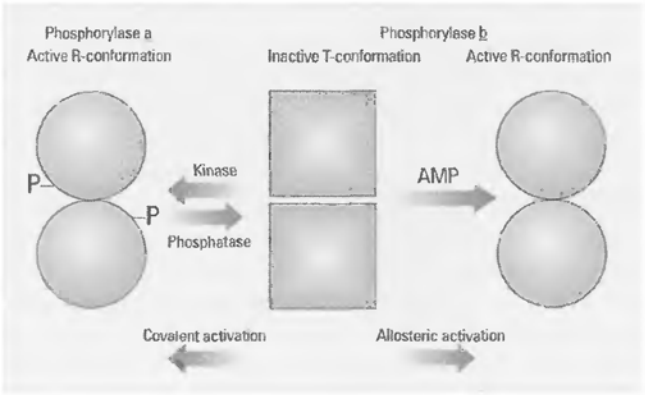


Fig. 7.9 Comparison of the allosteric regulation of glycogen phosphorylase with covalent regulation by phosphorylation dephosphorylation. The role of protein phosphatases in regulation was realized for the first time more than 50 years ago, as a cellular enzyme activity that converts the active *a* form of glycogen phosphorylase to an inactive *b* form.[23]

But, aside from phosphorylation, the nonphosphorylated *b* form of glycogen phosphorylase can also be activated allosterically[20] by binding AMP.[21] Allosteric activation of glycogen phosphorylase and covalent activation by phosphorylation are elicited in the cell by different signals (Fig. 7.9). Through interaction with the allosteric metabolite, 5′-AMP, phosphorylase can respond to the metabolic needs of the cell, whereas phosphorylation triggered by the second messenger cAMP and a cAMP-triggered phosphorylation cascade allows phosphorylase to respond to external signals (circulating hormones such as adrenaline) but also to internal signals (such as Ca^{2+}, released in the contracting muscle). Ca^{2+} signals through a Ca^{2+}-calmodulin-dependent kinase and a Ca^{2+}-triggered phosphorylation cascade which converts phosphorylase *b* to *a*.

Summary

Allosteric activation and covalent activation

The term 'Allosteric activation' was proposed by Jacques Monod,[20] who introduced and propagated this fundamental concept as the *deuxième loi de la biologie*, (the first law, I suppose, would be a molecular definition of the genes and the genetic code). Allosteric and covalent activation of proteins by phosphorylation are not only involved in regulation of key enzymes, they are also fundamental mechanisms in cellular signal transfer, because they apply as well to regulatory proteins without enzymic activity. For example, the first step in signal transmission involves allosteric activation: external signals, hormones, growth factors, and cytokines are allosteric activators. They bind to their cognate receptors in the plasma membrane and induce an 'allosteric' conformational transition, leading to activation and, in some cases, to oligomerization and autophosphorylation. Receptors, activated upon binding of the ligand, can then recruit adaptors and transmitters through specific autophosphorylated sites. Binding of hormones to G-protein-coupled receptors is a variation on the same theme. The ligand or agonist triggers a conformational transition that creates binding sites for heterotrimeric α,β,γ G proteins. The G proteins are activated and take over the role of the receptor and in turn activate allosterically regulated enzymes, such as adenylyl cyclase and phospholipases, which form second messengers, such as cAMP and phospholipid messengers. The second messengers turn

on enzymes, often kinases, again by ligand-triggered conformational transitions. The activated kinases then propagate the signal by phosphorylation and covalent modification, of components of signalling cascades. As we shall see (Chapter 11), lipophilic hormones, such as steroid hormones take a short cut and circumvent membrane-bound receptors, targeting genes directly.

Finally, an important point which distinguishes covalent activation and allosteric activation should be mentioned. That is their different mode of deactivation. In the first case we have a distinct class of enzymes, the phosphatases, which deactivate and reverse phosphorylation, whereas in the second case deactivation is regulated by dissociation of the activator or binding of inhibitors.

Phosphorylation dephosphorylation controls interconnectivity

We have seen that the prime function of phosphorylation in cellular signalling is modification of the interconnectivity of the components of signalling cascades. Since phosphorylation dephosphorylation can be regulated independently, the regulatory potential of this modification of interconnectivity is nearly unlimited. It endows cells with a versatile network, capable of transmitting a multitude of signals to different targets. The emergence of many enzymes catalysing phosphorylation dephosphorylation reactions parallels the evolution of more and more intricate and subtle ways of controlling the signals received by cells. These regulatory networks have given organisms the versatility to adapt to a changing environment, which is essential for their survival.

How phosphorylation dephosphorylation changes the properties of regulatory proteins

There are several ways by which phosphorylation dephosphorylation changes the properties of regulatory proteins.

1. Introduction of phosphates into proteins may expose new recognition surfaces, may change their location in the cell, promote transport into the nucleus, or modify interactions with genes. Moreover, the association of a regulatory protein with a modifying enzyme creates a new heterologous complex with new interaction surfaces, thus opening the way to still other modifications. This can create new opportunities for recruiting other partners with different structural and functional properties.
2. Targeting of kinases and phosphatases to other locations in the cell brings enzymes and substrates together.[14] Some enzymes and proteins have built-in recognition domains for cellular structures, or have acquired such recognition domains by attachment of fatty acyl- or glycosylphosphatidylinositol (GPI) anchors. Others require targeting subunits. Targeting subunits are either separate entities, regulatory subunits, or substrates of enzymes. The kinases, targeted to subcellular locations, have broad and often overlapping substrate specificities. Thus, to bring a quite unspecific kinase or phosphatase to cellular compartments, where only one kind of substrate is hidden, contributes to selectivity and specificity of the reactions. Conversely, relocation can also terminate the reaction by separating enzyme and substrate.
3. Scaffolding proteins may organize kinases and phosphatases in supramolecular multienzyme complexes, facilitating reactions and improving specificity by preventing cross-reactions.

4. Positive control by phosphorylation and negative control by dephosphorylation are equally important. Depending on the properties of the signalling proteins and the nature of signals, phosphorylation may be more important in some cases and dephosphorylation in others. Moreover, in some cases dephosphorylation may have positive and phosphorylation negative effects on the activity of regulators.

Does the receptor switch on and off the activity of the whole signalling cascade like the first enzyme in a chain of metabolic reactions?

Are the consequences of the activation of a receptor which provides the input signal and starts a phosphorylation cascade different from the consequences of activation of an enzyme, such as glycogen phosphorylase, which is the first enzyme feeding substrates into a chain of metabolic reactions and supplying glycolysis with fuel?

Growth factor and cytokine receptors with inherent or acquired tyrosine kinase activity are autophosphorylated on binding of their respective ligands. Phosphorylation is accompanied by receptor oligomerization. But, what comes first is not always clear. However, it is certain that binding of the ligand to the receptor turns on the whole signalling chain. If phosphorylation of the receptor is the pacemaker, triggering activation of the whole signalling chain, one would expect deactivation of the receptor to turn off the whole signalling cascade. This is highly unlikely, because the key regulators of the individual downstream steps in a signal-transfer cascade are usually regulated individually. Thus the control function of a receptor which delivers the input signal for a signalling cascade, differs from the role of the first enzyme in a chain of metabolic reactions. The main difference between an enzyme which is a pacemaker of a chain of metabolic reactions and of a receptor which regulates signal input, is that only in the first case, activation deactivation of an enzyme turns on or shuts off the whole chain of reactions, downstream of the first enzyme. Consider glycogen phosphorylase, which is the first enzyme of the glycolytic chain: phosphorylase feeds the first substrate into the glycolytic chain and when its activity ceases, the downstream metabolic reactions also come to a standstill, because of lack of fuel. On the other hand, receptor activity is often of shorter duration than the cellular response.

A signalling cascade is a steady-state system

A phosphorylation cascade, or quite generally a signalling cascade, can be looked at as a steady-state system with continous inflow and outflow of information. The idea of treating a regulatory cascade as a steady-state system was applied by Earl Stadtman and Boon Chock to the regulation of activity of metabolic enzymes by interconversion, more than 20 years ago.[24] Their farsighted treatment of the regulation of the interconversion of metabolic enzymes by cascades of converter enzymes, each catalysing a discrete covalent modification reaction, foresaw the possibility of driving a signalling cascade which may be regulated separately at many steps, by changing the input of signals and establishing a steady state of activity of the whole chain.

This model requires input of energy for the maintainenance of the steady state and control of the phosphorylated–dephosphorylated state of regulatory molecules, either by energy-consuming phosphorylation reactions or by negative feedback, often, but not exclusively, by dephosphorylation.

Sensitivity of signalling

An attempt was made recently to find out how sensitively the regulatory system that controls the cell cycle and cell proliferation responds to signal input. Ferrell *et al.*[25] found that, in intact oocytes, the response is 'ultrasensitive', a kinetic characterization introduced by Daniel E. Koshland Jr.[26] (Ultrasensitivity has been defined[27] as the response of an enzyme that is more sensitive to changes in the concentration of the substrate than an enzyme with a normal hyperbolic response, according to the Michaelis–Menten equation. One can also use the Hill coefficient (n_H) to indicate hyperbolic (Michaelis–Menten) sensitivity ($n_H = 1.0$), ultrasensitivity ($n_H > 1$), and subsensitivity ($n_H < 1$).[28]

The increase of the Hill coefficient of the oocyte cell cycle in response to the input of growth-promoting signals was very large ($n_H = 5.0$). Thus, the response was highly cooperative and ultrasensitive. The important point is that the ultrasensitive response is a property of the phosphorylation cascade as a whole, because the individual enzymes of the MAP kinase cascade which propagate the input signal, did not show cooperativity. D. E. Koshland Jr expects that studies with intact cells will reveal more such ultrasensitive responses, notably of transcription factors, to small fluctuations of signals, especially when cells differentiate or when they de-differentiate in the course of malignant transformation.

References

1. The reader may want to look up the, 'Protein Kinase Resource (PKR). This is a web-accessible compendium of information on the protein kinase family of enzymes.
2. E. Hafen. Kinases and phosphatases-a marriage is consummated. *Science*, **280**, 1212–1213, 1998. Also: M. Camps, A. Nichols, C. Gillerion, B. Antonsson, M. Muda, C. Chabert, *et al.* Catalytic activation of the phosphatase MKP-3 by ERK2 mitogen-activated protein kinase. *Science*, **280**, 1262–1265, 1998.
3. P. Cohen and P. T. Cohen. Protein phosphatases come of age. *J Biol Chem.*, **264** (36), 21435–21438, 1989. Also: P. Cohen. The structure and regulation of protein phosphatases. *Annu Rev Biochem*, **58**, 453–508, 1989.
4. D. Barford. Molecular mechanisms of the protein serine/threonine phosphatases. TIBS, **21**, 407–412, 1996.
5. M. P. Egloff, P. T. Cohen, P. Reinemer, and D. Barford. Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *Nature*, **376** (6543), 745–753, 1995.
6. C. R. Kissinger, H. E. Parge, D. R. Knighton, C. T. Lewis, L. A. Pelletier, A. Tempczyk, *et al.* Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. *Nature*, **378** (6557), 641–644, 1995.
7. D. Barford, A. J. Flint, and N. K. Tonks. Crystal structure of human protein tyrosine phosphatase 1B. *Science*, **263**, 397–404, 1994.
8. D. Barford. Molecular mechanisms of the protein serine/threonine phosphatases. *J Mol Biol.*, **254**, 942–959, 1995.
9. J. Goldberg, H.-B. Huang, Y.-G. Kwon, P. Greengard, A. C. Nairn, and J. Kuryan. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature*, **376**, 745–753, 1995.
10. J. P. Griffith, J. L. Kim, E. E. Kim, M. D. Sintchak, J. A. Thomson, M. J. Fitzgibbon, *et al.* X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. *Cell*, **82** (3), 507–522, 1995.
11. M. Hubbard and P. Cohen. On target with a mechanism for reversible phosphorylation. *TIBS*, **18**, 172–177, 1993.
12. Y. Nishizuka. The molecular heterogeneity of protein kinase C and its implications for cellular regulation. *Nature*, **334** (6184), 661–665, 1988.

13. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts and J. D. Watson. *Molecular Biology of the cell.*' Garland Publishing, New York, 1994.
14. M. C. Faux and J. D. Scott. More on target with protein phosphorylation: Conferring specificity by location. *TIBS*, **21**, 312–315, 1996.
15. C. F. Cori and G. T. Cori. Polysaccharide phosphorylase. In *Nobel lectures: Physiology or Medicine, (1942–1962)*, pp. 186–206, 1947.
16. P. D. Boyer and E. G. Krebs (ed.). *The Enzymes*, Vol 17 (3rd edn). Academic Press, 1986.
17. D. Barford, S. H. Hu, and L. N. Johnson. Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP. *J Mol Biol*, **218**, 233–260, 1991.
18. R. J. Fletterick. Glyogenphosphorylase: Plasticity and specificity in ligand binding. *Proc Robert A. Welch Found Conf Chem Res*, **27**, 173–220, 1992.
19. D. Palm, H. W. Klein, R. Schinzel, M. Bühner and E. J. M. Helmreich. The role of pyridoxal 5′-phosphate in glycogen phosphorylase catalysis. In: H. Neurath, ed. Perspectives in Biochemistry, vol. 2, Publ. Am. Chem. Soc. pp. 132–140, (1990).
20. J. Monod, J. P., Changeux and F, Jacob. Allosteric proteins and cellular control systems. *J. Mol Biol*, **12**, 88–118, 1963.
21. E. Helmreich, and C. F., Cori. The role of adenylic acid in the activation of phosphorylase. *Proc. Natl. Acad. Sci,* USA, **51**, 131–138, 1964.
23. E. H. Fischer and D. L. Brautigan. A phosphatase by any other name: from prosthetic group removing enzyme to phosphorylase phosphatase. *TIBS*, **7, (1)**, 3–4, 1982.
24. B. P. Chock and E. R. Stadtman. Covalently interconvertible enzyme cascade systems. *Meth Enzymol.*, **64**, 397–425, 1980.
25. C. Y. Huang and J. E. Ferrell Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci USA **93** (19), 10078–10083, 1996. Also: J. E. Ferrell Jr and E. M. Machleder. The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes [see comments]. *Science*, **280** (5365), 895–898, 1998.
26. D. E. Koshland Jr. The era of pathway quantification [comment]. *Science*, **280** (5365), 852–853, 1998.
27. A. Goldbeter and D. E. Koshland Jr. An amplified sensitivity arising from covalent modification in biological systems *Proc Natl Acad Sci USA*, **78** (11), 6840–6844, 1981. Also: D. E. Koshland Jr, A. Goldbeter and J. B. Stock. Amplification and adaptation in regulatory and sensory systems. *Science*, **217** (4556), 220–225, 1982.
28. A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, London, Chapter 1 1995: presents a discussion of enzymes and their role in metabolic pathways.

8

Regulation by a hormone: the insulin response

Is insulin a chimera; a growth factor and a hormone?

The action of insulin has always fascinated biochemists and physicians. The study of insulin and of diabetes, the disease caused by lack of the hormone or by a lack of response to the hormone, is like a mirror, reflecting the advances of biochemistry and medicine. Using the β-cells in the pancreas from animal sources, which were discovered by Langerhans in 1869, Banting and Best isolated insulin in 1921, and Abel crystallized it in 1928. In 1955 the primary sequence of insulin and in 1965 the crystal structure of the biologically active part of insulin which binds to the receptor were elucidated by Frederick Sanger and Dorothy Crowfoot Hodgkin, respectively. In 1965, H. Zahn in Germany, N. Katsoyannis in the USA and Niu Ching-I in Wang's laboratory in Shanghai independently synthesized insulin by chemical means. Today, human insulin is produced using bioengineered *E. coli* as a source. Donald Steiner discovered proinsulin in 1967, and elucidated the processing of pro-insulin, with the formation and the release of the insulin peptide C, in 1972 (Chapter 1). Moreover, the receptor for insulin was the first membrane-bound receptor to be identified (Pedro Cuatrecasas, 1968). But there is hardly any other problem in biological regulation in general, and endocrine regulation in particular, that has been so unyielding and intractable as that of the molecular mechanism of insulin action.

Part 1 concludes with a discussion of the actions of insulin. This provides an opportunity to see how regulatory mechanisms, such as phosphorylation dephosphorylation and control by second messengers, enable the body to respond to a hormone which, in many ways, behaves like a growth factor.

The consequences of a failure of insulin function

Insulin has an essential life-supporting function. It maintains glucose homeostasis. Glucose itself, or a metabolite of glucose, triggers the release of insulin from the insulin-producing β-cells.[1] If the response to insulin fails, or insulin is not supplied sufficiently, or both processes do not function properly, blood glucose regulation fails and the

hyperglycaemia that follows is the cause of a rather common, life-threatening disease, diabetes mellitus.[2] Human type 2 insulin-dependent diabetes mellitus (IDDM) is a condition characterized by an excess of glucose in the bloodstream. Both insulin resistance of insulin-responsive tissues and a lack of an adequate response of the β-cells to glucose, with consequently diminished insulin secretion, are responsible for the disease. While the cause of IDDM is not clear, MODY (maturity-onset diabetes of the young) is an autoimmune disease. It is also characterized by elevated blood glucose, due to a lack of insulin, requiring insulin substitution. Autoaggressive T cells recognize a self-protein, expressed in the β-islets of the pancreas, the enzyme, glutamic acid decarboxylase. The T cells infiltrate the pancreas and eventually destroy the β-islets (Fig. 8.1).

Glucose is toxic: protein glycation

The vital role of insulin is best appreciated when we look at the consequences of a lack of insulin and/or a lack of insulin response, leading to excess blood glucose. Hyperglycaemia is responsible for many manifestations of the disease, because it leads to non-enzymic glycation of proteins, as shown by Anthony Cerami[5] (of course, there are other important manifestations of the disease, such as the dysfunction of fatty acid metabolism and the formation of ketone bodies, which have not been considered here; see P. J. Randle *et al.*[6]).

Proteins with relatively short half-lives are glycated by rather acute increases in blood glucose, whereas chronic, small elevations of blood glucose lead to irreversible chemical modifications of long-lived proteins. These have been called 'advanced-glycation end-products' (AGEs)[7] (Fig. 8.2).

To sum up: the essential life-supporting function of insulin is the maintenance of glucose homeostasis. The question therefore is, how is insulin doing that? We shall focus therefore on transport and metabolism of glucose and its control by insulin. Thus the scope of this chapter is quite narrow, dealing only with insulin actions. Although the classical human type 2 diabetes mellitus is caused by both an impairment of insulin

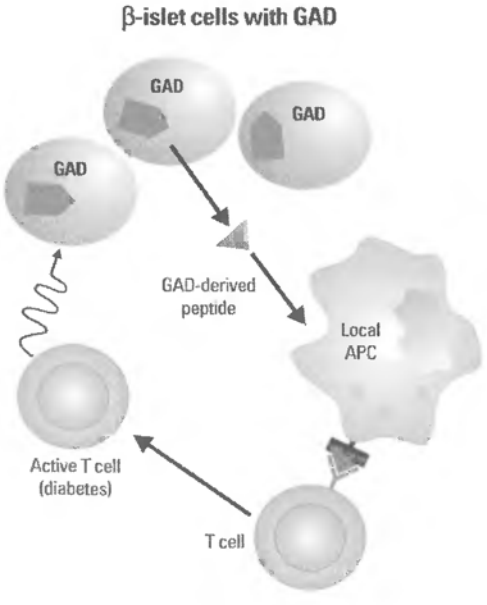


Fig. 8.1 Autoimmune diabetes develops in non-obese diabetic (NOD) mice, when peptides from glutamic acid decarboxylase (GAD) are expressed in the β-islet cells of the pancreas and are presented by local antigen-presenting cells (APC) to T cells (see Chapter 14). (This scheme is reproduced from a drawing in ref. 4, by permission of H. v. Boehmer and Adelaida Sarukhan and Science.)

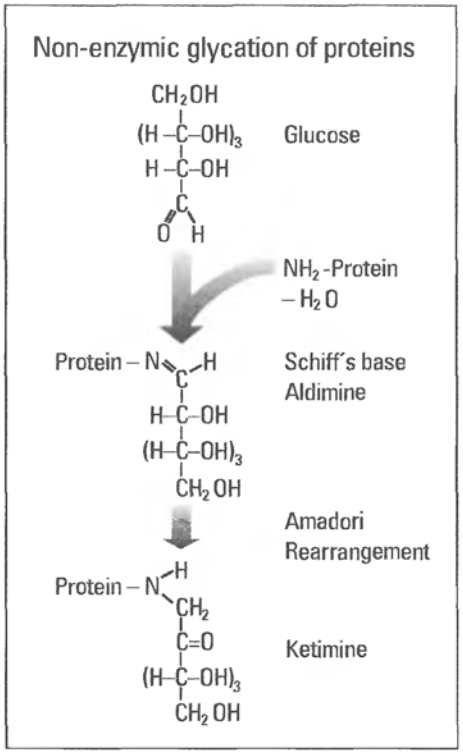


Fig. 8.2 Non-enzymic glycation of proteins. In the non-enzymic chemical modification of proteins by glucose, the sugar is added in a Schiff's base linkage to a free amine of an amino-acid side-chain, such as that of lysine and the adduct is rearranged by the Amadori reaction from an aldimine to a ketimine. Not shown is the subsequent conversion of the product of the Amadori reaction. The ketimine can be rearranged by successive dehydration and by β- elimination to an ene- dion which can react with the amino-acid-side chain of another protein, forming stable protein-protein crosslinks. An example is cross-linking of collagen, which has been related to kidney failure in diabetes. Attempts have been made to break these cross-links with chemicals, which cleave vicinal diketones. 'Advanced-glycation end-products' (AGEs), are bound to RAGE, which is an approximately 50 KD protein. RAGE is a member of the immunoglobulin superfamily and is expressed on the surface of endothelial cells. AGE bound to RAGE damages endothelial cells and interferes with vascular functions, although the signalling pathway activated by RAGE–AGE complexes is not known.[8] (RAGE is also expressed in neuronal cells,[9] and signalling by RAGE has been linked to manifestations of neurotoxicity. RAGE is overexpressed in the brains of Alzheimer patients and the Alzheimer amyloid-β peptide (Aβ) is a ligand of RAGE.[10] Aβ is neurotoxic and plays a role in the pathogenesis of Alzheimer's disease. The formation of Aβ–RAGE complexes in the brain is correlated with an increase of intracellular oxidants, which could upregulate RAGE, because the promoter of the *RAGE* gene contains two NF-κB sites, which respond to stress and oxidants (see Chapter 10.) Vascular complications of diabetes mellitus,[11,12] have also been attributed to an uncontrolled, overactive protein kinase Cβ in vascular smooth muscle, resulting from non-enzymatic glycation of the kinase.

secretion and of insulin actions, we shall not consider the regulation of insulin secretion.[13,14]

Insulin has in common with growth factors that the hormone signals through a tyrosine kinase receptor. But, insulin actions are unique in so far as they appear to fall into two distinct categories. First, fast, immediate effects, primarily on the removal of glucose from the blood and its utilization in the tissues. These effects are linked directly to glucose homeostasis. Secondly, slow, more subtle and prolonged systemic effects, more typical of growth factors, such as stimulation of DNA and protein synthesis, mitogenic effects, and promotion of cell growth and differentiation. The dilemma is that the hormonal metabolic effects and the growth-promoting effects cannot be distinguished as easily as one might expect. Thus, the pivotal question is, how can the role of insulin as a growth factor be reconciled with its function as a hormone, acutely regulating blood glucose? This question can only be answered once we know what the signalling pathways operated by insulin have in common with growth-factor signalling in general, and when we find out where the common signalling pathways, shared by insulin and other growth factors, intersect with the enzymes and proteins regulating specifically sugar uptake and utilization, the control of which is a unique function of insulin. Any discussion of insulin action in the context of growth-factor signalling must focus on these links.

Glucose homeostasis and insulin

Claude Bernard already realized in the middle of the nineteenth century that the liver is responsible for the homeostasis of blood glucose. Soskin and colleagues[15] found in 1938 that in a perfused liver preparation *in situ*, the uptake of glucose by the liver is determined by the concentration of glucose in the blood carried to the liver by the portal vein. When the concentration of glucose was high, it was taken up by the liver, and when the concentration of glucose was low, it was released from the liver into the blood. Thus, uptake and release of glucose by the liver were proportional to the degree of hyper- or hypoglycaemia, respectively.

Does insulin correct the overproduction of glucose in the liver?

Since insulin normalizes blood glucose, it was quite plausible to conclude that insulin acts primarily on the liver and corrects the overproduction of glucose. But Géry Hers[16] found in the 1970s that the control of glucose homeostasis by the liver is primarily a self-regulatory mechanism, carried out by glucose itself. Moreover, this self-regulatory mechanism alone cannot prevent the increase in blood glucose in diabetes. And, indeed, evidence for a role of insulin in the control of glucose homeostasis by the liver has never been unequivocal until today. This state of affairs gave rise to the view, favoured mainly by biochemists, (among them Carl F. Cori) that insulin actually corrects the underutilization rather than the overproduction of glucose. This shifted emphasis from the liver to the muscle, as the main target of insulin (see also ref. 17).

Does insulin correct the underutilization of glucose?

The most dramatic effect of insulin is the hormone-dependent uptake of glucose by muscle, as demonstrated by Levine and Goldstein.[18] This directed attention to the role of insulin on glucose uptake in muscle and adipose tissue, and eventually led to the elucidation of an insulin-responsive glucose-uptake mechanism in these tissues.[19] Furthermore,

it brought the second step to attention, the effect of insulin on the conversion of glucose into glycogen.[20] In all these insulin-responsive reactions, glucose is only the substrate to be utilized; it has no active regulatory role. Thus, the first step in glucose utilization is the uptake of the sugar from the blood and its transfer into the tissues. Insulin can make muscles and the adipose tissue permeable for glucose by recruiting a glucose carrier, capable of rapidly removing sizeable amounts of glucose from the bloodstream and supplying the insulin-responsive tissues with glucose. To the contrary, liver cells are freely permeable to glucose, even in the absence of insulin.

Insulin effects on glucose transport in muscle and adipose tissue

Since the average muscle mass in humans is about 40% of the body mass, and since resting muscle in the absence of insulin is impermeable to glucose, the sometimes too rapid and drastic lowering of glucose in the blood after application of insulin can be accounted for by the action of insulin on the uptake of glucose in muscle. The rapid removal of glucose from the blood is the most impressive effect of insulin. The facilitated insulin-dependent diffusion of glucose in tissues, such as striated muscle and adipose tissue, is accomplished by the GLUT4 glucose transporter.[21] (Muscle work also stimulates uptake of glucose in skeletal muscle by GLUT4, but apparently through a mechanism different from that involving insulin.[22])

Once the vesicles carrying GLUT4 have arrived at the plasma membrane, they must fuse with the membrane. For that purpose the cell uses a machinery, the components of which have been identified in recent years. Among the proteins that participate in the docking process, the SNARE protein family of receptors is of central importance, as was shown in J. Rothman's laboratory (Fig. 8.3).[23]

A search is now under way to find a target among the components of the SNARE machinery, which responds to the insulin signal and promotes insulin-dependent glucose transport. Although the problem is far from being solved, some interesting information has already been obtained by Alan R. Saltiel and his laboratory.[25] They found that in adipocytes insulin stimulates phosphorylation of a caveolar protein, caveolin which lines the cytoplasmic surface of caveolae.[26] It was suggested that in adipocytes, a phosphorylation cascade transmits the insulin signal to the glucose uptake mechanism via caveolin and caveolae. In adipocytes, which are highly responsive to insulin, a coupling partner for the insulin receptor was identified, CBL, the product of *cbl*, a proto-oncogene. c-*cbl* is related to v-cbl, the oncogene of the Cas-NS-1 retrovirus. (Cas is the abbreviation of Casitas B-lineage lymphoma.) The product of c-*cbl* is CBL. CBL is a linker. An adaptor CAP (the CBL-associated protein) binds to CBL and brings it to the insulin receptor, where it is phosphorylated by the receptor tyrosine kinase. The tyrosine-phosphorylated CBL then activates the Src-family tyrosine kinase, Fyn, which phosphorylates caveolin and helps to bring the glucose carrier to the membrane. Caveolae have a specific lipid composition, different from the bulk of the membrane, which is characterized by a core of cholesterol, glycosphingolipids, and sphingomyelin. They form in the Golgi and are shipped to the cell surface (Fig. 8.4). Caveolae are enriched in receptors and signal transducers. They function like rafts,[27] and collect and concentrate their cargo in the lipid bilayer. In that way, they could facilitate the assembly of signal chains in the membrane.

PtdIns-3,4,5-P_3, which is implicated as a second messenger in insulin action (see below), has been reported to activate membrane fusion of vesicles containing glucose transporters, but the molecular details have not yet been worked out. But it is to be expected

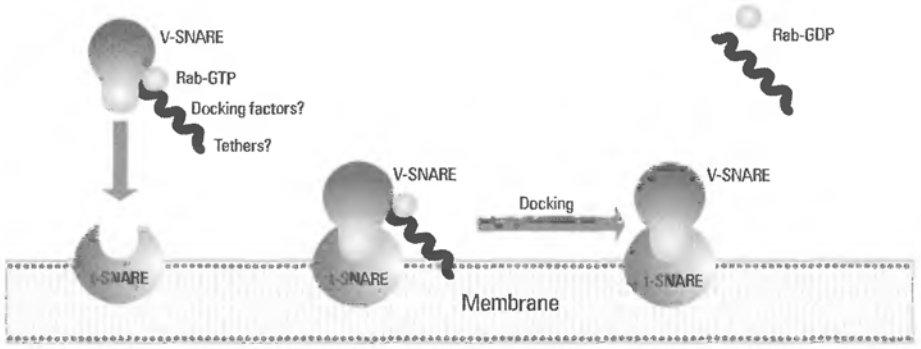


Fig. 8.3 v-SNAREs and t-SNAREs are the receptors on vesicles and organelles and on the target membrane, respectively. Docking factors are recruited to nascent transport vesicles, containing the v-SNARES (transport vesicles are derived from an organelle, such as the Golgi, or are vesicles destined for secretion). This process is controlled by the monomeric, Rab G-protein in its activated, GTP-bound form. The scheme shows how such a transport vesicle with v-SNAREs is targeted to the plasma membrane. The docking site in the membrane is recognized by the vesicle receptors, the v-SNAREs. Target specifiers may help to dock the vesicle with the v-SNAREs in the membrane. Additional factors then activate the docking machinery which links the v-SNAREs to the t-SNAREs on the target site on the membrane. Once docking is completed, the docking factors, targeting specifiers, and Rabs are released. Three distinct entities of the SNARE machinery are required: one is NSF, the *N*-ethylmaleimide-sensitive fusion protein (Sec18p in yeast) which is an ATPase and provides the energy. The others are the SNAPs (soluble NSF attachment proteins; Sec17p in yeast), and the membrane-associated SNAP-receptors, named SNAREs. A transport vesicle is loaded with proteins in one cellular compartment and then moves through the cytoplasm until it eventually reaches a docking site on another vesicle or the plasma membrane. The membranes of the transport vesicle and the target compartment fuse with each other and the cargo, together with the membrane of the transport vesicle, is incorporated into the target. (This scheme is based on Fig: 1 in ref. 24, with permission of Nature Cell Biology.)

that as the instruments which regulate protein traffic in the cell in general, and endo- and exocytosis in particular, are better characterized, we shall eventually also understand the actions of insulin on sugar transport in muscle and adipose tissues. The uptake of glucose is followed by its utilization, notably its conversion into glycogen. This seems to be a major point of insulin control. Many important advances have been made over the years in the control of this pathway by insulin, mainly by Philip Cohen and his laboratory.[28] In the following, we summarize what we know about this insulin-regulated phosphorylation cascade.

Insulin-signalling pathways

Some of the components participating in the insulin-regulated kinase cascade, play also a role in signalling by other growth factors and cytokines. A common reservoir of tools, shared by other growth factors, strengthens the argument that the action of insulin on

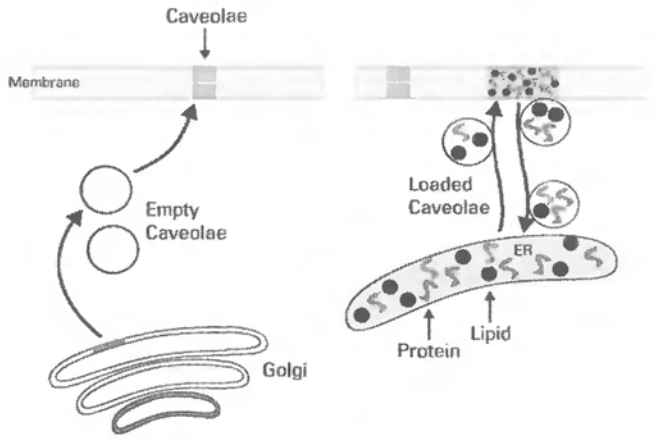


Fig. 8.4 The function of 'caveolae'. On the left: once the caveolae have arrived in the membrane, they begin to transport cholesterol and other lipids from the endoplasmic reticulum to the plasma membrane and back (shown on the right). (Reproduced from Fig. 3, in ref. 26, with permission of the author and Annu Rev. Biochem.)

glucose transport and glycogen metabolism is a variation of growth-factor-like signalling, but it also makes it more difficult to find the insulin-specific features.

Receptor substrates and the PI 3-kinase pathway

The insulin receptor is the prototype of a tyrosine kinase receptor with a constitutive, oligomeric structure (Chapter 1). The receptor has been cloned and characterized in detail, notably by Axel Ullrich, R. C. Kahn and colleagues, see: Chapter 1.[29] Binding of insulin stimulates the intrinsic receptor tyrosine kinase. A crystal structure of the tyrosine kinase domain of the insulin receptor was solved.[30] This leads to phosphorylation of tyrosine residues and to the recruitment and subsequent phosphorylation of substrates.

A quite unique feature of signalling through the insulin receptor is the participation of receptor-specific substrates.[31,32] The need for substrates is thought to reflect the weak interactions of the phosphorylated insulin receptor with SH2 domains of docking proteins. Consequently, the insulin receptor can not directly recruit adaptors and transducers, as do EGF and other growth-factor tyrosine kinase receptors. This is supported by the properties of a mutated insulin receptor which cannot interact with insulin receptor substrates (IRSs). This mutant receptor could not signal, although it was still autophosphorylated.

There are two major insulin receptor substrates, IRS-1 and IRS-2. IRS-1 has a (pleckstrin-like) homology domain, $IH1^{PH}$, and a phosphotyrosine-binding domain, $IH2^{PTB}$. $IH2^{PTB}$ binds to the activated, phosphorylated insulin receptor in a juxtamembrane region, where the crucial phosphotyrosyl residue 960 is located. (The crystal structure of the $IH2^{PTB}$ domain, alone and complexed with the juxtamembrane region of the insulin receptor, has been solved.[33])

The phenotype of mice deficient in IRS-1 indicated that other docking molecules and substrates must be involved in insulin signalling.[34] IRS-1-deficient mice were stunted in their growth and had impaired glucose tolerance, but the absence of IRS-1 did not completely eliminate the response to insulin. The need for another IRS also became apparent from the phenotypes of mice carrying insulin receptor mutants.[35–36] A more likely candidate as primary substrate of the insulin receptor is IRS-2. In contrast to mice deficient in IRS-1, mice that lack the insulin receptor substrate-2 gene are hyperglycaemic and

have a phenotype resembling human type 2 insulin-dependent diabetes mellitus.[38] Thus one might conclude that IRS-1 is either a universal linker or adaptor that participates in other growth-factor and cytokine receptor-mediated signalling pathways, or is a linker for an as yet not defined pathway for insulin signals, whereas IRS-2 seems to be more directly involved in insulin signalling. However, it should also be considered that mice might not be an adequate model for a human disease such as type 2 diabetes. We shall be confronted with that problem again in Part 4, when we discuss mice models for inherited cancer.

But there may be other insulin receptor substrates, which contribute to the pleiotropy of insulin effects. CBL in adipocytes has already been mentioned. Tyrosine-phosphorylated IRS-2 and other IRS-like proteins could recruit through their SH2 domains Grb2, which in turn binds through its SH3 domain the guanine nucleotide exchange factor, Sos, which helps to form the activated GTP-bound form of Ras. Thus, insulin signals might be transmitted through the Ras/MAP kinase or JNK pathway,[39] but connections of insulin to this or the JAK/STAT pathway or any other common cellular signalling route are rather ambiguous.

The most convincing evidence, presently available, points to a participation of the PtdIns 3-kinase pathway in insulin signalling. The PtdIns 3-kinase pathway and a phosphorylation cascade, possibly involved in insulin signalling, is shown in Fig. 8.5.

Protein kinase B, PKB has a PH domain to which PI- 3,4,5-P_3 binds (not shown in Fig. 8.5). Binding of PI- 3,4,5-P_3 converts PKB through a conformational change to a substrate of the upstream PDK kinase(s). Mutants of PKB with replacement of Arg25 and Trp99 in the PH-domain no longer bind the 3,4,5-phosphorylated PtdIns. Philip Cohen's laboratory has purified and characterized an upstream 3,4,5-phosphoinositide-dependent protein kinase (PDK), which phosphorylates and activates PKB.[40] PKB has two regulatory sites, Thr308 and Ser473.[42] PDK is likewise activated by PtdIns-3,4,5-P_3 and activates PKB by phosphorylation of these sites.[42] PKB activated on phosphorylation is released from the membrane. IP_3 has been implicated in the release of PKB from the membrane. PKB phosphorylates and inhibits glycogen synthase kinase-3 (GSK-3).[43] Consequently, the glycogen synthase *a* stays in the active, non-phosphorylated state and stimulates glycogen synthesis. PKB itself is shut off by dephosphorylation by protein phosphatase 2A, PP2A. The PDK/PKB phosphorylation cascade stimulates glycogen synthesis.

But it also is responsible for the growth-factor-like effects of insulin, promoting growth and proliferation and suppressing apoptosis. PKB may be involved in other cellular responses. It may activate the ribosomal $p70^{S6}$ kinase and stimulate protein synthesis. This would explain the stimulation of protein synthesis by insulin. Another important target of the PKB phosphorylation cascade may be glucose transport. This would be in accordance with the stimulation of glucose transport by insulin in skeletal muscle and adipocytes. But the role of PKB in the activation of these pathways is not yet clear.

To sum up: insulin signalling is initiated by activation and autophosphorylation of the insulin tyrosine kinase receptor upon binding the hormone. But whether the insulin receptor is primarily responsible for the lack of response to insulin in IDDM is questionable. The receptor phosphorylates its substrate and recruits the PtdIns 3-kinase, which forms PtdIns 3,4,5-P_3 (PIP_3) which activates the kinases PDK and PKB. PKB phosphorylates and inactivates GSK3, glycogen synthase kinase-3, thus promoting glycogen syn-

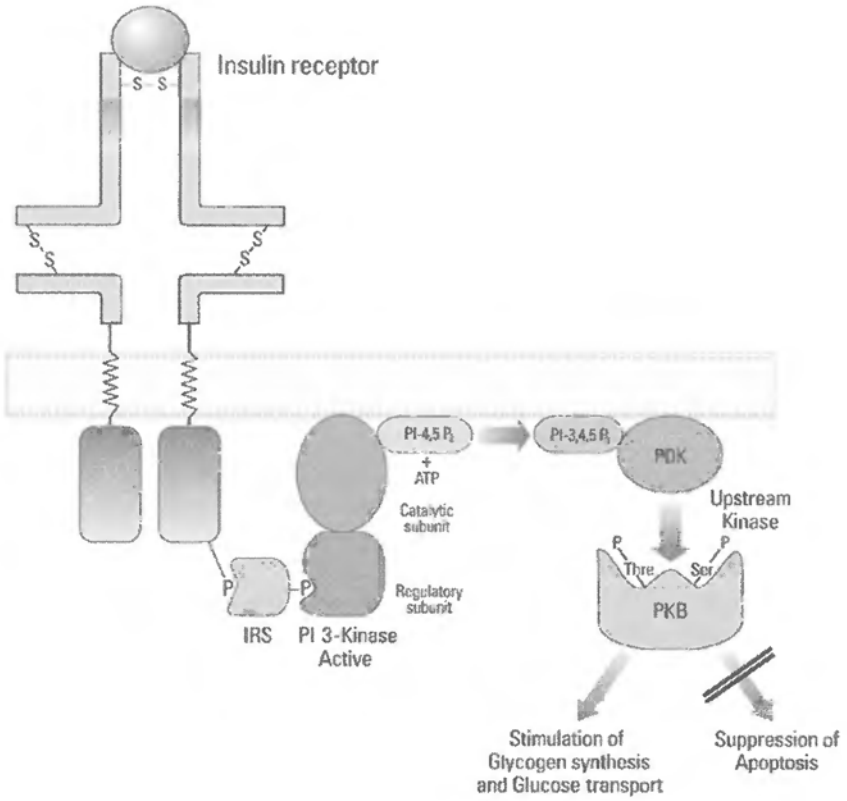


Fig. 8.5 PI is phosphatidylinositol. The PI_3-kinase docks on the receptor–substrate complex. The activated PI 3-kinase catalyses the synthesis of the second messenger PtdIns-3,4,5-P_3 from PtdIns-4.5-P_2 and ATP. The level of PtdIn-3,4,5-P_3 is controlled by PtdIns phosphatase which reverses the PI 3-kinase reaction (not shown). PI_3-kinase has two subunits, a catalytic subunit and a regulatory subunit. The latter has two SH2 domains, through which it interacts with the tyrosine-phosphorylated IR substrate, IRS-P. The interaction with the activated IR–IRS complex turns on the catalytic subunit of the PI 3-kinase. Inhibition of PI 3-kinase by a specific inhibitor, 'wortmannin', blocks insulin effects on glycogen metabolism. The downstream kinase, PKB (or Akt) has a central role in a phosphorylation cascade, initiated by an insulin-mediated activation of PI 3-kinase (According to ref. 41).

thesis. The final outcome is the formation of an active (dephosphorylated) form *a* of glycogen synthase which synthesizes glycogen, while at the same time, glycogen phosphorylase *b*-remains dephosphorylated and inactive, so that the breakdown of glycogen is turned off, preventing a futile cycle (see, Chapter 7). However, we have to keep in mind that PKB is activated not only by insulin, but is also activated rapidly (and nearly 50-fold) by growth factors such as PDGF, EGF, basic FGF and others. Thus, the main problem at present is to find out where signals from growth-factor receptors and signals from the insulin receptor separate and go their own ways, and how insulin instructs PKB to phosphorylate and inactivate GSK3 and stimulate glycogen synthesis.

Insulin could also promote glycogen synthesis by *activating* glycogen synthase phosphatase, rather than *inactivating* glycogen synthase kinase.[44] However, a pathway linking insulin receptor activation to the activation of serine/threonine phosphatases of the PP1 type has not yet been identified.

Cross-talk between insulin, TNF-α, and leptin

A characteristic feature of insulin signalling is its versatility and plasticity, due to the interconnectivity of the insulin signalling pathway with signalling pathways of growth factors and cytokines. To account better for the enormous versatility, R. Kahn prefers the concept of a signalling network for the propagation of insulin signals, rather than that of a signalling cascade.

Since the insulin signal travels over pathways shared with growth factors and cytokines, mutual interference of insulin and growth-factor signalling is actually expected. This kind of cross-reaction has been implicated in the downregulation of insulin activity in obese patients. Obesity is often associated with insulin resistance in late-onset diabetes mellitus of the aged. When normal mice are overfed, they have high levels of serum fatty acids and become insulin resistant and diabetic. Thus, obesity is a big risk factor for diabetics. One should remember that the only treatment available in the pre-insulin era was starvation. Several factors have been implicated in insulin resistance in obese rodents and humans. One is TNF-α, which is overexpressed in adipose tissue of obese rodents and humans. The cross-talk of insulin with TNF-α is an example of pathway interconnectivity. TNF-α interferes with the interaction of the insulin receptor with IRS-1.[45] Obese mice with loss of function mutations, in the gene encoding TNF-α and/or in the genes encoding the two TNF-α receptors, TNFR-I and TNFR-II, had an improved insulin response.[46] But, the mutual interference of insulin with TNF-α signalling is not the only cause of insulin resistance and obesity. Another factor that has been implicated is the peroxysome proliferator-activated receptor (PPAR αγ), a key regulator of adipocyte differentiation (Chapter 10).

PPARs are activated by a variety of fatty acids, and response elements addressed by PPARs have been located in the promoter regions of genes encoding enzymes participating in fatty-acid oxidation. Mutations of the human gene for PPARγ had effects suggesting a role of PPARγ in insulin resistance and obesity,[47] whereas uncoupling of obesity from insulin resistance through a targeted mutation in a gene encoding the adipocyte fatty-acid binding protein, aP2, indicated a link between fatty-acid metabolism, obesity, and insulin resistance.[48]

Of special interest in the control of obesity are factors that regulate appetite and satiety.

Insulin resistance in obesity: the role of leptins

Regulation of the body weight, is of great medical interest, because of the mutual interdependency of obesity and insulin resistance and its contribution to other systemic disease states. see ref. 49. Among the many factors that are involved in the complicated control of appetite and food consumption and the control of body weight, the leptins have received much attention.

Leptin is a small, 16 kDa cytokine, expressed in fat cells. Two severely obese children with very low serum leptin levels, who are members of a family with a highly consanguineous, pedigree were found to have a homozygous frame-shift mutation of the leptin gene.[50] Although in this rare human genetic disorder leptins seem to be important for the control of body weight, they are certainly not the only factors controlling body weight. Whether leptins are effective in the control of appetite in grossly obese persons who

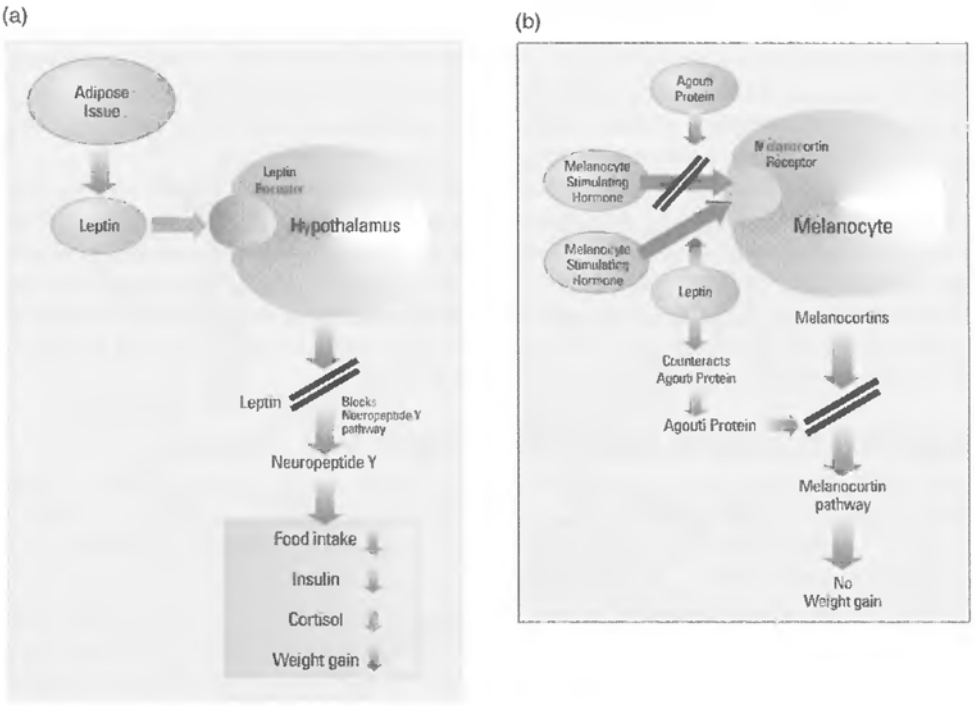


Fig. 8.6 (a) Leptin originates from adipose tissue and acts on the hypothalamus, where it inhibits synthesis and release of the neuropeptide Y in the brain. Neuropeptide Y (NPY) is a hypothalamic neurotransmitter. Leptin binds to a high-affinity receptor in the hypothalamus and downregulates the release of NPY. Mice with a disrupted leptin gene have very high levels of NPY. The consequences are increased food intake, increased insulin requirement, increased cortisol production, and weight gain. Leptin keeps the neuropeptide Y in check, but NPY is not the only factor involved in leptin regulation of obesity. (b) Leptin also prevents overweight through its effect on the melanocortin pathway. This effect is independent of NPY. Leptin relieves blockage of the melanocortin-4 receptor by a protein, named the 'agouty' protein. In the brain, melanocortins suppress appetite. But when the melanocortin receptor is blocked by the agouti protein, appetite is stimulated and weight is gained. Leptin counteracts the agouty protein and re-establishes the melanocortin pathway. (Reproduced from *Science*, ref. 51 with permission of the journal.)

overeat and consume excess food is doubtful. In addition to a lack of leptin, obesity might be due to dysfunction of other factors, to overproduction of neuropeptide Y (NPY), to a block in the melanocortin pathway, and/or to interference with insulin signalling by TNF-α or other neurotrophic factors that regulate the appetite in the hypothalamus (Fig. 8.6).

Recently, an uncoupling protein which is a member of the mitochondrial carrier family,[52] was shown to affect the metabolic rate and glucose homeostasis in mice. Transgenic mice that overexpress a human uncoupling protein in skeletal muscle eat more but weigh less than wild-type mice. These mice also have lower blood glucose and insulin levels and clear glucose faster from the blood.[53] These observations point to a central role of thermogenesis in weight control.

Gene transcription

Several insulin-responsive *cis*-acting gene elements have been identified, referred to as insulin response elements (IREs). The problem is, that the IREs include so many common targets, such as the serum response element (SRE) and immediate early genes, controlled by many growth factors. Thus, activation of these genes is not specific for insulin.

Caenorhabditis elegans has a specific insulin-like signalling pathway which controls transcription of key metabolic and developmental genes. The master switch of transcriptional activation in the worm is the DAF-16 transcription factor.[54] DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. A human homologue of DAF-16 has been implicated in transcriptional control of insulin-responsive genes. Lack of control of insulin-responsive genes has been thought to be responsible for some cases of NIDDM, the non-insulin-dependent form of late-onset diabetes.[55]

Signalling through an insulin-like receptor in *C. elegans*

In *C. elegans*, the *daf-2* gene codes for a homologue of the insulin receptor.[56] This insulin-like receptor is activated by an unknown, presumably insulin-like, ligand. The DAF-2 pathway cross-reacts synergistically with a pathway activated by DAF-7, a TGF-β-like protein (Chapter 6). This pathway controls transcription through activation of the DAF-4/DAF-1 receptor and phosphorylation and activation of SMADs. On the other hand, the DAF-2 receptor with its insulin-like ligand recruits a phosphatidylinositol phosphate kinase, AGE-1, which catalyses the synthesis of the lipid messenger, Ptdlns-3,4,5-P_3, initiating an insulin-like signalling cascade in the worm (Fig. 8.7).

This pathway determines the fate of the worm: when overpopulation leads to scarcity of food, *Caenorhabditis* stops feeding and a pheromone forces the worms to enter the 'Dauer' phase. This state of longevity can last for 2 months or longer. On the other hand, when food is plenty, the worms die after about 2 weeks. In this situation, in the absence of the pheromone, an insulin-like ligand activates the DAF-2 receptor and initiates an insulin-like signalling pathway and shuts off the longevity pathway. When DAF-2 signalling is interrupted and the insulin-like signalling pathway is shut off, the DAF-16-regulated longevity pathway takes over, increasing the life span of the worm. This has led to conjecture that the increase in the life span of *C. elegans* is analogous to mammalian longevity due to caloric restriction, which, likewise, reduces the insulin requirement.[58] Humans can endure starvation to some extent, thanks to gluconeogenesis, the formation of glucose from sources other than carbohydrates. Gluconeogenesis maintains the homeostasis of blood glucose in starvation. This is an evolutionary heritage which enabled humans to survive as hunters and gatherers, with sporadic periods of starvation and irregular, intermittent feeding. The more we became dependent on regular food intake, the greater became our dependency on insulin. It was suggested that famines in human history may have selected for human homologues of the *C. elegans daf* genes, the *db* or *ob* obesity genes.[59] Obese mice, expressing either *db* or *ob* genes, survive fasting better than wild-type controls. Such homologues of the *daf* genes in humans might have become recessive diabetes genes. The high frequency of late-onset diabetes in human populations may reflect selection of such 'diabetes' genes.

This is not the only example of a beneficial genetic endowment that can become a genetic burden under drastically changed conditions. I direct attention to the protection

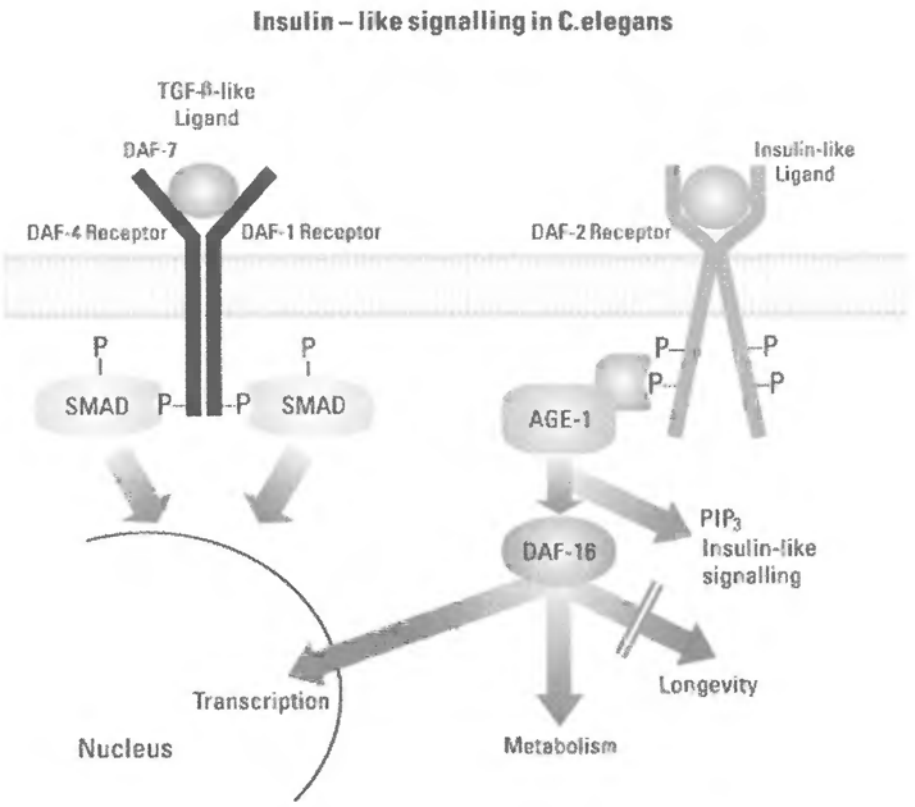


Fig. 8.7 A model of insulin-like signalling in *C elegans*. In the absence of the longevity pheromone, an unknown insulin-like ligand activates DAF-2, whereas a TGF-β-like signalling molecule, DAF-7, activates the DAF-1 and DAF-4 receptors. The activated, tyrosine phosphorylated DAF-2 receptor recruits AGE-1, and AGE-1 downregulates DAF-16, the master switch that controls transcription in the worm.[50] DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. AGE-1 is a PtdIns P_2, PIP_2 kinase that produces the second messenger, phosphatidylinositol-3,4,5-trisphosphate (PIP_3). Signalling by PIP_3 is analogous to signalling in insulin-reponsive mammalian and human tissues. As in mammals the second messenger, PIP_3, may regulate glucose uptake and metabolism in *C. elegans*. This model for regulation of the 'Dauer formation' pathway has been published in Fig. 4 of ref. 56,57. (It is reproduced with permission of the authors and Science.)

that glucose- 6-phosphate dehydrogenase deficiency in red blood cells (a sex-linked trait) can give a person against the dangerous falciparum malaria. However, the other side of the coin is the susceptibility to haemolysis after exposure to certain substances and food ingredients. Another case is the protection against malaria of people who carry a mutant sickle-cell haemoglobin in their red bood cells which, however, is shared with the complications arising at reduced oxygen tension.

Conclusions

The case for insulin being a growth factor, functioning as a hormone, is persuasive, but not proved. The identification of the steps which link insulin binding, activation and phosphorylation of its receptor and transfer of the insulin signal to changes in the activity of rate-limiting processes in the uptake and utilization of glucose, has proved to be difficult.[60,61] But, there is no doubt that the rather recent identification of insulin-responsive linkers, kinases, and phosphatases, and the role of lipid messengers has impressively advanced our understanding of insulin action. A further advance is the realization that different insulin-responsive signalling pathways may be compartmentalized in the cell.[62] (The role of compartmentalization in cellular regulation was discussed in Chapter 7).

If one assumes that insulin is a growth factor, because all the effects, as diverse as they are, originate from interactions of the hormone with its receptor, which is an RTK like other growth-factor receptors, one would attribute the different manifestations of insulin action in different cells to individual regulatory programmes in each of the terminally differentiated insulin-responsive cells and tissues in the body. Rapid and large effects in muscle, and slower and smaller effects in liver, could reflect differences in the properties, and in the response of the receptors and the downstream cellular signalling chains. Different responses of cells in differentiated tissues and bifurcations in signalling pathways might explain why both growth-factor-like and hormone-like effects occur in one and the same organ.

But, be that as it may, the quintessential function of insulin is to maintain blood glucose homeostasis. If insulin fails, blood glucose regulation fails and the hyperglycaemia that follows is responsible for the pathology of diabetes. Thus, the central problem is to find the bifurcation of a common growth-factor-like signalling pathway and a singular insulin pathway which directs the hormonal signal to the proteins responsible for the transfer of glucose from the blood into tissues and to the enzymes catalyzing the subsequent utilization of glucose and its deposition into glycogen. There is hope that this may be accomplished in the not too distant future. Once that goal is reached, it might turn out that the mechanisms and instruments of the insulin-signalling cascade are, in principle, the same as those of other growth-factor-dependent signalling cascades. Growth-factor-like and hormone-like effects could be two sides of the same coin. Thus, we need not yet give up hope that eventually a skeleton mechanism may evolve that can, in principle, explain the pleiotropic actions of insulin. But, whether the appealing hypothesis that insulin improves the utilization of glucose by a receptor-triggered phosphorylation–dephosphorylation cascade is correct, can only be decided when the missing links have been found and the details of these pathways are completely clarified.

References

1. P. Maechler and C. B. Wollheim. Mitochondrial glutamate acts as a messenger in glucose-induced insulin exocytosis. *Nature*, **402**, 685–689, 1999.
2. E. R. Froesch, J.-P. H. Assal, and C. Liniger. Diabetes mellitus: physiology and pathophysiology of glucose, fat and ketone metabolism. In Clinical Endocrinology. Theory and Practice, (ed.) A. Labhart, pp: 749–860. Springer Verlag, Berlin, Heidelberg, 1986.
3. H. von Boehmer and A. Sarukhan. GAD, a single autoantigen for diabetes. *Science*, **284**, 1135–1137, 1999. Also: J.-W. Yoon, C.-S. Yoon, H.-W. Lim, Q. Q. Huang, Y. Kang, K. H. Pyun, *et al.* Control of autoimmune diabetis in non-obese diabetic (NOD), mice by GAD. Expression or suppression in β cells. *Science*, **284**, 1183–1187, 1999.

4. H. von Boehmer and A. Sarukhan. GAD, a single autoantigen for diabetes. *Science*, **284**, 1135–1137, 1999.
5. V. M. Monnier, R. R. Kohn and A. Cerami. Accelerated age-related browning of human collagen in diabetes mellitus. *Proc Natl Acad Sci USA*, **81**, 583–587, 1984.
6. P. J. Randle. α-ketoacid dehydrogenase complexes and respiratoty fuel utilisation in diabetes. *Diabetologia*, **28**, 479–484, 1985. Also: P. J. Randle, P. B. Garland, C. N. Hales, and E. A. Newsholme. *Ciba Foundation Colloquia Endocrin*, **15**, 1966.
7. K. Drickamer. Breaking the curse of the AGEs, News and views. *Nature*, **382**, 211–212, 1996. Also: S. Vasan, X. Zhang, X. Zhang, A. Kapurniotu, J. Bernhagen, S. Teichberg, *et al.* An agent cleaving glucose-derived protein crosslinks *in vitro* and *in vivo*. *Nature*, **382**, 275–278, 1996.
8. A. M. Schmidt, M. Vianna, M. Gerlach, J. Brett, J. Ryan, J. Kao, *et al.* Isolation and characterization of two binding proteins for advanced glycosylation end products from bovine lung which are present on the endothelial cell surface. *J Biol Chem*, **267**, (21), 14987–14997, 1992. Also: M. Neeper, A. M. Schmidt, J. Brett, S. D. Yan, F. Wang, Y. C. Pan, *et al.* Cloning and expression of a cell surface receptor for advanced glycosylation end products of proteins. *J Biol Chem*, **267**, (21), 14998–15004, 1992.
9. O. Hori, J. Brett, T. Slattery, R. Cao, J. Zhang, J. X. Chen, *et al.* The receptor for advanced glycation end products (RAGE) is a cellular binding site for amphoterin. Mediation of neurite outgrowth and co-expression of rage and amphoterin in the developing nervous system. *J Biol Chem*, **270**, (43), 25752–25761, 1995.
10. Sh. D. Yan, X. Chen, J. Fu, M. Chen, H. Zhu, A. Roher, *et al.* RAGE and amyloid-β peptide neurotoxicity in Alzheimer's disease. *Nature*, **382**, 685–691, 1996.
11. M. Brownlee. In *Diabetes mellitus, theory and practice*, (ed.) H. Rifkin and D. Porte pp. 279–291. Elsevier, New York, 1990.
12. M. J. Stevens, E. L. Feldman, and D. A. Greene. The aetiology of diabetic neuropathy: the combined roles of metabolic and vascular defects *Diabet Med*, **12**, (7), 566–579, 1995.
13. F. M., Matschinsky. A lesson in metabolic regulation inspired by the glucokinase glucose sensor paradigm. *Diabetes*, **45**, 223–241, 1996.
14. C. B. Wollheim, J. Lang, and R. Regazzi. The exocytotic process of insulin secretion and its regulation by Ca^{2+} and G proteins. *Diabetes Rev*, **4**, 276–297, 1996.
15. S. Soskin, R. Levine. Carbohydrate Metabolism. Chicago, University Chicago Press, pp. 315, 1946.

    Also: S. Soskin. *Endocrinology*, **26**, 297–308, 1940.
16. H.-G. Hers. The control of glycogenmetabolism in the liver. *Annu Rev Biochem*, **45**, 167–190, 1976.
17. E. A. Newsholme and C. Start. *Regulation in metabolism*. J. Wiley and Sons, Chichester, pp. 247–292, 1981.
18. R. Levine and M. S. Goldstein. On the mechanism of action of insulin. *Rec Progr Hormone Res*, **11**, 343, 1955.
19. S. W. Cushman and L. J. Wardzala. Potential mechanism of insulin action on glucose transport in the isolated rat adipose cell. Apparent translocation of intracellular transport systems to the plasma membrane. *J. Biol. Chem.* **255**, 4758–4762, 1980.

    K. Suzuki and T. Kono. Evidence that insulin causes translocation of glucose transport activity to the plasma membrane from an intracellular storage site. *Proc. Natl. Acad. Sci. USA*, **77**, 2542–2545, 1980.
20. D. A. E. Cross, D. R. Alessi, P. Cohen, M. Andjelkovic and B. A. Hemmings. Inhibition of glycogen synthase kinase 3 by insulin mediated by protein kinase B. *Nature*, **378**, 785–789, 1996.
21. C. M. Wilson and S. W. Cushman. Insulin stimulation of glucose transport activity in rat skeletal muscle. Increase in cell surface GLUT-4 as assessed by photo labelling. *Biochem. J.*, **299**, 755–759, 1994.
22. S. Lund, G. D. M. Holman, O. Schmitz, and O. Pedersen. Contraction stimulates translocation of glucose transporter GLUT4 in skeletal muscle through a mechanism distinct from that of insulin. *Proc Natl Acad Sci USA*, **92**, 5817–5821, 1995.
23. T. Sollner, S. W. Whiteheart, M. Brunner, H. Erdjument Bromage, S. Geromanos, P. Tempst, and J. E. Rothman. SNAP receptors implicated in vesicle targeting and fusion [see comments]. *Nature*, 362 (6418), 318–324, 1993. Also: J. E. Rothman. Mechanisms of intracellular protein transport. *Nature*, **372** (6501), 55–63, 1994.
24. S. R. Pfeffer. Transport-vesicle targeting: Tethers before SNARE's. *Nature Cell Biology*, **1**, E17–E21, 1999.
25. C. C. Mastick, M. J. Brady, and A. R. Saltiel. Insulin stimulates the tyrosine phosphorylation of caveolin, *J Cell Biol*, **129**, 1523–1531, 1995.
26. R. G. W. Anderson. The caveolae membrane system. *Annu Rev. Biochem*, **67**, 199–225, 1998.
27. K. Simons and E. Ikonen. Functional rafts in cell membranes. *Nature*, **387**, 569–572, 1997.
28. P. Cohen. Dissection of the protein phosphorylation cascades involved in insulin and growth factor action. *Biochem Soc Trans*, **21**, 555–567, 1993.

29. A. Ullrich, J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petruzzelli, T. J. Dull, *et al.* Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature*, **313** (6005), 756–761, 1985.

30. S. R. Hubbard, L. Wei, L. Ellis and W. A. Hendrickson. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*, **372**, 746–754, 1994.

31. M. G. Myers and M. F. White. New frontiers in insulin receptor substrate signalling. *TEM*, **6**, 209–215, 1995.

32. M. F. White. The IRS signalling system in insulin and cytokine action. *Phil Trans Royal Soc London B, Biol Sci*, **351**, 181–189, 1996.

33. M. J. Eck, S. D. Dhe-Paganon, T. Trüb, R. Nolte, and S. E. Shoelson. Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell* **85**, 695–705, 1996.

34. M. E. Patti and C. R. Kahn. Lessons from transgenic and knockout animals about noninsulin dependent diabetes mellitus. *TEM*, **7**, 311–319, 1996.

35. R. L. Joshi, B. Lamothe, N. Cordonnier, K. Mesbah, E. Monthioux, J. Jami, and D. Bucchini. Targeted disruption of the insulin receptor gene in the mouse results in neonatal lethality. *EMBO J*, **15**, 1542–1547, 1996.

36. D. Accili, J. Drago, E. J. Lee, M. D. Johnson, M. Cool, P. Salvatore, *et al.* Early neonatal death in mice homozygous for a null allele of the insulin receptor gene. *Nature, Genet*, **12**, 106–109, 1996.

37. P. De Meyts and K. Seedorf. The mechanism of insulin receptor binding, activation and signal transduction. In *Contributions of physiology to the understanding of diabetes* (ed. G. R. Zahnd and C. B. Wollheim), pp. 89–107. Springer, Berlin-Heidelberg-New York, 1997.

38. D. J. Withers, J. S. Ciutierrez, H. Towery, D. J. Burks, J. M. Ren, S. Previs, *et al.* Disruption of IRS-2 causes type 2 diabetes in mice. *Nature*, **391**, 900–904, 1998.

39. S. Giorgetti-Peraldi, F. Peyrade, V. Baron, and E. Van Obberghen. Involvement of Janus kinases in the insulin signalling pathway *Eur J Biochem*, **234**, 656–660, 1996.

40. D. R. Alessi, S. R. James, C. P. Downes, A. B. Holmes, P. R. J., Gaffney, C. B., Reese and P. Cohen. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Bα. *Curr Biol*, **7**, 261–269, 1997.

41. B. A. Hemmings. Signal transduction. Akt signalling: Linking membrane events to life and death decisions. *Science*, **275**, 628–630, 1997.

42. B. A. Hemmings. PtdIns(3,4,5,) P_3 gets its message across. *Science*, **277**, 534, 1997. Also: D. Stokoe, L. R. Stephens, T. Copeland, P. R. J. Gaffney, C. B. Reese, G. F. Painter, *et al.* Dual role of phosphatidylinositol-3,4,5-trisphosphate in the activation of protein kinase B. *Science*, **277**, 567–570, 1997.

43. C. Sutherland and P. Cohen. The a-isoform of glycogen synthase kinase-3 from rabbit skeletal muscle is inactivated by p70 S6 kinase or MAP kinase-activated proteinkinase I in vitro. *FEBS Lett*, **338**, 37–42, 1994.

44. A. R. Saltiel. The paradoxical regulation of protein phosphorylation in insulin action. *FASEB J*, **8**, (13), 1034–1040, 1994. Also: J. A. Printen, M. J. Brady, A. R. Saltiel. PTG, a protein phosphatase 1-binding protein with a role in glycogen metabolism. *Science*, **275**, 1475–1478, 1997.

45. G. S. Hotamisligil, P. Peraldi, A. Budavari, R. Ellis, M. F. White and B. M. Spiegelman. IRS-1-mediated inhibition of insulin receptor tyrosine kinase activity in TNF-alpha- and obesity-induced insulin resistance. *Science*, **271** (5249), 665–668, 1996.

46. K. T. Uysal, S. M. Wiesbrock, M. W. Marino, and G. S. Hotamisligil. Protection from obesity-induced insulin resistance in mice lackung TNF-α function. *Nature*, **389**, 610–614, 1997.

47. I. Barroso, M. Gurnell, V. E. F. Crowley, M. Agostini, J. W., Schwabe, M. A. Soos, *et al.* Dominant negative mutations in human PPARγ associated with severe insulin resitance, diabetes mellitus and hypertension. *Nature*, **402**, 880–883, 1999.

48. G. S. Hotamisligil, R. S. Johnson, R. J. Distel, R. Ellis, V. E. Papaioannou, and B. M. Spiegelman. Uncoupling of obesity from insulin resistance through a targeted mutation in *aP2*, the adipocyte fatty acid binding protein. *Science*, **274**, 1377–1379, 1996.

49. M. A. Pelleymounter, *et al.* Effects of the obese gene product on body weight regulation in *ob/ob* mice. *Science*, **269**, 540–543, 1995. Also: J. L. Halaas, *et al.* Weight-reducing effects of the plasma protein encoded by the obese gene. *Science*, **269**, 543–546, 1995.

Also: Y. Zhang, *et al.* Positional cloning of the mouse *obese* gene and its homologue. *Nature*, **372**, 425–432, 1994.

50. C. T. Montague, I. S. Farooqi, J. P. Whitehead, M. A. Soos, H. Rau, N. J. Wareham, *et al.* Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature*, **387**, 903–908, 1997.

51. T. Gura. Obesity sheds its secrets: While the hormone leptin has dominated obesity research, new findings suggest it has help in regulating body weight. Results that may provide potential avenues for obesity therapies. *Science*, **275**, 751–753, 1997.

52. O. Boss *et al.* Uncoupling protein-3, a new member of the mitochondrial carrier family with tissue-specific expression. *FEBS Lett*, **408**, 39–42, 1999.

53. J. C. Clapham, J. R. S. Arch, H. Chapman. A. Haynes, C. Lister, G. B. T. Moore, *et al.* Mice overexpressing human uncoupling protein-3 in skeletal muscle are hyperphagic and lean. *Nature*, **406**, (6794), 415–418, 2000.

54. S. Ogg, S. Paradis, S. Gottlieb, G. I. Patterson, L. Lee, H. A. Tissenbaum, and G. Ruvkun. The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature*, **389**, 994, 1997.

55. C. R. Kahn, D. Vicent, and A. Doria. Genetics of non-insulin-dependent (type-II) diabetes mellitus. *Annu Rev Med*, **47**, 509–531, 1996.

56. W. Roush. Worm longevity gene cloned: A gene that helps control the life-span of the nematode, *C. elegans* encodes the worm version of the insulin receptor, thereby providing a possible link between aging and glucose metabolism. Research News. *Science*, **277**, 897–898, 1997. Also: K. D. Kimura, H. A. Tissenbaum, Y. Liu, G. Ruvkun. daf-2, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science*, **277**, 942–946, 1997.

57. S. Gottlieb and G. Ruvkun. daf-2, daf-16 and daf-23: genetically interacting genes controlling Dauer formation in *Caenorhabditis elegans*. *Genetics*, **137**, (1), 107–120, 1994.

58. C. Finch. *Longevity, senescence and the genome*. University of Chicago Press, Chicago, 1990.

59. D. L. Coleman, Obesity genes: beneficial effects in heterozygous mice. *Science*, **203**, (4381), 663–665, 1979.

60. R. A. Roth, B. Zhang, J. E. Chin, and K. Kovacina. Substrates and signalling complexes: the tortured path to insulin action. *J Cell Biochem*, **48**, 112–118, 1992.

61. A. R. Saltiel. Diverse signaling pathways in the cellular actions of insulin. *Am J Physiol*, **270**, (3), Pt 1, E375–385, 1996.

62. D. Mochly Rosen. Localization of protein kinases by anchoring proteins: a theme in signal transduction. *Science*, **268**, (5208), 247–251, 1995.

2

Cell signalling and gene transcription

9

Machinery of gene transcription

Signals of growth factors, cytokines, and hormones, received and transmitted by membrane receptors or nuclear receptors, are eventually targeted to genes. In the nucleus, they call up information essential for the regulation of growth, proliferation, and differentiation of cells. The read-out of the genetic information is tightly regulated at every level—at both the transcriptional and the translational level. Splice variants of pre-mRNAs enlarge the supply of regulatory proteins. Moreover, there are controls of mRNA lifetimes. Finally, the lifetime of the finished product, the protein, is controlled by adjusting rates of protein synthesis and degradation (Fig. 9.1).

Only gene transcription and its control will be considered here. What is discussed here and in the following chapters 10 and 11 on the regulation of gene transcription is only a brief survey of those aspects most relevant for cell signalling. In the genome era in which we live, the analysis of gene expression is the central issue. Excellent and comprehensive treatments of this field are available (ref. 1).

The transcriptional machinery

The transcriptional machinery responds to a multitude of signals, which trigger time-ordered expression of genes that control cellular proliferation and differentiation, the cell cycle, and, eventually, the death of a cell. Controlled gene expression determines the spatiotemporal developmental pattern of an organism and ultimately leads to the proper assembly of an embryo, its growth and maturation. A complex transcriptional machinery sorts and integrates the vast amount of converging and diverging signals. It manages to increase or decrease, at the right moment, pre-messenger RNA production from a particular gene.

The transcriptional apparatus

Nuclear genes encoding messenger RNAs for proteins are transcribed in eukaryotes by RNA polymerase II (Pol II). This is an enzyme complex comprising at least 12 subunits. Pol II, like the other DNA-directed RNA polymerases (I and III), cannot recognize

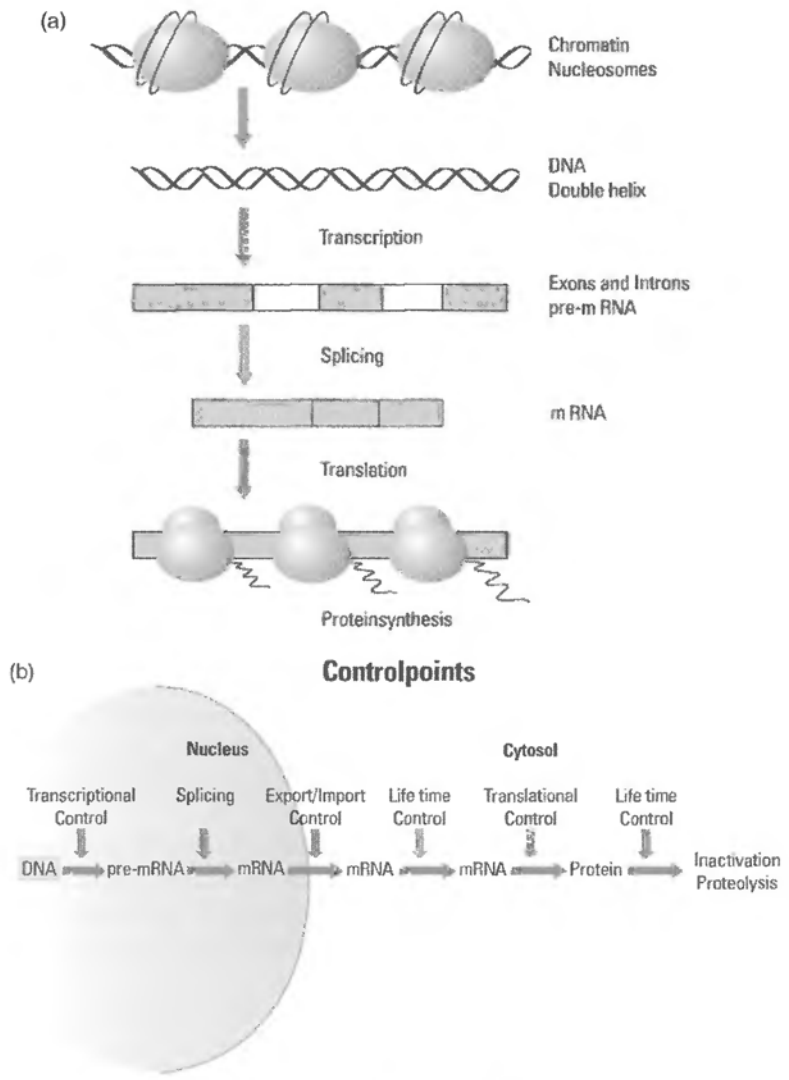


Fig. 9.1 (a) The sequence of steps from DNA to protein; (b) the points of control from DNA to protein.

directly a promoter sequence in the DNA. Promoter recognition and initiation of transcription requires at least 30 additional proteins.

Binding to the promoter and initiation of transcription

We distinguish between basic transcription factors and signal-regulated transcriptional activators. Each of these categories comprises a bewildering profusion and variety of factors and accessory co-factors, and more can be expected. Transcription is initiated by binding of the TATA-box-binding protein (TBP) to the TATA box in the gene promoter. TBP is a subunit of the general transcription factor complex, TFIID (Fig. 9.2).

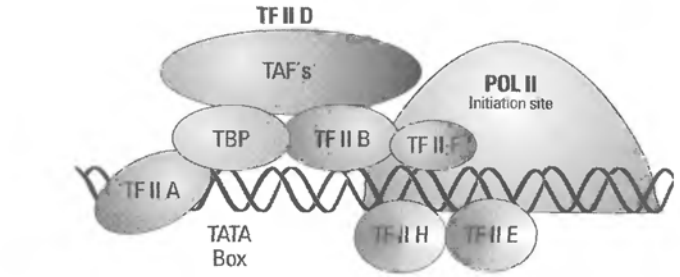


Fig. 9.2 The major components assembled in the pre-initiation complex, PIC. TFIIA and TFIIB help to bind the TFIID complex to the promoter, and then TFIIF, which is a constitutive part of the Pol II complex, directs Pol II to the promoter-bound TFIID complex. The TFIID complex is a target of transcriptional regulators, controlling mitosis. TFIIH is another multi-component complex which has enzymic activies. It has an ATPase, a protein kinase, and a helicase activity, and a DNA excision repair function. TFIIH helps to 'melt' the promoter DNA, making it accessible to the transcriptional machinery, and TFIIE assists the TFIIH helicase. Finally, the kinase activity of the transcriptional complex phosphorylates Pol II and starts transcription.The factors forming the pre-initiation complex belong to the family of basic, general transcription factors, which are evolutionary conserved in eukaryotes, from yeast to humans. These proteins, together with co-activators and repressors, are assembled sequentially, forming the pre-initiation complex. Co-activators are the TAFs (the TATA-box-activating factors). TAFs interact with the TATA-box-binding protein (TBP). TAFs and TBPs are parts of the TFIID complex, although the exact role of the TAFs is still a matter of discussion. (Based on information published in ref. 2,3 and elsewhere.)

Some of these factors are associated with chromatin, such as the HMG (high-mobility group) proteins, others have enzymatic activities, such as DNA topoisomerase and polyADP-ribose polymerase. The point is that the ordered assembly of the pre-initiation complex (PIC) brings Pol II to the transcriptional initiation region of a gene. This sets the stage for the interplay with specific, signal-responsive transcriptional regulators. The calling up of the genetic programme of a cell is conducted in a highly regulated, time-ordered sequence. Promoter and enhancer regions control the expression of genes. Many gene promoter regions in eukaryotic cells have a TATA box or another initiator sequence, to which PIC delivers Pol II (Fig. 9.3).

How do the distantly bound transcription factors contact the basic transcriptional machinery and instruct Pol II when to start transcription? Although there are linkers to establish these contacts, the chromatin complex itself helps to make the contacts. DNA is highly condensed, suprahelical, and often bends on binding a dimeric transcription factor, bringing the transcriptional regulators closer together (Fig. 9.4).

Interaction of the TATA-box-binding protein (TBP) with promoter DNA is rather inefficient and appears to be the rate-limiting step for the start of transcription. TBP must actually dissociate first from the TFIID complex before it can bind to the TATA-box DNA. Dissociation of TBP is facilitated by the dimeric structure of TFIID,[4] when it is not bound to DNA, and by the interaction of TBP with TFIIA.

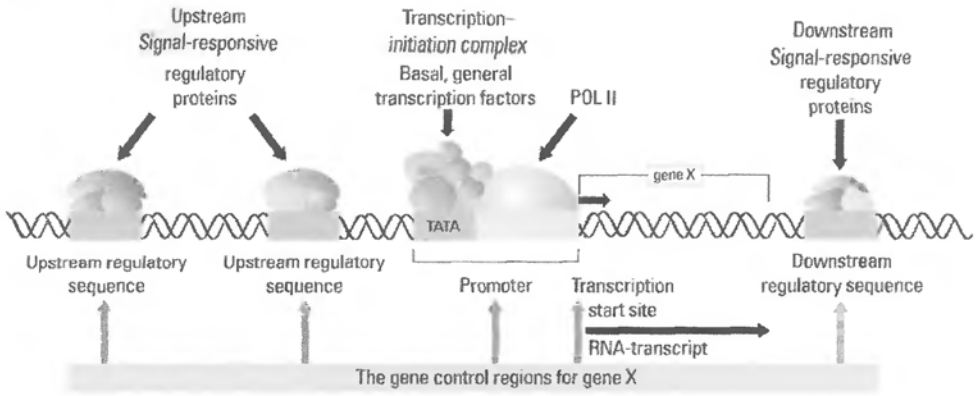


Fig. 9.3 The regions, upstream and downstream of the TATA box, that regulate initiation of gene expression. The TATA box, as the name indicates has a A+T-rich sequence and is positioned about 30–50 base pairs (bp) upstream of the transcriptional initiation site. Upstream from the TATA box are the upstream activating and enhancer sequences (UAS). Binding of PIC to the TATA box region is not sufficient to initiate transcription. Binding of transcriptional activators to several of these upstream sequences is also required. The scheme depicts the gene regions to which the basic, general transcription factors and Pol II bind and the regions to which the special, signal-responsive gene-regulatory transcription factors bind. It is indicated that signal-transducing transcription factors bind to quite distant DNA sites, upstream and downstream of the gene, the transcription of which is controlled. The gene control regions may be hundreds of base pairs removed from the transcriptional initiation sites.

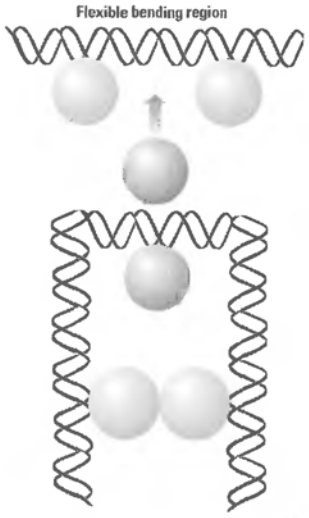


Fig. 9.4 How binding of a protein to a flexible, A-rich DNA region induces a bend of the DNA, which brings two subunits of a transcription factor together, so that they can dimerize.

The activation of Pol II and the lifetime of the active transcriptional complex is regulated by phosphorylation dephosphorylation. Pol II is initially underphosphorylated. When the PIC is formed, the polymerase is fully phosphorylated and transcription starts. When the transcribed RNA is elongated, the polymerase is dephosphorylated and ready for the next round of transcription. Phosphorylation of Pol II also links transcription to the cell cycle. Pol II is phosphorylated by a serine/threonine-specific cyclin-activated kinase (CAK), the activity of which is dependent on cyclin H. CAK is identical with the cyclin H-dependent kinase, Cdk7, and together with cyclin H is part of the TFIIH complex (Fig. 9.5).[5,6]

The link between gene transcription and the cell cycle is of pivotal importance for the control of cell proliferation. Many examples could be cited where cyclin/Cdks regulate transcription. The SRBs (suppressors of RNA polymerase II in humans, yeast, and *Drosophila*) are closely related to cyclin C and the corresponding Cdk8[7] (when the *srb* genes are deleted or mutated in yeast, transcription of all genes transcribed by Pol II is out of control).

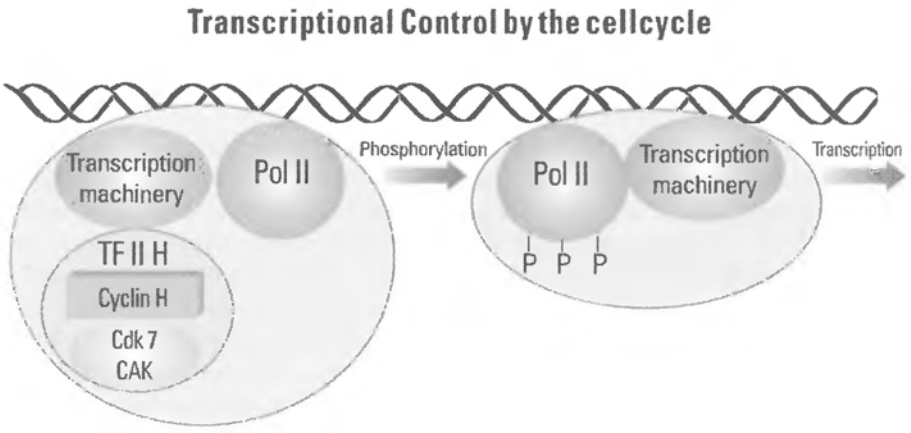


Fig. 9.5 A link between gene transcription and the cell cycle is shown, involving activation of Pol II through phosphorylation by the cyclin-dependent kinase, Cdk7/CAK. (Based on information from ref. 6.)

Pol II is not the only target of Cdks. Cell-cycle kinases also control transcription by phosphorylation of TBPs and TBP-associated factors, TAFs; although, at least in yeast, TAFs seem not to be obligatory for transcription.[8] Tumour suppressors, such as p53, which are controlled by phosphorylation by Cdks, link transcription and the cell cycle.[9,10] Breakdown of transcriptional control by tumour suppressors is related to cancer. This is discussed in Part 4.

Structural motifs of DNA-binding proteins

Just as proteins have specific structural motifs for the recognition of other proteins (SH2, SH3 domains and so forth, see Chapter 2), so those proteins binding to genes have specific structural recognition motifs for DNA. (see ref. 11). Most transcription factors interact with the major groove of DNA.

Max Perutz, in a lucid discussion of how proteins recognize genes, has pointed out that proteins and DNA 'learned' to recognize each other as palindromic sequences appeared during evolution.[12] Recognition of palindromic sequences also explains why many transcription factors are dimeric. Perutz pointed out that a palindromic sequence in a double-stranded DNA can accommodate the protein at each half of the two strands by a rotation of 180° around the twofold axis of symmetry. This brings two molecules of the protein into related positions on the DNA and facilitates dimer formation. Such an arrangement has several advantages: first, binding of a dimer to palindromic sequences would be co-operative and strong. Moreover, when a dimeric transcription factor combines, in response to signals from a cytokine or a hormone, with an accessory protein that changes its conformation either allosterically or by phosphorylation, the conformational change can be amplified and propagated through changes in the arrangement of the monomers in the dimer. Finally, binding of a dimeric transcription factor helps to bend the DNA.

X-ray crystallography has provided information about how these structural motifs recognize DNA. The first motif for DNA binding which was described in proteins was the helix-turn-helix (HTH) motif. It is the classical structure where an α-helix is inserted into

the major groove of DNA and binding is stabilized through a second helix. It is also the simplest, most common DNA-binding motif. Many transcriptional proteins use this motif to bind to DNA, including the homeodomain proteins, which have a central function in the regulation of developmental programmes.

Walter Gehring's group was the first to isolate and clone a homeotic gene, *antennapedia,* which determines the development of the head and segments of the thorax of *Drosophila* (for information see ref. 13). These genes are called homeotic (from the Greek *homoios*, alike) because they act alike, all regulating the expression of whole groups of genes necessary for the ordered development of the fly. Proteins homologous to the *Drosophila* homeodomain proteins regulate homeotic genes (the homeobox genes) in many eukaryotes, including humans. In yeast, similar genes code for proteins that determine the mating type of yeast. It is remarkable that the HLH motif of homeodomain proteins is always presented to the DNA in the same way, despite the fact that sequence identity among homeodomain proteins is rather low (only 16 out of 60 amino-acid positions are identical).

As example of a homeodomain protein, the structure of MATα2 will be described briefly. MATα2 is a repressor of the yeast mating type a. GCN-4 is a gene control factor, (Fig. 9.7), and MCM1 is a mating control factor in yeast. MATα2 combines with the transcriptional regulator MCM1 and binds cooperatively to the operator region of the homeotic gene that determines the mating type in the yeast, *Saccharomyces cerevisiae.* The structure of an MATα2/MCM1/DNA complex was solved recently by Tan and Richmond (Plate 15).[14] A small (only eight residues long) module in the linker region of MATα2 bound to DNA recruits MCM1. Bending of DNA brings the two proteins closer together, facilitating their interaction. A remarkable feature of the complex is that the linker is structurally extraordinarily pliable: In the *cis*-α2 monomer, the linker is a β-strand hairpin, and in *trans*-α2 it is helical. The linker is responsible for the cooperative interaction of MATα2 and MCM1 with a stretch of operator DNA 30–31 bp long. The participation of MCM1 increases the affinity of MATα2 for the DNA by 50- to 100-fold.

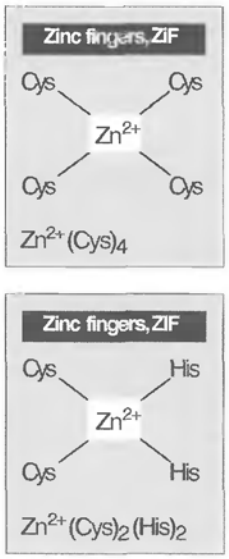


Fig. 9.6 The structures of common Zn^{2+}-binding motifs are shown. Above is a $Zn^{2+}(Cys)_4$ motif, and below a $Zn^{2+}(Cys)_2(His)_2$ finger. (The same structural schemes are shown in ref. 15 and are reproduced here thanks to the courtesy of Professor Gerhard Krauss.

Another common DNA-binding motif is the zinc finger (ZIF), found in many proteins that regulate gene transcription in eukaryotes (Fig. 9.6). As example, the crystal structure of a ZIF-containing peptide is shown, which binds to immediate early genes controlling cell growth in mice (Plate 16).[16] ZIFs direct α-helices to the major groove of DNA, where they interact with specific sites.

α-Helices with regular repeats of leucines or other hydrophobic amino-acid residues on one side interact with each other by means of hydrophobic interactions, forming coiled coils. Transcription factors with such a leucine-zipper motif form homo- and heterodimers. An example of a leucine zipper complexed with DNA containing a binding site for the transcription factor Jun/Fos is shown in Fig. 9.7.

Another common structural arrangement is the basic helix-loop-helix (bHLH) motif, (Plate 17), first recog-

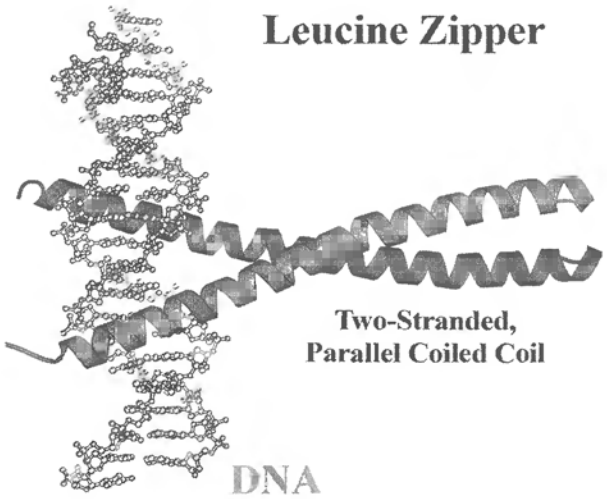


Fig. 9.7 The structure is of a peptide corresponding to the leucine zipper of the yeast transcriptional activator, GCN4. The α-helices wrap around each other in a left-handed supercoil, forming a two-stranded coiled coil. The α-helices are packed as in the model proposed by Francis Crick in 1953. The coiled coil of α-helices was one of the first protein structures deduced from X-ray diffraction studies. Transcription factors with a leucine zipper (bZIP) have a high degree of amino-acid similarity in the DNA-binding domain, and often address similar DNA sites. Transcription factors with bZIPs include such common transcription factors as c-Jun, c-Fos, CREB (cyclic AMP reponse-element-binding protein) and GCN4, a transcriptional activator which controls genes for amino-acid biosynthesis in yeast. The zipper-like interactions are instrumental in dimerization, but the contacts with DNA are made by adjacent regions, rich in basic residues. (The ribbon model has been reconstructed with permission of the authors and Cell on the basis of the data in ref. 17.)

nized by David Baltimore and his colleagues, more than 10 years ago.[18] HLH transcription factors have two amphipathic helices, separated by a loop of variable length and a basic region that is responsible for sequence-specific DNA recognition. The prototype of a helix-loop-helix protein is the transcription factor Myc. c-Myc is a common, signal-responsive transcription factor, like Jun/Fos. These factors control cell growth and proliferation. Myc has a partner, Max (the Myc-associated factor X), with which it forms a heterodimer with a common leucine-zipper motif.[21–22] On dimerization, a four-helix bundle is formed, of which the two basic ends bind to DNA. Myc–Max heterodimers control Myc-driven cell proliferation. The transcription factors of the oncogenic papilloma virus are a variation on this structural theme (see Part 4). These viral factors are also dimeric and have an alternatively spliced helix-loop-helix (HLH) motif. They recognize palindromic sequences and bind cooperatively with high affinity to short DNA sequences of only a few base pairs.

Both the HLH motif and $Zn^{2+}(Cys)_4$ motifs are found in the DNA-binding domain (DBD) of nuclear steroid-hormone receptors (see Chapter 11).

An important eukaryotic transcription factor is the nuclear factor, NF-κB (NF-xB). This factor binds to DNA through β-sheets. The β-sheets of the dimeric transcription factor wrap around the DNA like a mantle. Contacts are made, as in the case of α-helices, with the major groove of the DNA. The structure is shown in Plate 18.

Having introduced some of the most common DNA-binding motifs, let us try to envision what the pre-initiation complex bound to DNA actually looks like.

A structural view of the assembly of the pre-initiation complex

Fortunately, the complicated process of the assembly of the many components of PIC at transcription initiation sites can be simplified by substituting the multicomponent TFIID complex by TBP. The TBP–TATA-box complex represents a reduced PIC which can actually carry out transcription *in vitro*.

The crystal structure of a human TATA-box-binding protein (hTBP), complexed with TATA-box DNA, may be compared with that of a yeast TBP/TATA-box complex and a similar isoform of TBP from *Arabidopsis thaliana*, bound to a TATA box in an adenoviral promoter (Plate 19).[25–27]

TFIIB assists in the assembly of the TFIID–DNA complex, of which TBP is a part. Therefore, it would be useful to know what role TFIIB has in the interaction of TBP with the TATA box. The structure of a ternary complex of TFIIB with TBP and the TATA-box DNA shows the positions of the residues in TBP that are crucial for the interaction with TFIIB. Plate 20 shows that TFIIB is attached to the phosphoribose backbone of the DNA, both upstream of the TATA box and also at the transcriptional start site, which is downstream of the TATA box.[28]

Finally, in Plate 21, the structure of the yeast TFIIA–TBP–DNA complex is shown. TFIIA interacts with TBP and the TFIID complex and stimulates transcription of the Pol II gene.[29–32] TFIIA has two characteristic structural motifs. One is a six-stranded β-sandwich, the other is a left-handed four-helix bundle. The β-sandwich domain of TFIIA is alone responsible for all of the interactions with the DNA, whereas its helix-bundle domain projects away and is free to interact with signal-responsive transcription factors. These interactions are important for regulation of transcription. Little conformational change occurs in TBP when it binds to TFIIA. The main difference between the TFIIA–TPB complex and the TFIIB–TBP complex is that TFIIA binds upstream of the TATA box, away from the transcriptional start site, whereas TFIIB binds downstream of the TATA box. Moreover, TFIIB is positioned on the side opposite to TFIIA.

Conclusions: what have we learnt from the structures?

We have learnt from the structures that the TBP/TATA complex is a nucleoprotein scaffold upon which other factors, such as TFIIA and TFIIB attach with high affinity through a combination of stereospecific and electrostatic interactions. TFIIB is positioned at the transcriptional start site, between the attachment sites for TBP and Pol II. On the other hand, TFIIA has no contacts with the DNA downstream of the TATA box and does not interact with the transcriptional start site and/or any of the components of the basal transcription machinery, all of which are located downstream of the TATA box. Thus, TFIIA and TFIIB can bind simultaneously, without mutual interference. Moreover, TFIIA is accessible to specific, signal-responsive regulatory transcription factors. Its location upstream of the TATA box also enables TFIIA to absorb and scavenge transcriptional inhibitors, making them ineffectual.

To sum up:

1. The specific contacts of TBP with DNA are quite modest; however, the structural changes of the DNA on binding of TBP are dramatic.

2. The entire top side of TBP bound to the TATA-box DNA is out of the way and is free to interact with other factors. This leaves a generous surface on TBP available for interactions with the multitude of factors which are parts of the transcription initiation complex. The same is true for the TFIIA/TFIIB–TBP-complexes. The extensive surfaces displayed by the TFIIA/TFIIB–TBP–DNA complexes represent potential sites for binding basal initiation factors, signal-responsive transcriptional activators, co-activators and mediators, and, most importantly, leave room for Pol II.

Structural organization of nuclear DNA

Gene expression is still responsive to some controls, even when the DNA is naked. However, the range of regulation is much reduced *in vitro* as compared to conditions *in vivo*, where the chromatin structure is intact (as in the chromosomes in the nucleus of the cell). Therefore, the structural organization of native nuclear DNA is important for the regulation of gene transcription.

The nucleosome

The nucleosome manages to cram 2 m of human DNA into a space just a few micrometres in diameter. The basic DNA packaging unit is called the nucleosome core particle. It contains DNA wrapped around a histone octamer $(H_2A, H_2B, H_3, H_4)_2$. The nucleosome particles were first described by Roger Kornberg and Aaron Klug in 1974. They are connected, much like beads on a string, interrupted by stretches of naked DNA. An average cell nucleus contains 25 million of these particles, each consisting of a core of the histone octamer, around which 146 bp of the DNA are wrapped (Fig. 9.8).

Structure of the nucleosome

T. Richmond and his team, see also ref. 34, have obtained a X-ray structure that provides a detailed picture of the core particle. This is a remarkable feat, because with a molecular weight of 206 000 Da, about half of it protein and half DNA, the nucleosome core structure is, so far, the largest DNA– protein complex to have been resolved at atomic resolution (Plate 22). The structure shows how the histone proteins wrap around the DNA and how the DNA is curved at some positions more than at others. As pointed out above, bending the DNA brings factors that regulate gene transcription closer together.

Fig. 9.8 The nucleosome

For DNA replication and repair, and for gene transcription, the DNA must be at least partially unwrapped. Unwrapping of the DNA involves an ATP-dependent disruption of nucleosomes and is part of gene transcription and a point of control. It allows transcriptional activators access to specific gene sites. Richmond's structure helps to envisage how large enzyme complexes, such as the DNA

Table 9.1 The properties of histones

Histones	Properties
H_1	Linker, binds to methylated genes
H_2	Conserved structure
H_2B	Core of nucleosome
H_3	
H_4	Acetylated

replicase or the RNA polymerase, can travel along the DNA strand without completely displacing the nucleosome. It seems that the histone octamer may never be totally removed from the DNA and that an enzyme such as DNA replicase transiently peels off 30 or 40 bp long stretches of the DNA from the histones. When the enzyme has done its job and has passed the site, DNA snaps back and, together with the histones, reforms the original nucleosome. Covalent modification of the histones may facilitate unwrapping and re-formation of nucleosomes.

Modification of histones

Histones are among the most highly conserved proteins; Table 9.2 summarizes their properties. The tails of the histones stick out of the nucleosome, making them accessible to enzymes for modification. Modifier enzymes, which play an important role in regulating transcription, are the histone acetyl transferases and deacetylases. In transcriptionally active chromatin, lysine residues in the NH_2-terminal tail of nucleosomal histones are acetylated. Acetylation and deacetylation of histones are supposed to modulate the nucleosomal conformation and open specific gene sites for transcription. The activity of many transcription factor complexes is enhanced through histone acetylation of specific promoter sites. Thus, histone acetylation–deacetylation is linked to transcription.[35,36] Figure 9.9 shows how modification of histones by acetylation, methylation, and phosphorylation may help to open up the chromatin for gene transcription.

Two types of histone acetyltransferases have been distinguished: the type A histone acetyltransferases seem to have a regulatory role, whereas the type B histone acetyltransferases may have a more structural role in chromatin assembly. Type A histone acetyltransferases acetylate a particular set of lysines in histones H_3 and H_4. These sites are different and do not overlap with the sites targeted by the type B histone acetyltransferases.

A simple way to control transcription would be acetylation–deacetylation reactions carried out by the transcription factors themselves. Transcriptional regulators with histone acetylation–deacetylation activities have been identified.[39–40] For example, the transcriptional repression and anti-cancer activity of a human DNA-binding protein (Sin3) has been ascribed to its histone deacetylase activity.[41]

Acetylation is not the only covalent modification of histones. Histones can be methylated, ADP-ribosylated, and phosphorylated. Each of these covalent modifications has been implicated in packaging DNA, in the replication of DNA, and in the regulation of gene transcription.

To sum up: for a gene to be transcribed, the chromatin wrapped in the nucleosome must become accessible to basic and signal-responsive regulatory transcription factors,

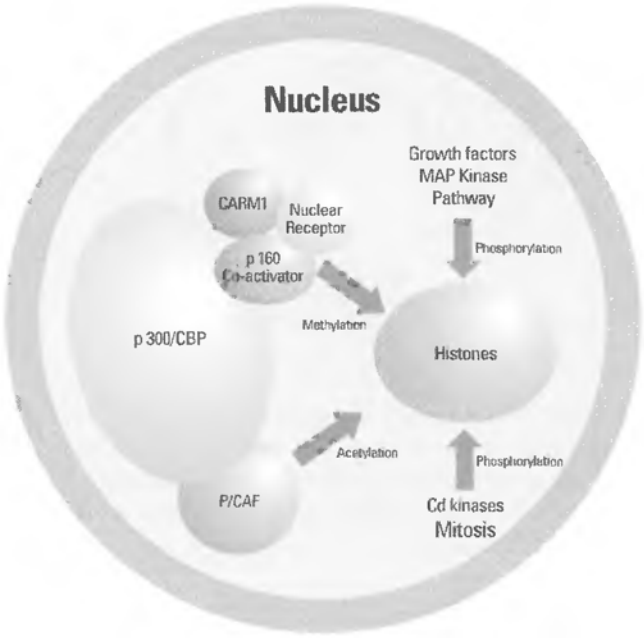


Fig. 9.9 Some nuclear receptors and co-activators covalently modify certain histones by methylation, and certain transcriptional regulators modify them by acetylation. Growth-factor and stress signals, and other external signals, lead through the MAP kinase pathway to histone phosphorylation. Histone phosphorylation by cyclin-dependent kinases is also involved in chromatin condensation in the course of chromosome packaging for cell division. CARM 1 is a co-activator-associated arginine (R) methyltransferase.[37] P/CAF is a member of the family of nuclear histone acetyltransferases. p300/CBP is an important transcriptional co-activator system (see Chapter 10). (The information in this scheme is taken from ref. 38.)

and to the transcriptional machinery, including DNA-dependent RNA polymerase II. Only when all these factors are in place, bound to promoter and enhancer regions, can gene transcription start.

Methylation of genes

A DNA modification which changes gene transcription is methylation of nucleotide bases. Heterochromatin contains heavily methylated DNA which cannot be transcribed, Genes in euchromatin are less methylated and are transcribable. Methylation at C5 of cytosine in CpG sequences by a methyltransferase, with *S*-adenosyl methionine (SAM) as methyl donor, silences genes in heterochromatin. Methylation may affect the higher-order structure of DNA and may impede access of the transcriptional machinery to DNA. Members of the histone H_1 family have been associated with transcriptional repression, because these histones bind preferentially to methylated CpG sequences (Fig. 9.10).[41]

Although the role of methylation in the control of gene expression is not yet resolved, there is a situation where methylation plays an important role, that is genomic imprinting, the differential transcription or silencing of genes inherited from the mother or the father. Mammalian cells are diploid. One set of chromosomes is inherited from the mother, the other from the father. Normally the two alleles of a gene, inherited from the mother or from the father, are transcribed equally, without bias. But there are about 200 genes which are only expressed either when inherited from the mother, or from the father. These genes are called 'imprinted' genes (Fig. 9.11).

Several human diseases are linked to imprinted genes.[43] Genomic imprinting has also been demonstrated in transgenic mice. For example, transgenic mice with the gene for

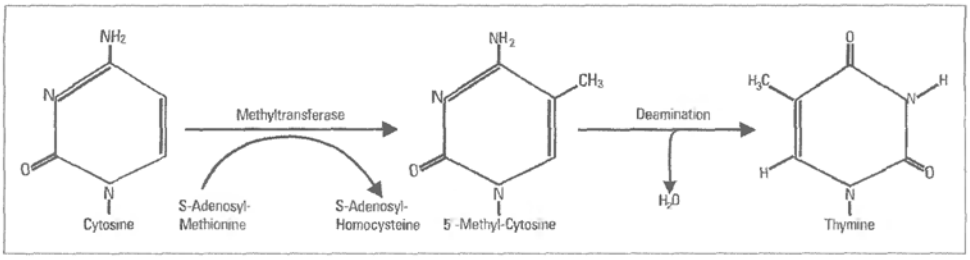


Fig. 9.10 Methylation of cytosine.

the insulin-like growth factor 2 (IGF-2) inactivated by mutation developed normally when the mutated gene was inherited from the mother, but when the same gene was paternally inherited, the transgenic mice were stunted in their growth. The reason for the difference in the phenotype is that only the paternally inherited IGF-2 gene was expressed in these mice, whereas the maternal gene was silent.

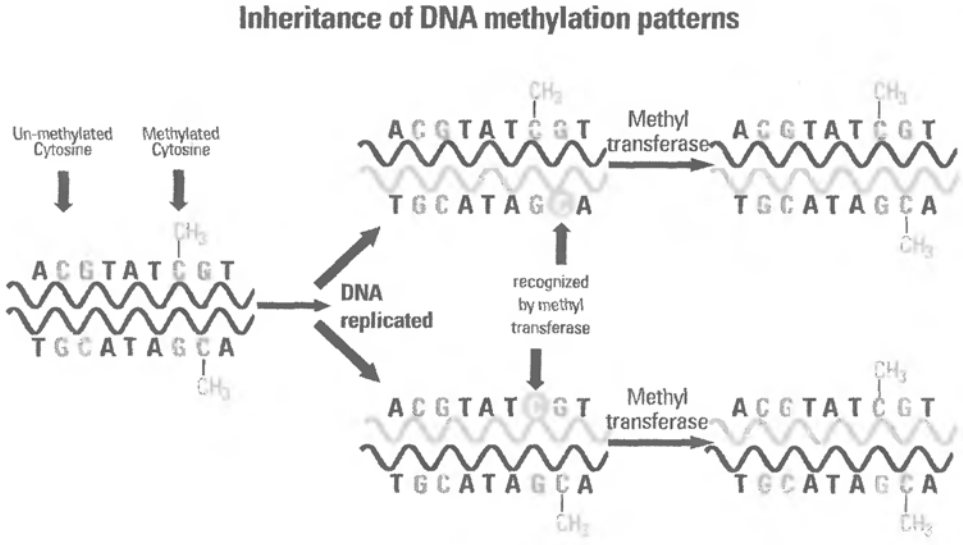


Fig. 9.11 In the DNA of vertebrates, a proportion of the cytosines in the CG sequence are methylated. The DNA methylation pattern is faithfully maintained in the course of embryogenesis and cell differentiation, because normally the discriminating, selective action of the methyltransferase maintains the pattern of DNA methylation, once it is established. Not so in imprinted genes, for which the lack of a methylating enzyme is thought to be responsible. (This information is from ref. 42, Fig. 9–68, p. 450 and is shown with permission of Taylor and Francis, Inc.)

Summary

More than 30 basic, general transcription factors (GTFs), in addition to the RNA polymerases, participate in gene transcription. The remarkable progress made in the past few years in the identification and structural characterization of essential parts of the preinitiation complex and of the nucleosome gives hope that this intricate puzzle can be solved and that the whole transcription machinery may be put together. A detailed picture may eventually emerge, that helps us to understand how some 30 000 or more genes in the human cell are expressed and controlled. Of course, many mechanistic details need to be clarified; to mention only one: the question of whether the DNA and RNA polymerases move along the DNA in the course of replication or transcription, or whether these enzymes are immobilized and remain fixed to the template, like parts of a factory through which the DNA passes (Fig. 9.12).

Having introduced the components of the basic transcriptional machinery, we can now return to our primary concern, namely to explain how growth factors, cytokines, and hormones manage to regulate transcription of specific genes. But, for signal-responsive transcriptional regulators to bind to sites in the control region of specific genes, certain requirements must be fulfilled:

(1) the specific target site in the chromatin must be unwrapped and made accessible, so that

(2) distantly bound transcriptional activators can be linked to the transcriptional machinery.

In the next chapter, we shall see how signal-responsive transcription factors control gene transcription. Moreover, we shall see in Chapter 11 that lipophilic hormones, such as steroid hormones and retinoids, use a more direct route to the gene, because these hormones bind to diffusible cytosolic receptors that are also transcriptional activators.

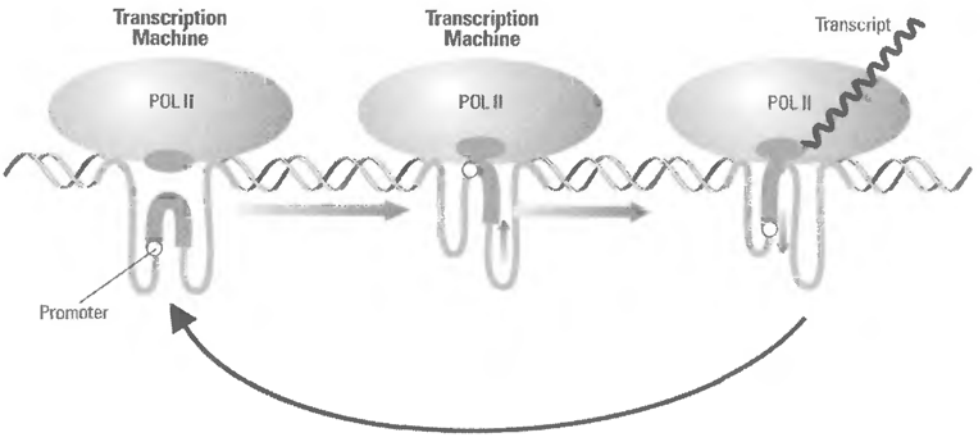


Fig. 9.12 A transcription cycle, according to P. R.Cook. A loop of chromatin is attached to the transcriptional machinery. The transcript is synthesized, while the template slides through the polymerase. When transcription is completed, the chromatin template detaches and a new cyle of transcription can start. (Based on Fig. 5 in ref. 44 and shown with permission of Science.)

References

1. M. Carey and S. T. Smale. Transcriptional Regulation in Eukaryotes: Concepts, Strategies and Techniques. Cold Spring Harbor Laboratory Press, New York, 640 pp, 2000.
2. T. Oelgeschlager, C.M. Chiang, and R.G. Roeder. Topology and reorganization of a human TFIID-promoter complex. *Nature*, **382** (6593), 735–738, 1996.
3. N. Segil, M. Guermah, A. Hoffmann, R.G. Roeder, and N. Heintz. Mitotic regulation of TFIID inhibition of activator-dependent transcription and of changes in cellular localization. *Genes Dev*, **10**, 2389–2400, 1996.
4. A. K. Taggart and B. F. Pugh. Dimerization of TFIID when not bound to DNA. *Science*, **272** (5266), 1331–1333, 1996.
5. M. Rossignol, I. Kolb-Cheynel, and J.-M. Egly. Substrate specificity of the Cdk-activating kinase (CAK), is altered upon association with TFIIH. *EMBO J*, **16**, 1628–1637, 1997.
6. E. A. Nigg. Cyclin-dependent kinase 7: At the cross roads of transcription, DNA repair and cellcycle control? *Curr Opin Cell Biol*, **8**, 312–317, 1996.
7. P. Rickert, W. Seghezzi, F. Shanahan, H. Cho, and E. Lees. Cyclin C/Cdk8 is a novel CTD kinase associated with RNA polymerase II. *Oncogene*, **12**, 2631–2640, 1996.
8. D. M. Chao and R. A.Young. Activation without a vital ingredient, News and Views. *Nature*, **383**, 119–120, 1996.

   Also: Z. Moqtaderi, Yu. Bai, D. Poon, P. A. Weil, and D. Struhl. TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature*, **383**, 188–191, 1996.

   Also: S. S. Walker, J. C. Reese, L. M. Apone, and M. R. Green. Transcription activation in cells lacking TAF II. *Nature*, **383**, 185–188, 1996.
9. J. R. Bischoff, P. N. Friedman, D. R. Marshak, C. Prives, and D. Beach. Human p53 is phosphorylated by p60-cdc2 and cyclin B-cdc2. *Proc Natl Acad Sci, USA*, **87**, 4766–4770, 1990.
10. H.-W. Sturzbecher, *et al.* p53 interacts with p34/cCdc2 in mammalian cells: Implications for cell cycle control and oncogenesis. *Oncogene*, **5**, 795–801, 1990.
11. C. O. Pabo and R. T. Sauer. Transcription factors: Structural families and principles of DNA recognition. *Annu Rev Biochem*, **61**, 1053–1095, 1992.
12. M. Perutz. *Protein structure. New approaches to disease and therapy*, Chapter 3. W.H. Freeman, New York, 1992.
13. W. J. Gehring, M. Affolter, and T. Buerglin. Homeodomain proteins. *Annu Rev Biochem*, **63**, 487–526, 1994.

   Also: W. J. Gehring, *et al.* Homeodomain-DNA recognition. *Cell*, **78**, 211–223, 1994.
14. S. Tan and T. J. Richmond. Crystal structure of the yeast MATa 2/MCM1/DNA ternary complex. *Nature*, **391**, 660, 1998.
15. G. Krauss. *Biochemie der Regulation der Signaltransduktion*. Wiley-VCH, 1997.
16. N. P. Pavletich and C. O. Pabo. Zinc finger–DNA recognition: crystal structure of a Zif268–DNA complex at 2.1 Å. *Science*, **252**, 809–817, 1991.
17. T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α-1 helices: Crystal structure of the protein–DNA complex. *Cell*, **71**, 1223, 1992.
18. C. Murre, P. S. McCaw, and D. Baltimore. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell*, **56**, 777–783, 1989.
19. H. Weintraub. The Myo D family and myogenesis: redundancy, networks and thresholds. *Cell*, **75**, 1241–1244, 1993.
20. P. C. M. Ma, M. A. Rould, H. Weintraub, and C. O. Pabo. Crystal structure of myoD BLH-domain bound to DNA: Perspectives on DNA recognition and implications for transcriptional activation. *Cell*, **77**, 451, 1994.
21. L. M. Facchini and L. Z. Penn. The molecular role of Myc in growth and transformation: Recent discoveries lead to new insights. *FASEB J*, **12**, 633–651, 1998.
22. M. Henriksson and B. Luscher. Proteins of the Myc network: essential regulators of cell growth and differentiation. *Adv Cancer Res*, **68**, 109–182, 1996.
23. G. Ghosh, G. van Duyne, S. Ghosh, and P. B. Sigler. Structure of the NF-xB p50 homodimer bound to a κB site. *Nature*, **373**, 303–310, 1995.

24. C. W. Müller, F. A. Rey, M. Sodeoka, G. L. Verdine, and S. C. Harrison. Structure of the NF-κB p50 homodimer bound to DNA. *Nature*, **373**, 311–317, 1995.
25. J. L. Kim and S. K. Burley. 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nature Struct Biol*, **1** (9), 638–653, 1994.
26. D. B. Nikolov, H. Chen, E. D. Halay, A. Hoffman, R. G. Roeder, and S. K. Burley. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci, USA*, **93**, 4862, 1996.
27. Y. Kim, J. H. Geiger, S. Hahn, and P. B. Sigler. Crystal structure of a yeast TBP/TATA-box complex [see comments]. *Nature*, **365** (6446), 512–520, 1993.
28. D. B. Nikolov, H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, *et al.* Crystal structure of a TFIIB–TBP–TATA complex. *Nature*, **377**, 119–128, 1995.
29. J. J. Kang, D. T. Auble, J. A. Ranish, and S. Hahn. Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. *Mol Cell Biol*, **15** (3), 1234–1243, 1995.
30. J. A. Ranish, W. S. Lane, and S. Hahn. Isolation of two genes that encode subunits of the yeast transcription factor IIA. *Science*, **255** (5048), 1127–1129, 1992.
31. S. Tan, Y. Hunziker, D. F. Sargent, and T. J. Richmond. Crystal structure of a yeast TFIIA/TBP/DNA complex [see comments]. *Nature*, 381 (6578), 127–151, 1996.
32. J. H. Geiger, S. Hahn, S. Lee, and P. B. Sigler. Crystal structure of the yeast TFIIA/TBP/DNA complex [see comments]. *Science*, **272** (5263), 830–836, 1996.
33. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond. Crystal structure of the nucleosome core particle at 2.8 Å resolution [see comments]. *Nature*, **389** (6648), 251–260, 1997.
34. C. Featherstone. X-Ray crystallography researchers get their first good look at the nucleosome. *Science*, **277**, 1763–1764, 1997.
35. A. P. Wolffe. Histone deacetylase: a regulator of transcription [comment]. *Science*, **272** (5260), 371–372, 1996.
36. M.-H. Kuo, J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, *et al.* Transcription-linked acetylation by gcn5p of histones H3 and H4 at specific lysines. *Nature*, **383**, 269–272, 1996.
37. D. Chen, H. Ma, H. Hong, S.S., Koh, S.-M- Huang, B. T. Schurter, *et al.* Regulation of transcription by a protein methyltransferase. *Science*, **284**, 2174–2177, 1999.
38. M. Hagmann. How chromatin changes its shape. News Focus. *Science*, **285**, 1200–1203, 1999.
39. A. P. Wolffe. Transcriptional control. Sinful repression [news; comment]. *Nature*, **387** 6628, 16–17, 1997.
40. L. Alland, R. Muhle, H. Hou Jr, J. Potes, L. Chin, N. Schreiber Agus and R. A. DePinho. Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression [see comments]. *Nature*, **387** (6628), 49–55, 1997.
41. A. Bird. The essentials of DNA methylation. *Cell*, 70, 5–8, 1992.
42. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell*. Garland Publishing, New York, 1994.
43. J. G. Hall. Genomic imprinting: nature and clinical relevance. *Annu Rev Med*, **48**, 35–44, 1997.
44. P. R. Cook. The organization of replication and transcription. *Science*, **284**, 1790–1795, 1999.

10

Regulation of gene transcription by growth factors and cytokines

General transcription factors have been discussed above, and now we turn to specific, signal-responsive transcriptional regulators. Specific transcription factors link extracellular signals, received by receptors and transmitted through cellular signalling pathways to the gene. Considering the central role of transcriptional activation for growth and differentiation of all cells and tissues, it is hardly surprising that thousands of factors control the expression of about 30 000 human genes. Therefore, the discussion in this chapter is, by necessity, simplistic and selective. Moreover, bear in mind that new factors are discovered all the time and that there is still a way to go until all facets of gene regulation are understood.[1–3]

Signal-regulated transcriptional activators

This chapter focuses on control of gene transcription by growth factors and cytokines. This is part of our effort to find out how specific and unique cellular responses to growth promoters are accomplished.

Growth-factor- and cytokine-regulated transcription

First, we consider properties of transcriptional regulators that activate genes, involved in the control of cell growth. A prominent member of this group is the Jun-family of transcription factors: *v-jun* is the oncogene of an avian sarcoma virus, ASV 17; c-*jun* is the cellular proto-oncogene that encodes the trancription factor, activating protein 1 (AP1). The product of the c-*jun* gene, Jun, dimerizes with Fos, forming an heterodimeric transcription factor, Jun/Fos, which binds to the AP1 promoter site. c-*fos* is a proto-oncogene; v-*fos* is the oncogene of a murine Fujinami-osteosarcoma virus and an avian sarcoma virus.

The Jun/Fos transcription factors

The c-*jun* gene is rapidly induced by a great variety of signals, ranging from growth factors and cytokines to lipid messengers and G proteins.[4,5] Each signal may result in a

somewhat different cellular response. Overexpression or mutations of the c-*jun* and c-*fos* genes can cause cancer[6] (see Part 4). The products of the c-*jun* and c-*fos* genes control their own expression. These genes belong to the group of 'immediate early' genes, i.e. genes for which transcription is induced immediately in response to growth-promoting signals.

Transcription factors bind to *cis*-acting and *trans*-acting sites. cis-acting regulatory sites are on the same side of the gene that is transcribed, *cis* means on this side, whereas *trans*-responsive sites can be on either side. *Trans* means across, beyond, on the other side. In genetics *trans* also means a regulatory gene element whose functions are insensitive to its position.[7] An example of common *trans*-responsive DNA elements, engaged in the induction of the c-*jun*/c-*fos* genes are the upstream activating sequences (UASs) and the serum response elements (SREs). An example of a *cis*-responsive element is CRE, the cAMP response element. Activation of CRE brings the c-*jun*/c-*fos* genes under the control of the second messenger, cAMP. Productive interaction of the serum response element, SRE, with the serum response factor (SRF) requires, like activation of CRE by CREB (the CRE-binding protein), the assistance of an auxiliary factor, the ternary complex factor (TCF). Both, activation of CRE and SRE are controlled by phosphorylation. More distant enhancer elements are involved in the response of the c-*jun*/c-*fos* gene to stimuli from cytokine receptors.

In Fig. 10.1 a scenario is shown, where different signals travelling over different signalling pathways eventually all converge on the promoter of the c-*fos* gene, but each signal addresses a different specific element in the gene promoter. This selective response

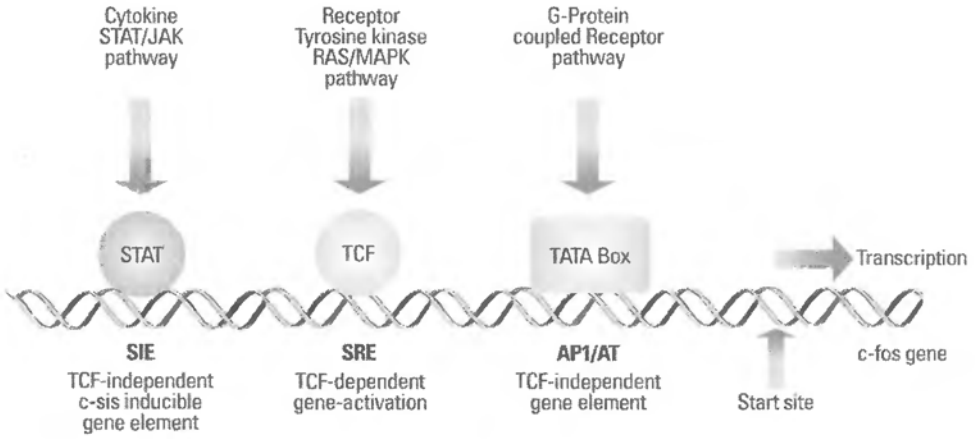


Fig. 10.1 The Raf–Ras–MAPK (ERK) pathway and other kinase cascades bring signals to specific transcriptional activators with the help of the ternary complex factor (TCF), and address the SRE-element. JNK pathways or G-protein linked pathways may induce the *c-fos* gene by TCF-independent mechanisms, addressing the AP1/AT promoter site, proximal to the SRE site and close to the TATA-box initiation site. The AT site is addressed by many eukaryotic transcription factors which interact with Pol II promoters. Cytokine signals, transmitted by the JAK/STAT pathway adress the *c-sis* inducible element, SIE, located distally to the SRE, in the *c-fos* promoter region. *c-sis* is a proto-oncogene that encodes the B chain of PDGF. *v-sis* is an oncogene from simian osteosarcoma virus. Coordinate activation of all these elements may be necessary for full induction of early response genes.

is made possible through interactions with TCFs[8] and other factors. For example, productive interaction of the serum response element (SRE) with the serum response factor (SRF), a transcription factor, requires assistance by an auxiliary factor, the ternary complex factor (TCF), which forms a ternary complex with SRF and SRE. Activity of TCFs is controlled by phosphorylation. Although phosphorylation activates transcription, it has no effect on binding of these factors to DNA.

The structure of the heterodimeric transcription factor Fos/Jun bound to DNA has been solved.[9] Jun contains a basic leucine zipper (bZIP) motif. Formation of a stable heterodimer is necessary for binding to DNA (see below).

CREB, a cyclic AMP- and Ca^{2+}-regulated transcription factor

A number of hormones and growth factors activate second messenger pathways and transmit signals to the gene through activation of second-messenger-dependent protein kinases, such as cAMP-dependent protein kinase A (PKA). These signals are targeted to the cAMP response element (CRE) by CREB (the CRE-binding protein) which receives the cAMP signals. This response element controls induction of numerous cellular genes. Transcriptional regulation by cAMP is illustrated in Fig. 10.2.

NF-AT and T-cell proliferation

The transcription factor, NF-AT (nuclear factor of activated T cells),[11,12] is essential for T-cell proliferation in the thymus,[13] because it regulates expression of the *IL-2* gene (see Chapter 14). NF-AT is representative of a cell-specific factor, controlling transcription in functionally differentiated cells. There are four known mammalian NF-AT genes and the corresponding NF-AT factors. Several sites within the DNA region, regulating expression of the *IL-2* gene, are addressed by NF-AT.[14] To activate these sites effectively, NF-AT cooperates with the Fos/Jun transcription factors.[15] There is an AP1 (Fos/Jun)-responsive promoter site immediately downstream and adjacent to the NF-AT sites in the promoter region of the *IL-2* gene.

The three-dimensional structure of the DNA-binding domains of NF-AT, Fos, and Jun in a quaternary complex with a DNA fragment from the *IL-2* gene promoter provides an explanation why more than one transcriptional regulator is necessary to give a productive

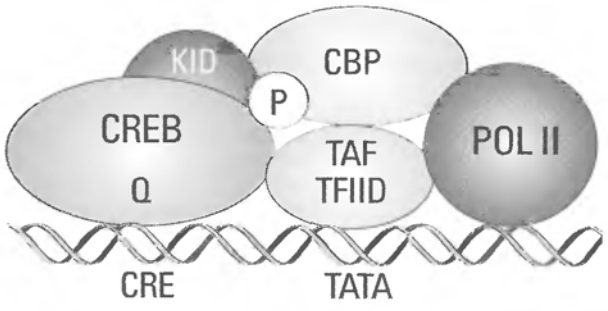


Fig. 10.2 CREB stimulates transcription of cAMP-responsive genes. These signals are transmitted through phosphorylation of the KID domain of CREB by the cAMP-responsive PKA, which introduces a phosphate into Ser133. The phosphoseryl residue binds to and recruits CBP the CREB-binding protein. CBP induces transcription by its association with the glutamine-rich Q2 region of CREB, which interacts with TAF, a component of TFIID, which hands the signal to the Pol II complex. CREB also accepts Ca^{2+} signals. (Reproduced with permission of the author and Annu. Rev. Biochem. from Fig. 3 in ref. 10.)

response (Plate 23).[16–20] The reason is that the interactions between the adjacently bound transcriptional proteins are cooperative and gene activation is synergistic. The quaternary complex consists of the bZIP elements of Fos and Jun, the Rel homology regions (RHRs)[21,22] of the DNA-binding domain of human NF-AT1, and a DNA sequence of the murine IL-2 promoter. *v-rel* is a retroviral oncogene, that causes reticulo-endotheliosis in birds (turkeys). The product of the proto-oncogene, Rel, is a transcription factor, localized in the nucleus. The DNA bound to the transcriptional complex contains the distal antigen-receptor response element 2 (ARRE-2).

NF-AT is controlled by phosphorylation. It is also a substrate of the Ca^{2+}-calmodulin-activated phosphatase, calcineurin,[23] and a target of immunosuppressive drugs.[24]

The DNA-binding region and the bZIP elements of Fos/Jun form a tight complex with each other and with the DNA. Although the segments of Fos and Jun that interact with DNA are single α-helices which stretch across the DNA major groove (just like other DNA complexes with bZIP proteins), when complexed with NF-AT, Fos/Jun is oriented differently with respect to the DNA. In the complex with NF-AT, the Fos/Jun heterodimer, AP-1, is oriented in a way that Jun can bind to a site closer to the NF-AT site.[17] This requires bending of the two α-helices of Fos and Jun and bending of the DNA to accommodate the Fos/Jun–NF-AT complex with DNA (compare with the information in refs 9 and 20).

NF-κB and NF-xB

Another member of the Rel family of transcription factors is NF-κB. In Plate 23 the structure of a NF-κB B-p52 complex with DNA is shown.[19] NF-κB is a multifunctional transcription factor. Related to NF-κB is NF-xB. NF-κB transmits neurotrophin- and NGF-signals.[27] It participates in the immune response, (Chapter 14), and plays a role in the nervous system.[25] The NGF/NF-κB pathway is involved in nerve-cell regeneration, and in the response to inflammation, NGF-$p75^{NTR}$ signals and cytokines activate the NF-κB signalling pathway.[28,29] NF-κB/NF-xB transmit IL-2 and IL-6, interferon-β, and neurotrophin signals to the genes.[26]

GATA

GATA transcription factors are representative of factors controlling cell differentiation. GATA-1 controls haematopoietic stem cells, GATA-3 is involved in the control of differentiation of thymocytes, and GATA-4 is a heart-specific transcription factor which has been implicated in the activation of genes responsible for cardiac hypertrophy.

Cell-specific regulation is carried out in cooperation with a common transcriptional activator, the CREB-binding protein (CBP) (see above) which cooperates with the cell-specific transcriptional activator, FOG (friend of GATA-1), and controls the activity of the transcription factor, GATA-1.

The sterol regulatory-element-binding proteins (SREBPs)

Tissue-specific transcription factors of great medical interest are the liver-specific, sterol regulatory-element-binding proteins 1 and 2, SREBP-1 and SREBP-2, which regulate expression of genes encoding the low-density lipoprotein (LDL) receptors. High fat intake downregulates the transcription of LDL receptor genes in the liver, and as a result, plasma levels of LDL rise and the risk of coronary heart disease greatly increases. The mechanism of action of these transcription factors was unravelled in the laboratory of M. Brown and J. Goldstein in Dallas.[30] The SREBPs are localized in the endoplasmic

reticulum, where they are anchored to the membrane by an hairpin-like segment in the middle of the protein. SCAP (the SREBP-cleavage-activating protein),[31] is a carrier protein that binds to SREBPs and brings them to the Golgi, where two proteases cleave the SREBPs sequentially, freeing the amino-terminal part of the SREBP. This has a basic helix-loop-helix motif (bHLH) and is the real transcription factor.[32,33] It now moves to the nucleus, where it activates gene transcription. The formation of the transcription factor is controlled by steroids. When sterols accumulate in the liver, proteolysis of the SREBPs is blocked, and the SREBPs remain in the ER. Consequently, the genes for the LDL receptor are not transcribed. Although the details of the regulation of SCAPs and SREBPs have not yet been worked out, transcriptional control by SREBPs is an impressive example of gene regulation by cellular translocation and of the role of protein processing by controlled proteolysis.

The MADs and SMADs

A key target of the MADs in Drosophila, (see Chapter 6), and SMADs in the regulation of morphogenesis are closely related homeobox early genes (*mix1* and *mix2*). SMAD 2 and SMAD 4, respectively MAD, and a SMAD 4 homologue in *Drosophila*, recruit a transcription factor, FAST 1, and form ARF (the activin response factor complex).[34] ARF addresses an activin-response element (ARE) in the promoter of the *mix2* gene. Figure 10.3 shows the crystal structure of the conserved C-terminal, MH2 domain of SMAD 4. The functional form of SMAD 4 is trimeric. Missense mutations in the MH2 region are found in cancer.

The STATs

STATs (signal transducers and activators of transcription) transmit signals primarily from cytokines (see Chapter 6).[36] Seven mammalian *stat* genes have been identified up to now and localized on chromosomes. Splicing of the pre-messenger RNA may lead to additional variants. STAT 2 participates (jointly with STAT 1) only in IFN-α/β signalling.[37] STAT 2 is present in nearly all cells, whereas the distribution of STAT 4 and STAT 5 is more restricted. STAT 4 is expressed in myeloid cells, where it is induced by IL-4 signals and where it participates in early myeloid differentiation.[38,39] STATs have

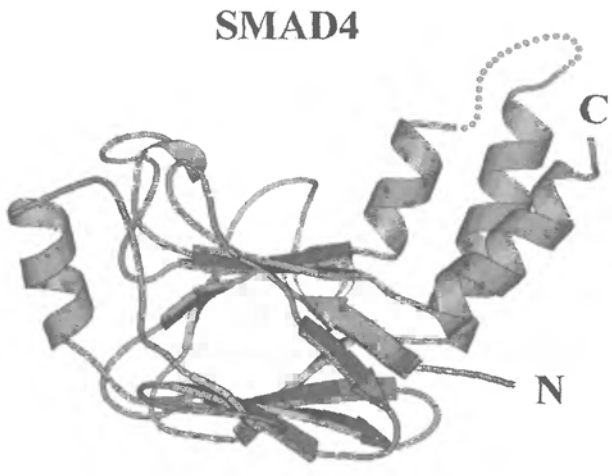


Fig. 10.3 The C-terminal domain (residues 319–552) of human SMAD 4 at 2.5 Å resolution It is a three-helix bundle attached to a β-sandwich scaffold (compare with the STAT 4 structure in Fig. 10.4). (Reproduced with permission of the authors and Nature from data in ref. 35, and the data available in the protein databanks.)

more than 50% sequence homology between residues 600 and 700. They have SH2 domains near the carboxy-termini and the region between amino acids 500 and 600 resembles SH3 domains. STATs form dimers and are regulated by phosphorylation. They bind cooperatively to promoter elements in target genes (Fig. 10.4).

STAT 1 transmits interferon-γ signals. STAT 1 and most other STATs recognize a DNA consensus sequence, the GAS sequence (the interferon-γ-activated sequence), 5′-TTN$_5$AA-3′ (where N represents any nucleotide) (Fig. 10.5a). Differences in the length of spacers between TT and AA in the GAS-element are thought to be responsible for discrimination between different STAT transcription factors. Interferon-α/β signals are transmitted through heterodimeric STAT 1/STAT 2, which form a complex with a protein, p48, and recruit ISGF3 (the interferon-stimulated gene factor 3), and address the interferon-stimulated response element (ISRE) (Fig. 10.5b). Without the assistance of the additional modifiers, p48 and ISGF3, the STAT 1 and 2 homo- and heterodimers could not recognize ISRE and would address the GAS sequence. Therefore, modifier proteins can change gene recognition.[42] There are more examples where ISGFs or IRFs (interferon-response factors) cooperate with STATs and other proteins and specify gene recognition.[43] For example, in response to stimulation by the glucocorticoid receptor, IRF-1 directs STAT 5 to the promoter of the β-casein gene in the lactating mammary gland. Another example is the cooperation of STAT 1 with the transcription specifier protein, SP1, in the activation of the promoter of the *I-CAM* gene (I-CAM is a cell adhesion molecule and the major receptor for rhinovirus). However, there are also interferon-stimulated signalling pathways that are not dependent on IRFs for gene targeting and expression.[44]

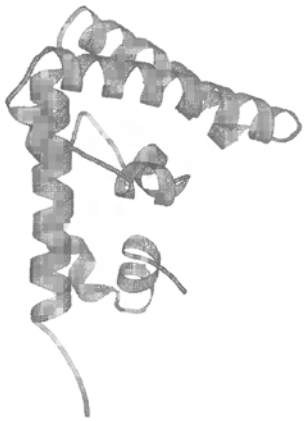


Fig. 10.4 The crystal structure, at 1.45 Å resolution, of the conserved NH_2-terminal domain, comprising the first 123 residues of STAT 4. The N-domain is instrumental in dimerizing the STAT molecules. The N-domain of one monomer is shown. Attention is directed to the hook-like structure of the dimer-forming domain, which consists of helices assembled in a way to form an extensive interface for polar interactions. Dimerization enables STATs to bind DNA cooperatively and selectively.[40] Mutagenesis of an invariant tryptophan residue at the heart of this interface abolished dimerization and cooperative DNA binding *in vitro* and reduced the transcriptional response to cytokine stimulation *in vivo*. (Reproduced with permission of the authors and Science from the data in ref. 41.)

Since even tyrosine kinase receptors (RTKs) and receptors for glucocorticoids occasionally engage STATs, and considering that the STATs address different genes in response to these diverse signals, it is unlikely that recognition of common response elements, such as GAS and ISRE, is enough to guarantee a specific response. For example, STAT 1 homodimers address GAS elements in response to INF-α/β, but also in response to INF-γ, although the cellular responses to INF-α/β and INF-γ are different. How can one and the same STAT bring about a different response, depending on the activating signal?

One possibility is an additional covalent modification, such as phosphorylation of serine 727 of STAT 1, which occurs only in response to INF-γ.[45] Dephosphorylation may also contribute to the specificity of INF signalling, because the tyrosine phosphatase, PTP1D, which is associated with IFNAR1 (the interferon-α receptor1) participates in the response to INF-α/β.[46]

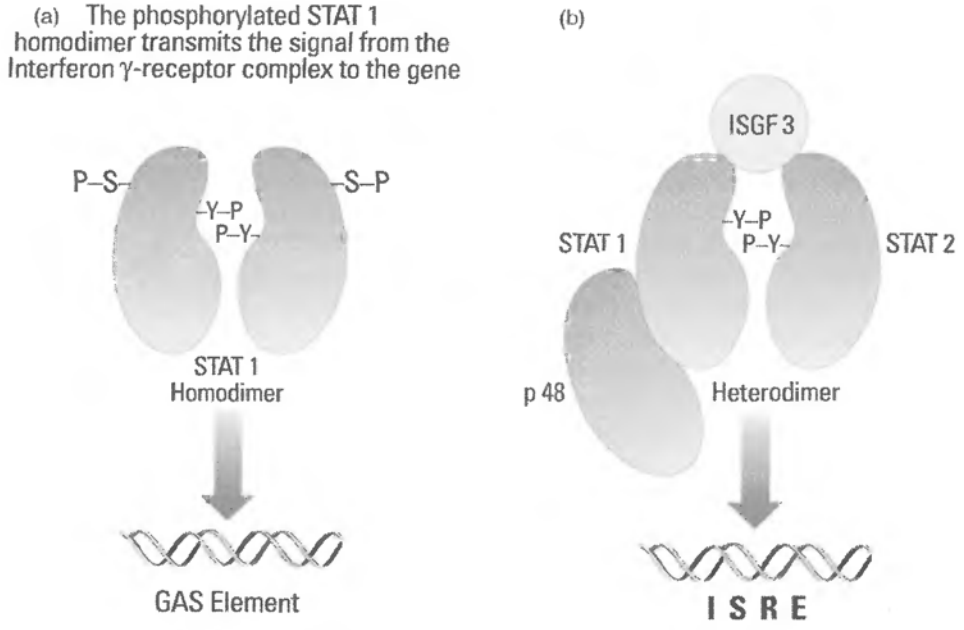


Fig. 10.5 (a) Homodimeric, phosphorylated STAT 1 and STAT 1/STAT 2 heterodimers bind in response to interferon-γ to the same DNA sequence element, GAS. (b) Interferon-α/β activates the transcription factor complex containing p48/ISGF3 and STAT 1/STAT 2 heterodimers. This transcription-factor complex addresses the interferon-stimulated regulatory element, ISRE.

Another way to improve specificity of gene recognition is by cooperation of STATs with additional factors. Accessory factors enlarge the gene recognition repertoire of STATs. STATs complexed with accessory proteins assume different conformations, changing or improving gene recognition. These transcriptional complexes may have different kinetic parameters, altering onset, duration, and efficacy of the gene response. Some proteins that help STATs (and SMADs), such as p300 and CBP, relax the chromatin structure near the transcription start sites and increase the rate of transcription. But these accessory factors also help STATs to recognize promoter elements and turn on a specific gene. For example, tyrosine-phosphorylated STAT 2 interacts with the large (*c.* 2500 amino acids long) nuclear protein, p300. However, when the same STAT 2 is serine/threonine phosphorylated, it binds to another accessory factor, CBP. This allows one and the same STAT to address different gene elements in response to different signals, causing different cellular responses. A further example is the interaction of STAT 1 with p300, which is essential for the antiviral action of INFs. Since STAT 1 can bind with the same carboxy-terminal site to both p300 and the adenoviral protein E1A , access of E1A to STAT 1 will be hindered in the presence of large-enough concentrations of p300, and the rate of transcription of genes encoding adenoviral proteins will be reduced. More factors modifying or enhancing gene recognition can be expected (Fig. 10.6).

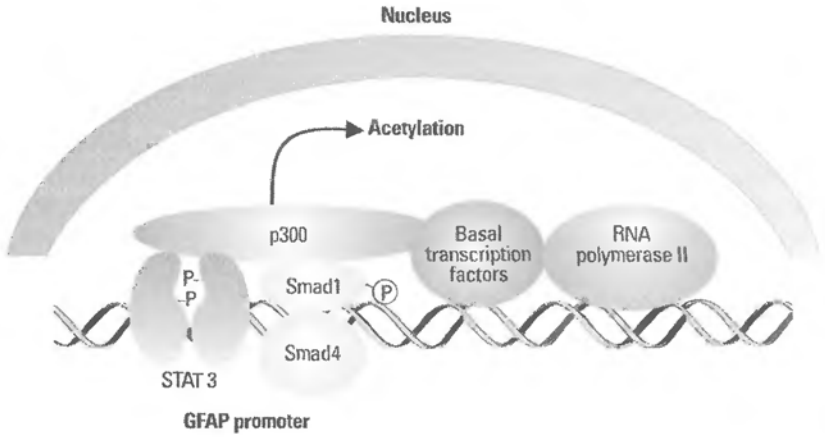


Fig. 10.6 How p300 in the nucleus may couple and integrate signals transmitted through STATs and SMADs. The example is the synergistic integration of the signalling pathways for the cytokines LIF (leukaemia-inhibiting factor) and BMP 2 (bone morphogenetic protein 2). The LIF signal is transmitted and propagated through the JAK/STAT pathway: STAT 3 is phosphorylated and dimerized and moves to the nucleus, where it binds to the promoter for the glial fibrillary acidic protein (GFAP). The BMP 2 signal is transmitted through a phosphorylated SMAD 1/SMAD 4 complex which also binds to the GFAP promoter. STAT 3 and SMAD 1 jointly recruit p300 which integrates the STAT/SMAD signals and activates transcription, by means of its ability to acetylate histone proteins and to loosen the chromatin structure. (Reproduced with permission of the authors and Science from the figure in ref. 47.)

Stress-activated transcription factors

Nearly every step in signal transfer from the cell membrane to the nucleus is sensitive to oxidants, because of the likelihood of oxidative damage of proteins.[48] ARE, a *cis*-acting genetic antioxidant response element, with the core sequence 5′-GTGAC$_3$GCA-3′, mediates gene activation in response to oxidative stress and xenobiotics.[49] An ARE is located in the 5′ flanking region of the glutathione *S*-transferase (GST) gene in the rat and thus regulates reductive detoxification.[50] The ARE core sequence is recognized by heterodimeric transcription factors, which are modified by oxidative stress. NF-κB and Jun–Fos are transcription factors which respond to oxidative stress signals.[51,52] But the basic leucine-zipper transcription factors, Nrf1 and Nrf2 (nuclear response factors), are essential for the cellular response to oxidative stress-inducing agents. The reason is that Nrfs have a key role in the induction of ARE and the expression of detoxifying enzymes, such as the enzymes glutathione *S*-transferase,[53] γ-glutamylcysteine synthetase,[54] or NADPH:quinone oxidoreductase.[55] Oxidative stress signals are often oxygen radicals, generated by cellular metallo-catalysed enzyme systems.[48,56] A stress-induced p38 SAP/MAP kinase has been implicated in the induction of drug-metabolizing, detoxifying enzymes.[57] A gene, *chop*, also known as growth arrest and DNA damage (GADD)-inducible gene, is expressed in response to stress of all kinds and in all cells. The product of the *chop* gene (CHOP) is activated by phosphorylation. A stress-induced p38 SAP/MAP kinase has also been implicated in CHOP phosphorylation, but this has not been proved.[58] (The cytosolic c-Abl tyrosine kinase is also activated in response to DNA

damaging agents.[59,60]). CHOP forms heterodimers with cell-specific transcription factors and blocks their function. Overexpression of CHOP can have detrimental consequences. Chaperones and heat-shock proteins also play an important role in stress responses.

To sum up: the response to oxidative stress is medically of great importance. It plays a role in ageing and in the protection against xenobiotics and cancerogenic agents, such as cigarette smoke. However, induction of detoxifying enzymes is also a problem of great interest for the pharmaceutical industry, because a too rapid elimination of therapeutic agents, especially of anti-cancer drugs, can make them ineffectual.

Summary

Structural information has provided us with a perception of how transcription factors cooperatively activate specific target sites on the DNA. The main mechanisms for the regulation of signal-responsive transcription factors are homo- and heterodimerization, formation of complexes with accessory factors, and phosphorylation–dephosphorylation. Phosphorylation also controls the translocation of transcription factors from the cytosol to the nucleus.

Control mechanisms: phosphorylation

Since each monomer of the Jun/Fos heterodimer can be phosphorylated separately, each can make an independent contribution to the transcriptional activity of the complex. For example, exclusive phosphorylation of a single threonine, Thr232, located in the N-terminal region of Jun by JNK (the Jun amino-terminal kinase), gives a characteristic and specific stress response.

Another important aspect of phosphorylation of transcription factors is the control of degradation. Phosphorylation of Jun by MAP kinase decreased ubiquitination, forestalling ubiquitin-dependent proteolysis, thus prolonging the lifetime of the transcription factor.[61] This could prolong transcription, with obvious consequences.

Sites of control

In the chain of signalling, originating from the receptor on the cell surface and ending at the gene, there are three major points of control:

(1) at the cell surface, at the level of ligand–receptor interactions;

(2) inside the cell, where many interactions with adaptors, linkers, and transducers occur and where transcriptional factors are directed to the nucleus; and

(3) at the level of the gene, where transcription factors and accessory regulatory factors must find specific gene promoter elements and enhancer sites.

Control at the receptor level

The lifetime of the active receptor complex determines the length of signalling. Thus, each process that affects the activation state of the receptor, by whatever means—processing of receptor and ligand, receptor oligomerization, phosphorylation, relocation, and other modifications—can, in principle, control signalling. Ligand–receptor coupling and activation are unique for each ligand–receptor pair, in terms of structure, kinetics, and thermodynamics. Thus, different responses are expected when a different cytokine or growth factor binds to a different receptor. This may explain why transcription, in

response to IFN-α, of the *irf-1* gene, encoding the interferon-response factor IRF-1, lasts for 2 hours, whereas the same gene is transcribed for 16 hours in response to INF-γ. Thus, even when the same gene is addressed, each of these different receptors may signal for a different length of time, although each receptor may use the same signalling path and engage the same transducers and transcriptional activators.

The possibilities of control of cellular signalling are nearly unlimited, because the activity of each component in a signalling cascade is subject to separate regulatory controls. Cross-talk between different signalling routes may mutually interfere, positively or negatively, with signalling, enhancing or reducing the final transcriptional response.

Control at the intracellular level: nuclear import

Most transcriptional activators are cytosolic proteins. Therefore, transport to the nucleus and export from the nucleus are points of transcriptional control in cells. Moreover, the time a transcriptional protein resides in the nucleus is a critical factor in the regulation of gene transcription.

Inward nuclear transport is carried out by proteins—the importins–that bring their cargo through the nuclear pore assembly into the nucleus. Here the cargo is unloaded and nuclear cargo is taken up by exportins and delivered to the cytoplasm. The monomeric G protein, Ran, controls nuclear transport. In the nucleus, Ran is in the GTP-bound form, and in the cytoplasm it is in the GDP-bound form. Outside the nucleus, Ran–GDP might help proteins to dock on the transportins, which then carry them through the nuclear-core complexes into the nucleus, where they are unloaded. Inside the nucleus, where the concentration of GTP is higher than in the cytoplasm, GDP bound to Ran is exchanged for GTP and Ran–GTP helps to load the carrier with nuclear cargo and, together with the exporter, it exits the nucleus and moves into the cytoplasm. Outside the nucleus Ran–GTP is converted to Ran–GDP by a Ran GAP and Ran is now ready for a new round of the transport shuttle (Fig. 10.7).[62]

Ran also regulates intracellular processing of ribosomal 5.8S RNA, suggesting that it may also participate in endoplasmic transfer of macromolecules, just as it is participating in nuclear–cytoplasmic transfer. Quite likely, regulation of nuclear transport may turn out to be more complicated, involving additional import and export signals and other factors, besides Ran.

Transcriptional proteins have specific nuclear localization sites (NLSs). For nuclear import, strategic sites in the transcription factor are modified by phosphorylation. As example, let us consider the transport from the cytosol into the nucleus of the transcription factors, NF-AT4 and NF-AT3. These transcriptional proteins are located in the cytoplasm. NF-AT4 is expressed in high levels in thymocytes. In dormant cells, their NLSs are masked by phosphorylation of specific serines. Translocation to the nucleus is controlled by Ca^{2+}. An increase in intracellular Ca^{2+} activates the protein phosphatase, calcineurin, which dephosphorylates and unmasks the nuclear localization site and initiates nuclear import of NF-AT4.[64,65] The fully dephosphosphorylated form of NF-AT4 is transcriptionally active. As long as calcium is supplied, calcineurin remains active and the dephosphorylated, transcriptionally active NF-AT4 transcription factor complex persists in the nucleus. But, when Ca^{2+} is no longer supplied, the activity of calcineurin declines and phosphorylation takes over. Consequently, NF-AT4 is exported from the nucleus and NF-AT4-mediated transcription stops. Nuclear export is controlled by NES (nuclear export sites or signals) and is carried out by an export receptor that recognizes

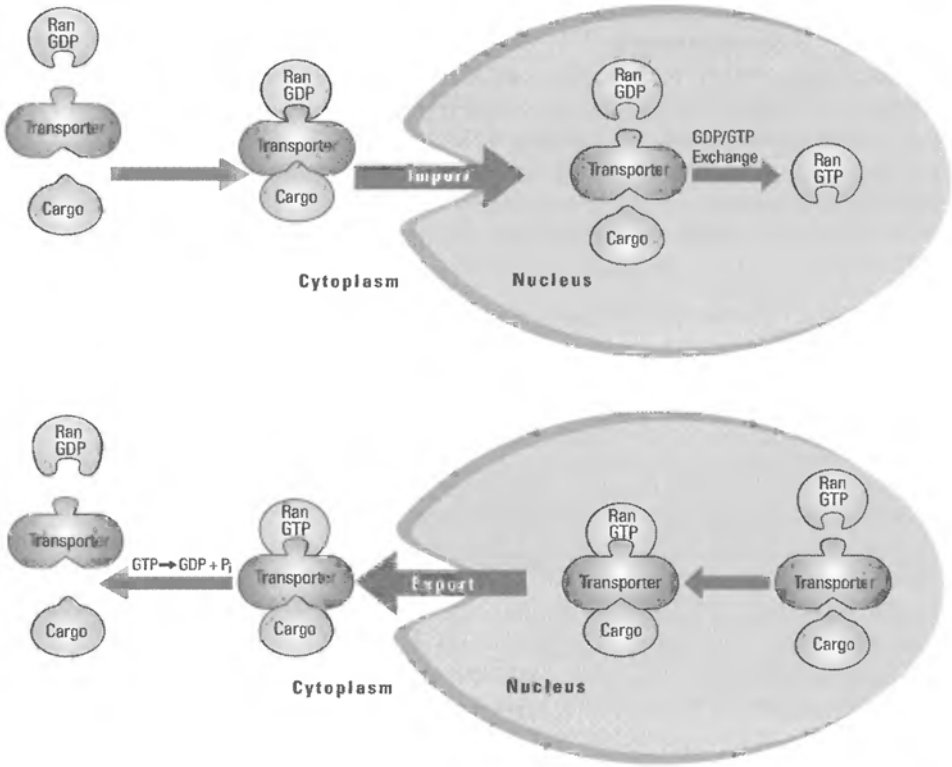


Fig. 10.7 Nuclear transport proceeds through core complexes and is mediated by transport receptors of the importin β-family. These receptors interact with the nuclear core complexes and shuttle between the nucleus and the cytoplasm. (Reproduced with permission of Elizabeth Pennisi and Science from the model in ref. 63.)

the NES. The export receptor and calcineurin compete with each other for binding to the transcription factor and determine the residence time of NF-AT4 in the nucleus (Fig. 10.8).[66,67]

NF-AT transcription factors are targets of immunosuppressive drugs, such as cyclosporin A and FK506, because they inhibit the calcineurin protein phosphatase and gene activation in T cells (see Chapter 7).

NF-AT3 also cooperates with GATA-4, a heart-specific transcription factor. When Ca^{2+} levels rise in a stressed heart, NF-AT3 moves to the nucleus and binds to GATA-4 and both turn on genes responsible for cardiac hypertrophy. The clinical importance of cardiac hypertrophy, which eventually causes heart failure, makes the development of drugs that act like immunosuppressives, preventing nuclear import of NF-AT3, an active area of research into control of cardiac hypertrophy.[69]

Control at the level of the gene

When one compares the most abundant proteins in cells from many different cell types, one finds that the majority have similar structure and function. Only perhaps 1000 are different and are responsible for the large differences in phenotype, morphology, and

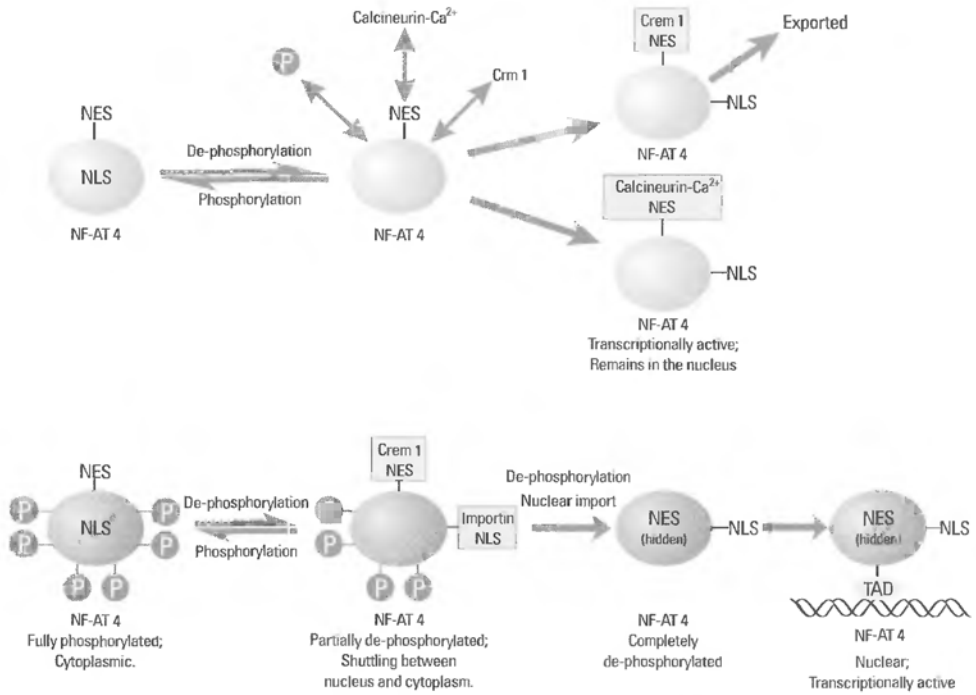


Fig. 10.8 Above, import of the transcription factor NF-AT4 into the nucleus. In activated cells, import is initiated by calcineurin-mediated dephosphorylation of NF-AT4. Dephosphorylation unmasks the nuclear-localization signal (NLS), and at the same time blocks the nuclear export signal (NES). The NES is recognized by the exportin protein (Crm1). Nuclear export is an active process. Moreover, nuclear export requires rephosphorylation of the NF-AT4 transcription factor. It is indicated that dephosphorylation by calcineurin and nuclear export are mutually exclusive, because calcineurin and Crm 1 compete for a common binding site on NES. When NES binds to Crm1, NT-AT4 is exported from the nucleus, and when calcineurin binds to NES, NF-AT4 remains in the nucleus and forms a transcriptionally active complex. Below, how the extent of dephosphorylation controls the transcriptional activity of NF-AT4. When NF-AT4 is fully phosphorylated, NLS is hidden and the transcription factor remains in the cytoplasm. When NF-AT4 is only partially dephosphorylated, NLS is exposed and can interact with importin a/b which promote nuclear import, and at the same time, NES can interact with the exportin Crm1, which promotes nuclear export. The consequence is that the transcription factor shuttles between the nucleus and the cytoplasm and is not transcriptionally active. In order to become transcriptionally fully active, NF-AT4 must be completely dephosphorylated. This prevents export from the nucleus by blocking NES, and may increase the affinity of the transcription factor for DNA by exposure of its *trans*-activating domain (TAD). (The entire scheme is reproduced with permission of Drs Patrick G. Hogan and Anjana Rao and Nature from Fig. 1 in ref. 68.)

function of differentiated cells. This agrees with the fact that only a relatively small number of DNA sites have been found which are unique for each different cell line. Sites which are likely points of control in cell-specific gene expression are the 'locus control'

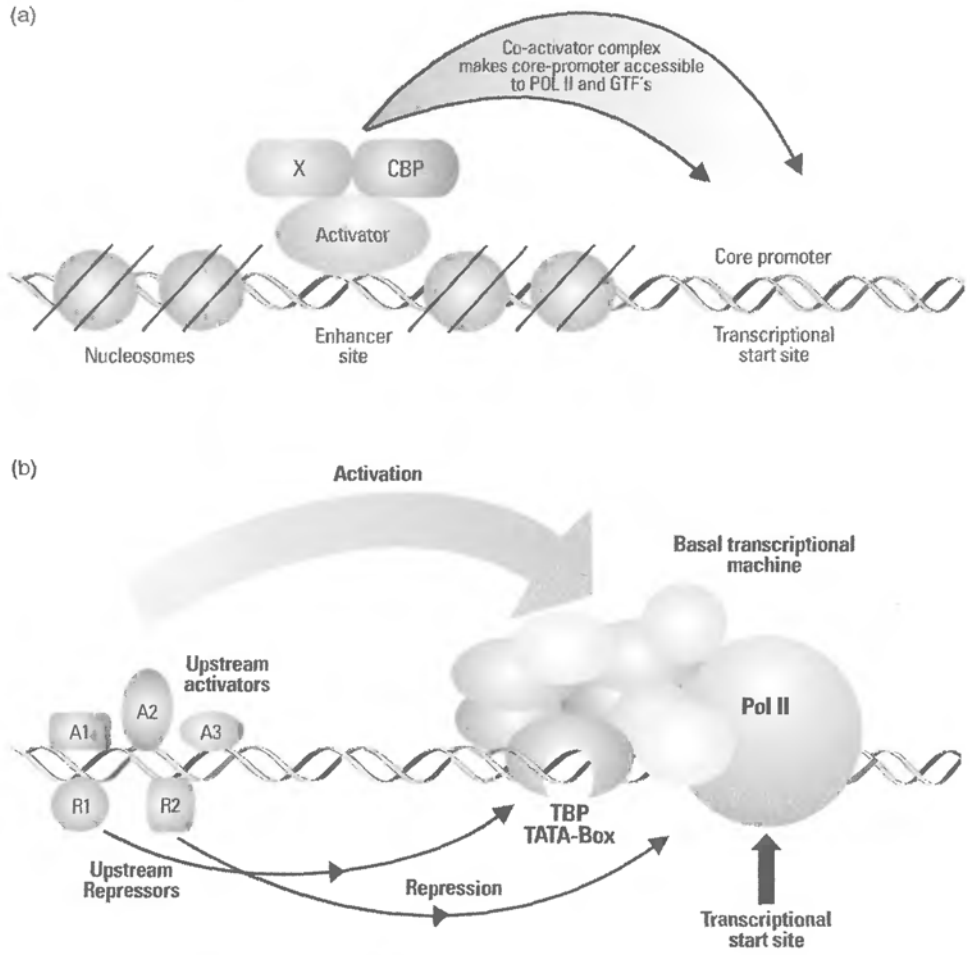


Fig. 10.9 (a) A transcriptional regulator protein bound to a distal enhancer recruits the co-activator CBP and possibly other co-activators, indicated by X. An example is the synergistic activation of transcription by CBP and the tumour suppressor p53.[71] CBP is a subunit of a larger co-activator complex, together with p/CAF, the CBP-associated acetyl transferase. The CBP complex is a histone acetylase which may decondense the chromatin and may make the core promoter accessible to the Pol II and the general basal transcription factors (GTFs). The distal enhancer site, to which the signal-dependent transcriptional activator binds, can be over 1 kb apart from the transcriptional start site. (Reproduced with permission of the authors and Science from Fig. 1 in ref. 71.) (b) How various upstream activators, A1, A2, A3, and repressors, R1, R2, bind to upstream regulator sequences and signal to different components of the basal transcriptional apparatus. Positive signals of the activators bring Pol II to the transcriptional start site, and negative signals prevent this. The scheme is simplified. Probably, activators and repressors first recruit various co-activators and co-repressors, which then regulate the basal Pol II transcriptional complex. Signals may be integrated distantly from the transcriptional start site, at the level of *trans*-regulatory DNA elements, as shown here, or they may be integrated, close to the transcriptional start site, at the level of *cis*- regulatory DNA elements. (Reproduced with permission of the authors and Science from Fig. 3 in ref. 71.)

sites in certain genes. The term 'locus control' was originally introduced for the immunoglobin genes. 'Locus control' regions are regulatory DNA sequences. They have effects over large distances, greater than 50 kb, and control distant genes. The core of 'locus control' regions contains CACC- and GATA-like sequences, such as those recognized by GATA and Fos–Jun-type transcription factors. Although these sequences are important probably no one particular sequence in the promoter region of any particular gene is sufficient. Many factors must cooperate, in a time-coordinated manner, to achieve cell-specific gene expression.

Accessory proteins that couple to transcriptional activators improve the specificity of gene recognition. Cooperation of transcription factors with co-activators is necessary, because each receptor does not enlist its own unique transcriptional activator. Although the number of transcription factors being discovered is rising, it probably will remain far less than the number of receptors from which they receive signals. This means that the same transcription factors are turned on by different receptors, but with different outcomes. Therefore, to make differential gene expression and a specific response possible, under these circumstances, requires cooperation with a multitude of modifying factors. Many different factors must cooperate, in a time-coordinated manner and in a partly positive and partly negative fashion, in order to realize specific gene expression and unique cellular responses, notably in the regulation of developmental processes (Fig. 10.9). A particularly interesting example of gene control at the level of the transcription factor are the *p*ituitary-specific *t*ranscription factors, the PIT's. PIT1 has a domain, (the POU domain, quite common among transcription factors) and binds to the 5′-TAAT-3′ consensus sequence. PIT1 with a pituitary-specific regulatory POU domain activates the growth hormone, GH, gene in the GH-producing cells, (somatotropic cells in the pituitary). But, surprisingly, PIT suppressed the GH gene in the prolactin-producing, lactotropic cells. The differences in the interaction area of PIT with the promoter sequences of the GH genes, in the two cell types, the somatotropic and the lactotropic cells involved only the spacing of two base pairs, as shown by X-ray crystallography. This small change in the rapprochement of the POU domain of PIT to the promoter sequence was enough to trigger a conformational transition in the transcription factor that altered the accessibility of PIT for co-repressors, resulting in the recruitment of *N-CoR*, the nuclear receptor co-repressor and persistent inactivation of the GH gene in prolactin-producting cells. (See [272]).

An important control mechanism of transcriptional activity involves the chromatin structure. Transcription is inhibited during mitosis, because transcription factors are displaced from promoter sequences as well as from bulk chromatin. Displacement of transcription factors from mitotic chromosomes is one of the means by which the cell cycle resets transcriptional programmes[73] (we shall come back to this in Part 3). In cell differentiation programmes, specific timing mechanisms are operative in gene expression, just like those governing the cell cycle (see Chapter 12). The time-controlled on/off switches of the activity of genes are still not understood fully. All the components participating in these control circuits must be identified, and of equal importance will be a more quantitative analysis of protein–protein and protein–DNA interactions in terms of kinetics, affinities, and amplitude of response. Although the structural and functional consequences of the interaction of transcription factors with specified gene sites are just beginning to be understood, there is every reason to expect that the time-ordered and controlled expression of the genetic information that decides the fate of the cell will be clarified in due course.

The future belongs to proteomics

The completion of the genome sequences of *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Drosophila*, and finally of the human genome,[74] has shifted emphasis more and more to the functional identification of the encoded proteins. To emphasize this, in the yeast genome about one-third of the open reading frames (ORFs) code for proteins with unknown function[75] (open reading frames contain that part of the genetic information which can be expected to have meaningful information about proteins). This is the place to make a few remarks on the state of the art of the techniques,[76–78] for the identification of proteins and the interactions between proteins. One approach uses genetic methods, such as two-hybrid analysis, where each open reading frame is excised and inserted into a transcriptional activation domain in the genome. In this way 6000 yeast transformants have been engineered and then screened for proteins. The proteins were identified by sequence analysis, and about 1000 of them were found to interact with other proteins, which were also sequenced. These techniques are now complemented by new biochemical methods, employing mass spectrometry for identification and screening, and separation techniques that do not destroy protein– protein interactions. Together, these powerful methods, combined with bioinformatics and usually lumped together and called 'proteomics', should eventually be able to identify the functions of all individual proteins, together with their interacting partners, and in that way help to identify their role in a living cell. Only when this has been achieved, will the effort made in deciphering the human genome bring the hoped-for pay-off.

References

1. R. G. Roeder. Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. *Cold Spring Harbor Symp*, **63**, 201–218, 1998.
2. R. G. Roeder, J. Fondell, W. Gu, M. Guermah, M. Ito, U. Kim, *et al.* The role of ubiquitous and cell specific coactivators in transcription regulation. *FASEB J*, **12** (8), A 1318, 1998.
3. M. Ptashne and A. Gann. Transcriptional activation by recruitment. *Nature*, **386**, 569–577, 1997.
4. C. S. Hill and R. Treisman. Differential activation of c-fos promoter by serum lysophosphatidic acid, G proteins, and polypeptide growth factors. *EMBO J*, **14**, 5037–5047, 1995.
5. T. Kerppola and T. Curran. Transcription. Zen and the art of Fos and Jun [news]. *Nature*, **373** (6511), 199–200, 1995.
6. D. S. Ladtman. *Eukaryotic transcription factors*, Chapter 7, Transcription factors and cancer, pp. 153–176. Academic Press, London, 1991.
7. A. D. Smith (managing editor). *Oxford dictionary of biochemistry and molecular biology*, Revised edition. Oxford University Press, Oxford, 2000.
8. R. Treisman. Ternary complex factors: growth factor regulated transcriptional activators. *Curr Opin Genet Dev*, **4**, 96–101, 1994.
9. J. N. Glover and S. C. Harrison. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature*, **373**, 257–261, 1995.
10. M. Montminy. Transcriptional regulation by cAMP. Annu Rev Biochem, **66**, 807–822, 1997.
11. A. Rao, C. Luo, and P. G. Hogan. Transcription factors of the NF-AT family: regulation and function. Annu Rev Immunol, **15**, 707–747, 1997.
12. J. P. Northrop, *et al.* NF-AT components define a family of transcription factors targeted in T-cell activation. *Nature*, **369**, 497–502, 1994.
13. E. S. Masuda, *et al.* NF-ATx, a novel member of the nuclear factor of activated T cells family that is expressed predominantly in the thymus. *Mol Cell Biol*, **15**, 2697–2706, 1995.
14. T. Hoey, Y. L. Sun, K. Williamson, and X. Xu. Isolation of two new members of the NF-AT gene family and functional characterization of the NF-AT proteins. *Immunity*, **2**, 461–472, 1995.

15. J. Jain, P. G. McCaffrey, V. E. Valge-Archer, and A. Rao. Nuclear factor of activated T cells contains Fos and Jun. *Nature*, **356**, 801–804, 1992.

16. L. Chen, J. N. M. Glover, P. G. Hogan, A. Rao, and S. C. Harrison. Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature*, **392**, 42–48, 1998.

17. L. Chen, *et al.* Only one of the two DNA-bound orientations of AP-1 found in solution cooperates with NFATp. *Curr Biol*, **5**, 882–889, 1995.

18. M. Chytil and G. L. Verdine. The Rel family of eukaryotic transcription factors. *Curr Opin Struct Biol*, **6**, 91–100, 1996.

19. P. Cramer, C. J. Larson, G. L. Verdine, and C. W. Muller. Structure of the human NF-κB p52 homodimer–DNA complex at 2.1 Å resolution. *EMBO J*, **16**, 7078–7090, 1997.

20. P. König and T. J. Richmond. The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. *J Mol Biol*, **233**, 139–154, 1993.

21. S. A. Wolfe, *et al.* Unusual Rel-like architecture in the DNA-binding domain of the transcription factor NFATc. *Nature*, **385**, 172–176, 1997.

22. J. Jain, E. Burgeon, T. M. Badalian, P. G. Hogan, and A. Rao. A similar DNA-binding motif in NFAT family proteins and the Rel homology region. *J Biol Chem*, **270**, 4138–4145, 1995.

23. J. Jain, *et al.* The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. *Nature*, **365**, 352–355, 1993.

24. P. G. McCaffrey, *et al.* Isolation of the cyclosporin-sensitive T cell transcription factor NFATp. ATp. *Science*, **262**, 750–754, 1993.

25. J. Kuriyan and D. Thanos. Structure of the NF-kappa B transcription factor: a holistic interaction with DNA. *Structure*, **3** (2), 135–141, 1995.

26. U. Otten and R. A. Gadient. Neurotrophins and cytokines—intermediaries between the immune and nervous systems. *Int J Dev Neurosci*, **13** (3–4), 147–151, 1995.

27. B. D. Carter, C. Kaltschmidt, B. Kaltschmidt, N. Offenhauser, R. Bohm Matthaei, P. A. Baeuerle and Y. A. Barde. Selective activation of NF-kappa B by nerve growth factor through the neurotrophin receptor p75 [see comments]. *Science*, **272** (5261), 542–545, 1996.

28. C. J. Woolf, B. Safieh Garabedian, Q. P. Ma, P. Crilly, and J. Winter. Nerve growth factor contributes to the generation of inflammatory sensory hypersensitivity. *Neuroscience*, **62** (2), 327–331, 1994.

29. G. R. Lewin and L. M. Mendell. Nerve growth factor and nociception. *Trends Neurosci*, **16** (9), 353–359, 1993.

30. M. S. Brown and J. L. Goldstein. The SREBP pathway: Regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell*, **89**, 331–340, 1997.

31. X. Hua, A. Nohturfft, J. L. Goldstein, and M. S. Brown. Sterol resistance in CHO cells traced to point mutation in SREBP cleavage acrivating protein, (SCAP). *Cell*, **87**, 415–426, 1996.

32. J. Sakai, R. B. Rawson, P. J. Espenhade, D. Cheng, A. C. Seegmiller, J. L. Goldstein, and M. S. Brown. Molecular identification of the sterol-regulated luminal protease that cleaves SREBP's and controls lipid composition of animal cells. *Molecular Cell*, **2**, 505–514, 1998.

33. E. A. Duncan, U. P. Dav, J. Sakai, J. L. Goldstein, and M. S. Brown. Second-site cleavage in sterol regulatory element binding protein occurs at transmembrane junction as determined by cysteine panning. *J Biol Chem*, **273**, 17801–17809, 1998.

34. F. Liu, C. Pouponnot, and J. Massague. Dual role of the Smad4/DPC4 tumor suppressor in TGFbeta-inducible transcriptional complexes. *Genes Dev*, **11** (23), 3157–3167, 1997.

35. Y. Shi, A. Hata, R. S. Lo, J. Massagué, and N. P. Pavletich. A structural basis for mutational inactivation of the tumoursuppressor SMAD 4. *Nature*, **388**, 87, 1997.

36. Y. Han, D. W. Leaman, D. Watling, N. C. Rogers, B. Groner, I. M. Kerr, *et al.* Participation of JAK and STAT proteins in growth hormone-induced signaling. *J Biol Chem*, **271** (10), 5947–5952, 1996.

37. J. E. Darnell Jr, I. M. Kerr, and G. R. Stark. Jak-STAT pathways and transcriptional activation in response to IFN's and other extracellular signalling proteins. *Science*, **264**, 1415–1421, 1994.

38. J. Hou, U. Schindler, W. J. Henzel, T. C. Ho, M. Brasseur, and S. L. McKnight. An interleukin-4-induced transcription factor: IL-4 Stat. *Science*, **265**, 1701–1706, 1994.

39. K. Yamamoto, F. W. Quelle, W. E. Thierfelder, B. L. Kreider, D. J. Gilbert, N. A. Jenkins, *et al.* Stat 4 a novel gamma interferon activation site- binding protein expressed in early myeloid differentiation. *Mol Cell Biol*, **14**, 4342–4349, 1994.

40. X. Xu, Y.-L. Sun, and T. Hoey. Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain. *Science*, **273**, 794–797, 1996.

41. U. Vinkemeier, I. Moarefi, J. E. Darnell Jr, and J. Kuriyan. Structure of the amino-terminal protein interaction domain of STAT-4. *Science*, **279**, 1048–1052, 1998.

42. G. R. Stark, I. M. Kerr, B. R. Williams, R. H. Silverman and R. D. Schreiber. How cells respond to interferons. *Annu Rev Biochem*, **67**, 227–264, 1998.

43. H. A. R. Bluyssen, R. Muzaffar, R. J. Vliestra, A. C. J. van der Made, S. Leung, G. R. Stark, *et al.* Combinatorial association and abundance of components of interferon-stimulated gene factor 3 dictate the selectivity of interferon responses. *Proc Natl Acad Sci, USA*, **92**, 5645–5649, 1995.

44. S. J. Haque and B. R. Williams. Identification and characterization of an interferon (IFN)-stimulated response element-IFN-stimulated gene factor 3-independent signaling pathway for IFN-alpha. *J Biol Chem.* **269** (30), 19523–19529, 1994.

45. Z. Wen, Z. Zhong, and J. E. Darnell Jr. Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell*, **82** (2), 241–250, 1995.

46. M. David, G. Zhou, R. Pine, J. E. Dixon, and A. C. Larner. The SH2 domain-containing tyrosine phosphatase PTP1D is required for interferon alpha/beta-induced gene expression. *J Biol Chem*, **271** (27), 15862–15865, 1996.

47. J. Janknecht and T. Hunter. Nuclear fusion of signalling pathways. *Science*, **284**, 443–444, 1999.

    Also: K. Nakashima, M. Yanagisawa, H. Arakawa, N. Kimura, T. Hisatsune, M. Kawabata, et al. Synergistic signalling in fetal brain by STAT 3-Smad 1 complex bridged by p300. *Science*, **284**, 479–482, 1999.

48. E. R. Stadtman. Protein oxidation and aging. *Science*, **257** (5074), 1220–1224, 1992.

49. W. W. Wasserman and W. E. Fahl. Functional antioxidant responsive elements. *Proc Natl Acad Sci, USA*, **94** (10), 5361–5366, 1997.

50. S. Liu and C. B. Pickett. The rat liver glutathione S-transferase Ya subunit gene: Character of the binding properties of a nuclear protein from Hep G2 cells that have affinity for the antioxidant response element. *Biochemistry*, **35** (35), 11517–11521, 1996.

51. P. C. Fernandez, J. Machado Jr, V. T. Heussler, C. Botteron, G. H. Plamer, and D. Dobbelaere. The inhibition of NF-κB activation pathways and the induction by thiocarbamates in T cells are blocked by the glutathione precursor *N*-acetyl-L-cysteine. *Biol Chem*, **380** (12), 1383–1394, 1999.

52. J. M. Muller, M. A. Cahill, R. A. Rupec, P. A. Baeuerle, and A. Nordheim. Antioxidants as well as oxidants activate c-fos via Ras-dependent extracellular-signal-regulated kinase 2 and Elk-1. *Eur J Biochem*, **244** (1), 45–52, 1997.

53. T. Nguyen, T. H. Rushmore, and C. B. Pickett. Transcriptional regulation of glutathione S-transferase: Analysis of the antioxidant response element and its activation by 12-O-tetradecanoylphorbol-13-acetate. *J Biol Chem*, **269** (18), 13656–13662, 1994.

54. M. Kwong, Y. W. Kan, and J. Y. Chan. The CNC basic leucine zipper factor Nrf1, is essential for cell stimulation to oxidative stress-inducing agents, Role for Nrf1 in γ-glutamylcysteine synthetase expression in mouse fibroblasts. *J Biol Chem*, **274** (52), 37491–37498, 1999.

55. R. Venugopal and A. K. Jaiswal. Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate anti-oxidant response element mediated expression of NAD(P)H: quinone oxidoreductase1 gene. *Proc Natl Acad Sci, USA*, **93** (25), 4960–4965, 1996.

56. C. K. Sen and L. Packer. Antioxidant and redox regulation of gene transcription [see comments]. *FASEB J*, **10** (7), 709–720, 1996.

57. R. Yu, S. Mandlekar, W. Lei, W. E. Fahl, T. H. Tan, and A. T. Kong. p38 mitogen-activated protein kinase negatively regulates the induction of drug-metabolizing enzymes that detoxify carcinogens. *J Biol Chem*, **275** (4), 2322–2327, 2000.

58. X. Z. Wang and D. Ron. Stress-induced phosphorylation and activation of the transcription factor CHOP (GADD153) by p38 MAP Kinase. *Science*, **272** (5266), 1347–1349, 1996.

59. S. Kharbanda, R. Ren, P. Pandey, T. D. Shafman, S. M. Feller, R. R. Weichselbaum, and D. W. Kufe. Activation of the c-Abl tyrosine kinase in the stress response to DNA-damaging agents. *Nature*, **376** (6543), 785–788, 1995.

60. Z.-M. Yuan, Y. Huang, Y. Whang, C. Sawyers, R. Weichselbaum, S. Kharbanda and D. Kufe. Role for c-Abl tyrosine kinase in growth arrest response to DNA damage. *Nature*, **382**, 272–274, 1996.

61. A. M. Musti, M. Treier, and D. Bohmann. Reduced ubiquitin-dependent degradation of c-Jun after phosphorylation by MAP kinases. *Science*, **275**, 400–402, 1997.

62. F. Melchior and L. Gerace. Two-way trafficking with Ran. *Trends Cell Biol*, **8**, (5), 175–179, 1998.

63. E. Pennisi. The nucleus's revolving door. *Science*, **279**, 1129–1131, 1998.

64. C. R. Beals, N. A. Clipstone, S. N. Ho, and G. R. Crabtree. Nuclear localization of NF-ATc by a calcineurin-dependent, cyclosporin-sensitive intramolecular interaction. *Genes Dev*, **11** (7), 824–834, 1997.

65. F. Shibasaki, E. R. Price, D. Milan, and F. McKeon. Role of kinases and the phosphatase calcineurin in the nuclear shuttling of transcription factor NF-AT4. *Nature*, **382**, 370–373, 1996.

66. M. Ohno, M. Fornerod, and I. W. Mattaj. Nucleocytoplasmic transport: the last 200 nanometers. *Cell*, **92** (3), 327–336, 1998.

67. J. Zhu and F. McKeon. NF-AT activation requires suppression of Crm 1-dependent export by calcineurin. *Nature*, **398**, 256–260, 1999.

68. P. G. Hogan and A. Rao. Modification by nuclear export? News and Views. *Nature*, **398**, 200–201, 1999.

69. M. Barinaga. Signaling path may lead to better heart-failure therapies; Research News. *Science*, **280**, 383, 1998.

70. W. Gu, X. L. Shi, and R. G. Roeder. Synergistic activation of transcription by CBP and p53. *Nature*, **387** (6635), 819–882, 1997.

71. M. Mannervik, Y. Nibu, H. Zhang, and M. Levine. Transcriptional co-regulators in development. *Science*, **284**, 606–609, 1999.

72. K. M. Scully, E. M. Jacobson, K. Jepsen, V. Lunyak, H. Viadiu, C. Carrière, D. W. Rose, F. Hooshmand, A. K. A. Aggarwal, M. G. Rosenfeld. Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification, *Science*, **290**, 1127–1131, 2000.

73. M. A. Martinez Balbas, A. Dey, S. K. Rabindran, K. Ozato, and C. Wu. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell*, **83** (1), 29–38, 1995.

74. A. Goffeau, B. G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver, Life with 6000 gene, *Science*, **274**, (5287) 563–567, 1996.

75. P. Uetz, *et al.* A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623–627, 2000.

76. M. Fromont-Racine, J. C. Rain, and P. Legrain. Toward a functional analysis of the yeast genome through exhaustive two hybrid screens. *Nature Genet.*, **16**, 277–282, 1997.

77. A. Flores *et al.* A protein–protein interaction map of yeast RNA polymerase III. *Proc Natl Acad Sci, USA*, **96**, 7815–7820, 1999.

78. S. Fields and O. Song. A novel genetic system to detect protein–protein interactions. *Nature*, **340**, 245–246, 1989.

11

Regulation of gene transcription by hormones

Steroid hormone receptors

Receptors for lipophilic hormones are transcription factors

Water-soluble hormones, growth factors, and cytokines must bind first to cell-surface receptors, but lipophilic, fat-soluble hormones can pass through the membrane lipid bilayer. Once inside the cell, they are recognized by diffusible receptor proteins which bring them to the nucleus where they control gene transcription. Thus, regulation of gene transcription by lipophilic hormones seems quite simple. There is, in principle, no need for transducers, linkers, and adaptors for transmission of a signal.

The classical model of steroid hormone action postulates that binding of the hormone to a diffusible, intracellular receptor induces a conformational transition of the binding protein, making it capable of binding with high affinity to specific DNA sites, the hormonally regulated target genes.

Nuclear hormone receptors

The cDNAs for the glucocorticoid and the oestrogen receptors were isolated more than 15 years ago. They were among the first genes, coding for transcriptional gene activators, to be identified. The family of nuclear receptors is the largest family of transcription factors. Until now, more than 150 different members of the superfamily of nuclear receptors, from worms to insects to humans, have been described. The discovery of an insect receptor for a steroid hormone, ecdysone, indicated that this kind of receptor must have evolved prior to the separation of vertebrates and invertebrates.

Nuclear receptors are distinguished on the basis of their structural properties, whether they are monomers, homo- or heterodimers, and with respect to the kind of DNA sequences they recognize. Based on these criteria, the nuclear receptor superfamily can be divided broadly into four classes as shown in Table 11.1:

1. Class I receptors are homodimers. Dimerization is ligand-induced and these receptors bind to DNA half-sites with inverted, palindromic repeats.

Table 11.1 Nuclear receptor superfamily

Class	Receptor	Receptor Structure	Ligand	HPE
I	Glucorticoid receptor (GR)	Ligand-induced homodimer	Cortisol	Palindromic inverted repeat
	Mineralocorticoid receptor (MR)	Ligand-induced homodimer	Aldosterone	Palindromic inverted repeat
	Progesterone receptor (PR)	Ligand-induced homodimer	Progesterone	Palindromic inverted repeat
	Androgen receptor (AR)	Ligand-induced homodimer	Testosterone	Palindromic inverted repeat
	Oestrogen receptor	Ligand-induced homodimer	Oestrogen	Palindromic inverted repeat
II	Thyroid receptor (TR)	Heterodimer	Triiodothyronine, thyroxin	Direct repeat
	Vitamin D_3 receptor (VDR)	Heterodimer	1,25-Dihydroxy vitamin D_3	Direct repeat
	9-*cis*-retinoic acid receptor (R×R)	Heterodimer	9-*cis*-retinoic acid	Direct repeat
	All-*trans*-retinoic acid receptor (RAR)	Heterodimer	All-*trans*-retinoic acid	Direct repeat
	Chicken ovalbumin, Upstream promotor transcription factor (COUP-TF)	Heterodimer	?	Direct repeat
	Peroxysome proliferation activating receptor (PPAR)	Heterodimer	?	Direct repeat
	Ultra spinacle protein (USP)	Heterodimer	?	Direct repeat
III	9-*cis*-retinoic acid receptor (R×R)	Homodimer	9-*cis* retinoic acid	Direct repeat
	Orphan receptor	Homodimer	?	Direct repeat
IV	Receptors forming complexes with other transcription factors	Monomer	?	DNA extended core site

2. Class II receptors are heterodimers. To this group belong receptors for non-steroidal compounds, such as the retinoids. The heterodimers are composed, for example, of a 9-*cis*-retinoic-acid-bound monomer combined with an all-*trans*-retinoic-acid-bound monomer or with a thyroid-hormone-bound monomer. These heterodimeric receptors bind to hexameric DNA repeats.
3. Class III receptors bind, like class II receptors, to direct DNA repeats, but differ from class II receptors because they can bind as homodimers. Examples are homodimeric 9-*cis*-retinoic acid receptors (RXRs). Most of the orphan nuclear receptors also belong to this class.
4. Class IV receptors are monomers which bind in concert with other proteins to extended DNA core sites.

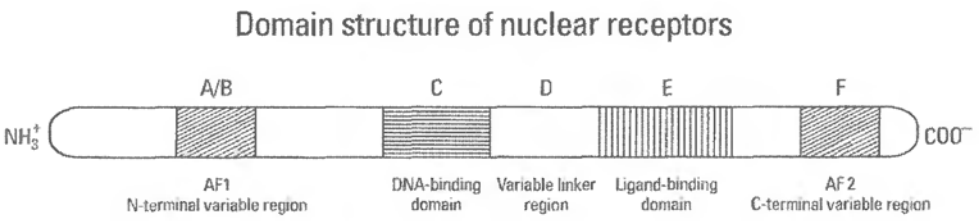


Fig. 11.1 Functions assigned to the structural regions of nuclear receptors. Six regions (A–F) are distinguished. Regions C, E, and D are highly conserved, whereas regions A/B and F are more diverse. A/B contains a transcriptional activation function, AF-1. The conserved central region (66 amino acids long), C, contains the core of the DNA-binding domain (DBD) with two zinc-finger-like binding motifs. The 220-amino-acid-long region E, in the C-terminal half of the polypeptide chain, is also highly conserved and contains the ligand-binding domain (LBD), and adjacent to it a ligand-responsive transcriptional activation function (AF-2) and a dimerization surface. Region D is located between the DBD and the LBD regions. (Derived from information in ref. 1.)

Functions assigned to the structural regions of nuclear receptors are shown in Fig. 11.1. The ligand-binding (LB) domain is a conformational switch that, on ligand binding, shifts the receptor to a transcriptionally active state. Whether DNA-binding (DB) and ligand-binding (LB) domains evolved independently from each other, and whether the first members of this family of proteins were constitutive, non-regulated transcription factors, without a LBD, is a matter of speculation.

Nuclear receptors in *Drosophila*: the ecdysone receptor

Nuclear receptors for the steroid hormone ecdysone in *Drosophila* deserve special consideration, for several reasons.[2] First, the insect hormone ecdysone (Fig. 11.2) was the first steroid hormone shown to act at the level of the gene, because it induced puffs in the giant chromosomes of the fruitfly. Secondly, ecdysone activates the developmental programme of *Drosophila*,[3] and, finally, the sequences of the DNA-binding domain (DBD) of the *Drosophila* nuclear ecdysone receptor and of receptor homologues (for example, the COUP-TFs; chicken ovalbumin upstream promoter transcription factors) are highly conserved and nearly identical in vertebrates and humans.

There are three ecdysone receptor isoforms: ECR-A, ECR-B1, and ECR-B2. All have identical DBDs and LBDs, but differ in the amino-terminal sequences. Functional

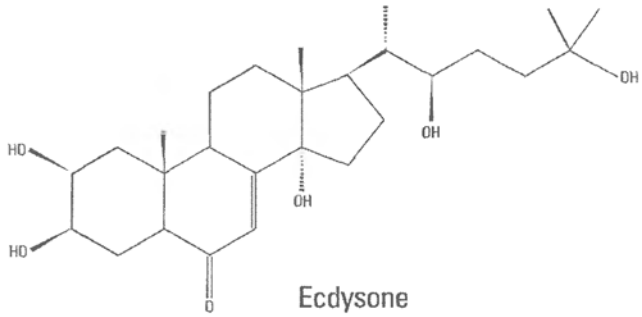


Fig. 11.2 Structure of the steroid hormone, ecdysone.

differences between these isoforms have not been established clearly. The ecdysone receptor is a heterodimer. One half of the protein is encoded by the *ecr* gene, which is ecdysone-responsive; thus the ligand regulates the synthesis of its own receptor. The other half of the ecdysone heterodimer is coded by a different gene which expresses a completely different protein, the ultraspinacle protein (USP).

ECRs become functional only upon heterodimerization with USP. Ecdysone, the ligand, promotes formation of the ECR–USP heterodimer, which has a high affinity for ecdysone-responsive DNA elements (EcREs). Since USP is a homologue of the vertebrate and mammalian 9-*cis*-retinoic acid receptor (RXR), *Drosophila* USPs can form heterodimers with many mammalian non-steroidal DNA-binding proteins, including retinoic acid, thyroid hormone, and vitamin D receptors.

Transcription factors that bind to the same element in an ecdysone-responsive gene to which ECR/USP heterodimers bind, can interfere with gene activation by the ECR/USP receptor complex. Negative control is also possible at the level of the ECR complex itself. Monomeric orphan receptors, which are unable to bind to DNA, but still can form heterodimers with ECR, interfere with the formation of transcriptionally active ECR/USP complexes. Negative control also works the other way around, by scavenging and removing USP. This kind of negative control is also operative in mammalian cells, for example, heterodimerization in various combinations occurs between retinoic acid receptors and other nuclear receptors, such as the nerve growth-factor-induced nuclear receptor, NFI-B, or with COUP-TF. These nuclear receptors negatively control RXR signalling by forming heterodimers with RXR. This reminds us how important a strictly ordered, temporal pattern of expression of such co-receptors, different from that of the receptor is, to prevent mutual interference of competing transcriptional activators (see Chapter 2).

Nuclear receptors in metamorphosis and embryogenesis

To date, 16 genes that code for nuclear receptors have been identified in *Drosophila*, and half of them are either induced or repressed by ecdysone. At the end of the larval stage, ecdysone is released in pulses, building up transiently high levels of the hormone and inducing a small set of early genes. At a later stage, more than 100 late, secondary response genes are induced. The products of the late genes are thought to guide each ecdysone-sensitive target tissue through metamorphosis. In *Drosophila* the different kinds of developmental genes controlling the organization of the embryo are expressed in a hierarchical order. For each determinant a different gradient is set up at different times. Moreover, the determinants expressed by the developmental genes are localized in different parts of the embryo, where they control the formation of a regular pattern of segmental and subsegmental units. For more information, the *Molecular Biology of the Cell*[4] should be consulted. I consider only the *homeotic selector* (*hom*) genes, which bear the information for the global organization of the body. Mutations in homeotic genes can rearrange a whole body segment. The time-ordered expression of the *hom* genes and of the *segment polarity* genes determines the segmental subdivision and patterning of the body. The *hom* genes have a highly conserved, characteristic sequence, the homeobox. Practically all animals have *hom* genes. Mammals, including humans, have *Hox* genes, which specify the position of newly formed parts in the body. The mammalian *Hox*-gene family contains 39 developmental control genes, divided into 13 groups and located in four chromosomal gene clusters. Each group is made up of three similar, 'paralogous' genes. Thus, mammals have three closely

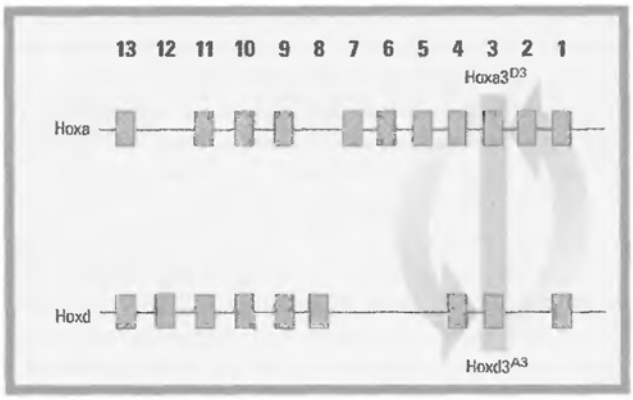


Fig. 11.3 The complexes of the group 3 paralogous genes of *Hox-a* and *Hox-d* are shown. Genes are indicated by the boxes. The group 3 genes are highlighted and the arrows indicate the swap between the *Hox-a3* and *Hox-d3* coding sequences carried out by Greer *et al.*, leading to the two novel alleles, *Hox a3* ^{D3} and *Hox d3* ^{A3}. (The information used for this figure was taken from ref. 5 with permission of Dr. D. Duboule and Nature.)

related (or paralogous) copies for each single developmental gene found in flies or in worms. Thus, the vertebrate *Hox* genes must have gone through several rounds of duplication in the course of evolution. Figure 11.3 shows the functional relationship between the products of the paralogous *Hox* genes: *Hox-a3*^{D3}, and *Hox-d3*^{A3}. This experiment of Greer *et al.*[5] gives an answer to the puzzle of why expression of the *Hox-a3* and *Hox-d3* genes has unique and different consequences, although all paralogous group 3 genes encode identical proteins. Greer *et al.* exchanged the protein-coding region of the *Hox-a3* gene with that of the *Hox-d3* gene in an embryonic stem cell, leaving the rest of the transcriptional unit with the domains regulating the expression of these genes intact. The outcome was a new *Hox-a3*^{D3} locus in the genome of the cell with its original regulatory DNA sequences. The gene put in this new location could now express a Hox-D3 protein which, when it was expressed effectively, could substitute for the missing Hox-A3 protein. This saved the life of the mouse embryo, which dies when the Hox-A3 protein is not expressed. The *Hox-d3* locus in the genome was engineered in the same way to *Hox-d3*^{A3} allowing it to produce a Hox-A3 protein in its new locus. However, expression of two allelic copies were required, because transgenic mice were not rescued if only one allele was expressed.

These experiments suggested (1) that the proteins encoded by paralogous *Hox* genes are functionally not identical; and (2) that it is the amount of a Hox protein which is expressed that matters (see also refs 6,7). When the quantity of a Hox transcriptional regulator is important, this could have developmental consequences, since Hox-A3 and Hox-D3 proteins are actually produced in different quantities in different developmental situations. For example, neural crest and mesoderm cells may each require different threshold concentrations of a particular *Hox* group 3 protein to initiate their developmental programmes. If further experiments confirm this, one would have to revise ideas about the function of *Hox* genes in development: rather than ascribing a unique individual regulatory role to each Hox protein in a certain environment, a more global role in developmental control, being accomplished by gene dosage and supply of different concentrations of the *Hox* proteins, would have to be considered.

Among other evolutionary conserved developmental genes in *Drosophila* are the *svp* (seven up) genes, in the central nervous system, the *pax* genes, paired homeobox genes, expressed in the photoreceptor precursor cells in the eye, and the *fringe* genes, which encode the WNT proteins (the 'wing' nuclear transcription factors, which control wing formation).[8] In *Drosophila*, the message of the product of the *fringe* genes is transmitted by a nuclear receptor, named '*frizzled*' (Dfz2). Homologues of Dfz2 exist in birds, fish, and mammals (see Part 4). Other conserved, developmental determinants are the Tll-protein, coded by the *tll* gene

in the fly, and its homologue in the mouse, the TLX protein, coded by the *tlx* gene, and the human homologue, the HNF4 protein (hepatic nuclear factor), expressed only in the intestine and the liver. (Developmental genes in the fly are usually named according to the phenotype where these genes were mutated or deleted).

To sum up: determinants that control the developmental programme of *Drosophila* are highly conserved nuclear receptors. They have similar DBDs with a DNA-recognition helix. The DNA recognition site, the P-box, has a typical $Zn^{2+}(Cys)_4$ motif (these motifs do not bind directly to DNA, but help to direct the recognition helix to the major groove of the DNA).

Other ecdysteroids and the juvenile hormone

Although ecdysone is the only steroid hormone in *Drosophila* of which the function is known, the haemolymph of the fly contains many other ecdysteroids, the function of which remains to be clarified. The haemolymph also contains the sesquiterpenoid, juvenile hormone (JH). In many insects, JH regulates ecdysone release. JH has structural features in common with the retinoids and may signal through a nuclear RXR receptor.

Properties of human steroid hormone receptors (SHRs)

Mammalian nuclear receptors for steroid hormones (SHRs) are of great importance in physiology and medicine, because they control not only developmental pathways but also regulate central physiological and metabolic functions in the adult organism. Steroid hormones and vitamin D are derivatives of cholesterol. Structures of vitamin D_3 and of

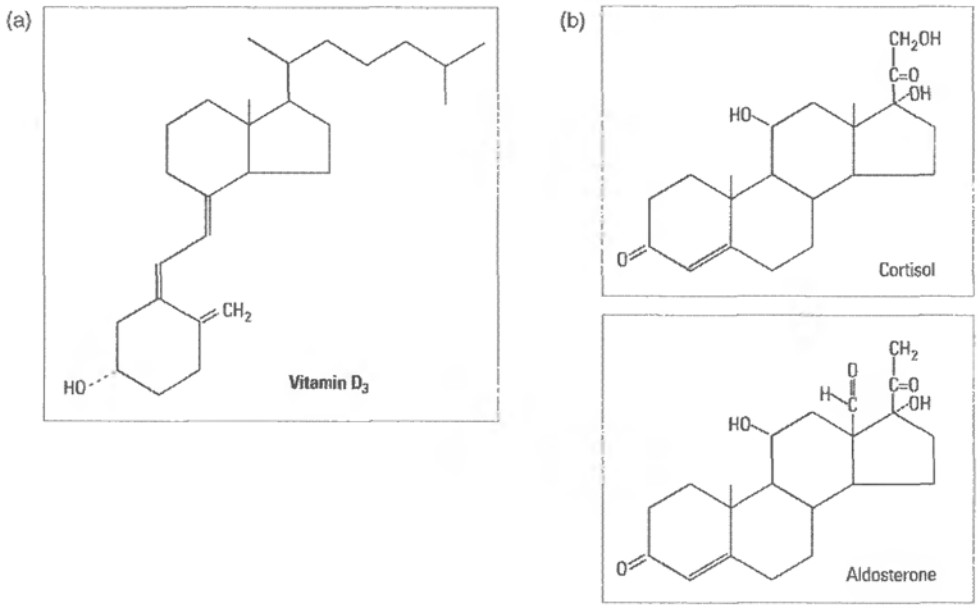


Fig. 11.4 (a) Structure of vitamin D_3. Vitamin D_3 is derived from cholesterol by UV irradiation. (b) Structures of cortisol and aldosterone. Cortisol is a major glucocorticoid, and aldosterone is a major mineralocorticoid. Both are derived by a series of hydroxylation reactions from progesterone, from which androgens and oestrogens are also synthesized.

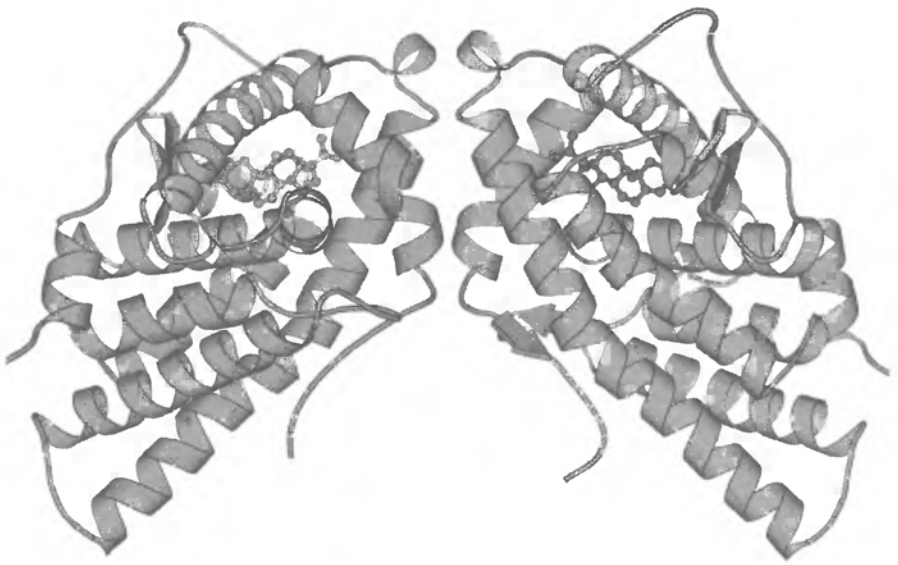


Fig. 11.5 The ligand-binding domain of the dimer of the progesterone receptor. Each receptor LBD-chain binds one molecule of the steroid hormone. The LBD of the progesterone receptor 1 is on the left and the LBD of the progesterone receptor 2 is on the right. The steroid hormone is shown as a ball-and-stick structure. The LBDs were expressed as GST (glutathione *S*-transferase) fusion proteins. The liganded progesterone–LBD contains 10 α-helices, arranged as a helical sandwich. Although the ligand-binding domain of the progesterone receptor shares only 15% sequence identity with other nuclear receptors, the principal features of the fold of the LBD are conserved in the oestrogen receptor and in other members of the nuclear steroid hormone receptor superfamily. (Reproduced with permission of the authors and Nature from data in Fig. 3 of ref. 11, and data available in protein databanks.)

cortisol and aldosterone are shown in Fig. 11.4 (Lubert Stryer's textbook of biochemistry should be consulted with respect to synthetic pathways;[9] information on the central role of SHRs in humans, in health and disease, can be found in textbooks of endocrinology[10]). Figure 11.5 shows the structure of the ligand-binding domain of the progesterone receptor, a representative nuclear SHR receptor.

Structure of SHR homodimers

SHRs, activated upon hormone binding, bind as homodimers to palindromic DNA sequences in hormone response elements (HREs). At the gene, SHRs interact both with the basic transcription machinery and with specific, SH-responsive transcriptional regulators. To summarize the characteristic, structural properties of SHRs: the DNA-binding domain of the glucocorticoid receptor is a globular structure, which can be subdivided into two parts, the P-box and the D-box. Each module has a zinc-finger and an amphipathic helix. The P-box starts with a short segment of an antiparallel β-sheet which helps to orient the amino-acid residues that contact the phosphate backbone of the DNA, but the important contacts with the nucleotides in the major groove of the DNA are made by the P-box α-helix. The other module, the D-box, is more important for the contacts with the phosphate backbone of the DNA and for dimerization of the receptor.

The role of chaperones

A unique property of nuclear steroid hormone receptors is that they associate with a large multicomponent protein complex of chaperones, the Hsp90 (heat-shock protein) system.[12] (Chaperones which associate with nascent polypeptide chains and assist them to fold are named chaperonins, whereas chaperones are proteins which associate with polypeptide chains, post-translationally) (Fig. 11.6).

Figure 11.6a is a simplified scheme, because the receptor/chaperone complex is in equilibrium with an other complex where the SHR is already in a metastable, nearly mature state and is capable of binding the hormone. This complex (not shown) is different and lacks Hsp70 and the co-chaperones, but it has accquired new proteins, including peptidyl-prolyl *cis/trans* isomerases. Proteins with prolyl isomerase activity bind to immunosuppressive agents, such as FK506. Another name for these proteins is FKBPs, FK-binding proteins; see Chapter 7). The Hsp90 chaperone protects several nuclear hormone receptors and eukaryotic protein kinases, many of which are proto-oncogenic and play a prominent role in cancer.

The geldanamycin antibiotic which binds to Hsp 90 has antiproliferative and antitumour effects. It inhibits the Hsp90-mediated conformational maturation reaction, and consequently the proteins bound to the Hsp 90 are now degraded. The structure of the geldanamycin-binding domain of Hsp90 has a pronounced pocket, 15 Å deep, that is highly conserved across species. Geldanamycin binds inside this pocket, adopting a compact structure. The important point is that a polypeptide chain can bind to the same pocket. This is supported by similarities of the structures of geldanamycin and peptides. Thus, the geldanamycin-binding pocket, shown in Fig. 11.6c, is the polypeptide-substrate-binding site of Hsp 90. The pocket is conserved in heat-shock proteins and chaperones, supporting the notion that the pocket binds a portion of the polypeptide substrate and participates in the conformational maturation/refolding reaction.

The group of SHRs that require the assistance of the Hsp90 chaperone system for maintaining a conformation in the unliganded state, capable of binding the hormone any time it arrives, also includes receptors for androgens and oestrogens and for mineralocorticoids.[16] Association with chaperones protects and prevents degradation of the unliganded receptors and keeps them responsive and poised for action. In an ATP-dependent reaction, an initial complex is formed between the SHR and the chaperones. Whether the receptor is phosphorylated at this stage, and where, is not clear. Upon binding the hormone, the SHR and the chaperones separate and the SHR is now ready to function as a transcription factor, possibly after further phosphorylation.[17] (Steroid hormone receptors are not the only targets of chaperones. Other targets are regulatory protein kinases. For example, the Raf kinase forms a complex with Hsp90; disruption of this complex destabilizes Raf-1 and prevents Raf–Ras association;[18] see Chapter 3).

Receptors for ovarian hormones and signalling by oestrogens

The role of ovarian hormones in breast cancer deserves special attention (see also Part 4). The oestrogen receptor (ERα) has, like all SHRs, a modular structure with two transcriptional activation functions, AF-1 and AF-2. AF-1 is in the NH_2-terminal and AF-2 in the carboxy-terminal region. Besides ERα, there exists another ER, ERβ. The existence of ERβ explains why tissues, such as the ovaries and the urogenital tract, which have no ERα, are still responsive to oestrogen. (see ref. 19). More than one ER also explains why mice without one of the oestrogen receptors have no oestrogen-related

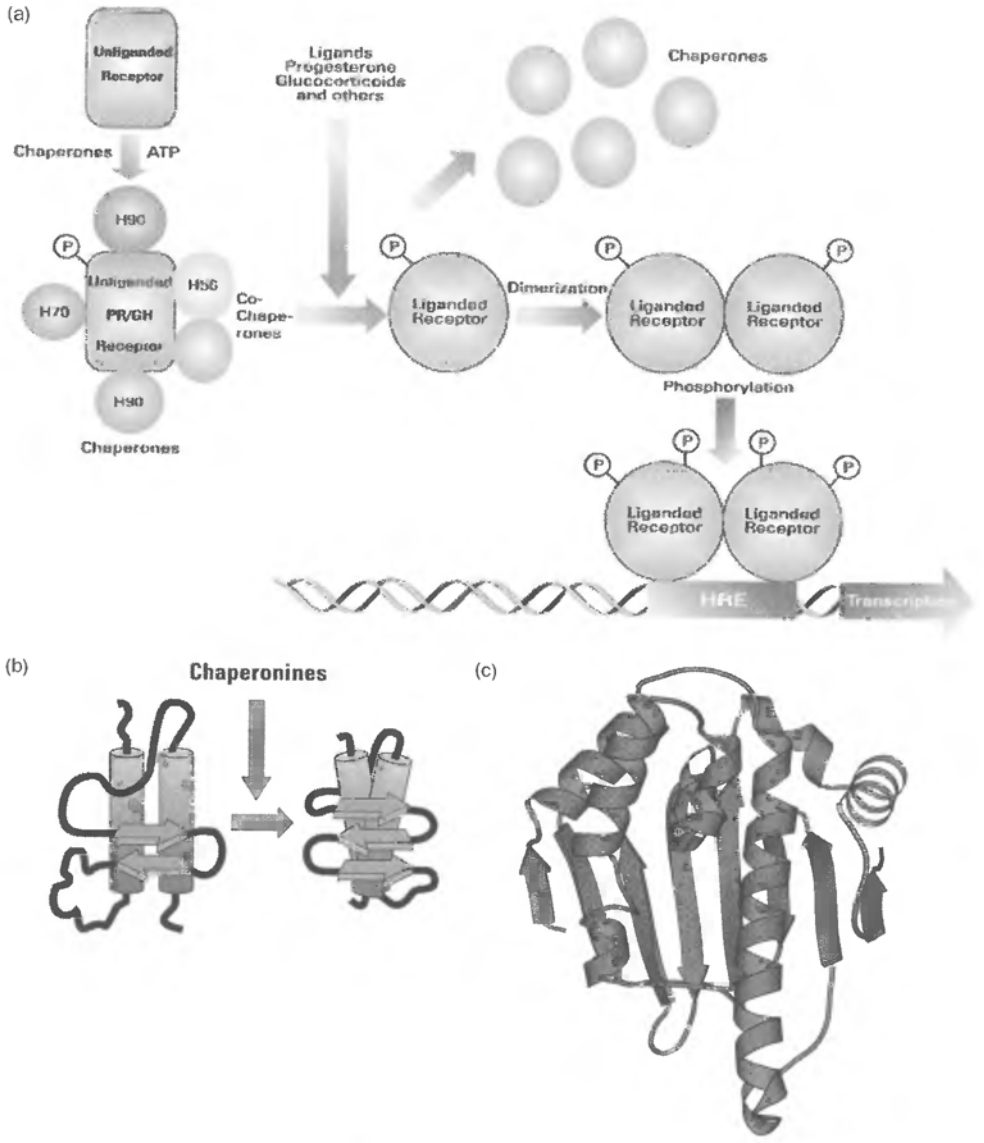


Fig. 11.6 (a) The role of chaperones. Unliganded receptors for glucocorticoids and for progesterone, are associated with heat-shock proteins. The unliganded progesterone/glucocorticoid receptor, PR/GR, is phosphorylated and associates with Hsp90, Hsp70, and at least with two co-chaperones, one of which is Hsp56. In this configuration, the receptors do not bind to DNA and have no transcriptional activities. On binding ligands, the heat-shock proteins dissociate and the receptors dimerize, and bind to hormone-responsive gene elements, HRE, and activate target genes. Gene recognition and transcriptional activation are regulated by phosphorylation and dimerization. (Drawn, with permission of Dr. B. O'Malley, using the information in Fig. 10 in ref. 13.) (b) How chaperonins assist in folding nascent proteins. The protein on the left side already has a secondary structure, but no ordered tertiary structure, whereas the protein on the right side has attained, with the help of chaperonins, a final well-defined secondary and tertiary structure. (Drawn, with permission of Dr. Ulrich Hartl, using the information in Fig. 1 of ref. 12.) (c) The crystal structure of the binding pocket of the Hsp90 heat-shock protein. The binding pocket of Hsp90 is the substrate-binding site. (In the diagram the ligand, geldanamycin, has been omitted). (The ribbon model is reproduced with permission of the authors, and Cell using the information in ref. 14, which is also available in protein databanks.)

defects. Both ERα and ERβ are activated by the same hormone, 17β-oestradiol (Fig. 11.7).

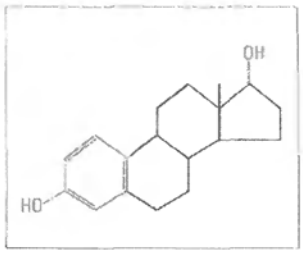


Fig. 11.7 Structure of oestradiol.

ERα and ERβ are both targeted to AP1 gene sites (to which the growth-promoting transcription factors, Jun/Fos, bind) and to the cyclin D1 gene (Chapter 12).[20] When the cyclin D1 gene is disrupted and the cell cycle is blocked, the ERα-dependent proliferation of the mammary gland during pregnancy is prevented. ERα enhances transcription at the AP1 site, but ERβ inhibits transcription of immediate early genes. Consequently, ERα upregulates proliferation of uterine and mammary tissues and ERβ dampens the response. Therefore, the oestrogen receptor has become the target of anti-oestrogens such as tamoxifen, which are used in the treatment of breast cancer (Fig. 11.8). However, the differences between ERα and ERβ explain why the action of anti-oestrogens depends on the ER that is expressed. Tamoxifen and related anti-oestrogens are ER ligands, they are thought to block the transcriptional activating domain, AF-2, of the receptor. However, when ERβ binds tamoxifen, the drug actually activates transcription at the AP1 site, imitating a typical oestrogen-like effect. This may be related to undesired oestrogen-like effects of anti-oestrogens.

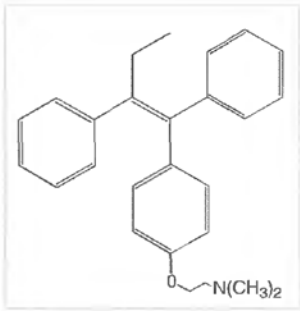


Fig. 11.8 Structure of tamoxifen.

Breast cancer cells, like most other cancer cells, are exposed to excess growth factors which activate the Ras/MAPK pathway, promoting cellular proliferation. Consequently, ERs are phosphorylated, increasing their transcriptional activity. The uncontrolled transcriptional activity of phosphorylated ERs is implicated in breast cancer, the most common tumour in females. Downregulation of immediate early genes, responsible for cellular proliferation, is one of the objectives of breast cancer therapy (see Part 4).

A possible interaction of many oestrogenic chemicals in the environment with the human oestrogen receptor has received publicity. A warning was actually issued that the synergistic interaction of mixtures of several low-dose environmental chemicals with the hER may have negative consequences on fertility. However, these fears seem to be unfounded.[21]

How is a hormonal signal controlled?

How do steroid hormone receptors accomplish such different biological actions as those exemplified by glucocorticoids, mineralocorticoids, by progesterone, oestrogens, and androgens?

Several points must be considered:

1. The formation, release, activation and inactivation, and the lifetime of each hormone are strictly regulated in accordance with the needs of the body.
2. Interaction of the receptor with the ligand is hormone-specific.[22]
3. Finally, gene recognition and gene expression by the hormone bound to its receptor are also specific. Specificity is improved and fine tuned with the help of auxiliary transcription factors, co-activators and repressors. Moreover, some SHRs interact directly with other transcription factors.

Mechanisms of control

Efficient induction of gene expression by steroid hormones requires assistance by co-activators. Since the role of co-activators has already been introduced, we can be brief. There are two kinds of interactions: those which initiate transcription and those which regulate transcriptional activity. Transcription is initiated by the two *trans*-activating regions of SHRs, the AF-1 and AF-2 domains. These domains establish the contact with the HRE (the hormone-responsive gene element) and with the basic transcriptional machinery and Pol II. However, full expression of the transcriptional activity of all members of the steroid hormone receptor superfamily requires a separate co-activator and the formation of the SHR–co-activator-1 (SRC-1) complex.[23] Subsequently, the SHR–SRC-1 complex binds CBP and forms a ternary complex. CBP has two binding sites; the amino-terminal site binds to all nuclear receptors, whereas the carboxy-terminal site is more discriminating and binds to different nuclear receptors with different affinity. For instance, this site interacts with high affinity with SHRs, but binds only weakly to the RXR- and not at all to RAR- and thyroid hormone-receptors (see below). The ternary complex formed with SHR, a co-activator, and CBP binds to Jun and prevents the formation of a transcriptionally active Fos/Jun-complex, thus blocking transcription at the AP1 site.[24] The role of CBP becomes apparent when CBP concentrations become limiting within the cell. This is the case in the Rubinstein–Taybi syndrome, an autosomal dominant, inheritable disease, where only one allele of the *CBP* gene is expressed. This disease is associated with mental retardation and skeletal and facial abnormalities.

SHR activity can also be inhibited directly. An example is the interaction of oestrogen receptors (ERs) with an inhibitory protein, the ER-inhibitor protein, RIP140. The receptor inhibitor interacts only with transcriptionally active ERs.

Finally SHRs are subject to phosphorylation, although the role of phosphorylation is not yet clear in all cases. SHRs are substrates of serine/threonine kinases, and, at least in the case of oestrogen receptors, also of tyrosine kinases. Moreover, transcriptionally active SHRs are phosphorylated by a DNA-dependent kinase. SHRs can also respond indirectly to hormones that signal through the second messenger, cAMP, because the co-activator CBP is a target of the cAMP-dependent serine/threonine kinase, PKA. Phosphorylation of CBP enhances SHR activity.

Nuclear receptors for non-steroids: thyroid hormone and retinoic acid receptors, their role in morphogenesis

In Table 11.2 the properties of two families of nuclear receptors, A and B, are compared. To family A belong the steroid hormone receptors and to family B, the non-steroidal receptors,[25] the thyroid hormone receptors, the retinoic acid receptors, the vitamin D_3 receptor, the peroxysome proliferator activating receptor, and several orphan receptors.

Thyroid hormone receptors

Representatives of non-steroidal ligands, binding to nuclear receptors are T_4 (tetra-iodothyronine; thyroxine) and T_3 (tri-iodothyronine) (Fig. 11.9).

Two genes, *tra* and *trb*, encode thyroid hormone receptors (THRs). Mutations of the *trb* gene have been found in patients with a dominant inherited general refractoriness to

Table 11.2 The two families of nuclear receptors and their properties are compared. Group A are the steroid hormone receptors and group B the non-steroidal nuclear receptors

Groups	Receptors	Properties
A	Glucocorticoid receptors, GR	Long A/B domains: associate with chaperones.
	Androgen receptors, AR Progesterone receptor, PR Mineraloconticoid receptor, MR Oestrogen receptor, ER	Unliganded receptors do not bind to DNA
B	Thyroid hormone receptor, THR All-*trans* retinoic acid receptor, RAR Vitamin D_3 receptor, VDR 9-*cis* retinoic acid receptor, RXR, Peroxysome proliferator activating receptor, PPAR Orphan receptors	Short A/B domains; do not associate with chaperones. Unliganded receptors bind to DNA

thyroid hormone, the 'GRTH syndrome'. Point mutations in the region of the *trb* gene, coding for the LBD impair or abolish ligand binding, but leave DNA binding intact. Therefore, such a mutated THR can still form heterodimers with receptors for 9-*cis* retinoic acid (RXRs) and interfere with RXR signalling. This explains peculiarities of the phenotype of the GRTH syndrome, which are difficult to rationalize by a loss of response to thyroid hormone alone.

Quite generally, all proteins that associate with nuclear receptors can interfere, negatively or positively, with their function, provided they are present in the cell in high-enough concentrations. An example is the human and murine SHP proteins (e.g. small heterodimerization partners). SHPs are conserved proteins which form heterodimers with several members of the nuclear receptor superfamily, including thyroid hormone receptors, and negatively control their activity. Other proteins belonging to this group are orphan receptors, of which the ligand is not known, as the name implies.[26] Orphan receptors, forming heterodimers with nuclear receptors, can also have positive effects and can function as co-activators. An example of an orphan receptor, functioning as a co-activator is Dax-1. It hybridizes with the FTZ-F1 nuclear receptors in *Drosophila* and their homologues in vetebrates and mammals. Loss-of-function mutations of either the

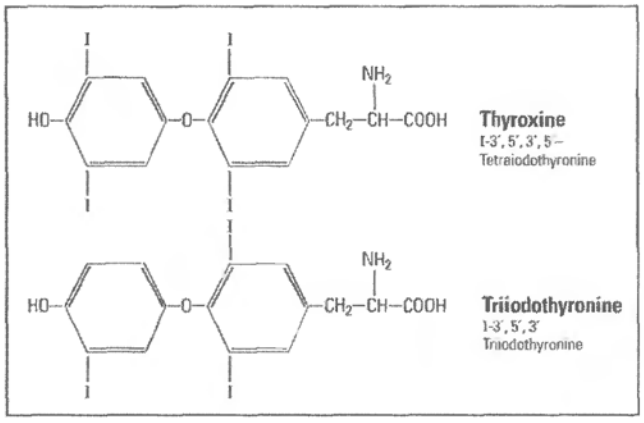


Fig. 11.9 Structures of T_4- and T_3-iodothyronines.

receptor or of the orphan co-receptor cause severe developmental defects. In *ftz-f1* null mice, the hypothalamic nucleus, the gonads, and the adrenals are not formed, and loss-of-function mutations of *dax-1* cause adrenal hypoplasia and hypogonadism in humans. A tissue-specific nuclear orphan receptor in mammals and humans is the hepatocyte nuclear factor (HNF4). The *hnf4* gene is only expressed in liver, kidney, and the intestine. The HNF4 nuclear factor activates several liver-specific genes. Mice lacking a functional HNF4 nuclear receptor, die as early embryos.

Nuclear receptors which are not *sensu stricto* orphan receptors are the peroxysome proliferator-activated receptors (PPARs, α, β, γ), and NGFI-B the nerve growth factor-induced receptor B. PPARs induce proliferation of peroxysomes in the liver, a process that plays a role in carcinogenesis of the liver. Although ligands for these receptors are known, they are quite heterogeneous and not well defined. For instance, PPARs are activated by a variety of fatty acids, and gene response elements addressed by PPARs have been located in the promoter regions of genes encoding enzymes participating in the oxidation of fatty acids. However, surprisingly, enzymes catalysing fatty acid oxidation were not impaired in dominant negative PPAR-mutant mice.

Retinoic acid and retinoid receptors: RARs and RXRs

Two classes of retinoic acid receptors are distinguished—RARs and RXRs. The former are activated by all-*trans* (tRA) and by 9-*cis* retinoic acid (9cRA), whereas the RXR family is activated exclusively by 9cRA. The RAR and RXR family consists each of three main α, β, γ types and additional variants: α1, α2, β1–β4, and γ1, γ2, and so on (Fig. 11.10).

Function of RARs

A typical RA-receptor is the receptor for vitamin A (all-*trans* retinol). Vitamin A is essential for pre-and postnatal development. Congenital vitamin A deficiencies are accompanied by a broad spectrum of malformations. In adult life, vitamin A is indispensable for growth, maintenance, and survival of many tissues. It is necessary for vision and reproduction. A lack of vitamin A in the adult causes night blindness and photoreceptor degeneration. But the devel-

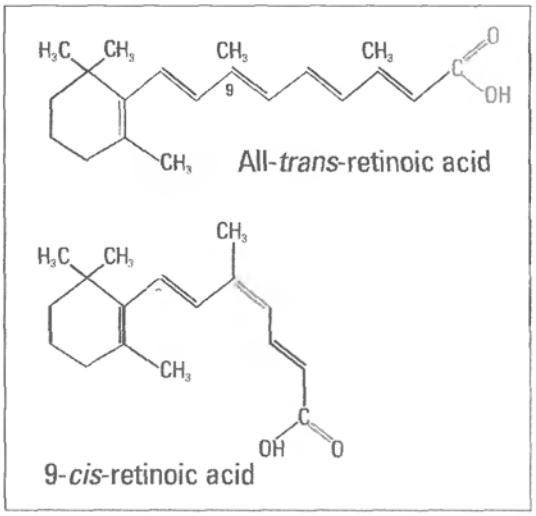


Fig. 11.10 The two acid derivatives of vitamin A: all-*trans* retinoic acid and 9-*cis* retinoic acid, which participate in activation and signalling of the retinoic acid receptors (all-*trans* retinoic acid is vitamin A acid).

oping eye is most affected in vitamin A deficiency. Vitamin A and its cognate RAR are essential for eye morphogenesis. The master gene in the eye developmental programme is the *Pax 6* gene, which is conserved in *Drosophila,* mice, and humans—an impressive manifestation of evolutionary conservation. Since the eye has different structure and function in invertebrate and vertebrate species, convergent evolution and different genetic programmes for eye determination in invertebrates and vertebrates have been invoked by evolutionary biologists. Eye development was even thought to be an exception to classical Darwinian evolution by selection. Now the existence and the conservation of the *Pax 6* gene in invertebrates and vertebrates, including humans, proved mainly by work of Walter Gehring and his laboratory at the Biocenter in Basle, requires revision of this view (for more information, see refs 27–30). The *Pax-6-eyeless* gene is the key regulator of eye development. Mammals, fruitflies, and squid (*Loligo opalescens*), all carry the *Pax 6* gene, strengthening the idea that the genetic programme for eye development evolved only once, some 500 to 600 million years ago, before vertebrates and invertebrates diverged. The different morphological and functional properties of the eye in invertebrates and vertebrates, show that eye development, like any other developmental programme, has been modified in each species by particular, tissue-specific features and components, acquired in the course of evolution.

The role of vitamin A in the morphogenesis of the eye is only one example of the role of retinoic acid derivatives in the implementation of developmental programmes. RARs are also involved in the determination of the body axis, the spine. Mutants of RARγ (and, to a lesser degree, of RARα), are associated with malformations of vertebrae. Moreover, administration of excess RA (vitamin A acid), can induce extra limb formation in the embryo, and loss-of-function mutations of RA α/γ receptors produce disordered limb morphogenesis. RA signals through RARα and RARγ are obligatory for differentiation and maintenance of epithelia, of the male accessory glands, and spermatogenesis. RA may also be needed for skin formation and maintenance in the adult.

The phenotype of homozygous double-null mutant RAR mice demonstrates the crucial role of RARs in the development of the mesenchyme. These mice have malformations of the mesenchymal neural crest and a dramatically reduced viability and life expectancy. Since the appearance of neural crest cells (NCCs) coincided with the emergence of vertebrates, it was reasoned that RARs may have co-evolved with NCCs and that RARs and vertebrates have appeared together. Along this line, a major role is ascribed to RARs in the modification of old ancestral developmental programmes. Such modifications are crucial for the development of new structures in the course of evolution and are indispensable for the appearance of vertebrates.

Retinoids in cancer

In normal cells, the expression of RARs is enhanced by retinoids. Many cancer cells lose the ability to respond to retinoids because they lack COUP-TF (a mammalian ecdysone-like transcription factor). Expression of COUP-TF in cancer cells can restore the retinoid response. Apparently, COUP-TF is required for the interaction of the RARs with the transcriptional co-activator CBP, enhancing DNA binding.[31]

Function of RXRs

While the biological role of RARs is undisputed, it is less certain what, if any, biological process is regulated by homodimeric RXRs, since RXRα, β, γ knock-out mice were viable.[32] Moreover, it has proved difficult to exclude the contribution of heterodimeric RAR/RXRs to developmental functions ascribed to RXRs alone. This dilemma is apparent in the role

ascribed to RXR in malformations of the myocardium and the eye.[33] Originally, an RXRα mutant was thought to be solely responsible for a 'spongy myocardium' and cardiac failure. But since a milder form of this developmental defect is also observed in some RARα or γ mutants, a heterodimeric partnership of the RXRα mutant with RARα or γ seems more likely. The same uncertainty applies with respect to an exclusive role of RXRα mutants in abnormalities in eye development. Although RXRα-null mutant mice display ocular malformations, here again an exclusive role of RXRα is questionable, since the effects were strongly synergetic when both RXRα and RARγ genes were mutated. This leaves in doubt whether RXR homodimers have a developmental role of their own.[34] Chen *et al.*[35] have shown with RXR- and RAR-specific agonists and antagonists that RARα in the heterodimer is decisive, because it directs the RXR–RARα heterodimer to DNA, and only when the heterodimer is bound to DNA can RXR be activated by an RXR-specific agonist. Thus without RARα RXR cannot function.

Synthetic retinoids, either specific for RAR or RXR, such as those introduced by Chen *et al.*[35] could considerably enlarge therapeutic possibilities.

Structural motifs of retinoid receptors

Structural information about retinoid receptors is available. This information has contributed greatly to the elucidation of the structural basis of ligand binding, dimerization, and recognition of DNA response elements, for retinoid receptors in particular, and nuclear receptors in general.

RARs and the RXRs have a modular structure, like all nuclear receptors, although the structures of RARs and RXRs do not have much in common with other members of the nuclear receptor superfamily. Moreover, RARs and RXRs do not need the help of chaperones. Each RAR of a given species has a unique A region, whereas the amino-terminal B regions of all three RAR isoforms are quite the same. A region, located between the DBD and the LBD, is well-conserved in all RARs and RXRs, even from different species. It is a hinge region, which may have a functional role that is not yet known. The carboxy-terminal regions of the three RAR subtypes are different and not related to each other. They also have no counterpart in the RXRs. Their role, if any, is not known.

Most gene response elements (REs) for retinoic acid receptors are direct repeats of a core motif, $PuG(GT)CA(X)_n$ or closely related motifs, where Pu is purine and X is any base. REs with palindromic arrangements, and with more complicated arrangements of two or more hexameric motifs with variable orientation and/or variable spacings, have also been found. Binding of RAR homodimers and RXR/RAR and RXR/THR heterodimers to the response elements in the target gene is cooperative and synergistic, and involves both dimerization interfaces in the DBDs of RXR/RAR (but only one dimerization surface in the case of RXR/THR).[36] What is remarkable is the fidelity with which dimeric receptor DBDs recognize the REs. The DNA recognition helix of the DBD covers only 4–5 bp, and fits exactly to the response element.

Thanks mainly to the efforts of the Strasbourg group of D. Moras and Pierre Chambon, and Paul Sigler's laboratory at Yale University, three-dimensional structural information is now available in the case of heterodimeric RXR/TR and RXR/RAR DBDs bound to direct DNA repeats (DR4 and DR5), respectively (Plate 24 and Fig. 11.11). This is one of the still rather rare cases where one can profitably explain, on a molecular basis, the interactions of a receptor with the ligand and with target genes.

The structure of Rastinejad *et al.*[37] (Plate 24) shows how the DNA provides a scaffold for the asymmetrical dimerization of the DBDs of RXR and THR. This interface involves the carboxy-terminal extension of the DNA-binding domain of THR. This arrangement gives a clue to how nuclear receptor heterodimers recognize the spacings between DNA repeats and can distinguish between closely related response elements.

Binding of either ligand, 9cRA or thyroid hormone, had no effect on the interaction of RXR/THR heterodimers *in vitro* with their REs. However *in vivo*, in cells, ligand binding was necessary for the interaction of the receptor with the promoter regions of the gene. This example demonstrates how important the natural chromatin environment in the nucleus is for ligand-responsive interactions of receptors and transcription factors with target genes.

If one compares the unliganded LBD with the liganded LBD of RARγ, one sees that the overall fold of the liganded LBD is more compact than the empty LBD. In the case of the thyroid hormone receptor, the ligand is completely buried within the hydrophobic core of the LBD.[40] This conformational change, triggered by the ligand, has been compared to a trap sprung by a mouse. In both cases, mouse or ligand are trapped. Binding of the ligand affects the hydrophobic core of the receptor and aligns secondary-structure elements. These structural changes upon ligand binding may be representative for nuclear receptors in general.[41]

The most significant change of the receptor on ligand binding is the formation of new recognition surfaces. These newly formed sites enable the receptor not only to recognize DNA elements, but also to interact with co-activators and co-repressors.

A notable feature of retinoic acid nuclear receptors is that their DBDs and LBDs can communicate with each other: retinoic acid receptors bound to DNA response elements modulate ligand binding through a kind of feedback mechanism. This allows fine tuning of ligand-dependent transcriptional activity.

To sum up what we have learned from the structures:

1. Binding of ligand to the dimeric receptor is cooperative.
2. Ligand binding moulds new interfaces for interactions with the DNA response element. These interfaces fit the receptor exactly to the REs. Thus, ligand binding guarantees the fidelity of gene recognition, and communication between LBD and DBD controls transcriptional activity.

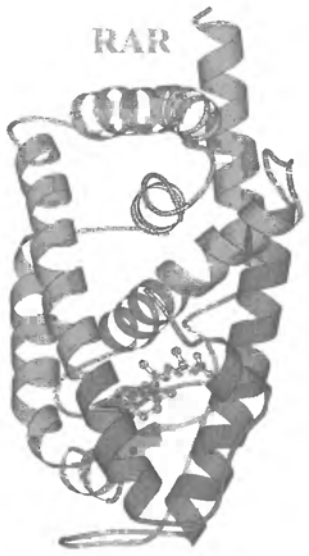


Fig. 11.11 Structure of a fraction (residues 178–423) of the ligand-binding domain of the human retinoic acid receptor, RARγ bound to all-*trans* retinoic acid. The ligand, all-*trans* retinoic acid, is shown as a stick-and-ball structure. The structure is typical of the ligand-binding domains of nuclear receptors in general. The 2.0 Å crystal structure reveals the ligand-binding interactions. The overall fold of the binding pocket is similar to that of the human RXRα LBD. The ligand-binding pockets of the α,β,γ- isoforms of RARs differ only in three residues. These three residues seem to be responsible for the differences between the three RAR subtypes. (Reproduced with permission of Professor D. Moras and Nature from the data in Fig. 1 of ref. 38, and the corresponding data available in protein databanks. See also ref. 39.)

3. The interaction surfaces, formed on ligand binding, also provide a structural scaffold for interaction with co-activators and co-repressors.

Regulation

Retinoic acid receptors can, like other nuclear receptors (for example the oestrogen receptors), entertain cross-talk with growth-factor- or hormone-regulated signalling pathways. They can interfere with growth-factor signalling and shut off immediate early genes under the control of the Jun/Fos transcription factors. Moreover, as already been discussed, retinoic acid receptors are like other nuclear receptors regulated by co-activators and co-repressors.[42,43]

Therefore only a special case is mentioned. This is the role of transcriptional silencing by unliganded nuclear receptors in the regulation of development, differentiation, and carcinogenesis.[44] Unliganded RA/THR heterodimers strongly *trans*-repress the basal activity of the target gene. Silencing of gene expression by the unliganded receptor requires assistance by co-repressors. Two such co-repressors partake in the silencing effects. These factors have been named N-CoR (nuclear co-repressor) and SMRT (silencing mediator for RAR and THR). These co-repressors bind effectively only to unliganded RA/THR receptors. The silencing effect of the receptor–co-receptor– SMRT–DNA interaction is abolished by the ligand, but the release of the repressor on binding the ligand depends on the gene element, RE, to which the receptor–repressor complex is bound. When the repressor is not released upon ligand binding, constitutive repression of the gene can be the consequence.

Which ligand determines the specificity of signalling in heterodimeric receptors?

An important point is which signal—the retinoic acid or the thyroid hormone or the vitamin D signal, or all of them—is transmitted by heterodimeric RXR/THR or RXR/VDR (vitamin D receptor)? The answer is that each subunit in the heterodimer can bind its cognate ligand and transmit its message. Heterodimerization is a safeguard against complete loss of function, when one of the subunits is non-functional. This explains also why single knock-outs of only one monomer of a heterodimer had less effect than expected. Double knockouts of both subunits are often necessary for a noticeable change in phenotype. Pierre Chambon[45] has pointed out that even when heterodimeric retinoid receptor chimeras should respond also to thyroid hormone, vitamin D, or other ligands, it would not necessarily mean that the retinoic acid ligand has become ineffectual, because in the body all the different ligands are not expected to be delivered simultaneously and at the same concentration to the receptor. Instead, each ligand is provided in accordance with its own characteristic, highly regulated, spatiotemporal pattern of synthesis and release.

Conclusions

The structural design of retinoic acid receptors explains, at least in part, the pleiotropy of this class of receptors. It is a consequence of the combinatorial possibilities offered by heterodimerization. The many different functional units that can be formed by heterodimerization make it understandable why so many different gene response elements can be recognized. This, and a distinct spatiotemporal pattern of expression, both in the developing embryo and in differentiated tissues, suggests that each RAR and RXR isoform and each heterodimer formed may have a unique function. Pierre Chambon has directed attention to the large repertoire of combinatorial arrangements. These combina-

torial arrangements, together with a rising number of co-activators and co-repressors, make clear that there are many more possibilities to account for pleiotropic effects than are actually realized. Pleiotropy is part of evolution. It is important biologically and should not be mistakenly equated with redundancy.[46]

References

1. Ming-Jer Tsai and B. W. O' Malley. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu Rev Biochem*, **63**, 451–486, 1994.
2. M. Beato, P. Herrlich and G. Schütz. Steroid hormone receptors: many actors in search of a plot. *Cell*, **83**, 851–857, 1995.
3. C. S. Thummel. From embryogenesis to metamorphosis: The regulation and function of *Drosophila* nuclear receptor superfamily members. *Cell*, **83**, 871–877, 1995.
4. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell.* Garland Publishing, New York, 1994.
5. J. M. Greer, J. Puetz, K. R. Thomas, and M. R. Capecchi. Maintenance of functional equivalence during paralogous Hox gene evolution. *Nature*, **403**, (6770) 661–665, 2000.

   For a comment, see D. Duboule. News and views: A Hox by any other name. *Nature*, **403**, 607–610, 2000.
6. D. Duboule and A. S. Wilkins. The evolution of 'bricolage'. *Trends Genet*, **14**, (2) 54–59, 1998.
7. D. Acampora, V. Avantaggiato, F. Tuorto, P. Barone, M. Perera, D. Choo, D. Wu, G. Corte, and A. Simeone. Differential transcriptional control as the major molecular event in generating Otx1-/- and Otx2-/- divergent phenotypes. *Development*, **126**, (7) 1417–1426, 1999.
8. J. Y. Wu, L. Wen, W.-J. Zhang, and Y. Rao. The secreted product of *Xenopus* Gene lunatic Fringe, a vertebrate signaling molecule. *Science*, **273**, 355–358, 1996.

   Also W. Roush. Research News; Receptor for vital protein finally found. *Science*, **273**, 309, 1996.
9. L. Stryer. *Biochemistry*, (3rd edn). W.H. Freeman, New York, 1988.
10. A. Labhart. *Clinical endocrinology, theory and practice.* Springer Verlag, New York, 1986.
11. S. P. Williams and P. B. Sigler. Atomic structure of progesterone complexed with its receptor. *Nature*, **393**, 392, 1998.
12. F. Ulrich Hartl. Molecular chaperones in cellular protein folding. *Nature*, **381**, 571–580, 1996 (see Fig. 1).
13. M.-J. Tsai and W. O' Malley. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu Rev Biochem*, **63**, 451–486, 1994 (see Fig. 10, p. 474).
14. C. E. Stebbins, A. A. Russo, C. Scheneider, N. Rosen, F. U. Hartl, and N. P. Pavletich. Crystal structure of an HSP90–Geldamycin complex: Targeting of a protein chaperone by an antitumor agent. *Cell*, **89**, 239, 1997.
15. D. F. Smith, L. Whitesell, S. C. Nair, S. Y. Chen, V. Prapanich, and R. A. Rimerman. Progesterone receptor structure and function altered by geldanamycin, an Hsp90-binding agent. Mol Cell Biol, **15**, 6804–6812, 1995.
16. D. Picard, B. Khursheed, B. Garabedian, M. G. Fortin, S. Lindquist, and K. R. Yamamoto. Reduced levels of Hsp 90 compromise steroid receptor action *in vivo*. *Nature*, **348**, 166–168, 1990.
17. D. F. Nathan and S. Lindquist. Mutational analysis of Hsp90 function: Interactions with a steroid receptor and a protein kinase. Mol Cell Biol, **15**, 3917–3925, 1995.
18. T. W. Schulte, M. V. Blagosklonny, C. Ingui, and L. Neckers. Disruption of the Raf-1-Hsp 90 molecular complex results in destabilization of Raf-1 and loss of Raf-1-Ras association. *J Biol Chem*, **270**, 24585-24588, 1995.
19. E. Pennisi. Differing roles found for Estrogen's two receptors. *Science*, **277**, 1439, 1997.
20. K. Paech, P. Webb, G. G. J. M. Kuiper, S. Nilsson, J.-Å. Gustafsson, P. J. Kushner, and T. S. Scanlan. Differential ligand activation of estrogen receptors ERα and ERβ at AP1 sites. *Science*, **277**, 1508–1510, 1997.
21. S. F. Arnold, D. M. Klotz, B. M. Collins, P. M. Vonier, L. J. Guillette Jr, and J. A. McLachlan. Synergistic activation of estrogen receptor with combinations of environmental chemicals [see comments]. *Science*, **272** (5267), 1489–1492, 1996 [retracted by J. A. McLachlan. *Science*, **277** (5325), 462–463, 1997].
22. R. V. Weatherman, R. J. Fletterick, and T. S. Scanlan. Nuclear receptor ligands and ligand binding domain. *Annu Rev Biochem*, **68**, 576, 1999.

23. S. A. Ofiate, S. Y. Tsai, M.-J. Tsay, M.-J., B. W. O'Malley. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science*, **270**, 1354–1357, 1995.

24. D. Chakravarti, V. J. LaMorte, M. C. Nelson, T. Nakajima, I. G. Schulman, H. Juguilon, *et al.* Role of CBP/P300 in nuclear receptor signalling. *Nature*, **383**, 99–103, 1996.

    For a comment see: R. Janknecht, and T. Hunter. A growing coactivator network, News and views. *Nature*, **383**, 22–23, 1996.

25. P. Kastner, M. Mark, and P. Chambon. Nonsteroid nuclear receptors: What are genetic studies telling us about their role in real life? *Cell*, **83**, 859–869, 1995.

26. W. Seol, H. S. Choi, and D. D. Moore. An orphan nuclear hormone receptor that lacks a DNA binding domain and heterodimerizes with other receptors. *Science*, **272** (5266), 1336–1339, 1996.

27. G. Halder, P. Callaerts, and W. J. Gehring. New perspectives on eye evolution. *Curr Opin Genet Develop*, **5**, 602–609, 1995.

28. W. J. Gehring. The master control gene for morphogenesis and evolution of the eye. *Genes to Cells*, **1**, 11–15, 1996.

29. S. I. Tomarev, P. Callaerts, L. Kos, R. Zinovieva, G. Halder, W. Gehring, and J. Pitiagorsky. Squid Pax-6 and eye development. *Proc Natl Acad Sci, USA*, **94**, 2421–2426, 1997.

30. P. Callaerts, G. Halder, and W. J. Gehring. Pax 6 in development and evolution. *Annu Rev Neurosci*, **20**, 483–532, 1997.

31. B. Lin, G. Q. Chen, D. Xiao, S. K. Kolluri, X. Cao, H. Su, and X. K. Zhang. Orphan receptor COUP-TF is required for induction of retinoic acid receptor beta growth inhibition, and apoptosis by retinoic acid in cancer cells. *Mol. Cell. Biol.* **20**, (3) 957–970, 2000.

32. W. Krezel, V. Dupe, M. Mark, A. Dierich, P. Kastner, and P. Chambon. RXR gamma null mice are apparently normal and compound RXR alpha +/–/RXR beta –/–/RXR gamma –/– mutant mice are viable. *Proc Natl Acad Sci, USA*, **93** (17), 9010–9014, 1996.

33. P. Kastner, J. M. Grondona, M. Mark, A. Gansmuller, M. LeMeur, D. Decimo, *et al.* Genetic analysis of RXR alpha developmental function: convergence of RXR and RAR signaling pathways in heart and eye morphogenesis. *Cell*, **78** (6), 987–1003, 1994.

34. P. Kastner, M. Mark, N. Ghyselinck, W. Krezel, V. Dupe, J. M. Grondona, and P. Chambon. Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development*, **124** (2), 313–326, 1997.

35. Y. Chen, J. Clifford, C. Zusi, J. Starrett, D. Tortolani, J. Ostrowski, *et al.* Two distinct actions of retinoid-receptor ligands. *Nature*, **382** (6594), 819–822, 1996.

36. B. Roy, R. Taneja, and P. Chambon. Synergistic activation of retinoic acid (RA)-responsive genes and induction of embryonal carcinoma cell differentiation by an RA receptor alpha (RAR alpha)-, RAR beta-, or RAR gamma-selective ligand in combination with a retinoid X receptor-specific ligand. *Mol Cell Biol*, **15** (12), 6481–6487, 1995.

37. F. Rastinejad, T. Perlmann, R. M. Evans, P. B. Sigler. Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature*, **375**, 203, 1995.

38. J.-P., Renaud, N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer, and D. Moras, Crystal structure of the RAR-γ ligand-binding domain bound to all-*trans* retinoic acid. *Nature*, 378, 681–689, 1995.

39. W. Bourguet, M. Ruff, P. Chambon, H. Gronemeyer, and D. Moras. Crystal structure of the ligand binding domain of the human nuclear receptor RARγ. *Nature*, **375**, 377–382, 1995.

40. R. L. Wagner, J. W. Apriletti, M. E. McGrath, B. L. West, J. D. Baxter, and R. J. Fletterick. A structural role for hormone in the thyroid hormone receptor. *Nature*, **378**, 690–697, 1995.

41. J. M. Wurtz, W. Bourguet, J. P. Renaud, V. Vivat, P. Chambon, D. Moras, and H. Gronemeyer. A canonical structure for the ligand binding domain of nuclear receptors. *Nature Struct Biol*, **3**, 87–94, 1996.

42. R. Kurokawa, M. Soderstrom, A. Horlein, S. Halachmi, M. Brown, M. G. Rosenfeld, and C. K. Glass. Polarity-specific activities of retinoic acid receptors determined by a co-repressor [see comments]. *Nature*, **377** (6548), 451–454, 1995.

43. J. D. Fondell, M. Guermah, S. Malik, and R. G. Roeder. Thyroid hormone receptor-associated proteins and general positive cofactors mediate thyroid hormone receptor function in the absence of the TATA box-binding protein-associated factors of TFIID. *Proc Natl Acad. Sci, USA*, **96** (5), 1959–1964, 1999.

44. A. J. Horlein, A. M. Naar, T. Heinzel, J. Torchia, B. Gloss, R. Kurokawa, *et al.* Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor [see comments]. *Nature*, **377** (6548), 397–404, 1995.

    Also: J. D. Chen and R. M. Evans. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*, **377** (6548), 454–457, 1995.

45. P. Chambon. A decade of molecular biology of retinoic acid receptors. *FASEB J*, **10**, 940–954, 1996.

46. H. Chiba, J. Clifford, D. Metzger, and P. Chambon. Distinct retinoid X receptor-retinoic acid receptor heterodimers are differentially involved in the control of expression of retinoid target genes in F9 embryonal carcinoma cells. *Mol Cell Biol*, **17** (6), 3013–3020, 1997.

3

Global cell regulatory programmes

12

Regulation of the cell cycle

Cells have universal programmes to control growth, proliferation, and death. One of these programmes is the cell cycle which controls cell division. Progression through the cycle is controlled by external and internal signals. The cell cycle has a S phase, in which DNA is replicated, and an M phase, a mitotic phase, in which the replicated genetic information is divided into equal parts and passed on to the daughter cells. The interval between the S phase and the M phase in the cycle, in the right-handed direction of the clock, is the G_2 phase, followed by the interval between the M phase and the next S phase, which is the G_1 phase (Fig. 12.1; for more information, see ref. 1). However, in some cases, cited by Kim Nasmyth,[2] biochemical and structural transitions of the cell-cycle machinery do not coincide with the cell-cycle phases to which they are assigned.

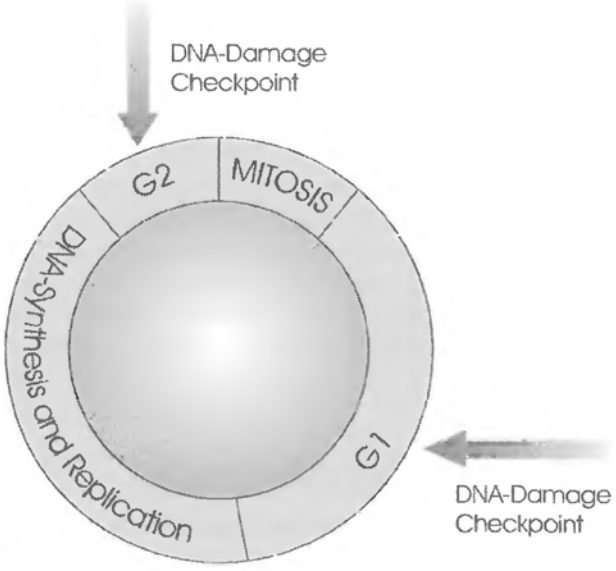


Fig. 12.1 The phases between the end and the start of mitosis are the interphases. G_1 and G_2. In clockwise direction, between the end of G_1 and the start of G_2 comes the S-phase, when DNA is synthesized and replicated. After mitosis cells may enter a quiescent phase, G_0, or proceed to G_1 and eventually to the S-phase. Checkpoint controls sense DNA damage and arrest the cell cycle in either G_1 or G_2. These controls prevent DNA synthesis and replication and mitosis, either when DNA is damaged or when DNA has not been properly replicated.

We shall focus on control of the cell cycle in multicellular organisms. The length of the cycle in cells of multicellular organisms is variable. Some cells, such as nerve cells for example, have an active metabolism and can grow, but they do not divide. Likewise, skeletal muscle cells do not divide, but they increase their size greatly on training. However , some new cells may be formed in the adult organisms from stem cells. On the other hand, epithelial cells in the outer layers of the skin and of the intestine divide twice a day or more. These variations reflect the length of time the cells remain arrested in G_0 before starting a new round of DNA replication. Thus, cell growth can be uncoupled from the cell cycle and from cell division. The cell can decide whether to enter G_1, DNA synthesis, and mitosis or to stay dormant in G_0. But once the decision has been made to enter mitosis, the time that elapses from the beginning of the S phase through mitosis is quite constant, ranging from 12 to 24 hours in adult, mammalian cells (for more information, see ref. 3).

The molecular basis of the cell cycle

Although nearly all cellular processes have spatiotemporal patterns and are timed in one way or the other, the outstanding example of a built-in clock is the cell cycle. The wheels and cogs of the clock that time the cell cycle are the cyclins and the cyclin-dependent kinases (Cdks).

The cyclins and the cyclin-dependent kinases

Cyclins are the regulatory moieties of Cdks, the cyclin-dependent kinases. Without cyclins, Cdks are inactive. Each cyclin has a Cdk as partner and each cyclin/Cdk pair has a different function in the cell cycle. Whereas the concentration of the Cdks is constant, the cellular concentrations of the cyclins vary with the cell-cycle phases, as the name implies. These fluctuations are of regulatory significance, and so are association/dissociation of cyclin/Cdk complexes.

Cyclins in higher eukaryotes are classified and divided into eight families, from A to H, based on sequence relationships and the time of appearance in the cell cycle. Although cyclins are a family of diverse proteins, they all have a common region of about 100 amino acids in the C-terminal part. This domain is the cyclin box.

Figure 12.2a shows some of the players and their targets in the G_1 and S phases, and Fig. 12.2b those in the G_2 and M phases. There exist at least seven different Cdks in mammalian cells. Each of them is called up and activated in a particular phase of the cell cycle. Cyclins of the E and D families and the corresponding Cdks transmit positive growth-factor signals in the G_1 phase, and help the cell to start a S phase and synthesize DNA. After a certain stage in the G_1 phase, the cell has reached the point of no return and becomes unresponsive to external signals and is committed to enter the S phase. This point is the restriction point. From that point on the cell cycle is controlled by internal signals. These are the checkpoint controls.

Mechanisms for timing of cell-cycle transitions

As cells progress through the cycle, they undergo several discrete transitions. Each transition is an irreversible shift from one level of activities to an other. The most remarkable

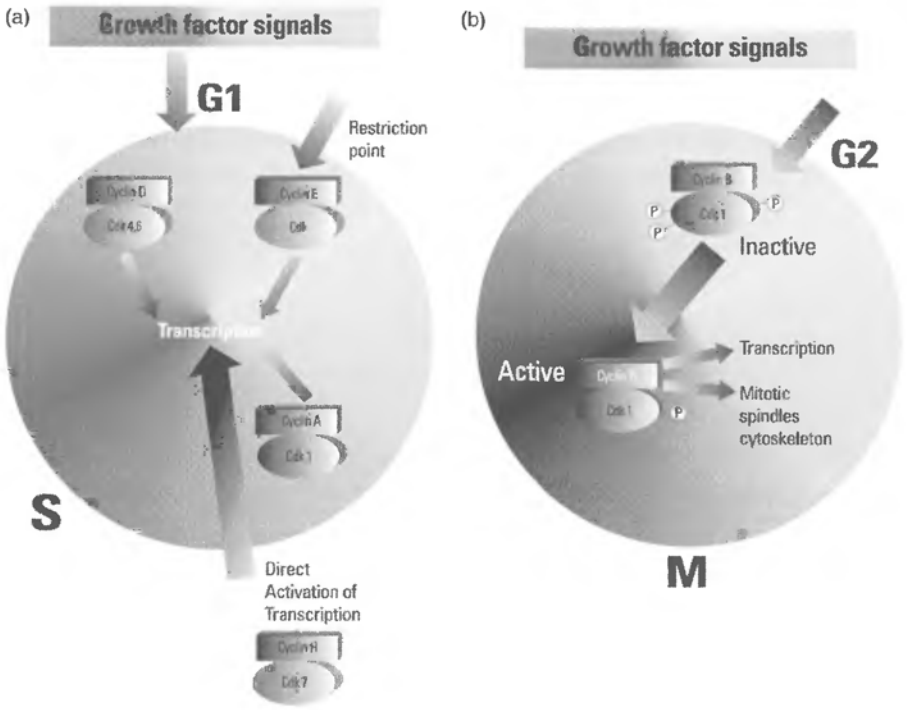


Fig. 12.2 (a) The D-type cyclins (D1, D2, and D3) and the cyclin D-dependent kinases, Cdk4/6, control the mammalian cell-cycle clock during the G_1 phase and promote the G_1/S transition. Each of the D cyclins has its own Cdk and CdkI (Cdk inhibitor), and each D-cyclin accepts signals from a different growth factor. Moreover, the three D-type cyclins are differently expressed in differentiated cells and tissues, indicating that they participate in a cell-type-specific control of proliferation. In transgenic mice where the cyclin D1 gene is disrupted, proliferation is suppressed only in the retina and in the mammary gland epithelium, whereas disruption of the cyclin D2 gene suppresses proliferation of ovarian granulosa cells, making cyclin D2-deficient female mice sterile and non-responsive to the follicle-stimulating hormone (FSH). Overexpression of cyclin D2 has been related to cancers of the ovaries and testes in humans.[4] Cyclin E–Cdk is also active in the G_1 phase. It guides the cycle through the S phase and prepares the entry into the G_2 phase. Cyclin A–Cdk1 is targeted to general and specific transcription factors and controls transcription in the S phase. Cyclin H–Cdk7 promotes transcription in the S phase by directly activating polymerase II, through phosphorylation of the carboxy-terminal part of Pol II (see Chapter 9). (b) Cyclin B/Cdk1 is active in the M phase and controls expression of proteins required for the construction of the mitotic spindles and the reorganization of the cytoskeleton in mitosis. The activity of Cdk1 is dependent on its phosphorylation state. (Cyclin F may also be active in the G_2 phase, but its partners are not known.)

property of the cell cycle is that these transitions occur at a precise time and in an orderly and coordinated fashion. The cell instructs the cycle when to begin DNA replication and when to enter mitosis and divide. A checkpoint monitors DNA replication. When damage has been sensed, the cell cycle is stopped.[5] This implicates operation

of a timing device that ensures that the next event does not start before the preceding one is completed. This is like a substrate–product relationship in a multi-enzyme chain, such as glycolysis, where the preceding reaction ceases when the substrate is used up.

Cell-cycle transitions are driven by the cyclin-dependent kinases

Cyclin–Cdks were appropriately named the cell cycle's 'engines'.[6] The cyclin-dependent protein kinases (Cdks) are soluble serine/threonine kinases of 34–40 kDa. The Cdks share with other serine/threonine protein kinases sequence similarities, including a subset of residues that is essential for catalytic activity. Cdks contribute the catalytic subunit, whereas the regulatory subunit is contributed by a cyclin. Cyclins control the kinase activity, determine the substrate specificity and the subcellular location of Cdks. Each of these processes is a potential site of regulation. The major substrates of the Cdks are proteins regulating gene transcription. Cdks can be controlled in three major ways:

1. by phosphorylation dephosphorylation,
2. by Cdk inhibitors (CdkIs); and
3. by dissociation of the active cyclin–Cdk complex and subsequent proteolysis of the cyclin.

Control of Cdks

CAKs are Cdk-activating kinases. They are apparently constitutively active, so that activation of Cdk by CAKs is controlled primarily by the accessibility of phosphorylation sites in the cyclin–Cdk complex to CAKs. Activation is the consequence of phosphorylation of

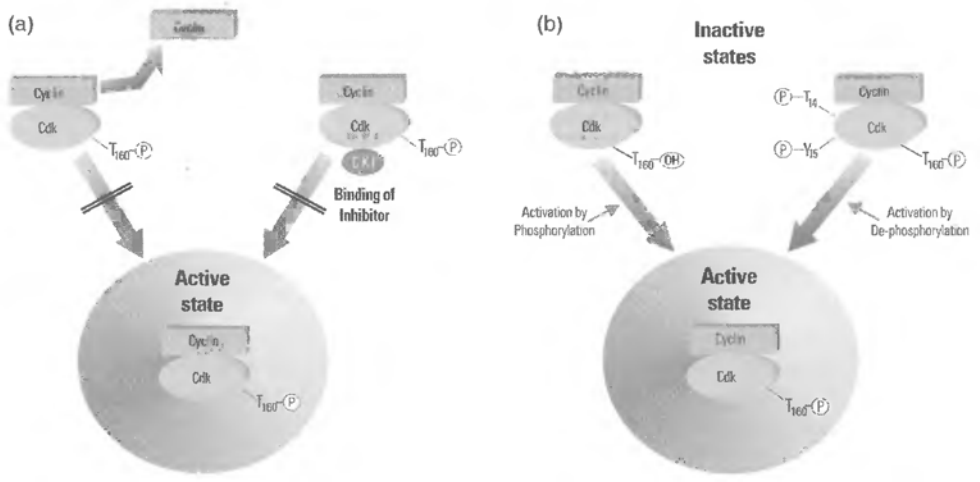


Fig. 12.3 (a) On the right, blockage of activation by inhibition of Cdk–cyclin by the Cdk inhibitor (CKI) and, on the left, de-activation by dissociation of the regulatory cyclin moiety. (b) Activation of Cdk1, Cdc2) by phosphorylation of threonine 160, (threonine 161 in human Cdk1) is indicated on the left, and activation by dephosphorylation of the inhibitory threonylphosphate 14 and tyrosylphosphate 15 by the Cdc25 phosphatase is shown on the right. The regulation of the Cdc25 phosphatase is explained later in Fig. 12.16.

a strategically located single threonine residue in the active-site cleft of the Cdk. The phosphate bound to the threonine triggers a conformational transition and makes the kinase accessible to the cyclin and stabilizes the 1 : 1 Cdk–cyclin complex. There are activating and inhibiting sites. Whereas the Thr160 site in the yeast Cdk is an activating site, phosphorylation at Thr14 and Tyr15 (Y15), within the ATP-binding domain, is inhibitory. Therefore, phosphorylation of Thr160 and dephosphorylation of the phospho-Thr14 and phospho-Y sites lead to activation. Thus, control of Cdk activity by dephosphorylation is as important as activation by phosphorylation. Three Cdk-specific phosphatases have been identified in eukaryotic cells (Cdc25A, B, and C). These phosphatases are dual-specificity phosphatases, dephosphorylating phospho-Ser/Thr and phospho-Tyr residues. The Cdc25 phosphatases are themselves subject to control by phosphorylation dephosphorylation on Ser/Thr sites. Phosphorylation by cyclin B/Cdk1 leads to activation of the phosphatase. The Cdc25 phosphatases are de-activated when the activating phosphothreonyl sites are dephosphorylated by the serine/threonine-specific phosphatases, PP1 and PP2A. Like the cyclin–Cdks, each of the modifier phosphatases is expressed in a different phase of the cell cycle. Major control mechanisms of Cdks are summarized in Fig. 12.3.

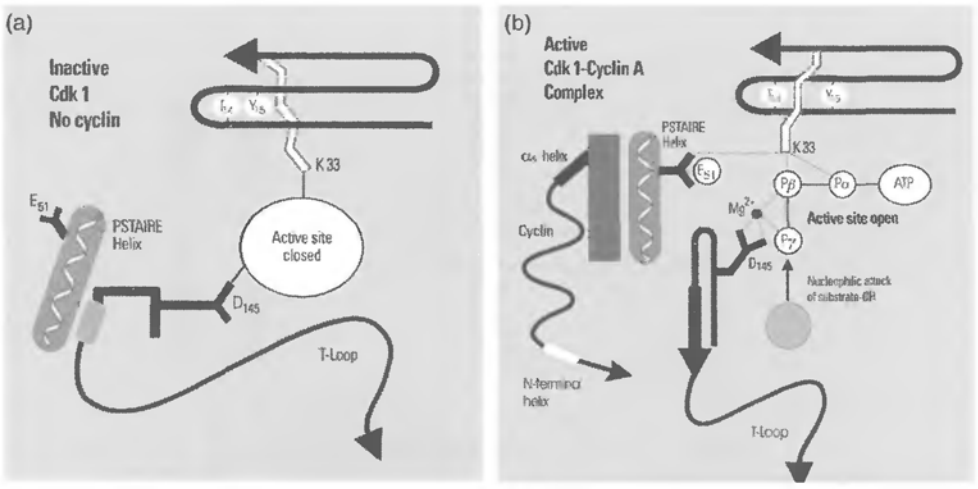


Fig. 12.4 (a) In the inactive state of Cdk, a helix of 16 amino acids, the 'PSTAIRE' helix, is positioned above the active-site cleft and disorients the side-chains, forming the ATP-binding site. Moreover, a long loop, the T-loop, rising from the carboxy- terminal lobe, impedes cleavage of ATP and phosphoryl transfer to the substrate. (b) The binding of cyclin gets the T-loop out of the way. The catalytic cleft becomes now accessible. The change in the orientation of the 'PSTAIRE' helix brings E51 in the right position with respect to K33 and D145. These residues (Glu51, Lys33 and Asp145) align the βγ-bond of the Mg^{2+}–ATP complex with the hydroxyl group of the serine/threonine of the protein, which is accepted as substrate and phosphorylated. This arrangement makes an 'in-line' nucleophilic attack of the substrate serylhydroxyl on the βγ-bond of ATP possible. Whereas the Cdk undergoes large conformational changes on binding cyclin, the cyclin conformation does not change much. The cyclins apparently serve as a rigid scaffold to accommodate and support the Cdks. (Drawn with permission of Dr. J. Pines and Nature, using information in refs 7. See also ref. 8))

The best-characterized cyclin-dependent kinase in mammals is Cdk1, also named Cdc2. High-resolution structures are available of the inactive Cdk1 without cyclin, and a partly active Cdk1–cyclin A complex. A schematic, simplified comparison of the essential features of the inactive and the active Cdk gives a clue as to how the enzyme is activated (Fig. 12.4a, b). Activation results from two events:

1. On binding to the Cdk, the α5 helix of cyclin A re-orients the 'PSTAIRE' helix of Cdk1, opening the active site cleft. (PSTAIRE is proline, serine, threonine, alanine, isoleucine, arginine, glutamic acid).
2. The interaction of the Cdk with cyclin gets the T-loop out of the way. This orients the γ-phosphate of ATP and permits phosphoryl transfer to the serine/threonine-OHs of the Cdk substrate.

Whereas growth factors activate the cell-cycle machinery, TGF-β sends signals to halt the cycle (Fig. 12.5). For the inhibitory response, kinase inhibitors such as $p27^{KIP-1}$ (kinase-inhibiting protein-1 with molecular mass 27 kDa) are responsible. $p27^{KIP-1}$ stops the G_1/S transition at the restriction point. It binds to the activated, phosphorylated Cdk1–cyclin A complex and perturbs the binding site for ATP, shutting down the kinase activity completely.

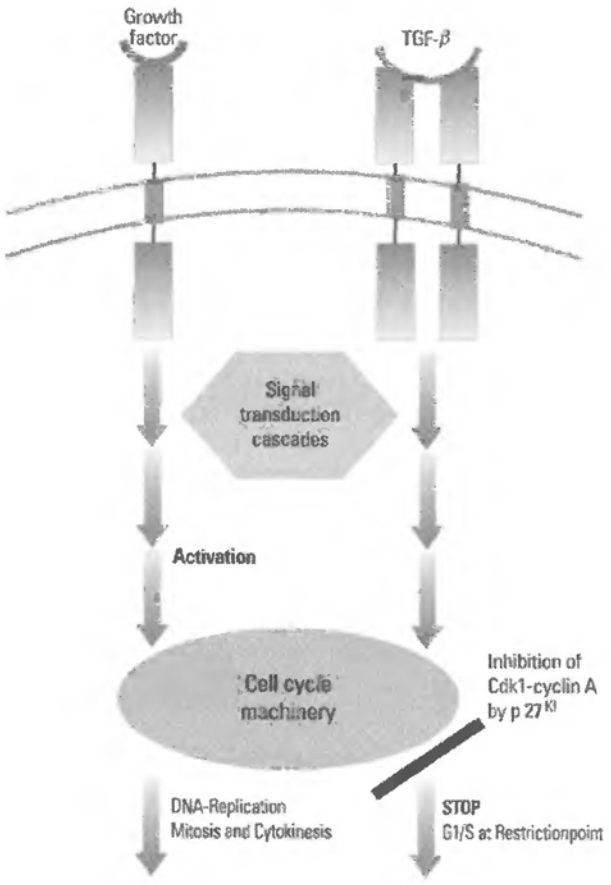


Fig. 12.5 Positive control by growth factors and negative control by TGF-β are juxtaposed. The negative signals are transmitted to Cdk1–cyclin A by the inhibitor, $p27^{KIP-1}$.

Plate 25, ref. 9 shows, on the left, the structure of the $p27^{KIP-1}$ cyclin-dependent-kinase inhibitor protein-1, bound to the cyclin A–Cdk1 complex, and on the right is introduced the cyclin-box motif of cyclin A, to which the $p27^{KIP-1}$ inhibitor binds. The key binding motif of inhibitory proteins for cyclin–Cdks appears to be conserved. Disruption of the $p27^{KIP-1}$ gene in mice removes the negative control and leads to cell proliferation in many tissues.

More CKIs have been identified. Each of these proteins is a potential regulator of the cell cycle. Since malfunctions in the cyclin–Cdk control of the cell cycle lead to uncontrolled proliferation and may cause cancer, interest in the development of effective CKIs is great.

Among substrates of cyclin Cdks are the pRB and E2F transcription factors, which can combine with each other and which play a central role in cyclin–Cdk control of transcription and in the regulation of cell-cycle progression. Deregulation may result in malignant transformation. Their role is indicated in Fig. 12.6.

E2F binds to cell-cycle-regulated DNA elements (CDEs) and regulates the expression of cyclin genes. Positive and negative control of cell-cycle-regulated DNA elements and cyclin genes by E2F is shown in Fig. 12.7; (see also refs 10, 11). Positive and negative controls of the G_1/S transition of the cell cycle are explained in more detail in Fig. 12.8. Loss-of-function mutations of the RB gene deregulate the cell cycle and promote tumour formation, (we shall come back to this in Part 4). And finally a more indirect mechanism by which the cell cycle controls the transcription machinery is shown in Fig. 12.9.

Regulation of Cdks by proteolysis

Control by phosphorylation dephosphorylation is a mechanism with which we are familiar, but probably less so with control by proteolysis of key regulatory proteins.

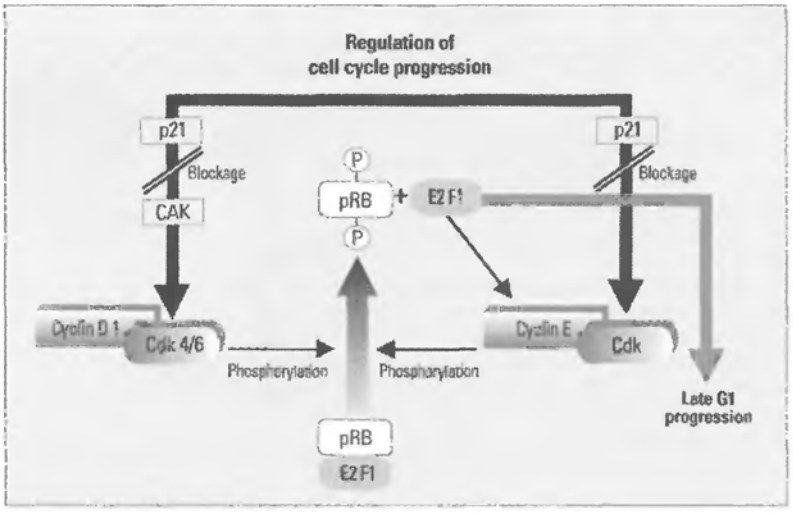


Fig. 12.6 Regulation of cell-cycle progression. pRB is the retinoblastoma suppressor protein and E2F1 is a transcription factor, active in late G_1. E2Fs are a family of proto-oncogenic transcription factors. CAK is the Cdk-activating kinase and p21 is a CKI which blocks Cdk-activation. CAK activates and phosphorylates Cdk4/6 and other Cdks. In mid-G_1, pRB is phosphorylated by cyclin D1/Cdk4/6. The phosphorylated pRB releases the transcription factor E2F1, which in turn initiates transcription of genes, required for late G_1 progression and entry into the S phase. This leads to the activation of cyclin E/Cdk which in turn activates E2F.

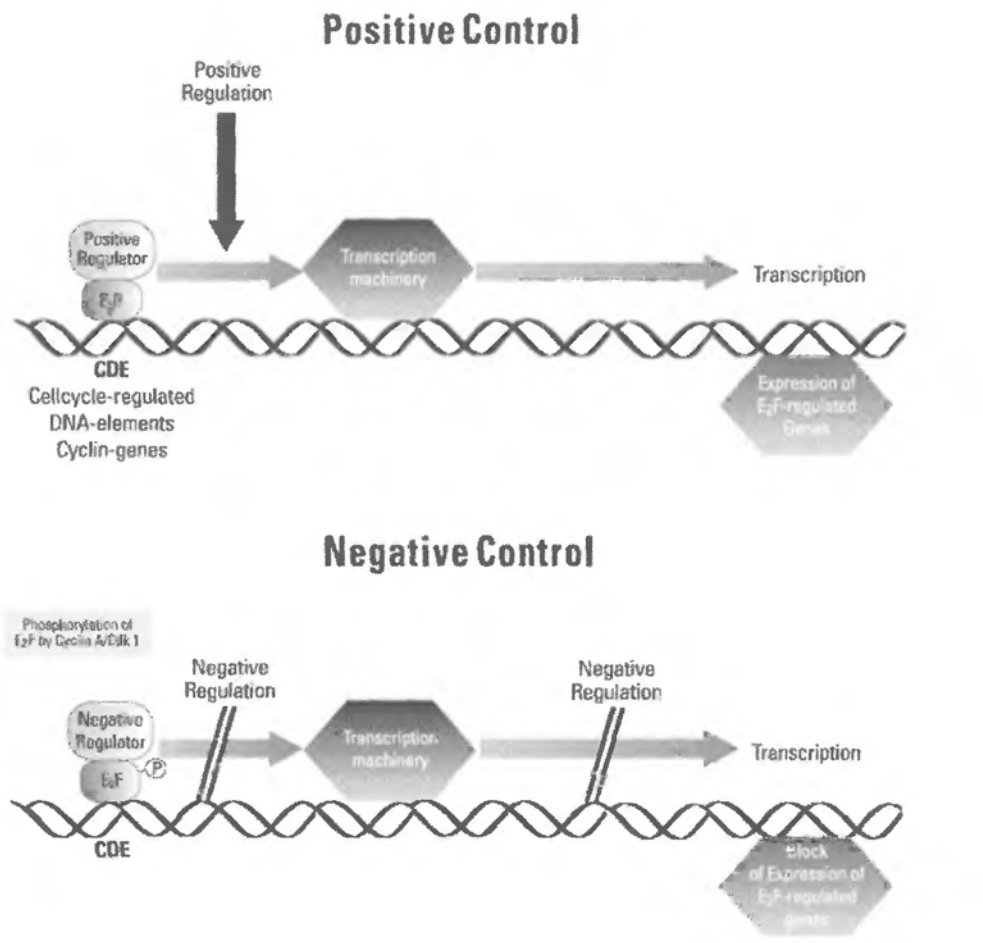


Fig. 12.7 CDEs are transcribed in G_1. Positive regulators promote activation of transcription of CDEs by E2F (positive control, above). Negative control (below) results from cyclin A/Cdk1 phosphorylation of the regulatory subunit of E2F. Phosphorylation reduces binding of E2F to DNA and leads to decreased transcriptional activity and blockage of the expression of E2F-regulated genes. (The scheme is simplified. Additional regulators of transcription must be involved, because E2F$^{-/-}$ knock-out mice have no defects in cell proliferation and develop normally, and only later in life develop a wide range of tumours. There are many CDE-binding factors, that recognize CDEs.

Although we concentrate on the role of proteolysis in the control of the cell cycle and of cell death (see Chapter 13), it is realized more and more that proteolysis of key regulators is involved in other pathways, regulating cell growth, differentiation, and embryogenesis. An example is regulation of embryonic pattern formation in *Xenopus laevis* by proteolysis of SMADs, transmitting BMP/activin signals. When, Smurf 1 (SMAD ubiquitination regulatory factor 1), a ubiquitin ligase, is expressed ectopically in *Xenopus* embryos it inhibits BMP signalling and affects pattern formation (*hSmurf 1* is a human gene, encod-

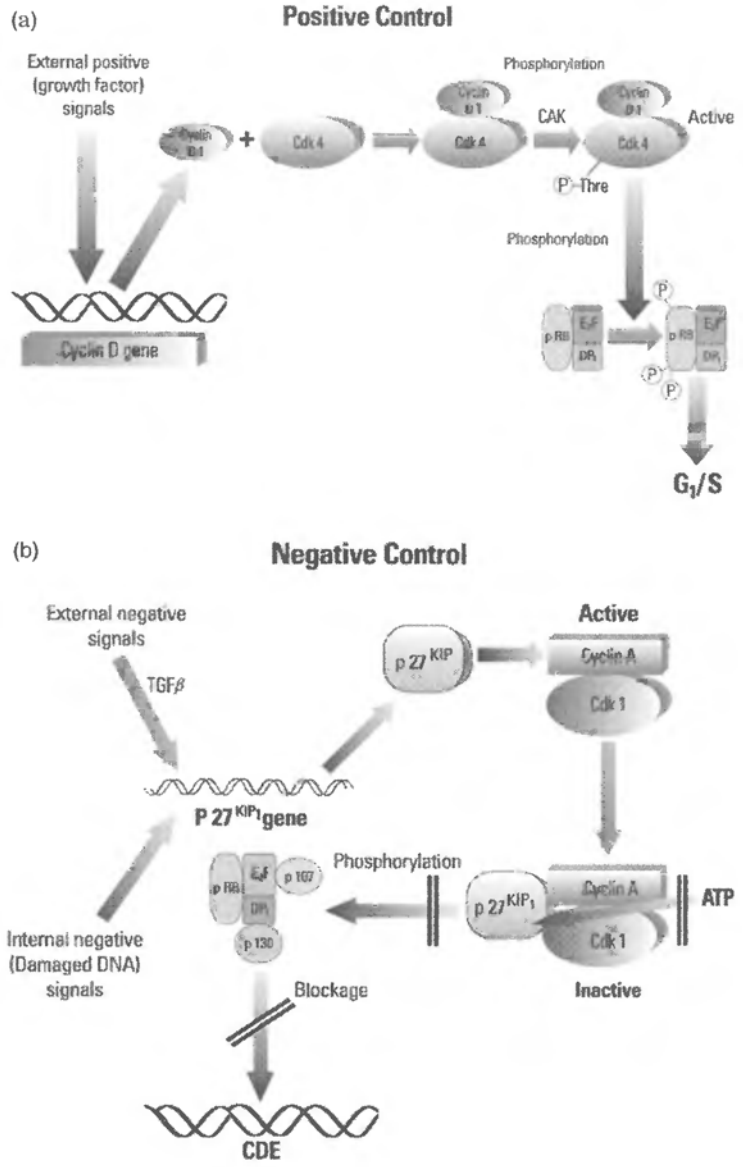


Fig. 12.8 (a) Positive control of the G_1/S transition. Positive signals promote expression of the genes encoding D cyclins. Cyclin D1 combines with Cdk4 and forms a cyclin–Cdk complex. The complex is activated by phosphorylation of Cdk4 by CAK. The active complex adresses the transcription factor pRB–E2F complex and phosphorylates pRB. The consequence is promotion of the G_1/S transition. E2F proteins are heterodimeric DNA-binding proteins which have a DP-1 regulatory subunit and binding sites for the tumour suppressors, pRB and p53, and presumably for other transcription-regulating proteins. The E2F complex has histone kinase activity and functions as a transactivating transcription factor. It binds to promoters of genes which regulate the transition of G_1 to the S phase. (b) How negative control stops the G_1/S transition. Negative, external or internal, signals lead to the expression of the Cdk inhibitor $p27^{KIP-1}$. Control by CKIs (KIPs) is dependent on cellular concentrations. $p27^{KIP-1}$ binds to the cyclin A/Cdk1 complex and inhibits Cdk1 by blocking access of ATP to the active site. Inactivation of cyclin A–Cdk1 by CKI prevents phosphorylation of the DP-1 regulatory subunit of E2F and forestalls activation of the pRB/E2F transcription factor complex and prevents activation of CDEs. p107 and p130 are additional regulator proteins which are required for the regulation of the pRB/E2F transcription factor complex.

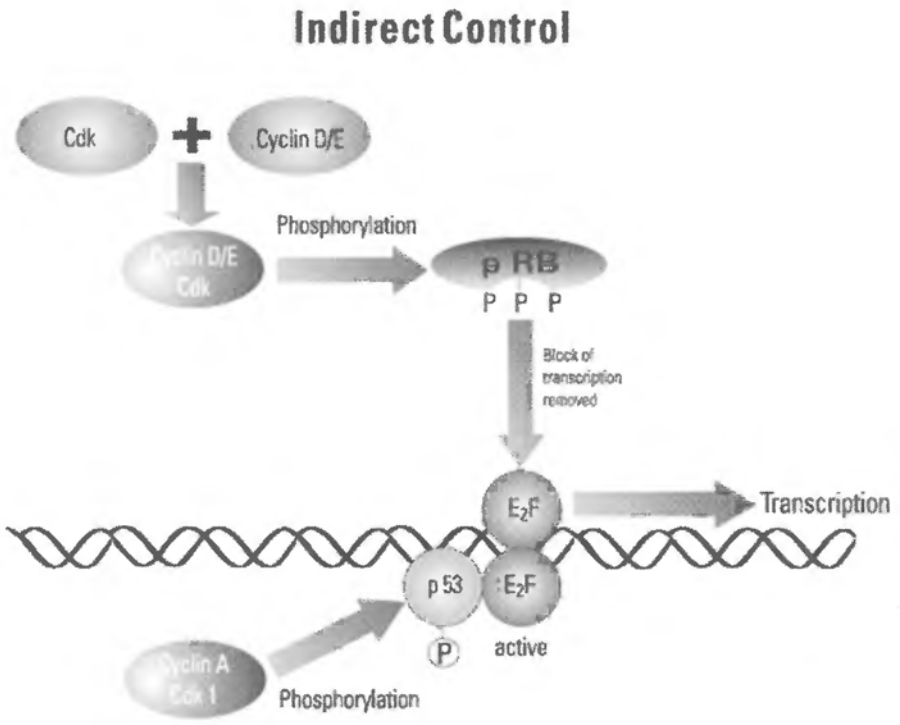


Fig. 12.9 Indirect control of transcription by cyclin–Cdks involves interactions of transcription factor E2F with the tumour suppressors p53 and pRB. Cyclin D- and cyclin E-associated kinases phosphorylate pRB and lift the block of transcription. Appropriate changes of the phosphorylation state of p53 improve the specificity of DNA binding and promote the transcriptional activity of E2F. This effect is amplified by phosphorylation of p53 by cylin A–Cdk1. (Cyclin H/Cdk7/CAK and cyclin C/Cdk8 complexes can activate transcription directly, by association with the basal transcription machinery and phosphorylation of RNA polymerase II, Pol II; see also Chapters 9 and 10).

ing a closely related homologue).[12] Limited proteolysis also plays a central role in the processing of antigens (see Chapter 14).

Proteolysis of cyclins is synchronized with the cycle.[13] Different cyclins are degraded with different timing. Proteolysis is responsible for oscillations in the activity of cyclin-dependent kinases during progression through the cycle. An example is the removal of cyclin A by proteolysis in late S phase, which accounts for the sharp decline of Cdk1 activity, making possible exit from mitosis.

Proteins targeted for proteolysis have destruction boxes, i.e. specific recognition sequences for the proteolysis machinery, similar to the death domains of proteins destroyed in apoptosis (Chapter 13).

The proteolysis machinery involved in cell-cycle transitions comprises an ubiquitinating enzyme complex,[14–16] and a multicomponent protein complex with a multicatalytic 26S proteasome. Proteasomes are abundant multisubunit particles. They comprise up to 1% of all cellular proteins.[17,18] The proteolysis machinery is known in the

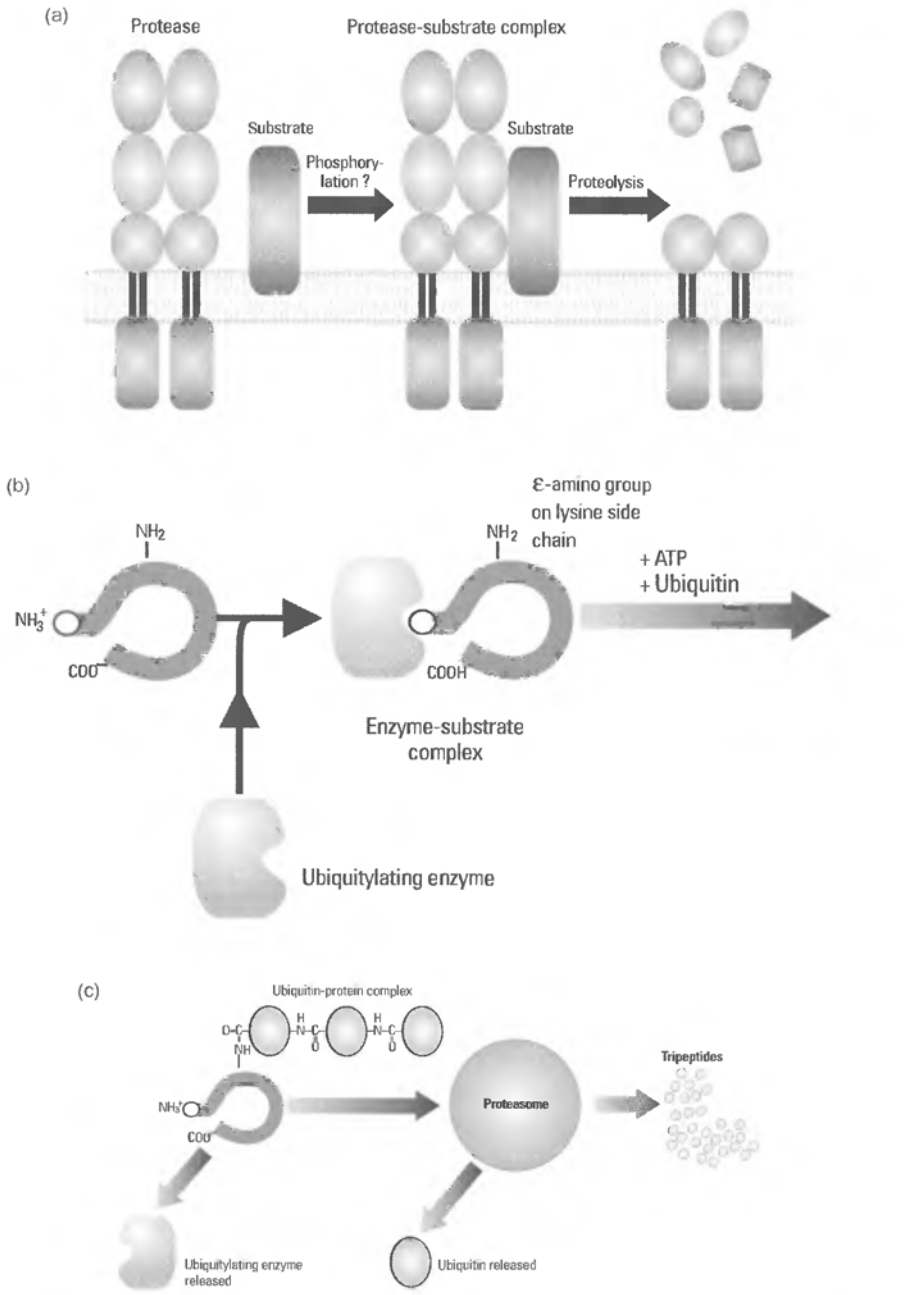


Fig. 12.10 (a) The interaction of a membrane-bound protease with its substrate and how phosphorylation might help recognition of a substrate by a protease is shown. (b) A protein destined for proteolytic destruction is recognized by the ubiquitinating enzyme and forms an enzyme–substrate complex. Subsequently, ubiquitin molecules are joined together and attached by a peptide bond to an ε-amino group of a lysine side-chain of the target protein and a ubiquitin–protein complex is formed. (c) The ubiquitinated protein is delivered to the proteasome. The protein is degraded into small peptides and the ubiquitin is recovered. (Figure 5–39 in Chapter 5 of *Molecular Biology of the Cell* by B. Alberts *et al.*[3] served as a model for schemes (b) and (c), which are reproduced with permission of Taylor and Francis, Inc.)

cell-cycle field as the 'cyclosome' or the 'anaphase promoting complex', (APC). In Fig. 12.10 the ubiquitinated proteins are directed to the 26S proteasome for proteolysis. The regulatory mechanisms which time proteolysis are still a matter of speculation. This is a challenge for future research. Points of regulatory control need not necessarily be the proteasome itself but could be any of the steps preparing the protein for proteolysis, including ubiquitination. Fig. 12.10(a), (b). Complexes of cyclins and Cdks and associated activators and inhibitors might be subject to covalent modifications, such as phosphorylation. Subsequent conformational changes might help the ubiquitinating enzymes and the protease to recognize the substrate.

There are marked differences in the rate and duration of proteolysis of cyclins in different organisms. At the early stages of amphibian embryogenesis, cyclin destruction is only of brief duration and cyclin concentrations are kept high, as one would expect, because they are needed. To the contrary, in cultured fibroblasts, proteolysis continues for several hours and is only terminated at the onset of the S phase, when all B-type cyclins have been completely removed.

Control of chromosomal DNA replication and of chromosome duplication

After having described the major instruments and the basic, biochemical mechanisms which accomplish cell-cycle transitions, we now turn to the control of chromosomal DNA replication. which guarantees that DNA is replicated only once per cycle. For the purpose of orientation, DNA replication and mitosis are related in Fig. 12.11 to the actual morphological changes of the cell, starting with interphase and ending with cytokinesis and cell division. The replicated DNA is condensed into two complete sets of chromosomes. The chromosomes are exactly partitioned with the help of mitotic spindles, composed of microtubules, to the poles of the dividing cell. The cell is then cleaved, forming two new, separate cells.

The regulatory factors that control initiation of replication of chromosomal DNA are the 'replication licensing factors' (RLFs), which consist of two major components,

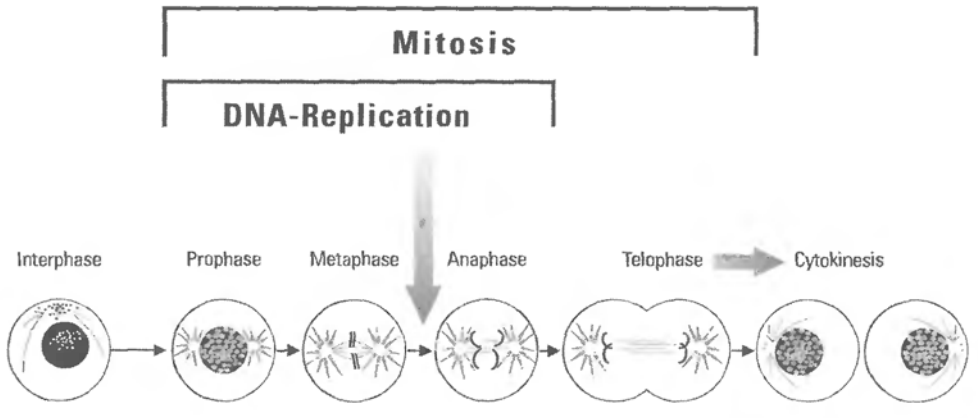


Fig. 12.11 DNA replication and mitosis are related to the actual morphological changes of the cell, starting with the interphase and ending with cytokinesis.

RLF-M and RLF-B. RLF-M is a complex of six proteins. RLF-B directs, with the help of other factors, RLF-M to the chromatin, where it 'licenses' the origins of replication. The 'licensing' factors set the stage for DNA replication and start the S phase. Once replication has started, the licensing complex is removed. Each step of the replication of chromosomal DNA, starting from the origins of replication (the pre-replication complexes, pre-RCs), is controlled by Cdks. The S phase is initiated on formation of a complex containing cyclin A–Cdk1 and the transcription factor E2F (and an additional regulatory protein, p107). E2F directs the complex to cell-cycle-regulated DNA sites, located upstream of genes which encode the enzymes required for DNA replication. (see: Ref: 19)

Control of chromosome duplication in eukaryotic cells must guarantee the faithful duplication of the genetic information of a cell, before it is transferred to the next generation. Formation of a pre-replicative state (the pre-RC complex) requires a highly ordered assembly of many proteins which together form the origin recognition complex (ORC) at the origins of DNA replication. Once the pre-RCs are formed, the appropriate Cdks are activated. Activation is turned on by proteolytic removal of a Cdk inhibitor. This guarantees that new pre-RCs can only be formed again when all the M-phase cyclins have been degraded and the cell has entered G_1.

To sum up: A central point of intervention of cell-cycle-regulated protein kinases, is to prevent further DNA replication in the post-replicative state.

The telomerase life-cycle clock

Chromosome duplication in eukaryotic cells requires the enzyme telomerase which replicates the chromosome ends, the telomeres, a task that cannot be carried out by DNA replicases. (Telomerase is a ribonucleoprotein where the RNA moiety serves as template for addition of short nucleotides to the 3' end of the chromosome. The telomerase contains reverse transcriptase motifs which are essential for the duplication of the telomeres. For further information see ref. 20.). The cell-cycle clock is tightly coupled to the telomerase 'clock'.

All finally differentiated eukaryotic cells lack telomerase activity. Thus, when cells divide, the ends of the chromosomes can not be duplicated and the telomeres get shorter and shorter with every cell division, until the cell finally dies. Thus, the lack of telomerase activity may determine the number of divisions a cell can go through. And indeed, disruption of the telomerase gene in yeast leads to senescence and death (see Fig. 12.12). The ageing process in the yeast cells was reversed when human telomerase was expressed. Expression of the catalytic subunit of the human telomerase holoenzyme (hTERT) also enabled human cells to avoid senesce and to multiply indefinitely. This raised hopes to immortalize human cells, without interfering with their differentiation. (For more information, see ref. 21 and Chapter 17).

If the telomerase activity alone were to determine the life span of a cell, one would expect that cells such as muscle cells and neurons, that do not divide, never age. But, this is obviously not the case. This paradox shows that the relationship between control of senescence and cell death is much more complicated. Recent work of T. Kiyono *et al.*[22], reviewed by Robert Weinberg,[23] with two human epithelial cell types (keratinocytes and mammary epithelial cells), showed that making cells immortal requires not only expression of the *hTERT* gene and of the telomerase, but also inactivation of the retinoblastoma (RB) tumour-suppressor pathway (the RB pathway halts cells in the G_1 phase of the cell cycle in response to a wide range of inhibitory signals, see Fig. 12.12).

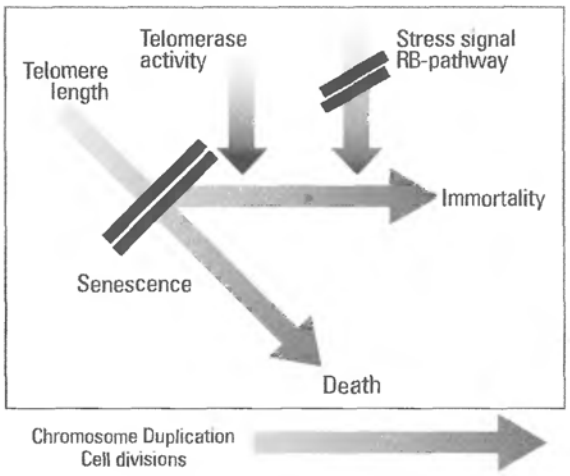


Fig. 12.12 With each cell division, the telomeres at the end of the chromosomes become shortened. The model shows that in order to stop senescence and death, both telomerase activity and blockage of the RB pathway are needed. Telomerase activity and blockage of the RB pathway are intimately linked, because the factors that block the RB pathway (the *myc* oncogene and the E6 oncogene of the human papillomavirus, see Part 4) also upregulate expression of hTERT.[24] In contrast to normal differentiated cells, cancer cells have telomerase activity. In this way, cancer cells overturn the life-span-controlling clock and become immortal. Since, about 90% of all human cancers are derived from epithelial cells, it follows that in epithelial cancer cells the RB pathway has also been inactivated. To the contrary, when the tumour-suppressor p 53 accumulates, the cells age prematurely.[25]

Cytokinesis

Once in every turn of the cycle, the cell divides and splits into two daughter cells. The duplicated chromosomes must be separated with perfect precision. Moreover, all of

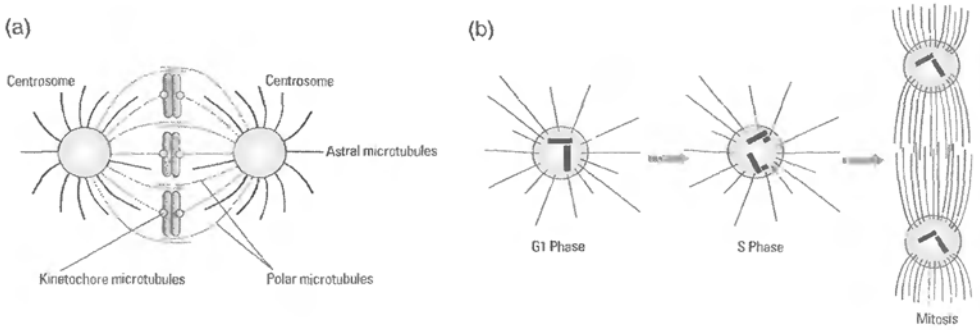


Fig. 12.13 (a) We distinguish the centrosomes and the kinetochores. There are three classes of microtubules in a mitotic spindle. Attached to the kinetochores are the kinetochore-microtubules and attached to the centrosomes are the astral and the polar microtubules, which become the spindle poles. (b) The centrosomes undergo characteristic changes during the cell cycle. In the S phase, daughter centrioles begin to form. Finally, the centrosome divides to form the mitotic spindle poles.

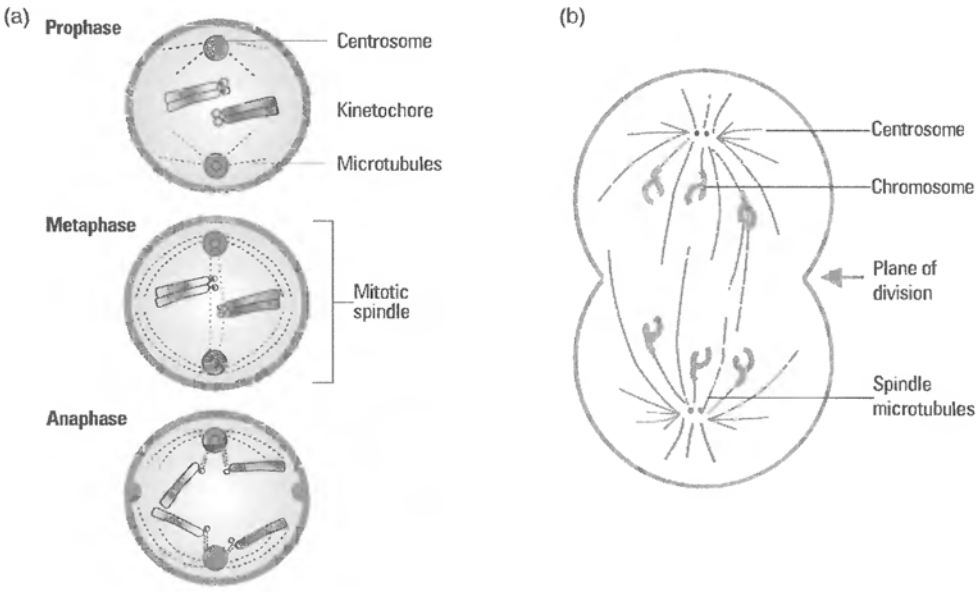


Fig. 12.14 (a) In prophase the spindle begins to form, and the microtubules are organized by the centrosomes, at either end of the cell. In metaphase, the chromosomes align themselves along the spindle axis and are attached to a region called the kinetochore and form the mitotic spindle. In anaphase, the sister chromatids are attached to microtubules, each pair is aligned, and the chromosomes move with the shortening microtubules to the opposite poles of the cell. (Reproduced with permission of Professor R. A. Weinberg and Nature from Fig. 1 of ref. 26.) (b) It may be seen that the DNA after having been replicated has condensed into two complete sets of chromosomes. One set is distributed to each of the two newly forming cells by the mitotic spindle, composed of microtubules, which are polymers of tubulin.

the other cell contents—the nuclear plasma and the membranes, the cellular organelles, the mitochondria, and the cytoplasm—must also be divided and properly reapportioned. This process is known as cytokinesis. The different steps of cytokinesis are visualized in Figs 12.13 and 12.14. Figure 12.13a is a diagram of a mitotic spindle, and the changes of the centrosomes during S phase, when the spindle poles in mitosis form, are shown schematically in Fig. 12.13b. In Fig. 12.14a the formation of the mitotic spindle during the transition from prophase–metaphase to anaphase is shown, and in Fig. 12.14b a cell in mitosis is shown.

Cell cleavage and chromosome alignment during metaphase are driven by disruptive forces, applied by the microtubules, which are balanced by opposing forces that hold the sister chromatids together.[27] The microtubules in the spindle receive positional information from the cell cortex and, on the basis of this information, direct the movement of the contractile proteins. Initially, the actomyosin-cytoskeleton generates tension over the whole cell surface, but late in mitosis the contractility becomes more localized, forming the cleavage furrow.

Control mechanisms,[28] shared by all eukaryotic cells, ensure that the separation of sister chromatids does not start before all pairs of sister chromatids have been properly

aligned on the mitotic spindle in anaphase. Anaphase terminates mitosis, and the proteolysis machine cleans up the workshop and removes the cyclins and other factors, preparing the cell for cytokinesis and cell division. Thus, the anaphase-promoting complex (APC or cyclosome) prevents further chromosome replication before all sister chromatids are separated.

Ran and microtubule organization

Ran, like most other monomeric GTPases, has many functions (Chapter 4). One of them is its role in microtubule organization in mitosis. A number of Ran-binding proteins have been isolated and found to link Ran to microtubules. Therefore, T. Nishimoto *et al.*[29] have speculated that Ran may coordinate the transfer of macromolecules in the cell with cell-cycle progression. One of these Ran-binding proteins is a nucleotide-exchange protein, RCC1, the others, Ran BP1 and Ran BMP, somehow help a Ran-GAP to convert Ran-GTP to Ran-GDP. In *Xenopus* egg extracts, depleted of RCC1, Ran-GTP accumulates and induces organization of the microtubular network which forms the mitotic spindle by drawing the chromosomes to the opposite poles of the dividing cell (Fig. 12.15).

After having discussed the cell-cycle transitions and the major steps in cytokinesis, the controls, (checkpoints) that watch over cell-cycle transitions and arrest the cycle when a defect has been spotted will be discussed. These cell-cycle checkpoints are of great importance. Malfunctions of these controls result in reduced fidelity of the timing of the cell cycle. They cause chromosomal instabilities (CINs) and result in uncontrolled cellular proliferation, and eventually cancer (see ref. 31 and Part 4).

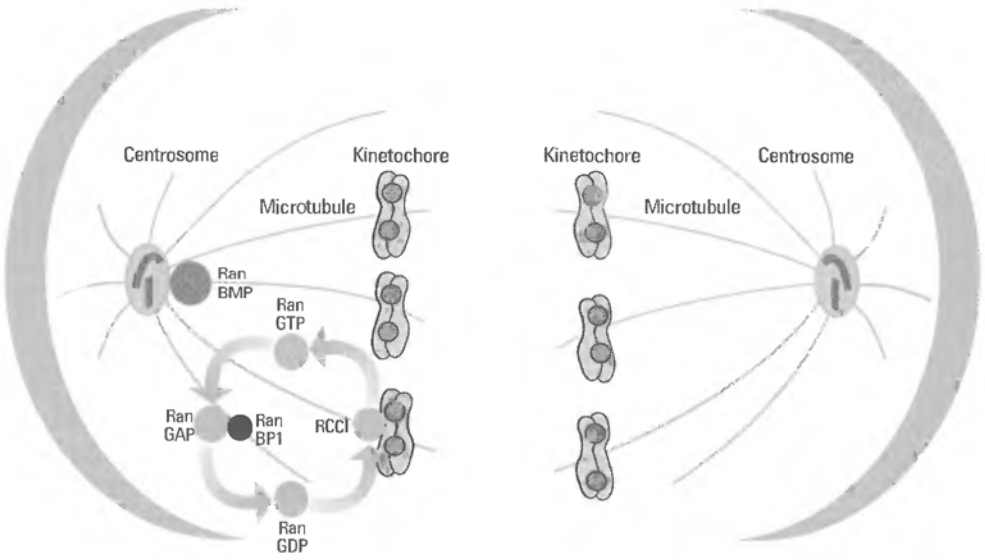


Fig. 12.15 Apparently, the Ran-GAP, RCC1, and Ran BP1 must all cooperate in order to propagate spindle formation. RCC1 converts Ran-GDP to Ran-GTP. A Ran-GAP, with the help of Ran BP1, reconverts Ran-GTP to the GDP form. Ran BMP accumulates on the centrosomes, from which the mitotic spindle starts to grow. (Drawn using information from ref. 30 with permission of Science.)

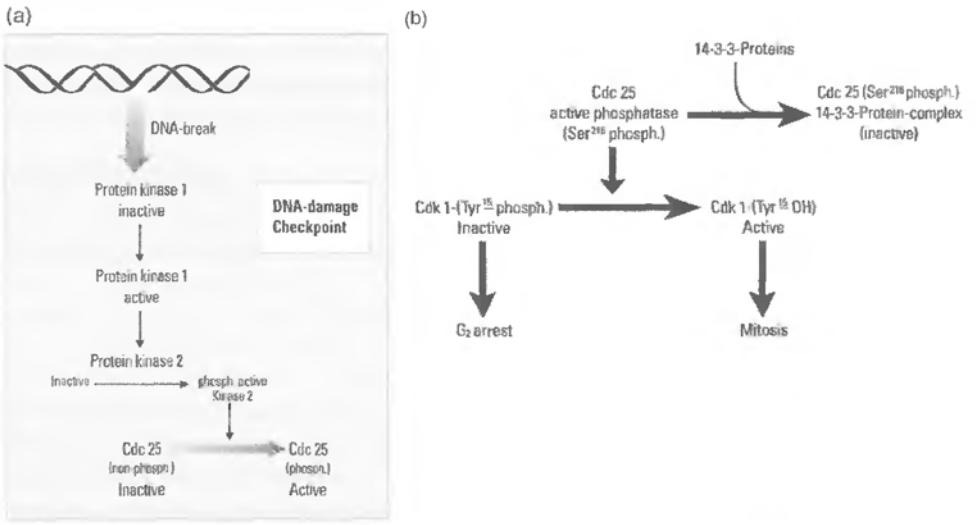


Fig. 12.16 (a) A single-strand DNA break is recognized by a protein kinase 1, which is activated.[35] The active protein kinase 1 phosphorylates and activates protein kinase 2 which, in turn, phosphorylates and activates the Cdc25 protein phosphatase. (b) The role of the Cdc25 phosphatase in the regulation of cyclin B/Cdk1. The activity of the Cdc25 phosphatase is itself regulated by phosphorylation through the cyclin B/Cdk1 complexes in yeast and mammals. Phosphorylation activates the phosphatase. The active, phosphorylated Cdc25 tyrosine phosphatase dephosphorylates Tyr15 of the inactive cyclin-dependent kinase 1, (Cdk1), and removes the block to mitosis. The active phosphatase also amplifies the cyclin B/Cdk1 kinase activity. The positive activation loop of the kinase and the phosphatase initiates mitosis. Through its phosphorylated serine site, the Cdc25 phosphatase can bind to 14-3-3 proteins, an interaction that inactivates and removes the phosphatase and forestalls its further action. (The checkpoint kinases are serine/threonine kinases, which were identified in fission yeasts and in human cells; 14-3-3 proteins are a family of linker proteins that interact with several regulatory proteins, including phosphatases.) A similar phosphorylation cascade is responsible for cell-cycle arrest in G_1. But this may not be the only way to arrest the cell cycle. Human cells, expressing a mutated, uncontrolled, non-phosphorylatable Cdk1 are not completely ineffective in arresting the cycle in G_2, suggesting the requirement for additional mechanisms. Sequestration of cyclin B could play a role. (for further information see refs 36,37). (Both (a) and (b) were drawn with permission of Science using information from ref. 5.)

Cell-cycle checkpoints

All eukaryotic cells have checkpoint regulation.[32] A deletion in a gene encoding a checkpoint regulator abolishes the strict follow-up of one cell-cycle event after the other. Kim Nasmyth,[33] therefore wants the term 'checkpoint' to be applied only to those controls that ensure the ordered sequence of cell-cycle events. Following his definition, checkpoints are in essence inhibitory signalling pathways that keep a cell-cycle transition in check until the preceding operation is completed.

As example of a checkpoint control, the response of the cell cycle in yeast to DNA damage will be described. Damaged DNA activates a checkpoint kinase (chk1) which is targeted to

the mitotic inducer, the Cdc25 phosphatase. The inducer generates a signal that induces the expression of genes coding for DNA-repair enzymes.[34] When the repair fails, the cell dies. DNA fragmentation signals apoptosis, linking cell-cycle checkpoint control to the apoptotic programme (Chapter 13). How DNA damage is sensed and how a phosphorylation cascade arrests the cell cycle in G_2 is shown in Fig. 12.16

Checkpoint controls are attractive targets for cancer therapy. Most agents that kill cancer cells do so by affecting checkpoint functions.[38] Therefore, the more we know about checkpoint pathways and specific checkpoint-controlling proteins, the better our chances to treat cancers more effectively.

Evolution of the cell cycle

The emergence of the cell cycle was probably linked to the increase in the genome in the course of evolution. The first, primitive eukaryotes may have carried out DNA replication and mitosis, S and M phases, simultaneously rather than one after the other, just like bacteria. But as genomes became more complex and multicellular organisms developed, cells had to couple DNA replication and mitosis to cell division. The next step in the evolutionary progression may have been the emergence of checkpoint mechanisms to stop the previous stage until transition to a new stage was completed.

Cyclins and Cdks are conserved among all multicellular eukaryotes. Regulation of transcription of genes encoding components of the cell-cycle machine by cyclin–Cdks, in cooperation with the E2F transcription factors, are, in principle, the same in vertebrates and invertebrates. Furthermore, at least one of the two families of Cdk inhibitors in vertebrates is also present in *Drosophila*.

How the enormous proliferation and formation of up to 10^{13} cells is coordinated with the morphogenetic programme in multicellular vertebrates is a central question in developmental biology. In mammals, patterning and differentiation are coordinated with cell duplication. During early stages in development in many species the cell cycle runs at maximum speed, subdividing the fertilized egg into more and smaller cells. As development progresses, controls are activated coordinating the cell cycle with morphogenesis. Eventually proliferation is halted altogether to allow for terminal cell differentiation.

The study of the regulation of cyclin-dependent kinases (Cdks) in *Drosophila* has provided insights into how the developmental programme of the fly is controlled at the molecular level. The first phase in the development of the *Drosophila* embryo is unique and not found in mammals.[39] Insects (and plants), in contrast to mammals, pass through *endoreduplication* cycles. The endoreduplication cycle in *Drosophila* is set up by the 'fizzy-related' (*fzr*) gene.[40] Endoreduplication is a sensible adaptation to the situation in *Drosophila* where the first steps in embryonic development take place in the egg where the genome is enlarged, first by formation of new nuclei by a series of rapid nuclear divisions, and then followed by formation of the first cells. Development in these large eggs takes place in a space which has about the same size in the beginning as at the end. Only at later stages does *Drosophila*, like mammals, adjust cell growth and cell division. This phase of *Drosophila* development occurs in 'imaginal discs', where immature 'imaginal' cells assemble, of which growth and proliferation are controlled by a cycle that is linked to patterning processes. It is controlled by signalling molecules, quite like in mammals. These signals appear to be nearly as conserved in vertebrates and invertebrates as the cell-cycle control apparatus itself, suggesting that the mechanisms that coordinate growth,

cell proliferation, and patterning in developing tissues have ancient origins. Eventually, proliferation of cells is halted altogether, to allow for terminal differentiation under the control of homeotic genes (for further reading see ref. 41).

Conclusions

The regulatory control that matters most for a cell is the control of the cell cycle. The main cell-cycle checkpoints have been identified and the principles of cyclin–Cdk regulation are understood. The cycle is controlled by phosphorylation dephosphorylation, positive and negative regulatory factors, and time-controlled proteolysis. So we are beginning to understand how the wheels of the clockwork are turning, and are getting an idea of what kind of signals tell the cell when to divide. Although we do have a considerable amount of information about the parts of the clock and how they function, the controllers of the machinery are just now being identified. These regulators are probably not much different from other regulators in cellular signalling, such as monomeric GTP-binding proteins, with on- and off-switch functions, kinases, phosphatases, and so forth.

References

1. Resources on the World Wide Web:

   An On-Line Biology Book, by M. J. Farabee, includes a chapter on Cell Division: Binary Fission and Mitosis. This chapter describes the phases of the cell cycle and provides background information on cell-cycle regulation. Several diagrams are included and links to related Web resources are available.

   The Dictionary of Cell Biology (Academic Press, London, 1995), defines some of the terms used in this article.

   Cell Division and the Cell Cycle , developed at the University of Alberta, describes the cell cycle and includes a glossary of terms.

   Mitosis, developed at McGill University, provides high-quality animations of mitosis and cell division.

   DNA Replication illustrates the biosynthesis of DNA.

   Virtual Mitosis, developed by the Department of Biological Sciences at the University of Cincinnati, presents photomicrographs of the mitotic process and defines terms. It also offers a link to Virtual Meiosis, in which the phases of meiosis are illustrated.

   Cell Cycle is a page in the Cytogenetic Terms Index. Cell cycle is defined with links to descriptions of cell division and the phases of mitosis.

   Diagrams of the phases of meiosis and mitosis are offered in the About Biotech Graphics Gallery, a component of Access Excellence, a national educational programme sponsored by Genentech, Inc.

   The MIT Biology Hypertextbook, developed by the Experimental Study Group at the Massachusetts Institute of Technology, provides background information on the biology of cells and nucleic acids. Mitosis, a section within Cell Biology, provides descriptions and diagrams of the phases of the cell cycle.

2. K. Nasmyth. Viewpoint: Putting the cell cycle in order. *Science*, **274**, 1643–1645, 1996.
3. B. Alberts, *et al. The Molecular Biology of the Cell.* Garland Publishing, New York , 1994.
4. P. Sicinski, J. L. Donaher, Y. Geng, S. B. Parker, H. Gardner, M. Y. Park, *et al.* Cyclin D2 is an FSH-responsive gene involved in gonadal cell proliferation and oncogenesis. *Nature*, **384**, 470–474, 1996.
5. T. Weinert. A DNA damage checkpoint meets the cell cycle engine. *Science*, **277**, 1450–1451, 1997.
6. E. Lees. Cyclin dependent kinase regulation. *Curr Opin Cell Biol*, **7** (6), 773–780, 1995.
7. J. Pines. Cyclins and cyclin-dependent kinases: a biochemical view. *Biochem J*, **308**, 697–711, 1995.
   Also: J. Pines. Conformational change. *Nature*, **376**, 294, 1995.

8. G. Krauss. *Biochemie der Regulation und Signaltransduktion*, Wiley-VCH, Abb. 14.9, p. 405, 1997.
9. A. A. Russo, P. D. Jeffrey, A. K. Patten, J. Massagué, and N. P. Pavletich. Crystal structure of the $p27^{kip1}$ cyclin-dependent-kinase inhibitor bound to the cyclin A–Cdk1 complex. *Nature*, **382**, 325–331, 1996.
   Also: D. O. Morgan. Under arrest at atomic resolution, News and views. *Nature*, **382**, 295–296, 1996.
10. T. Jacks and R. A. Weinberg. The expanding role of cell cycle regulators. *Science*, **280**, 1035–1036, 1998.
    Also F. Di Cunto, G. Topley, E. Calautti, J. L. Hsiao Ong, P. K. Seth, and G. P. Dotto. Inhibitory function of p21Cip1/WAF1 in differentiation of primary mouse keratinocytes independent of cell cycle control. *Science*, **280**, (5366) 1069–1072, 1998.
11. N. Dyson. pRB, p107, and the regulation of the E2F transcription factor. *J Cell Sci (Suppl.)*, **18**, 81–87, 1994.
12. H. Zhu, P. Kavsak, S. Abdollah, J. L. Wrana, and G. H. Thomsen. A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. *Nature*, **400**, 687–693, 1999.
13. W. Seufert, B. Futcher, and S. Jentsch. Role of ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature*, **373**, 78–81, 1995.
14. S. Jentsch and S. Schlenker. Selective protein degradation: A journey's end within the proteasome. *Cell*, **82**, 881–884, 1995.
15. A. Herschko and A. Ciechanover. The ubiquitin system. *Annu Rev Biochem*, **76**, 425–479, 1998.
16. G. C. Turner, F. Du, and A. Varshavsky. Peptides accelerate their uptake by activating a ubiquitin-dependent pathway. *Nature*, **405**, 579–583, 2000.
17. W. Baumeister, J. Walz, F. Zühl, and E. Seemüller. The proteasome paradigm of a self-compartimentalizing protease. *Cell*, **92**, 367–380, 1998.
18. J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, and R. Huber. Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution *Science*, **268**, (5210) 533–539, 1995.
    Also: E. Seemuller, A. Lupas, D. Stock, J. Löwe, R. Huber, and W. Baumeister. Proteasome from Thermoplasma acidophilum: a threonine protease. *Science*, **268**, (5210) 579–582, 1995.
19. C. W. Greider, K. Collins, and C. Autexier. In *DNA replication in eukaryotic cells*, (ed. M. L. De Pamphilis), pp. 619–638. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1996.
20. T. S. Lendvay, D. K. Morris, J. Sah, B. Balasubramanian, and V. Lundblad. Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. *Genetics*, **144** (4), 1399–1412, 1996.
    Also: J. Lingner, T. R. Hughes, A. Shevchenko, M. Mann, V. Lundblad and T. R. Cech. Reverse transcriptase motifs in the catalytic subunit of telomerase [see comments]. *Science*, **276** (5312), 561–567, 1997.
21. A. G. Bodnar, M. Ouellette, M. Frolkis, S. E. Holt, C. P. Chiu, G. B. Morin, *et al.* Extension of life-span by introduction of telomerase into normal human cells [see comments]. *Science*, **279** (5349), 349–352, 1998.
    Also: H. Vaziri and S. Benchimol. Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span. *Curr Biol*, **8** (5), 279–282, 1998.
22. T. Kiyono, S. A. Foster, J. I. Koop, J. K. McDougall, D. A. Galloway, and A. J. Klingelhutz. Both Rb/p16INK4a inactivation and telomerase activity are required to immortalize human epithelial cells [see comments]. *Nature*, **396** (6706), 84–88, 1998.
23. R. A. Weinberg. Telomeres: Bumps on the road to immortality. *Nature*, **396**, 23–24, 1998.
24. A. J. Klingelhutz, S. A. Foster, and J. K. McDougall. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature*, **380** (6569), 79–82, 1996.
25. M. Serrano, A. W. Lin, M. E. McCurrach, D. Beach, and S. W. Lowe. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*, **88** (5), 593–602, 1997.
26. T. L. Orr-Weaver and R. A. Weinberg. A checkpoint on the road to cancer. *Nature*, **392**, 223–224, 1998.
27. J. A. Theriot and L. L. Satterwhite. New wrinkles in cytokinesis, News and Views. *Nature*, **385**, 388–389, 1997.
    Also: K. Burton and L. Taylor. Traction forces of cytokinesis measured with optically modified elastic substrata. *Nature*, **385**, 450–454, 1997.
28. B. Stillman. Cell cycle control of DNA replication. *Science*, **274**, 1659–1664, 1996.
    Also: R. W. King, R. J. Deshaies, J.-M. Peters, and M. W. Kirschner. How proteolysis drives the cell cycle. *Science*, **274**, 1653–1659, 1996.
29. T. Ohba, M. Nakamura, H. Nishitani, and T. Nishimoto. Self-organization of microtubule asters induced by *Xenopus* egg extracts by GTP-bound Ran. *Science*, **284**, 1356–1358, 1999.
    Also: A. Wilde and Y. Zheng. Stimulation of microtubule aster formation and spindle assembly by the small GTPase Ran. *Science*, **284**, 1359–1362, 1999.

30. E. Pennisi. Nuclear transport protein does double duty in mitosis. *Science*, **284**, 1260–1261.

31. C. Lengauer, K. W. Kinzler, and B. Vogelstein. Genetic instabilities in human cancers. *Nature*, **396**, 643–649, 1998.

32. S. J. Elledge. Cell cycle checkpoints: Preventing an identity crisis. *Science*, **274**, 1664–1672, 1996.

33. K. Nasmyth. Control of the yeast cell cycle by the Cdc28 protein kinase. *Curr Opin Cell Biol*, **5**, (2) 166–179, 1993.

34. B. Furnari, N. Rhind, and P. Russell. Cdc25 mitotic inducer targeted by chk1 DNA damage checkpoint kinase [see comments]. *Science*, **277** (5331), 1495–1497, 1997.

    Also: C. Y. Peng, P. R. Graves, R. S. Thoma, Z. Wu, A. S. Shaw, and H. Piwnica-Worms. Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216 [see comments]. *Science*, **277** (5331), 1501–1505, 1997.

35. N. C. Walworth and R. Bernards. Rad-dependent response of the chk1-encoded protein kinase at the DNA damage checkpoint [see comments]. *Science*, **271** (5247), 353–356, 1996.

    Also: M. J. O'Connell, J. M. Raleigh, H. M. Verkade, and P. Nurse. Chk1 is a wee1 kinase in the G2 DNA damage checkpoint inhibiting cdc2 by Y15 phosphorylation. *EMBO J.*, **16**, (3), 545–549, 1997.

36. F. al Khodairy and A. M. Carr. DNA repair mutants defining G2 checkpoint pathways in *Schizosaccharomyces pombe*. *EMBO J*, **11** (4), 1343–1350, 1992.

    Also: J. C. Ford, F. al Khodairy, E. Fotou, K. S. Sheldrick, D. J. Griffiths, and A. M. Carr. 14-3-3 protein homologs required for the DNA damage checkpoint in fission yeast. *Science*, **265** (5171), 533–535, 1994.

    Also: N. Rhind, B. Furnari, and P. Russell. Cdc2 tyrosine phosphorylation is required for the DNA damage checkpoint in fission yeast. *Genes Devel*, **11** (4), 504–511, 1997.

    Also: T. H. Lee and M. W. Kirschner. An inhibitor of p34cdc2/cyclin B that regulates the G2/M transition in *Xenopus* extracts. *Proc Natl Acad Sci, USA*, **93** (1), 352–356, 1996.

37. A. Amon, M. Tyers, B. Futcher, and K. Nasmyth. Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins. *Cell*, **74** (6), 993–1007, 1993.

    Also: L. Hartwell, T. Weinert, L. Kadyk, and B. Garvik. Cell cycle checkpoints, genomic integrity, and cancer. *Cold Spring Harb Symp Quant Biol*, **59**, 259–263, 1994.

38. L. H. Hartwell and M. B. Kastan. Cell cycle control and cancer. *Science*, **266**, 5192, 1821–1828, 1994.

39. C. F. Lehner and M. E. Lane. Cell cycle regulators in *Drosophila*: downstream and part of developmental decisions. *J Cell Sci*, **110**, 523–528, 1997.

    Also: B. A. Edgar and C. F. Lehner. Developmental control of cell cycle regulators: a fly's perspective. *Science*, **274** (5293), 1646–1652, 1996.

40. S. J. Sigrist and C. F. Lehner. *Drosophila* fizzy-related down-regulates mitotic cyclins and is required for cell proliferation arrest and entry into endocycles. *Cell*, **90** (4), 671–681, 1997.

41. Cell Cycle Genes is a chapter of the Interactive Fly, an Internet guide to *Drosophila* genes and their role in development. The cell cycle in *Drosophila* is described and the role of the cyclins in regulating the cell cycle is outlined. A biological overview and a discussion of evolutionary homologues refer to cell-cycle regulation in other organisms.

13

Regulation of cell death

The cell cycle is complemented by a programme controlling cell death. This programme orders cells to destroy themselves when they are no longer needed or when they are damaged and dysfunctional. Cell death is characterized by blebbing and vacuole formation, chromatin condensation, and DNA fragmentation. To survive, cells from multicellular organisms (metazoans) must keep the suicide programme under control and prevent its activation. The signalling pathways receiving and transmitting 'death signals' involve cell–cell interactions. The first appearance of programmed cell death (PCD) may go back to the emergence of multicellular, metazoan organisms. An example is stalk formation in the slime mould, *Dictyostelium discoideum*, where kind of a CD programme directs the assembly of aggregated cells, slugs and stalk formation.[1]

Because apoptosis decides the life or death of a cell, any malfunction of this regulatory programme can do harm. Unregulated apoptosis of certain neurons in the brain contributes to loss of function in cerebral ischaemia, stroke, Alzheimer's, and Parkinson's disease.[2] The fact that many key regulators in PCD are potential oncogenes links this pathway to cancer. Rapidly dividing cells with damaged DNA, which fail to activate their apoptotic programme, become cancer cells.

Signalling pathways in apoptosis

We begin our discourse on cell death with a summary of the tools and the components of the machinery that carry out the CD programme. We shall see that, despite a bewildering variety of new names, regulation of the apoptosis pathways has basic features in common with other cellular signalling pathways.

Factors promoting cell death

The apoptotic machinery is controlled, like the cell cycle, by internal and external signals. External signals are received by receptors. The receptors that trigger apoptosis are called 'death' receptors. Although they may participate in other cellular signal transduction pathways, when they interact with 'death' ligands they activate 'death' effectors. The effectors are the caspases, a family of cysteine proteases which eventually destroy the cell.

Death receptors

Death receptors are members of the TNF-receptor family; they all share similar, cysteine-rich extracellular domains. Moreover, death receptors have a homologous cytoplasmic sequence, the 'death domain', through which they link up with the apoptotic machinery. Linkers have matching 'death effector domains' (DEDs). Linkers transmit apoptotic signals to the suicide machinery. Prototype 'death' receptors are CD95 (also named Fas or Apo1) and TNFR-I.[3,4] The death domain of TNFR-I is TRADD (the tumour necrosis factor receptor-associated death domain), and the corresponding CD95 /Fas 'death' domain is FADD. The ligands for death receptors are TNF-related homotrimeric molecules. On the basis of the crystal structure of a complex of a death ligand, (lymphotoxin α), with TNFR-I, one can conclude by analogy that each CD95-ligand trimer, (CD95L), binds to three CD95 receptor molecules. Since the CD95L is a membrane-attached protein, expressed for example, on activated T lymphocytes, interaction with the similarly membrane-bound receptor couples cells together (more about the role of cell–cell interactions in the Chapter 14).

Like other receptors, 'death' receptors are activated by binding ligands. The activated death receptor–ligand complex is linked with its 'death domain' (DD) to a homologous DED (death effector region), repeated in tandem, and recruits a zymogen form of a protease, pro-caspase 8 (Fig. 13.1). But the linker is, like other linkers, versatile. Besides coupling the receptor–ligand complex to pro-caspase 8, it can interact with other cytoplasmic regulator molecules, which may re-orient the receptor, disengaging it from the apoptotic pathway and directing it to the JNK (Jun NH_2-terminal kinase) pathway, which transmits stress signals. Upon activation, the pro-caspase oligomerizes and becomes active through autocatalytic self-cleavage, like other zymogens, forming an initiator caspase, which then initiates a proteolytic cascade, eventually leading to apoptosis. Thus, caspase activation commits the cell to die.

TNFR-I signalling is as diverse as CD95 receptor signalling. The ligand, TNF (the tumour necrosis factor) is produced in response to infection by activated macrophages and T cells. TNF binds to the TNFR-I and activates the transcription factors NF-κB and Jun/Fos, which turn on many proliferation-promoting genes, including genes required for the response of the body to inflammation. Expression of these genes suppresses apoptosis, because these genes express linkers, the TNF-receptor-associated factors (TRAFs), which direct the TNF signal to other pathways than the apoptotic pathway. Therefore only when *NF-κB*[6] and c-*Jun*/c-*Fos* genes are silenced and these linkers are not produced, does apoptosis occur. In this case, the TNF receptor-I/TNF complex unmasks its TRADD (TNFR-associated death domain) and activates pro-caspase 8.[7] Thus, the choice between apoptosis and other non-apoptotic, cellular responses is made by linkers, the TRAFs (Fig. 13.2).

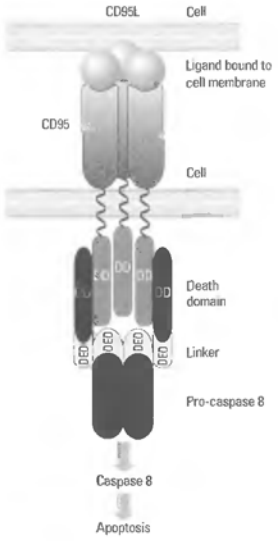


Fig. 13.1 The CD95-liganded death receptor with the intracellular death domains (DDs) binds to the death effector domain (DED) which links the receptor–ligand complex to the pro-caspase 8. Cell denotes the cell boundary. (Based on the information in Fig. 1 of ref. 5 and reproduced with permission of the authors and Science.)

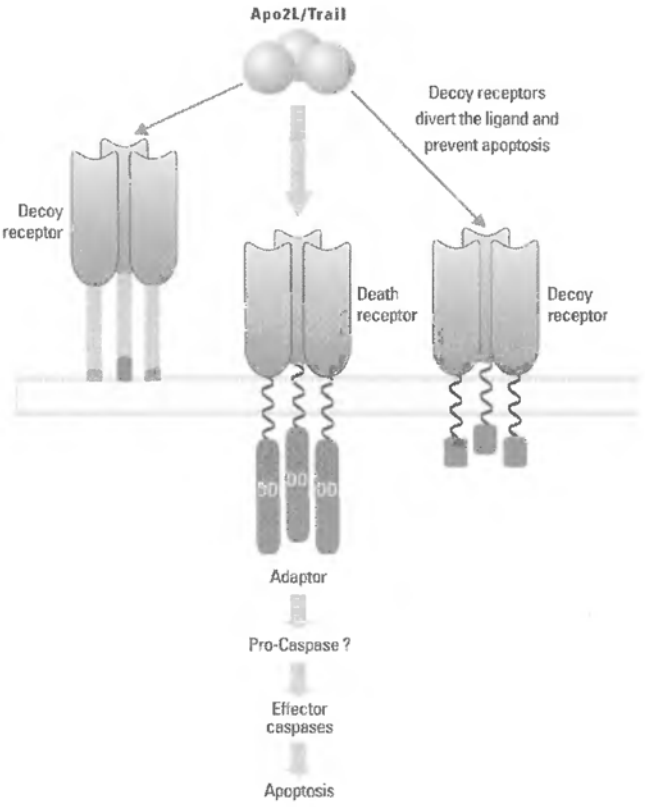


Fig. 13.2 A TNF-related factor, which has the greatest similarity with the CD95 ligand, is the ligand Apo2L or Trail. Apo2L binds to the death receptors (DRs) DR4 and DR5.[8–10] Induction of apoptosis by Apo2L also involves an adaptor and caspase activation. The decoy receptors 1 and 2 are glycosylphosphatidylinositol (GPI)-anchored membrane proteins that have a ligand-binding domain, just like DR4 or DR5, but lack a signalling cytoplasmic tail. They cannot signal, but scavenge and remove the apoptotic ligand, Apo2L, preventing its productive interaction with the death receptors.[11–13] (Based on the information in Fig. 3 of ref. 5 and reproduced with permission of the authors and Science.)

The X-ray crystal structure of the complex of the extracellular domain of the human 55 kDa tumour necrosis factor (TNF) receptor with human TNF-β has been determined at 2.85 Å resolution. It provides a model for TNF receptor activation. The complex contains three receptor molecules bound symmetrically to one TNF-β trimer. The TNF-β subunits form a groove into which the receptor is inserted. The structure of the receptor–ligand complex also determines its orientation with respect to the cell membrane. The TNF-receptor structure is likely to be representative of the TNF-receptor family as a whole, including the NGF receptor (see Chapter 1).

Figure 13.3 shows the structure of the soluble, extracellular domain of the type I TNF receptor, without the ligand, TNF-α or -β. The structure is of interest because, contrary to expectations (see Chapter 2), the unliganded receptor domains are already associated.

A class of receptors similar to TNFR-I are the death receptors (DRs) 3, 4, and 5. Activation of DRs is prevented by the expression of so-called decoy receptors, which compete with the DRs for the same ligands. The genes encoding the death receptors DR4, DR5, and the decoy receptors (DcR1 and DcR2), map together on human chromosome 8p21–22, suggesting that they are derived from an ancestral gene.

Cell-death proteases

Proteolysis plays a central role in apoptosis, as it does in the control of the cell cycle (Chapter 12). Apoptotic signals activate a cascade of proteolytic enzymes which degrade

proteins essential for the maintenance of cellular integrity. The result is the destruction of the cell. The instruments of the apoptotic machinery are the 'caspases'. Caspases are cysteine metalloproteases that cleave proteins after an aspartic acid residue. About a dozen caspases have been identified up to now. The first caspase identified was the product of the cell-death gene, *ced-3*, in *Caenorhabditis elegans*. Its mammalian homologue is the interleukin-1β-converting enzyme (ICE), or caspase 1.[16] Caspases are activated by proteolysis of precursor 'zymogens', the pro-caspases. Autocatalysis initiates a caspase cascade, whereby subsequent caspases are activated by specific cleavage by the upstream caspase, finally forming the effector caspases. One distinguishes between initiator caspases, 8 and 9, which start the caspase cascade by activation of the downstream caspases and effector caspases, 3, 6, and 7, which destroy cellular proteins in the course of apoptosis. The caspase cascade resembles the blood-clotting cascade (see textbooks of biochemistry). Figure 13.4 summarizes the structural properties of caspases.

Caspases are among the most specific proteases known.[17] Thanks to the crystal structures of active caspases (caspase 1 and 3), we have some idea how these enzymes are activated. The structure of the interleukin-1β–converting enzyme (ICE), which processes an inactive interleukin precursor to interleukin-1β, is shown in Plate 26.[18]

Four or more amino acids, NH_2-terminal to the aspartate after which the polypeptide chain is cleaved, make up the recognition site in pro-caspases. The sequence of the recognition-site tetrapeptide differs among different pro-caspases. The only conserved residue is aspartate.

Proteolysis of enzymes (and non-enzymic regulatory proteins) by caspases often results in deregulation of these proteins, when limited proteolysis disengages regulatory from effector functions. A case in question is the deregulation of enzymes and proteins involved in DNA repair and replication. Among the targets of caspases is a DNAse inhibitor, I^{CAD} (caspase-activated deoxyribonuclease inhibitor), which is proteolytically activated. Other targets of caspases are anti-apoptotic proteins, such as the Bcl-2 proteins (see below), which protect cells against apoptosis (see ref. 19). Finally, proteins essential for assembly and integrity of the cytoskeleton are substrates of effector caspases. As more substrates of the proteolysis machinery are identified, it should eventually be possible to assign to each effector caspase a role in one of the many apoptotic events leading to cell death, such as DNA fragmentation, chro-

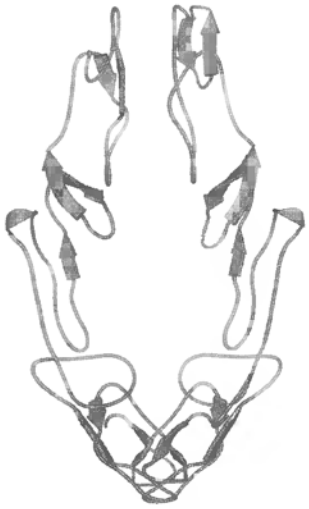


Fig. 13.3 In the crystal structure at 2.25 Å the soluble, extracellular domains of the type I TNF receptor, without the ligand (TNF-α or -β), are associated, antiparallel to each other and in a way that overlaps the TNF-ligand-binding site at the bottom of the picture. The monomers are related to each other by a twofold axis of symmetry. The receptor domains are elongated, and arranged in a nearly linear order. If intact TNF receptors in the cell were to associate in that manner, their cytoplasmic domains would be more than 100 Å apart. The consequences would be obvious: dimerization would prevent signalling without the ligand. (This ribbon diagram was reproduced with permission of Dr S. Sprang from data in ref. 14 and J. Biol. Chem. and J. Mol. Biol. and the corresponding data available in protein databanks.) Another form of parallel dimers (not shown) was also observed, where the dimer interface is well separated ftrom the binding site for the ligand, TNF. These receptor dimers could cluster on binding TNF. The formation of receptor dimers buries large areas of the protein surface. Thus the dimerization interfaces are large.

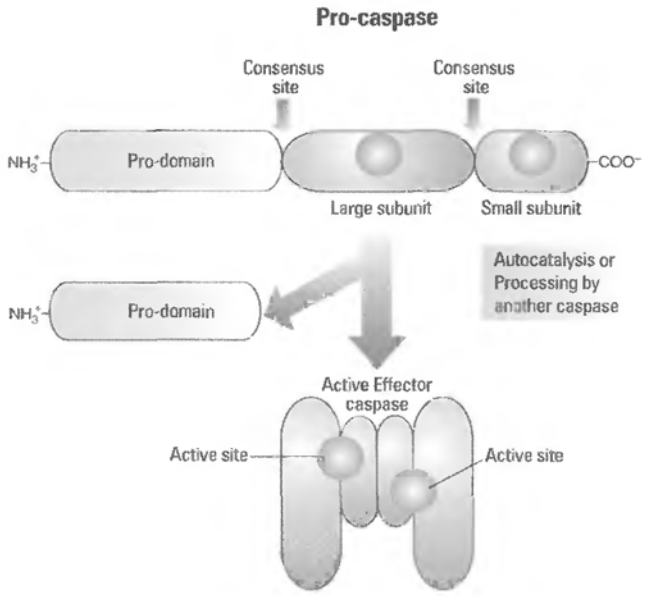


Fig. 13.4 Common structural properties of caspases. Pro-caspases have a smaller (10 kDa) subunit, a larger (20 kDa) subunit and an amino-terminal pro-domain. Cleavage of consensus sites in the pro-enzymes separates the two subunits from the remaining NH_2-terminal Pro-domain, which controls processing. Processing is either by autocatalysis or by another caspase with compatible specificity. After processing is completed, the two subunits, the large and the small one, combine, and associate to a tetramer with two catalytic sites, each functioning independently.

matin condensation, cell shrinkage, membrane blebbing and disassembly, and conversion into membrane-enclosed vesicles (apoptotic bodies), which are then engulfed and removed by phagocytic cells.

Activation of caspases is irreversible, because it involves peptide-bond cleavage. This is unlike most other protein modifications which play a role in cellular regulation. Therefore, proteolysis is involved only in unidirectional, irreversible processes, such as the cell cycle and cell death. But, the possibilities to regulate irreversible reactions are rather limited. In a cascade of proteolytic reactions, the first enzyme in the chain is the most likely point of control. This is the initiator caspase. The signals controlling initiator caspases vary, there are both external and internal signals (Fig. 13.5). Several mechanisms control the irreversible activation of caspases, including phosphorylation, separation, and compartmentalization of pro-caspases and positive and negative regulators.

Another way to regulate caspases is by interaction with inhibitors. For example, fibroblasts that overexpress the growth-promoting transcription factor c-Myc die, because the caspase inhibitors (IAPs), inhibitors of apoptosis, are inactivated by phosphorylation. In Fig. 13.6 is shown how phosphorylation of a caspase inhibitor controls caspase activity and apoptosis. When cells are driven to proliferate by an overactive transcriptional activator, such as c-Myc, an unregulated active caspase, no longer restrained by the inhibitor, can overcome the effect of Bcl-2 and the cell will eventually die. Only when Bcl-2 is overexpressed simultaneously with c-Myc, can cell death be prevented. Thus, Bcl-2 keeps the c-Myc signals in check.[22,23]

Intracellular apoptotic signals

Mitochondria play a critical role in the control of cell death, because they provide a major intracellular apoptotic signal, cytochrome *c*.[24] There are several, probably interrelated, ways in which mitochondria shed cytochrome *c*.[25] Early in apoptotic cells, mitochondria undergo a permeability change that disrupts electron transport. This situation is charac-

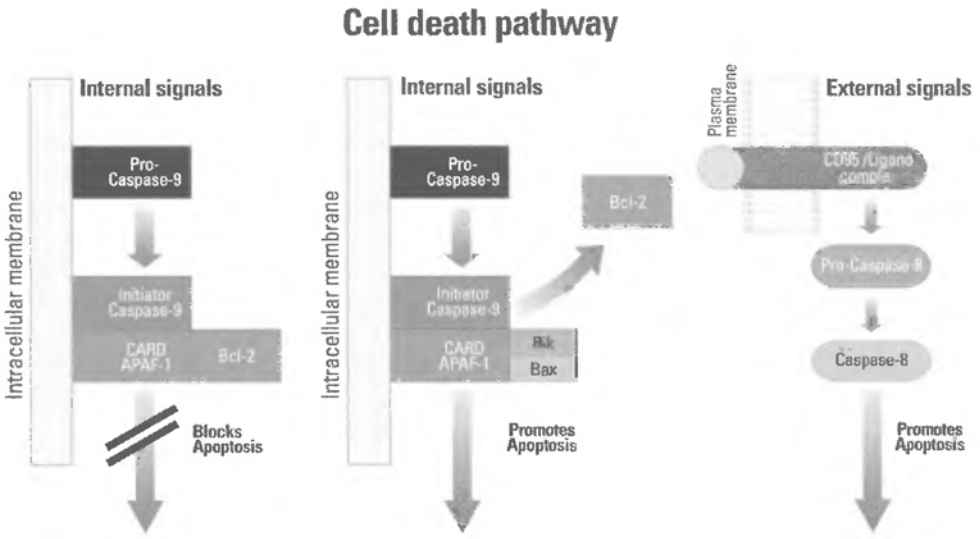


Fig. 13.5 External signals activate pro-caspase 8. Pro-caspase 8 is activated by the CD95/CD95L death receptor–ligand complex, internal signals activate another pro-caspase, pro-caspase 9. Activation of pro-caspase 9 requires the cofactor, APAF-1 (apoptosis-activating factor), which combines with the caspase through its CARD (caspase-recruiting domain).[21] There are two major classes of proteins that determine whether a cell turns on the apoptosis pathway, leading eventually to cell death. A class of proteins that prevents apoptosis are the Bcl-proteins. A prototype of a regulator that prevents apoptosis is Bcl-2, Bcl-X_L, whereas the Bax proteins, Bax, Bad, Bak, Bik , and others promote apoptosis. In mammalian cells, the Bcl-2 proteins check the various cytotoxic signals transmitted by the pro-apoptotic Bax/Bik proteins.

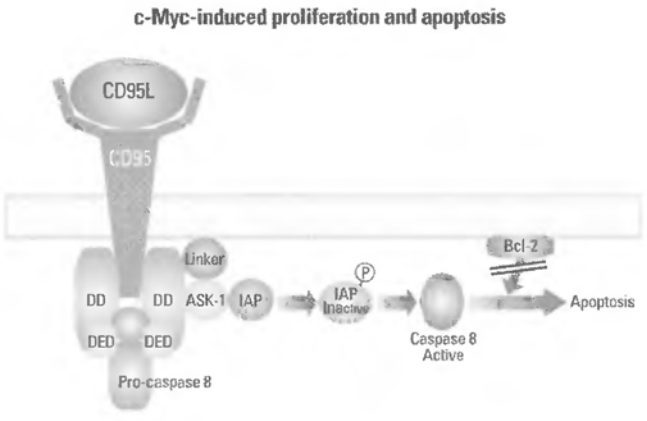


Fig. 13.6 Control of caspases by inhibitors in c-Myc-induced proliferation. Caspase inhibitors (IAPs) are inactivated by phosphorylation by ASK-1. ASK-1 is an apoptotic signal-regulated protein kinase, which binds to the CD95L/CD95 ligand–receptor complex with the help of a special linker protein. Interaction with the receptor–ligand complex activates the kinase and shuts off the caspase inhibitors (IAPs), by phosphorylation. The caspase is activated and blocks Bcl-2, (which normally prevents apoptosis), sending the cell towards apoptosis.

terized by a breakdown of the membrane potential. As a consequence, mitochondrial proteins are shed including cytochrome *c*, which combines with the apoptosis-activating factor, APAF-1 and pro-caspase 9, forming a proteolytically active 'apoptosome' (Fig. 13.7). Novel regulators of apoptosis, originating from mitochondria are being discovered. One of them is Smac, mitochondrial activator of apoptosis.[26] Smac interacts with the BIR3 domain of IAP and blocks its inhibitory action on caspases. (Bir is an abbreviation of baculoviral IAP repeat. The abbreviations of the many other pro- and anti-apoptotic factors are not explained. They usually are derived from the viral oncogenes from which they originate). A high resolution crystal structure of a complex of Smac with the BIR3 domain of IAP could help to find small peptides that inactivate IAP in cancer cells and induce apoptosis. Another possibility is that the dysfunction of electron transport and the subsequent decline of cellular energy (ATP) eventually turns on the apoptotic programme. (The mitochondrial ATP/ADP carrier has also been implicated.) The details of the mitochondrial change are still a matter of discussion.

Another consequential, internal apoptotic signal is damaged DNA (Fig. 13.8 and Chapter 12). DNA damage is sensed by a signalling pathway which decides between DNA repair and cell survival, and apoptosis.

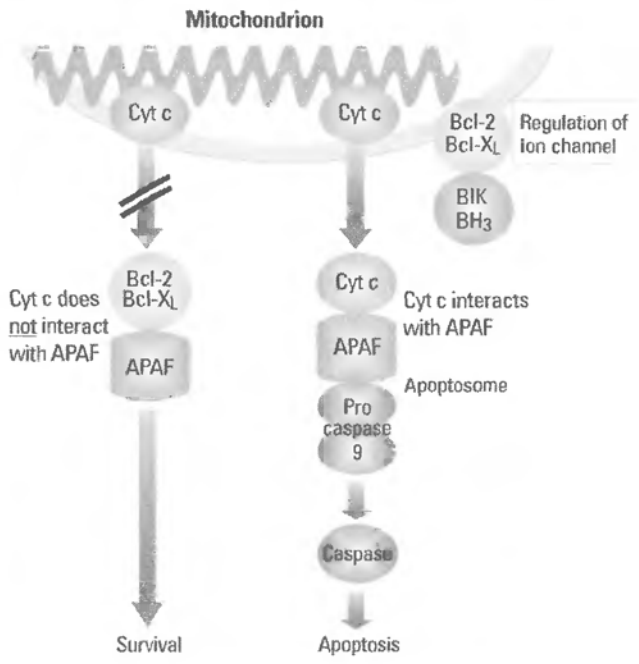


Fig. 13.7 Cytochrome *c*, released from mitochondria, turns on apoptosis by participating in the formation of the apoptosome and activating the caspase cascade. Release of cytochrome c from mitochondria is inhibited by the survival factor Bcl-2/Bcl-X_L. Bcl-2/Bcl-X_L combine with APAF, preventing its interaction with cytochrome *c*. On the other hand, binding of BH_3/BIK to Bcl-2/Bcl-X_L blocks the anti-apoptotic action of Bcl-2/Bcl-X_L (Bcl-X_L, is an anti-apoptotic member of the Bcl-2 family of survival proteins). Whereas, the 'good proteins', Bcl-2/Bcl-X_L, bind to APAF and prevent it from activating pro-caspase 9, a death signal provokes interaction of BH_3/BIK with Bcl-X_L, thus preventing the interaction of Bcl-X_L with APAF, freeing APAF which now binds to cytochrome c, forming an apoptosome and promoting apoptosis. In the presence of ATP, and when cytochrome *c* is released from mitochondria, APAF becomes active and binds to pro-caspase 9, promoting its dimerization and autocatalytic activation. Caspase 9 then initiates activation of the downstream effector caspases.

Cells with fragmented and damaged chromosomes express the *bax* gene, which starts the apoptotic programme, although details are not yet known. It is speculated that imperfect chromosome pairing in meiosis triggers apoptosis. An example is spermatogenesis, where three-quarters of the developing spermatocytes are destroyed. Male mice with inactivated *bax* genes carry an excess of damaged germ cells, which would normally have been eliminated by apoptosis.

Factors preventing cell death

The ratio of Bcl-2 and Bax in a cell determines its fate. When the Bcl-2 level is higher than the Bax level, apoptosis is suppressed and, alternatively, apoptosis is promoted when the Bcl-2 level is lower than the Bax level.

How, can the 'good' proteins prevent the activity of the 'bad' ones? Genetic and biochemical studies of *C. elegans* have provided clues to the pro-survival function of Bcl-2/ Bcl-X_L. Bcl-2/Bcl-X_L has two effects:

1. It prevents the action of APAF-1 and interferes with the formation of the apoptosome and the activation of pro-caspase 9.
2. Bcl-2/Bcl-X_L proteins are localized on the outer mitochondrial membrane. They are assumed to function as ion channels, maintaining the integrity of the mitochondrion and preventing shedding of cytochrome *c*. The structure of Bcl-X_L provides support for this idea (Plate 27). (Refs. 27–37 provide more information on the properties of pro- and anti-apoptotic proteins which control the fate of the cell).

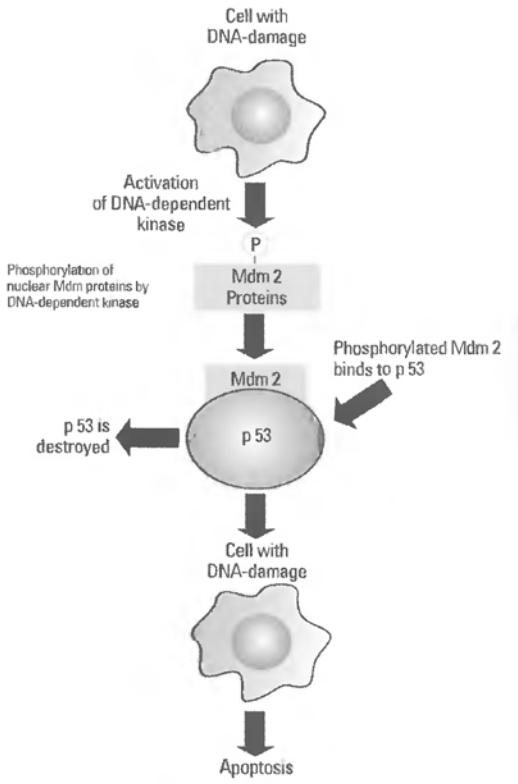


Fig. 13.8 When damaged DNA is sensed, the Mdm-2 proteins receive a signal from a DNA-dependent protein kinase. DNA-damage-induced phosphorylation of either Mdm-2 proteins or p53 promotes apoptosis.

The structure also gives a clue to an other, central function of the Bcl-2 family of proteins, namely their ability to form heterodimers with APAF and other regulatory proteins. There are at least 15 members of the Bcl-2 family and all possess at least one of four conserved motifs, known as Bcl-2 homology domains (BH-$_1$–BH_4).[38] These domains are essential for dimerization.[39] The fact that the pro-survival proteins of the Bcl-2 family and the pro-apoptotic Bax family proteins combine readily with each other, shows how structurally similar these functionally so different proteins are. This is also the reason why a survival protein, Bcl-X_L, can interact productively with a small peptide (16 amino acids long) of a pro-apoptotic, Bak protein (Plate 27). This gives hope for the development of small peptides to modulate the survival functions of their partners.

The interaction between Bcl-X_L and Bad is regulated by phosphorylation by the serine/threonine protein kinase Akt (or protein kinase B), a member of the PtdIns-3,4,5-P_3-regulated pathway. This scenario is summarized in Fig. 13.9. Akt

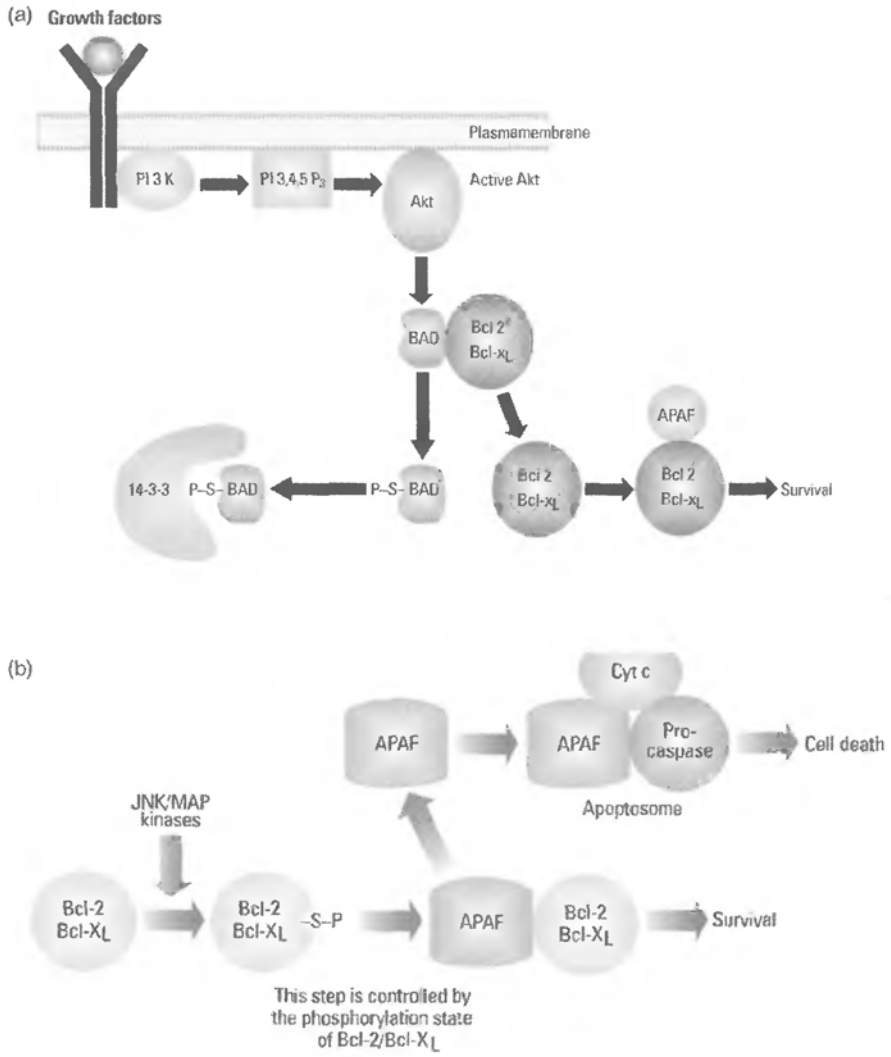


Fig. 13.9 (a) The protein kinase Akt is activated on the membrane by PtdIns-3,4,5-P_3, formed by the PI 3-kinase. Akt phosphorylates Bad on a serine, which causes its release from the complex with Bcl-2/Bcl-X_L.[40–41] The phosphorylated Bad is then sequestered by 14-3-3 proteins[42] and neutralized, whereas Bcl-2/Bcl-X_L is now free to bind to and remove APAF and suppress the caspase cascade, preventing cell death. (b) Bcl-2/Bcl-X_L proteins are controlled by phosphorylation through JNK and MAP kinases, each targeted to a different serine. This brings Bcl-2/Bcl-X_L under the control of JNK and MAP kinases, transmitting stress and growth-promoting signals. These signals regulate the interaction of APAF with Bcl-2/Bcl-X_L.

is a major player in the control of apoptosis (We have already come across Akt in Chapter 8. Akt is the human homologue of the protein kinase encoded by the v-Akt oncogene of the transforming retrovirus Akt.

Conclusions

An important point of control is the expression of several pro-survival genes, encoding members of the Bcl-2 family, which is under the control of cytokines and growth factors. This links growth of a cell to its life expectancy.

Moreover, each differentiated cell may have its own specific apoptosis programme. This is suggested by the phenotype of *bcl-2*$^{-/-}$ knock-out mice. These mice develop normally, with apoptosis occurring only in some tissues and organs and not in others. This raises the question whether pro-and anti-apoptotic proteins do have their own, cell-specific response pattern.[43-44]

Many questions remain unanswered. One puzzle is that the pro-survival proteins of the Bcl-2 family and the apoptotic Bax family proteins appear to be structurally so similar, but functionally so different. More structural information may help to find the functionally decisive structural differences.

The role of apoptosis in the life cycle of cells

No multicellular (metazoan) organism could exist without balance between cell survival and cell death. Unregulated proliferation of cells in multicellular organisms will drive apoptosis. The reason is that rapidly proliferating cells run short of growth factors. This starts a vicious circle, because these cells compensate for the lack of growth factors by overexpression of c-*myc*, which in turn activates the apoptotic programme.[45,46] This is summarized in Fig. 13.10.

Apoptosis, cancer, and development

To sum up: cells expressing anti-apoptotic Bcl-2 have problems entering the cell cycle and proliferating, whereas expression of the pro-apoptotic Bcl-2 antagonist, Bax, has the opposite effect and accelerates cell-cycle progression. p53 has a Janus-headed function: it can promote cell proliferation and drive cells into apoptosis, but it can also stop cell proliferation by arresting the cell cycle. While, normally, proliferation is held in check by apoptosis, oncogenic mutations or overexpression of Ras and Myc drive proliferation in an unregulated fashion, upsetting the balance of proliferation and apoptosis. This potentially catastrophic situation, uncontrolled proliferation without counterbalance by apoptosis, eventually causes malignant transformation.

Provided that cell proliferation and apoptosis are balanced, malignant transformation can be prevented. Cell proliferation promoted by c-Myc and Ras is balanced by apoptosis, and apoptosis is poised by cell survival with the help of Bcl-2 and p53, which arrest growth and block the cell cycle.[50] Thus, the actions of each of these growth regulators is balanced by opposing actions. This mutual interdependency may explain why the many rapidly dividing cells in every animal are not transformed more often to cancer cells. Only when the regulatory balance between PCD and growth and proliferation is upset do cells become malignant.

Survival factors are 'Janus'-headed

One side of survival factors is benevolent. This is the side of their beneficial function, the protection of the integrity of cells. However, protection of the cell and prevention of apoptosis can also bring disaster, because pro-survival mechanisms may eventually cause malignant transformation.[51]

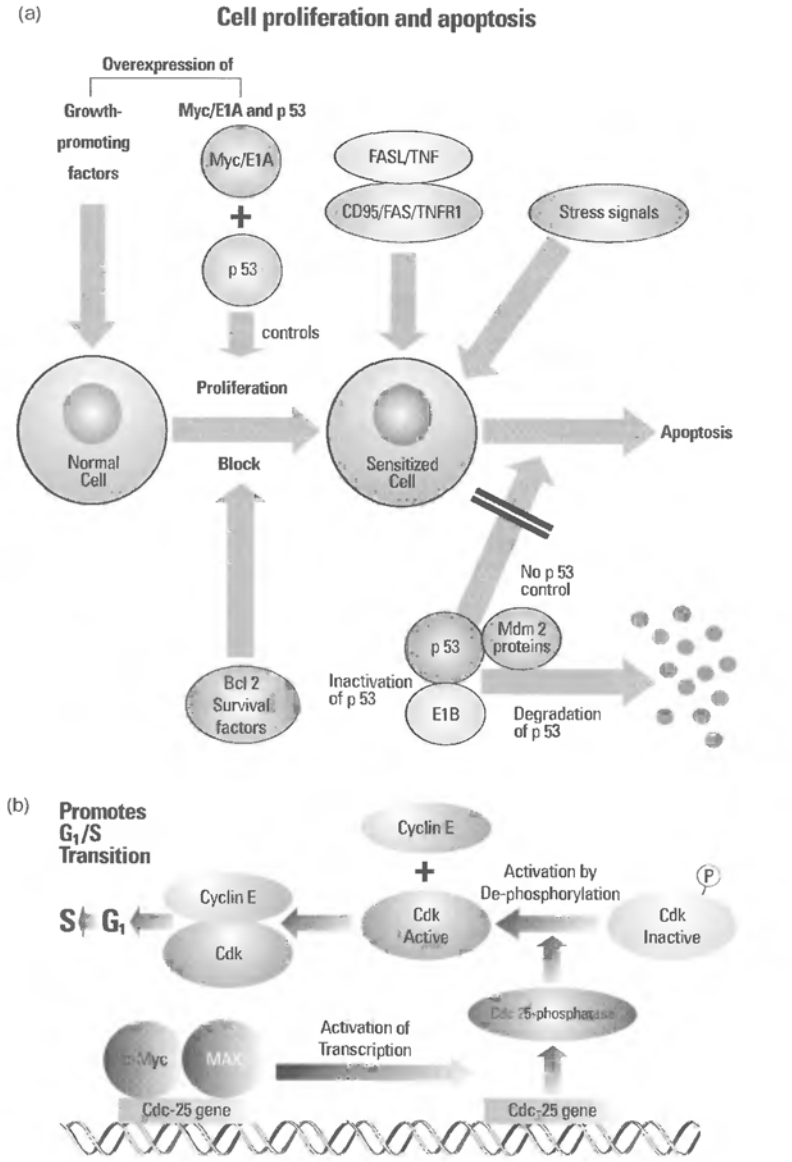


Fig. 13.10 Cell proliferation and apoptosis. (a) A normal cell is sensitized by overexpression (or mutations) of growth-promoting factors. These factors deregulate the control of cell growth, making the cell responsive to apoptotic signals. For example, Myc induces expression of genes that upregulate the CD95 ligand or the CD95 receptor, or both, thus enhancing the CD95-mediated death signals.[47] Another potent growth-promoting and eventually apoptotic factor is the Myc/E1A complex. (E1A and E1B are the products of growth-promoting adenoviral oncogenes; see Part 4). E1A causes accumulation of the tumour suppressor, p53, which may eventually drive cells into apoptosis. Myc-driven cell proliferation is counteracted by Bcl-2 survival factors and the pro-apoptotic action of p53 is counteracted by E1B which binds and inactivates p53. Another factor that keeps p53 levels low are the Mdm-2 proteins, which give the signal for degradation of p53. (Attention is directed to the role of the tumour suppressor p53, which can promote survival or apoptosis of cells, depending on the conditions). (b) A still quite hypothetical explanation has been advanced for c-Myc-driven apoptosis:[48]

c-Myc forms heterodimers with Max,[49] and the Myc/Max complex activates transcription of the *cdc25A* gene. The *cdc25A* gene codes for the Cdc25 protein phosphatase that activates the cyclin-E-dependent Cdk kinase by dephosphorylation, which forces the transition from the G_1 to the S phase, thus driving the cell cycle

B-cell lymphomas are an example. Approximately 85% of all B-cell lymphomas have a chromosomal translocation that brings the *bcl-2* gene in juxtaposition to a transcriptionally active locus. This translocation causes overexpression of the anti-apoptotic survival factor Bcl-2, prevents PCD, and prolongs the life span of the lymphocytes. But it is also responsible for malignant transformation of the lymphocytes, because in aged cells mutations accumulate that would normally be eliminated by PCD. Thus, proliferating B cells in mice overexpressing human Bcl-2 eventually change to an aggressive malignant lymphoma cell type (Bcl-2 levels are also often elevated in myeloid leukaemia). Thus, a deficiency of Bax and an overshoot of pro-survival Bcl-2 is potentially oncogenic.[52]

The ambiguous role of a survival factor, in this case, the tumour suppressor p53, becomes apparent when this survival factor and tumour suppressor is non-functional. When normal mouse embryos were irradiated or treated with carcinogens about 60% of them died, but only about every fifth pup that survived had malformations. However, in transgenic mice that lacked p53, only 7% died when the embryos were exposed to the same level of radiation, but two-thirds of the survivors had defects at birth.[53] The striking inverse relationship between fetal lethality and the incidence of birth defects reveals a kind of 'proof-reading' function of p53. The tumour suppressor p53 recognizes cells with DNA damage and helps to eliminate these cells. 'Cellular proof-reading' is comparable to DNA proof-reading by DNA polymerases that excise any incorrectly paired bases.

The role of p53 in promoting apoptosis of cells with chromosomal aberrations has led Brash,[54] to believe that the first function of suppressors, such as p53, may have been the prevention of the emergence of aberrant cell lineages. He suggests that p53-driven apoptosis may have helped to eliminate new variant cell lineages in the course of development. Newly evolving aberrant cell lineages would be expected to carry chromosomal changes more frequently, which are likely to be recognized by p53 and deleted by apoptosis.

Two examples follow, to show how essential cell death is for the development of multicellular organisms.

Controlled cell death in the nervous system

We have already alluded to the fact that the NGF-activated $p75^{NTR}$ receptor is a potential death receptor, like other members of the TNF family of receptors. Evidence is now becoming available that NGF-dependent activation of $p75^{NTR}$s can initiate cell death and apoptosis, whereas NGF and neurotrophin-binding to TRKs (growth-factor tyrosine-kinase-type receptors) has the opposite effect and prevents cell death. Thus, early in development, endogenous NGF causes the death of neurons that express $p75^{NTR}$, whereas cells that express TRK receptors are not affected and survive. Thus, the fate of developing neurons depends on the type of the receptor expressed.[55] This points to the central role of programmed cell death in the development of the nervous system.[56]

Apoptosis and immune surveillance

Cell death is equally important for the immune system.[57,58] The immune system produces killer T-cells, which eliminate infected cells and cancer cells (Chapter 14). 'Killer T-cells' express CD95L on the cell surface. The CD95 ligand on T cells binds to the CD95 receptor on infected cells and cancer cells and calls up their CD programme. The next step in the elimination of infected cells and tumour cells (but also of normal cells in the course of lymphocyte development) is the recruitment of phagocytes, which are tethered by a

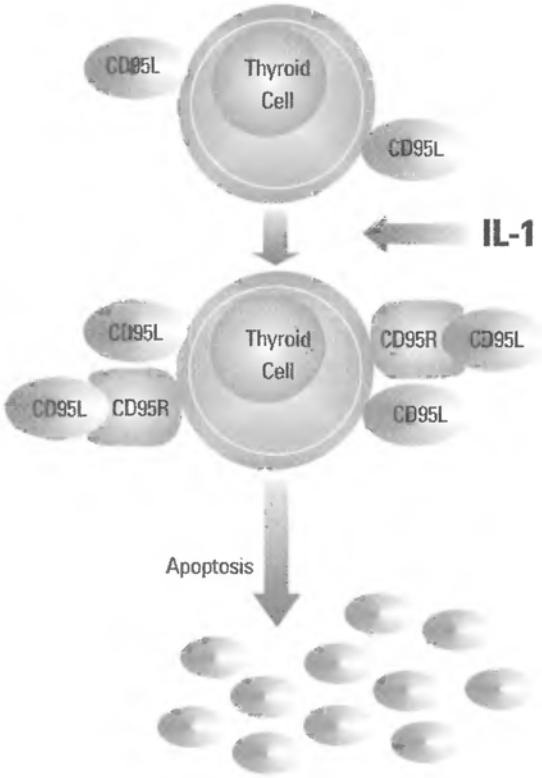


Fig. 13.11 Thyroid cells express the CD95 ligand, but express the CD95 receptor only when stimulated by IL-1. When this happens the death programme is activated in these cells by an autocrine activation loop, resulting in apoptosis, eventually leading to Hashimoto's disease. (Based on the information in ref. 61.)

glycosylphosphatidylinositol (GPI)-anchored protein to the cells condemned to die. Consequently, the cell death genes (the *ced* genes) are activated and proteins are expressed which rearrange the cytoskeleton, enabling the phagocyte to engulf the dying cell.[59]

Cell–cell interactions also help the immune system to get rid of autoreactive cells in the thymus (see Chapter 14), but this beneficial protection may have unfortunate consequences. This is exemplified by the fate of thyroid cells in patients with Hashimoto's disease (Fig. 13.11).[60]

References

1. R. A. Firtel. Integration of signaling information in controlling cell-fate decisions in Dictyostelium. *Genes Dev*, **9**, (12) 1427–1444, 1995.
   Also: R. R. Kay. Differentiation and patterning in Dictyostelium. *Curr Opin Genet Dev*, **4**, (5) 637–641, 1994.
2. M. Barinaga. Stroke-damaged neurons may commit cellular suicide. *Science*, **281**, 1302–1303, 1998.
   Also: M. Barinaga. Is apoptosis key in Alzheimer's disease. *Science*, **281**, 1303–1304, 1998.
3. C. A. Smith, T. Farrah, and R. G. Goodwin. The TNF receptor superfamily of cellular and viral proteins: activation, costimulation, and death. *Cell*, **76**, 959–962, 1994.
4. L. A. Tartaglia and D. V. Goeddel. Two TNF receptors. *Immunol Today*, **13** (5), 151–153, 1992.
5. A. Ashkenazi and V. Dixit. Death receptors: Signaling and modulation. *Science*, **281**, 1305–1308, 1998.
6. M. A. Degli Esposti, W. C. Dougall, P. J. Smolak, J. Y. Waugh, C. A. Smith, and R. G. Goodwin. The novel receptor TRAIL-R4 induces NF-kappaB and protects against TRAIL-mediated apoptosis, yet retains an incomplete death domain. *Immunity*, 7, (6), 813–820, 1997.

7. S. Nagata. Apoptosis by death factor. *Cell*, **88**, 355–365, 1997.
8. A. M. Chinnaiyan, K. O'Rourke, G. L. Yu, R. H. Lyons, M. Garg, D. R. Duan, L. Xing, R. Gentz, J. Ni and V. M. Dixit. Signal transduction by DR3, a death domain-containing receptor related to TNFR-1 and CD95. *Science*, **274** (5289), 990–992, 1996.
9. J. Kitson, T. Raven, Y. P. Jiang, D. V. Goeddel, K. M. Giles, K. T. Pun, *et al.* A death-domain-containing receptor that mediates apoptosis. *Nature*, **384** (6607), 372–375, 1996.
10. J. L. Bodmer, K. Burns, P. Schneider, K. Hofmann, V. Steiner, M. Thome, *et al.* TRAMP, a novel apoptosis-mediating receptor with sequence homology to tumor necrosis factor receptor 1 and Fas(Apo-1/CD95). *Immunity*, **6** (1), 79–88, 1997.
11. G. Pan, J. Ni, Y. F. Wei, G. Yu, R. Gentz, and V. M. Dixit. An antagonist decoy receptor and a death domain-containing receptor for TRAIL [see comments]. *Science*, **277** (5327), 815–818, 1997.
12. J. P. Sheridan, S. A. Marsters, R. M. Pitti, A. Gurney, M. Skubatch, D. Baldwin, *et al.* Control of TRAIL-induced apoptosis by a family of signaling and decoy receptors [see comments]. *Science*, **277** (5327), 818–821, 1997.
13. J. Mongkolsapaya, A. E. Cowper, X. N. Xu, G. Morris, A. J. McMichael, J. I. Bell, and G. R. Screaton. Lymphocyte inhibitor of TRAIL (TNF-related apoptosis-inducing ligand): a new receptor protecting lymphocytes from the death ligand TRAIL. *J Immunol*, **160** (1), 3–6, 1998.
14. D. W. Banner, A. D'Arcy, W. Janes, R. Gentz, H. J. Schoenfeld, C. Broger, H. Loetscher, and W. Lesslauer. Crystal structure of the soluble human 55 kD TNF receptor- human TNF-β complex: implications for TNF receptor activation. *Cell*, **73** (3), 431–445 , 1993.

    Also: J. H. Naismith, T. Q. Devine, B. J. Brandhuber, and S. R. Sprang. Crystallographic evidence for dimerization of unliganded tumor necrosis factor receptor. *J Biol Chem*, **270**, 13303–13307, 1995.

    Also: L. E. Rodseth, B. J. Brandhuber, T. Q. Devine, M. J. Eck, K. Hale, J. H. Naismith, and S. R. Sprang. Two crystal forms of the extracellular domain of type 1 tumor necrosis factor receptor. *J Mol Biol*, **239**, 332, 1994.
15. N. A. Thornberry and Y. Lazebnik. Caspases: Enemies within. *Science*, **281**, 1312–1316, 1998.
16. N. P. Walker, R. V. Talanian, K. D. Brady, L. C. Dang, N. J. Bump, C. R. Ferenz, *et al.* Crystal structure of the cysteine protease interleukin-1 beta-converting enzyme: a (p20/p10)2 homodimer. *Cell*, **78** (2), 343–352, 1994.
17. N. A. Thornberry, T. A. Rano, E. P. Peterson, D. M. Rasper, T. Timkey, M. Garcia Calvo, *et al.* A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. *J Biol Chem*, **272** (29), 17907–17911, 1997.
18. K. P. Wilson, J. A. Black, J. A. Thomson, E. E. Kim, J. P. Griffith, M. A. Navia, *et al.* Structure and mechanism of interleukin-1 beta converting enzyme [see comments]. *Nature*, **370** (6487), 270–275, 1994.
19. A. G. Porter, P. Ng, and R. U. Janicke. Death substrates come alive. *Bioessays*, **19** (6), 501–517, 1997.

    Also: V. Cryns and J. Yuan. Proteases to die for. *Genes Dev*, **12** (11), 1551–1570, 1998.
20. J. P. Medema, C. Scaffidi, F. C. Kischkel, A. Shevchenko, M. Mann, P. H. Krammer, and M. E. Peter. FLICE is activated by association with the CD95 death-inducing signaling complex. (DISC). *EMBO J*, **16** (10), 2794–2804, 1997.
21. P. Li, D. Nijhawan, I. Budihardjo, S. M. Srinivasula, M. Ahmad, E. S. Alnemri, and X. Wang.

    Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell*, **91** (4), 479–489, 1997.
22. R. P. Bissonnette, F. Echeverri, A. Mahboubi, and D. R. Green. Apoptotic cell death induced by c-*myc* is inhibited by *bcl-2*. *Nature*, **359** (6395), 552–554, 1992.
23. A. Fanidi, E. A. Harrington, and G. I. Evan. Cooperative interaction between c-*myc* and *bcl-2* proto-oncogenes. *Nature*, **359** (6395), 554–556, 1992.
24. D. R. Green and J. Reed. Mitochondria and apoptosis. *Science*, **281**, 1309–1312, (1998).
25. N. Zamzami, P. Marchetti, M. Castedo, D. Decaudin, A. Macho, T. Hirsch, *et al.* Sequential reduction of mitochondrial transmembrane potential and generation of reactive oxygen species in early programmed cell death. *J Exp Med*, **182** (2), 367–377, 1995.
26. Z. Liu, C. Sun, E. T. Olejniczak, R. P. Meadows, S. F. Betz, T. Oost, J. Herrmann, J. C. Wu and S. W. Fesik. Structural basis for binding of Smac/DIABLO to the XIAP BIR3 domain. *Nature*, **408**, 1004–1008, 2000.

    Also: G. Wu, J. Chai, T. L. Suber, J.- W.- Wu, C. Du, X. Wang and Y. Shi: Structural basis of IAP recognition by Smac/DIABLO. *Nature*, **208**, 1008–1012. 2000.

27. M. Sattler, H. Liang, D. Nettesheim, R. P. Meadows, J. E. Harlan, M. Eberstadt, *et al.* Structure of BCL-XL-BAK peptide complex: Recognition between regulators of apoptosis. *Science*, **275** (5302), 983–986, 1997.

28. S. W. Muchmore, M. Sattler, H. Liang, R. P. Meadows, J. E. Harlan, H. S. Yoon, *et al.* X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. *Nature*, **381** (6580), 335–341, 1996.

29. E. Yang, J. Zha, J. Jockel, L. H. Boise, C. B. Thompson, and S. J. Korsmeyer. Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and promotes cell death. *Cell*, **80** (2), 285–291, 1995.

30. J. M. Boyd, G. J. Gallo, B. Elangovan, A. B. Houghton, S. Malstrom, B. J. Avery, *et al.* Bik, a novel death-inducing protein shares a distinct sequence motif with Bcl-2 family proteins and interacts with viral and cellular survival-promoting proteins. *Oncogene*, **11** (9), 1921–1928, 1995.

31. E. H. Cheng, B. Levine, L. H. Boise, C. B. Thompson, and J. M. Hardwick. Bax-independent inhibition of apoptosis by Bcl-XL. *Nature*, **379** (6565), 554–556, 1996.

32. C. Borner, I. Martinou, C. Mattmann, M. Irmler, E. Schaerer, J. C. Martinou and J. Tschopp. The protein bcl-2 alpha does not require membrane attachment, but two conserved domains to suppress apoptosis. *J Cell Biol*, **126** (4), 1059–1068, 1994.

33. T. W. Sedlak, Z. N. Oltvai, E. Yang, K. Wang, L. H. Boise, C. B. Thompson, and S. J. Korsmeyer. Multiple Bcl-2 family members demonstrate selective dimerizations with Bax. *Proc Natl Acad Sci, USA*, **92** (17), 7834–7838, 1995.

34. J. M. Adams and S. Cory. The Bcl-2 protein family: Arbiters of cell survival. *Science*, **281**, 1322–1326, 1998.

35. S. J. Korsmeyer. Regulators of cell death. *Trends Genet*, **11** (3), 101–105, 1995.

36. M. O. Hengartner. Death cycle and Swiss army knives. *Nature*, **391**, 441–442, 1998.

37. E. Yang and S. J. Korsmeyer. Molecular thanatopsis: a discourse on the BCL2 family and cell death. *Blood*, **88** (2), 386–401, 1996.

38. X. M. Yin, Z. N. Oltvai, and S. J. Korsmeyer. BH1 and BH2 domains of Bcl-2 are required for inhibition of apoptosis and heterodimerization with Bax [see comments]. *Nature*, **369** (6478), 321–323, 1994.

39. M. Hanada, C. Aime Sempe, T. Sato, and J. C. Reed. Structure–function analysis of Bcl-2 protein. Identification of conserved domains important for homodimerization with Bcl-2 and heterodimerization with Bax. *J Biol Chem*, **270** (20), 11962–11969, 1995.

40. S. R. Datta, H. Dudek, X. Tao, S. Masters, H. Fu, Y. Gotoh, and M. E. Greenberg. Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell*, **91** (2), 231–241, 1997.

41. T. F. Franke and L. C. Cantley. A Bad kinase makes good. *Nature*, **390**, 116–117, 1997.

42. J. Zha, H. Harada, E. Yang, J. Jockel, and S. J. Korsmeyer. Serine phosphorylation of death agonist BAD in response to survival factor results in binding to 14-3-3 not BCL-X. *Cell*, **87**, (4), 619–628, 1996.

43. D. J. Veis, C. M. Sorenson, J. R. Shutter, and S. J. Korsmeyer. Bcl-2-deficient mice demonstrate fulminant lymphoid apoptosis, polycystic kidneys, and hypopigmented hair. *Cell*, **75** (2), 229–240, 1993. [Comments] See: Cell 87, (4), 619–628, 1996.

44. S. Kamada, A. Shimono, Y. Shinto, T. Tsujimura, T. Takahashi, T. Noda, *et al.* bcl-2 deficiency in mice leads to pleiotropic abnormalities: accelerated lymphoid cell death in thymus and spleen, polycystic kidney, hair hypopigmentation, and distorted small intestine. *Cancer Res*, **55** (2), 354–359, 1995.

45. D. R. Green. A Myc-induced apoptosis pathway surfaces. *Science*, 278, 1246–1247, 1997.

46. G. I. Evan, A. H. Wyllie, C. S. Gilbert, T. D. Littlewood, H. Land, M. Brooks, *et al.* Induction of apoptosis in fibroblasts by c-myc protein. *Cell*, **69** (1), 119–128, 1992.

47. A. O. Hueber, M. Zörnig, D. Lyon, T. Suda, S. Nagata, and G. I. Evan. Requirement for the CD95 receptor–ligand pathway in c-Myc-induced apoptosis. *Science*, **278**, 1305–1309, 1997.

48. A. Sommer, K. Bousset, E. Kremmer, M. Austen, and B. Luscher. Identification and characterization of specific DNA-binding complexes containing members of the Myc/Max/Mad network of transcriptional regulators. *J Biol Chem*, **273** (12), 6632–6642, 1998.

49. B. Amati, T. D. Littlewood, G. I. Evan, and H. Land. The c-Myc protein induces cell cycle progression and apoptosis through dimerization with Max. *EMBO J*, **12** (13), 5083–5087, 1993.

50. R. U. Janicke, F. H. Lee, and A. G. Porter. Nuclear c-Myc plays an important role in the cytotoxicity of tumor necrosis factor alpha in tumor cells. *Mol Cell Biol*, **14** (9), 5661–5670, 1994.

51. J. C. Reed. Double identity for proteins of the Bcl-2 family. *Nature*, **387**, 773–776, 1997.

52. M. E. McCurrach, T. M. Connor, C. M. Knudson, S. J. Korsmeyer, and S. W. Lowe. bax-deficiency promotes drug resistance and oncogenic transformation by attenuating p53-dependent apoptosis. *Proc Natl Acad Sci, USA*, **94** (6), 2345–2349, 1997.

53. T. Norimura, S. Nomoto, M. Katsuki, Y. Gondo, and S. Kondo. p53-dependent apoptosis suppresses radiation-induced teratogenesis [see comments]. *Nature Med*, **2** (5), 577–580, 1996.

54. D. E. Brash. Cellular proofreading [comment]. *Nature Med*, **2** (5), 525–526, 1996.

55. J. M. Frade, Rodriguez-Tébar, and Y.-A. Barde. Induction of cell death by endogenous nerve growth factor through its p^{75} receptor. *Nature*, **383**, 166–168, 1996.

56. T. M. Michaelidis, M. Sendtner, J. D. Cooper, M. S. Airaksinen, B. Holtmann, M. Meyer, and H. Thoenen. Inactivation of bcl-2 results in progressive degeneration of motoneurons, sympathetic and sensory neurons during early postnatal development. *Neuron*, **17** (1), 75–89, 1996.

57. S. Cory. Regulation of lymphocyte survival by the bcl-2 gene family. *Annu Rev Immunol*, **13**, 513–543, 1995.

58. D. T. Chao and S. J. Korsmeyer. BCL-2 family: regulators of cell death. *Annu Rev Immunol*, **16**, 395–419, 1998.

59. J. Savill. Apoptosis, phagocyte docking without shocking. *Nature*, **392**, 442–443, 1998.

60. C. Giordano, G. Stassi, R. De Maria, M. Todaro, P. Richiusa, G. Papoff, *et al*. Potential involvement of Fas and its ligand in the pathogenesis of Hashimoto's thyroiditis. *Science*, **275**, 960–963, 1997.

61. N. Williams. Research News, Thyroid disease: A case of cell suicide. *Science*, **275**, 926, 1997.

14

Regulation of the immune response

The immune response is complex and has many facets. It is an essential cell-specific function, guaranteeing our survival. This chapter concentrates on a few topics centred around signalling in the immune system. The processing of the ligands, the antigens, and their presentation to antigen receptors will be discussed, focusing on cell (T-cell) mediated immune responses. Outstanding textbooks on immunology and cell biology are available,[1,2] which the reader could consult.

Signalling in the immune system involves cell–cell interactions, but this is not unique to the immune system. Cell–cell interactions are also involved in signalling and cellular differentiation in the nervous system, establishing the interconnectivity of neurons.

The immune cells

Lymphocytes originate from pluripotent haematopoietic stem cells, which also give rise to all the other blood cells—red blood cells, white blood cells, and platelets. The stem cells are located primarily in the haematopoietic tissues (in the liver in fetuses and in the bone marrow in adults). T cells (thymus-derived cells) develop in the thymus from precursor stem cells that migrate in from the haematopoietic tissues via the blood. B cells (bone-marrow-derived cells) develop in mammals from stem cells in the bone marrow. The tissues where the lymphocytes develop are called the primary lymphoid organs. Most lymphocytes generated in a primary lymphoid organ die. But those that survive mature and migrate to the secondary, peripheral lymphoid organs—the lymph nodes, the spleen and the lymphoid tissues in the gastrointestinal tract, the respiratory tract, and the skin. It is in the secondary lymphoid organs that the immune response occurs and where B cells and T cells react with foreign antigens (Fig. 14.1).

The immune response

Antigen-specific activation of T and B lymphocytes follows, in principle, a pattern not much different from activation of cells by growth factors and cytokines. The antigen is the

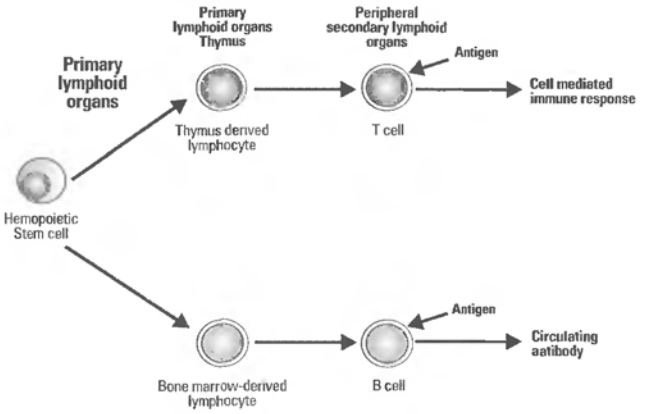


Fig. 14.1 Primary and secondary lymphoid organs. The primary lymphoid organs are the tissues where the lymphocytes develop from haematopoietic stem cells. The encounter with antigen, the immune response, occurs in the peripheral, secondary lymphoid organs. T cells are responsible for the cell-mediated immune response and B cells for the antigen-stimulated synthesis and secretion of circulating antibodies.

ligand, recognized by a membrane-bound receptor. Once bound, it elicits a specific cellular response. However, antigen–receptor interactions in the cell-mediated immune response involve an elaborate system of cell–cell interactions. The ligand molecule is processed and prepared by one kind of cell and presented to a receptor located on an other cell. The same applies to some extent to the cytokines that regulate the maturation and proliferation of immune cells. These signalling molecules are also sometimes produced by one kind of cell, but act on another type of cell, although in some cases cytokines activate by autocrine stimulation the same cell, which has produced them.

The major difference between T cells and B cells is that the former recognize antigens only when they are displayed on the surface of another cell, whereas B cells accept cell free, particulate and soluble antigens. Here the focus will be on T cells.

The MHCs, the major histocompatibility complex molecules, have a central role in the presentation of antigens. Mature ($CD4^+$ helper) T cells in the peripheral lymphoid organs, in the spleen, and in the lymph nodes recognize antigenic peptides associated with MHC class II molecules, displayed on the surface of bone-marrow-derived antigen-presenting cells (APCs), whereas ($CD8^+$) killer T-cells recognize foreign antigens bound to class I MHC molecules which are displayed on the surface of cells that are destined to be destroyed.

Processing and presentation of antigen

Each antigen presents many structurally different epitopes, antigenic sites, to which the immune system responds by producing many clones of activated lymphocytes, but each directed specifically against only one of the many sites of the antigen. Although the response of cells to an antigen is polyclonal, each clone produces only one kind of antibody. 'One cell one antibody'. In that way, activation by antigen leads to a huge expansion of the reservoir of immune competent cells.[3]

Differentiation of lymphocytes produces cells with different functions, such as B cells, and T helper and killer cells. The former produce and secrete antibodies, releasing them into lymph and blood, the latter are responsible for the cellular immune response. The response of T cells is antigen-specific, just like the response of B cells. But, T cells interact with another cell in the body, the antigen-presenting cells (APCs), which have taken up and processed the foreign antigen. Helper T cells react with APCs and killer T cells

interact with and kill infected cells or cancer cells, which present on their surface a fragment of a cell-specific antigenic peptide that they have produced (Plate 28). A classic example is recognition of virus-infected cells by killer T-lymphocytes, sensing viral antigens on the cell surface.[4]

R. Zinkernagel and P. Doherty first recognized that T lymphocytes accept foreign cellular antigens only when they are embedded in complexes with major histocompatibility proteins. (Moreover, they found that an antigenic peptide, presented on the surface of a cell and recognized by T cells, is combined with an MHC protein always originating from only one allele of the MHC gene (the MHC restriction principle)[5] (see also ref. 6).

MHC, the major histocompatibility complex

MHC molecules are polymorphic. Each individual has five or more gene loci encoding MHC molecules. Each dimeric MHC molecule binds only one ligand. Although there is only a limited number of structures of MHC molecules, perhaps 20, MHC molecules are able to bind to every conceivable structure, to a very large number of different peptides and carbohydrates and unnatural, synthetic molecules. The X-ray crystallographic structure helps to explain this (Fig. 14.2).[7–10]

Class I MHC molecules, recognized by killer T-cells are expressed on virtually all nucleated cells. This makes sense, because cytotoxic, killer T-cells must be able to identify all cells that can become infected. Class II MHC molecules are more restricted and are confined to antigen-presenting cells (APCs).[11]

Figure 14.3 shows that when a self-antigen binds with low affinity to a T cell in the thymus, a primary lymphoid organ, the T cell will survive and can move on to the peripheral lymphoid organs. (positive selection). But, when the self-antigen is accepted by the T cell with high affinity, the T cell will die (negative selection). In that way, autoimmune-reactive T cells are eliminated. The selection takes place in different regions of the thymus and is carried out by different cells. Positive selection takes place in the cortex of the thymus, whereas negative selection occurs in the medulla of the thymus.[12–14]

The antigen-processing machinery is not yet completely understood. The current interest in antigen processing reflects the importance of the recognition of antigens for transplantation medicine and for an understanding of autoimmune diseases.

Selection and clonal expansion of T cells

A unique and characteristic property of immune cells is their differentiation, proliferation, and clonal expansion in response to antigenic stimulation.[16] Intimately related with this programme is the capability of the immune system to distinguish between self and foreign antigens. The selection processes that control the development of $CD4^+$ helper T-cells that recognize MHC class II-peptide complexes, and of $CD8^+$ cytotoxic T-cells that recognize MHC class I-peptide complexes, are similar.

The encounter between self-antigens and MHC molecules in the thymus is responsible for the selection. The experience gained in the encounter with self-antigens, binding to MHC molecules, helps to select T cells trained to react with every conceivable antigen, except its own. Encounter with a foreign antigen then leads to the expansion and formation of a huge and diverse repertoire of T cells.

The role of cytokines

T-cell differentiation to effector cells and expansion of the cell pool is controlled by cytokines, quite like in other haematopoietic cell lines. Expression of the IL-2 gene is an

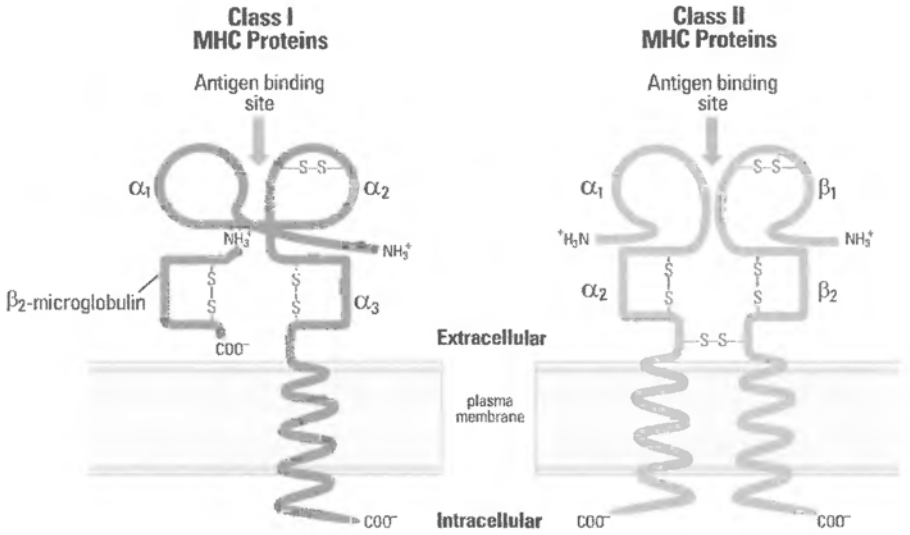


Fig. 14.2 Class I and Class II MHC Proteins. On the left is shown a class I MHC protein. It has a α-chain, containing three extracellular domains, α_1, α_2, α_3. The α_2 and α_3 domains are held together by disulphide bonds. Non-covalently associated with the α-chain is a smaller, separate peptide chain, the β_2-microglobulin. The β_2-microglobulin is a separate entity. It is not encoded within the MHC gene locus. The β_2-microglobulin and the α_3 domain are Ig-like domains. But, whereas the β_2-microglobulin is invariant, the α-chain is very polymorphic, mainly in the α_1 and α_2 domains. A single antigen-binding site is on the NH_2-terminal, extracellular site of the molecule. The antigen-binding site consists of a deep groove formed by two long α-helices derived from the nearly identical α_1 and α_2 domains. The base of the groove is formed by eight β-strands, derived from the same two α-domains. The groove can accomodate an extended peptide chain of about 10 amino-acid residues. On the right is a class II MHC molecule. Class I and class II molecules are structurally similar. The MHC II protein has an α- and a β-chain, which are both polymorphic. Both the α_2 and the β_2 domains are immunoglobulin-like and held together by disulphide bonds. Again the two outermost, extracellular domains, α_1 and β_1 form the antigen-binding groove. β_1 is held together by disulphide bonds. All chains are glycosylated, with the exception of the β_2-microglobulin.

important early event in T-cell proliferation. The functions of cytokines are summarized in Table 14.1, which shows the cells that produce interleukins, involved in the immune response, and the cells targeted by interleukins, and their responses. Although expression of the IL-2 gene is an important event in T-cell proliferation it is not the only cytokine involved, as is apparent in Table 14.1.

How IL-2 controls T-cell differentiation is not yet fully resolved.[17,18] IL-2 has apparently growth-promoting and growth-terminating effects. Early, when IL-2 concentrations are low, the proliferative effect may be dominant, but later, when the cytokine concentration (and that of other stimuli) increase, IL-2 may terminate the response. Inhibition of lymphocyte proliferation, notably by IL-2 and TGF-β, is clinically important, because immunosuppressive cytokines, such as TGF-β, prevent the inflammation accompanying immune responses.[19]

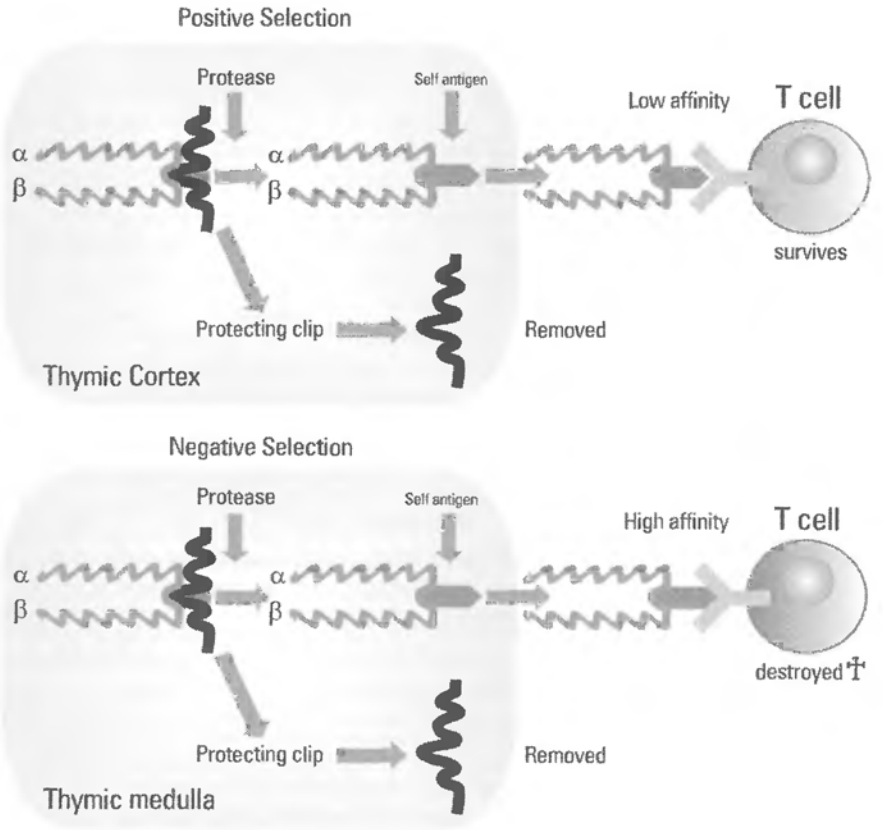


Fig. 14.3 Selection of T-cells by self-antigen. The unoccupied, antigen-binding site on the α,β-chains of a MHC class II molecule is protected against proteolysis during procession of the antigen by a place holder, called CLIP (MHC-class II-associated invariant chain peptide). CLIP is a transmembrane glycoprotein which remains bound until the peptide fragment from the self antigen is processed and bound to MHC-II, ready to interact with the T-cell receptor. Finally, the released invariant chain fragment is degraded by cysteine proteases (cathepsins). (The scheme is based on the information in ref. 15, and shown with permission of Science).

The phenotype of IL-2-deficient mice shows that, besides IL-2, additional factors are required for differentiation of T cells into functionally different subsets, because disruption or deletion of the IL-2 gene does not lead to immunodeficiency, but rather to uncontrolled accumulation of activated lymphocytes, resulting from a block of apoptosis.[20] This points to the participation of other cytokines, such as IL-10, in the immune response.[21] All signals converge eventually on the transcription factor NF-ATc (nuclear factor for activation of T cells), a member of the NF-AT family of transcription factors (see Chapter 10). NF-ATc is targeted to the promoter of the IL-2 gene. The nuclear import of NF-ATc is controlled by Ca^{2+} and calcineurin (a calcium–calmodulin-activated serine/threonine phosphatase) and the main target of immunosuppressive drugs in clinical use which are calcineurin inhibitors (see also Chapter 10).[23–24]

Table 14.1 Properties of some interleukins

Interleukin	Produced by	Target	Functions
IL-1	Antigen-presenting cells	Helper T-cells	Activation
IL-2	Helper T-cells	All activated T cells and B cells	Proliferation
IL-3	Helper T-cells	Various haematopoietic cells	Proliferation
IL-4	Helper T-cells	B cells	Stimulation of proliferation and maturation; antibody class switching
IL-5	The same helper T-cells that make IL-4	B cells, eosinophils	Promotion of proliferation and maturation
IL-6	Helper T-cells and macrophages	Activated B cells, T cells	Promotion of B cell maturation to Ig-antibody secreting cells; activation of T cells
γ-Interferon	The same helper T-cells that make IL-2	B cells, macrophages, endothelial cells	Activation of various MHC genes; activation of macrophages

The information in this table is taken from Table 23-4 of ref. 2, with permission of Taylor and Francis, Inc.

The antigen receptors

T-cell activation involves interactions of membrane-bound immunoglobulin-like receptors with membranous antigen–MHC complexes. The antigen receptors of B and T lymphocytes are complex membrane-bound multisubunit proteins. Antigen receptors are immunoglobulin (Ig)-like antibody molecules. B cells do not secrete the first antibodies that they make, instead, they insert them into the plasma membrane, where they now serve as receptors for antigen. Each B cell has approximately 10^5 such molecules in the plasma membrane.

The kinship of cell-bound antigen receptors and soluble antibody immunoglobulin molecules becomes apparent when one compares their structures (Fig. 14.4).

T-cell receptors have two disulphide-linked polypeptide chains (α, β), each of which has a variable amino-terminal region and a constant carboxyl-terminal region. Variability is created by genetic rearrangements and somatic mutations (see ref. 25). The variable regions form the one antigen-binding site of the membrane-bound receptor, just as in the case of a soluble antibody molecule with its two binding sites. The affinity of the T-cell receptor for antigen is rather low ($K_a \sim 10^4$ litres/mole). The low affinity is compensated by cell–cell interactions with the antigen-presenting cell. T cells form clusters with antigen-presenting cells. This cell–cell interaction has been called an immunological synapse.[26] The formation of a cell cluster is the key event in T-cell proliferation.[27] For the initiation of the cellular immune response, a host of co-activators and co-receptors is necessary.

Structures of complexes of a human T-cell receptor and a human MHC molecule, presenting a viral peptide (Plate 29, ref 28), and of a T-cell receptor β-chain with a bound superantigen,[29] and structures of α- and β-chains of the TCR have been solved. Superantigens (SAgs) are viral or bacterial proteins. They have been implicated in the toxic shock syndrome. The *Staphylococcus aureus* enterotoxins SEC2 and SEC3 cause toxic shock and food poisoning. The SAg seems to drive a wedge between the TCR and

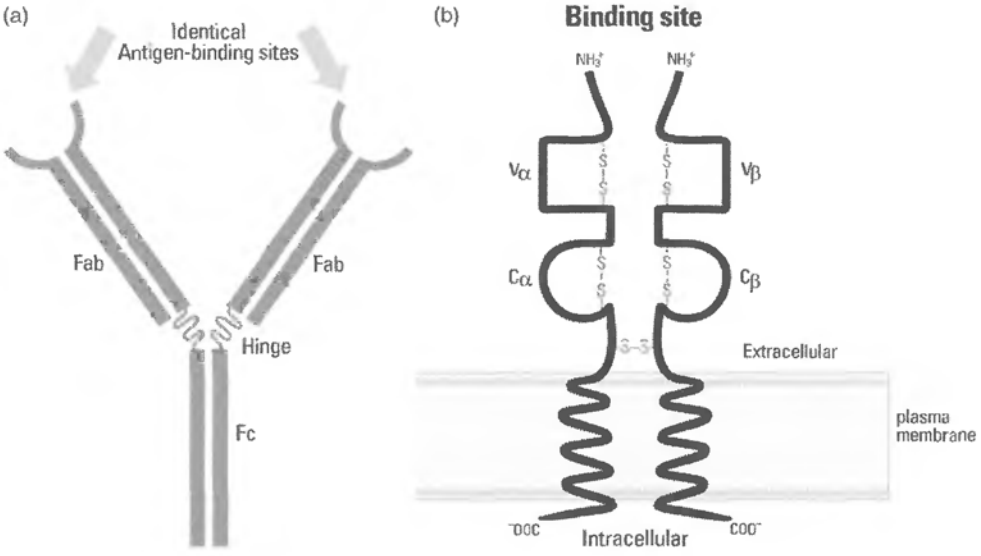


Fig. 14.4 (a) An antibody immunoglobulin molecule has two identical antigen-binding sites, formed by the Fab chains. Two of the Fab chains are linked through a flexible hinge region to the two Fc chains, which carry the antigenic determinants of an antibody protein. Each half of the Fab part is made up of two chains, a light chain and a heavy chain. The heavy chains extend into the Fc region. Each light chain and each heavy chain has variable and constant regions. The variable regions form the antigen-binding sites. (b) The T-cell antigen receptor (TCR) is constructed like an antibody, but has only one binding site for antigen. The transmembrane T-cell antigen receptor has an α-chain and a β-chain with variable and constant regions.The variable parts of each chain, V_α and V_β form the antigen-binding site.

the MHC. The consequence is that superantigens prevent access of a MHC-bound antigenic peptide to the TCR. Thus, the SAgs circumvent the normal mechanism of recognition of specific peptide/MHC complexes by TCRs.

An important message from the structural information is that no gross conformational changes of the TCR occur when it is complexed with the antigen–MHC molecule. Thus, interactions of the MHC-bound antigenic peptides with the TCR alone seem not to be enough for signal transduction. Therefore, for productive signalling, additional proteins on the surface of T cells (and of the partner cells) are required. These co-receptors are not involved directly in antigen binding, but their transmembrane polypeptide chains with cytoplasmic tails are required for signalling and linking the antigen receptor to a tyrosine kinase phosphorylation cascade.

The signalling complex

Some of these accessory proteins are summarized in Table 14.2. All T cells express a set of invariant transmembrane polypeptide chains, the so-called CD3 complex. CD is the abbreviation of cluster of differentiation, because each of the CD proteins was originally

Table 14.2 Accessory proteins on the surface of T cells

Protein		Family	Expressed on	Interacts with	Functions
CD3	γ chain	Ig	all T cells	–	Signal transduction
	δ chain	Ig			
	ε chain	Ig			
	φ chain	–			
CD4	one chain	Ig	helper T-cells	class II MHC	Adhesion to antigen-presenting cells
CD8	(homodimer or heterodimer)	Ig	cytotoxic T-cells	class I MHC	Adhesion to infected target cells
CD28	(homodimer)	Ig	many helper and cytotoxic T-cells	B7	Secondary signals

The information in this table is taken from Table 23-3 in ref. 2, with permission of Taylor and Francis, Inc.

defined as a T-cell-differentiation antigen and a marker to distinguish the various types of effector T-cells. Helper T-cells and cytotoxic, killer T-cells bear additional specific marker proteins, the CD4 and CD8 proteins, respectively (Table 14.2).

Signalling in T and B cells

Signalling in T and B cells is controlled by phosphorylation and dephosphorylation, whereby Src-related cytosolic tyrosine kinases play a major role.

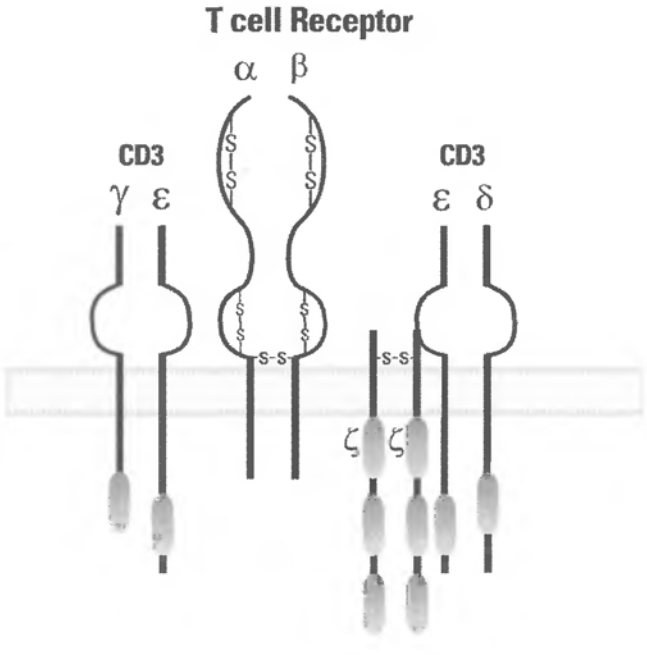


Fig. 14.5 The TCR complex is associated with a number of T-cell specific membrane-spanning proteins.[31] The antigen receptors and the co-receptors, CD3, CD4, and CD8, together with associated enzymes, tyrosine kinases and phosphatases, form the actual signalling complex. The cytoplasmic chains of the co-receptor molecules have characteristic consensus sequences, the ITAMs (immunoreceptor tyrosine activation motifs). Each of the invariant ζ-chains and the CD3-γ,δ, ε chains, contain 1–3 copies of the ITAM motifs. The structure of the TCR-signalling complex still needs to be clarified.

Phosphorylation and dephosphorylation

A cytosolic tyrosine kinase, the Lc (lymphocyte) kinase, phosphorylates the ζ chains of the signalling T-cell receptor co-receptor complex. The intracellular ARAM motif (antigen-recognition and activation motif) of the ζ chains serves as docking site for the SH2 domain of the ZAP-70 kinase (ζ-associated protein 70 kinase).[32,33] Patients carrying mutations in both alleles for Zap-70 suffer from a rare human immunodeficiency disease, where $CD8^+$ T-cells are not produced, and the remaining peripheral $CD8^+$ T-cells are unresponsive.[34] ZAP-70 activates PLC- γ1. This leads to the formation of lipid messengers and activation of the PtdIns pathway. A phosphatidylinositol phosphatase, that regulates formation of phosphatidylinositol-lipid messengers serves as negative regulator in lymphocytes. Fyn, another cytosolic tyrosine kinase of the src family activates a phosphatidylinositol kinase and responds to T-cell activation with activation of the IP_3 receptor, which controls the intracellular IP_3-gated calcium channel.[35] This leads to release of intracellular Ca^{2+}.

The attenuation of cell proliferation of T and B cells is an impressive example of the role of phosphatases in the control of growth and expansion of cell populations. The activity of the tyrosine kinases (Lc, Fyn, and ZAP) is balanced by tyrosine phosphatases. The tyrosine phosphatase, SHP-2 (see Chapter 3), binds to a T-cell surface protein, CTLA-4 (cytotoxic T lymphocyte-associated protein-4). Binding to CTLA-4 activates the phosphatase, which now dephosphorylates and removes the pTyr-docking sites on CD3ζ for the adaptor Sos/GRB2. But, without coupling to Sos/GRB2, the activated T-cell cannot turn on the Ras/MAPK pathway and the lymphocytes can not proliferate. On the other hand, when phosphorylation remains turned on, T cells proliferate in an uncontrolled fashion. This is the case in *ctla-4* knock-out mice, where the activity of the kinase is no longer poised and balanced by a tyrosine phosphatase. A variation of the theme is the action of the related tyrosine phosphatase SHP-1, which counteracts phosphorylation by the tyrosine kinase ZAP-70.[36] Mutant T cells, without SHP-1, respond hypersensitively to antigen and are stimulated in an uncontrolled fashion.

Despite significant differences in antigen presentation between T and B cells, cellular signalling is similar. The role of the tyrosine kinase ZAP-70 in T cells is carried out in B cells in an analogous fashion by Syk, a closely related cytosolic tyrosine kinase.[37] In brief, on activation of the B-cell receptor, the two SH2 domains of the Syk kinase bind to two phophotyrosyls in the conserved ITAM (immunoreceptor tyrosine activation motif) domain of the cytoplasmic extensions of the B-cell receptor (BCR) complex. This leads to activation of the Syk kinase. Associated with the BCR complex are co-receptors and a SHP1 tyrosine phosphatase which de-activates the co-receptor, CD22, and turns down activation. Thus, phosphatases downregulate cell proliferation in B cells, in principle, like in T cells, demonstrating that signalling in B cells and T cells is regulated in the same way by phosphorylation dephosphorylation.

Co-stimulation by 'second signals'

Although T- and B-cell activation is specified by the antigen—the 'first signal'—additional 'second signals' are necessary for a productive immune response. The central function of co-stimulators is the promotion of interactions of T cells with antigen-presenting cells and also of T and B cells. Typical co-stimulators, mediating second signals, are CD28 and members of the B7 family, B7-1 (CD80) and B7-2 (CD86) (Fig. 14.6).[38]

Activation and proliferation of T cells does not occur when either B7 or CD28 is missing. When CD28-deficient T cells encounter an antigen-presenting cell, they are not stimulated to proliferate and expand, instead they rapidly undergo apoptosis. But, when B7 interacts with CD28 and turns on tyrosine phosphorylation and the PtdIns 3-kinase, signalling is productive. Activation of the PtdIns 3-kinase pathway leads to the formation of the PtdIns-3,4,5-P_3 lipid messenger, although this is not the only signalling pathway involved in T-cell activation. While co-stimulation promotes interaction between helper T-cells and antigen-presenting cells, a second system of co-stimulation promotes interaction of B cells and T-helper cells, which is necessary for efficient B-cell activation. This is the CD40 pathway which establishes contacts between helper T-cells and B cells and other cells and macrophages. Still another co-stimulator is C3D, located on the surface of all T cells, which participates in complement activation. C3D engages as partners, CD21/CD22 proteins on the surface of B cells and in concert with the antigen, enhances B-cell proliferation.[39] The critical role of complement as a second signal for B cells is demonstrated by the virtual absence of antibody responses in mice lacking C3D or the CD21 receptor.

A practical consequence is that for highly immunogenic skin and cardiac allografts to be accepted, both the CD40 and the CD28 pathways must be blocked simultaneously.[40] Therefore, co-stimulation by 'second signals' has become a major target for drugs developed for the prevention of allograft rejection. Many efforts are being made to better understand the second-signalling routes, hoping to find eventually more effective strategies to forestall allograft rejection. Clinical trials of antagonists of co-stimulators, blocking pathological autoimmune responses and responses to allografts, are currently under way.

Summary

The essential features of lymphocyte activation resemble receptor-mediated signalling by cytokines. In both cases, the receptors lack kinase activities, and signalling from the receptor on the cell surface to transmitters, and eventually to the nucleus and the genes, is mediated and controlled by cytosolic protein kinases and phosphatases. Recognition

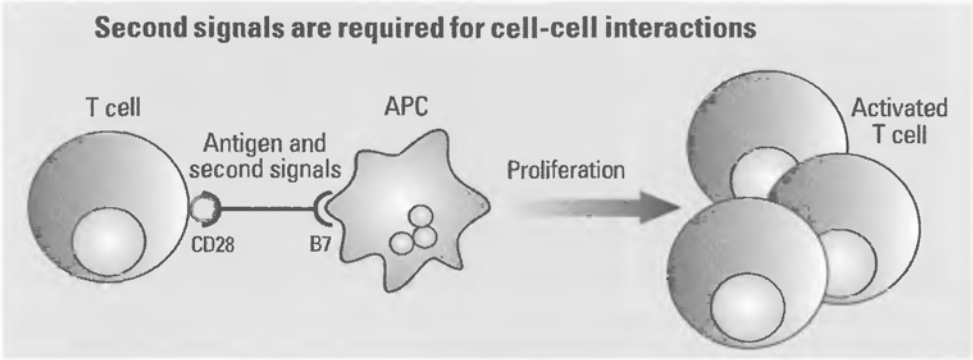


Fig. 14.6 The B7 and CD 28 molecules belong to the immunoglobulin (Ig) family of proteins. B7 is located on the surface of antigen-presenting cells (APCs) and docks on the CD28 receptor of T cells.

between regulatory components is through familiar structural motifs, such as SH2 domains, and phosphotyrosyl docking sites.[41] Finally, the control of differentiation of lymphocytes into a variety of effector cells with different functions has much in common with developmental pathways of other cell lineages, notably other haematopoietic cells. One group of transcription factors is quite universally required for differentiation of many haematopoietic cell lineages, whereas the other group is restricted to T-cell development. A transcription factor specific for the earliest stages in thymocyte development is GATA-3, a Zn-finger-type transcription factor which recognizes the GATA-motif in the DNA (see Chapter 10). Several T-cell-specific early genes have been implicated as targets of GATA-3, among them the gene for the T-cell-receptor β-subunit, TCR-β. A homozygous mutation in the GATA-3 gene in murine embryonic haematopoietic stem cells causes an early developmental defect in the GATA-$3^{-/-}$ thymocytes.[42]

The complexity of control of T and B cells reflects the necessity to handle so many different and complex tasks: cell–cell interactions, interactions of receptors with antigens, prevention of recognition of self-antigens, transmission of primary and secondary signals, leading eventually to clonal selection and expansion and a productive immune response (Fig. 14.7).

In the next section, the events after the encounter with antigen has taken place will be discussed. Of special interest in this context are signals controlling cell death and their role in establishing and maintaining self-tolerance.

Events after antigenic stimulation

Following the antigen response, some lymphocytes develop into long-lived, functionally quiescent memory cells (the induction of a high level of memory cells is an indication of

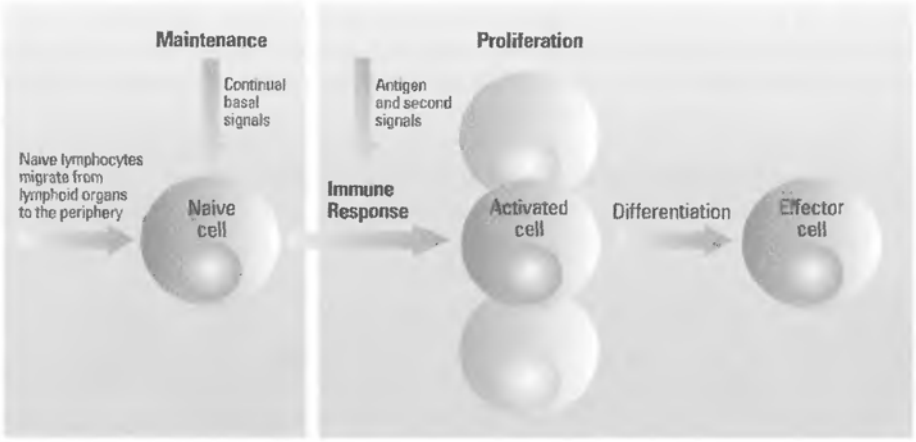


Fig. 14.7 Stages in the immune response: naive lymphocytes migrate from the primary lymphoid organs where they are formed to the periphery, where they are maintained, waiting for an antigen. On exposure to antigen and second signals they are activated. Activation of naive T and B cells is followed by an enormous proliferation and clonal expansion and ends with their maturation and terminal differentiation to effector cells.

a successful vaccination), but the majority of the lymphocytes which are produced in response to antigenic stimulation are either used up or removed (Fig. 14.8). Consequently, the reservoir of lymphocytes must be replenished and brought up to about the same number of cells as before the response to the antigen. Once this stage is reached, the immune system goes into a steady state, where the number of lymphocytes entering the pool is about equal to the number of cells that are removed. But, even in that quiescent state, the cells are kept on alert so that they can again respond to a new antigenic challenge.

The preservation and maintenance of an adequate pool of lymphocytes, ready to respond to an antigen, is somehow dependent on continuous signals.[43] These signals are apparently presented in the form of major histocompatibility complex-encoded molecules on the surface of dendritic cells.[44–45] Only with these signals can virgin cells, which have not yet encountered an antigen, survive ready for action, for long periods of time, up to several months. In the absence of signals, lymphocytes die by apoptosis (Fig. 14.9).

As summarized in Fig. 14.9, continuous exposure to signals is essential for the survival of dormant T cells. The cells are maintained in this state because the death programme is checked and apoptosis is forestalled. The baseline level of lymphocytes in the resting state is low. Upon antigenic stimulation, this level can increase 1000-fold, to 10^3 or more lymphocytes within a week, and then return to baseline levels again within 4–12 weeks, after the antigen has been removed.[50]

To sum up: the absence of stimulation by antigens and second signals promotes passive cell death. Maintenance, homeostasis, is only guaranteed when the cell continuously

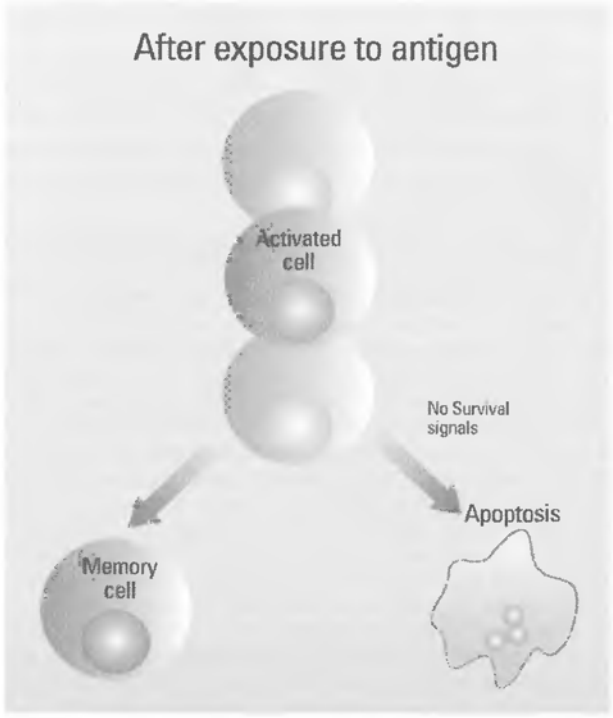


Fig. 14.8 The events after exposure to antigen. After the immune response, most of the activated cells are removed, in the absence of survival signals, by apoptosis, but some cells survive and become memory cells.[46] However, the more often the activated effector cells divide, fewer and fewer memory cells are formed. When these cells are stimulated a second time with the same antigen, the response is much greater than the first time. (Information from Fig. 2 of ref. 46 used, with permission of the author and Nature.)

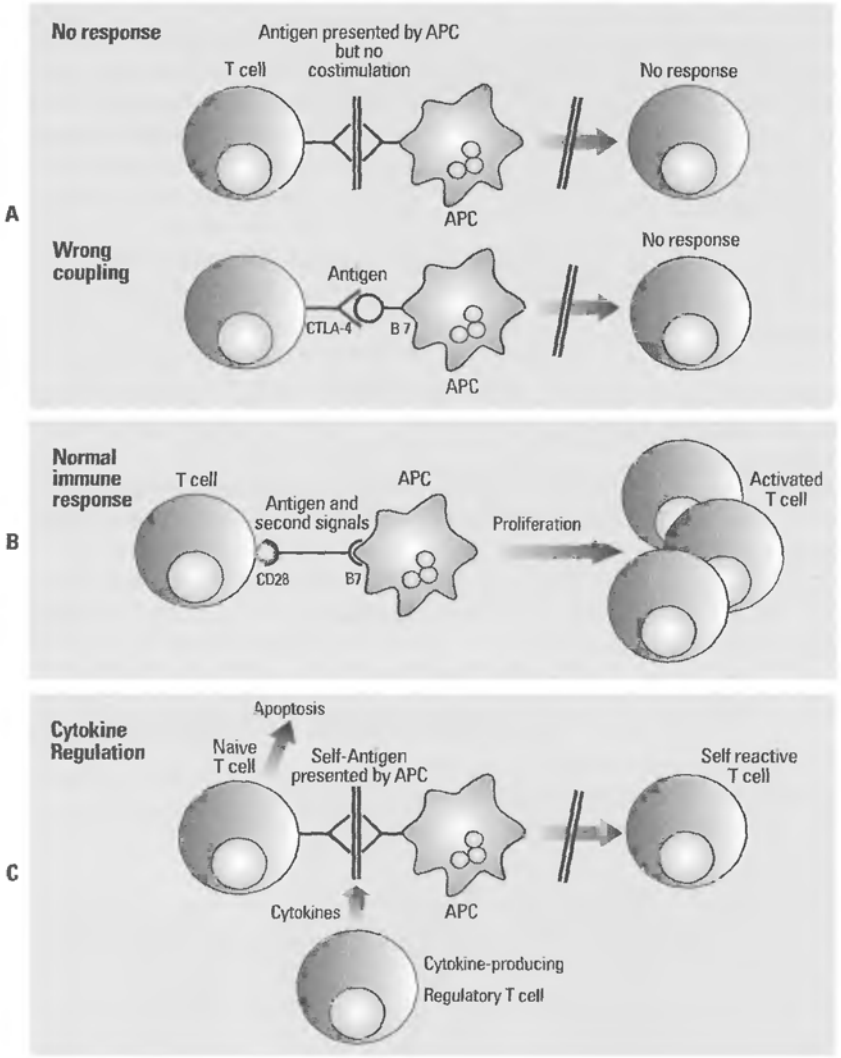


Fig. 14.9 (a) In the absence of stimulation by second signals, or when the coupling of T cells with antigen-presenting cells (APCs) is incorrect, no response occurs. The T cells are anergic and will die a passive cell death. Passive cell death is prevented when subliminal stimulation by antigens and second signals is supplied. In the stimulated cells anti-apoptotic proteins (Bcl-2 and Bcl-X_L) are expressed.[47] Lymphocytes which die a passive death are deficient of anti-apoptotic proteins. Unresponsive lymphocytes have a signalling-competent receptor complex, but they do not receive second signals provided by co-stimulators. This is due either to a lack of B7/CD28-mediated co-stimulation, or to wrong coupling (that is when the CTLA-4-co-receptor complex, rather than the legitimate CD28 T-cell co-receptor, interacts with B7 molecules on the cells presenting antigens).[48] As mentioned above, CTLA-4 carries with it the tyrosine phosphatase SHP-2. Thus, when CTLA-4 is bound to the surface of T cells and cross-linked to B7 on antigen-presenting cells (APCs), the consequence is that dephosphorylation prevents expression of the IL-2 gene and consequently blocks the progression of T cells through the cell cycle and their proliferation.[49] Although this scenario is quite plausible, it is still speculative. (b) A normal immune response occurs when both antigen stimulation and co-stimulation by second signals is provided. (c) IL-2 and cytokine-producing regulatory T cells can prevent the recognition of antigen-presenting cells (APCs), and prevent the proliferation and differentiation of naive T cells to mature effector cells and an immune response. Active cell death of naive T cells occurs on stimulation by self-antigen. It leads to expression of the death receptor, Fas (CD95), and its ligand, FasL, and to autocrine activation of the death pathway. Activation-induced cell death prevents the emergence in peripheral tissues of mature, self-reactive T cells from naive T cells and maintains self-tolerance.

receives survival signals. The maintenance of the lymphocyte pool in homeostasis and the prevention of passive cell death requires expression of anti-apoptotic proteins, Bcl-2 and Bcl-X_L. Whereas the lymphocyte pool is controlled by passive cell death, active cell death protects against reactivity towards self-antigens and establishes tolerance. Stimulation by self-antigens promotes expression of pro-apoptotic proteins, Fas/FasL and IL-2, and active cell death by apoptosis

Self-tolerance

Self-tolerance protects us against an immune response to our own self-antigens. Normally, lymphocytes that have left the primary lymphoid organs, where they were generated, are tolerant and maintain tolerance even when they encounter self-antigens in peripheral tissues. How mature lymphocytes manage to remain unresponsive to self-antigens, although they are equipped with receptors capable of recognizing self-antigens, is still a matter of speculation. But it is clear that cells stimulated by antigens (including self-antigens) eventually must die. It follows that a breakdown of the mechanisms controlling active cell death should result in autoimmune diseases. This is the case: knock-out mice, lacking the death-promoting CD95L/CD95 receptor develop autoimmune diseases.

As in other cells, the CD95L/CD95 receptor (FasL / FasR) pathway is critical for terminating immune responses by apoptosis. In humans, mutations of the gene for the death receptor, the CD95R (Fas-receptor) gene, are associated with an autoimmune syndrome.[51] But this is not the only cause of an autoimmune response. Mice also develop an autoimmune syndrome when the *ctla-4* gene (the gene for the cytotoxic lymphocyte-associated protein) is disrupted.[52] Cytokines also play a role in the control of active cell death and autoimmunity, this is apparent from the phenotype of mice lacking functional genes for IL-2 or the IL-2 receptor. These mice develop lymphadenopathy and various manifestations of autoimmunity, including autoimmune haemolytic anaemia. A locus that maps close to the *ctla-4* gene in mice and humans, and also polymorphism of the IL-2 gene, have been linked to the insulin-dependent form of diabetes mellitus in the young (MODY, maturity onset diabetes of the young), which is an autoimmune disease (see Chapter 8 and refs 53,54).

Conclusions and outlook

The identification of the pathways in lymphocytes that control the response to foreign and self antigens has provided a solid base from which the problem of self-tolerance can be attacked and eventually solved. However, there may still be formidable obstacles along this road. Van Parijs and A. K. Abbas[50] have pointed out that one of the complications in the prevention of autoimmune diseases could be that each self-antigen may have its own specific programme to establish self-tolerance. This would mean that the processing and signalling pathways for each self-antigen must be identified in order to understand the pathogenesis of a related autoimmune disease and eventually find ways and means of preventing it.

References

1. Richard, A. Goldsby, Thomas, J. Kindt, Barbara Osborne. 'Kuby Immunology', Fourth Edition, W. H. Freeman, New York, 2000

   Also: A. K. Abbas, A. H. Lichtman, and J. S. Pober. Cellular and Molecullar Immunology, Saunders, Philadelphia, 1991.

2. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell.* Garland Publishing, New York, 1994.

3. F. M. Burnet. *The clonal selection theory of acquired immunity,* Cambridge University Press, Cambridge, 1959.

   Also: N. K. Jerne. Antibodies and learning: selection versus instruction. In G. C. Quarton, T. Melnechuk, and F. O. Schmitt, (eds). The Neurosciences. A study Program, Rockefeller University Press, New York, NY, 1967.

4. R. M. Zinkernagel. Immunology taught by viruses [see comments]. *Science*, **271** (5246), 173–178, 1996.

5. R. M. Zinkernagel and P. C. Doherty. Restriction of *in vitro* T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*, **248**, 701–702, 1974.

6. P. Parham. Pictures of MHC restriction [news; comment]. *Nature*, **384** (6605), 109–110, 1996.

7. D. R. Madden, J. C. Gorga, J. L. Strominger, and D. C. Wiley. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell*, **70** (6), 1035–1048, 1992.

8. D. R. Madden.The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol*, **13**, 587–622, 1995.

9. D. H. Fremont, M. Matsumura, E. A. Stura, P. A. Peterson, and I. A. Wilson. Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb [see comments]. *Science*, **257** (5072), 919–927, 1992.

10. L. J. Stern and D. C. Wiley. Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure*, **2** (4), 245–251, 1994.

11. J. Pieters. MHC class II restricted antigen presentation. *Curr Opin Immunol*, **9** (1), 89–96, 1997.

12. P. Cresswell. Proteases, processing, and thymic selection [see comments]. *Science*, **280** (5362), 394–395, 1998.

13. S. M. Alam, P. J. Travers, J. L. Wung, W. Nasholds, S. Redpath, S., Jameson, and N. R. J. Gascoigne. T-cell receptor affinity and thymocyte positive selection. *Nature*, **381**, 616–620, 1996.

14. C. A. Janeway Jr. Thymic selection: two pathways to life and two to death. *Immunity*, **1** (1), 3–6, 1994.

15. T. Nakagawa, W. Roth, P. Wong, A. Nelson, A. Farr, J. Deussing, *et al.* Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus. *Science*, **280** (5362), 450–453, 1998.

16. N. K. Jerne. The immune system. *Sci. Amer.*, **229**, (1), 55–60, 1973.

17. J. Jain, C. Loh, and A. Rao. Transcriptional regulation of the IL-2 gene. *Curr Opin Immunol*, **7** (3), 333–342, 1995.

18. E. Serfling, A. Avots, and M. Neumann. The architecture of the interleukin-2 promoter: a reflection of T lymphocyte activation. *Biochim Biophys Acta*, **1263** (3), 181–200, 1995.

19. M. M. Shull, I. Ormsby, A. B. Kier, S. Pawlowski, R. J. Diebold, M. Yin, *et al.* Targeted disruption of the mouse transforming growth factor-beta 1 gene results in multifocal inflammatory disease. *Nature*, **359** (6397), 693–699, 1992.

20. B. Kneitz, T. Herrmann, S. Yonehara, and A. Schimpl. Normal clonal expansion but impaired Fas-mediated cell death and anergy induction in interleukin-2-deficient mice. *Eur J Immunol*, **25** (9), 2572–2577, 1995.

21. K. W. Moore, O. G. A. R. de Waal Malefyt, P. Vieira, and T. R. Mosmann. Interleukin-10. *Annu Rev Immunol*, **11**, 165–190, 1993.

22. L. A. Timmerman, N. A. Clipstone, S. N. Ho, J. P. Northrop, and G. R. Crabtree. Rapid shuttling of NF-AT in discrimination of Ca^{2+} signals and immunosuppression. *Nature*, **383** (6603), 837–840, 1996

23. S. L. Schreiber and G. R. Crabtree. The mechanism of action of cyclosporin A and FK506. *Immunol Today*, **13** (4), 136–142, 1992.

24. K. T. Shaw, A. M. Ho, A. Raghavan, J. Kim, J. Jain, J. Park, *et al.* Immunosuppressive drugs prevent a rapid dephosphorylation of transcription factor NFAT1 in stimulated immune cells. *Proc Natl Acad Sci, USA*, **92** (24), 11205–11209, 1995.

25. P. Leder. The genetics of antibody diversity. *Sci Am.*, **246**, (5), 72–83, 1982.

26. A. Grakoul, S. K. Bromley, C. Sumen, M. M. Davis, A. S. Shaw, P. M. Allen, and M. L. Dustin. The immunological synapse: A molecular machine controlling T cell activation. *Science*, **285**, 221–227, 1999.

27. M.C. Raff. Cell surface Immunology, *Sci Amer.*, **234**, (5), 30–79, 1982.

28. D. N. Garboczi, P. Ghosh, U. Utz, Q. R. Fan, W. E. Biddison, and D. C. Wiley. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2 [comment]. *Nature*, **384** (6605), 134–141, 1996.

29. B. A. Fields, E. L. Malchiodi, H. Li, X. Ysern, C. V. Stauffacher, P. M. Schlievert, *et al.* Crystal structure of a T-cell receptor beta-chain complexed with a superantigen [see comments]. *Nature*, **384** (6605), 188–192, 1996.

30. K. C. Garcia, M. Degano, R. L. Stanfield, A. Brunmark, M. R. Jackson, P. A. Peterson, *et al.* An alpha beta T cell receptor structure at 2.5 Å and its orientation in the TCR–MHC complex [see comments]. *Science*, **274** (5285), 209–219, 1996.

31. A. Weiss. T cell antigen receptor signal transduction: a tale of tails and cytoplasmic protein-tyrosine kinases. *Cell*, **73** (2), 209–212, 1993.

32. M. Iwashima, B. A. Irving, N. S. van Oers, A. C. Chan, and A. Weiss. Sequential interactions of the TCR with two distinct cytoplasmic tyrosine kinases. *Science*, **263** (5150), 1136–1139, 1994.

33. E. Arpaia, M. Shahar, H. Dadi, A. Cohen, and C. M. Roifman. Defective T cell receptor signaling and CD8+ thymic selection in humans lacking zap-70 kinase. *Cell*, **76** (5), 947–958, 1994.

34. A. C. Chan, T. A. Kadlecek, M. E. Elder, A. H. Filipovich, W. L. Kuo, M. Iwashima, *et al.* ZAP-70 deficiency in an autosomal recessive form of severe combined immunodeficiency. *Science*, **264** (5165), 1599–1601, 1994.

35. T. Jayaraman, K. Ondrias, E. Ondriasova, and A. R. Marks. Regulation of the inositol 1,4,5-trisphosphate receptor by tyrosine phosphorylation. *Science*, **272** (5267), 1492–1494, 1996.

36. D. R. Plas, R. Johnson, J. T. Pingel, R. J. Matthews, M. Dalton, G. Roy, *et al.* Direct regulation of ZAP-70 by SHP-1 in T cell antigen receptor signaling. *Science*, **272**, 1173–1176, 1996.

37. A. L. DeFranco. Transmembrane signaling by antigen receptors of B and T lymphocytes. *Curr Opin Cell Biol*, **7** (2), 163–175, 1995.

38. D. J. Lenschow, T. L. Walunas, and J. A. Bluestone. CD28/B7 system of T cell costimulation. *Annu Rev Immunol*, **14**, 233–258, 1996.

39. M. C. Carroll. The role of complement and complement receptors in induction and regulation of immunity. *Annu Rev Immunol*, **16**, 545–568, 1998.

40. C. P. Larsen, E. T. Elwood, D. Z. Alexander, S. C. Ritchie, R. Hendrix, C. Tucker Burden, *et al.* Long-term acceptance of skin and cardiac allografts after blocking CD40 and CD28 pathways. *Nature*, **381** (6581), 434–438, 1996.

41. G. R. Crabtree and N. A. Clipstone. Signal transmission between the plasma membrane and nucleus of T lymphocytes. *Annu Rev Biochem*, **63**, 1045–1083, 1994.

42. C. N. Ting, M. C. Olson, K. P. Barton, and J. M. Leiden. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature*, **384** (6608), 474–478, 1996.

43. A. W. Goldrath and M. J. Bevan. Selecting and maintaining a diverse T-cell repertoire. *Nature*, **402**, 255–262, 1999.

44. T. Brocker. Survival of mature CD4 T lymphocytes is dependent on major histocompatibility complex class II-expressing dendritic cells. *J Exp Med*, **186** (8), 1223–1232, 1997.

45. J. Kirberg, A. Berns, and H. von Boehmer. Peripheral T cell survival requires continual ligation of the T cell receptor to major histocompatibility complex-encoded molecules. *J Exp Med*, **186** (8), 1269–1275, 1997.

46. B. Rocha. Looking out for memory T cells. *Nature*, **399**, 531–532, 1999.

    Also: J. Sprent and D. F. Tough. Lymphocyte life-span and memory. Science, 265, 1395–1400, 1994.

47. S. Cory. Regulation of lymphocyte survival by the bcl-2 gene family. *Annu Rev Immunol*, **13**, 513–543, 1995.

48. L. Van Parijs, D. A. Peterson, and A. K. Abbas. The Fas/Fas ligand pathway and Bcl-2 regulate T cell responses to model self and foreign antigens. *Immunity*, **8** (2), 265–274, 1998.

    Also: G. Nunez, D. Hockenbery, T. J. McDonnell, C. M. Sorensen, and S. J. Korsmeyer. Bcl-2 maintains B cell memory. *Nature*, **353** (6339), 71–73, 1991.

49. V. L. Perez, L. Van Parijs, A. Biuckians, X. X. Zheng, T. B. Strom, and A. K. Abbas. Induction of peripheral T cell tolerance in vivo requires CTLA-4 engagement. *Immunity*, **6** (4), 411–417, 1997.

50. L. Van Parijs and A. K. Abbas. Homeostasis and self-tolerance in the immune system: turning lymphocytes off. *Science*, **280** (5361) 243–248, 1998.

51. L. Van Parijs and A. K. Abbas. Role of Fas-mediated cell death in the regulation of immune responses. *Curr Opin Immunol*, **8** (3), 355–361, 1996.

52. E. A. Tivol, F. Borriello, A. N. Schweitzer, W. P. Lynch, J. A. Bluestone, and A. H. Sharpe. Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity*, **3** (5), 541–547, 1995.

53. T. J. Vyse and J. A. Todd. Genetic analysis of autoimmune disease. *Cell*, **85** (3), 311–318, 1996.

54. P. Denny, C. J. Lord, N. J. Hill, J. V. Goy, E. R. Levy, P. L. Podolin, *et al.* Mapping of the IDDM locus Idd3 to a 0.35-cM interval containing the interleukin-2 gene. *Diabetes*, **46** (4), 695–700, 1997.

4

Loss of regulatory control and its consequences

15

Transformation of normal cells to tumour cells

By now the reader should have gained a perception of the basic design of cellular signal transduction pathways. In the following chapters he or she will find out how helpful this information is for understanding carcinogenesis because, in principle, malignancy is caused by a failure of the genetic programme controlling cell growth, proliferation, and differentiation. Genes encoding nearly every regulatory protein are normal, constitutive, cellular proto-oncogenes. They may become oncogenes by mutation and cause cancer. They may also cause cancer when their transcription is out of control and they are overexpressed. Keeping in mind the role of proto-oncogenes in cellular regulation, it is not surprising that mutation or overexpression of the key components of signal-transduction chains cause regulatory chaos and lead to uncontrolled proliferation and de-differentiation of cells.

But even more important than the activation of oncogenes are defects in tumour-suppressor genes. This is hardly surprising, considering the central role of tumour suppressors in the control of the cell cycle and cell death. In this case, mutations lead to a loss rather than a gain of function. The fact that, in most cases, a number of mutations must accumulate to transform a normal cell to a cancer cell shows how robust the signalling networks are.

In this chapter, the emphasis is on deregulation and malfunction of cellular signalling and how it may cause cancer.[1]

Oncogenes

Overexpression or mutations of any of the many proto-oncogenes could possibly cause deregulation, uncontrolled cell proliferation, de-differentiation, and malignant transformation. Although oncogenes originating from cellular proto-oncogenes are found in 10–30% of all human cancers, expression of one, or even several, of them may not be enough to transform a normal cell into a cancer cell.

Our understanding of how oncogenes alter cellular growth and proliferation, eventually leading to malignant transformation, has gained immensely from the more and more detailed information on structure and function of the normal regulatory factors encoded by the homologous cellular proto oncogenes (Table 15.1). (There is ample information available on oncogenes in books and reviews. (See: 1–3).

Table 15.1 Some regulatory proteins that are the products of potential oncogenes, and their functions

Proto-oncogenes	Functions
PDGF	Growth factors
EGF	
CSF (colony-stimulating factor)	
PDGF receptor	Growth factor receptors with tyrosine kinase activity
EGF receptor (erb B)	
MCSF receptor (Receptor for the macrophage-colony-stimulating factor)	
Src protein kinase	Cytoplasmic, tyrosine-specific and serine/threonine-specific protein kinases
Raf protein kinase and others	
Ras proteins H-(Harvey) Ras K-(Kirsten) Ras	Monomeric GTP-binding proteins
Thyroid hormone receptor	Nuclear receptors
Steroid hormone receptors	
Myc	Nuclear transcription factors
Fos/Jun: bind to the AP1 site and other gene sites	
Rel: binds to gene-sequences related to NF-$_\kappa$ B-binding sites	

For further information see Chapter 24 of ref. 1.

Constitutive, cellular genes (proto-oncogenes) control growth of normal cells. Retroviruses acquire these cellular genes through a chance recombination with the DNA of the host cell and, in turn, can return this information by infection. Retroviruses are RNA viruses, therefore their genetic information must first be transcribed into DNA, before it is incorporated into the host cell genome. First, the viral reverse transcriptase makes a DNA copy of the viral RNA, forming a DNA/RNA hybrid. Next, a second DNA strand is formed, and the RNA strand is degraded and replaced by the DNA strand, giving a DNA double helix. The DNA double helix is integrated into the host chromosome by a viral integrase, Thus, the viral genome is integrated into the host cell chromosome and is replicated by the host cell's DNA-dependent RNA polymerase. As new retroviruses are formed and released from the cell, they carry with their viral genome a part of the cellular genome which they can transfer to other cells.

Since many cellular oncogenes were originally identified in transforming retroviruses, and since some retroviruses cause rapid malignant transformation in animals, notably chicken, it seemed for a while that most human cancers are caused by infection with retroviruses. But that is not the case. The number of human cancers which originate from retroviral infections is actually rather small (Table 15.2). In many cases, proto-oncogenes can become oncogenes by mutation or overexpression in human cells, without retroviral assistance.

Table 15.2 Human cancers originating from retroviral infection

Oncogene	Source of virus	Function of the cellular proto-oncogene	Virus-induced tumour
abl	Mouse, cat	Protein tyrosine kinase	B-cell leukaemia; sarcoma.
erb-B	Chicken	Protein tyrosine kinase, truncated EGF-receptor	Erythroleukaemia; fibrosarcoma
fms	Cat	Protein tyrosine kinase, receptor for the macrophage colony-stimulating factor, (MCSF)	Sarcoma
kit	Cat	Protein tyrosine kinase, receptor for the Steel-factor	Sarcoma
src	Chicken	Cytosolic, protein tyrosine kinase	Sarcoma
raf	Chicken/mouse	Protein serine/threonine kinase, (Ras-dependent)	Sarcoma
sis	Monkey	B chain of PDGF, (platelet-derived growth factor)	Sarcoma
fos/jun	Mouse/chicken	Combine and form the AP1 transcriptional regulating complex	Fibrosarcoma; osteosarcoma
myc	Chicken	Transcriptional HLH-type regulator	Sarcoma; carcinoma.
H-*ras*	Rat	Monomeric p21 G-protein	Sarcoma
K-*ras*	Rat	Monomeric p 21 G-protein.	Erythroleukaemia

Information from Table 24-4 of ref. 1.

Table 15.2 lists some oncogenes found in transforming retroviruses and the function of the corresponding cellular proto-oncogenes. The difference between, for example, a v-*src* oncogene and its cellular counterpart, the c-*src* cellular proto-oncogene, which is not oncogenic, is that the introns present in the cellular c-*src* have been removed in the v-*src* oncogene. Moreover, the v-*src* oncogene encodes a tyrosine kinase which carries mutations in the kinase domain that deregulate the src-tyrosine kinase and make it overactive.

Cellular regulatory proteins with oncogenic potential

Every proto-oncogene encoding a normal constituent of a signalling pathway can become an oncogene when mutated or overexpressed. Only a few examples are given, because the properties and functions of the normal cellular proteins encoded by proto-oncogenes have already been discussed.

Growth factors

When a cell bearing the corresponding receptor produces the matching growth factor, it can be stimulated through an autocrine activation loop. Autocrine self-stimulation by growth factors promotes rapid cell proliferation. This may set the stage for malignant transformation, because in a rapidly proliferating cell population, the chances of spontaneous mutations and errors in DNA replication increase. Thus overexpression of growth factors may cause malignancies in birds and in mammals, including humans. For example, when a normal human gene/proto-oncogene, for example the c-*sis* gene that encodes

the B-chain of PDGF, is overexpressed in cultured mouse fibroblasts, the cells become malignant. The same is the case when glial cells, which normally do not express PDGF, acquire the genetic information and express this growth factor (ectopic expression). The glial cells are then transformed to glioblastomas, a viciously malignant glial cell tumour. The association of uncontrolled expression of the growth factor FGF with human stomach cancer and Kaposi sarcoma is an other example.

Growth-factor receptors

Genes encoding growth-factor receptors are often overexpressed in cancer cells. An example are the genes for members of the EGF receptor family. The c-*neu* gene and the c-*erb-B1* gene are overexpressed in up to 80% of squamous cell carcinomas. v-*erb-B* is the retroviral homologue of c-*erb-B*. The avian Erb virus, erythroblastosis virus, transmits its v-*erb-B* gene to chicken by infection and causes erythroleukaemia. The v-*erb-B* oncogene encodes a truncated EGF receptor that is deregulated and constitutively active when expressed. Transgenic mice expressing the product of a mutated *erb-B* gene, with a replacement of a valine by a glutamic acid, developed adenocarcinomas of the breast. Another mutant form of a receptor that is constitutively active, is the receptor for the macrophage colony stimulating factor (MCSF-1), which is expressed in myeloid leukaemia cells.

Cytosolic regulatory factors and linkers

Figure 15.1 shows, in a simplified fashion, that deregulation of any key regulatory factor, such as the enzymes controlling phosphorylation dephosphorylation in the cytosol or the proteins controlling gene transcription in the nucleus, can contribute to cancer.

The role of cytosolic tyrosine kinases in the pathogenesis of human leukaemia has made them a target for drugs. The so-called 'Tyrphostins', developed by A. Levitzki and

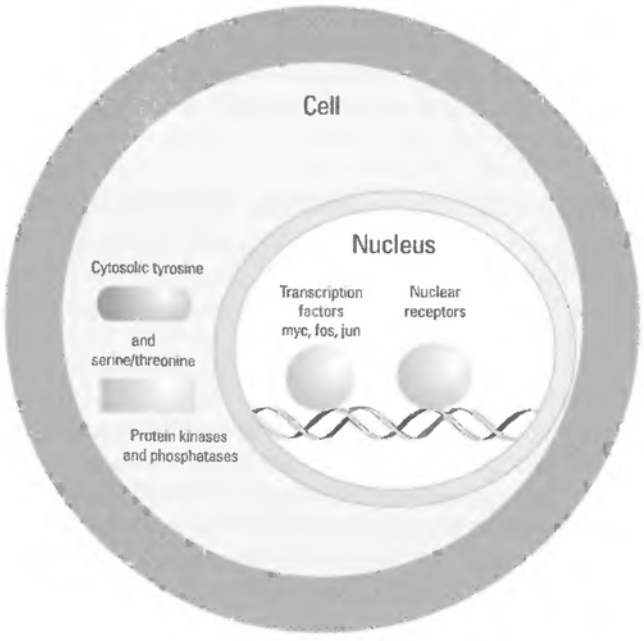


Fig. 15.1 Bearing in mind the central role of phosphorylation–dephosphorylation reactions in the control of signalling and in the regulation of the cell cycle and the role of transcription factors controlling gene expression in growth and proliferation, it is hardly surprising that cancer-causing oncogenes are mainly cytosolic phosphotyrosine and phosphoserine/threonine protein kinases and phosphatases and transcriptional regulators.

colleagues,[4] have been found to be effective in the treatment of certain forms of leukaemia.

But non-enzymic components of regulatory pathways also play a role in oncogenesis. An example of a proto-oncogene that encodes a non-enzymic linker is the CBL protein, which is encoded by the c-*cbl* proto-oncogene, the cellular homologue of the murine retroviral v-*cbl* oncogene of the Cas NS-1 retrovirus. The removal of a leucine-zipper-like domain makes v-*cbl* a transforming oncogene.[5–6] The Cbl protein has an evolutionarily conserved N-terminal region, Cbl-N, which is a tyrosine-kinase-binding domain (TKB).[7] The TKB domain of Cbl is responsible for the downregulation of cytosolic and receptor tyrosine kinases. On binding to RTKs, Cbl initiates receptor downregulation, removal from the cell surface, and transfer to intracellular vesicles, as a prelude to ubiquitination and eventual proteolytic degradation.[7,8]

G proteins

Prominent representatives of cellular proto-oncogenes are members of the *ras* gene family.[9] Their viral homologues are the avian Harvey (H) and Kirsten (K) *ras* oncogenes. Many human cancers contain sequences in their DNA that are homologous to viral *ras* oncogenes. Mutated p21 Ras G proteins have a prolonged active state. A permanent GTP-bound 'on state', whatever its cause, results in uncontrolled growth, eventually leading to malignant transformation. Not only overexpression and mutations of the *ras* proto-oncogene, but also mutations and overexpression of any of the many regulatory factors controlling the 'on and off' switch of the Ras-GTPases are implicated in tumorigenesis (see Chapter 3). This has made Ras and factors regulating the GTPase activity of Ras and related small, monomeric G proteins, prime targets for anticancer drugs.

Mutant forms of the α-subunit of the heterotrimeric G protein G_s, which are in the GTP-bound 'on' state and constitutively active, have been found in pituitary tumours and in pseudohypoparathyroidism, a disease characterized, among other things, by resistance to parathyroid hormone (see Chapter 5).[10]

Transcription factors

Many transcription factors involved in the control of growth and the cell cycle are encoded by proto-oncogenes.[11–12] However, the complexity of the regulation of gene expression makes it difficult to make a convincing case for a cause–effect relationship of a change in the activity of any single transcriptional regulator (see Chapter 10).

How proto-oncogenes become cancer genes

Proto-oncogenes become oncogenes:

(1) by mutation;

(2) by gene amplification and overexpression;

(3) by chromosomal translocation and gene rearrangement;

see Fig. 15.2. Since these mechanisms are described elsewhere in textbooks, we only mention a few examples, relevant to cellular signalling.

Mutational activation

Ras, the first oncogene found to be associated with a human tumour, an urinary bladder cancer, had a point mutation, changing a valine to a glycine. Other point mutations in *ras*

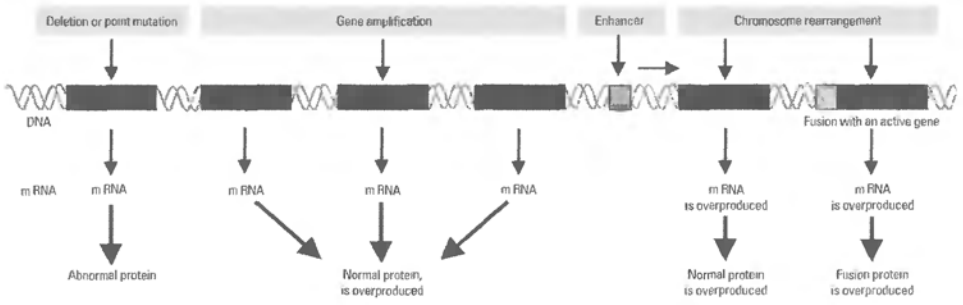


Fig. 15.2 Common mechanisms for transformation of proto-oncogenes to cancer genes. A fourth mechanism, recombination between a cellular DNA sequence and a retroviral gene is not shown, because the consequences are like those of chromosomal rearrangement. (See also Fig. 24–27 in ref. 1, reproduced with permission of Taylor and Francis, Inc.).

have been found in colorectal cancers and adenocarcinomas of the lung and other human cancers. But the presence of *ras* mutants is not of much predictive value, because they are extremely common and cells may live happily with a mutated Ras, because there are many ways to bypass a dysfunctional Ras.

Activation by chromosomal translocation, gene rearrangement, and gene fusion

A classic example of a chromosomal translocation of a gene and its consequence is Burkitt's lymphoma. In 75% of patients with Burkitt's lymphoma, the c-*myc* proto-oncogene, encoding the transcription factor c-Myc, is translocated from its normal location on chromosome 8 to a site on chromosome 14. This translocation positions the c-*myc* gene adjacent to the very active promoter-enhancer of the immunoglobulin heavy chains. This chromosomal translocation is responsible for the overproduction of normal c-Myc. High cellular concentrations of c-Myc promote formation of Myc homodimers, thus preventing oligomerization with partner proteins such as the Myc-associated proteins Max and MAD. The failure to form active Myc complexes impairs the transcriptional activity of Myc and leads to an increased responsiveness of the cell to mitotic signals. In B lymphocytes this causes uncontrolled proliferation and eventually formation of neoplastic monoclonal B-cells. Persistent overexpression of the c-*myc* gene has also been found in neuroblastoma, small cell cancers of the lung, breast and colon cancer, and leukaemia.

Another event causing overexpression of transcription factors is juxtaposition of these genes to a gene locus for antigen receptors in T lymphocytes. Nearly half of the genes that are translocated to antigen receptor loci in human leukaemia cells, notably in T-cell acute lymphoblastic leukaemia (T-ALL), encode helix-loop-helix (HLH) transcription factors (such as Max and MAD) (see Chapter 10 and 17).

Genes encoding homeodomain transcriptional regulators are also activated by translocation. In T-ALL, the gene for the Hox-11 transcription factor is expressed ectopically in normal, terminally differentiated effector T-cells. When the Hox-11 gene is expressed,

the T cells revert to an immature cell type, where a set of genes, which are normally dormant in mature T cells are expressed, causing malignant transformation. This reversion to an early developmental programme is quite often the cause of cancer.

Finally, gene deletion and replacement by another gene can have the same effect as chromosomal translocation, leading to the expression of normally dormant genes. For example in T-ALL, the control regions of the *tal-1* gene, which is silent in normal T cells, are replaced with another regulatory gene sequence, that of the *sil* proto-oncogene, which is normally expressed in T cells. This leads to the ectopic expression of the *tal-1* gene in T cells.

Gene fusion

When the 'breakpoints' where a gene is excised are located in introns, the excised gene, when it is translocated, combines more readily with an adjacent gene (introns are non-transcribed DNA sequences). Such fused genes may code for chimeric proteins. The fusion of the *bcr–abl* genes is an example. The *bcr* gene (bcr stands for breakpoint cluster region) codes for a serine/threonine kinase and the c-*abl* gene for a tyrosine kinase (see Chapter 3). The c-*abl* gene is activated when translocated from chromosome 9 to the brc of chromosome 22, where it is fused with the *bcr* gene. The fused Bcr–Abl protein now has a tyrosine kinase activity which is more active than that of the original Abl kinase. The fusion of the *bcr–abl* genes is known as the 'Philadelphia chromosome translocation'. It is associated with leukaemia (Fig. 15.3).

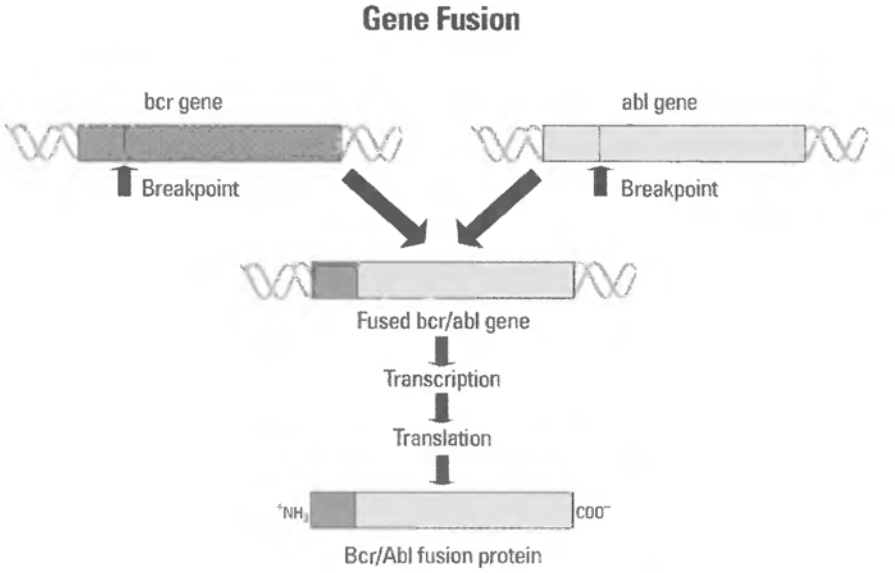


Fig. 15.3 The fusion protein has the amino-terminal part of the Bcr kinase and the carboxy-terminal part of the Abl tyrosine kinase. The result is that the Abl kinase in the fused protein is hyperactive and drives proliferation of haematopoietic cells in the bone marrow. (See also Figs 24–25 of ref. 31, reproduced with permission of Taylor and Francis, Inc.)

Other examples are fusions of the *e2a* gene, which encodes the E2A transcription factor, a cell-cycle-regulating factor. The *e2a* gene fuses with genes that encode other transcription factors, such as HLF (hepatocyte leukaemia factor). This causes overexpression of the chimeric transcription factor. Juxtaposition of the *trans*-activation domains of E2A and HLF in the E2A–HLF fusion protein results in mismatched heterodimerization and in competition between dimerization partners of E2A–HLF. These gene fusions are the most common changes in the chromosomes of children with acute lymphoblastic leukaemia (ALL).

Gene amplification

Gene amplification manifests itself in homogeneous DNA-staining regions and abnormal banding patterns of chromosomes. Homogeneous staining regions, all representing c-*myc* genes, have been found in several tumours, including neuroblastomas, where they predict rapid progression and give a poor prognosis. Gene amplification and overexpression of the *erb-B* gene has been found in about a third of all breast and ovarian cancers.

Tumour-suppressor genes

In 1969, Henry Harris, George Klein, and colleagues, reported that cancerous cells, which could produce tumours in animals, lost their tumorigenicity when they were fused with non-cancerous normal cells. The loss of tumorigenicity was then perpetuated from one generation to the next. This experiment was a landmark in cancer research. Many experiments that followed demonstrated that this condition applied to all kinds of cancers—virally induced, chemically induced, and spontaneous tumours—and to a variety of cell types—epithelial cells, fibroblasts, and lymphocytes.

Thus it seems that the normal, healthy cell contributed one or more genes to tumour cells, which code for protein(s) that are missing in the tumour cell and which suppress their tumorigenicity. The genes transduced from normal cells appeared to restore a defective function in the cancer cell, making the cancer cell again responsive to the regulatory constraints in normal cells. These genes are the 'tumour-suppressor genes'. They are inactivated when a normal cell is transformed to a malignant cell.

Thus, a characteristic property of a tumour cell is a loss of tumour-suppressor genes. Tumour-suppressor genes and oncogenes have in common that both code for proteins that normally function in the regulation of orderly cell growth. Tumour suppressors are transcriptional regulators.

Tumour-suppressor genes are more and more being incriminated in human cancers. Usually, both alleles of a tumour-suppressor gene must be deleted or mutated to make the tumour suppressor non-functional. (For more information we refer to periodicals[13] and ref. 14.)

The *RB* gene

For a rare (1 in 30 000) inherited human childhood tumour, retinoblastoma, a defective tumour suppressor gene is responsible, the *RB* gene (defects in the *RB* gene locus occur also in 70% of all osteosarcomas; the *RB* gene is located on the human chromosome 13q14).[15]

In a patient with retinoblastoma, the *RB* gene is either mutated in both alleles or, when only one allele is mutated, the other allelic gene is deleted. This occurs quite often in

Table 15.3 Ways in which the *RB* tumour-suppressor gene is lost and the consequences

Changes of *RB* gene	Consequences
Loss of only one copy of the *RB* gene	Healthy cell
One copy is mutated and the normal healthy copy is deleted	Loss of function of the tumour suppressor
Point mutations in both alleles of the *RB* gene	Loss of function of the tumour suppressor
Elimination of the healthy *RB* gene copy together with adjacent region in the chromosome	Loss of function of the tumour suppressor
Defect in recombination of maternal and paternal allelic genes	Loss of heterozygosity Both alleles are defective

Information taken from ref. 16.

tumour-suppressor genes. A cell that has lost only one of its two copies of the *RB* tumour-suppressor gene is usually normal and healthy. But there are several ways in which a cell can lose the other, healthy copy of a tumour-suppressor gene. For example, the allelic copy may be lost by a large chromosomal deletion or it may be inactivated by a point mutation, just like the first copy. Or, the healthy copy may be replaced by recombination with the defective copy in the course of meiosis. The result is loss of heterozygosity (meaning that both alleles are defective);[17] this decides the fate of a patient carrying defective *RB* genes. Loss of heterozygosity is characteristic for cancer and is used to identify and clone tumour-suppressor genes. Table 15.3 summarizes the ways in which the *RB* tumour-suppressor gene is lost.

The Rb protein

The Rb proteins (pRBs) are transcriptional regulators, controlling genes for the cell-cycle machinery. pRB senses and integrates a multitude of proliferative and antiproliferative signals and regulates the genes that allow progression from the G_1 state to the S phase. The tumour-suppressor function of pRB and pRB-like proteins manifests itself in growth inhibition due to arrest of the cell cycle in G_1.[18] Thus, the Rb protein watches over the commitment of a cell to enter the DNA-synthetic S phase of the cycle (Chapter 12).[19]

The *RB* gene encodes a transcriptional co-repressor which forms a complex with the cell-cycle-regulating transcription factor E2F (members of the E2F family of transcriptional activators control the genes involved in the G_1/S transition and in DNA replication;[20] see Chapter 12). The RB–E2F transcriptional complex contains additional factors, such as the histone deacetylase, HDAC1, which facilitates access of the transcriptional complex to E2F-responsive gene promoters, such as the cyclin E promoter. pRB and HDAC1 act in concert and repress the cyclin E promoter.[21] Hence the retinoblastoma protein (pRB) is a tumour-suppressor that represses gene expression in concert with a histone deacetylase, by modulating the architecture of chromatin (Fig. 15.4). The pRB–E2F transcriptional complex also represses transcription of RNA polymerase III[22] (see also Chapter 9).

pRB has also been implicated in TGF-β-mediated growth inhibition (Chapter 6). Without a functional pRB, TGF-β is no longer capable of shutting off c-Myc-controlled transcription of genes which promote cell growth. Control of transcription by TGF-β involves the common SMAD 4, which acts like a tumour suppressor.[23] SMAD 4 is also known as DPC4 ('deleted in pancreatic carcinoma') because it is non-functional in nearly

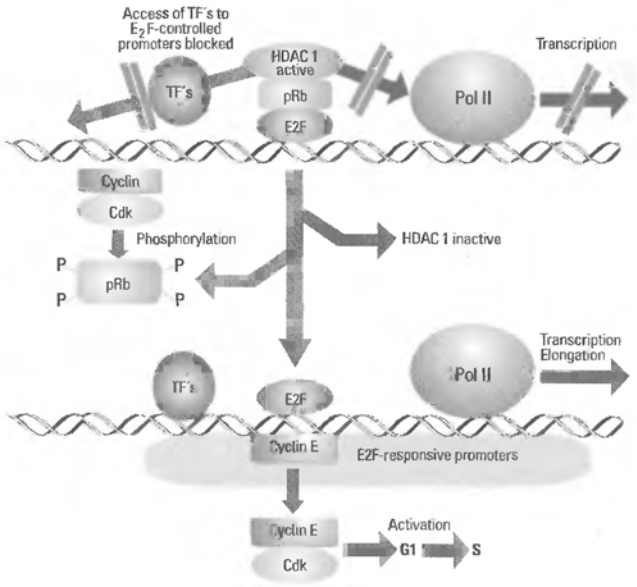


Fig. 15.4 Brehm *et al.*[25] and Magnaghi-Jaulin *et al.*[26] have found that RB forms a complex with the E2F protein and the histone deacetylase HDAC1. The complex represses expression of E2F-controlled promoters, such as the cyclin E promoter. HDAC1 may facilitate the removal of highly charged acetyl groups from core histones, causing a tighter association of DNA in nucleosomes, preventing transcription factors (TFs) from gaining access to the DNA. The repression is then released in G_1 in response to proliferative signals by phosphorylation of RB by cyclin-dependent kinases (Cdks). Phosphorylation dissociates RB from the complex, allowing transcription of E2F-responsive promoters. (The information for this figure comes from Fig. 1 of ref. 27 with permission of the author and Nature.)

half of all pancreatic carcinomas and to a lesser extent in a variety of other cancers. As the tumour progresses, more and more functional SMAD 4 is lost. Restoration of SMAD 4 by transfection re-establishes growth inhibition by TGF-β[24] (see Chapters 6 and 10).

pRB is regulated by phosphorylation, which is synchronized with the cell cycle. There are more than half a dozen serine/threonine phosphorylation sites in pRB. The phosphorylation state of pRB changes with cell-cycle transitions and in response to proliferative signals. The cyclin-dependent kinases Cdk4 and Cdk6 are activated in G_1 and phosphorylate the RB component of the RB/E2F transcriptional complex. At mid-to-late G_1, RB becomes hyperphosphorylated and is subsequently released from the promoter-bound E2F, freeing E2F-regulated genes from the constraint of the suppressor pRB and allowing expression of genes that are essential for cell-cycle progression. This commits cells irreversibly to enter the S phase.

A model of regulation of pRB was proposed by Robert Weinberg. The Weinberg model postulates that in a normal cell, pRB is dephosphorylated after mitosis. Dephosphorylated pRB interacts with transcriptional regulators and suppresses growth.[18] Growth inhibition lasts until growth factors order the cell to proceed through G_1, when cyclin-dependent kinases (Cdks) are activated which re-phosphorylate pRB, removing the blockade of growth. But in rapidly growing tumour cells, pRB is underphosphorylated, as in quiescent normal cells. To rationalize this paradoxical situation, Robert Weinberg has reasoned that mutated pRB in tumour cells is no longer phosphorylatable; even though pRB is properly underphosphorylated, the mutation has changed the structure of the tumour suppressor in a way that the mutated pRB has lost control over the cell cycle.

To sum up:

1. The tumour suppressor, pRB, is a transcriptional co-regulator (co-repressor), the activity of which is regulated by phosphorylation. Phosphorylation is by cyclin-dependent kinases, thus linking pRB to the cell cycle.

2. pRB dampens growth by blocking expression of genes which promote cell-cycle progression and proliferation. When pRB is not functioning, expression of these genes is de-repressed, causing a situation quite like that when proto-oncogenes are overexpressed or activated by mutation. In each case, the consequence is uncontrolled cell proliferation, driving the cells into their eventual transformation to cancer cells. When pRB does not function properly, the histone deacetylase recruited by pRB cannot do its job and the histone acetylase (HAC) takes over. This leads to the adverse change. HAC locally destabilizes the chromatin and deregulates and enhances expression of critical E2F-regulated genes, such as the cyclin E gene. Expression of cyclin E results in activation of Cdk, which in turn phosphorylates the linker histone, H_1.[28] Linker histones are involved in packaging of chromatin. Accordingly, phosphorylation of H_1 helps to open the chromatin template. This facilitates DNA replication and eventually cell proliferation.[29]

3. Because of its role in cell-cycle control and proliferation, pRB is an attractive target for agents designed to enforce cell-cycle checkpoint controls that block malignant transformation.

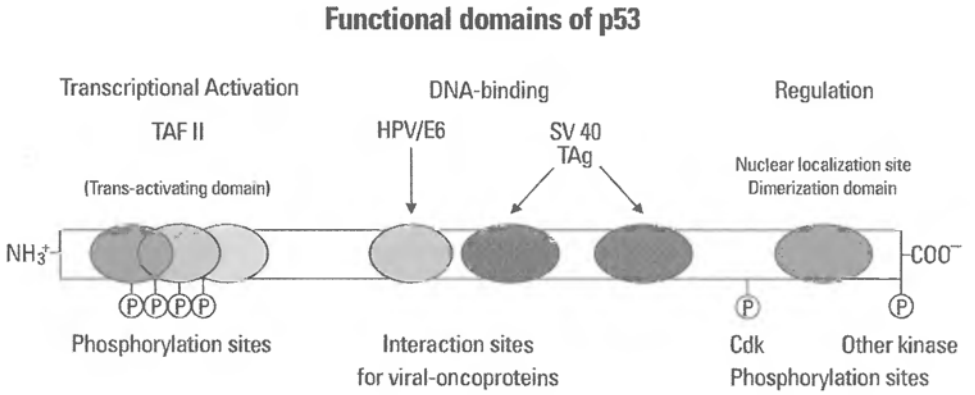


Fig. 15.5 The NH_2-terminal activation domain contains phosphorylation sites for serine/threonine kinases, among them a DNA-dependent serine/threonine kinase. This domain interacts with TAFII, a component of TFIID in the PIC, the transcriptional pre-initiation complex (see Chapter 9). The central DNA-binding domain is the port of entry for viruses. p53 interacts with viral oncoproteins, such as the large T-antigen (TAg) of the SV40 polyoma virus, and the oncoprotein, E6, of the human papilloma virus (HPV). The COOH-terminal regulatory domain contains the phosphorylation sites for cyclin-dependent Cdk and other kinases. This domain also contains a nuclear translocation signal (NLS) and a dimerization site.[30] (The information for this figure is from Abb. 15.7, p. 444 in ref. 33, with permission of Professor Gerhard Krauss and Wiley-VCH.)

The tumour suppressor p53

The tumour suppressor p53 is a very versatile transcriptional regulator. Among the many genes controlled by p53 are cyclin genes, genes for an inhibitor of cyclin-dependent kinases, and the *bax* gene, which promotes apoptosis. This repertoire of genes controlled by p53 links this tumour suppressor to cell-cycle control and apoptosis[30,31] (see Chapters 12 and 13).

p53 has three functional domains: an NH_2-terminal transcriptional activation domain, a central DNA-binding domain, and a COOH-terminal, regulatory domain (Fig. 15.5). A 2.2 Å crystal structure of a p53–DNA complex is available (Plate 30, ref. 34).

Mutations of the p53 suppressor gene are the most frequent suppressor gene mutations in human cancers

Mutations of the p53 gene are the most common suppressor gene mutations in human cancers.[35] Mutations in p53 or in proteins that regulate p53 have been found in more than 80% of all human tumours. Whereas mutations in other tumour-suppressor genes, such as in *RB* and *APC* genes, are mostly nonsense mutations that cause truncation or instability of the expressed protein, more than 90% of the p53 gene mutations are either missense or point mutations, the latter leading to changes of only one amino acid. The high frequency of mutations of p53, each affecting a different amino acid, explains the great variety of human cancers carrying such mutations. The mutations and accompany-

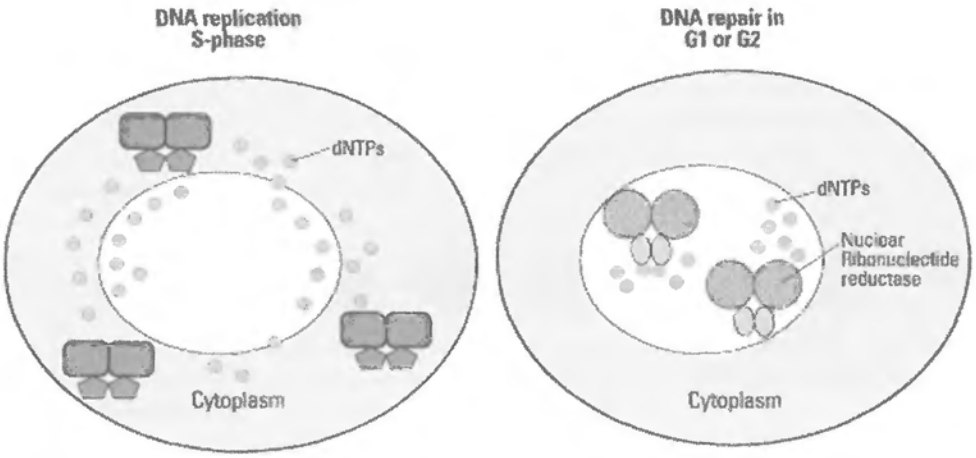


Fig. 15.6 Differential localization of ribonucleotide reductase (RNR) during DNA replication and repair. *Left:* During the S phase of the cell cycle, the cytoplasmically localized RNR synthesizes deoxyribonucleotide triphosphates (dRNTPs) in the cytoplasm, which diffuse into the nucleus for DNA replication. The larger and the smaller subunits are the regulatory, R1, and the catalytic, R2, subunits of the normal RNR tetramer. *Right:* DNA repair in a cell in the G_1 or G_2 phase is shown. Tanaka *et al.*[36] show that, in the presence of DNA damage, the *p53R2* gene is expressed. Nuclear p53R2 (together with an unknown large subunit protein, probably containing the missing subunits of the nuclear RNR complex), then produce dRNTPs in the nucleus for DNA repair. (Reproduced with permission of the authors and Nature from Fig. 1 in ref. 37.)

ing conformational changes may increase or decrease the stability of p53, affect sequence-specific DNA binding, consequently altering its activity as a transcriptional regulator. Deregulation of p53 is accompanied by changes in the phosphorylation pattern by cell-cycle kinases that operate in the G_1 phase. Therefore the function of p53 is directly linked to cell-cycle regulation.

An essential function of the tumour suppressor p53 is to prevent propagation of cells with damaged DNA. In that way p53 suppresses tumour formation. Cells with damaged DNA cannot enter the S phase. They are arrested in the G_1 phase. This gives cells time for DNA repair. p53 may also keep them arrested permanently, preventing proliferation of damaged cells.

But p53 also plays a role in the repair of damaged DNA. Among p53-regulated genes is a gene, *p53R2*, which encodes the catalytic subunit of a nuclear form of the enzyme ribonucleotide reductase (RNR). RNR catalyses the rate-limiting step in the synthesis of deoxyribonucleotide triphosphates (dRNTPs), the building blocks both for DNA replication and repair. RNR is an interesting enzyme, identified and characterized in Peter Reichard's laboratory in Stockholm.[38] It is a tetramer with two different subunits. The larger one is the regulatory subunit, R1, the smaller one is the catalytic subunit, R2 (R2 contains an iron centre and a tyrosyl free radical that is essential for catalysis). Not only is p53 regulated by the cell cycle, but also the mammalian ribonucleotide reductase. The catalytic subunit is synthesized before DNA replication starts and it disappears in late S phase or early G_2 phase, after DNA has been synthesized. Whereas the concentration of the catalytic R2 is controlled by the cell cycle, the regulatory R1 subunit does not change during the cell cycle.

The *p53R2* gene is expressed in response to DNA damage and the p53R2 protein takes part in DNA repair. p53R2 is like a normal R2, but it differs in one important aspect, namely in its location. Whereas the normal RNR2 is in the cytoplasm,[39] where it synthesizes dRNTPs that then diffuse into the nucleus, p53R2 is located in the nucleus (Fig. 15.6). This led Tanaka *et al.*[36] to speculate that for DNA replication in the S phase, the cytoplasmically synthesized dRNTPs may suffice, whereas DNA repair in G_2 may need a supply of more concentrated dRNTPs, supplied by p53 RNR2 in the nucleus, right at the DNA sites, which must be repaired.

The DNA damage-response pathway that controls induction of *p53R2* genes is conserved in yeast and humans and is regulated by a phosphorylation cascade (Fig. 15.7).[40]

In cancer, the genes for p53R2 and other components of the DNA damage-repair pathway are often mutated. Cells lacking p53R2 are more affected by DNA damage than

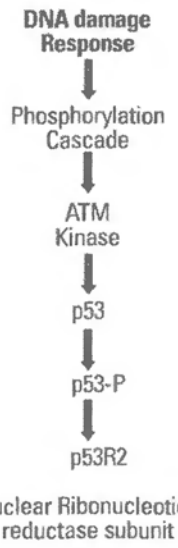


Fig. 15.7 The DNA damage-response pathway. DNA damage is sensed by protein kinases. In mammals and in yeast the activation of the ATM kinase and of the DUN1 kinase, respectively, is controlled by a phosphorylation cascade, responsive to DNA-damage signals. (ATM is a kinase and the product of the *AT* gene, the ataxia telangiectasia gene). p53 is phosphorylated and in humans and yeast p53-P removes the block at promoters of RNR genes and the nuclear *p53R2* is expressed. (Reproduced from Fig. 2 in ref. 37 with permission of the authors and Nature.)

cells with functional p53R2. The remaining, residual repair capability of cells lacking nuclear p53R2 was attributed to the remaining supply of dRNTPs, produced by the normal cytoplasmic ribonucleotide reductase. Therefore, the possibility was considered that inhibitors, specific for the cytoplasmic form of the RNR enzyme, might sensitize cancer cells, carrying already non-functional nuclear p53R2 mutants, to DNA-damaging chemotherapeutic agents, because these cells would now no longer have DNA repair capabilities.

The importance of the repair of damaged DNA in keeping the cells alive and preventing cancer is also apparent from mutations in genes involved in nucleotide excision repair. The enzymes involved in nucleotide excision repair are other tools that cells have to repair mismatched DNA. Defects in DNA nucleotide excision repair can cause skin cancer (xeroderma pigmentosa) and colon cancer.

References

1. For general information, see:

   B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell.* 3rd ed. Garland Publishing, New York, 1994.

   Also: H. Lodish, D. Baltimore, A. Berk, S. L. Zipursky, P. Matsudaira, and J. Darnell (ed.). *Molecular cell biology*, (3rd edn). W.H. Freeman, New York, 1995.
2. R. Hesketh. *The oncogene handbook.* Academic Press, London, 1994.
3. E. Rubin and L. Farber (ed.) *Pathology*, (2nd edn). Lippincot, Philadelphia, 1993.

   Also: Robbins Pathologic Basis of Disease by Ramzi S. Cotran, Vinay Kumar, Tucker Collins and Stanley L. Robbins, 6th edition, W B Saunders Co, Philadelphia, 1999.

   Also: L. M. Franks and N. M. Teich (eds). *Introduction to the cellular and molecular biology of cancer*, (3rd edn). Oxford University Press, Oxford, 2000.
4. A. Levitzki and A. Gazit. Tyrosine kinase inhibition: An approach to drug development. *Science*, **267**, 1782–1788, 1995.
5. T. J. Blake, M. Shapiro, H. C. Morse, III, and W. Y. Langdon. The sequence of the human and mouse c-cbl proto-oncogenes show v-cbl was generated by a large truncation encompassing a proline-rich domain and a leucine zipper-like motif. *Oncogene*, **6**, (4) 653–657, 1991.
6. W. Meng, S. Sawasdikosol, S. J. Burskoff, and M. J. Eck. Structure of the amino-terminal domain of Cbl complexed to its binding site on ZAP-70 kinase. *Nature*, **398**, (6722), 84–90, 1999.
7. P. S. Lee, Y. Wang, M. G. Dominguez, Y.-G. Yeung, M. A. Murphy, D. D. Bowtell, and E. R. Stanley. The Cbl protooncoprotein stimulates CSF-1 receptor multiubiquitination and endocytosis, and attenuates macrophage proliferation. *EMBO J*, **18**, (13), 3616–3628, (1999).
8. G. Levkowitz, H. Waterman, E. Zamir, Z. Kam, S. Oved, W. Y. Langdon, L. Beguinot-Geiger, and Y. B. Yarden. c-Cbl/Sli-1 regulates endocytic sorting and ubiquitination of the epidermal growth factor receptor. *Genes Dev*,**12**, (23), 3663–3674, 1998.
9. J. L. Bos. The *ras* family and human cancerogenesis. *Mutation Res*, **195**, (3) 255–271, 1988.
10. C. A. Landis, S. B. Masters, A. Spada, A. M. Pace, H. R. Bourne, and L. Vallar. GTPase inhibiting mutations activate the α-chain of Gs and stimulate adenylyl cyclase in human pituitary tumours. *Nature*, **340**, 292–296, 1989.
11. M. L. Cleary. Transcription factors in human leukaemias. Oncogenes in the development of leukaemia. *Imperial Cancer Research Fund. Cancer surveys*, **15**, 89–104, 1992.
12. D. S. Latdman. *Eukaryotic transcription factors. Chapter 7: Transcription factors and cancer*, pp. 153–176. Academic Press, London, 1991.
13. A. Levine (ed.) Tumour suppressor genes, the cell cycle and cancer. *Cancer Surveys*, **12**. Cold Spring Harbor Press, New York, 1992.

    Also: G. F. Vande Woude and G. Klein eds. Advances in Cancer Research, Academic Press, San Diego, California. This classic and essential series presents critical overviews on select aspects of both cancer research and the basic underlying sciences.

14. D. Haber and E. Harlow. Tumour-suppressor genes: evolving definitions in the genomic age [news]. *Nature Genet.* **16** (4), 320–322, 1997.

15. U. Francke. Retinoblastoma and chromosome 13. *Cytogenet Cell Genet*, **16** (1–5), 131–134, 1976.

16. W. K. Cavenee, T. P. Dryja, R. A. Phillips, W. F. Benedict, R. Godbout, B. L. Gallie, A. L. Murphree, L. C. Strong, and R. L. White. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature*, **305**, (5937), 779–784, 1983.

17. A. G. Knudson Jr. Mutation and cancer: Statistical study of retinoblastoma. *Proc Natl Acad Sci, USA*, **68** (4), 820–823, 1971.

18. R. A. Weinberg. The retinoblastoma protein and cell cycle control. *Cell*, **81**, 323–330, 1995.

19. W. R. Sellers, J. W. Rodgers, and W. G. Kaelin Jr. A potent transrepression domain in the retinoblastoma protein induces a cell cycle arrest when bound to E2F sites. *Proc Natl Acad Sci, USA*, **92** (25), 11544–11548, 1995.

20. J. R. Nevins. E2F: a link between the Rb tumor suppressor protein and viral oncoproteins. *Science*, **258** (5081), 424–429, 1992.

21. A. P. Wolffe. Histone deacetylase: a regulator of transcription [comment]. *Science*, **272** (5260), 371–372, 1996.

22. R. White, D. Trouche, K. Martin, S. Jackson, and T. Kouzarides. Repression of nRNA polymerase III transcription by retinoblastoma protein. *Nature*, **382**, 88–90, 1996.

23. Y. Zhang, X.-H. Feng, R.-Y. Wu, and R. Derynck. Receptor-associated Mad homologues synergize as effectors of the TGF-β response. *Nature*, **383**, 169–172, 1996.

24. M. P. de Caestecker, P. Hemmati, S. Larisch Bloch, R. Ajmera, A. B. Roberts, and R. J. Lechleider. Characterization of functional domains within Smad4/DPC4. *J Biol Chem*, **272** (21), 13690–13696, 1997.

25. A. Brehm, E. A. Miska, D. J. McCance, J. L. Reid, A. J. Bannister and T. Kouzarides. Retinoblastoma protein recruits histone deacetylase to repress transcription [see comments]. *Nature*, **391** (6667), 597–601, 1998.

26. L. Magnaghi Jaulin, R. Groisman, I. Naguibneva, P. Robin, S. Lorain, J. P. Le Villain, *et al.* Retinoblastoma protein represses transcription by recruiting a histone deacetylase [see comments]. *Nature*, **391** (6667), 601–605, 1998.

27. R. A. De Pinho. The cancer–chromatin connection. *Nature*, **391** (6667), 533–536, 1998.

28. R. E. Herrera, F. Chen, and R. A. Weinberg. Increased histone H1 phosphorylation and relaxed chromatin structure in Rb-deficient fibroblasts. *Proc Natl Acad Sci. USA*, **93** (21), 11510–11515, 1996.

29. H. Matsushime, D. E. Quelle, S. A. Shurtleff, M. Shibuya, C. J. Sherr, and J. Y. Kato. D-type cyclin-dependent kinase activity in mammalian cells. *Mol Cell Biol*, **14** (3), 2066–2076, 1994.

30. S. P. Linke, K. C. Clarkin, and G. M. Wahl. p53 mediates permanent arrest over multiple cell cycles in rresponse to gamma-irradiation. *Cancer Res.*, **57**, (6), 1171–1179, 1997.

31. T. M. Gottlieb, and M. Oren. p53 and apoptosis. *Semin Cancer Biol.*, **8**, (5), 359–368, 1998.

32. L. J. Ko and C. Prives. p53: puzzle and paradigm. *Genes Devel*, **10**, 1054–1072, 1996.

33. G. Krauss. *Biochemie der Regulation und Signaltransduktion*. Wiley-VCH, 1997.

34. Y. Cho, S. Gorina, P. D. Jeffrey, and N. P. Pavletich. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346–355, 1994.

35. C. C. Harris. p53: At the crossroads of molecular carcinogenesis and risk assessment. *Science*, **262**, 1980–1981, 1996.

36. H. Tanaka, H. Arakawa, T. Yamaguchi, K. Shiraishi, S. Fukuda, K. Matsui, Y. Takei and Y. Nakamura. A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage. *Nature*, **404**, 42–49, 2000.

37. G. Lozano and S. J. Elledge. Cancer: p53 sends nucleotides to repair DNA. *Nature*, **404**, 24–25, 2000.

38. A. Jordan and P. Reichard. Ribonucleotide reductases. *Annu Rev Biochem*, **67**, 71–98, 1998.

39. Y. Engstrom, B. Rozell, H. A. Hansson, S. Stemme, and L. Thelander. Localization of ribonucleotide reductase in mammalian cells. *EMBO J*, **3**, 863–867, 1984.

40. S. Elledge, Z. Zhou, and J. Allen. Ribonucleotide reductase: regulation, regulation, regulation. *Trends Biochem Sci*, **17**, (3), 119–123, 1992.

16

Loss of developmental controls in cancer

What we have considered up to this point, may have created the impression that only deregulation and dysfunctions of cellular proliferation cause cancer. Of course this is not true. Reactivation of developmental programmes operative in the embryo and the emergence of immature cell types play an important role in the genesis of tumours.

An understanding of developmental controls could help us to understand what has gone wrong when immature, malignant cells arise. Whereas normal cells are capable of differentiating into mature cells, which can form terminally differentiated cells and tissues, some cells, e.g. teratocarcinomas, seem to have lost the ability to differentiate normally. Because immature teratocarcinoma cells divide indefinitely without changing their character, and when isolated and propagated in tissue culture behave like immortal cancer cells, it was first thought that this cancer-like property of teratocarcinoma cells was the result of mutations of control genes, as in other cancer cells. But it was realized that immature teratocarcinoma cells are more like embryonic stem cells, which fail to differentiate. What has gone wrong is that the normal programme of differentiation is blocked in these cells. This block is due to a defect in communication of cells with their neighbours. Therefore, when these cells are transferred to a normal, embryonic environment, the block is removed and the cells differentiate normally. This explains why teratocarcinoma cells, when introduced in a 8-cell-stage mouse embryo blastocyst, form a healthy chimeric mouse (Fig. 16.1). Recently this phenomenon has received wide attention, ever since Ian Wilmut and colleagues were able to clone the sheep 'Dolly' by transferring the nucleus from a terminally differentiated cell, isolated from the mammary gland of an adult sheep to a totipotent embryonic egg cell from which the nucleus was removed. In the meantime, the same procedure has been used successfully for cloning mice, pigs, and cattle. I do not want to recapitulate the consequences of these successful cloning experiments, because they have already received wide publicity. I only want to recall that 277 fetus had to be sacrificed until one transfer was successful. The other 276 fetal cells did not develop properly, for whatever reason. This is a sobering and forbidding thought for all those entertaining the idea of cloning humans. But I want to direct attention to the really surprising fact that emerged from cloning Dolly, and which

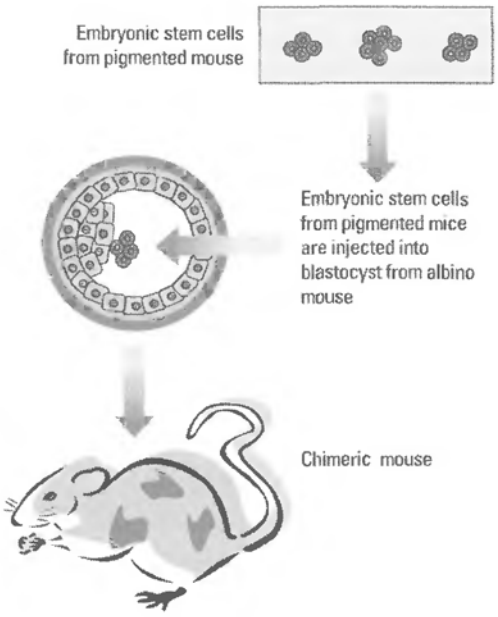


Fig. 16.1 Both teratocarcinoma cells and normal embryonal stem cells behave alike. Both types of cells are accepted and incorporated into the inner layer of the blastocyst, when they are injected into the blastocyst cavity. They behave like normal embryonic stem cells, and the progeny of these cells are found in practically every cell and tissue of the chimeric mouse, where they differentiate normally. Moreover, they even form normal germ cells. Thus, teratocarcinoma stem cells, when separated from their undifferentiated cancerous daughter cells and transplanted into a compatible host, not only survive, they also do no harm to the host and do not make him cancerous. Incidentally, the capability of embryonic stem cells to be accepted by a recipient blastocyst is the basis of producing gene knock-out, chimeric mice, where a normal gene is replaced by an altered, engineered version of the gene. These chimeric mice then carry the mutation in their stem cells and transmit them to differentiated cells, where they are expressed. (With permission of Taylor and Francis, Inc. See Fig. 21–32 in ref. 1.)

is relevant here in the context of a discussion of the behaviour of teratocarcinoma cells, i.e. that an enucleated embryonic egg cell could provide all the regulatory growth and differentiation factors needed to re-programme the genetic information removed from an already terminally differentiated mammary gland cell.

Cancer cells can arise from malignant stem cells, very early in the embryo. The emergence of leukaemic cells in the course of lymphocyte development is a good example (Chapter 14). In the course of lymphocyte development the sequential expression of genes for surface proteins and immunoglobulin-like receptors is tightly controlled, but in leukaemic cells this control is lost. The leukaemic cell appears to be arrested in a stage, through which each normal lymphocyte passes. Consequently, the leukaemic cell is unable to complete the normal programme of lymphocyte maturation. Thus, certain leukaemias and lymphomas are not caused by uncontrolled proliferation, rather their differentiation is blocked. The cause of malignancy in these cells is a block in the pathway to maturation, preventing the formation of fully differentiated cell lines with specific functions; an example of considerable medical interest is mutation of the *APC* gene and its consequences.

The adenomatous polyposis coli tumour suppressor gene

Somatic mutations of the *APC* gene are found in more than 70% of all adenomatous polyps and carcinomas of the colon and rectum. It has been estimated that about 50% of the population in the Western world will develop colorectal polyps during a normal life span, as a consequence of *APC* mutations. (See: ref: 2–6).

More than one gene must be altered by mutations to cause colorectal carcinoma. Although mutations and loss of the *APC* gene are an early event, they alone are not enough to cause cancer, but they seem to increase cell proliferation and speed up the renewal of cells in the colon, without altering their differentiation. What comes next are mutations of proto-oncogenes, such as the K(Kirsten)-*ras* proto-oncogene. Colon cells carrying these additional gene defects have already defects in the differentiation pattern, leading to intermediate adenoma. When these changes are followed by mutations of other genes, such as the *DCC* gene (deleted in colon carcinoma) and the *p53* gene, late adenoma and eventually carcinoma develop. Since the loss of each of the tumour suppressors, APC, DCC, and p53, requires two mutations (both copies of the gene must be inactivated, as was discussed above), a total of seven mutations, including the Ras mutation paves the way to malignant transformation.[7]

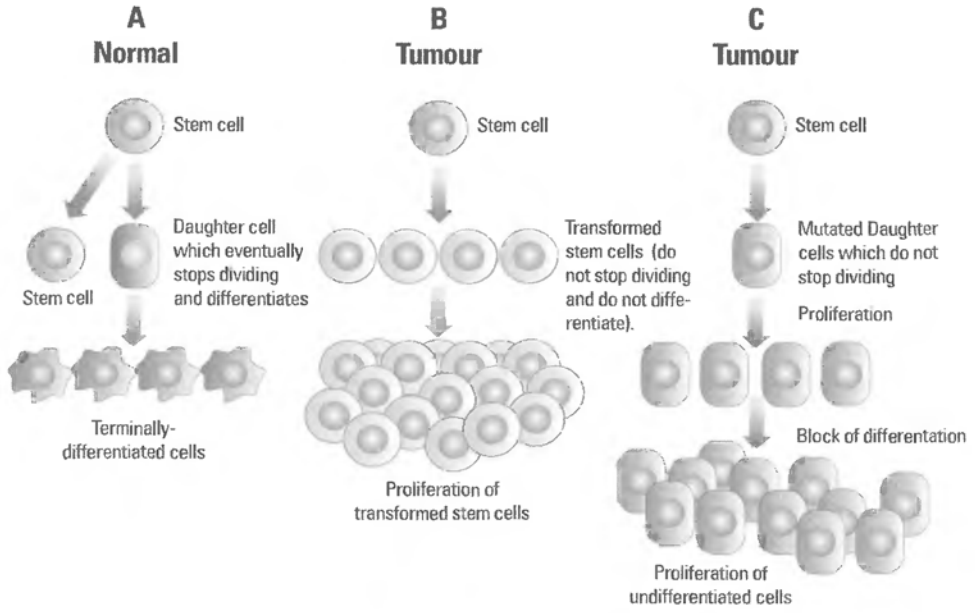


Fig. 16.2 (a) Normally, a stem cell produces with each cell division another stem cell, but also a daughter cell that is destined to differentiate terminally. (b) When a stem cell divides, producing new stem cells but no terminally differentiated cells which stop dividing, the result is a stem-cell-derived tumour, such as a teratocarcinoma. (c) A tumour also emerges when the non-stem-cell daughter cells fail to differentiate but do not stop dividing. This leads to proliferation of undifferentiated cells.

Epithelial cells in the intestine, as in many other epithelia such as the skin, renew themselves continually, by shedding terminally differentiated cells from the surface and replacing them with new cells, derived from the stem cells in the deeper basal layers (Fig. 16.2). But, for malignant transformation, it is not enough that stem cells just divide more rapidly, because as stem cells divide more rapidly, more terminally differentiated cells will be produced too. This would merely increase cell turnover. Immortal cancer cells are formed only when this balance is upset, either when the stem cell itself is transformed and fails to produce a cell which will differentiate and eventually die, or when the cell originating from the stem cell is mutated and fails to differentiate normally and instead divides indefinitely, without dying.

Cancers arising from a single transformed stem cell are monoclonal.[8] The prime example is multiple myeloma, where each neoplastic plasma cell from the same patient produces only one unique kind of immunoglobulin. Another example is B-cell lymphoma, where each lymphocyte carries either κ- or λ-immunoglobulin light chains on its surface, but not both. Still another example, often cited in favour of a monoclonal origin of cancer cells, is the expression of the genes for the two isozymes of glucose 6-phosphate dehydrogenase, which are located on the X chromosomes of the mother. These genes are distributed randomly in both X chromosomes of the mother. Thus, some stem cells will carry one and others the other isoenzyme gene, and the same applies to the differentiated cells arising from these stem cells. When tumour cells arise from these cells they express the particular isozyme that was present in the precursor cell from which the tumour cell originated. A monoclonal origin of cells from colorectal tumours has been questioned.[9] But whether a polyclonal origin is a peculiar property of APC or is also found in other tumours, remains to be seen.

The APC protein and β-catenin

The APC protein is large, roughly 300 kDa. It forms homodimers and interacts with several cellular proteins, among them β-catenin.[10] Through interaction with β-catenin, APC controls cell–cell interactions and cell migration. It has long been known that malignant, transformed cells lose contact inhibition and form irregular arrays of cells, with one cell piled upon the other (this was one of the first characteristic properties of transformed cells in culture to be recognized). It is still used as a diagnostic tool to distinguish between normal and cancer cells in culture (Fig. 16.3).

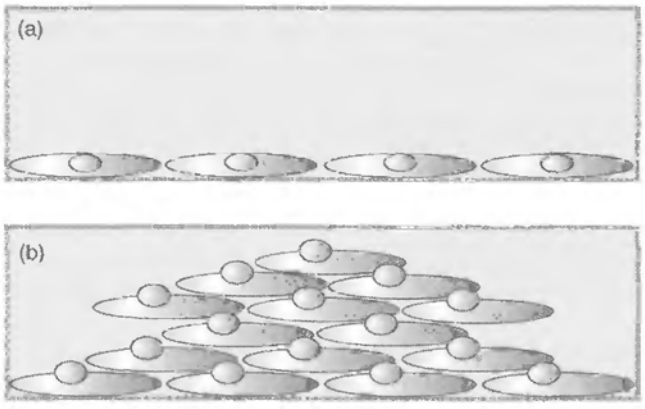


Fig. 16.3 Cancer cells, unlike most normal cells, continue to grow and pile up on top of each other after they have formed a confluent monolayer. The reason is that they have lost contact inhibition. (a) Normal cells; (b) cancer cells.

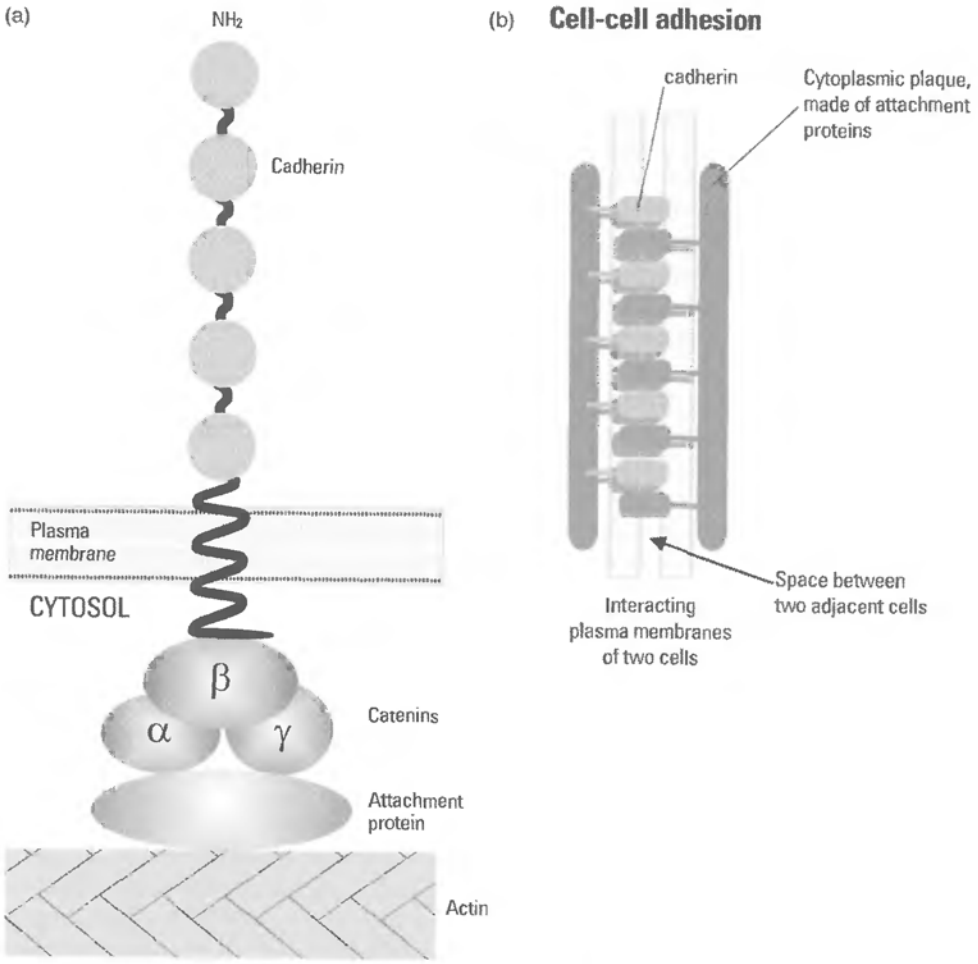


Fig. 16.4 (a) A cadherin molecule. The extracellular part of the protein is folded into five similar domains, three of which contain Ca^{2+}-binding sites. This part of the molecule is involved in cell–cell adhesion. The cytoplasmic part is connected to the actin cytoskeleton by catenins, such as α- E-catenin (which is structurally similar to vinculin), and a number of attachment proteins. The cytoplasmic domain which establishes the contact with catenins and the actin filaments is necessary for cell–cell adhesion. Cadherin molecules lacking the cytoplasmic tail can no longer hold cells together. There are several different cadherin molecules, which interact with the corresponding cadherins located on the surface of other cells. Cell–cell adhesion by cadherins is controlled by Ca^{2+}. (Fig. 19–24, in *Molecular biology of the cell*, B. Alberts *et al.*[1], is a model of this Fig, shown with permission of Taylor and Francis, Inc.) (b) The assembly of cadherins in bridging two adhesive cells with the help of attachment proteins. Each cytoplasmic attachment plaque is associated with a network of actin filaments (not shown).

β-Catenin is a multifunctional protein. It is a component of the cadherin cell-adhesion complex and is present in cell–cell adhesive junctions and desmosomes. Cadherins are calcium-binding transmembrane proteins. They are bound to vinculin-like molecules, the catenins, which link them through attachment proteins to the actin cytoskeleton of the cell (Fig. 16.4) (a short introduction to the regulation of cellular contractility is given in Chapter 4).

β-Catenin is a member of the Armadillo family of proteins (the *Drosophila* homologue of β-catenin is named 'Armadillo'). Armadillo/β-catenin participates in the Wingless/Wnt signalling cascade which regulates patterning and morphogenesis in *Drosophila* and humans.[11,12]

In the larvae of the fruitfly, cells must find their way to the body segment to which they belong. The control of patterning, for example the assembly of the wing in *Drosophila*, is carried out by a series of key signalling molecules, such as Wingless (Wg) and Hedgehog (Hh), and other morphogens. Hh is a representative example: it is released from cells in the posterior compartment with the help of the protein Dispatched (Disp). Assisted by Disp, Hh then moves to the anterior compartment, where it recruits another protein, Patched (Ptc), which keeps it there in a restricted area, where it can signal over short distances to a signal transducer, Smoothened. For release Hh is covalently modified in an unusual way, namely by attachment of cholesterol.[13]

A common denominator of patterning and embryogenesis is the movement of cells from one compartment to another and communication between cells.

Signalling by the Wnt pathway

Normally, cell proliferation and cell death are in a steady state, where the rate of proliferation is matched by the rate of apoptosis. Colon cancer, like other cancers, is caused by mutations in key components of regulatory pathways; in this case mutations of the tumour-suppressor gene *APC*, the adenomatous polyposis coli gene. When APC is inactivated and turned off, the Wnt pathway, normally operative only in the embryo, is turned on. The Wnt pathway also plays a role in other cancers, for example in mammary tumours in mice and humans (for further information and literature references, see ref. 14).

Signalling is initiated by binding of Wnt-factors to the Wnt or 'Frizzled' receptors. These receptors resemble remotely heptahelical, G-protein-coupled receptors (Chapter 5). In Wnt signalling, the Tcf (T-cell factor) family of transcriptional regulators plays a key role in the transmission of the Wnt signals to the genes.[15,16] Tcf combines with β-catenin, forming a complex, with Tcf contributing a DNA-binding domain and β-catenin a transactivating domain. In colon cancer cells, and in melanoma cell lines, high levels of free β-catenin accumulate and drive formation of complexes of β-catenin with Tcf-4 or Lef-1 (the lymphoid enhancer factor), which activate gene expression. The activated genes express regulatory proteins which stimulate cell proliferation and inhibit apoptosis. Thus, a key event is the increase of β-catenin, but what increases the level of β-catenin in the colon cell? The answer is, mutations of *APC* or mutations of β-catenin. These mutants have in common that they have lost their ability to block Wingless/Wnt signalling in colon cells. Thus, under these circumstances colon cells revert to embryonic cells.

The Wingless–Wnt pathway and β-catenin–APC signalling have much in common. For example, when the Wingless–Wnt pathway is activated in colon cells, the serine/

threonine-specific glycogen synthase kinase (GSK-3β) is inactivated, and free β-catenin accumulates in the nucleus, where it interacts with a transcription factor complex containing Tcf and Lef, activating gene transcription.[17] Thus, the situation in colon carcinoma cells resembles Wingless–Wnt signalling in embryonic cells, where phosphorylation of β-catenin is also blocked and free β-catenin monomers accumulate. For the accumulation of free β-catenin monomers, the loss of functional APC seems to be responsible, because re-introduction of wild-type APC reduced the level of cytoplasmic β-catenin. (Abnormally high levels of β-catenin were also found in human melanoma cell lines.) Normally, APC destroys β-catenin by recruiting and activating glycogen synthase kinase 3, (GSK-3β), which in turn phosphorylates β-catenin, initiating its proteolytic destruction. A key component participating in the destruction of β-catenin is axin,[18,19] so called because in mice it is required for the development of the dorsal–ventral body axis. Axin has a RGS (regulator of G-protein signalling) domain (Chapter 5), with which it binds to APC. How axin actually helps to destroy β-catenin is not clear, and likewise the intriguing possibility that G proteins may participate in Wnt signalling is not resolved. Apart from axin, other proteins cooperate with APC in the regulation of β-catenin, but their roles are not yet clearly defined. However, it is certain that in normal, healthy colon cells APC shuts down Wnt signalling and initiates the phosphorylation of catenin, leading to its destruction (Fig. 16.5).

It is clear how normal, wild-type APC downregulates the level of β-catenin, but why are high levels of β-catenin in the cytoplasm of the cell cancerogenic? The answer is that when APC is mutated in the colon cell and no longer can order the phosphorylation and destruction of β-catenin, the monomeric form of β-catenin accumulates first in the cytoplasm and subsequently moves to the nucleus, where it interacts with Tcf and Lef, forming transcriptionally active complexes. Uncontrolled activation of Tcf-controlled genes, by the β-catenin–Tcf–Lef complexes, is a likely cause of cancer, as shown in melanoma cell lines. Malignant transformation of colon cells also occurs when β-catenin is mutated and becomes unresponsive to GSK-3β-mediated phosphorylation. In this case, β-catenin remains unphosphorylated and cannot be destroyed. Thus, a mutant β-catenin, which is not properly phosphorylated, may cause cancer even when a functional, wild-type APC is present. (β-Catenins with abnormal phosphorylation patterns have been found in cancer cells.)

Summary

Migration is essential for colon cells. Normal epithelial colon cells migrate from crypt to villi, where they die and are sloughed off. Accumulation of β-catenin prevents cell migration. Thus, in order to permit cell migration and prevent adenomatous polyposis coli, the APC protein must order β-catenin degradation. In embryonic cells β-catenins are stabilized by Wingless/Wnt signals. These signals keep the cell in the right place and promote formation of β-catenin–Tcf–Lef complexes, activating the expression of Wingless/Wnt-responsive genes, which are essential for embryonic development. The accumulation of β-catenin in early colon cancer and melanoma cells points to similarities between embryonic signalling pathways and pathways which become operative in differentiated cells in the course of transformation to cancer cells. Hence, what is good for embryonic cells is bad for adult, differentiated cells.

The APC/β-catenin complex is only one of many proteins controlling cell–cell interactions. The integrin system has already been discussed (Chapter 4). Another factor,

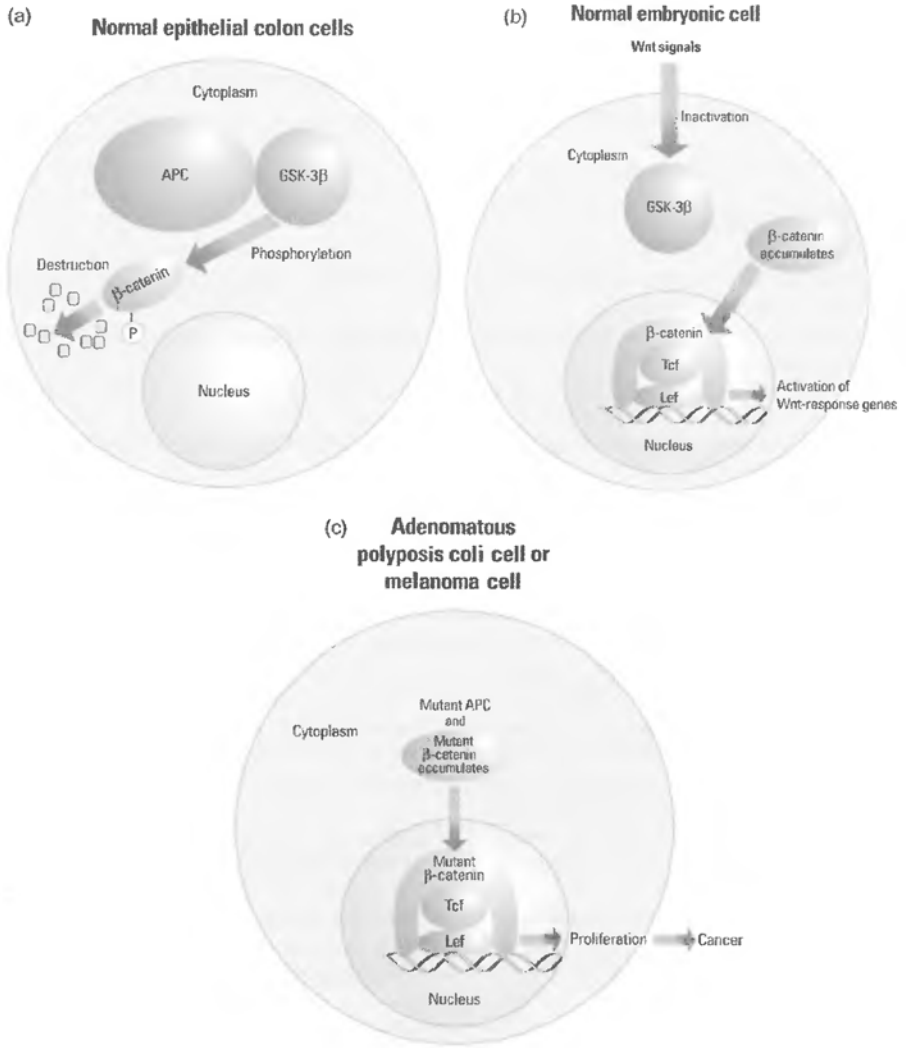


Fig. 16.5 The APC protein regulates β-catenin levels in normal cells. β-Catenin is an abundant cellular protein and much of it is bound to the cytoplasmic domain of the E-cadherin cell–cell adhesion complex. Mutations in APC or β-catenin lead to accumulation of β-catenin in the cytoplasm of cancer cells. Moreover, mutations in APC or β-catenin deregulate cell growth via T-cell factor 4 (Tcf-4). (a) In normal colon cells, glycogen synthase kinase 3β (GSK-3β) and APC promote degradation of β-catenin, probably as a result of phosphorylation of the NH_2-terminal sequences of β-catenin. The consequence is that β-catenin is not available for forming complexes with cadherin, promoting cell adhesion. Therefore, normal cells do not adhere to each other and do not pile up. Instead, they are free to migrate and are shed. (b) In normal embryonic cells, Wingless (Wnt) signals inactivate GSK-3β. The consequence is that β-catenin is not degraded and therefore accumulates. It moves to the nucleus, where it combines with Tcf/Lef, stimulating expression of Wingless/Wnt-responsive genes. (c) Point mutations and small deletions in either β-catenin or APC in early stages of colon cancer or melanoma lead to accumulation of β-catenin in mature, differentiated cells, preventing migration of these cells. This situation should be compared with the situation in embryonic cells (b). The difference between this situation and that in normal cells (a), is that mutated β-catenin is not phosphorylated by GSK-3β and consequently is not degraded. Therefore, β-catenin accumulates and forms with Tcf/Lef transcriptional complexes, as in embryonic cells. Although in adult cells the Wingless/Wnt signals are missing, the β-catenin/Tcf/Lef transcriptional complex also activates genes, just like the Wingless/Wnt signals in embryonic cells, promoting cell proliferation and blocking apoptosis, and directing the cells along the road to cancer. (Information taken mainly from refs 19–22.)

regulating growth by density-inhibition in normal cells, is a membrane-bound receptor, CiR. Ci is the plasma membrane glycoprotein, contactinhibin. Binding of Ci to its receptor is modulated by phosphorylation.[23] What relationships, if any, exist between Ci/CiR and APC/β-catenin is not known.

References

1. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell.* Garland Publishing, New York, 1994.
2. C. Lengauer, K. W. Kinzler, and B. Vogelstein. Genetic instabilities in human cancers. *Nature*, **396**, (6712), 643–649, 1998.
3. K. W. Kinzler and B. Vogelstein. Lessons from hereditary colorectal cancer. *Cell*, **87** (2), 159–170, 1996.
4. D. F. Ransohoff and C. A. Lang. Screening for colorectal cancer [see comments]. *N Engl J Med*, **325** (1), 37–41, 1991.
5. J. Jen, S. M. Powell, N. Papadopoulos, K. J. Smith, S. R. Hamilton, B. Vogelstein and K. W. Kinzler. Molecular determinants of dysplasia in colorectal lesions. *Cancer Res*, **54** (21), 5523–5526, 1994.
6. W. F. Bodmer, C. J. Bailey, J. Bodmer, H. J. Bussey, A. Ellis, P. Gorman, *et al.* Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature*, **328**, 614–616, 1987.
7. E.R. Fearon and B. Vogelstein. A genetic model for colorectal tumorigenesis. *Cell*, **61**, (5), 759–767, 1990.
8. P. C. Nowell. The clonal evolution of tumor cell populations. Science.194, (4260), 23–28.1976.
9. M. R. Novelli, J. A. Williamson, I. P. Tomlinson, G. Elia, S. V. Hodgson, I. C. Talbot, *et al.* Polyclonal origin of colonic adenomas in an XO/XY patient with FAP. *Science*, **272** (5265), 1187–1190, 1996.
10. B. Rubinfeld, B. Souza, I. Albert, O. Muller, S. H. Chamberlain, F. R. Masiarz, *et al.* Association of the APC gene product with beta-catenin. *Science*, **262** (5140), 1731–1734, 1993.

    Also: L. K. Su, B. Vogelstein, and K. W. Kinzler. Association of the APC tumor suppressor protein with catenins. *Science*, **262** (5140), 1734–1737, 1993.
11. M. van-de-Wetering, R. Cavallo, D. Dooijes, M. van Beest, J. van Es, J. Loureiro, *et al.* Armadillo coactivates transcription driven by the product of the *Drosophila* segment polarity gene dTCF. *Cell*, **88** (6), 789–799, 1997.
12. B. M. Gumbiner. Signal transduction of beta-catenin. *Curr Opin Cell Biol*, **7**, (5), 634–640, 1995.
13. R. Burke, D. Nellen, M. Bellotto, E. Hafen, K. A. Senti, B. J. Dickson, and K. Basler. Dispatched, a novel sterol-sensing domain protein dedicated to the release of cholesterol-modified hedgehog from signaling cells. *Cell*, **99**, (7), 803, 1999.
14. A. Wodarz and R. Nusse. Mechanisms of Wnt signaling in development. *Annu Rev Cell Dev Biol*, **14**, 59–88, 1998.

    Also: the Wnt gene homepage: http://www. stanford.edu/~rnusse/wntwindow.html and http://www. stke.org

    Also: M. Peifer and P. Polakis. Wnt signalling in oncogenesis and embryogenesis – a look outside the nucleus. *Science*, **287**, 1606–1609, 2000.

    Also: P. Polakis. The oncogenic activation of beta-catenin. *Curr Opin Genet Dev*, **9**, 15, 1999.
15. J. Riese, X. Yu, A. Munnerlyn, S. Eresh, S. C. Hsu, R. Grosschedl, and M. Bienz. LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. *Cell*, **88** (6), 777–787, 1997.
16. E. Brunner, O. Peter, L. Schweizer, and K. Basler. Pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in *Drosophila*. *Nature*, **385** (6619), 829–833, 1997.
17. B. Rubinfeld, I. Albert, E. Porfiri, C. Fiol, S. Munemitsu, and P. Polakis. Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly [see comments]. *Science*, **272** (5264), 1023–1026, 1996.
18. M. J. Hart, R. de los Santos, I. N. Albert, B. Rubinfeld, and P. Polakis. Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. *Curr Biol*, **8**, (10), 573–581, 1998.

    Also: C. Sakanaka and L. T. Williams. Functional domains of axin. Importance of the C terminus as an oligomerization domain. *J Biol Chem*, **274**, (20), 14090–14093, 1999.

19. S. Munemitsu, I. Albert, B. Souza, B. Rubinfeld, and P. Polakis. Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proc Natl Acad Sci, USA*, **92** (7), 3046–3050, 1995.

20. V. Korinek, N. Barker, P. J. Morin, D. van Wichen, R. de Weger, K. W. Kinzler, *et al.* Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC–/– colon carcinoma [see comments]. *Science*, **275** (5307), 1784–1787, 1997.

21. P. J. Morin, A. B. Sparks, V. Korinek, N. Barker, H. Clevers, B. Vogelstein, and K. W. Kinzler. Activation of beta-catenin–Tcf signaling in colon cancer by mutations in beta-catenin or APC [see comments]. *Science*, **275** (5307), 1787–1790, 1997.

22. P. Polakis. Loss of β-catenin regulation by mutations in the β-catenin gene or by inactivation of the APC tumor suppressor. Third international colloquium on 'Cellular signal recognition and transduction', 7–11 October 1997, Berlin. Germany.

    Also: B. Rubinfeld, P. Robbins, M. El-Gamil, I. Albert, E. Porfiri, and P. Polakis. Stabilization of β-catenin by genetic defects in melanoma cell lines. *Science*, **275**, 1790–1792, 1997.

    Also: M. Peifer. β-Catenin as oncogene: The smoking gun. *Science*, **275**, 1752–1753, 1997.

23. G. Gradl, D. Faust, F. Oesch, and R. J. Wieser. Density-dependent regulation of cell growth by contactinhibin and the contactinhibin receptor. *Curr Biol*, **5**, 526–535, 1995.

17

The causes of cancer

Whatever the cause of cancer, chemical, physical, viral, or spontaneous, all cancers have in common changes of genes controlling the programme that regulates growth and differentiation of cells. Probably, the first person to see clearly the causal relation between defects in the genetic information of a cell, dysregulation of cellular proliferation, and malignant transformation was the eminent biologist, Theodor Bovery, Professor at the University of Würzburg. He wrote 1914,[1] 'In those cases, where the unconstrained tendency of cells to grow is the only phenomenon characterizing the disease, probably only the chromosomes (genes) which regulate cell division are either absent or are overactive'.[2]

Genetics of cancer

Only about 1–2 % of all cancers are hereditary. More than 20 different hereditary cancers have been characterized and attributed to specific germline mutations in more than 25 genes.[3]

Inherited cancer genes

Proteins encoded by inherited cancer genes are often tumour suppressors and transcription factors regulating gene expression, controlling the cell cycle, apoptosis, and cell differentiation, or receptors, regulating cell proliferation. Others are DNA damage-repair enzymes, maintaining the integrity of the genomic information. Inherited cancer mutations are often found in genes encoding components of a multi-component, highly complex assembly, the 'DNA mismatch repair system' (MMR). The large functional repertoire of these proteins makes it understandable why it is so difficult to find the gene that is responsible for the dysfunction that actually is the cause of the malignant disease. For example, a hereditary predisposition to melanoma is due to a defect in the control of the Cdk4, a cyclin-dependent cell-cycle kinase.[4] The expression and activity of the Cdk4 kinase is normally under the control of a tumour suppressor. Thus, one would expect that melanoma is associated with a germline mutation in the gene for the tumour suppressor. However, it turned out that the mutation is not in the tumour-suppressor gene itself, but

in the *Cdk4* kinase gene, making expression of the *Cdk4* kinase gene insensitive to the damping action of the suppressor protein.

To prove that a cancer is inherited in a true Mendelian way is difficult. One of the reasons is that the cancer phenotype of patients who do have an inherited cancer gene may be clinically indistinguishable from that of patients who do not carry such an inherited cancer gene. For example, the breast cancer associated proteins, BRCA1 and BRCA2, are encoded by two different genes, but since the proteins are both involved in the repair of double-stranded DNA breaks, mutations in each of these different genes manifest themselves as the same clinical phenotype, namely breast cancer. Thus, it is difficult to find the germline mutation in the gene which is responsible for breast cancer. Similarly, melanomas have mutations in different genes, but all of these genes participate in cell-cycle control, again making it difficult to find the gene with an inherited germline mutation. Furthermore, even when the same gene is mutated but at different sites, the clinical manifestations may be different. A case in question involves mutations in the gene, coding for a transmembrane tyrosine kinase receptor, the glial-derived neurotrophic factor receptor, GDNFR (Chapter 1).[5] Although all mutations of the receptor gene increase the activity of the receptor tyrosine kinase in one way or the other, each mutation causes a different effect; either on substrate affinity or on catalytic parameters (V_{max}). These differences lead to different clinical manifestations.

These examples show how difficult it is to relate results of gene testing to the actual clinical manifestation. Moreover, as pointed out by Eric R. Fearon,[3] mutations in genes encoding enzymes driving cellular metabolism might make individuals react differently to environmental and dietary agents, thus increasing indirectly an inherited risk of cancer.

This shows how careful and detailed the genetic and clinical analysis of a suspected hereditary cancer must be, before a risk prognosis of cancer, with its inevitable emotional consequences, can be made. Finally, testing for inherited cancer is hampered due to the lack of an animal model of human cancer. Transgenic mice bearing germline cancer genes often do not develop the cancer syndromes seen in humans carrying the same germline defects.[3]

Somatic mutations in proto-oncogenes and tumour-suppressor genes

The vast majority of mutations in cancer are not hereditary. They are somatic and are found only in the cells of an individual patient. A common event is mutation in one allele of a proto-oncogene, followed by deletion of the other normal allelic gene, leaving only the mutant allele. This is called a loss of heterozygosity (LOH). The appearance and clinical manifestation of a human tumour is the outcome of a complex, multi-step process, involving somatic mutations in many cellular genes. Somatic mutations in proto-oncogenes and in tumour-suppressor genes accumulate throughout the lifetime of an individual. Therefore the crucial question is, how does a tumour eventually arise, despite the fact that we tolerate so many mutant genes during our lifetime? What is responsible for a burst in mutability, eventually causing cancer? One would expect a mutation causing wide-ranging, global effects, such as chromosomal instabilities (CINs). These changes could affect checkpoint genes, controlling the chromosome condensation, kinetochore assembly, and the cohesion of the chromatids.

There are checkpoints controlling replication and separation of chromosomes during mitosis. A DNA-damage checkpoint controls chromosome replication and stops cells

before entering prophase when an error in DNA replication is sensed. In the pro(meta)phase, the spindle begins to form and the microtubules are organized by the centrosomes at either end of the cell. In metaphase, as the spindle forms, the chromosomes attached to the microtubules align themselves, forming the kinetochore. In anaphase, the chromosomes move along the shortening microtubules, to opposite poles of the cell. As was discussed in Chapter 12, normally a 'checkpoint' controls the cellular machinery that pulls the chromatids apart, and monitors the proper assembly of the mitotic spindle. Checkpoints control chromosome condensation, kinetochore assembly, and the cohesion of the chromatids. All processes involved in the alignment of chromosomes, the microtubule dynamics and assembly of microtubules, centrosome replication, and mitotic spindle formation are controlled before the cell enters the anaphase. When the checkpoint control does not function correctly, the improper allocation of the chromosomes to the two daughter cells leads to aneuploidy (irregular number of chromosomes) and to chromosomal instability (CIN). Therefore, mutations silencing cell-division checkpoints are likely to cause cancer, because such mutations would cause increased mutability. Loss of a cell-division checkpoint control can result in the removal of whole chromosomes and changes in chromosome numbers. The resulting aneuploidy is often found in cancer cells (Fig. 17.1).

Mutations silencing a cell-division checkpoint cause aneuploidy in cancer cells

Cahill *et al.*[7] (discussed by Weaver and Weinberg[8]) have shown that the mechanisms controlling the proper separation of chromosomes during cell division[9] are defective in colorectal cancer cells, causing aneuploidy.

Genes encoding proteins that prevent entry into anaphase when the spindle assembly is in disarray, have been identified in yeast. These are the *MAD* genes (mitotic arrest-deficient genes) and the *BUB* genes (which got their name because they make budding in yeast unresponsive to benomyl, hence budding uninhibited by benomyl; benomyl is a drug that causes the microtubules to depolymerize and to fall apart[10]). The spindle assembly checkpoint monitors the microtubule attachment, senses signals from unattached kinetochores, and sends out a stop signal when the microtubules are not properly attached (Fig. 17.2).

When the checkpoint control has sensed a defect in the micotubular attachments, the cell is arrested in metaphase. Mutants of the human *hBUB1* gene have been detected

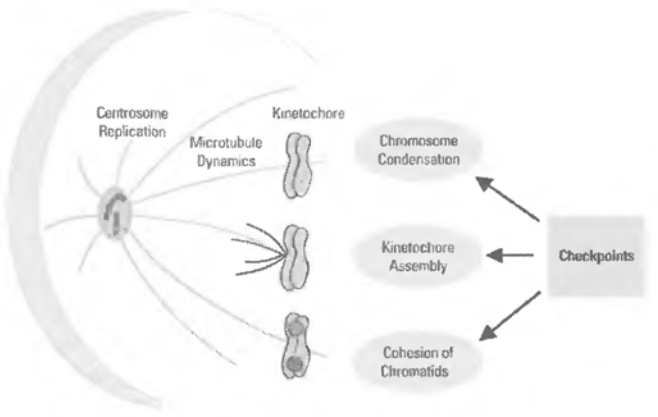


Fig. 17.1 The mitotic-spindle checkpoints ensure that chromosomes are aligned correctly before anaphase. If chromosomes are not attached stably to the microtubules, the checkpoint control prevents entry into anaphase, the irreversible step in cell division.[6]

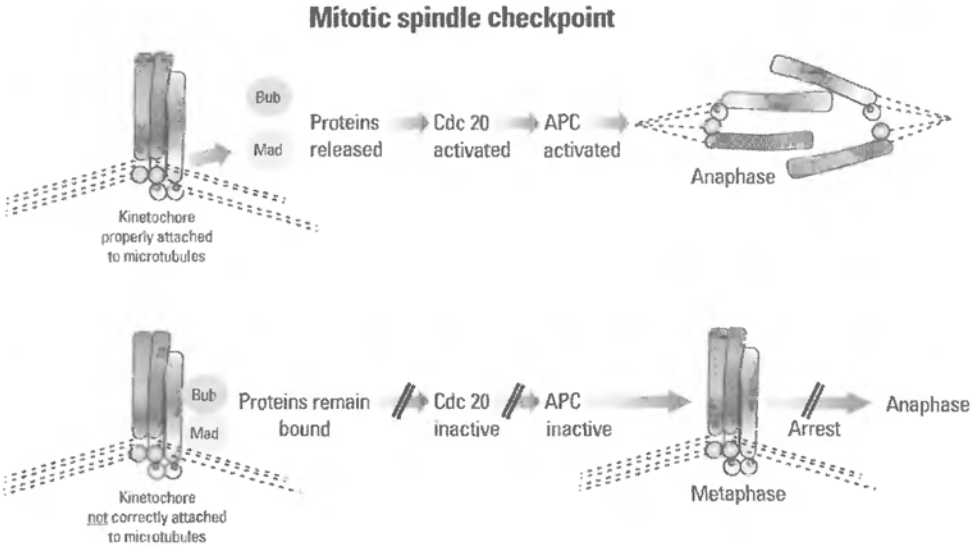


Fig. 17.2 *Top*: The normal situation: the Bub and Mad proteins dissociate from the kinetochore region when the microtubules are properly aligned. The released Bub/Mad proteins activate a protein (Cdc20) which regulates entry into mitosis and activates the anaphase-promoting complex (APC) and gives the go-ahead for entry into the anaphase. *Below*: The situation when the kinetochore is not attached properly to the microtubules. In this case, the Bub and Mad proteins remain attached to the kinetochore, Cdc20 remains inactive, and APC is not activated. This arrests the cell in metaphase. Thus, the Bub and Mad and Mps-1 proteins in yeast and mammals, respectively, respond to improper assembly of the spindle by arresting cells. (Reproduced with permission of Professor R. A. Weinberg and Nature from Fig. 2 in ref. 8.)

in 2 of 19 colorectal tumour cell lines with high rates of aneuploidy. Transfer and expression of mutant *hBUB1* genes in normal euploid cells was shown by Cahill *et al.*[11] to disrupt mitotic checkpoint controls.

How does aneuploidy promote malignant transformation? It probably accelerates the loss of heterozygosity and increases the rate at which tumour-suppressor genes are lost.

The role of DNA methylation in cancer

Hypermethylation of the tumour-suppressor gene, p53, has been found in cancerous mice. Methylation of DNA has been assumed to be irreversible and the irreversibility has been linked to a possible mutagenic role of methylation in cancer. However, recently a mammalian demethylating enzyme has been identified,[12] but how methylation patterns and gene expression are exactly related is not yet clear (Chapter 10). However, it can be expected that we shall learn more about the role of methylation of DNA in cancer in the future.

The role of telomerase in the transformation of normal human cells to cancer cells

As early as 1983, Robert Weinberg and colleagues[13] had been able to convert primary cells from rodents (mice) into immortal tumour cells by the expression of oncogenes.

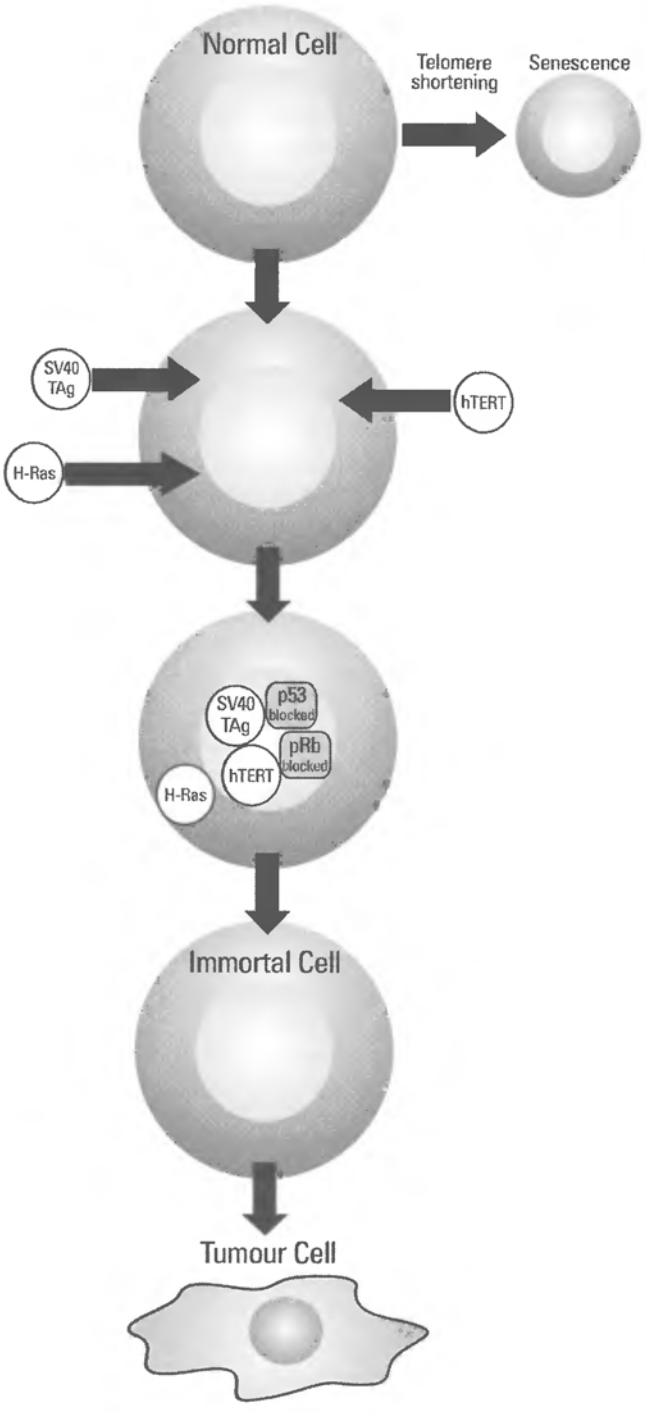


Fig. 17.3 The expression of two oncogenes, SV40 TAg and H-*ras*, together with the catalytic subunit of human telomerase, hTERT, transforms normal human cells to immortal cells and eventually to tumour cells. (Information from Hahn *et al.*;[14] see also ref. 15.)

However, all attempts to immortalize human cells have failed. Recently, however, Robert Weinberg's laboratory showed that the expression of the catalytic subunit of the human telomerase (hTERT), together with two oncogenes, H-*ras* and a SV40, simian virus oncoprotein (the large T antigen, TAg), converted normal, human epithelial and fibroblast cells into cancer cells.[14] This explains the previous failure which was a consequence of the difference between rodent and human cells, because human cells lack telomerase activity. Thus, while the cessation of telomerase activity in senescent human cells limits their life span (Chapter 12), it also brings the cell in a quiescent state and guards it against transformation to a cancer cell (Fig. 17.3).

Oncogenic DNA and RNA viruses

About 20% of all cancers have been traced to tumour viruses, although the number of people infected is probably much larger than the number of people who actually develop cancer. Moreover, cancer-causing viruses may only be one of many contributing factors. DNA tumour viruses are probably more important as cancer-causing agents than RNA retroviruses (Table 17.1).

DNA tumour viruses and RNA retroviruses differ fundamentally in the way they cause malignant transformations. DNA tumour viruses infect a host cell, in a way not different from other DNA viruses infecting cells. Having entered the cell, the virus uses the enzymes of the host cell to transcribe the viral genome and replicate. A DNA virus is programmed to produce a large number of new, infectious virus particles and is not primarily programmed to transform the host cell into a cancer cell. However, occasionally the viral DNA is stably inserted in the host-cell chromosome and transforms the cell. On the other hand, the cellular, c-proto-oncogenes, the homologues of retroviral genes, the v-oncogenes, are commonly found in the genome of normal cells, whereas the genes derived from DNA viruses have no homology with any known cellular gene. The genes of DNA viruses produce their own, unique transforming proteins in the cell. One way in which viral proteins, such as TAg (the large T antigen) of the polyoma SV40 (simian virus 40) or the oncoproteins of the adenovirus E4orf6, transform cells is by binding to tumour suppressors, p53 or pRB, blocking their actions,[16] or by interfering with the function of

Table 17.1 Human cancer viruses

Virus	Associated tumour
DNA viruses	
Polyoma viruses SV40 (Simian virus)	?
Papova (papilloma) viruses	Carcinoma of the uterine cervix
Hepatitis-B virus	Liver cancer
Herpes virus	Burkitt's lymphoma
Epstein–Barr virus	B-lymphocyte cancer
RNA retroviruses	
Human T-cell leukaemia virus, type I (HTLV-I)	Adult T-cell leukaemia
HIV virus	Karposi sarcoma
HIV-1 (human immunodeficiency virus)	Endothelial cell cancer of blood and lymph vessels

From information reported in Table 24-6 of ref. 17 and reproduced with permission of Taylor and Francis, Inc.

other transcriptional activators.[18] Another target of adenoviral oncoproteins is the histone transacetylase and the histone deacetylase (HDAC1), which are part of the transcriptional activation complex (see Part 2). After adenoviral infection, the adenoviral protein E1A binds to pRB and blocks the action of HDAC1. (The interaction of p53 with the adenoviral E1B protein is quite like that of E1A.) These events drive the cell cycle, DNA replication, and proliferation, and set the stage for malignant transformation. However, this is not the only way that DNA viruses can cause malignant transformation of cells. For example, the middle part of the large T-antigen (TAg) of the polyoma simian virus (SV40) expressed in virus-infected cells, also binds to cytosolic tyrosine kinases of the c-Src family. About a 20-fold increase in the activity of the kinase when bound to the SV40 polyoma protein most likely contributes to the malignant transformation of the virus-infected cell. Moreover, the Myc transcriptional proteins share structural features with the SV40 large T-antigen and with adenoviral E1A transcription factors, pointing to common DNA-interaction sites and to the possibility of mutual interference.[19] Since adenoviruses are used as vehicles for gene delivery in somatic gene therapy in humans, care must be taken in these cases to remove the interaction domains of the adenoviral proteins with tumour suppressors and transcriptional activators. Another problem is that these viral gene vectors may be antigenic and elicit an immune response.

Summary

The cancer cell is out of control. All the genes that play a role in malignant transformation encode essential components of the machinery regulating the cell cycle and cell division, controlling cellular proliferation and differentiation, and maintaining the integrity of the genome and safeguarding DNA replication. The following are mainly responsible for deregulation:

1. Several somatic mutations of tumour-suppressor genes, controlling cell-cycle and cell-division checkpoints, together with mutations or overexpression of proto-oncogenes. Such mutations lead to an abnormal gain or loss of function of key regulatory elements.
2. More than one crucial hit is required. Defects in cell-cycle and cell-division checkpoints lead to aneuploidy and hypermutability.
3. The small number of germline inherited cancers are caused by loss-of-function mutations of genes encoding tumour suppressors.
4. In cases such as that exemplified by colon cancer cells, the tumour cell regresses and returns to an embryonic developmental programme.

Outlook

Cells use common tools and identical basic mechanisms and programmes of regulation which they have acquired in the course of evolution. These evolutionarily conserved tools allow them to respond to a multitude of external and internal signals, and enable them to deal effectively with the challenges of the environment and therefore survive. The elucidation of the structure and function of the instruments keeping the intricate signalling network in a cell under control should give new possibilities to manipulate and control the

tools regulating growth and differentiation of cells. The more specific and effective the means become to correct regulatory dysfunctions by repairing the instruments of regulation, either at the level of the gene or at the level of the protein, the greater the chances of treating cancers more effectively.

References

1. T. Boveri. *Zur Frage der Entstehung maligner Tumoren*. Jena, 1914.
2. I am grateful to my colleague, Professor Koschel from the Department of Immunology of the University of Würzburg, for directing my attention to this citation.
3. E. R. Fearon and B. Vogelstein. A genetic model for colorectal tumorigenesis. *Cell*, **61** (5), 759–767, 1990.

   Also: Online Mendelian Inheritance in Man (OMIM), a catalog, edited by V. A. McKusick, of human genes and genetic disorders. For each gene, OMIM provides general informationon the function of the gene, references derived from MEDLINE, and links to gene maps and molecular sequence data.

   Also: E. R. Fearon. Human cancer syndromes: clues to the origin and nature of cancer. *Science*, **278**, 1043–1048, 1997.
4. M. Serrano, G. J. Hannon, and D. Beach. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4 [see comments]. *Nature*, **366** (6456), 704–707, 1993.
5. C. Eng. The RET proto-oncogene in multiple endocrine neoplasia type 2 and Hirschsprung's disease. Seminars in medicine of the Beth Israel Hospital, Boston. *N Engl J Med*, **335** (13), 943–951, 1996.

   Also: L. M. Mulligan, J. B. Kwok, C. S. Healey, M. J. Elsdon, C. Eng, E. Gardner, *et al.* Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. *Nature*, **363** (6428), 458–460, 1993.

   Also: B. Pasini, I. Ceccherini, and G. Romeo. RET mutations in human disease. *Trends Genet*, **12** (4), 138–144, 1996.

   Also: L. M. Mulligan, C. Eng, T. Attie, S. Lyonnet, D. J. Marsh, V. J. Hyland, *et al.* Diverse phenotypes associated with exon 10 mutations of the RET proto-oncogene. *Hum Mol Genet*, **3** (12), 2163–2167, 1994.

   Also: T. Attie, A. Pelet, P. Edery, C. Eng, L. M. Mulligan, J. Amiel, *et al.* Diversity of RET proto-oncogene mutations in familial and sporadic Hirschsprung disease. *Hum Mol Genet*, **4** (8), 1381–1386, 1995.

   Also: M. Angrist, S. Bolk, B. Thiel, E. G. Puffenberger, R. M. Hofstra, C. H. Buys, *et al.* Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung disease. *Hum Mol Genet*, **4** (5), 821–830, 1995.
6. C. L. Rieder, R. W. Cole, A. Khodjakov, and G. Sluder. The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *J Cell Biol*, **130** (4), 941–948, 1995.
7. D. P. Cahill, C. Lengauer, J. Yu, G. J. Riggins, J. K. Willson, S. D. Markowitz, *et al.* Mutations of mitotic checkpoint genes in human cancers [see comments]. *Nature*, **392** (6673), 300–303, 1998.
8. T. L. Orr-Weaver and R. A. Weinberg. A checkpoint on the road to cancer. *Nature*, **392**, 223–224, 1998.
9. C. L. Rieder, R. W. Cole, A. Khodjakov, and G. Sluder. The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *J Cell Biol*, **130** (4), 941–948, 1995.
10. K. G. Hardwick, E. Weiss, F. C. Luca, M. Winey, and A. W. Murray. Activation of the budding yeast spindle assembly checkpoint without mitotic spindle disruption. *Science*, **273** (5277), 953–956, 1996.

    Also: L. H. Hwang, L. F. Lau, D. L. Smith, C. A. Mistrot, K. G. Hardwick, E. S. Hwang, *et al.* Budding yeast Cdc20: a target of the spindle checkpoint [see comments]. *Science*, **279** (5353), 1041–1044, 1998 [published erratum appears in *Science*, **280** (5368), 1331, 1998].

    Also: S. H. Kim, D. P. Lin, S. Matsumoto, A. Kitazono, and T. Matsumoto. Fission yeast Slp1: an effector of the Mad2-dependent spindle checkpoint [see comments]. *Science*, **279** (5353), 1045–1047, 1998.
11. D. P. Cahill, C. Lengauer, J. Yu, G. J. Riggins, J. K. Willson, S. D. Markowitz, *et al.* Mutations of mitotic checkpoint genes in human cancers [see comments]. *Nature*, **392** (6673), 300–303, 1998.
12. S. K. Bhattacharya, S. Ramchandani, N. Cervoni, and M. Szyf. A mammalian protein with specific demethylase activity for mCpG DNA. *Nature*, **397**, 579–583, 1999.

13. H. Land, L. F. Parada, and R. A. Weinberg. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature*, **304**, 596–602, 1983.

14. W. C. Hahn, C. M. Counter, A. S. Lundberg, R. L. Beijersbergen, M. W. Brooks, and R. A. Weinberg. Creation of human tumour cells with defined genetic elements. *Nature*, **400**, 464–468, 1999.

15. J. B. Weitzman and M. Yaniv. Rebuilding the road to cancer. *Nature*, **400**, 401–402, 1999.

16. T. Dobner, N. Horikoshi, S. Rubenwolf, and T. Shenk. Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor. *Science*, **272** (5267), 1470–1473, 1996.

17. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular biology of the cell*. Garland Publishing, New York, 1994.

18. X. J. Yang, V. V. Ogryzko, J. Nishikawa, B. H. Howard, and Y. Nakatani. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature*, **382** (6589), 319–324, 1996.

19. J. Figge, T. Webster, T. F. Smith, and E. Paucha. Prediction of similar transforming regions in simian virus 40 large T, adenovirus E1A, and myc oncoproteins. *J Virol*, **62** (5), 1814–1818, 1988.

Glossary

ABC transporters belong to a large family of membrane proteins that carry many different molecules through the membrane. Transport is driven by the hydrolysis of ATP.

abl is an oncogene first found in the Abelson strain of mouse leukaemia virus. Mutations of the corresponding, cellular *c-abl* protooncogene are found in chronic granulocytic leukaemia of humans. The product of *c-abl* is a tyrosinekinase.

***a*-helix,** first described by Pauling and Corey, 1951. It is a spiral configuration of a polypeptide chain in which the turns are determined by hydrogen bonds formed between the amide peptide bonds of the constituent aminoacids of the chain. The *a*-helix has 3.6 residues per turn. Each aminoacid residue is translated relative to the other by 1.5 Å along the axis of the helix. The translation per one complete turn is 5.4 Å. The *a*-helix can be left- or right-handed, but right-handed helices are much more common.

ACE, Angiotensin Converting Enzyme is a peptidyl carboxy peptidase which forms the octapeptide angiotensin II from angiotensin I. Angiotensin II is a powerful pressor agent responsible for essential hypertension. Therefore, compounds that inhibit ACE are used in the treatment of hypertension.

Actinfilaments are helical proteins, formed by the polymerization of actin monomers. They are major components of the cytoskeleton of eukaryotic cells and part of the contractile machinery of skeletal muscle.

Actin-binding proteins are members of a group of diverse proteins that bind to actinfilaments and affect their function. Among them are *a*-actinin, myosin, profilin, paxillin, talin and many others.

ADAM. A Disintegrin And MetalloProteinase family of proteins. Disintegrins, as the name implies, are proteins which interfere with interactions of cells with proteins in the extracellular matrix. An example are the inhibitors of the interaction of blood platelets with fibrinogen. Disintegrins are found in snake venom. Metalloproteinases are a family of proteases which need a bivalent cation for catalysis. MMP's are matrix metalloproteases. They are associated with the extracellular matrix.

Adhesionmolecules are expressed on the surface of cells and mediate the adhesion among cells or of cells with the extracellular matrix. Adhesion molecules bind to transmembrane receptors, the integrins. There are several groups of adhesion molecules for example, the selectins on leucocytes and the immunoglobulin superfamily, to which the ICAM's, the intercellular adhesionmolecules,the cadherins and CD 44, a cell adhesion molecule, belong. CD44 is constitutively expressed on the surface of many cells, including B and T lymphocytes. CD44 binds to collagen and to glycosaminoglycans, (*e.g:* Hyaluronate, a ß 1→4-linked D-glucuronyl-(ß 1→3)-N-acetyl D-glucosamine).

Allele, one of two alternative forms of the same gene in a diploid cell. Each allele has the same locus on the same homologous chromosome but may differ in coding sequences.

Allelic exclusion is the process by which a cell uses either the gene from the maternal chromosome or the allelic gene from the paternal chromosome, but not both. Allelic exclusion occurs in genes encoding antibodies and T cell receptors.

Allosterism is a property of an enzyme or any protein, of which the reactivity is modified by binding of an effector at a site other, (allo), than the site where the substrate or the reactant binds. This leads to a conformational change, altering positively or negatively the biologically-relevant activity of the protein. The concept of allosterism was introduced by Jacques Monod.

Amadori rearrangement, an acid- or base-catalyzed non-enzymic chemical reaction in which a protein- N-aldosylamine adduct is converted into the corresponding substituted 1-amino 1-deoxy 2-ketose-adduct. It occurs in the Maillard reaction, a nonenzymic reaction where a reducing aldose sugar combines at its anomeric carbon with an aminogroup of an aminoacid, peptide or protein, to form first a Schiff's base and then by an Amadori rearrangement the correspoding ketoamine. This reaction is responsible for the browning of sugar-containing foods and the non-enzymic glycosylation of proteins in diabetes.

Anaphase , the stage of mitosis where the two sets of chromosomes separate and move away to the opposite poles of the cell.

Ångstrom,(Å) is equal to 10^{-10} meter or 0.1 nm.

Ankyrin is a 200 kDa protein that links cytoskeletal proteins to membrane proteins. It is bound tightly to the cytoplasmic surface of the erythrocyte plasma membrane to which it attaches the cytoskeletal protein spectrin.

APC, antigen-presenting cells are bone-marrow-derived cells, macrophages and dendritic cells which process foreign antigens and present them to T cells, OR, adenomatous polyposis coli protein, enoded by the human *apc* gene.

Apoptosis, programmed cell death is a vital cellular program which keeps the number of cells in metazoa within the physiological range. The biochemical and morphological sequelae of apoptosis of a cell include cleavage of DNA by endonucleases, chromatin condensation, nuclear fragmentation, proteolytic destruction of the cytoskeleton, blebbing of the plasmamembrane and cell shrinkage.

Archaea, or archaebacteria are one of the three kingdoms of cells. The other kingdoms are the Eukarya and the Bacteria.

Autocrine stimulation. Stimulation by a molecule that activates the same cell where it is produced.

Bacteriorhodopsin, is a retinal-containing protein in the purple membrane of a halophilic, (salt-loving) archaebacterium, *Halobacterium halobium,* which pumps protons out of the cell on activation by light.The three-dimensional structure of bacteriorhodopsin resembles that of rhodopsin in the eye.

B cells are the principal cells of the islets of Langerhans in the pancreas. B cells produce insulin, store the hormone in granules and secrete it on stimulation by glucose.

Blastocyte is an undifferentiated toti-potent embryonic cell.

Blastomeres are formed by cleavage of the fertilized egg.

Blastula is an early stage of the embryo, before gastrulation.

B lymphocytes , B cells or bone marrow-derived lymphocytes are formed in humans in the bone marrow by stem cells. On antigenic stimulation they develop into plasma cells and synthesize and secrete circulating antibodies.

Blood clotting is a process forming a blood clot, consisting of aggregated platelets and a mesh of fibrin molecules. The clot plugs the wound and stops bleeding, Blood clotting is a complex sequence of reactions where 13 factors are participating. It is a cascade of interactions of proteinases with their substrates and of interactions of other proteins, cofactors, lipids and Ca^{2+}.

Bombesin is a tetradecapeptide amide. It is a peptidehormone which activates a G protein-coupled receptor and causes release of gastrin and cholecystokinin in the intestine.

Burkitt's lymphoma is a human tumour of mature immunoglobulin-producing B lymphocytes.

C-terminal and N-terminal are the end and the start respectively, of a peptide chain in a protein. The N-terminal aminoacid in a protein has a free amino group. This group is often acylated. When it is free and protonated and positively charged at neutral pH, it may be also called NH_2-terminal or NH_3^+ -terminal. At the C terminus, the carboxylgroup is free. Since it is mostly unprotonated and uncharged, the C-terminus may also be named carboxyterminal.

Cadherins are calcium-dependent cell adhesion proteins that mainly participate in interactions with cadherins, located on other cells. Cadherins have five similar extracellular domains some of which have calcium-binding sites, and an intracellular C-terminal domain that interacts with the actin cytoskeleton. Different cadherins are named according to the tissues were they were first found. E-cadherins were found in many epithelial cells. N-cadherins are present in nerves and muscle and P-cadherin are present in placental cells.

Calcineurin is the serine/threonine protein phosphatase 2B. It binds calmodulin and its activity is regulated by Ca^{2+} .

Calmodulin is a small Ca^{2+} -binding regulator protein that brings the activity of several cellular proteinkinases and phosphatases and other enzymes under the control of Ca^{2+}. It has an EF-hand and binds four calcium ions per molecule.

CAM, is the name for a cell adhesion molecule on the surface of animal cells. N-CAM is located on neurons.

Carcinomas are cancers of epithelial cells. They are the most common human cancers.

Caveolae. are closed vesicles, filled with lipids and proteins which are attached to plasma membranes. They are enriched with glycosyl phoshatidylinositol-, GPI-anchored proteins and are assumed to transport lipids and proteins.

Caveolin is a protein that lines the cytoplasmic surface of caveolae.

CD markers are antigenically distinct proteins on the surface of cells which are used for the characterization of such cells. More than 80 individual markers have been identified.

For example, CD3 is present on all T cells and is part of the T cell antigen receptor-complex. CD4 occurs on T helper cells which participate in MHC class II-mediated reactions. CD8 occurs on cytotoxic killer T cells which carry out MHC class I-mediated reactions.

Cellcortex is an actin-rich layer on the inner, cytosolic face of the cell membrane, responsible for the plasticity of the cell surface.

Cellcycle is the sum of the processes occuring when a mother cell divides, forming two daughter cell. The cellcycle is divided into G1, the period between the end of cell division and the start of DNA synthesis in the S phase. The S phase is followed by the G2 phase , which lasts until the M phase, mitosis, begins. Differentiated cells often enter a G_0 phase, a quiescent phase where they withdraw from the cell cycle and rest.

Celldivision in eukaryotic cells is divided into division of the nucleus, mitosis, and division of the cell, cytokinesis. Mitosis continues from the early prophase to the late telophase, when cytokinesis begins.

Cdc genes, cell division cycle genes. A family of genes, encoding proteins participating in cell division. They were originally identified in yeast, using temperature sensitive mutants. Homologues of these genes have been found in many other species, including man. The proteins encoded by these genes are classified as Cdc's followed by a number

Cdk's are cyclin dependent serine threonine protein kinases. There are several Cdk's. Cdk1 is identical with cdc2, the kinase encoded by cdc2. Mammalian Cdk1 is associated with both A and B type cyclins.

Centrosome is a structure in eukaryotic cells made up mainly of microtubules. It is made up of a pair of cylindric centrioles, the satellite bodies, and a cytoplasmic zone close to the nucleus where the microtubules during mitosis are organized forming the spindle pole. The structure of the centrosome appears to change continually during cell division.

Centromere is that site on the chromosome that is attached to the mitotic spindle during divison of the nucleus in mitosis. The centromere is also the point in the condensed chromosome where the two sister chromatids are connected. In the late prophase, the centromere has two kinetochores, one for each sister chromatid.

Chemotaxis. The movement of cells in response to concentration gradients of attractants.

Chaperonines are proteins that help nascent proteins to fold correctly.

Chaperones are proteins, often heat shock and stress proteins, that protect post-translationally proteins against misfolding and denaturation.

Chlathrin. Chlathrin is a large, 180kD protein which is the main constituent of the coat of the chlathrin-coated pits in the plasma membrane, where they are involved in eukaryotic cells in transfer of proteins between the plasmamembrane and endoplasmic vesicles.

Choleratoxin is an enterotoxin produced by the bacterium *Vibrio cholerae*. On infection, one of the toxin's subunits, the B subunit binds to a ganglioside, GM1, in the membrane of sensitive cells. On binding, the other subunit, A, is released. The A subunit is an ADP-ribosyltransferase which transfers ADP-ribose from NAD^+ to the

α-subunit of a heterotrimeric G protein, Gs. ADP-ribosylation inhibits the GTPase activity of αs and consequently the G(s= stimulating) α protein remains irreversibly in the active, GTP-bound state which activates adenylylcylase. The product of the reaction, cAMP, which is produced in large amounts and in an uncontrolled manner is then responsible for the severe disturbance of the water and the electrolyte balance in the intestine of the patient.

Chromatid is one copy of a newly replicated chromosome, still attached at the centromere to the other sister copy.

Chromatin, is a complex of DNA, histones and other non-histone proteins. The chromosomes in the nucleus of eukaryotic cells exist as chromatin complexes.

Chromosomes are composed of a very long DNA molecule with associated proteins. They carry the genetic information of eukaryotic cells. A diploid somatic human cell carries 46 chromosomes, 23 pairs, of which one pair are sex chromosomes, either XY or XX.

Cis means on this side. In the chemical nomenclature *cis* means that two substituents of a molecule lie on the same side of the reference plane. In genetics *cis* means that a regulatory element, a promoter or operator is located on the same side, or on the same chromosome, where the regulated genes lie. Accordingly, *trans* means on the other side, and a *trans*-acting regulatory gene element can lie on the other side or on an other chromosome, where the regulated genes are. In other words, the location of a *trans*-acting gene regulator is irrelevant for its effects on the regulated genes.

Clonal selection, is the essence of a theory to explain the ability of the immune system to produce specific antibodies against practically any conceivable structure of an immunogen, (antigen). Sir Francis Macfarlane Burnet is given credit for this theory. According to his theory, the immune system has preexisting clones of cells, eqipped with receptors that can recognize every conceivable antigenic structure. On encounter with the antigen, the preexisting clone, equipped with the matching receptor is selected by the immunogen and proliferates, producing a great number of the same cell, producing the same antibody. (One cell, one antibody. Clonal expansion). A few cells of the clone remain after the encounter with the antigen. They are the memory cells, which are preconditioned and remain responsive. On a second encounter with the same immunogen, they react immediately and more vigorouly, accounting for a greater secondary response than the first response of the naive lymphocytes. Such a secondary response occurs for example after a successsfull vaccination.

Clone , is a group of genetically identical cells.

Collagen is a group of fibrous proteins of high tensile strenght that are the main-components of connective tissue in animals. They are also present in the extracellular matrix. They are glycine- and proline-rich proteins. There are many classes and types of collagen. Class I collagens are all fibre-forming. They are expressed in skin, tendon, and bone. Other collagens are found in cartilage tissue.

Complement is a complex system of serum proteins, mostly β-and γ-globulins. They are designated by the letter C and a number. They are activated by antigen-antibody complexes. The complement system is part of the immune response and helps to eliminate pathogenic organisms either by directly causing their lysis or by initiating phago-

cytosis. It is also involved in immunological tissue inflammation and injury.

Consensus sequence is a DNA, RNA or protein sequence which is functionally important and conserved and present in many related molecules.

CRE, the cyclic AMP response element is present in many gene promoters. It has the sequence; GTGACGT[A/G][A/G]. When it associates with the CRE-binding protein, CREB, the genes controlled by the CRE- promoter are turned on.

CREB , the CRE-binding protein is encoded by the *crb's* , a family of genes, encoding transcription factors. The CREB's are activated by phosphorylation by cAMP-dependent kinase and bind as dimers to CRE.

Crk is a protooncogene, related to *v-crk,* the oncogene of the avian sarcoma virus CT10. Crk is the abbreviation of CT regulator of kinase. Although, Crk induces tyrosine phosphorylation, it is not a tyrosine kinase. It is actually a linker, containing SH2 and SH3 domains which links up with tyrosine kinases.

Cyclophilin is a small protein that binds to the immunosuppressant cyclosporin A, a cyclic undecapeptide.

Cytokinesis is the sum of all processes involved in cell division that follow mitosis and division of the nucleus.

Cytoskeleton is defined as the sum of the various filamentous proteins of eukaryotic cells that remain after the cells are extracted with a mild detergent. The cytoskeleton includes actin filaments, two-stranded helical polymers, which form the microfilaments and the actin-binding proteins. Other components are microtubules and intermediate filaments. The cytoskeleton has not only a role in maintaining the shape of cells, it is also actively engaged in cell division, in the organisation and the dynamic movement of cell organelles and in the movement of cells in chemotaxis.

Cytokeratin belongs to a family of intermediate filament proteins in epithelial cells.

Dalton. Dalton, Da, is a molecular mass unit, widely used in Biochemistry. A dalton is equivalent to the mass of a hydrogen atom, 1.66×10^{-24} g. The relative molecular mass or molecular weight, Mr, is a mass ratio, expressed by a number without dimensions. Mr 10.000 would be equivalent to 10 000 Da or 10 kDa.

Dendrogram is a tree-like diagram that shows relationships between organisms or other entities.

Dendritic cells have protoplasmic protuberances with which they contact other cells.They play important roles in the immune and the nervous system.

Dephosphorylation of protein-bound phosphates is catalyzed by specific protein phosphatases which transfer phosphate groups, covalently bound to tyrosine-OH- or to serine /threonine OH-groups to H_2O.

Desmosomes are intercellular junctions, consisting of two cell membranes, separated by an interspace. The interspace contains dense fibrillar plaques rich in cell adhesion proteins of the cadherin type. Desmosomes are important for cell-cell adhesion and communication. Epithelial tissue in the body stretched mechanically is rich on desmosomes.

Dictyostelium discoideum is a cellular slime mold, an amoeba. It is a preferred object for the study of cell locomotion, chemotaxis and differentiation.

Dissociation constant, Kd, is a measure of the tendency of a complex to dissociate. For an equilibrium, A+B→AB, Kd is given by the ratio of [A][B]/[AB]. The smaller Kd the tighter the binding between A and B.

Disulfide bonds , -S-S- bonds are covalent bonds between the sulfhydryl, SH groups of cysteines. Intermolecular -S-S- bonds link two separate polypeptide chains and intramolecular -S-S- bonds are within the same polypeptidechain.

DNA cloning involves insertion of a DNA molecule into a vector. (Bacterial plasmids are often used as cloning vectors. Plasmids are extrachromosomal double stranded DNA molecules). The host cells are transfected with the vector and the plasmid-infected cells are selected and propagated. The DNA to be cloned may be a fragment of a genomic DNA or a complementary DNA, cDNA. (cDNA is a single stranded DNA molecule that has a complementary base sequence to a messenger RNA. It is produced from isolated mRNA by reverse transcriptase. cDNA unlike genomic DNA contains no introns).

Downstream. are coding sequences more towards the 3′ end of the DNA strand. By convention DNA sequences are red from the 5′ end to the 3′ end.

Drosophila melanogaster, is a fruit fly and the preferred object of genetic studies of development

Dynamin is a microtubule-associated GTP-binding protein. Dynamins are activated by proteinkinase C. They have a PH, a pleckstrin-homology domain and bind inositol phospholipids which activate the GTPase of dynamin. Dynamins are involved in microtubule assembly and vesicular traffic.

Ecdysone is a steroid that is formed and secreted by the prothoracic glands of immature insects but also by the ovaries of adult female insects. It was first isolated from the silk worm, *Bombix mori*. Ecdysone is the prohormone of the moulting hormone ecdysterone, but the prohormone has also effects on its own at other stages of insect development.

ECM, extracellular matrix is a complex network made up of polysaccharides, such as glycosaminoglycans or cellulose and proteins such as collagen, which are secreted by cells.

Ectoderm is the outermost layer of the three germ layers in the animal embryo. It develops mainly into the epidermis, the skin and the neuronal tissues.

EF-hand EF-hand is a helix-turn helix motif which is found in many Ca^{2+}-binding proteins. To envisage an EF-hand, look at the inside of your right hand with thumb and forefinger extended at 90°. The thumb will then point towards the C-terminus of the helix F of the reference protein, parvalbumin, and the forefinger will point along helix E of parvalbumin in direction to the N-terminus, whereas the clenched fingers represent the loop, connecting E and F which wraps around the bound Ca^{2+}.

Endoderm is the innermost germ layer in the animal embryo which develops into the lining of the gut from pharynx to rectum, including the liver, the pancreas and related organs.

Endoglin is a major glycoprotein of the vascular endothelium. It is a homodimer of disulfide-linked subunits. It forms a complex with the TGF-ß receptor. It has a RGD, Arg-Gly-Asp-sequence which is recognized by integrins. Therefore, endoglin may not only play a role in the regulation of TGF-ß signalling but it may also mediate cell adhesion by promoting the binding of endothelial cells to integrin.

Enhancer is a regulatory DNA sequence to which proteins bind that regulate the rate of transcription of structural genes. Enhancer sequences are often many thousand base pairs away from the regulated genes.

EPH: A subfamily of RTK's, receptor tyrosine kinases. Related proteins are ELK , EPH-Like Kinase, ERK, ELK-Related Kinase. and ECK. epithelial Cell Kinase. Overexpresion of EPH causes cancer.

epigenetic are all processes, modifying gene expression without changing the coding sequence. Epigenetic are posttranscriptional modifications of proteins by phosphorylation, glycosylation and so forth.

Epitope. Originally, an epitope was described as a structural element, recognized by an antibody. Now it means any structural determinant on the surface of a protein which is recognized by an other protein.

ERB A and B are genes originally found in an avian erythroblastosis virus. *ERB A* is fused to *ERB B* . *ERB B* encodes a truncated EGF-receptor.

Erythropoietin is a potent stimulator of growth and differentiation of erythroblasts, erythrocyte precursor cells.

Expressionvectors are artificial constructs containing foreign DNA-genes which can be expressed and transcribed when introduced into host cells. To promote expression of the foreign genes, vectors usually contain potent promoter sequences.

Fallopian tube is the oviduct in the mammalian ovary.

FGR is a member of the Src family nonreceptor tyrosine kinases. The oncogenic FGR is a product of the *v-fgr* gene of the feline sarcoma virus Gardner-Rasheed.

FK506 is the propriety name given to an immunosuppressant by a japanese company. FK506 is a macrolide antibiotic, isolated from streptomyces tsukubaensis. It is a potent inhibitor of T cell activation, preventing allograft rejection. FK506 binds to FK-binding proteins, the FKBP's.

FKBP , for example, the FK506-binding protein is a member of the family of immunophilins. Immunophilins are receptor-like small binding proteins that participate in T cell activation. These small proteins have peptidyl-prolyl-isomerase activity. Binding of the complex of FK506 with its FKBP blocks the action of calcineurin, a phosphatase which is involved in T cell activation.

Focal adhesion plaque is the region on the surface of a cell which anchors it to the extracellular matrix. The attachment is made by transmembrane proteins, such as the integrins.

Fos was originally found in murine osteosarcoma viruses. The product of the cellular *fos* proto-oncogene, Fos, is a growth-promoting transcriptionfactor, found in the nucleus of cells, stimulated by growth factors.

Gametes are haploid reproductive male or female cells. Gametes with opposite sex fuse with each other, forming a zygote.

Gastrula is an early stage of development of an animal embryo, following the blastula stage. In the gastrula, the cells begin to invaginate to form the rudiments of a gut cavity.

GATA or **ERYF 1** is a transcriptional activator for the development of erythroid cells. It is named GATA, because it binds to the consensus sequence [A/T] GATA-[A/G] in genes expressed in erythroid cells.

Genome is the whole of the genetic information of an organism.

G proteins are GTP-binding proteins. There are two major classes of G proteins. Monomeric and heterotrimeric G proteins. An example of a monomeric G protein is ras, the product of *c-ras*, a cellular protooncogene, related to the rat sarcoma oncogene. The other family of G proteins are the hetrotrimeric α,β,γ–G proteins which couple to heptahelical receptors. Both exist in an active GTP-bound state and an inactive GDP-bound state. On binding to a ligand-activated receptor, heterotrimeric α,β,γ–G proteins dissociate into a GTP-bound α-subunit and a β,γ -complex. Both transmit signals to targets, either independently or synergistically. The GTP-bound α-subunit is slowly converted by its intrinsic GTPase activity to the inactive GDP-bound state, whereas the monomeric G proteins have practically no intrinsic GTPase activity. Therefore, regulation of the activity of monomeric G proteins requires a separate set of GTPase-activating proteins, the GAP's. The conversion of the inactive GDP-bound state to the active GTP-bound state is then regulated by other factors, which either accelerate or inhibit the dissociation of GDP and its replacement by GTP. These factors are the GDP exchange factors, the GEF's or the GDI's, the GDP dissociation inhibitors. In the case of the receptor-coupled heterotrimeric G proteins, it is the activated receptor which carries out most effectively GDP-GTP exchange.

G protein-coupled receptors are a large class of transmembrane receptors. They have seven transmembrane helices. Therefore, they are called heptahelical or serpentine receptors. They bind watersoluble hormones, such as adrenaline but also peptides and accept sensory signals, light, odorants and some taste stimuli. On binding the ligand they transmit the signal to heterotrimeric, α,β,γ–G proteins.

General transcription factors belong to a group of proteins which assemble at the TATA box or a similar region in the promoter.They are required for the initiation of transcription by the DNA-dependent RNA polymerase in eukaryotic cells.

Golgi apparatus is a membrane-enclosed organelle in the endoplasmic reticulum of eukaryotic cells where proteins are modified post-transcriptionally, mainly by glycosylation, and where proteins and lipids are sorted for transport to their destinations.

Glycosylaminoglycan is a long chain consisting of pairs of sugars of which one is always an aminosugar. Examples are heparin, hyaluronic acid and chondroitin sulfate. Glycosylaminoglycans are covalently linked to proteins, forming proteoglycans in cartilage and connective tissues.

Hematopoiesis is the process generating bloodcells. It takes place mainly in the bone marrow.

Heterochromatin is that part of the chromatin in the eukaryotic cell where the chromosomes remain condensed and transcriptionally inactive in the interphase, even when

the cell is not in mitosis. Heterochromatin is replicated at a later stage of the cell divison cycle than euchromatin, the part of the chromatin with the transcriptionally-active chromosome.

Heterozygote is any diploid cell or organism, containing two different alleles of a gene.

Hill coefficient is a measure of cooperativity of ligandbinding sites in a protein. The Hill coefficient is estimated from a plot of the Hill equation, where log [y/(1-y)] is plotted against log[x], and where y is the fractional saturation of the protein with ligand and [x] gives the concentration of the free, (un-bound), ligand. The slope, h, is the Hill coefficient. A value of h greater than 1 gives an estimate of the extent of cooperation among the ligand-binding sites on the molecule. (The Hill coefficient is named after the English physiologist A. V. Hill).

Histones belong to a group of small conserved proteins, rich in basic aminoacids, arginine and lysine, which are associated with DNA in eukaryotic cells.

Homozygote is any diploid cell or organism, containing two identical alleles of a gene.

Homeobox is a short, 180 base pair long DNA sequence present in homeotic genes. The homeobox encodes transcriptional regulators with characteristic DNA-binding motifs.

Homeotic genes, homeobox-containing genes, abbreviated *Hox* genes. Homeotic genes were first discovered in *Drosophila.* which has two clusters of *Hox* genes. In vertebrates there are four such clusters of 9-11 *Hox* genes, each cluster located on a separate chromosome, spanning about180 kbases of DNA. The extent of homology of the *Hox* genes in different species is high. They code for the homeodomain- or homeobox-recognizing proteins. These are transcription factors. Both structure of homeobox-recognizing proteins and their interaction with DNA has been determined. The *Hox* genes and the corresponding proteins are expressed at particular stages in embryogenesis. Homeotic genes control the formation of whole parts of the body in the course of embryonic development. Mutations of *Hox* genes therefore result in the conversion of one part of the body into an another, for example in *Drosophila* in replacement of an antenna by a leg. The expression of *Hox* genes is controlled by retinoids.

hGH, human growth hormone. To the family of genes encoding hGH, belong also the genes coding for choriomammotropin and prolactin. Growth hormone is identical with somatotropic hormone, STH, or somatotropin.

Hypothalamus is a region of the human brain lying below the thalamus at the floor of the third ventricle. It lies just above the pituitary gland which it supplies with various regulatory factors.

Hyperpolarization is due to changes in the charge difference across the membrane and results in a change of the resting membrane potential.

Immortalization is the formation of a cell line which can divide forever. Cancer cells are an example.

Imprinting is a measure of the extent to which maternally- and paternally-inherited genes are expressed. Imprinted genes have different patterns of methylation of cytosines in maternally- and paternally-inherited genes. Imprinting of genes plays a role in the inheritance of diseases, for example in Huntington's disease, Huntington's chorea.

Immunophilins are proteins that participate in T cell activation and bind immunosuppressants such as FK506 and rapamycin. A related group of proteins are the cyclophilins which bind cyclosporin A. On binding immunosuppressants the immunophilins block T cell activation.

Induced fit is the essence of a theory proposed by Daniel E. Koshland jr. It postulates that a substrate (or a ligand) induces on binding a conformational change in the enzyme or protein to which it binds. This ligand-induced reorientation fits the ligand to the binding site and makes the protein reactive. The induced fit is a prerequisite for an enzyme's action on the substrate or for ligand-dependent activation of a protein, in general.

Inflammatory response of a tissue is caused by invasion of white blood cells, which release histamine and other substances.

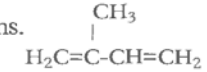
Integrins are members of a large family of transmembrane proteins that act as receptors for cell-adhesion molecules. Integrins are heteromeric proteins. The type of integrin expressed on the surface of a cell determines what kind of adhesion molecules and hence which cells can be bound.

Interconvertible enzymes are enzymes that exist in at least two structurally and functionally different forms that are interconvertible. These forms may arise from covalent attachment of phosphoryl-or other groups, of which attachment and removal is catalyzed by separate specific enzymes. Different forms may also represent different conformations formed on reversibly binding a ligand at an 'allosteric' site, a site other, than the catalytic or substrate-binding site. A prime example are the different forms of hemoglobin, in the unliganded de-oxy- and the liganded oxy-state, as shown by X-ray crystallography by the English molecular biologist, Max Perutz.

Interferons , IFN's belong to a group of cytokines of Mr in the range of 15.000 to 30.000. There are three types of interferons, INF-α, INF-β and INF-γ. They are released by cells on exposure to inducing agents, such as viruses. A special property of the INF's is their anti-viral activity.

Interleukins , IL's belong to a heterogeneous group of cytokines. Up to now there are about 15 IL's and more can be expected. They, notably IL-2, have an essential role in the regulation of proliferation of B and T cells in the immune system.

Islets of Langerhans are clusters of endocrine cells in the pancreas. One distinguishes the glucagon-secreting A cells, the insulin-secreting B cells and the somatostatin-secreting D cells. (Named after Paul Langerhans who first described such clusters of cells as a medical student in 1869).

Isoprenes are unsaturated hydrocarbons with 5 C atoms.  They are the structural unit of isoprenoids.

Jun is a gene family encoding transcriptionfactors. v-jun is an oncogene from the avian sarcoma virus, ASV17. (Since 17 is *junana* in japanese, this oncogene was named *jun*.) The product of this gene is Jun which forms heterodimers with Fos which like the homodimers of Jun bind to the AP1 consensus site of early genes which promote cell growth and proliferation. Jun is a typical DNA-binding leucin zipper protein. DNA-binding and dimerization of Jun is regulated by serine /threonine phosphorylation, dephosphorylation.

Juvenile hormones are insect hormones that keep the insect in the juvenile, (larval) stage. They are farnesoic acids, acids of the sesquiterpene alcohol farnesol.

Kinetochore is a complex structure, formed in mitosis from chromatin. The kinetochore recruits microtubules from the mitotic spindle which move the chromosome to the poles of the cell.

Kit is a family of genes, encoding tyrosinekinase receptors. The product of the *c-kit* protooncogene, KIT is a receptor for the stemcell factor, or steel factor. Binding leads to activation of the receptor, autophosphorylation and dimerization and association with phosphatidylinositol 3-kinase and phospholipase Cγ. KIT and its ligand are highly expressed in small cell lung cancer and in testicular germ cell tumours.

Leptin is a protein, encoded by the *ob* gene, first identified in obese mice. The *ob* gene is apparently expressed exclusively in white adipose tissue. Leptin binds to a receptor in the hypothalamus.

Linkage means usually that two genetic loci that lie near each other on the same chromosome are inherited together.

Locus is the position of a gene on a chromosome. Alleles of the same gene occupy the same locus on the chromosomes.

Loss of heterozygosity, (LOH), is a situation which develops, when one allele of a gene has been lost. Sometimes the remaining allele is duplicated. When this involves large sequences which are simultaneously duplicated, a long stretch of DNA may be formed with many copies of the same, homozygous allele.

Mating of yeast. There are two mating types in the yeast, *saccharomyces cerevisiae*, a and **a,** which are both necessary for mating, the fusion of haploid cells and the formation of diploid cells. The mating type, MAT, is determined by the Mat locus on the DNA. The MATα2 protein encoded by the corresponding MAT locus resembles a homeobox-recognizing protein. It is a master regulator, controlling gene expression in yeast.

MAD is a family of genes, regulating mitosis and cell proliferation. (Mitotic arrest deficiency genes). The MAD proteins encoded by these genes form complexes with the transcription factor Myc.

MAX is a family of genes coding for Max which forms heterodimers with Myc and controls cell proliferation. Max has a helix turn helix motif.

Meiosis is a special form of cell division when sperm and egg cell combine. Meiosis results in the reduction of genetic material. It is the sum of the processes which are involved in the division of a diploid germ cell and the formation of four daughter cells, each containing half the number of chromosomes of the parent cell. Therefore, this process is also called reduction division. Meiosis involves two rounds of cell division. In the 'leptotene', each chromosome is duplicated, forming two identical sister chromatids. In the 'zygotene' each chromatid pairs with its homologue, forming the synaptonemal complex. When synapsis of the chromosomes is completed, the 'pachitene' starts. During this phase crossing over occurs between the partner chromatids. The final phase is diakinesis, where two daughter cells are produced. Although at this stage each cell has a copy of each chromosome but each chromosome is still made up of two sister chromatids. Therefore a second division follows where the two sister chromatids are separated, but without replication.

Melanocortins. This is a generic name for the peptide hormones, melanotropin and corticotropin, (ACTH, adrenocorticotropic hormone), because both hormones are formed in the anterior pituitary gland from the same melanocortin precursor. Melanotropin controls melanocyte growth and pigmentation. Corticotropin stimulates the production of glucocorticoids and mineralocorticoids, such as aldosteron, in the adrenal cortex. The Melanocortin-pathway is somehow involved in the control of appetite and body weight.

Melanocytes are cells that produce the dark pigment, melanin.

Melanocyte-stimulating hormone. MSH, is produced in the intermediary lobe of the pituitary. Its function in mammals and humans is not clear. MSH binds to the melanocortin receptor. These receptors are G protein coupled and activate adenylylcyclase.

Membranepotential is the electric potential across a membrane. It results from the difference in charges at either side of the membrane. The resting potential varies in animal cells, but is usually about -70mV. A Na^+/K^+-ATPase exchanges intracellular Na^+ for extracellular K^+ and is responsible for the maintenance of low intracellular Na^+ concentrations.

MEN , multiple endocrine neoplasias, including thyroid carcinomas and tumours of the adrenal medulla,(phaeochromocytoma).

Mendelian inheritance implies that every cell with the exception of germ cells has a pair of allelic genes for each character. The two genes in each pair separate in meiosis, so that each gamete has only one gene. Separation of any pair of allelic genes is not linked to separation of any other pair, except the genes are linked.

Mesoderm the middle of the three primary germ layers of an animal embryo. It is the source of bone, muscle, connective tissue and other structures.

Metaphase is the stage of mitosis at which chromosomes are already attached to the mitotic spindle at the equator, but have not segregated and are not yet assembled at the opposite poles of the cell.

Metazoa is a kingdom to which all multicellular animals belong whose cells are organized into tissues with specific functions and which are controlled by a nervous system.

MHC is the abbreviation of major histocompatibiliy complex, a complex of genes that encode a family of cell surface proteins, (iso-antigens). In humans, these proteins are named leukocyte-associated antigens; HLA. These iso- antigens are of central importance in tissue grafting. HLA- incompatible grafts are more likely to be rejected. The HLA- or MHC- proteins are glycoproteins. Two classes, Class I and Class II are distinguished:

MHC Class I proteins are present on the surface of most nucleated somatic cells. The 45 kDa α-subunit is a trans-membrane protein. It is a single polypeptide chain, with a short hydrophobic region inserted in the membrane and a short hydrophilic C-terminal region, extending into the interior of the cell, and a large, glycosylated N-terminal region at the outside. The external region is folded into three separate domains and is bound to the 11.5 kDa β-subunit. The β-subunit is also named β_2-microglobulin. The β-subunit is homologous to a single immunoglobulin like Ig-domain.

MHC Class II proteins are more restricted and are found only on the surface of certain cell types, such as B lymphocytes, some T lymphocytes and some macrophages and macrophage-like cells. They have a 33 kDa α-chain and a 28 kDa β-chain. Both α- and β-subunits have N-terminal extracellular, transmembrane and C-terminal intracellular domains.

Michaelis constant, Km is named after the German biochemist, Leonor Michaelis. $Km = (k_{-1} + k_{cat}) / k_{+1}$; When $k_{cat} << k_{+1}$, the Km approximates Ks, the enzyme-substrate dissociation constant.

Micron is a unit of measurement, often applied to cells. 1 micron is equal to 10^{-6} meters.

Mitogen is a factor that promotes mitosis and proliferation of cells.

Mitosis is the division of the nucleus of eukaryotic cells to form two daughter nuclei that contain the same number of chromosomes and are genetically identical with the parent nucleus.

MYC is a family of genes that encode transcription factors. The cellular protooncogene *c-myc* is one of seven related genes.

Myo D is the gene for the myoblast determining protein. It is related to MYC. It also encodes a transcription factor.

Nerve growth factor, NGF, was the first member of a family of growth factors of nerve cells. It was discovered by the Italian neurobiologist Rita Levi-Montalcini. Its active form is a dimer of the β-subunits.

Neu. The neu oncogene was found in a rat neuroblastoma. The neu-protooncogene encodes a EGF-like receptor with tyrosinekinase activity. The Neu receptor binds EGF and a specific Neu-activating factor. Neu may play a role in neurogenesis.

Neurofibromin. Neurofibromin is a GAP, a GTPase-activating protein for Ras. It is implicated in the pathogenesis of neurofibromatosis.

Neuropeptide Y. Neuropeptide Y is identical with melanostatin. It is one of the most abundant neuropeptides. It stimulates food intake. Leptin blocks Neuropeptide Y gene expression and its release from the hypothalamus. But whether this effect is responsible for a role of leptin in weight control and prevention of obesity is not clear.

Neuropeptide Y receptor. The receptors for neuropeptide Y are G protein-coupled receptors.

Neurotrophins, a group of neurotrophic growth factors.

NMR, nuclear magnetic resonance is the resonance due to the capture of electromagnetic radiation by the atomic nuclei at a specific frequency in a magnetic field. The resonant absorption by the nuclei is due to the flipping of the orientation of their magnetic dipole. The NMR spectrum gives information on the chemical environment of the nuclei. Two dimensional NMR is used to determine the solution structure of small proteins, up to about 20kDa.

Notch. The *notch* gene in *Drosophila* encodes a receptor-like transmembrane protein with EGF-like repeats. Notch is essential for ectoderm differentiation.

Nuclear receptors. Intracellular DNA-binding proteins which bind lipophilic ligands which pass the membrane without assistance. Examples are the steroid hormone receptors and the receptors for non-steroidal ligands, such as retinoids and the hormone thyroxin.

Nucleosome is a fundamental subunit of chromatin in eukaryotes. It contains a short stretch of DNA wrapped around a core of histones

Oncogene is a gene causing cancer. The first oncogene was isolated from a transforming retrovirus, *src*, the Rous sarcoma virus, by Peter Duisberg and Peter Vogt in 1970. *src* encodes a 60 kDa proteintyrosinekinase, pp60src. M.J. Bishop and H.E. Varmus then found that oncogenes originate from normal cellular genes, the protooncogenes by chance recombination on retroviral infection. Protooncogenes are mostly involved in growth control. When cellular protooncogenes are mutated or overexpressed they become oncogenes.

Orphan receptors: Receptors of which the ligands are not known.

Palindromic sequences are nucleotide sequences in a DNA strand that are identical to the sequences in the complementary DNA strand, when both are read in the same direction, either from the 5′ or the 3′ direction. (Note that in double-stranded DNA one strand runs from 5′ to 3′ and the other complementary strand from 3′ to 5′).

Parazoa are sponges.

Peroxisome is a cellular organelle where proteins and other molecules are oxidized. It contains enzymes that form and degrade hydrogen peroxide, (H_2O_2).

Phagocytosis is the process by which phagocytic cells ingest particulate materials by endocytosis.

Pheromones are factors that bind to G protein coupled receptors and regulate the cell cycle and mating in haploid yeast cells.

Phospholipid is the major family of lipid molecules that are found in biological membranes. They usually contain two fattyacids, one saturated and the other unsaturated, esterified with hydroxyl groups of glycerophosphate, which is bound, also through a hydroxyl group to one of a variety of polar head groups, such as choline, ethanolamine, serine and sphingomyelin.

Phosphorylasekinase phosphorylates and activates glyogen phosphorylase which degrades phosphorolytically glycogen to glucose-1-phosphate. It catalyses the reaction; 4 ATP+ 2 phosphorylase ***b*** dimers ⇒ 4 ADP + phosphorylase ***a*** tetramer. The activity of the kinase is regulated by phosphorylation by the cAMP-dependent protein kinase or by dephosphorylation by the serine/threonine proteinphosphatase 1 and by Ca^{2+}. As shown by Philip Cohen, Ca^{2+}-regulation is mediated by the δ-subunit of the kinase which is identical with Ca^{2+}-binding calmodulin.

Phosphorylation is the process of introducing enzymatically a phosphoric acid residue into a protein from ATP as a phosphoryldonor. This results in the formation of a covalent phosphoric acid ester with the OH-groups of serine/threonine or with the phenolic group of tyrosine residues, or the formation of a phosphamide in N-phosphohistidine.

Photon is an elementary particle of light and other electromagnetic radiation.

Platelet or thrombocyte is a blood cell, lacking a nucleus. It binds to injured vessel walls and initiates bloodclotting.

Pleckstrin is a protein in platelets and a substrate of the serine/threonine protein kinase C. It contains a structural motif, the PH-, the pleckstrin-homology domain.

Point mutation is the change in a single nucleotide in a coding region of the DNA, often leading to the exchange of a single aminoacid by another aminoacid in the protein.

Polyploid is a diploid cell, containing more than two sets of homologous chromosomes.

Positional information determines the position of cells relative to other cells in a multicellular, metazoan organism.

Posttranslational modifications are enzyme-catalyzed covalent modifications of a mature protein after it has been synthesized. Examples of posttranslational modifications are phosphorylation, glycosylation, sulfation, methylation and prenylation. Especially those modifications that are reversible, such as phosphorylation by de-phosphorylation through the action of phosphatases are important in regulation.

Promoter is the DNA sequence 5′ to the gene which is transcribed. It is the sequence to which the RNA polymerase binds and where transcription starts.

14-3-3 protein is any member of a highly conserved family of proteins, involved in signaltransduction in higher eukaryotes. The name comes from the position of the bands of these proteins in gel electrophoresis. Members of this group of proteins have been implicated in the regulation of proteinkinase C and in interactions with the Raf kinase and other kinases.

Proteasome is a large multisubunit complex in the cytosol that degrades proteins, marked by ubiquitination. Proteolysis requires ATP.

Proteomics. The elucidation of the complete genetically determined expression profile of an organism, including composition, structure and function of all proteins expressed in the adult organism.

Proteinkinase A, cyclic AMP-dependent serine/threonine proteinkinase. Activation by cAMP produces two active catalytic monomers and a regulatory dimer, binding four molecules of cAMP. Type I and type II kinases are distinguished. In the case of the type II kinase, the regulatory chains mediate membrane association and interaction with MAP kinase and AKAP, the A kinase anchoring protein.

Proteinkinase C , PKC, a family of serine/threonine proteinkinases which require an anionic phospholipid for activity and are regulated by diacylglycerol and Ca^{2+}. Phorbolesters which are tumour promoters can substitute for diacylglycerol.

Proteoglycans are proteins with glycosylaminoglycan chains.

Protozoa are eukaryotic single-celled organisms.

Quaternary structure is the three-dimensional arrangement of several identical or different polypeptide chains in a multi-subunit protein complex. One distinguishes between homologous and heterologous complexes and between dimeric, trimeric, tetrameric and oligomeric structures.

RAF's are serine/threonine kinases, encoded by protooncogens. They share common structural properties with proteinkinase C, both in their conserved C-terminal catalytic

and their regulatory, N-terminal domains. The c-Raf 1 kinase is a 74 kDa protein and has been found in all mammalian cells.

RAFT's are regions in the plasma membrane which are thought to transport and concentrate lipids and proteins. They are enriched with lipids, notably cholesterol, and proteins. They could be scaffolds facilitating the assembly of signalling complexes in the membrane.

Recessive is that member of a pair of alleles that is not expressed in an organism which expresses only a dominant allele.

Recoverin. Recoverin is a small, 26kDa Ca^{2+} binding protein with a typical EF-hand. It is localized in the retina and the pineal gland and has been implicated in the activation of rod guanylate cyclase by rhodopsin and the formation of cGMP. Recoverin is an autoantigen in an autoimmune retinopathy.

REL is a family of genes, encoding transcription factors of the κB family. *v-rel* is the oncogene of an avian reticuloendothelial virus. Rel encodes proteins which form homo- and heterodimers that bind to genes, which are also adressed by NF-κB.

RET is a human protooncogene which encodes the RET receptor. There are two forms of receptors with tyrosinekinase activity. They are orphan receptors with unknown ligands. Oncogenic variants are found which are constitutively phosphorylated on tyrosines and form homo-and heterodimers. These variants have been found in endocrine disorders, MEN, and in Hirschsprung's disease.

Retinitis pigmentosa is a genetic disease which affects 1 in 4000 individuals. Tunnelvision and eventual blindness are caused by the premature death of photoreceptors.

Retinoblastoma is a childhood cancer of the developping retina. The deleted gene, responsible for the disease encodes a protein, pRB, which is like a tumoursuppressor. It interacts with various transcription factors and suppresses cellular proliferation.

Retinoic acid receptors bind retinoic acid-, (vitamin A acid), derivatives and regulate gene expression.

RTK. Membrane-bound receptor with tyrosine kinase activity. These receptors pass the membrane only once. On binding ligand they autophosphorylate their own tyrosine residues and form dimers. The phosphorylated tyrosine residues are recognized by intracellular proteins. Examples are growthfactor receptors.

Saccharomyces are yeasts that are used for beer brewing and baking. Yeasts are unicellular fungi. They reproduce either by budding or by conjugation and are widely used as objects for the study of genetics and eukaryotic cell biology.

SHC is named a gene family that encodes SH2-containing proteins which are mostly linkers.

Somatic cell, any cell in an animal or human other than a germ cell.

Sos , *son of sevenless* , is a Drosophila gene that encodes a guanine nucleotide release factor, which activates Ras. It is involved in neuronal development in Drosophila.

Spectrin is the major constituent of the erythrocyte cytoskeleton.

Splicing is the enzymic process in eukaryotic cells by which non-coding introns are excised from the primary, heterogeneous RNA, (hn RNA) in the nucleus, after transcription of RNA from DNA is completed. The cut ends are then rejoined, spliced, forming messenger RNA which is translated into protein. Alternative splicing of the primary RNA transcript of DNA produces different messenger RNA's, when different exons are spliced together.

Src is a oncogene first found in the Rous sarcoma virus where it encodes the $pp60^{v\text{-}src}$ kinase. The corresponding cellular kinase is $pp60^{c\text{-}src}$. The Src kinases are regulated by phosphorylation de-phosphorylation. Members of the Src family of cytosolic tyrosine kinases with important funtions in cell signalling are for example Yes, Fyn, Lck, Blk, Btk, Csk, Hck, Fgr and Yrk. Fyn associates with the p85 subunit of 1-phosphatidyl inositol 3-kinase and promotes growth of fibroblasts. Lck is a lymphocyte-specific tyrosinekinase. All these tyrosinkinases have a number of properties in common: They are myristoylated on a N-terminal site. They have SH domains, the Src-homology domains, SH2 and SH3.

Stem cells. Embryonic and adult stem cells are distinguished. Embryonic stem cells are taken from an early stage of the embryo, such as from blastocytes. They are undifferentiated and totipotent. Their potential to differentiate and to form different cell lines is unlimited. Adult stemcells are taken from the blood forming bone marrow, from epithelial cells from the skin and other sources. They are pluripotent. Both, embryonic totipotent and adult pluripotent stem cells can replace functionally differentiated cells and tissues in the body. Stem cells can divide. After division, they may form again a stem cell or proceed to a final, fully differentiated state.

Stem cell factor, SCF, also called steel factor or mast cell growth factor is a cytokine that stimulates the proliferation of myeloid and lymphoid hemopoietic progenitor cells. SCF is a ligand of the receptor tyrosine kinase, c-KIT, a protooncogene. The KIT oncogene is the gene of a transforming feline sarcoma virus.

Structural genes are also called household genes, coding for proteins which form cellular structures in the body and for the enzymes, driving the cell's metabolism. They are usualy distinguished from those gene regions which regulate gene expression in development and in the adult organism.

TATA box is a consensus sequence in the promoter region of eukaryotic genes, where general transcription factors and the RNA polymerases assemble and where transcription starts.

Telomers seal the end of a chromosome. The telomers consist of simple, tandemly-repeated sequences. Typically one strand is G-rich and the other C-rich. A 15-residue long G-rich strand is added to the 3' end of the chromosomal DNA by the telomerase. This allows a primer to bind and initiate synthesis on the other strand.The telomerase maintains the lenght of the chromosome and prevents the loss of coding sequences.

Teratoma. occur most frequently in the testis and the ovary. They are tumours containing cells and tissues which do not belong to that particular organs. These dislocated tissues arise from totipotent embryonic stem cells.

TGF-α, a transforming growth factor which belongs to the EGF family and binds to tyrosine kinase receptors.

TGF-β, a factor, unrelated to TGF-α which binds to a receptor with serine/threonine kinase activity.

Thioester-bond is an energy-rich bond between an acid group and a thiol.

T lymphocytes, thymus-derived lymphocytes or T cells are formed in the thymus. They carry out immune reactions involving cell-cell interactions and are responsible for cell-mediated immunity. One distinguishes specific and non-specific cytotoxic killer T cells, (NK cells, natural killer cells), and helper T cells which cooperate with antigen-presenting cells, APC's, in the initiation of an immune response. Suppressor T cells dampen the action of helper T cells.

Trans- See: cis-oriented.

Transcription is the copying of one strand of DNA into a complementary RNA strand by a DNA-dependent RNA polymerase.

Transcriptionfactors are proteins which initiate or regulate transcription. The former are the basic, general transcriptionfactors, the latter are the signal-dependent, special regulatory transcriptionfactors.

Transfection is the introduction of a foreign DNA molecule into an eukaryotic cell. The introduced genes can then be expressed in the cell.

Transgenic mice carry in their germline a gene that has been introduced *in vitro* into an egg or into embryonic cells at an early stage of development.The cells are then transferred into a foster mother where they develop to maturity until finally a new mouse is born. Transgenic means the result of a gene transfer across the borders of species or strains. In modern usage, transgenic animals are animals carrying genes, which were introduced arteficially into the genome of the early embryo. The newly accquired genes are called transgenes.

Translation is the process, where the RNA transcript of the DNA, the messenger RNA is translated as specified by the genetic code and new proteins are synthesized from aminoacids by the ribosome which is a 'ribozyme'.

Tumoursuppressors are transcription factors and cell cycle regulators. Since mutations of tumoursuppressor genes are commonly found in cancer cells, the products of the normal genes must be important in the prevention of cancers. The first tumoursuppressor gene was discovered 1971 by Albert Knudson and colleagues who discovered the gene for p53, a 53 kDa phosphoprotein. Mutations of the p53 gene are among the most commonly found mutations in human cancers. p53 is a transcription factor that controls a set of genes, of which the products are required for cell division. One of these genes encodes an inhibitor of cyclin-dependent proteinkinases. p53 inhibits the malignant transformation of cells in culture by oncogenes.

Ubiquitination marks proteins for proteolysis by the proteasome. It involves the covalent attachement of a small protein, ubiquitin, expressed in all eukaryotic cells, to lysines of proteins. Ubiquitination is catalyzed and controlled by enzymes.

Upstream. The nucleotide sequence of a gene is read by convention from the 5′ end to the 3′ end. Therefore, upstream means more toward the 5′ end, or the beginning of the coding sequence.

VEGF. Vascular endothelial growth factor is produced by epithelial cells, smooth muscle cells and macrophages. It plays a role in angiogenesis. It is related to PDGF. VEGF binds to the VEGF receptor, a tyrosinekinase receptor with seven immunoglobulin folds in the extracellular domain which is related to the PDGF receptor.

Vinculin is a cytoskeletal protein, associated with the cytoplasmic side of focal adhesion plaques. Vinculin associates with talin and binds integrins to the cytoskeleton.

Vmax is the maximum rate of an enzyme reaction when the substrate is present in saturating amounts. Vmax is derived from the Michaelis-Menten equation, where v is the initial velocity and P is the product of the reaction. $v=V$ [S]/(Km+[S]=d [P]/dt.

Vitamin A is a fat-soluble vitamin. Chemically it is identical with the alcohol retinol, whereas the corresponding vitamin A aldehyd is retinal and the acid derivative of vitamin A is retinoic acid. A sufficient supply of vitamin A is required for embryonic growth and differentiation, for reproduction, including spermatogenesis, oogenesis and placental development and for vision in the developping and the adult organism.

Wild type is the naturally-occuring phenotype of a species. It is distinguished from a mutant phenotype of members of the same species.

White blood cells are those nucleated blood cells, which lack hemoglobin. They are also called leukocytes. To this group belong lymphocytes, neutrophiles, eosinophiles, basophiles and monocytes.

X-ray crystallography is a technique determining the three-dimensional arrangement of atoms in a protein from the diffraction pattern of X-rays, passing through a crystal of the protein or an other molecule.

Xenopus laevis , the South African clawed toad, is a frog, not a toad. It is often used as a model organism for the study of early vertebrate development.

Zinc finger is a common structural motif in DNA-binding proteins. It is a hairpin bend of the protein held together by the Zn atom.

Zygote is a diploid cell, such as a fertilized egg cell, formed by the fusion of the gametes, the sperm with the egg.

Index

Note: When they fall on pages not covered by the text reference, references to illustrations are indicated in **bold**. Page numbers followed by t indicate tables.

Abl tyrosine kinases 38
activins 102, 103t
ADAM proteins 5–6
adaptor proteins and linkers
  ankyrin repeats 37, 38
  G proteins 43–5
  notch repeats 37
  PH domain plate 4 (2p), 35–6
  phosphotyrosine phosphatases 37–8, 41–3
  PTB domain 35–6, 37
  Ras 46–9, **50**
  SH2 domain plate 2 (1p), 33–4
  SH3 domain plate 3 (2p), 34–5
  sterile α-motif (SAM) 37, **38**
  tyrosine kinases plate 5 (3p), 37–40
adenomatous polyposis coli (APC) gene/protein 286–92
adenylate cyclase, adenylyl cyclase—G_s-α complex plate 10 (6p)
β_2 adrenergic receptor 77–8
  desensitisation 81–2
  signalling pathway 78
aldosterone, structure 195
allosteric activation, of proteins 132–3
ankyrin repeats 37, **38**
antibodies, structure 255–6
antigen presentation plate 28 (15p), 251–2, **253**, **254**
antigen receptors plate 29 (16p), 255–7
apoptosis
  and cancer 243–5
  and immune system function 245, 246
  and nervous system development 245
  prevention plate 27 (15p), 241–3
  promotors 234–41
  and self-tolerance 263
arrestin
  and G-protein coupled receptor signalling 83
  and receptor desensitisation 81–2, **83**

β-catenin, and cancer 287–92
B cells
  activation 250–1
  antigen receptors 255
  signalling pathways 258–9
  survival and death 261–3
basic helix-loop-helix (bHLH) motif 162–3
  structure plate 17 (10p)
Bax
  and apoptosis 241
  and cancer 243–5
Bcl-2–X_L, structure plate 27 (15p)
Bcl-2
  and apoptosis 241–3
  and cancer 243–5
bone morphogenetic proteins (BMPs) 102, **103**, 103t, **104**
brain-derived neurotrophic factor (BDNF) 14, 15t
breast cancer, and oestrogen receptor 197, 199

calcineurin, structure plate 13 (7p), 126
cAMP
  and G-protein coupled receptors 78, **79**
  in olfaction 95–6
cAMP response-element-binding protein (CREB) 174
cancer
  and adenomatous polyposis coli (*APC*) gene 286–7
  and adenomatous polyposis coli (APC) protein 287–92
  and apoptosis 243–5
  and DNA viruses 299–300
  inherited cancer genes 294–5
  and loss of developmental control 284–5
  oncogenes 269–76
  proto-oncogene mutations 295–9
  and retroviral infection 270–1, 299
  tumour-suppressor genes 276–82
caspases plate 26 (14p), 237–8, **239**
caveolae, and glucose transport 141, **143**
CBL protein, and cancer 273
CD (cluster of differentiation) receptors, and immune response 256–7
CD95 receptor, and apoptosis 235
Cdc42 GTPase 64–5
  control of phospholipases 70–2
  and cytoskeleton assembly 65–70
  and Wiskott–Aldrich syndrome protein (WASP) 73
cell adhesion
  and cancer 287–9
  role of integrins 68t, 69–70
cell cycle
  checkpoint control 228–30, 295–9

cell cycle (*cont.*)
chromosome duplication 225, **226**
cyclin-dependent kinases *see* cyclin-dependent kinases (Cdks)
cyclins 214, **215**
cytokinesis 226–8
DNA replication 224–5
evolution of 230
phases 213
cell death, programmed *see* apoptosis
cell lineage differentiation, cytokine functions 117–19
cGMP, in visual response 90–2
chaperones, and steroid hormone receptors 197, **198**
chaperonins, and steroid hormone receptors 197, **198**
chemotaxis 70–2
G protein control of 89–90
chromosome duplication 225, **226**
colony-stimulating factors (CSFs) 117, 118t
processing 7
cortisol, structure 195
covalent activation, of proteins 132–3
CREB (cAMP response-element-binding protein) 174
Crk adaptor protein 31–2
Csk kinase 40
cyclin-dependent kinases (Cdks) 214, **215**
and cell-cycle transitions 216–24
cyclin-dependent-kinase inhibitor (CKI) plate 25 (14p)
and DNA replication 224–5
and evolution of the cell cycle 230
cyclins 214, **215**
see also cell cycle
cytochrome *c*, and apoptosis 238, 240
cytokines
antiviral effects **118**, 119
and cell lineage differentiation 117–19
and embryonic development 116–17
JAK/STAT signalling pathway 111–16
receptors 110–11, **112**, **113**, **114**
and T cell differentiation 252–4, 255t
cytokinesis 226–8
cytoskeleton assembly, role of GTPases 65–70
cytosolic kinases 123

DAF-2
receptor 17–18
signalling pathway 148–9
death receptors 235–6, **237**
development
embryonic
cytokine functions 116–17
STATs 116–17
loss of control of, and cancer 284–5
of nervous system, and apoptosis 245
and steroid hormone receptors 193–5
diabetes 137–8
DNA
damage 240–1
histone modification 166–7
methylation 167–8, 297
nucleosomes plate 22 (12p), 165–6
replication 224–5
synthesis 62–3
DNA-binding proteins, structural motifs plate 15 (9p), plate 16 (9p), plate 17 (10p), plate 18 (10p), plate 19 (11p), plate 20 (11p), plate 21 (12p), 161–4
DNA tumour viruses 299–300

ecdysone receptor 192–3
ecdysteroids 192
EGF *see* epidermal growth factor (EGF)
embryonic development, cytokine functions 116–17
Eph-like receptors 9t, 18, 37, **38**
ephrins 9t
epidermal growth factor (EGF)
EGF receptor and cancer 272
signal transduction mechanisms 9t, 10–11
erythropoietin (EPO) receptor, ligand-dependent dimerization 25–7
estrogen receptor *see* oestrogen receptor

fibroblast growth factors (FGFs), signal transduction mechanisms 9t, 12–14
focal adhesion kinase (FAK), integrin-FAK signalling system 68–70
Fos/Jun transcription factors 172–4

G-protein-coupled receptors
control of hormonal signalling 79–81
and olfaction 92–6
and Ras/MAP kinase pathway 83–5
receptor desensitisation 81–2, **83**
rhodopsin activation 85–8
second messenger signalling plate 10 (6p), plate 11 (6p), 78, **79**
structure 76–8
and taste transduction 96–7
and visual response 90–2
G proteins
$\alpha\beta\gamma$-holocomplex structure plate 7 (4p)
$\beta\gamma$-subunit functions 79–80
and cancer 273
G-α subunit structure plate 6 (4p)
GDP/GTP cycle 43–5
G_i-α1 subunit plate 8 (5p)
heterotrimeric G proteins 88–90
monomeric G proteins 89–90
Ras/GAP complex structure plate 9 (5p)
Ras/MAP kinase pathway 57–64
Rho/Rac/Cdc42 GTPases 65–73
in visual response plate 12 (7p)
GATA transcription factors 175
GDNF (glial-cell-line-derived neurotrophic factor) 14–17

GDP/GTP cycle
  $\alpha\beta\gamma$-holocomplex structure plate 7 (4p)
  and cell signalling 43–5
  G-α subunit structure plate 6 (4p)
  G_i-α1 subunit plate 8 (5p)
  Ras/GAP complex plate 9 (5p)
gene transcription see transcription
genomic imprinting 167–8
glial-cell-line-derived neurotrophic factor (GDNF) 14–17
glucose
  homeostasis, insulin actions 140–2
  protein glycation 138, 139
glutamic-acid-rich proteins (GARPs), in visual response 91
glycation, of proteins 138, 139
glycogen phosphorylase
  regulation by phosphorylation dephosphorylation 130–2
  structure plate 14 (8p)
Grb2 (growth-factor-receptor-binding protein), SH2 and SH3 domains 32, 34–5
growth factor receptors
  and cancer 272
signal transduction mechanisms 8–14
growth factors
  and cancer 271–2
  processing 5–8
growth hormone (GH) receptor
  binding domain structure plate 1 (1p)
  ligand-dependent dimerization 24–5
GTPase-activating proteins (GAPs), and control of Ras activity 46–8
GTPases 43–5
  control of phospholipases 69, 70–2
  and cytoskeleton assembly 65–70
  Ras superfamily 64–5
  signalling pathways 72–3
guanine nucleotide release proteins (GNRPs), and control of Ras activity 48–9, 50
gusducin, and taste transduction 96–7

haematopoietic cells, differentiation, cytokine functions 117–19
heat-shock protein (Hsp) system, and steroid hormone receptors 197, 198
hepatocyte growth factor (HGF) 9t, 18
HGF-like factor 9t
histones 165–6
  modification 166–7
  nucleosome core particle structure plate 22 (12p)
homeodomain proteins 162, 163
  MATα2/MCM1/DNA complex structure plate 15 (9p)
homeotic genes 162, 193–5
5–HT (serotonin) receptor 77

immune response 250–1, 259–60
  antigen presentation plate 28 (15p), 251–2
  antigen receptors plate 29 (16p), 255–7
  lymphocyte survival and death 261–3
  lymphoid organs 250, 251
  signalling pathways in T cells and B cells 258–9
  T cell differentiation 252–4, 255t
  T cell selection 252
immune surveillance, and apoptosis 246
immunoglobulins, structure 255–6
imprinting, of genes 167–8
insulin
  functions 137–40
  and glucose homeostasis 140–2
  as growth factor 137, 142–50
  history 137
  processing 8
  signal transduction mechanisms 9t
  signalling pathways 142–50
insulin-like factor, signal transduction mechanisms 9t
insulin receptor
  homologues 17–18
  insulin receptor substrates (IRS) 143–6
  signal transduction mechanisms 9t, 17
  structure 11
integrin-FAK signalling system 68–70
integrins 65–70
interferons 118, 119
interleukin-1β-converting enzyme (ICE) plate 26 (14p), 237
interleukins, and T cell differentiation 252–4, 255t

JAK/STAT signalling pathway, and cytokines 111–16
JNKs (Jun N-terminal kinases) 60
Jun/Fos transcription factors 172–4
juvenile hormone (JH) 195

keratinocyte growth factors (KGFs) 9t, 12
KIT receptor 9t, 14
KL ligand 9t, 14

leptin, cross-talk with insulin signalling pathways 148–50
leucine zipper motif 162, 163
ligands, growth factor processing 5–8
linker proteins see adaptor proteins
lymphocytes see B cells; T cells

macrophage colony-stimulating factor-1 (MCSF-1) 9t
  signal transduction mechanisms 11
MAD proteins 176
major histocompatibility complex (MHC)
  and antigen presentation 251–2, 253, 254
  and antigen receptor, structure plate 29 (16p)
MAP kinase pathway see Ras/MAP kinase pathway
MAP kinases 60–2
  role in growth and proliferation 62–3
MAPKKs (MEKs) 60

mast cell growth factor 9t, 14
MATα2/MCM1/DNA complex, structure plate 15 (9p)
MCSF-1 (macrophage colony-stimulating factor-1) 9t
metalloproteinases, and growth factor processing 5–6
methylation, of genes 167–8
  and cancer 297
MHC see major histocompatibility complex (MHC)
mitochondria, and apoptosis 238, 240
mitogen-activated protein kinases see MAP kinases
morphogenesis, TGF-β signalling pathway 108–9
multi-enzyme organelles, phosphorylation cascades 129–30

nerve growth factor (NGF)
  and apoptosis 245
  p75NGFR 15–16, 245
  signal transduction mechanisms 14–17
nervous system development, and apoptosis 245
neuregulin 9t
neurofibromin 1, and control of Ras activity 47–8
neurotrophins 14–17
  and apoptosis 245
NF-AT transcription factor 174–5
  NF-AT/Jun/Fos/DNA complex plate 23 (13p)
NF-κB transcription factor 164, 175
  NF-κB/p52/DNA complex plate 23 (13p)
  structure plate 18 (10p)
NF-xB transcription factor 175
NGF see nerve growth factor (NGF)
notch repeats 37
nuclear receptors
  for non-steroids see retinoic acid receptors (RARs); retinoid receptors (RXRs); thyroid hormone receptors
  for steroid hormones see steroid hormone receptors
nuclear transport, and transcriptional control 181–2, 183
nucleosomes 165–6
  core particle structure plate 22 (12p)

oestradiol, structure 199
oestrogen receptor, and breast cancer 197, 199
olfaction
  neuronal connections 92–3
  odorant receptors 93–5
  odorants 93
  signal transduction in olfactory neurons 95–6
oncogenes
  and proto-oncogene functions 269–71
  and signalling pathways 271–3
  transformation from proto-oncogenes 273–6
oxidative stress, stress activated transcription factors 179–80

p53
  and cancer 243–5, 279, 280–2
  structure plate 30 (16p)
p75NGFR
  and apoptosis 245
  signal transduction mechanisms 15–16
PDGF (platelet-derived growth factor), signal transduction mechanisms 9t, 11–12
PH domain plate 4 (2p), 31–2, 35–6
phosducin plate 12 (7p), 90–1
phosphatidylinositol 3-kinase pathway 59–60
  cytosolic kinase functions 123
  and insulin action 144, 145
phospholipases, control by G proteins 69, 70–2
phosphoprotein phosphatases
  calcineurin structure plate 13 (7p)
  localization and targeting 123–7
phosphorylation cascades
  as multi-enzyme organelles 129–30
  protein kinases 128–9
  receptor kinases 123
phosphorylation
  dephosphorylation control of signalling pathways 122–3
  glycogen phosphorylase a, structure plate 14 (8p)
  phosphorylation cascade organization 123–30
  regulation of enzymes and signalling pathways 130–5
phosphotyrosine-binding domain see PTB domain
phosphotyrosine phosphatases 37–8, 41–3
platelet-derived growth factor (PDGF), signal transduction mechanisms 9t, 11–12
pleckstrin-homology domain see PH domain
PP1 phosphatase 126, 127
programmed cell death see apoptosis
prolactin (PRL) receptor, ligand-dependent dimerization 24–5
protein glycation 138, 139
protein kinases, localization and targeting 128–30
protein tyrosine phosphatase (PTP) 41–2
proteinases, and growth factor processing 5–7
proteolysis
  and apoptosis 236–8, 239
  regulation of cyclin-dependent kinases 219–20, 222–4
proteomics 186
proto-oncogenes
  functions 269–71
  and signalling pathways 271–3
  transformation to oncogenes 273–6
PTB domain 31–2, 35–6, 37
Pyk kinases, in Ras/MAP kinase pathway 58–9

Rac GTPase 64–5
  control of phospholipases 70–2
  and cytoskeleton assembly 65–70
Raf kinase, in Ras/MAP kinase pathway 57–8
Ran GTPase, and cytokinesis 228
Rap 1, and G-protein coupled receptor signalling 84–5
Ras
  control of activity 46–9, 50, 123

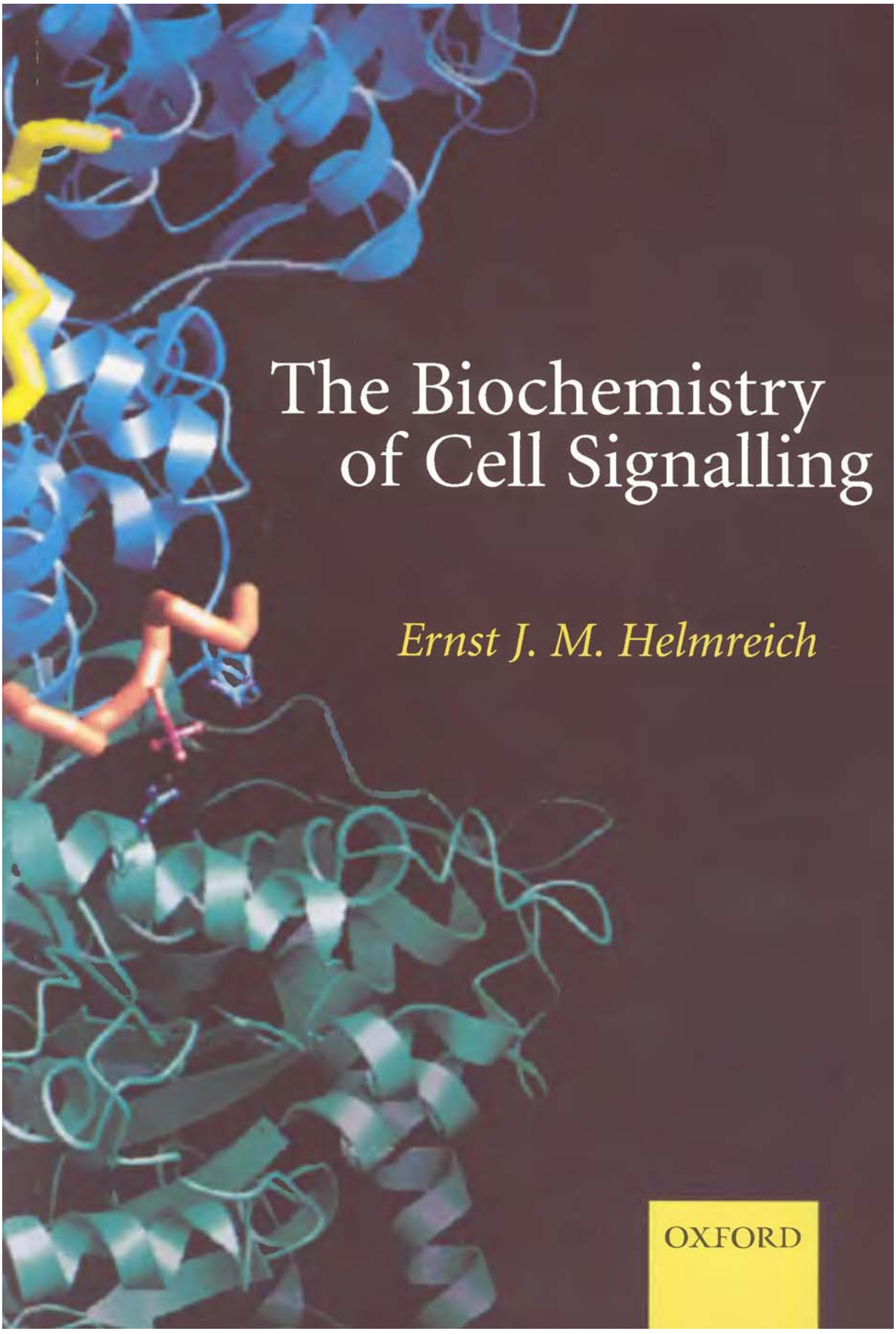
and Rho/Rac/Cdc42 signalling pathways 72–3
Ras GTPase superfamily 64–5
Ras/MAP kinase pathway 32, 33
and G-protein coupled receptor signalling 83–5
MAP kinases 60–2
MAPKKs 60
phosphatidylinositol 3–kinases 59–60
Pyk kinases 58–9
Raf kinase 57–8
role in growth and proliferation 62–3
specificity 63–4
ras oncogene 273, 274
Rb gene, and cancer 276–7
Rb protein 277–9
receptor kinases, phosphorylation 123
receptor protein tyrosine phosphatase (RPTP) 43
receptor tyrosine kinases (RTKs), signal transduction mechanisms 8–18
receptors
ligand-dependent dimerization 24–30
plasticity 28–30
properties 3–5
retinoblastoma, Rb 276–9
retinoic acid receptors (RARs)
functions 202–3
regulation 206
structure plate 24 (13p), 204–6, 206–7
retinoid receptors (RXRs)
functions 203–4
structure plate 24 (13p), 204–6, 206–7
retroviruses, and cancer 270–1, 299
Rho GTPase 64–5
control of phospholipases 70–2
and cytoskeleton assembly 65–70
and G-protein coupled receptor signalling 83–4
rhodopsin
activation 85–8
in visual response 90–2
RNA polymerase II, regulation of 158–61
RNA polymerase II transcription factors *see* transcription factors, general
RNA synthesis, role of MAP kinases 62–3
SAM (sterile α-motif) 37, 38
SAPKs (stress-activated protein kinases) 60
scatter factor (SF) 9t, 18
second messenger signalling, and G-protein-coupled receptors plate 10 (6p), plate 11 (6p), 78, 79
self-tolerance, and apoptosis 263
serine/threonine phosphatases 124–7
calcineurin structure plate 13 (7p)
serotonin (5–HT) receptor 77
SH2 domain plate 2 (1p), 31–2, 33–4
SH3 domain plate 3 (2p), 31–2, 34–5
Shc adaptor protein 31–2
sheddases, and growth factor processing 5–7
SMADs 176
and TGF-β signalling 102–9
SNARE proteins, and insulin action 141, 142
Src-homology 2 domain *see* SH2 domain
Src-homology 3 domain *see* SH3 domain
Src tyrosine kinases 38–40
structure plate 5 (3p)
SREBPs (sterol regulatory-element-binding proteins) 175–6
STATs 176–8, 179
and embryonic development 116–17
JAK/STAT signalling pathway 111–16
steel factor 9t, 14
stem cell factor 9t, 14
sterile α-motif (SAM) 37, 38
steroid hormone receptors
chaperones and chaperonins 197, 198
control of signalling 199–200
and developmental control genes 193–5
ecdysone receptor 192–3
ligand structures 195–6
oestrogen receptor and breast cancer 197, 199
receptor classes 190–2
structure plate 24 (13p), 196, 201t
sterol regulatory-element-binding proteins (SREBPs) 175–6
stress activated transcription factors 179–80
tamoxifen, structure 199
T cells
activation 250–1
and antigen presentation plate 28 (15p), 251–2, 254
antigen receptors plate 29 (16p), 255–7
differentiation 252–4, 255t
selection 252
signalling pathways 258–9
survival and death 261–3
taste transduction, G protein functions 96–7
TATA box plate 19 (11p), 164
and initiation of transcription 158–9, 160
TATA-box-binding protein (TBP) 164
and initiation of transcription 158–9
structure plate 19 (11p), plate 20 (11p), plate 21 (12p)
telomerase
and cancer 297–9
and chromosome duplication 225, 226
teratocarcinoma 284–5
TGF-α
processing 7
signal transduction mechanisms 9t, 10
TGF-β
growth factor superfamily 102–5
and morphogenesis 108–9
receptors 102–5
signalling pathway 105–7
thyroid hormone receptors 200–2
structure plate 24 (13p)
TNF-α
cross-talk with insulin signalling pathways 148
processing 7
TNF-α converting enzyme (TACE) 5–6
TNF-β, processing 7
TNF receptors
and apoptosis 235–6, 237
oligomerization 27–8
tolerance, of self 263
transcription
DNA-binding proteins plate 15 (9p), plate 16 (9p), plate 17 (10p), plate 18 (10p), plate 19 (11p), plate 20 (11p), plate 21 (12p), 161–4

initiation 157–61
pre-initiation complex structure 164–5
transcription factors 179–80
and cancer 273
control of 180–5
CREB 174
GATA 175
general plate 20 (11p), plate 21 (12p), 158–61, 164–5
Jun/Fos 172–4
MAD proteins 176
NF-AT transcription factor 174–5
NF-AT plate 23 (13p), 174–5
NF-κB 175
NF-κB/p52/DNA complex plate 23 (13p)
NF-κB 175
SMAD proteins 176
SREBPs 175–6
STAT proteins 176–8, 179
stress activated transcription factors 179–80
transducin 86–7
G-α subunit structure plate 6 (4p)
G protein αβγ-holocomplex structure plate 7 (4p)
transforming growth factors *see* TGF-α; TGF-β
tumour necrosis factors *see* TNF-α; TNF-β
tumour-suppressor genes
adenomatous polyposis coli (*APC*) gene 286–7
p53 279, 280–2
Rb gene 276–7
tyrosine kinases plate 5 (3p), 37–40
and cancer 272

vascular endothelial growth factor (VEGF)
processing 7–8
signal transduction mechanisms 9t, 11
Vav adaptor protein 31–2
visual response, G protein functions plate 12 (7p), 90–2
vitamin A receptor 202–3
vitamin D_3, structure 195

Wiskott–Aldrich syndrome protein (WASP) 73
Wnt signalling pathway, and cancer 289–90

zinc finger motif 162
structure plate 16 (9p)



The Biochemistry of Cell Signalling

Ernst J. M. Helmreich

OXFORD

The biochemistry of cell signalling

The biochemistry of cell signalling

Ernst J. M. Helmreich

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2001
Reprinted 2002

A Catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data

Helmreich, E. J. M. (Ernst J. M.)
The biochemistry of cell signalling/Ernst J. M. Helmreich.
Includes index.
1. Cellular signal transduction. I. title.
QP517.C45 H455 2001 572'.4—dc21 2001016302

ISBN 0 19 850820 4 (Pbk.)

10 9 8 7 6 5 4 3 2 1

Printed in Great Britain
on acid-free paper by
T. J. International Ltd
Padstow, Cornwall

This book is dedicated to the memory of Carl F. Cori. His wisdom and grasp of the principles of cellular regulation has profoundly influenced my own approach to the biochemistry of cellular signalling.

Preface

Alles Gescheite ist schon gedacht worden, man muss nur versuchen, es noch einmal zu denken. [Everything worthwile and intelligent has already been thought. One can only try to think it over again.]

Johann Wolfgang von Goethe (Wilhelm Meisters Wanderjahre)

The complexity of cellular regulation is overwhelming. This is mainly due to the large repertoire of players that are parts of the regulatory network in a cell, each with a different name, usually not explained, each with a different structure and function. In addition there is a multitude of genes which encode regulatory proteins. If one wants to keep an overview of such a central and rapidly developing and expanding field, the question is how to reconcile seemingly conflicting intentions, namely to keep updated and at the same time not lose sight of the whole. Faced with that decision I chose to focus on basic features of signalling from the membrane to the nucleus.

To make things simpler, I have directed attention to common functional and structural properties of regulatory proteins. I have focused on phosphorylation, dephosphorylation, protein–protein recognition, on- and off-switching by G proteins, proteolysis of regulatory proteins, protein–DNA recognition in the control of gene transcription, and a few other fundamental aspects. To make things intelligible, I have relied on graphic presentations, as much as on text. Moreover, wherever possible, I have included structural information to explain function.

The other objective was to give the reader a chance to find out how the cell uses these instruments to accept and react to a multitude of intra- and extracellular signals. Sensory perception was chosen as an example to explain signalling by G-protein-coupled receptors. The role of phosphorylation dephosphorylation was exemplified in the response to the hormone insulin, and the recognition of cell-bound signals was introduced in terms of the immune response. The global regulatory programmes of a cell, the cell cycle and programmed cell death (apoptosis), were linked to the control of cell proliferation, the selection of cell lines, and the prevention of autoimmune responses. The chapter concluding the book then tries to explain what goes wrong in a cell that has lost control over proliferation and differentiation and becomes a cancer cell.

I have not tried to be comprehensive. Important aspects of cellular regulation have been left out. A complete treatment would have to be encyclopaedic, considering the breadth of the field and the rapidity of the progress being made. Moreover, excellent textbooks of biochemistry and cellular and molecular biology, and a wealth of detailed reviews on every aspect of signal transduction, are available. What I present are fundamental aspects of cellular regulation, which are usually not treated in this particular context. But, that could not be done without relying on basic facts, described in textbooks of biochemistry and molecular biology. Should the reader not be familiar with these and does not find an explanation in my book, he or she must make use of the information available elsewhere.[1] I know that this is a problematic compromise, but it had to be made to avoid repetition of information, competently and comprehensively presented elsewhere.

The literature, on which description and interpretation in this book is based, is cited in reference lists at the end of each chapter, overall containing about 1000 entries. It can be found in the Web site, where proper credit is given to all the colleagues who have done the experiments that are discussed in this book. The reader is invited to use this information, which should help him or her to extend and clarify the information presented. The references are added to the benefit of those readers who might feel that I have left out too much, or may even have presented in some cases already outdated concepts. Such mishaps are unavoidable in such a rapidly developing field.

This book is adressed to all who are curious and who want to understand fundamental biological processes that ensure our survival. The person who might be interested is probably someone just starting research who wants to learn about aspects of regulation, outside his chosen field. It could become a supplementary text for graduate students and post-docs in biochemistry and molecular biology, and for joint PhD–MD programmes. But also a physician who wants to understand the molecular aetiology of the many diseases that are the consequence of regulatory disorders, and the chemist who wants to learn about cellular regulation to find agents to treat related diseases, might also profit from this book.

Ernst J. M. Helmreich
Summer 2000

Notes

1. Related resources on the World Wide Web:

   The Dictionary of Cell Biology (Academic Press, London, 1995), defines some of the terms used in this article.

   The MIT Biology Hypertextbook, developed by the Experimental Study Group at the Massachusetts Institute of Technology, provides background information on the biology of cells and nucleic acids. Mitosis, a section within Cell Biology, provides descriptions and diagrams of the phases of the cell cycle.

   Pedro's BioMolecular Research Tools is a collection of WWW links to information and services useful to molecular biologists. It provides links to molecular biology search and analysis tools; bibliographic, text, and Web search services; guides and tutorials; and biological and biochemical journals and newsletters.

The World Wide Web Virtual Library: Biosciences points to virtual library pages for Biological Molecules, and Biochemistry and Molecular Biology. Each of these pages presents a long list of Web resources. The World Wide Web Virtual Library Biological Molecules covers molecular sequence and structure databases, metabolic pathway databases, and other lists of Web resources. The World Wide Web Virtual Library: Biochemistry and Molecular Biology is a list of resources, listed by provider.

Cell and Molecular Biology Online is a well-organized list of Web resources for cell and molecular biologists. For each resource, a brief description is provided.

CSUBIOWEB, the California State University Biological Sciences Web server, provides links to other Web sites on cell biology and molecular biology.

Acknowledgements

The idea to write a book on cell signalling was an offspring of my stay at the NIH in Bethesda in the early nineties as a Fogarty Scholar in Residence. Exposed to a stimulating environment, I became aware of the exciting research going on in signal transduction, outside my own niche. It was primarily my own curiosity that got me to write my first book, a book on the biochemistry of cell signalling. Therefore, I wish to express my gratitude to the Fogarty Centre for giving me the opportunity to stay for a while at the NIH.

This book could never have been written without the help and the criticisms of my colleagues. I am greatly indebted to Martin Lohse, Professor of Pharmacology, Walter Sebald, Professor of Biochemistry, and Klaus Koschel and Thomas Hünig, Professors of Immunobiology, at the University of Würzburg, and to F. Ulrich Hartl, Professor at the Max Planck Institute for Biochemistry, Martinsried bei München, partly for their very valuable criticisms of the book and partly for allowing me to include some of their data. I deeply appreciate their efforts and their help. Other colleagues who have read critically one or other chapters of the book, were Felix Wieland, Professor of Biochemistry, University of Heidelberg; Kurt Kochsiek, Professor of Internal Medicine, Hans K. Müller Hermelink, Professor of Pathology, Christoph Jungwirth, Professor of Virology and Immunobiology, all at the University of Würzburg; and Louise Johnson, FRS, University of Oxford, UK. Thanks are due to my friend and colleague Kurt Kochsiek for allowing me to retreat to his chalet in the mountains to collect my thoughts and progress with my writing. My special thanks are due to Werner Schnepp and Cornelius Krasel in Würzburg and to Martin Noble in Oxford for contributing most of the structures presented in this book.

To Mrs. Jutta Schwab I am much indebted for the art work.

Special thanks are due to the group of dedicated people at Oxford University Press who have been engaged in the many aspects involved in publication a book. It was a pleasure to work with them.

Finally, I wish to thank my wife, Mylein, for her patience and tolerance in the light of my preoccupation with writing this book.

Contents

Abbreviations xv

Part 1 **The machinery of signal transduction** 1

1 Molecular basis of signal transduction 3

2 Activation of receptors by oligomerization 24

3 Components of signalling networks: linkers and regulators 31

4 Signal transduction pathways through small monomeric G proteins 57

5 Signal transduction pathways through heterotrimeric G proteins: transmission of hormonal and sensory signals 76

6 Signal transduction pathways controlling morphogenesis and haematopoiesis 102

7 Control of signalling by phosphorylation and dephosphorylation 122

8 Regulation by a hormone: the insulin response 137

Part 2 **Cell signalling and gene transcription** 155

9 Machinery of gene transcription 157

10 Regulation of gene transcription by growth factors and cytokines 172

11 Regulation of gene transcription by hormones 190

Part 3 **Global cell regulatory programmes** 211

12 Regulation of the cell cycle 213

13 Regulation of cell death 234

14 Regulation of the immune response 250

Part 4 **Loss of the regulatory control and its consequences** 267

15 Transformation of normal cells to tumour cells 269

16 Loss of developmental controls in cancer 284

17 The causes of cancer 294

Glossary 303

Index 323

Abbreviations

ACE	angiotensin-converting enzyme
ACK	activin-receptor-like kinase
ADAM	A disintegrin and metalloproteinase (family)
ADP-PNP	adenosine 5′-[β, γ-imido]triphosphate
ADP-PCP	adenosine 5′-[β, γ-methylene]triphosphate
AFAP	actin-fibre-associated protein
AGE	advanced-glycation end-product
AKAP	A-kinase-associated protein
ALL	acute lymphoblastic leukaemia
AMH	anti-Müllerian hormone
AP	activating protein
APAF	apoptosis-activating factor
APC	antigen-presenting cell OR anaphase-promoting complex OR adenomatous polyposis coli
β-APP	β-amyloid precursor protein
β_2-AR	β_2-adrenergic receptor
ARAM motif	antigen recognition and activation motif
ARE	activin-reponse element OR antioxidant response element
ARF	activin response factor complex
β-ARK	β-adrenergic receptor kinase
ARRE	antigen-receptor response element
ASV	avian sarcoma virus
ATP	adenosine triphosphate
bcr	breakpoint cluster region
BCR	B-cell receptor
BDNF	brain-derived neurotrophic factor
BH	Bcl-2 homology domain
bHLH	basic helix-loop-helix
BMP	bone morphogenetic protein
BUB	genes making budding in yeast unresponsive to benomyl
bp	base pair
BP	binding protein

BTK	Bruton's agammaglobulinaemia tyrosine kinase
Btk	gene for Bruton's tyrosine kinase
bZIP	basic leucine zipper
CAH	carbonic anhydrase
CAK	cyclin-activated kinase
CAM	cell-adhesion molecule
cAMP	cyclic adenosine monophosphate
CAP	CBL-associated protein
CARD	caspase-recruiting domain
CARM	co-activator-associated arginine methyltransferase
Caspr	contactin-associated protein
CBL	product of the *cbl* proto-oncogene
CBP	CREB-binding protein OR cAMP-binding protein
CD	cell death
cdc	cell division control gene
CDE	cell-cycle-regulated DNA element
CDEF	CDE-binding factor
CDK	cyclin-dependent kinase
CDKI	cyclin-dependent kinase inhibitor
CDMP	cartilage-derived morphogenic protein
cDNA	complementary DNA
CFC	colony-forming cell
cGMP	cyclic guanosine monophosphate
chk	checkpoint kinase
Ci	contactinhibin
CINs	chromosomal instabilities
CIS	cytokine-inducible SH_2-(containing protein)
CLIP	MHC-class II-associated invariant chain peptide
CNG	cyclic nucleotide-gated
CNGC	cyclic nucleotide-gated channel
CNTF	ciliary neurotrophic factor
Co-SMAD	common SMAD
COUP-TF	chicken ovalbumin upstream promoter transcription factor
CPS	carbamoyl phosphate synthase
9cRA	9-*cis* retinoic acid
CRE	cAMP response element
CREB	cAMP response-element-binding protein
CRIB	Cdc42/Rac interactive binding domain
Crk	CT10 virus regulator of kinase
CSF	colony-stimulating factor
Csk	carboxy terminal src-related kinase
CTLA	cytotoxic lymphocyte-associated protein
DAG	diacylglycerol

DBD	DNA-binding domain
DD	death domain
DED	death effector domain
DIF	differentiation-inducing factor
DISC	death-inducing signalling complex
DPC	deleted in pancreatic carcinoma
dRNTPs	deoxyribonucleotide triphosphates
dsRNA	double-stranded RNA
EBP	erythropoietin binding protein
ECD	extracellular hormone or ligand-binding domain
ECM	extracellular matrix
ECR	ecdysone receptor
EcRE	ecdysone-responsive DNA element
EGF	epidermal growth factor
EGFR	EGF receptor
eIF-4E	eukaryotic translation initiation factor-4E
EMP1	EPO-mimetic peptide 1
EPAC	exchange protein directly activated by cAMP
EPO	erythropoietin
ER	endoplasmic reticulum OR oestrogen receptor
ERK	extracellular signal-responsive kinase
FADD	Fas 'death' domain
FAK	focal adhesion kinase
FAT	focal adhesion targeting (domain)
FGF	fibroblast growth factor
FGFR	FGF receptor
FKBP	FK-binding protein, immunophilin
fMLP	*N*-formyl-Met-Leu-Phe
FOG	friend of GATA-1
FPP	farnesyl diphosphate
FSH	follicle-stimulating hormone
FTase	farnesyl transferase
Gab-1	Grb2-associated binder-1
GABA	γ-aminobutyric acid
GAD	glutamic acid decarboxylase
GADD	growth arrest and DNA damage
GAP	GTPase-activating protein
GARP	glutamic-acid-rich protein
GAS sequence	interferon-γ-activated sequence
GBD	GTPase-binding domain
GCRK	G-protein-coupled receptor kinase
G-CSF	granulocyte-colony-stimulating factor
GDI	GDP-dissociation inhibitor
GDF	growth and differentiation factors

GDNF	glial-cell-line-derived neurotrophic factor
GDNFR	glial-derived neurotrophic factor receptor
GDP	guanosine diphosphate
GEF	GDP-exchange factor
GFAP	glial fibrillary acidic protein
GGPP	geranylgeranyl diphosphate
GH	growth hormone
GM-CSF	granulocyte-macrophage colony-stimulating factor
GMP	guanosine monophosphate
GNRP	guanine nucleotide release protein
GPCR	G-protein-coupled receptor
GPI	glycosylphosphatidylinositol
GPRP	G-protein-regulated receptor serine/threonine phos phatase
GR	glucocorticoid receptor
Grb	growth-factor-receptor-binding protein
GRK	G-protein-coupled receptor kinase
GRTH	general refractoriness to thyroid hormone
G_s	stimulatory G protein
GSK	glycogen synthase kinase
GST	glutathione *S*-transferase
GTF	general transcription factor
GTP	guanosine triphosphate
GTPase	guanosine triphosphatase
HAC	histone acetylase
HB-EGF	heparin-binding EGF-like growth factor
HC	heretidary chondrodysplasia
hCG	human chorionic gonadotrophin
Hck	haematopoietic cell kinase
HGF	hepatocyte growth factor
hGH	human growth hormone
hGHbp	human growth hormone binding protein
HHT	hereditary haemorrhagic telangiectasia
HLF	hepatocyte leukaemia factor
HLH	helix-loop-helix
HMG	high-mobility group (proteins)
HNF	hepatic nuclear factor
hPRL	human prolactin
HPV	human papilloma virus
HRE	hormone-responsive element
Hsp	heat-shock protein
HSPGs	heparan sulphate proteoglycans
5-HT	5-hydroxytryptamine (serotonin)
hTERT	human telomerase holoenzyme

IAP	inhibitor of activated protease (caspase)
I^{CAD}	caspase-activated deoxyribonuclease inhibitor
ICE	interleukin-1β-converting enzyme
IDDM	insulin-dependent diabetes mellitus
IFN	interferon
IFNAR1	interferon-α receptor1
Ig	immunoglobulin
IGF	insulin-like growth factor
IL	interleukin
IP_3	inositol trisphosphate
IR	insulin receptor
IRE	insulin response element
IRF	interferon-response factor
IRS	insulin receptor substrate
ISGF	interferon-stimulated gene factor
ISRE	interferon-stimulated response element
ITAM	immunoreceptor tyrosine activation motif
JAB	JAK-binding protein
JAK	Janus kinase
JH	juvenile hormone
JNK	Jun N-terminal kinase
KGF	keratinocyte growth factor
LAP	latency-associated peptide
LBD	ligand-binding domain
Lck	lymphocyte kinase
LDL	low-density lipoprotein
LFA	lymphocyte-function-associated protein
LIF	leukaemia inhibitory factor
LIFR	LIF receptor
LIP	lymphocyte inhibitory protein
LMW	low molecular weight
LOH	loss of heterozygosity
LPA	lysophosphatidic acid
LTBP	latent TGF-β binding protein
MAD	mitotic arrest-deficient gene
MAP	mitogen-activated protein
MAPK	MAP kinase
MAPKK	MAP kinase kinase
MAPKKK	MAP kinase kinase kinase
MAT	mating-type protein
Max	Myc-associated factor X
MCSF-1	macrophage colony-stimulating factor-1
MEK	mitogen-extracellular signal-responsive kinase
MEN	multiple endocrine neoplasias

MHC	major histocompatibility complex
MIS	Müllerian inhibiting substance
MMPs	matrix metalloproteinases
MMR	DNA mismatch repair system
MODY	matury-onset diabetes of the young
MyoD	Myoblast determining gene or protein
NCC	neural crest cell
N-CoR	nuclear co-repressor
NES	nuclear export signal
NF	nuclear factor
NF1	neurofibromatosis type 1 or the neurofibromin GAP
NF-AT	nuclear factor of activated T cells
NF-ATc	nuclear factor for activation of T cells
NGF	nerve growth factor
NGFI-B	nerve growth factor-induced receptor
NIDDM	non-insulin-dependent diabetes mellitus
NLS	nuclear localization site
NMDA	*N*-methyl D-alanine
NOD	non-obese diabetic
NPY	neuropeptide Y
NSF	*N*-ethylmaleimide-sensitive fusion protein
NT	neurotrophin
OBP	odourant-binding protein
OMP	olfactory marker protein
OP-1	osteogenic protein-1
ORC	origin recognition complex
ORF	open reading frame
PA	phosphatidic acid or plasminogen activator
PAH	phosphatidic acid hydrolase
PAK	p21 src-activated kinase
p53BP2	p53-binding protein 2
PCD	programmed cell death
PDB	protein database
PDE	phosphodiesterase
PDGF	platelet-derived growth factor
PDGFR	platelet-derived growth factor receptor
PDK	3-phosphoinositide-dependent protein kinase
PH	pleckstrin-homology (domain)
PhlP	phosducin-like protein
Pi	inorganic phosphate
PIC	pre-initiation complex
PICK	perinuclear inhibitor of C kinase
PIP_2	$PtdInsP_2$
PIP_3	phosphatidylinositol-3,4,5-phosphate

PIP 5-kinase	phosphatidylinositol phosphate 5-kinase
PKB	protein kinase B
PKC	protein kinase C
PLC	phospholipase C
PLD	phospholipase D
PMDS	persistent Müllerian duct syndrome
Pol II	RNA polymerase II
PP	protein phosphatase
PPAR	peroxysome proliferator-activated receptor
PRL	prolactin
PRPP	phosphoribosyl pyrophosphate
PTB	phosphotyrosine-binding (domain)
PtdIns	phosphatidylinositol
PtdIns-3-P	phosphatidylinositol-3-phosphate
$PtdInsP_2$	phosphatidylinositol bisphosphate
PtdIns-3,4-P_2	phosphatidylinositol-3,4-bisphosphate
PtdIns-3,4,5-P_3	phosphatidylinositol-3,4,5-trisphosphate
PTP	protein tyrosine phosphate or protein tyrosine phosphatase
pTyr	phosphotyrosine
RA	retinoic acid
RACK	receptor-associated C-kinase
RAR	retinoic acid receptor
Rb	retinoblastoma
RC	replication complex
RE	response element
REL	oncogene of avian reticuloendothelial virus
RET	human transforming oncogene, in T-cell lymphoma
RGS	regulator of G-protein signalling
RHR	Rel homology region
RLF	replication licensing factor
RNR	ribonucleotide reductase
RPTP	receptor protein tyrosine phosphatase
R-SMAD	receptor-regulated SMAD
RTKs	receptor tyrosine kinases
RXR	9-*cis*-retinoic acid receptor
S1P	sphingosine-1-phosphate
SIRP	signal regulatory protein
SAgs	superantigens
SAM	sterile *a*-motif OR *S*-adenosyl methionine
SAPKs	stress-activated protein kinases
SCAP	SREBP-cleavage-activating protein
SCID	severe combined immunodeficiency
SF	scatter factor

SHC	SH_2-containing (proteins)
SH2	Src-homology 2 (domain)
SH3	Src-homology 3 (domain)
SHP	small heterodimerization partners
SHIP	SH_2 domain-containing inositol 5-phosphatase
SHR	steroid hormone receptor
SIRP	signal-regulatory protein
SMAD	SMA-MAD-related
SMRT	silencing mediator for RAR and TR
Smurf 1	SMAD ubiquitination regulatory factor 1
SNAP	soluble NSF attachment protein
SNARE	SNAP receptor
SOCS	SH_2 domain-containing suppressors of cytokine signalling
SRB	Suppressor of RNA polymerase B
SRC	SHR-co-activator or Src, a Rous sarcoma virus oncogene
SRE	serum response element
SREBP	sterol regulatory-element-binding protein
SRF	serum response factor
SS-1	STAT-induced STAT inhibitor 1
STAT	signal-transducing transcriptional activator (signal transducer and activator of transcription)
T_3	tri-iodothyronine
T_4	tetra-iodothyronine (thyroxine)
TACE	TNF-α converting enzyme
TAD	*trans*-activating domain
TAF	TATA-box-activating factor
TAg	large T-antigen
TAK	TGF-activated kinase
T-ALL	T-cell acute lymphoblastic leukaemia
TAR	aspartate receptor
TBP	TATA-box-binding protein
Tcf	T-cell factor
TCF	ternary complex factor
TCR	T-cell antigen receptor
TGF	transforming growth factor
THR	thyroid hormone receptor
TIMPs	tissue inhibitors of metalloproteinases
TKB	tyrosine-kinase-binding domain
TNF-α	tumour necrosis factor-α
TNFR	tumour necrosis factor receptor
tRA	all-*trans* retinoic acid
TRADD	tumour necrosis factor receptor-associated death domain

TRAF	TNF-receptor-associated factors
TRAIL	TNF-related apoptosis-inducing ligand
TRAMP	TNF-related apoptosis-mediating protein (receptor)
TRK	growth-factor tyrosine-kinase
TSF	serine chemotaxis receptor
UAS	upstream activating sequence
USP	ultraspinacle protein
UTP	uridine triphosphate
UV	ultraviolet
VDR	vitamin D receptor
VEGFs	vascular endothelial growth factors
VGF	vaccinia virus growth factor
WASP	Wiskott–Aldrich syndrome protein
ZAP-70 kinase	ζ-associated protein 70 kinase
ZNF	zinc finger
ZIP	zinc finger protein

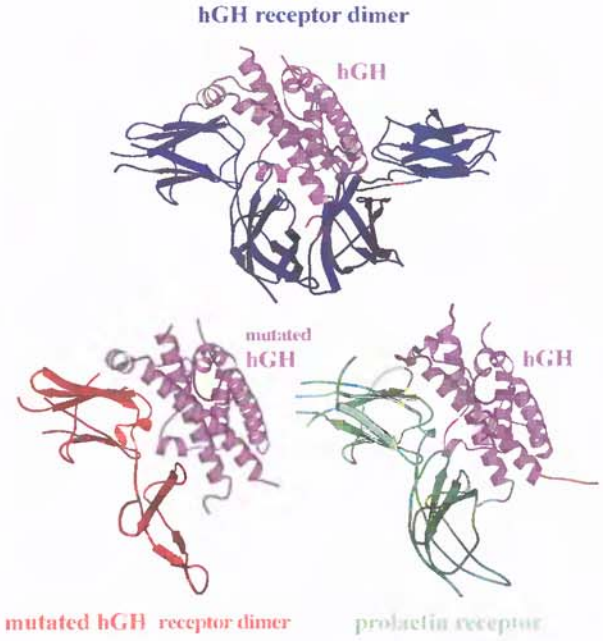


Plate 1 *Top*: A ribbon presentation of hGH complexed with the dimeric extracellular binding domain of the hGH receptor. hGH is in purple and the dimer of hGHbp is in blue. hGH is a four-helix bundle. Each hGHbp binding domain is composed of two β-sheets, one amino- and one carboxy-terminal. Each sheet contains seven β-strands. Shown is a hGhbp dimer.
Right: Ribbon presentation of the hGH–hPRL-receptor complex. The four helices of the hGH are in purple and the β-sheets of the two binding domains of the prolactin receptor are in green.
Left: The hGHbp is now in red. A critical tryptophan at position 104 in the hGHbp was deleted, enlarging the binding cavity (compare with the top picture). Mutations of five residues, K168R, D171T, K172Y, E174A, and F176Y in hGH gave a ligand molecule that could fill the cavity and restore binding. (K is lysine, R is arginine, D is aspartic acid, T is threonine, Y is tyrosine, E is glutamic acid, A is alanine, and F is phenylalanine). (Reproduced from the data in ref. 1, Fig. 1 and in ref. 2 of Chapter 2, with permission of the authors and Science and Nature.)

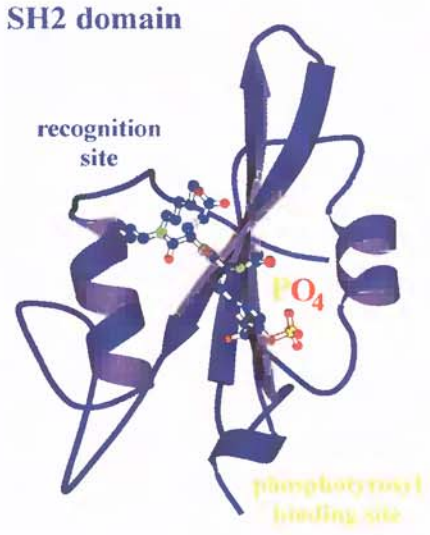


Plate 2 Structure of a typical SH2 domain, bound to a phosphotyrosyl containing peptide. The SH2 domain belongs to the cytosolic pp60 src tyrosine kinase. The peptide is a synthetic acyl formyl phosphotyrosyl-glu-(*n*-*N*-dipentenyl) amine. Only, the recognition site and the phosphotyrosyl interaction site are shown. (The ribbon model is reconstructed with permission of the authors and Biochemistry from data in ref. 22 of Chapter 3, deposited in data banks.)

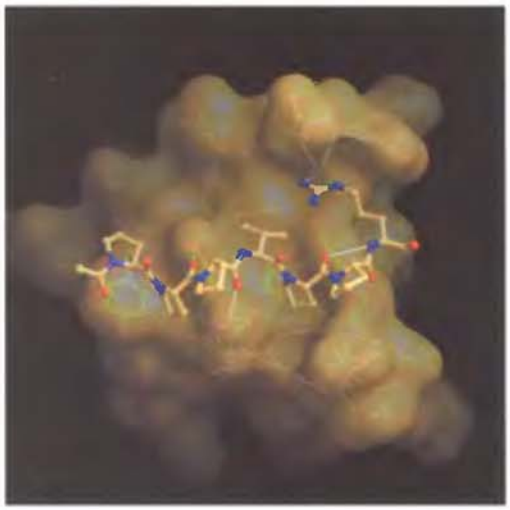


Plate 3 A view of the crystal structure of the SH3 domain of the human, cytosolic tyrosine kinase Fyn, complexed with a proline-rich SOS peptide (see ref. 35 of Chapter 3). The picture presented shows that a SH3 domain is a globular structure of approximately 60 amino acids with a well-conserved β-sheet structure (see ref. 36 of Chapter 3). (This picture was generously contributed by Dr Martin Noble, University of Oxford.)

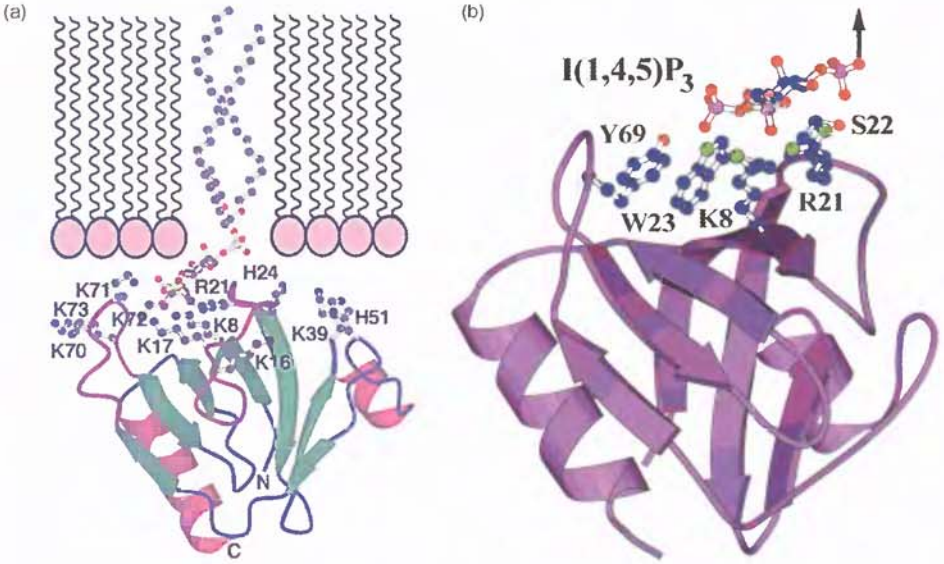


Plate 4 (a) The structure characteristic of PH-domain-containing proteins. PH domains contain roughly perpendicular anti-parallel β-sheets, arranged around a C-terminal amphipathic α-helix. Both slabs of β-sheets are relatively flat and together form a wedge. The carboxy-terminal α-helix closes the wedge on the back side. The β-sheets and the α-helix show highest sequence homology among PH-domains, whereas the loops between the β-strands are quite variable. The interaction of PH-domain proteins with the membrane is indicated schematically. This scheme was generously contributed by Professor Hartmut Oschkinat and is shown with permission of the authors and the EMBO Journal, (ref. 47, Chapter 3). (b) Structure of the binding site for inositol-1,4,5-trisphosphate (I (1,4,5)P_3) in the PH domain of β-spectrin. The structure of IP_3 is shown as a ball and stick model. PH domains in proteins are electrostatically polarized and have a positively charged surface, facilitating interaction with negatively-charged phospholipid headgroups in membranes. Phospholipids, such as phosphatidyl inositol-4,5-bisphosphate and IP_3, have been shown to associate with PH domains. (ref. 49 of Chapter 3). The aminoacids in the PH domain which are responsible for this interaction are Y (tyrosine), W (tryptophan), K (lysine), R (arginine), and S (serine). In many PH-domains there is a group of lysines and arginines. (This ribbon model was reproduced with permission of the authors and Nature from data in ref. 48 of Chapter 3, available in protein data banks.)

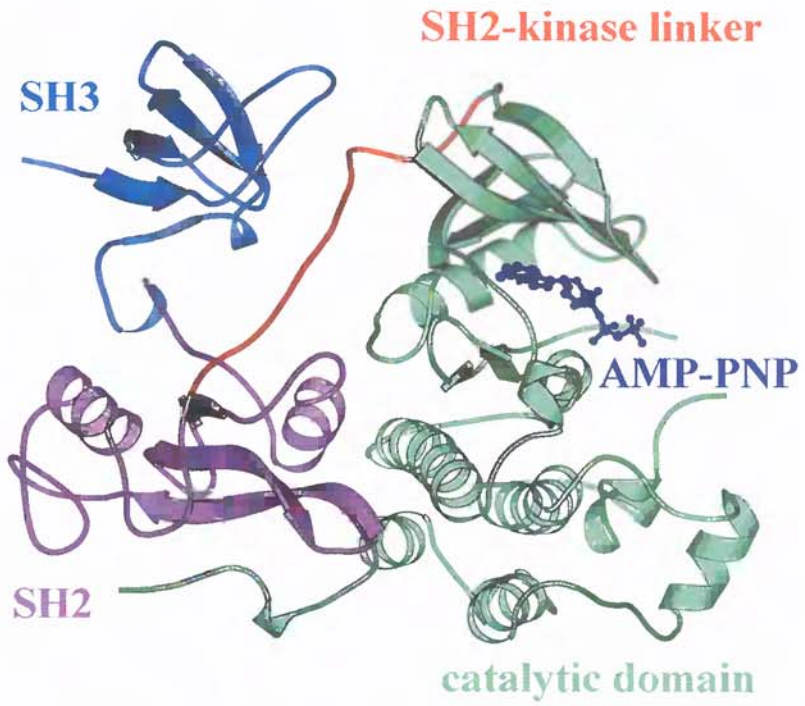


Plate 5 Structure of the tyrosine kinase Hck complex with AMP-PNP bound to the active site. The cytosolic Hck kinase is a member of the c-Src kinase family. The molecular surfaces of the SH3 domain of the kinase are in blue, the SH2 domain is in purple and the linker and the catalytic region are in green. The non-hydrolysable ATP analogue, AMP-PNP is a stick and ball model in dark purple. (ref. 65, Chapter 3; see also refs 66 and 67 of Chapter 3, and for mechanistic comparisons the crystal structure of the catalytic subunit of the cAMP-dependent kinase, protein kinase A, ref. 68 of Chapter 3). Comparing the crystal structures of the catalytic domains of the inactive, c-Src- and Hck-kinases and the active form of a related Src family kinase, such as Lck, revealed that in the inactive conformation Tyr 527 in the carboxy terminal tail is phosphorylated, whereas in the active state, it is unphosphorylated and instead Tyr 416 in the activation loop of the kinase domain is phosphorylated. (refs 65, 67, and 69 of Chapter 3). Phosphorylation of tyrosine 527 in the tail directs the SH3 domain to a position where it can bind to the kinase linker region. This is the key interaction which blocks the kinase activity, because the linker region can now contact the small lobe of the kinase catalytic domain and block it. When Y527 is unphosphorylated, the SH3 domain loses its contact to helix C and the SH2 kinase linker becomes flexible. The helix rotates, the block is removed and the kinase becomes active. Phosphorylation of Tyr 416 in the activation segment re-orients Arg 385 and forms the ATP-binding site. (The ribbon model is reproduced with permission of the authors and Nature from Fig. 2 in ref. 65 of Chapter 3.)

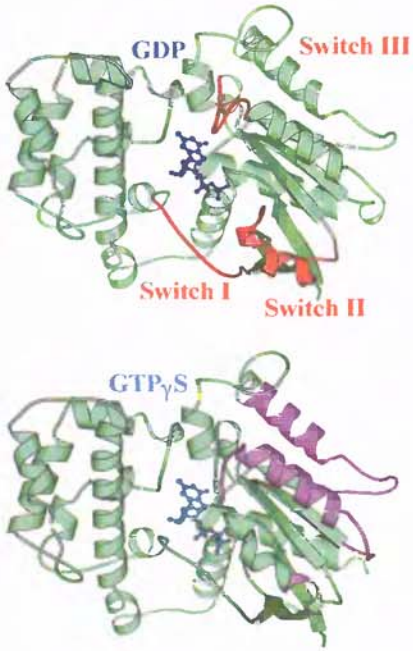


Plate 6 Comparison of structures of activated GTPγS-bound- and inactive GDP-bound G-α-subunits of transducin. (Reproduced with permission of the authors and Nature from Fig: 1*a* and *b*, in ref. 84 of Chapter 3) GTPγS is a non-hydolysable GTP analogue, where one of the oxygens of the γ-phosphate is replaced by a thiol group. Binding of GTPγS permanently activates G-α-subunits. GTPγS and GDP are presented as ball and stick models. The Switch regions involved in the transition from the 'off' to the 'on' state and the face of the GTPγS-activated G-α subunit of transducin where it interacts with the effector cGMP phosphodiesterase are shown in purple. The interaction domains of heterotrimeric G proteins with their effectors were characterized by H.E. Hamm and colleagues. They are pointed out on the basis of the information in ref. 85 of Chapter 3, with permission of Professor Heidi Hamm and her colleagues.

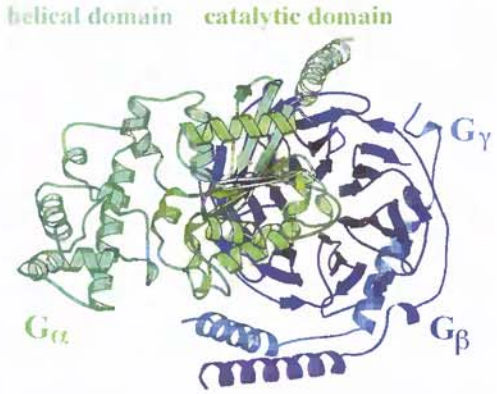


Plate 7 Ribbon presentation of the α,β,γ-holocomplex of transducin. The GTPase activity of heterotrimeric G proteins, resides in the larger, 42–45 kDa, α-subunits with the helical domain, (left). The Gα subunit is in green. The G β-subunit is in dark blue and the G γ-subunit is in light blue. (Reproduced with permission of the authors and Nature from ref. 86 in Chapter 3.)

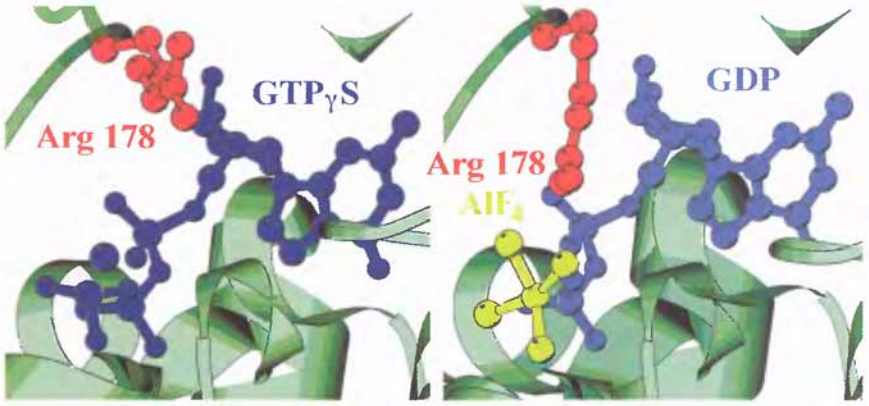


Plate 8 G_i-α1 is the α subunit of a heterotrimeric G protein coupled to β-adrenergic receptors. Gi-α1 is assumed to regulate the activity of the target enzyme, adenylylcyclase, which produces the second messenger, cAMP. On the left is shown the complex of the α-subunit, G_i-α1, with GTPγS, and on the right with GDP-AlF_4^-. (AlF_4^-, in the presence of GDP, is a powerful activator of the GTPase of α-subunits of heterotrimeric G proteins, because AlF_4^- mimics the γ-phosphate of GTP and adds the missing γ-phosphate to the GDP-enzyme complex.) The conserved arginine residue, Arg 178, is in red, AlF_4^- is in yellow, GDP is in light blue and GTP γS is in dark blue. (These ribbon models are based on the information in ref. 89 of Chapter 3, available in the databanks, and reproduced with permission of the authors and Science.)

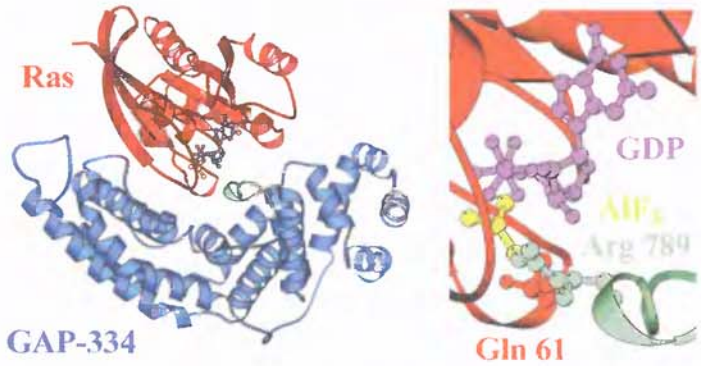


Plate 9 The crystalline structure of a complex of a part, (GAP-334), of the p120 neurofibromin, NF- GAP, with GDP-AlF_3^- -bound p21 H-Ras. GAP contributes the essential Arg residue, (Arg^{789}), missing in Ras, and consequently, the Ras-GAP complex acquires GTPase activity. The structure of a part of the GAP, (GAP 334), with p21 H-Ras is on the left, (Ras is in red and GAP 334 is in blue). On the right is shown the active site of Ras with bound GDP and AlF_3 in the complex with GAP 334. Only in the Ras–GAP complex is the GTPase reaction of Ras turned on by AlF_3, because the GAP contributes the essential Arg residue, (Arg^{789}), that stabilizes the transition state and neutralizes the charge. Moreover, GAP stabilizes the switch II region and the active conformation of Ras. (AlF_3 is in yellow and the Arg^{789} contributed by the GAP is in green).The provision of the essential Arg-residue is the key to the activation of the GTPase activity. It also explains why the Ras-GTPase, when complexed with GAP, becomes responsive to GDP–AlF_3, because only the Ras–GAP complex can position AlF_3, just like the γ-phosphate of GTP is positioned. This can be deduced from the structure of the α-subunit of G_i complexed with GDP–AlF_4^-, which shows a drastic reorientation of Arg 178 (see Plate 8). (These ribbon models are based on the information in ref. 97 of Chapter 3, reproduced with permission of the authors and Science).

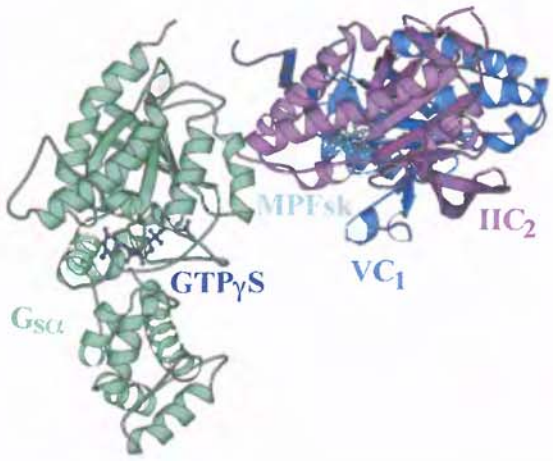


Plate 10 Part of the complex of G_s-α with the cytoplasmic domains of a reconstructed soluble, catalytically-active form of adenylylcyclase. VC_1 is a fragment of canine adenylylcyclase, type V and IIC_2 is the C_2 domain of rat adenylylcyclase type II. The VC_1 and the IIC_2 domains form an active complex with G_s-α–GTPγS. MPFsk is forskoline, a potent activator of adenylylcyclase. It is shown as a ball and stick figure. Activation promotes the formation of a complex between the VC_1 and the IIC_2 domains. The activated complex binds a molecule of the non-hydrolysable ATP analogue, Ap(CH_2)pp. The complex is shown with the soluble C_1 domain of adenylylcyclase in blue and the C_2 domain in purple. G_s α is in green and GTPγS is depicted as a ball and stick figure. The catalytic part of G_s α with the binding site for GTPγS is shown on top, and the GAP domain below. The data are from ref. 13 and the Brookhaven protein data bank. They are reproduced as a ribbon model with permission of Professors Sprang and Gilman and Science. The structure allows some speculation about the mechanism of adenylyl cyclase. These speculations are based on the similarity of the α,β-helical folds of adenylylcyclase with the catalytic site domain of DNA polymerase I. The resemblance between the adenylylcyclases and the polymerases includes the acidic residues which coordinate metalions (ref. 14, 15 of Chapter 5). Based on their structure, the authors propose that binding of G_s-α to adenylylcyclase induces a change in the relative orientation of the C_1 and C_2 domains, forming an active, catalytic site.

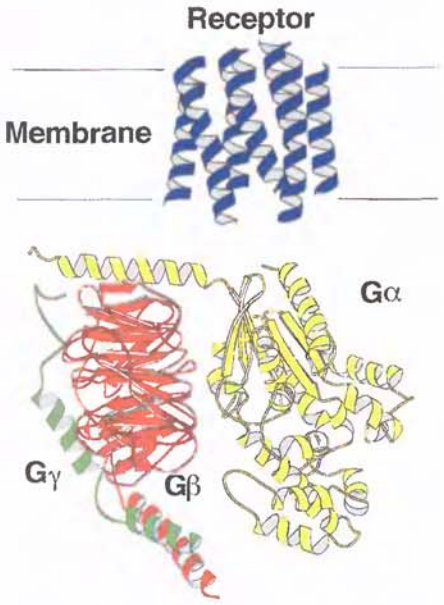


Plate 11 A model visualizing one possible position of a heterotrimeric G protein in relation to a G protein-coupled heptahelical receptor, GPCR, and the plasmamembrane. The receptor is in dark blue. The Gα subunit is in yellow, the Gβ subunit in red and the Gγ subunit in green. The model was constructed by Dr. Cornelius Krasel, Department of Pharmacology, University of Würzburg and is shown with his permission.

Plate 12 (*left*) The structure of Phosducin. Phosducin is shown in red. The Gβ-subunit of a heterotrimeric G protein is in dark blue and the Gγ-subunit is in light blue. The Gβ subunit is a protein with WD-repeats, forming β-propellers. (W is trytophan, D is aspartic acid). Phosducin has two domains: the N-terminal domain is helical and binds to the β-subunit of the $\beta\gamma$-complex. Since the interaction site of the β-subunit with phosducin is the same with which the β-subunit binds to the α-subunit in the heterotrimeric $\alpha\beta\gamma$-complex, it follows that binding of the β-subunit to phosducin precludes formation of the $\alpha\beta\gamma$-complex. The C-terminal domain of phosducin is juxtaposed to that part of the β-subunit that interacts with the γ-subunit of the heterotrimeric, $\alpha\beta\gamma$-G protein. This site is also in the neighbourhood of the farnesyl moiety which is attached to the γ-subunit. The N- and the C-terminal domains cover the site and the tip of the β-subunit propeller like a blanket. The end of the N-terminal domain is flexible. It contains the Ser 73 phosphorylation site. The C-terminal domain is a 5-stranded β-sheet, surrounded by α-helices on both sides. Such a structure is characteristic for proteins with a 'thioredoxin fold'. There are two binding regions for β,γ-subunits. One is a large (63 or 105 aminoacids residues long) N-terminal region, the other is a small region at the C-terminus of phosducin which also binds $\beta\gamma$-subunits with comparable affinity. This explains why the phosducin-like proteins, which are homologues of phosducin, but lack an N-terminal domain, still bind $\beta\gamma$-subunits. What remains a puzzle is how phosducin also manages to bind to the α-subunit. (Reproduced with permission of the authors and Cell from ref. 69, Chapter 5; see also ref. 70, Chapter 5).

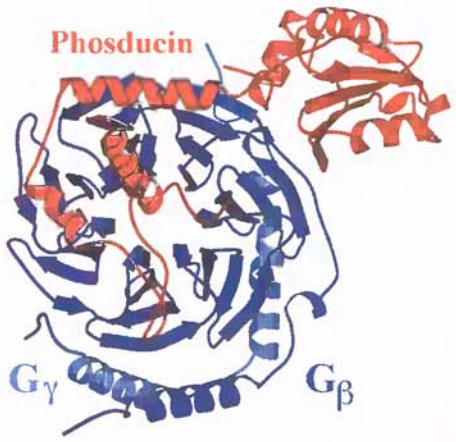


Plate 12

Plate 13

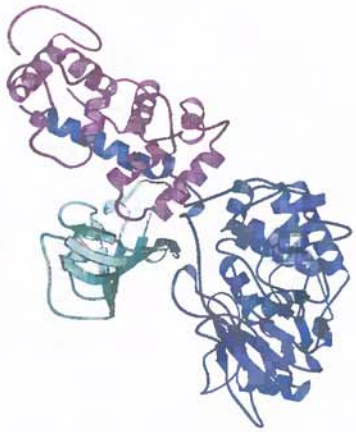


Plate 13 (*right*). The Ser/Thr phosphatase calcineurin consists of two tightly bound subunits; a catalytic subunit, (A), and a regulatory subunit (B). The catalytic subunit A shares over 40% sequence identity with other Ser/Thr phosphatases. The structure of the ternary complex of a fragment of the calcineurin phosphatase with the FK-binding protein, FKBP 12 and the immunosuppressant drug FK 506 has been determined and a ribbon model of the catalytic and regulatory subunits of the calcineurin phosphatase together with the FKBP is shown. The catalytic subunit is in blue and the regulatory subunit is in purple. The binding protein, FKBP, to which the immunosuppressant FK 506 binds is in cyan green. (The ligand of the binding protein is not shown). The catalytic subunit is the largest part of the complex. It is characterized by a β-sandwich motif. The most striking feature of the complex is the 22 residue α-helix of the regulatory subunit which links the regulatory B subunit and the catalytic A subunit which are 40 Å apart from each other. The regulatory B component contains on the upper surface two calmodulin like domains. The FK-binding protein interacts with the exposed underside of the α-helix of the regulatory subunit B and does not contact the active site in A. The FKBP is more than 10 Å removed from the catalytic site. With the immunosuppressant, FK 506 bound, it is even further away, 25 Å. Thus, the immunosuppressant can not inhibit the phosphatase directly. But, the ternary complex is so arranged that substrates can not gain access to the active site of the phosphatase. (Reproduced with permission of the authors and Cell from ref. 10 of Chapter 7).

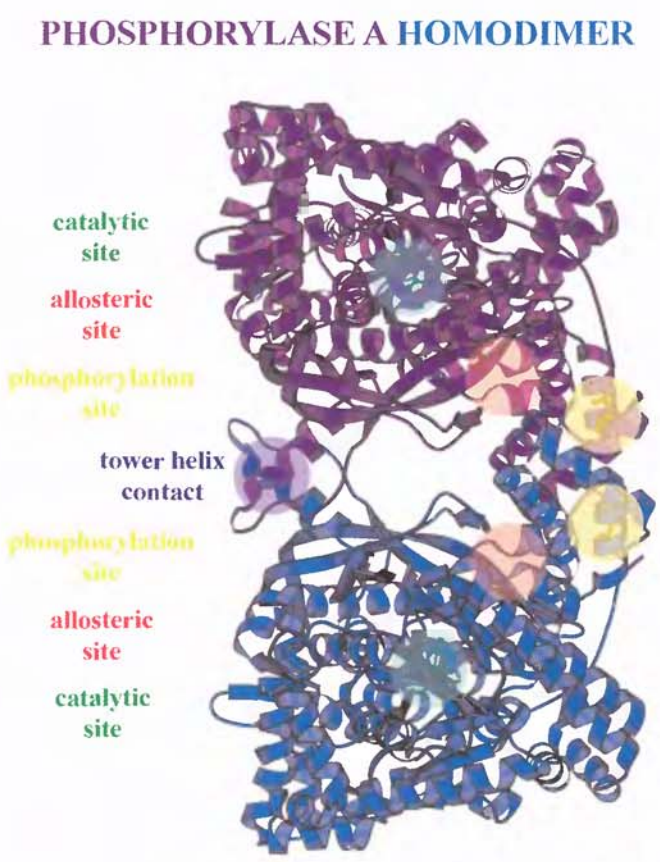


Plate 14 The active, dimeric R-state of rabbit skeletal muscle glycogen phosphorylase a. (The terms R- and T-states were introduced by Jacques Monod, see below ref. 20, Chapter 7, R is the relaxed conformation which is active and T is the tight conformation which is inactive). The catalytic sites are indicated in green. (Pyridoxal-5' phosphate, ref. 19, Chapter 7, which participates in catalysis is shown as a ball and stick structure). The allosteric sites to which 5' AMP binds are in pink. The phosphorylation site on Ser 14 is in yellow. Both R- and T-state structures have a similar topology, but different tertiary and quaternary structures. Comparison of both structures revealed the conformational changes resulting from phosphorylation of Ser 14. The N-terminal tail residues are poorly ordered in the T-state. The transition from the T-state phosphorylase *b* to the R-state phosphorylase *a* results in a change of the N-terminal residues to an ordered coiled conformation. The consequence of these changes is that the N-terminal residues make *intrasubunit* contacts in the T-state conformation of phosphorylase *b*, but now make *intersubunit* contacts in the R-state conformation of phosphorylase *a*. One of the important subunit-subunit interactions in the dimer involves the contacts between the α 2-helix of one subunit with the cap-region of the other subunit in the dimer. (The cap residues form a loop between the first and the second α-helices of the other subunit). The interactions made by the N-terminal residues induce, (1.2 to 2.4-Å), structural changes at the cap region. The tower helix contacts are central to the transition. They are shown in purple. The tertiary conformational changes on transition to the active R conformation lead to extensive quaternary structural changes in the orientation of the subunits in the dimer. This rearrangement consists of a 10° rotation of one subunits to the other, about an axis approximately perpendicular to the 2-fold axis of the dimer. The tertiary conformational changes are dampened and reduced to 0,5–1.3 Å by the simultaneous quaternary changes. The correlation between tertiary and quaternary conformational changes provides also a clue to the structural basis for non-covalent 'allosteric' activation by binding the allosteric effector, 5'-AMP. (Reproduced with permission of Professors Louise N. Johnson, and David Barford and J. Mol, Biol. from ref. 17 of Chapter 7. See also the phosphorylase structure on the cover of this book which was generously provided by Professor Louise N. Johnson, University of Oxford).

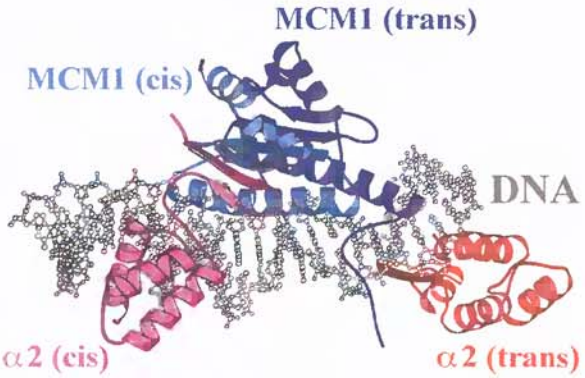


Plate 15 Crystal structure of the yeast MAT-α 2/MCM1/DNA ternary complex. MAT-α 2 is the mating-type protein α2 in the yeast. It is a transcriptional repressor. MAT-α2 interacts with MCM1, a signal-(pheromone)-responsive transcriptional regulator. The DNA is from an operator (STE6), which controls a homeobox DNA. A ribbon representation of the ternary complex is shown. The structure shows a slab of MCM1, flanked by two copies of MATα 2 (abbreviated α 2). One copy of MCM1 in light blue, is in *cis*-position, relative to DNA, the other in dark blue, is in *trans*-position (*cis* and *trans* describe the relative positions of the protein on the DNA. Cis is on the same side as the transcription initiation site). The MCM1/α2 complex covers the DNA extensively, making contact with 24 bps. The α-helices are located parallel to the minor groove of the DNA and extend to the major groove. This causes the DNA to bend around MCM1. The homeodomain of the MATα2 protein binds to the major groove of the DNA and the N-terminal arm inserts into the minor groove. In the *cis*- configuration, α2 is in dark red and in the trans-configuration in light red. In the cis-configuration, α2 is extended. This is made possible through bending of the DNA. The trans interaction of α is essentially like the cis interaction). (The ribbon structure was reconstructed with permission of the authors and Nature from data in figure 1a of ref. 14 of Chapter 9).

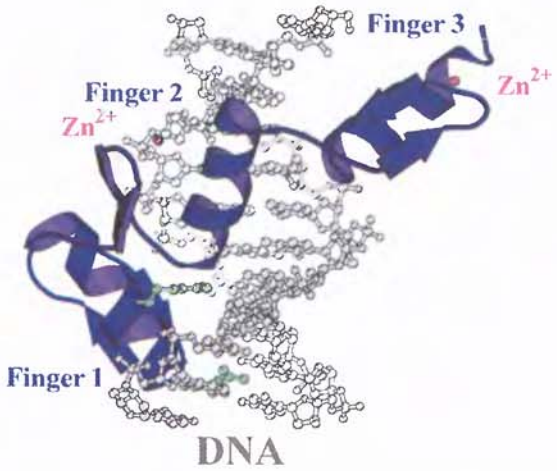


Plate 16 DNA recognition by a zinc-finger protein. Structure of a murine immediate early gene DNA-binding protein complex with an 11 bp DNA fragment. The protein ZIF 268 contains three Zn fingers. The DNA-binding Zn-finger protein, shows a characteristic feature of Zn-finger proteins, namely a modular structure. Individual Zn fingers are repeated in a way that each Zn-finger can contact 3bps, (basepairs). The contacts are made with a single strand of DNA. Arginine–guanine contacts seem to be most important. (The ribbon structure is reconstructed from data in ref. 16 of Chapter 9 deposited in the protein database, with permission of the authors and Science).

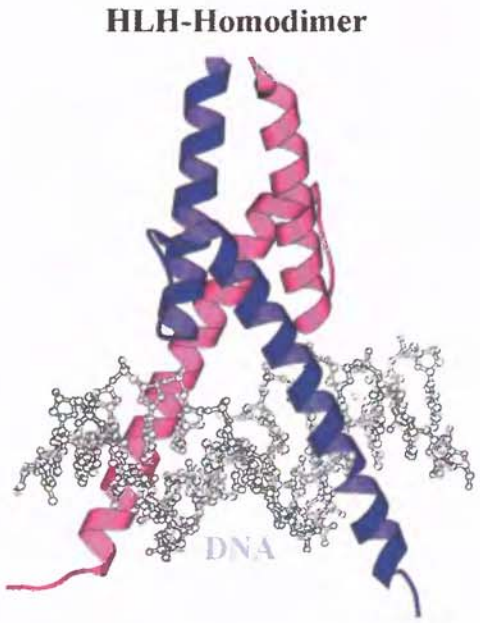


Plate 17 A mutant mouse basic helix-loop-helix domain (bHLH) of a Myo D protein (residues 102–166, with cysteine 135 replaced by serine). The bHLH domain is complexed with 5'-deoxy-P-TCAACAGCTTGA-3'DNA. The HLH homodimer is presented. One HLH-motif is in blue and the other is in pink. Members of the Myo D family of proteins are transcriptional activators, that play a central role in muscle development and in the formation of specific muscle cells. (The ribbon structure was reconstructed with permission of the authors and Cell from data in ref. 20 of Chapter 9).

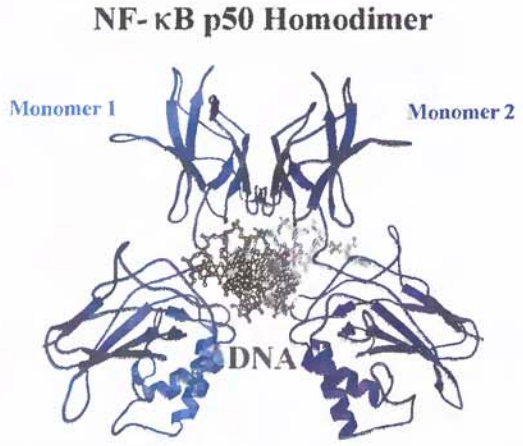
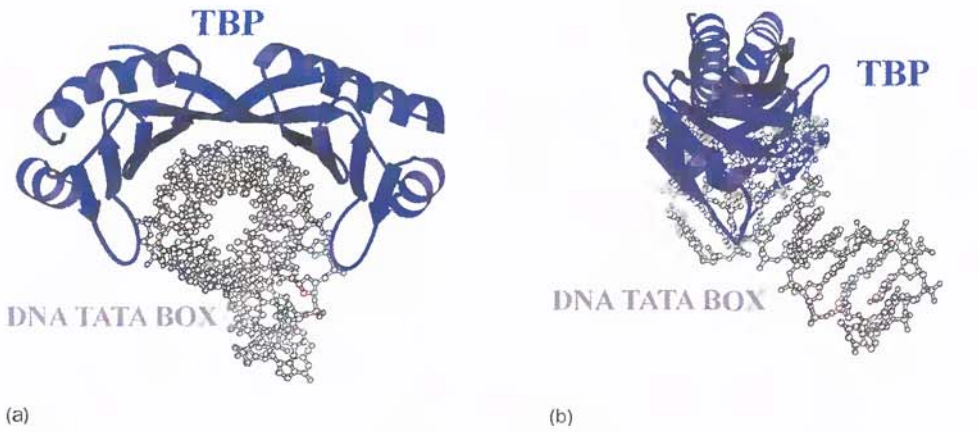
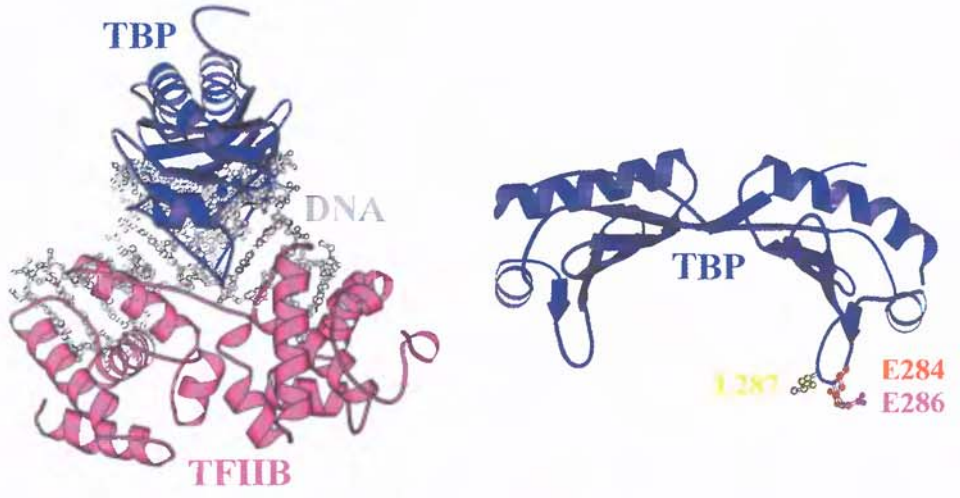


Plate 18 Structure of the NF-κB p50 homodimer bound to a 10 base pair κB- responsive DNA site. The structure of the DNA-binding p50 subunit of NF-κB complexed with DNA is an example how β-sheets make specific contacts with the DNA. The p50 NF-κB subunit dimerizes. Both halves are folded as a β-sandwich. The dimer wraps around the major groove of the unfolded DNA double helix, and the β-sheets make the specific contacts. The N-terminal part of NF-κB p50 is similar to the core domain of the tumour suppressor, p53 (see Chapter 15). (The ribbon structure was constructed with permission of the authors and Nature from data in ref. 23 of Chapter 9; see also ref. 24 of Chapter 9.)



(a) (b)

Plate 19 (a) The crystal structure of a human TATA-box-binding protein (hTBP) complexed with DNA of an adenoviral promoter containing the TATA box. (Reproduced from ref. 26 of Chapter 9, with permission of Professors R.G. Roeder and S.K. Burley and Proc. Natl Acad. Sci. USA). See also: Ref. 25. The core DNA-binding domain of hTBP, contains residues 155-333 of the holo-TBP and the DNA contains 16 bp. The hTBP is similar to a yeast TATA box binding protein and to the *Arabidopsis thaliana* TBP. (b) The structure of a yeast TATA-box-binding protein. (Reproduced from ref. 27 of Chapter 9 with permission of the authors and Nature). TBP binds to DNA with its β-sheets. TBP has a saddle-shaped structure formed by a curved antiparallel β-sheet and binds to the minor groove of the TATA-box DNA. Binding of the protein to the DNA is by an induced fit mechanism and is accompanied by a major distortion of the DNA, including a kink, for which a phenylalanine of TBP is responsible, whereas the conformational change in the protein is quite subtle. The final stabilization of the protein–DNA complex requires bending and partial unwinding of the DNA.



(a) (b)

Plate 20 (a) Structure of a TFIIB–TBP/TATA element DNA ternary complex. (b) MOLSCRIPT structure of a TBP (*Arabidopsis thaliana* TBP-2), with the position of residues which are crucial for the interaction of a TBP with TFIIB. L is leucine and E is glutamic acid. (Reproduced with permission of Professors R. G. Roeder and S.K. Burley and Nature from data in ref. 28 of Chapter 9).

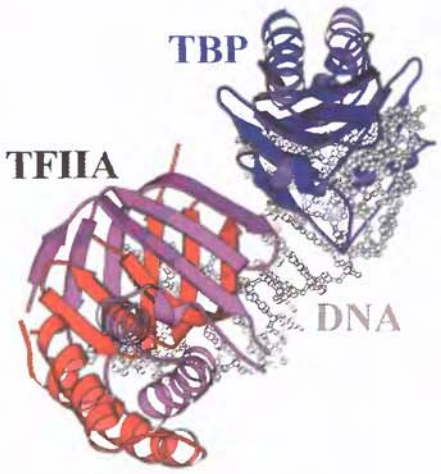


Plate 21 Structure of a yeast TFIIA–TBP–DNA complex. TFIIA binds as a tightly associated dimer, composed of a small and a larger subunit (ref. 30 of Chapter 9). The interface between them forms a large, solvent-excluded area of 6000 Å^2 which connects the two subunits, securely, making it unlikely that either of them can exist without the other as a separate structural and functional entity. The structure shown is that of yTFIIA (*y* is yeast), but is representative for hTFIIA (h is human) because TFIIA is a structurally and functionally highly conserved molecule. Therefore, yTFIIA can be substituted for human TFIIA. in *in vitro* transcription. (Reproduced with permission of the authors and Nature from data in ref. 32 of Chapter 9; see also ref. 31 of Chapter 9.)

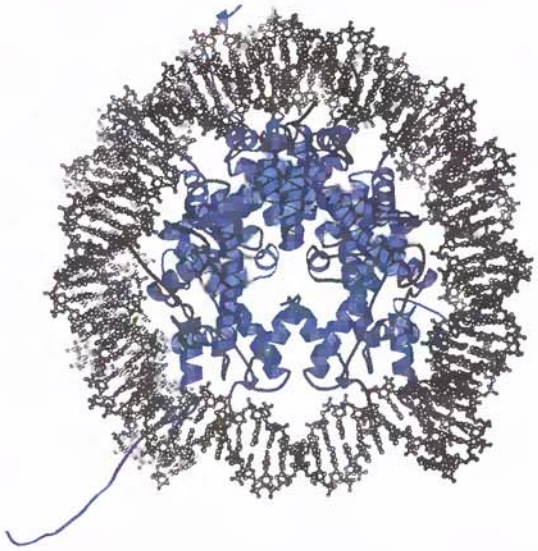


Plate 22 Core particle of the nucleosome. The 146 bp DNA superhelix is on the outside and the eight histone protein chains are in blue and on the inside. The view is down the DNA superhelix. The structure shows in atomic detail how the histone protein octamer is assembled and how 146 base pairs of DNA are organized into a superhelix around it. Histone amino-terminal tails pass over and between the gyres of the DNA superhelix to contact neighbouring particles. The lack of uniformity between multiple histone–DNA-binding sites causes the DNA to deviate from ideal superhelix geometry. (Reproduced with permission of the authors and Nature from ref. 33 of Chapter 9; see also ref. 34 of Chapter 9 for comments.)

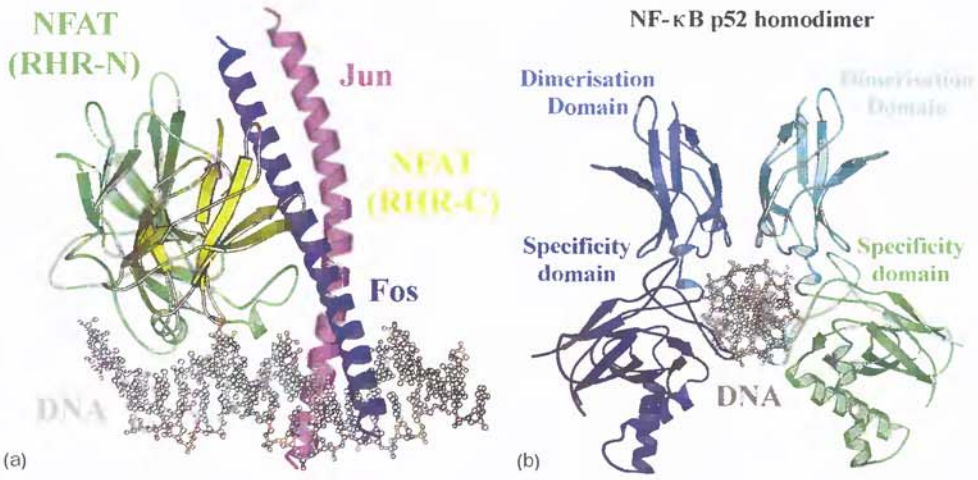


Plate 23 The NFAT-Jun-Fos-DNA complex: (a) The N-terminal Rel homology region, (RHR-N), of the recognition domain of NF-AT is in green and the C-terminal Rel homology region, (RHR-C) is in yellow, Fos is dark blue and Jun is purple. The structure shows the interface between NF-AT and the Jun-Fos heterodimer. This helps to explain how NF-AT and the Jun-Fos heterodimer together activate synergistically the expression of immune-response genes in the ARRE-2 region. The tight association of the three proteins with DNA covers 15 base pairs of ARRE 2. An arginine (Arg 285), on the Jun helix is involved in the interactions of NF-AT and Jun-Fos with the DNA. The cooperative interaction of NF-AT and Jun-Fos with the DNA is facilitated not only by bending of the Fos component, but also and equally important by bending of DNA. (Reproduced with permission of the authors and Nature from ref. 16 of Chapter 10 and from the corresponding data deposited in protein data banks.) (b) Another member of the Rel-family of transcription factors is NF-κB. The structure of the human NF-κB p52 homodimer–DNA complex is shown. This provides a good view of how a dimeric transcription factor is arranged on the DNA in a way that allows the loops of the specificity domains of both halves of the dimer to interact with specific target regions on the DNA. The role of dimerization of a transcription factor in target recognition on the DNA should be noted. (Reproduced from Fig. 2a in ref. 19 of Chapter 10 with permission of the authors and the EMBO Journal. See also: Refs: 17, 20 in Chapter 10.)

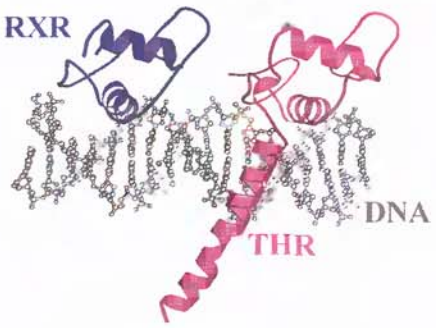


Plate 24 The structure provides information about the remarkable fidelity with which nuclear receptors recognize response elements in the DNA. The structural determinants which direct nuclear receptor heterodimers, consisting of the 9-*cis*-retinoic acid receptor RXR and the thyroid hormone receptor THR to the DNA are indicated. The receptor heterodimers recognize response elements, composed of two direct repeats of the sequence, 5'-AGGTCA-3', separated by 1–5 bp. The protein–DNA interface is extensive. Two-thirds of the accessible surface is buried alone by the THR–DNA interaction. The interaction of the DBD, (DNA binding domain), of the THR with the minor groove of the DNA makes possible an extensive interaction of the core DBD of THR with the major groove of the DNA. The contacts of the DBDs of each receptor bridge the minor groove of the DNA double helix. The residues contributing to the dimer interface are remarkably specific, allowing only the formation of certain heterodimers. (This ribbon structure was reproduced with permission of the authors and Nature from data in Fig. 5 of ref. 37 of Chapter 11, also available in protein databanks.)

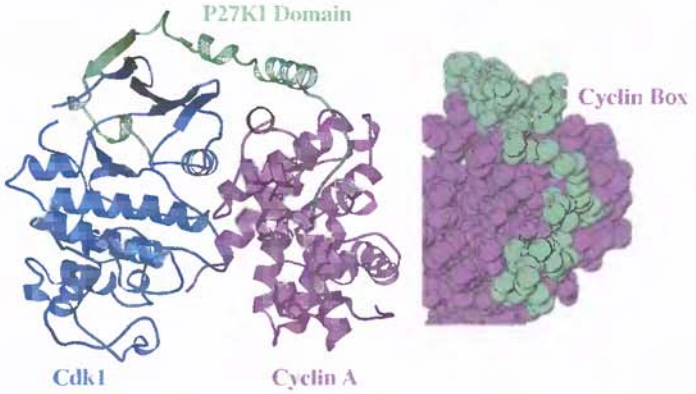


Plate 25 *Left*: Structure of the $p27^{KIP-1}$ cyclin-dependent-kinase inhibitor (CKI) bound to the cyclin A–Cdk1 complex. The structure shows that the CKI (in green) interacts both with cyclin A (in purple) and with Cdk1 (in blue). The amino-terminal region of $p27^{KIP-1}$ binds to cyclin, and the carboxy-terminal region interacts with Cdk1. On cyclin A, it binds in a groove, formed by conserved cyclin box residues (see below). On Cdk 1, the inhibitor binds to the amino-terminal lobe and disturbs it. It also inserts in the catalytic cleft of the kinase, mimicking ATP and inhibiting the Cdk. The interaction is synergistic. explaining why $p27^{KIP-1}$ binds more avidly to the Cdk 1–cyclin A complex, than to each separate component. (Reproduced with permission of the authors and Nature from ref. 9 of Chapter 12 and from the corresponding data available in protein data banks.) *Right*: Structural scheme of the cyclin-box motif of cyclin A to which the $p27^{KIP-1}$ binds. The peptide-binding groove of cyclin A. in purple, is filled by the green balls of the $p27^{KIP-1}$ protein. The cyclin box is a rather common, stable structural fold, which comprises a bundle of five tandemly repeated α-helices. On the surface of the cyclin box are clusters of conserved residues, which, in the case of the cyclin A–Cdk1 pair, have been shown to be part of the interaction surface of the cyclin with the cognate kinase.

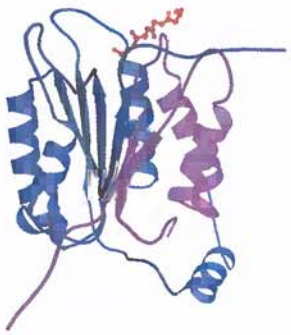


Plate 26 The structure of human ICE in complex with an inhibitor. This is a stereo-ribbon drawing of the ICE heterodimer. The larger (p20) unit is in blue and the smaller (p10) unit is in purple. The active site is on top, bound to a tetrapeptide inhibitor. The inhibitor, acetyl-Tyr-Val-Ala-Asp is in red, as a ball-and-stick structure. The p20 and the p10 subunits are tightly associated. The enzyme core is a six-stranded β-sheet with five parallel strands and one antiparallel strand. The antiparallel strand is at the edge of the β sheet. Six α-helices lie parallel to the β–strands and prevent access of solvent to the β–sheet. The last seven, carboxy-terminal residues of p20 and the first, seven aminoterminal residues of p10 stickout of the compact structure and form two antiparallel β–strands that interact extensively with each other. It was proposed that the active form of ICE in solution is a tetramer where the p10 subunit from one ICE monomer forms a complex with a p20 subunit of another ICE. This facilitates processing and proteolytic activation of ICE. because two of the cleavage sites, involved in the activation of ICE come to lie adjacent to each other in the tetrameric arrangement. Finally, the structure confirms the relationships between human ICE and cell-death proteases in other organisms, which also associate to form a tetramer. (Reproduced with permission of the authors and Nature from the data in ref. 18 of Chapter 13, and corresponding data available in protein data banks.)

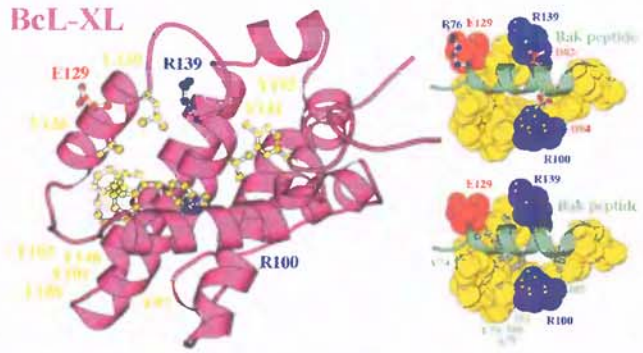


Plate 27 The 16-aminoacid Bak-peptide, complexed with a biologically active truncated mutant of Bcl-XL The structure of the complex is similar to the X-ray and NMR structure of uncomplexed Bcl-XL. On the left are shown the sidechains in the binding site of Bcl-X_L, involved in the binding of the Bak peptide. Hydrophobic sidechains facing the peptide are in yellow. Sidechains of positively charged residues are in blue and negatively charged residues in red. R is arginine; L is leucine, E is glutamic acid, V is valine, F is phenylalanine, Y is tyrosine, A is alanine, I is isoleucine, and D is aspartic acid. The peptide side chains are given as atomic ball-and-stick structures. On the right are two surface representations of the binding pocket of Bcl-X_L, to which the Bak peptide is bound. The sidechains of residues of the Bak peptide, involved in binding are in green. The Bak peptide binds to a hydrophobic cleft of Bcl-X_L. When bound, it forms an α-helix, although it is a random coil in solution. Again, hydrophobic residues of Bcl-X_L in contact with the peptide are in yellow. Sidechains of positively charged residues are in dark blue and negatively charged residues in red. There are two central hydrophobic α-helices surrounded by amphipathic helices. The two central helices, α-1 and α-2 are connected by a long flexible loop, which contains a phosphorylation site. It may be a negative control region. The fifth and sixth of the amphipathic α-helices participate together with the hydrophobic α-helices in forming the mitochondrial channel and the region, spanning the helices α-2 to α-7 is the domain interacting with other proteins. The Bcl-proteins contain additional homology domains, some of them are involved in dimerization. Finally, the carboxy-terminus carries a membrane anchor. (These ribbon diagrams are reproduced with permission of Dr. S.W. Fesik and colleagues and Science and Nature from data in refs 27 and 28 of Chapter 13 and the corresponding data stored in protein data banks.) For further information, the original literature should be consulted (See refs in Chapter 13).

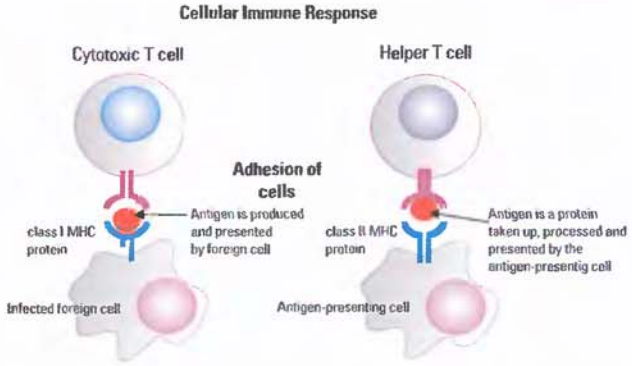


Plate 28 T cells react with antigen presented by another cell. Therefore, the immune response by T cells involves cell–cell interactions. An antigen bound to a class I MHC molecule is part of an infected cell. It is a marker recognized by cytotoxic T cells which kill the antigen-presenting cell. On the other hand an antigen associated with a class II MHC complex is a foreign compound that has first been taken up by the antigen-presenting cell, APC, dendritic cells, and macrophages, where it is processed proteolytically, before it is presented to helper T-cells. Helper T-cells are involved in transplantation reactions. (This is reproduced with permission of Taylor and Francis, Inc. Fig. 23–47 of Chapter 14 in ref. 2 served as a model.)

Plate 29 (*left*) The crystal structure of the ternary complex between a soluble, human T-cell antigen receptor, specific for a peptide, TAX, of the human T-cell lymphotropic virus (HTLV-1), bound to the human MHC-molecule, HLA-A2. The structure of the TCR-TAX-HLA-complex shows that the TCR with its variable V_α and V_β loops is orientated diagonally over the MHC–peptide complex. The TCR buries almost all the exposed peptide, presented by the MHC, including a large area of the MHC molecule, itself. The TCR with a constant, (C), β chain region and with a variable, (V), β chain region and variable, (V), α chains is on top. C_β and V_β are in magenta. V_α is in purple. The peptide embedded in the blue MHC molecule is indicated by the tiny black trace. (The viral nonapeptide, TAX, has the sequence LLFGYPVYV; leucine, leucine, phenylalanine, glycine, tyrosine, proline, valine, tyrosine, valine)). Below is shown the β_2 microglobulin domain in magenta, and the α_3 domain of the MHC in blue. The structure with the TCR on top and the peptide–MHC complex on the bottom, shows how a TCR on a cytotoxic T-cell could span the space between the T cell and a virally infected target cell presenting the MHC complex with viral peptides. The binding mode, observed in the TCR-TAX-HLA-A2 complex may be shared by all class I MHC molecules and all $\alpha.\beta$-TCRs. (These ribbon presentations are reproduced with permission of the authors and Nature from ref. 28 of Chapter 14.)

Plate 29

Plate 30

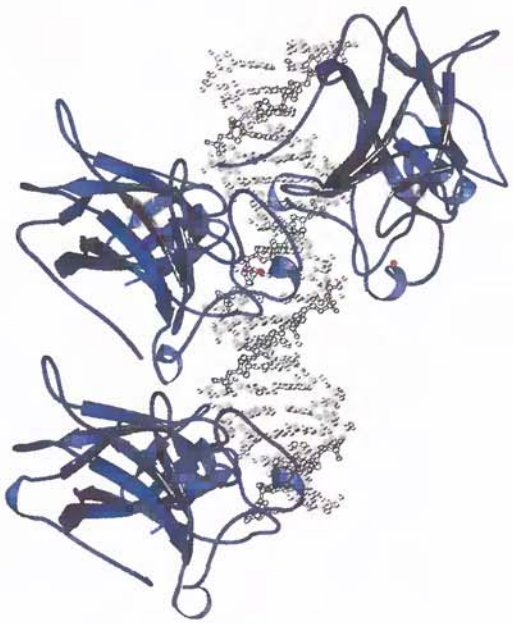


Plate 30 (*right*) A ribbon model of a p53–DNA complex. Only three subunits of the p53 tetramer are shown in this view of the structure. (Zn atoms in the loops are indicated as tiny ball and sticks, (a ball and two sticks). The loops that interact with the DNA are indicated in magenta purple. p53 binds to DNA sites that typically contain four copies of a consensus pentamer seqence. Thus each p53 monomer binds one pentamer sequence. Tetramer formation of p53 on the DNA is compatible with the structural arrangement. The key conclusion from the structure is that the majority of the cancer-causing mutations of the tumour suppressor, p53, occur in that part of the structure that is involved in DNA binding. Arginine residues, (Arg248), in the contact regions of p53 with the DNA are often mutated in tumours. Some mutations appear to eliminate critical DNA contacts, whereas others may destabilize the loops and the loop–sheet–helix motif. The loop–sheet–helix motif directs p53 to the bases in the large groove of DNA, and an additional loop makes contacts with the minor groove (see Chapter 10). (This ribbon model is reproduced with permission of the authors and Science from the data in ref. 34 of Chapter 15, which are also available in protein databanks).

The Biochemistry of Cell Signalling deals in depth with the principles of cell signalling, concentrating on structure and mechanism. It will serve as a reliable map through the maze of cell signalling pathways and help the reader understand how malfunctions in these pathways can lead to disease.

The book is divided into four parts.

Part 1 describes the machinery of signal transduction starting with the properties of signals, receptors (including receptor activation), regulators, and the molecules that link receptor and regulator. The design of signalling cascades is explained by describing central signalling pathways: the Ras-regulated MAPK and PI-3 pathways; the Rho/Rac/Cdc 42 pathway controlling chemotaxis and regulating the cytoskeleton; the G protein coupled receptor cascades in response to sensory and hormonal signals; signalling by TGF-β in morphogenesis; cytokine signalling that controls haemopoiesis. There is also a discussion of the insulin response. As phosphorylation–dephosphorylation is involved in nearly all cellular regulatory processes, Part 1 concludes with a synopsis of its role in signalling.

Part 2 describes the implementation of the signalling cascades focusing on the effect on gene transcription. After a brief description of the transcriptional machinery the regulation of transcription by cytokines and growth factors in the control of cell growth and the mechanisms and sites of control are discussed in detail. The regulators discussed include Jun/Fos, NF-AT, SMADs, and STATs. The next two chapters cover gene regulation by nuclear receptors, including both the steroid hormone receptors and non-steroid nuclear receptors e.g. the retinoic acid receptors RAR and RXR.

Part 3 studies the global cellular regulatory programs for the control of cell growth and proliferation. The first chapter concerns the regulation of the cell cycle and the role of the cyclin-dependent kinases, telomerase, Ran, and cell cycle checkpoints. The next topic is the signalling pathways in apoptosis: the TNF-receptor family death receptors, caspases, the intracellular apoptosis signals, and the role of apoptosis in the lifecycle of cells. Part 3 ends with a discussion of the signal pathways involved in the immune response, focusing on the involvement of cell–cell interactions.

Part 4 considers loss of regulatory control and its consequences with respect to the molecular basis of cancer. It first describes the cellular regulatory proteins that have oncogenic potential, how they can become oncogenic and cause the transformation of normal cells to cancerous cells. Next is an analysis of the loss of developmental controls, the APC protein, β-catenin, and the Wnt pathway, that lead to mature terminally differentiated cells reverting to immature embryonic cells. The book ends with a summary of the molecular and cellular causes of cancer and an outlook for novel therapies.

Throughout the text, the emphasis is on structure and mechanism and is well illustrated with 200 figures. *The Biochemistry of Cell Signalling* will be an invaluable companion to all graduate students studying cell signalling.

Cover image was generously contributed by Professor Louise N. Johnson, F.R.S., Director of the Laboratory of Molecular Biophysics of the University of Oxford. It shows glycogenphosphorylase with the N-terminal peptide in the active phosphorylated conformation in comparison with the inactive non-phosphorylated conformation.

OXFORD
UNIVERSITY PRESS

ISBN 0-19-850820-4
9 780198 508205

www.oup.com