Textbook of Cell Signalling in Cancer

An Educational Approach

Jacques Robert

Springer

Textbook of Cell Signalling in Cancer

Jacques Robert

Textbook of Cell Signalling in Cancer

An Educational Approach

Springer

Jacques Robert
Institut Bergonié
Bordeaux
France

Updated translation from the French language edition 'Signalisation cellulaire et cancer' by J. Robert © Springer-Verlag France, Paris, 2010; ISBN: 978-2-8178-0027-1

Additional material to this book can be downloaded from http://extras.springer.com

ISBN 978-3-319-14339-2 ISBN 978-3-319-14340-8 (eBook)
DOI 10.1007/978-3-319-14340-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2015931090

Printed on acid-free paper

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

Foreword

There is a crucial need for oncologists to know cell–cell signalling pathways, to understand their interactions, their role in cell survival, proliferation and motility, and their involvement in cancer. The practice of clinical oncology has become dependent upon the knowledge of cancer biology. Not knowing that KRAS is downstream EGFR in one of the major proliferation signalling pathway would lead the clinician to prescribe useless, toxic and expensive treatments to half of the patients suffering from metastatic colon cancer.

This book aims at presenting the scientific bases of signalling pathways to readers involved in the practice of oncology. I have not tried to hide the complexity of these pathways, of their interactions, of their connections; however, it is not necessary to be confused for being accurate, and simplification is always required when covering a large field. All along the writing of this book, I have tried not to swamp the reader in unnecessary details and tortuous paths, but to present essentially accurate and proven facts, to show their interest in clinical oncology, and to provide didactic drawings that the reader can easily reproduce by hand to better memorise the paths of information.

The main cell signalling pathways are presented in apparently independent chapters, each of them being devoted to a 'signalling system' represented by the inseparable pairing of a receptor type and its cognate ligands (although in some cases the description of downstream pathways required specific chapters). The biochemical substratum has not been avoided but I have tried to keep it unobtrusive. In order to remind the reader of the basis of molecular biology, that he has certainly learnt before but may have forgotten, the general mechanisms controlling DNA replication and repair, gene expression and protein activity have been presented in three annexes, together with the alterations of these mechanisms that contribute to oncogenesis.

Each chapter should be self-sufficient, but the crosstalks between signalling pathways are so varied and important that many cross-references from one chapter to another have been introduced. In each chapter, after the presentations of the actors of the play comes the description of the course of information within the pathway and the implementation of effector action; negative and positive regulations are then mentioned. Each chapter and annex is completed by a brief description of the oncogenic alterations of the pathway studied and by a short inventory of the pharmacological targets that may be taken in consideration for therapeutics.

Beyond the scope of this book would have been the description of the actual molecules available or in development for cancer treatment.

I have tried to use standard shape and colour symbols in all graphs, so that the reader can recognise at first sight a G-protein–coupled receptor or a small GTPase, for instance. I have priviledged the legibility of the graphs to the detriment of exhaustivity. Everyone knows the beautiful posters that contain so many arrows that it becomes impossible to discern what is crucial from what is not: such representations show the complexicity but do not bring usable information. I have tried to adopt in all figures a homogeneous colour code but this was not always possible. As a general rule, kinases are blue, phosphatases red, extracellular ligands beige, transcription factors yellow, G-proteins green, adapter proteins purple, etc., but the rainbow does not contain enough shades for all purposes!

I have limited the bibliography, at the end of each chapter, to a series of up-to-date and easily accessible general reviews. The reader who wants to deepen his or her knowledge on a specific pathway would certainly take advantage of these reviews before reading the original articles where science was made. I have not referenced every assertion or description: this would have been impossible in an educational book; a *PubMed* research '*cell signalling* AND *cancer*' retrieves more than 120,000 papers, among which 25,000 reviews. Moreover, because of the trivialisation of *PubMed*, the reader can easily find the origin of every assertion by using the adequate keywords. For instance, when I mention (p. 16): 'A t(8;9)(p11;q33) translocation of the *FGFR1* gene is responsible for the spontaneous dimerisation of the receptor in myeloproliferative syndromes', the reader can immediately find ten references on PubMed after typing '*translocation t*(*8*;*9*)(*p11*;*q33*) AND *FGFR1*': there is no need to mention any of them in this book.

There are several possibilities for the study of cell signalling in oncology. One can start from the presentation of the actors that intervene in all pathways: ligands, receptors, protein kinases, small G-proteins, transcription factors, etc., and describe their multiple functions. One can also choose a 'horizontal', topological integration level: plasma membrane, cytoplasm, mitochondria, nucleus. I have preferred a 'vertical' level of integration, defining each pathway by the type of receptor involved and describing the available corresponding data: ligands, receptor activation, transduction of the information, implementation of effectors.

Despite the complexicity and the multiple connectivity of signalling pathways, there exist homogenous receptor families, able to interpret the signals brought by ligands belonging to definite families: the book is organised around this ligand–receptor axis. This choice allows a didactic presentation of signalling modules but induces some difficulties at the level of the 'common final pathways' of cell signalling, i.e. the transcription factors, that are activated in multiple pathways and play multiple roles. It would have been possible to write a special chapter on transcription factors, but I have preferred to spread their description with that of the signalling pathway which is predominantly involved in their activation. Maybe also a special chapter of small G-proteins mechanisms of activation and function would have been useful; in fact, I have described them within the most relevant chapters.

I have identified a definite number of ligand–receptor axes, but others may exist. I have not individualised the TNF–TNFR (*tumor necrosis factor–tumor necrosis factor receptor*) superfamilies because their study was required in the chapters on apoptosis and on T-cell receptors, so that a dedicated chapter would have been redundant. Some ligand–receptor axes of lesser importance have been briefly described in connection with other pathways: the tyrosine phosphatase receptors at the end of the chapter on tyrosine kinase receptors, and the guanylyl cyclase receptors after cytoplasmic guanylyl cyclases in the chapter on nitric oxide signalling. I have not written a special chapter on cell adhesion molecules and on the control of cell adhesion and motility: this is briefly reviewed at the end of the chapter on semaphorins. Sensory signalling has not been studied (it could have been mentioned in the chapter on G-protein–coupled receptors) and I have only touched on the vast domain of nerve and muscle signalling in the chapter dedicated on ion channels–coupled receptors. I found in these pathways no obvious relationship with cancer.

I hope that this book will fulfil its role in providing essential basic information to the oncologist during his or her training. I am aware that, very soon, several chapters will be obsolete (some are probably obsolete right now) but in the meantime, this textbook may accompany the reader within the maze of the novel therapies that target oncogenic mechanisms, and enable young oncologists to situate any new proposal of the pharmaceutical industry at its right place in the jungle of cell signalling pathways that I tried not to make inextricable.

Cellular and molecular biology makes an immoderate consumption of abbreviations and acronyms and it is sometimes difficult to recognise and interpret them. Some are very simple acronyms (EGFR for *epidermal growth factor receptor*, for instance); other are abbreviations, such as AREG for *amphiregulin*. Many of these became common names for genes and proteins and the original meaning is often forgotten; who remembers (and cares) that RAS was forged from *rat sarcoma* and RAF from *rat fibrosarcoma*? Finally, some abbreviations look like acronyms but have no meaning: this is the case for AKT, for instance.

I have followed the usual practice for protein names, with sometimes a temptation toward rationalisation and simplification, eliminating hyphens as much as possible (why CD45 and IL-12?), lower case letters, etc., but this was not always possible because of the weight of history. We do not have an official nomenclature system for proteins, but we have the HUGO (*human genome organisation*) nomenclature for genes: I have used it as frequently as possible, in italics for the gene itself, in roman characters for the corresponding protein. I have tried to always mention the name of the gene when it is different from the usual name of the protein (p38α and *MAPK14*, $p16^{INK4a}$ and *CDKN2A*, for instance); when there is no special indication for the gene name, the reader can infer that it is the same as that of the protein (*TRAF6* is the gene encoding TRAF6, *ICOS* the one encoding ICOS, for instance).

A recurrent problem is that numerous proteins may belong to small subfamilies encoded by distinct homologous genes but displaying similar functions; should we write ‘the AKT protein’ or ‘AKT1, AKT2 and AKT3 proteins’? ‘the RAS protein’ or KRAS, HRAS and NRAS proteins’? I have tried, when the distinction was not

required, to use generic names (AKT, RAS). Similarly, a single gene has often multiple protein products (isoforms), resulting from alternative splicing. The individual function of each isoform is exceptionally known and I did not even mention isoform existence, with few exceptions, such as the *BCL2L1* gene products, $BCLX_L$ and $BCLX_S$.

Acknowledgements

This work originates from the courses I teach for many years at the university of Bordeaux to medicine and pharmacy students. My students will easily recognise the graphs that were first drawn for them and perfected year after year. My ambition would be to convince the best of them to enter a career in cancer research or cancer medicine and participate actively to the progresses that are required to understand and conquer what is perhaps the most fascinating challenge in medicine.

This book was first published in French in 2010 and I am indebted to my American friends, especially Dean Brenner, at the University of Michigan, for having convinced me to translate it into English. I have been helped in this matter by my wife Alice and by a colleague and friend, Jacques Bonnet: both of them are much more fluent than I am, and I am very grateful for the time they spent. This was nevertheless a hard task and I hope that native English readers will forgive me for the many stylistic defects that may have been left.

Several friends and colleagues were kind enough to read various chapters of the original edition, and I would like to thank them for their encouragements and their suggestions or corrections, especially Jacques Bonnet, Abdelkader Bounaama, Patricia de Cremoux, and Valérie Le Morvan. But I am the only one responsible for the errors or misinterpretations that have been left! I am also grateful to the members of my research group who have so kindly accepted that I spend so much of my time to this book, which has diverted me from the follow up of their research work.

The French professors of oncology welcomed me among them 25 years ago with much generosity. I have worked with them in many instances and cooperations (teaching, research, seminars, editorial boards, evaluation and recruitment committees). I hope that this book will be of help in their teaching activities. My special thanks are for François Eschwège, Michel Marty and Maurice Schneider who have generously opened for me the doors of oncology. It is now my turn to open these doors to younger oncologists, and I try to give them as much as I received from the older colleagues.

This book has been made possible thanks to the constant support of my Springer editor, Nathalie L'Horset-Poulain, and I take this opportunity to express my gratitude to her and the whole team of people who has worked on this book, both in its French version and its English one.

Contents

Introduction: General Principles of Cell Signalling From Receptors to Effectors Cell signalling and Cancer

Cell communication is indispensable for multicellular organisms: cells must necessarily exchange the information required for coordinating their activities. It also exists in unicellular organisms such as yeast cells, which also must communicate, if only to find sexual partners. The principles of cell communication are universal: signalling molecules are emitted by a given cell and recognised by another one, which in response activates a transduction pathway, leading to an effector system able to carry out the corresponding tasks. The achievement of this information transfer requires from the source cell the encoding of the message in a way that could be correctly interpreted, and from the cell receiving the message the corresponding decoding systems (Fig. 1). The variety of signal transduction systems is considerable, at the level of both signal reception and task execution. However, it is certainly possible to identify general patterns and common structures of organisation if sought for with enough patience.

At the level of signal reception, the mechanisms are roughly dependent upon the chemical structure of the messengers:

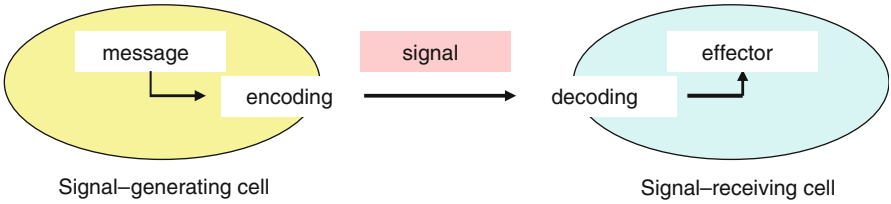


Fig. 1 From the message to the effector. A cell generating a signal must first encode the message in a way that could be correctly interpreted; the signal diffuses then in the extracellular space, sometimes the blood stream; and the cell receiving the message should perform the corresponding decoding so that the effectors can respond to the instruction

- Hydrophilic messengers (aminoacids and derivatives, peptides, proteins) cannot enter the cells since they cannot easily cross the membranes; a membrane receptor is required to receive and understand the message and to transduce the information hereafter.
- Lipophilic messengers (steroids, fatty acids and derivatives, etc.) and very simple compounds (oxygen, nitric oxide) are able to diffuse inside cell membranes and to reach directly their intracellular targets, in the cytoplasm or the nucleus.
- Ionic compounds (Na^+, K^+, Cl^- and Ca^{2+}) are able to induce the opening or closing of membrane channels allowing the generation of transmembrane currents which are used for the transmission of nervous impulses but are also responsible for various intracellular events.

The transduction of the signals received by the receptors follows many different processes but the general mechanisms are not numerous, the main ones being:

- Recruitment of adapter proteins able to interact with other proteins and induce conformational (3-dimensional) changes, and thus protein activity.
- Phosphorylation and dephosphorylation reactions, catalysed by kinases and phosphatases, respectively, which modify the protein tridimensional structure.
- Small G-protein activation, through processes of exchange and hydrolysis of guanyl nucleotides, which also induce changes in protein conformation.
- Generation of second intracellular messengers, which relay the information brought to the membrane by the extracellular first messengers.

Finally, the effectors are also diverse, but there again it is possible to group them together in a few entities:

- Transcriptional regulators, currently called *transcription factors* which command target genes transcription; these are the most general effectors and the most often used downstream the signal transduction pathways.
- Translational regulators acting on protein synthesis, which are directly involved in some signalling pathways.
- Proteins of the cytoskeleton or the extracellular matrix, which control the phenomena of cell adhesion, motility and migration.
- Ionic channels, mentioned here as effectors, involved especially (but not uniquely) in synaptic transmission.

The cell, therefore, has a ‘standard toolbox’ from which it can draw the adequate tools to understand the information received and execute the orders that are given. This may be this ‘do-it-yourself’ aspect which is disconcerting for the student who enters the field of cell signalling: he will find, in apparently distant and independent pathways, common molecules, identical protein domains, and similar transcription factors. For instance, ‘EGF motifs’ are found in the proteins of the growth factors pathway, but also in the Notch and in the integrins pathways. That does not mean that everything is everywhere, and reciprocally!

Cell signalling may operate at highly variable distances: the first signalling system that has been identified, at the end of the nineteenth century, is the endocrine system, in which a specialised organ (an endocrine gland) secretes, into the blood flow, specific molecules (hormones) aimed at distant cells and organs. When it was discovered that molecules produced in a given cell could be delivered to the neighbouring cells, was forged the word 'paracrine'; then the term 'juxtacrine' for signal transmitted between jointed cells; and eventually the term 'autocrine' when the molecules, after transiting in the extracellular space, exert their effect on the same cell. A special mention should also be given on the synaptic transmission of informations between nerve cells or between nerve and muscle cells.

Besides the intercellular signalling exists an intracellular signalling, and both are tangled up and hardly separable. This mode of communication is in particular carried out by the second messengers mentioned earlier, and also by a number of proteins, which are synthesised in response to all kinds of requests. In addition, the multiple signals received by a cell from other cells and from its environment must be integrated inside the cell, which must take into account their synergistic or divergent meanings.

For a schematic representation, the informations conveyed from one cell to another can be classified into six large types of instructions to be executed, which form opposite pairs: to proliferate or to differentiate; to remain attached or to migrate; to survive or to die. One can distinguish, as a first approximation, 'proliferation pathways', 'motility pathways' and 'survival pathways'. But, very often, the same signals can convey these different instructions: proliferation signals can become differentiation signals and survival signals can become death signals, according to the receptor equipment of the cell receiving the message.

When studying cell signalling, the linear model established by the endocrinologists of the end of the nineteenth century – a unique key for a unique lock (*one* ligand, *one* receptor, *one* message, *one* physiological action) – must be abandoned. Indeed, the same ligand can bind various receptors (which may have opposite roles); and the same receptor can receive information from different ligands; the intracellular signals transmitted after receptor activation are generally multiple and the physiological actions are variable from one tissue to another. The individual equipment of each cell or tissue in receptors and transduction systems is responsible for the ultimate behaviour of the cells which receive a message; this explains why the same message can induce cell division in a given tissue and cell differentiation in another. The effect of the signals exchanged between cells is pleiotropic and depends upon the 'cellular context'.

Cell Signalling and Cancer

We will not describe here the mechanisms involved in oncogenesis but present some characteristics of malignant cells to show how the molecular alterations occurring in cell signalling pathways constitute the bases of oncogenesis. The cancer cell is by definition genetically unstable, and is consequently able to explore the potentialities

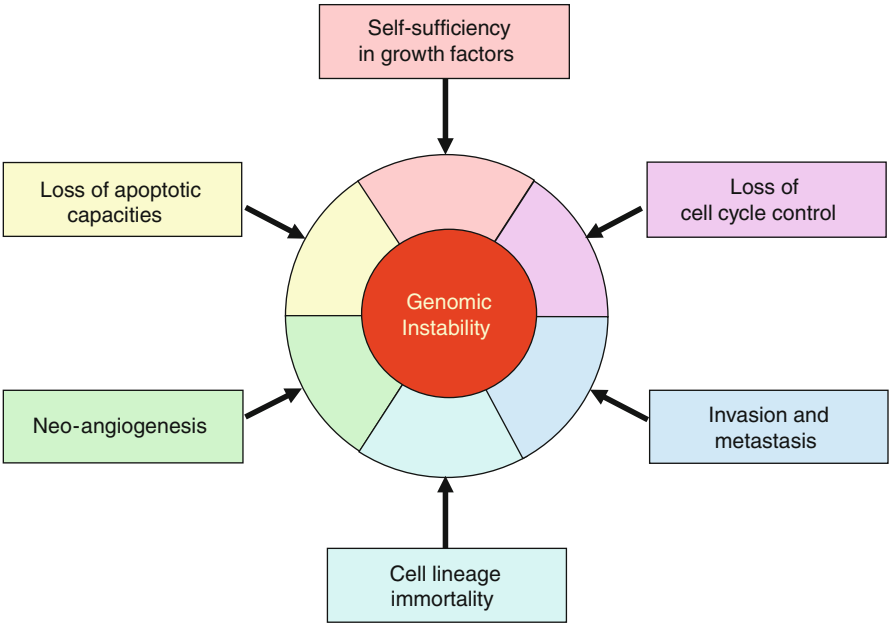


Fig. 2 The 'hallmarks' of cancer. This is an adapted reproduction of the scheme originally published by Hanahan and Weinberg in 2000. The six core families of mechanisms are represented around the cancer cell, which can pick up specific alterations in each of them. This is possible thanks to the basic property shared by all cancer cells and might well be the primum movens of all oncogenetic mechanisms, genome instability

of the whole genome, to select any proliferative or migratory advantage, and to transmit it to its descent. Proliferative, because any tumor is a neoformation requiring active cell reproductive ability; migratory, not to say invasive, because the malignancy of cancers lies in their capacity to disseminate in the organism. All signalling pathways involved in cell proliferation and differentiation, in cell adhesion and migration, in cell survival and death, may pave the way to oncogenic alterations. In that sense, cancer appears as a disease of cell signalling.

In a general review elaborated 25 years after the discovery of the first oncogene, Hanahan and Weinberg have classified the mechanisms of oncogenesis in six large families; their contribution, with some personal adaptations, is still valid 15 years later, although (Fig. 2).

- Self-sufficiency in growth signals
- Insensitivity to anti-growth signals (that I reformulate as 'loss of cell cycle control')
- Apoptosis evasion
- Limitless replicative potential
- Sustained angiogenesis
- Tissue invasion and metastasis

One can add a last general mechanism which underlies all others: genomic instability. For all mechanisms, it is always possible to detect an alteration in information transfer originating from a mutation in a protein involved in signal transduction, leading to a disease of cell signalling. An 'actualisation' of the Hanahan and Weiberg paper in 2011 does not actually bring more concepts and needlessly adds new 'basic mechanisms' that cannot be considered as such, in my view.

The cancer cell way to grow, divide and disseminate certainly constitutes the best model to study Darwinian evolution at the cellular level. It would be so simple if the cancer cell would only divert the main mechanisms that control epithelial cell proliferation and migration; but in addition, it is able to seek unsuspected pathways involved in inflammation, cell polarity, axonal migration, intercellular junctions and many other specialised processes that seemed limited to a particular tissue or to a precise step of embryonic development. To understand the hijacking of cell signalling pathways exerted by cancer cells, one should not be limited to study obvious processes, but seek for the unusual, unrecognised, even exceptional features.

The understanding of oncogenetic mechanisms and of signalling pathways alterations in cancers has permitted the emergence of the concept of targeted therapies. Whereas the drugs used in classical chemotherapy target only the effectors of cell multiplication, normal or neoplastic, this new approach aims at targeting the true mechanisms of oncogenesis and cancer development, and should have, therefore, a high degree of selectivity toward cancer cells. In targeted therapies, the targets are the proteins involved in the *control* of cell proliferation, migration and survival, and no longer those involved in the *execution* of these programmes, in other words the *cause* of the abnormality generating a cancer, and not the *effect* resulting of this abnormality. These therapies have emerged when the understanding of cancer biology has reached a sufficient level and when the tools allowing to interfere with malignant transformation have become available.

One of the first bibliographic connections that can be found between cell signalling and cancer therapy resides in the Acts of a symposium held in Cambridge in 1989 and published in the journal *Cancer Chemotherapy and Pharmacology*. It was already mentioned that '*the idea that cell membranes and intracellular signal cascades could become the target of antitumor drugs has come of age*'. The idea was however fairly new and 10 more years were required for the emergence of the first targeted therapies.

Further Reading

Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100:57–70.
Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.

Growth Factors and Tyrosine Kinase Receptors

1

Abstract

Signalling pathways originating from the interaction between growth factors (GF) and tyrosine kinase receptors (TKRs) are certainly the best known because of their leading role in oncogenesis. These pathways are multiple and will be studied in several chapters; the present chapter exclusively concerns GFs and TKRs, whereas Chaps. 2 and 3 will present two major signalling pathways downstream these interactions. We adopt here a definition of 'growth factors' restricted to those activating TKRs. Other classes of receptors will be studied in other chapters: receptors activating cytoplasmic tyrosine kinases (Chap. 4), serine/threonine kinase receptors (Chap. 5), etc.

More than 100 GF and 58 TKRs, distributed among 20 families, have been identified and a large part of them may play a role in oncogenesis. Activation of TKRs occurs according to a unique mechanism, involving dimerisation and autophosphorylation of the receptors, through their own tyrosine kinase activity, and constitutes the starting point of the signalling pathway. These dimeric associations of TKRs establish the link between GFs and cellular response; there exists a complex combinatorial pattern of these associations, which determines the type of action exerted on the cell receiving the signal. The number of possible combinations is high, which explains why the same signal may generate distinct consequences in the target cells: for instance, proliferation and differentiation may be induced by the same signal acting on different cells. Lastly, the signalling pathways downstream the GF–TKR interaction are multiple and depend upon the proteins expressed by the target cell, which is often referred to as 'the cellular context'.

1.1 General Features of the GF–TKR Interaction

1.1.1 Overview

The various GFs and TKRs can be grouped in families of homologous proteins: for instance, the GFs of the epidermal growth factor (EGF) family interact with the TKRs of the epidermal growth factor (EGFR) family, and so on. The denominations

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_1

'epidermal' or fibroblastic' growth factors have more to do with the circumstances of their discovery than with the specificity of their target: epithelial tissues may express fibroblastic growth factors (FGF) and non-epithelial tissues may express epidermal growth factors.

Growth factors are polypeptides transmitting mitogenic messages, i.e. messages able to trigger mitosis. They are secreted by numerous cell types, not organised in individualised glands, and act mostly in the close vicinity of the cell of origin. By analogy (and opposition) to the *endocrine* system, comprised of hormones circulating in blood, growth factors are said to be *paracrine* factors. They may also act as juxtacrine factors (when the cell of origin and the target cell are linked by tight junctions) and as autocrine factors (when they directly act on the cell which secreted them). Whereas the cellular response to GFs, as mediated by TKRs, has been deciphered to a high degree of molecular understanding, as presented below, the regulation of the expression and secretion of GFs by the cells of origin has not been studied in depth, and we have no comparable knowledge of the pathways involved, which would be of utmost importance in oncogenesis.

Table 1.1 presents the 20 classes of TKRs and the corresponding GFs. All TKRs possess a single transmembrane domain; a homologous intracellular domain bearing the catalytic centre, of limited structural variety; and an extracellular part which is highly variable from one class to another. This extracellular part contains various protein domains characterised by specific amino acid sequences (cysteine-rich domains, acidic domains, immunoglobulin-like domains, etc.), which are presented on Fig. 1.1. The mechanism of activation of all TKRs is the same: after GF binding occurs a step of dimerisation, which triggers the autophosphorylation of tyrosine residues located in the intracellular domain, close to the *C*-terminal part of the receptor. The presence of phosphotyrosine residues constitutes the true signal that permits the implementation of a transduction pathway leading to the activation of intracellular effectors.

These phosphotyrosine residues are recognised by proteins equipped with binding domains called 'SH2 domains' (*SRC homology domains 2*) and 'PTB domains' (*phosphotyrosine-binding domains*). These proteins are able to recognise in a very specific manner the different phosphorylated tyrosine residues of the activated receptor. More than 100 proteins bearing a SH2 domain and 35 bearing a PTB domain have been identified. Two types of signalling can be implemented:

- Cytoplasmic proteins with a SH2 or a PTB domain can be phosphorylated on tyrosine residues, thanks to the catalytic activity of the TKR, and hence display a new biological activity resulting from this phosphorylation. Some of them are cytoplasmic tyrosine kinases, such as SRC (from *sarcoma*) or JAK (*Janus kinase*, Chap. 4), tyrosine phosphatases such as SHP1 (*SH2 domain-containing phosphatase 1*) or other protein types regulated by phosphorylation, such as phospholipase C gamma (PLCγ).
- Adapter (docking) proteins, also bearing a SH2 or a PTB domain, can be recruited after TKR activation without subsequent phosphorylation but simply because they can specifically recognise and bind the phosphotyrosine residues of the activated receptor. This is, for instance, the case of GRB2 (*growth factor receptor-binding protein 2*), which is one of the main entries into the MAP kinase pathway (Chap. 2), and of p85 (gene *PIK3R1*), which is one of the main entries into the phosphoinositide 3-kinase pathway (Chap. 3).

Table 1.1 Families of tyrosine kinase growth factor receptors

Class	Family	Receptors	Ligands
Class I	EGF receptors	EGFR ERBB2 (no ligand) ERBB3 (no TK activity) ERBB4	EGF TGFα HBEGF Epigen (EPGN) Epiregulin (EREG) Betacellulin (BTC) Amphiregulin (AREG) Neuregulins 1 to 4 (NRG1 to NRG4)
Class II	Insulin receptors	INSR IGF1R [IGF2R][a] INSRR	Insulin IGF1 IGF2
Class III	PDGF receptors and related receptors	PDGFRα, PDGFRβ KIT FLT3 CSF1R	PDGFA to PDGFD SCF (KITLG) FLT3LG CSF1, IL34
Class IV	FGF receptors	FGFR1 FGFR2 FGFR3 FGFR4	FGF1 to FGF14 FGF16 to FGF23
Class V	VEGF receptors	FLT1 (VEGFR1) KDR (VEGFR2) FLT4 (VEGFR3)	VEGFA, VEGFB, PlGF VEGFA VEGFC, VEGFD
Class VI	HGF receptors	MET MST1R (RON)	HGF (SF) MST1
Class VII	NTRK receptors	NTRK1 NTRK2 NTRK3	NGF BDNF NT3, NT4
Class VIII	Ephrin receptors	EPHA1 to EPHA8 EPHB1 to EPHB6	EFNA1 to EFNA5 EFNB1 to EFNB3
Class IX	Récepteurs TAM	AXL MERTK TYRO3	GAS6 PROS1
Class X	LTK/ALK receptors	LTK ALK	PTN MDK
Class XI	Angiopoietin receptors	TIE1 TEK (TIE2)	– ANGPT1-4
Class XII	ROR receptors	ROR1 ROR2	WNT proteins
Class XIII	DDR receptors	DDR1 DDR2	Collagen
Class XIV	RET receptor	RET	GDNF

(continued)

Table 1.1 (continued)

Class	Family	Receptors	Ligands
Classes XV–XX	ROS receptor	ROS1	?
	RYK receptor	RYK	WNT proteins
	MUSK receptor	MUSK	?
	Lemur receptors	AATK, LMTK2, LMTK3	?
	PTK7 receptor	PTK7 (no TK activity)	WNT proteins
	STYK1 receptor	STYK1	?

[a]IGF2R is not a TKR

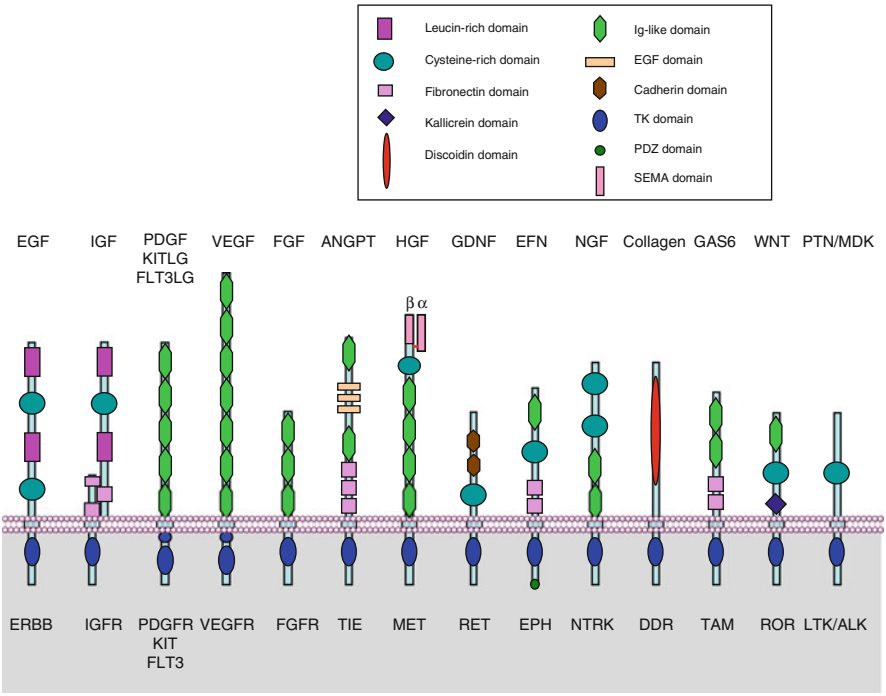


Fig. 1.1 The tyrosine kinase receptors family. Fourteen of the main TKR families are represented schematically. For each receptor, the abbreviated names are indicated below and the ligands above. The various TRKs differ essentially by their extracellular segment which contain characteristic protein domains able to carry out original functions and involved in ligand binding: leucine-rich domains, cysteine-rich domains, immunoglobulin-like domains, cadherin domains, fibronectin III-like domains, EGF-like domains, etc. The intracellular segment especially contains the catalytic tyrosine kinase domain, the *C*-terminal tyrosine residues serving as phosphorylation substrates and, in the case of ephrin receptors, a PDZ domain

1.1.2 Oncogenic Alterations

Mutations and overexpressions of growth factors certainly play an important role in oncogenesis, but the primary role of GF molecular alterations in oncogenesis has not often been demonstrated; in contrast, germline inactivating mutations of growth factor genes are responsible for many hereditary diseases such as various forms of dwarfism. In contrast, at the level of TKRs, numerous alterations have been found in cancers which play a major role in oncogenesis. The main alterations will be detailed along the presentation of the different TKR families.

Growth factors and their receptors constitute one of the major research fields to identify pharmacological targets in oncology; the development of targeted therapies has been, for a large part, based upon the research of molecules able to inhibit the proliferation signalling pathways from their first steps. The main successes obtained up to now precisely concern growth factors and tyrosine kinase receptors.

1.1.3 Pharmacological Targets

Growth factors can be targeted with molecules that are able to trap them and block their activity in the extracellular space. Soluble receptors, aptamers, antibodies and peptide constructs can be used to this goal. The major success in this respect concerns a humanised monoclonal antibody, bevacizumab, which binds and inactivates a proangiogenic growth factor, VEGFA; the antiangiogenic activity of this antibody is currently used in the treatment of various cancers in the palliative setting, especially colorectal cancers. Other antibodies directed against various growth factors are presently in development.

Tyrosine kinase receptors can be targeted either at the extracellular level, thanks to monoclonal antibodies, or at the intracellular level by small molecules able to inhibit their tyrosine kinase activity. Several receptors have been thus successfully targeted, especially the epidermal growth factor receptor (EGFR) and the vascular endothelial growth factors (VEGFRs), against which several drugs have been marketed; several other TKRs have been taken in consideration for targeting and appropriate drugs are under development. Whereas monoclonal antibodies are by definition highly specific of a given TKR, this is not so for tyrosine kinase inhibitors (TKIs), which may present a preferential binding on a given receptor but are also active on other TKRs with significant activity. Some TKIs are, therefore, multi-targeting compounds, and it may be sometimes difficult to identify the receptor whose inhibition is actually responsible for pharmacological effect. Table 1.2 presents the main TKIs that have been developed and the target receptor(s) that they inhibit.

1.2 The Paradigmatic Example of the EGF (ERBB) Family

1.2.1 Growth Factors and Their Receptors

Eleven growth factors of the EGF family have been identified: EGF itself, TGFα (*transforming growth factor α*), amphiregulin (AREG), epiregulin (EREG), HBEGF (*heparin-binding EGF*), β-cellulin (BTC), epigen or epithelial mitogen

Table 1.2 Some tyrosine kinase inhibitors (TKIs)

TKI name	Code number	Pharmaceutical laboratory	Target	Present status
Gefitinib	ZD1839	AstraZeneca	EGFR (reversible)	Phase III (Iressa®)
Erlotinib	OS-774	Roche	EGFR (reversible)	Approved (Tarceva®)
Lapatinib	GW572016	GSK	EGFR, ERBB2	Approved (Tyverb®)
Canertinib	CI-1033	Pfizer	Pan-ERBB (irreversible)	Phase II
Afatinib	BIBW-2992	Boehringer Ingelheim	EGFR, ERBB2 (reversible)	Phase III
Dacomitinib	PF-00299804	Pfizer	Pan-ERBB (irreversible)	Phase III
	HM781-36B	Hanmi	Pan-ERBB (irreversible)	Phase I
Neratinib	HKI-272	Wyeth	EGFR, ERBB2, (irreversible)	Phase II
Pelitinib	EKB-569	Wyeth	EGFR, ERBB2, (irreversible)	Phase II
Vandetanib	ZD6474	AstraZeneca	EGFR, FGFR1, VEGFR2, RET	Approved (Zactima® or Caprelsa®)
Sunitinib	SU-11248	Pfizer	VEGFR, PDGFR, KIT, FLT3, FGFR1	Approved (Sutent®)
Sorafenib	BAY-439006	Bayer	B-RAF, RET, VEGFR, FGFR1, FLT3	Approved (Nexavar®)
Vatalanib	PTK787/ ZK222584	Novartis	VEGFR, KIT	Phase III
Axitinib	AG-013736	Pfizer	VEGFR, PDGFR, KIT	ATU (Inlyta®)
Semaxanib	SU-5416	Pharmacia	KIT, FLT3, VEGFR2	Stopped
Cediranib	AZD-2171	AstraZeneca	VEGFR, PDGFR	Phase III
Motesanib	AMG-706	Amgen	VEGFR, PDGFR, RET, KIT, FLT3	Phase II
Pazopanib	GW-786034	GSK	VEGFR, PDGFR	Phase III
Imatinib	STI-571	Novartis	BCR-ABL, KIT, PDGFR	Approved (Glivec®)
Nilotinib	AMN-107	Novartis	BCR-ABL, KIT, PDGFR, RET	Approved (Tasigna®)
Dasatinib	BMS-354825	BMS	BCR-ABL, KIT, PDGFR	AMM (Sprycel®)
Masitinib		AB Science	FGFR3, PDGFR, KIT	Phase III
Lestaurtinib	CEP-701	Cephalon	FLT3, NTRK, JAK2	Phase III
Crizotinib	PF-02341066	Pfizer	ALK, MET	Approved (Xalkori®)
Regorafenib	BAY-73-4506	Bayer	VEGFR, KIT, RET	Approved (Stivarga®)

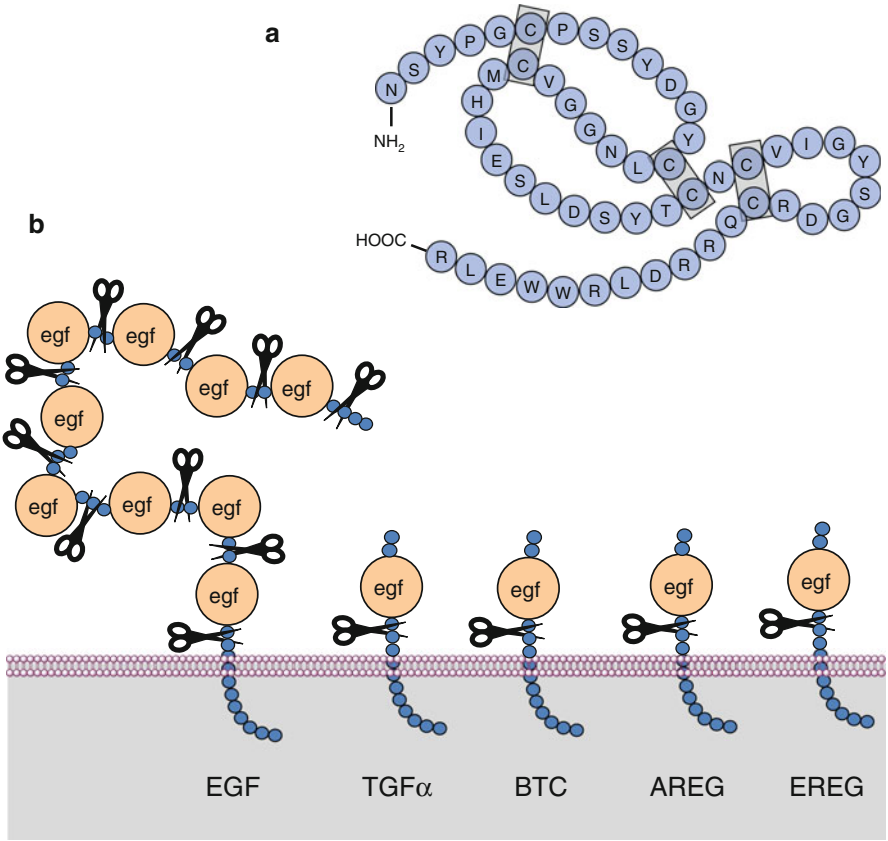


Fig. 1.2 EGF family growth factors. (**a**) Primary structure of epidermal growth factor (EGF), comprised of 53 amino acids with three intramolecular disulphide bridges. (**b**) General structure of the ligands of the EGF family. EGF itself is synthesised as a transmembrane precursor bearing 9 egf motifs which can be cleaved by a metalloproteinase such as ADAM17. The other growth factors are also synthesised as transmembrane precursors but contain only one cleavable EGF motif

(EPGN) and four neuregulins (NRG1-4). These factors play a major role in the proliferation capacity of many tissues. EGF itself is a small soluble molecule of 53 amino acids, with three intramolecular disulphide bridges (Fig. 1.2a). It derives from a voluminous transmembrane protein of 1,207 amino acids containing a series of analogous egf motifs whose release by cleavage in the extracellular space provides soluble factors. The membrane insertion of native EGF permits its action on adjoining cells, explaining thus a juxtacrine activity. The other factors of the EGF family are also synthesised as transmembrane precursors but contain only one egf extracellular motif. The cleavage of the transmembrane precursors into soluble diffusible growth factors is carried out by metalloproteinases such as TACE (*TNF-alpha-converting enzyme*) or ADAM17 (*a disintegrin and metalloproteinase*) (Fig. 1.2b).

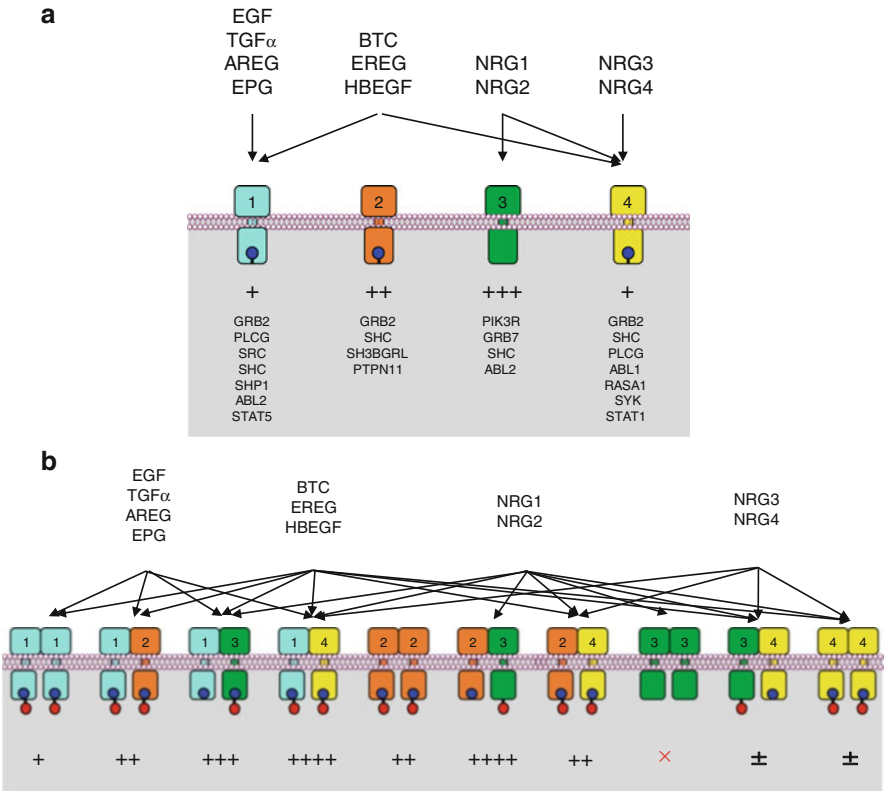


Fig. 1.3 Growth factors and receptors of the ERBB family. The 11 ligands and 4 receptors of the ERBB family are associated in a precise combinatory pattern, according to the respective affinity of each ligand for each receptor. The combinatorial pattern of ligand–receptor interactions is represented in (**a**) before dimerisation and in (**b**) after dimerisation (ten possible combinations). The resulting signal is of variable intensity according to the respective affinity of the diverse phosphotyrosine residues of the receptor for the SH2 or PTB domain-containing proteins. The ERBB2 receptor does not recognise any ligand, and its homodimerisation only occurs when it is abundant enough in the membrane, especially due to gene amplification. The ERBB3 receptor displays no kinase activity, and its homodimerisation induces no intracellular signal. The main proteins that recognise the phosphotyrosine residues via their SH2 or PTB domain are indicated below each receptor, and the quantitative importance of the interaction is represented by + signs

These 11 growth factors are recognised by 4 distinct ERBB (class I) receptors (EGFR or ERBB1, ERBB2 or HER2, ERBB3, ERBB4). There is a complex combinatorial pattern of interaction between the 11 growth factors and the 4 receptors able to recognise them and activate a signalling pathway. This interaction leads to the homo- or heterodimerisation of two molecules of receptor (Fig. 1.3). EGF, TGFα, amphiregulin and epigen only recognise and bind EGFR itself; HBEGF, epiregulin and β-cellulin recognise and bind EGFR and ERBB4; neuregulins 1 and 2 recognise and bind ERBB3 and ERBB4, and neuregulins 3 and 4 ERBB4. There are two peculiarities: the ERBB2 receptor has no ligand ('deaf' receptor)

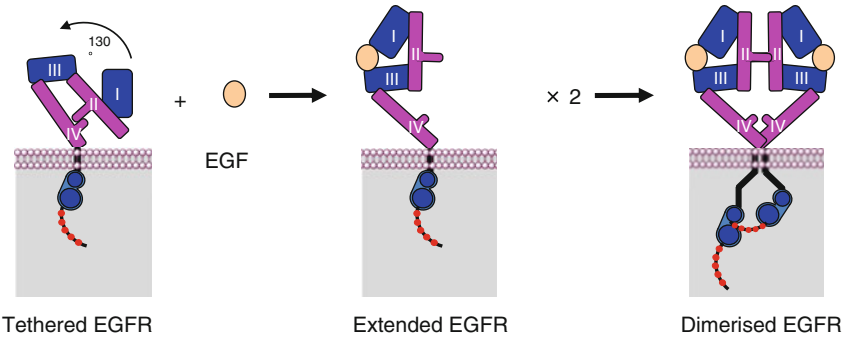


Fig. 1.4 EGFR dimerisation. In the absence of ligand, EGFR adopts a tethered conformation and cannot be dimerised; in the presence of a ligand such as EGF, the conformation of the receptor becomes extended, which unveils an affinity site at the level of the extracellular domain II. Dimerisation induces a change in the relative arrangement of the intracellular catalytic domains of the receptor, allowing autophosphorylation to occur

and cannot undergo homodimerisation in normal conditions; the ERBB3 receptor has no tyrosine kinase activity ('dumb' receptor) and its homodimers are inactive and cannot transduce a signal.

1.2.2 ERBB Receptors Activation

EGFR dimerisation is permitted by the unmasking of a binding site between two molecules of receptor, this unmasking being induced by the binding of each molecule of ligand to a molecule of receptor. At the inactive state, the receptor adopts a 'tethered' conformation, not allowing the interaction between two molecules of receptor (Fig. 1.4). The binding of EGF, at the level of the extracellular domains 1 and 3, allows the conversion to an 'extended' conformation, which unveils a binding site for another molecule of receptor having identically bound a molecule of ligand. The ERBB2 receptor has no ligand because it is always under extended conformation and, therefore, can be dimerised with other receptors without prior activation.

The dimerisation step allows the contact between two intracellular catalytic domains. Their interaction induces the opening of the active centre of the enzyme and allows ATP to access this centre, thus called the *ATP pocket*; the catalytic tyrosine kinase activity is then operative. The orientation of the two intracellular domains is asymmetric, one of the catalytic domains being able to phosphorylate the tyrosine residues of the *C*-terminal end of the second one (Fig. 1.5). The phosphotyrosine residues may then be recognised and bound by a great variety of proteins bearing an SH2 (*SRC homology-2*) or a PTB (*phosphotyrosine-binding*) domain. The two types of signalling described above can be then implemented, firstly through the activation of kinases or phosphatases such as SRC or SHP1, which are phosphorylated and activated by the tyrosine kinase activity of the receptor, and secondly through adapter proteins such as GRB2 or p85α (gene *PIK3R1*). The receptor phosphotyrosine

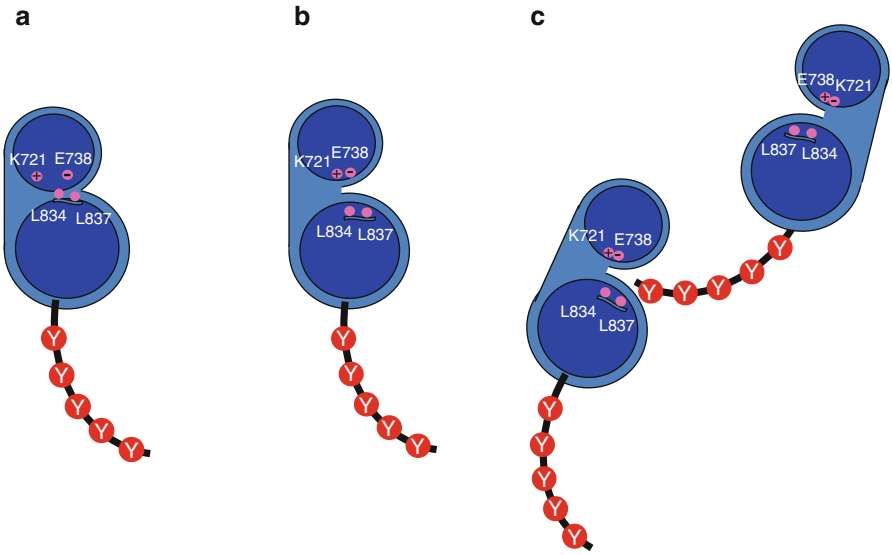


Fig. 1.5 EGFR autophosphorylation. (**a**) Before receptor activation, the catalytic site is not accessible to ATP because of the presence of the aliphatic amino acid residues Leu^{833} and Leu^{838}. (**b**) Upon receptor activation, the catalytic site is open to ATP and ATP can interact with charged amino acid residues (Lys^{716} and Glu^{734}). (**c**) The phosphorylation of the *C*-terminal tyrosine residues (Y) of one of the EGFR monomers is carried out by the catalytic activity of the second one

residues that are recognised by each SH2- or PTB-domain-containing protein have been identified: for instance, SRC recognises and binds with high affinity the EGFR phosphotyrosine residue at position Tyr^{974}, SHP1 the residue Tyr^{1173}, GRB2 the residues Tyr^{1148} and Tyr^{1173}, etc. In addition, the various ligands that activate EGFR induce the phosphorylation of different tyrosine residues: as a result, TGFα or AREG may induce specific effects, distinct from those induced by EGF.

The various ERBB receptors are activated by their ligands and activate in turn the downstream effectors according to the same general mechanisms, involving homo- or heterodimerisation and autophosphorylation. In addition to the various combinations already mentioned (homodimers and heterodimers) and the various growth factors that induce their activation, Fig. 1.3 presents the main proteins that are activated by each of the receptors and the intensity of the resulting signal. Since proteins with SH2 or PTB domains have variable affinities for the different phosphotyrosine residues of the receptors, the nature and the intensity of the signal vary according to the dimer that has been formed, the tyrosine residues that have been phosphorylated and the SH2- or PTB-domain-containing proteins available in the cell.

ERBB receptors are internalised by endocytosis and are subsequently either recycled or deactivated by transfer into the lysosome. Ligand binding accelerates internalisation and decreases thus the availability of the receptor at the surface of the cell. Internalisation occurs within clathrin-containing vesicles and can be followed, after transfer into endosomes, by ubiquitinylation by an E3 ubiquitin ligase (see Annex C) called CBL (*Casitas B-lineage lymphoma*), which displays a binding domain for phosphotyrosine residues.

1.2.3 Oncogenic Alterations

The amplification of the *EGFR* and *ERBB2* genes are oncogenic events; for *EGFR*, this amplification occurs in several types of epithelial cancers, but its importance in oncogenesis is still discussed; for *ERBB2*, an amplification occurs in 15–25 % of breast cancers and constitutes a characteristic oncogenic event, defining the ERBB2-positive cancers and allowing specific targeted therapies. *ERBB2* amplification also occurs in gastric cancers and some other malignancies. The mechanism of gene amplification remains incompletely understood.

Two types of oncogenic mutations can be found in the *EGFR* gene. The first one concerns the extracellular part of the protein, a part of which is deleted, leading to a permanent activation of the receptor in the absence of ligand. The vIII mutation is the most frequent one (with a large deletion of amino acids 24–297) and is found especially in glioblastoma. The second type concerns the intracellular catalytic site of the receptor (tyrosine kinase activity). These mutations have been identified essentially in a small proportion (5–10 %) of non-small-cell lung cancers and allow the permanent accessibility of the catalytic site to ATP: they permit the autophosphorylation of the tyrosine residues in the absence of a mitogenic signal and of receptor dimerisation. These mutations are found on exon 18 (G719X, 5 % of activating mutations), exon 19 (various deletions in codons 747–751, 45 %), exon 20 (rare insertions) and exon 21 (L858R, 50 %, the most frequent one). Since they maintain the ATP-binding domain open, they are also responsible for the sensitivity of tumour cells to inhibitors of the tyrosine kinase activity of the receptor, being thus simultaneously *oncogenic* and *drug-sensitising* mutations.

These mutations must be clearly distinguished from other mutations, also occurring close to the catalytic domain of EGFR, which conversely block the access of the TKI to the active site and lead to resistance to these drugs; the most frequent of these mutations is the T790G mutation.

1.2.4 Pharmacological Targets

EGFR can be recognised and blocked by monoclonal antibodies; two of them have been marketed: cetuximab (chimeric) and panitumumab (human); others are in development. The ERBB2 receptor can be also recognised and blocked by monoclonal antibodies; trastuzumab has proven for 10 years an important activity in *ERBB2*-amplifying breast tumours, both in metastatic and adjuvant situations. Pertuzumab has been marketed more recently; since it does not bind the same epitopes as trastuzumab (and therefore may not target the same ERBB complexes), it displays a synergistic action with trastuzumab.

Several inhibitors of the tyrosine kinase activity (TKI) of EGFR have been developed. Gefitinib and erlotinib were the first ones; they have been marketed for the treatment of *EGFR*-mutated non-small-cell lung cancers. Some TKIs, such as lapatinib, display activity on both EGFR and ERBB2 and some others, such as dacomitinib, on all ERBB receptors. Afatinib and neratinib are characterised by an irreversible covalent binding to the EGFR and ERBB2 receptors and are expected to

display a higher activity than the classical TKIs. Other TKIs such as vandetanib also have a marked activity of VEGF receptors, and it may be difficult to decide whether their activity is due to ERBB receptor inhibition or to angiogenesis inhibition.

1.3 Other Families of Growth Factors and Growth Factor Receptors

1.3.1 Platelet-Derived Growth Factors (PDGF) and Related Growth Factors

Four growth factors named PDGFA to PDGFD have been identified. They are active as homodimers (AA, BB, CC and DD) or heterodimers (AB). The association of the monomers is achieved through disulphide bridges. They are recognised by two receptors, PDGFRA or α and PDGFRB or β, which in turn form homodimers αα and ββ or heterodimers αβ, thanks to the fact that their ligands are already dimerised, allowing thus the association of two receptor molecules. The combinatory pattern is not completely elucidated; it is generally thought that ligand dimers PDGFAA and CC activate the formation of receptor dimers PDGFRαα and ligand dimers PDGFBB (and possibly DD) the formation of all types of receptor dimers (αα, ββ and αβ). On the other hand, the heterodimeric ligand PDGFAB would induce only the formation of the heterodimeric receptor PDGFRαβ and possibly that of the homodimeric receptor αα (Fig. 1.6c). The messages that are transmitted depend on the equipment of the target cells in every variety of receptor.

Because of structural relationships, this family contains also four other growth factors, each of them able to bind one cognate receptor: SCF (*stem cell factor*), also known as KITLG (*KIT ligand*), whose receptor is KIT or SCFR; FLT3LG (*FMS-like tyrosine kinase-3 ligand*) whose receptor is FLT3; and CSF1 (*colony-stimulating factor 1*, also known as *macrophage colony-stimulating factor*, MCSF) and interleukin-34 (IL34), which share a receptor called CSF1R.

PDGFB is the homologue of the murine oncogene *Sis*; oncogenic mutations of this gene have been identified in meningioma and a reciprocal translocation (t22;7) in a rare form of skin sarcoma. The *PDGFRA* and *PDGFRB* genes are altered in several cancer types: myeloproliferative syndromes, mastocytosis and gastrointestinal stromal tumours (GIST). These alterations may be translocations leading to fusion genes or activating mutations at the level of exons 12 (juxtamembrane region), 14 (kinase domain) or 18 (activating loop) of the receptors.

The *KIT* receptor gene can be mutated also at the level of several exons (11, 13, 17) and these mutations are the most characteristic alterations encountered in GISTs. *KIT* amplification is also found in these tumours. The *FLT3* gene is mutated in one third of acute myeloid leukaemia, especially at the level of the juxtamembrane domain, which plays a negative regulatory role in receptor activity, and at the level of the kinase domain (amino acids Asp^{835} and Ile^{836}).

The structural relationship between the TKRs of the VEGFR and PDGFR families explains why most TKIs display cross activity on both types of receptors, so that the antitumour and the antiangiogenic activities cannot be easily distinguished. Imatinib, originally developed for the treatment of chronic myeloid leukaemia because of its inhibiting activity of the cytoplasmic chimeric tyrosine kinase BCR-ABL, has also a marked activity on the KIT receptor, so that it is primarily prescribed in GIST treatment. In this indication, other TKIs such as nilotinib and dasatinib have been developed, with different properties according to the precise mutations of the *KIT* gene (either oncogenic and drug-sensitising or leading to imatinib resistance). The elaboration of treatment strategies aimed at optimising the prescription of the various available TKIs is an important challenge to treat this disease.

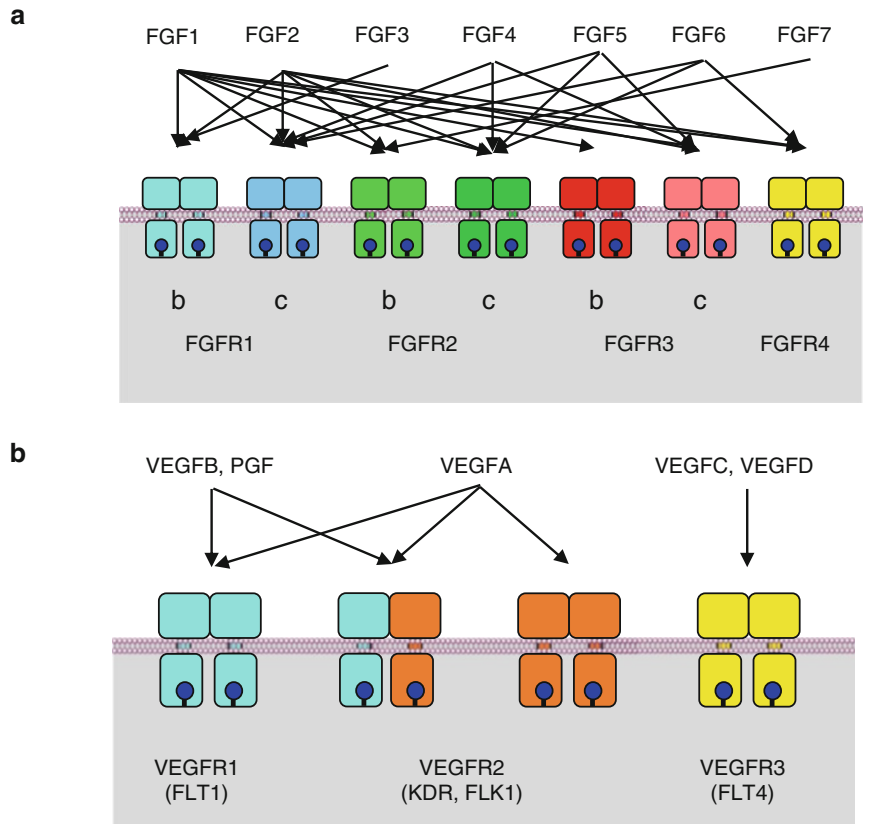


Fig. 1.6 Combinatorial pattern of some growth factors and TKRs. The combinatorial pattern of the interaction between growth factors and their cognate receptors is presented for some members of the FGF family (**a**) and for the members of VEGF (**b**), PDGF (**c**) and IGF (**d**) families

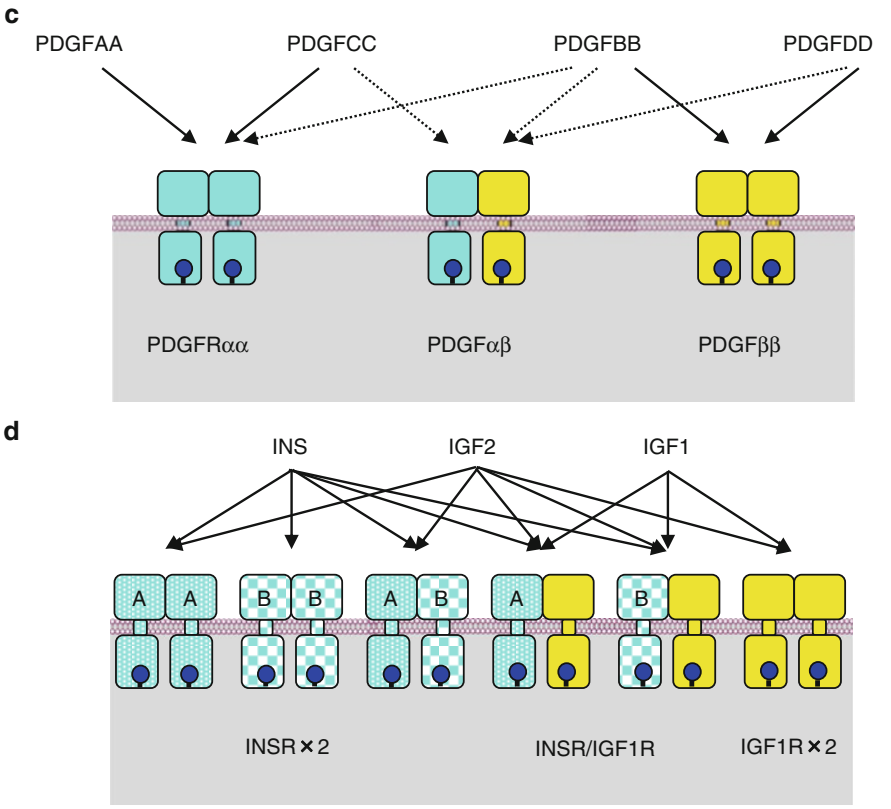


Fig. 1.6 (continued)

1.3.2 Insulin-Like Growth Factors (IGF) and Their Receptors

Insulin is a well-known pancreatic hormone with hypoglycaemic properties and shares several properties with growth factors, especially the fact that it recognises and binds a TKR. Analogues of insulin, called *insulin-like growth factors* (IGF1 and IGF2) or somatomedins, are produced essentially by the liver and also work through the activation of a TRK. The insulin receptor (INSR) exists under two tissue-specific isoforms, A and B, due to an alternative splicing. Only insulin can induce the homodimerisation of INSR-B, expressed only in specific tissues, which leads to the classical effects of this hormone on carbohydrate metabolism. Insulin and IGF2 can activate other INSR dimerisation processes, involving the A isoform, expressed in a variety of epithelial tissues. IGF1 and IGF2 can induce the homodimerisation of another ubiquitous receptor, IGF1R, as well as the heterodimerisation of INSR (A or B) with IGF1R, which is also obtained by insulin itself (Fig. 1.6d). INSR and IGF1R display an original structure among TKRs, with two polypeptide chains attached through disulphide bonds, one chain being only extracellular and the other mostly intracellular. There also exists a false receptor, IGF2R, which can bind IGF2

without being able to transmit a signal and behaves, as a consequence, as an inhibitor of this signalling pathway.

The activation of the target proteins occurs through the preliminary phosphorylation of docking proteins, the insulin receptor substrates (IRS), which are subsequently recognised by the SH2 domain of the regulatory subunit of PI3 kinase; the PI3 kinase pathway is the main pathway activated downstream insulin and IGF binding to their cognate receptors (Chap. 3).

Concerning oncogenesis, it has been suggested that high circulating levels of IGF1 could be associated to the risk of cancer. INSR (isoform A) and IGF1R are frequently overexpressed in cancer cells, mostly in the absence of gene amplification; no oncogenic mutations have yet been identified. IGF2R would rather play a tumour-suppressor function, through binding to IGF2 without signal generation, although no loss-of-function mutation has been identified.

Monoclonal antibodies against the growth factors IGF1 and IGF2 are in preclinical development; in addition, several monoclonal antibodies against IGF1R, such as figitumumab or dalotuzumab, have entered clinical trials. Tyrosine kinase inhibitors active against IGF1R have also been identified, such as linsitinib; targeting the receptor with monoclonal antibodies appears more promising than with TKIs since they have a high specificity toward the receptor, whereas TKIs may inhibit the INSR with possible unwanted metabolic side effects.

1.3.3 Fibroblastic Growth Factors (FGF) and Their Receptors

There exist 18 growth factors of this family, numbered after FGF1 (aFGF, acidic FGF or FGFα) and FGF2 (bFGF, basic FGF or FGFβ). FGFs play multiple roles in the development of numerous tissue types; in particular, FGF2 is an angiogenic factor. FGFs are recognised by 4 distinct receptors, FGFR1 to FGFR4, with the exception of a group related to FGFs named FHF (*FGF homologous factors*), which have no receptor-binding site. FGF activity is often dependent upon the presence of a membrane glycosaminoglycan, heparan sulphate, which plays a coreceptor role for the FGFs. FGFR1, FGFR2 and FGFR3 each exist under two isoforms, b and c, resulting from alternative splicing and expressed either in mesenchymal (conjunctive) tissues or in epithelial tissues, with different affinities to their substrates.

As for EGFR, FGF receptor dimerisation is required for activation. However, it seems that another mode of dimerisation would operate; it would involve firstly the existence of two binding sites for the growth factor, one with high and the other with low affinity, and secondly the existence of a binding site of the receptor with heparan sulphate. The combinatorial pattern of FGFs with FGFRs is highly complex and Fig. 1.6a only tries to illustrate this complexity for the seven first FGFs. Finally, the activation of the effector substrates of the tyrosine kinase activity of FGFRs often requires the tyrosine phosphorylation of intermediate docking proteins named FRS (*FGF receptor substrates*).

Concerning oncogenesis, *FGF3* is the homologue of the murine oncogene *Int2* and is amplified in several types of human tumours. Oncogenic alterations of

FGFRs have been also described in human tumours, essentially through activating mutations of the tyrosine kinase domain of these receptors: this is the case for *FGFR1* in glioblastoma, for *FGFR2* in endometrial cancers, for *FGFR3* in multiple myeloma and bladder cancers and for *FGFR4* in rhabdomyosarcoma. A t(8;9)(p11;q33) translocation of the *FGFR1* gene is responsible for the spontaneous dimerisation of the receptor in myeloproliferative syndromes. *FGFR1* overexpression is an important oncogenic event found in a majority of prostate cancer cells, and *FGFR1* amplification is found in hormone-resistant breast cancers and in non-small-cell lung cancers.

Among the TKIs, few molecules exclusively target FGFRs; for instance, dovitinib also targets the VEGF receptors. Antibodies against FGFR3 and protein traps against FGFs are developed for the treatment of cancers presenting an oncogenic activation of the FGF–FGFR pathways.

1.3.4 Hepatocyte Growth Factor (HGF) and the MET Receptor

A tyrosine kinase receptor named MET (*mesenchymal-epithelial transition factor receptor tyrosine kinase*) is activated by HGF, also known as *scatter factor* (SF). MET is comprised of two chains, α and β, associated by a disulphide bridge, but it is synthesised as a unique polypeptide chain that is later cleaved by a proprotein convertase called *furin*. MET contains a SEMA domain analogous to those found in semaphorins (Chap. 11) and four immunoglobulin-like domains. HGF is also an αβ heterodimer resulting from endopeptidic cleavage of a precursor by a serine proteinase, *matriptase*; it contains two MET-binding domains.

MET activation is a frequent oncogenic event in several cancer types and plays a major role in metastasis initiation through epithelial-to-mesenchymal transition (EMT). This activation can be obtained by *HGF* or *MET* gene overexpression, *MET* gene rearrangements and mutations, and generation of autocrine stimulation of cancer cells. A related receptor is MST1R (*macrophage stimulating-1 receptor*) also known as RON (*récepteur d'origine nantaise*), and its ligand is MST1 (*macrophage stimulating-1*) or HGFL (*hepatocyte growth factor-like*).

Targeting the MET pathway can be obtained with monoclonal antibodies directed against HGF or MET and with tyrosine kinase inhibitors with variable specificity toward MET; tivantinib has a good specificity against MET whereas cabozantinib is also active on the VEGF and TIE receptors and crizotinib is also active on the ALK receptor.

1.3.5 Glial Cell Line-Derived Neurotrophic Factors and the RET Receptor

The RET (*rearranged during transfection*) receptor is involved in the development of the neural crest. It is structurally related to cadherins (Chap. 11). It is activated by a protein complex consisting of GFL (*GDNF family ligand*) and a coreceptor, GFRα

(*GDNF family receptors alpha*). Germinal mutations of RET are associated to type II multiple endocrine neoplasia, and somatic mutations are found in medullary thyroid carcinomas. Several TKIs have shown activity against RET and vandetanib has been marketed for the treatment of medullary thyroid cancer.

1.3.6 Anaplastic Lymphoma Kinase and Leukocyte Receptor Tyrosine Kinase Receptors

These receptors, ALK (*anaplastic lymphoma kinase*) and LTK (*leukocyte receptor tyrosine kinase*), were first identified in lymphoma, especially ALK, whose gene is rearranged in anaplastic large T-cell lymphomas (t2;5 translocation, leading to the chimeric protein NPM-ALK). It is also oncogenic in a small subgroup (around 5 %) of non-small-cell lung tumours through gene rearrangement with *EML4* (*echinoderm microtubule-associated protein-like 4*). The ligands of these receptors are small proteoglycans bound to heparan sulphate, pleiotrophin (PTN) and midkin (MDK).

Several efficient tyrosine kinase inhibitors directed against ALK have been developed and marketed: crizotinib and ceritinib are available for the treatment of the small-cell lung tumours which bear ALK rearrangements.

1.3.7 Vascular Endothelial Growth Factors and Their Receptors

Four VEGFs, named VEGFA to VEGFD, have been identified in addition to a closely related placental growth factor, PGF or PlGF. They are secreted as homodimers, the monomers being associated by two disulphide bridges. There exist three cognate tyrosine kinase receptors, VEGFR1 (gene *FLT1*, *FMS-like tyrosine kinase 1*), VEGFR2 (gene *KDR*, *kinase insert domain protein receptor 2*) and VEGFR3 (gene *FLT4*, *FMS-like tyrosine kinase 4*). These growth factors stimulate mainly the proliferation and maturation of endothelial cells, which are normal, non-cancerous cells, contributing to oncogenesis through the formation of the blood vessels required for tumour nutrition.

VEGFA recognises and binds VEGFR1 and VEGFR2, VEGFB and PGF recognise and bind only VEGFR1, whereas VEGFC and D recognise and bind VEGFR3 (Fig. 1.6b). Neuropilins 1 and 2 (NRP1 and 2) are coreceptors to VEGFA, able to bind the growth factor without transmitting a signal. They are associated to the VEGFR1 and 2 receptors and appear to be required for their activity (see Chap. 11). VEGFA is the key regulator of arterial and venous angiogenesis, with VEGFR2 as the principal mediator of its proangiogenic activity, while VEGFR3, activated by VEGFC and D, is the mediator for lymphangiogenesis. VEGFB and PGF are not indispensable to angiogenesis; they are thought to amplify the angiogenic signal brought by VEGFA by blocking its binding site on VEGFR1, which has only low tyrosine kinase activity and cannot contribute to angiogenesis. Due to the low expression of VEGFR1, homodimers have less chance to occur than VEGFR1–VEGFR2

heterodimers, which have limited action on endothelial cells' proliferation and have essentially a role in angiogenesis modulation. In addition, a soluble form of VEGFR1, without tyrosine kinase activity, can sequester the VEGFs and plays a role in angiogenesis downregulation.

Several molecular forms of VEGFA have been identified, resulting from alternative splicing (Annex B). The principal one, $VEGF_{165}$, lacks the exon 6 product. The $VEGFA_{121}$ isoform, lacking the domain encoded by exons 6 and 7, has a very low affinity to glycosaminoglycans and has, consequently, a much higher diffusibility in the extracellular space, whereas the $VEGFA_{186}$ and $VEGFA_{201}$ isoforms keep high affinity for these compounds and remain at the surface of the cells. The affinity of these isoforms for neuropilins also differs, $VEGFA_{121}$, for instance, having no binding site for these coreceptors.

VEGF production by tumour cells, which is stimulated by various signals such as hypoxia and inflammation, is responsible for the neo-angiogenesis required by tumours to grow beyond a few millimetres in diameter. The endothelial cell does not participate itself to the cancer phenotype, and because of its genetic stability, the VEGFRs present in its membrane do not display activating mutations as other tumour cell receptors such as EGFR do.

An anti-VEGFA monoclonal antibody, bevacizumab, has been marketed, and its antiangiogenic activity is being used in the treatment of colorectal cancers and other malignancies. Another approach to target this growth factor utilises a pseudo-antibody, VEGF-trap or aflibercept, whose Fab portion has been replaced by the sequences of VEGFR1 and VEGFR2 involved in binding to VEGFA, VEGFB and PGF; this results in a high affinity of this trap to all three growth factors. Antibodies against VEGFR2 are also of interest and ramucirumab has been recently approved for the treatment of gastric cancers.

Numerous tyrosine kinase inhibitors have been developed, with variable specificity for the VEGFRs and generally cross inhibition of other receptors, especially of the PDGFR family (see above). It is not known whether the multiple targeting displayed by these TKIs is an advantage or a drawback for cancer treatment. Sunitinib and sorafenib, which are approved mainly for the treatment of kidney cancer and hepatocarcinoma, have especially low specificity to VEGFRs. Newer compounds display better specificity and can better be considered as antiangiogenic drugs: this is the case of axitinib, vatalanib, pazopanib and regorafenib (Table 1.2), which show cross-reactivity, however, with the KIT and PDGFR receptors.

1.3.8 Angiopoietins (ANG) and the TIE Receptors

Angiopoietins are, as VEGFs, endothelial growth factors and play a role in angiogenesis, explaining their potential interest in oncology. As for VEGF genes, angiopoietin gene transcription is activated, among other factors, by hypoxia, via the activation of the HIF transcription factors (Chap. 16). There exist two angiopoietin receptors, TIE1 (*tyrosine kinase with immunoglobulin-like and EGF-like domains 1*) and TIE2, also known as TEK (*tyrosine kinase, endothelial*), and three ligands

(angiopoietins 1, 2 and 4 or ANG1, 2 and 4 [genes *ANGPT1, 2* and *4*]). In addition, one should mention seven related proteins, the angiopoietin-like factors (ANGPTL 1 to 7), having either a positive or a negative effect on angiogenesis, but whose receptor remains unknown.

The TIE2 receptor, which is present only in endothelial cells and stimulates cell growth, is recognised, bound and homodimerised by ANG1, ANG2 and ANG4, whereas the TIE1 receptor, which is inactive as a homodimer, is able to block TIE2 (and the angiopoietin-induced proliferation signal) after ANG2-induced TIE1–TIE2 heterodimerisation. Thus, ANG1 and ANG4 appear as agonists of endothelial cell proliferation, while ANG2 is either agonist or antagonist according to the type of receptor expressed by the cell (TEK only or both receptors).

Several pharmacological approaches have been developed to inhibit angiogenesis through the angiopoietin pathway; the interaction between ANG1 or ANG2 and the TIE2 receptor may be inhibited by a chimeric peptibody, AMG-386 (trebananib), obtained by fusion of the Fc portion of a monoclonal antibody and the peptide structure of the ANG2-binding domain of TIE2. Several anti-ANG2 monoclonal antibodies have been selected for further research, as well as tyrosine kinase inhibitors of variable specificity. One of them, regorafenib, now marketed, appears as active on both VEGF receptors and TIE2.

1.3.9 Ephrins and Their Receptors

The ephrins (EFN) and their receptors (EPH) constitute the most important group in the superfamily of TKRs. They are separated in two subfamilies, A and B, EFNAs (five members) recognising and binding EPHAs (9 members), whereas EFNBs (3 members) recognise and bind EFNBs (5 members). The combinatorial pattern of EFN–EPH associations has not been completely deciphered. Ephrins A are associated to the plasma membrane of the producing cell through a glycosylphosphatidylinositol anchor, whereas ephrins B have a true transmembrane helix domain. In both cases, the interaction between ephrins and their receptors requires cell–cell contact (juxtacrine signalling).

Ephrins are involved in cell migration, axon guidance and vascular development. Their intracellular activities are mediated in particular by the cytoplasmic tyrosine kinase ABL and by the G-protein exchange factors active on G-proteins of the RHO family. They positively or negatively regulate cell proliferation and migration, as a function of the events occurring after receptor activation, which depend on the SH2 domain-containing proteins recruited by the activated receptor.

Alterations in ephrin receptor expression are found in several cancer types, and they are involved in oncogenesis, either as tumour promoters (overexpression) or suppressors (loss of expression). Mutations in the EPH genes have also been described. As they are involved in angiogenesis, they may represent relevant targets in oncology. Antibodies against EPHA2 and against EPHB4 are under development for cancer therapy.

1.3.10 Other Growth Factor: Growth Factor Receptor Couples

The other TKRs are less known than the one presented above; their role in oncogenesis is less obvious and their therapeutic targeting is still in infancy.

Neurotrophic Tyrosine Kinase Receptors or Tropomyosin Receptor Kinases

Three receptors (NTRK1, 2 and 3 or TRKA, B and C) and four ligands (NGF [*nerve growth factor*], BDNF [*brain-derived neurotrophic factor*], NT [*neurotrophin*] 3 and 4) are essentially expressed in neural tissue. NGF is also recognised by a lower-affinity receptor of the tumour necrosis factor receptor (TNFR) superfamily, $p75^{NTR}$ (gene *TNFRSF16*), which certainly plays a modulating role in NGF signalling through its high-affinity TK receptor. Germline mutations of NTRKs are observed in several types of hereditary neurological diseases and mental deficiency.

NTRK1 and NTRK3 are overexpressed in good-prognosis neuroblastomas and disappear at advanced stages. Conversely, NTRK2 and its ligand, BDNF, are overexpressed in high-grade neuroblastomas, concomitantly with *MYC* gene amplification. Rearrangements of the *NTRK* genes have been observed in papillary thyroid carcinomas and some childhood sarcomas. *NTRK* overexpression has been also noticed in breast and prostate cancers.

Lestaurtinib, a TKI originally selected for FLT3 inhibition, is also an inhibitor of the NTRKs and blocks their activation; it is currently in clinical trials against neuroblastomas.

TYRO3, AXL AND MERTK Receptors

This subfamily of TKRs called 'TAM receptors' is comprised of three receptors, AXL (from *anexelekto*, which means 'not controlled' in Greek), MERTK (*monocytes, epithelial and reproductive tissues tyrosine kinase*) and TYRO3 (*tyrosine kinase 3*), and two ligands, GAS6 (*growth arrest-specific gene 6*) and PROS1 (*protein S*).

The genes of these receptors have been initially cloned from leukaemia cells and considered as proto-oncogenes: their level of expression appears to be related to the malignancy of solid tumours, especially glioblastomas, but gene mutations or rearrangements have never been found in tumours. Monoclonal antibodies and TKIs are in development for targeting these receptors.

Discoidin Domain Receptor Receptors

Discoidin domain receptors, DDR1 and DDR2, have a unique extracellular domain, homologous to *Dictyostelium discoideum* discoidin. They are in fact collagen receptors, similarly to integrins (Chap. 10), and play a role in epithelial cell adhesion to the extracellular matrix, which explains their designation as 'cell adhesion kinases' (CAK). They are overexpressed in several carcinoma types and are involved in tumour cell invasivity.

Tyrosine Kinase-Like Orphan Receptors

These two receptors, ROR1 and ROR2, have often been considered as 'orphans', i.e. without known ligand. They recognise in fact the WNT proteins (Chap. 7) and act as annex receptors for non-canonical pathways. They play a role in cell motility and polarity and in skeletal and neuronal development. *ROR1* is overexpressed in acute lymphoblastic leukaemia, and mutations have been found in dysmorphic hereditary diseases.

ROS Receptor

This receptor is the product of the murine *Ros* oncogene and has no known ligand. Gene rearrangements of *ROS* have been observed in glioblastomas and lung cancers, and gene overexpression has been found in several types of solid tumours.

PTK7 Receptor

This receptor is devoid of tyrosine kinase activity and might have a regulatory function in WNT proteins signal transmission (Chap. 7). Overexpression of this gene has been noticed in colorectal cancers.

Muscle, Skeletal, Receptor Tyrosine Kinase Receptor

This receptor, MUSK, is located at the level of neuromuscular synapses. Germline mutations are observed in myasthenic syndromes.

Apoptosis-Associated Tyrosine Kinase Receptor and Lemur Tyrosine Kinases

The apoptosis-associated tyrosine kinase receptor (AATK) and the lemur tyrosine kinase receptors (LMTK2 and 3) are three poorly known tyrosine kinase receptors are involved in neuronal differentiation.

Receptor-Like Tyrosine Kinase

As the ROR receptors, RYK is an atypical receptor is involved in WNT signalling (Chap. 7); it binds and activates the WNT proteins and participates in the regulation of axon guidance.

Serine–Threonine–Tyrosine Kinase Receptor

This receptor, STYK1, remains very poorly known.

1.4 Tyrosine Phosphatase Receptors

In comparison to tyrosine kinase receptors, tyrosine phosphatase receptors (PTPR or RPTP) appear as forgotten, or at least neglected, signalling molecules. This family of 21 proteins with tyrosine phosphatase activity display receptor functions and mediates a variety of cellular processes related to cell proliferation and differentiation. PTPRs contain an extracellular segment with varied interaction domains, especially immunoglobulin-like and fibronectin III-like domains (Fig. 1.7), a single

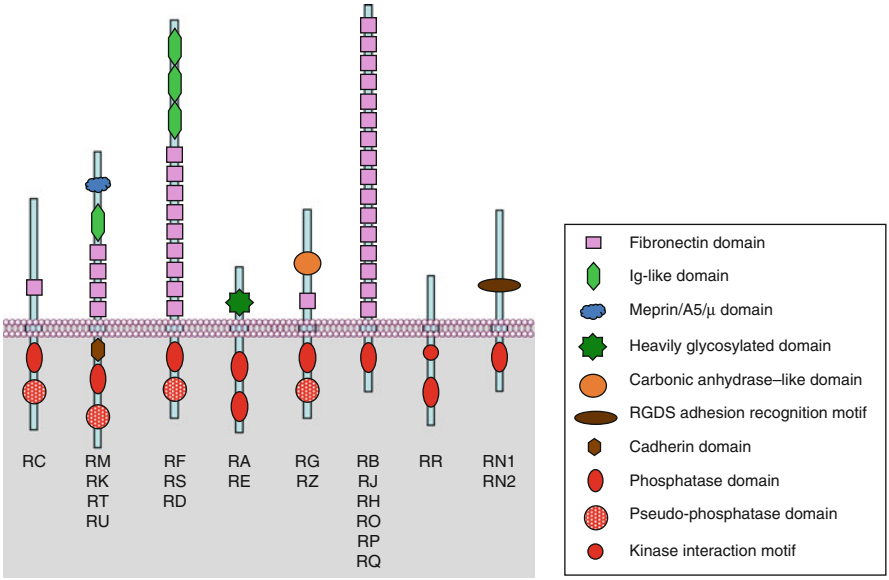


Fig. 1.7 The tyrosine phosphatase receptors family. There are 21 tyrosine phosphatase receptors, distributed in 8 families. Several nomenclatures have been proposed; we have chosen the HUGO gene nomenclature, using the prefix PTPR (protein tyrosine phosphatase receptor type) followed by a letter from A to U, with few omissions. Another nomenclature uses the prefix RPTP followed by a Greek letter, and some tyrosine phosphatase receptors have received common names such as DEP1 (*density-enhanced phosphatase 1*, gene *PTPRJ*) or LAR (*leukocyte antigen-related*, gene *PTPRF*). All of them have a transmembrane domain and two intracellular phosphatase domains, the distal one being inactive. They carry various motifs and domains in their extracellular portion, justifying the variety of potential ligands. Their cognate ligands remain yet poorly known

transmembrane segment and an intracytoplasmic catalytic domain, which is duplicated in tandem in 12 PTPRs, the distal transmembrane segment being most often catalytically inactive. The active site is defined by the sequence HCX_5R, the cysteine residue being essential for catalysis. These proteins function in a coordinated manner with protein tyrosine kinases and control signalling pathways involved in a broad spectrum of basic physiological processes. Non-receptor tyrosine phosphatases (PTPN) constitute another group of phosphatases that share some functions with PTPRs but have no role in extracellular signal processing.

Signal transduction through PTPRs occurs through dephosphorylation of tyrosine residues of target proteins. Whereas cytoplasmic tyrosine phosphatases may be, for a large part of them, of dual specificity (acting either on phosphotyrosine or phosphoserine/phosphothreonine residues), all tyrosine phosphatase receptors have single specificity. PTPRs exist in the membrane as monomers or dimers and are inactive when dimerised: a helix-turn-helix domain of one monomer inhibits as a wedge the catalytic site of the other one, and reciprocally. This contrasts with tyrosine kinase receptors, which are activated by dimerisation. Several non-exclusive models have been proposed for PTPR regulation. In a first model, PTPRs are

constitutively monomeric and active and downregulate tyrosine phosphorylation of various proteins; ligand binding induces receptor dimer formation, which deactivates the catalytic activity of the receptor. In a second model, PTPRs are constitutively organised as inactive dimers; ligand binding induces the loss of dimerisation and unmasks the phosphatase activity.

Another activation mechanism of PTPR dimerisation, and consequently activity, is provided by oxidation–reduction reactions. It was shown in vitro that hydrogen peroxide (H_2O_2) could stabilise the dimeric, inactive form of PTPRs. The cysteine residues of the catalytic sites can serve as sensors of reactive oxygen species and can form intermolecular disulphide bonds which are responsible for dimer stabilisation. Such monomer covalent binding can also occur between distinct PTPR monomers. This redox mechanism is reversible, through reduction of the disulphide bond which restores the monomeric state.

Until now, the cognate extracellular ligands and intracellular substrates of most PTPRs remain largely unknown. The identified PTPR ligands are mostly heparan sulphate proteoglycans (HSPGs), such as syndecans (SDC), glypicans (GPC), agrin (AGRN) or pleiotrophin (PTN), a ligand of the ALK and LTK tyrosine kinase receptors. Other PTPR ligands are proteins of the extracellular matrix (ECM) such as collagen XVIII (gene *COL18A1*) or laminin–nidogen complexes. Several PTPRs present the characteristics of cell adhesion molecules of the ICAM family (Chap. 11), and they interact with each other through homophilic interactions: the ligand of one PTPR molecule being the receptor of the other one, and reciprocally.

Effector proteins downstream PTPRs are certainly as varied as those located downstream TKRs. Only some of them have been identified, and only for some PTPRs. For instance, RPTPζ (gene *PTPRZ1*) dimerisation (thus inactivation) under the action of pleiotrophin is accompanied by an increase in the level of phosphorylation of β-catenin (Chap. 7) and, therefore, its destruction in the proteasome; it also increases the level of phosphorylation of β-adducin and $p190^{RHOGAP}$, two proteins involved in cytoskeletal functions, the second one being the GTPase activator of small G-proteins of the RHO family (see Chaps. 10 and 11). Another example is the PTPR called LAR (*leukocyte antigen-related*) (gene *PTPRF*), which is activated by syndecans, and dephosphorylates auto-phosphorylated cytoplasmic tyrosine kinases, such as SRC, FAK and ABL, as well as TKRs such as EGFR, ERBB2 or MET. Similarly, CD45 (gene *PTPRC*), when activated, limits the degree of phosphorylation of the cytoplasmic tyrosine kinase LYN; preventing dimer formation of CD45 leads to lymphoproliferative syndromes, since LYN acts as a negative regulator of the B-cell receptor of lymphocytes (Chap. 13).

Because of their dephosphorylation activity of auto-activated tyrosine kinases, several PTPRs play a tumour suppressor role. Mutations and deletions of DEP1 (gene *PTPRJ*) have been found in colon, breast and lung cancers, and convergent evidence suggests that RPTPκ (gene *PTPRK*) mutations are involved in central nervous system lymphomas, whereas RPTPγ (gene *PTPRG*), LAR (gene *PTPRF*) and RPTPρ (gene *PTPRT*) are potential tumour suppressors in colorectal cancers.

Further Reading

Acquaviva J, Wong R, Charest A. The multifaceted roles of the receptor tyrosine kinase ROS in development and cancer. Biochim Biophys Acta. 2009;1795:37–52.
Appleman LJ. MET signaling pathway: a rational target for cancer therapy. J Clin Oncol. 2011;29:4837–8.
Beenken A, Mohammadi M. The FGF family: biology, pathophysiology and therapy. Nat Rev Drug Discov. 2009;8:235–53.
Blume-Jensen P, Hunter T. Oncogenic kinase signalling. Nature. 2001;411:355–65.
Blumenschein Jr GR, Mills GB, Gonzalez-Angulo AM. Targeting the hepatocyte growth factor-cMET axis in cancer therapy. J Clin Oncol. 2012;30:3287–96.
Bose R, Zhang X. The ErbB kinase domain: structural perspectives into kinase activation and inhibition. Exp Cell Res. 2009;315:649–58.
Brodeur GM, Minturn JE, Ho R, Simpson AM, Iyer R, Varela CR, Light JE, Kolla V, Evans AE. Trk receptor expression and inhibition in neuroblastomas. Clin Cancer Res. 2009;15:3244–50.
Brooks AN, Kilgour E, Smith PD. Molecular pathways: fibroblast growth factor signaling: a new therapeutic opportunity in cancer. Clin Cancer Res. 2012;18:1855–62.
Burgess AW, Cho HS, Eigenbrot C, et al. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol Cell. 2003;12:541–52.
Chiarle R, Voena C, Ambrogio C, Piva R, Inghirami G. The anaplastic lymphoma kinase in the pathogenesis of cancer. Nat Rev Cancer. 2008;8:11–23.
Ciardiello F, Tortora G. EGFR antagonists in cancer treatment. N Engl J Med. 2008;358:1160–74.
Citri A, Yarden Y. EGF-ERBB signalling: towards the systems level. Nat Rev Mol Cell Biol. 2006;7:505–16.
Dieci MV, Arnedos M, Andre F, Soria JC. Fibroblast growth factor receptor inhibitors as a cancer treatment: from a biologic rationale to medical perspectives. Cancer Discov. 2013;3:264–79.
Ellis LM, Hicklin DJ. VEGF-targeted therapy: mechanisms of anti-tumour activity. Nat Rev Cancer. 2008;8:579–91.
Ferguson KM. Structure-based view of epidermal growth factor receptor regulation. Annu Rev Biophys. 2008;37:353–73.
Ferrara N. VEGF and the quest for tumour angiogenesis factors. Nat Rev Cancer. 2002;2:795–803.
Green JL, Kuntz SG, Sternberg PW. Ror receptor tyrosine kinases: orphans no more. Trends Cell Biol. 2008;18:536–44.
Gschwind A, Fischer OM, Ullrich A. The discovery of receptor tyrosine kinases: targets for cancer therapy. Nat Rev Cancer. 2004;4:361–70.
Hallberg B, Palmer RH. Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. Nat Rev Cancer. 2013;13:685–700.
Harari PM, Allen GW, Bonner JA. Biology of interactions: anti-epidermal growth factor receptor agents. J Clin Oncol. 2007;25:4057–65.
Harper SJ, Bates DO. VEGF-A splicing: the key to anti-angiogenic therapeutics? Nat Rev Cancer. 2008;8:880–7.
Harris RC, Chung E, Coffey RJ. EGF receptor ligands. Exp Cell Res. 2003;284:2–13.
Heuckmann JM, Rauh D, Thomas RK. Epidermal growth factor receptor (EGFR) signaling and covalent EGFR inhibition in lung cancer. J Clin Oncol. 2012;30:3417–20.
Holbro T, Civenni G, Hynes NE. The ErbB receptors and their role in cancer progression. Exp Cell Res. 2003;284:99–110.
Holbro T, Hynes NE. ErbB receptors: directing key signaling networks throughout life. Annu Rev Pharmacol Toxicol. 2004;44:195–217.
Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors. Nat Rev Cancer. 2005;5:341–54.
Jones RB, Gordus A, Krall JA, MacBeath G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature. 2006;439:168–74.
Jorissen RN, Walker F, Pouliot N, Garrett TPJ, Ward CW, Burgess AW. Epidermal growth factor receptor: mechanisms of activation and signalling. Exp Cell Res. 2003;284:31–53.

Julien SG, Dubé N, Hardy S, Tremblay ML. Inside the human cancer tyrosine phosphatome. Nat Rev Cancer. 2011;11:35–49.
Kumar A, Petri ET, Halmos B, Boggon TJ. Structure and clinical relevance of the epidermal growth factor receptor in human cancer. J Clin Oncol. 2008;26:1742–51.
Leahy DJ. A molecular view of anti-ErbB monoclonal antibody therapy. Cancer Cell. 2008;13:291–3.
Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010;141:1117–34.
Linger RM, Keating AK, Earp HS, Graham DK. TAM receptor tyrosine kinases: biologic functions, signaling, and potential therapeutic targeting in human cancer. Adv Cancer Res. 2008;100:35–83.
Mohebiany AN, Nikolaienko RM, Bouyain S, Harroch S. Receptor-type tyrosine phosphatase ligands: looking for the needle in the haystack. FEBS J. 2013;280:388–400.
Nikolaienko RM, Agyekum B, Bouyain S. Receptor protein tyrosine phosphatases and cancer: new insights from structural biology. Cell Adh Migr. 2012;6:356–64.
Nunes-Xavier CE, Martín-Pérez J, Elson A, Pulido R. Protein tyrosine phosphatases as novel targets in breast cancer therapy. Biochim Biophys Acta. 1836;2013:211–26.
Olayioye MA, Neve RM, Lane HA, Hynes NE. The ErbB signaling network: receptor heterodimerisation in development and cancer. EMBO J. 2000;19:3159–67.
Olsson AK, Dimberg A, Kreuger J, Claesson-Welsh L. VEGF receptor signalling – in control of vascular function. Nat Rev Mol Cell Biol. 2006;7:359–71.
Ostman A, Hellberg C, Böhmer FD. Protein-tyrosine phosphatases and cancer. Nat Rev Cancer. 2006;6:307–20.
Pasquale EB. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. Nat Rev Cancer. 2010;10:165–80.
Phay JE, Shah MH. Targeting RET receptor tyrosine kinase activation in cancer. Clin Cancer Res. 2010;16:5936–41.
Pollak M. Insulin and insulin-like growth factor signalling in neoplasia. Nat Rev Cancer. 2008;8:915–28.
Roskoski Jr R. The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacol Res. 2014;79:34–74.
Schlessinger J. Ligand-induced, receptor-mediated dimerisation and activation of EGF receptor. Cell. 2002;110:669–72.
Schneider MR, Wolf E. The epidermal growth factor receptor ligands at a glance. J Cell Physiol. 2009;218:460–6.
Scott LM, Lawrence HR, Sebti SM, Lawrence NJ, Wu J. Targeting protein tyrosine phosphatases for anticancer drug discovery. Curr Pharm Des. 2010;16:1843–62.
Seshacharyulu P, Ponnusamy MP, Haridas D, Jain M, Ganti AK, Batra SK. Targeting the EGFR signaling pathway in cancer therapy. Expert Opin Ther Targets. 2012;16:15–31.
Sequist LV, Lynch TJ. EGFR tyrosine kinase inhibitors in lung cancer: an evolving story. Annu Rev Med. 2008;59:429–42.
Sorkin A, Goh LK. Endocytosis and intracellular trafficking of ErbBs. Exp Cell Res. 2009;315:683–96.
Tebbutt N, Pedersen MW, Johns TG. Targeting the ERBB family in cancer: couples therapy. Nat Rev Cancer. 2013;13:663–73.
Tonks NK. Protein tyrosine phosphatases: from genes, to function, to disease. Nat Rev Mol Cell Biol. 2006;7:833–46.
Tonks NK. Protein tyrosine phosphatases – from housekeeping enzymes to master regulators of signal transduction. FEBS J. 2013;280:346–78.
Verstraete K, Savvides SN. Extracellular assembly and activation principles of oncogenic class III receptor tyrosine kinases. Nat Rev Cancer. 2012;12:753–66.
Vlacich G, Coffey RJ. Resistance to EGFR-targeted therapy: a family affair. Cancer Cell. 2011;20:423–5.
Weiner HL, Zagzag D. Growth factor receptor tyrosine kinases: cell adhesion kinase family suggests a novel signaling mechanism in cancer. Cancer Invest. 2000;18:544–54.

Wilson KJ, Gilmore JL, Foley J, Lemmon MA, Riese DJ. Functional selectivity of EGF family peptide growth factors: implications for cancer. Pharmacol Ther. 2009;122:1–8.

Wykosky J, Debinski W. The EphA2 receptor and ephrinA1 ligand in solid tumors: function and therapeutic targeting. Mol Cancer Res. 2008;6:1795–806.

Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol. 2001;2:127–37.

Zha J, Lackner MR. Targeting the insulin-like growth factor receptor-1R pathway for cancer therapy. Clin Cancer Res. 2010;16:2512–7.

Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell. 2006;125:1137–49.

MAP Kinase Pathway

2

Abstract

The MAP kinase pathway (or, better, the MAP kinase pathway*s*) is certainly one of the main proliferation pathways, the best known and the one whose oncogenic alterations have elicited the highest number of studies devoted to these alterations. This pathway is activated downstream numerous membrane receptors, especially those of the EGF receptor family, and leads ultimately to the activation of transcription factors called MAPs (*mitogen-activated proteins*), which govern the transcription of numerous genes required for DNA replication and cell division. Besides the classical pathway, the best known, which will be presented in detail, several parallel MAP kinase pathways have been identified. They are activated by other stimuli and lead to the activation of other transcription factors involved in stress response, inflammation and development, rather than in cell proliferation.

In the previous chapter, we showed that the phosphotyrosine residues of the activated receptor could be recognised by a large number of adapter proteins bearing an SH2 or a PTB domain. We will start this chapter with some of these proteins, which associate receptor activation to a GTP exchange protein, which in turn activates a small G-protein of crucial importance, the RAS protein. RAS activation triggers a cascade of phosphorylations leading to the activation of transcription factors.

2.1 From Receptor Activation to RAS Activation

The *growth factor receptor-bound protein 2* (GRB2) is an adapter protein with SH2 and SH3 domains. Its SH2 domain recognises specific phosphotyrosine residues of activated TKRs (Chap. 1). This allows the enrolment at the membrane of a GDP–GTP exchange factor (GEF) called SOS1 (*son of sevenless homologue 1*), which replaces GDP by GTP as a ligand of the membrane protein RAS. The RAS

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_2

proteins belong to the large protein family of *small G-proteins*, and GDP–GTP exchange on a small G-protein is commonly mediated by GEFs. Whereas the pathway EGFR–GRB2–SOS1–RAS is the best studied, there exist multiple other pathway downstream activations of tyrosine kinase receptors: other SH2 domain-containing adapter proteins are able to bind phosphotyrosine residues of tyrosine kinase receptors. This is especially the case for FGF receptors, for which a protein called FRS (*FGF receptor substrate*) is phosphorylated by the receptor and then recognised, at the level of its phosphotyrosine residues, by GRB2 or another adapter protein.

Three RAS protein homologues play similar roles in various tissues: KRAS (Kirsten RAS), HRAS (Harvey RAS) and NRAS (neuroblastoma RAS). All small G-proteins function similarly: they display high affinity for GDP, which is replaced by GTP via a GEF. This exchange induces a change of conformation, triggering downstream activity (for instance, RAS activation by GTP leads in particular to its binding to a serine/threonine kinase called RAF). Small G-proteins possess an intrinsic GTPase activity, allowing the retroconversion of GTP to GDP by elimination of the GTP gamma phosphate group. This GTPase activity may be stimulated by a GTPase-activating protein (GAP), which constitutes therefore the third member of the functional triad GEF–small G-protein–GAP, which is involved in many cellular events outside the control of proliferation studied in this chapter.

There are other possibilities of RAS activation. In the pathway opened by TKR activation, the GEF is SOS1 and the GAP is RASGAP ($p120^{RASGAP}$, gene *RASA1*). Another GEF for RAS is RASGRP (*RAS guanyl nucleotide-releasing protein*), which can be activated by other molecules: inositol trisphosphate (IP3) through Ca^{2+} mobilisation and diacylglycerol (DAG). These two molecules can be formed from phosphatidylinositol by phospholipase C gamma (PLCγ), which possesses a SH2 domain allowing its activation by a TKR (see Chap. 1). IP3 and DAG are 'second messengers' that can also be formed by phospholipase C beta (PLCβ), which is activated via another signalling pathway (Chap. 6) involving receptors with seven transmembrane domains called GPCR (*large heterotrimeric G-protein-coupled receptors*).

The RAS proteins are anchored to the plasma membrane through covalent bonds with a branched hydrophobic moiety constituted of 15 carbon atoms (farnesyl) or 20 carbon atoms (geranylgeranyl). These prenyl moieties, which are intermediates of cholesterol biosynthesis, are added as post-translational modifications at the level of the endoplasmic reticulum. Their binding to the RAS proteins occurs on a cysteine residue of the polypeptidic chain, borne by a –CAAX motif (C for cysteine, A for aliphatic amino acid (Val, Leu or Ile), X for Met or Ser), the three amino acids AAX being eliminated after prenylation by RAS-converting endopeptidase (gene *RCE1*) (Fig. 2.1). In addition, some RAS proteins undergo palmitoylation in the Golgi apparatus. This modification is a covalent binding of palmitic acid, a fatty chain of 16 carbon atoms, at the level of a cysteine residue located upstream the one that is prenylated. It is generally admitted that the recruitment of SOS1 to the membrane,

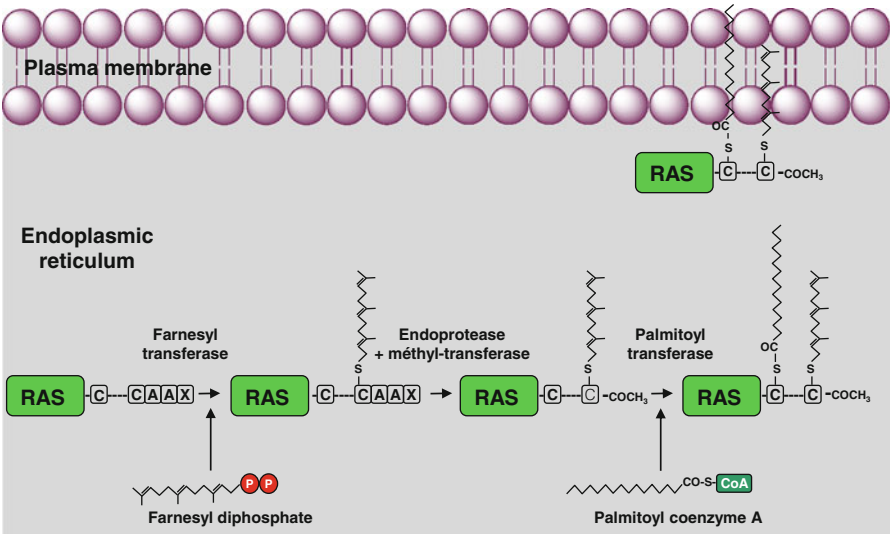


Fig. 2.1 Post-translational modifications of RAS proteins. The RAS proteins undergo, in the endoplasmic reticulum, covalent binding to a prenyl (farnesyl or geranylgeranyl) moiety on a cysteine residue included in a –CAAX box. The amino acids AAX are then deleted and replaced by an acetyl moiety at the *C*-terminal end. Later, a palmitic acid is bound to RAS proteins on another cysteine residue located some amino acids upstream of the prenylated cysteine. These modifications are required for the plasma membrane insertion of RAS proteins

thanks to its affinity for GRB2, enables its interaction with its RAS protein substrate and, as a consequence, RAS activation. This membrane anchoring is required for RAS activity.

When RAS becomes bound to GTP, it undergoes a conformational modification that allows the recruitment to the membrane of a serine/threonine kinase which is a RAF protein in the MAP kinase pathway. Besides being the main activator of the MAP kinase pathway described below, RAS has many effector proteins (Fig. 2.2): RAS can directly activate the catalytic subunit of PI3 kinase (Chap. 3), as well as other enzymes such as phospholipase C epsilon (PLCε). It can also regulate the activity status of other small G-proteins by activating their GEFs or GAPs: this is the case for RAL, which plays a role in endocytosis, and for RHO family G-proteins, which are involved in the control of cytoskeleton movements. The RAS proteins are, therefore, at an important crossroad of signalling pathways. RAS protein activation is a rapidly reversible, since intrinsic RAS GTPase activity, stimulated by the RASGAP protein, rapidly deactivates RAS by leading it back to its inactive form, RAS-GDP, which is unable to recognise RAF and attract it to the plasma membrane. RASGAP presents an SH2 domain allowing it to be recruited at the membrane level by a TKR to exert its action. The successive steps of the MAP kinase pathway are described below (Fig. 2.3).

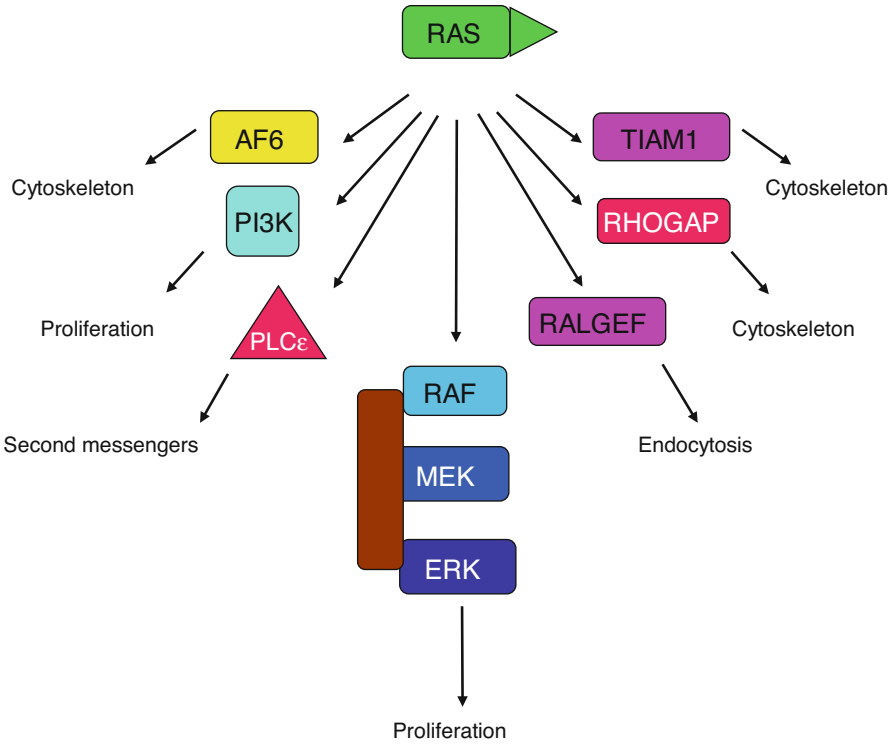


Fig. 2.2 Effectors of the RAS proteins. The RAS proteins have a multiplicity of targets and can activate a large variety of signalling pathways. The MAP kinase pathway is central in this respect, but activation of phospholipase Cε (*PLCε*) and of PI3 kinase also concurs to proliferation, while the activation of G-protein exchange factors (GEFs) and GTPase-activating proteins (GAPs) entails several roles on cell processes such as endocytosis and cytoskeleton organisation

2.2 Kinase Cascade

The direct target of RAS in the MAP kinase pathway is a serine/threonine kinase called RAF. There are three distinct RAF proteins, ARAF, BRAF and CRAF (gene *RAF1*). RAF binding to RAS–GTP induces a conformational change which unveils an activation loop allowing their phosphorylation and consequently their activation. The activation process of the RAF proteins is complex and not completely deciphered; it involves RAF dimerisation and phosphorylation on various sites by several kinases such as cyclic AMP-activated protein kinase A (PKA), which is brought into play following GPCR activation (Chap. 6); the cytoplasmic tyrosine kinase SRC, activated by several types of stimulus; the JAK kinase activated by cytokine receptors (Chap. 4) or the kinase named PAK1 (*p21-activated kinase 1*), which operates after activation of the integrin (Chap. 10) or semaphorin (Chap. 11) pathways.

RAF proteins may undergo phosphorylation on several distinct sites. Two of them are in the *N*-terminal domain, on an SSYY motif where the first serine residue

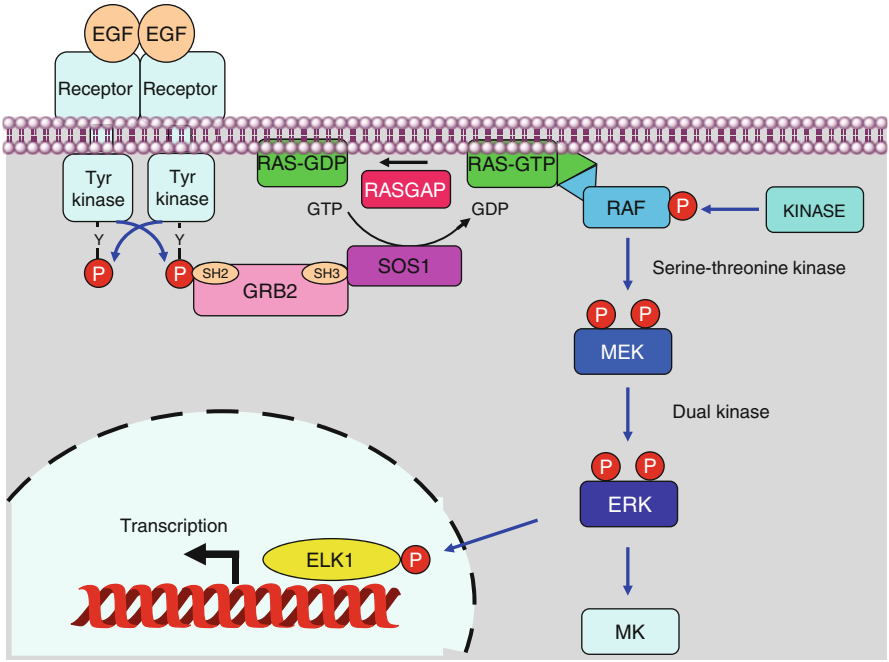


Fig. 2.3 MAP kinase pathways. In the canonical signalling pathway, the activation of a TKR enables the recruitment of an SH2 domain-containing protein, GRB2; the SH3 domain of GRB2 is recognised by the SOS1 protein, a GDP–GTP exchange factor for RAS proteins. RAS–GTP can recruit to the plasma membrane a RAF protein, which thus can be phosphorylated by a variety of other kinases (PKA, PAK, SRC) activating its kinase function. RAS–GTP is then deactivated by its own GTPase activity stimulated by a RASGAP (*RAS GTPase-activating protein*) protein. RAF phosphorylates and activates a MEK protein, which in turn activates an ERK protein, which in turn phosphorylates and activates a transcription factor such as ELK1 or a kinase of the MK family

and the second tyrosine residue must be phosphorylated; however, for BRAF, the first serine residue is constitutively phosphorylated, and the two tyrosine residues are replaced by aspartic acid residues whose negative charge is constitutively 'phosphomimetic'. Two other phosphorylation sites are located on a threonine and a serine residue (Thr^{599} and Ser^{602} for BRAF). Whereas ARAF and CRAF require four phosphorylated residues for activity, BRAF requires only two; the consequences upon RAF-mediated oncogenesis will be detailed below.

Downstream RAF kinases are the kinases MEK1 and MEK2, and downstream MEK1/2 are the kinases ERK1 and ERK2 (*extracellular signal-regulated kinases*). This constitutes, therefore, a three-step kinase system, the RAF–MEK–ERK cascade. The RAF kinases are also called MAP3 kinases (MAP3K or MAPKKK), the MEK kinases are MAP2K or MAPKK and the ERK kinases are the proper MAP kinases (MAPK). It will be shown later that this RAF–MEK–ERK 'module' is only one among others and that other parallel MAP kinase pathways exist, with similar three levels of phosphorylation. They are acting in stress response or inflammation rather than in cell proliferation. A total of 25 MAP3K, 7 MAP2K and 13 MAPK

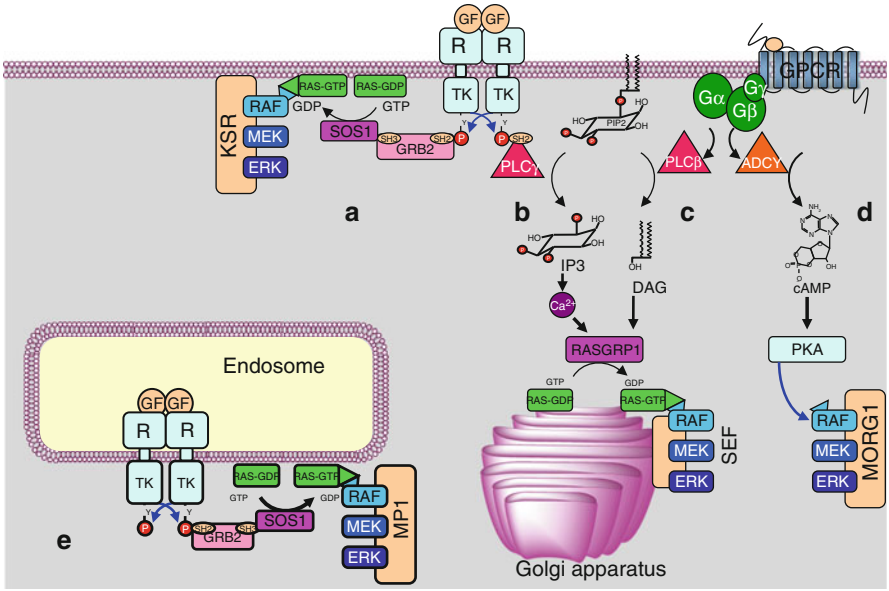


Fig. 2.4 Kinases organisation and scaffold proteins. The three MAP kinases acting sequentially are not scattered in the cytoplasm but assembled by scaffold proteins. (**a**) The signals brought by a growth factor (GF) to a tyrosine kinase receptor (TKR) activate the RAS–RAF–MEK–ERK pathway at the plasma membrane level, thanks to the scaffold protein KSR (*kinase suppressor or RAS*). (**b**) They can also activate phospholipase C gamma (PLCγ) leading, through the second messengers inositol trisphosphate (IP3) and diacylglycerol (DAG), to the activation of the RAF–MEK–ERK cascade in the Golgi apparatus, thanks to the scaffold protein SEF (*similar expression to FGF gene*). (**c**, **d**) Signals activating a G-protein-coupled receptor (GPCR) can activate the RAF–MEK–ERK pathway, either by generating the second messengers IP3 and DAG produced by PLCβ (**c**), or by generating the second messenger cAMP (cyclic AMP), which activates the protein kinases A (PKA), using the scaffold protein MORG1 (*MAPK organiser 1*) (**d**). (**e**) The activation of a TKR at the endosomal level leads to a classical activation of a RAS protein, then to the sequence RAF–MEK–ERK using the scaffold protein MP1. Other systems may exist, involving the same kinase cascade RAF–MEK–ERK and other scaffold proteins such as β-arrestin or paxillin

have been identified in the superfamily of serine/threonine kinases. The official gene names of the MEK and ERK proteins studied here differ from their common denomination: *MAP2K1* and *MAP2K2* for MEK1 and MEK2, *MAPK3* and *MAPK1* for ERK1 and ERK2, respectively.

MEK1/2 activation is obtained by phosphorylation of two close serine residues: Ser^{218} and Ser^{222} for MEK1 and Ser^{222} and Ser^{226} for MEK2, located in the activation loop of the catalytic centre. The most active kinase on MEK1/2 is BRAF, followed by CRAF, while ARAF has lower activity. MEK1/2 is in turn able to phosphorylate ERK1/2. MEKs are dual kinases, able to phosphorylate both a threonine and a tyrosine residue, also at the level of the activation loop, following each other as TEY (Thr^{202}–Glu^{203}–Tyr^{204} for ERK1 and Thr^{185}–Glu^{186}–Tyr^{187} for ERK2). Kinase specificity to the kinase of lower level depends on interactions between protein–protein binding domains rather than on different catalytic activities.

The three successive steps of kinase action are not scattered in the cytoplasm. Scaffold proteins are able to maintain the three sequentially acting kinases as a functional complex. For the RAS–RAF–MEK–ERK pathway, this scaffold protein is KSR1 (*kinase suppressor of RAS-1*) and operates at the level of the plasma membrane (Fig. 2.4). Other scaffold proteins operate for this MAP kinase module at other sites: endosomes, Golgi apparatus and focal adhesions; the scaffold protein of the Golgi apparatus is recruited by the exchange factor RASGRP1, activated by the second messengers generated by the PLCs γ and β, as mentioned earlier.

Downstream ERK1 and 2 are two types of substrates: a large set of transcription factors involved in proliferation and various kinases of the MK (*MAPK-activated protein kinases*) family; they are presented in paragraph 4. Negative regulators of the MAP kinase pathway consist in a series of dual phosphatases, called MKP (*MAP kinase phosphatases*) or DUSP (*dual specificity phosphatases*). They operate at the level of phosphothreonine and phosphotyrosine residues, with variable substrate specificity, at the level of the activating sequences TXY.

2.3 Other Signalling Modules

As mentioned earlier, there are several signalling modules operating downstream RAS proteins, organised as phosphorylation cascades, with distinct MAP3Ks, MAP2Ks, MAPKs and scaffold proteins. These are the JNK module, the p38 module and the ERK5 module. Figure 2.5 presents these pathways, which are parallel to the RAF–MEK–ERK canonical pathway. It is important to note that their cellular effects involve differentiation or apoptosis rather than cell proliferation, exemplifying the difficulty to assign a unique function to a given signalling pathway: the 'cellular context', i.e. the conditions in which these effects are exerted, plays a crucial role.

There are three JUN *N*-terminal kinases: JNK1, 2 and 3 (genes *MAPK8*, *9* and *10*), also known as SAPKs (*stress-activated protein kinases*). They are activated in response to various stresses such as heat shock, ionising radiations or DNA damage rather than by growth factors. TNF (*tumour necrosis factor*) and its analogues as well as the WNT ligands (Chap. 7) also activate this pathway. They are phosphorylated on a TPY sequence (Thr^{183}–Pro^{184}–Tyr^{185} for JNK1) by MEK4 and MEK7 (genes *MAP2K4* and *MAP2K7*), which are themselves phosphorylated by diverse MAP3K at positions Ser^{257} and Thr^{261} for MEK4 and Ser^{271} and Thr^{275} for MEK7 (see Fig. 2.5). The JUN kinases phosphorylate transcription factors such as JUN, MYC and ELK1, which in turn regulate the expression of specific genes involved in cell death or survival, differentiation or proliferation. They also phosphorylate several kinases of the MK family.

There are four p38 kinases: p38 α, β, γ and δ (genes *MAPK14*, *11*, *12* and *13*), which are also activated in response to stress (hypoxia, UV radiations, osmotic shock, inflammation) rather than by growth factors. They are activated by phosphorylation on a TGY sequence (Thr^{180}–Gly^{181}–Tyr^{182} for p38α) by MEK3 and

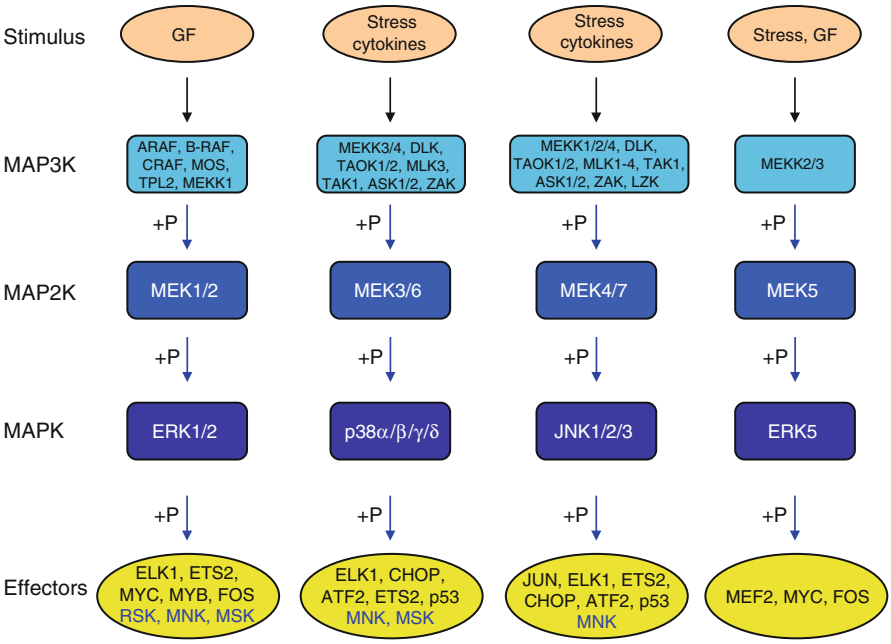


Fig. 2.5 The different modules of kinase cascades. Five main modules of multistep kinases have been identified; all of them originate from a stimulus (growth factor, cytokine, stress, etc.) which activates a MAP3K, which phosphorylates and activates a MAP2K, which phosphorylates and activates a MAPK, which phosphorylates and activates an effector (transcription factor or another serine/threonine kinase). The initial stimulus may originate from a small G-protein such as RAS or from phosphorylation by a fourth-level kinase (MAP4K). From left to right, the classical ERK1/2 pathway, the p38 pathway, the JNK pathway and the ERK5 pathway

MEK6 (genes *MAP2K3* and *MAP2K6*) themselves phosphorylated by various MAP3K at positions Ser^{253} and Thr^{257} for MEK3 and Ser^{241} and Thr^{245} for MEK6. They also phosphorylate transcription factors and MK kinases among which several are common to the ERK1/2 and JNK1/2/3 pathways. They exert a positive effect on apoptosis and a negative effect on cell proliferation; their role is likely to prevent cells that have undergone stress from entering the cell cycle. They also have a positive role on cell motility.

Finally, ERK5 or BMK1 (*big mitogen-activated protein kinase 1*, gene *MAPK7*), activated by MEK5 (gene *MAP2K5*), is a MAPK module involved in activating cell proliferation, as the ERK1/2 module. Other MAP kinases such as MAPK4 and MAPK6 remain poorly studied.

For each of these modules, there are distinct scaffold proteins, different phosphatases and various regulatory proteins: there is an important information redistribution from the original activation of a MAP3K, explaining the pleiotropic aspect of their physiological actions. The activation of these pathways generally originates from small G-proteins, whose archetype is RAS, which recruit the first kinase of the cascade for phosphorylation. Other systems of activation exist, with fourth-level kinases (MAP4K) activated in response to various signals.

2.4 MAP Kinase Substrates

MAP kinases have the basal property of phosphorylating numerous transcription factors, such as ELK1, ETS, FOS, JUN, MYC and SP1 which induce the expression of numerous genes involved in cell cycle and cell proliferation; most of them were already known as proto-oncogenes. MAP kinases can also phosphorylate a series of kinases with various biological properties, known as MKs (*MAPK-activated protein kinases*), comprising RSKs (ribosomal S6 kinases), MSKs (*mitogen and stress-activated kinases*) and MNKs (*MAP kinase-interacting proteins*). These kinases exert various effects in the cell, especially at the level of cell proliferation and adhesion.

2.4.1 MAP Kinases-Activated Transcription Factors

MYC Family

The transcription factors of this family are characterised by an HLH (helix–loop–helix; see Annex B) motif, which allow their dimerisation with a common partner, MAX (*MYC-associated factor X*). Their activation is finely regulated and their half-life is short; MAX is in contrast stable, ubiquitous and permanently expressed. The MAX–MAX homodimers have no transcriptional activity, and only heterodimers are active. The proteins able to bind MAX belong to several groups: the MYC group, originally identified from the oncogenes c-*myc*, N-*myc* and L-*myc* (named after the avian myelocytomatosis) and comprising the MYC, MYCN and MYCL proteins; the MAD or MXD (*MAX dimerisation protein*) group and the MNT (*MAX-interacting protein*) and MGA (*MAX gene associated*) proteins.

The MYC–MAX complexes are transcriptional activators that bind promoter sequences containing the hexanucleotide CACGTG of the E-box and recruiting the histone acetyltransferases required for chromatin relaxation (see Annex B). They are also able to repress the transcription of other genes when binding the initiator elements (INR) of other gene promoters. Numerous genes are activated by MYC at the transcriptional level: genes involved in cell growth and proliferation and in cell cycle engagement (cyclins, CDKs (Chap. 17), telomerase (Annex A), translation-initiating factors (Annex C), etc.). Among the genes whose transcription is repressed by MYC are negative regulators of the cell cycle ($p21^{CIP1}$, $p27^{KIP1}$, $p15^{INK4b}$), as well as the gene encoding p53. Paradoxically, MYC is able to induce apoptosis through BAX activation (see Chap. 18).

In contrast, the MAD–MAX complexes are transcriptional repressors at the level of the same gene promoters. This repression involves the corepressor SIN3 (*silencing homologue 3*), which belongs to the histone deacetylase complex. MAD activity induces a blockade of cell proliferation and an inhibition of transformation, as well as the induction of differentiation in several tissues.

FOS–JUN Family

The transcription factors JUN and FOS (named after the corresponding oncogenes) belong to a family called AP1 (*activator protein 1*) and are associated to

cell proliferation and differentiation. They operate after dimerisation. There are three JUN proteins (JUN or c-JUN, JUNB and JUND) and four FOS proteins (FOS or c-FOS, FOSB, FRA1 and FRA2). The JUN proteins can form homodimers or heterodimers with the FOS proteins, but the FOS proteins can only dimerise with JUN proteins. Heterodimerisation with the transcription factors of the MAF (*musculoaponeurotic fibrosarcoma*) and ATF (*activating transcription factor*) families is also possible. The MAP kinases ERK, as well as those of the JNK and p38 modules, are able to phosphorylate the components of the AP1 complexes. Other phosphorylations are possible: for instance, JUNB can be phosphorylated by the mitosis entry cyclin B–CDK1 complex (Chap. 17) which drives it to proteolysis.

The various dimers that can be formed may have opposite physiological roles: the JUN–FOS dimer induces the transcription of proliferation genes, such as the genes of cyclin D (*CCND*) or of *CDK4* and *CDK6* (Chap. 17), whereas the JUN–JUNB or JUN–JUND dimers are unable to do so; they appear, therefore, as MAP inhibitors. They induce, in contrast, the transcription of negative regulators of cell cycle. This explains why the different MAP kinase modules, which differentially activate the various components of the AP1 complexes, may have opposite effects on cell proliferation and differentiation. Among the target genes of the AP1 complexes are also those encoding cytokines (Chap. 4) and TNF-related factors.

ETS Family

This family comprises a series of transcription factors, among which are ETS1, ETS2, ELK1, SAP1, SAP2, ERF, NET and ELK3. They can combine with other transcription factors to generate functional dimers. A first group (ETS1 and 2) can combine with AP1 family elements; the group comprising ELK1 and SAP1 and 2 (collectively called TCF [*T-cell factor*]) is combined with transcription factors of the SRF family (*serum responsive element-binding transcription factor*) to activate SRE (*serum responsive element*) sites on DNA. A third group is exclusively responsible for transcription repression.

The ETS family factors target genes encoding other transcription factors, such as MYC and FOS, cell cycle regulatory proteins (cyclin D1, p53, $p21^{CIP1}$; Chap. 17), antiapoptotic proteins (BCL2; Chap. 18), cytokines (Chap. 4), growth factors (Chap. 1), etc.

FOXO Family

The FOXO (*forkhead box class O*) transcription factors, and especially FOXO3, are involved in the induction of apoptosis and cell cycle arrest by stimulating the transcription of genes encoding p21, p27 (Chap. 17), BIM and TRAIL (Chap. 18). ERK1/2 phosphorylates FOXO3 on Ser^{344}, Ser^{294} and Ser^{425}. These phosphorylations favour the interaction of FOXO3 with an E3 ubiquitin ligase, MDM2 (Chap. 17), which drives it to the proteasome. AKT (Chap. 3) is also involved in FOXO3 phosphorylation, on different serine or threonine residues, leading also to FOXO3 inactivation via another mechanism.

2.4.2 MAP Kinase-Activated Kinases

There are four RSK (*ribosomal protein S6 kinases*) and two MSK (*mitogen and stress-activated protein kinases*) (genes *RPS6KA1* to *RPS6KA6*), which can be phosphorylated and activated by ERK1/2, JNK and p38 kinases, as well as, for some of them, by PDK1 (*phosphoinositide-dependent kinase 1*) (Chap. 3); they play a role in cell growth, survival and proliferation and phosphorylate numerous substrates, such as FOS, JUN (see above) and other transcription factors (CREB, ATF, NFκB), as well as the $p27^{KIP1}$ and MYT1 proteins (Chap. 17), IκBα (Chap. 12), BAD (Chap. 18), LKB1 (gene *STK11*) (Chap. 3) and SOS1 (see above). Despite their denomination, the S6 ribosomal protein is only an accessory substrate in comparison with its activation downstream the mTOR pathway by three other RSK kinases (genes *RPS6KB1* to *RPS6KB3*) (Chap. 3).

Two MNK kinases (*MAP kinase-interacting kinases*, genes *MKNK1* and *2*) and three MAP kinase-activated kinases (genes *MAPKAPK2*, *3* and *5*), also activated by ERK1/2, JNK and/or p38, are mainly involved in translation control and admit protein synthesis elongation factors as substrates.

2.5 Oncogenic Alterations

The MAP kinase pathway is one of the principal oncogenic pathways, and several oncogenes and tumour suppressor genes have been identified along this pathway for a long time. Whereas the adapter protein GRB2 and the GDP–GTP exchange factor SOS1 do not appear to contain oncogenic mutations, the three RAS proteins, KRAS, HRAS and NRAS, harbour recurrent mutations in numerous cancer types. For instance, KRAS is mutated in about 40 % of colorectal cancers, 20 % of non-small-cell lung cancers, 70 % of pancreatic cancers and 5 % of breast cancers. This is one of the most frequently mutated proteins in human cancers. The activating mutations are mainly located in codons 12 and 13 and replace glycine residues by other amino acids. The mutated protein is still able to bind GDP, to accept the GTP exchange and to recognise and activate a RAF protein, but it becomes unable to perform its enzymatic GTPase function and generate GDP from GTP: it remains permanently active to impose a proliferation instruction. It is often said that this mutation 'transforms a one-shot gun into a machine-gun'. Less frequent mutations are found at the level of codon 61 (Gln), mutations that refrain the GTPase-stimulating activity of RASGAP. The functional consequences are similar, the RAS intrinsic GTPase activity appearing insufficient for a rapid deactivation.

For the same reasons, the loss-of-function mutations of the proteins which stimulate RAS GTPase activity (or some related factors such as the NF1 [*neurofibromin 1*] protein) are tumour suppressors; the germinal alteration of NF1 is responsible for a neurological disease, neurofibromatosis, which is accompanied by cancer predisposition.

The RAF proteins are also the sites of activating oncogenic mutations. The main one is encountered on the *BRAF* gene where it replaces a valine residue at position

600 by a glutamic acid residue. No *ARAF* mutation and very few *RAF1* mutations have been identified. This is due to the fact that several mutations in these proteins would be required to get constitutive activation, whereas BRAF, already 'prepared' at the level of the *N*-terminal phosphorylation sites, only requires a negative charge in the vicinity of the amino acids 599–602 to mimic the presence of an activating phosphate moiety. The V600E mutation, which introduces a glutamic acid residue close to the catalytic domain, fulfils this office. These mutations are encountered mainly in malignant melanomas and thyroid cancers and in other cancers such as ovarian and colorectal cancers.

The MEK and ERK proteins, in contrast, do not carry frequent and recurrent activating mutations. However, with the development of exome sequencing in a large number of tumours, activating mutations of MEK have been found, such as F129L or C121S, but no oncogenic mutations in ERK1 or ERK2 have ever been identified. These proteins are often overexpressed in cancer and likely bring a contribution to oncogenesis through this mechanism. The degree of activation of the MAP kinase pathway in cancer can be evaluated by the degree of phosphorylation of the MEK and ERK proteins, thanks to the availability of monoclonal antibodies that are specific of the phosphorylated forms of these proteins.

In contrast to the MEK1/2–ERK1/2 pathway, the MAP2K and MAPK of the p38 and JNK modules may behave as suppressors of tumours, and loss-of-function mutations of MEK4 have been identified in various cancer types, although with low frequency (pancreas, colon, lung, breast, etc.). This MAP2K is underexpressed in 75 % of ovarian cancers. The level of activity of p38 kinases and MEK6 is very low in hepatocarcinomas. However, an increase in the phosphorylation level of p38α has been associated to an increased level of malignancy of various cancers, maybe because this pathway is involved in the regulation of the expression of inflammatory cytokines such as IL6 (Chap. 4) or IL1 (Chap. 12). In addition, the JNK and p38 pathways stimulate angiogenesis and tumour invasivity because of their stimulating activity for cell motility and metalloproteinase expression. One of the main effectors of the JNK pathway, the JUN transcription factor, which is part of the AP1 complex, is oncogenic when heterodimerised with FOS factors and not when homodimerised.

Finally, the MAP transcription factors, downstream the MAPK pathway, are also important proto-oncogenes. Among the transcription factors that are phosphorylated and activated by MAPK are the proteins MYC, MYB, JUN, FOS and several others of the ETS family whose oncogenic role is well known. Their contribution to oncogenesis is generally found at the level of their expression rather than at the level of mutational alterations, that is, in the amount of the MYC or JUN available protein for MAX or FOS binding, respectively. Conversely, the MAD proteins appear as tumour suppressors. MYC is amplified in several tumour types, and the amplification of *MYCN* is a marker of poor prognosis in children's neuroblastomas.

In contrast to the MAPs, the FOXO transcription factors that are phosphorylated by the MAP kinases ERK1/2 are tumour suppressor genes; their translocation has been observed in leukaemias and sarcomas.

2.6 Pharmacological Targets

All the steps of this signalling pathway have been explored from a pharmacological point of view, aiming at the discovery of molecules leading to its inhibition which might, therefore, present antiproliferative properties for the treatment of cancer.

The adapter proteins with SH2 domains could theoretically be inhibited by compounds that mimic protein–protein interactions, i.e. peptides or peptidomimetics. This interesting approach is impeded by the fact that peptides, which can be handled in vitro, cannot enter cells easily and are poor candidates for pharmacological development.

The RAS proteins have been the subjects of intense research, principally oriented to the inhibition of its membrane insertion. Molecules able to inhibit farnesyl transferase or geranylgeranyl transferase activities have been identified and have been subjected to clinical trials. Four classes of inhibitors have been selected: farnesyl-pyrophosphate analogues, which are false substrates for farnesyltransferase; peptidomimetics of the CAAX box; bi-substrate inhibitors; and natural products selected by high-throughput screening. During preclinical exploration, several compounds displayed important antiproliferative activity. Four compounds have entered clinical trials, among which tipifarnib and lonafarnib, that are non-peptidic peptidomimetics of the CAAX box. Unfortunately, these compounds, with low activity and high toxicity, led to disappointing results. They seem to be more active on other small G-proteins, such as those of the RHO family, involved in other signalling pathways than on RAS proteins. Some clinical trials are pursued in combination with chemotherapy or radiotherapy, because these compounds have shown some radiosensibilising properties.

It should be noted that the activating RAS mutations disconnect the MAP kinase pathway from TKR activation, allowing thus the tumour cells bearing such mutations to self-sufficiency regarding growth factors, one of the main families of oncogenic mechanisms. It has been shown that, very consistently, the activating mutations of the RAS proteins induce a complete insensitivity to anti-EGFR antibodies (Chap. 1).

The proteins of the MAP kinases cascade have also been subjected to pharmacological research. An original RAF inhibitor, sorafenib, had been identified by screening; this compound is also a potent inhibitor of the VEGF receptors (Chap. 1) and appears as an antiangiogenic compound rather than as a tumour growth inhibitor. Another RAF inhibitor, vemurafenib, is now available for the treatment of malignant melanoma. It is especially active on tumours bearing the mutated form of BRAF (V600E) and should be reserved for such tumours, because it behaves paradoxically as an activator of the MAP kinase pathway in wild-type BRAF tumours. Other BRAF inhibitors are under development, such as dabrafenib. It should be noticed that resistance to BRAF inhibitors rapidly occurs, especially through activating mutations of MEK1/2.

Several inhibitors of MEK1/2 are being evaluated in the clinics, especially in combination with BRAF inhibitors in the treatment of malignant melanomas bearing the V600E or another activating mutation, in order to prevent the

development of resistance to BRAF inhibitors. The fact that ERK1 and ERK2 are the sole substrates of MEK1 and MEK2 and have no activating mutations encourages MEK targeting, which could be conceived as a general way of targeting the whole 'classical' MAP kinase pathway and not the parallel pathways. The MEK inhibitors that are the most advanced in this development are trametinib and selumetinib. In contrast to many other kinase inhibitors, they do not act as ATP-competitive molecules but interfere with an allosteric site.

The JNK and p38 pathways behave rather as tumour suppressor pathways than as pro-oncogenic pathways. The inhibition of these pathways has been developed for the treatment of inflammatory syndromes, but the molecules available for kinase inhibition do not seem to present direct anticancer effects. However, as chronic inflammation is a factor stimulating tumour progression, these compounds may interfere with oncogenesis, and they have been tried in oncology. In addition, these compounds may interfere with the proangiogenic properties of the JNK and p38 pathways and, thus, deserve attention.

Finally, even if the transcription factors, which represent the ultimate part of the MAP kinase pathway, constitute potent bona fide targets, the development of inhibitors is still in infancy. Peptidomimetics targeting the protein–protein interactions at the level of MYC–MAX or JUN–FOS, for instance, might reveal therapeutic interest in oncology.

Further Reading

Berndt N, Hamilton D, Sebti SM. Targeting protein prenylation for cancer therapy. Nat Rev Cancer. 2011;11:775–91.

Boutros T, Chevet E, Metrakos P. Mitogen-activated protein (MAP) kinase/MAP kinase phosphatase regulation: roles in cell growth, death, and cancer. Pharmacol Rev. 2008;60:261–310.

Chang F, Steelman LS, Lee JT, et al. Signal transduction mediated by the Ras/Raf/MEK/ERK pathway from cytokine receptors to transcription factors: potential targeting for therapeutic intervention. Leukemia. 2003;17:1263–93.

Dhanasekaran DN, Kashef K, Lee CM, Xu H, Reddy EP. Scaffold proteins of MAP-kinase modules. Oncogene. 2007;26:3185–202.

Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. Oncogene. 2007;26:3279–90.

Grandori C, Cowley SM, James LP, Eisenman RN. The Myc/Max/Mad network and the transcriptional control of cell behavior. Annu Rev Cell Dev Biol. 2000;16:653–99.

Karnoub AE, Weinberg RA. Ras oncogenes: split personalities. Nat Rev Mol Cell Biol. 2008;9:517–31.

Konstantinopoulos PA, Karamouzis MV, Papavassiliou AG. Post-translational modifications and regulation of the RAS superfamily of GTPases as anticancer targets. Nat Rev Drug Discov. 2007;6:541–55.

Leicht DT, Balan V, Kaplun A, Singh-Gupta V, Kaplun L, Dobson M, Tzivion G. Raf kinases: function, regulation and role in human cancer. Biochim Biophys Acta. 2007;1773:1196–212.

Lito P, Rosen N, Solit DB. Tumor adaptation and resistance to RAF inhibitors. Nat Med. 2013;19:1401–9.

McCubrey JA, Steelman LS, Abrams SL, et al. Targeting survival cascades induced by activation of Ras/Raf/MEK/ERK, PI3K/PTEN/Akt/mTOR and Jak/STAT pathways for effective leukemia therapy. Leukemia. 2008;22:708–22.

McKay MM, Morrison DK. Integrating signals from RTKs to ERK/MAPK. Oncogene. 2007;26:3113–21.
Michaloglou C, Vredeveld LC, Mooi WJ, Peeper DS. BRAF (E600) in benign and malignant human tumours. Oncogene. 2008;27:877–95.
Montagut C, Settleman J. Targeting the RAF-MEK-ERK pathway in cancer therapy. Cancer Lett. 2009;283:125–34.
Owens DM, Keyse SM. Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases. Oncogene. 2007;26:3203–13.
Pritchard AL, Hayward NK. Molecular pathways: mitogen-activated protein kinase pathway mutations and drug resistance. Clin Cancer Res. 2013;19:2301–9.
Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D. RAS oncogenes: weaving a tumorigenic web. Nat Rev Cancer. 2011;11:761–74.
Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene. 2007;26:3291–310.
Roux PP. Blenis J.ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. Microbiol Mol Biol Rev. 2004;68:320–44.
Santarpia L, Lippman SM, El-Naggar AK. Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. Expert Opin Ther Targets. 2012;16:103–19.
Schubbert S, Shannon K, Bollag G. Hyperactive Ras in developmental disorders and cancer. Nat Rev Cancer. 2007;7:295–308.
Sebolt-Leopold JS, Herrera R. Targeting the mitogen-activated protein kinase cascade to treat cancer. Nat Rev Cancer. 2004;4:937–47.
Steelman LS, Abrams SL, Whelan J, et al. Contributions of the Raf/MEK/ERK, PI3K/PTEN/Akt/mTOR and Jak/STAT pathways to leukemia. Leukemia. 2008;22:686–707.
Turjanski AG, Vaqué JP, Gutkind JS. MAP kinases and the control of nuclear events. Oncogene. 2007;26:3240–53.

Phosphatidylinositol 3-Kinase Pathway 3

Abstract

The phosphatidylinositol 3-kinase (PI3K) pathway is one of the signalling pathways activated following the interaction of a growth factor with a tyrosine kinase receptor (TKR). It follows a parallel path to the MAP kinase pathway, and similarly, it operates after binding a phosphotyrosine residue of the receptor to an adapter protein. It consists of sequential activations of kinases and leads to multiple effects on cell growth, proliferation and survival.

This pathway is especially interconnected with the MAP kinase pathway through the RAS proteins; it is, in addition, able to integrate metabolic and nutritional signals which allow to coordinate cell growth and proliferation to nutrient availability. This is the pathway adopted by insulin, which can be considered by some aspects as a growth factor, as mentioned in Chap. 1. Multiple oncogenic alterations have been identified in this pathway, which contains for this reason numerous potential targets for therapeutic approaches.

3.1 From Phosphatidylinositol 3-Kinase to AKT Proteins

Phosphatidylinositol 3-kinases (PI3Ks) are a family of lipid kinases related to serine/threonine kinases. They function as heterodimers consisting of a catalytic subunit and a regulatory subunit. In the main PI3Ks, responsible for the proliferation pathway described in this chapter, the catalytic subunit (p110) is encoded by the *PIK3CA* gene, and the regulatory subunit (p85) is encoded by the *PIK3R1* gene. This regulatory subunit bears a SH2 domain allowing binding to phosphotyrosine residues of activated TKRs, which induce the activation of the catalytic subunit (Fig. 3.1). In some cases, especially for the growth factors of the insulin family (Chap. 1), p85 activation occurs via an adapter protein, IRS1 or 2 (*insulin receptor substrate 1 or 2*); this protein is phosphorylated by the IGF1R receptor after binding

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_3

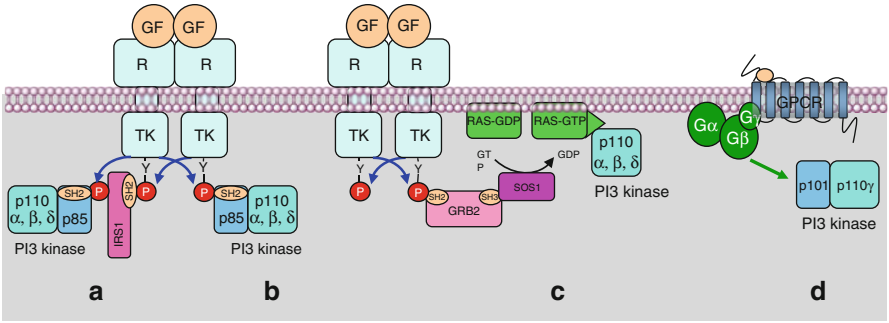


Fig. 3.1 Activation of PI3 kinase by different ways. The regulatory subunit of class I PI3 kinases (p85) (α, β or δ enzymes) can recognise a phosphotyrosine residue of an activated tyrosine kinase receptor (TKR), either directly (**b**) or via an adapter protein such as IRS1 (*insulin receptor substrate 1*) (**a**). The catalytic subunit of PI3 kinase (p110) can also be activated by an activated (GTP-bound) RAS protein (**c**) or by a heterotrimeric large G-protein activated in response to a GPCR (**d**). In that case, this is the γ enzyme that is activated

to phosphotyrosine residues, and the phosphotyrosine residues thus generated on IRSs are recognised by the SH2 domain of p85.

PI3 kinase can also be activated by the RAS protein of the MAP kinase pathway (Chap. 2). The catalytic subunit bears, in its *C*-terminal portion, a binding site for RAS–GTP. This is a major interconnection between the two signalling pathways, and its consequences are important in the field of targeted therapies. Finally, some PI3 kinases can be activated by large heterotrimeric G-protein-coupled receptors (GPCR) (Chap. 6) (Fig. 3.1).

PI3 kinases catalyse the phosphorylation at position 3 of the inositol moiety of a membrane lipid, phosphatidylinositol 4,5-bisphosphate, to generate phosphatidylinositol 3,4,5-trisphosphate (PIP3) (Fig. 3.2). In contrast to the important signalling pathway involving phospholipases Cβ or γ, which generate 1,4,5-trisphosphoinositol (IP3) as a second messenger (Chap. 6), the message resides in the presence of a phosphate group at the position 3 of the intact lipid, because this phosphate group can be recognised and bound by proteins equipped with a particular domain called *pleckstrin homology domain* (PH).

There are indeed three different classes of PI3 kinases, and four members in class I, of interest in this chapter. They are noted PI3Kα, β, γ and δ and have distinct tissue distribution and distinct precise phosphoinositide substrate specificity. Their catalytic subunit is called p110α, β, γ or δ (genes *PIK3CA* to *PIK3CD*), and they are activated by diverse regulatory subunits, p85α, p85β or p55 (genes *PIK3R1* to *PIK3R3*) for PI3Kα, β and δ, whereas distinct regulatory subunits, p101 and p84/87 (genes *PIK3R5* and *PIK3R6*), are required for PI3Kγ. These class I PI3 kinases differ according to their capacity of being activated by TKRs, GPCRs, cytokines, integrins and/or RAS proteins. Class II and III PI3 kinases are much less known, and their functions rather concern membrane traffic and receptor internalisation.

The phosphorylation catalysed by PI3K is counterbalanced by dephosphorylation mediated by a lipid phosphatase, PTEN (*phosphatase and tensin homologue*).

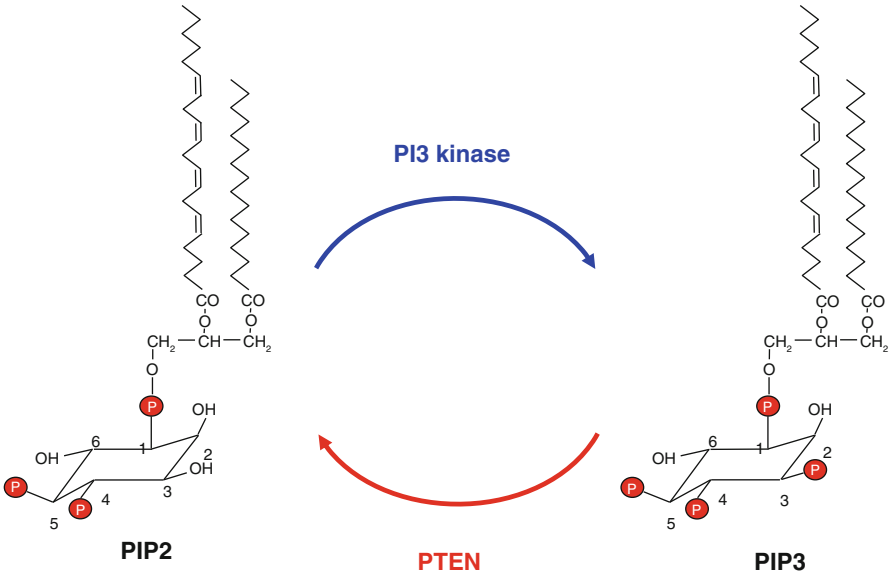


Fig. 3.2 Catalytic activities of PI3 kinase and PTEN. PI3 kinase converts phosphatidylinositol 4,5-bisphosphate (PIP2) into phosphatidylinositol 3,4,5-trisphosphate (PIP3); PTEN catalyses the reverse reaction. Phosphatidylinositol consists of a glycerol skeleton esterifying two fatty acid moieties, generally palmitic acid (C16:0) at position 1 and arachidonic acid (20:4) at position 2 and an inositol (hexahydroxy-cyclohexane) -1-phosphate moiety at position 3

This phosphatase ensures the negative regulation of this signalling pathway. It also intervenes as protein phosphatase in other signalling pathways. Figure 3.3 presents the whole PI3 kinase pathway, from the activated receptor to the ultimate effectors.

The presence of a phosphate moiety at position 3 of an inositol-containing lipid enables to recruit to the membrane PH domain-containing protein kinases, especially PDK1 (*phosphoinositide-dependent kinase 1*, gene *PDPK1*), and the AKT proteins, also known as PKBs (*protein kinase B*). Likely because of a particular arrangement of these two kinases when bound to phosphoinositides, PDK1 can phosphorylate and activate AKT; this occurs on a threonine residue, Thr^{308} (Thr^{309} for AKT2 and AKT3), located in the catalytic centre. Another PH domain-containing kinase is Bruton kinase (BTK), which is specifically expressed in B lymphocytes (Chap. 13).

3.2 AKT Proteins and Their Substrates

There are three AKT proteins, AKT1, 2 and 3, the most common (and the most commonly involved in cancer) being AKT1. They have a catalytic function of serine/threonine kinase, they are activated by phosphorylation and they activate or

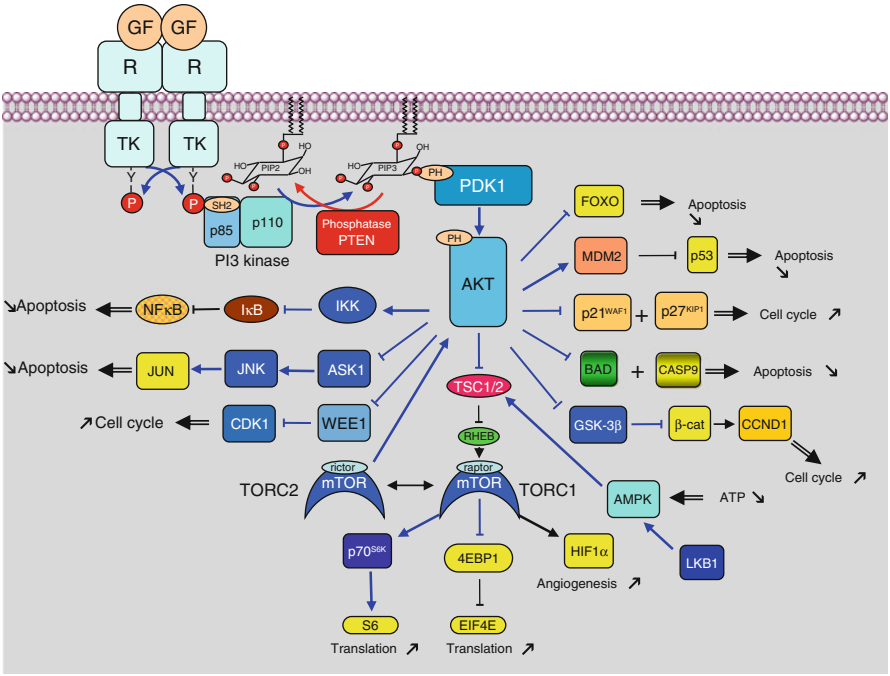


Fig. 3.3 PI3 kinase pathway after activation by a TKR. After activation of a tyrosine kinase receptor (TKR) by a growth factor (GF), the PI3 kinase regulatory subunit (p85) binds a phosphotyrosine residue of the activated receptor via an SH2 domain. The PI3 kinase catalytic subunit (p110) then phosphorylates PIP2 into PIP3, while the PTEN phosphatase catalyses the reverse reaction. The phosphate group at position 3 of the inositol moiety of PIP3 is recognised and bound by serine/threonine kinases equipped with a PH domain, PDK1 and AKT, with the first one phosphorylating the second. AKT is activated by this phosphorylation and can catalyse the phosphorylation of a large array of protein substrates, which are thus activated or inhibited. AKT substrates include TSC2, which is responsible for the control of mTOR activation (which will be detailed in the next figure), BAD, MDM2, $p21^{CIP1}$, $p27^{KIP1}$, GSK3β, IKK, WEE1, ASK1, FOXO3 and others. All these proteins are involved in cell survival and proliferation, through the control of apoptosis, cell cycle entry or protein synthesis

deactivate by phosphorylation a large variety of proteins, including other serine/threonine kinases. Among the main substrates of AKT proteins, and without the possibility of being exhaustive, one can mention (Fig. 3.3):

- The TSC2 (*tuberous sclerosis complex 2*) protein, also known as tuberin. This protein is inhibited by AKT-induced phosphorylation on Ser^{939} and Thr^{1462} residues; once combined with TSC1 (hamartin), the complex constitutes a GTPase-activating protein (GAP) for a small G-protein of the RAS family called RHEB (*RAS homologue enriched in the brain*); this pathway leads to mTOR activation (see below).
- BAD (*BCL2-associated agonist of cell death*), which is a 'BH3-only' protein involved in the positive regulation of mitochondrial apoptosis (Chap. 18); this protein is inactivated by phosphorylation.

- MDM2 (*murine double-minute p53 binding*) protein, activated by phosphorylation, which is the E3 ubiquitin ligase of protein p53 (see Chap. 17); MDM2 activation is thus in opposition to the p53 effects on cell cycle arrest and apoptosis induction.
- $p21^{CIP1}$ (gene *CDKN1A*) and $p27^{KIP1}$ (*CDKN1B*), which are CDK (*cyclin-dependent kinase*) inhibitors and therefore are negative regulators of the advancement of cells in the G1 phase of the cycle (see Chap. 17); AKT-mediated phosphorylation inhibits these CDK inhibitors.
- GSK3β (*glycogen synthase kinase 3β*), which integrates signals of metabolic origin (see below) and which plays an important role in the Wnt pathway (see Chap. 7). This protein is inhibited by AKT-mediated phosphorylation; as a consequence, GSK3β can no longer phosphorylate β-catenin, and this last protein can relocalise in the nucleus and activate transcription programmes leading to cell proliferation and cell cycle entry, instead of being directed to the proteasome for proteolysis.
- IKK (*IκB kinase*), activated by phosphorylation, which phosphorylates and inhibits the IκB (inhibitor of NFκB) protein, an inhibitor of the transcription factor NFκB (Chap. 12); as a consequence, AKT positively regulates NFκB, which is involved in cell survival, proliferation, invasion, angiogenesis and chemoresistance.
- WEE1, inhibited by this phosphorylation, which is a CDK1-inhibiting kinase, a master component of mitosis triggering (Chap. 17);
- ASK1 (*apoptosis signal-regulating kinase 1*), inhibited by AKT-mediated phosphorylation, which is an MAP3K activating the JNK pathway (Chap. 2) involved in apoptosis induction in response to proapoptotic signals and to endothelial reticulum stress (ERS) (Chap. 18 and Annex C).
- FOXO (*forkhead box class O*), already mentioned in Chap. 2, which are apoptosis-inducing transcription factors acting both on the extrinsic pathway (TRAIL, FASL) and on the mitochondrial pathway (BH3-only proteins) (Chap. 18); their phosphorylation by AKT on Thr^{32}, Ser^{253} and Ser^{315} induces their cytoplasmic relocalisation and the loss of their transcriptional functions.

All AKT actions, therefore, drive the cells toward proliferation and survival.

3.3 mTOR Protein and the TORC Complexes

The mTOR protein (*mammalian target of rapamycin*, now *mechanistic target of rapamycin*) was named by analogy with a yeast protein that is inhibited by a natural product, rapamycin, after binding to a small protein, FKBP12 (*FK506-binding protein 12*). This is a serine/threonine kinase operating downstream AKT and is involved in the phosphorylation of a variety of substrates. mTOR, as well as the PI3Ks, belongs to the family of PIKKs (*phosphoinositide 3-kinase-related protein kinases*), like ATM and ATR involved in DNA repair (Chap. 17). A scheme of the various domains of the protein is presented in Fig. 3.4. mTOR is not directly

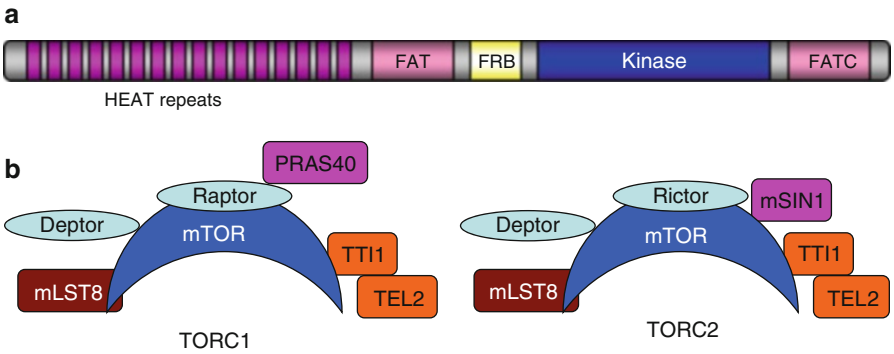


Fig. 3.4 Structural organisation of the mTOR protein and of the TORC complexes. (**a**) Structural organisation of the mTOR protein. The *N*-terminal part of the protein contains a tandem repeat of HEAT (*antiparallel α-helices found in huntingtin, elongation factor 3, PP2A and TOR*) domains, which are involved in protein interactions; an FAT (*FRAP, ATM and TRAPP*) domain; an FRB (*FK506-binding protein 12-rapamycin binding*) domain, which is the binding site for the rapamycin-FKBP12 complex; the classical kinase domain, organised in two lobes as for all kinases; and a second FAT domain in the *C*-terminal part (FATC). (**b**) Protein composition of the two complexes involving mTOR, TORC1 and TORC2. The two complexes share, in addition to mTOR, the proteins mLST8, TTI1 and TEL2 and the protein DEPTOR (*DEP domain-containing mTOR-interacting protein*). They differ by their association with RAPTOR (*regulatory associated protein of TOR*) in TORC1 and RICTOR (*rapamycin-insensitive companion of TOR*) in TORC2 and of the associated proteins, PRAS40 (or AKT1S1) and mSIN1 (or MAPKAP1), which are substrates of AKT and of a MAPK, respectively

activated by AKT: its activation results from the action of the small G-protein RHEB, mentioned earlier, which is active when combined to GTP. The double inhibition of TSC2 by AKT and of RHEB by TSC2 ultimately results in an activation of mTOR by AKT (Fig. 3.5).

In addition to AKT-induced activation, mTOR is able to integrate metabolic and nutritional signals. The 'purpose' of that is the fact that cell proliferation requires a favourable nutritional state for the supply of ATP and amino acids necessary for the synthesis of new cellular components (proteins, nucleic acids, etc.). Hypoxia, hypoglycaemia, increased energy consumption by the muscle, etc., generate an inhibition of mTOR. AMP-dependent kinase (AMPK) constitutes a sensor of the nutritional state of the cell, represented by the AMP–ATP ratio. AMPK activates TSC2 by phosphorylation on residues Thr^{1227} and Ser^{1345}, which accelerates the deactivation of RHEB by stimulation of its GTPase activity. This leads to mTOR inhibition. Nutritional wealth, in contrast, leads to TSC2 inhibition, with opposite consequences.

AMPK has also various metabolic effects which reinforce glucose entry and metabolism in the cell; it also has autophagy-inducing properties, always aiming at energy sparing. AMPK can be activated by various kinases, especially a serine/threonine kinase named LKB1 (*liver kinase B1*) (gene *STK11*). LKB1 is active after combination with two proteins, STRAD (*STE20-related adaptor*) and MO25 (*mouse protein 25*) to form a complex. In addition to its action on AMPK, LKB1

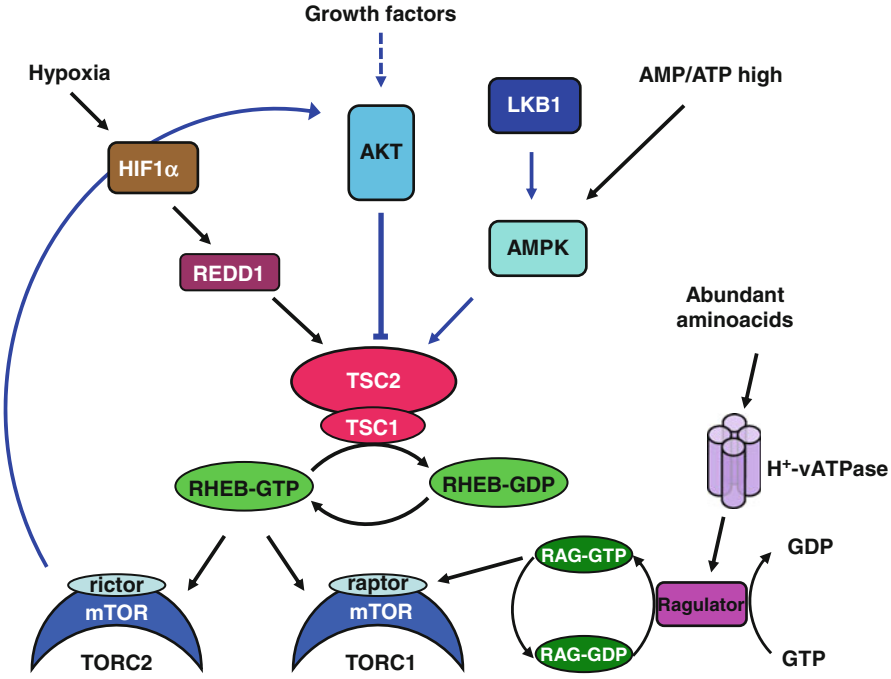


Fig. 3.5 Activation of the mTOR protein. mTOR can be first indirectly activated by AKT. AKT phosphorylates and inactivates the tuberous sclerosis complex (TSC) which consists of the association between TSC2 (tuberin) and TSC1 (hamartin) and which possesses GAP (*GTPase-activating protein*) activity on the small G-protein called RHEB (*RAS homologue enriched in the brain*). RHEB exists under two forms, an inactive one when bound to GDP and an active one when bound to GTP. The GTP-bound form activates mTOR but is rapidly deactivated by its intrinsic GTPase activity, which is stimulated by TSC. In addition to AKT control, TSC is able to integrate metabolic and nutritional signals which thus modulate mTOR activation. The decrease of ATP concentrations activates AMPK (*AMP-dependent kinase*), which in turn activates TSC2 by phosphorylation. In addition, the abundance of amino acids in lysosomes activates a vacuolar proton pump (H^+-ATPase) which activates a protein complex called *ragulator*, equipped with a GEF (*guanyl nucleotide exchange factor*) activity. The ragulator activates the RAPTOR protein via the activation of small G-protein called RAGA and RAGB, through GDP–GTP exchange

phosphorylates and activates various proteins involved in cell polarisation, indicating an unsuspected connection between energy metabolism and epithelial cell polarity.

Another nutritional sensor is the level of free amino acids generated by the lysosomal digestion of proteins. In the lysosome, they induce the activity of a vacuolar H^+-ATPase which activates a protein complex called *ragulator*. This complex displays a guanyl nucleotide exchange (GEF) activity toward small G-proteins called RAGA and RAGB. When bound to GTP (activated state), RAGA and RAGB activate a protein called RAPTOR (*regulatory associated protein of TOR*), which can recruit mTOR at the level of the lysosomal membrane where it can interact with the small G-protein responsible for its activation, RHEB (Fig. 3.5).

mTOR can be engaged in two types of protein complexes (Fig. 3.4):

- The TORC1 complex, in which mTOR is combined with RAPTOR and other proteins. This complex is sensitive to rapamycin and can phosphorylate numerous substrates (see below).
- The TORC2 complex, in which mTOR is combined to a protein called RICTOR (*rapamycin-insensitive companion of TOR*) and other proteins. This complex is insensitive to rapamycin and is able to phosphorylate a small number of protein substrates of the AGC kinase family, including AKT itself, which is activated by phosphorylation on Ser^{473}, at the *C*-terminal level, which constitutes thus a positive feedback loop reinforcing mTOR actions.

At the basal state, the TORC1 complex is combined to a translation initiation factor, EIF4E (*eukaryotic translation initiation factor 4E*), via a repressor protein, 4EBP1 (gene *EIF4EBP1*, *EIF4E-binding protein 1*). In the absence of phosphorylation of 4EBP1, the EIF4E factor cannot be released to recruit the ribosomes required to initiate the translation process of mRNA into proteins. Upon activation, mTOR (TORC1) can phosphorylate 4EBP1, which releases the translation factor EIF4E, which can therefore exert its action on protein synthesis (Fig. 3.6).

mTOR (in the TORC1 complex) is also able to phosphorylate and activate a series of RSK (*ribosomal S6 kinase*) proteins, distinct from the RSK proteins activated by the MAP kinases (Chap. 2), which are known as $p70^{S6K}$ kinases and are encoded by the *RPS6KB1*, *2* and *3* genes. These kinases belong to the AGC family as AKT does; they are involved in the phosphorylation and activation of the S6 ribosomal protein, which is required for ribosome structure and translational activity. Other substrates of the S6 kinases are other kinases such as PDK1, the activator of AKT; the proapoptotic protein BAD, which is inhibited by phosphorylation (Chap. 18); the p53 inhibitor MDM2; the translation elongation factor EEF2 (*eukaryotic translation elongation factor 2*), required in the process of protein synthesis; and the adaptive protein IRS, intermediate between TRKs and PI3 kinase activation (see above), generating thus a negative feedback loop.

mTOR has an important effect on the expression of several proteins, through mechanisms that have not always been elucidated but which may involve the phosphorylation of their transcription factors. mTOR phosphorylates and activates STAT3, a transcription factor downstream cytokine signalling (Chap. 4); phosphorylation occurs on Ser^{727}, in addition to STAT3 phosphorylation by JAK2 on Tyr^{705}. mTOR also activates SREBPs (*sterol regulatory element-binding proteins*) and PPARα and γ (*peroxisome proliferator-activated receptors*), all required for adipogenesis; in addition to protein synthesis, mTOR thus activates also lipid synthesis.

mTOR plays also a role in angiogenesis, through the activation of HIF1α (*hypoxia-induced factor 1α*), a transcription factor involved in neo-angiogenesis signalling (see Chap. 16); conversely, HIF1α is able to indirectly activate TSC2 and is, therefore, a negative regulator of mTOR. Finally, mTOR plays a role in

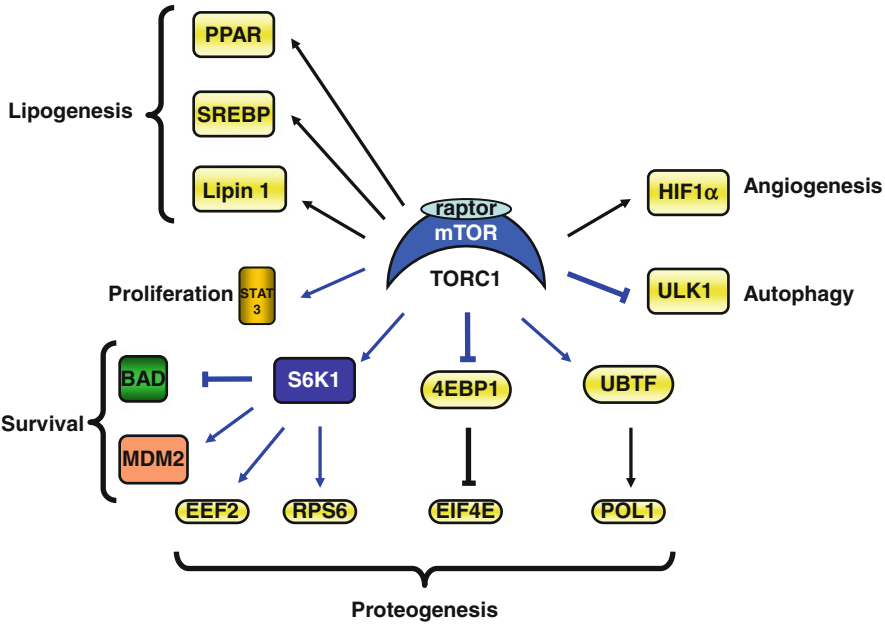


Fig. 3.6 Downstream signalling of mTOR: some of the main targets of mTOR as mediated by the TORC1 complex. mTOR can activate several proteins by phosphorylation, such as the $p70^{S6K}$ kinases and the EIF4E-binding protein 1 (4EBP1), which, once phosphorylated, releases EIF4E (*eukaryotic translation initiation factor 4E*). $p70^{S6K}$ can phosphorylate other substrates such as BAD, MDM2 or EEF2, in addition to the ribosomal S6 protein. mTOR can also indirectly enhance the transcription of several genes, such as HIF1α or STAT3. TORC1 thus regulates a number of cellular processes such as cell survival and proliferation, protein synthesis and lipidogenesis, angiogenesis and autophagy

autophagy, through the inhibitory phosphorylation of proteins involved in the autophagy programme, especially the ULK1–ATG13–FIP200 complex, since ULK1 (*unc-51-like kinase 1*) is a substrate of mTOR.

As a conclusion, all the actions of mTOR in the TORC1 complex converge, directly or indirectly, toward enhancement of protein and lipid synthesis, facilitation of cell proliferation and consequently cell and tissue growth.

3.4 Oncogenic Alterations

The PI3 kinase pathway contains an important number of proto-oncogenes and tumour suppressor genes whose mutational and non-mutational alterations can lead to cancer. PI3 kinase is itself a major oncoprotein; mutations in the *PIK3CA* gene are frequent and belong to the oncogenic landscape of breast and colorectal cancers. The most frequent activating mutations are located on amino acids Glu^{542}, Glu^{545} and His^{1047}. Gene amplification or protein overexpression may also be found in

tumours. The regulatory subunit may also be subjected to activating mutations, mainly found in lymphoproliferative disorders and colorectal cancers.

PTEN phosphatase, which reverses PI3 kinase action, is the product of a tumour suppressor gene of the gatekeeper type, whose alterations are found in endometrial, colorectal and ovarian cancers, as well as in glioblastomas, melanomas, etc. PTEN mutations in the germline are found in a syndrome of cancer predisposition called Cowden disease. PTEN inactivation may occur through a large variety of events: invalidating mutations, deletions, gene underexpression associated to promoter methylation and all negative transcriptional regulation mechanisms.

AKT-activating mutations are known, but remain exceptional and are not commonly found in human cancers. In contrast, AKT gain of function by gene amplification, overexpression and/or hyperphosphorylation is frequently oncogenic. Constitutive AKT phosphorylation on Ser^{473} has been observed in several cancer types (breast, pancreas, endometrium, prostate, etc.).

TSC1 and 2 are thus called after a developmental disease associated with genetic predisposition to cancer and tuberous sclerosis. Germline mutations, followed by loss of heterozygosity in a somatic cell lineage, classically explain the involvement of these proteins in sporadic cancers such as angiosarcomas.

Among the numerous substrates of AKT, some activate pro-oncogenic pathways, such as those leading to NFκB (Chap. 12), and some other anti-oncogenic pathways such as those leading to the FOXO transcription factors, which are tumour suppressors involved in rhabdomyosarcomas and certain leukaemias: chromosome rearrangements induce a loss of function of these proapoptotic factors.

LKB1 was initially identified as the product of a tumour suppressor gene; its germline mutations are accompanied by the formation of hamartomas and polyps, especially in the intestines, through haploinsufficiency. This syndrome of predisposition to cancer is known as Peutz–Jeghers disease. Indeed, the role of this protein in cell polarity explains its tumour suppressor effects, rather than its role in energy metabolism control.

This is only recently that several oncogenic mutations of mTOR were identified. They are accompanied by kinase hyperactivity and, at least for some of them, by an exquisite sensitivity to rapamycin-derived anticancer drugs (see below). Most of them are located in different domains of the *C*-terminal portion of the protein (for instance, E2014K, I2017V, A2020V and E2419K).

3.5 Pharmacological Targets

Most proteins of the PI3 kinase pathway can be considered as potential targets for anticancer pharmacological development. At present time, only rapamycin-derived mTOR inhibitors have reached the market.

Two natural products, quercetin and wortmannin, were used as leads for the elaboration of PI3 kinase inhibitors. The crystallisation of the p110α–p85 complex was helpful for optimising several leads. These molecules have reduced specificity because of the common structure of serine/threonine kinase catalytic centres, especially in the PIKK family, which comprises the various PI3 kinases, mTOR and the proteins involved in DNA damage detection, ATM (*ataxia telangiectasia-mutated*),

ATR (*ataxia telangiectasia and Rad3-related*) and DNA–PK (*DNA-dependent protein kinase*). Several molecules have entered clinical trials. Some compounds, such as GDC-0941, XL-147, BKM-120 (buparlisib) and BAY-806946, are active against the α, γ and δ enzymes and are less active on the β enzyme. Selective compounds against the α enzyme were obtained (BYL-719, GDC-0032) and some against the δ enzyme, which is involved in chronic lymphoid leukaemia (idelalisib or CAL-101). Due to the structural proximity of the various PIKK, several PI3 kinase inhibitors are also active on mTOR (see below).

The PDK1 serine/threonine kinase can also be targeted by inhibitors. UCN-01, a derivative of staurosporine, has been proposed for PDK1 inhibition, but this compound is a weak and non-specific inhibitor. Other compounds are in development: OSU-03012 (AR-12), BX-795, BX-912 and GSK-2334470.

AKT proteins can also be targeted, first by inhibitors of their kinase activity and second by compounds interacting with their PH domain. AKT kinase inhibitors have been developed, and as for all serine/threonine kinase inhibitors, problems of specificity were raised. Most identified compounds are reversible ATP mimetics, such as GSK-690693, MK-2206, GSK-2141795, GDC-0068 and VQD-002. Some compounds have also an activity against the $p70^{S6K}$ proteins, downstream mTOR, which have analogies with AKT. PH domain-interacting compounds are phosphoether lipids such as miltefosine, which has been long ago as a topical drug for cutaneous metastases of breast cancer, and perifosine, which is present in development.

The mTOR protein benefits of a highly specific natural inhibitor, rapamycin, which acts in an allosteric manner and not at the level of the kinase catalytic centre. Rapamycin and its derivatives, called *rapalogues*, are specific to the TORC1 complex and require the binding of rapamycin to a small protein, FKBP12 (*FK506-binding protein 12*). Rapamycin, under the name of sirolimus, has been used for a long time as an immunosuppressor; analogues were preferred for cancer treatment, such as temsirolimus, everolimus and deforolimus. These compounds are used in the treatment of renal advanced cancers and hepatocarcinomas. They are subjected to molecularly driven clinical trials in tumours bearing molecular alterations of the PI3 kinase pathway. An exquisite activity of rapalogues in tumours bearing activating mutations of mTOR has been noticed.

Outside the field of rapalogues, mTOR can be targeted with serine/threonine kinase inhibitors. It should be underlined that these compounds are active on both TORC1 and TORC2 complexes. They can be either specific for mTOR (torin-1 and torin-2, AZD-8055, AZD-2014, OSI-027, CC-223 and INK-128) or with mixed activity against PI3 kinase and mTOR, as already mentioned. None of them has yet been approved for the clinics.

Finally, the $p70^{S6K}$ proteins are also serine/threonine kinases that can be targeted with inhibitors of variable specificity: BI-D1870, H 89, PF-4708671 and AT-7867.

Further Reading

Abraham RT, Gibbons JJ. The mammalian target of rapamycin signaling pathway: twists and turns in the road to cancer therapy. Clin Cancer Res. 2007;13:3109–14.

Bader AG, Kang S, Zhao L, Vogt PK. Oncogenic PI3K deregulates transcription and translation. Nat Rev Cancer. 2005;5:921–9.

Bartholomeusz C, Gonzalez-Angulo AM. Targeting the PI3K signaling pathway in cancer therapy. Expert Opin Ther Targets. 2012;16:121–30.
Burris 3rd HA. Overcoming acquired resistance to anticancer therapy: focus on the PI3K/AKT/mTOR pathway. Cancer Chemother Pharmacol. 2013;71:829–42.
Chalhoub N, Baker SJ. PTEN and the PI3-kinase pathway in cancer. Annu Rev Pathol. 2009;4:127–50.
Courtney KD, Corcoran RB, Engelman JA. The PI3K pathway as drug target in human cancer. J Clin Oncol. 2010;28:1075–83.
Cully M, You H, Levine AJ, Mak TW. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. Nat Rev Cancer. 2006;6:184–92.
Dancey J. MTOR signaling and drug development in cancer. Nat Rev Clin Oncol. 2010; 7:209–19.
Engelman JA, Luo J, Cantley LC. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nat Rev Genet. 2006;7:606–19.
Fu Z, Tindall DJ. FOXOs, cancer and regulation of apoptosis. Oncogene. 2008;27:2312–9.
Grabiner BC, Nardi V, Birsoy K, et al. A diverse array of cancer-associated mTOR mutations are hyperactivating and can predict rapamycin sensitivity. Cancer Discov. 2014;4:554–63.
Guertin DA, Sabatini DM. Defining the role of mTOR in cancer. Cancer Cell. 2007;12:9–22.
Ihle NT, Powis G. Take your PIK: phosphatidylinositol 3-kinase inhibitors race through the clinic and toward cancer therapy. Mol Cancer Ther. 2009;8:1–9.
Inoki K, Corradetti MN, Guan KL. Dysregulation of the TSC-mTOR pathway in human disease. Nat Genet. 2005;37:19–24.
Inoki K, Kim J, Guan KL. AMPK and mTOR in cellular energy homeostasis and drug targets. Annu Rev Pharmacol Toxicol. 2012;52:381–400.
Jewell JL, Russell RC, Guan KL. Amino acid signalling upstream of mTOR. Nat Rev Mol Cell Biol. 2013;14:133–9.
Keniry M, Parsons R. The role of PTEN signaling perturbations in cancer and in targeted therapy. Oncogene. 2008;27:5477–85.
Laplante M, Sabatini DM. Regulation of mTORC1 and its impact on gene expression at a glance. J Cell Sci. 2013;126:1713–9.
Liu P, Cheng H, Roberts TM, Zhao JJ. Targeting the phosphoinositide 3-kinase pathway in cancer. Nat Rev Drug Discov. 2009;8:627–44.
Meric-Bernstam F, Gonzalez-Angulo AM. Targeting the mTOR signaling network for cancer therapy. J Clin Oncol. 2009;27:2278–87.
Rodon J, Dienstmann R, Serra V, Tabernero J. Development of PI3K inhibitors: lessons learned from early clinical trials. Nat Rev Clin Oncol. 2013;10:143–53.
Sengupta S, Peterson TR, Sabatini DM. Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. Mol Cell. 2010;40:310–22.
Sparks CA, Guertin DA. Targeting mTOR: prospects for mTOR complex 2 inhibitors in cancer therapy. Oncogene. 2010;29:3733–44.
Vanhaesebroeck B, Stephens L, Hawkins P. PI3K signalling: the path to discovery and understanding. Nat Rev Mol Cell Biol. 2012;13:195–203.
Yang JY, Hung MC. A new fork for clinical application: targeting forkhead transcription factors in cancer. Clin Cancer Res. 2009;15:752–7.
Yap TA, Garrett MD, Walton MI, Raynaud F, de Bono JS, Workman P. Targeting the PI3K-AKT-mTOR pathway: progress, pitfalls, and promises. Curr Opin Pharmacol. 2008;8:393–412.
Yuan TL, Cantley LC. PI3K pathway alterations in cancer: variations on a theme. Oncogene. 2008;27:5497–510.

Cytokines Pathway

4

Abstract

An important group of signalling molecules operate through a pathway which is slightly different from that described in Chap. 1: their receptors do not present tyrosine kinase activity but are coupled to a cytoplasmic tyrosine kinase which ensures, after receptor activation, the transduction of a signal leading to the transcription of target genes. The main signalling pathway downstream the activation of these receptors by their ligands involves cytoplasmic JAK (*just another kinase*, now *Janus kinase*) proteins and STAT (*signal transducer and activator of transcription*) transcription factors, but the activation of other kinases is sometimes possible, with connections with the MAP kinases (Chap. 2) and the PI3 kinase (Chap. 3) pathways.

Whereas the word 'cytokine' designates often a large variety of signalling proteins, including growth factors (Chap. 1), TNF-related factors and chemokines (Chap. 6), we will adopt a mechanistic definition and restrict this denomination to the ligands of membrane receptors devoid of tyrosine kinase activity but coupled with cytoplasmic tyrosine kinases. This includes most interleukins and interferons, but excludes many signalling molecules that activate other receptors, such as tyrosine kinase receptors, death receptors, interleukin-1 and toll-like receptors, etc. The cytokines envisaged in this chapter may be involved in multiple cell processes: proliferation and differentiation of blood cells, immunity and inflammation, body weight and size control, antiviral defences, etc.

4.1 Cytokines

About 40 cytokines operating through the activation of a cytoplasmic tyrosine kinase have been identified and can be distributed in two types according to their general structure and that of their receptors (Table 4.1). Type I, characterised by the presence of four α helixes, can be roughly subdivided in several families:

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_4

Table 4.1 Main cytokines, cytokine receptors and downstream corresponding effectors

Type	Cytokine	Specific receptor	Family-specific receptor	Kinase	Transcription factor
Type I	IL2	IL2Rα/β	IL2Rγc (*IL2RG*)	JAK1 + JAK3	STAT5
	IL7	IL7Rα			
	IL9	IL9Rα			
	IL15	IL15Rα			
	IL4	IL4Rα			STAT6
	IL21	IL21Rα			STAT1, 3
	GMCSF (*CSF2*)	GMCSFRα (*CSF2RA*)	IL3Rβc (*CSF2RB*)	JAK2	STAT5
	IL3	IL3Rα			
	IL5	IL5Rα			
	IL6	IL6Rα	gp130 (*IL6ST*)	JAK1	STAT1, 3, 5
	IL11	IL11Rα/β			
	IL25	IL17RB			
	CNTF	CNTFRα			
	LIF	LIFR			
	OSM	OSMR			
	IL27 (IL30)	IL27Rα			
	IL12	IL12Rβ2	IL12Rβ1	JAK2 + TYK2	STAT4
	IL23	IL23Rα			STAT3
	GH	GHR		JAK2	STAT5
	PRL	PLR			
	OBS (*LEP*)	OBR (*LEPR*)			
	EPO	EPOR			
	TPO (*THPO*)	TPOR (*MPL*)			
	GCSF (*CSF3*)	GCSFR (*CSF3R*)			
Type II	IFNα, β, ω, ε	IFNαR1	IFNαR2	JAK1 + TYK2	STAT1, 2, 3, 5
	IFNγ	IFNγR1	IFNγR2	JAK1 + JAK2	STAT1, 3, 5
	IL10	IL10Rα	IL10Rβ	JAK1 + TYK2	
	IL22	IL22Rα1/2			
	IL26	IL20Rα			
	IL19, 20	IL20Rα	IL20Rβ		
	IL24	IL22Rα			

- Interleukin-2 (IL2) family: IL2, IL4, IL7, IL9, IL15 and IL21.
- Interleukin-6 (IL6) family: IL6, IL11, IL27, LIF (*leukaemia inhibiting factor*), CNTF (*ciliary neurotrophic factor*) and OSM (*oncostatin M*).
- Interleukin-3 (IL3) family: IL3, IL5 and GMCSF (*granulocyte-macrophage colony-stimulating factor receptor*).

- Interleukin-12 (IL12) family: IL12 and IL23.
- Hormones and other haematopoietic growth factors family: GCSF (*granulocyte colony-stimulating factor*), EPO (erythropoietin), TPO (thrombopoietin, gene *THPO*), GH (*growth hormone*), prolactin (PRL) and leptin (LEP or OBS).

Type II contains interferons and other interleukin subfamilies: IL10 family (IL10, IL22, IL26) and IL20 family (IL19, IL20, IL24).

The cytokines each have many functions and their physiological actions are not limited to a unique tissue but are often ubiquitous and pleiotropic. These functions are also often redundant, so that several cytokines may exert the same effect of some target tissues or cells. The precise role of all cytokines will not be presented in detail, and the reader is referred to immunology and haematology handbooks. We will essentially focus this chapter on haematopoietic growth factors because of their role in cell proliferation and haematological malignancies and on immunity- and inflammation-related processes because of their potential role in oncogenesis.

4.2 Cytokine Receptors and JAK Activation

Cytokine receptors are membrane proteins with a single transmembrane domain and an extracellular *N*-terminal portion. They are characterised by the presence of cysteine residues in the distal part of the extracellular domain and by a motif containing two tryptophan and two serine residues (WSXWS) on the proximal part of this domain. The intracellular domain does not contain a catalytic centre but several well-conserved domains, in particular box 1, close to the transmembrane domain, characterised by a proline-rich motif of eight amino acids, and box 2, consisting of a group of hydrophobic amino acid residues in the prolongation of the α helix of the transmembrane domain, followed by a series of charged amino acid and tyrosine residues able to bind a phosphate moiety. The two boxes are involved in receptor binding to a JAK protein. One of these receptors, CNTFR, has no transmembrane and intracellular domains and is anchored to the plasma membrane through a molecule of glycosylphosphatidylinositol (Annex C). Several receptors present soluble variant isoforms generated by alternative splicing: this is the case for the receptors of IL4, IL5, IL7, IL9, LIF, GCSF, CNTF, GH and PRL. These soluble forms behave as cytokine transporters and some of them merely as traps, which bind a cytokine without transmitting a signal.

Cytokine receptors must be dimerised to activate the signalling pathway. The dimers and their ligand constitutes the ‘reception complex’ characteristic of the message to be transmitted. This can be a heterodimerisation between a ligand-specific monomer (α type) and a common, ligand family-specific, monomer (β or γ type). The most frequently used ligand family-specific receptors are IL3Rβ or βc for the GMCSF family, IL2Rγ or γc for the IL2 family and gp130 or IL6ST (*IL6 signal transducer*) for the IL6 family (Table 4.1). Heterodimerisation also occurs for the receptors of type II cytokines of the IL10 and IL20 families, with IL10Rβ and IL20Rβ as common receptors for these interleukins, respectively. In contrast, for the

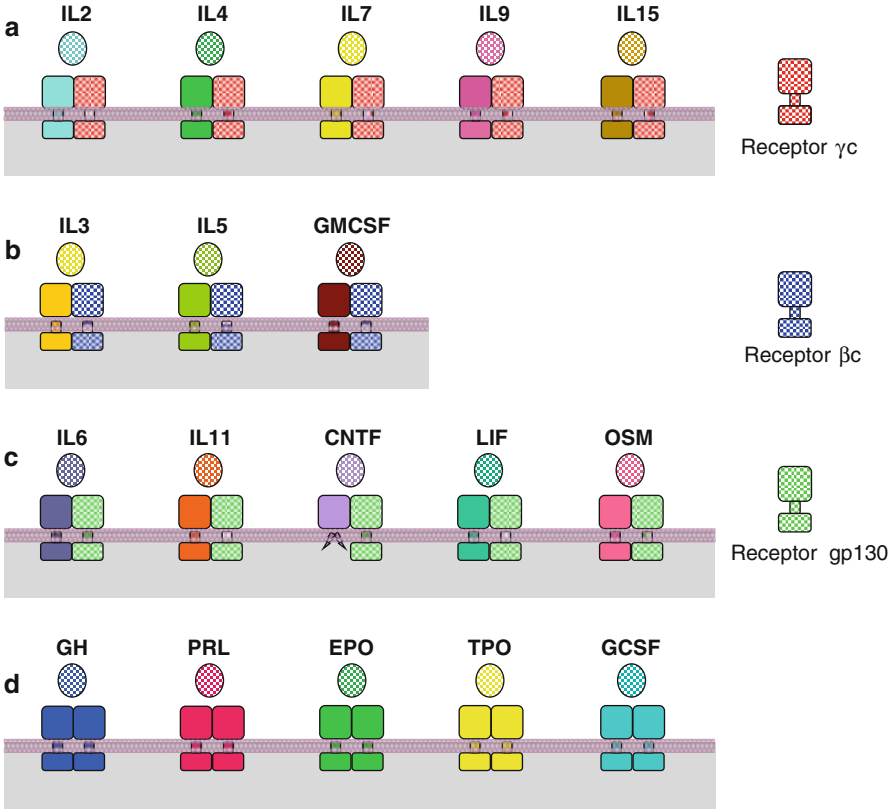


Fig. 4.1 Signal reception complexes of the main type I cytokines. The receptor dimers formed after recognition and binding of a cytokine are diverse. (**a**) Dimerisation occurs between an α receptor (specific to the cytokine) and a γc receptor (IL2Rγc, common to the IL2 cytokine family). (**b**) Dimerisation occurs between an α receptor (specific to the cytokine) and a βc receptor (IL3Rβc, common to the IL3 cytokine family). (**c**) Dimerisation occurs between an α receptor (specific to the cytokine) and a gp130 receptor (IL6ST) (common to the IL6 cytokine family). (**d**) Dimerisation occurs between two identical cytokine-specific receptors

hormone (GH, PRL, etc.) receptors and the other haematopoietic cytokine (EPO, TPO) receptors, there exists only one type of receptors that is homodimerised upon ligand binding. A total of 36 reception complexes have been thus identified and Fig. 4.1 presents some of them.

Receptor homodimerisation occurs sequentially: the cytokine binds to a first monomer, which induces a conformation change allowing the binding of a second monomer and the interaction between the two monomers. Receptor heterodimerisation is more complex and may follow various types of association between the ligand-specific monomer and the common, family-specific, monomer, βc, γc or gp130.

Receptors are constitutively associated to JAK kinases. Receptor dimerisation allows the activation of these kinases, which undergo autophosphorylation

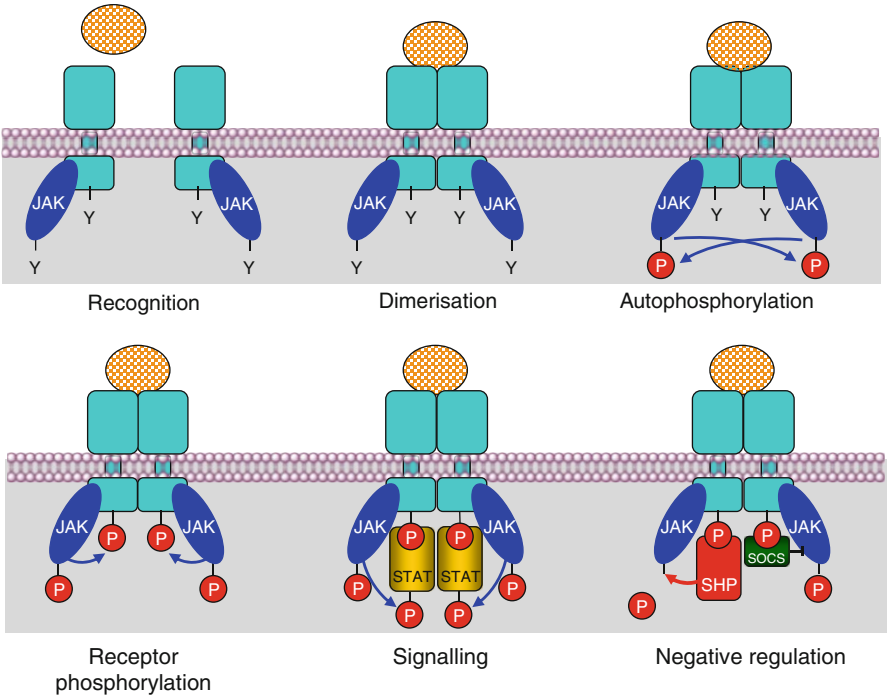


Fig. 4.2 Cytokine receptor activation via the JAK kinases. JAK proteins are constitutively associated to the receptor complexes. The binding of a cytokine enables receptor dimerisation and autophosphorylation of the associated JAK proteins. These kinases then phosphorylate the receptor on *C*-terminal tyrosine residues, which are thus recognised by diverse proteins, especially the STAT transcription factors, which are phosphorylated by the JAK kinases. The negative regulation of the signalling pathway is ensured in particular by SOCS proteins, which compete with STAT proteins for receptor binding on phosphotyrosine residues and inhibit the JAK proteins, and by SHP phosphatases which bind the receptor complexes and dephosphorylate the JAK kinases

and then become able to phosphorylate the receptor at the level of tyrosine residues of box 2 (Fig. 4.2). Four JAK proteins have been identified, called JAK1, JAK2, JAK3 and TYK2. The association between a JAK protein and a receptor occurs as described in Table 4.1. Some receptors are associated to two identical JAK proteins, JAK1 or JAK2, while other receptors are associated to distinct JAK proteins: JAK1/JAK2, JAK1/JAK3, JAK1/TYK2 or JAK2/TYK2. No receptor associated to two JAK3 or two TYK2 proteins, or to JAK2/JAK3 or JAK3/TYK2 combinations, has been identified. After autophosphorylation and receptor phosphorylation, the JAK proteins can phosphorylate a series of transcription factors of the STAT family, which are recruited to the membrane thanks to their SH2 recognition domains of the phosphotyrosine residues of the activated receptor.

The JAK proteins are characterised by the presence of several homologous domains JH (JAK homology domains) (Fig. 4.3): a JH1 *C*-terminal domain, bearing

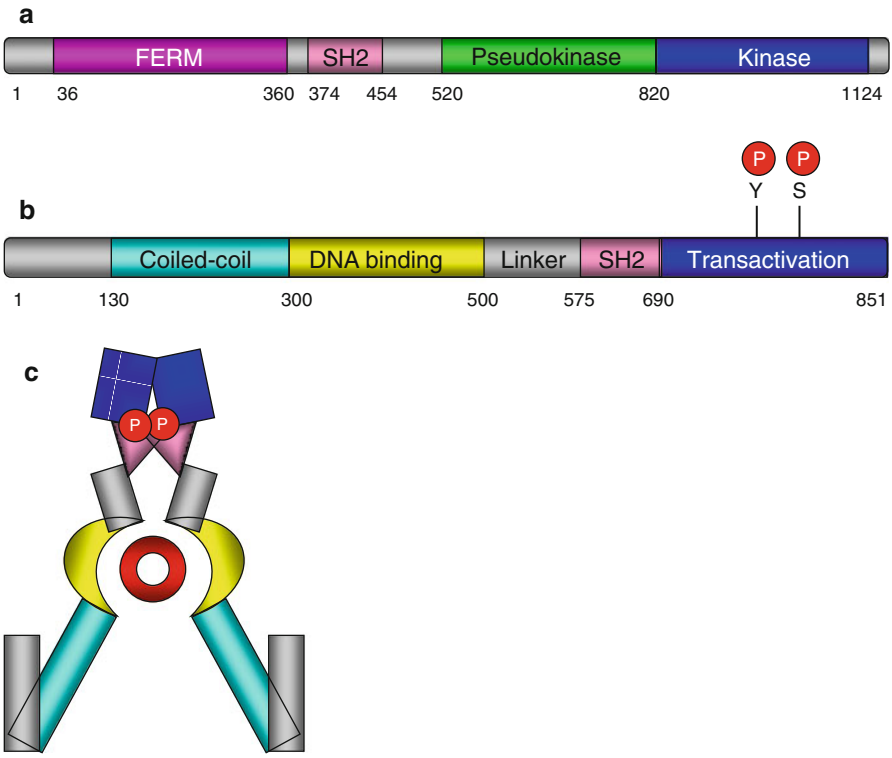


Fig. 4.3 General structure of the JAK and STAT proteins. (**a**) JAK proteins contain a kinase domain, a pseudokinase domain, a SH2 domain for binding to the receptor phosphotyrosine residues and a FERM *N*-terminal domain for receptor interaction. (**b**) STAT proteins contain a transactivation domain bearing the tyrosine residue that has been phosphorylated by JAK and a regulatory serine residue, a SH2 domain for binding the phosphotyrosine residues of the activated receptor, a DNA-binding domain and an *N*-terminal coiled-coil domain. (**c**) Schematic representation of a STAT protein after dimerisation, from crystallography data; the DNA-binding domains enclose a DNA sequence (*red*); the SH2 domains of each monomer are associated to a phosphotyrosine residue of the other monomer

the tyrosine kinase activity; a JH2 pseudokinase domain able to negatively regulate kinase activity; and a FERM (*four-point-one, ezrin, radixin and moesin*) *N*-terminal domain involved in receptor binding.

The different JAK proteins are redundant but may play specific roles. JAK1, JAK2 and TYK2 are ubiquitous, while JAK3 is specifically expressed in the haematopoietic tissues and is activated in response to cytokines of the IL2 family. JAK1 is activated in response to numerous cytokines, especially of the IL2, IL6 and type II cytokines. JAK2 is characteristic for the response to haematopoietic factors (EPO) and hormones. TYK2 is mainly involved in type II cytokines response.

4.3 Signal Transduction

The JAK kinases ensure the phosphorylation of STAT transcription factors. Seven STAT proteins have been identified, denominated STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B and STAT6. STAT expression is ubiquitous; all have a well-conserved SH2 domain allowing their recruitment by the activated receptor on its phosphotyrosine residues and a *C*-terminal tyrosine residue which is the site of phosphorylation by JAK (Fig. 4.3). This induces the formation of phosphorylated STAT protein homodimers or heterodimers, thanks to reciprocal interactions between the SH2 domain of the first one and the phosphotyrosine residue of the second one. These dimers can then migrate to the nucleus where they bind the promoter sequences of target genes, thanks to their DNA-binding domain (Fig. 4.4). One or several distinct STAT proteins can be activated in response to the binding of a cytokine ligand to its receptor, depending on the cellular context (tissue type, differentiation state, ligand concentration).

The physiological actions of the STAT proteins are not redundant: STAT1 is rather specific of the IFNs; STAT3 is preferentially activated by IL6 family

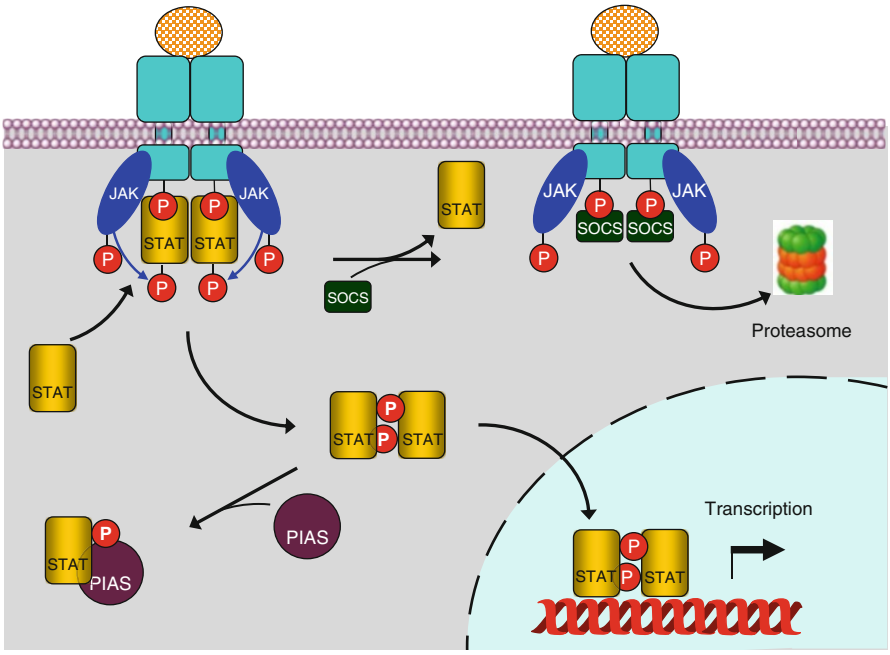


Fig. 4.4 The JAK–STAT signalling pathway. After phosphorylation, STAT dimers migrate to the nucleus, where they recognise the promoters of target genes and activate their transcription. PIAS proteins can however bind STAT molecules and prevent their dimerisation. SOCS proteins inhibit signal transmission by binding the receptor to prevent binding and phosphorylation of STAT proteins at these sites; they inhibit JAK proteins and drive the complex to proteasomal degradation

cytokines; STAT4 is activated by IL12, STAT6 by IL4 and STAT5 by IL2 and IL3 families interleukins, hormones and haematopoietic growth factors.

One of the main regulation processes of cytokine signal transduction is operated by SOCS (*suppressor of cytokine signalling*) proteins, also known as SSI (*STAT-induced STAT inhibitor*) or JAB (*JAK-binding protein*) proteins. There are eight SOCS proteins, denominated SOCS1 to SOCS7 and CIS (*cytokine-inducible SH2 domain-containing protein*). These proteins bear a SH2 domain, which can bind the receptor phosphotyrosine residues and prevent STAT binding to the same activated receptor, and a JAK inhibition domain called KIR (*kinase-inhibitory region*); they also have a SOCS box which recruits E3 ubiquitin ligases (see Annex C), allowing the destruction of the reception complexes and maybe of the STAT proteins themselves. The SOCS genes are target genes for the STAT transcription factors, which allow negative retrocontrol of this signalling pathway. Another negative regulation is ensured by PIAS (*protein inhibitor of activated STAT*) proteins, which are able to bind the phosphorylated STAT proteins, thus preventing their dimerisation. Finally, SHP (*SH2 domain-containing*) phosphatases or PTPN (*protein tyrosine phosphatases, non-receptor*), equipped with SH2 domains, are able to dephosphorylate the JAK kinases.

The target genes of the STAT transcription factors are numerous, but they remain incompletely identified. STAT3 is involved in cell survival and proliferation, through the transcriptional activation of genes such as *BCL2* (Chap. 18), *CCND1* (cyclin D1, Chap. 17), *VEGFA* (Chap. 1), *MYC*, *JUN* and *FOS* (Chap. 2). In contrast, STAT3 is a repressor of *TP53* transcription (Chap. 17). It may seem paradoxical to observe that there is a multiplicity of signals able to activate the JAK–STAT pathway, as well as a large variety of ligand–receptor association complexes, whereas these various messages lead to the activation of a small number of distinct STAT factors.

There are numerous interconnections between the JAK–STAT pathway and the other signalling pathways. The JAK kinases are equipped with a SH2 domain and thus are able to recognise and bind phosphotyrosine residues of the tyrosine kinase receptors (TKRs) described in Chap. 1: they can thus activate a STAT protein in response to growth factor signalling and so can do the STAT proteins themselves. In addition, the JAK proteins are able to activate the pathways that are classically located downstream TRK activation, such as the MAP kinase (Chap. 2) and the PI3 kinase (Chap. 3) pathways.

4.4 Oncogenic Alterations

Numerous oncogenic alterations have been identified in this signalling pathway; many of them are related to inflammation and immunity: immune deficiencies, inflammatory and autoimmune diseases. For instance, there is an association between *JAK3* germline mutations and severe combined immunodeficiency (SCID) syndromes, in mice as in humans. Also, germline alterations of the leptin receptor or the growth hormone receptor generate various congenital hereditary endocrine diseases. We will not describe the immune or endocrine consequences of the

pathological alterations found in this pathway and focus only on the oncogenic alterations; these can be found at all the steps of this signalling pathway: haematopoietic growth factor receptors, JAK kinases, STAT transcription factors and their SOCS inhibitors.

The IL3Rα receptor is overexpressed in blast cells in more that 80 % of acute myeloid leukaemias, inducing an increase in STAT5 activation. In other cases, the common receptor IL3Rβ is expressed as a truncated form, which renders it insensitive to ubiquitinylation and subsequent proteolysis. Similarly, a truncated form of the GCSF receptor (GCSFR) is associated to STAT5 overstimulation, and point mutations in this receptor are also found in acute myeloid leukaemias. Mutations in the thrombopoietin receptor (TPOR) have been found in myeloproliferative neoplasias.

The discovery of the V617F JAK2 mutations in several myeloproliferative neoplasias has represented an important step in the knowledge of the genesis of these malignant haematological diseases (*polycythaemia vera* or Vaquez disease, essential thrombocythaemia, idiopathic myelofibrosis). This mutation occurs in the JH2 domain of JAK2 and is responsible for a constitutive activation of its tyrosine kinase activity, leading to receptor hypersensitivity. Other molecular alterations of JAK2, such as K539L, have been described in V617E-negative myeloproliferative neoplasia, as well as in some acute lymphoblastic leukaemias and acute megakaryocytic leukaemias. Also, activating mutations of JAK1 have been found in acute T-cell lymphoblastic leukaemias and JAK3 mutations in acute megakaryocytic leukaemias.

Although no mutation of STAT genes have ever been found in malignant diseases, these genes often behave as proto-oncogenes. Constitutive activation of STAT1 has been observed in acute myeloid leukaemias, B-cell acute lymphoblastic leukaemias and erythroleukaemias; that of STAT3 in Hodgkin disease and some acute myeloid leukaemias as well as in several solid tumours (prostate and breast cancer, hepatocellular carcinomas); and that of STAT5 in erythroleukaemias, acute myeloid leukaemias, acute lymphoblastic leukaemias, acute megakaryocytic leukaemias and chronic myeloid leukaemias. However, the STAT proteins can play an ambiguous role: STAT1 is a promoter of apoptosis in interferon signalling and could behave as a tumour suppressor gene in late stages of oncogenesis by inducing apoptosis. Similarly, STAT3 is activated by IL10, which is an anti-inflammatory cytokine and could reduce inflammation-related oncogenetic processes.

The SOCS proteins behave as tumour suppressors; SOCS1 promoter methylation, leading to gene inactivation, has been observed in nearly 60 % of acute myeloid leukaemias, as well as in chronic myeloid leukaemias and in hepatocarcinomas. Similarly, SOCS3 promoter methylation is found in hepatocarcinomas as well as in various squamous cell carcinomas.

4.5 Pharmacological Targets

Since 1992, recombinant IL2 (aldesleukin) has been introduced in the treatment of advanced renal cancers and malignant melanomas, two cancers that are highly resistant to chemotherapy and appear to be associated to immune deficiency. IL2 is

physiologically produced by $CD4^+$ T lymphocytes. It induces an increase of lymphocyte proliferation and the clonal expansion of natural killer cells, which constitutively express IL2 receptors. IL2 also stimulates the production of other factors such as tumour necrosis factor (TNF), interleukin 1 (IL1) and interferon γ. Conversely, an anti-IL2 approach could reveal interest in the treatment of the cancers for which IL2 may play a pathogenic role: T-cell leukaemias and lymphomas, Hodgkin disease, follicular B-cell lymphomas and hairy cell leukaemias.

Interferon α is also used in therapeutics because of its stimulatory activity of natural killer cells and antigen-presenting cells. Outside the field of B and C hepatitis, it is prescribed in chronic myeloid leukaemias, hairy cell leukaemias, Kaposi sarcoma and metastatic malignant melanomas. Other interferons (interferons β and γ) are utilised outside the field of oncology.

The haematological toxicity of high-dose chemotherapy has encouraged the development of therapeutic use of haematopoietic growth factors to accelerate the production of progenitor cells. Erythropoietin for the treatment of anaemia, thrombopoietin for the treatment of major platelet deficiencies and GCSF and GMCSF for the prevention and treatment of neutropenia have been thus introduced in the oncological armamentarium.

Targeting JAK2 activity by tyrosine kinase inhibitors (TKI) lies on the various successes that have been obtained in targeting other tyrosine kinase receptors (EGFR) or cytoplasmic tyrosine kinases (mainly BCR–ABL). Ruxolitinib has been approved for the treatment of myeloproliferative neoplasia, without taking into account the presence of the JAK2 V617F mutation. Other compounds are in development, but the problem of substrate specificity is certainly one of the most difficult to solve. Several JAK3-targeting TKIs are also in development for the treatment of rheumatoid arthritis, autoimmune diseases and transplant rejection.

Targeting STAT transcription factors, especially STAT3 and STAT5, is under study but remains a difficult task. Peptidomimetics targeting the STAT SH2 domains, STAT mRNA antisense oligonucleotides inhibiting STAT synthesis, decoy oligonucleotides that mimic the DNA-binding site of STAT transcription factors and small molecules identified by screening represent the various tracks that can be followed. All of them remain far from current therapeutic usage.

Further Reading

Benekli M, Baumann H, Wetzler M. Targeting signal transducer and activator of transcription signaling pathway in leukemias. J Clin Oncol. 2009;27:4422–32.

Constantinescu SN, Girardot M, Pecquet C. Mining for JAK–STAT mutations in cancer. Trends Biochem Sci. 2007;33:122–31.

Gadina M, Hilton D, Johnston JA, Morinobu A, Lighvani A, Zhou YJ, Visconti R, O'Shea JJ. Signaling by type I and II cytokine receptors: ten years after. Curr Opin Immunol. 2001;13:363–73.

Germain D, Frank DA. Targeting the cytoplasmic and nuclear functions of signal transducers and activators of transcription 3 for cancer therapy. Clin Cancer Res. 2007;13:5665–9.

Guo Y, Xu F, Lu T, Duan Z, Zhang Z. Interleukin-6 signaling pathway in targeted therapy for cancer. Cancer Treat Rev. 2012;38:904–10.

Haan C, Kreis S, Margue C, Behrmann I. Jaks and cytokine receptors – an intimate relationship. Biochem Pharmacol. 2006;72:1538–46.
Ihle JN. Cytokine receptor signalling. Nature. 1995;377:591–4.
Ihle JN, Witthuhn BA, Quelle FW, Yamamoto K, Silvennoinen O. Signaling through the hematopoietic cytokine receptors. Annu Rev Immunol. 1995;13:369–98.
Kishimoto T, Taga T, Akira S. Cytokine signal transduction. Cell. 1994;76:253–62.
Miklossy G, Hilliard TS, Turkson J. Therapeutic modulators of STAT signalling for human diseases. Nat Rev Drug Discov. 2013;12:611–29.
Moutoussamy S, Kelly PA, Finidori J. Growth-hormone-receptor and cytokine-receptor-family signaling. Eur J Biochem. 1998;255:1–11.
Mullally A, Ebert BL. Janus reveals another face: the biologic rationale for targeting Janus kinase 2 in lymphoma. J Clin Oncol. 2012;30:4168–70.
Murray PJ. The JAK-STAT signaling pathway: input and output integration. J Immunol. 2007;178:2623–9.
O'Shea JJ, Gadina M, Schreiber RD. Cytokine signaling in 2002: new surprises in the Jak/Stat pathway. Cell. 2002;109:S121–31.
O'Sullivan LA, Liongue C, Lewis RS, Stephenson SEM, Ward AC. Cytokine receptor signaling through the Jak–Stat–Socs pathway in disease. Mol Immunol. 2007;44:2497–506.
Pesu M, Laurence A, Kishore N, Zwillich SH, Chan G, O'Shea JJ. Therapeutic targeting of Janus kinases. Immunol Rev. 2008;223:132–42.
Quintás-Cardama A, Verstovsek S. Molecular pathways: Jak/STAT pathway: mutations, inhibitors, and resistance. Clin Cancer Res. 2013;19:1933–40.
Renauld JC. Class II, cytokine receptors and their ligands: key antiviral and inflammatory modulators. Nat Rev Immunol. 2003;3:667–76.
Santos CI, Costa-Pereira AP. Signal transducers and activators of transcription-from cytokine signalling to cancer biology. Biochim Biophys Acta. 1816;2011:38–49.
Taniguchi T. Cytokine signaling through nonreceptor protein tyrosine kinases. Science. 1995;268:251–5.
Tayal V, Kalra BS. Cytokines and anti-cytokines as therapeutics – an update. Eur J Pharmacol. 2008;579:1–12.
Yoshimura A. Signal transduction of inflammatory cytokines and tumor development. Cancer Sci. 2006;97:439–47.
Zhang J, Li H, Yu JP, Wang SE, Ren XB. Role of SOCS1 in tumor progression and therapeutic application. Int J Cancer. 2012;130:1971–80.

TGFβ Pathway

5

Abstract

Transforming growth factor beta (TGFβ) is the lead of a large family of factors studied in this chapter, which comprise several stricto sensu TGFβs and a subfamily of related factors, and a subfamily of bone morphogenetic factors (BMPs). The receptors of these factors display serine/threonine kinase activity, which distinguishes them from those presenting (or are associated to) tyrosine kinase activity (Chaps. 1 and 4). They are involved in many cellular processes, especially during development, and play a role in cell survival, proliferation and differentiation, as well as in cell adhesion and motility, and its alterations are found in cancer, inflammation and fibrosis. Overall, the TGFβ pathway generally works in opposition to cell proliferation and plays a tumour suppressor role counteracting oncogenesis, but the role of this pathway is ambivalent since it frequently enhances cell motility and migration, participates to the epithelial-to-mesenchymal transition and, consequently, may favour metastasis dissemination.

The factors of the TGFβ family are active as homodimers and are responsible of the formation, at the level of the target cell, of a hexameric complex made of two molecules of ligand, two molecules of type I receptors and two molecules of type II receptors, in addition to coreceptors required for signalling. Once activated by ligand binding, type II receptors phosphorylate type I receptors, which in turn phosphorylate transcription factors of the SMAD family. This denomination comes from the names of the ortholog genes *sma* (small) of *Caenorhabditis elegans* and *mad* (*mothers against decapentaplegic homologue*) of drosophila. The SMAD factors bind to a common interactant, also of the SMAD family, and form a transcriptional activator or repressor complex which is able to relocate to the nucleus and regulate the transcription of target genes, which are more or less specific of the extracellular factor generating the signal.

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_5

5.1 TGFβ Family Ligands

Table 5.1 inventories the main factors of the lato sensu TGFβ family, whose total number exceeds 30. These factors and their receptors may have very variable tissue distribution and are expressed at definite periods of life: some of them are expressed

Table 5.1 Ligands of the TGFβ family (lato sensu) and their receptors

	Ligand	Type I receptor	Type II receptor	Coreceptor	SMAD factor
TGFβ subfamily	TGFβ 1, 2, 3	TβR1, ALK5 (*TGFBR1*)	TβR2 (*TGFBR2*)	β-glycan (*TGFBR3*)	SMAD2, SMAD3
	TGFB1, *2*, *3*				
	Activins A, B, C (*INHBA*, *B*, *C*)	ALK2 (*ACVR1A*)	*ACVR2A*		
		ALK4 (*ACVR1B*)	*ACVR2B*		
	Myostatin (*MSTN*, GDF8)	ALK4 (*ACVR1B*)	*ACVR2A*		
			ACVR2B		
	Nodal (*NODAL*)	ALK4 (*ACVR1B*)	*ACVR2A*	Cripto (*CFC1, TDGF1*)	
		ALK7 (*ACVR1C*)	*ACVR2B*		
	GDF 1, 3 (*GDF1*, *3*)	ALK4 (*ACVR1B*)	*ACVR2A*	Cripto (*CFC1, TDGF1*)	
		ALK7 (*ACVR1C*)	*ACVR2B*		
	Inhibin (*INHA*)	ALK4 (*ACVR1B*)	*ACVR2A*	β-glycan (*TGFBR3*)	
			ACVR2B		
	Lefty 1, 2 (*LEFTY1*, *2*)	ALK4 (*ACVR1B*)	*ACVR2A*		
		ALK7 (*ACVR1C*)	*ACVR2B*		
BMP subfamily	BMP 2, 4 (*BMP2*, *4*)	ALK3 (*BMPR1A*)	*BMPR2*	*RGMA*	SMAD1, SMAD5, SMAD8
	GDF 5, 6, 7 (*GDF5*, *6*, *7*)	ALK6 (*BMPR1B*)	*ACVR2A*	*RGMB*	
			ACVR2B		
	BMP 6, 7 (*BMP6*, *7*)	ALK2 (*ACVR1A*)	*BMPR2*	*RGMA*	
		ALK3 (*BMPR1A*)	*ACVR2A*	*RGMB*	
		ALK6 (*BMPR1B*)	*ACVR2B*		
	BMP 9, 10 (*BMP9*, *10*)	ALK1 (*ACVRL1*)	*BMPR2*	Endoglin (*ENG*)	
	Anti-Mullerian hormone (*AMH*)	ALK3 (*BMPR1A*)	*AMHR2*		
		ALK2 (*ACVR1A*)			

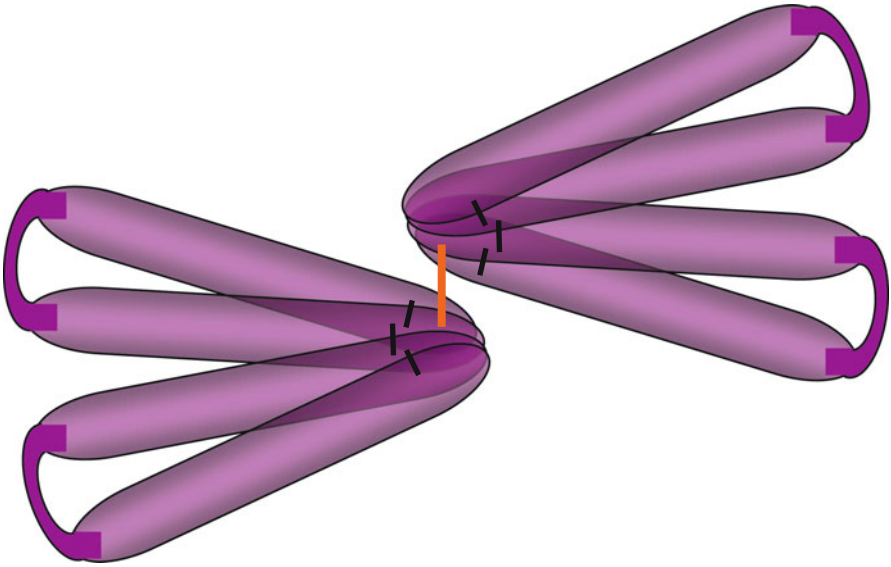


Fig. 5.1 Schematic representation of a TGFβ dimer. TGFβ structure reveals the presence of four 'fingers' linked by intrachain disulphide bridges (*black lines*) at the level of the 'cysteine knot'. Two TGFβ molecules are bound by an interchain disulphide bridge (*orange line*)

only in some cell types and/or during a very short period of embryonic life, while others are ubiquitous and/or expressed all life long. In the TGFβ subfamily are found three TGFβs stricto sensu; a series of proteins called inhibins or activins, nodal, lefty and myostatin; as well as GDF (growth differentiation factors) 1 and 3. In the BMP subfamily are found BMP 2, 4, 5, 6, 7, 9 and 10; GDF 5, 6 and 7; and AMH (*anti-Mullerian hormone*) also known as MIS (*Mullerian inhibitory substance*).

These factors are synthesised as precursor latent forms, which are later submitted to proteolytic cleavage releasing the active *C*-terminal part of the protein. The proteolytic enzymes in charge of this activation are convertases of the subtilisin family, such as furin, and collectively called PCSKs (*proprotein convertases subtilisin/kexin*) or PACEs (*paired basic amino acid cleaving enzymes*) (see Annex C). After cleavage, the three TGFβs remain non-covalently attached to the *N*-terminal peptide, called the *latency-associated peptide*, which is not the case for other factors of the TGFβ and BMP subfamilies. All TGFβs and BMPs are secreted as dimers and are characterised by the presence of a characteristic cysteine knot at the *C*-terminal level, which allows the formation of three intramolecular disulphide bridges; a last disulphide bridge is used for dimerisation (Fig. 5.1).

All these factors are characterised by an antiparallel dimeric structure, which is formed before receptor interaction, allowing the recruitment of two molecules of type II receptors and of two molecules of type I receptor (Fig. 5.2). The ligands of the TGFβ subfamily have higher affinity for type II receptors, and this is only afterwards that two molecules of type I receptor are recruited and bound to the

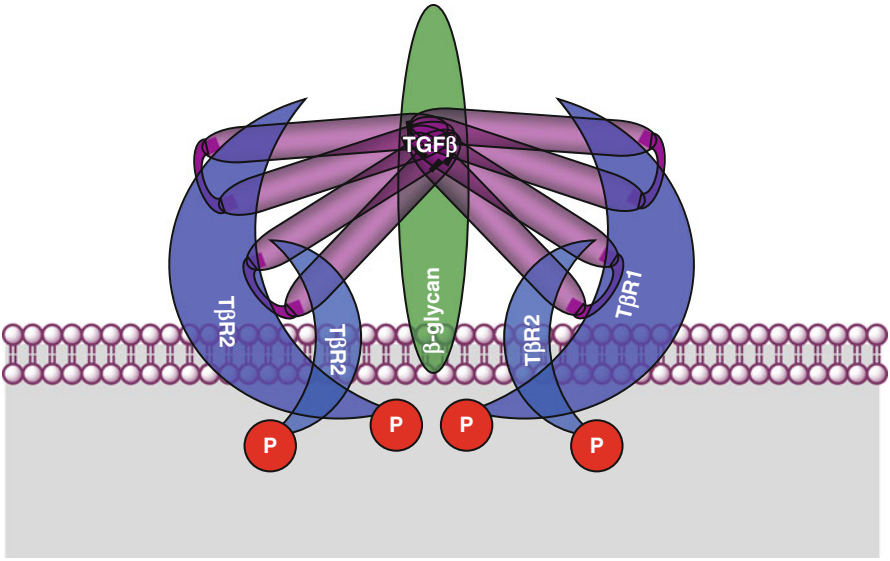


Fig. 5.2 TGFβ reception complex. Schematic representation of the reception complex, which comprises the TGFβ dimer, the two molecules of TβR1 receptor, the two molecules of TβR2 receptor and the coreceptor, β-glycan

ligand-type II receptor complex. Some ligands of the BMP subfamily can bind indifferently to any receptor type, the complex thus formed recruiting afterwards the receptor of the other type. Dimerisation generally occurs between two identical molecules (homodimerisation); however, the dimerisation between an inhibin (INHA) and an activin (INHB) is possible, the heterodimer having a negative effect on receptor recruitment.

In the extracellular space, ligands of both subfamilies are protected from receptor interaction by antagonistic proteins, which are generally secreted proteins, each of them being more or less specific of the factors. Some of them are associated to glycans to form proteoglycans, such as decorin; others are structural analogues of the ligands and inhibit their interaction with their receptor through a mechanism of negative dominance (cerberus, gremlin, sclerostin); others are binding proteins whose progressive diffusion in the extracellular space generates gradients, which play a major role in morphogenetic processes: this is the case of follistatin, chordin or noggin.

5.2 Receptors and Activation

At the level of the target cells of this signalling pathway, the receptors of the TGFβ and BMP subfamilies are membrane proteins with a single transmembrane domain, an extracellular portion aimed at ligand recognition and binding and an intracellular

part carrying serine/threonine kinase activity. These receptors are of two types, I and II, which especially differ by the presence of a glycine–serine (GS) box in the juxtamembrane domain of type I receptors. In addition, membrane proteins of the receptor cell are required for signal transduction and serve as coreceptors. Several nomenclatures have been used, which adds to the natural complexity of this system. Table 5.1 tentatively recapitulates the available data.

- There are seven type I receptors, called ALK (*activin receptor-like kinase*), from ALK1 to ALK7. The type I receptor of the three TGFβs is unique and is called TβR1 or ALK5 (gene *TGFBR1*). The type I receptors of the other members of the TGFβ subfamily are ALK2, ALK4 and ALK7 (genes *ACVR1A*, *ACVR1B* and *ACVR1C*) and those of the BMP subfamily are most often ALK1 (gene *ACVRL1*), ALK3 and ALK6 (genes *BMPR1A* and *BMPR1B*); the type I receptor of the anti-Mullerian hormone is ALK2 or ALK3.
- There are five type II receptors; similarly, the type II receptor of the TGFβs is unique (TβR2, gene *TGFBR2*); the type II receptors of the other members of the TGFβ subfamily are ACVR2A and ACVR2B (same gene names); the type II receptors of the BMP subfamily are either the same ACVR2A and ACVR2B or BMPR2; the type II receptor of the anti-Mullerian hormone is AMHR2.
- All the coreceptors, sometimes called type III receptors, have not all been identified; for the TGFβs, this is a proteoglycan called β-glycan or TβR3 (gene *TGFBR3*); for the TGFβ subfamily, this is a protein called cripto (genes *TDGF1*, for *teratocarcinoma-derived growth factor 1*, and *CFC1* for *cripto*, *FRL1*, *cryptic family 1*); for the BMP subfamily, these are either RGM (*repulsive guidance molecule*) A and B or endoglin (gene *ENG*) for BMP9 and 10. The mechanism of action of the coreceptors is not fully understood and may vary according to the receptor complexes and the ligands.
- Some decoy receptors have been identified, such as the BAMBI (*BMP and activin membrane-bound inhibitor homologue*) protein; they have structural analogy with type I receptors but only a reduced intracytoplasmic portion, devoid of kinase activity. These molecules are negative regulators of the signalling pathway.

Most of the dimeric factors of the TGFβ and BMP subfamilies bind first to two molecules of type II receptor, which then recruits two molecules of type I receptor; this is the reverse for some BMP factors such as BMP2 and 4. In all cases, the active complex is a hexameric complex that has been crystallised for some of them, allowing X-ray diffraction structural studies. The variety of the possible combinations explains the variety of the messages that can be transduced by this signalling pathway. The phosphorylation of type I receptor occurs at the level of the GS box, upstream the catalytic domain. This phosphorylation induces a conformational change which reveals the catalytic site of the type I receptor, which then can phosphorylate the SMAD proteins. The protein FKBP12 (*FK506-binding protein*) can bind the GS box when it is not phosphorylated and thus prevents receptor activation in the absence of the TGFβ signal.

5.3 Signal Transmission in the TGFβ Pathway

In contrast to the complexity of the combinatorial pattern of ligands and receptors, there is a relative simplicity of the signal transduction pathway (Fig. 5.3). The TGFβ subfamily factors are responsible for the activation by phosphorylation of SMAD2 and SMAD3 transcription factors, while those of the BMP subfamily induce the same for SMAD1, SMAD5 and SMAD8. The binding of these 'primary' SMAD proteins (R-SMADs) to SMAD4 (co-SMAD) induces their migration to the nucleus where they recognise and bind the promoter sequences of their respective target genes. Two proteins of the same family, SMAD6 and SMAD7 (I-SMADs), are competitive inhibitors of R-SMADs at the level of the activated receptors and prevent

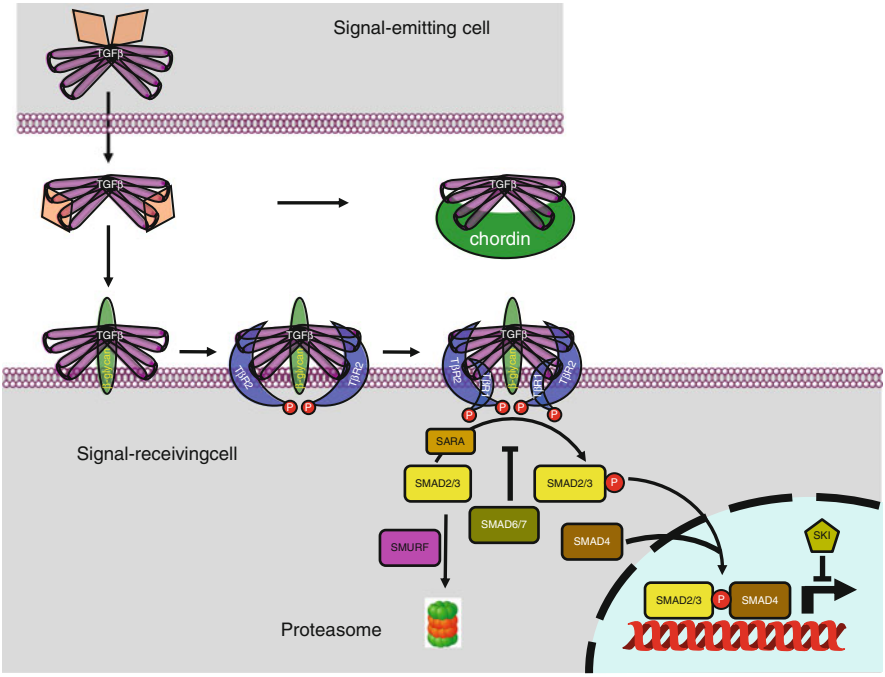


Fig. 5.3 TGFβ signalling pathway. The TGFβ precursor is cleaved in the endoplasmic reticulum but the *N*-terminal part (*orange diamonds*) remains non-covalently bound to TGFβ. In the extracellular space, it can be inactivated through binding proteins such as chordin. It is recognised by a type II receptor, TβR2 (gene *TGFBR2*), which undergoes dimerisation and autophosphorylation. Type II receptor recruits and phosphorylates two molecules of type I receptor, TβR1 or ALK5 (gene *TGFBR1*). The hexameric complex is associated to a coreceptor such as β-glycan (gene *TGFBR3*) or endoglin (gene *ENG*). Type I receptor can recruit and phosphorylate a transcription factor, SMAD2 or SMAD3, thanks to an anchor protein, SARA (*SMAD anchor for receptor activation*). This phosphorylation can be inhibited by another SMAD protein, SMAD6 or SMAD7. The activated transcription factor is associated to SMAD4, which allows its relocalisation in the nucleus, where it recruits coactivators or corepressors and triggers the transcription of various target genes. The SMAD proteins can also be driven to the proteasome via an E3 ubiquitin ligase, SMURF (*SMAD ubiquitinylation regulating factor*)

their phosphorylation. The specificity of ligand–receptor interactions is not as strict as mentioned above or in Table 5.1. ALK1 can be recruited by TGFβ when expressed at sufficient levels in endothelial cells, leading to the activation to SMAD1, 5 or 8 instead of SMAD2 or 3 and, consequently, to different, even opposite actions on the target gene promoters.

Several proteins are involved in the regulation of SMAD protein phosphorylation by the activated reception complexes; especially, SARA (*SMAD anchor for receptor activation*) (gene *ZFYVE9*) recruits certain R-SMADs to the activated receptor and appears to be required for their phosphorylation. The destruction of the reception complexes is ensured by endocytosis and proteolysis in the proteasome after the action of an E3 ubiquitin ligase (see Annex C) called SMURF (*SMAD ubiquitinylation regulating factor*).

The SMAD proteins contain two functional domains, MH1 (*MAD homology domain 1*) and MH2 (*MAD homology domain 2*), separated by a linker domain (Fig. 5.4). The MH1 domain is responsible for DNA binding, but the domain in charge of the transcriptional activity is MH2 and is also responsible for the interactions with the other SMAD proteins and the type I receptor. This domain contains the SSXS sequence at the *C*-terminal extremity, which is the phosphorylation target of the SMAD proteins by the ALK receptors. This phosphorylation induces a conformational change which releases domain MH1 from domain MH2, activating thus transcriptional activity. The SMAD transcription factors recruit transcriptional coactivators and corepressors so that, according to the context (i.e. the availability of nuclear proteins of the transcriptional machinery), the TGFβ stimulus can activate or repress the transcription of different genes. Nuclear inhibitors of TGFβ

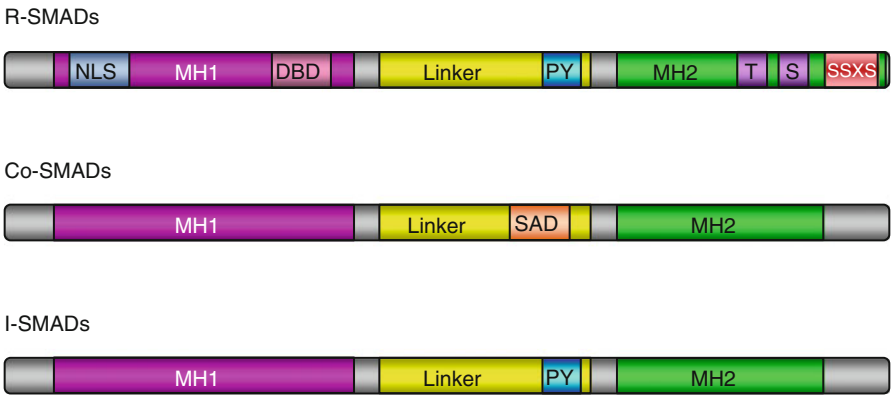


Fig. 5.4 Structural organisation of SMAD proteins. All SMAD proteins are characterised by the presence of two MAD homology domains, MH1 and MH2, separated by a linker with variable sequence. R-SMADs (SMAD1, 2, 4, 5 and 8) bear in addition a nuclear localisation sequence (NLS), a DNA-binding domain (DBD) and a *C*-terminal phosphorylation site SSXS. R-SMAD MH2 domain contains in addition a series of interaction sites with various interactants, included the TβR1 receptor (T) and the protein SARA (S). Co-SMAD (SMAD4) bears a SMAD activation domain (SAD). R-SMADs and I-SMADs (SMAD6 and 7) have a SMURF binding sequence, PPXY (PY), localised in the linker

signalling operate at the SMAD-induced transcription level; these are the proteins SKI (*Sloan-Kettering Institute oncogene*) and SKIL (or SNO).

Besides the canonical signalling pathway described above, which is obtained through SMAD protein activation, the factors of the TGFβ family are able to activate the MAP kinase pathways (Chap. 2), likely through the activation of MAP3Ks such as TAK1 (gene *MAP3K7*), leading to the activation of the ERK, p38 or JNK pathways. Conversely, the SMAD transcription factors can be phosphorylated by MAP kinases, especially ERK1/2, on target sites that can induce their activation or inhibition.

5.4 Oncogenic Alterations

The pathways activated by TGFβ and its analogues play multiple roles in development and morphogenesis, so that the germline mutations of the genes encoding the proteins of the pathway may generate numerous kinds of hereditary diseases and congenital malformations. Their roles in oncogenesis are complex and ambivalent. The TGFβ pathway is rather an anti-oncogenic pathway in the early stages, because it acts against proliferation and induces differentiation and apoptosis. However, in the late stages, it may favour invasion and metastasis through its actions on tumour cell adhesion and migration, on the tumour microenvironment, on angiogenesis and on immune survey.

In most studies, the different players of the TGFβ pathway behave as tumour suppressor genes. Invalidating germline mutations of *BMPR1A* (ALK3), of *SMAD4*, and more rarely of endoglin have been found in a syndrome of predisposition to digestive cancers called juvenile polyposis. In addition, some functional polymorphisms of TGFβs and their receptors are accompanied by an increase in susceptibility to some cancers. Breast and lung cancer metastases are frequently associated to the loss of *TGFBR3* gene expression. Mutations of *TGFBR2* and *SMAD4* are frequently observed in colon, pancreas and lung cancers. *TGFBR2* mutations are related to the presence of a microsatellite within the coding sequence, which is frequently altered when microsatellite instability occurs in colorectal cancers, either sporadic or with hereditary predisposition. *TGFBR2* loss of expression has been observed in various cancer types.

The positive role of TGFβ on metastasis formation has been detected experimentally: the inhibition of the kinase activity of TβR1 inhibits the formation of lung metastases from human mammary tumours xenografted in *nude* mice. Such a role is suspected in human cancers, because TGFβ is a potent inducer of the epithelial-to-mesenchymal transition, through the action of SMAD on the expression of *SNAI1* (SNAIL), *SNAI2* (SLUG) and *TWIST*.

During angiogenesis, the protease activation of TGFβ latent forms of endothelial cells stimulates their differentiation into mural cells and, therefore, contributes to neovessel formation. In contrast, BMP9 (*GDF2*) inhibits the proliferation of endothelial cell and induces the vascular quiescence characteristic of the phase of neovessel maturation.

5.5 Pharmacological Targets

Targeting the TGFβ pathway should be conceived with much caution because of the ambivalent effect of this pathway, which schematically inhibits proliferation and stimulates metastatic dissemination. Macromolecular inhibitors of TGFβ (monoclonal antibodies such as fresolimumab, trapping proteins, antisense oligonucleotides such as trabedersen [directed against TGFβ2], soluble receptors, sequestration proteoglycans mimicking β-glycan) have undergone preclinical and early clinical studies, especially in glioblastomas and malignant melanomas. The TGFβ receptors can also be targeted, and antibodies against ALK5 (TβR1), TβR2, ALK1 and endoglin have entered clinical trials.

Small molecules inhibiting the serine/threonine kinase activity of ALK receptors have also been developed with, as usual for such compounds, cross-reactivity with other kinases such as p38. However, ALK5 (TβR1) inhibitors with acceptable specificity have been obtained, with an inhibitory effect on epithelial-to-mesenchymal transition in preclinical models. LY-2157299 has entered clinical trials in hepatocarcinomas, pancreatic cancers and glioblastomas. Inhibiting the TGFβ pathway may, as expected, inhibit metastatic dissemination but may also lead to stimulation of tumour immunogenicity and, therefore, contribute to the therapeutic effect.

The SMAD proteins are much more difficult to target. Aptamers able to inhibit the active domains of these proteins have been selected and appear to be functional in vitro.

Further Reading

Akhurst RJ, Hata A. Targeting the TGFβ signalling pathway in disease. Nat Rev Drug Discov. 2012;11:790–811.

Arjaans M, Oude Munnink TH, Timmer-Bosscha H, et al. Transforming growth factor (TGF)-β expression and activation mechanisms as potential targets for anti-tumor therapy and tumor imaging. Pharmacol Ther. 2012;135:123–32.

de Caestecker M. The transforming growth factor-β superfamily of receptors. Cytokine Growth Factor Rev. 2004;15:1–11.

Dong M, Blobe GC. Role of transforming growth factor-β in hematologic malignancies. Blood. 2006;107:4589–96.

Drabsch Y, ten Dijke P. TGF-β signalling and its role in cancer progression and metastasis. Cancer Metastasis Rev. 2012;31:553–68.

Gordon KJ, Blobe GC. Role of transforming growth factor-superfamily signalling pathways in human disease. Biochim Biophys Acta. 2008;1782:197–228.

Heldin CH, Landström M, Moustakas A. Mechanism of TGF-β signaling to growth arrest, apoptosis, and epithelial–mesenchymal transition. Curr Opin Cell Biol. 2009;21:166–76.

Ikushima H, Miyazono K. TGFβ signalling: a complex web in cancer progression. Nat Rev Cancer. 2010;10:415–24.

Leivonen SK, Kähäri VM. Transforming growth factor-β signaling in cancer invasion and metastasis. Int J Cancer. 2007;121:2119–24.

Levy L, Hill CS. Alterations in components of the TGF-β superfamily signaling pathways in human cancer. Cytokine Growth Factor Rev. 2006;17:41–58.

Massagué J, Seoane J, Wotton D. Smad transcription factors. Genes Dev. 2005;19:2783–810.
Massagué J. TGFβ in cancer. Cell. 2008;134:215–30.
Pickup M, Novitskiy S, Moses HL. The roles of TGFβ in the tumour microenvironment. Nat Rev Cancer. 2013;13:788–99.
Rahimi RA, Leof EB. TGF-β signaling: a tale of two responses. J Cell Biochem. 2007;102: 593–608.
Shi Y, Massagué J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell. 2003;113:685–700.
Smith AL, Robin TP, Ford HL. Molecular pathways: targeting the TGF-β pathway for cancer therapy. Clin Cancer Res. 2012;18:4514–21.
Wakefield LM, Hill CS. Beyond TGFβ: roles of other TGFβ superfamily members in cancer. Nat Rev Cancer. 2013;13:328–41.
Zavadil J, Böttinger EP. TGF-β and epithelial-to-mesenchymal transitions. Oncogene. 2005;24: 5764–74.

G-Protein-Coupled Receptors

6

Abstract

G-protein-coupled receptors (GPCRs) constitute a huge family of membrane receptors to numerous hormones, neurotransmitters and diverse compounds. A large part of cardiovascular and neurological pharmacology is based upon the knowledge of these receptors and on the identification of molecules able to interfere with them, in an agonistic or antagonistic way. We will only briefly consider this vast domain, so as to focus on some of these receptors, those having a possible link with oncogenesis. In a first part, we will present the overall way of operating of these receptors and the main events resulting from their activation. We will present in a special part a family of GPCR and cognate ligands, which are susceptible to be involved in oncogenesis and metastasis, the chemokine–chemokine receptor families.

As a general feature, the functioning of these receptors involves the activation of heterotrimeric protein complexes called G-proteins, whose α subunit is able to bind and hydrolyse GTP. These proteins activate the membrane enzymes, especially adenylyl cyclase and phospholipase C, which produce second messengers (cyclic AMP, diacylglycerol, inositol trisphosphate, Ca^{2+}). These second messengers ultimately activate various types of serine/threonine kinases, especially protein kinases A (PKAs), B (AKT) and C (PKCs). Heterotrimeric G-proteins may also activate small G-proteins through the activation of their guanine nucleotide exchange factors (GEFs).

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_6

6.1 Structure and Mechanism of Action of G-Protein-Coupled Receptors

6.1.1 Receptors

GPCRs are comprised of a single polypeptidic chain equipped with seven transmembrane domains; these domains are made of a series of 22–25 hydrophobic amino acids, which allow the formation of an α-helix inserted within the aliphatic chains on membrane phospholipids. The *N*-terminal part of the receptors is extracellular while the *C*-terminal part is intracellular; three loops separating the transmembrane domains are present on each side of the membrane. The polypeptidic chain thus winds in the membrane as a snake, and these receptors are sometimes called 'serpentine receptors'. The number of receptors in this family is higher than 800, which represents a significant part of the genome; they are divided in several families and subfamilies as a function of their structure, their ligands and the type of binding. With some simplification, one can identify:

- The *rhodopsin receptors* family, the most abundant (about 700 members). Among them are receptors to small molecules with intercellular messenger function: acetylcholine, catecholamines and other biogenic amines; nucleosides and nucleotides; lipid mediators such as retinal, prostaglandins, leukotrienes and thromboxanes, free fatty acids and endocannabinoids; and protein and peptide hormones (corticotropin [ACTH], gonadotropins [FSH, LH], thyrotropin [TSH], oxytocin [OXT], angiotensin [AGT], somatostatin [SST], chemokines, etc.). This family contains also all the olfactory and gustatory receptors, which represent more than 400 different proteins. The binding between receptor and ligand involves deep transmembrane structures of the receptors. Some of these receptors are activated by thrombin- or trypsin-induced proteolysis, which cleaves the *N*-terminal part of the receptor, which then behaves as an autonomous ligand; they are called protease-activated receptors (PAR).
- The *secretin receptors* family (15 members), which comprises the receptors to various peptide hormones: glucagon, parathormone, calcitonin, and hypothalamic releasing factors (CRH [corticotropin-releasing hormone], GNRH [gonadotropin-releasing hormone] or LHRH [luteinising hormone-releasing hormone], TRH [thyrotropin-releasing hormone]). The binding between ligand and receptor involves the transmembrane domains and the *N*-terminal part of the receptor.
- The *glutamate receptors* family (22 members), comprising the metabotropic receptors (as opposed to ionotropic receptors, Chap. 15) of several neurotransmitters (glutamate, GABA), which bind to the *N*-terminal part of the receptor. This family also contains a Ca^{2+} 'sensor' (CASR); the calcium ion binds to an extension of the *N*-terminal part of the receptor, which then interacts with the deep transmembrane domains and behaves as an 'auto-ligand'.
- The *adhesion receptors* family, whose receptors have a bulky *N*-terminal domain allowing them to interact with extracellular molecules, but whose true ligands are mainly proteoglycans.

- The *frizzled receptors* family, which comprises the 10 WNT protein receptors (Chap. 7) and the secondary receptor of the Hedgehog pathway, called *smoothened* (SMO) (Chap. 9). These receptors present structural analogies with GPCR, but the G-protein activation characteristic for GPCR has not been always demonstrated.

In all cases, ligand–receptor binding induces a conformational change which generates the message. One can consider that the receptors exist under two forms, an active and an inactive form, and that their ligands stabilise one form or the other one. Some receptors have been shown to be activated after dimerisation, but this may not be the case for all receptors.

6.1.2 G-Proteins

The common final way for GPCR activation resides in the involvement of large heterotrimeric G-proteins. All G-proteins are able to bind and hydrolyse GTP (they are therefore GTPases); there is a distinction between 'small G-proteins' such as RAS (Chap. 2), RHEB or RAG (Chap. 3) and 'large heterotrimeric G-proteins' which are studied here and are transducers of the information received by GPCR. Large heterotrimeric G-proteins are comprised of three polypeptidic chains, called α, β and γ. The β and γ chains are tightly bound in a unique entity βγ. G-proteins are anchored to the membrane by two lipid moieties, the first one being a linear chain (myristoyl or palmitoyl group), which is added post-translationally at the *N*-terminal extremity of the α subunit, and the second one being a branched chain (farnesyl or geranylgeranyl group), which is post-translationally added to the *C*-terminal part of the γ subunit. The α subunit contains the binding domain of guanyl nucleotides, which is occupied by GDP at basal state. Large heterotrimeric G-proteins are stably associated with an effector enzyme and can transiently associate with a GPCR.

GPCR ligand binding induces this transient association between a GPCR and a large G-protein, and this binding induces the replacement of GDP by GTP at the level of the α subunit. The α subunit then detaches from the βγ subunit and realises the GTP cleavage thanks to its GTPase activity. During the brief period when the α subunit is bound to GTP, the associated effector enzyme is activated (sometimes inhibited when it has constitutive activity) and can exert (or no longer exert) its catalytic activity. As soon as GTP is hydrolysed, this regulation ceases, the effector enzyme returns to basal state and the α subunit reassociates to the βγ subunit. Similarly, the βγ subunit can activate other effectors before reassociating with the GDP-bound α subunit. Large G-proteins thus function as switches giving one signal only; they must reassociate to a newly activated receptor molecule and exchange GDP against GTP to give a novel impulse to their effectors. Figure 6.1 presents the operating scheme of heterotrimeric G-protein activation.

In front of hundreds of receptors, there is a total of 16 genes encoding the α subunits (genes *GNAx)*, 5 for β subunits (genes *GNB1* to *GNB5*) and 12 for the γ

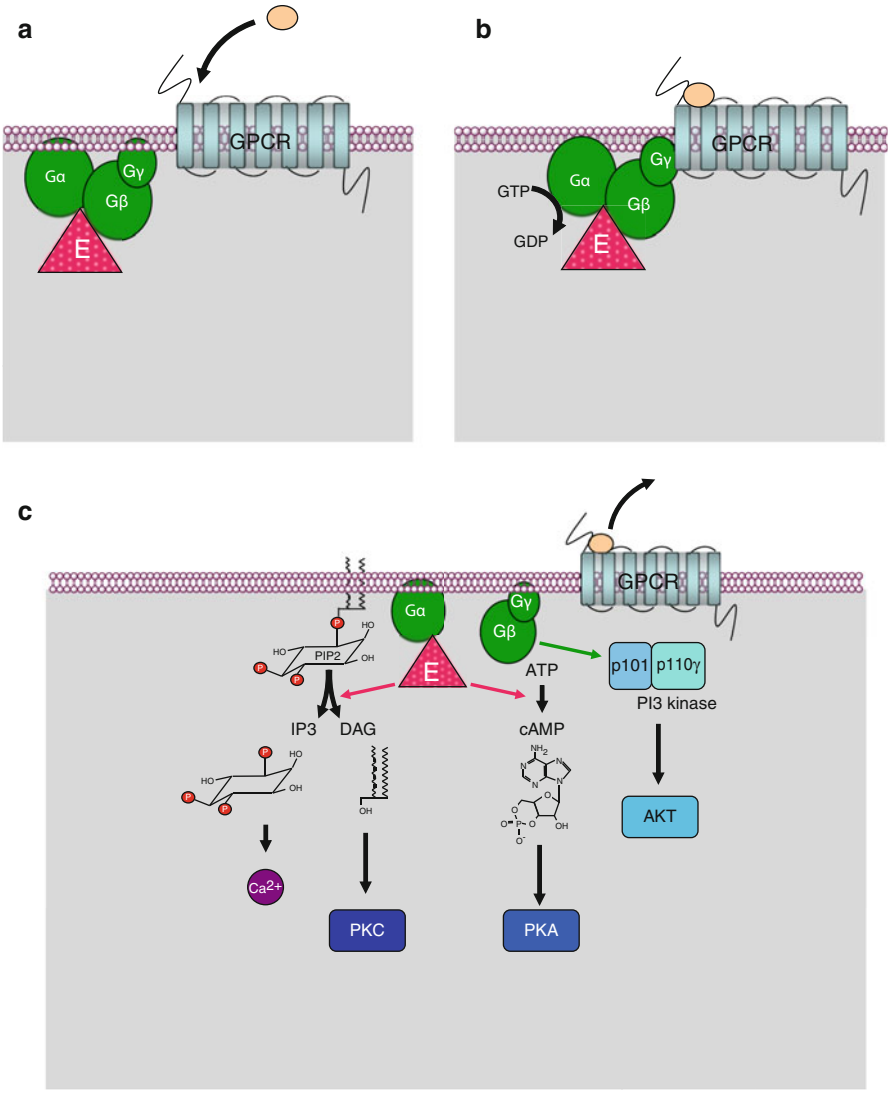


Fig. 6.1 Activation and function of large heterotrimeric G-proteins. At the basal state (**a**), the 7-transmembrane-domain receptor and the heterotrimeric G-protein complex are dissociated; the α subunit is bound to GDP and to its effector enzyme, adenylyl cyclase (ADCY) or phospholipase C beta (PLCβ). Upon ligand binding (**b**), receptor and Gαβγ complex are transiently associated; the α subunit exchanges its GDP ligand against GTP. At the active state (**c**), the GTP-bound α subunit can dissociate from the receptor and the βγ subunit and can activate (or deactivate in the case of an α_i subunit) the effector enzyme. cAMP is then generated from ATP by ADCY (or no longer synthesised), and DAG and IP3 are generated by phospholipase C beta (PLCβ), which induces especially protein kinase A (PKA) activation by cAMP, protein kinase C (PKC) activation by DAG and Ca^{2+} release by IP3. In addition, the βγ subunit is able to activate the regulatory p101 subunit of phosphoinositide 3-kinase gamma (PI3Kγ), which in turn can activate AKT

subunits of G-proteins (genes *GNG1* to *GNG12*). The α subunits reflect the tissue and functional diversity of G-proteins, whereas the βγ subunits are rather involved in controlling G-protein activity. There are several types of α subunits, the main ones being α_s (gene *GNAS*), α_i (genes *GNAI1* to *GNAI3*), α_q (genes *GNAQ*, *GNAL* and *GNA11*) and α_{12} (gene *GNA12* and *GNA13*), which differ accordingly to the effector enzyme they can activate: GTP-bound α_s subunits activate adenylyl cyclases (genes *ADCY1* to *ADCY10*), while α_i subunits inhibit this activity; GTP-bound α_q subunits activate phospholipases C beta (PLCβ, genes *PLCB1* to *PLCB4*) and GTP-bound α_{12} subunits activate small G-proteins of the RHO family. Adenylyl cyclase and PLCβ generate second messengers, which are common intermediates of the signals brought by GPCR ligands. The βγ subunit, transiently detached from the α subunit during activation, also has proper effectors, especially the regulatory subunit of PI3 kinase gamma (p101, gene *PIK3R5*), as well as some types of ionic channels.

The activity of large heterotrimeric G-proteins is regulated by RGS (*regulator of G-protein signalling*) proteins, which stimulate GTPase activity and are therefore GAPs (*GTPase-activating proteins*). As a consequence, they deactivate large G-proteins that are brought back to GDP-bound, inactive state. There are about 20 RGS proteins with variable specificity toward G-proteins. Large heterotrimeric G-proteins can also be regulated by phosphorylation: the βγ subunit is able to recruit a kinase (GRK, *GPCR kinase*) which phosphorylates the activated receptor and accelerates its desensitisation through the formation of a β-arrestin-binding site, inducing receptor internalisation and recycling.

6.2 Second Messengers of GPCR Activation

6.2.1 Adenylyl Cyclases and Cyclic AMP

Adenylyl cyclases (*ADCYs*) are high-molecular-weight membrane enzymes, comprised of two homologous parts, each made of six transmembrane helices and a large *C*-terminal domain called the C-domain. The two C-domains are associated and present an ATP-binding site and a catalytic site for its conversion into cAMP. These enzymes are controlled by heterotrimeric large G-proteins; those equipped with an α_s subunit activate adenylyl cyclase when bound to GTP, whereas those equipped with an α_i subunit inhibit adenylyl cyclase. They are also controlled by multiple factors and integrate much information, in particular through Ca^{2+} binding and protein kinase C-induced phosphorylation.

Cyclic APM (cAMP) (Fig. 6.2) is the second messenger of a large number of signals received by GPCR: hormones, neurotransmitters, etc. It is involved in the regulation of numerous enzymes, especially in metabolic pathways. cAMP targets are protein kinases A (PKAα, β and γ, genes *PRKACA*, *B* and *G*), which are tetrameric serine/threonine kinases containing two catalytic and two regulatory subunits. They phosphorylate their targets on the serine residues of an RRXS consensus motif. cAMP-induced PKA activation is cooperative, which allows a rapid and wide

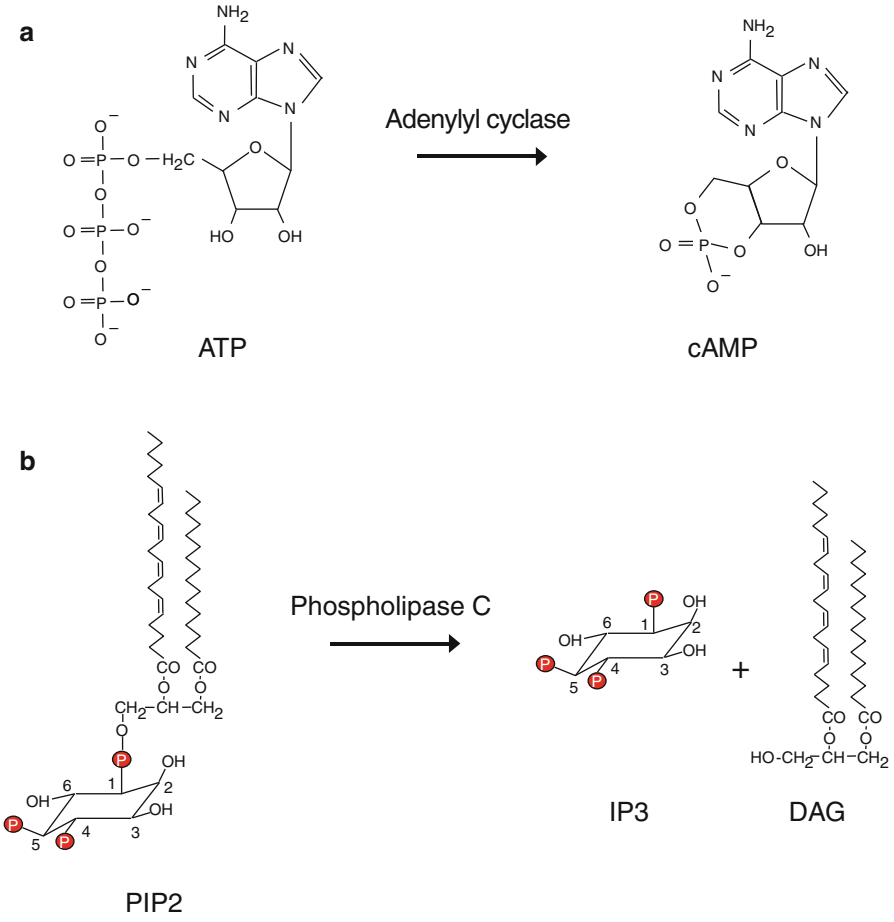


Fig. 6.2 Formation of second messengers. (**a**) Adenylyl cyclase can convert ATP into cyclic AMP (cAMP). (**b**) Phospholipase C beta hydrolyses phosphatidylinositol 4,5-bisphosphate (PIP2) into inositol trisphosphate (IP3) and diacylglycerol (DAG)

effect, rapidly terminated by the fact that PKAs phosphorylate and activate cAMP phosphodiesterases (genes *PDE4A* to *PDE4D*), which hydrolyse cAMP. PKAs target numerous proteins, among which are glycogen phosphorylase kinase (PHK), which is activated, and glycogen synthase (GYS), which is deactivated, in response to this phosphorylation. PKAs also exert an effect at the gene transcription level, by activating through phosphorylation a transcription factor called CREB (*cAMP response element-binding protein*), which relocalises into the nucleus and recognises specific promoter sequences of several target genes, called CRE (*cAMP response element*). cAMP is also able to directly activate EPAC (*exchange protein directly activated by cAMP*), a GTP–GDP exchange factor of a small G-protein, RAP, which activates other intracellular signalling pathways.

6.2.2 Phospholipases C, Diacylglycerol and Inositol 1,4,5-Trisphosphate

Phospholipases C (PLCs) are membrane enzymes equipped with a PH (*pleckstrin homology*) domain allowing them to recognise and bind phosphoinositides. PLCγ can be activated by a tyrosine kinase receptor (TKR, Chap. 1), whereas PLCβs are activated by a GPCR through the action of large heterotrimeric G-proteins equipped with an α_q subunit. They act on phosphatidylinositol 4,5-bisphosphate and generate a lipophilic molecule, diacylglycerol (DAG), and a hydrosoluble molecule, inositol 1,4,5-trisphosphate (IP3) (Fig. 6.2). These two molecules are the second messengers of numerous GPCR ligands, hormones and various transmitters.

DAG is the principal physiological activator of protein kinases C (PKCs). This family of nine serine/threonine kinases (genes *PRKCx*), of variable tissue expression, has numerous substrates with relatively low specificity and exerts various intracellular effects. Before DAG-induced activation, PKCs must be phosphorylated by PDK1 (*phosphoinositide-dependent kinase 1*), which enables their autophosphorylation on two distinct sites. DAG interaction enables PKC anchoring in the plasma membrane, which is required for catalytic activity. Several other lipid compounds, produced by enzymes other than phospholipase C, such as phospholipases A2 and D and sphingomyelinase, are also able to activate PKCs: lysophosphatidic acid (LPA), arachidonic acid, sphingosine, etc. PKCs are able to bind numerous proteins that direct its action to precise protein targets: STICKs (*substrates that interact with C-kinase*), such as vinculin (VCN), talin (TLN) and annexin A (ANXA), PICKs (*proteins interacting with C-kinase*) and RACKs (*receptors for activated C-kinase*). These proteins play major roles in cell adhesion and polarity and intracellular contacts and explain the multiple actions of PKCs in these processes.

PKCs can be divided in several subfamilies: class A ('conventional': PKCα, PKCβ and PKCγ, activated by DAG and Ca^{2+}), class B ('novel': PKCδ, PKCε, PKCθ and PKCη) and class C ('atypical': PKCι/λ and PKCζ). PKCs phosphorylate many substrates, with variable specificity, depending upon the different members of the family. They are involved in various pathways, especially in cell proliferation, cell cycle entry, cell survival, cell adhesion and cell migration (Fig. 6.3). Among the PKC substrates are the IKK (*IκB kinase*) proteins, which generate NFκB (Chap. 12), the CRAF (*RAF1*) MAP3 kinase in the MEK–ERK MAP kinase pathway (Chap. 2), and GSK3β, which is inhibited by this phosphorylation on Ser^9 and which is known as a tumour suppressor in the Wnt (Chap. 7) and Hedgehog (Chap. 9) pathways.

IP3 is one of the main factors responsible for Ca^{2+} mobilisation from intracellular storage compartments toward the cytosol. Cytosolic Ca^{2+} concentration is very low (0.05–0.1 μM), and Ca^{2+}-dependent ATPase pumps throw it out continuously to the extracellular space or the endoplasmic reticulum. IP3 allows a rapid mobilisation of Ca^{2+}, thanks to the activity of IP3 receptors (IP3R1, 2 and 3; genes *ITPR1*, *2* and *3*), which are present at the surface of endoplasmic reticulum membranes. These receptors are tetrameric calcium channels of which each component is a protein with six transmembrane domains and two intracellular cytosolic extremities (see Chap. 15). The *N*-terminal part harbours the IP3-binding site as well as several regulation

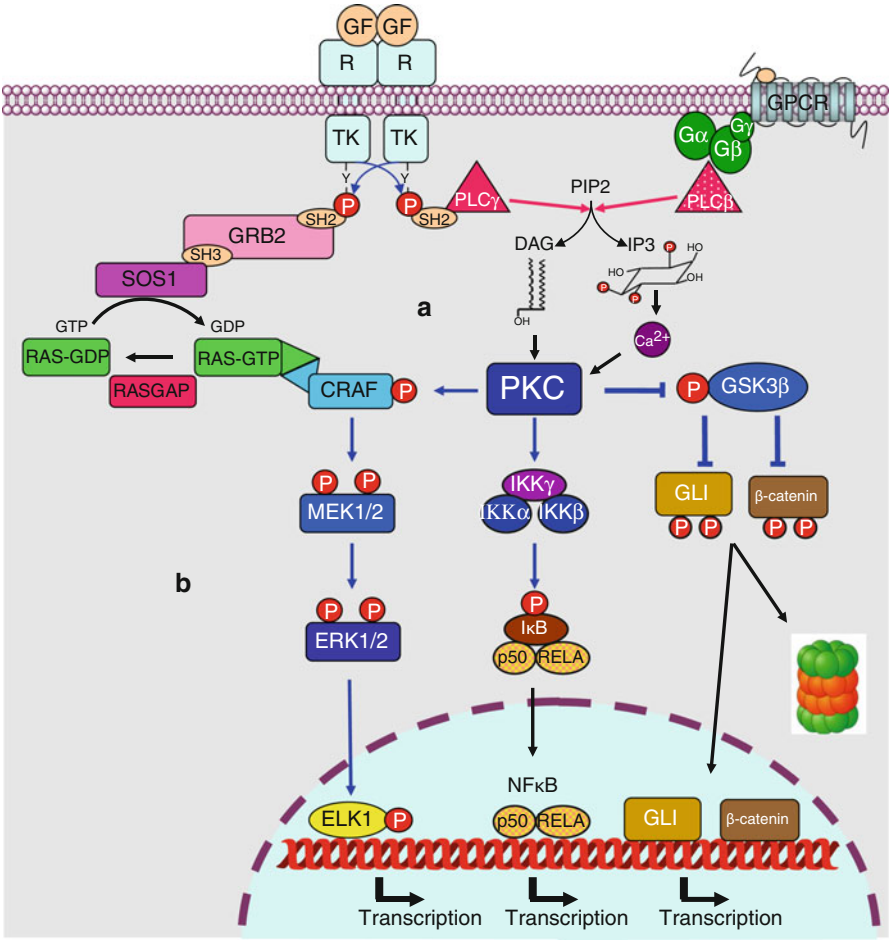


Fig. 6.3 PKC-activating and PKC-activated pathways. (**a**) PKC-activating pathways. PKCs are activated by diacylglycerol (DAG) and, for some of them, by Ca^{2+}, two second messengers obtained by hydrolysis of phosphatidylinositol 4,5-bisphosphate by a phospholipase C, which can be activated either by tyrosine kinase receptors (PLCγ) or by G-protein-coupled receptors (PLCβ). (**b**) PKC-activated pathways. Among numerous protein targets, only some of them are indicated. Activating phosphorylation of CRAF and IKK leads to the activation of MAP and NFκB transcription factors, respectively. The phosphorylation of GSK3β is in contrast inhibitory; this kinase phosphorylates GLI (Hedgehog pathway, Chap. 9) and β-catenin (Wnt pathway, Chap. 7), which are transcription factors or transcriptional activators, which leads them to ubiquitinylation and proteasome destruction. This double inhibition explains the activation of these pathways by PKC

domains. These receptors are activated by an increase in the local Ca^{2+} concentration, which realises a positive feedback called CICR (*Ca^{2+}-induced Ca^{2+} release*), but their Ca^{2+} sensitivity progressively decreases and they are rapidly deactivated.

Ca^{2+} is involved in numerous cell processes (Chap. 15). As a divalent cation, it is able to bind two negative charges borne by proteins or other macromolecular

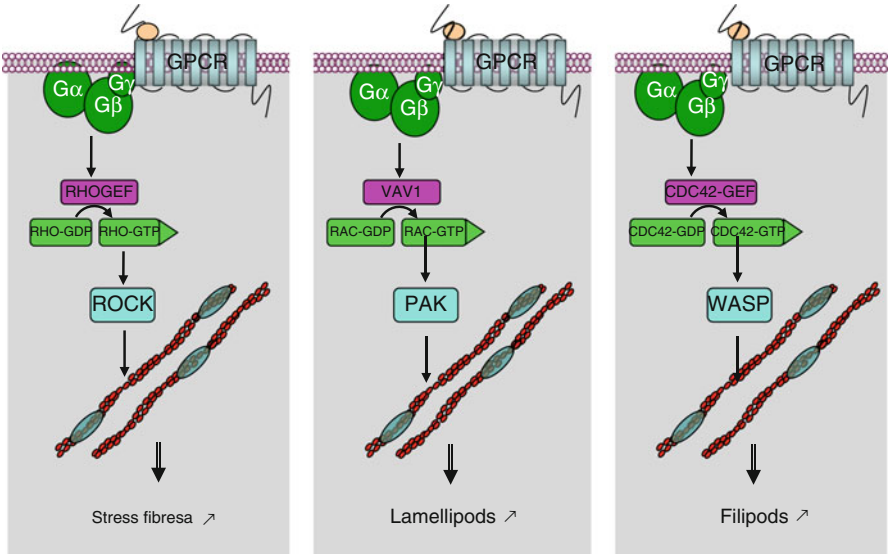


Fig. 6.4 Small G-protein activation downstream *GPCR* signalling. A series of small G-proteins of the RHO family are activated in response to GEF activation by a large heterotrimeric G-protein. There are 16 small G-proteins in this family, including RHOA, RHOD, RND1-3, RAC1-3, CDC42 and others, as well as 12 distinct GEFs. These small G-proteins are responsible for the activation of serine/threonine kinases, which in turn are able to phosphorylate myosin or actomyosin cytoskeleton-interacting proteins, such as cofilin, profilin and others

edifices and to play a major role at the level of cell adhesivity. Most Ca^{2+} effects are mediated by its association to a small protein, calmodulin. The activity of many proteins is regulated by the Ca^{2+}–calmodulin complex and thus by Ca^{2+} availability, especially proteins involved in signalling, such as PKCs, PLCs, adenylyl cyclases, phosphodiesterases, NO synthase, etc. Other proteins display Ca^{2+}-binding sites and can be directly activated by this ion: these are particularly cytoskeleton proteins involved in cell shape maintenance and cell motility.

6.2.3 RHO Family Small G-Protein Activation

This is the third type of signalling downstream GPCR activation, which can promote the activation of GDP–GTP exchange proteins (GEFs) other than RAS; the small G-proteins of the RAP family are activated by a cAMP-activated GEF called EPAC (*exchange protein directly activated by cAMP*); the small G-proteins of the RHO family (RHOA-D, RND1-3, RAC1-3, CDC42, etc.) are activated by various GEFs, via the α_{12} subunit of large heterotrimeric G-proteins. The activation of these small G-proteins leads to multiple intracellular actions, at the level of cell proliferation, cytoskeleton organisation, motility, polarity and adhesion (Fig. 6.4). The mechanism of GEF activation by the α_{12} subunits of large heterotrimeric G-proteins is not completely elucidated but seems to involve direct interaction.

6.2.4 Connections with Other Signalling Pathways

The MAP kinase and the PI3 kinase pathways can be activated in different ways following GPCR activation (Chaps. 2 and 3). cAMP-activated PKAs are able to play the role of MAP4 kinases and to phosphorylate MAP3 kinases (Fig. 2.3) such as BRAF, which initiates the MEK–ERK MAP kinase module. DAG-activated PKCs are able to phosphorylate JUN *N*-terminal kinase (JNK) in another MAP kinase module. In conjunction with Ca^{2+}, DAG is able to activate a RAS protein GDP–GTP exchange, RASGRP1, initiating a MAP kinase module at the level of the Golgi apparatus (Fig. 2.3). Regulatory subunits of phosphatidylinositol 3-kinase gamma (PI3Kγ), namely, p101 (gene *PIK3R5*) and p84/p87 (gene *PIK3R6*), are also susceptible of activation by a large heterotrimeric G-protein, through the βγ subunit, leading to AKT pathway activation. All these interconnections show that GPCR-activated pathways participate to cell proliferation and explain why pharmacological DAG analogues, phorbol esters, exert a co-carcinogenic effect of tumour growth promotion.

6.3 Oncogenic Alterations and Pharmacological Targets

GPCR-activated pathways are not as frequently altered in cancers as the TKR-activated pathways are (Chaps. 1, 2 and 3). It is however interesting to identify and describe such alterations, which may represent potential targets for innovative approaches. We have mentioned in the previous section the fact that, in a general way, GPCR-activated pathways could contribute to cell proliferation; we mentioned especially the tumour-promoting role of PKCs through their pharmacological alterations by phorbol esters. Several specific alterations of these pathways can be mentioned for their contribution to oncogenesis. The special case of chemokines and their receptors will be studied below, in Sect. 6.4.

Some GPCR ligands are known mitogenic factors: thrombin, lysophosphatidic acid (LPA), sphingosine-1-phosphate (S1P), gastrin-releasing peptide (GRP), endothelin and prostaglandins stimulate cell proliferation through the activation of their cognate receptors; they are overexpressed in several cancer types, but their driver role in oncogenesis is hard to assign. This action on cell proliferation may be mediated through PKC activation or by the classical proliferation pathways (MAP kinase and PI3 kinase), interconnected with GPCR-activated pathways via small G-proteins. Other GPCR ligands may also be considered as cancer promoters: this is the case for catecholamines, and the use of β-blockers has been proposed in cancer therapeutics; this is also the case for purinergic receptor (P2Y type) ligands. Conversely, several GPCR ligands can induce experimentally cancer growth inhibition and apoptosis, so that their use has been proposed as part of cancer therapy: this is the case of (endo) cannabinoids, enkephalins and adenosine receptor (A3 type) agonists.

At the receptor level, MAS1, related to angiotensin receptors, has been presented as an oncogene, but no somatic mutations have been identified. Most oncogenic mutations found in GPCRs concern the TSH (*thyroid-stimulating hormone*), FSH

(*follicle-stimulating hormone*) and CCK (*cholecystokinin*) receptors in thyroid, ovary and colon cancers, respectively. Mitogenic factor receptors of the GPCR family are overexpressed and sometimes mutated in cancers; this is the case for S1P and LPA receptors, as well as for PAR receptors: the thrombin receptor, F2A, is highly overexpressed in invasive cancers of various organs. Outside the rhodopsin family receptors, mutations have been found in glutamate metabotropic receptors, adhesion receptors and SMO. Small-cell lung cancers and endocrine tumours in general could be sustained by paracrine or autocrine activation of the receptors of the peptides that they produce in large excess: bombesin (*GRP*), bradykinin (*KNG1*), neuromedin B (*NMB*), cholecystokinin (*CCK*), galanin (*GAL*), neurotensin (*NTS*) and vasopressin (*AVP*). The same occurs for carcinoid tumours secreting vasomotor substances. Activation of endothelin (*EDN*), bradykinin, bombesin and angiotensin (*AGT*) receptors has been observed in prostate cancers. One can hypothesise that a therapeutic approach targeting these receptors or their ligands could be efficient in the treatment of these rare but often lethal cancers.

At the level of large heterotrimeric G-proteins, it has been experimentally shown that several α subunits could be oncogenic. However, only some oncogenic mutations have been described in human cancers; they concern the GTPase activity of α_s and α_{12} subunits in thyroid, adrenal and hypophysis tumours and the α_q and α_{11} subunits in malignant melanomas. More recently, large-scale tumour genome sequencing has revealed that α subunits were mutated in various digestive cancers, such as colon cancer, hepatocellular carcinoma and biliary tract and pancreatic tumours.

Downstream many GPCRs, PKCs have been involved in many pathological states, including cardiology, neurology and oncology. The oncogenic role of PKCs is multiple: PKCα is involved in proliferation and PKCβ and PKCδ in angiogenesis; PKCε is overexpressed in breast and colon cancers, PKCη in non-small-cell lung cancers and glioblastomas and PKCθ in gastrointestinal sarcomas. Several therapeutic approaches have been developed to target PKCs, unsuccessfully until now: kinase inhibitors, such as staurosporine or enzastaurin, DAG mimetics blocking PKC activation, PKC-interacting proteins such as RACK proteins, peptidomimetic inhibitors able to compete with definite PKC substrates and antisense oligonucleotides (ASO) able to inhibit the expression of specific PKCs, such as aprinocarsen for PKCα.

6.4 Chemokine Pathway

6.4.1 Chemokines and Chemokine Receptors

Chemokines constitute a class of small proteins of 8–15 kDa characterised by the presence of four cysteine residues generating two disulphide bridges. The two first cysteine residues can be either consecutive (CCL chemokines) or separated by another amino acid residue (CXCL chemokines) (Fig. 6.5). There are in addition two rare chemokine types, CX3CL and XCL. The chemokine receptors are GPCRs

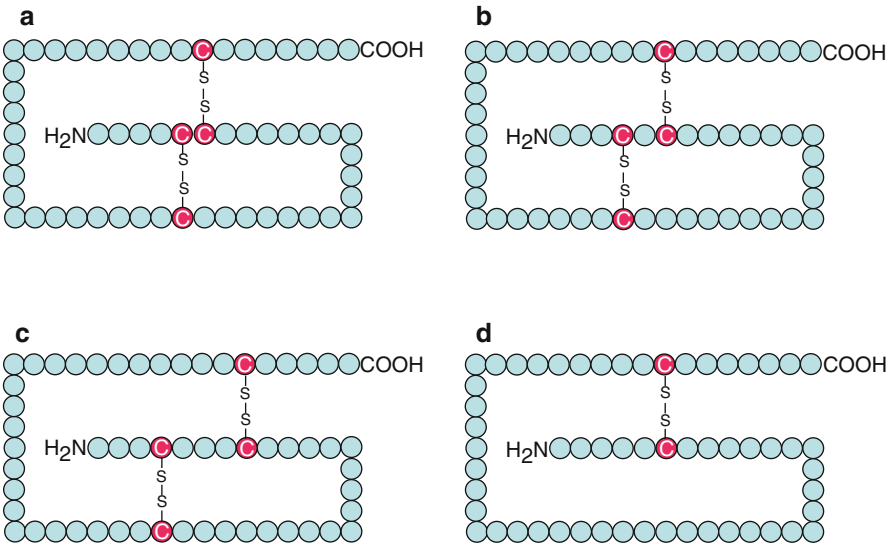


Fig. 6.5 Schematic representation of the various chemokine types. Adjacent or nearby cysteine residues can form, with other cysteine residues, intramolecular disulphide bridges. (**a**) CCL type: the polypeptidic chain contains adjacent cysteine residues; (**b**) CXCL type: the two cysteine residues are separated by one amino acid residue; (**c**) CX3CL type: the two cysteine residues are separated by three amino acids; (**d**) XCL type: a single disulphide bridge is present in the polypeptidic chain

distributed into four families according to the cognate chemokines and are referred to with the same rule as chemokines (CCR, CXCR, CX3CR and XCR). About 46 chemokines able to activate 18 chemokine receptors have been identified (Table 6.1).

Chemokine–chemokine receptor interactions present an important pleiotropic character: each chemokine may have multiple intracellular effects and in addition a high level of redundancy; each receptor can be activated by several ligands, and the same ligand can often recognise and bind several receptors. Chemokines are mostly involved in the migration of various cell types, especially lymphocytes, via chemotactic interactions aiming at attracting the cells equipped with the cognate receptors. Chemokines play a role as mediators of inflammation, with CXCL being responsible for neutrophil recruitment and CCL for macrophage recruitment. Two chemokine receptors, CXCR4 and CXCR5, are responsible for HIV entry in T cells and macrophages, contributing thus to virus infection. Chemokines also exert an effect on immunity and on tumour cell migrations, explaining their role in metastasis dissemination.

6.4.2 Chemokine Signalling

The type of signalling induced by chemokines is not different from that of other GPCR ligands: signalling involves large heterotrimeric G-proteins and effector

enzymes, especially adenylyl cyclase, phospholipase C and GDP–GTP exchange factors for small G-proteins. The ultimate targets of this signalling pathway are cell adhesion and cytoskeleton proteins that are implicated in cell motility. This explains the chemotactism processes that are operated after chemokine–chemokine receptor coupling. Chemokines are secreted in the extracellular space around a source cell, and they bind the cognate receptors of circulating target cells such as lymphocytes. These cells will in turn exert, at this location, their specific functions downstream receptor activation. These phenomena are not exclusive of the immune system where it was first discovered: chemokine action may occur at multiple levels, such as angiogenesis, development, etc.

Table 6.1 Chemokines and chemokine receptors

	Chemokine	Alias	Chemokine receptor
CCL chemokines (β type)	CCL1		CCR4, CCR8
	CCL2		CCR1, CCR2
	CCL3		CCR1, CCR5
	CCL3L1		CCR5
	CCL3L3		CCR5
	CCL4		CCR5, CCR8
	CCL4L1		CCR5
	CCL4L2		CCR5
	CCL5	RANTES	CCR1, CCR3, CCR5
	CCL7		CCR1, CCR2, CCR3, CCR5
	CCL8		
	CCL10		CCR1
	CCL11	Eotaxin	CCR3, CCR5
	CCL13		CCR1, CCR2, CCR3, CCR5
	CCL14		CCR1, CCR5
	CCL15		CCR1, CCR3
	CCL16		CCR1, CCR2
	CCL17	Dendrokin	CCR4
	CCL18		NI
	CCL19		CCR7, CCR11
	CCL20		CCR6
	CCL21	ECL	CCR3, CCR7, CCR11
	CCL22		CCR4
	CCL23		CCR1
	CCL24	Eotaxin 2	CCR3
	CCL25		CCR9, CCR11
	CCL26	Eotaxin 3	CCR3, CCR10
	CCL27	Eskine	CCR10
	CCL28	Skinkine	

(continued)

Table 6.1 (continued)

	Chemokine	Alias	Chemokine receptor
CXCL chemokines (α type)	CXCL1	MGSAα	CXCR2
	CXCL2	MGSAβ	
	CXCL3		
	CXCL4	*PF4*, platelet factor 4	CXCR3 (splice variant B)
	CXCL4L1	*PF4V1*	
	CXCL5		CXCR1, CXCR2
	CXCL6		
	CXCL7		
	CXCL8	IL8	
	CXCL9		CXCR3
	CXCL10		
	CXCL11		
	CXCL12	SDF1	CXCR4, CXCR7
	CXCL13		CXCR5
	CXCL14	Bolekine	
	CXCL15	Lungkine	
	CXCL16		CXCR6
CX3CL chemokines (δ type)	CX3CL1	Fractalkine	CX3CR1
XCL chemokines (γ type)	XCL1	Lymphotactin α	XCR1
	XCL2	Lymphotactin β	

Chemokines CCL6, CCL9, CCL10, CCL12 and CXCL15 were inferred from mouse orthologs and have not been identified in the human genome. There are some discrepancies in the correspondence chemokine–chemokine receptor in the literature, in relation to the level of affinity of the interactions
Abbreviation: ND, not determined

6.4.3 Oncogenic Alterations and Pharmacological Targets

Chemokines are involved in multiple pathological processes, in the fields of infectiology, pneumology or rheumatology. In oncology, chemokines are mostly involved in metastasis. An important role for the chemokines that are overexpressed by tumour cells is to recruit lymphocytes and macrophages to the tumour, so as to maintain a chronic inflammatory state. CCL5 or RANTES (*regulated on activation, normal T cell expressed and secreted*) is a chemokine involved in macrophage recruitment. Macrophage activation induces the production of metalloproteinases (MMP) whose activity is crucial for metastasis development, because they hydrolyse the mesenchymal matrix of tumour stroma. Another tumour chemokine, CXCL12 or SDF1 (*stromal cell-derived factor 1*), is able to recruit endothelial cells that contribute to neovessel development (angiogenesis), because the progenitors of these cells express the SDF1 receptor, CXCR4. Finally, the CXCL4 (*PF4*)–CXCR3 axis plays a major angiostatic role.

Another aspect of chemokine function in oncology concerns tumour cell motility, thanks to the chemokine receptors that they often express to a high level, such

as CXCR4 and CCR7. Receptor activation enables these cells to also secrete metalloproteinases. In addition, CXCL12, which recognises and binds CXCR4 and CXCR7, is highly expressed in the lungs, liver and bone marrow, which suggests that it may be responsible for tumour cell metastasis to these organs, especially because CXCR7 appears to be only expressed in tumour cells. CCL21 or ECL (*efficient chemoattractant for lymphocytes*), which is the main ligand of CCR7, is in contrast highly expressed in lymph nodes and thus would be responsible for tumour cell attraction to them. CXCL8 (IL8, *interleukin 8*) and its receptors CXCR1 and CXCR2 are involved in oncogenesis and metastasis development and so are the CXCL16–CXCR6, CXCL1/2–CXCR2 and CCL2–CCR1/2 axes.

Finally, chemokine receptor activation is susceptible to favour their multiplication: we have mentioned earlier that GPCR activation could activate PI3Kγ pathway as well as some MAP kinase modules.

Several pharmaceutical laboratories have engaged into the development of drugs targeting chemokines or chemokine receptors. This research benefits from the fact that chemokines are involved in several pathological processes and are of interest in several therapeutic domains. However, the problems due to the pleiotropic and redundant aspects of this pathway have slowed down this development. A first approach consisted in the synthesis of cyclic peptides mimicking CXCL12 and blocking its receptor, CXCR4. This was initiated in the AIDS field with antileukinate but was further developed in oncology; several molecules are presently in clinical trials. Another approach is the identification, from high-throughput screening, of small molecules able to inhibit chemokine receptors such as CXCR4, CXCR1, CXCR2 and CXCR7. Such molecules have shown preclinical activity, and some have entered clinical trials: reparixin, which prevents CXCR1 binding to its ligand, CXCL8 (IL8), is presently evaluated in early breast cancer. Targeting chemokines themselves is conceivable with aptamers able to trap them before receptor interaction and with small molecules such as plerixafor, an anti-CXCL12 molecule. Finally, in view of the success of many antibodies as therapeutic agents, chemokine and chemokine receptor targeting by monoclonal antibodies should provide a series of active antimetastatic agents; carlumab is an anti-CCL2 monoclonal antibody evaluated in prostate cancer, and BMS-936564 is an anti-CXCR4 monoclonal antibody presently in phase I.

Further Reading

Allen SJ, Crown SE, Handel TM. Chemokine: receptor structure, interactions, and antagonism. Annu Rev Immunol. 2007;25:787–820.

Audet M, Bouvier M. Restructuring G-protein–coupled receptor activation. Cell. 2012;151: 14–23.

Balakin KV, Ivanenkov YA, Tkachenko SE, Kiselyov AS, Ivachtchenko AV. Regulators of chemokine receptor activity as promising anticancer therapeutics. Curr Cancer Drug Targets. 2008;8:299–340.

Balkwill F. Cancer and the chemokine network. Nat Rev Cancer. 2004;4:540–50.

Berridge MJ. Inositol trisphosphate and calcium signalling mechanisms. Biochim Biophys Acta. 2009;1793:933–40.

Deng L, Chen N, Li Y, Zheng H, Lei Q. CXCR6/CXCL16 functions as a regulator in metastasis and progression of cancer. Biochim Biophys Acta. 1806;2010:42–9.
Dhanasekaran N, Tsim ST, Dermott JM, Onesime D. Regulation of cell proliferation by G proteins. Oncogene. 1998;17:1383–94.
Domanska UM, Kruizinga RC, Nagengast WB, et al. A review on CXCR4/CXCL12 axis in oncology: no place to hide. Eur J Cancer. 2013;49:219–30.
Dorsam RT, Gutkind JS. G-protein-coupled receptors and cancer. Nat Rev Cancer. 2007;7:79–94.
Entschladen F, Zänker KS, Powe DG. Heterotrimeric G protein signaling in cancer cells with regard to metastasis formation. Cell Cycle. 2011;10:1086–91.
Fraser CC. G protein-coupled receptor connectivity to NF-kappaB in inflammation and cancer. Int Rev Immunol. 2008;27:320–50.
Garin A, Proudfoot AE. Chemokines as targets for therapy. Exp Cell Res. 2011;317:602–12.
Heasley LE. Autocrine and paracrine signaling through neuropeptide receptors in human cancer. Oncogene. 2001;20:1563–9.
Houshmand P, Zlotnik A. Therapeutic applications in the chemokine superfamily. Curr Opin Chem Biol. 2003;7:457–60.
Karnoub AE, Weinberg RA. Chemokine networks and breast cancer metastasis. Breast Dis. 2006–2007;26:75–85.
Kehrl JH. Chemoattractant receptor signaling and the control of lymphocyte migration. Immunol Res. 2006;34:211–27.
Khan SM, Sleno R, Gora S, et al. The expanding roles of Gβγ subunits in G protein-coupled receptor signaling and drug action. Pharmacol Rev. 2013;65:545–77.
Lagerstrom MC, Schioth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. Nat Rev Drug Discov. 2008;7:339–57.
Lappano R, Maggiolini M. G protein-coupled receptors: novel targets for drug discovery in cancer. Nat Rev Drug Discov. 2011;10:47–60.
Marinissen MJ, Gutkind JS. G-protein-coupled receptors and signaling networks: emerging paradigms. Trends Pharmacol Sci. 2001;22:368–76.
O'Hayre M, Vázquez-Prado J, Kufareva I, et al. The emerging mutational landscape of G proteins and G-protein-coupled receptors in cancer. Nat Rev Cancer. 2013;13:412–24.
Olson TS, Ley K. Chemokines and chemokine receptors in leukocyte trafficking. Am J Physiol. 2002;283:R7–28.
Pisanti S, Picardi P, D'Alessandro A, Laezza C, Bifulco M. The endocannabinoid signaling system in cancer. Trends Pharmacol Sci. 2013;34:273–82.
Teicher BA, Fricker SP. CXCL12 (SDF-1)/CXCR4 pathway in cancer. Clin Cancer Res. 2010;16:2927–31.
Wang Z, Huang H. Platelet factor-4 (CXCL4/PF-4): an angiostatic chemokine for cancer therapy. Cancer Lett. 2013;331:147–53.
Waugh DJ, Wilson C. The interleukin-8 pathway in cancer. Clin Cancer Res. 2008;14:6735–41.
Zlotnik A, Yoshie O, Nomiyama H. The chemokine and chemokine receptor superfamilies and their molecular evolution. Genome Biol. 2006;7:243.
Zlotnik A. Chemokines and cancer. Int J Cancer. 2006;119:2026–9.

Wnt Pathway

7

Abstract

The term Wnt was forged by contraction of ***Win****gless-type mouse mammary tumour virus* ***int****egration site* after the discovery on a viral oncogene in mice which was homologous to a mutant *Drosophila* gene. The Wnt pathway, especially under its canonical form (Wnt–β-catenin pathway), is a signalling pathway involved in vertebrate and invertebrate development. It plays a major role in embryogenesis and morphogenesis and in stem cell commitment to differentiation or proliferation. This pathway is relatively complex, and all its processes and physiological roles are not completely deciphered. The proteins involved in this pathway are called 'Wingless', 'Frizzled', 'Dishevelled', which originally designated the mutants of Drosophila development and demonstrates the universality of this pathway in the animal kingdom. Germline alterations of this pathway in humans determine congenital hereditary diseases, and some somatic and germinal alterations are associated with oncogenesis.

Briefly, protein messengers called WNT activate GPCR-related membrane receptors called Frizzled (FZD), and this activation leads to the stabilisation of the cytoplasmic form of a protein involved in intercellular junctions of epithelial tissues, β-catenin. The stabilised protein can then be transferred into the nucleus where it activates transcription programmes of numerous genes, especially those encoding the proteins required for cell proliferation and cell cycle entry (cyclin D, MYC, etc.).

7.1 WNT Ligands and Their Receptors

Nineteen distinct WNT proteins are found in humans. The first of these genes, WNT1, had been identified at the integration site of a retrovirus inducing mammary tumours in mouse and called *Int2*, showing thus a potential role of this gene in normal development and oncogenesis; an ortholog of this gene, called *Wingless*, had

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_7

been thereafter identified in *Drosophila*: the Wnt name results from the contraction of *Wingless* and *Integration*.

WNT proteins are cysteine-rich glycoproteins secreted by diverse cell types; their genes are organised in several clusters localised on different chromosomes and present high homology. Basic gene sequences are well conserved in all metazoans. WNT proteins are expressed in a tissue-specific manner, but the precise role of each WNT protein is yet poorly understood.

WNT ligands are recognised by *Frizzled* receptors (FZD); there are 10 FZD receptors in humans, but the combinatorial pattern between the 19 ligands and the 10 receptors is not entirely known. FZD receptor activation requires the presence of a phosphorylable coreceptor, LRP5 or LRP6 (*LDL receptor-related protein*). FZDs are seven-transmembrane domain receptors or GPCRs (*G-protein-coupled receptors*, Chap. 6); the concomitant activation of an FZD receptor with LRP5/6 coreceptor leads to the formation of a complex with a cytoplasmic protein called *Dishevelled* (DVL), an adapter protein whose activation allows the break of a cytoplasmic complex including the central element of this pathway, β-catenin. Three DVL proteins, DVL1, 2 and 3, exist in humans. It was believed that the FZD receptors function without coupling with a large heterotrimeric G-protein, contrasting to other GPCRs. It seems in fact that such G-proteins operate between FZD and DVL activation, similarly as they operate downstream the activation of all GPCR.

7.2 Wnt–β-Catenin Pathway

β-catenin (gene *CTNNB*) is a molecule involved in both cell adhesion and signalling. In the first function, β-catenin is bound, in the cytoplasm, to the actin cytoskeleton via a molecule of α-catenin and, in the extracellular space, to the junctional structures of E-cadherin (CDH1) (Fig. 7.1) (Chap. 11). β-catenin plays, therefore, a major role in adherens junctions of epithelial cells and in the maintenance of tissue architecture.

Free cytoplasmic β-catenin, which is in equilibrium with bound β-catenin at adherens junctions, is maintained in an inactive state, in the absence of WNT signals, thanks to its phosphorylation, which drives it to proteasomal destruction. This phosphorylation is ensured within a degradation complex, in which intervene adapter proteins, axin and APC (*adenomatous polyposis coli*), and two serine/threonine kinases, CK1α (*casein kinase 1α*, gene *CSNK1A*) and GSK3β (*glycogen synthase kinase 3β*, gene *GSK3B*), which can add phosphate groups on two different sites. β-catenin phosphorylation enables its binding to β-TRCP (*beta-transducing repeat-containing protein*), which is its cognate E3 ubiquitin ligase and drives it to proteasomal destruction (see Annex C) (Fig. 7.2a).

Frizzled receptor-induced DVL activation, likely through a large heterotrimeric G-protein, induces the phosphorylation of the coreceptor LRP5/6, on a PPPSPXS motif, by a membrane casein kinase (CK1γ, gene *CSNK1G1*) and by GSK3β, which can no longer phosphorylate β-catenin (Fig. 7.2b). This phosphorylation generates

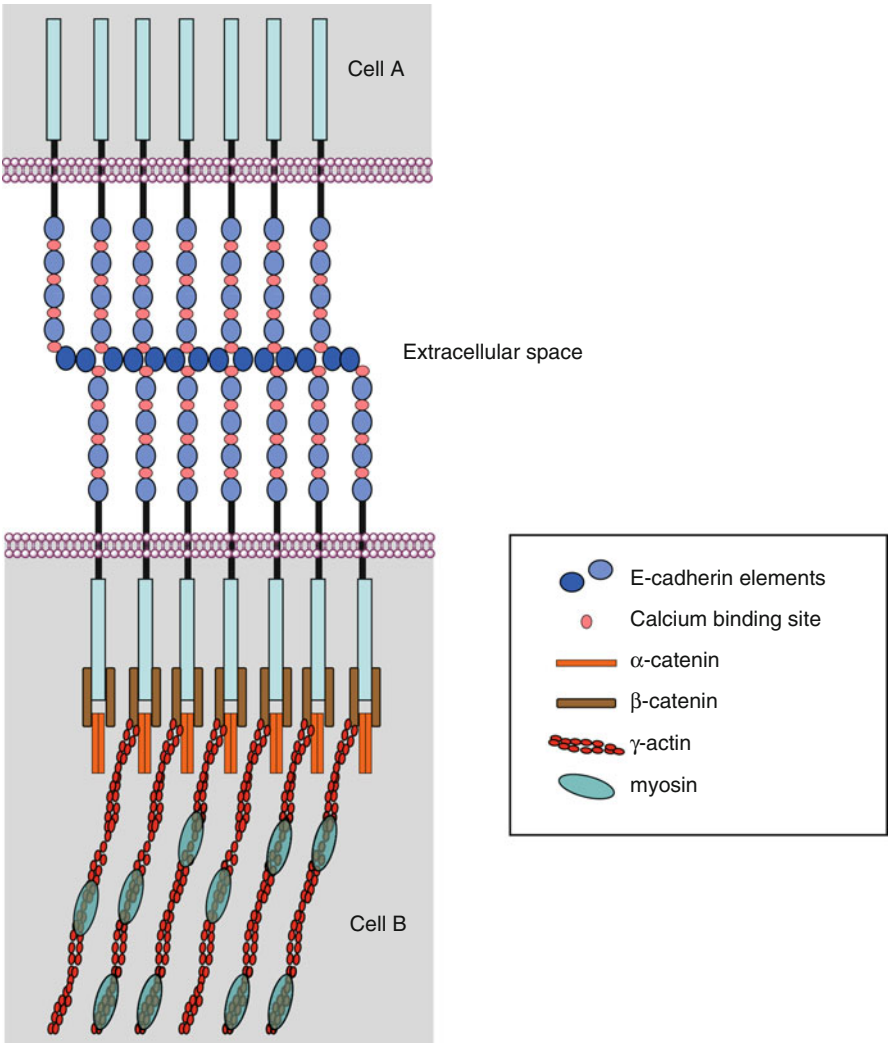


Fig. 7.1 Intercellular junctions involving E-cadherin. Two types of epithelial cell intercellular junctions involve E-cadherin: adherens junctions and desmosomes. These intercellular junctions mobilise β-catenin, which allow the anchoring of E-cadherin to the actin cytoskeleton, via a molecule of α-catenin. The E-cadherin molecules are associated in the extracellular space via Ca^{2+} ions; their intracellular domain is attached to β-catenin and α-catenin molecules, which tie them to the actin cytoskeleton

a signal that induces the relocalisation of axin to the membrane, releasing then β-catenin from the cytoplasmic destruction complex which was until then driving it to the proteasome in the absence of receptor activation. This Wnt pathway, called the 'canonical pathway', appears to be most often activated after receptor binding of specific WNT ligands, WNT1, WINT3A and WNT8.

Free, non-phosphorylated β-catenin can then enter the nucleus, where it is able to move a transcription factor of the TCF/LEF (*T-cell factor/lymphoid enhancer factor*) family from its DNA binding sites and thus trigger the transcription of genes that were until then repressed (Fig. 7.2b). Among the genes that are thus transcribed are *MYC*, which encodes a transcription factor involved in cell proliferation (Chap. 2), and cyclin D1 (*CCND1*), which is involved in cell cycle entry (Chap. 17). When it gets out of the nucleus, β-catenin is recaptured by APC and, in the absence of a new WNT signal, reinserted into a destruction complex.

Other systems appear to be able to indirectly activate β-catenin pathway, by inducing the break of the β-catenin–E-cadherin bond and the increase in cytoplasmic β-catenin availability for nucleus entry; this is achieved by tyrosine phosphorylation of β-catenin by tyrosine kinase receptors such as EGFR, ERBB2 or MET (Chap. 1).

More than hundred proteins are involved, with variable importance, in Wnt signalling. In addition to those already mentioned, one can mention:

- FRAT (*frequently rearranged in advanced T-cell lymphoma*) proteins, which inhibit β-catenin phosphorylation by GSK3β.

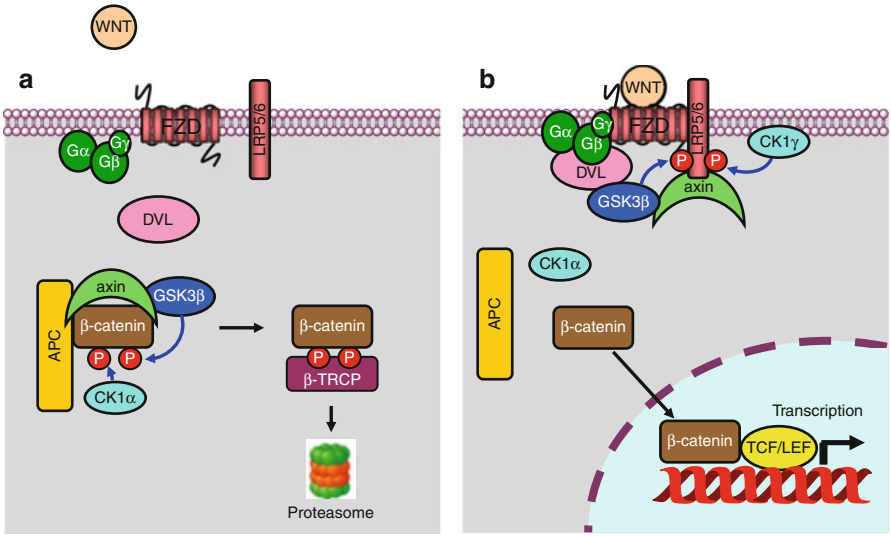


Fig. 7.2 The Wnt–β-catenin pathway. (**a**) In the absence of stimulation of the Frizzled (*FZD*) receptor by a WNT ligand, β-catenin is phosphorylated in a destruction complex-containing axin, APC and the GSK3β and CK1α kinases. This phosphorylation leads β-catenin to its E3 ubiquitin ligase, β-TRCP, and then to the proteasome. (**b**) In the presence of a WNT ligand bound to the FZD receptor, a dishevelled (DVL) protein is recruited, likely after activation of a large heterotrimeric G-protein, allowing thus the phosphorylation of the LRP5/6 coreceptor by CK1γ and GSK3β and the recruitment of axin to the membrane. The destruction complex is dissociated, and non-phosphorylated β-catenin can enter the nucleus, where it associates to a transcription complex including the TCF/LEF factors. The derepression of this complex allows the transcription of numerous genes involved in cell proliferation, among which MYC and CCND1 are relevant examples

- SFRP (*secreted Frizzled-related proteins*, genes *SFRP1* to *SFRP5*), which are decoy receptors for WNT ligands in the extracellular space; they divert the ligands from the receptor target and negatively regulate the pathway.
- WIF1 (*Wnt inhibitory factor 1*), which binds to the WNT proteins in the extracellular space and inhibit them.
- DKK (*Dickkopf homologues*, genes *DKK1* to *DKK4*) proteins, which also are secreted proteins interacting with the coreceptors LRP5/6, inducing their endocytosis and the deactivation of FZD receptors.

7.3 'Non-canonical' Wnt Pathways

Non-canonical Wnt pathways are signalling pathways originating from the interaction of WNT proteins with an FZD receptor but not involving β-catenin. In particular, these pathways are operated downstream the WNT5A and WNT11 ligands. The first one activates phospholipase C pathway after activation of a large heterotrimeric G-protein (subunit α_q), as commonly seen with GPCRs (Chap. 6), with the involvement of a DVL protein. This induces the cytosolic release of Ca^{2+} via IP3 synthesis and the activation of protein kinase C via DAG formation (Fig. 7.3a), as described in Chap. 6.

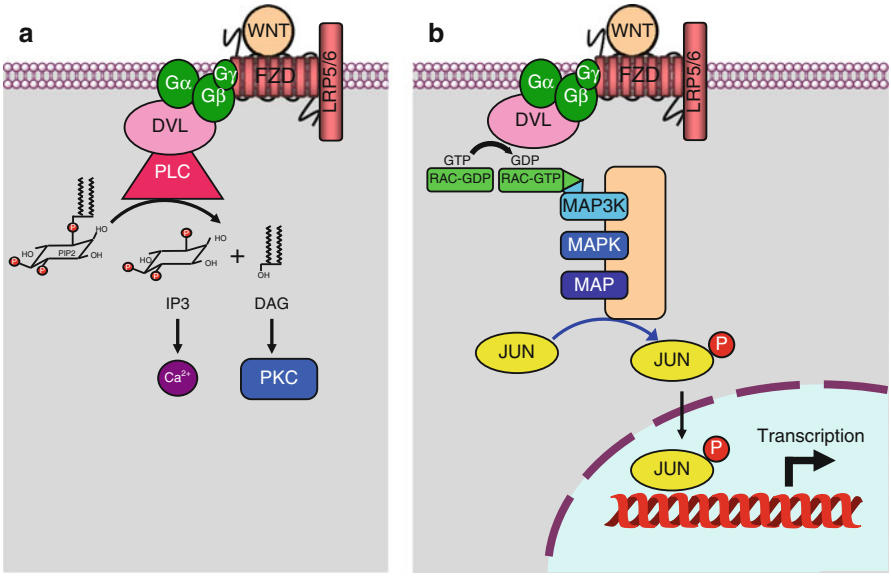


Fig. 7.3 Non-canonical Wnt pathways. These signalling pathways are activated by a WNT signal but do not involve β-catenin. (**a**) DVL induces the activation of phospholipase C, which generates from phosphatidylinositol 4,5-bisphosphate the second messengers diacylglycerol (DAG) and trisphosphoinositol (IP3), activating the pathways described in Chap. 6. (**b**) DVL induces the activation of small G-proteins such as RAC1, which allows the activation of a MAP kinase cascade leading to the phosphorylation of transcription factors and to the activation of proliferation genes (Chap. 2)

The second non-canonical pathway utilises the adapter protein DVL and a large heterotrimeric G-protein with an α_{12} subunit, for the activation of JUN *N*-terminal kinase (JNK), involved in one of the MAP kinase modules (Chap. 2), showing thus a good example of signalling pathway interconnections. This activation is done via small G-proteins of the RHO family, such as RHOA, RHOU, RAC or CDC42 (Fig. 7.3b) and governs cell orientation and planar cell polarity (PCP). Unusual tyrosine kinase receptors, ROR and RYK (Chap. 1), are coreceptors, or annex receptors, to the WNT proteins and are able to activate or deactivate, through ligand sequestration, this non-canonical pathway.

7.4 Oncogenic Alterations

There are numerous germline alterations in this pathway, leading to abnormalities of the embryonic and foetal development and determining congenital malformations. Oncogenic alterations may concern positive regulators of this pathway, which behave as proto-oncogenes or metastasis promoters (WNT ligands, especially WNT1, WNT3A and WNT7A, FZD receptors, LRP5/6 coreceptors), and also negative regulators, which behave as tumour suppressor genes (other WNT ligands, for instance, WNT5A and WNT11, APC, AXIN, GSK3β, DKKs, SFRPs).

In human cancers, numerous somatic alterations have been identified; some are well-defined mutations, in the β-catenin gene *CTNNB*, in *APC* or *AXIN2*; others are alterations in gene expression, especially concerning WNT ligands, FZD receptors or DVL adapter proteins, whose driver role in oncogenesis has not always been demonstrated.

The *APC* gene has been identified long ago as a tumour suppressor gene playing a major role in colorectal oncogenesis. Its germline mutations determine familial *adenomatous polyposis coli*, one of the major syndromes of colorectal cancer hereditary predisposition. When somatic, its mutations constitute, in the Vogelstein model, the most frequent initiating events. Generally, these are nonsense mutations leading to a truncated protein. It seems that the mutation of only one allele would be enough to constitute this cancer-initiating event, by haploinsufficiency.

Colorectal cancers are especially dependent, for their genesis as for their development, upon the β-catenin pathway. In addition to the *APC* mutations, those occurring in *AXIN2* or in the β-catenin gene *CTNNB* are frequent. Molecular alterations of this pathway are also found in hepatocarcinomas, ovary cancers and desmoid and fibromatous sarcomas. The Wnt pathway is especially involved in stem cell proliferation, but also in angiogenesis and metastasis, especially by controlling the equilibrium between cell adhesion and migration, a process where β-catenin plays a major role.

7.5 Pharmacological Targets

The β-catenin pathway represents a major potential target in oncology, in particular for colorectal cancers which keep, all along their evolution, a marked addiction to this pathway. We will not mention untargeted approaches and the potential role of

anti-inflammatory nonsteroid compounds or of vitamins A and D on the possible inhibition of this pathway.

A first target can be found at the level of WNT protein production, which is under the control of a protein called porcupine (gene *PORCN*), an acyltransferase required for their secretion. A small molecule inhibitor of porcupine has been identified by screening; IWP-1 (*inhibitor of WNT production*) and others have been synthesised for clinical trials.

Targeting WNT ligands and FZD receptors can be achieved with monoclonal antibodies. However, considering the facts that there are 19 WNT ligands, 10 FZD receptors and 2 LRP coreceptors and that their combinatorial pattern is not completely deciphered, it might be difficult to identify valuable anticancer compounds. In preclinical models, anti-WNT3A, anti-FZD7 and anti-LRP6 monoclonal antibodies have shown anticancer activity. Vantictumab, an anti-FZD7 antibody, has entered clinical trials. Peptidomimetics able to play an agonistic or an antagonistic role are also in development; this is the case for mimetics of SFRP1, which trap WNT factors in the extracellular space, and for FZD2 and FZD7 mimetics, which inhibit Wnt signalling.

At the intracellular level, small molecules identified by high-throughput screening have been selected for further studies. Several potential targets can be identified at the different levels of the signalling pathway: Iinterference with LRP5 adapter role, inhibition of FZD–DVL interaction, stabilisation of the β-catenin destruction complex, inhibition of β-catenin transport to the nucleus and prevention of TCF–β-catenin association are among the various tracks that are followed by the pharmaceutical industry.

Further Reading

Anastas JN, Moon RT. WNT signalling pathways as therapeutic targets in cancer. Nat Rev Cancer. 2013;13:11–26.

Angers S, Moon RT. Proximal events in Wnt signal transduction. Nat Rev Mol Cell Biol. 2009;10:468–77.

Arce L, Yokoyama NN, Waterman ML. Diversity of LEF/TCF action in development and disease. Oncogene. 2006;25:7492–504.

Barker N, Clevers H. Mining the Wnt pathway for cancer therapeutics. Nat Rev Drug Discov. 2006;5:997–1014.

Clevers H, Nusse R. Wnt/β-catenin signaling and disease. Cell. 2012;149:1192–205.

Fodde R, Brabletz T. Wnt/beta-catenin signaling in cancer stemness and malignant behavior. Curr Opin Cell Biol. 2007;19:150–8.

Gavert N, Ben-Ze'ev A. beta-Catenin signaling in biological control and cancer. J Cell Biochem. 2007;102:820–8.

Herr P, Hausmann G, Basler K. WNT secretion and signalling in human disease. Trends Mol Med. 2012;18:483–93.

Huang H, He X. Wnt/beta-catenin signaling: new (and old) players and new insights. Curr Opin Cell Biol. 2008;20:119–25.

Jeanes A, Gottardi CJ, Yap AS. Cadherins and cancer: how does cadherin dysfunction promote tumor progression? Oncogene. 2008;27:6920–9.

Katoh M, Katoh M. WNT signaling pathway and stem cell signaling network. Clin Cancer Res. 2007;13:4042–5.

Kikuchi A, Yamamoto H. Tumor formation due to abnormalities in the beta-catenin-independent pathway of Wnt signaling. Cancer Sci. 2008;99:202–8.
Kimelman D, Xu W. beta-catenin destruction complex: insights and questions from a structural perspective. Oncogene. 2006;25:7482–91.
Klaus A, Birchmeier W. Wnt signalling and its impact on development and cancer. Nat Rev Cancer. 2008;8:387–98.
Mikels AJ, Nusse R. Wnts as ligands: processing, secretion and reception. Oncogene. 2006;25: 7461–8.
Niehrs C. The complex world of WNT receptor signalling. Nat Rev Mol Cell Biol. 2012;13: 767–79.
Polakis P. The many ways of Wnt in cancer. Curr Opin Genet Dev. 2007;17:45–51.
Takahashi-Yanaga F, Kahn M. Targeting Wnt signaling: can we safely eradicate cancer stem cells? Clin Cancer Res. 2010;16:3153–62.
Takebe N, Harris PJ, Warren RQ, Ivy SP. Targeting cancer stem cells by inhibiting Wnt, Notch, and Hedgehog pathways. Nat Rev Clin Oncol. 2011;8:97–106.
Wang Y. Wnt/Planar cell polarity signaling: a new paradigm for cancer therapy. Mol Cancer Ther. 2009;8:2103–9.
Wend P, Holland JD, Ziebold U, Birchmeier W. Wnt signaling in stem and cancer stem cells. Semin Cell Dev Biol. 2010;21:855–63.

Notch Pathway

8

Abstract

The Notch pathway, as the Wnt–β-catenin pathway, is an important signalling pathway in vertebrate and invertebrate development. It is involved in embryogenesis and morphogenesis and plays a major role in the fate of stem cells to differentiation. The loss of function of some of its actors leads to hereditary malformations, and activating mutations and/or overexpression are found in malignant diseases.

Briefly, protein messengers called DSL (*Delta*, *Serrate*, *Lag2*, after the names of the ligands found, respectively, in mammals, *Drosophila* and *Caenorhabditis elegans*) recognise and activate membrane receptors called NOTCH, this activation consisting in receptor cleavage. The cleavage products can then migrate into the nucleus and activate the transcription of target genes through inhibition of transcriptional repressors. One of the important features in this pathway is the juxtacrine interaction of ligands and receptors, both ligands and receptors being transmembrane proteins: the presentation of a ligand to a neighbour cell induces the activation of the receptor cell.

8.1 DSL Ligands

Two families of DSL ligands can recognise the NOTCH receptors in mammals: DLL (*delta-like ligands*) and JAG (*jagged*), which correspond to the *Serrate* ligands of *Drosophila*. These ligands are type I transmembrane proteins (i.e., their *N*-terminal part is in the extracellular space).

DSL ligands have a short intracytoplasmic domain, a single transmembrane domain and a large characteristic extracellular domain, made of 6–14 tandem repeats of EGF-like domains, a DOS (*delta and OSM-11-like*) domain absent from DLL3 and DLL4 and a DSL domain. DLL3 and DLL4, devoid of a DOS domain, require a DLK (*delta-like protein homologue*) coligand, equipped with DOS and

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_8

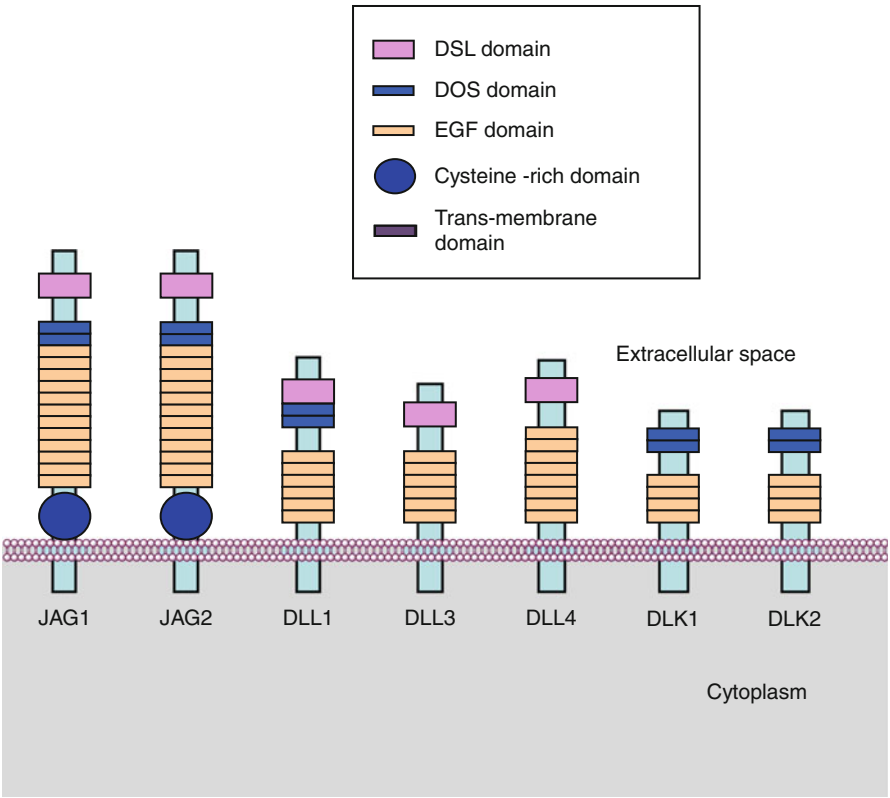


Fig. 8.1 DSL ligands. DSL ligands contain a series of domains that are essential for their function: the DOS and DSL recognition domains and 6–14 EGF-like domains. The JAG ligands contain in addition a cysteine-rich domain. DLL3 and DLL4 have no DOS domain and depend upon the presence of the coligands DLK1 and DLK2, which harbour a DOS domain but no DSL domain

EGF-like domains but devoid of a DSL domain. The simultaneous presence of DOS and DSL domains appears necessary for the ligand–receptor interaction, the DOS domain being either directly brought by the ligand (for JAG1, JAG2 and DLL1) or by a coligand (for DLL3 and 4). The protein structure of ligands and coligands of the Notch pathway is presented on Fig. 8.1.

8.2 NOTCH Receptors

DSL ligands are recognised by NOTCH receptors; four NOTCH receptors exist in humans, which are also type I transmembrane proteins. Whereas the protein sequence in *Drosophila* is continuous and covers the extra- and intracellular parts, NOTCH receptors are heterodimeric in mammals, the monomers being generated in the Golgi apparatus by post-translational cleavage of a precursor containing the

whole sequence, followed by non-covalent reassociation of the two monomers. The cleavage is realised by furin, a member of the subtilisin-like proprotein convertase family at a site called S1, and the resulting monomers are called NECD (*notch extracellular domain*) and NTMICD (*notch transmembrane and intracellular domain*).

In their extracellular part, the NOTCH receptors consist of 29–36 tandem repeats of EGF-like motifs, which contain *O*-glycosylation sites for fucose or glucose residues. The EGF-like motifs are followed by a NRR (*negative regulatory region*) sequence consisting of three tandem repeats of LNR (*lin12 and notch repeats*) motifs and a heterodimerisation (HD) domain where the S1 cleavage site is located. After the transmembrane domain, the intracellular part contains a RAM (RBPJK [*recombination signal binding protein for immunoglobulin kappa J region*]–*association module*) motif, followed by nuclear localisation sequences (NLS), ankyrin repeats (ANK) and a transactivation domain harbouring PEST (*proline–glutamic acid–serine–threonine-rich*) domains. Figure 8.2 presents the general structure of the NOTCH receptors.

8.3 NOTCH Receptor Activation and Signal Transmission

Ligand–receptor binding occurs by homotypic interaction of the EGF-like domains and induces the proteolytic cleavage of the receptor by a metalloproteinase, ADAM10 (*a disintegrin and metalloproteinase*) or ADAM17 (also known as TACE, *TNF-alpha converting enzyme*), at the level of the site S2 located within the heterodimerisation domain, thus in the extracellular part of the receptor, 12 amino acid residues ahead of the transmembrane domain. This cleavage is possible thanks to a conformational change releasing the S2 site from the protection exerted by the LNR domains. The intermediate protein is called NEXT (*notch extracellular truncation*) and is the substrate of the proteolytic activity exerted by the γ-secretase complex, at two transmembrane sites called S3 and S4. After cleavage, the NOTCH intracellular domain (NICD) is able to migrate to the nucleus, thanks to the unmasking of the NLS. γ-secretase is a multimeric transmembrane complex harbouring presenilins 1 and 2 (genes *PSEN1* and *2*), nicastrin (gene *NCSTN*) and activation and stabilisation factors of these proteases.

In the nucleus, NICD interacts, thanks to its RAM domain, with a DNA-binding protein of the CSL (CBF [*Core-binding factor-1*] */su(H)/lag-1*) family, which is RBPJK in humans. The ANK domain then recruits a coactivator called MAML (*mastermind-like protein*), which in turn recruits the MED8 (*mediator of RNA polymerase II transcription*, *subunit 8*) factor, allowing the transactivation of target genes. Figure 8.3 presents the steps of the post-translational maturation of the NOTCH receptors, interaction with DSL ligands, proteolysis, migration of the active part to the nucleus and interaction with activating and repressing transcriptional complexes.

The NOTCH receptor, which must be cleaved and whose intracellular part constitutes in fact the second messenger, can be used only once: ligand and receptor

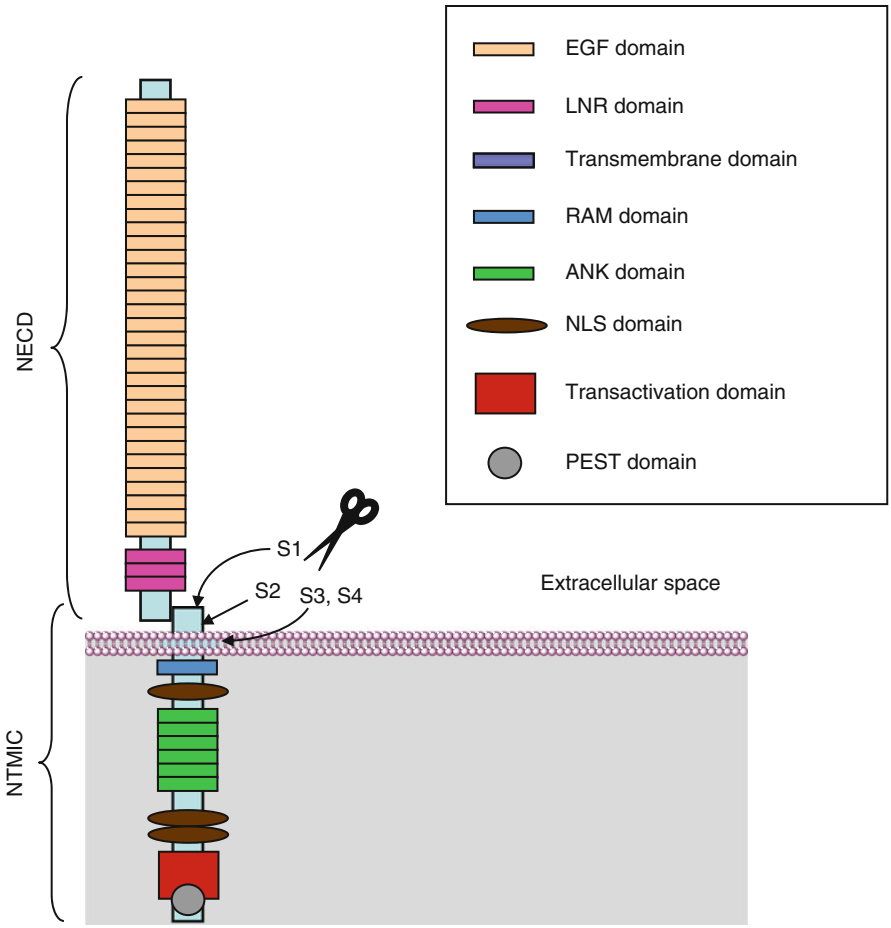


Fig. 8.2 NOTCH receptors. The four NOTCH heterodimeric receptors are built on the same model and associate an extracellular domain (*NECD*) and a transmembrane and intracellular domain (NTMICD), tied together via a non-covalent interaction at the level of a heterodimerisation domain HD; the two monomers result from the cleavage of a unique precursor protein, which is cleaved within the HD domain in the Golgi apparatus by furin. NECD contains 29–36 EGF-like domains and three LNR domains involved in the negative regulation of the Notch signalling pathway by protecting the S2 cleavage site of the HD domain. NTMICD contains a RAM domain, involved in the interaction with DNA-binding proteins, together with nuclear localisation sequences NLS, ankyrin repeats ANK involved in the recruitment of the transcriptional activator MAML and a transactivation domain containing PEST motifs

availability at the surface of the two cells exchanging a message is, therefore, a decisive factor for signalling. Synthesis and degradation of the ligands and receptors are the key elements of this regulation. Ubiquitinylation plays a crucial role for ligand and receptor endocytosis and turnover (see Annex C).

The heterodimerisation HD domain, containing the successive proteolytic cleavage sites, is certainly the most sensitive domain of the NOTCH receptors. Glycosylation of

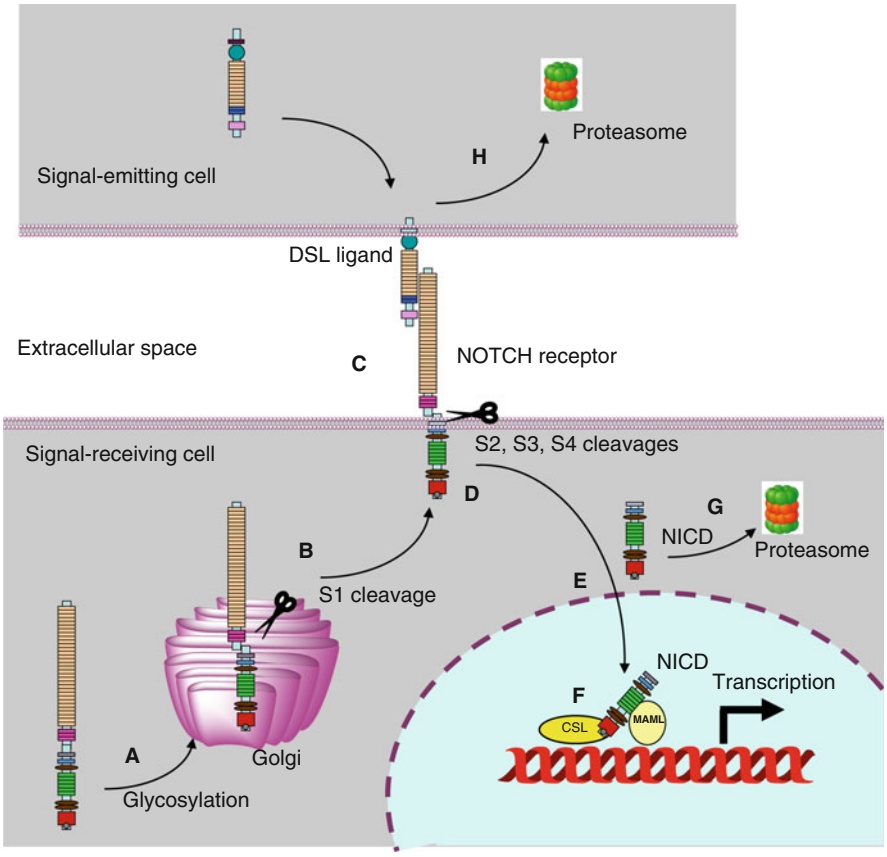


Fig. 8.3 Notch signalling pathway. (**A**) After synthesis, the NOTCH receptor is fucosylated and glycosylated on serine and threonine residues of EGF-like domains in the endoplasmic reticulum. (**B**) In the Golgi apparatus, it is cleaved by furin at the S1 site of the HD domain, but the two cleavage products, NECD and NTMICD, remain associated by non-covalent bonds localised at the level of the HD domain and thus form a heterodimer. (**C**) Once in the plasma membrane, the NOTCH receptor can be recognised by the DSL ligands of a neighbour cell by homotypic interaction involving the EGF-like domains. (**D**) This recognition induces the unmasking of the cleavage site S2 of the HD domain of NTMICD by an ADAM proteinase, allowing then a new cleavage at the S3 and S4 transmembrane sites by the γ-secretase proteolytic complex. (**E**) The truncated receptor, NICD, can then migrate to the nucleus thanks to its nuclear localisation sites (NLS). (**F**) NICD is recognised in the nucleus, at the level of its RAM domain, by a CSL protein, which binds it to DNA sequences. (**F**) This allows the recruitment of a MAML protein, which binds to the ANK domains of NICD, so that transcription of target genes can occur. (**G**) NICD is then destroyed in the proteasome after ubiquitinylation, thanks to the PEST domains localised on its **C**-terminal part. (**H**) Similarly, the DSL ligands are digested in the proteasome after ubiquitinylation

the EGF-like motifs of the NOTCH receptors also represents an important feature for the regulation of this signalling pathway. Finally, numerous crosstalks with other signalling pathways have been described at the level of the transcription of target genes: for instance, the JAG1 ligand is the product of a gene activated by β-catenin (Chap. 7).

8.4 Oncogenic Alterations

The Notch pathway plays a crucial role in the multiplication of stem cells in intestinal villi and breast cancers; its inhibition allows the reorientation of these cells toward differentiation. Among the NOTCH target genes are *SNAI1* (SNAIL) and *SNAI2* (SLUG), which are transcriptional repressors of E-cadherin, an epithelial adhesion protein whose expression must be inhibited to allow cell migration (Chap. 7). The Notch pathway is thus responsible for the crucial steps of the epithelial-to-mesenchymal transition (EMT). The Notch pathway is also involved in an important step of angiogenesis: the formation of vascular ramifications. The endothelial tip cells synthesise, upon VEGF impulse, a DSL ligand able to activate the proliferation of another type of endothelial cell, the stalk cells, which constitute the new vessel branches and their anastomoses.

The Notch pathway appears as an essential oncogenic pathway, and its role in oncogenesis is certainly far from being completely understood. Gain-of-function alterations in the Notch signalling pathway have been initially observed in T-cell lymphomas and lymphoblastic leukaemias. Thereafter, such alterations have been identified in other malignancies, especially colon and prostate cancers. The role of the Notch pathway in the fate of stem cells is certainly a key factor for the interest of this pathway in oncology.

A rare t(7;9) translocation in T-cell lymphoblastic leukaemias allowed the identification, as early as 1991, of the oncogenic role of a truncated, constitutively active NOTCH receptor. Later, numerous activating mutations have been identified, especially at the level of the HD and PEST domains. Mutations in the HD domain induce weakening of the interaction between NECD and NTMICD: this facilitates the subsequent cleavage at the S2 and S3–S4 sites for spontaneous release of NICD, without DSL ligand interaction. Furthermore, mutations and deletions of the *C*-terminal region may remove the NICD binding sites for its E3 ubiquitin ligase and stabilise the active form of the receptor at the level of its nuclear site of action.

8.5 Pharmacological Targets

Pharmacological interventions on the successive steps of the NOTCH receptor activation aim at the inhibition of this signalling pathway in stem cell reproduction, oncogenesis and angiogenesis. Potential targets are found at the level of ligand–receptor interactions, receptor glycosylation, initial cleavage generating NECD and NTMICD, subsequent ADAM- and γ-secretase-mediated cleavages, receptor or ligand ubiquitinylation, NICD interaction with nuclear proteins, etc.

Up to now, the essential target for pharmacological development has been the transmembrane cleavage mediated by the γ-secretase complex. γ-secretase inhibitors had been designed and developed for the treatment of Alzheimer disease, such as semagacestat and avagacestat, because the proteolytic activities generating the β-amyloid peptides are also catalysed by γ-secretase. Several clinical trials with original γ-secretase inhibitors have been undertaken in oncology. Another approach

concerns the development of monoclonal antibodies targeting the NRR domain of the NOTCH receptors and the DLL ligand more specifically involved in angiogenesis. Demcizumab (anti-DLL4), OMP-52M51 (anti-NOTCH1) and OMP-59R5 (anti-NOTCH2) have entered clinical trials.

Further Reading

Al-Hussaini H, Subramanyam D, Reedijk M, Sridhar SS. Notch signaling pathway as a therapeutic target in breast cancer. Mol Cancer Ther. 2011;10:9–15.

Aster JC, Blacklow SC. Targeting the Notch pathway: twists and turns on the road to rational therapeutics. J Clin Oncol. 2012;30:2418–20.

Brou C. Intracellular trafficking of Notch receptors and ligands. Exp Cell Res. 2009;315: 1549–55.

Fortini ME. Notch signaling: the core pathway and its posttranslational regulation. Dev Cell. 2009;16:633–47.

Guruharsha KG, Kankel MW, Artavanis-Tsakonas S. The Notch signalling system: recent insights into the complexity of a conserved pathway. Nat Rev Genet. 2012;13:654–66.

Koch U, Radtke F. Notch and cancer: a double-edged sword. Cell Mol Life Sci. 2007;64: 2746–62.

Kopan R, Ilagan MX. The canonical Notch signaling pathway: unfolding the activation mechanism. Cell. 2009;137:216–33.

Nefedova Y, Gabrilovich D. Mechanisms and clinical prospects of Notch inhibitors in the therapy of hematological malignancies. Drug Resist Updat. 2008;11:210–8.

Qiao L, Wong BC. Role of Notch signaling in colorectal cancer. Carcinogenesis. 2009;30:1979–86.

Ranganathan P, Weaver KL, Capobianco AJ. Notch signalling in solid tumours: a little bit of everything but not all the time. Nat Rev Cancer. 2011;11:338–51.

Rizzo P, Osipo C, Foreman K, Golde T, Osborne B, Miele L. Rational targeting of Notch signaling in cancer. Oncogene. 2008;27:5124–31.

Shao H, Huang Q, Liu ZJ. Targeting Notch signaling for cancer therapeutic intervention. Adv Pharmacol. 2012;65:191–234.

Tien AC, Rajan A, Bellen HJ. A Notch updated. J Cell Biol. 2009;184:621–9.

Wang Z, Li Y, Banerjee S, Sarkar FH. Emerging role of Notch in stem cells and cancer. Cancer Lett. 2009;279:8–12.

Yan M, Plowman GD. Delta-like 4/Notch signaling and its therapeutic implications. Clin Cancer Res. 2007;13:7243–6.

Hedgehog Pathway

9

Abstract

The Hedgehog pathway is an important signalling pathway for the development of vertebrates and invertebrates. It is involved in embryogenesis and morphogenesis and plays a major role in the renewal and differentiation of stem cells. Ligands and receptors are normally not expressed beyond embryonic life, except in stem cells. Loss of function of the actors of this pathway determines hereditary malformations, and activating mutations or gene overexpression is found in various malignancies.

Protein signals are issued in the vicinity of target cells equipped with adequate receptors. The ligands, first identified in Drosophila and called *hedgehog* (HHG) (by reference to the Drosophila larvae when affected by invalidating mutations of this gene), have been afterwards found in mammals. The membrane receptor of these ligands is called *patched* (PTCH). Receptor activation by the ligand induces the removal of a constitutive inhibition exerted by this receptor on a membrane protein called *smoothened* (SMO). This is followed by a cascade of events leading to the activation of transcription factors called GLI (for glial cells) and consequently the transcription of target genes required for proliferation.

9.1 Hedgehog Ligands

Three Hedgehog proteins are found in mammals; these are *sonic hedgehog* (gene *SHH*), *Indian hedgehog* (gene *IHH*) and *desert hedgehog* (gene *DHH*). They are first synthesised as a 45 kDa precursor, which undergoes an autocatalytic cleavage releasing a 19 kDa *N*-terminal peptide involved in signalling and a 45 kDa *C*-terminal peptide, which is involved in cleavage and bears a catalytic activity of cholesterol transferase.

HHG proteins are subjected to post-translational modifications that are required for their activity; on the *C*-terminal side, they are covalently bound to a cholesterol

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_9

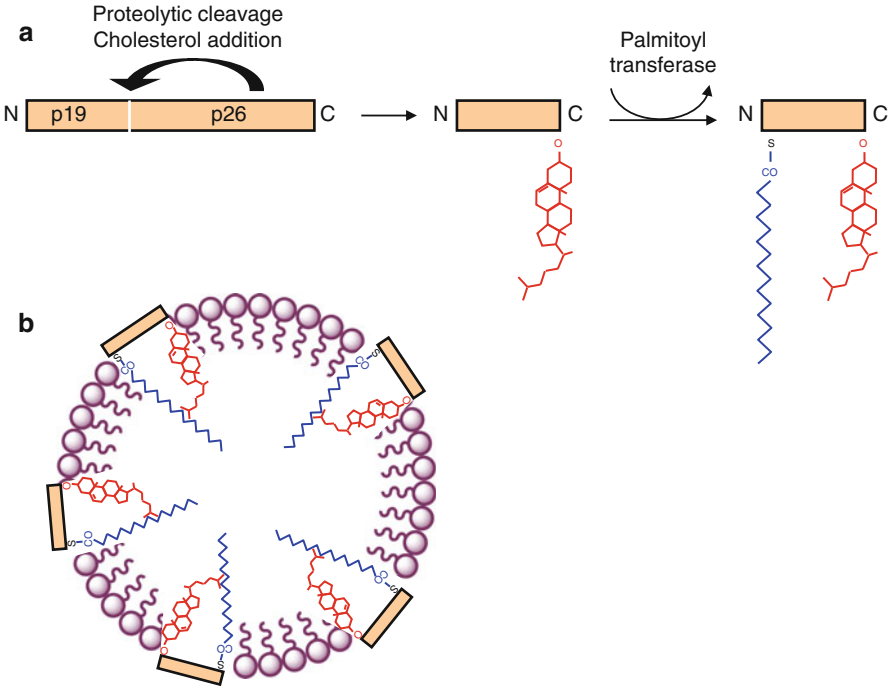


Fig. 9.1 HHG ligands and their post-translational modifications. (**a**) Native HHG proteins (45 kDa) are autocatalytically cleaved to generate a signalling fragment, which is covalently bound to a cholesterol molecule on its *C*-terminal serine. Afterwards, a palmitoyltransferase HHAT adds a palmitic acid to a cysteine residue located at the *N*-terminal extremity. (**b**) The molecule, with two lipid substituents, is inserted in phospholipid micelles, as part of lipoprotein structures

molecule; on the *N*-terminal side, a specific palmitoyltransferase, HHAT (*hedgehog acyltransferase*) or SKI (*skinny hedgehog*), binds them to a palmitic acid moiety at the level of a cysteine residue. When equipped with these two hydrophobic ligands, HHG proteins can be inserted in the peripheral phospholipid layer of a lipoprotein and thus form a multimeric complex (Fig. 9.1).

The mechanism of secretion of the HHG ligands is not precisely known. It involves a 12-transmembrane-domain protein called *dispatched* (DISP), which is indispensable and behaves as a membrane transporter of HHG proteins. This protein contains a sterol-binding domain, which should intervene in the processes of recognition and secretion of the HHG proteins.

9.2 Patched Receptors and Their Activation

The HHG ligands with their two lipid substituents are recognised on a target cell by a receptor called *patched* (PTCH). There are two distinct PTCH proteins, encoded by the genes *PTCH1* and *PTCH2*, which seem to have different functions, the first

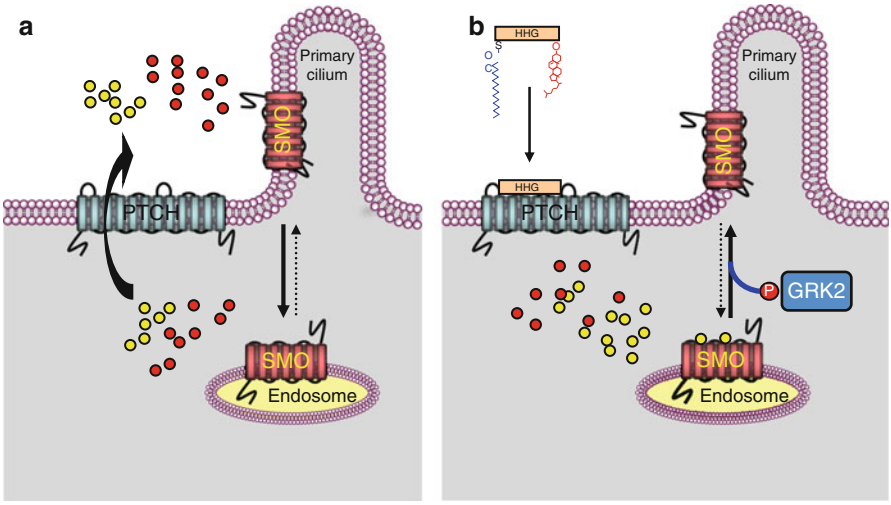


Fig. 9.2 PTCH receptor information transduction to SMO proteins. This figure presents a possible model for SMO activation by the PTCH receptor following HHG ligand binding. The SMO membrane protein is inactive in its endosomal localisation and is activated after migration to the primary cilium. In a first hypothesis, a substrate of PTCH transport activity (oxysterol?) would favour SMO migration to primary cilia and stabilise the active form of SMO. In the absence of HHG ligand (schema **a**), PTCH would pump oxysterol outside the cell, preventing thus its positive action on SMO ciliary localisation. In the presence of the HHG ligand (schema **b**), PTCH pumping activity would be inhibited, and the oxysterol would be available for maintaining SMO under its active form. In a second hypothesis, the substrate of the transport activity of PTCH (vitamin D3?) would be an inhibitor of SMO and would favour its inactive, endosomal localisation. In the absence of HHG ligand (schema **a**), PTCH would pump this sterolic molecule outside the cell where they would inhibit SMO. In the presence of HHG ligands (schema **b**), PTCH pumping activity would be inhibited, and SMO would no longer be destabilised by the inhibitory sterol. The two hypotheses are presented on the same schemas, the putative SMO activator (oxysterol?) as yellow circles and the putative SMO inhibitor (vitamin D3?) by red circles

one being the best known and the only one to be considered here. PTCH belongs to the same transporter family as DISP and also presents 12 transmembrane domains and a cholesterol-binding domain. It also presents two large extracellular domains, which ensure HHG binding. PTCH is only an intermediary in HHG signal reception, since it transfers the information thus received to another membrane protein called *smoothened* (gene *SMO*), by releasing the constitutive inhibition exerted in this protein. The mechanism of this release is not intimately known, but it seems that no direct interaction exists between PTCH and SMO. PTCH would only be a transport protein and would control the localisation of a small inhibitory or excitatory molecule that remains to be identified. In the absence of signal, PTCH would pump an SMO inhibitory molecule from the inside to the outside of the cell; when bound to HHG, PTCH would cease pumping and SMO would then be functional; the reverse would be true if the molecule could conversely activate SMO (Fig. 9.2). Some authors hypothesise that the small molecule transported by PTCH and interacting with SMO could be vitamin D3 or an oxysterol.

SMO is a membrane protein with seven transmembrane domains, belonging to the fifth family of GPCRs (*G-protein-coupled receptors*) (Chap. 6) as FZD receptors (Chap. 7). When PTCH binds an HHG ligand, SMO accumulates in primary cilia, which are subcellular organelles present in all vertebrate cells and form protrusions at the surface of the plasma membrane, devoid of motility and based upon a microtubule cytoskeleton. SMO migration to primary cilia is required for Hedgehog signalling. This migration is accompanied by SMO phosphorylation by GRK2 (*GPCR kinase 2*, gene *ADRBK1*), which is thought to be required for migration. Proteins other than PTCH are able to bind HHG ligands and modulate the signal brought to the target cell: CDO (*cell adhesion molecule-related/downregulated by oncogenes*) and BOC (*brother of CDO*), which are members of the immunoglobulin superfamily and function as accessory receptors to HHG; it is not clear whether or not they are required for signal transmission. They operate via a glycosaminoglycan (heparan sulphate)-binding site. Another protein, HHIP (*hedgehog-interacting protein*), would be a decoy receptor for HHG ligands and would thus inhibit this signalling pathway.

9.3 HHG Signal Transduction

Whereas the previous steps of Hedgehog pathway signalling are highly similar in *Drosophila* and mammals, the subsequent steps of signal transduction, downstream SMO activation, appear to be different, and the proteins identified in *Drosophila* have no exact orthologs in humans. In the absence of HHG signal, zinc-finger transcription factors called GLI (*glioblastoma-associated oncogene*) (genes *GLI1*, *GLI2* and *GLI3*) are sequentially phosphorylated in the primary cilia by serine/threonine kinases: PKA (*protein kinase A*), CK1 (*casein kinase 1*) and GSK3β (*glycogen synthase kinase 3β*), organised with an adapter protein called SUFU (*suppressor of fused*) in a destruction complex; these phosphorylations lead the GLI transcription factors to the proteasome. In the presence of a HHG signal, the kinases are inhibited, and the GLI transcription factors can migrate to the nucleus and activate their target genes (Fig. 9.3).

PKA inhibition occurring upon SMO activation could be due to the recruitment by SMO of a large heterotrimeric G-protein with an α_i subunit, whose activation inhibits adenylyl cyclase and thus cAMP formation and cAMP-induced PKA activation (see Chap. 6). In *Drosophila*, an assembly protein, *cos2*, is involved in the concerted activation of the kinases; its recruitment by SMO prevents this activation. In mammals, there is no *cos2* ortholog and this function could be assumed by the kinesin KIF7; in fact, a protein called SUFU (*suppressor of fused*) is the main negative regulator of the GLI transcription factors in mammals, whereas it plays a minor role in Hedgehog signal transduction in *Drosophila*. In addition, a truncated variant of GLI3 could play a negative regulation of Hedgehog signalling by negative dominance.

The target genes of the GLI transcription factors are involved in morphogenesis, especially Homeobox (HOX) gene family members. *PTCH1* and *HHIP* are also

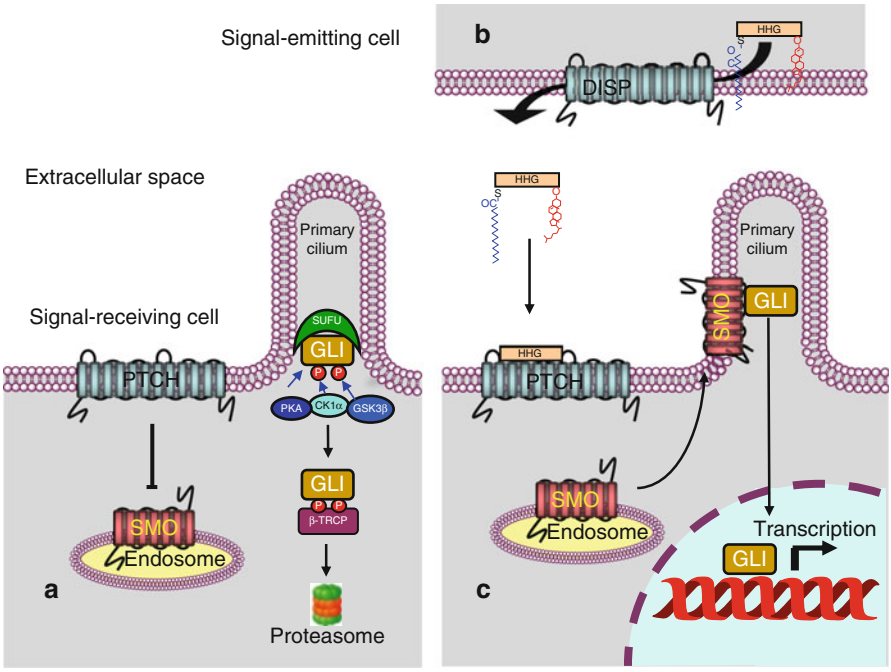


Fig. 9.3 The Hedgehog signalling pathway. (**a**) In the absence of HHG ligands, PTCH inhibits SMO and prevents it to migrate to the primary cilia. The GLI transcription factors are phosphorylated in the primary cilia by several kinases (CK1, PKA, GSK3β), a process that drives them to the proteasome. (**b**) The HHG ligands are matured after translation and leave the producing cell via the membrane protein DISP. (**c**) In the presence of HHG ligands on PTCH, SMO can migrate to the primary cilia, preventing thus the phosphorylation of the GLI transcription factors. These factors can then migrate to the nucleus and activate transcription programmes. One of the major negative regulators of this pathway is the SUFU protein. A positive regulator (not represented here) could be a large heterotrimeric G-protein with α_i subunit, which would inhibit cAMP production and consequently PKA activity

target genes, which explains positive and negative retroaction control of this pathway. Genes encoding proteins involved in cell proliferation, such as BCL2 (*B-cell lymphoma 2*, Chap. 18), cyclin D1 (Chap. 17) or PDGFRα (*platelet-derived growth factor receptor α*, Chap. 1), are also target genes of the Hedgehog signalling pathway.

9.4 Oncogenic Alterations

The Hedgehog pathway was first shown to be involved in a syndrome of predisposition to skin basocellular carcinomas (BCCs), called Gorlin syndrome. A heterozygous mutation in the *PTCH1* gene has been identified in normal cells of these patients, and an additional loss of the non-mutated allele has been detected in

tumour cells. These invalidating mutations induce the loss of PTCH-mediated SMO inhibition, which is in turn responsible for the increased transcription of the GLI target genes. Similarly, sporadic BCCs also present *PTCH1* inactivating mutations or *SMO* activating mutations. *PTCH1* and *SMO* appear, therefore, as a tumour suppressor gene and a proto-oncogene, respectively. Other malignancies, such as medulloblastomas and rhabdomyosarcomas, also present such mutations, as well as less frequent *SUFU* inactivating mutations.

In other cancers, alterations can be found at the level of the HHG ligands. HHG overexpression, with subsequent increase in GLI activation, is observed in pancreatic and small-cell lung cancers, but the driver role of these alterations has not been proved; autocrine loops favouring cell proliferation could well be operating in such cases. Stromal cell interactions with the Hedgehog pathway in tumour cells have been evidenced. The Hedgehog pathway appears, therefore, as involved in basic mechanisms of oncogenesis and metastasis, especially because of its role in stem cell proliferation and epithelial-to-mesenchymal transition.

9.5 Pharmacological Targets

A natural inhibitor of the Hedgehog pathway, cyclopamine, was discovered early; this sterol-related alkaloid induces alterations in lamb foetal development when pregnant ewes were fed with contaminated grass. This SMO inhibitor likely interferes with an oxysterol-binding site. Cyclopamine itself does not seem usable in therapeutics because of its very low bioavailability, but other compounds, such as vismodegib, have been identified and recently patented, with marked antitumour effect in metastatic invasive BCCs and medulloblastomas. *PTCH1* mutations, such as W844C, inducing a constitutive activation of the signalling pathway, have been identified in vismodegib responders; conversely, secondary *SMO* mutations in treated patients, such as D473H, induce resistance to vismodegib. Vitamin D3 could behave as an SMO inhibitor, which would be in agreement with its known role on skin trophicity.

Another potential target is represented by the GLI transcription factors, downstream the Hedgehog pathway. Several small molecules have been identified and may likely enter clinical trials. Finally, upstream the PTCH–SMO interaction, monoclonal antibodies targeting the HHG proteins have shown activity in preclinical models.

Further Reading

Amakye D, Jagani Z, Dorsch M. Unraveling the therapeutic potential of the Hedgehog pathway in cancer. Nat Med. 2013;19:1410–22.

Briscoe J, Thérond PP. The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol. 2013;14:418–31.

Eaton S. Multiple roles for lipids in the Hedgehog signalling pathway. Nat Rev Mol Cell Biol. 2008;9:437–45.

Epstein EH. Basal cell carcinomas: attack of the hedgehog. Nat Rev Cancer. 2008;8:743–54.
Jiang J, Hui C. Hedgehog signaling in development and cancer. Dev Cell. 2008;15:801–12.
Kasper M, Jaks V, Fiaschi M, Toftgård R. Hedgehog signalling in breast cancer. Carcinogenesis. 2009;30:903–11.
Low JA, de Sauvage FJ. Clinical experience with Hedgehog pathway inhibitors. J Clin Oncol. 2010;28:5321–6.
Lum L, Beachy PA. The Hedgehog response network: sensors, switches, and routers. Science. 2004;304:1755–9.
Ma G, Xiao Y, He L. Recent progress in the study of Hedgehog signalling. J Genet Genomics. 2008;35:129–37.
McMillan R, Matsui W. Molecular pathways: the hedgehog signaling pathway in cancer. Clin Cancer Res. 2012;18:4883–8.
Rohatgi R, Scott MP. Patching the gaps in Hedgehog signalling. Nat Cell Biol. 2007;9:1005–9.
Rubin LL, de Sauvage FJ. Targeting the Hedgehog pathway in cancer. Nat Rev Drug Discov. 2006;5:1026–33.
Scales SJ, de Sauvage FJ. Mechanisms of Hedgehog pathway activation in cancer and implications for therapy. Trends Pharmacol Sci. 2009;30:303–12.
Tang JY, So PL, Epstein EH. Novel Hedgehog pathway targets against basal cell carcinoma. Toxicol Appl Pharmacol. 2007;224:257–64.
van den Brink GR. Hedgehog signaling in development and homeostasis of the gastrointestinal tract. Physiol Rev. 2007;87:1343–75.
Varjosalo M, Taipale J. Hedgehog signaling. J Cell Sci. 2007;120:3–6.
Watson S, Serrate C, Vignot S. Voie de signalisation Sonic Hedgehog : du développement embryonnaire aux thérapies moléculaires ciblées. Bull Cancer. 2010;97:1477–83.
Yauch RL, Dijkgraaf GJ, Alicke B, et al. Smoothened mutation confers resistance to a Hedgehog pathway inhibitor in medulloblastoma. Science. 2009;326:572–4.

Integrins

10

Abstract

Integrins constitute a family of heterodimeric transmembrane receptors, whose ligands are essentially the proteins of the extracellular matrix (collagens, laminins, fibronectin, etc.). They play a structural role by recruiting inside the cell the cytoskeleton-associated multiprotein complexes responsible for cell–matrix junctions; they also play a functional role, as their activation leads to proliferation and migration signals, via cytoplasmic kinases and small G-proteins. One of the original features of integrins is the bidirectional signalling that characterises their function: from outside to inside the cell, as most signalling pathways, and from inside the cell to outside, aiming at modulating cell attachment to the extracellular matrix as well as the activity of matrix metalloproteinases.

Integrins are, therefore, primarily involved in cell adhesion and motility mechanisms and, via signalling crosstalks, in cell proliferation processes. They play a central role, therefore, in epithelial-to-mesenchymal transition during development. Several integrins are involved in angiogenesis, as they mediate endothelial cells migration and multiplication. A subclass of integrins is essentially involved in immune processes, as they mediate lymphocyte attachment to their target cells. The alterations of integrins are found in diseases affecting the extracellular matrix (rheumatoid arthritis, multiple sclerosis, psoriasis) as well as in the invasiveness and metastatic capacity of endothelial and tumour cells; they represent, therefore, a major potential target in oncology and clinical trials are ongoing.

10.1 Integrins and Integrin Ligands

10.1.1 Structural Organisation of Integrins

Integrins are heterodimeric receptors with a single transmembrane domain per monomer, able to bind the components of the extracellular matrix (ECM) through

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_10

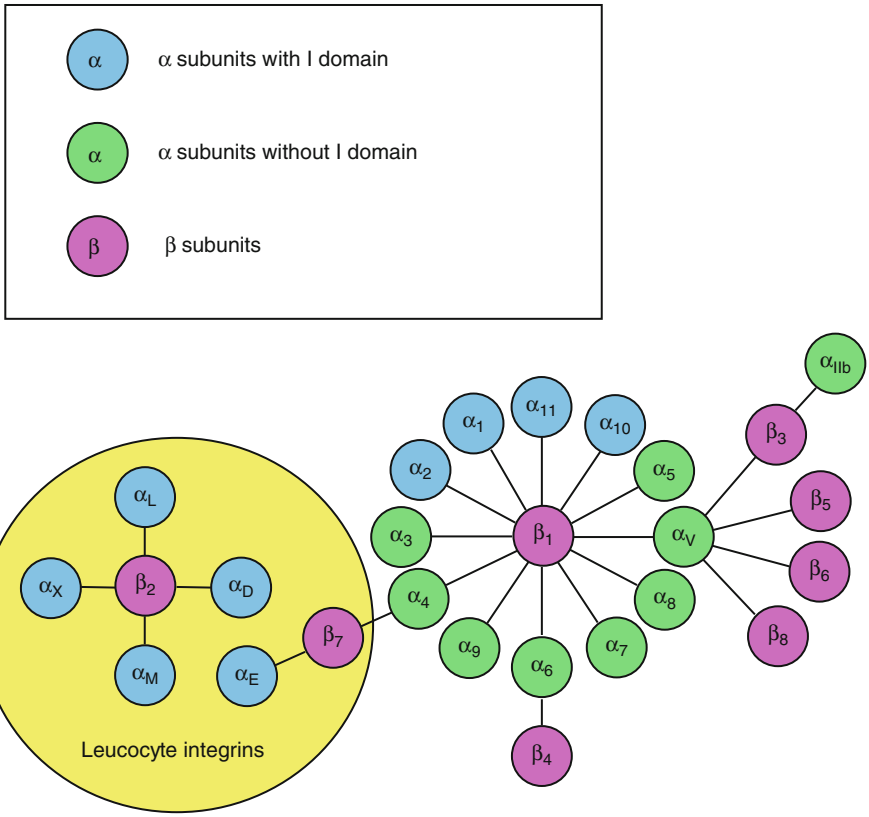


Fig. 10.1 Integrin assembly from their α and β subunits. The 18 α chains and 8 β chains can assemble in diverse ways to form the 24 integrins. Some integrins are specifically expressed in leukocytes and serve as immunoglobulin receptors. Some α chains have an interaction domain with the extracellular matrix (ECM); others do not and signalling operates only with the β-chain interaction (βI) domain

their extracellular part. With 18 distinct α subunits (genes *ITGA*) and 8 β subunits (genes *ITGB*), a total of 24 different integrins can be assembled (Fig. 10.1). The extracellular part of integrins is considerably larger than their intracellular part; it can be represented as an *N*-terminal globular head above two legs.

The globular head of the α chains consists of seven identical domains arranged as the blades of a propeller. Some integrins contain in addition, upstream the propeller domain, a domain analogous to type A von Willebrand factor (αI integrins). This domain enables the fixation of a divalent cation, Ca^{2+} or Mg^{2+}, which interacts with the ECM; it is called the *metal ion-dependent adhesion site* (MIDAS). Downstream the propeller domain, the α chains contain domains sequentially called thigh, genu, calf1 and calf2, then the transmembrane domain and the intracellular domain (Fig. 10.2a).

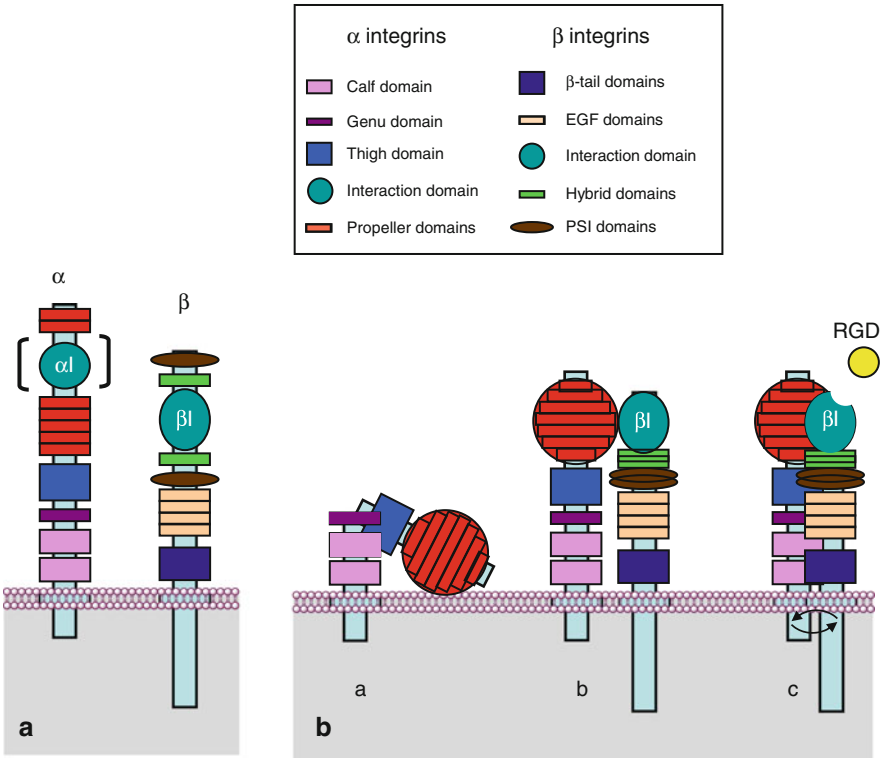


Fig. 10.2 Integrin structure and spatial conformation of α and β chains. (**a**) Schematic representation of the characteristic domains of the α and β chains of integrins along the polypeptidic sequence. The αI domain of interaction with ECM is not present on all α chains. (**b**) Spatial representation of integrins. The propeller domains of the α chains are organised in a globular part, the head; the βI domains are exposed following the folding of the hybrid domains over the PSI domains. (*a*) Integrins adopt a folded conformation in the absence of stimulation; (*b*) under the effect of an intracellular stimulation (inside-out signalling), integrins adopt an extended conformation; (*c*) they can thus unmask a binding site with the extracellular matrix (ECM), which allows the interaction between the cytoplasmic domains and the generation of a signal (outside-in signalling)

The β chains also contain a globular head, a βI domain analogous to the αI domain but constant, bearing a MIDAS that follows a folding of the *N*-terminal extremity; this domain is inserted between two homologous hybrid domains, themselves inserted between two homologous PSI (*plexins, semaphorins and integrins*) domains. Downstream are four EGF domains, a tail domain (TD), the transmembrane domain and the intracytoplasmic part (Fig. 10.2a).

The intracytoplasmic domains of α and β chains are short but are functionally important; the interaction between these two cytoplasmic domains is stabilised by an ionic bond between an arginine residue of the α chain and an aspartic acid residue of the β chain.

In the absence of stimulation, integrins present a bent conformation, with the globular head lying down on the membrane (Fig. 10.2b). Intracellular messages induce their straightening, allowing them to interact with extracellular ligands. Ligand–integrin binding takes place at the level of the αI domain, for those having such a domain, and at the level of the βI domain for the others. These domains exist under two conformations, one closed, with low affinity for the ligand, and the other one open, with a MIDAS bound to Ca^{2+} or Mg^{2+}, enabling the formation of a ionic bond with an aspartic acid residue of the ligand (Fig. 10.2b). Binding of integrins to their ligands induces a conformational change of the cytoplasmic domains of the α and β chains. The α chains are constitutively phosphorylated, whereas the β chains display phosphorylation sites allowing the regulation of their activity. Diverse protein kinase C (PKC) enzymes are involved in the phosphorylation of β chains serine/threonine residues, but it appears that tyrosine residues may also be phosphorylated.

An important feature in inside-out signal transduction is the aggregation of several integrin dimers; such a clustering is required in order to increase the attachment of the cell to the ECM. All along the polypeptidic chains of the matrix proteins are present numerous integrin binding sites, requiring a sufficient number of cell adhesion molecules for efficient gripping; these clusters of integrin dimers attached to the ECM are called *focal adhesions*. Signal transduction arising out of these clusters depends on both *affinity* changes (conformational switch of the intracytoplasmic domain) and *avidity* changes (clustering of integrin dimers).

10.1.2 Integrin Ligands

The main extracellular ligands of integrins, considered as receptors, are the ECM proteins: fibrinogen (FG), fibronectin (FN), von Willebrand factor (VWF), collagens (COL), laminin (LAM) and vitronectin (VTN). Some integrins display high ligand specificity, while others are in contrast of low specificity. Particularly, a tripeptidic sequence present in some ECM proteins is specifically recognised by α_V integrins: the RGD (arginine–glycine–aspartic acid) sequence. This sequence can serve as a guide for the design of competitive inhibitors of integrins; some snake venoms take advantage of this sequence for the disruption of the ECM. Other integrins recognise different tripeptidic sequences, such as IET or LDV, and some integrins recognise more complex tridimensional structures.

A subgroup of integrins (see Fig. 10.1) are selectively present at the surface of leukocytes, where they participate to immune functions. These integrins, such as $\alpha_L\beta_2$ or LFA1 (*lymphocyte function-associated antigen 1*) and $\alpha_M\beta_2$ or MAC1 (from *macrophage*), can recognise molecules of the immunoglobulin superfamily, such as ICAMs (*intercellular adhesion molecules*) and VCAM (*vascular cell adhesion molecule*). The five ICAMs (ICAM1 to ICAM5) are leukocyte transmembrane immunoglobulins whose binding to integrins allows cell–cell adhesion (see Chap. 11). Other extracellular integrin ligands are bacterial polysaccharides and viral proteins.

On the intracellular side, integrins are able to interact with many proteins involved in cytoskeleton structure regulation; these proteins bind the integrin β chain on the one hand and actin on the other hand, linking thus the ECM to the cytoskeleton. The

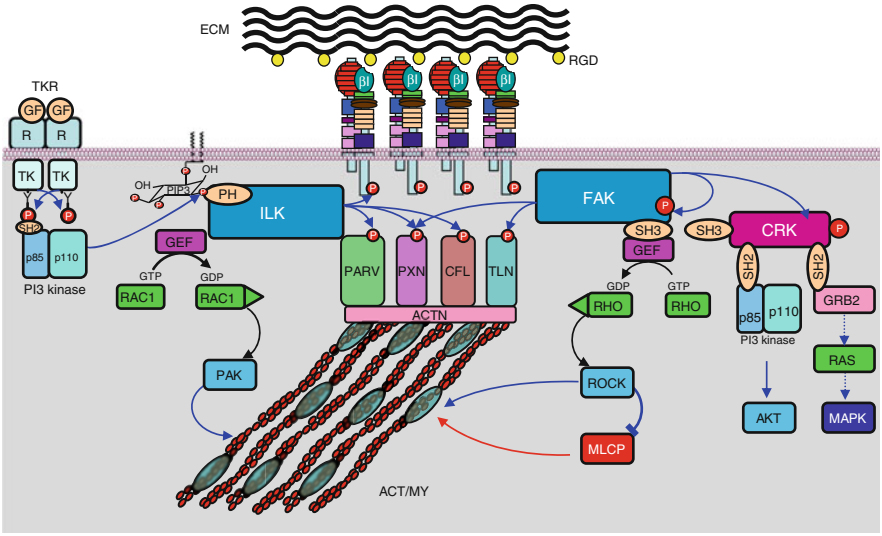


Fig. 10.3 Integrin-activated signalling pathways. A focal adhesion is schematised in the middle of the graph, showing the association between extracellular matrix (ECM) and intracellular actomyosin cytoskeleton (ACT/MY), via the assembly between integrins and several interactants (parvin, paxillin, cofilin, talin, actinin), some of which only are represented here. Two signalling pathways generated by integrin activation are shown: On the left, ILK (*integrin-linked kinase*), recruited, for instance, by PI3 kinase, phosphorylates the integrin β chains as well as several focal adhesion proteins; ILK also activates exchange factors (GEF) of small G-proteins of the RHO family, such as CDC42 and RAC1, which, once activated by GTP binding, can activate in turn some kinases like PAK, which can directly act on cytoskeleton proteins. On the right, FAK (*focal adhesion kinase*), recruited through integrin activation, phosphorylates focal adhesion proteins as well as SRC family kinases (SFK). Adapter proteins are able to activate proliferation pathways such as the AKT and the MAP kinases pathway. FAK also activates, via GEFs, small G-proteins of the RHO family, which, once activated, can activate in turn some kinases like ROCK (*RHO-activated kinase*), which can act on cytoskeleton proteins, either directly or through the phosphatase MLCP (*myosin light chain phosphatase*)

most important of these proteins are talins (TLN), kindlins (FERMT, for *fermitin family homologues*), α-actinins (ACTN), filamins (FLN), cytohesins (CYTH), parvins (PARV) and tensins (TNS). These proteins not only are structural proteins, but they display diverse functions which are involved in transduction of signals received and transmitted by integrins. Adapter proteins are also involved in signal transduction: vinculin (VCL), paxillin (PXN) and various 14-3-3 proteins. The layout of integrin clusters with these proteins constitutes the focal adhesions (Fig. 10.3).

10.2 Signalling Pathways Arising from Integrins

Integrins operate via two signalling ways: *inside-out* signalling, allowing intracellular messages to induce extracellular ligand recognition and binding via integrin activation, and *outside-in* signalling, as all receptors, allowing extracellular

messages to induce specific cell responses. Integrins are therefore bidirectional receptors. These two activities are however inseparable: intracellular signals originating especially from TKR or GPCR activation activate first the receptor function of integrins by inducing their straightening, which governs ligand recognition; afterwards, this binding with extracellular ligands activates the formation of focal adhesions clusters and generates various intracellular messages. Integrin activation leads especially, according to cell type and integrin dimer, to the activation of cytoplasmic kinases, among which the tyrosine kinase FAK (*focal adhesion kinase*) (gene *PTK2*, *protein tyrosine kinase 2*) and the serine/threonine kinase ILK (*integrin-linked kinase*). These kinases are committed to various signalling pathways, leading either to the activation of transcription factors or to the activation of small G-proteins. These pathways are presented on Fig. 10.3.

FAK contains an *N*-terminal FERM (*four-point-one*, *ezrin*, *radixin and moesin*) domain, a central domain bearing the tyrosine kinase activity and a *C*-terminal FAT (*focal adhesion targeting*) domain, allowing FAK binding to paxillin and talin, and consequently its recruitment to focal adhesions. Integrin-mediated FAK activation operates through autophosphorylation, leading to the recruitment of other cytoplasmic tyrosine kinases of the SRC family (SFK, *SRC family kinases*), which complete FAK phosphorylation and are phosphorylated by FAK in return. FAK can also phosphorylate other proteins at the level of focal adhesions, such as α-actinin and paxillin.

FAK bears domains recognised by the SH2- and SH3-binding sites of various proteins:

- Via its SH3 domain-interacting sites, FAK induces the activation of GDP–GTP exchange proteins (GEF) of small G-proteins of the RHO family, such as $p190^{RHOGEF}$ or RGNEF (*RHO-guanine nucleotide exchange factor*), which are involved in cell motility through the control of the actin cytoskeleton.
- Via its phosphotyrosine residues recognised by proteins with SH2 domains, FAK activates adapter proteins such as GRB2, the classical activator of the MAP kinase proliferation pathway (Chap. 2); p85, the regulatory subunit of PI3 kinase (Chap. 3), also governing proliferation and survival pathways; NCK (*non-catalytic region of tyrosine kinase*); CRK (*chicken tumour virus regulator of kinase*) and its associate CAS (*CRK-associated substrate*), which are involved in cell motility; and finally the SHC (*SH2 domain-containing*) proteins, which are involved in multiple cell processes.

ILK bears an *N*-terminal domain with ankyrin (ANK) repeat sites, allowing binding to PINCH (*particularly interesting new Cys–His-rich*) proteins, equipped with LIM (*LIN-11, ISL-1 and MEC-3*) domains and also known as LIMS (*LIM and senescent cell antigen-like domains*) proteins, to a phosphatase called ILKAP (*ILK-associated protein phosphatase*) and to integrin β chains. ILK also bears a central PH (*pleckstrin homology*) domain, allowing its activation by phosphatidylinositol 3,4,5-trisphosphate (see Chap. 3), and a *C*-terminal domain of binding with parvins (PARV), which are themselves actin-binding proteins and constitute a complex with

ILK and PINCH, known as IPP (*ILK, PINCH and PARV*). ILK belongs to focal adhesions and is responsible for the phosphorylation of its partners, the integrin β chain, parvins, cofilins (CFL), paxillin, etc.

Beyond the focal adhesions, ILK main substrates are the serine/threonine kinases AKT (Chap. 3) and GSK3β (Chaps. 7 and 9), both involved in cell survival and proliferation. ILK also activates small G-proteins of the RHO family, such as RHOA, RHOD, CDC42 (*cell division cycle 42*), RND1 and RAC1 (*RAS-related C3 botulinum toxin substrate 1*), involved in cell motility as mentioned above.

Downstream the two integrin-activated kinases, FAK and ILK, one can find other kinases involved in cell proliferation and survival as well as the activation systems of small G-proteins of the RHO family, involved in cell motility. These proteins interact with the actomyosin cytoskeleton and induce the modifications required for cell migration. They promote actin polymerisation and the assembly of polymerised filaments, allowing thus the formation of filipods and lamellipods involved in cell movements. CDC42 and RAC1 activate various proteins such as WASP (*Wiskott–Aldrich syndrome protein*) and PAK (*p21-activated kinase*). RHOA induces the assembly and contraction of actomyosin fibres, especially through the activation of ROCK (*RHO-associated kinase*), which inhibits MLCP (*myosin light chain phosphatase*).

There exist multiple connections between the integrin pathway and the pathways induced by tyrosine kinase receptors (TKR, Chap. 1), especially their common activation of the MAP kinase pathway (Chap. 2) and the PI3 kinase pathway (Chap. 3), via the integrin-dependent kinases, FAK and ILK. NCK appears as an important link between the two signalling pathways; this adapter protein bears SH2 and SH3 domains. The integrin pathway is as important for the control of cell proliferation and survival as for the control of cell adhesion and motility; the cytoplasmic kinases FAK and ILK and the small G-proteins of the RHO family are the preferred mediators of these effects.

Some integrins are able to induce apoptosis when they are not bound to their ligand. They behave as dependence receptors (Chap. 18); this mechanism is called IMD (*integrin-mediated death*) and operates through caspase 8 activation. It especially concerns integrin $\alpha_V\beta_3$, a cell proliferation-associated integrin. This integrin-mediated death mechanism can be inhibited through the recruitment of cytoplasmic tyrosine kinase of the SRC family (SFK), which induce cell proliferation and survival via the activation of appropriate transcription factors.

10.3 Oncogenic Alterations

Various integrins are overexpressed in cancers, such as $\alpha_6\beta_4$, $\alpha_V\beta_3$, $\alpha_V\beta_5$, $\alpha_5\beta_1$ and $\alpha_4\beta_1$, and they appear to play a major role in metastatic tumour development. These integrins cooperate with growth factors of the EGF and PDGF families (Chap. 1) and exert a positive effect on cell proliferation. Other integrins are in contrast negatively regulated in cancers, such as $\alpha_2\beta_1$, which activates the p38 MAP kinase

pathway (Chap. 2), or $\alpha_V\beta_6$ and $\alpha_V\beta_8$, which contribute to TGFβ activation (Chap. 5); these integrins exert, therefore, a downregulation of cell proliferation.

Because of their effects on cell migration, integrins are central to the epithelial-to-mesenchymal transition, which allows tumour cells of epithelial origin to acquire a mesenchymal phenotype capable to migrate and disseminate. These are again $\alpha_V\beta_6$ and $\alpha_V\beta_8$ integrins that can favour tumour invasiveness through TGFβ activation. Germline mutations of some integrin genes are found in some systemic diseases such as *epidermolysis bullosa* (integrin $\alpha_6\beta_4$) and congenital muscular dystrophy (integrin $\alpha_7\beta_1$), but they have not been found in cancers.

Endothelial cells express integrins in response to proangiogenic factors secreted by tumour cells; the binding of these integrins to ECM proteins (collagens, fibronectin, vitronectin, etc.) controls proliferation and migration of these endothelial cells, contributing thus to vessel formation and maturation. The α_V integrin chains can associate with various β chains; integrin $\alpha_V\beta_3$, equipped with a RGD-binding domain present on fibronectin, vitronectin and fibrinogen, is more abundant in tumour vessels than in normal tissue vessels. It is overexpressed during wound healing and inflammation, suggesting that inflammation could well be at the origin of the angiogenic switch of tumours. Integrin $\alpha_V\beta_3$ plays thus a major role in survival and migration of endothelial cells; its antagonists induce endothelial cells apoptosis in vitro as in vivo.

Several other integrins have a major effect on angiogenesis. Integrin $\alpha_V\beta_5$, closely related to $\alpha_V\beta_3$ and which binds vitronectin, is induced by VEGF. Integrin $\alpha_5\beta_1$, the ligand of which is fibronectin, is mainly induced by bFGF in endothelial cells. Integrin $\alpha_4\beta_1$, stimulated by VEGF and bFGF (FGF1), recognises VCAM (*vascular cell adhesion molecule*) as a ligand, which enables the attachment of endothelial cells and pericytes to the smooth muscle cells that express this cell adhesion molecule on their surface. Integrins $\alpha_1\beta_1$ and $\alpha_2\beta_1$ have opposite roles in angiogenesis; they recognise mainly the collagens and their expression in endothelial cells is controlled by VEGFA and VEGFC, which suggests a role in both blood and lymphatic vessels.

10.4 Pharmacological Targets

Because of the major role played by integrins in cell adhesion and migration, they constitute a potential therapeutic target in oncology. Three approaches have been developed for integrin targeting: peptides and peptidomimetics that can imitate their ligand, monoclonal antibodies and small molecules interfering with their activity.

The RGD tripeptide is the docking place on the ECM for several integrins, especially $\alpha_V\beta_3$; peptides that mimic this sequence are susceptible to block cell signalling events induced by integrin activation when they are attached to the matrix. Natural peptides originating in particular from snake venoms and are called disintegrins have been tested as anticancer agents, but synthetic peptides are preferred, such as cilengitide, which entered clinical trials but was shown to be devoid of activity.

Monoclonal antibodies directed against specific integrins have been developed outside the field of oncology, and this approach is promising. The assembly of the α and β chains raises the problem of the optimal strategy: should specific chains be targeted or rather the assembly of two chains in a definite integrin? In addition, the extended conformation of integrins seems a better target than its folded conformation. In oncology, several antibodies have entered clinical trials: etaracizumab (targeting $\alpha_V\beta_3$), volociximab ($\alpha_5\beta_1$), intetumumab (α_V subunit) and natalizumab (α_4 subunit). Finally, small non-peptidic molecules able to interact with integrins at the level of their docking sites on ECM have been identified; they can mimic, for instance, the RGD structure of ECM proteins interacting with $\alpha_V\beta_3$ and other integrins (GLPG-0187) or the interaction of the α_4 subunit with VCAM (AJM-300) or the interaction of $\alpha_5\beta_1$ with fibronectin (JSM-6427).

Downstream integrin receptors, the FAK and ILK proteins represent bona fide targets of the integrin pathway. The FAK-encoding gene, *PTK2*, is overexpressed in several cancer types, especially during metastatic spreading, but no activating mutations have been identified. The mechanism of this overexpression remains unknown, but the fact that the *PTK2* gene promoter harbours sites repressed by p53 and activated by NFκB suggests the existence of crosstalks between these transcription factors and FAK expression. FAK targeting is conceivable, as it displays tyrosine kinase activity: ATP-competitive inhibitors have entered clinical trials. The compounds in development display cross-reactivity with a closely related cytoplasmic tyrosine kinase, PYK2 (*proline-rich tyrosine kinase 2*) (gene *PTK2B*), and with IGF1R, a receptor of the insulin-like growth factor subfamily (Chap. 1).

The ILK protein is activated in cancers, especially when a mutational or transcriptional defect of PTEN is present, which usually constitutes a potent brake for all proliferation effects depending on phosphatidylinositol 3,4,5-trisphosphate. This activation is associated to ILK overexpression, and ILK activating mutations have not yet been identified in cancers. ILK inhibition by antisense approaches or by serine/threonine kinase inhibitors is under study, because it induces in vitro important effects on cell proliferation and migration, especially in cells harbouring an invalidating *PTEN* gene mutation. However, in view of the redundant character of the signalling pathways activated by ILK, one can doubt about the therapeutic efficiency of ILK targeting.

Further Reading

Arnaout MA, Goodman SL, Xiong JP. Structure and mechanics of integrin-based cell adhesion. Curr Opin Cell Biol. 2007;19:495–507.

Arnaout MA, Mahalingam B, Xiong JP. Integrin structure, allostery, and bidirectional signaling. Annu Rev Cell Dev Biol. 2005;21:381–410.

Cabodi S, Camacho-Leal M, Di Stefano P, Defilippi P. Integrin signalling adaptors: not only figurants in the cancer story. Nat Rev Cancer. 2010;10:858–70.

Cox D, Brennan M, Moran N. Integrins as therapeutic targets: lessons and opportunities. Nat Rev Drug Discov. 2010;9:804–20.

Desgrosellier JS, Cheresh DA. Integrins in cancer: biological implications and therapeutic opportunities. Nat Rev Cancer. 2010;10:9–22.

Eble JA, Haier J. Integrins in cancer treatment. Curr Cancer Drug Targets. 2006;6:89–105.
Gahmberg CG, Fagerholm SC, Nurmi SM, Chavakis T, Marchesan S, Grönholm M. Regulation of integrin activity and signalling. Biochim Biophys Acta. 2009;1790:431–44.
Goodman SL, Picard M. Integrins as therapeutic targets. Trends Pharmacol Sci. 2012;33:405–12.
Guo W, Giancotti FG. Integrin signalling during tumour progression. Nat Rev Mol Cell Biol. 2004;5:816–26.
Hehlgans S, Haase M, Cordes N. Signalling via integrins: implications for cell survival and anti-cancer strategies. Biochim Biophys Acta. 2007;1775:163–80.
Hood JD, Cheresh DA. Role of integrins in cell invasion and migration. Nat Rev Cancer. 2002; 2:91–100.
Huveneers S, Truong H, Danen HJ. Integrins: signaling, disease, and therapy. Int J Radiat Biol. 2007;83:743–51.
Legate KR, Montañez E, Kudlacek O, Fässler R. ILK, PINCH and parvin: the tIPP of integrin signalling. Nat Rev Mol Cell Biol. 2006;7:20–31.
Luo BH, Carman CV, Springer TA. Structural basis of integrin regulation and signaling. Annu Rev Immunol. 2007;25:619–47.
Mitra SK, Hanson DA, Schlaepfer DD. Focal adhesion kinase: in command and control of cell motility. Nat Rev Mol Cell Biol. 2005;6:56–68.
Mitra SK, Schlaepfer DD. Integrin-regulated FAK-Src signaling in normal and cancer cells. Curr Opin Cell Biol. 2006;18:516–23.
Rathinam R, Alahari SK. Important role of integrins in the cancer biology. Cancer Metastasis Rev. 2010;29:223–37.
Schwock J, Dhani N, Hedley DW. Targeting focal adhesion kinase signaling in tumor growth and metastasis. Expert Opin Ther Targets. 2010;14:77–94.
Vicente-Manzanares M, Choi CK, Horwitz AR. Integrins in cell migration–the actin connection. J Cell Sci. 2009;122:199–206.
Zhao J, Guan JL. Signal transduction by focal adhesion kinase in cancer. Cancer Metastasis Rev. 2009;28:35–49.

Semaphorins and Adhesion Molecules 11

Abstract

Semaphorins constitute a group of signalling proteins originally described in the central nervous system where they are involved in the growth of the axonal cone and in axonal guidance. They have in fact pleiotropic roles during development of multiple organs and can stimulate or inhibit the same function according to the receptors activated in the target tissues. Semaphorin receptors are especially present at the surface of endothelial cells, which implies that they play a role in angiogenesis. Since most semaphorins are membrane integral proteins, the signalling pathways are exclusively implemented in juxtacrine situations. Semaphorins play a major role in cell adhesion and motility, so that the alterations of this signalling pathway contribute to neurodegenerative diseases and cancer.

Only some of the signalling pathways involving semaphorins are presently known from ligand binding to intracellular effectors: we will present them in this chapter. Many things remain to be discovered in the semaphorin field. This chapter is completed by a short presentation of various adhesion molecules, outside the semaphorin field, which are involved in various types of cell–cell and cell–matrix contacts, in addition to those already studied in Chaps. 10 and 11.

11.1 Semaphorins and Semaphorin Receptors

Semaphorins, usually abbreviated as SEMA, constitute in vertebrates a group of 20 proteins distributed in 5 classes, from class 3 to class 7. Those belonging to class 3 are secreted, whereas the other ones are membrane-bound, either through the existence of a single transmembrane domain (semaphorins of classes 4, 5 and 6) or by anchoring with a glycosylphosphatidylinositol membrane lipid (semaphorin 7A). All semaphorins have a characteristic *N*-terminal domain called the SEMA domain, 500 amino acids long, which consists of a seven-blade β-propeller structure and contains dimerisation and receptor-binding sites. They also contain other specific domains (Fig. 11.1). Semaphorins operate as homodimers associating two polypeptidic chains bound by a disulphide bridge.

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_11

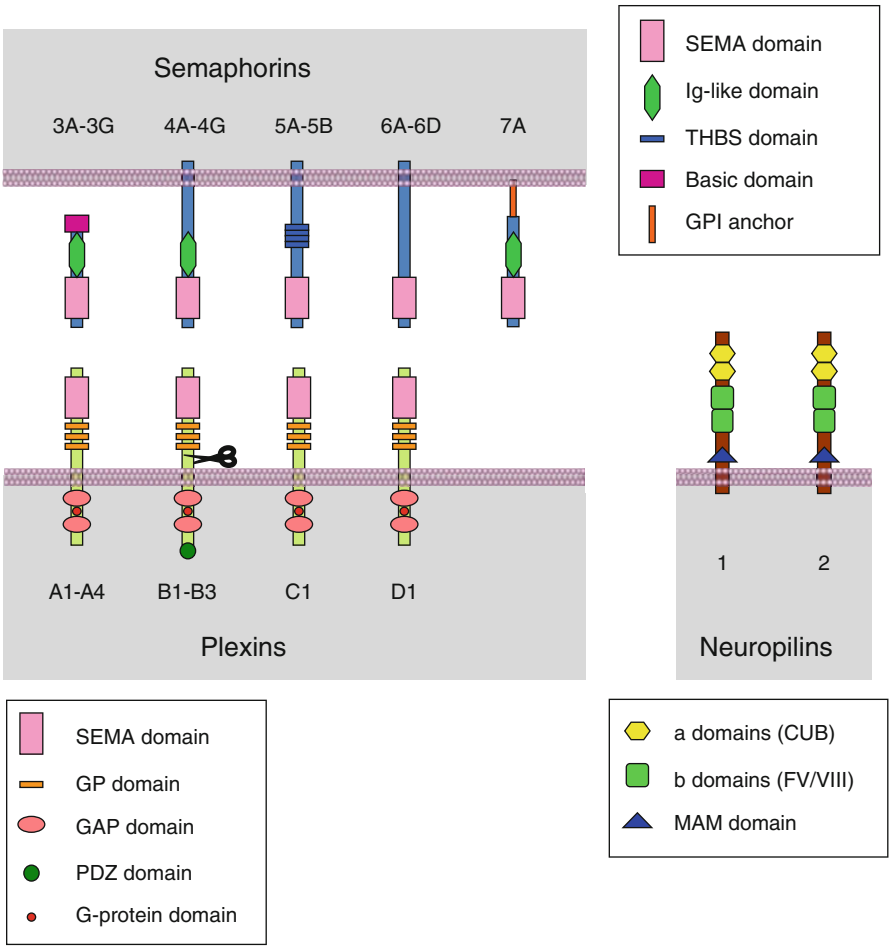


Fig. 11.1 General structure of semaphorins, plexins and neuropilins. Semaphorins (SEMAs) are a family of 21 proteins distributed in 5 subfamilies, from semaphorins 3 to semaphorins 7. They are either secreted (semaphorins 3) or bound to the plasma membrane. Membrane insertion is obtained either by a transmembrane domain (semaphorins 4, 5 and 6) or anchoring via a glycosylphosphatidylinositol (GPI) group (semaphorins 7). All semaphorins contain a SEMA domain allowing receptor binding and dimerisation and, according to the class they belong to, an immunoglobulin-like (IgG) domain, a thrombospondin-like (THBS) domain or a basic domain. Plexins (PLXNs) are a family of 9 proteins distributed in 4 subfamilies, from plexins A to plexins D. These are transmembrane proteins equipped, on the extracellular side, with a SEMA domain and glycine- and proline-rich (GP) domains and, on the intracellular side, with a small G-protein binding domain, a GAP (*GTPase activating protein*) domain and, for plexins B, a PDZ (*postsynaptic density-95*, *disc-large and zonula occludens*) domain and a GEF (*guanyl nucleotide exchange factor*) domain. Neuropilins (NRPs) are a group of 2 proteins, neuropilin 1 and neuropilin 2. These are transmembrane proteins, with a very short intracellular domain and, on the extracellular side, two domains *a* for complement binding (CUB, *complement protein*, *urchin embryonic growth factor and bone morphogenic protein 1*), two domains *b* homologous to the coagulation factor V/VIII and a domain *c* called MAM (*meprin*, *A5 protein*, *μ protein tyrosine phosphatase*), involved in neuropilin dimerisation and interactions with other receptors

Table 11.1 Combinatorial pattern of semaphorins and semaphorin receptors

Receptor	Ligand	Coreceptor
PLXNA1	SEMA6C, SEMA6D	
PLXNA2	SEMA6A, SEMA6B	
PLXNA3		
PLXNA4	SEMA6A, SEMA6B	
PLXNB1	SEMA4D	CD72
PLXNB2	SEMA4A, SEMA4C, SEMA4D	
PLXNB3	SEMA5A, SEMA5B	SYN3
PLXNC1	SEMA7A	Integrin $\alpha_1\beta_1$
PLXND1	SEMA3E	NRP1, NRP2
	SEMA4A	TIM2[a]
NRP1	SEMA3A, SEMA3B, SEMA3C, SEMA3D	PLXNA1-PLXNA4
NRP2	SEMA3B, SEMA3C, SEMA3D, SEMA3F, SEMA3G	

[a]TIM2 is a mouse protein whose human closest equivalent is TIMD1 (*T-cell immunoglobulin and mucin domain-containing protein*)

The semaphorin receptors of membrane-bound semaphorins (classes 4–7) and semaphorin 3E are the plexins (PLXNs): plexins A for semaphorins 6, plexins B for semaphorins 4 (except semaphorin 4A) and semaphorins 5, plexin C1 for semaphorin 7A and plexin D1 for semaphorins 3E and 4A. A given semaphorin can bind several plexin receptors, and a given plexin can bind several semaphorin ligands (Table 11.1). Secreted semaphorins (class 3, but with the exception of semaphorin 3E) use neuropilins (NRPs) as primary receptors, together with plexins A, which are responsible for signal transduction. In lymphocytes, SEMA4A also uses TIM2 (*T-cell immunoglobulin and mucin domain-containing protein*), as a receptor or coreceptor, while SEMA4D can use CD72, a membrane lectin expressed in B-cells. Diverse membrane proteins may be associated to semaphorin reception: tyrosine kinase receptors (TKR, Chap. 1), integrins (Chap. 10), through the interaction of a RGD motif borne by SEMA7A, as well as TREM2 (*triggering receptor expressed on myeloid cells 2*) and DAP12 (*DNAX-activation protein 12*), a TRK-binding protein encoded by the gene *TYROBP* (*TYRO protein tyrosine kinase binding protein*). Proteoglycans with heparan sulphate (HSPG) or chondroitin sulphate (CSPG) moieties are also associated with several semaphorin receptors.

Plexins, which appear as the main semaphorin receptors, constitute a family of 9 proteins distributed in 4 classes (A to D). These proteins display a single transmembrane domain and, on the intracytoplasmic side, a binding domain to small G-proteins of the RHO subfamily such as RHOD, RND1 or RAC1, and they exert on these small G-proteins a GTPase stimulating activity (GAP). Plexins B present in addition, on the *C*-terminal side, a PDZ (*postsynaptic density 95/disc-large/zona occludens*) domain for binding a GDP–GTP exchange factor (GEF) for small G-proteins. The extracellular part of plexins contain a SEMA domain involved in semaphorin binding, and class B plexins have in addition a proteolytic domain similar to that of the subtilisin protein convertase, furin (Fig. 11.1). Due to the dimeric structure of the active ligands, plexins are thought to also act as dimers.

Neuropilins (2 proteins in mammals) are the primary receptors of class 3 semaphorins (excepted SEMA3E), but signalling requires the intervention of a class A plexin. They contain two domains for complement binding (CUB [*complement protein, urchin embryonic growth factor and bone morphogenic protein 1*]), two domains homologous to the coagulation factor V/VIII and a domain called MAM (*meprin*, *A5 protein*, *μ protein tyrosine phosphatase*), which is important for their dimerisation and their interaction with other receptors. Their intracellular domain is very short and does not seem able of signal transduction, which explains the requirement of a plexin molecule as receptor. Neuropilins can also be used as coreceptors for some forms of VEGF (*vascular endothelial growth factor*) (Chap. 1) and they interact with VEGF receptors. They may also interact with other TKR, such as MET, FGFR2 and PDGFRB, and serve as coreceptors to several growth factors (Chap. 1). Furthermore, they interact with adhesion factors of the CAM (*cell adhesion molecules*) family and integrins (Chap. 10).

11.2 Semaphorin-Induced Signal Transmission

The diversity of semaphorins and semaphorin receptors explains why the signalling pathways they activate are only partially known. Only the pathways originating from semaphorins 3A and 4D have been deciphered and will be presented here as representative examples. It appears that the different complexes semaphorin–plexin can each activate several transduction pathways involving various systems: small G-proteins, cytoplasmic tyrosine kinases, integrins, etc. The action of semaphorins on cytoskeleton, cell adhesion and motility is pleiotropic and certainly redundant.

11.2.1 Semaphorin 3A Signalling

The reception complex NRP1–PLXNA1, at the basal inactive state, is associated to a protein called FARP2 (*FERM*, *RHO-GEF and pleckstrin domain protein 2*) with GEF (*guanine nucleotide exchange factor*) activity (Fig. 11.2a). When bound to SEMA3A, the reception complex releases FARP2, unmasking its GEF activity and allowing it to activate RAC1, a small G-protein, through GDP–GTP exchange. RAC1 can then sequentially activate the serine/threonine kinases PAK1 (*p21-activated kinase 1*) and LIMK1 (*LIM domain kinase 1*) and cofilin (CFL), a protein controlling actin polymerisation (Fig. 11.2b). RAC1 facilitates also the association of RND1, another small G-protein of the RHO subfamily, with plexin A1, which stimulates the GAP activity of the plexin. The target of this action of plexin A1 is RRAS, another small G-protein also involved in integrin signalling (Fig. 11.2c). This results in the inactivation of integrin β_1 subunit, inducing the detachment of cells from the extracellular matrix. The interaction of RND1 with plexin A can be

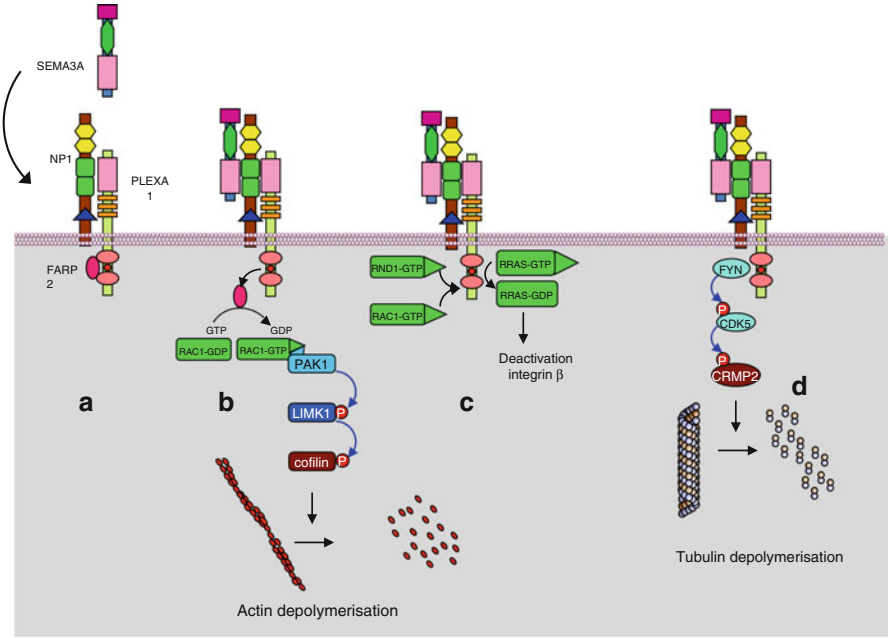


Fig. 11.2 Semaphorin 3A signalling pathway. (**a**) SEMA3A recognises, on the target cell, an inactive reception complex constituted of NPL1 and PLXNA1. This plexin is bound to protein FARP2 (*FERM*, *RHO-GEF and pleckstrin domain protein 2*), which is a guanyl nucleotide exchange factor (GEF). (**b**) The interaction of SEMA3A with the reception complex induces the release of FARP2, which can exchange the GDP bound to the small G-protein RAC1 against GTP; RAC1 is thus activated and induces a cascade of kinases, PAK1 and LIMK1, leading to the activation of cofilin (CFL), which depolymerises the actin cytoskeleton. (**c**) RAC1 and RND, other small G-proteins of the RHO subfamily, activate the GAP (*GTPase activating protein*) function of PLXNA1, which deactivates another small G-protein, RRAS. This induces the deactivation of integrin β_1 subunit and the detachment of the cell from the extracellular matrix. (**d**) Once activated, PLXA1 can in turn activate cytoplasmic tyrosine kinases such as FYN, FES or FER. They recruit in turn and phosphorylate the serine/threonine kinase CDK5, allowing thus the activation of CRMP2 (*collapsin response mediator protein*, gene *DPYSL2*, *dihydropyrimidinase-like 2*), a microtubule-depolymerising agent

inhibited by another small G-protein, RHOD, which can bind the same site of plexin A1.

Upon activation by SEMA3A, the NRP–PLXNA1 reception complex can activate cytoplasmic tyrosine kinases such as FYN, FES or FER. This activation allows the phosphorylation of plexin A1 by CDK5 (*cyclin-dependent kinase 5*), which induces the recruitment of a protein called CRMP2 (*collapsin response mediator protein 2*) (gene *DPYSL2* [*dihydropyrimidinase-like 2*]) to be phosphorylated (Fig. 11.2d). CRMPs act as microtubule-depolymerising agents and are thus mediators of SEMA3A-induced effects on the cytoskeleton.

11.2.2 Semaphorin 4D Signalling

Plexin B1, the semaphorin 4D receptor, is dimeric in the absence of signal. Binding SEMA4D induces the activation of small G-proteins of the RHO subfamily (RHOD, RND1, RAC1), giving similar effects to those elicited by SEMA3A on the NRP1–PLXNA1 reception complex. However, SEMA4D may also have opposite effects: the activation of its receptor may induce, on the one hand, the activation of a RHO-GAP protein (and therefore RHO deactivation); and on the other hand the activation of a RHO-GEF protein (and therefore RHO activation).

11.3 Oncogenic Alterations and Pharmacological Targets

Semaphorins play a major role in cell adhesion and motility and consequently in invasion, metastasis, angiogenesis and immune processes. Semaphorin signalling can induce opposite effects as a function of the ligand–receptor system involved and may give positive or negative instructions concerning cancer cell growth and metastatic dissemination. Semaphorins 3B and 3F are considered as tumour suppressors; the loss of function of SEMA3B, related to promoter hypermethylation or gene polymorphism, is frequently found in non-small-cell lung cancers. This semaphorin has experimentally an inhibitor effect on cell growth. Similar observations have been made on SEMA3F, which displays antiangiogenic properties and inhibits cell attachment to the extracellular matrix; this semaphorin could be, therefore, antimetastatic, as SEMA3A. Semaphorins 3C and 3E have been shown in contrast to favour angiogenesis and consequently tumour progression. SEMA4D, which induces the activation of the MET and ERBB2 tyrosine kinase receptors, via PLXNB1, favours tumour invasiveness; it has also been shown to display proangiogenic properties; the same is true for the couple SEMA5B–PLXNB3.

Neuropilin overexpression has been observed in numerous cancer types and is associated to tumour progression in digestive, mammary and pulmonary cancers, among others. The experimental inhibition of NRP2 is able to slow down the development of colorectal tumours. Plexins A are overexpressed in several cancer types and *PLXNB1* mutations have been detected in prostate cancers. However, some conflicting results have been published, with a decrease in the expression of plexin B1 in some cancer types and an overexpression in others.

Because of their pleiotropic and redundant effects, the semaphorins appear as hardly druggable for cancer therapy. The monoclonal antibody technology offers several relevant tracks for semaphorin, plexin and neuropilin inhibition. Especially, the pro-oncogenic semaphorins, SEMA3C, SEMA3E and SEMA4D, could be candidates for antibody targeting, as well as the plexin they activate, PLXNB1. Conversely, the tumour-suppressing semaphorins of the SEMA3 family could be used for the development of gene replacement or peptide mimicking therapies.

11.4 On Some Other Adhesion Proteins

Cell–cell adhesion is required for the cohesion of epithelial tissues. Several types of cell–cell junctions have been described, with distinct molecular organisations and functions: tight junctions (*zonula occludens*), adherens junctions, desmosomes and gap junctions. In addition, cell–matrix adhesion is required in both epithelial and mesenchymal tissues for the attachment of cells to the stroma. Both types of adhesion should be disrupted to enable cells to migrate: this is required for a multiplicity of physiological events, especially during embryo development, and this is used by cancer cells to evade their original location and implant metastases. Adhesion and motility are connected to signalling processes, first because they are triggered by messages received by the cells and also because they trigger in turn the activation of signalling pathways. Integrins and semaphorins represent archetypes of cell–matrix and cell–cell adhesion molecules, respectively, and of the connections between adhesion and signalling. There exist in addition a number of cell–cell adhesion molecules which are not directly involved in cell signalling but whose alterations are associated to oncogenesis and metastatic processes.

11.4.1 Cell–Cell Junctions

Tight junctions (TJ) are constituted of several types of proteins involved in cell–cell contacts at the apical side of epithelial and endothelial cells. They contain integral transmembrane proteins (notably claudins [CLDNs], occludin, tricellulin, MARVEL domain-containing proteins, junctional adhesion molecules [JAMs] of the membrane immunoglobulin family and some others) and peripheral anchoring proteins located on the cytoplasmic membrane surface and linking tight junctions to the cytoskeleton (*zonula occludens* [ZO] proteins [TJPs], membrane-associated guanylyl kinases [MAGIs] and some others). In addition to their barrier and fence functions, tight junctions are involved in cell proliferation and migration and therefore in oncogenesis and metastasis.

Adherens junctions and desmosomes are constituted with proteins belonging to the same families but with different members: cadherins (CDHs), armadillo repeat-containing proteins and plakins. In adherens junctions, E-cadherin extracellular portions of two adjacent cells are bound together thanks to Ca^{2+} affinity sites and to the actin cytoskeleton through β-catenin and α-catenin (Fig. 7.1). Desmosomal cadherins are called desmogleins (DSGs) and desmocollins (DSCs). They are bound to the armadillo repeat-containing proteins called junction plakoglobin (JUP), desmoplakin (DSP) and plakophilins (PKPs), which bridge the desmosomal cadherins to the cytoskeleton, as catenins do in adherens junctions. Overexpression of some of these proteins and reduced expression of other are found in various epithelial cancers.

Gap junctions (GJ) are transmembrane channels connected between adjacent cells, which enable the passage of small molecules (less than 1,000 Da) from one cell to another, including metabolites (glucose, amino acids, nucleotides) and signalling molecules such as the second messengers IP3, Ca^{2+}, cAMP and cGMP. They

are constituted of two hexameric transmembrane protein complexes called *connexons* or hemichannels. About 20 different proteins with some tissue specificity may enter in the constitution of connexons and are called *connexins* (CXs, genes *GJx*). Connexins are transmembrane proteins with four transmembrane helices. Most connexons are heterohexameric, but homohexameric associations are also found; in addition, two opposite hemichannels may be constituted of identical or different connexins. This variety in connexin associations and connexon structure likely explains the variety of gap junction functions in various tissues. As a general feature, there is loss of gap junctions and downregulation of CXs in cancers, which suggests a tumour suppressor role for these proteins.

11.4.2 Claudins

Claudins (CLDNs) constitute a family of 24 tight-junction transmembrane proteins of 200–300 amino acids containing four transmembrane helices with *N*- and *C*-terminal ends in the cytoplasm. The *C*-terminal end contains a PDZ (*postsynaptic density 95/discs-large/zonula occludens*) motif that interacts with corresponding motifs in the peripheral cytoplasmic ZO proteins. Claudins are also phosphorylated in the *C*-terminal region by various kinases with different consequences. They are expressed in a tissue-specific way, although the same tissue or cell type generally expresses several different claudins. They may be associated to various processes according to the tissue and to the claudin molecule; for instance, some of them (CLDN2, CLDN15) behave as cation channels and others (CLDN4, CLDN7) as anion channels. Also, some of them (CLDN1, 4, 7) increase transepithelial resistance and others (CLDN2 and 10) decrease this parameter. The complete understanding of individual claudin functions in epithelial tissues remains to be established.

Several alterations of claudin expression have been found in cancers: upregulation of *CLDN3* and *CLDN4* is observed in various epithelial cancers, while *CLDN1* and *CLDN7* are downregulated in invasive cancers, possibly in relation to the acquisition of metastatic properties, since the epithelial-to-mesenchymal transition genes *SNAI1* (SNAIL) and *SNAI2* (SLUG) repress claudin expression.

11.4.3 Cadherins

Cadherins (CDHs) constitute a large family of adhesion molecules which play a major role at the level of cell–cell junctions, in connection with Ca^{2+} ions. E-cadherin (CDH1) is characteristic of epithelial cells, P-cadherin (CDH3) of placenta, N-cadherin (CDH2) of mesenchymal tissues, etc. Cadherins allow homophilic interactions at the level of adherens junctions and desmosomes. Cadherins are transmembrane proteins with a voluminous extracellular segment and a small intracellular segment (Fig. 11.3). They are characterised at the extracellular level by a repeat of five Ca^{2+} interaction domains; the association between cadherin molecules takes place at their *N*-terminal extremity, where HAV (His–Ala–Val) domains and

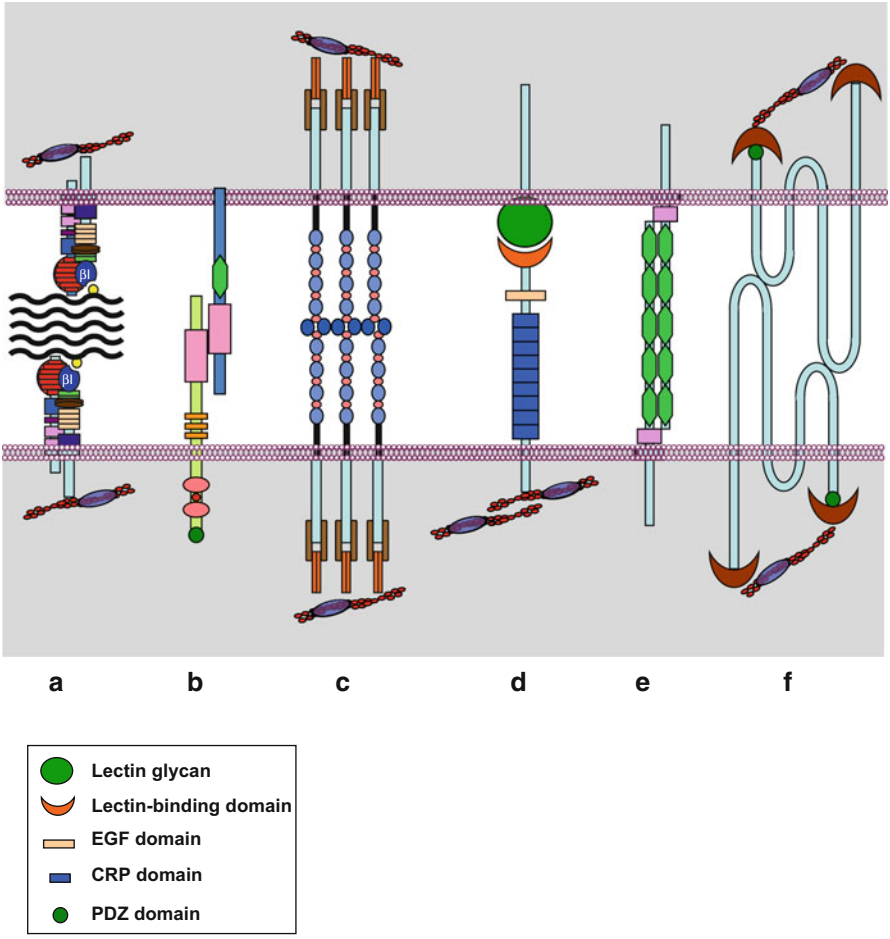


Fig. 11.3 Cell adhesion molecules. General organisation of five main types of adhesion molecules: (**a**) integrins; (**b**) semaphorins; (**c**) E-cadherin; (**d**) selectins; (**e**) IgCAM; (**f**) tetraspanins

tryptophan residues interact with hydrophobic domains. Their intracellular segment is connected to the actin cytoskeleton via molecules of α-, β-, γ- and δ-catenin, constituting together a cytoplasmic cell adhesion complex (CCC).

E-cadherin behaves as a tumour suppressor or rather a metastasis suppressor. Its expression is repressed during the epithelial-to-mesenchymal transition, which is the initiating event of tumour invasion; the transcriptional regulators SNAIL (gene *SNAI1*) and SLUG (gene *SNAI2*) play a major role in this repression. E-cadherin is then replaced by mesenchymal cadherins such as N-cadherin during a *cadherin switch*. The role of E-cadherin in the cytoplasmic availability of β-catenin has been presented in Chap. 7. E-cadherin intervenes also during δ-catenin-induced activation of RHO family small G-proteins, RAC1 and CDC42, through the activation of various GEFs such as VAV2 and TIAM1. The activation of these small G-proteins

releases α-catenin, which induces the disruption of CCCs and the rearrangement of cytoskeleton proteins: CDC42 would be involved in filipod formation, RAC1 in that of lamellipods and RHOA in that of stress fibres. Other pathways of GEF activation or GAP deactivation of RHO proteins have been mentioned in Chaps. 6 (chemokines) and 10 (integrins).

Direct targeting of N-cadherin can be obtained through various approaches: monoclonal antibodies recognising the ectodomain; cyclic and linear peptides reproducing the structure of the homophilic interaction domain, such as exherin; and small molecules mimicking this structure.

11.4.4 Selectins

Selectins are transmembrane proteins with an *N*-terminal extracellular segment, containing a lectin-type domain, an EGF-like domain and a CRP (*complement regulatory protein*) domain; a single transmembrane segment; and a *C*-terminal intracytoplasmic domain (Fig. 11.3). Three types of selectins have been identified in tissues: endothelium (selectin E, gene *SELE*), platelet (selectin P, *SELP*) and leukocytes (selectin L, *SELL*). Via their lectin domain, they are able to bind several leukocyte membrane glycoproteins, especially blood group-specific proteins bearing a terminal fucose residue. P-selectin ligand (gene *SELPLG*) is the best known ligand, present on leukocyte cell membranes. All three selectins bind to SELPLG, and binding is facilitated by sulphation of the *N*-terminal part of the protein. Another common selectin ligand is the leukocyte membrane glycoprotein CD44, which interacts with various components of the extracellular matrix, including glycosaminoglycans such as hyaluronic acid.

Selectin ligands are often abnormally expressed in cancer cells (annex C) and their interaction with selectins may be involved in tumour progression. Several interaction levels should be considered: tumour cell interactions with platelets and leukocytes to constitute vascular emboli; and tumour cell interactions with endothelial cells to modulate tumour cells extravasation (and consequently metastatic migration) and to stimulate endothelial cell proliferation (and consequently neoangiogenesis). Targeting selectins for cancer treatment has been proposed, especially with inhibitors of mucin *O*-glycosyltransferases and fucosyltransferases, with peptides mimicking the interaction site with selectins and with anti-selectin monoclonal antibodies.

11.4.5 Cell Adhesion Molecules of the Immunoglobulin Family

A large family of membrane immunoglobulins are cell adhesion molecules; we already mentioned ICAM and VCAM as integrin ligands; there are also two neuronal NCAMs, a melanoma-derived MCAM, a platelet/endothelium PECAM and some others. Some IgCAM have a proper transmembrane domain; others are attached to the membrane thanks to a glycosylphosphatidylinositol anchor. They are

involved in immune response, cerebral development, epithelial and vascular tissue morphogenesis. They are characterised by the presence, at the extracellular level, of immunoglobulin-like domains and fibronectin III domains (Fig. 11.3). They ensure cell adhesion through *trans* homophilic interactions and by heterophilic interactions with other molecules such as integrins, N-cadherin and tyrosine kinase receptors.

NCAMs are especially involved in intracellular signalling, through the activation of cytoplasmic tyrosine kinases like FYN, which activate FAK. NCAMs behave as tumour suppressors and lower expression has been associated with oncogenesis and metastatic dissemination; in combination with N-cadherin and FGFR4, they induce a stimulation of integrin attachment to the extracellular matrix (inside-out signalling). Other CAMs are also involved in cancer progression, exerting sometimes a pro-oncogenic role, such as MCAM, and sometimes an anti-oncogenic role. A special class of IgCAM, called nectins (genes *CADM1* to *CADM4*), play an important role in adherens junctions and cell motility, in conjunction with integrin signalling.

Several CAM-targeting antibodies have been developed up to clinical trials: for instance, edrecolomab and catumaxomab against EPCAM, this last compound being bispecific and targeting also CD3, a component of the T-cell receptor (Chap. 13); lorvotuzumab against NCAM; and labetuzumab against CEACAM, these two last antibodies being coupled with toxins or radionuclides, so that it is difficult to know whether they actually inhibit their target or simply direct the toxin to this target.

11.4.6 Tetraspanins

Tetraspanins constitute a family of 33 transmembrane proteins with four membrane-spanning domains (Fig. 11.3). They have short *N*- and *C*-terminal intracellular segments, a very short intracellular loop and two longer extracellular loops, especially the second one (70–130 amino acids), with recognition sites for various interactants and four intramolecular disulphide bridges. They are linked to juxtamembrane palmitic acid residues through covalent binding with cysteine residues. They are organised as clusters gathering, in addition to tetraspanins, several types of membrane proteins: other adhesion molecules such as VCAM, ICAM or EPCAM, integrins and tyrosine kinase receptors. These clusters are called tetraspanin-enriched microdomains (TEM).

Some tetraspanins are expressed in lymphocytes and play a role in the immune response at the level of lymphocyte adhesion and migration; most tetraspanins are expressed at the surface of endothelial and epithelial cells and are also involved in cell adhesion and motility. Tetraspanins 8 and 24 (TSPAN8 and CD151, respectively) have clear pro-metastatic and proangiogenic effects through their interactions with integrins $\alpha_3\beta_1$ and $\alpha_6\beta_4$, inducing FAK activation. Tetraspanins 27, 29 and 30 (CD9, CD82 and CD63, respectively) rather appear as metastasis suppressors, both by their inhibitor effects on maturation of endothelial cells and on intravasation of tumour cells.

Antibodies against tetraspanins, especially CD151, have been developed, but have not yet demonstrated an anticancer activity.

Further Reading

Bagri A, Tessier-Lavigne M, Watts RJ. Neuropilins in tumor biology. Clin Cancer Res. 2009;15:1860–4.
Bielenberg DR, Pettaway CA, Takashima S, Klagsbrun M. Neuropilins in neoplasms: expression, regulation, and function. Exp Cell Res. 2006;312:584–93.
Blaschuk OW, Devemy E. Cadherins as novel targets for anti-cancer therapy. Eur J Pharmacol. 2009;625:195–8.
Cavallaro U, Christofori G. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. Nat Rev Cancer. 2004;4:118–32.
Cavallaro U, Dejana E. Adhesion molecule signalling: not always a sticky business. Nat Rev Mol Cell Biol. 2011;12:189–97.
Ch'ng ES, Kumanogoh A. Roles of Sema4D and Plexin-B1 in tumor progression. Mol Cancer. 2010;9:251.
Coupland LA, Parish CR. Platelets, selectins, and the control of tumor metastasis. Semin Oncol. 2014;41:422–34.
Dusek RL, Attardi LD. Desmosomes: new perpetrators in tumour suppression. Nat Rev Cancer. 2011;11:317–23.
Gaur P, Bielenberg DR, Samuel S, et al. Role of class 3 semaphorins and their receptors in tumor growth and angiogenesis. Clin Cancer Res. 2009;15:6763–70.
Goshima Y, Sasaki Y, Yamashita N, Nakamura F. Class 3 semaphorins as a therapeutic target. Expert Opin Ther Targets. 2012;16:933–44.
Gu C, Giraudo E. The role of semaphorins and their receptors in vascular development and cancer. Exp Cell Res. 2013;319:1306–16.
Guttmann-Raviv N, Kessler O, Shraga-Heled N, Lange T, Herzog Y, Neufeld G. The neuropilins and their role in tumorigenesis and tumor progression. Cancer Lett. 2006;231:1–11.
Hemler ME. Tetraspanin proteins promote multiple cancer stages. Nat Rev Cancer. 2014;14:49–60.
Kandouz M, Batist G. Gap junctions and connexins as therapeutic targets in cancer. Expert Opin Ther Targets. 2010;14:681–92.
Kwon MJ. Emerging roles of claudins in human cancer. Int J Mol Sci. 2013;14:18148–80.
Läubli H, Borsig L. Selectins promote tumor metastasis. Semin Cancer Biol. 2010;20:169–77.
Martin TA. The role of tight junctions in cancer metastasis. Semin Cell Dev Biol. 2014;36C:224–31.
Neufeld G, Kessler O. The semaphorins: versatile regulators of tumour progression and tumour angiogenesis. Nat Rev Cancer. 2008;8:632–45.
Neufeld G, Sabag AD, Rabinovicz N, Kessler O. Semaphorins in angiogenesis and tumor progression. Cold Spring Harb Perspect Med. 2012;2:a006718.
Ouban A, Ahmed AA. Claudins in human cancer: a review. Histol Histopathol. 2010;25:83–90.
Parsons JT, Horwitz AR, Schwartz MA. Cell adhesion: integrating cytoskeletal dynamics and cellular tension. Nat Rev Mol Cell Biol. 2010;11:633–43.
Pellet-Many C, Frankel P, Jia H, Zachary I. Neuropilins: structure, function and role in disease. Biochem J. 2008;411:211–26.
Rehman M, Tamagnone L. Semaphorins in cancer: biological mechanisms and therapeutic approaches. Semin Cell Dev Biol. 2013;24:179–89.
Roth L, Koncina E, Satkauskas S, Crémel G, Aunis D, Bagnard D. The many faces of semaphorins: from development to pathology. Cell Mol Life Sci. 2009;66:649–66.
Runkle EA, Mu D. Tight junction proteins: from barrier to tumorigenesis. Cancer Lett. 2013;337:41–8.
Sakurai A, Doçi CL, Gutkind JS. Semaphorin signaling in angiogenesis, lymphangiogenesis and cancer. Cell Res. 2012;22:23–32.
Suzuki K, Kumanogoh A, Kikutani H. Semaphorins and their receptors in immune cell interactions. Nat Immunol. 2008;9:17–23.
Takai Y, Miyoshi J, Ikeda W, Ogita H. Nectins and nectin-like molecules: roles in contact inhibition of cell movement and proliferation. Nat Rev Mol Cell Biol. 2008;9:603–15.

Tamagnone L. Emerging role of semaphorins as major regulatory signals and potential therapeutic targets in cancer. Cancer Cell. 2012;22:145–52.
Tran TS, Kolodkin AL, Bharadwaj R. Semaphorin regulation of cellular morphology. Annu Rev Cell Dev Biol. 2007;23:263–92.
Zhou Y, Gunput RA, Pasterkamp RJ. Semaphorin signaling: progress made and promises ahead. Trends Biochem Sci. 2008;33:161–70.

Toll-Like Receptors, Interleukin 1 and NFκB

12

Abstract

Toll-like receptors (TLR) constitute a family of receptors involved in immune and inflammatory processes, which recognise molecules derived from infectious agents (bacteria, viruses) and endogenous substances. They are studied in this chapter together with interleukin 1 (IL1) and its receptors (IL1R) and with related interleukins such as IL18 and IL33, with which they share a general activation mechanism. The activation of these receptors leads, after several steps involving cytoplasmic kinases, to the activation of transcription factors of the NFκB (*nuclear factor κB*) family, which activate the inflammatory response. NFκB can be activated in response to other signalling pathways, especially AKT in the PI3 kinase pathway (Chap. 3) or tumour necrosis factor (TNF). The signals generated by TLR/IL1R activation induce cell survival and proliferation: this is why the corresponding pathways can be diverted by oncogenesis processes.

The immune system contains two types of receptors able to recognise bacterial and viral components, called *pattern-recognition receptors* (PRR): membrane toll-like receptors and intracellular receptors called NLR (*NOD-like receptors*, NOD meaning *nucleotide binding and oligomerisation domain*). Activation of NLR activates in turn IL1 family interleukins, which, once secreted, induce an inflammatory response when recognised by cells equipped with their cognate receptors.

12.1 IL1 Family Interleukins and Their Receptors

The interleukins of this family are distinct from those studied in Chap. 4, because they are structurally different and induce a different signalling pathway. The IL1 family comprises 11 members distributed in four groups. In the IL1 group are IL1α (IL1F1) and IL1β (IL1F2) (genes *IL1A* and *IL1B*) as well as IL1RA (*IL1 receptor antagonist*) or IL1F3 (gene *IL1RN*), which cannot activate the signalling pathway

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_12

and is therefore an inhibitor of this pathway. IL18 (IL1F4) and IL33 (IL1F11) each constitute a group; the IL36 group, recently identified, comprises IL36α (IL1F6), IL36β (IL1F8), IL36γ (IL1F9), IL37 (IL1F7), IL38 (IL1F10) and IL36RN (IL1F5), which is an antagonistic ligand as IL1RA. IL1 family interleukins are synthesised as precursors which are activated by proteolytic cleavage, by calpain for IL1α and by caspase 1 or ICE (*IL1-converting enzyme*) for IL1β, IL18, IL33 and the other ones.

There are two IL1 receptors (IL1R1 and IL1R2) and two IL18 receptors (IL18R1 and IL18BP, for *IL18-binding protein*), but there is only one for IL33 (IL33R or ST2L, gene *IL1RL1*) and one for the group of IL36α, IL36β, IL37 and IL38 (IL36R, gene *IL1RL2*). IL1R1 contains a large cytoplasmic domain and is able to transduce appropriate signalling, whereas IL1R2 has lost the major part of this domain and cannot transduce any message; it behaves as a decoy receptor and can be removed from the membrane and adopt a soluble extracellular form; this is the same for IL18BP toward IL18. In addition to these receptors, there are accessory proteins (ILRAP), which are coreceptors able to form heterodimers with the primary receptors, constituting thus ternary complexes IL–ILR–ILRAP required for signal transduction. IL1α and β, IL33 and IL36 utilise IL1RAP, while IL18 utilises IL18RAP. The combinatory pattern of ligands and receptors of the IL1 family is presented Fig. 12.1a. Other 'orphan' receptors have also been identified. All these receptors, collectively called ILRs, contain, at the extracellular level, immunoglobulin-like domains and, for the active receptors, a domain called TIR (*toll-interleukin receptor*).

IL1 family interleukins are basic mediators of immunity and inflammation. Most of them have a proinflammatory role, against which act the antagonistic ligands, IL1RN and IL36RN, as well as the decoy receptors IL1R2 and IL18BP. Caspase 1 activation is the initiating event of IL1 activation. This occurs at the level of supramolecular platforms, called inflammasomes, where adapter proteins enable the autoproteolytic activation of procaspase 1 into an active caspase, similarly as apoptosomes enable the activation of procaspase 9 (Chap. 18). The physiological roles of IL1 will not be described here; we will only mention the involvement of IL1β in autoimmune diseases, in which the inflammatory aspect is important and which are called auto-inflammatory diseases. Numerous strategies for blocking IL1, its activators and its effectors, have been developed in therapeutics.

12.2 Toll-Like Receptors and Their Ligands

Toll-like receptors (TLRs) have been originally discovered in *Drosophila* and then researched and identified in mammals. They are characterised by intracytoplasmic domains that are homologous to those of IL1 family receptors, but their extracellular domains are different: they contain leucine-rich domains and no immunoglobulin-like domains. The signalling pathways downstream toll-like and IL1 receptors are mostly common. There are nine TLRs (TLR1 to TLR9), among which five are present on the plasma membrane and four on endosomal membranes (Fig. 12.1b). Their activation results from dimerisation, generally homodimerisation, sometimes heterodimerisation as the one occurring between TLR2 and its partners TLR1 and TLR6.

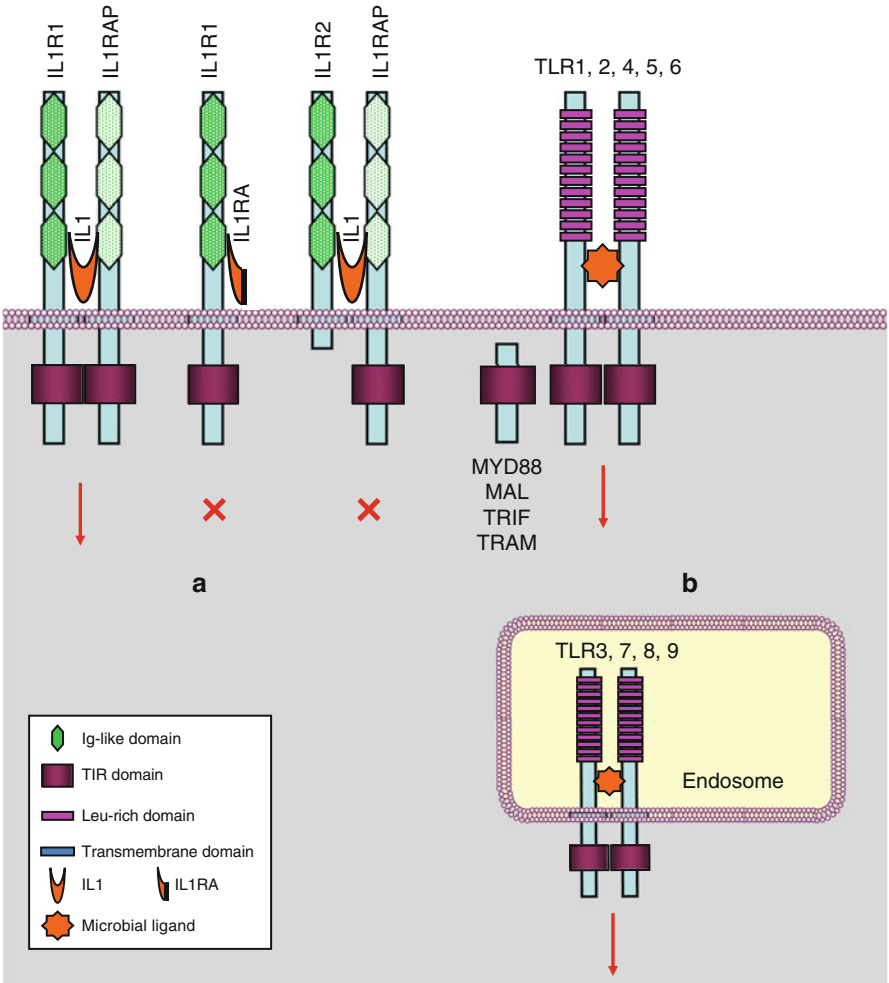


Fig. 12.1 Interleukin 1 receptors and toll-like receptors. (**a**) Interleukin 1 receptors (IL1R) transmit a signal (*red arrow*) when IL1 is bound to a complex IL1R1–IL1RAP (*IL1 receptor accessory protein*). IL1RA (*IL1 receptor antagonist*) cannot transmit a signal. IL1 binding on an IL1R2–IL1RAP complex cannot transmit a signal. (**b**) Toll-like receptors are localised on the plasma membrane or at the level of endosomes. They induce signalling when homodimerised, sometimes heterodimerised (TLR1 or 6 with TLR2)

The ligands able to activate TLRs are from extracellular or intracellular origin. These are essentially molecules of microbial origin, generically called PAMPs (*pathogen-associated molecular patterns*): lipopolysaccharides (LPS) of Gram-negative bacteria for TLR4; lipoproteins and lipoteichoic acid for TLR1, 2 and 6; flagellin for TLR5 and viral nucleic acid fragments, essentially RNAs, for the intracellular TLR3, 7, 8 and 9. Numerous endogenous ligands are also able to activate TLRs, mainly TLR2 and 4: heat-shock proteins (HSPs), HMGB1 (*high mobility*

group box 1) protein, uric acid crystals, surfactants, glycosaminoglycans, S100A9 protein and extracellular matrix proteins, such as fibrinogen and fibronectin.

Without entering into the details of the physiological roles of TLRs in immunity and inflammation, one can mention that TLRs, at the surface of epithelial cells, are at the forefront of recognition of infectious agents. Their roles in the fight against infection are varied: they are able to induce an inflammatory reaction, to activate NADPH oxidase which generates reactive oxygen and nitrogen species, to recruit leukocytes and macrophages, to allow the activation of cytokines such as interleukin 12 and interferons (Chap. 4) or interleukin 1, etc.

12.3 Signal Transduction from TLRs and ILRs

12.3.1 General Aspects

Several signalling pathways can be opened from ILRs and TLRs: the MAP kinase pathways (ERK, p38, JNK, Chap. 2); the NFκB pathway, which is presented below; as well as the IRF3 (*Interferon regulatory factor 3*) pathway, which regulates the activity of cytokines (Chap. 4). NFκB activation occurs following the activation of most ILRs and TLRs, with the exception of TLR3; IRF3 activation mainly occurs downstream TLR3 receptor activation.

The signals generated by ILRs and TLRs are overall similar, because both receptor families contain a common TIR (*toll-interleukin receptor*) domain in their intracellular segment. TIR domains are also present in the sequence of a series of adapter proteins which are recruited by homophilic interactions. The MYD88 (*myeloid differentiation primary response gene 88*) protein is an adapter protein common to all ILRs and TLRs; other adapter proteins are specific for TLR2 and/or TLR4, especially MAL/TIRAP (*MYD88 adapter-like/TIR domain-containing adapter protein*), TRIF (*TIR domain-containing adapter-inducing interferon β*) and TRAM (*TRIF-related adapter molecule*). Receptor activation occurs through dimerisation and is followed by the recruitment of these proteins, which will in turn recruit cytoplasmic kinases, called IRAKs (*IL1 receptor-associated kinases*).

12.3.2 NFκB Pathway

In the pathway leading to NFκB (*nuclear factor of kappa light polypeptide gene enhancer in B-cells*) downstream the activation of IL1R, TLR4 or the endosomal TLRs, the first step is the autophosphorylation of IRAK1, stimulated by IRAK4, which allows the recruitment of the adapter protein TRAF6 (*TNF receptor-associated factor 6*), a member of the TRAF family, and of another adapter protein, pellino 1 (gene *PELI1*), which together form the complex 1 (Fig. 12.2). The cytoplasmic constituents of complex 1 detach from the receptor and can interact with a membrane-bound MAP3 kinase (Chap. 2), TAK1 (*TGFβ-activated kinase 1*) (gene *MAP3K7*) and with two other adapter proteins, TAB1 and TAB2 (*TGFβ-activated*

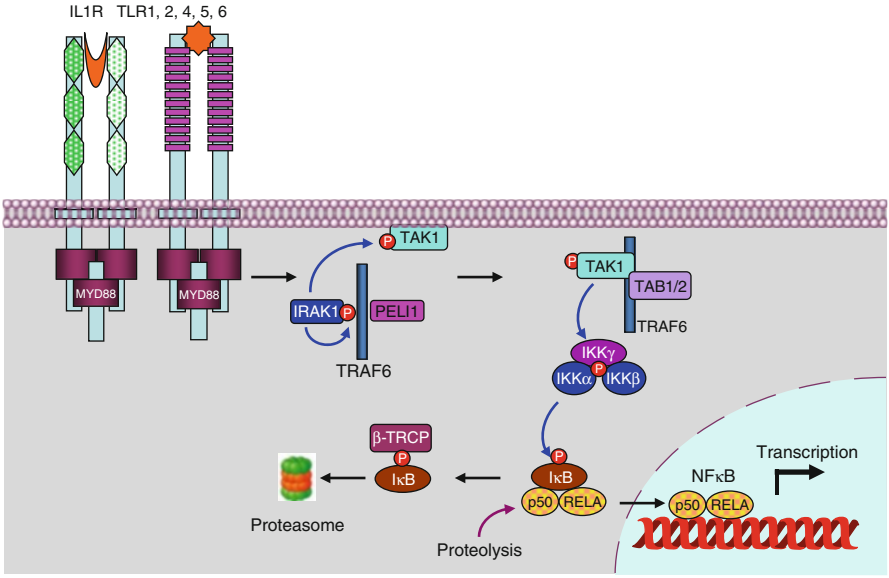


Fig. 12.2 The main signalling pathway downstream TLRs and ILRs. After the intervention of an adapter protein (MYD88), IRAK1 can autophosphorylate and recruit other adapter proteins (TRAF6, PELI1) and form a complex. This complex can activate the membrane MAP3 kinase, TAK1, through phosphorylation by IRAK1 and form another complex after recruitment of the adapter proteins TAB1 and TAB2. This complex can leave the membrane, and TAK1 can phosphorylate IKK, which is comprised of two kinase subunits, IKKα and β, and a regulatory subunit, IKKγ or NEMO. IKK can then phosphorylate the proteins IκB, which leads them to proteasome. The transcription factors of the NFκB–REL family are released, and their nuclear localisation sequence unmasked; they can thus translocate into the nucleus and exert their function on target genes. In parallel to this canonical pathway, other modalities can lead to NFκB or IRF3 activation; they require other adapter proteins and other kinases for the phosphorylation of IκB proteins, but the general scheme remains the same

kinase 1 binding protein) to form complex 2. At the level of complex 2, comprising IRAK1, TRAF6, TAK1, TAB1 and TAB2, IRAK1 can phosphorylate TAK1, which allows the dissociation of this complex from the membrane and its release in the cytosol as complex 3 (which comprises thus TRAF6, TAK1, TAB1 and TAB2). TRAF6 is an E3 ubiquitin ligase (see Annex C) able to perform auto-ubiquitinylation; this does not lead it to the proteasome, but on the contrary enables it to activate complex 3. TAK1 can thus phosphorylate a serine/threonine kinase called IKK (*IκB kinase*), which is responsible for the formation of NFκB. As TAK1 is a MAP3 kinase, it can also phosphorylate the MAP2 kinases of the p38 and JNK pathways (Chap. 2).

IKK is comprised of three protein subunits, two of which presenting kinase activity (IKKα or IKBKA [gene *CHUK*] and IKKβ [gene *IKBKB*]), while the third one presents kinase regulatory activity (IKKγ [gene *IKBKG*], also known as NEMO, for *NFκB essential modulator*). In this complex, IKK phosphorylates, on two serine residues, a group of proteins called IκB (α, β, δ, ε and ζ) (*inhibitors of NFκB*), which

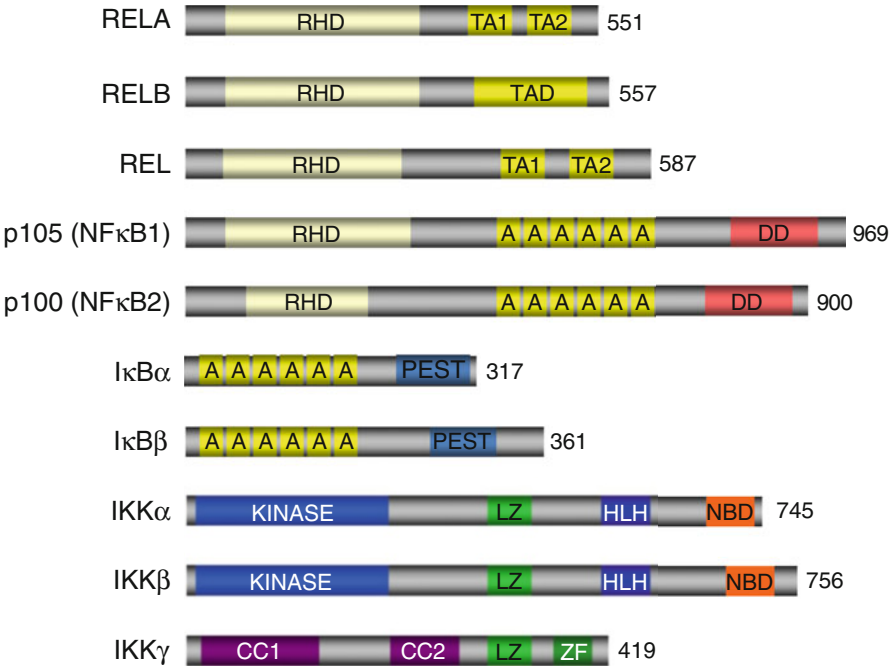


Fig. 12.3 General structure of NFκB, IκB and IKK proteins. The three protein families are presented, with their amino acid length and their characteristic domains. RHD, REL homology domain; TAD, transactivation domain (sometimes subdivided in TA1 and TA2); A, ankyrin repeat; DD, death domain; PEST, proline–glutamic acid–serine–threonine-rich domain; LZ, leucine zipper domain; ZF, zinc finger domain; HLH, helix–loop–helix domain; NBD, NEMO (IKKγ)-binding domain; CC, coiled-coil domain

are associated via ankyrin domains to the precursors of NFκB, which also display such domains. This allows IκB to be recognised by an E3 ubiquitin ligase, β-TRCP (*transducing repeat-containing protein*), of the SCF (*SKP1/cullin/F-box*) family, which drives it to ubiquitinylation and proteasomal proteolysis. The destruction of IκB is accompanied by the proteolysis-induced activation of NFκB precursors by unmasking a nuclear localisation sequence (Annex C), which drives NFκB transcription factors to the nucleus and to the recognition of their target gene promoters. IKK proteins also have functions other than those leading to NFκB activation, but which will not be envisaged here.

Active NFκB factors are heterodimers associating: (i) a strictly speaking NFκB factor NFκB1 (p50, gene *NFKB1*) or NFκB2 (p52, gene *NFKB2*); both bear IκB-interacting ankyrin domains and are activated by proteolysis of inactive precursors, p105 and p100; and (ii) a REL (*reticuloendotheliosis viral oncogene homolog*) factor, REL (c-REL), RELA (p65, NFκB3) or RELB. All these factors are characterised by a dimerisation and DNA-binding domain called RHD (*REL homology domain*). REL factors display in addition transcription activation domains (Fig. 12.3). It is therefore necessary that an NFκB factor and a REL factor be combined and that this heterodimer

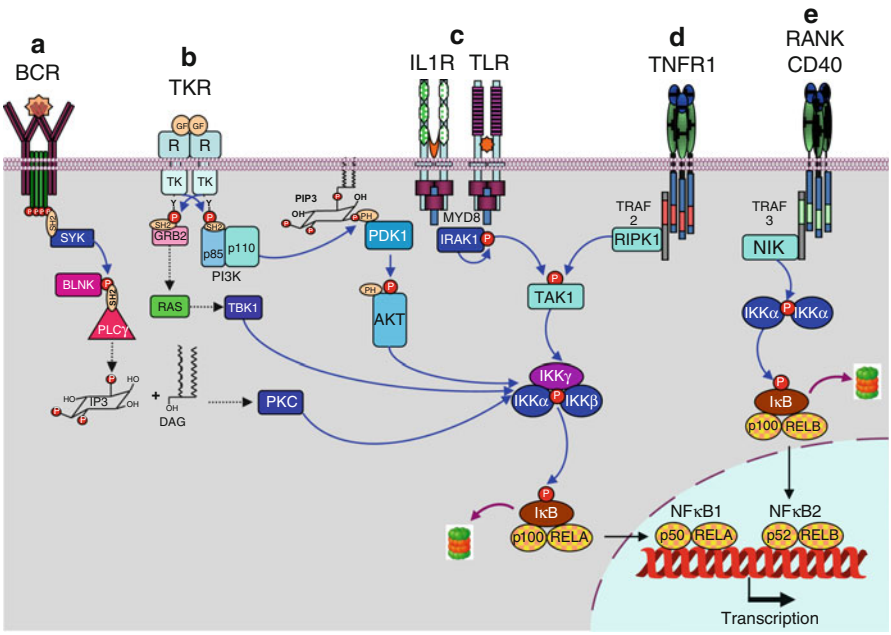


Fig. 12.4 NFκB activation pathways. Various receptor types are able to induce NFκB activation, from left to right: (**a**) lymphocyte receptors (BCR, TCR), via the activation of phospholipase Cγ and afterwards of protein kinase C; (**b**) some tyrosine kinase receptors (TKR), either via RAS activation and phosphorylation of TBK1 (*TANK-binding kinase 1*) or via the activation of PI3 kinase and AKT; (**c**) IL1 receptor (IL1R) and related interleukin (IL18, IL33) receptors and toll-like receptors (TLR), via the activation of TAK1 (*TGFβ-activated kinase 1*, gene *MAP3K7*); (**d**) receptors of the tumour necrosis factor (TNF) superfamily (TNFSFR), via the activation of RIPK1 (*receptor-interacting serine/threonine kinase 1*). All these activated pathways lead to the phosphorylation of the IKKα–IKKβ–IKKγ complex, which phosphorylates IκB. This phosphorylation drives it to the proteasome and enables the activating proteolysis of the NFκB1 precursor (p50) which, associated to RELA, can be translocated to the nucleus and exert its transcription factor function. A special event (**e**) concerns some receptors of the TNF superfamily (CD40, RANK), which activate a homodimeric complex of IKKα via NIK (*NFκB-inducing kinase*, gene *MAP3K14*) and then allow the activation of NFκB2 (p52) associated to RELB. Other adapter proteins are not included in this simplified schema

migrates in the nucleus, for rendering possible the transcription of target genes. NFκB factors are the common final outcome of the ILR and TLR pathways. They are also produced downstream the PI3K pathway, via AKT (Chap. 3), via the lymphocyte receptor pathway (Chap. 13) and by the TNF receptor pathway (Fig. 12.4), following activation schemes close to those operating after TLR or ILR activation. In addition, some kinases such as NIK (*NFκB-inducing kinase*) (gene *MAP3K14*) or RIPK1 (*TNF receptor-interacting serine/threonine kinase 1*), downstream TNF-family receptors activation, are able to phosphorylate IKK via other TRAF adapter protein.

NFκB transcription factors activate the transcription of numerous genes, such as those involved in the negative control of apoptosis (especially those encoding the antiapoptotic proteins BCL2, BCLXL, CIAP1, CIAP2, XIAP, survivin, Chap. 18);

in cell proliferation (cyclin D, MYC, Chap. 2); in cell adhesion and motility (ICAM, VCAM, fibronectin, metalloproteinases); in inflammation (IL6, Chap. 4); and in angiogenesis (VEGF, COX2). NFκB thus exerts positive effects on cell cycle entry, survival and proliferation. NFκB also induces the synthesis of IκB, which keeps it under inactive state, realising thus a negative feedback.

12.3.3 IRF3 Pathway

In the pathway leading to the transcription factor IFR3, the adapter protein TRAM is recruited upon activation of the TLR3 receptor at the level of endosomes and of TLR4 at the level of plasma membrane. Viral RNAs are the best-known ligands of these receptors. As in the pathway leading to NFκB, successive complexes are formed, which detach from the membrane to activate the transcription factor IRF3. To these complexes belong the adapter proteins TRIF and TRAF3, of the TRAF family, and kinases which play a role analogous to that of IKK, called TBK1 (TANK [*TRAF-associated NFκB activator*]-*binding kinase 1*) and IKKε (IKBKE). IRF3 is a transcription factor responsible for the transcription of the genes encoding the type I interferons (IFNα4, IFNβ, Chap. 4) and the chemokines RANTES (*regulated on activation*, *normal T-cell expressed and secreted*) or CCL5 and CXCL10 (Chap. 6).

12.4 Oncogenic Alterations and Pharmacological Targets

12.4.1 At the Receptor Level

TLRs open a major signalling pathway for response to bacterial and viral infections; with ILRs, these are essential mediators of the inflammatory reaction. At the level of the intestinal and pulmonary mucosae, these receptors play a role of epithelium protection against the ingested or breathed chemicals. They are also important for tissue homoeostasis, especially for tissue repair and regeneration. TLR activation may play a beneficial role for mucosae protection against carcinogenic compounds and for tumour cell lysis, thanks to the recruitment of immunologically competent cells. In addition, TLR3 activation induces proapoptotic pathways. The activity of certain antitumour vaccines could be mediated through TLR4 activation to induce antitumour immune response of T-cells.

These tumour suppressor effects of the TLR/ILR pathway are counterbalanced by obvious pro-oncogenic effects. TLR stimulation of tumour cell lines leads in vitro to an increase in cell survival and proliferation. Experimentally, tumour progression can be induced by TLR agonists, and carcinogenesis can be facilitated by downstream adapter proteins such as MYD88. Furthermore, there is a positive effect of these signalling pathways on angiogenesis, which certainly contributes to their pro-oncogenic properties. Oncogenic mutations of TLRs or ILRs have not yet been identified; however, some gene polymorphisms borne by TLRs are associated to differences in susceptibility to epithelial cancers (prostate, breast, colon, nasopharynx).

TLR stimulation may be used in therapeutics for the treatment of viral infections and cancers. Imiquimod, a derivative of imidazoquinoline, is an agonist for TLR7 and has been developed for the topical treatment of papillomavirus-associated cancers, such as vulvar intraepithelial neoplasia, as well as skin basocellular carcinomas. TLR9 agonists, such as agatolimod, a synthetic oligodeoxynucleotide mimicking single-strain viral DNA molecules, have also entered clinical trials in basocellular carcinomas, cutaneous T lymphomas and metastatic malignant melanomas. Similarly, flagellin preparations stimulating TLR5, and poly(A:U) and poly(I:C) (rintatolimod) stimulating TLR3, are in clinical trials. These compounds can induce immune responses but display moderate antitumour activity.

12.4.2 At the NFκB Level

In the absence of recurrent mutations of the kinases involved in the transduction of signals received by TLRs or ILRs, NFκB factors appear as the main modulators of the pro-oncogenic effects of this signalling pathway, because of their effects favouring cell survival and proliferation. Several mechanisms of NFκB activation in human tumours or cell lines have been described:

- Overexpression or hyperactivity of the different receptors (EGFR, ERBB2, MET, TNFR, integrins, cytokine receptors), whose activation especially induces NFκB activation.
- Overexpression or hyperactivity of the ligands of such receptors (IL1β, cytokines, TNF, HGF, etc.).
- Hyperactivity of cytoplasmic kinases such as JAK, ABL (in connection with the BCR–ABL translocation) and AKT (in connection with the oncogenic character of the PI3 kinase pathway).
- Mutations of the genes encoding the IκB proteins, especially IκBα (*NFKBIA*), more rarely IκBε (*NFKBIE*) in Hodgkin disease, together with the loss of the non-mutated allele, according to the classical process of activation of tumour suppressor genes.
- Mutations of the genes encoding the various components of NFκB, especially *REL*, whose amplifications and point mutations have been described in B-cell lymphomas. Rearrangements of the *NFKB2* gene have also been found in various haematopoietic malignancies.
- One can mention in addition germinal mutations of the genes encoding NEMO (IKKγ, gene *IKBKG*), IκBα (gene *NFKBIA*) and IRAK4, which lead to congenital hereditary diseases affecting the immune system.

Targeting NFκB could find multiple therapeutic applications, principally in rheumatology and oncology but, as a general feature, targeting transcription factors in a difficult task. The inhibition of several proteins upstream NFκB activation (AKT, PI3 kinase, MAP kinases, etc.) contributes to its downregulation. Indirectly, proteasome inhibition (Annex C) seems to slow down IκB destruction following its

phosphorylation, therefore decreasing NFκB activation. This could be one of the main mechanisms of action of bortezomib in the treatment of myelomas. Also upstream NFκB, targeting the serine/threonine kinase activity of IKKs is conceivable; small molecules extracted from natural products, such as curcumin, resveratrol or genistein, would act at this level, and an original inhibitor of IKK, CHS-828, has entered clinical trials. RNA interference (Annex B) is experimentally able to inhibit the synthesis of IKK, TAK1 or NFκB itself. Peptides mimicking the structure of NFκB or NEMO have also shown potential activity. Downstream NFκB, histone deacetylase inhibitors (Annex B) could prevent the execution of transcriptional programmes driven by NFκB. Finally, as some effectors of NFκB are involved in inflammation, corticosteroids, nonsteroid anti-inflammatories and COX2 inhibitors, developed as anticancer drugs, might eventually act on the effectors of this signalling pathway. In all cases, the use of such compounds will find its clinical applications in the potentialisation of chemotherapy rather than as anticancer drugs per se.

Further Reading

Arend WP, Palmer G, Gabay C. IL-1, IL-18, and IL-33 families of cytokines. Immunol Rev. 2008;223:20–38.

Baud V, Karin M. Is NF-kappaB a good target for cancer therapy? Hopes and pitfalls. Nat Rev Drug Discov. 2009;8:33–40.

Connolly DJ, O'Neill LA. New developments in toll-like receptor targeted therapeutics. Curr Opin Pharmacol. 2012;12:510–8.

Courtois G, Gilmore TD. Mutations in the NF-kappaB signaling pathway: implications for human disease. Oncogene. 2006;25:6831–43.

Dinarello CA. Immunological and inflammatory functions of the interleukin-1 family. Annu Rev Immunol. 2009;27:519–50.

El-Omar EM, Ng MT, Hold GL. Polymorphisms in toll-like receptor genes and risk of cancer. Oncogene. 2008;27:244–52.

Gay NJ, Gangloff M. Structure and function of toll receptors and their ligands. Annu Rev Biochem. 2007;76:141–65.

Hayden MS, Ghosh S. Shared principles in NF-kappaB signaling. Cell. 2008;132:344–62.

Hennessy EJ, Parker AE, O'Neill LA. Targeting toll-like receptors: emerging therapeutics? Nat Rev Drug Discov. 2010;9:293–307.

Kawai T, Akira S. TLR signaling. Semin Immunol. 2007;19:24–32.

Lamkanfi M, Dixit VM. Inflammasomes: guardians of cytosolic sanctity. Immunol Rev. 2009;227:95–105.

Li F, Sethi G. Targeting transcription factor NF-kappaB to overcome chemoresistance and radioresistance in cancer therapy. Biochim Biophys Acta. 2010;1805:167–80.

Lin Y, Bai L, Chen W, Xu S. The NF-kappaB activation pathways, emerging molecular targets for cancer prevention and therapy. Expert Opin Ther Targets. 2010;14:45–55.

Nakanishi C, Toi M. Nuclear factor-kappaB inhibitors as sensitizers to anticancer drugs. Nat Rev Cancer. 2005;5:297–309.

O'Neill LA. The interleukin-1 receptor/toll-like receptor superfamily: 10 years of progress. Immunol Rev. 2008;226:10–8.

O'Neill LA, Bryant CE, Doyle SL. Therapeutic targeting of toll-like receptors for infectious and inflammatory diseases and cancer. Pharmacol Rev. 2009;61:177–97.

Perkins ND. Integrating cell-signalling pathways with NF-kappaB and IKK function. Nat Rev Mol Cell Biol. 2007;8:49–62.

Perkins ND. The diverse and complex roles of NF-κB subunits in cancer. Nat Rev Cancer. 2012;12:121–32.

Rakoff-Nahoum S, Medzhitov R. Toll-like receptors and cancer. Nat Rev Cancer. 2009;9:57–63.

Sethi G, Sung B, Aggarwal BB. Nuclear factor-kappaB activation: from bench to bedside. Exp Biol Med. 2008;233:21–31.

Shen HM, Tergaonkar V. NF-kappaB signaling in carcinogenesis and as a potential molecular target for cancer therapy. Apoptosis. 2009;14:348–63.

Tamura T, Yanai H, Savitsky D, Taniguchi T. The IRF family transcription factors in immunity and oncogenesis. Annu Rev Immunol. 2008;26:535–84.

Lymphocyte Receptor Pathways

13

Abstract

The two types of lymphocytes, B and T, which are, respectively, responsible for humoral and cell-mediated immunity, express membrane receptors in relation to their immune functions: their task involves antigen recognition and elaboration of appropriate responses. Toward a multitude of possible antigens, a huge diversity in recognition structures is necessary: it is represented by the membrane immunoglobulins in charge of their binding. Contrasting to this diversity is the relative uniformity of receptor-associated proteins that are in charge of the elaboration of the response and are bound to the receptor itself by non-covalent bonds. Downstream the receptor complex, several pathways that have been studied in other chapters may be implemented to regulate lymphocyte proliferation and migration; they mainly involve the activation of cytoplasmic kinases, the generation of second messengers and the activation of small G-proteins.

We cannot summarise all immunology in a single chapter, but only extract some information related to the modalities of lymphocyte signalling, from receptors to effectors, trying to identify the features that cancer cells can utilise to reach their own goals and how these features can serve as potential targets for anticancer treatments.

13.1 B-Cell Receptors

13.1.1 B-Cell Receptor Activation

B-cell receptors (BCRs) are multiprotein complexes comprising a membrane immunoglobulin (Ig), which can bind the antigen and may be of infinite diversity, and a signalling element comprised of two constant proteins bound together via disulphide bridges, Igα (CD79A) and Igβ (CD79B). The antigen-binding element is a complete Ig, associating two heavy H chains and two light L chains, which is inserted in the plasma membrane through its Fc portion, while the *N*-terminal Fab portions are presented to the outside (Fig. 13.1). These Fab portions present a huge variety of sequences, obtained by gene recombination and somatic hypermutations; they determine the clonal specificity

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_13

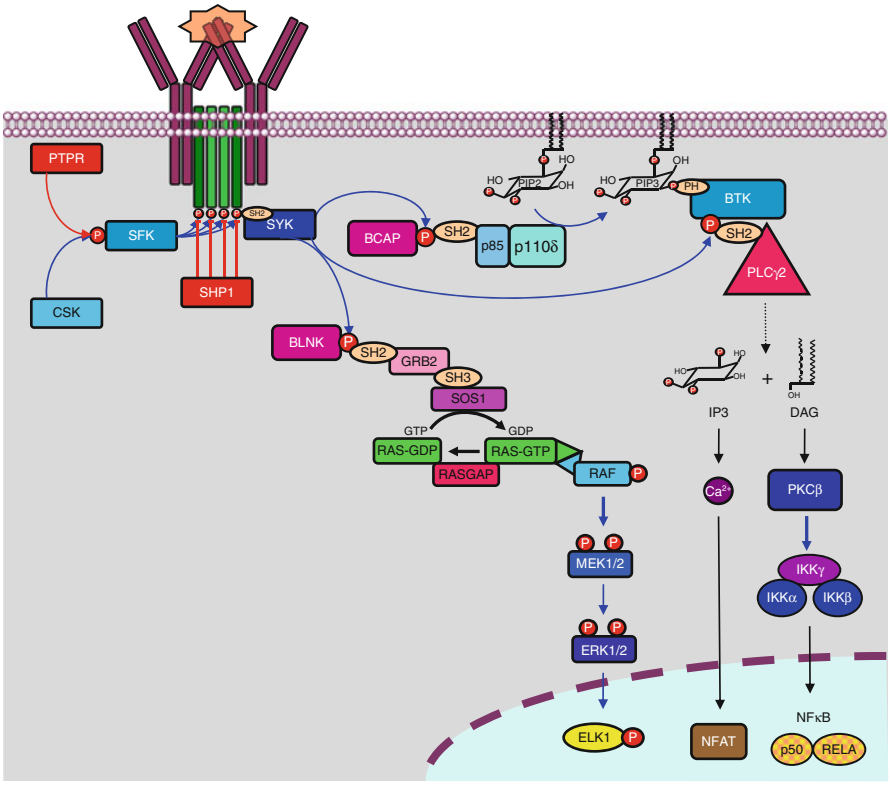


Fig. 13.1 B-cell receptors and associated signalling pathways. B-cell receptors are comprised of two immunoglobulin chains for ligand (soluble antigen) recognition and a complex formed of two α and two β chains for signal transduction. Phosphorylation of tyrosine residues of the α and β chains by LYN, an SRC family kinase, equilibrated by SHP phosphatases, allows the recruitment of another tyrosine kinase, SYK, which phosphorylates the adapter protein BLNK. BLNK induces the activation of proliferation pathways, thanks to the recognition of its phosphotyrosine residues by SH2 domain-containing proteins, of which an example is given (the MAP kinase pathway). Activated via BCAP (*B-cell cytoplasmic adapter for PI3K*), PI3Kδ plays a major role in activating the Bruton kinase, BTK, which in turn activates PLCγ2. PLCγ2 catalyses the formation of PIP3 and DAG from phosphatidylinositol 4,5-bisphosphate, and DAG activates PKCβ, a serine/threonine kinase at the origin of the production of active NFκB transcription factors through the phosphorylation of an adapter protein, CARD11, which forms a complex with two proteins, MALT1 and BCL10 (not shown). IP3 activates Ca^{2+} release from intracellular stores, which activates, among many other effects, the NFAT transcription factor

of antibodies, and each of them is able to recognise and bind an original antigen. The signalling element, Igα and Igβ, is associated to them through non-covalent bonds. Igα and Igβ both contain an activation motif containing two tyrosine residues (ITAM, *immunoreceptor tyrosine-based activation motif*) on the cytoplasmic side.

Cytoplasmic tyrosine kinases are in charge of the phosphorylation of these ITAM tyrosine residues (Fig. 13.1); this allows subsequent phosphotyrosine recognition by SH2 domain-containing proteins (see Chap. 1). These kinases belong to the SRC family kinases (SFK), especially LYN (*Yamaguchi sarcoma viral related oncogene*

homologue), FYN (*FES-YES-related novel gene*), BLK (*B-lymphoid tyrosine kinase*) and LCK (*lymphocyte-specific protein tyrosine kinase*). These kinases are anchored to the plasma membrane thanks to a myristoyl chain added after translation. Their activity is facilitated by the local aggregation of the receptors in lipid rafts; this aggregation takes place upon receptor activation and enables kinase clustering. Phosphatases like SHP1 (*SH2-containing phosphatase 1*) can be recruited to counterbalance kinase activation and obtain an adequate level of phosphorylation of the ITAM tyrosine residues. Outside the reception complex, LYN is itself submitted to a balance between inactivating phosphorylation, mediated by CSK (*C-terminal SRC tyrosine kinase*), and activating dephosphorylation mediated by CD45 or PTPRC (*protein tyrosine phosphatase*, *receptor type*, *C*) (Chap. 1).

The ITAM phosphotyrosine residues of Igα and Igβ constitute docking sites for another tyrosine kinase, SYK (*spleen tyrosine kinase*), which afterward recognises and phosphorylates target substrates serving as effectors of BCR activation, among which is BLNK (*B-cell linker protein*) or SLP65, able to recruit various SH2 domain-containing proteins and to serve as a scaffold for diverse signalling molecules. The main early effector of BCR signalling pathway is a cytoplasmic kinase of the TEC family, called *Bruton kinase* (BTK). This kinase, which is equipped with SH2, SH3 and PH domains, is first attracted to the membrane by interaction of its PH domain with phosphatidylinositol 3,4,5-trisphosphate (PIP3) and then activated by phosphorylation on Tyr^{551} by SYK or SFKs. PIP3 is generated by PI3 kinase δ (Chap. 3), which is activated by the binding of the SH2 domains of its regulatory subunit, p55 (gene *PIK3R3*), to the phosphotyrosine residues of an adapter protein, BCAP (*B-cell cytoplasmic adapter for PI3K*, gene *PIK3AP1*), generated by SYK.

After recognition and binding, the BCR-activating antigen is internalised with the receptor and sent to endosomal compartments to be presented later to the major histocompatibility complex. Igα and Igβ are in contrast kept in the membrane so that they can sustain signalling in the absence of the reception complex. Figure 13.1 presents the initiation of BCR signalling.

13.1.2 BCR Signal Transduction Pathways

BCR activation leads to several signalling pathways (Fig. 13.1), especially aimed at cell proliferation required for the clonal expansion of B lymphocytes. We have already presented most of these pathways, and we refer the reader to the specific chapters where details are given:

- The PI3 kinase pathway (Chap. 3), already mentioned, is activated through multiple ways: one of them involves the adapter protein BCAP and another one the adapter protein GAB2; these adapter proteins are phosphorylated by SYK, and the resulting phosphotyrosine residues are recognised by the SH2 domain of the regulatory subunits of PI3K. CD19, a membrane protein considered as a coreceptor for the BCRs, is also involved in PI3 kinase activation pathway via the adapter protein BCAP; in addition to BTK activation, mentioned above, the PI3 kinase pathway activates AKT and mTOR, in lymphocytes as in other cells.

- The MAP kinase (ERK) pathway (Chap. 2) is activated through the recruitment of GRB2 by the adapter protein BLNK; GRB2 recruits in turn SOS1, the usual GEF for RAS proteins, which activates the MAP3 kinases of the RAF subfamily. Other GEFs activate the other MAP kinase pathways (p38, JNK) via the recruitment of the adequate MAP3 kinases.
- The IKK pathway (Chap. 12), which leads to the production of NFκB transcription factors, is activated by several means: IKK can be activated either by AKT, downstream PI3 kinase, or by TAK1 through the activation of small G-proteins by recruitment of an SH3 domain-containing GEF; IKK activation can also occur in lymphocytes via a more complex way: BTK can activate a phospholipase C, PLCγ2 (gene *PLCG2*), recruited via binding the phosphotyrosine residues of BLNK; PLCγ2-mediated cleavage of phosphatidylinositol 4,5-bisphosphate generates diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3) (Chap. 6). IP3 activates calcium release from storage compartments to the cytosol, which activates in particular the NFAT transcription factor (Chap. 15), while DAG can in turn activate PKCβ, at the origin of the production of active NFκB transcription factors through the phosphorylation of an adapter protein, CARD11, which forms a complex with two proteins, MALT1 and BCL10, able to activate the IKK complex.

All these pathways converge toward transcription factors involved in lymphocyte survival and proliferation, which are studied in the corresponding chapters, such as MYC, ELK, JUN and above all NFκB, which appears as the main transcription regulator of BCR activation. BLNK phosphotyrosine residues can also be recognised by GEFs such as VAV1, which activate the small G-proteins of the RHO family (CDC42, RAC1 and others), which control cell motility by their actions on the actin cytoskeleton, via proteins of the WASP family (see below).

13.1.3 Oncogenic Alterations and Pharmacological Targets

Since the B-cell reception system is exclusively expressed in B lymphocytes, the only malignancies that could be associated to the deregulation of this pathway are lymphomas, which can occur either through specific alterations occurring in lymphocyte signalling pathways or through common alterations occurring specifically in lymphocytes.

It is suspected for more than 50 years that antigen stimulation can contribute to the genesis of malignant non-Hodgkin lymphoma, owing to chronic inflammatory states. In addition, antigenic stimulation of BCRs is required for B lymphocyte survival: as a consequence, the expression of a functional BCRs would be the way through which original molecular alterations (translocations) could be expressed in a transformed lymphocyte lineage. Lymphocyte proliferation would be dependent upon antigen stimulation, and indeed very few malignant lymphoma cells no longer express BCRs. This would explain why anti-infectious or anti-inflammatory treatments contribute to lymphoma regression; this is the case for lymphomas associated to the hepatitis C virus or to the bacteria *Helicobacter pylori*: in these situations, the antiviral or antibacterial treatment represents an efficient therapeutic option. More generally, any approach aiming at the elimination of the antigens involved in tumour cell expansion would be efficient.

In addition to this non-specific mechanism explaining the global involvement of BCRs in lymphoma genesis, molecular alterations of specific proteins involved in BCR signalling have been related to oncogenesis. Activating mutations of *CD79A* and *CD79B* have been identified in B-cell lymphomas. The tyrosine kinases of the SRC family as well as SYK do not seem to harbour common activating mutations in B-cell malignancies. BTK is crucial for the survival and proliferation of B cells; loss-of-function germinal mutations lead to X-linked agammaglobulinaemia (XLA), an immunodeficiency disease, but activating mutations have not been identified in lymphocyte proliferations. The *CARD11* gene presents activating mutations that disconnect NFκB production from BCR activation, and the genes encoding its partners, *MALT1* and *BCL10*, are rearranged in some B-cell lymphomas. As in common epithelial cancers, activating mutations in the MAP kinases and PI3 kinase pathways, when they occur in lymphocytes or lymphocyte progenitors, can contribute to oncogenesis.

Several potential therapeutic targets can be found in the BCR-activated pathway. SRC family kinases can be inhibited by various tyrosine kinase inhibitors (TKIs), such as dasatinib, bosutinib and saracatinib, but these TKIs are certainly not specific for the first kinase of the BCR pathway, LYN. A TKI developed for specific SYK targeting, fostamatinib, has entered clinical trials in non-Hodgkin lymphomas. The Bruton kinase benefits from a selective and irreversible TKI, ibrutinib, which has revealed a marked activity in malignant lymphoma, but not in those harbouring an activating mutation downstream BTK, such as those occurring in the *CARD11* gene. Moreover, ibrutinib appears able to select resistant clones through mutations in *BTK* itself or in *PLCG2*. PKCβ is also a potential target in lymphomas; however, the clinical trials performed with enzastaurin or sotrastaurin have been disappointing. Of course, the non-lymphocyte-specific pathways activated downstream BCR activation can be used for lymphoma treatment; the corresponding inhibitors have been subjected to clinical trials but will not be detailed here: MEK inhibitors (Chap. 2), PI3 kinase inhibitors (Chap. 3), rapalogues (Chap. 3), NFκB downregulators (Chap. 12), etc.

Tumour B lymphocytes can also be targeted in therapeutics independently of the pathways studied in this chapter; normal and tumour B lymphocytes are characterised by the presence of membrane proteins whose role remains most often elusive: CD20, CD52, CD22 and CD23 are the best known. These membrane proteins can be targeted with appropriate monoclonal antibodies: rituximab and ofatumumab against CD20, alemtuzumab against CD52, are available for the treatment of malignant lymphomas. Several newer monoclonal antibodies bearing a toxin or a radioactive element to be delivered to the malignant cells are in development.

13.2 T-Cell Receptors

13.2.1 T-Cell Receptor Activation

T-cell receptors (TCR) are also multiprotein complexes dedicated to the binding of antigens presented by antigen-presenting cells (APC) of various types and to the elaboration of adequate responses, especially T-cell proliferation. They comprise an antigen-recognition element and a signalling element (Fig. 13.2). They are associated to coreceptors, CD4 or CD8, also of the immunoglobulin superfamily,

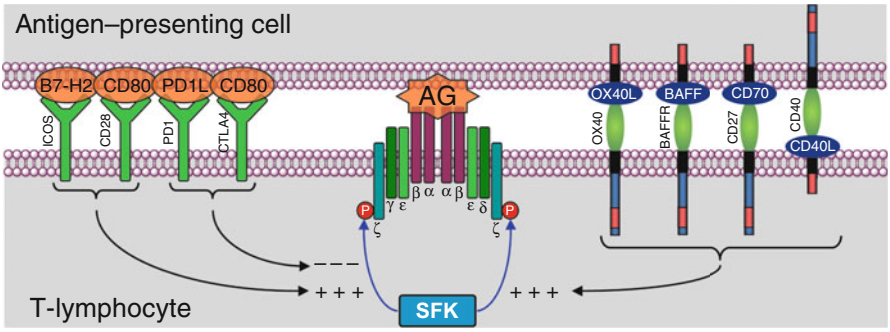


Fig. 13.2 T-cell receptors and their modulation. T-cell receptors are comprised of two groups of immunoglobulin αβ chains for ligand (antigen-presenting cells) recognition, and a CD3–CD247 complex associating two ε, two γ, two δ and two ζ chains. The phosphorylation of the ζ chains by FYN, an SRC family kinase (SFK), generates downstream signalling (Fig. 13.3). Accessory receptors of the immunoglobulin superfamily (left) allow the modulation of T-cell stimulation via the B7 proteins of the antigen-presenting cells, either positively (CD28, ICOS) or negatively (CTLA4, PD1). Modulators of the TNF receptor superfamily (right) (OX40, BAFFR, CD27, CD40) positively regulate T-cell receptor activation. TNFR superfamily members have been represented as monomers, whereas they are trimerised when activated by their ligand

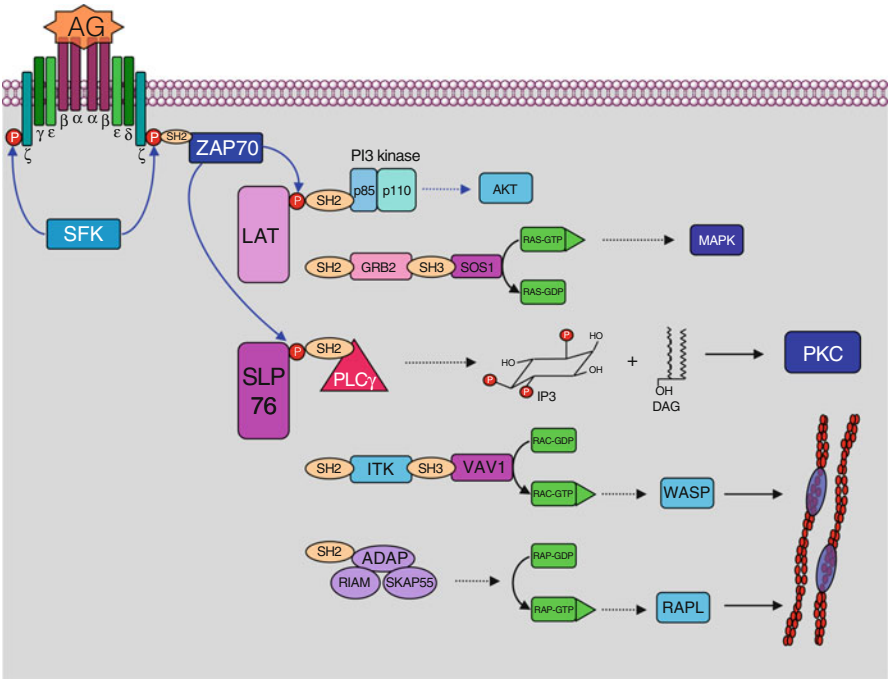


Fig. 13.3 T-cell receptor signalling pathways. The phosphorylation of the ζ chains by FYN allows the recruitment of the tyrosine kinase ZAP70, which phosphorylates adapter proteins such as LAT and SLP76. These proteins induce the activation of various pathways, thanks to the recognition of their phosphotyrosine residues by SH2 domain-containing proteins. Several examples are given: the MAP kinase pathway, the PI3 kinase pathway and the PLC pathway, with DAG and IP3 as second messengers. The TCR pathway also allows the activation of small G-proteins of the RHO family, which play a role in cell adhesion and migration, owing to their actions on cytoskeleton and their potential activation of integrins

which mediate T-cell binding to the molecules of the major histocompatibility complex (MHC), either of class I (CD8) or of class II (CD4), reinforcing thus the docking between T cells and APC. Another coreceptor, LAG3 or CD223, is in contrast with an inhibitor of TCR activation.

- The antigen-recognition element is comprised of two transmembrane polypeptides, α and β, whose glycosylated extracellular part is analogous to the Fab immunoglobulin portions; each of them contains a constant part and a variable part associated through a disulphide bond, a single transmembrane domain and a very short intracellular domain (4–10 amino acid residues). This element bears therefore the antibody diversity characteristic for immunoglobulins but is devoid of the Fc portion of immunoglobulins.
- The signalling element is comprised of transmembrane heterodimeric polypeptides, CD3, and a transmembrane homodimeric polypeptide, CD247. Three distinct molecules, CD3γ, CD3δ and CD3ε, encoded by the genes *CD3G*, *CD3D* and *CD3E*, are able to generate two CD3 heterodimers, CD3ε–CD3γ and CD3ε–CD3δ. CD247 is made of two ζ chains having a short extracellular domain. The intracellular segment of the CD3 proteins contains about 50 amino acid residues and that of the CD247 protein contains 110 residues. The ζ chains contain motifs with tyrosine residues susceptible of phosphorylation, the ITAMs.

The stoichiometry of the reception complex comprises two TCR heterodimers (TCRα–TCRβ), two CD3 heterodimers and one CD247 homodimer [$(\alpha\beta)_2$:εγ:εδ:ζζ]. These reception complexes are aggregated as multimeric clusters localised within the lipid rafts of the T-cell plasma membrane.

In addition to the main receptor, TCR, T lymphocytes display accessory or co-signalling receptors (Table 13.1), which are transmembrane immunoglobulins or members of the tumour necrosis factor receptor (TNFR) superfamily. Among those of the immunoglobulin superfamily, some have a stimulatory effect, such as CD28 or ICOS (*inducible T-cell co-stimulator*), whereas others have an inhibitory effect, such as CTLA4 (*cytotoxic T-lymphocyte antigen 4*) and PD1 (*programmed death 1*). TCR activation requires the binding, on these receptors, of co-signalling ligands collectively called B7 proteins (Table 13.1), also brought by antigen-presenting cells (APC). These ligands are also membrane proteins of the immunoglobulin superfamily. All the accessory co-signalling receptors of the TNFR superfamily, such as CD40, CD27 or CD137, are stimulators, and their cognate ligands belong to the TNF superfamily (Table 13.1).

As for BCRs, the first step of the signalling pathway is the activation of SRC family kinases such as FYN and LCK. These kinases phosphorylate the ITAM borne by CD247, which allows the recruitment of another cytoplasmic kinase, ZAP70 (*zeta-chain-associated protein kinase*, *70 kDa*). Adapter proteins with SH2 domains are then attracted to the membrane: LAT (*linker for the activation of T cells*) and SLP76 (*SH2 domain-containing leukocyte phosphoprotein of 76 kDa*) or LCP2 (*lymphocyte cytosolic protein 2*). These proteins are the required intermediates for the recruitment of the effector proteins.

Table 13.1 Ligands and co-signalling receptors of T lymphocytes

Superfamily	Effect on TCR	Receptor	Ligand
Immunoglobulins	Activator	*ICOS* (CD278)	*ICOSLG* (B7-H2, CD275)
		CD28	B7.1 (*CD80*) B7.2 (*CD86*)
	Inhibitor	*CTLA4* (CD152)	
		PD1 (*PDCD1*, CD279)	PDL1 (B7-H1, *CD274*)
			PDL2 (*PDCD1LG2*, B7-DC, CD273)
		TLT2 (*TREML2*)	B7-H3 (*CD276*)
		HVEM[a] (*TNFRSF14*, CD270)	*BTLA* (CD272)
		VISTA (PD1H)	B7-H4 (VTCN1)
		NCR3 (CD337)	*NCR3LG1* (B7-H6)
TNFSF TNFRSF	Activator	*CD40* (TNFRSF5)	*CD40LG* (CD154, TNFSF5)
		OX40 (CD134, *TNFRSF4*)	OX40L (CD252, *TNFSF4*)
		CD27 (TNFRSF7)	*CD70* (CD27L, TNFSF7)
		CD137 (4-1BB, *TNFRSF9*)	CD137L (4-1BB-L, *TNFSF9*)
		GITR (CD357, *TNFRSF18*)	GITRL (*TNFSF18*)
		BCMA (CD269, *TNFRSF17*/13A)	APRIL (CD256, *TNFSF13*)
		BAFFR (CD268, *TNFRSF13C*)	BAFF (CD257, *TNFSF13B*)
		TACI (CD267, *TNFRSF13B*)	

Official gene names are in italics

Abbreviations: APRIL, A proliferation-inducing ligand; BAFF, B-cell-activating factor; BCMA, B-cell maturation antigen; BTLA, B- and T-lymphocyte attenuator; CTLA4, cytotoxic T-lymphocyte antigen 4; GITR, glucocorticoid-induced TNFR-related protein; HVEM, herpesvirus entry mediator A; ICOS, inducible T-cell co-stimulator; NCR3, natural cytotoxicity triggering receptor 3; PD1, programmed cell death 1; TACI, transmembrane activator and CAML [calcium modulating ligand] interact; TLT2, TREM1 [triggering receptor expressed on myeloid cells]-like transcript 2; VISTA, V-domain Ig suppressor of T-cell activation; VTCN1, V-set domain-containing T-cell activation inhibitor 1

[a]HEVM is a receptor of the TNFR superfamily whose ligand is BTLA, of the immunoglobulin superfamily

A general inhibitor of T-cell activation is a protein called CBL (*Casitas B-lineage lymphoma*), which is an E3 ubiquitin ligase able to recognise phosphotyrosine residues; this molecule induces immune tolerance and plays a major role in the control of T-cell activation, especially in order to avoid self-recognition and autoimmune phenomena.

In addition to the T-lymphocyte-specific receptors, T cells express the channel receptor of Ca^{2+}, CRAC (*calcium release-activated calcium modulator*), which allows the activation of calcineurin, a phosphatase that activates the transcription

factor NFAT (*nuclear factor of activated T cells*). This transcription factor, despite of its name, is not specific for T cells and is studied in Chap. 15.

13.2.2 TCR Signal Transduction Pathway

A number of effector proteins can exert their effects through their recruitment by the adapter proteins LAT and SLP76. Most of the pathways activated by T-cell receptor stimulation are involved in cell proliferation and survival, in order to maintain the immune function of T cells and expand the adequate clones; and in cell adhesion and motility, in order to enable T-cell migration to their sites of action. Several proteins that activate these pathways have already been presented in other chapters, as they are not specific for T-cell activation:

- GRB2, which activates the MAK kinase pathway through RAS activation (Chap. 2).
- p85 (*PIK3R1*), the regulatory subunit of PI3 kinases, which opens the PI3 kinase pathway (Chap. 3) by recruiting the catalytic subunit PIK3CA.
- PLCγ, which allows the formation of DAG and IP3, and therefore opens Ca^{2+} signalling as well as PKC activation (Chap. 6).
- PKCθ, which induces the formation of a complex called CBM, specific of lymphocyte activation; this complex comprises the proteins CARMA1 (*CARD* [*caspase recruitment domain*] *and membrane-associated guanylate kinase*), BCL10 (*B-cell lymphoma 10*), an adapter protein with a CARD domain, and MALT1 (*mucosa-associated lymphoid tissue lymphoma translocation gene 1*). This complex enables the activation of NIK (*NFκB-inducing kinase*), an IKK-activating kinase in charge of NFκB activation (Chap. 12).
- ITK (*IL2-induced tyrosine kinase*), a serine/threonine kinase of the TEC family, which allows the recruitment of VAV1, a GEF activating small G-proteins of the RHO family, such as CDC42 and RAC1; these proteins are involved in cell adhesion and migration, through proteins interacting with the actin cytoskeleton, called WASP (*Wiskott–Aldrich syndrome proteins*). This pathway enables the intracellular activation of integrins by promoting their aggregation (Chap. 10). ITK contains a PH domain explaining its attraction to the membrane and its activation of the AKT pathway.
- RAP1 (*regulator of adhesion and polarisation 1*), which is a small G-protein activated by a trimeric complex recruited by SLP76: ADAP (*adhesion and degranulation-promoting adapter protein*), RIAM (*RAP1-interacting adapter molecule*) and SKAP55 (*SRC kinase-associated phosphoprotein of 55 kDa*); RAP1 is also a mediator of intracellular activation of integrins (Chap. 10) and plays a role in cell migration.

As BCR activation by soluble antigens, TCR activation by APC results in both clonal expansion and migratory capacities of the corresponding lymphocytes.

13.2.3 Oncogenic Alterations and Pharmacological Targets

T lymphocytes play a major role in the immune surveillance of tumours and the implementation of antitumour immunity; their capacity of survival and activation of their targets, especially IL2, is crucial for the inhibition of tumour cell proliferation in several cancer types, in particular malignant melanomas. The immune response toward tumour cells is most often inefficient or insufficient; most tumour antigens are not specific to tumour cells, and this prevents or delays the development of immune therapies targeting such antigens. In addition, cancers develop a variety of mechanisms to escape the host immune surveillance: mutations modifying the epitopes recognised by T cells, downregulation of co-stimulatory molecules, secretion of immunosuppressive cytokines, etc.

The therapeutic approaches in this field have long been disappointing, but the landscape of immunotherapy has completely changed within a few years: the inhibition of the co-inhibitory signalling molecules, both at the ligand and at the receptor levels, is possible thanks to monoclonal antibodies, and such treatments are now marketed or in late clinical evaluation for the treatment of metastatic malignant melanomas and advanced kidney cancers, and much hope is placed in other cancers. An anti-CTLA4 monoclonal antibody, ipilimumab, is available for the treatment of malignant melanomas, and another one, tremelimumab, is in development. Anti-PD1 (nivolumab, pembrolizumab, pidilizumab) and anti-PDL1 (BMS-936559) monoclonal antibodies will also be soon available. An original approach to stimulate the immune targeting of tumour cells consists in bispecific monoclonal antibodies, one target being a tumour antigen such as ERBB2 (breast cancers) or CD19 (B-cell lymphomas), the other target being CD3: ertumaxomab and blinatumomab are the respective *bispecific T-cell engagers* (BiTE) in development in these malignancies. In contrast, the co-signalling of the TNFR superfamily is all stimulatory and can only be targeted with agonist antibodies; dacetuzumab and lucatumumab are such agonist antibodies of CD40.

Cell therapy using genetically engineered T cells is an option that has been considered for a long time and has elicited numerous clinical trials. This can be conceived either by transferring to T-cell-specific genes encoding tumour antigen receptors, so that T cells will recognise the antigen-bearing cancer cells once reintroduced in the host or by transferring to T cells some genes involved in general capacities for survival, proliferation or migration, so that T cells will optimise their functions against cancer cells.

Another approach consists in combatting the immunosuppressive effects of CBL by small molecules interfering with its RING domain and in the future through interfering RNAs.

Further Reading

Anders K, Blankenstein T. Molecular pathways: comparing the effects of drugs and T cells to effectively target oncogenes. Clin Cancer Res. 2013;19:320–6.

Brezski RJ, Monroe JG. B-cell receptor. Adv Exp Med Biol. 2008;640:12–21.

Burbach BJ, Medeiros RB, Mueller KL, Shimizu Y. T-cell receptor signaling to integrins. Immunol Rev. 2007;218:65–81.

Ceeraz S, Nowak EC, Noelle RJ. B7 family checkpoint regulators in immune regulation and disease. Trends Immunol. 2013;34:556–63.

Cronin SJ, Penninger JM. From T-cell activation signals to signaling control of anti-cancer immunity. Immunol Rev. 2007;220:151–68.

Dal Porto JM, Gauld SB, Merrell KT, Mills D, Pugh-Bernard AE, Cambier J. B-cell antigen receptor signaling. Mol Immunol. 2004;41:599–613.

Gomez-Rodriguez J, Readinger JA, Viorritto IC, Mueller KL, Houghtling RA, Schwartzberg PL. Tec kinases, actin, and cell adhesion. Immunol Rev. 2007;218:45–64.

Hendriks RW, Yuvaraj S, Kil LP. Targeting Bruton's tyrosine kinase in B cell malignancies. Nat Rev Cancer. 2014;14:219–32.

Katzav S. Flesh and blood: the story of Vav1, a gene that signals in hematopoietic cells but can be transforming in human malignancies. Cancer Lett. 2007;255:241–54.

Kershaw MH, Westwood JA, Darcy PK. Gene-engineered T cells for cancer therapy. Nat Rev Cancer. 2013;13:525–41.

Kirkwood JM, Tarhini AA, Panelli MC, Moschos SJ, Zarour HM, Butterfield LH, Gogas HJ. Next generation of immunotherapy for melanoma. J Clin Oncol. 2008;26:3445–55.

Küppers R. Mechanisms of B-cell lymphoma pathogenesis. Nat Rev Cancer. 2005;5:251–62.

Peggs KS, Quezada SA, Allison JP. Cell intrinsic mechanisms of T-cell inhibition and application to cancer therapy. Immunol Rev. 2008;224:141–65.

Rojo JM, Bello R, Portolés P. T-cell receptor. Adv Exp Med Biol. 2008;640:1–11.

Smith-Garvin JE, Koretzky GA, Jordan MS. T-cell activation. Annu Rev Immunol. 2009;27:591–619.

Xue S, Gillmore R, Downs A, Tsallios A, Holler A, Gao L, Wong V, Morris E, Stauss HJ. Exploiting T-cell receptor genes for cancer immunotherapy. Clin Exp Immunol. 2005;139:167–72.

Yao S, Zhu Y, Chen L. Advances in targeting cell surface signalling molecules for immune modulation. Nat Rev Drug Discov. 2013;12:130–46.

Young RM, Staudt LM. Targeting pathological B cell receptor signalling in lymphoid malignancies. Nat Rev Drug Discov. 2013;12:229–43.

Nuclear Receptor Pathways

14

Abstract

Nuclear receptors constitute a special class of transcription factors that are activated by an extracellular signal. This signal is a hydrophobic substance, such as a steroid hormone, which can cross the plasma membrane and does not require a surface receptor to recognise it and transmit to the nucleus the information it conveys. The intracellular receptor is in charge of this function and of the regulation of target gene expression, after having recognised and bound its ligands. This is the most obvious difference between steroid hormones and polypeptidic hormones, which are recognised by G-protein-coupled receptors (GPCRs) (Chap. 6).

Nuclear receptors are the targets of many drugs aimed at treating metabolic and endocrinal diseases, drugs that can be agonists or antagonists of receptor activity. Two types of frequent cancers are hormone dependent: breast and prostate cancers. The nuclear receptors of oestrogens and androgens are, therefore, of major importance in oncology, because their alterations can participate to oncogenesis and cancer development and because they constitute major therapeutic targets. The hormonal treatments of breast cancer were historically the first 'targeted therapies' of cancers, but other nuclear receptors also present a major interest in oncology.

14.1 Structure and Function of Nuclear Receptors

There are about 50 nuclear receptors distributed in three major classes (Table 14.1); one of them (NR3 for *nuclear receptor 3*) contains especially the steroid hormones receptors: oestrogens (ERα and ERβ, genes *ESR1* and *ESR2*), progesterone (PGR), androgens (AR), glucocorticoids (GR) and mineralocorticoids (MR). Another class (NR1) gathers the receptors of thyroid hormones TRα and TRβ (genes *THRA* and *THRB*), vitamin D (VDR), retinoic acid (RARα, RARβ and RARγ; genes *RARA*, *RARB* and *RARG*), oxysterols (LXR, *liver X receptor*), biliary acids (FXR, *farnesoid X receptor*), fatty acids and eicosanoids (PPARα, PPARβ and PPARγ, *peroxisome proliferator-activated receptor*, genes *PPARA*, *PPARD*, *PPARG*) and various

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_14

Table 14.1 Main nuclear receptors and their ligands

Usual names	Nomenclature	Main ligands
TRα, TRβ	NR1A1, NR1A2	Thyroxine, triiodothyronine
RARα, RARβ, RARγ	NR1B1, NR1B2, NR1B3	All-*trans* retinoic acid
PPARα, PPARβ, PPARγ	NR1C1, NR1C2, NR1C3	Fatty acids, leukotrienes, prostaglandins
RORα, RORβ, RORγ	NR1F1, NR1F2, NR1F3	Cholesterol, cholesterol sulphate
LXRα, LXRβ	NR1H3, NR1H2	Oxysterols
FXRα, FXRβ	NR1H4, NR1H5	Biliary acids
VDR	NR1I1	Vitamin D
PXR, CAR	NR1I2, NR1I3	Xenobiotics
RXRα, RXRβ, RXRγ	NR2B1, NR2B2, NR2B3	9-*cis*-retinoic acid
ERα and ERβ (*ESR1* and *ESR2*)	NR3A1 et NR3A2	Oestradiol, tamoxifen
ERRα, β, γ	NR3B1, NR3B2, NR3B3	Diethylstilbestrol
GR	NR3C1	Cortisol
MR	NR3C2	Aldosterone
PR (*PGR*)	NR3C3	Progesterone
AR	NR3C4	Testosterone
AHR		Aromatic hydrocarbons (dioxin)

N.B. Only the physiological ligands are mentioned. Some receptors also accept various drugs as ligands. There are also numerous orphan receptors not listed here. AHR (*aryl hydrocarbon receptor*) or dioxin receptor is related to the nuclear receptors

xenobiotics (PXR, *pregnane X receptor*, gene *NR1I2*; and CAR, *constitutive androstane receptor*, gene *NR1I3*). The last class (NR2) only contains the accessory receptors of retinoic acid (RXRα, RXRβ and RXRγ, *retinoid X receptors*, genes *RXRA*, *RXRB* and *RXRG*). Each class contains in addition some related 'orphan' receptors, whose ligand remains unknown, such as ERRα, ERRβ and ERRγ (*oestrogen receptor-related receptors*). The structure of some ligands of nuclear receptors is presented on Fig. 14.1.

Nuclear receptors contain an *N*-terminal domain of transcription activation in the absence of ligand (AF1, *activation function 1*), a central domain for DNA binding (DBD, *DNA-binding domain*), owing to two zinc finger domains, a hinge domain and a *C*-terminal domain for ligand recognition and binding (LBD, *ligand-binding domain*), containing the AF2 site for dimerisation and ligand-activated transcription (Fig. 14.2). Depending on the case, nuclear receptors can act as monomers, homodimers or heterodimers and, in this case, always with a RXR receptor. They recognise on DNA a sequence called HRE (*hormone-responsive element*). This sequence is either repeated in tandem, forming two half-sites separated by a variable number of

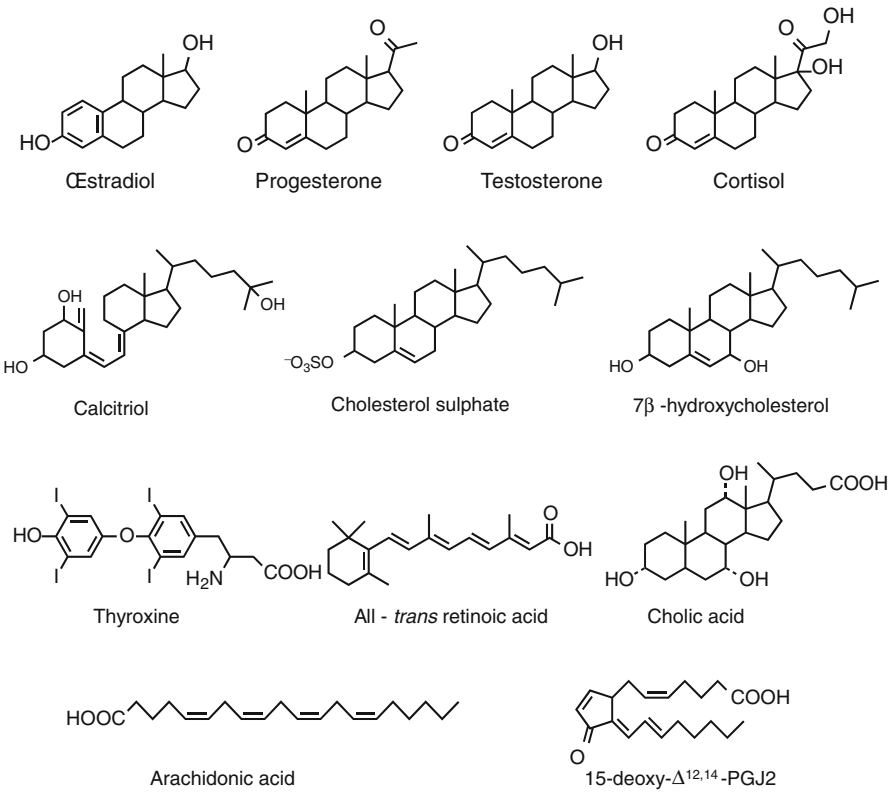


Fig. 14.1 Structure of the main ligands of some nuclear receptors. The ligands presented here recognise the following receptors: oestradiol, oestrogen receptor (ER, gene *ESR1*); progesterone, progesterone receptor (PR, gene *PGR*); testosterone, androgen receptor (AR); cortisol, glucocorticoid receptor (GR, gene *NR3C1*); calcitriol, vitamin D receptor (VDR); cholesterol sulphate, RAR-related orphan receptor (RORα, gene *RORA*); 7β-hydroxycholesterol, liver X-receptor (LXR, gene *NR1H2*); thyroxine, thyroid hormones receptor (TRα, gene *THRA*); all-*trans* retinoic acid (ATRA), retinoic acid receptor (RARα, gene *RARA*); cholic acid, biliary acid receptor (FXR, gene *NR1H4*); arachidonic acid and prostaglandin 15-deoxy-$\Delta^{12,14}$-PGJ2, peroxisome proliferator–activated receptor (PPARα, gene *PPARA*)

nucleotides (most frequently 3), or repeated in an inverted, palindromic way, the two half-sites being also separated by a variable number of nucleotides. In both cases, each receptor monomer binds on one of the half-sites.

Ligand–receptor binding can occur in the cytoplasm; the translocation of the activated receptor to the nucleus is possible thanks to the unmasking of a nuclear localisation sequence (Annex C). In other cases, the receptor is permanently localised in the nucleus; in the absence of ligand, the LBD is bound to a corepressor NCOR (*nuclear receptor corepressor*); when the LBD has recognised and bound a ligand, the transcription activation domain is activated owing to the intervention of a coactivator NCOA (*nuclear receptor coactivator*). The function of corepressors and

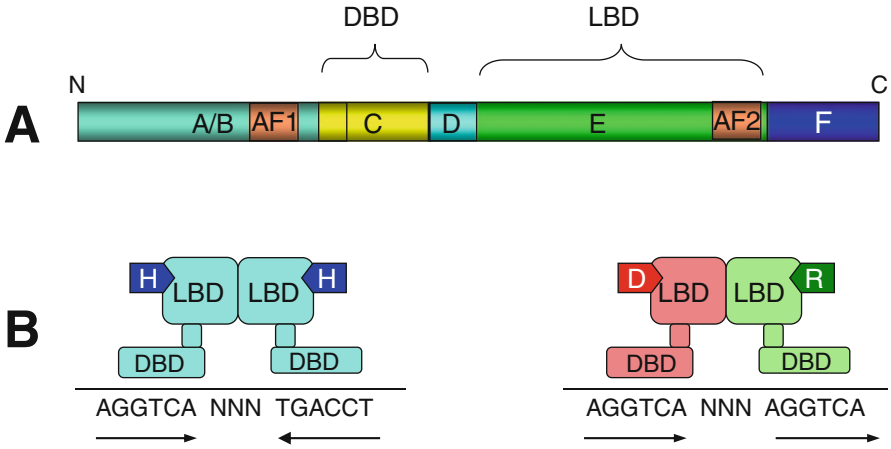


Fig. 14.2 General structure of nuclear receptors. (**a**) The polypeptidic sequence can be subdivided in *A*/*B*, *C*, *D*, *E* and *F* domains. The *A*/*B* domain contains a sequence AF1 (*activation function 1*) which enables the receptor to exert weak activity in the absence of ligand. *C* is the DNA-binding domain (*DBD*) which contains the zinc finger motifs recognising the hormone-responsive elements (HRE) on the target genes. *D* is a hinge domain allowing the receptor to fold after ligand binding and dimerisation. *E* is the ligand-binding domain (*LBD*), which also contains the dimerisation site and the AF2 sequence for ligand-dependent transactivation. (**b**) The interaction between ligand and receptor leads to the binding of the activated receptor on a DNA sequence. *F*or NR3 receptors (*left*), there is homodimerisation of the hormone (H)-activated receptor and recognition of a palindromic sequence; for NR1 receptors (*right*), there is heterodimerisation of the specific receptor (VDR in this example) activated by its ligand (vitamin D) and the common RXR receptor activated by 9-*cis*-retinoic acid, with recognition of a tandem sequence

coactivators is to recruit histone-modifying enzymes, which control the degree of methylation and acetylation of histones (Annex B). Histone deacetylation induces a compaction of the chromatin that does not allow target gene transcription, whereas their acetylation facilitates transcription (Fig. 14.3). Nuclear receptors can exert other effects, independently of their transcription function: some of them are able, in the cytoplasm, to activate signalling pathways such as the MAP kinases or the PI3 kinase pathways (Chaps. 2 and 3) or the pathways related to Ca^{2+} release in the cytosol (Chap. 6).

It is impossible to establish a list, even succinct, of the genes whose transcription is activated by nuclear receptors. Several hundreds of genes harbour, in their promoter sequence, responsive elements specific for one or several activated receptors. The research of target genes can be realised in silico, thanks to algorithms of identification of these promoter sequences over the whole genome, or by analysing the gene expression profiles obtained before and after treatment by a given ligand. The genes transcribed by an activated receptor reflect the characteristics of the cognate hormones, in the fields of cell metabolism, proliferation and differentiation. Oestrogens and androgens, for instance, induce a proliferation of the cells which express oestrogens and androgen receptors, respectively, which can be used by hormone-dependent cancers.

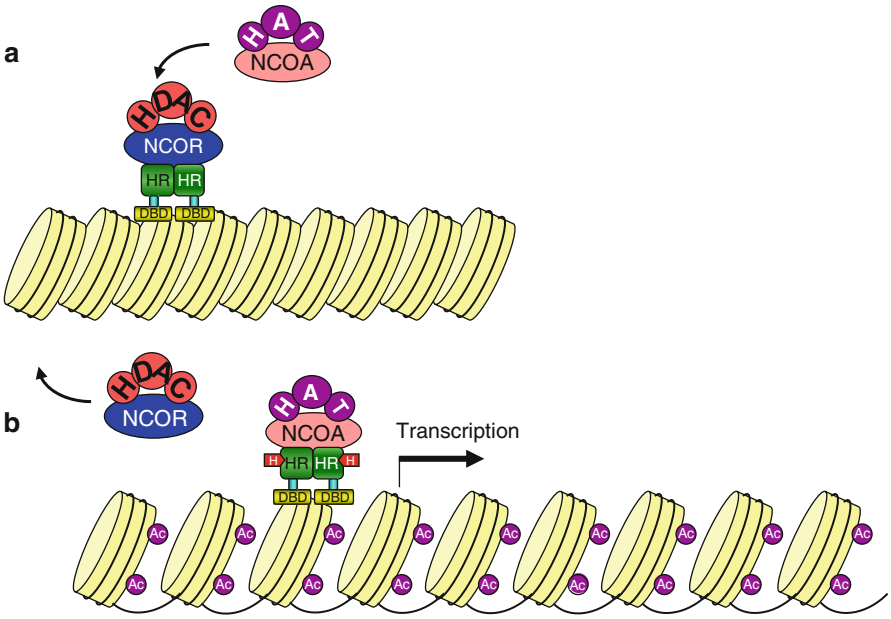


Fig. 14.3 Transcription activation by nuclear receptors. (**a**) In the absence of ligand, the transcriptional activity of the nuclear receptor is inhibited by a corepressor NCOR, which is associated to histone deacetylase activity HDAC. Chromatin is compact and does not enable transcription. (**b**) After ligand binding, the receptors can be combined with a coactivator NCOA, which recruits a histone acetyltransferase HAT; chromatin is thus relaxed and the transcription of target genes becomes possible

14.2 Steroid Hormones Receptors

Five classes of steroid hormones can activate nuclear receptors: oestrogens, progesterone, androgen, glucocorticoids and mineralocorticoids. In the absence of ligand, these receptors are associated in the cytoplasm to chaperone HSPs (*heat shock proteins*) which maintain them in inactive state. Ligand binding induces the removal of the HSP, the homodimerisation of the receptor and its translocation into the nucleus. The dimer recognises the target sequence HRE and triggers the transcription of the corresponding genes.

14.2.1 Oestrogens and Progesterone Receptors

The activity of oestrogens on breast cancer growth has been known for a long time, since castration was proposed more than one century ago as breast cancer treatment. Oestrogens favour the proliferation of tissues expressing ERα (NR3A1, gene *ESR1*), which occurs in 75 % of breast cancers. Oestrogens, used as hormone replacement therapy of menopause, have induced an increase in breast cancer incidence, which

is now decreasing, since an international warning on their use has been outspread. The second oestrogen receptor, ERβ (NR3A2, gene *ESR2*), is a negative regulator of ERα in normal breast tissue, in controlling its transcriptional activity through heterodimerisation, which decreases the transcription rate.

Since the 1970s, a hormonal therapy targeting this relationship between oestrogens and breast cancer growth has been implemented: tamoxifen and its derivatives, under the generic name of SERM (*selective oestrogen receptor modulators*), display an antagonistic effect to oestradiol, which is in some tissues partially agonistic. After binding to the receptor, they prevent some oestradiol transcriptional effects, but not all. In contrast, SERDs (*selective oestrogen receptor downregulators*), such as fulvestrant, bind to ERα with the same affinity as oestradiol but without any agonistic activity and inhibit receptor expression at the transcriptional level. In addition to the compounds which bind ERα, another therapeutic approach is available for hormone-dependent breast cancers: the inhibition of aromatase, the enzyme that produces C18 steroids (oestrogens) from C19 steroids (androgens).

Hormone treatments can only be efficient when breast tumour tissue expresses ERα. However, about 30 % of cancers expressing ERα are hormone resistant. For more than 20 years, a biochemical technique evaluating the binding capacity of oestradiol to its receptor in tumour cytosolic extracts was used to restrict tamoxifen prescription to the sole patients whose tumours expressed the receptor. This assay has been replaced by an immunohistochemical technique, more rapid and less expensive, but certainly much less informative. ERβ does not seem to be frequently expressed in breast cancers; anyway, it is not systematically evaluated like ERα. Its expression should be theoretically associated to an absence of hormonal dependency of cancer cells. Progesterone receptor PR (NR3C3, gene *PGR*) is also routinely evaluated in breast cancers. PR expression is highly correlated to that of ERα, but the two receptors are sometimes dissociated; it appears that the simultaneous presence of both receptors would be required for tamoxifen activity, to which ER^+–PR^- tumours are less often sensitive than ER^+–PR^+ tumours.

Even if the presence of hormone receptors in the tumour is an oncogenesis-promoting phenomenon, it appears difficult to consider this presence as an 'oncogenic alteration' since it occurs in the normal gland. Rare mutations of ERα accompany tamoxifen resistance. A variant form of ERα, in which the initial sequences of transcription activation in absence of ligand are deleted, would stimulate cell proliferation via activation of the MAP kinases and PI3 kinase pathways. In addition to receptor alterations, some alterations in the coreceptors of ERα have been identified: this is the case for the coactivator NCOA3, which was called AIB1 (*amplified in breast cancer*). A decrease in the expression of corepressors could, conversely, accompany tamoxifen resistance.

Loss of expression of ERα in breast cancers reveals a poor prognosis, in addition to the loss of hormone sensitivity. Several mechanisms have been suggested to explain this absence of expression, among which are promoter methylation, change in the level of activity of the transcription factors of the *ESR1* gene, mRNA degradation and proteasomal receptor destruction. The inhibition of the proliferation pathways downstream tyrosine kinase receptors such as ERBB2 (Chap. 1) seems able to restore ERα expression.

14.2.2 Androgen Receptors

Similarly to breast cancers, prostate cancers are stimulated by androgens because of the high expression of the androgen receptor (AR, NR3C4) in prostate tissue. These cancers are, therefore, also sensitive to hormonal treatments: castration, which eliminates androgen secretions; analogues and antagonists of the hypothalamic hormone LH-RH (gonadoliberin, gene *GNRH1*), which lead to an exhaustion of gonadic androgen secretions (leuprorelin, buserelin and others); anti-androgen compounds, which compete with androgens on the receptor-binding sites (cyproterone, flutamide, more recently enzalutamide); inhibitors of CYP17A1, the enzyme in charge of the 17α-hydroxylation of progesterone and pregnenolone, which leads to androgens (abiraterone). However, the antitumour effect of these treatments progressively decreases, and the tumour becomes hormone resistant when mutant forms of the receptor appear, which replace androgen stimulation by other stimuli.

A particular form of breast cancers, called apocrine tumours, is developed from special breast glands which expresses the androgen receptor and whose proliferation can also be inhibited by anti-androgens. The molecular identification of these tumours may enable such treatments in breast cancers.

14.2.3 Glucocorticoid Receptors

Glucocorticoids, via their interaction with GR (NR3C1), present essentially anti-inflammatory and immunosuppressive activities that are widely used in therapeutics. They also have the capacity of inducing cancer cell apoptosis in malignant haematological diseases and belong for a long time to the armamentarium of leukaemias, lymphomas and myelomas. Activated GR binding to a GRE induces the activation of the transcription of target genes involved in immunity and inflammation. Furthermore, activated GR can interact with transcription factors such as AP1 (Chap. 2) and NFκB (Chap. 12), forming transcriptional repression complexes, which lead to the inhibition of the pathways activated by these transcription factors, essentially cell survival and proliferation. More than 50 genes involved in the negative regulation of apoptosis (see Chap. 18) can thus be involved in the glucocorticoid-induced massive apoptosis of leukaemia or lymphoma cells. The loss of some of these pathways explains glucocorticoid resistance. GR mutations have been described in leukaemia cell lines, but the proof of their involvement in clinical resistance seems to be still lacking.

14.3 Thyroid Hormones Receptors

Thyroid hormones, thyroxin (T4) and triiodothyronine (T3), are ligands of TRα (NR1A1, gene *THRA*) and TRβ (NR1A2, gene *THRB*), these receptors being able to exist under two isoforms owing to alternative splicing. The TRα2 isoform is a false receptor which does not bind ligands and behaves as an inhibitor of hormonal action.

These receptors are heterodimerised with RXR; in the absence of ligand, they are bound to DNA at the level of their responsive element, TRE, and repress transcription thanks to the corepressor NCOR2 or SMRT (*silencing mediator of retinoic and thyroid receptor*). After ligand binding, TRs recruit coactivators which replace corepressors, and target gene transcription is possible after chromatin remodelling. There are also 'negative' TREs, which repress target gene transcription once the ligand is bound. Finally, thyroid hormones exert effects that are not mediated by TRE, especially via membrane interactions with other signalling systems.

TRs are the homologs of the viral oncogene *erb-A*, which is co-responsible, with the viral oncogene *erb-B*, of avian erythroblastosis and is able to block the differentiation of erythroid progenitors. Genetic and epigenetic alterations of TRs can be found in human cancers, leukaemia and carcinomas: point mutations, gene rearrangements, promoter methylation and loss of heterozygosity. Hepatocarcinomas present especially mutant forms of *THRA* or *THRB*, which behave as negative dominants. Some authors have described an inhibitory effect of TRs upon RAS-driven signalling. Taken together, these observations suggest that TRs play a tumour suppressor role, despite the fact that T3 has a mitogenic effect on the liver. The pharmacological use of these observations seems to be difficult because of the anticipated interactions with the potent hormonal effects of thyroid hormones.

14.4 Vitamin D Receptors

The active metabolite of vitamin D is a true hormone, which is formed, after photoactivation of sterol precursors brought by diet, by successive hydroxylations on carbons #1 and 25 to obtain 1α,25-dihydroxycholecalciferol or calcitriol; this compound exerts numerous physiological effects, especially on intestinal absorption of calcium and bone mineralisation. Calcitriol binds a nuclear receptor, VDR or NR1I1, which must heterodimerise with RXR to recognise target sequences on DNA. In the absence of ligand, the dimer is bound to VDRE (*vitamin D-responsive element*) and the transcription of target genes is inhibited, thanks to its association to the corepressor NCOR2 or SMRT. In this situation, histones are deacetylated and chromatin compaction does not allow gene transcription (Annex B.1). Ligand binding induces the replacement of the corepressor by a coactivator, the activation of histone acetyltransferases and the transcription of target genes, those having a VDRE in their promoter.

Several observations suggest the involvement of VDR in cell proliferation:

- Among the target genes of calcitriol are several genes involved in cell cycle arrest: $p21^{CIP1}$ (*CDKN1A*) and GADD45 (*growth arrest and DNA damage inducible transcript*).
- Experimentally, calcitriol has antiproliferative effects in vitro, which has been also observed in in vivo studies.
- *VDR* gene knockout in mice is accompanied by an increase of sensitivity to carcinogenic agents.

- In human tumours, an overexpression of the enzymes of vitamin D catabolism (CYP24A1) is observed, together with a decrease in the expression of biosynthesis enzymes (CYP27A1, CYP27B1).
- Although *VDR* mutations or rearrangements have not been observed in human tumours, an overexpression of *NCOR2* has been observed in breast, colon and prostate cancers.
- From an epidemiologic point of view, there is an association between vitamin D deficiency or low sunlight level and the risk of cancer.
- Several *VDR* gene polymorphisms are associated with a decrease or a loss of receptor activity and an increase in cancer risk.

The pathway initiated by vitamin D appears thus as playing an anti-oncogenic role, and associations of vitamin D and histone deacetylase inhibitors (HDAC, Annex B) have been proposed for the treatment of squamous cell carcinomas or myelodysplastic syndromes. This is a difficult approach, because pharmacological doses of vitamin D have deleterious effects owing to induced hypercalcaemia. However, since these effects seem to be mediated by membrane actions of vitamin D, independently of receptor binding, it may be possible to research vitamin D analogues able to bind with its nuclear receptor without displaying the toxic effects associated with the increase in intestinal calcium absorption. Nonsteroid modulators of calcitriol–VDR interaction are also under study, in combination with cytotoxic drugs.

14.5 Retinoic Acid Receptors

All-trans retinoic acid (ATRA) and, although at a lesser level, 9-cis-retinoic acid (9-cis-RA) are the principal active forms of vitamin A. They play a major role in growth and differentiation of many tissues. There are two types of retinoic acid receptors: those of the RAR type (RARα, RARβ and RARγ or NR1B1, NR1B2 and NR1B3; genes *RARA*, *RARB* and *RARG*) and those of the RXR type (RXRα, RXRβ and RXRγ or NR2B1, NR2B2 and NR2B3; genes *RXRA*, *RXRB* and *RXRG*). The first group represents high-affinity receptors for ATRA and 9-cis-RA, whereas the second group comprises ‘adopted’ receptors, able to bind only 9-cis-RA and with a lower affinity. These receptors are dimerised either as heterodimers RAR–RXR or as homodimers RXR–RXR, the first ones on DNA sequences called RARE (*retinoic acid responsive elements*) and the other ones on DNA sequences called RXRE (*retinoic X responsive elements*). In the absence of ligand, receptors repress gene transcription thanks to the intervention of NCORs that maintain the histones in deacetylated form, which prevents the accessibility of transcription initiation sites. When the ligand is bound, corepressors are replaced by coactivators, which recruit the histone acetyltransferases able to remodel chromatin and initiate transcription.

Retinoic acid exerts on normal and tumour cells an antiproliferative effect characterised by cell cycle arrest in G1 and induction of differentiation. RARβ expression, at the level of both mRNA and protein, is decreased in many cancer types.

Restoration of this expression enables tumour cell sensitisation to ATRA and tumour growth inhibition. This is the same for RARα and RARγ in some cancers. With the exception of acute promyelocytic leukaemia (next paragraph), the mechanism inducing RAR expression downregulation is not completely understood. RARβ promoter methylation and loss of heterozygosity at the level of the 3p24 locus have been suggested as mechanisms without definitive evidence of their involvement.

The role of the RARα receptor in acute promyelocytic leukaemia (APL) is in contrast well established. It was deciphered after the discovery of the therapeutic efficiency of ATRA in this disease, which represents one of the first successes of therapies targeted to a definite oncogenic mechanism. APL pathogenesis results from *RARA* gene translocation to diverse genomic sequences, leading to the synthesis of fusion proteins X-RAR or RAR-X. The most frequent fusions occur with the gene *PML* (*promyelocytic leukaemia*). In all cases, the fusion proteins exert a negative dominant effect on the normal function of the RAR, preventing the release of the corepressors, and consequently the transcription of retinoic acid target genes. The administration of ATRA at pharmacological doses allows to overflow this block and to normalise this signalling pathway. This results in the differentiation of promyelocytic cells that were blocked in their normal programming. In contrast, little is known concerning the precise target genes whose expression is required for normal cell growth. Disease progression occurs when cells become resistant to ATRA; resistance is associated to the occurrence of additional mutations of the fusion protein at the level of its ATRA-binding sequences.

14.6 Peroxisome Proliferator-Activated Receptors

Peroxisomes are intracytoplasmic organelles in charge of the detoxification of molecular oxygen and reactive oxygen species (ROS) (Chap. 16). They are also involved in fatty acids β-oxidation, cholesterol biosynthesis and degradation and glycerolipid synthesis. A series of endogenous and exogenous molecules exert an important effect on peroxisome proliferation; these are essentially lipids: polyunsaturated fatty acids (PUFA), prostaglandins and leukotrienes, which derive from PUFA, and some drugs, mainly the fibrates, which inhibit cholesterol biosynthesis. The PPARα, PPARβ/δ and PPARγ receptors (NR1C1, NR1C2 and NR1C3; genes *PPARA*, *PPARB* and *PPARG*), which were 'orphan' receptors, are activated by these molecules that they can bind with relatively low affinity and are consequently now 'adopted' receptors.

PPARs are heterodimerised with RXR for DNA binding. As the other nuclear receptors, they exert, in the absence of ligand, a repression of transcription at the level of PPREs (*peroxisome proliferator responsive elements*) through adequate NCORs, which are removed after ligand binding, thanks to the intervention of NCOAs. As the last resort, the level of chromatin acetylation regulates the transcriptional activity of these receptors. PPRE are found in the promoters of genes involved in fatty acid transport and lipid metabolism, as well as genes involved in cell proliferation, such as PDK1, *phosphoinositide-dependent kinase* (Chap. 3), and adhesion,

such as ILK, *integrin-linked kinase* (Chap. 11). The tissue distribution, ligand structure and physiological roles of the various PPARs are distinct but will not be discussed here.

PPARγ has been particularly studied in relation with cancer; its involvement in oncogenesis is still controversial: it is anti-oncogenic in most models but pro-oncogenic in some cases. Thiazolidinediones, which are PPARγ agonists and are used in the treatment of type II diabetes, have a preventive effect on tumour occurrence in mice. PPARγ mutations have been observed in colorectal cancers, as well as translocations leading to fusion proteins with negative dominant effect in thyroid cancers and homozygous mutations in prostate cancers. In contrast, thiazolidinediones may stimulate tumour evolution of APC-mutated colic adenomas (Chap. 7), which must lead to coloscopic surveillance of diabetic patients treated with this drug. Indirectly, cyclooxygenase 2 (COX2), which produces potent PPARγ agonists, is involved in the pro- or anti-oncogenic effects of these nuclear receptors. The combination of COX2 inhibitors and PPARγ agonists could exert a synergistic effect against tumour growth.

The oncogenic role of PPARβ/δ is even less clear than that of PPARγ; its activation may have pro-oncogenic consequences (resistance to apoptosis, increase of migratory capacities) and anti-oncogenic consequences (decrease in cell proliferation), according to the tissue or cell context. PPARα rather appears pro-oncogenic; its pharmacological agonists are fibrates, which induce in vitro and in vivo stimulation of hepatocyte proliferation and inhibition of apoptosis. These drugs have a long-term liver carcinogenic effect in rodents, but this has not been observed in humans.

14.7 Xenobiotics Receptors

The xenobiotics receptors, PXR (*pregnane X receptor*, gene *NR1I2*) and CAR (*constitutive androstane receptor*, gene *NR1I3*) were considered initially as orphan receptors since they do not bind any endogenous compound with high affinity. They recognise in fact a large variety of ligands and play a major role against the harmful effects of the accumulation of toxic exogenous and endogenous compounds. They share some agonistic ligands but not some others; and they also share the same DNA-binding sequences. They are mainly expressed in the liver, kidney and intestine. PXR and CAR target the genes encoding drug metabolising enzymes, in particular the cytochromes P450, and drug transport proteins, such as P-glycoprotein and other ABC transporters. They play a major role in the detoxification of xenobiotics and of toxic endogenous products (bilirubin, biliary acids). In oncology, these ‘xenoreceptors’ can modulate cancer risk as well as the activity of anticancer drugs.

As the other receptors of the NR1 family, they heterodimerise with an RXR receptor, bind in the absence of ligand to XREs (*xenobiotic-responsive elements*) and repress the transcription of target genes; in the presence of a ligand, this repression is alleviated and transcription can occur. Among the target genes that have been well characterised are cytochrome P450 3A4 (CYP3A4), which ensures the

oxidative metabolism of a large variety of drugs; UDP-glucuronosyltransferase 1A1 (UGT1A1), which conjugates drugs and endogenous compounds to glucuronic acid; and P-glycoprotein (Pgp, MDR1 for *multidrug resistance 1*, gene *ABCB1*), a transporter involved in the elimination of many xenobiotics out of the organism. Xenoreceptor agonists have not always been identified precisely; they had been identified initially as inducers of cytochromes and ABC transporters: rifampicin, phenobarbital, hyperforin, paclitaxel, ritonavir, etc. These compounds behave as potent inhibitors of the activity of oral drugs because of this induction phenomenon.

PXR and CAR expression is regulated by promoter methylation, which is higher in tumour cells than in normal cells, at least for *NR1I2*; by micro-RNAs at the level of the 3' UTR of the genes; and by the negative dominant effect of a truncated variant of the protein, obtained by alternative splicing.

AHR (*aryl hydrocarbon receptor*) is a xenoreceptor not belonging to the nuclear receptor family, but playing a similar role as PXR and CAR in xenobiotics metabolism. It is bound in the cytoplasm to chaperone proteins such as HSP90; when it binds to a ligand, it migrates to the nucleus, leaves its chaperone and binds to an intermediate protein called ARNT (*aryl hydrocarbon nuclear translocator*). Once in the nucleus, it binds DNA sequences called DRE (*dioxin response elements*) and activates the transcription of various genes, among which are *CYP1A1*, *CYP1A2* and *CYP1B1* and *UGT1A1*, *GSTP1*, *ALDH3A1* and *NQO1*. AHRR is a negative regulator of AHR expression.

Further Reading

Ali S, Coombes RC. Endocrine-responsive breast cancer and strategies for combating resistance. Nat Rev Cancer. 2002;2:101–12.

Altucci L, Leibowitz MD, Ogilvie KM, de Lera AR, Gronemeyer H. RAR and RXR modulation in cancer and metabolic disease. Nat Rev Drug Discov. 2007;6:793–810.

Aranda A, Pascual A. Nuclear hormone receptors and gene expression. Physiol Rev. 2001;81:1269–304.

Aranda A, Martínez-Iglesias O, Ruiz-Llorente L, García-Carpizo V, Zambrano A. Thyroid receptor: roles in cancer. Trends Endocrinol Metab. 2009;20:318–24.

Chen Y, Tang Y, Guo C, et al. Nuclear receptors in the multidrug resistance through the regulation of drug-metabolizing enzymes and drug transporters. Biochem Pharmacol. 2012;83:1112–26.

Collingwood TN, Urnov FD, Wolffe AP. Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription. J Mol Endocrinol. 1999;23:255–75.

Conzen SD. Nuclear receptors and breast cancer. Mol Endocrinol. 2008;22:2215–28.

de Cremoux P. Hormonothérapie des cancers du sein. Bull Cancer. 2011;98:1311–9.

Deeb KK, Trump DL, Johnson CS. Vitamin D signalling pathways in cancer: potential for anticancer therapeutics. Nat Rev Cancer. 2007;7:684–700.

Francis GA, Fayard E, Picard F, Auwerx J. Nuclear receptors and the control of metabolism. Annu Rev Physiol. 2003;65:261–311.

Frankfurt O, Rosen ST. Mechanisms of glucocorticoid-induced apoptosis in hematologic malignancies: updates. Curr Opin Oncol. 2004;16:553–63.

González-Sancho JM, García V, Bonilla F, Muñoz A. Thyroid hormone receptors/THR genes in human cancer. Cancer Lett. 2003;192:121–32.

Gronemeyer H, Gustafsson JA, Laudet V. Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov. 2004;3:950–64.
Handschin C, Meyer UA. Induction of drug metabolism: the role of nuclear receptors. Pharmacol Rev. 2003;55:649–73.
Hu X, Lazar MA. Transcriptional repression by nuclear hormone receptors. Trends Endocrinol Metab. 2000;11:6–10.
Lee JS, Kim KI, Baek SH. Nuclear receptors and coregulators in inflammation and cancer. Cancer Lett. 2008;267:189–96.
Massard C, Fizazi K. Targeting continued androgen receptor signaling in prostate cancer. Clin Cancer Res. 2011;17:3876–83.
Michalik L, Desvergne B, Wahli W. Peroxisome-proliferator-activated receptors and cancers: complex stories. Nat Rev Cancer. 2004;4:61–70.
O'Malley BW, Kumar R. Nuclear receptor coregulators in cancer biology. Cancer Res. 2009;69:8217–22.
Pascussi JM, Gerbal-Chaloin S, Duret C, et al. The tangle of nuclear receptors that controls xenobiotic metabolism and transport: crosstalk and consequences. Annu Rev Pharmacol Toxicol. 2008;48:1–32.
Privalsky ML. The role of corepressors in transcriptional regulation by nuclear hormone receptors. Annu Rev Physiol. 2004;66:315–60.
Rosenfeld MG, Glass CK. Coregulator codes of transcriptional regulation by nuclear receptors. J Biol Chem. 2001;276:36865–8.
Ryan CJ, Tindall DJ. Androgen receptor rediscovered: the new biology and targeting the androgen receptor therapeutically. J Clin Oncol. 2011;29:3651–8.
Soprano DR, Qin P, Soprano KJ. Retinoic acid receptors and cancers. Annu Rev Nutr. 2004;24:201–21.
Tenbaum S, Baniahmad A. Nuclear receptors: structure, function and involvement in disease. Int J Biochem Cell Biol. 1997;29:1325–41.
Thorne J, Campbell MJ. The vitamin D receptor in cancer. Proc Nutr Soc. 2008;67:115–27.
Wang T, Xu J, Yu X, Yang R, Han ZC. Peroxisome proliferator-activated receptor gamma in malignant diseases. Crit Rev Oncol Hematol. 2006;58:1–14.

Ion Channel-Coupled Receptors

15

Abstract

Conduction of nerve impulse is a major signalling pathway in multicellular organisms and requires ion channels situated at the level of plasma membranes of excitable cells. Opening and closing of these channels are controlled by the level of membrane polarisation, and this explains their generic name of *voltage-operated channels* (VOC). At the level of synapses, nerve impulse transduction from a neuron to another one or to a muscle cell is operated through the release, in the synaptic cleft, of signals that activate the opening of other ion channels. These are called *receptor-operated channels* (ROC), ionotropic receptors or *ligand-gated ion channels* (LGIC). Their ligands are neurotransmitters: acetylcholine, serotonin, amino acids and purine nucleotides (ATP). They can in addition react with intracellular proteins (small G-proteins, cytoplasmic tyrosine kinases) and with cytoskeletal proteins. Outside the nervous tissues, they can behave as signalling molecules in many cell types.

Receptor-activated ion channels are also present within the cells and enable especially the transfer of Ca^{2+} ions from one compartment to another one. Calcium signalling has been mentioned in Chap. 6: the second messenger resulting from GPCR (*G-protein-coupled receptors*) activation leads to Ca^{2+} release in the cytosol, which enables this 'third messenger' to activate numerous effectors, which play a major role in muscle contraction and secretion processes.

We will not detail here all the aspects of the signalling operated by ionic fluxes, which pertain to neurosciences rather than to oncology; after a simplified presentation of LGIC in excitable cells, we will focus on two special types of channel receptors: purinergic receptors and receptors enabling intracellular Ca^{2+} mobilisation.

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_15

15.1 Activation of Ligand-Gated Ion Channels

Ionotropic receptors or LGICs are ion channels allowing the transfer through the plasma membrane of excitable cells of anions (Cl^-) or cations (Na^+, K^+, Ca^{2+}). They are distinct from voltage-gated ion channels acting upstream or downstream their activation. The main LGICs are the acetylcholine nicotinic receptor, the serotonin receptor, certain receptors of γ-aminobutyric acid (GABA), glutamic acid and glycine, and ATP receptors called purinergic or P2X receptors. For most of these ligands, other receptors exist, which belong to the GPCR family; they are called *metabotropic* receptors to distinguish them from ionotropic receptors: the acetylcholine muscarinic receptor, amino acid receptors and a second kind of purinergic receptors, called P2Y, which also recognise adenylyl nucleotides, distinct from the adenosine receptors P1 or ADORs.

LGICs have a similar general structure, comprising transmembrane proteins that are assembled together to constitute a pore through which ions (anions or cations according to the LGIC) are transferred. There exist about 70 protein subunits constituting precise functional units, localised especially at the level of synapses. Some have a pentameric structure (acetylcholine, serotonin, GABA, glycine receptors), others a tetrameric structure (glutamate receptors) or a trimeric structure (purinergic receptors). The diversity of subunit assembly in each type of receptor (Fig. 15.1) enables the specificity of ligand recognition and binding and determines the ion to be transferred, the conductance of the channel, the speed of opening and closing of the channel, etc.

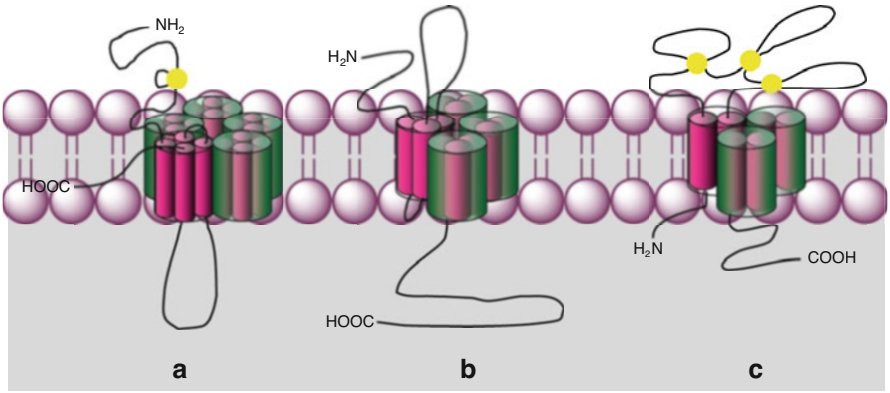


Fig. 15.1 General organisation of ion channel-coupled receptors. Plasma membrane ligand-gated ion channels (LGIC) are of several types. The *N*-terminal and *C*-terminal domains can be extracellular or cytoplasmic. Disulphide bridges are indicated in yellow. (**a**) Pentameric receptors, each monomer being constituted of four transmembrane helices: acetylcholine nicotinic receptor, serotonin, GABA and glycine receptors; (**b**) tetrameric receptors, each monomer being constituted of three transmembrane helices: glutamate receptor; (**c**) trimeric receptors, each monomer being constituted of two transmembrane helices: purinergic receptors, P2X type

Acetylcholine and serotonin receptors are Na^+, K^+ and Ca^{2+} channels; they induce membrane depolarisation and stimulate the opening of the voltage-dependent ion channels that convey nerve impulses: they are present at the level of *excitatory* synapses. Glycine and GABA receptors, in contrast, are Cl^- channels, which induce membrane hyperpolarisation when open; they have a negative effect on nerve impulse conduction and are therefore present at the level of *inhibitory* synapses. Finally, the ionotropic glutamate receptors are calcium channels involved in special modalities of synaptic transmission that will not be described here.

The nicotinic receptor of acetylcholine has been one of the first receptors identified, both at the physiological and at the molecular levels; it has clearly left its mark on the history of cell signalling and synaptic transmission at the level of the neuromuscular synapse. It comprises five assembled subunits: two of α type and one of each other type (β, γ and δ). There are nine different genes encoding the α subunits, four the β subunits and one the γ and δ subunits. Each subunit is a transmembrane protein with four membrane-spanning helices (M1 to M4) and extracellular *N*-terminal and *C*-terminal domains. The M2 helices interact together to delineate the pore used for cation flux, and the *N*-terminal domains also interact to delineate the binding site of acetylcholine. The ionotropic receptors of serotonin, glycine and GABA are built according to the same general scheme.

15.2 Purinergic Receptors

There are seven purinergic P2X receptors, P2X1 to P2X7, which are formed of three subunits, identical or different, each made of two transmembrane domains separated by a long extracellular loop containing always ten cysteine residues bound by disulphide bonds, and intracytoplasmic *N*-terminal and *C*-terminal domains. Ligand-driven receptor activation induces the entry of Na^+ or Ca^{2+} ions inside the cells and the efflux of K^+ ions, as a function of their concentration gradient and without marked selectivity. This results in membrane depolarisation, which is able to activate voltage-dependent channels and to generate an action potential. Purinergic receptors are localised in excitable cells (neurons, nonstriated muscle cells) and also in epithelial cells (lung, intestine, etc.) and endothelial cells.

It had been observed long ago that ATP could display antiproliferative activity. One can hypothesise that this effect is mediated by purinergic receptor (GPCR or LGIC) activation but possibly also by the synthesis of adenosine by membrane ecto-ATPases, since adenosine can bind distinct GPCRs called ADORs, or even by a nutritional effect of ATP. P2X and P2Y receptors are expressed in many cancer types, especially P2X5 and P2X7, which are the most frequently expressed. ATP treatment of cultured cells that express these receptors induces apoptosis, as detected by caspase 3 activation (Chap. 18), although the mechanisms underlying this involvement have not been described. In contrast, metabotropic (P2Y) receptors have a positive effect on cell signalling, which are mediated, as always for GPCR, by the second messengers, cyclic AMP, inositol trisphosphate IP3 and Ca^{2+} (Chap. 6).

Clinical trials have explored the feasibility of ATP perfusions, but it was not possible to conclude, beyond that step, whether these perfusions could be efficient in oncology. However, therapeutic tools more stable than ATP itself might reveal some interest.

15.3 Ca^{2+} Signalling

The particular importance of Ca^{2+} in signal transduction arises from the considerable difference existing between its extracellular concentration (about 1.3 mM) and its cytosolic concentration (around 100 nM). The existence of important Ca^{2+} stores in cytoplasmic organelles such as endoplasmic reticulum, lysosomes and mitochondria enables the generation of intracellular signals enabling the implementation of rapid and transient reactions. Four successive steps are involved in Ca^{2+} signalling:

- Various stimuli generate signals for Ca^{2+} mobilisation from extra- or intracellular sources.
- These signals activate ion channels that transfer a critical amount of Ca^{2+} in the cytosol.
- Ca^{2+} activates a series of cytoplasmic effectors which are sensitive to its concentration.
- Various active transporters pump out Ca^{2+} out of the cytosol to restore the initial state.

This general scheme operates in multiple ways within cells as a function of the equipment in adequate proteins at each of these four steps.

15.3.1 Ca^{2+} Mobilisation Signals

Several types of signals can induce Ca^{2+} mobilisation from the extracellular space or from the intracellular stores (Fig. 15.2). These signals induce the opening of channels enabling a 5–10-fold increase in Ca^{2+}cytosolic concentration (from 100 nM to 500–1,000 nM).

At the intracellular levels, these signals are inositol trisphosphate (IP3, Chap. 6), which acts on specific receptors (IP3R or ITPR) and on cyclic ADP-ribose (cADPR, Fig. 15.3), the physiological activator of ryanodine receptor (RYR). Other signals, such as sphingosine 1-phosphate (S1P) and nicotinic acid adenine dinucleotide phosphate (NAADP), are also active on these poorly characterised receptors. In all cases, Ca^{2+} itself induces its own release in cytosol thanks to a positive retrocontrol called CICR (*Ca^{2+}-induced Ca^{2+} release*). The true action of IP3 and cADPR is to increase receptor sensitivity to Ca^{2+} effects. The same cells can harbour different types of receptors, so that they can respond similarly to different stimuli.

IP3 generation is essentially obtained through the activation of G-protein-coupled receptors (GPCR) by various ligands (Chap. 6). Briefly, this second

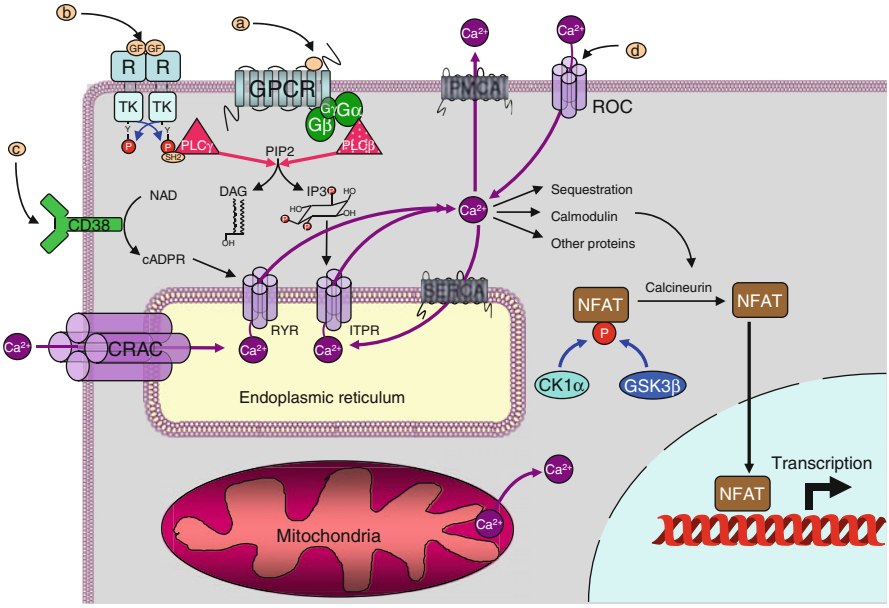


Fig. 15.2 Ca^{2+} signalling. Cytosolic Ca^{2+} (100 nM) can come from the extracellular space (1.3 mM) via receptor-activated channels (called ROC or LGIC) or from the endoplasmic reticulum via channels activated by IP3 (ITPR) or cAPDR (RYR). It is pumped back via ATPase transporters of the plasma membrane (PMCA) or the endoplasmic reticulum (SERCA). IP3 is produced especially after activation of a GPCR or a TKR, through the cleavage of phosphatidylinositol 4,5-bisphosphate by PLCβ or PLCγ. cAPDR is produced from NAD by membrane ecto-enzymes, CD38 and CD157. The Ca^{2+} stores of the endoplasmic reticulum can be reconstituted, when depleted, with extracellular Ca^{2+}, by a direct channel, CRAC, working in connection with an endoplasmic reticulum Ca^{2+} sensor, STIM. Thus, the signals that can modulate the cytosolic Ca^{2+} concentration are GPCR ligands (**a**), TKR ligands (**b**), CD38 ligands (**c**) and LGIC ligands (**d**), CRAC activation being realised by the depletion of endoplasmic reticulum stores. Among its multiple actions, Ca^{2+} can activate, via calmodulin, a phosphatase called calcineurin, which activated the NFAT transcription factors

messenger is produced from phosphatidylinositol 4,5-bisphosphate by phospholipase C beta (PLCβ) after activation by a large heterotrimeric G-protein. IP3 can also be produced by PLCγ after activation by tyrosine kinase receptors (Chap. 1), by PLCε after activation by small G-proteins of the RAS and RHO families or by PLCδ after activation by variations in the cytosolic Ca^{2+} concentration.

cADPR and NAADP are formed from NAD by membrane receptor enzymes, CD38 and CD157 (gene *BST1*), characterised by ADP-ribosyl cyclase and cADP-ribose hydrolase activities. CD38 is present in lymphoid tissues, especially chronic lymphocytic leukaemia cells. The messengers that activate CD38 and CD157 are not precisely known. CD38 activation can occur after interaction with membrane receptors expressed by lymphoid cells, especially B-cell and T-cell receptors (Chap. 13). Targeting CD38 could be of interest in chronic lymphocytic leukaemia.

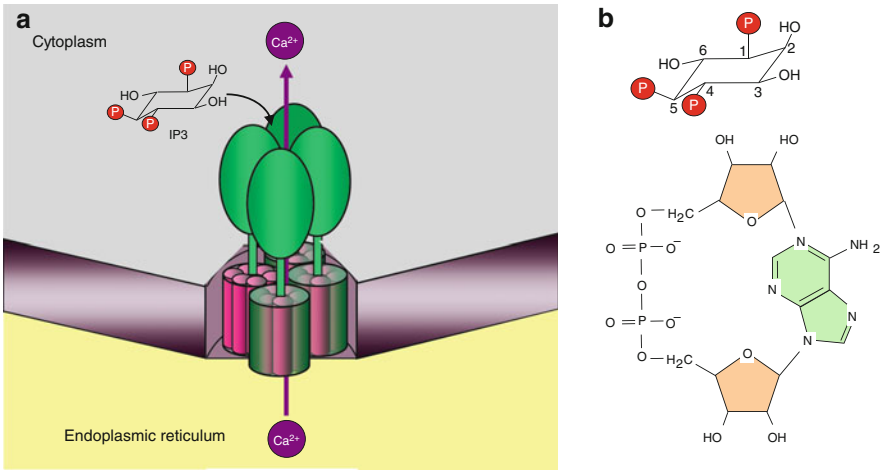


Fig. 15.3 Intracellular Ca^{2+} channels and their ligands. (**a**) General structure of the IP3 receptor (ITPR). This receptor is constituted of four subunits, each comprising six transmembrane helices, and delineates a pore allowing the transfer of Ca^{2+} from the endoplasmic reticulum to the cytoplasm. Large *N*-terminal cytoplasmic domains contain especially the IP3 recognition site. The ryanodine receptor is built in a similar way. (**b**) Structure of cyclic ADP-ribose (cADPR)

At the level of the plasma membrane, Ca^{2+} is mobilised through the activation of *voltage-operated channels* (VOCs); of *ionotropic receptor-operated channels* (ROCs), activated by GABA and other ligands (see Sect. 15.1); and of a third variety of channels of slow activation, SOCs (*store-operated channels*), which aim at restoring the Ca^{2+} stores of the endoplasmic reticulum (ER). VOCs are of five different types (L, N, P/Q, R, T) and are activated by membrane depolarisation, induced in excitable cells by nerve impulses. Their activity is modulated by phosphorylation for type L channels and by small G-proteins for the other channel types. ROCs of nerve cells are localised at the postsynaptic level and transduce inhibitory messages. SOCs are present in various cell types, especially T lymphocytes and endocrine glands (pancreas, hypophysis, etc.). The main SOC is a plasma membrane Ca^{2+} channel, CRAC (*calcium release-activated calcium modulator*, gene *ORAI1*), which is topologically and functionally coupled to an ER Ca^{2+} sensor, STIM (*stromal interaction molecule*, gene *STIM1*), so that Ca^{2+} can be transferred directly from the extracellular space to ER to restore the intracellular stores.

Mitochondria actively participate to the process of Ca^{2+} storage, when they are localised in the neighbourhood of the channel receptors ITPR and RYR, thanks to a special transporter, the *mitochondrial calcium uniporter* (MCU), associated with Ca^{2+} mitochondrial uptake (MICU) proteins to avoid mitochondrial Ca^{2+} overload.

15.3.2 Opening of Ca^{2+} Channels

Whereas ROCs (and SOCs in excitable cells) are responsible for Ca^{2+} entry in the cytosol from the extracellular space, Ca^{2+} cytosol entry from the ER is realised thanks to the opening of Ca^{2+}-selective membrane channels (Fig. 15.2), which are in fact ITPR and RYR themselves. ITPRs (Chap. 6) are actual intracellular ionotropic receptors (*ligand-gated ion channels*). There are three of them (ITPR1, 2 and 3), assembled as homo- or heterotetramers forming a pore of about 1.2 MDa. There is a large variety of possible combinations whose exact role remains to be deciphered. Each monomer contains six transmembrane helices and a large number of interaction domains with other proteins as well as various post-translational modification sites (Fig. 15.3). The *N*-terminal and *C*-terminal domains of each monomer are localised in the cytosol, and the IP3-binding domain is on the *N*-terminal side. Channel opening is mainly regulated by three ligands: Ca^{2+} itself, IP3 and ATP, which exerts an allosteric regulation of variable intensity according to the tetramer. ITPR phosphorylation by protein kinase C (Chap. 6), AKT (Chap. 3) and cytoplasmic tyrosine kinases constitutes a major regulatory mechanism.

Ryanodine receptors (RYRs) were first identified owing to the activity of this alkaloid as receptor agonist. Their physiological ligand, in addition to Ca^{2+} itself, is cADPR. There are also three of them (RYR1, 2 and 3) with tissue specificity, assembled as high molecular weight (2 MDa) homotetramers to enable Ca^{2+} transfer. Each monomer also contains six transmembrane helices forming the pore, and the *N*-terminal domain bears numerous regulation sites. As for ITPRs, this is in fact Ca^{2+} itself which is the main agonist of these receptors for CICR function.

The exit of Ca^{2+} from ER can take several aspects: sometimes, it consists of single emissions from the channel receptor, which are called *blips* for ITPR and *quarks* for RYR. The aggregation of receptor channels enables their concerted activation and the emission of *puffs* (ITPR) or *sparks* (RYR). Finally, the self-sustained entry of large amounts of Ca^{2+} in the cytosol enables the propagation of calcium waves, which rapidly and transiently extend to the whole cell and can reach neighbour cells through intercellular junctions. Ligand-mediated GPCR activation, which generates IP3, determines the type of Ca^{2+} response implemented by the effectors.

15.3.3 Ca^{2+}-Dependent Intracellular Activities

When Ca^{2+} enters the cytosol, it immediately binds to trapping proteins, such as parvalbumin, calsequestrin or calreticulin, which play a buffering role; a small fraction remains free and is available to activate specific proteins, which are sensitive to the variations of Ca^{2+} concentration and are called *Ca^{2+} sensors*. Two types of protein domains are able to bind Ca^{2+}: *EF-hand* motifs and *C2* motifs. The principal

Ca^{2+}-sensitive protein is calmodulin (CALM), the activation of which by Ca^{2+} induces a conformational change enabling downstream the control of several processes: smooth muscle contraction, interaction between signalling pathways, gene transcription and ion channel modulation.

Calmodulin contains four Ca^{2+}-binding sites in EF-hand motifs; once Ca^{2+} is bound, calmodulin can bind and activate a variety of proteins, among which:

- Kinases such as glycogen phosphorylase kinase and glycogen synthase kinase, which reinforce the action of PKA for glycogen mobilisation (Chap. 6)
- Enzymes such as adenylyl cyclase and phosphodiesterase (Chap. 6), which generate a brief pulse of cAMP
- Phosphatases such as calcineurins (CALN; genes *PPP3CA*, *PPP3CB* and *PPP3CC*; and *PPPCR1* and *PPPCR2*), which enable the activation of the transcription factor NFAT (see paragraph Sect. 15.3.5)
- Nitric oxide synthase (NOS), which is responsible for $NO^{\bullet}$ formation and, beyond, of cGMP formation, with many vascular effects (Chap. 16)
- Calcium/calmodulin-dependent protein kinases (CAMK), a series of serine/threonine kinases involved in multiple physiological processes, especially apoptosis (Chap. 18)

Proteins other than calmodulin are able to directly respond to Ca^{2+} concentration changes, as they harbour EF-hand or C2 domains:

- Protein kinase C beta (PKCβ), involved in the production of NFκB and interfering with apoptosis (Chaps. 12 and 18)
- Calpains (CAPN), cysteine proteases especially involved in IL1α activation (Chap. 12)
- Synaptotagmins (SYT) and other proteins involved, at the presynaptic level, in neurotransmitter release from storage vesicles
- Diacylglycerol kinases (DAGK), which phosphorylate DAG, the second messenger synthesised from phosphatidylinositol 4,5-bisphosphate by PLCβ (Chap. 6) or PLCγ (Chap. 2); these enzymes enable the resynthesis of inositol phospholipids
- Cytoskeleton proteins, such as α-actinins (ACTN) and gelsolin (GSN)

As a brief statement, cytosolic Ca^{2+} may exert effects on cell proliferation and differentiation, adhesion and motility, synaptic transmission, muscle contraction, carbohydrate metabolism, hormone secretion, apoptosis, chemotactism and various signalling pathways.

15.3.4 Restoration of the Initial State

Ca^{2+} is rapidly expelled out of the cytosol and sent back outside the cell or to intracellular storage sites, once the Ca^{2+} signal has been emitted. There are ATPase

pumps in the plasma membrane called Ca^{2+}-ATPases or PMCAs (*plasma membrane calcium ATPases*), encoded by genes of the ATP2B subfamily, with developmental and tissue-specific variants; these enzymes are able to efflux Ca^{2+} out of cells against an important concentration gradient (Fig. 15.2). In the ER, similar pumps are expressed, especially in muscle cells, called SERCA (*sarcoplasmic/endoplasmic reticulum calcium ATPases*), encoded by genes of the ATP2A subfamily. In addition, Na^{+}/Ca^{2+} exchangers called NCX (genes *SLC8A1* to *SLC8A3*) are in charge of Ca^{2+} pumping back to the endoplasmic (or sarcoplasmic) reticulum and/or to the extracellular medium. They have relatively low affinity for Ca^{2+} but high transport activity. Other Ca^{2+} exchangers exist in other organelles harbouring Ca^{2+} stores, such as the mitochondria, which express an NCLX protein (gene *SLC8B1*) with similar function of returning to the resting state as NCX in the ER.

15.3.5 NFAT, a Transcription Factor Activated by Ca^{2+} Entry

NFAT (*nuclear factor of activated T cells*) is a small family of five transcription factors first described in T lymphocytes (Chap. 13), but which were identified thereafter in many cell types. They are initially present in the cytoplasm at the phosphorylated state; calcineurin, upon activation by cytosolic Ca^{2+} release, coming especially from the extracellular space thanks to the channel receptor CRAC, removes this phosphate group. As a consequence, NFATs can migrate to the nucleus, where they can exert the transcription control of their target genes.

These five homologous transcription factors are structurally related to the REL factors of the NFκB pathway (Chap. 12); they contain a REL (*reticuloendotheliosis viral oncogene homologue*) domain and a transcription transactivation domain called NHD (*NFAT homology domain*). NFATs bind DNA as homodimers or heterodimers; they cooperate with other transcription factors such as AP1 (JUN–FOS, Chap. 2) and can form tetrameric assemblies.

NFATs are phosphorylated in the cytoplasm; there are several sites of phosphorylation, some of them being clustered in a serine-rich region (SRR), and others more spread. The cytoplasmic phosphorylation of NFATs is carried out by several serine/threonine kinases, protein kinases A (PKAs), casein kinase 1 (CK1) and glycogen synthase kinase 3 (GSK3β). As already mentioned, the opening of the CRAC plasma membrane Ca^{2+} channel is required for the activation of calcineurins (phosphatase activity), which dephosphorylate NFAT and induce its relocalisation in the nucleus, thanks to the unmasking of a nuclear localisation sequence. Once in the nucleus, nuclear kinases can rephosphorylate NFAT, thus inducing its way back to the cytoplasm.

The individual NFATs may play distinct or even opposite roles. NFAT1 acts against cell proliferation and induces the transcription of apoptosis genes, whereas NFAT2 has opposite effects on cell proliferation and survival. As a consequence,

NFAT1 appears as a tumour suppressor gene and *NFAT2* as a proto-oncogene. However, NFAT1 also induces the transcription of migration and invasion genes, which correspond to pro-metastatic and pro-angiogenic effects. A high level of NFAT2 activation is observed in numerous cancer types, especially, but not only, B- and T-cell lymphomas, with a permanent localisation of the factor in the nucleus. The mechanisms underlying hyperactivation in cancers are not presently known.

A potential target of NFAT is the action of calcineurin and other Ca^{2+}-activated proteins on compounds such as cyclosporine A and FK-506. These compounds have marked immunosuppressive activity and are used especially to treat transplant rejection. However, inducing immunosuppression does not seem desirable in cancer treatment, and more selective targets should be sought for. Peptides mimicking the interaction between calcineurin and NFAT have been identified, as well as small molecules with NFAT2-specific inhibition properties, based upon the structure of cyclosporine A or FK-506, and devoid of immunosuppressive effects.

Further Reading

Berridge MJ. Inositol trisphosphate and calcium signalling mechanisms. Biochim Biophys Acta. 2009;1793:933–40.

Berridge MJ. Calcium signalling remodelling and disease. Biochem Soc Trans. 2012;40: 297–309.

Berridge MJ, Lipp P, Bootman MD. The versatility and universality of calcium signalling. Nat Rev Mol Cell Biol. 2000;1:11–21.

Berridge MJ, Bootman MD, Roderick HL. Calcium signalling: dynamics, homeostasis and remodelling. Nat Rev Mol Cell Biol. 2003;4:517–29.

Clapham DE. Calcium signaling. Cell. 2007;131:1047–58.

Collingridge GL, Olsen RW, Peters J, Spedding M. A nomenclature for ligand-gated ion channels. Neuropharmacology. 2009;56:2–5.

Deli T, Csernoch L. Extracellular ATP and cancer: an overview with special reference to P2 purinergic receptors. Pathol Oncol Res. 2008;14:219–31.

Fliegert R, Gasser A, Guse AH. Regulation of calcium signalling by adenine-based second messengers. Biochem Soc Trans. 2007;35:109–14.

Foskett JK, White C, Cheung KH, Mak DO. Inositol trisphosphate receptor Ca^{2+} release channels. Physiol Rev. 2007;87:593–658.

Hamilton SL. Ryanodine receptors. Cell Calcium. 2005;38:253–60.

Keramidas A, Moorhouse AJ, Schofield PR, Barry PH. Ligand-gated ion channels: mechanisms underlying ion selectivity. Prog Biophys Mol Biol. 2004;86:161–204.

Lee HC. Physiological functions of cyclic ADP-ribose and NAADP as calcium messengers. Annu Rev Pharmacol Toxicol. 2001;41:317–45.

Malavasi F, Deaglio S, Funaro A, Ferrero E, Horenstein AL, Ortolan E, Vaisitti T, Aydin S. Evolution and function of the ADP ribosyl cyclase/CD38 gene family in physiology and pathology. Physiol Rev. 2008;88:841–86.

Mancini M, Toker A. NFAT proteins: emerging roles in cancer progression. Nat Rev Cancer. 2009;9:810–20.

Parekh AB. Store-operated CRAC channels: function in health and disease. Nat Rev Drug Discov. 2010;9:399–410.
Prakriya M. The molecular physiology of CRAC channels. Immunol Rev. 2009;231:88–98.
Prevarskaya N, Skryma R, Shuba Y. Calcium in tumour metastasis: new roles for known actors. Nat Rev Cancer. 2011;11:609–18.
Sheng M, Pak DT. Ligand-gated ion channel interactions with cytoskeletal and signaling proteins. Annu Rev Physiol. 2000;62:755–78.
White N, Burnstock G. P2 receptors and cancer. Trends Pharmacol Sci. 2006;27:211–7.

Signalling by Oxygen and Nitric Oxide 16

Abstract

Oxygen is, first of all, the combustive agent required for life and the provider of energy required by all heterotrophic organisms, those which cannot directly use solar energy as chlorophyllous plants do. Oxygen supply must be finely tuned, because insufficiency is as damaging as excess: hypoxia endangers cellular life through lack of energy supply, and reactive oxygen species (ROS) also endangers life because of their toxicity. In order to inform the organism of hypoxia or oxidative stress, signalling pathways are implemented and are studied in this chapter. In relation to oxidative stress, a special oxidised form of nitrogen, nitric oxide, is also a true intracellular messenger with multiple effects and will also be studied here.

The tumour cell is exquisitely sensitive to the effects of hypoxia and oxidative stress; it reacts to hypoxia by stimulating tumour vascularisation (angiogenesis), oxygen-independent energy supply and protein synthesis downregulation; it reacts to oxidative stress by activating signalling pathways still not completely deciphered. Reactive oxygen species appear as a double-edged sword, involved in carcinogenesis but perhaps useful to combat cancers.

16.1 Hypoxia

The main signalling pathway involved in hypoxia response depends on the activation of a transcription factor, HIF (*hypoxia-inducible factor*), which induces the expression of a large set of genes aimed at remedying this situation. The available oxygen amount is by itself the signal enabling HIF activation; oxygen thus appears as a signalling molecule, a messenger inducing the generation of appropriate responses.

16.1.1 Activation and Role of HIF

HIF is a heterodimeric complex consisting of an oxygen-sensitive α subunit and an oxygen-insensitive β subunit. There exist three distinct α and two β subunits. In the presence of adequate oxygen concentration, several amino acids of the α subunit

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_16

undergo post-translational hydroxylation, which drives HIFα to ubiquitinylation and proteasomal destruction (Annex C). A decrease in the availability of oxygen induces a decrease of this hydroxylation and a relocalisation of the α subunit to the nucleus; in the nucleus, the α subunit can be associated to the β subunit and the complex can stimulate the transcription of target genes (Fig. 16.1). There are many signalling pathways which activate HIF transcription, especially the proliferation pathways opened by MAP kinases and PI3 kinase (Chaps. 2 and 3). Several transcription factors, such as MYC or NFκB, thus activate the synthesis of HIF mRNAs.

The α subunits contain a domain called ODDD (*oxygen-dependent degradation domain*), with two proline residues that can be hydroxylated by prolyl 4-hydroxylases (PHD, genes *EGLN*) and a *C*-terminal domain containing an asparagine residue that can be hydroxylated in the nucleus by asparagine hydroxylase, an enzyme also known as FIH (*factor inhibiting HIF*, gene *HIF1AN*). These enzymes are non-haem dioxygenases, which use molecular oxygen, α-ketoglutarate and Fe^{2+} as a cofactor. Their activity is decreased when their substrate O_2 is not abundant enough (hypoxia). The hydroxylated proline residues are recognised by the VHL (*von Hippel–Lindau*) protein, which belongs to an E3 ubiquitin ligase complex called VBC (*VHL–elongin B–elongin C*), which enables HIF1α to be directed to the proteasome.

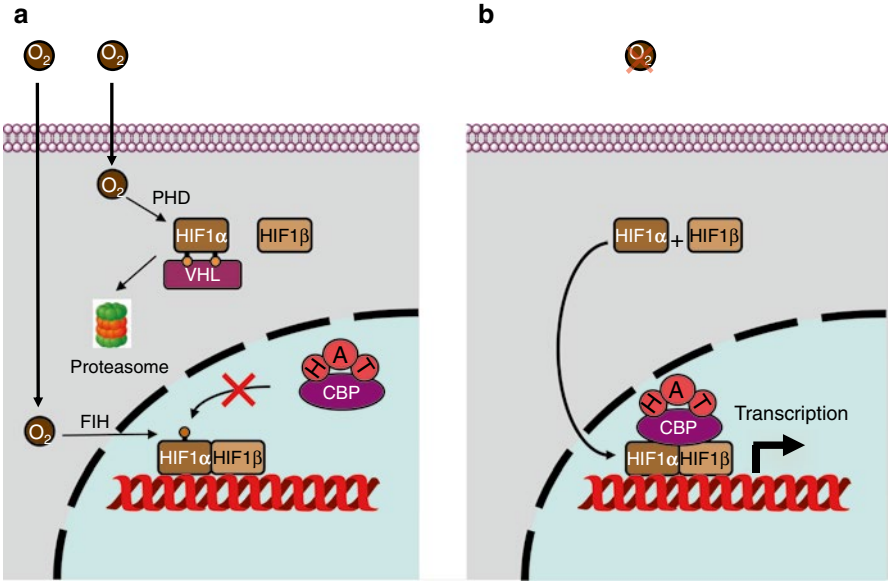


Fig. 16.1 Hypoxia signalling. (**a**) In the presence of oxygen, the transcription factor HIF1α is hydroxylated in the cytoplasm on two proline residues; this enables its recognition by the protein VHL, which is part of an E3 ubiquitin ligase complex, leading to proteasomal degradation. In addition, HIF1α is hydroxylated in the nucleus as an asparagine residue, which prevents its recognition by the transcription activator CBP (*CREB-binding protein*) that enables histone acetylation. (**b**) In the absence of oxygen, these hydroxylations cannot take place; HIF1α is stabilised and migrates to the nucleus, recognises its partner HIF1β and initiates transcription of target genes via the binding of the transcription activator CBP

In the nucleus, asparagine hydroxylation prevents HIFα binding to a coactivator of transcription, CBP (*CREB-binding protein*), which acts as a histone acetyltransferase for chromatin relaxation (Annex B). In the nucleus, the HIFα and HIFβ subunits are combined through their *N*-terminal domains and bind DNA, via a HLH (*helix–loop–helix*, Annex B) motif, on a hypoxia-responsive element (HRE) belonging to the promoter of the target genes. In the *C*-terminal domain lays the transactivation domain, TAD, divided, for HIF1α and HIF2α, into a hydroxylation-independent domain (*N*-TAD) and a domain inhibited by the FIH-mediated hydroxylation of an asparagine residue, *C*-TAD. The HIF3α subunit could rather play a negative dominant role on the transcription induced by the other HIFα subunits. The main target genes of HIF proteins are those encoding the VEGFs (*vascular endothelial growth factors*), especially *VEGFA* (Chap. 1). HIFs appear, therefore, as fundamental regulators of angiogenesis. Several dozens of other target genes are under the dependence of HIFs, involved especially in energy metabolism, erythropoiesis, vasodilatation and autophagy: for all these processes, oxygen availability is a crucial limiting factor.

16.1.2 Consequences of Hypoxia on Tumour Angiogenesis

Hypoxia constitutes the more important angiogenesis-promoting signal. Angiogenesis is indispensable to tumour growth beyond some millimetres in diameter, because only vessels can bring to the tumour the required nutriments and oxygen. In the absence of vascularisation, the deep tumour regions do not benefit of this supply and become rapidly necrotic, which limits tumour growth. Neo-angiogenesis is one of the major oncogenic processes and its targeting, which had been suggested 40 years ago as susceptible to bring original anticancer weapons, has eventually allowed the development of therapeutic approaches.

HIF overexpression is commonly observed in human cancers, sometimes in association with poor prognosis. Whereas there are no known activating mutations of the genes encoding the various HIF subunits, a loss-of-function mutation frequently occurs in the *VHL* gene, especially in renal cancers in which it appears recurrent. In the absence of this E3 ubiquitin ligase activity, the stability of the HIFα subunits is increased and their effects on tumour growth are increased. Antiangiogenic agents such as bevacizumab, an anti-VEGFA monoclonal antibody (Chap. 1), have brought proof of activity in this type of cancer as in several others.

16.1.3 Consequences of Hypoxia on Tumour Energy Metabolism

Among the target genes of HIFs, several genes are involved in glucose transport and anaerobic metabolism. In order to provide ATP, glucose can be metabolised according either to the mitochondrial pathway of oxidative phosphorylation, which requires oxygen, or to the cytosolic anaerobic glycolysis pathway. This pathway is less energetically profitable (two molecules of ATP are formed per molecule of

glucose instead of 38), but it can present an utmost interest for hypoxic cells. By favouring glucose uptake and commitment to glycolysis, HIFs enable hypoxic cell survival. It has been known for a long time that tumour cells have a high rate of anaerobic glycolysis, even in normoxic conditions (this is the Warburg effect); this may represent a preadaptation to the risk of hypoxia, which may have been genetically selected during repetitive hypoxia events followed by re-oxygenation.

Whereas there are presently no ways to target tumour cells by agents that would selectively kill the cells using anaerobic glycolysis rather than oxidative phosphorylation, this peculiarity of tumour cells is utilised by an imaging technique, positron-emission tomography: 2-deoxy-2-^{18}F-glucose, a non-metabolised glucose analogue, serves as a glucose substitute to detect tumour cells.

16.1.4 Other Consequences of Hypoxia

Among the various hypoxia-regulated genes is *DDIT4* (*DNA damage-inducible transcript 4*), which encodes REDD1/RTP801 (*regulated in development and DNA damage response*). This protein activates the complex TSC1–TSC2, an inhibitor of mTOR activation by the small G-protein called RHEB (Chap. 3); mTOR is involved in protein synthesis, lipogenesis, cell growth and autophagy. Hypoxia appears therefore as a way to reduce protein synthesis and to protect cell survival by autophagy.

Other potential targets of HIFs are the genes involved in apoptosis triggering (Chap. 18), via the opening of the transition pore that enables the efflux of cytochrome *c* from the mitochondria: proteins of the BNIP (*BCL2-interacting protein*) family are induced by HIFs. It seems that apoptosis induction can be operated when the hypoxic cell can no longer envisage survival.

Finally, many proteins playing a fundamental role in cell adhesion and migration are also transcriptionally induced by HIFs: vimentin (VIM), fibronectin (FN), keratins (KRT), matrix metalloproteinases (MMPs), cathepsins (CTSs), etc. Indirectly, E-cadherin (CDH1, Chaps. 7 and 11), which plays a major role in inhibiting epithelial-to-mesenchymal transition, is repressed during HIF activation. Hypoxia is, therefore, heavily implicated in metastatic processes.

16.1.5 Pharmacological Targeting of Hypoxia

Since tumour cells generally live and grow in a hypoxic environment, it would be interesting to develop drugs which would selectively target hypoxia. An old method for targeting hypoxic cells is the use of bioreductive drugs, which are activated by hypoxia by monoelectronic reduction catalysed by DT-diaphorase (NQO1) or cytochrome P450 reductase (POR). Mitomycin C is a good example of such drugs. Another approach is the use of nitroimidazoles, which have a radiosensibilising effect at the level of hypoxic regions of tumours; after reduction, these compounds stabilise DNA lesions. Misonidazole is used in positron emission tomography to visualise hypoxic tumour regions.

Upstream HIF1α, hypoxic cells targeting may consist in mTOR inhibition, which would induce a decrease in HIF1α expression (Chap. 3). Rapalogs and mTOR serine/threonine kinase inhibitors can be considered as antiangiogenic treatments and are especially active in renal cancers for which VHL mutations is an oncogenic driver. Targeting HIF1α itself is a difficult challenge. Indirect approaches with heat shock proteins (HSPs) inhibitors, such as geldanamycin, are considered. Direct approaches are represented by antisense oligodeoxynucleotides or by small molecules or monoclonal antibodies targeting HIF1α-induced proteins, such as carbonic anhydrase IX (CA9) or the glucose transporter GLUT1.

16.2 Oxidative Stress

16.2.1 Generation of Reactive Oxygen Species (ROS)

ROS are endogenously produced from molecular oxygen in several organelles, especially mitochondria, which is the main place where oxygen is utilised. ROS is the generic name for three chemical entities: the superoxide radical $O_2^{-\bullet}$, the hydroxyl radical $OH^{\bullet}$ and hydrogen peroxide H_2O_2. Superoxide ions are formed by reduction of molecular oxygen during a reaction catalysed by several enzymes, especially the oxidoreductases of the mitochondrial electron transport chain and membrane NADPH oxidases (NOX):

$$NADPH + O_2 \rightarrow O_2^{-\bullet} + NADP^+$$

Superoxide ions are detoxified by superoxide dismutases (SOD), Cu^{2+}-Zn^{2+}–SOD in the intermembrane mitochondrial fluid and Mn^{2+}–SOD in the mitochondrial matrix. The dismutation reaction generates hydrogen peroxide as follows:

$$2O_2^{-\bullet} + 2H^+ \rightarrow H_2O_2 + O_2$$

Hydrogen peroxide itself is detoxified by various enzymes: catalases (CAT) and peroxiredoxins (PRDX). However, in the presence of some metal ions such as Fe^{2+}, but also copper, chromium and cobalt, hydrogen peroxide can participate to the Fenton reaction, which generates hydroxyl radicals, $OH^{\bullet}$.

$$H_2O_2 + Fe^{2+} \rightarrow OH^- + OH^{\bullet} + Fe^{3+}$$

The Haber–Weiss reaction is also able to generate $OH^{\bullet}$ radicals:

$$O_2^{-\bullet} + H_2O_2 \rightarrow OH^- + OH^{\bullet} + O_2$$

This reaction combines in fact the Fenton reaction with Fe^{3+} reduction by superoxide ions:

$$Fe^{3+} + O_2^{-\bullet} \rightarrow Fe^{2+} + O_2$$

Hydroxyl radicals $OH^{\bullet}$ are highly reactive chemical species, with a half-life in aqueous media of less than 1 ns. They react therefore instantaneously very close

from their production sites and indifferently attack all molecular electron-rich regions in lipids, proteins and nucleic acids, especially because there are no rescue mechanisms for $OH^{\bullet}$ detoxification. As a consequence, peroxide radicals $ROO^{\bullet}$ are formed in these molecules. These peroxides can be detoxified by peroxiredoxins, thioredoxin (TXN) and other enzymes such as glutathione peroxidases (GPX), which catalyse the following reaction (Fig. 16.2a):

$$G-SH+ROO^{\bullet} \rightarrow \%\!\!\%G-SS-G+ROH$$

ROS can also generate protein mild oxidations and induce the formation of disulphide bridges that glutathione or thioredoxin can also detoxify (Fig. 16.2b).

16.2.2 ROS Contribution to Carcinogenesis and Anticancer Therapy

At the level of nucleic acids, ROS generate several oxidation products of purine and pyrimidine bases, the most frequent being 8-oxoguanine (Annex A). ROS appear, therefore, as mutagenic and carcinogenic agents; protection against ROS may represent a cancer prevention strategy. Several antioxidants such as ascorbic acid or β-carotene, and ROS scavengers such as resveratrol, *N*-acetylcysteine or quercetin have been proposed as nonconventional cancer therapies without any success. ROS are also formed by the action of ionising radiations on the water molecule: they are considered as one of the major modalities of cell death induction by radiotherapy. The effect of ionising radiations is highly increased by the oxygenation of irradiated tissues (what has been called the 'oxygen effect'); much research has been devoted to the identification of ways of increasing tumour tissue oxygenation, aiming at potentialising radiotherapy efficiency. In addition, ROS-generating agents at the tumour level have been proposed in therapeutics.

16.2.3 ROS as Signalling Agents

Independently of their deleterious effects, ROS may operate as signalling molecules and participate to the regulation of several physiological processes: cell proliferation and differentiation, adhesion and migration, apoptosis and survival. Upstream, various growth factors and cytokines are able to induce ROS formation by activating the transcription of NADPH oxidases; downstream, ROS can oxidise cysteine residues of various kinases and phosphatases and modulate their activity. Regulation of these processes can be exerted by ROS detoxification proteins such as SOD.

A first example is provided by ROS activation of the MAP3 kinase ASK1 (*apoptotic signal-regulated kinase 1*, gene *MAP3K5*), upstream the JNK and p38 pathways (Chap. 2). This kinase is kept inactive through binding to thioredoxin.

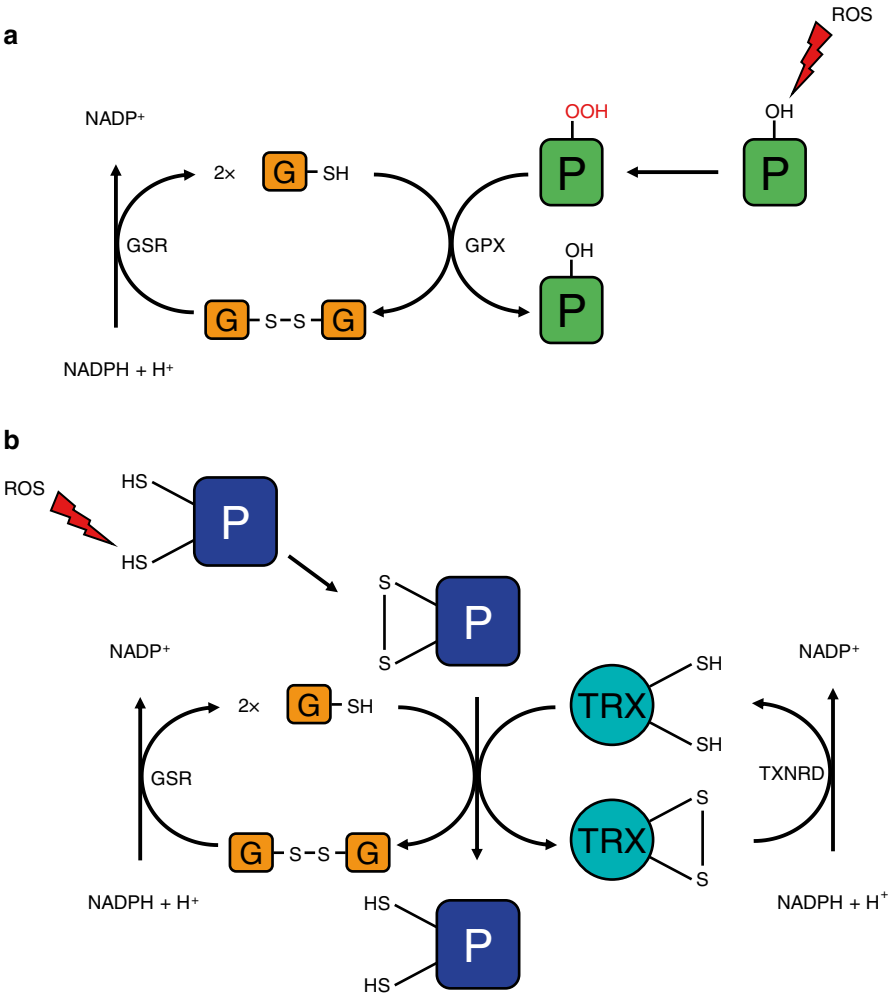


Fig. 16.2 Detoxification of oxidation products generated by ROS. (**a**) ROS can oxidise the thiol functions of cysteine residues of proteins (*P*) and form disulphide bridges. Reduced glutathione G–SH is able to reduce these disulphide bonds to thiols, thus forming oxidised glutathione G–S–S–G. Reduced glutathione is regenerated, thanks to a flavoprotein utilising reduced NADP, glutathione reductase (*GSR*). Thioredoxin can play the same role as glutathione; it is regenerated by thioredoxin reductase (*TXNRD*). The formation of disulphide bridges in a given protein, included thioredoxin, can initiate a signalling pathway owing to conformational changes of the protein. (**b**) ROS can oxidise hydroxyl functions borne by proteins (*P*) to form peroxides. Glutathione peroxidases (*GPX*) are able to reduce the peroxides –OOH into hydroxyls –OH by oxidising reduced glutathione G–SH into oxidised glutathione G–S–S–G

Thioredoxin is oxidised by ROS to form a disulphide bridge, which releases ASK1 and enables the initiation of the MAP kinases cascade. Another example is that of the NFκB transcription factors (Chap. 12), which are deactivated by ROS-induced oxidation of cysteine residues to form disulphide bridges. Another group of transcription factors, FOXOs (*forkhead box class O*, Chaps. 2 and 3), are specifically activated by hydrogen peroxide and induce cell death or a cell quiescence state characterised by tolerance to oxidative stress. Peroxiredoxins and thioredoxin are able to reduce these oxidised transcription factors and thus participate to the regulation of these signalling pathways.

The question is raised of the specificity of such highly reactive signalling molecules; whereas it is possible, because of their longer half-life, that hydrogen peroxide or superoxide ion may have some specificity, this is not conceivable for the hydroxyl radical. ROS-mediated signalling remains controversial. Bacteria have developed ROS response systems with high-affinity receptors that have been well characterised. This does not appear to be the same in mammals, for which ROS responses seem more diffuse and less specific, being only possible for some transcription factors such as MYC, NFκB, FOXO, p53 and a PPAR coactivator called PGC1 (*PPAR gamma coactivator*). The presently available data are insufficient to present an overall well-systematic view of ROS signalling in mammals.

16.3 Nitric Oxide

16.3.1 $NO^{\bullet}$ Generation

Nitric oxide or nitrogen monoxide $NO^{\bullet}$ is a gaseous radical compound, highly diffusible, of low molecular weight, with a half-life of some seconds in aqueous media, all features that have shed strong doubts on its status of intracellular messenger. It is produced by the three NO synthases (NOS): neuronal (nNOS, gene *NOS1*), inducible (iNOS, gene *NOS2*) and endothelial (eNOS, gene *NOS3*), from arginine as nitrogen donor, molecular oxygen and NADPH as electron donor. NOSs are haem proteins, sharing some similarities with cytochromes, and operate as dimers. They undergo several post-translational modifications (acylation, phosphorylation; see Annex C). They are activated by Ca^{2+} via calmodulin (Chap. 15) and by kinases such as AKT. iNOS is induced in response to various transcriptional stimuli, brought by some cytokines (Chap. 4) or NFκB (Chap. 12) and, conversely, repressed by other transcription factors such as p53.

$NO^{\bullet}$ is generated in response to endothelial or nerve cell stimulations; it is able to rapidly diffuse and exert, therefore, paracrine effects on cells present in the immediate vicinity of those where it was produced, frequently smooth muscle cells. $NO^{\bullet}$ plays numerous physiological roles at the vascular level (flux, permeability, angiogenesis, etc.) and in neuronal transmission. It is cytotoxic at high concentrations and exerts pro- and antiproliferative effects in different contexts. It is able to react with the superoxide ion to generate the peroxynitrite ion, $ONOO^{-}$, highly reactive and toxic, as follows:

$$NO^{\bullet} + O_2^{-\bullet} \rightarrow ONOO^{-}$$

as well as with haem proteins such as guanylyl cyclases (see below):

$$Fe^{2+} + NO^{\bullet} \rightarrow Fe^{2+} - NO$$

on which it can compete with another haem ligand, carbon monoxide CO. Finally, it is also able to induce protein nitrosylation at the level of cysteine residues, forming thus an –SNO adduct.

16.3.2 Cell Responses to $NO^{\bullet}$

The main targets of $NO^{\bullet}$ are cytoplasmic guanylyl cyclases (GUCYs), enzymes able to convert GTP into cyclic GMP (cGMP), which it is itself catabolised into GMP by a phosphodiesterase. This is similar to the conversion of ATP into cAMP by adenylyl cyclases (Chap. 6). GUCYs are αβ heterodimers, encoded by genes *GUCY1A2* and *1A3* for the α subunit and *GUCY1B2* and *1B3* for the β subunit. $NO^{\bullet}$ can bind the haem Fe^{2+}, which activates the enzyme.

cGMP is able to regulate cationic channels, protein kinases as well as some phosphodiesterases. A group of serine/threonine kinases, called PKG or PRKG, are specifically activated by cGMP. Effects of cGMP of smooth muscle relaxation or platelet adhesion and aggregation are now well characterised and have been used in pharmacology for a long time: trinitroglycerol, able to generate nitric oxide, is used for more than 150 years as a coronary vasodilator, and sildenafil, a phosphodiesterase inhibitor, is used for some years as a vasodilator of erectile tissues.

Another process engaging $NO^{\bullet}$ is protein nitrosylation, a still poorly known process. Caspase 3 nitrosylation inhibits its activity, whereas RAS nitrosylation activates the MAP kinases cascade. HIFα nitrosylation stabilises this factor and mimics therefore the effects of hypoxia.

16.3.3 $NO^{\bullet}$ Contribution to Carcinogenesis and Anticancer Therapy

Owing to the endothelial localisation of eNOS, $NO^{\bullet}$ appears as an important mediator of angiogenesis. Thanks to cGMP and PKG, it is able to activate endothelial cell proliferation using the MAP kinases pathway (Chap. 2), either after activation of RAF proteins or through RAS nitrosylation.

In tumour cells, iNOS expression may have the same consequences on cell proliferation and migration as those exerted by eNOS on endothelial cells. However, controversial results have been published, and it appears difficult to attribute an oncogenic character to the $NO^{\bullet}$ pathway.

Two pharmacological approaches have been developed: the first one aims at decreasing $NO^{\bullet}$ production by NOS inhibitors; there are several possible ways, especially peptides mimicking caveolin structure; the second one aims at generating high intratumour $NO^{\bullet}$ concentrations in order to benefit of its cytotoxic effect.

16.4 Guanylyl Cyclase Receptors

In addition to $NO^{\bullet}$-activated cytoplasmic guanylyl cyclases, there exist a small group of guanylyl cyclase receptors which offer an original signalling pathway for a cognate small group of peptides. These proteins serve as receptors for the natriuretic peptides (NPPs), namely, atrial natriuretic peptide (ANP, gene *NPPA*), brain natriuretic peptide (BNP, gene *NPPB*) and CNP (gene *NPPC*). These factors display natriuretic, diuretic and vasorelaxant properties and are involved in blood pressure regulation. ANP and BNP behave rather as endocrine hormones and CNP as a paracrine factor in the heart. These receptors are GCA or GUCY2A (gene *NRP1*) for ANP and BNP, and GCB or GUCY2B (gene *NRP2*) for CNP. Both receptors have an extracellular part for ligand binding, with cysteine-rich domains, a single membrane-spanning domain and an intracellular part with a kinase homology domain, an oligomerisation domain and a *C*-terminal catalytic guanylyl cyclase domain. They are phosphorylated in the basal state and undergo desensitisation upon dephosphorylation of specific sites in their kinase homology domain. A third receptor (NRP3), with very short intracellular domain, mainly serves as a decoy receptor for clearing extracellular NPPs.

Guanylyl cyclase activators, guanylin (gene *GUCA2A*) and uroguanylin (gene *GUCA2B*), serve as ligands for an intestinal endotoxin receptor with guanylyl cyclase activity (GCC, gene *GUCY2C*), and two other membrane guanylyl cyclases, GUCY2D and GUCY2F, are present in olfactory and photoreceptors, respectively, but are not known to display specific receptor functions. Stimulation of GCC promotes colon cancer cell growth arrest, and this protein appears as a tumour suppressor. Agonist activation of GCC has been proposed as a prevention approach for colorectal cancer.

Downstream guanylyl cyclase receptors, cyclic GMP plays the same roles as those already mentioned for cytoplasmic soluble guanylyl cyclases.

Further Reading

Block K, Gorin Y. Aiding and abetting roles of NOX oxidases in cellular transformation. Nat Rev Cancer. 2012;12:627–37.

Brahimi-Horn C, Pouysségur J. The role of the hypoxia-inducible factor in tumor metabolism growth and invasion. Bull Cancer. 2006;93:E73–80.

D'Autréaux B, Toledano MB. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. Nat Rev Mol Cell Biol. 2007;8:813–24.

Fang J, Seki T, Maeda H. Therapeutic strategies by modulating oxygen stress in cancer and inflammation. Adv Drug Deliv Rev. 2009;61:290–302.

Foster MW, Hess DT, Stamler JS. Protein S-nitrosylation in health and disease: a current perspective. Trends Mol Med. 2009;15:391–404.

Francis SH, Busch JL, Corbin JD, Sibley D. cGMP-dependent protein kinases and cGMP phosphodiesterases in nitric oxide and cGMP action. Pharmacol Rev. 2010;62:525–63.

Giaccia A, Siim BG, Johnson RS. HIF-1 as a target for drug development. Nat Rev Drug Discov. 2003;2:803–11.

Gorrini C, Harris IS, Mak TW. Modulation of oxidative stress as an anticancer strategy. Nat Rev Drug Discov. 2013;12:931–47.
Grek CL, Tew KD. Redox metabolism and malignancy. Curr Opin Pharmacol. 2010;10:362–8.
Hancock JT. The role of redox mechanisms in cell signalling. Mol Biotechnol. 2009;43:162–6.
Kaelin Jr WG. The von Hippel-Lindau tumour suppressor protein: O2 sensing and cancer. Nat Rev Cancer. 2008;8:865–73.
Kamata T. Roles of Nox1 and other Nox isoforms in cancer development. Cancer Sci. 2009; 100:1382–8.
Keith B, Johnson RS, Simon MC. HIF1α and HIF2α: sibling rivalry in hypoxic tumour growth and progression. Nat Rev Cancer. 2011;12:9–22.
Leiper J, Nandi M. The therapeutic potential of targeting endogenous inhibitors of nitric oxide synthesis. Nat Rev Drug Discov. 2011;10:277–91.
Miyata T, Takizawa S, van Ypersele de Strihou C. Hypoxia. 1. Intracellular sensors for oxygen and oxidative stress: novel therapeutic targets. Am J Physiol Cell Physiol. 2011;300:C226–31.
Murad F. Nitric oxide and cyclic GMP in cell signaling and drug development. N Engl J Med. 2006;355:2003–11.
Nogueira V, Hay N. Molecular pathways: reactive oxygen species homeostasis in cancer cells and implications for cancer therapy. Clin Cancer Res. 2013;19:4309–14.
Ruan K, Song G, Ouyang G. Role of hypoxia in the hallmarks of human cancer. J Cell Biochem. 2009;107:1053–62.
Semenza GL. Hypoxia-inducible factors in physiology and medicine. Cell. 2012a;148:399–408.
Semenza GL. Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. Trends Pharmacol Sci. 2012b;33:207–14.
Shen C, Kaelin Jr WG. The VHL/HIF axis in clear cell renal carcinoma. Semin Cancer Biol. 2013;23:18–25.
Trachootham D, Alexandre J, Huang P. Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? Nat Rev Drug Discov. 2009;8:579–91.
Wilson WR, Hay MP. Targeting hypoxia in cancer therapy. Nat Rev Cancer. 2011;11:393–410.
Wu WS. The signaling mechanism of ROS in tumor progression. Cancer Metastasis Rev. 2006;25:695–705.

Cell Cycle Control

17

Abstract

Cell cycle consists of the sequential events occurring between the emergence of a new cell and that of the two daughter cells generated by the first one. Cell cycle integrates a continuous growth cycle (increase in cell mass) and a discontinuous division chromosomal cycle (DNA replication and mitosis, i.e. distribution of the genome between the two daughter cells). Cell cycle entry represents, therefore, the execution of the cell proliferation programme, in response to the messages brought by growth factors, as studied in the previous chapters. The transduction of these messages leads, among other events, to the transcription of the genes required for cell cycle entry, the first one being cyclin D1. Cell cycle control is of crucial importance in oncogenesis, as the loss of some of its essential checkpoints characterises malignant transformation.

We present here essentially the pathways used for triggering and controlling cell cycle progression, and not the events themselves (DNA replication, mitosis). After a brief description of the main events that occur during the different phases of the cycle, we will present the effectors of these events, then the control processes themselves, before the description of the alterations of cell cycle regulation in cancer. In the early days, cell cycle was only defined by two phases, mitosis and interphase; DNA replication was afterwards recognised as an active phase, separated from mitosis by two *gaps*; these gaps are in no way resting phases, but important periods during which are actively prepared the events leading to DNA replication and mitosis.

17.1 Cell Cycle Phases

17.1.1 G1 Phase

G1 phase, between mitosis and DNA synthesis, has the most variable length among cell cycle phases. If nutriments are lacking or if the cells receive an antiproliferative signal, or a signal inducing terminal differentiation, the cells can delay their

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_17

progression in the cycle or leave the cycle and enter a quiescent phase, called G0. G1 progression is controlled by a checkpoint on DNA integrity; once this checkpoint has been cleared, the cycle is engaged without any possibility of stopping. DNA replication is prepared in G1 through derepression of transcription factors, which enable the transcription of the genes required for replication. This derepression involves their release from proteins that retain them and are called *pocket proteins*. The synthesis of the purine and pyrimidine nucleotides aimed at DNA incorporation requires numerous enzymes such as dihydrofolate reductase, thymidylate synthase and many others; DNA replication itself also requires numerous enzymes such as diverse polymerases, primase, helicases, topoisomerases, etc. The promoters of many of the genes encoding these proteins are activated during the G1 phase.

17.1.2 S Phase

DNA replication is simultaneously initiated on numerous different sites, called *origins of replication*. Each DNA region that is replicated from a given origin is called a *replicon*. Each group of replicons is replicated at a characteristic time of the S phase, generally at its beginning for the actively transcribed regions, whereas the regions containing transcriptionally inactive heterochromatin are transcribed later.

During the G1 phase, the chromosomes are modified in order to obtain the 'authorisation' to replicate by binding special proteins at the origins of replication, so that they form a pre-replication complex. During replication, these complexes are inactivated, preventing thus the chromosomal regions to replicate two times during the cycle. Phase S steps (see Annex A) comprise the opening of the double helix (requiring, for instance, helicases, topoisomerase 2, etc.) and the synthesis of DNA (requiring, for instance, primase, DNA polymerases, etc.). In addition to these enzymes, several proteins are required, associated to the initiation, elongation and termination processes, as well as histones for DNA packaging in nucleosomes.

17.1.3 G2 Phase

During the G2 phase, between DNA synthesis and mitosis, cells can make corrections to DNA structure, thanks to DNA repair processes, and prepare to mitosis. If damaged or non-replicated DNA is detected at the G2 checkpoint, a cascade of protein kinase activities is triggered, leading to the inactivation of cell cycle progression. Furthermore, protection mechanisms prevent the cell from initiating a new DNA replication process as long as mitosis has not occurred.

17.1.4 M Phase

During M phase and the subsequent cytodieresis (often improperly called cytokinesis), chromosomes and cytoplasm are divided to form two daughter cells. Chromosome segregation is controlled by the metaphase checkpoint, which delays chromatid separation until chromosomes are correctly in alignment on the equatorial plate of the mitotic spindle. Mitosis is divided in five phases (Fig. 17.1):

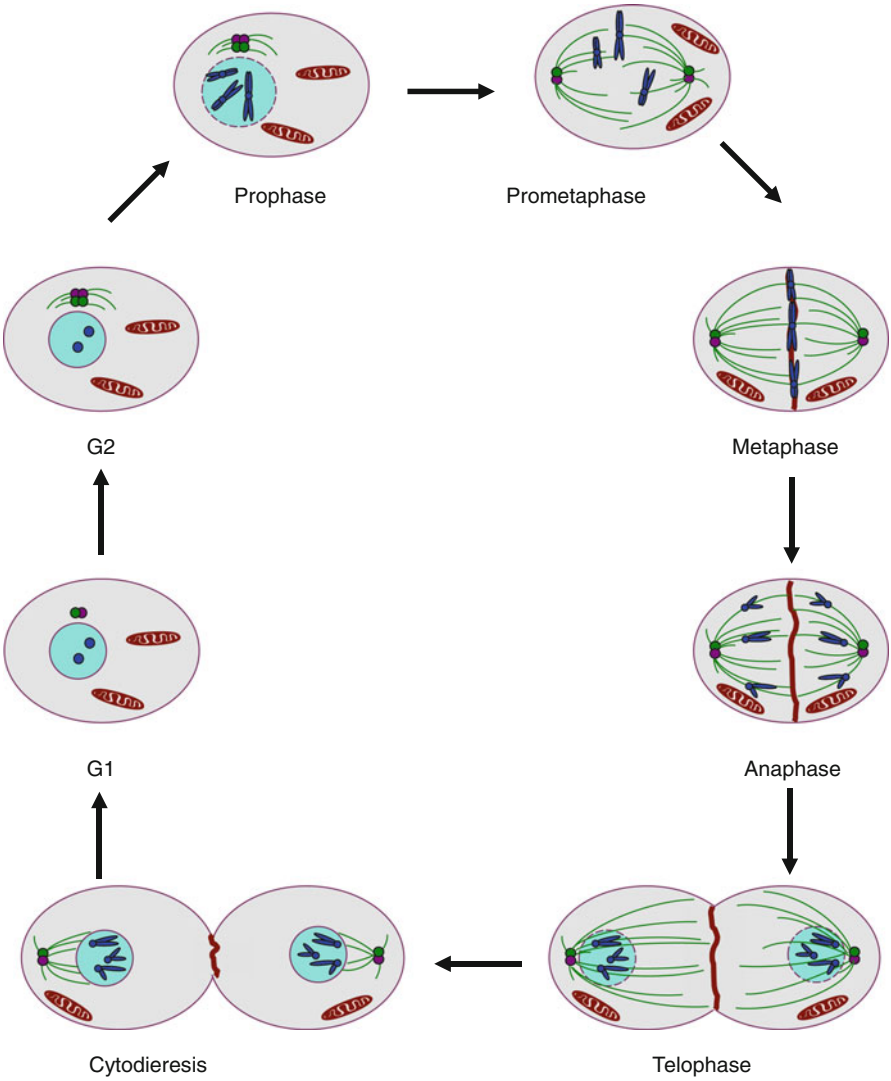


Fig. 17.1 Schematic representation of a mitosis. *Prophase*: chromatin is condensed to form chromosomes that gather the two sister chromatids; the centrosomes, duplicated in G2, get organised and serve as nucleation starters for microtubules. *Prometaphase*: the nuclear envelope is disrupted, the centrosomes migrate at two opposite poles of the cell and the microtubules get organised as a spindle on which chromosomes are attached at the level of their kinetochore. *Metaphase*: the chromosomes get aligned along the equatorial plate after their 'capture' by microtubules. *Anaphase*: the two sister chromatids are separated by APC/C (*anaphase-promoting complex*); the chromatids are directed to opposite poles and the microtubular spindle is spread in the entire cell. *Telophase*: the two copies of the genome are distributed at each pole of the cell, the nuclear envelope is reconstituted and a pinch is formed between the two opposite poles. *Cytodieresis*: a contractile ring is formed between the two nuclei and the daughter cells can separate

- *Prophase*: chromatin is condensed into chromosomes; centrosomes, duplicated during the G2 phase, are organised and serve as nucleation origins for microtubules.
- *Prometaphase*: the nuclear envelope is disrupted, the centrosomes migrate at the two poles of the cell and the microtubules are organised as a spindle where chromosomes are attached, at the level of their kinetochore.
- *Metaphase*: chromosomes align at half-distance of the two poles, on an 'equatorial plate', after they have been 'captured' by the microtubules.
- *Anaphase*: the homologous chromosomes leave the equatorial plate after deactivation of cohesins and activation of the separase complex enabled by the degradation of securin (gene *PTTG1*) in the proteasome, where it is driven by the APC/C–CDC20 (*anaphase-promoting complex/cyclosome–cell division cycle 20*) complex; homologous chromosomes migrate in opposite directions to the poles.
- *Telophase*: the two copies of the genome are distributed at each pole of the cell, the nuclear envelope is reconstituted and a tightening occurs between the two nuclei.
- *Cytodieresis*: a contractile actin ring is assembled at half-distance of the two poles and tightens the equatorial zone. This process enables the separation of the daughter cells.

17.2 Effector Proteins of Cell Cycle Control

17.2.1 Cyclins and Cyclin-Dependent Kinases

Cyclins (CCNs) are proteins of 35–90 kDa with no catalytic activity, but indispensable to the activity of serine/threonine kinases which are thus called cyclin-dependent kinases (CDKs). They are constituted of two central symmetric domains comprising five α helices. One of these domains (cyclin box) is highly conserved and represents the structural characteristic of these proteins. Cyclins undergo cyclic production and are present in the cell at definite periods of the cell cycle: cyclins D during G1 phase, cyclin E at the end of the G1 phase, cyclin A during S and G2 phases and cyclins B during M phase. Synthesis and degradation of cyclins follow thus a precise regulation, which requires definite transcription factors for synthesis and E3 ubiquitin ligases for their proteasomal destruction.

Cyclin-dependent kinases (CDKs) are generally smaller than cyclins (30–40 kDa) and display serine/threonine kinase activity only after binding with a cyclin. Cyclin binding to CDKs induces a conformational change enabling the ATP present in the active centre to react with protein substrates. CDKs are present in constant amounts all along the cell cycle: this is indeed the binding to a stimulatory cyclin that constitutes the decisive event for them to play their catalytic role. CDK4 or CDK6 binds cyclin D during the G1 phase, CDK2 binds cyclin E at the end of the G1 phase and cyclin A during the S phase; CDK1 binds cyclin A during the G2 phase and cyclin B during the M phase. Figure 17.2 presents the intervention of the cyclin–CDK complexes during cell cycle progression.

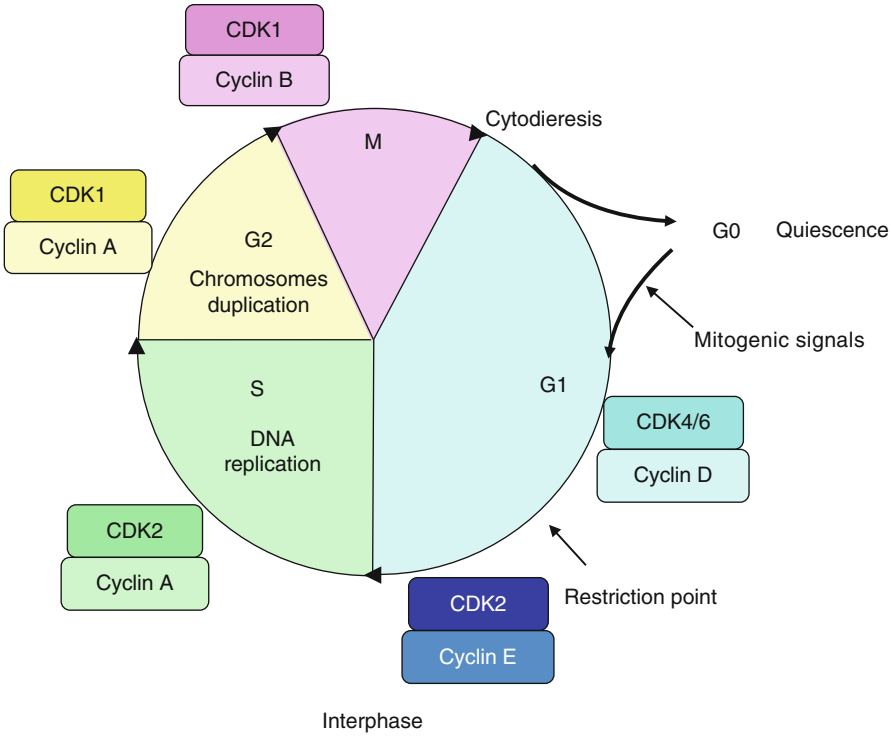


Fig. 17.2 Cyclin–CDK complexes and their intervention during the cell cycle. Cyclins are produced during precise periods of the cell cycle; by associating to cyclin-dependent kinases (CDK), they enable these kinases to phosphorylate the proteins required for cell cycle progression. The cyclin D–CDK4/6 complexes are involved in cell cycle progression in G1; the cyclin E–CDK2 complex allows cell entry into S phase; the cyclin A–CDK2 and cyclin A–CDK1 complexes are active during the S and G2 phases, respectively; the cyclin B–CDK1 complex appears at the end of the G2 phase and constitutes the triggering factor of mitosis

Each cyclin–CDK couple presents phase specificity and substrate specificity for the proteins they can phosphorylate. Some of these substrates are well identified, for instance the RB1 (*retinoblastoma 1*) protein for the complex cyclin D–CDK4/6 (see below), but all substrates have not yet been identified. As a general feature, the cyclin–CDK substrates are either enzymes that intervene in a specific way to catalyse a step required for cell cycle progression, or transcription factors enabling the synthesis of the proteins required for these processes. A special cyclin–CDK couple operate as activators of other cyclin–CDK couples: this is the cyclin H–CDK7 couple which, when associated to a third protein, MAT1 (*Ménage à trois 1*), phosphorylates the cyclin D–CDK4/6 and the cyclin B–CDK1 complexes on a threonine residue (Thr^{161} for CDK1), rendering them up to 300-fold more active. This is why the cyclin H–CDK7 is also known as CAK (*CDK-activating kinase*). This complex also operates in DNA repair (Annex A) and cannot activate a new mitosis as long as this repair has not been carried out.

17.2.2 Inhibitory Kinases WEE1 and MYT1

In addition to the cell cycle-activating kinase complexes, several inhibitory kinases can delay cell cycle progression. WEE1 (from the name of the 'petite' yeast mutation in Scottish) can phosphorylate CDKs on a tyrosine residue (Tyr^{15} for CDK1) while MYT1 (*myelin transcription factor 1*) phosphorylates CDKs on a threonine residue (Thr^{14} for CDK1), at the level of the ATP-binding site. They inhibit thus the kinase activity of the target CDKs, which will require reactivation thanks to phosphatases of the CDC25 subfamily (see below). WEE1 and MYT1 are deactivated by phosphorylation catalysed by AKT, a key element of the PI3 kinase pathway (Chap. 3), as well as by ubiquitinylation that drives them to the proteasome (Annex C).

17.2.3 Phosphatases

In addition to the two kinase types, stimulatory and inhibitory, two types of phosphatases, conversely inhibitory and stimulatory, hydrolyse the phosphate moieties brought by the kinases. Inhibitory phosphatases operate thus against CDK function and remove the phosphate brought by CAK (Thr^{161} for CDK1); it has been called KAP or KAC (gene *CDKN3*) by inverting the acronym CAK. Stimulatory phosphatases remove the phosphate groups brought by WEE1 and MYT1; these are called CDC25s, a small family of three different enzymes (CDC25A, B and C) with overlapping specificities.

CDC25A seems rather involved in the regulation of the G1 → S transition and CDC25B and CDC25C in the G2 → M transition. CDC25 are major targets for the control checkpoint of DNA integrity in G2. The activation of cyclin–CDK complexes by CDC25 phosphatases constitutes the triggering event for these complexes. CDC25 phosphatases can be activated by phosphorylation catalysed by PLK1 (*Polo-like kinase 1*) and by cyclin–CDK complexes, in a positive retroactive loop, and deactivated by phosphorylations on other residues by the checkpoint kinases CHK1 and CHK2 (genes *CHEK1* and *CHEK2*), two serine/threonine kinases activated by the kinases in charge of DNA damage detection, ATM and ATR (see below).

17.2.4 Protein Inhibitors of CDKs

In addition to the regulation system involving phosphorylation–dephosphorylation processes, a system of regulation through protein–protein interactions has been identified. It involves two families of small protein inhibitors of CDKs that are called CKIs. The first family is INK4 (*inhibitors of kinases*) and comprises four proteins with molecular weights of 16, 15, 18 and 19 kDa, respectively encoded by genes INK4a, INK4b, INK4c and INK4d (now *CDKN2A*, *CDKN2B*, *CDKN2C* and *CDKN2D*), and usually called $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$ and $p19^{INK4d}$. These

proteins inhibit the binding of cyclins D to CDK4 or 6 and modify the conformation of the CDKs and their ATP-binding site. The second family is called CIP/KIP and comprises three proteins: $p21^{CIP1}$, $p27^{KIP1}$ and $p57^{KIP2}$, which indicates both their molecular weights and the (older) names of the genes that encodes them, which now are *CDKN1A*, *CDKN1B* and *CDKN1C*, respectively. These proteins bind the cyclin E–CDK2 complex and strongly inhibit its function, firstly by modifying the conformation of the CDK, secondly by binding the ATP site. The cyclin E–CDK2 complex phosphorylates the KIP proteins, enabling their transfer to the proteasome by the E3 ubiquitin ligase complex SKP1–SKP2 (*S phase kinase-associated protein 1/2*). The AKT protein of the PI3 kinase pathway phosphorylates p21 and p27, ensuring thus their inhibition (Chap. 3).

CKI plays a fundamental role in cell growth regulation during the G1 and G0 phases, causing cell cycle arrest in response to antiproliferative signals, while this inhibition is alleviated in the presence of proliferation signals.

17.2.5 Mitotic Kinases

Two other serine/threonine kinase families are involved in cell cycle, the *aurora kinase* family (genes *AURKA*, *AURKB* and *AURKC*) and the *polo-like kinase* family (genes *PLK1* to *PLK5*); they are mainly involved in mitosis control. The aurora kinases can autophosphorylate when activated by appropriate protein interactions. AURKA controls mitosis entry by activating CDC25B and PLK1 by phosphorylation; it then ensures centrosomes separation and maturation, recruiting the proteins required to the assembly of the mitotic spindle. AURKB is localised at the level of the chromosomal kinetochores and ensures the attachment of the kinetochores on the spindle. AURKC principally intervenes during meiosis.

The polo-like kinases are present at the level of the spindle poles and of chromosome kinetochores. PLK1 is phosphorylated by AURKA, then by AURKB, and phosphorylates a series of substrates whose activation or inactivation is required to ensure their functions during mitosis, such as cohesin and separase during anaphase. PLK1 ensures thus the coordination of the successive steps of the mitosis. The other PLKs have been less studied.

17.2.6 Checkpoint Kinases and DNA Integrity Control Kinases

As mentioned earlier, there are two checkpoint kinases, CHK1 and CHK2 (genes *CHEK1* and *CHEK2*), which are involved in the control of cell cycle progress at the level of the G1 → S and of the G2 → M transitions, respectively. They inhibit the corresponding CDC25 phosphatases by phosphorylation and can therefore delay cell cycle progression. They are under the control of other kinases that check DNA integrity, ATM (*ataxia telangiectasia mutated*), ATR (*ataxia telangiectasia and Rad-3 related*) and DNA-dependent protein kinase, DNAPK (gene *PRKDC*).

17.2.7 Biochemical Mechanisms of Cell Cycle Regulation

Several mechanisms of regulation of the controlling elements of the cell cycle have been described here: (1) protein–protein interactions (between cyclins and CDKs, between CKIs and cyclin–CDK complexes); (2) phosphorylation–dephosphorylation processes (kinase and phosphatase activities, either on tyrosine or on serine/threonine residues); (3) proteasomal destruction of proteins such as cyclins, WEE1 and $p27^{KIP1}$ after ubiquitinylation. Other regulatory mechanisms can operate to control the activity of the cyclin–CDK complexes, such as their translocation from the cytoplasm where they are synthesised to the nucleus where they must operate. Cyclin phosphorylations appear to be responsible for these translocations. Figure 17.3 presents the main positive and negative regulatory processes that ensure cell cycle progression.

17.3 Control of Cell Cycle Progression

17.3.1 G1 → S Transition and DNA Synthesis

The mechanisms ensuring cell cycle progression in G1, with the objective of DNA replication, are well understood now. The G1 phase is not a simple gap, but an active

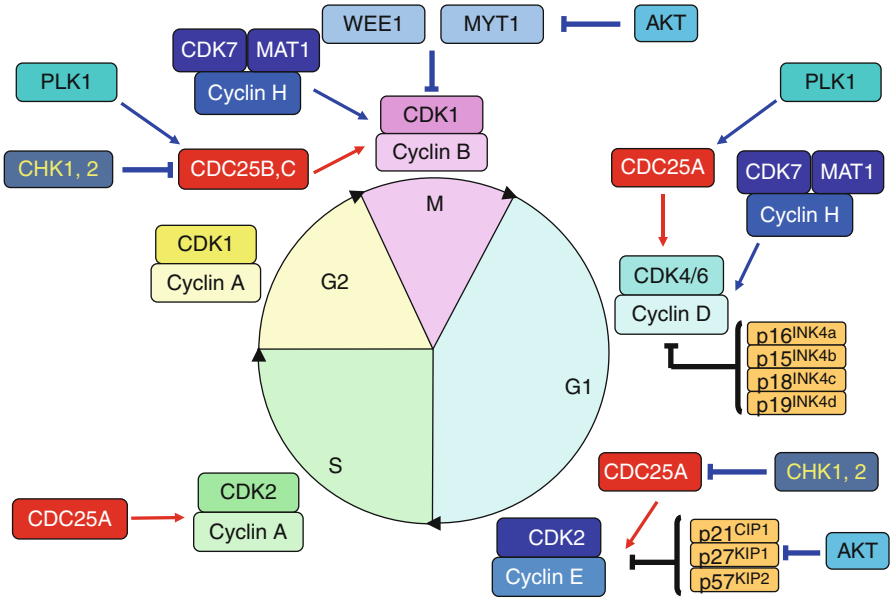


Fig. 17.3 Positive and negative regulations exerted on cell cycle. The cyclin–CDK complexes can receive a stimulatory phosphorylation on a threonine residue (Thr^{161} for CDK1) from the cyclin H–CDK7–MAT1 complex also known as CAK (*CDK-activating kinase*) and inhibitory phosphorylations on tyrosine and threonine residues (Tyr^{14} and Thr^{15} for CDK1) from WEE1 and MYT1. CDC25s are tyrosine phosphatases that activate the cyclin–CDK complexes by removing the inhibitory phosphate group. They are deactivated through phosphorylation by the checkpoint kinases, CHK1 and CHK2. The INK4 and CIP/KIP protein families are inhibitors of the cyclin–CDK complexes

phase during which decisive events occur toward well multiplication. These events are schematised on Fig. 17.4.

Cyclins D are produced thanks to the intervention, on the promoter of their genes, of the transcription factors activated through the MAP kinases and PI3 kinase proliferation pathways (Chaps. 1, 2 and 3). Cyclins D will then meet and bind CDK4 or CDK6 molecules. These complexes, still inactive owing to inhibitory phosphorylations by WEE1 and MYT1, are activated first by CAK and then by CDC25 phosphatases, enabling them to exert their catalytic function. CDC25A is the major activator of the G1 → S checkpoint; it is inhibited by CHK1-mediated phosphorylation, CHK1 being in turn is controlled by ATM.

Among the main substrates of the cyclin D–CDK4/6 complex are the retinoblastoma protein RB1 and its homologues RBL1 and RBL2 (*retinoblastoma-like 1* and *2*) (p107 and p130), also known as *pocket proteins*. These proteins regulate the activity of transcription factors of the E2F family. At the basal state (non-phosphorylated), RB1 is bound to E2F, which renders it inactive, maintaining cells in G0 or G1; after phosphorylation on various sites by the cyclin D–CDK4/6 complex, RB1 progressively loses its affinity for E2F, which can thus, once released, activate the transcription of a series of genes involved in DNA replication: DNA polymerases, primase, helicases, etc. (Annex A). The last phosphorylation of RB1, which allows the maximal release of E2F, is ensured by another cyclin–CDK complex: the cyclin E–CDK2 complex, which allows phase S entry.

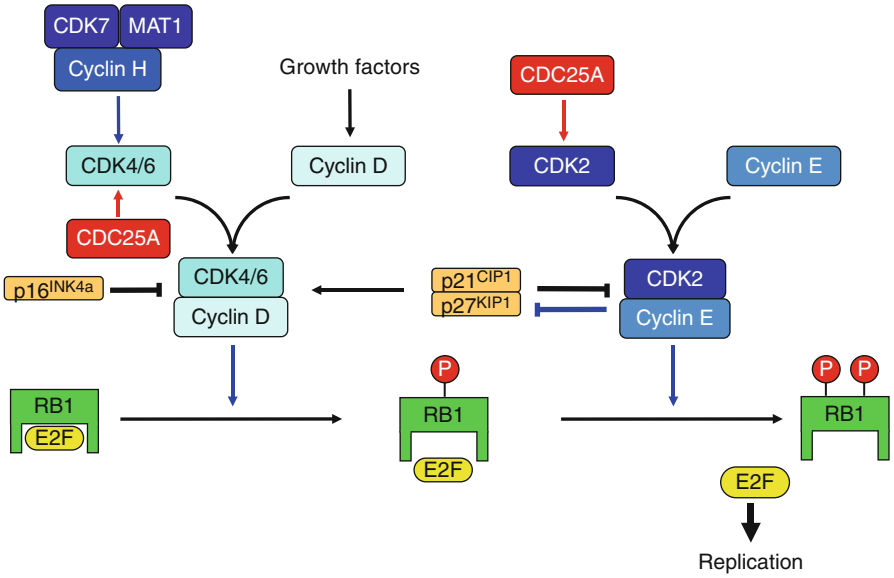


Fig. 17.4 G1 → S transition: phosphorylation of RB proteins. One of the major roles of the cyclin D–CDK4/6 and cyclin E–CDK2 complexes is to phosphorylate the RB1 protein, which is then dissociated from the transcription factors E2F, these factors becoming thus functional and activate the transcription of the genes required for DNA synthesis. Several phosphorylation sites are present on RB1, some of them being phosphorylated by the cyclin D–CDK4/6 complex and the last one by the cyclin E–CDK2 complex, which triggers the maximum release of E2F

This process is negatively regulated by INK4 family proteins (*CDKN2*), at the level of the cyclin D–CDK4/6 complexes, and by CIP/KIP family proteins (*CDKN1*), at the level of the cyclin E–CDK2 complex. Interestingly, these proteins also bind, but do not inhibit, CDK4: they facilitate cyclin D binding and activate the translocation of the complex to the nucleus. Thus, $p27^{KIP1}$ allows the formation of the cyclin D–CDK4 complex while delaying the activation of the cyclin E–CDK2 complex.

DNA synthesis is initiated through anchoring, on the origins of replication, of protein factors assembled as an ORC (*origin recognition complex*), then through the binding of the protein complex called MCM (*mini-chromosome maintenance*). In addition to the cyclin E–CDK2 complex, which phosphorylates a factor enabling the binding of DNA polymerase α to DNA, CDC45, another kinase complex of the same type, associating a cyclic protein, DBF4, to a kinase, CDC7, enables the stimulatory phosphorylation of the MCM proteins. DNA replication can then take place, according to the classical scheme described in Annex A.

17.3.2 G2 → M Transition and Mitosis

The mechanisms of activation of the cyclin B–CDK1 complex, once known as MPF (*mitosis-promoting factor*), then CDC2, are schematised of Fig. 17.5. Inhibitory phosphorylations, catalysed by the WEE1 and MYT1 kinases, take place on Tyr^{15} and Thr^{14}, whereas the CAK complex (cyclin H–CDK7–MAT1) carries out a stimulatory phosphorylation on Thr^{161}. WEE1 and MYT1 are then deactivated by AKT and driven to the proteasome by an E3 ubiquitin ligase complex, associating TOME1 (*trigger of mitotic entry*, gene *CDCA3*) to the factor SKP1 (*S phase kinase-associated protein*). Once bound to cyclin B, the complex is inactive as long as a CDC25 tyrosine phosphatase has not released the phosphate on the Tyr^{15} residue. This G2 → M checkpoint, ensured by CDC25B or C, is controlled through the inhibitory phosphorylation of CDC25 by CHK2, which is itself under the control of ATR.

Thanks to interactions of cyclin B with other proteins, the cyclin B–CDK1 complex can migrate to the nucleus and ensure all the phosphorylations required for mitosis (histones involved in chromosome decondensation, lamins, nucleolin, condensin, kinesins, peroxiredoxin, etc., as well as proteins implicated in cytoskeleton reorganisation during mitosis). Other regulatory controls of the activity of the cyclin B–CDK1 complex are exerted by kinases such as PLK1, which phosphorylates and activates cyclin B, and AURKA (which activates PLK1 by phosphorylation), B and C.

The active cyclin B–CDK1 complex exerts a positive retroaction on its own production through phosphorylation of CDC25 phosphatases and another positive action on the system in charge of cyclin B destruction by driving it to the proteasome; this double mechanism explains the rapidity with which the active complex cyclin B–CDK1 reaches a high concentration and then disappears at the end of mitosis. Finally, the cyclin B–CDK1 ensures an inhibitory phosphorylation of the WEE1 kinase, which also constitutes a positive retroaction. During its short existence, the cyclin B–CDK1 complex realises a high number of phosphorylations of the effector proteins of mitosis.

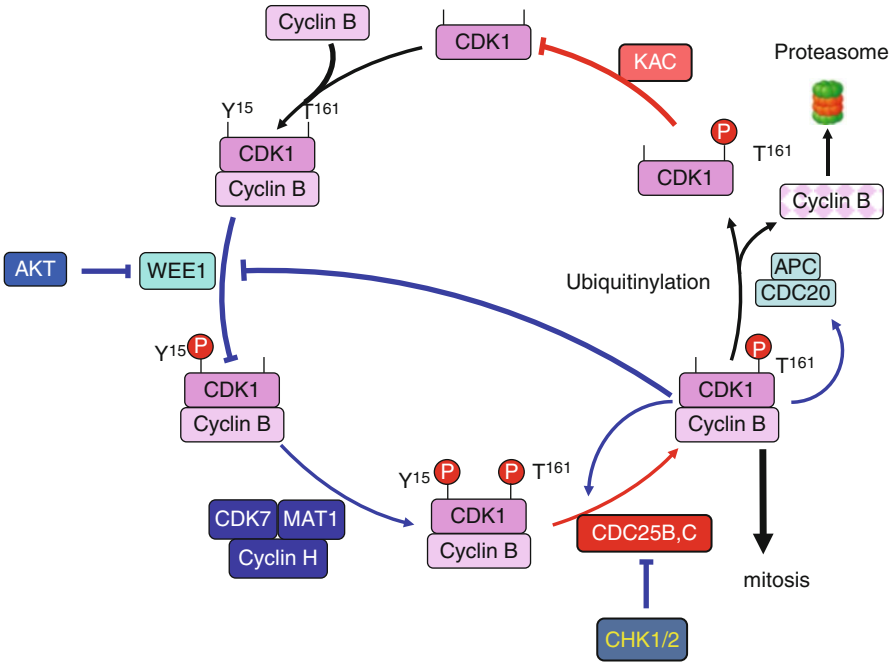


Fig. 17.5 G2 → M transition: activation of the cyclin B–CDK1 complex. To be active, the cyclin B–CDK1 complex must be phosphorylated on Thr^{161} by the cyclin H–CDK7–MAT1 complex (CAK) and dephosphorylated on Tyr^{14} by CDC25B/C after having received this phosphate group from WEE1. This complex stimulates its own activating dephosphorylation by CDC25 and its own proteasomal destruction by stimulating the engagement of cyclin B to the proteasome, enabling thus the formation of a high but transient peak of concentration of the active complex, which can thus phosphorylate its many substrates, including the APC/C–CDC20 complex. Kinase activity is represented by *blue lines* and that of the phosphatases by *red lines*. Substrate activation is represented by an *arrow* and its inhibition by a *bar*

We will only describe here some crucial events during mitosis control (Fig. 17.6):

- *Chromosomes condensation* is realised by a protein complex called condensin, activated by phosphorylation by the cyclin B–CDK1 complex; it also requires chromatin rearrangements obtained through post-translational modifications of histones, especially phosphorylations.
- *Formation of the mitotic spindle* occurs by polymerisation of αβ tubulin dimers and is regulated by diverse proteins that are activated by phosphorylation, kinesins on the + side and dyneins on the – side. Chromosomes bind the spindle on the equatorial plate and those not yet bound generate a 'waiting signal' serving as a mitotic checkpoint ensured by MAD2 (*mitotic arrest deficient 2*) and BUB1 (*budding uninhibited by benzimidazoles 1*), which inhibit a protein complex associating two proteins, APC/C (*anaphase-promoting complex/cyclosome*) and CDC20.
- *Chromosome separation* is carried out through inactivation of cohesins, owing to PLK1-mediated phosphorylation, PLK1 being itself activated by AURKA, and by

the activation of a protease that maintains it inactive, securin. Once the 'waiting signal' is alleviated, the APC/C–CDC20 complex acts an E3 ubiquitin ligase and leads securin to the proteasome. Separase and APC/C are activated by phosphorylation by the cyclin B–CDK1 complex.

- *Migration of the chromosomes* is realised when all chromosomes are attached to the mitotic spindle by their kinetochore. MAD2 sequesters CDC20 as long as it remains a chromosome not attached to the equatorial plate. The MAD2–CDC20 interaction is regulated by phosphorylation. The APC/C–CDC20 complex also governs the destruction of cyclin B in the proteasome, thus terminating mitosis.

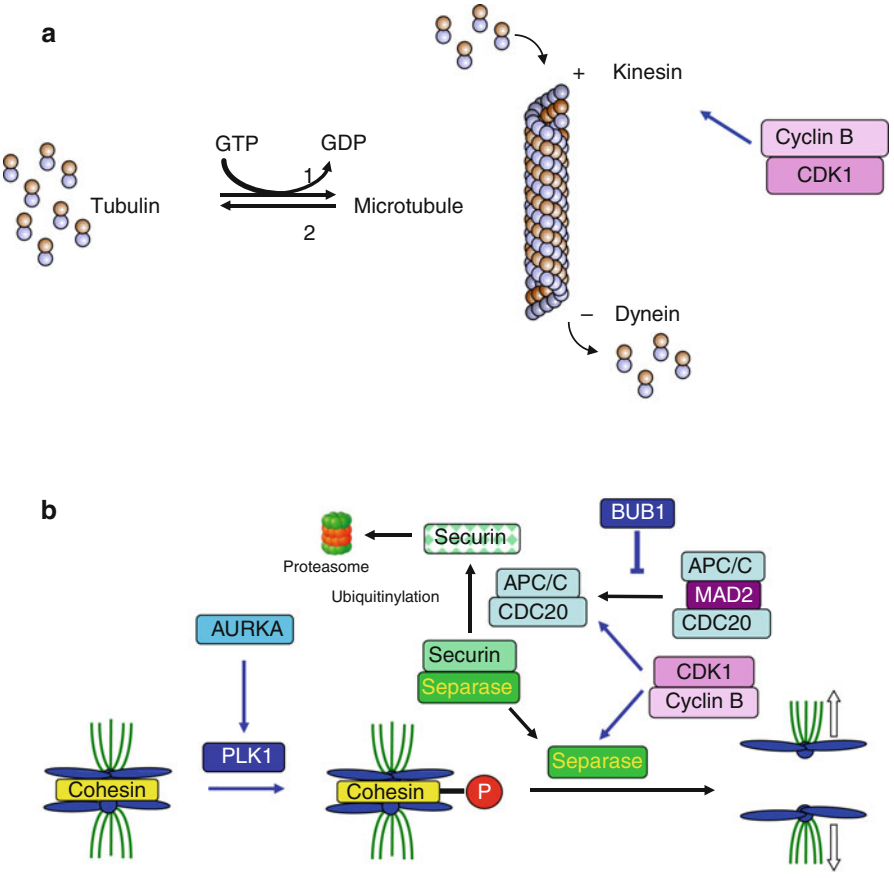


Fig. 17.6 Two aspects of mitosis: metaphase and anaphase. (**a**) Metaphase. To enter in metaphase, it is necessary to orientate the balance between tubulin and microtubules toward polymerisation, which is in part carried out by kinesins, which are activated by phosphorylation by the cyclin B–CDK1 complex. (**b**) Anaphase. Chromosome separation during anaphase is allowed by the deactivation of cohesins through phosphorylation by PLK1, and by activation of a protease complex, called separase, ensured by the destruction of a protein that maintain it inactive, securin (gene *PTTG1*). The APC/C–CDC20 complex drives securin to the proteasome. Separase and APC/C are activated by phosphorylation by the cyclin B–CDK1 complex

17.3.3 Control of DNA Integrity

The cell cycle is controlled at the level of several checkpoints, which can stop its progression; these checkpoints are under the supervision of the proteins that survey the integrity of newly synthesised DNA and assess the absence of lesions that would lead to hereditary abnormalities (Fig. 17.7). These proteins are serine/threonine kinases whose deleterious mutations lead to some diseases that are characteristic of DNA fragility such as *ataxia telangiectasia*: ATM, ATR and DNAPK. These kinases activate by phosphorylation the checkpoint kinases CHK1 and CHK2; ATM preferential substrate is CHK1 and ATR preferential substrate is CHK2. CHK1 and CHK2 can phosphorylate in turn the CDC25 phosphatases (CDC25A for CHK1 and CDC25B/C for CHK2), but these phosphorylations inhibit CDC25 activity, preventing thus the activation of the cyclin–CDK complexes required for cell cycle progression through the G1 → S transition or through the G2 → M transition. CHK2 activates in addition the inhibitory kinase WEE1, which reinforces the inhibition of CDK1 and maintains the cells in G2 phase. In addition, CHK2 inhibits PLK1 and prevents thus the realisation of the early mitotic events, since PLK1 is involved in centrosomes formation, chromosomes separation and mitotic spindle set up.

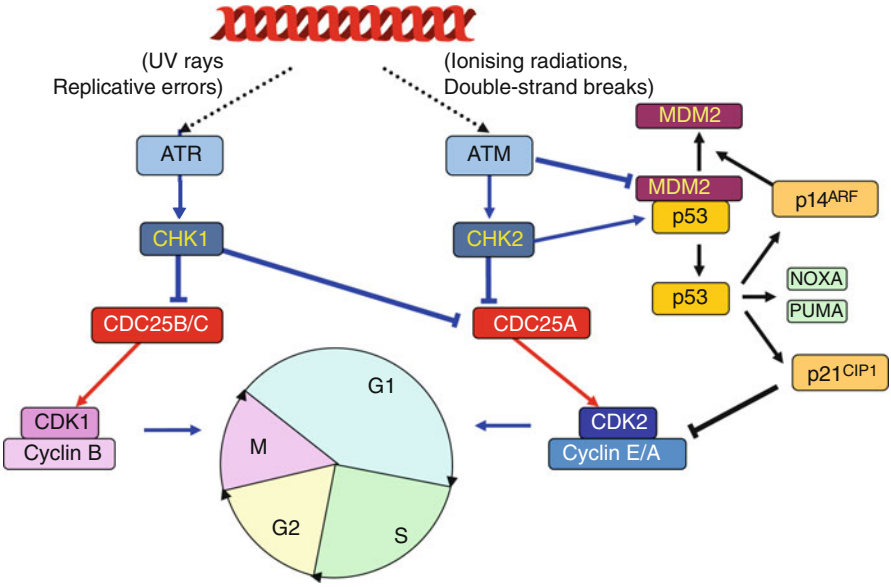


Fig. 17.7 Cell cycle checkpoints and control of DNA integrity. CHK1 and CHK2 kinases, which hinder cell cycle progression by phosphorylating and inhibiting CDC25 phosphatases, are activated by ATM and ATR kinases, which control DNA integrity and are activated mostly by single-strand breaks for ATR or by double-strand breaks for ATM. In addition, CHK2 phosphorylates and activates p53 while ATM removes p53 from MDM2 binding, the E3 ubiquitin ligase that normally leads p53 to the proteasome. In response to DNA damage, p53 can exert its function of transcription factor and drive cells to cell cycle arrest by activating the transcription of $p21^{CIP1}$, an inhibitor of the cyclin D–CDK4/6 complexes, or to apoptosis by activating the transcription of the BH3-only proteins, PUMA and NOXA

CHK2 is also able to activate by phosphorylation the transcription factor p53, which is thus at the crossroads between the information received from DNA integrity and the implementation of the pathways aiming at preventing the propagation of mutagenic DNA damage. p53 (gene *TP53*) is normally present in the cell at very low concentrations, owing to its binding to its E3 ubiquitin ligase, MDM2 (*murine double-minute homologue*), which drives it to the proteasome. p53 activation by CHK2-mediated phosphorylation releases it from MDM2 binding and enables it to play, after homotetramerization, its role of transcription factor. More than hundred genes are transcriptionally activated by p53. This protein, of utmost importance in oncology as a tumour suppressor, is able to delay cell cycle progression in the G1 phase, by stimulating the production of $p21^{CIP1}$, a CKI that inhibits cyclin D–CDK4/6 complexes. It is also able to induce apoptosis through the transcription of the BH3-only proapoptotic proteins PUMA and NOXA (Chap. 18). p53 availability is reinforced by the fact that MDM2 is directly inactivated by phosphorylation by ATM. MDM2 is also inactivated by an alternative product of the *CDKN2A* gene, which also encodes a CKI, $p16^{INK4a}$. This alternative protein, ARF (*alternate reading frame*) or $p14^{ARF}$, has no structural similarity with $p16^{INK4a}$, but is involved in the same tumour suppressor function. $p14^{ARF}$ production is activated by p53, which inhibits therefore its own MDM2-driven proteasomal destruction.

17.4 Oncogenic Alterations in Cell Cycle Control

The delicate machinery of cell cycle is susceptible to be altered in many ways that may contribute to the loss of its control and to oncogenesis. Numerous proteins that regulate cell cycle progression can behave as oncogenes products or as tumour suppressors. Without drawing up an exhaustive list, one can consider as proto-oncogenes the genes encoding CDK4 and CDK6, whose amplification is noticed in sarcomas, or cyclin D1, whose constitutive (and no longer cyclic) expression follows the t(11;14) translocation occurring in certain types of lymphomas. CDC25 genes also appear as oncogenic, owing to their overexpression in various cancer types, associated with poor prognosis. MDM2 also behaves as an oncoprotein: an amplification of the *MDM2* gene, which is located on the same amplicon as CDK4 (chromosome 12q14-15), occurs in liposarcomas and other sarcomas. Without being directly oncogenic, the aurora and the polo-like kinases are often overexpressed in cancers.

One can consider as tumour suppressors $p16^{INK4a}$ and $p14^{ARF}$, the gene they share being mutated or invalidated in many leukaemias and solid tumours, as well as the other members of the INK4 or the CIP/KIP families. Loss of function of RB1 is associated to various cancer types. This protein is in a crucial gatekeeper for cell engagement into DNA replication. Its germinal mutations are characteristic of the familial form of retinoblastoma. CHK1 and CHK2, as well as the proteins that activate these checkpoint kinases, ATM, ATR and DNAPK, behave as tumour suppressors of the caretaker type and protect the cell against inopportune reproduction when DNA lesions have not yet been cleared. Last but not least, p53 is the protein whose

invalidating mutations are the most frequent in human cancers; p53 mutations are encountered in about 50 % of all cancers and generally occur as a late and pejorative event in cancer evolution. These mutations are extremely varied and their consequences may differ according to the alteration of the process of promoter recognition by the altered transcription factor. Among many other effects, p53 mutations prevent it to induce cell cycle arrest (via $p21^{CIP1}$) or apoptosis induction (via NOXA and PUMA) in response to DNA damage: the resulting genomic instability is a factor of acceleration of cancer progression.

As already mentioned, there are two major checkpoints of cell cycle progression; their oncogenic alterations enable the cells to pursue their proliferation despite the presence of DNA lesions, which should lead to cell cycle arrest or cell death:

- The G1 → S transition, which is especially under the control of the INK4 and the CIP/KIP family; since $p21^{CIP1}$ transcription is dependent upon p53, this protein appears as the main factor responsible for this checkpoint.
- The G2 → M transition, which is mainly under the control of WEE1 and MYT1, whose inhibition allows to relieve the mitosis block.

17.5 Pharmacological Targets

Cell cycle control appears, therefore, as a major element of the regulation of cell proliferation, and its alteration contributes to oncogenesis. Because of their motor role in cell cycle progression and of their pro-oncogenic properties, CDKs are considered as valuable pharmacological targets for cancer treatment, and several laboratories are devoted to the identification of specific inhibitors of their kinase activity. Screening of natural products has gathered numerous molecules of potential interest, belonging to various chemical families: purines, pyrido-pyrimidines, indirubines, paullones, etc. However, these molecules are exceptionally specific of a single CDK, and the problem is indeed recurrent for serine/threonine kinase inhibitors. In addition, it appears that the various CDKs can replace each other in controlling the successive steps of cell cycle, with the exception of CDK1, which is not replaceable. More than a dozen of molecules have entered clinical trials, such as roscovitine (seliciclib), palbociclib or dinaciclib, which bear a purine moiety with various heterocyclic substitutions.

CDC25s also constitute interesting potential targets for the identification of antiproliferative molecules. Several chemical classes of phosphatase inhibitors, especially dicoumarins, are being evaluated. The checkpoint kinases CHK1 and CHK2 can also be targeted by kinase inhibitors, with the aim of sensitising the cells to the anticancer agents that induce DNA damage: these agents elicit an accumulation of cells in G1 and/or in G2, which limits their effects when the checkpoints of cell cycle progression are operational. Similarly, targeting the WEE1 kinase could prove to be a valuable approach, because it strictly controls the G2 checkpoint.

The aurora and polo-like kinases, which control the execution of the crucial phases of mitosis, may also be targeted with appropriate kinase inhibitors, which could inhibit mitosis triggering or progression through its successive phases. However, the compounds under study appear to behave as mitotic poisons (such as vinca-alkaloids or taxanes) rather than as targeted therapies. Several compounds have entered clinical trials, either targeting both AURKA and AURKB (tozasertib, danusertib) or presenting some specificity for one of the two kinases. Similarly, PLK1 inhibitors have also been developed, such as volasertib or rigosertib.

It is always difficult to target tumour suppressors, since this would require the restoration of a lost function and not to inhibit an excessive catalytic action; however, high-throughput screening of molecules able to inhibit p53–MDM2 binding has allowed the identification of a novel class of molecules, nutlins, that interfere with this association and stimulate the proapoptotic functions of p53. They are only active on cells not harbouring p53-invalidating mutations. Other types of p53 reactivating approaches are being developed.

The execution of the programmes of cell cycle progression can also be inhibited during the various phases and represent fundamental targets for cancer treatment, and numerous drugs targeting the DNA replication or mitosis programmes have been used for many years in cancer therapy. In that case, all proliferating cells are the targets of such drugs, and not only those that have undergone an oncogenic alteration of the control processes topoisomerase inhibitors and antimetabolites, which hinder DNA replication (S phase), and spindle poisons that counteract tubulin polymerisation (vinca-alkaloids) or inhibit microtubule depolymerisation (taxanes)—belong to the basic anticancer armamentarium.

Further Reading

Antoni L, Sodha N, Collins I, Garrett MD. CHK2 kinase: cancer susceptibility and cancer therapy – two sides of the same coin? Nat Rev Cancer. 2007;7:925–36.

Archambault V, Glover DM. Polo-like kinases: conservation and divergence in their functions and regulation. Nat Rev Mol Cell Biol. 2009;10:265–75.

Atherton-Fessler S, Hannig G, Piwnica-Worms H. Regulation of Cdc2 activity by phosphorylation at T14/Y15. Prog Cell Cycle Res. 1996;2:99–105.

Bassett EA, Wang W, Rastinejad F, El-Deiry WS. Structural and functional basis for therapeutic modulation of p53 signaling. Clin Cancer Res. 2008;14:6376–86.

Bell SP, Dutta A. DNA replication in eukaryotic cells. Annu Rev Biochem. 2002;71:333–74.

Boonstra J. Progression through the G1-phase of the on-going cell cycle. J Cell Biochem. 2003;90:244–52.

Boutros R, Lobjois V, Ducommun B. CDC25 phosphatases in cancer cells: key players? Good targets? Nat Rev Cancer. 2007;7:495–507.

Brown CJ, Lain S, Verma CS, Fersht AR, Lane DP. Awakening guardian angels: drugging the p53 pathway. Nat Rev Cancer. 2009;9:862–73.

Carassou P, Meijer L, Le Moulec S, Aoun J, Bengrine-Lefèvre L. Le cycle cellulaire et ses cibles: inhibition des CDK. Bull Cancer. 2012;99:163–71.

Chen HZ, Tsai SY, Leone G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. Nat Rev Cancer. 2009;9:785–97.

Cheok CF, Verma CS, Baselga J, Lane DP. Translating p53 into the clinic. Nat Rev Clin Oncol. 2011;8:25–37.
Dai Y, Grant S. New insights into checkpoint kinase 1 in the DNA damage response signaling network. Clin Cancer Res. 2010;16:376–83.
Dar AA, Goff LW, Majid S, Berlin J, El-Rifai W. Aurora kinase inhibitors–rising stars in cancer therapeutics? Mol Cancer Ther. 2010;9:268–78.
Diallo A, Prigent C. Les sérine/thréonine kinases contrôlant la progression du cycle cellulaire comme cibles thérapeutiques. Bull Cancer. 2011;98:1335–45.
Doree M, Galas S. The cyclin-dependent protein kinases and the control of cell division. FASEB J. 1994;8:1114–21.
Fisher RP. Secrets of a double agent: CDK7 in cell-cycle control and transcription. J Cell Sci. 2005;118:5171–80.
Grana X, Garriga J, Mayol X. Role of the retinoblastoma protein family, pRB, p107 and p130 in the negative control of cell growth. Oncogene. 1998;17:3365–83.
Hayles J, Nurse P. A review of mitosis in the fission yeast Schizosaccharomyces pombe. Exp Cell Res. 1989;184:273–86.
Jones SM, Kazlauskas A. Growth factor-dependent signaling and cell cycle progression. Chem Rev. 2001;101:2413–23.
Kaldis P. Another piece of the p27Kip1 puzzle. Cell. 2007;128:241–4.
Lapenna S, Giordano A. Cell cycle kinases as therapeutic targets for cancer. Nat Rev Drug Discov. 2009;8:547–66.
Lens SM, Voest EE, Medema RH. Shared and separate functions of polo-like kinases and aurora kinases in cancer. Nat Rev Cancer. 2010;10:825–41.
Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer. 2009;9:153–66.
Masui Y. Towards understanding the control of the division cycle in animal cells. Biochem Cell Biol. 1992;70:920–45.
Nilsson I, Hoffmann I. Cell cycle regulation by the Cdc25 phosphatase family. Prog Cell Cycle Res. 2000;4:107–14.
Obaya AJ, Sedivy JM. Regulation of cyclin-Cdk activity in mammalian cells. Cell Mol Life Sci. 2002;59:126–42.
O'Connor MJ, Martin NM, Smith GC. Targeted cancer therapies based on the inhibition of DNA strand break repair. Oncogene. 2007;26:7816–24.
Peters JM. The anaphase-promoting complex: proteolysis in mitosis and beyond. Mol Cell. 2002;9:931–43.
Prosperi E, Stivala LA, Scovassi AI, Bianchi L. Cyclins: relevance of subcellular localization in cell cycle control. Eur J Histochem. 1997;41:161–8.
Schwartz GK, Shah MA. Targeting the cell cycle: a new approach to cancer therapy. J Clin Oncol. 2005;23:9408–21.
Stark GR, Taylor WR. Control of the G2/M transition. Mol Biotechnol. 2006;32:227–48.
Stolz A, Ertych N, Bastians H. Tumor suppressor CHK2: regulator of DNA damage response and mediator of chromosomal stability. Clin Cancer Res. 2011;17:401–5.
Taylor WR, Stark GR. Regulation of the G2/M transition by p53. Oncogene. 2001;20:1803–15.
Vidal A, Koff A. Cell-cycle inhibitors: three families united by a common cause. Gene. 2000; 247:1–15.

Apoptosis Induction and Regulation

18

Abstract

Apoptosis is one of the main modalities of cell death, so important that it has outshined all others that only begin to be rediscovered. Apoptosis is an active, programmed way of cell death, implemented in response to intracellular or extracellular signals. Intracellular signals can originate from DNA lesions, mitosis defects, oxidative stress or other stresses, while extracellular signals correspond to death messages emitted by other cells. Apoptosis leads to the activation of proteases, called caspases, able to hydrolyse the cell constituents independently from the proteasome. Apoptosis plays many fundamental roles during embryology and tissue homoeostasis; to provide only one example, it is responsible for thymus regression during the passage from childhood to adulthood. Many authors have considered apoptosis as the obligate pathway to cell death induced by anticancer drugs, but in many cases it is only a phenotypic consequence, secondary to other mechanisms of drug-induced cell death, which only indirectly involve apoptosis. Apoptosis is often opposed to necrosis, but comparisons can also be made with senescence and mitotic cell death.

We present in this chapter the signalling pathways which lead to the execution of the cell death programme, but we neither describe the techniques that allow to detect or quantify apoptosis nor the execution of the apoptosis programme. In the first part, we present the most important proteins involved in apoptosis, from effectors to initiators. We present then the intracellular messages and the 'intrinsic' pathway that drive cells to death after they have undergone diverse types of stress and the extracellular signals that activate the 'extrinsic' pathway and drive cells to death on command. Various oncogenic alterations in apoptosis pathways have been identified, which lead to the identification of pharmacological targets.

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8_18

18.1 Proteins Involved in Apoptosis

18.1.1 Caspases

Caspases are the effector enzymes of apoptosis; these are proteases characterised by the presence of a cysteine residue in their active centre and by the capacity of hydrolysing polypeptide chains at the level of an aspartic acid residue, which explains the name of *casp*ase. Some caspases (class I) are able to activate signalling molecules through proteolysis, such as interleukin 1 (IL1) (Chap. 12). Other caspases (class II) are the apoptosis effector caspases and can hydrolyse and destroy a large variety of cell proteins; caspase 3 (CASP3) is the main class II caspase. Finally, class III caspases, such as caspase 8 and caspase 9, are apoptosis initiator caspases, which activate the class II caspases by proteolysis.

Protein activation by mild proteolysis on definite sites is one of the most important mechanisms of post-translational modulation of protein activity (Annex C). Cell proteins should normally be protected from proteolytic enzymes; most of them exist therefore under inactive forms (procaspases in the case of caspases), which themselves are activated by mild proteolysis. Of course, this requires an initiating event in order to carry out the first proteolytic activity of the cascade. In the case of apoptosis, this event consists in the formation of transient supramolecular edifices, sometimes called *platforms*, which arrange the procaspases, with the help of adapter proteins, so that they will become able to mutually hydrolyse themselves and thus become activated.

As a general mechanism, the activation of effector caspases requires initiator procaspases, adapter proteins enabling them to organise a platform and the effector procaspases awaiting proteolytic activation. In the case of the intrinsic pathway, the initiator procaspase is procaspase 9, and the activation platform is called the *apoptosome*; in the case of extracellular signal-triggered apoptosis, the initiator procaspase is procaspase 8, and the activation platform is called the DISC (*death-inducing signalling complex*). The activation of a procaspase into a caspase involves the elimination of an *N*-terminal prodomain, the break of a *C*-terminal domain that remains non-covalently attached to the rest of the molecule and the dimerisation of two rearranged caspase monomers (Fig. 18.1). Initiator and effector procaspases differ by the structure of their *N*- and *C*-terminal domains; initiator procaspases harbour in their *N*-terminal part specific recognition domains, which are called CARDs (*caspase recruitment domains*) or DEDs (*death effector domains*) as appropriate. All caspases have kept the same amino acid sequence in their active centre (QACXG) and the capacity, one activated and dimerised, to split polypeptide chains immediately ahead an aspartic acid residue. Figure 18.2 presents the two steps of the activation of an initiator procaspase into a caspase and of an effector procaspase into a caspase as well as the ultimate intervention of an activated effector caspase on a target protein.

18.1.2 BCL2 Family Proteins

BCL2 (*B-cell lymphoma 2*) proteins constitute a family of 22 proteins, which are either mitochondrial or able to associate to mitochondria during their activation process. They present sequence homologies explaining why they belong to the same

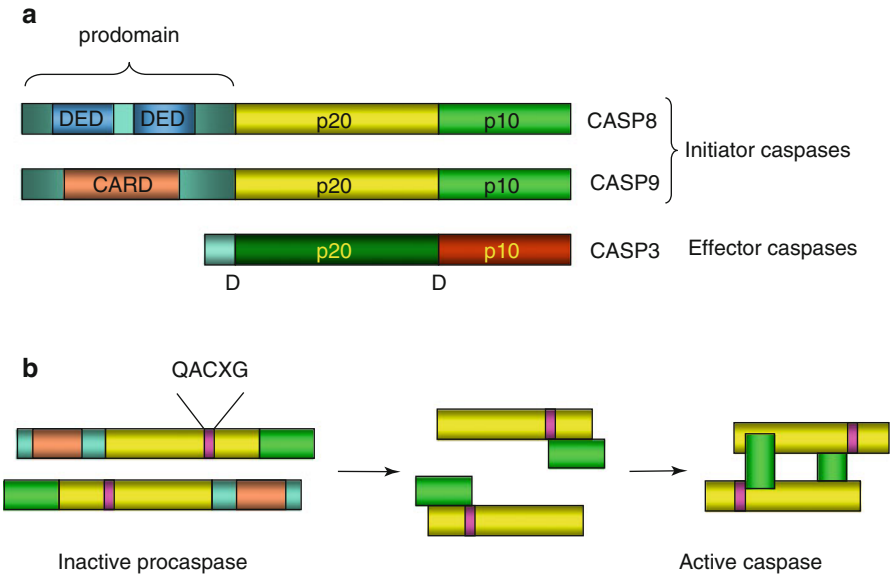


Fig. 18.1 Caspase structure and activation. (**a**) General structure of caspases. All procaspases contain two domains, p20 and p10, the cleavage of which, at the level of an aspartic acid residue D, allows the transition to an active conformation. They also contain a prodomain whose elimination is required for catalytic activity. Initiator procaspases have an especially long prodomain, which bears DED (caspase 8 and caspase 10) or CARD (caspase 9) interaction domains. (**b**) Caspase activation. The example represented here is that of caspase 9. The catalytic site responsible for cleavage is characterised by the QACXG sequence. The arrangement of procaspases in apoptosome or DISC enables the proteolytic cleavage of one molecule by another one at the level of the aspartic acid residue. After prodomain elimination, the rearrangement of p10 and p20 domains and the dimerisation of the protein activate the catalytic function

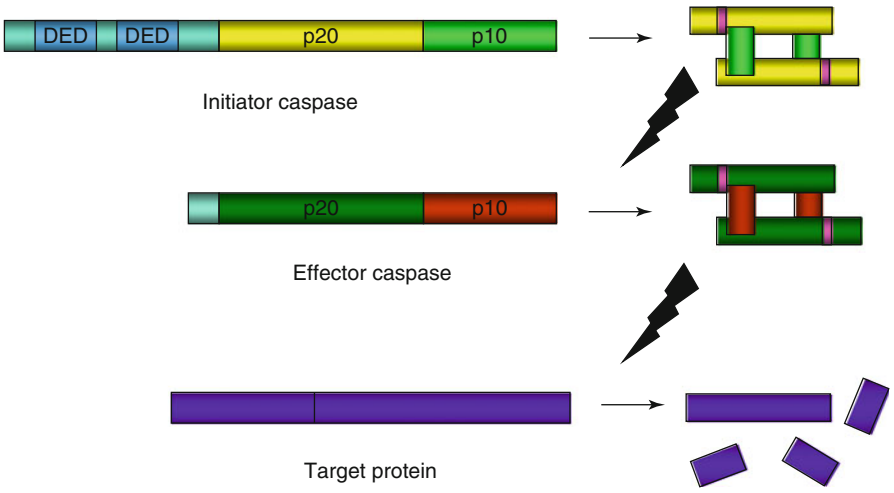


Fig. 18.2 Caspase proteolytic cascade. An initiator caspase (here caspase 8, with two DED domains for the recognition of the adapter proteins FADD and TRADD), which is itself autocatalytically activated, activates in turn an effector procaspase at the level of two aspartic acid residues. The cleavage products, after dimerisation, constitute the active caspase that can hydrolyse in turn various cell proteins, always at the level of an aspartic acid residue

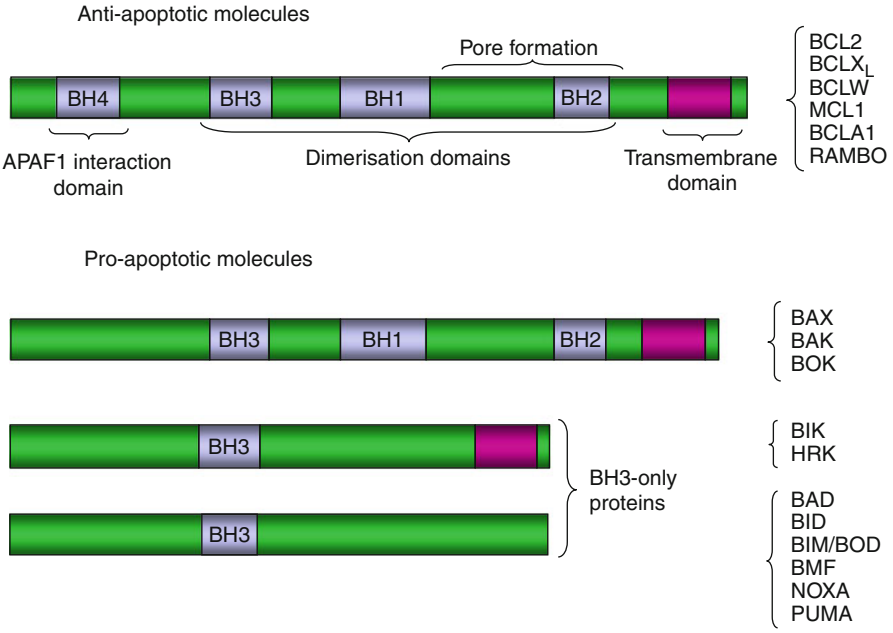


Fig. 18.3 BCL2 family proteins. These proteins can contain a transmembrane domain and up to four special interaction domains, involved in dimerisation (BH1, BH2 and BH3 domains), in the opening of the mitochondrial permeability transition pore (BH1 and BH2 domains) and in the interaction with APAF1 (BH4 domain). The BCL2-type antiapoptotic proteins are inserted in the mitochondrial outer membrane thanks to their *N*-terminal portion bearing the transmembrane domain and mobilise APAF1 in the cytosol through their *C*-terminal portion. The BAX-type proapoptotic molecules also have a transmembrane domain, dimerisation domains and protein–protein interaction domains. The BH3-only proteins have generally no transmembrane domain and are not inserted in the mitochondrial membrane but display mitochondrial tropism due to their BH3 domain, which can interact with the corresponding domains of the other BCL2 family proteins

family, but their functions may be quite different from one to another and it is necessary to divide this family into subfamilies. They are characterised by the presence of specific domains of different types called BH (*BCL2 homology*) domains. Some of them comprise four BH domains (BH1 to BH4) and a transmembrane domain enabling their insertion in the mitochondrial outer membrane. BH1 and BH2 domains are involved in the regulation of the permeabilisation of this membrane; the BH3 domain, in association with BH1 and BH2 domains, is a protein–protein interaction domain allowing their homo- or heterodimerisation; the BH4 domain, essentially intracytoplasmic, is an interaction domain for the adapter protein, APAF1 (*apoptotic peptidase-activating factor 1*). The BCL2 protein subfamilies are the following ones (Fig. 18.3):

- Proteins containing all four BH domains and a transmembrane domain: BLC2; $BCLX_{L/S}$ (gene *BCL2L1*), which are splicing variants with opposite functions; BCLW (gene *BCL2L2*); MCL1 (*myeloid cell leukaemia sequence 1*) or BCL2L3; BCL2A1 or BCL2L5; BCL-RAMBO (gene *BCL2L13*); BNIP3

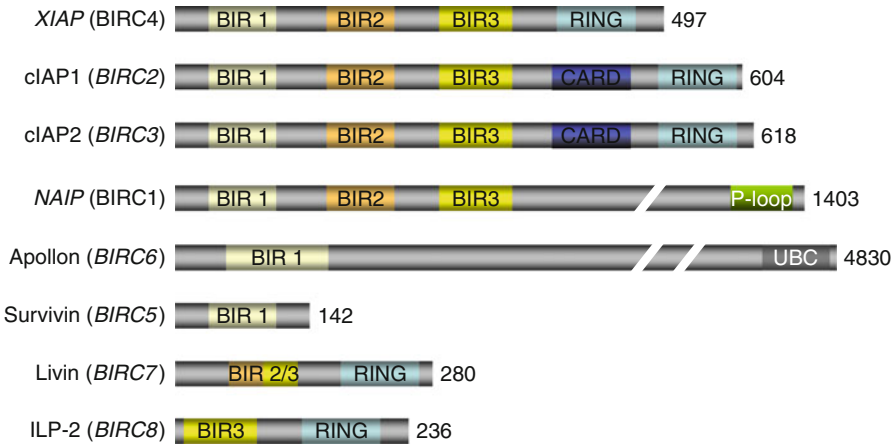


Fig. 18.4 IAP family proteins. These eight proteins with caspase inhibitory function are cytosolic. They harbour various interaction domains, especially a CARD domain enabling their interaction with caspases, several BIR domains, a RING domain for E3 ubiquitin ligase activity or a UBC domain for E2 ubiquitin-conjugating activity. Gene names are in italics. The number of constitutive amino acids in indicated on the right

(*BCL2/adenovirus-interacting protein 3*); and DIVA (*death-inducing vBCL2/Apaf1-binding protein*, gene *BCL2L10*); these proteins are antiapoptotic, except the $BCLX_S$ variant.

- Proteins containing three BH domains (the BH4 domain is lacking): BAX (*BCL2-associated X protein*) or BCL2L4, BAK (*BCL2 homologous antagonist/killer*) or BCL2L7 (gene *BAK1*), BCL2L12, BCL2L14 and BOK (*BCL2-related ovarian killer*) or BCL2L9; these proteins are proapoptotic.
- Proteins containing only a BH3 domain (*BH3-only proteins*), which are proapoptotic; some of them also contain a transmembrane domain: BIK (*BCL2-interacting killer*) and HRK (*Harakiri BCL2-interacting protein*), while others do not: BAD (*BCL2-associated agonist of cell death*) or BCL2L8; BID (*BH3 domain-interacting death agonist*); BIM (*BCL2-interacting mediator of cell death*), also known as BOD (*BCL2-related ovarian death agonist*) (gene *BCL2L11*); BMF (*BCL2-modifying factor*); BFK (*BCL2 family kin*, gene *BCL2L15*); NOXA (gene *PMAIP1*, *phorbol myristate acetate-induced protein 1*); and PUMA (*p53-upregulated mediator of apoptosis*) (gene *BBC3*, *BCL2-binding component*).

We describe the mechanisms by which these proteins are engaged in the control of apoptosis triggering in the part devoted to the intrinsic apoptosis pathway.

18.1.3 IAP Family Proteins

IAPs (*inhibitor of apoptosis proteins*) or BIRCs (*baculovirus IAP repeat-containing proteins*) constitute a group of eight cytoplasmic caspase inhibitors. These proteins harbour a series of characteristic domains (Fig. 18.4). Some are constant, the BIR1,

BIR2 and BIR3 domains, and allow protein–protein interactions, especially with caspase 3. Others are inconstant, the CARD (*caspase recruitment domain*) domains, which allow their interaction with initiator caspases or adapter proteins, and the RING (*really interesting new gene*) domains, which display an E3 ubiquitin ligase activity toward caspases, driving them to the proteasome and adding caspase destruction to caspase inhibition. In addition to their role in apoptosis inhibition, these proteins play an important role in cytodieresis, at the end of mitosis (Chap. 17), and in the regulation of NFκB pathway (Chap. 12).

The proteins equipped with a CARD domain, and therefore able to interact with CASP8, are cIAP1 and cIAP2 (*cellular inhibitor of apoptosis protein*) (genes *BIRC2* and *BIRC3*, respectively). Together with XIAP (*chromosome X inhibitor of apoptosis protein*), alias BIRC4 or ILP1, they also have a RING domain, indicating their function of E3 ubiquitin ligase toward caspases. Apollon or bruce (gene *BIRC6*) contains in contrast a ubiquitin-conjugating (UBC) domain and functions as an E2 ubiquitin-conjugating enzyme. NAIP (*NLR family apoptosis inhibitory protein*), alias BIRC1, seems to be confined to neural cells, preventing them from apoptosis. Survivin (gene *BIRC5*), livin (gene *BIRC7*) and ILP2 (*IAP-like protein 2*) (gene *BIRC8*) are involved in cancer cell apoptosis; livin and ILP2 harbour a RING domain.

Most of these proteins are themselves inhibited by a mitochondrial protein, SMAC (*second mitochondria-derived activator of caspase*) (gene *DIABLO*, *direct inhibitor of apoptosis-binding protein with low pI*). This protein acts as a dimer and binds the BIR3 domains of the IAP, so that their binding to caspases is inhibited.

18.1.4 Death Receptors and Their Ligands

Death signals are brought by proteins that are members of the TNF superfamily (TNFSF). TNFSF members that are type II transmembrane proteins (intracellular *N*-terminal extremity) can be cleaved and generate diffusible ligands. They recognise, on the plasma membrane of the target cell, cognate receptors of the TNF receptor superfamily (TNFRSF). There exists a total of about 17 TNFSF members and 26 TNFRSF members; only those involved in cell death induction are studied here. Three TNFSF members are involved in apoptosis signalling: FASL (*FAS ligand*) or TNFSF6 (gene *FASLG*), TNF (*tumour necrosis factor*) or TNFSF2 and TRAIL (*TNF-related apoptosis-inducing ligand*) (gene *TNFSF10*).

The TNFRSF members that recognise these death signals are called *death receptors*. These are type I transmembrane proteins that harbour an extracellular ligand-binding domain, a single transmembrane domain and an intracellular signalling domain (Fig. 18.5). Death receptors contain a *death domain* (DD), which enables their interaction with an adapter protein harbouring a homologous domain; there are four main death receptors: FAS (*fragment for apoptosis stimulation*), also known as CD95 or TNFRSF6; TRAILR1 or DR4 (gene *TNFRSF10A*); TRAILR2 or DR5 (gene *TNFRSF10B*); and TNFR1 or DR1 (gene *TNFRSF1A*). These receptors

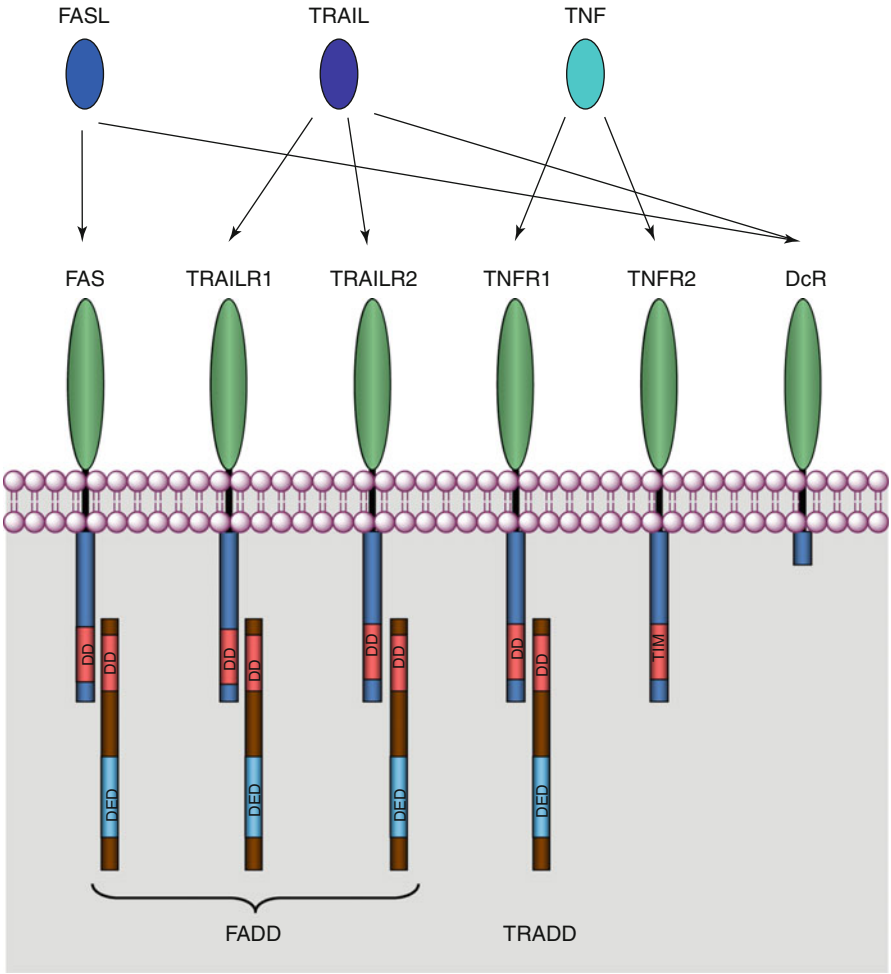


Fig. 18.5 Cell death receptors, ligands and adapter proteins. The ligands of the TNF superfamily (FASL, TNF and TRAIL) recognise and bind receptors of the TNFR superfamily (FAS, TNFR1, TRAILR1 and TRAILR2) as well as decoy receptors (DcR) devoid of binding domains with an adapter protein. Many other ligands members of the TNF superfamily can bind cognate receptors of the TNFR superfamily outside the field of apoptosis

recognise the extracellular ligands FASL, TRAIL and TNF, respectively. FAS, TRAILR1 and TRAILR2 bind to the same intracellular adapter protein, FADD (*FAS-associated death domain protein*), whereas TNFR1 binds to the adapter protein TRADD (*TNF receptor-associated death domain protein*). Both adapter proteins can activate an initiator protein, caspase 8.

In addition, some receptors harbour no intracellular binding domain and transduce no message. They have either a very short intracellular domain or are simply attached to the membrane by a GPI anchor. These are *decoy receptors*, DcR1 or

TRAILR3 (gene *TNFRSF10C*), DcR2 or TRAILR4 (gene *TNFRSF10D*) and DcR3 (gene *TNFRSF6B*), which are able to block the signals brought by TRAIL (for DcR1 and DcR2) or FAS (for DcR3) and divert the information received.

We describe the mechanisms by which these proteins are engaged in the control of apoptosis triggering in the part devoted to the extrinsic apoptosis pathway. A variety of other proteins also harbour interaction domains with the adapter proteins FADD and TRADD; some are found in the caspase family and in the IAP family and have been already mentioned; others will be described with the extrinsic apoptosis pathway.

18.2 Intrinsic (Mitochondrial) Apoptosis Pathway

18.2.1 Activation of the Intrinsic Pathway

The activation of the intrinsic apoptosis pathway depends upon the emission of signals that are sent to the mitochondria; the mitochondria integrate these signals and elaborate a response which leads to the activation of an initiator caspase, caspase 9, from which ensues the activation of the effector caspases, caspase 3 and caspase 7 (Fig. 18.6). The signals that reach the mitochondria consist of a variety of proteins of the BCL2 family called the *BH3-only* proteins, which are produced in response to various stresses by several types of mechanisms. Some BH3-only

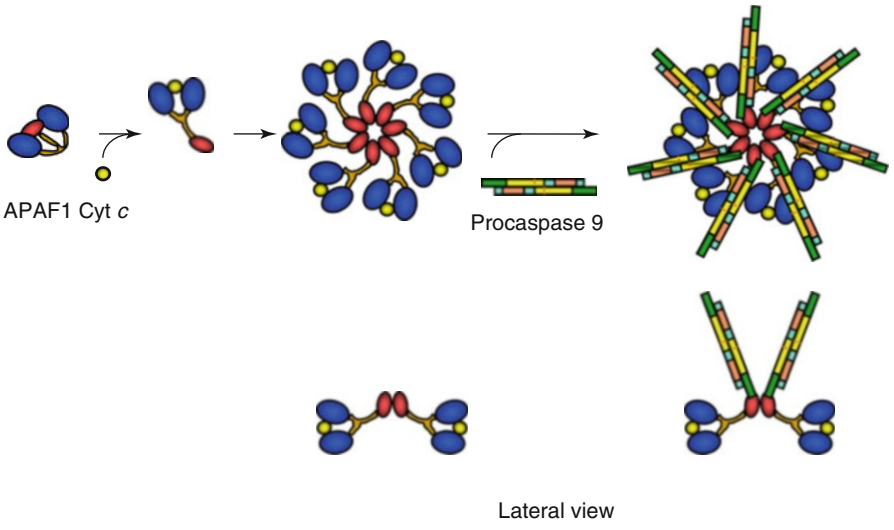


Fig. 18.6 Formation and activation of the apoptosome. APAF1, which is folded at the basal state, is able, in the presence of cytochrome *c*, to undergo a conformation change and form a heptameric structure, apoptosome, to which are associated procaspase 9 molecules. The APAF1 CARD domains can interact with the corresponding domains of caspase 9, allowing autocatalytic activation of caspase 9

proteins are simply regulated by transcriptional control, and others undergo post-translational modifications that govern their activity. DNA damage, which enables p53 activation (Chap. 17), leads in particular to the transcription of the genes encoding PUMA and NOXA. Similarly, microtubule damage leads to the transcription of the genes encoding BIM and HRK. In contrast, BAD deactivation is induced by AKT-mediated phosphorylation (Chap. 3) in response to growth factor signalling, and BID is activated by proteolysis carried out by caspase 8, the initiator caspase of the extrinsic pathway.

BH3-only proteins are able, thanks to their BH3 domain, to interact at the level of the mitochondria with other BCL2 family proteins equipped with transmembrane domains allowing their mitochondrial membrane insertion. These interactions between BCL2 proteins control the permeabilisation of the mitochondrial outer membrane; the proapoptotic proteins BAX and BAK are able to form oligomers constituting transmembrane channels or inducing its permeabilisation through lipid–protein interactions. The exact mechanisms used by antiapoptotic (BCL2 type) and proapoptotic proteins (BAX type) to regulate the permeability of the mitochondrial outer membrane are not fully understood and remain still controversial.

The 'rheostat theory' is based upon the equilibrium between these two types of proteins, which are subjected to homo- or heterodimerisation processes. In this scenario, the abundance of BCL2-type proteins favours BCL2–BCL2 dimers and prevents permeabilisation, whereas the abundance of BAX-type proteins favours BAX–BAX dimers and promotes mitochondrial permeabilisation, the BH3-only proteins interacting with the corresponding binding domains of BCL2 or BAX to orientate toward one type of dimerisation or the other. All BH3-only proteins do not have the same direct activator effect on BAX-type proteins (or inhibitor effect on BCL2-type proteins): there exists some specificity for their action. BID, BIM and PUMA interact with all BCL2-type proteins, whereas BAD interacts only with BCL2, $BCLX_L$ and BCLW, and NOXA with BCLA1 and MCL1. In addition, there is also some tissue specificity in BCL2-type (and possibly BAX-type) protein expression. These specific features of the interactions between the inducers of intracellular apoptosis (BH3-only proteins) and the mitochondrial proteins in charge of regulating mitochondrial membrane permeability (BCL2-type and BAX-type proteins) enable the adaptation of apoptosis implementation as a function of the cellular context.

There should be complementary hypotheses explaining the permeabilisation of the mitochondrial outer membrane, which implies the regulation of opening the mitochondrial permeability transition pore (PTP), also called the *megachannel*. PTP is a protein complex forming a channel of about 3 nm through the inner and outer mitochondrial membranes, which enables a transient direct communication between the mitochondrial matrix and the cytoplasm. It is comprised of several proteins, among which two membrane transporters that establish a contact, VDAC (*voltage-dependent anion channel*) at the level of the outer membrane and ANT (*adenine nucleotide translocator*) at the level of the inner membrane, in association with cyclophilin D (CYPD, gene *PPIF*, *peptidylprolyl isomerase F*) in the matrix. PTP is opened and contributes to mitochondrial membrane permeabilisation when the

electrochemical transmembrane potential is lost (oxidative stress, alterations of calcium homoeostasis), since in the usual conditions the inner membrane does not play any role in the permeabilisation of the mitochondrial outer membrane.

18.2.2 Cell Death Induced by the Intrinsic Pathway

Whatever the sequence of the events leading to this process, and the respective roles of the PTP and BCL2 family proteins, mitochondrial membrane permeabilisation enables the exit, from the intermembrane space, of several proteins that are known under the generic name of AIFs (*apoptosis-inducing factors*) (Fig. 18.6). The most important AIF is cytochrome *c*, an electron transporter of the mitochondrial chain. The release of cytochrome *c* out of the mitochondria represents the triggering event of apoptosis. It is responsible for the organisation, in the cytoplasm, of a 'platform' allowing the activation of procaspase 9 into active caspase 9 by proteolysis, according to an autocatalytic mechanism. This platform, called *apoptosome* (Fig. 18.7), is a transient supramolecular edifice following a seven-order symmetry and associates seven molecules of cytochrome *c*, seven molecules of an adapter protein called

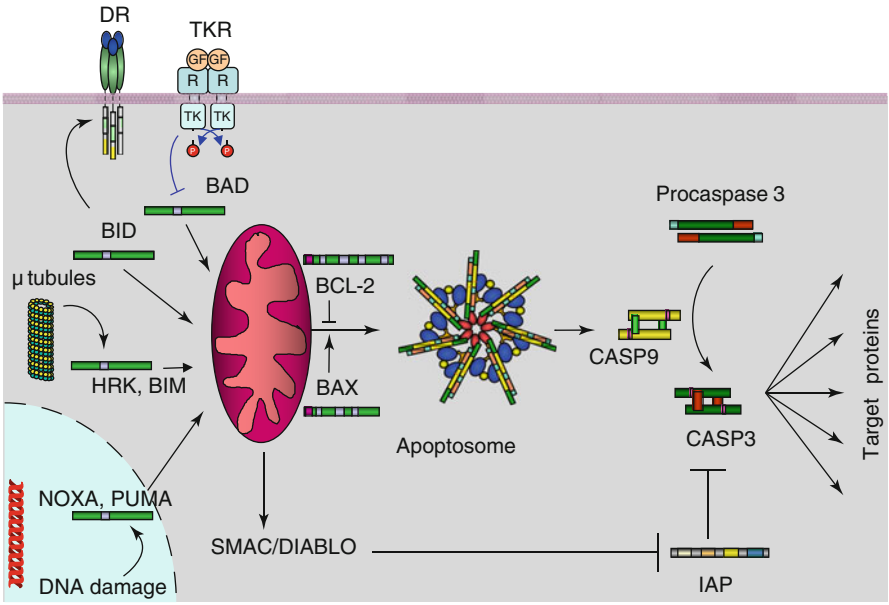


Fig. 18.7 Intrinsic (mitochondrial) apoptosis pathway. Various signals induced by intracellular or extracellular stresses are brought to mitochondria by BH3-only proteins and interact with the mitochondrial BCL2 family proteins, favouring the opening of the transition pore. The release of cytochrome *c* out of the mitochondria enables its interaction with APAF1, generating with procaspase 9 and dATP the heptameric structure called apoptosome. This results in the proteolytic autoactivation of procaspase 9 into active caspase 9, which in turn activates procaspase 3 into caspase 3 by proteolysis. Caspase 3 becomes thus able to hydrolyse its target proteins

APAF1 (*apoptotic peptidase-activating factor 1*), seven molecules of dATP and seven molecules of procaspase 9. This supramolecular structure enables a special disposition of the molecules of procaspase 9, such that they can mutually cleave each other and generate active caspase 9, which is then able to activate procaspase 3 into caspase 3, as previously described.

Another important factor able to exit out of mitochondria when it becomes permeable is SMAC. This protein is an inhibitor of the proteins that inhibit caspases (the IAPs). SMAC reinforces apoptosis induction ensured by cytochrome *c* exit from mitochondria by inhibiting the effector caspases of the cell death programme. Other AIFs are susceptible to be released by mitochondria during outer membrane permeabilisation, such as HTRA2 (*high temperature requirement protein A2*). Finally, some reactive oxygen species (ROS) can also leave mitochondria in the same conditions: the superoxide ion $O_2^{-\bullet}$, the radical $OH^{\bullet}$ and hydrogen peroxide H_2O_2. These toxic species are produced from molecular oxygen (Chap. 16), used in the electron transport chain and generate an oxidative stress within the cell.

18.3 Extrinsic (Death Receptor) Apoptosis Pathway

18.3.1 Activation of the Extrinsic Pathway

The apoptosis-inducing ligands described above, FASL, TNF and TRAIL, are able to specifically recognise their cognate receptors FAS, TNFR1 and TRAILR1 and TRAILR2, in addition to decoy receptors (Fig. 18.5). TNF can also bind TNFR2 and activate a distinct type of pathway described below. The ligands are originally trimeric transmembrane proteins whose *C*-terminal part is localised outside the cell; this part can be cleaved to generate the soluble forms, called 'death factors', by analogy with growth factors. Ligand–receptor interaction can occur in a juxtacrine way, when both partners are embedded in the plasma membrane, or in a paracrine way when the ligand has been cleaved and released in the extracellular space. The receptors harbour extracellular ligand-binding domains and intracellular domains able to transduce the death signals. However, the decoy receptors titrate the death signals without transducing the signal received and display antiapoptotic functions. The relative expression of true and decoy receptors modulates the death message brought by the ligands. Ligand binding to its cognate receptor induces its trimerisation. However, some receptors are already pre-associated as oligomers thanks to PLAD (*pre-ligand assembly domain*), so that the ligand does not actually induce trimerisation but rather stabilises the receptor trimers. Death receptor trimerisation appears as analogous to growth receptor dimerisation described in Chap. 1.

Receptor activation by trimerisation allows the homophilic interaction between their death domains (DD) and the corresponding death effector domains (DED) of the adapter proteins, FADD and TRADD, FADD binding to FAS, TRAILR1 and TRAILR2 and TRADD binding to TNFR1. The DED domains of the adapter proteins are in turn recognised by the corresponding domains of procaspase 8, so that a molecular edifice is constituted, a platform called DISC (*death-inducing signalling*

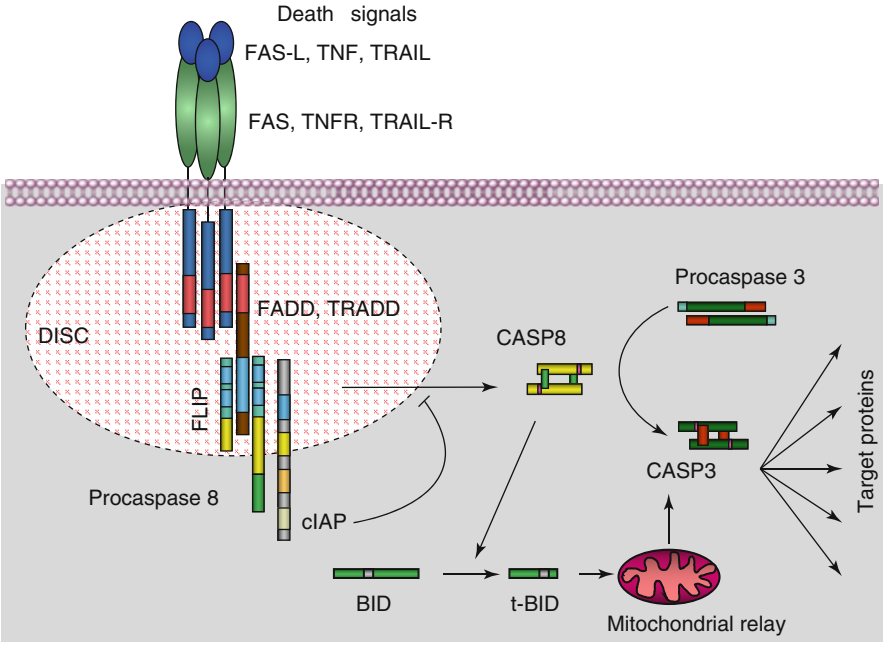


Fig. 18.8 Extrinsic (death receptors) apoptosis pathway. Various signals brought by cell death ligands (FASL, TNF, TRAIL) activate membrane receptors (FAS, TNFR, TRAILR1 and TRAILR2) by inducing their trimerisation. The intracytoplasmic portion of these receptors contains a death domain DD, which allows the recruitment of adapter proteins (FADD, TRADD) equipped with an analogous death effector domain (DED). Procaspase 8 molecules, also harbouring DEDs, are then recruited and constitute with the adapter proteins a supramolecular edifice called the DISC (*death-inducing signalling complex*). By autocatalytic proteolysis, procaspase 8 is activated into caspase 8, which in turn activates procaspase 3 into caspase 3. Caspase 3 becomes thus able to hydrolyse its target proteins. Caspase 8 is also able, in some cell types, to activate by proteolysis the BH3-only protein BID which, under its truncated t-BID form, interacts with BCL2 proteins at the level of the mitochondria, intensifying thus the extrinsic apoptotic signalling by a mitochondrial contribution. FLIP (*FADD-like inhibitory protein*) is homologous to caspase 8, with DED domains, but is devoid of caspase activity and acts as an inhibitor of apoptosis. cIAP1 and cIAP2 can inhibit caspase 8 and therefore the induction of apoptosis

complex), which plays in the vicinity of the plasma membrane the role played by apoptosome in the vicinity of mitochondria (Fig. 18.8). Several DISCs are associated together at the level of membrane lipid rafts and form macroaggregates.

18.3.2 Cell Death Induced by the Extrinsic Pathway

The activation of procaspase 8 in the DISC is performed by autocatalytic cleavage, which leads to the active form of the molecule. Caspase 8 is then able to activate procaspase 3 into caspase 3 as already described. In some cell types, caspase 8 is also able to activate a BH3-only protein, BID, called t-BID once truncated by

caspase 8. As other BH3-only proteins, t-BID interacts with BCL2 proteins and stimulates the permeabilisation of the mitochondrial outer membrane allowing cytochrome *c* exit from the mitochondria. This crosstalk between the extrinsic and intrinsic apoptosis pathways enables the amplification of the death messages; the cells that use this bypass are said to be 'type II' cells, by contrast with 'type I' cells, which do not.

In addition to the intracellular portion of the receptor, the adapter protein and the initiator caspase, another partner, may participate to the DISC; it is called FLIP (*FADD-like inhibitory protein*) (gene *CFLAR*, *caspase 8 and FADD-like apoptosis regulator*). This protein exists under a long form, with *N*-terminal DED domains and *C*-terminal sequence homology with caspase 8 (but without the catalytic cysteine residue), and a truncated form on the *C*-terminal end, this post-translational modification being mediated by caspase 8. FLIP acts at a competitive caspase 8 inhibitor.

18.3.3 Alternative Signalling Pathways Induced by TNF Superfamily Ligands

The activation of an initiator caspase and the implementation of apoptosis are not the only consequence of the activation of a death receptor by a ligand. The receptor death domains and the death effector domains of adapter proteins (FADD and TRADD), caspase 8 and FLIP, can be recognised by various proteins involved in cell signalling (Fig. 18.9).

- DAXX (*death-domain associated protein*) interacts in the cytoplasm with FAS via its death domain and activates the JUN *N*-terminal kinase pathway (JNK) via a MAP3 kinase, ASK1 (gene *MAP3K5*). It may in turn be activated by JNK, in a positive feedback loop. DAXX is able to reinforce the apoptosis-inducing function of FAS activation via the JNK pathway.
- RIP (*receptor-interacting protein*) kinases (RIPK1 to RIPK4) harbour CARD domains that can interact with the DED domains of TRADD, caspase 8 and FLIP. Via the MAP3 kinase TAK1, they activate both the IKK (*IκB kinase*) complex to generate NFκB (Chap. 12) and the JNK pathway (Chap. 2) to generate JUN.
- The apoptosis inhibitory proteins cIAP and cIAP2 (genes *BIRC2* and *BIRC3*) also harbour CARD domains enabling them to directly interact with caspase 8 to inhibit its proapoptotic function, as well as with the adapter protein FLIP. They are able to activate NIK (*NFκB-inducing kinase*) and thus the IKK complex, also generating NFκB.

As already mentioned, there is an important number of TNF and TNFR superfamily members, which are involved in various cellular processes such as apoptosis, immunity or inflammation. Death receptors are characterised by the presence of DDs, as described above, but other TNFRSF members, including a second TNF receptor, TNFR2 (gene *TNFRSF1B*), harbour another type of domain called TIM

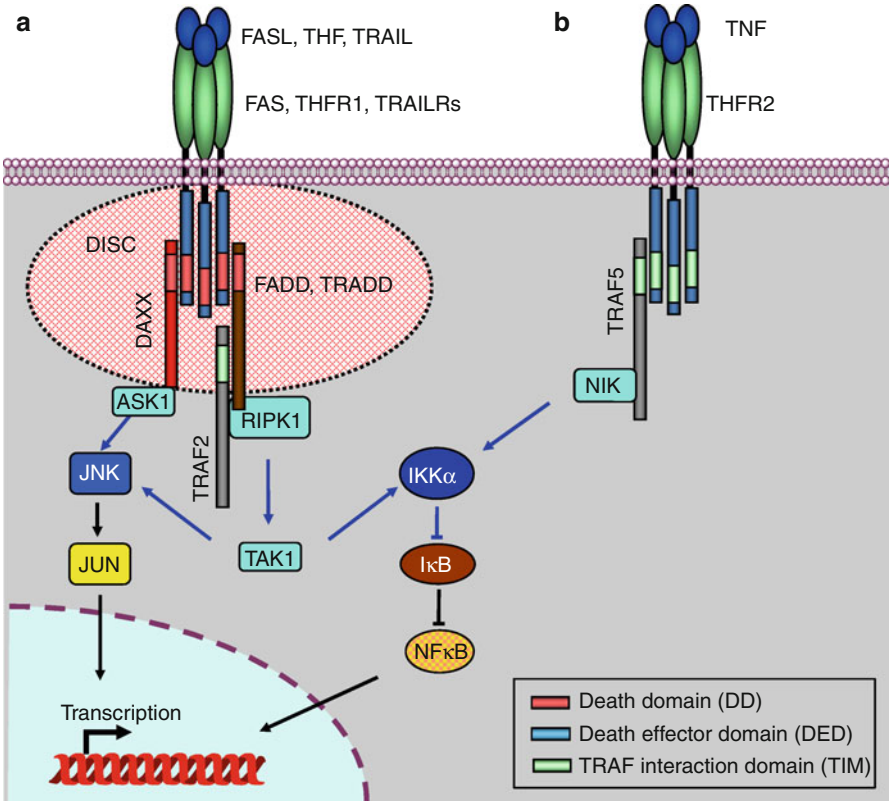


Fig. 18.9 Signalling pathways activated by death receptors and other TNFRSF members. (**a**) Death receptor-activated pathways. In addition to apoptosis induction (Fig. 18.8), these receptors are able to activate several pathways by recruiting adapter proteins equipped with a DD domain, which in turn can recruit effector proteins thanks to their DED or CARD domain: (1) thanks to the adapter protein DAXX, they can activate the JNK pathway via the MAP3 kinase ASK1 and thus reinforce apoptosis induction; (2) thanks to the adapter protein TRADD, with or without the intervention of TRAF2, they can activate RIPK1 (*receptor-interacting protein kinase 1*), which in turn activates by phosphorylation the MAP3 kinase TAK1 and subsequently the JUN kinase (JNK) pathway, but may also activate the IKK complex which generates NFκB. (**b**) Pathways activated by other TNFRSF members such as TNFR2. These receptors do not induce apoptosis, as they harbour no DD domain but a TIM (*TRAF interacting motif*) domain. They recruit the adapter TRAFs (*TNF receptor-associated factors*) such as TRAF5 in the case of TNFR2, which can activate several types of proteins such as NIK (*NFκB-inducing kinase*) and thus generate NFκB. This illustrates the ambivalent role that can be played by TNF superfamily members; these proteins behave as apoptosis inducers or protectors as a function of the cellular context, i.e. the cell equipment in receptors and adapter proteins. The trimerisation of the ligands and receptors is schematised, but not the subsequent trimerisation of their partners

(*TRAF-interacting motif*) and activate various adapter proteins called TRAFs (*TNF receptor-associated factors*). The full description of the ligand–receptor systems and of the pathways activated is beyond the scope of this book, and only an insight on the pathways opened by TRAFs is given below (Fig. 18.9).

TRAFs constitute a family of six proteins (TRAF1 to TRAF6), each of them interacting with various TNFRSF members with variable selectivity. TRAF6 is also involved downstream the activation of IL1 and TLR receptors (Chap. 12). TRAFs work as homo- or heterotrimers formed thanks to the trimerisation of the receptors with which they interact. Most TRAFs harbour a RING domain, suggesting their potential role as E3 ubiquitin ligases, which has been demonstrated only for TRAF6. TRAF2 is able to bind TIM domains of the TNFR2 receptor and of other TNFRSF members, but also the adapter protein TRADD downstream TNFR1 activation, as well as the antiapoptotic molecules cIAP1 and cIAP2. In addition, TRAFs contain domains allowing the activation of signalling molecules such as TANK (*TRAF-associated NFκB activator)* and NIK. These proteins activate the IKK (*inhibitor of NFκB kinase*) complex and thus enable the formation of NFκB (Chap. 12). Another pathway induced by TRAFs is the JNK pathway, via the involvement of a MAP3 kinase, TAK1 (gene *MAP3K7*), which reinforces apoptosis induction by TNFR activation. The same signals are therefore able to generate apoptosis, both through caspase 8 and JNK pathway activation, and survival through NFkB activation. This can be explained by the nature of the TRAF family member expressed in the cell receiving the signal.

18.4 Oncogenic Alterations of Apoptosis Pathways

Numerous alterations in the proteins in charge of apoptosis initiation and control are found in cancer cells. However, the question is not resolved of whether these alterations are among the causes or the consequences of apoptosis, in other words whether these alterations play a driver role in oncogenesis or intervene merely as a facilitator of cancer cell survival without being required for malignant transformation. We will not present here the oncogenic alterations of the inducing mechanisms of apoptosis, such as the mutations of p53, which have been presented in Chap. 17 (control of DNA integrity).

At the level of mitochondrial apoptosis, *BCL2* behaves as a proto-oncogene, especially in malignant lymphomas; its overexpression is most often due to the t(14;18) translocation which relocates it under the dependence of a strong promoter, that of the heavy chains of immunoglobulins. This overexpression is the initiator event of oncogenesis for follicular lymphomas. According to the tissues where they are specifically expressed, other genes of the BCL2 family, such as *MCL1*, may play an oncogenic role. It is not certain, however, that the genes encoding proapoptotic proteins of the BCL2 family, BAX and BAK, are tumour suppressor genes. Deletions and invalidating mutations of *BAX* have nevertheless been found in cancers.

At the level of extrinsic apoptosis, *FAS* mutations have been observed in non-Hodgkin lymphomas of various types, so that this gene is considered as a tumour suppressor gene. In addition, a decrease in expression of death receptors like FAS and an overexpression of decoy receptors are classically observed in

tumours, but the question of whether these changes are simply the consequence of oncogenesis or causal features is raised. It is difficult to assign to TRAIL and its receptors a tumour suppressor role outside some murine tumour models, although mutations have been observed in the TRAIL2 receptor in some human tumours.

Some authors have assigned a tumour suppressor function to initiator caspases, especially CASP8 in neuroblastomas, in which the promoter of its gene is hypermethylated. In contrast, some IAP may play an oncogenic role: *BIRC2* and *BIRC3* (proteins cIAP1 and 2) are amplified in various carcinomas; they can however be deleted in multiple myeloma, loosing thus their property of driving NIK to the proteasome and favouring the formation of the transcription factor NFκB (Chap. 12).

18.5 Pharmacological Targets

Even though there are few driver oncogenic alterations in the proteins involved in the apoptosis pathways, they constitute a potential reserve of targets, either to favour apoptosis induction in cancer cells or to amplify the action of the therapeutic agents that induce apoptosis through DNA damage.

At the level of mitochondrial apoptosis, molecules mimicking BH3-only proteins have been identified and developed up to clinical trials; they can induce apoptosis through binding the antiapoptotic proteins of the BCL2 family. Navitoclax and obatoclax are such BH3 mimetics that are being evaluated. In addition, a strategy targeting the mRNA of the BCL2 gene by antisense oligonucleotides has been conceived and developed up to phase III studies, but this drug, oblimersen, was eventually not approved.

At the level of death receptor-mediated apoptosis, analogues and agonists of the ligands have been identified. It is not possible to use FASL, which is highly hepatotoxic, but TNF itself, under the name of tasonermin, is being used to treat distal metastases of malignant melanomas and limb sarcomas, provided the drug is injected in the isolated perfused limb, because the passage of the molecule in the systemic circulation would be lethal. TNF is in this case generally associated to melphalan, a cytotoxic alkylator. TRAIL can be used in systemic injections, and a recombinant fragment of the protein has been introduced in clinical trials (dulanermin). Agonistic antibodies recognising the TRAIL receptors are also in development, such as mapatumumab for TRAIL1 and lexatumumab and tigatuzumab for TRAIL2; they are able to mimic the ligand action on the receptor and to stabilise and even aggregate the DISCs.

The last potential target is offered by the IAPs; all of them are not ubiquitously expressed, but some of them could be efficiently targeted, especially survivin. An antisense approach has validated the concept. Small molecules able to inhibit XIAP (embeline) or survivin (terameprocol) are in development. Also, SMAC-mimicking molecules have been identified and generically called SMC (*SMAC mimetic compounds*).

18.6 Dependence Receptors

A special variety of receptors is involved in apoptosis induction and is distinct from the classical death receptors: they are called *dependence receptors* or *addiction receptors*. In the presence of their ligand, they transduce, as all receptors do, a message of cell proliferation, differentiation, migration or else; but in the absence of ligand, these receptors induce apoptosis. They render thus the cell dependent on the ligand for survival. This particular situation prevents the cells from migrating beyond the area where the ligand is able to diffuse and thus control cell migration. They play a major role in organogenesis, especially in the nervous system (axon guidance) and in 'branched' organs (vessel arborescence, bronchial tree, etc.).

At least 15 receptors have been identified that fulfil the criteria for being dependence receptors; they belong to various families: in the tyrosine kinase receptor family (Chap. 1), this would be the case for RET, NTRK3 and ALK; in the TNFR superfamily, this is the case for NGFR (*nerve growth factor receptor*); this is also the case for some integrins (Chap. 10) and for the PTCH receptor of the Hedgehog pathway (Chap. 9). We focus here to a group of original receptors, the netrin 1 (NTN1) receptors, NTN1 and some homologous proteins being secreted proteins of the laminin family. These receptors are DCC (*deleted in colon cancer*) and the UNC5H1 to UNC5H4 receptors (genes *UNC5A* to *UNC5D*) (named after *uncoordinated*, the ortholog gene of *Caenorhabditis elegans*). NTN1 is an axon guidance factor, as probably also are the other NTN proteins; ligands and receptors are mainly expressed in the nervous system during development, but also, less specifically, in other organs such as the colon, lung, mammary gland or blood vessels.

UNC5H and DCC receptors are transmembrane proteins with a single transmembrane domain; they contain, in their extracellular portion, immunoglobulin-like domains and thrombospondin domains, which are replaced in DCC by fibronectin domains. At the intracellular level, UNC5H receptors harbour a death domain analogous to the death domains of FAS of TNFR and a protein interaction domain called ZU5 (*zona occludens homologue 5*), which are both absent in the DCC receptor. Receptor oligomerisation in the presence of NTN1 enables their activation and opens signalling pathways where cytoplasmic kinases such as FYN, FAK or ERK intervene and phosphorylate the receptor.

Apoptosis induction by dependence receptors is obtained, in the absence of ligand, through cleavage of the intracellular portion of the receptor by a caspase, at the level of an ADD (*addiction dependence domain*) (Fig. 18.10). In the case of DCC, caspase 9 is thus activated via adapter proteins. In the case of UNC5H receptors, the cleavage unmasks the death domain and the ZU5 domain. The death domain interacts with a serine/threonine kinase, DAPK (*death-associated protein kinase*), which contains a homologous death effector domain, responsible for apoptosis induction through caspase activation. The ZU5 domain recruits a protein called NRAGE (*neurotrophin receptor melanoma antigen homologue*, gene *MAGED1*), which is able to induce the JNK (*JUN N-terminal kinase*) pathway (Chap. 2) and inhibit the IAPs.

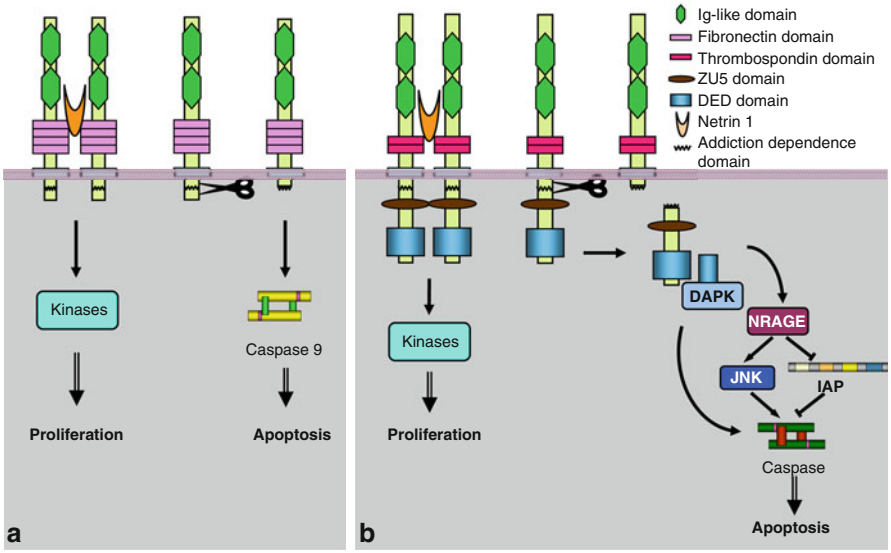


Fig. 18.10 Dependence receptors. (**a**) Activation of the DCC receptor. Left, the receptor is bound to netrin 1 (NTN1) and transduces a proliferation signal. Centre, in the absence of NTN1, it is cleaved by a caspase at the level of the ADD (*addiction dependence domain*). Right, it can thus activate caspase 9 and induce apoptosis. (**b**) Activation of a UNC5H receptor. Left, the receptor is bound to NTN1 and transduces a proliferation signal. Centre, in the absence of NTN1, it is cleaved by a caspase at the level of the ADD domain. Right, its ZU5 domain, via an adapter protein called NRAGE, can activate the JNK pathway and inhibit an IAP, whereas its death domain can activate DAPK and also induce apoptosis

DCC plays a major role in the renewal of the cells of the colic mucosa. NTN1 is produced by the cells located at the bottom of the crypts of the intestinal villi, while DCC is expressed all along the villi; proliferating cells at the bottom of the crypts are thus protected from apoptosis, while the differentiated cells, which migrate along the villi, move away from the NTN1 source and are thus progressively led to apoptosis. The NTN1 gradient would constitute a regulatory mechanism of the lifetime of intestinal cells, by eliminating the cells reaching the top of the villi, where they are submitted to mechanical and chemical aggressions in the intestinal lumen and should therefore be replaced.

Dependence receptors are bona fide tumour suppressor genes with *DCC* as a model. *DCC* is sometimes mutated, but more frequently lost in colon cancers, by deletion of the chromosomal region. NTN1 overexpression is associated to oncogenesis and tumour progression in several experimental models and is frequently observed in human cancers. By analogy, it seems that the *UNC5H* genes are also tumour suppressor genes.

Therapeutic approaches can be envisaged for targeting dependence receptors; NTN1 titration or sequestration can be obtained by small molecules. The role of UNC5H2 in blood vessel development suggests the possibility of developing original angiogenesis inhibitors.

Further Reading

Ashkenazi A. Targeting death and decoy receptors of the tumour-necrosis factor superfamily. Nat Rev Cancer. 2002;2:420–30.

Ashkenazi A. Directing cancer cells to self-destruct with pro-apoptotic receptor agonists. Nat Rev Drug Discov. 2008;7:1001–12.

Bao Q, Shi Y. Apoptosome: a platform for the activation of initiator caspases. Cell Death Differ. 2007;14:56–65.

Chipuk JE, Green DR. How do BCL-2 proteins induce mitochondrial outer membrane permeabilization? Trends Cell Biol. 2008;18:157–64.

Chipuk JE, Moldoveanu T, Llambi F, Parsons MJ, Green DR. The BCL-2 family reunion. Mol Cell. 2010;37:299–310.

Cory S, Adams JM. The BCL2 family: regulators of the cellular life-or-death switch. Nat Rev Cancer. 2002;2:647–56.

Croft M, Benedict CA, Ware CF. Clinical targeting of the TNF and TNFR superfamilies. Nat Rev Drug Discov. 2013;12:147–68.

Dempsey PW, Doyle SE, He JQ, Cheng G. The signaling adaptors and pathways activated by TNF superfamily. Cytokine Growth Factor Rev. 2003;14:193–209.

Fesik SW. Promoting apoptosis as a strategy for cancer drug discovery. Nat Rev Cancer. 2005;5:876–85.

Fulda S, Vucic D. Targeting IAP proteins for therapeutic intervention in cancer. Nat Rev Drug Discov. 2012;11:109–24.

Igney FH, Krammer PH. Death and anti-death: tumour resistance to apoptosis. Nat Rev Cancer. 2002;2:277–88.

Johnstone RW, Frew AJ, Smyth MJ. The TRAIL apoptotic pathway in cancer onset, progression and therapy. Nat Rev Cancer. 2008;8:782–98.

Kang MH, Reynolds CP. Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. Clin Cancer Res. 2009;15:1126–32.

Kelly PN, Strasser A. The role of Bcl-2 and its pro-survival relatives in tumourigenesis and cancer therapy. Cell Death Differ. 2011;18:1414–24.

Kruyt FAE. TRAIL and cancer therapy. Cancer Lett. 2008;263:14–25.

LaCasse EC, Mahoney DJ, Cheung HH, Plenchette S, Baird S, Korneluk RG. IAP-targeted therapies for cancer. Oncogene. 2008;27:6252–75.

Lejeune FJ, Rüegg C. Recombinant human tumor necrosis factor: an efficient agent for cancer treatment. Bull Cancer. 2006;93:E90–100.

Lessene G, Czabotar PE, Colman P. BCL-2 family antagonists for cancer therapy. Nat Rev Drug Discov. 2008;7:989–1000.

Li J, Yuan J. Caspases in apoptosis and beyond. Oncogene. 2008;27:6194–206.

Mahalingam D, Szegezdi E, Keane M, Jong S, Samali A. TRAIL receptor signalling and modulation: are we on the right TRAIL? Cancer Treat Rev. 2009;35:280–8.

Marzo I, Naval J. Bcl-2 family members as molecular targets in cancer therapy. Biochem Pharmacol. 2008;76:939–46.

Okada H, Mak TW. Pathways of apoptotic and non-apoptotic death in tumour cells. Nat Rev Cancer. 2004;4:592–603.
Ryan BM, O'Donovan N, Duffy MJ. Survivin: a new target for anti-cancer therapy. Cancer Treat Rev. 2009;35:553–62.
Villa-Morales M, Fernandez-Piqueras J. Targeting the Fas/FasL signaling pathway in cancer therapy. Expert Opin Ther Targets. 2012;16:85–101.
Vogler M, Dinsdale D, Dyer MJ, Cohen GM. Bcl-2 inhibitors: small molecules with a big impact on cancer therapy. Cell Death Differ. 2009;16:360–7.
Wiezorek J, Holland P, Graves J. Death receptor agonists as a targeted therapy for cancer. Clin Cancer Res. 2010;16:1701–8.
Yip KW, Reed JC. Bcl-2 family proteins and cancer. Oncogene. 2008;27:6398–406.
Youle RJ, Strasser A. The BCL-2 protein family: opposing activities that mediate cell death. Nat Rev Mol Cell Biol. 2008;9:47–59.
Yu JW, Shi Y. FLIP and the death effector domain family. Oncogene. 2008;27:6216–27.

Annex A: Control of DNA Replication and Repair

Abstract

We present in this annex some basic features concerning DNA structure, replication and repair. Cancers are gene diseases, associated to mutations or genes rearrangements, which generallay occur in a somatic cell, especially when exposed to environmental mutagenic agents (the term 'environment' being understood in a general sense). DNA replication is the target of several pharmacological approaches that aim at cancer treatment by preventing cell proliferation:

- *Alkylating and platinating agents* covalently bind to the nitrogen bases that constitute the DNA nucleotides.
- *Antimetabolites* inhibit nucleotides biosynthesis and disrupt DNA structure, frequently through fraudulent incorporation of abnormal nucleotides into DNA.
- *Topoisomerase inhibitors* stabilise the DNA breaks induced by specialised enzymes, topoisomerases, in charge of the regulation of the three-dimensional DNA structure.
- *Spindle poisons* interfere with mitosis, the ultimate step of cell reproduction, by disorganising the microtubule spindle where chromosomes attach.

All these compounds exert anticancer properties because they inhibit the multiplication of proliferating cells, cancerous or not, and are not studied here. We focus instead on the control of the events which govern DNA synthesis, on the one hand, and DNA damage repair, on the other hand. We also present the control processes that govern the protection of chromosome ends (telomeres) at the end of replication. All these processes are often altered in cancer cells, of which a fundamental feature is genomic instability. There is an important pharmacological research aiming at identifying relevant targets for cancer treatment in genomic instability.

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8

A.1 General Organisation of the Genome

We simply present in this section some basic data concerning the human genome, aiming at understanding the nature and role of oncogenic alterations.

A.1.1 Chromosomes and Chromatin

The human genome contains two times 3.2 billion nucleotide pairs distributed in 23 pairs of chromosomes. Figures A.1 and A.2 simply remind of base, nucleotide and DNA structure. Each chromosome contains a single DNA molecule under an extremely dense form (Fig. A.3). Cell chromosomes are only recognisable during mitosis, during which they are highly compact. Outside mitosis, this compaction is less, and chromosomes are no longer identifiable. In the chromosomes, DNA is tightly associated with numerous proteins, which together with DNA constitute *chromatin*. Heterochromatin is highly packed and cannot allow transcription process, while less dense euchromatin enables transcription to occur. Chromatin is comprised of regularly chained particles, nucleosomes, piled up one over each other. Each nucleosome contains a cylindrical protein core around which DNA is coiled up (Fig. A.4).

The protein constituents of nucleosomes are the histones, which are basic proteins interfering with DNA through ionic bonds. The nucleosome core is an octamer assembling two copies of histones belonging to the H2A, H2B, H3 ad H4 subfamilies, while H1 subfamily histones serve to internucleosomal interactions. Histones can undergo multiple post-translational modifications, which modulate their

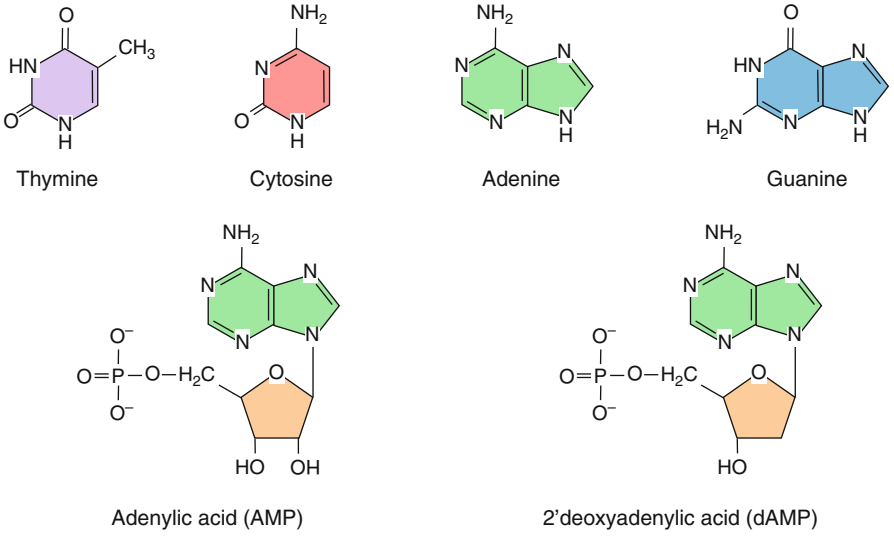


Fig. A.1 Structure of the DNA components. Structure of the four bases constituting DNA structure, of a 5′ phosphate nucleotide and a 5′ phosphate deoxynucleotide

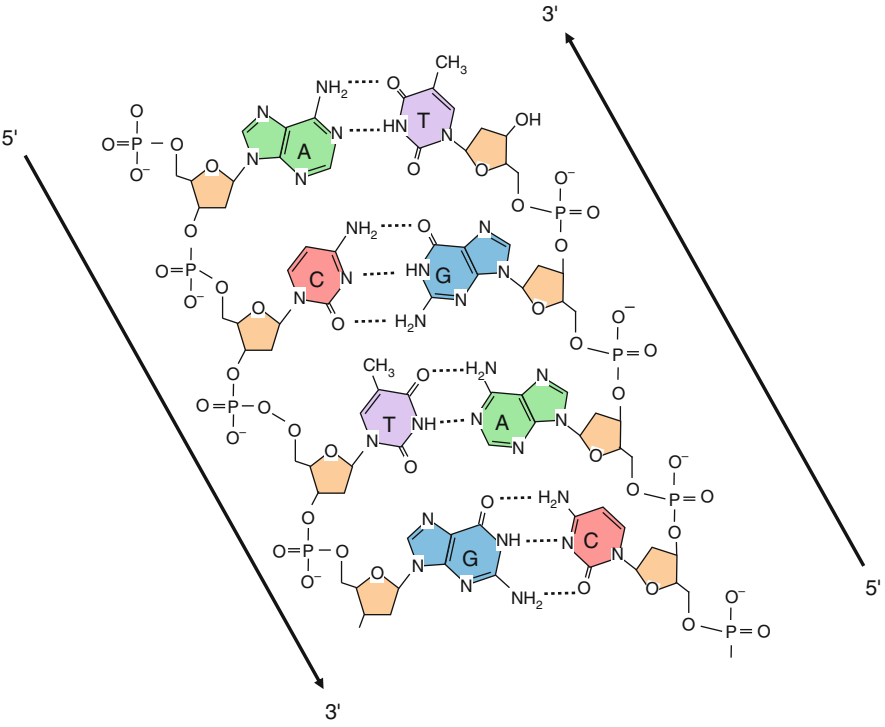


Fig. A.2 Primary structure of a DNA fragment. Base pairing through hydrogen bonds are presented as *dotted lines*

interaction with DNA and participate to transcriptional regulation (Annex B), especially methylations, acetylations and phosphorylations.

A.1.2 Genomic Sequences

Only a small part of chromosomal DNA sequences constitutes the genes, and among them, only a part encodes protein sequences. Close to the coding sequences, numerous transcription regulation sequences have been identified. Within the genome are encountered numerous non-coding and, as far as we presently know, not directly regulating sequences, among which a large part are repetitive and whose function remains unknown.

Genes

Genes contain the information required for protein synthesis. This information is discontinuous in the eukaryotic genome, and informative (coding) sequences called *exons* are separated by non-informative sequences called *introns* (Fig. A.5). The protein sequence is thus determined by the nucleotide sequences of the exons only. The junctions between exons and introns are short characteristic sequences that are recognised by specialised enzymes for the maturation of mRNAs (Annex B). On

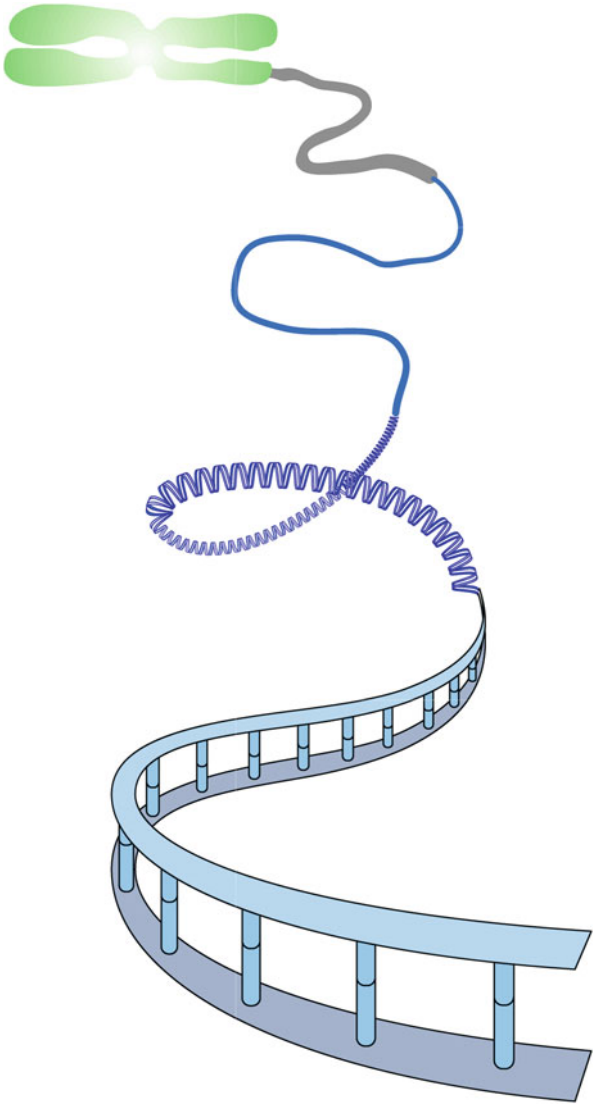


Fig. A.3 Chromosomal DNA packaging. Such a packaging is required to allow a molecule of 1 m long inside a cell nucleus of 1 μm diameter (Published with the courtesy of M. Donon)

each side of a group of coding sequences constituting a gene are found non-translated extremities, 5′UTR and 3′UTR, which play regulatory roles and are also eliminated during the maturation step.

These transcription-regulating sequences are often localised upstream the first exon, in the gene *promoter*, and are recognised by specific proteins called

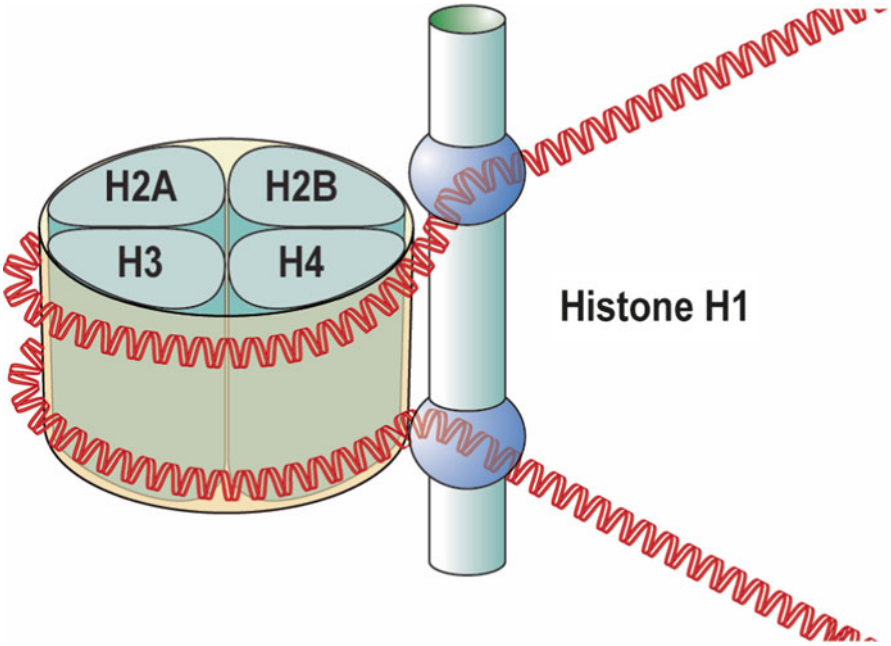


Fig. A.4 Nucleosomes structure. Nucleosomes are comprised of dimers of protein complexes associating four different histone molecules belonging to H2A, H2B, H3 and H4 classes. Histone H1 serves as internucleosomal scaffold (Published with the courtesy of M. Donon)

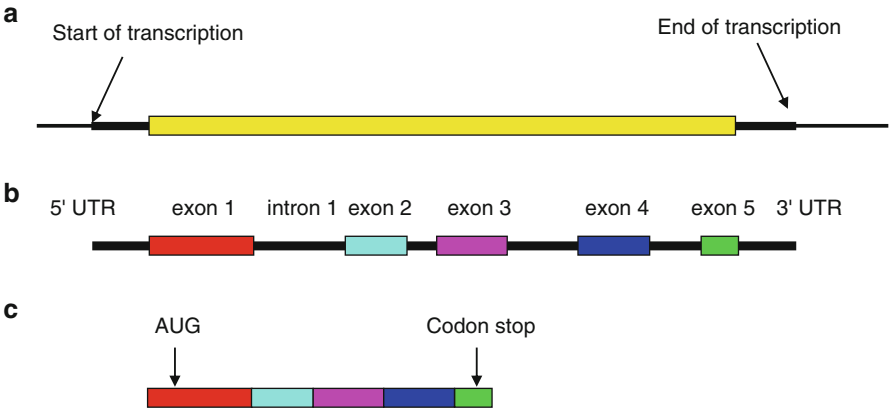


Fig. A.5 Discontinuous structure of eukaryotic genes. The gene (**a**) contains a long sequence of nucleotides that are transcribed into a primary mRNA (**b**), which is matured (**c**) so that only some fragments (*exons*) are translated into proteins. The 5′ and 3′ untranslated regions (UTR) and the introns are not translated into proteins. Introns are then eliminated by splicing (Annex B); specific nucleotide mRNA codons indicate the beginning (generally within exon 1) and the end (generally within the last exon) of translation (Annex C)

transcription factors. Other regulatory sequences are found sometimes at long distances from the gene they control and may be either *enhancers* or *silencers* of gene expression. Finally, regulatory sequences can also be found in the introns of a gene or in the 3′UTR, downstream the last exon, where micro-RNA binding domains can be found (Annex B).

Not all genes encode proteins: they may also encode RNAs devoid of messenger functions: ribosomal and transfer RNAs used for the machinery of protein synthesis (Annex C), micro-RNAs, RNAs contributing in the structure of ribonucleoproteins such as telomerase, etc.

Structural Chromosomal Sequences

Several chromosomal regions contain sequences that are essential for chromosomal functions: at the level of the centromere, where chromosomes are attached to the equatorial plate of the mitotic spindle (Chap. 17); at the level of the telomeres, the extremities of the chromosomes that must be protected from exonucleases (Sect. A.5); and at the level also of the origins of replication that are scattered all along the chromosomes (see Sect. A.3).

Repetitive Sequences

About 50 % of the genome is comprised of repetitive sequences, either one after the other (tandem repeats) or scattered all over the genome. They are of various lengths and made of repeats of a unit motif more or less complex. They can be distinguished according to the length of the motif and the number of repeats: (1) Microsatellites (*short tandem repeats* or STR), comprising some dozens of repeats of a short motif (two to five nucleotides), are found all over the genome, especially within introns where two-nucleotide CA repeats are frequent. The number of repeats is variable from an individual to another one, constituting a polymorphism that can be used as a genetic marker. (2) Minisatellites, with motifs of some dozens of nucleotides that are repeated several hundreds of times and are widespread in the genome, with highly polymorphic sequences. (3) Satellites, for which the motif of several hundreds of nucleotides is repeated 500,000–1,000,000 times.

Among the widespread sequences are:

- SINE (*short interspersed elements*) sequences, which contain 100–300 nucleotides. One of them, the Alu sequence, long of 282 nucleotides, is found in average every 4,000 base pairs, so that about 1,000,000 copies are to be found in the whole genome. They encode no protein and their role is unknown. They may derive from a gene encoding the RNA of a *signal recognition particle* whose transcripts would have been reinserted within the genome.
- LINE (*long interspersed elements*) sequences, which contain 5,000–7,000 nucleotides. They are thought to derive from certain transcripts from DNA polymerase II or from a retrotranscriptase, which would have been reinserted in the genome.

There are about 500,000 such sequences in the genome, often incomplete. Their insertion in a gene sequence can inactivate the transcription of this gene. They contain LTR (*long terminal repeat*) at their extremities, similar to those of viruses.

Transposable Elements

The LINE and SINE sequences are mobile sequences that can move randomly from a chromosomal location to another one; these sequences precisely encode enzymes able to catalyse their transposition, either directly (transposons) or after retrotranscription (retrotransposons). In normal cells, these enzymes are not expressed, owing to the high level of methylation of their promoters (Annex B). In several cancer types, however, these genes are reactivated, in relation to the loss of hypermethylation, but the mechanism of transcription reactivation of these genes remains poorly understood. It could be attributed to cancer genomic instability in relation to genotoxic stresses, including those elicited by anticancer agents. In addition, the transposition activity of the proteins thus synthesised amplifies genomic instability; according to the site of reinsertion of the transposed elements, transposition is able to inactivate tumour suppressor genes or even to activate proto-oncogenes. The inhibition of transposition activity in vitro induces the arrest of cell proliferation, so that a therapeutic approach could be conceived by inhibiting retrotranscriptase activity encoded by *LINE1*.

A.1.3 Main Genome Alterations

Various diseases, including cancers, are primarily associated to genome alterations. Some of them are point mutations, involving one single nucleotide: they can be identified by sequencing. Others correspond to gene rearrangements and can sometimes be detected by cytogenetics. In addition, common variations in DNA sequence are found all along the genome and responsible for the wide diversity of human species: these are called genetic polymorphisms.

Point Alterations: Mutations and Polymorphisms

Errors that occur during replication or that are induced by mutagenic agents can lead to the replacement of a nucleotide by another (*substitution*), to the loss of one or several nucleotide(s) (*deletion*) or by the addition of one or several nucleotides (*insertion*). The protein encoded by the gene may bear, therefore, a structural alteration that can lead to a decrease in its activity or the complete loss of its function. These mutations explain many hereditary diseases and malformations when they occur in a germinal cell and cancers when they occur in a somatic cell line. One can distinguish *silent mutations* (preservation of the amino acid in the polypeptide chain), *missense mutations* (replacement of an amino acid by another one) and *nonsense mutations* leading to truncated proteins (occurrence of a stop codon, alteration of a splicing site). Insertions and deletions of one or two nucleotides modify the reading frame and therefore the complete protein sequence: they are,

as a rule, nonsense alterations, whereas deletions and insertions of three (or a multiple of three) nucleotides preserve the reading frame, with loss or addition of one amino acid, and generally behave as missense alterations. The reader can find in Annex C (Table C.1) a reminder of the genetic code and of amino acid usual abbreviations.

In addition to these pathological alterations, rare but deleterious, there are numerous individual variations of the genome sequence that are called genetic polymorphisms. Their frequency may vary from one gene to another one; it is higher in introns than in exons. They are the support, in most cases, of minor phenotypic variations and explain the minor differences, from eye colour or skull shape to hereditary predisposition to some diseases or sensitivity to some drugs: as such, genetic polymorphisms are of major interest to the physician and the pharmacologist. Gene polymorphisms are not only simple replacements of a nucleotide by another one (that are called *single nucleotide polymorphisms* or SNPs): there are also polymorphisms involving insertions and deletions, as well as gene copy number (*copy number variations*, CNV), which may vary from an individual to another and may play a major role in gene expression level.

Mutations by substitutions and SNPs are, therefore, the same event from a biochemical point of view (replacement of a nucleotide by another), but they have different functional significance: one uses the word *mutation* for rare deleterious events and the word *polymorphism* for non-deleterious and frequent events (allele frequency of at least 10 %, which gives a frequency of at least 1 % of homozygous variants). As for mutations, one can distinguish silent or synonymous polymorphisms from non-synonymous polymorphisms, with replacement of an amino acid by another. Synonymous polymorphisms, nevertheless, may have consequences: the frequency of codon use for protein synthesis may vary, so that the corresponding transfer RNA population may not be adapted; as a possible consequence, protein synthesis may be slowed down and protein folding may be defective. In addition, the three-dimensional structure of the mRNA may be different after replacement of an amino acid by another one, and this may have consequences on mRNA stability and on translation efficiency, thus on the amount of protein produced.

Genome Rearrangements

Abnormal events during meiosis, and sometimes mitosis, may lead to important genome rearrangements. Trisomies are the most frequent rearrangements, with major pathological consequences. Deletions or insertions of large chromosome fragments may occur during homologous recombination or transposition processes. Gene duplication may also occur in the germinal cells during meiosis, but gene amplification, i.e. the large increase in gene copy number, essentially occurs in somatic cell lines and is generally observed in cancer cells. Finally, chromosome translocations, most often due to unequal recombination between homologous sequences between two different chromosomes, are also observed mainly in cancer cells, especially leukaemia cells.

A.2 DNA Replication Machinery

DNA replication takes place during the S phase of the cell cycle (Chap. 17). This is a semi-conservative process, as demonstrated more than 50 years ago by Meselson and Stahl: each of the two strands is used as template to be copied in a complementary way, by DNA polymerase(s). The constitutive elements of DNA, the four 2′-deoxynucleotides, are brought as trisphosphate esters, which provide both the required material and the energy required for their assembly. This polymerisation is carried out, on the two strands, in the 5′→3′ direction, in an antiparallel way and by pairing the homologous nucleotides, which are attached together by a phosphodiester bond. One of the two strands of the template is continuously copied from 3′ to 5′ to generate a *leading strand*, while the other one is also copied from 3′ to 5′ but discontinuously by fragments of about 200 deoxynucleotides called 'Okazaki fragments'. This new *lagging strand* requires afterwards the ligation of the fragments that have been thus generated (Fig. A.6a). The two new strands are simultaneously synthesised by DNA

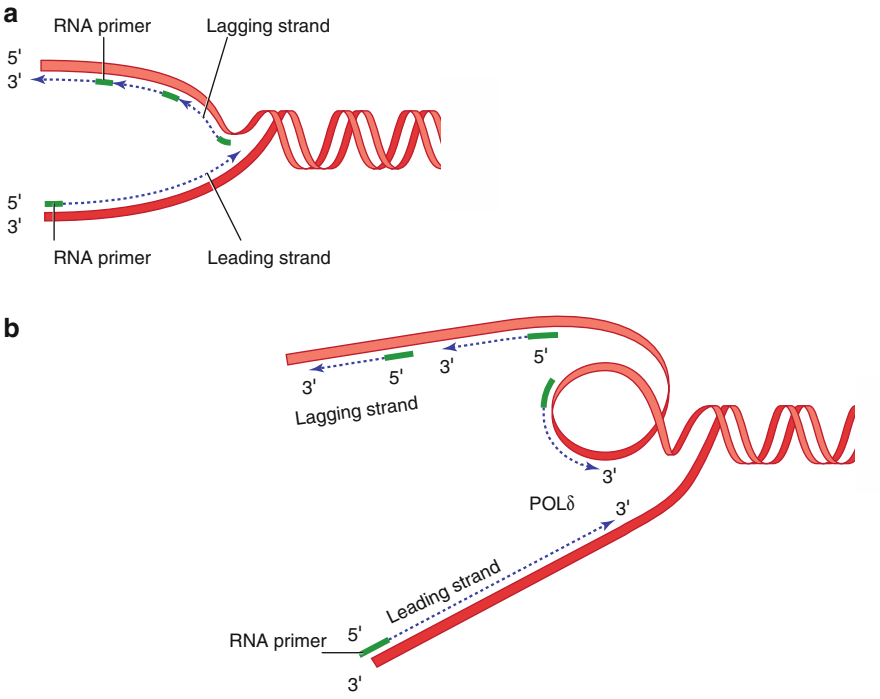


Fig. A.6 Replication process. (**a**) After the opening of the double helix, the two DNA strands are simultaneously copied in the 5′→3′ direction after the formation of a RNA primer which is replaced afterwards by DNA. This requires the continuous synthesis of one strand (leading strand) and the discontinuous synthesis of the other (lagging strand) by fragments of about 200 nucleotides (Okazaki fragments) that are bound together by ligase 1 (LIG1). (**b**) A loop formed by the lagging strands allows in fact the synthesis of the two strands by the same DNA polymerase dimer δ or ε (Published with the courtesy of M. Donon)

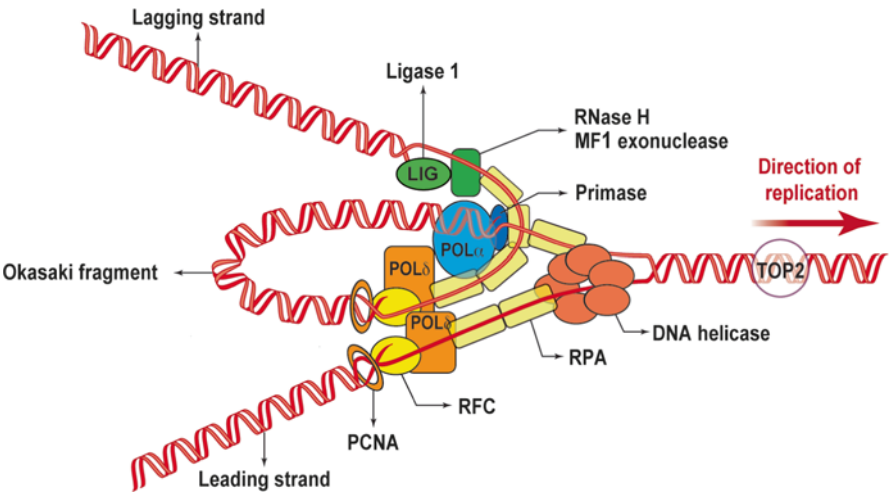


Fig. A.7 Replication machinery. The replication machinery involves a large number of different proteins. After intervention of topoisomerase II (TOP2), a helicase enables the opening of the double helix. Then RPA (*replication protein A*) is attached at the level of the origins of replication and an RNA strand is synthesised by an RNA polymerase called *primase*, because DNA polymerase cannot initiate DNA synthesis but only pursue synthesis from an RNA primer of about ten nucleotides. Once the RNA primer is synthesised, DNA polymerase α (POLA), in association with primase, takes over and begins DNA synthesis. After binding other DNA interaction proteins, RFC (*replication factor C*) and PCNA (*proliferating cell nuclear antigen*), the DNA polymerases δ or ε elongate the DNA chain. The RNA primer is then hydrolysed by RNase H and replaced by DNA via polymerases δ or ε (Published with the courtesy of M. Donon)

polymerases δ and ε (genes *POLD* and *POLE*), thanks to the fact that these enzymes act as dimers and that the DNA template constitutes a loop around them, in order to present in parallel the deoxynucleotides intended for copy (Fig. A.6b).

Replication starts simultaneously on several sites of each chromosome, called '*origins of replication*', characterised by precise sequences allowing their recognition by the replication machinery proteins. The opening of the template DNA molecule is carried out (1) by DNA topoisomerases II (TOP2), enzymes that realise transient double-strand breaks allowing the local relaxation of the double-helix supercoiling, and (2) by helicases allowing the opening of the double helix itself by transient breaks of the hydrogen bonds involved in base pairing. DNA synthesis is bidirectional and *replication forks* progress in both directions from the replication origins. After binding of RPA (*replication protein A*) at the level of the replication origins, an RNA strand is first synthesised by a heterodimeric RNA polymerase called *primase*, because DNA polymerase itself cannot initiate DNA synthesis but only pursue it from an RNA *primer* of about ten nucleotides. Once the RNA primer is synthesised, DNA polymerase α (gene *POLA*), in association with primase, takes over and begins DNA synthesis.

After binding other DNA-interacting proteins, RFC (*replication factor C*) and PCNA (*proliferating cell nuclear antigen*), the DNA polymerases δ and ε carry out DNA chain elongation (Fig. A.7). These polymerases, in addition to their DNA

synthesis function, display an editing (proofreading) function, i.e. the capacity of verifying that the last nucleotide added to the chain is the right one. This editing function is of crucial importance for genome stability and explains why the endogenous rate of DNA mutations (i.e. due to replication) is very low, about 10^{-10}. Transgenic mice expressing a mutant DNA polymerase devoid of editing function develop spontaneous cancers. The RNA primer is then destroyed by RNase H and replaced by DNA synthesised by DNA polymerases δ or ε. During DNA synthesis progress, ligase 1 (gene *LIG1*) enables the attachment of the fragments that are discontinuously synthesised on the lagging strand.

A.3 DNA Repair

Newly synthesised DNA can contain errors, despite the editing functions of the DNA polymerases δ and ε. It permanently also undergoes aggressions from physical and chemical agents: UV or ionising radiations, heat, mutagenic compounds, reactive oxygen species, etc. Several types of repair mechanisms must operate to restore DNA structure, each of them being directed toward given types of lesions, on the basis of the complementarity of the two DNA strands. When DNA damage cannot be repaired, the cell is normally driven to death; DNA damage tolerance characterises genomic instability of cancer cells and participates to oncogenesis: these lesions can induce an adaptive benefit from which cancer cells take advantage to proliferate and spread all over the body.

A.3.1 DNA Damage

DNA lesions are generally divided in two categories; Fig. A.8 presents some of these lesions:

Endogenous Damage

- Abasic sites whose formation most often results from the spontaneous breakdown of the glycosidic bond between the base and deoxyribose.
- Base oxidative alterations, leading in particular to 8-oxoguanine.
- Base deamination, inducing an alteration of the coding information, thus to mutations.
- Base mismatches resulting from non-corrected errors made by DNA polymerases.
- Accidental base methylations, leading for instance to N^7-methylguanine or N^3-methyladenine.

Exogenous Damage

- Pyrimidine or purine dimers, obtained by covalent bonds between two adjacent bases, especially induced by UV radiations.
- Polycyclic aromatic hydrocarbon adducts, such as benzopyrene, which attacks the electron-rich regions of nucleic bases.

- Adducts of other electrophilic compounds, which preferentially attack the N^7, N^2 and O^6 atoms atoms of guanine residues; some of these compounds are anticancer agents (alkylators, cisplatin).
- Double-strand breaks resulting from ionising radiations or drug-induced topoisomerase stabilisation.

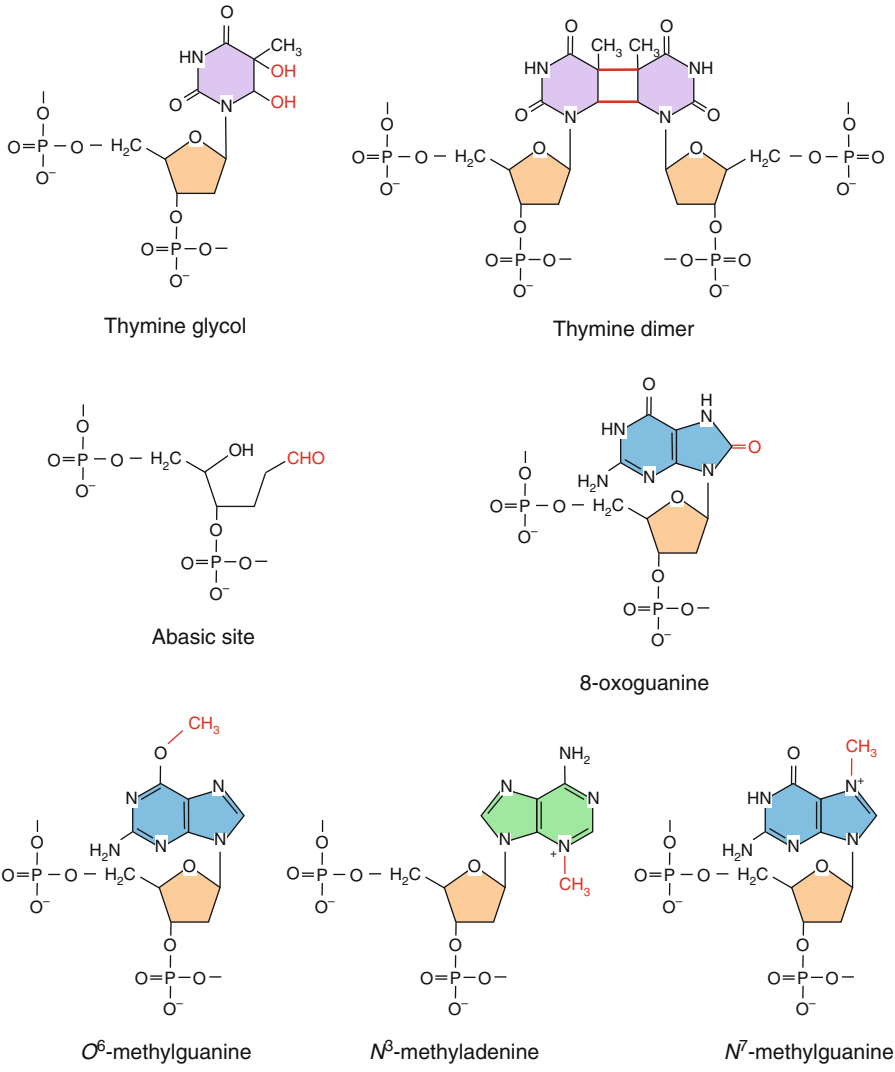
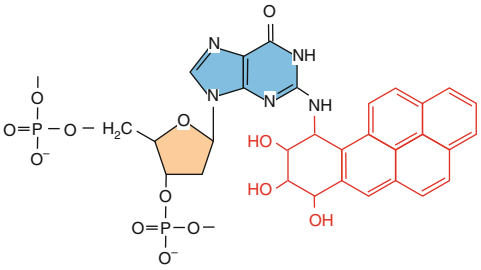


Fig. A.8 Some important types of DNA damage. The most frequent lesions originate from base excision, deamination, oxidation or alkylation. UV radiations are partly responsible for thymine dimers. More voluminous adducts can be formed through reaction with genotoxic molecules, carcinogenic or anticancer agents. Structural alterations are indicated in *red*



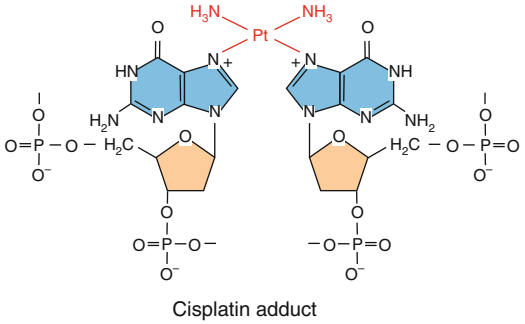


Fig. A.8 (continued)

A.3.2 DNA Repair Mechanisms

Each type of DNA damage can be carried out by a specialised repair mechanism. Table A.1 lists the main mechanisms and their involvement in the different types of

Table A.1 Main DNA lesions and repair mechanisms

Lesions	Repair mechanism
O^6-methylguanine	Repair by direct reversal
Base mismatch	Mismatch repair (MMR)
Abasic sites	Base excision repair (BER)
8-oxoguanine	
N^3-methyladenine	
N^7-methylguanine	
UV-induced lesions	Nucleotide excision repair (NER)
Thymine dimers	
Single-strand breaks	
N^3-methyladenine	
N^7-methylguanine	
Cisplatin adducts	
Dichloroethyl adducts (nitrogen mustards)	
Carcinogenic adducts (benzopyrene)	
Ionising radiation-induced lesions	Homologous recombination (HRR)
Double-strand breaks	Non-homologous recombination (NHEJ)

damage. The first step of all repair mechanisms is the recognition of the lesion and of its lethal character; this step is generally followed by cell cycle arrest, enabling either the implementation of the appropriate repair mechanism or the induction of a cell death programme if the damage cannot be repaired (Chap. 18). There is in addition a possibility of tolerating the lesions, allowing DNA polymerases to pursue DNA replication without DNA repair; this process is called *translesion synthesis* (TLS).

Direct DNA Repair

This mechanism (Fig. A.9) enables the repair of alkyl adducts on the O^6 of guanine residues by genotoxic agents and by some anticancer drugs such as temozolomide or nitrosoureas. These lesions are mutagenic because they induce G:C → A:T transitions during replication. Their repair involves an enzyme, *methylguanine methyltransferase* (MGMT), which transfers the alkyl group to the thiol group of a cysteine residue. This is a suicide enzyme, which cannot be detoxified and can only serve for the repair of a single lesion. This mechanism is thus dependent upon MGMT availability and consequently on its transcription and translation rates. The MGMT promoter is subjected to methylation of its CpG islands, which thus governs its expression level (Annex B). This expression is highly variable from a tissue to another and tumours present either the *Mer*$^+$ or the *Mer*$^-$ phenotype as a function of MGMT expression. Evaluating

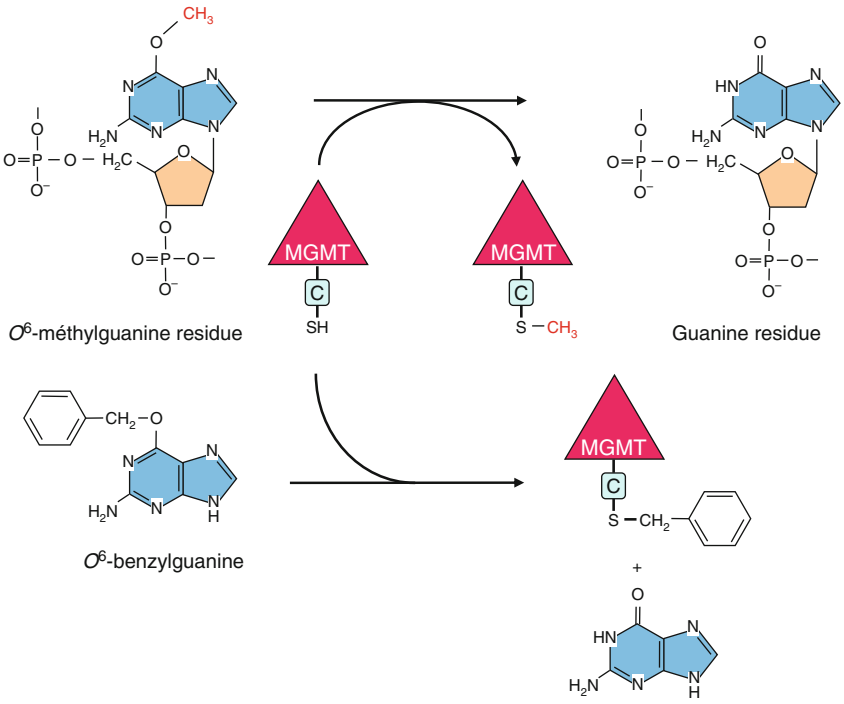


Fig. A.9 Direct repair. The methyl group of O^6-methylguanine is irreversibly transferred on a cysteine residue of the enzyme methylguanine methyltransferase (MGMT). Use of MGMT competitive inhibitors such as O^6-benzylguanine enables the saturation of its catalytic activity and potentiates the activity of alkylating agents

MGMT expression and/or promoter methylation can constitute a tool for predicting the activity of temozolomide in glioblastomas. An agent able to overcome MGMT activity is O^6-benzylguanine, which can serve as a decoy in diverting MGMT from the alkylated DNA adducts. Other derivatives are in development, such as lomeguatrib.

Mismatch Repair

This mechanism (MMR, Fig. A.10) aims at correcting the errors made during DNA synthesis by polymerases δ and ε, when these errors have not been corrected by their editing activity, which occurs with a frequency of about 10^{-10}. The introduction of the wrong base in front of the right one generates a *mismatch*. In addition, some DNA adducts, when nonrepaired, induce a confusion with another base: for instance, O^6 guanine can be paired with thymine; during the replication, adenine is inserted in front of thymine on one strand, but the adduct persists on the other, generating futile MMR cycles and accumulation of G:C → A:T mutations. MMR deficiency leads to resistance to agents that generate O^6 adducts, such as temozolomide.

Heterodimers formed by proteins MSH2 and MSH6, or MSH2 and MSH3 (MSH standing for *MutS homologue 2*), are able to recognise such mismatches and to recruit proteins specialised in their repair, MLH1 (*MutL homologue 1*) and PMS2 (*postmeiotic segregation increased*), forming all together a heterotetramer. The protein complex slides along the DNA to a short distance of the mismatch, then hydrolyses the strand that carries the mismatch over some dozens of nucleotides downstream, whereas the other strand is protected from exonucleases by RPA (*replication protein A*, gene *RPA1*). DNA resynthesis is then realised by DNA polymerase β (gene *POLB*) and strand continuity is ensured by a ligase.

This repair mechanism is deficient in certain cancers that present the phenotype of *microsatellite instability* (MSI), characterised by the fact that the dinucleotide repeats present in many microsatellites, which are difficult to replicate exactly the same during DNA replication, remain altered in the absence of functional MMR, so that the exact number of repeats is not maintained. This occurs especially in some forms of colon, ovarian and endometrium cancers, which present some phenotypic characteristics. This can be encountered both in sporadic cancers, when MLH or MSH mutations occur in a given somatic cell, and in cancers with hereditary predisposition when they occur in the germline. (HNPCC, *hereditary non-polyposis colon cancer*) Some microsatellites are more vulnerable than others and allele heterogeneity is evaluated for the molecular diagnosis of MSI.

Base Excision Repair

This mechanism BER, (Fig. A.11) mainly aims at replacing the bases that have been altered by an endogenous oxidative process, such as 8-oxoguanine and 5-hydroxycytosine. DNA glycosylases, which are diverse as a function of the damaged base (MPG [*N-methylpurine DNA glycosylase*], OGG1 [*8-oxoguanine DNA glycosylase*], etc.), are able to hydrolyse the *N*-osidic bond between the base and deoxyribose and thus to eliminate this base, generating an *abasic site*. This site may be recognised by the endonuclease APEX1 (*apurinic–apyrimidinic endonuclease 1*), which hydrolyses the phosphodiester bond and thus generates a single-strand break. Proteins XRCC1 (*X-ray repair cross-complementation group 1*) and PARP1 (*poly(ADP)ribose polymerase 1*) can recruit a *polynucleotide kinase* (PNK) that

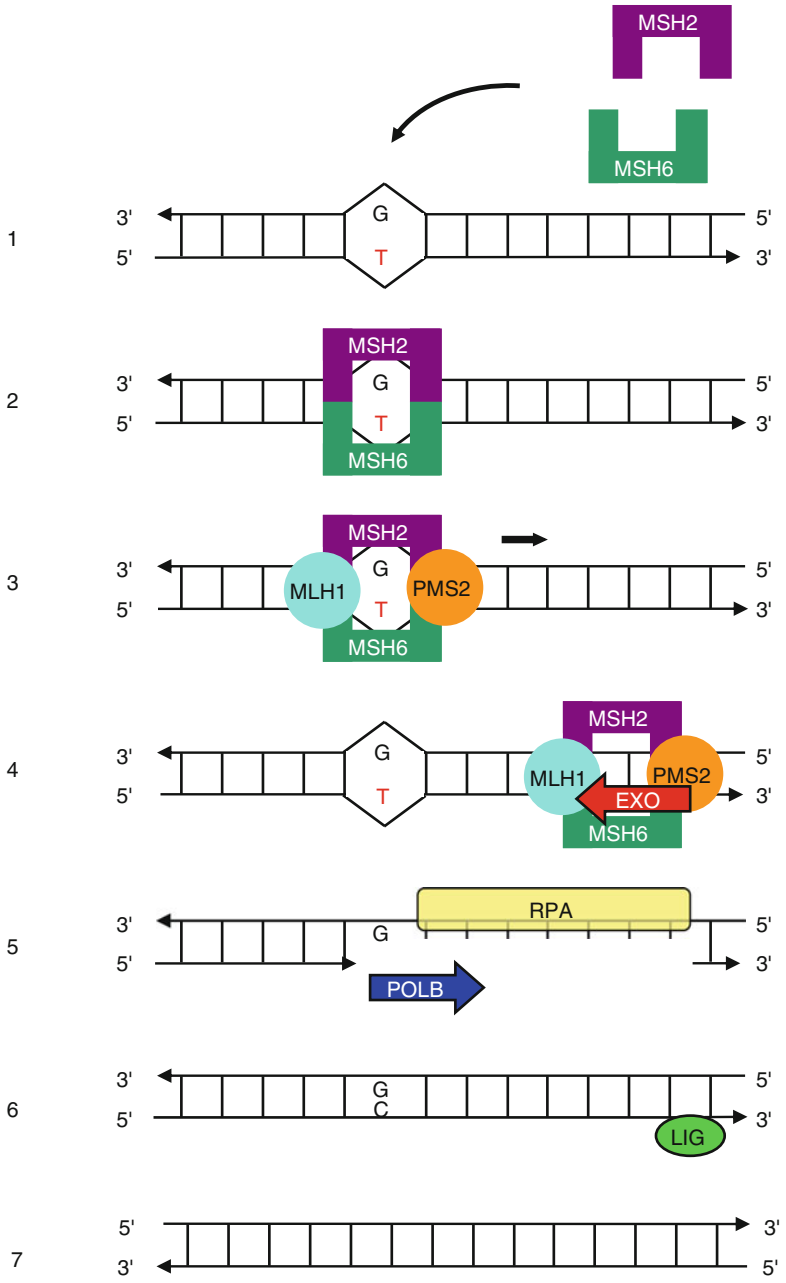


Fig. A.10 Mismatch repair (MMR). The G:T mismatch presented here is recognised by proteins of the MSH2–MSH6 complex, which then recruit the proteins MLH1 and PMS2. The tetramer thus formed slides downstream the mismatch and hydrolyses the mismatch-containing strand from there to the mismatch, thanks to an exonuclease (EXO). RPA protects the other strand to prevent nuclease degradation. DNA resynthesis is catalysed by DNA polymerase β and DNA continuity is restored by a ligase

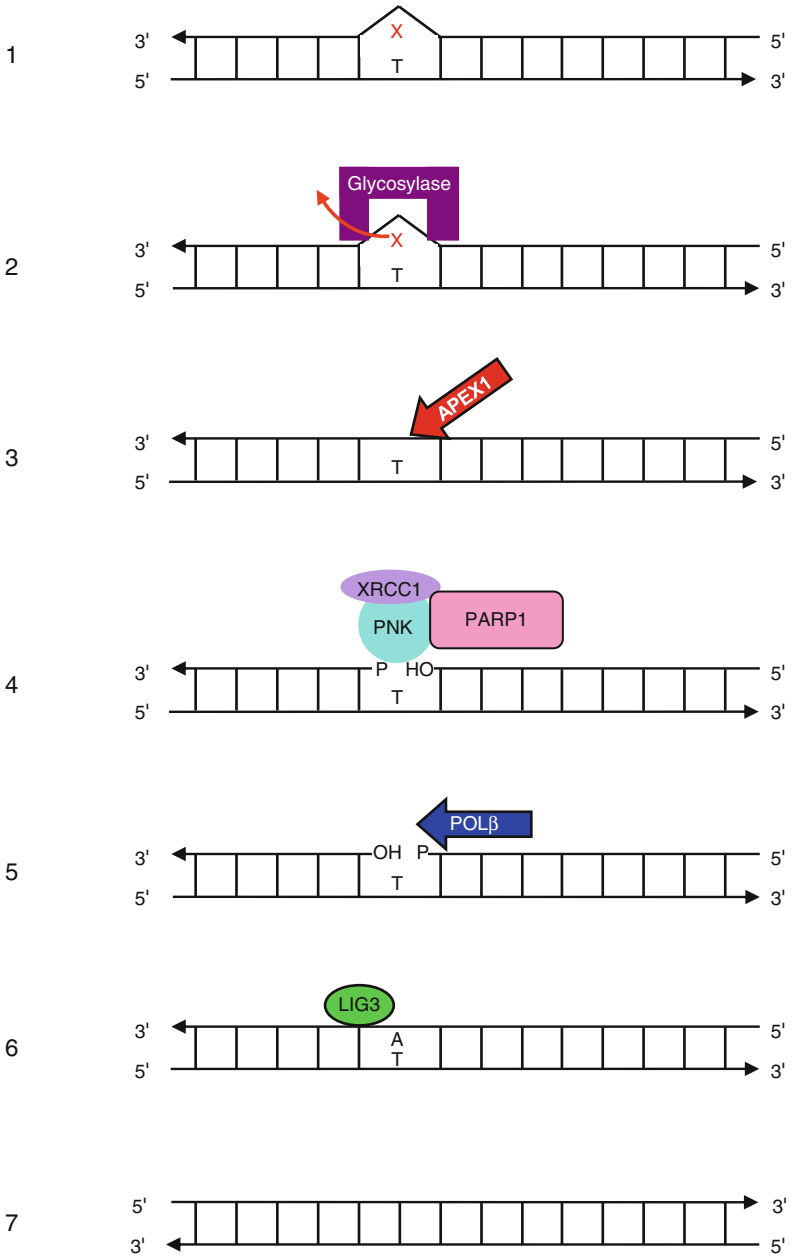


Fig. A.11 Base excision repair (BER). The damaged base (X) is eliminated by a DNA glycosylase specific for the type of damage, and DNA is incised by the endonuclease APEX1 at the level of the abasic site thus generated. The intervention of XRCC1 and PARP1 enables the recruitment of a polynucleotide kinase which phosphorylates the 5′OH end and dephosphorylates the 3′P end at the level of the break, a step required for the restoration of the phosphodiester bond by DNA polymerase. In the main BER pathway represented here, called *short patch*, the missing nucleotide is simply replaced by DNA polymerase β and the ligation is ensured by ligase 3 (LIG3)

phosphorylates the 5′OH end and dephosphorylates the 3′P end at the level of the break, a step required for the restoration of the phosphodiester bond by DNA polymerase β, sometimes polymerase θ or λ (genes *POLQ* and *POLL*). The sole missing nucleotide is replaced in the *short patch BER*, whereas five to six nucleotides are added at the level of the damage in the *long patch BER*, producing thus an overlap or a *flap*, which is removed by a *flap endonuclease*, FEN1 or GEN1. In both cases, ligase 3 (LIG3) must afterwards restore DNA continuity.

Nucleotide Excision Repair

This mechanism (NER, Fig. A.12) is mainly involved for UV-induced DNA damage but is also used to repair the damage generated by adducts formed with alkylating and platinating agents, especially at the level of the nitrogen atom #7 of guanine residues. NER utilises several XP proteins, owing their names because germinal mutations of their genes induce a group of diseases called *xeroderma pigmentosum*, characterised by UV-rays hypersensitivity and a hereditary predisposition to skin cancers. XP proteins are also often called ERCC (*excision repair cross-complementation*) proteins. Damage recognition can proceed in two different ways:

- In the case of *global genome NER* (GG-NER), the lesions are recognised, independently of their genomic localisation, by a protein complex XPC-RAD23B.
- In the case of *transcription-coupled NER* (TC-NER), only the lesions occurring in transcribed DNA regions are repaired, when RNA polymerase II must interrupt transcription, thanks to proteins called CSA and CSB (*Cockayne syndrome A* and *B*). This mechanism is carried out for a more rapid repair of transcribed genes in a given tissue.

In both cases, helicases called XPD (gene *ERCC2*) and XPB (gene *ERCC3*), with opposite polarities, associated to the protein transcriptional complex TFIIH, maintain the double helix in an open position to allow its accessibility to repair enzymes. XPA recognises and verifies the presence of the lesion, RPA binds to damaged DNA and the endonuclease XPG (gene *ERCC5*) hydrolyses the damaged strand on the 3′ side of the lesion. Then, the endonuclease XPF (gene *ERCC4*), associated to ERCC1, hydrolyses the damaged strand on the 5′ side of the lesion, thus releasing a DNA fragment of 24–32 nucleotides. DNA polymerase δ and ε reconstitute the excised sequence, bringing complementary nucleotides to the undamaged strand serving as a template, and ligase 3 (LIG3) restores DNA continuity.

NER is under the control of p53 (Chap. 17): tumours carrying a p53 mutation cannot activate this pathway, which results in an increased genomic instability. NER protein expression or activity plays a major role in the sensitivity to DNA damage-inducing anticancer drugs, especially platinum compounds. ERCC1 overexpression has been associated to cisplatin resistance in non-small-cell lung cancers, while some polymorphisms of ERCC2 have been associated to oxaliplatin efficacy in colorectal cancers. The development of drugs aimed at inhibiting NER to potentiate the activity of alkylating and platinating agents is currently explored. A particular

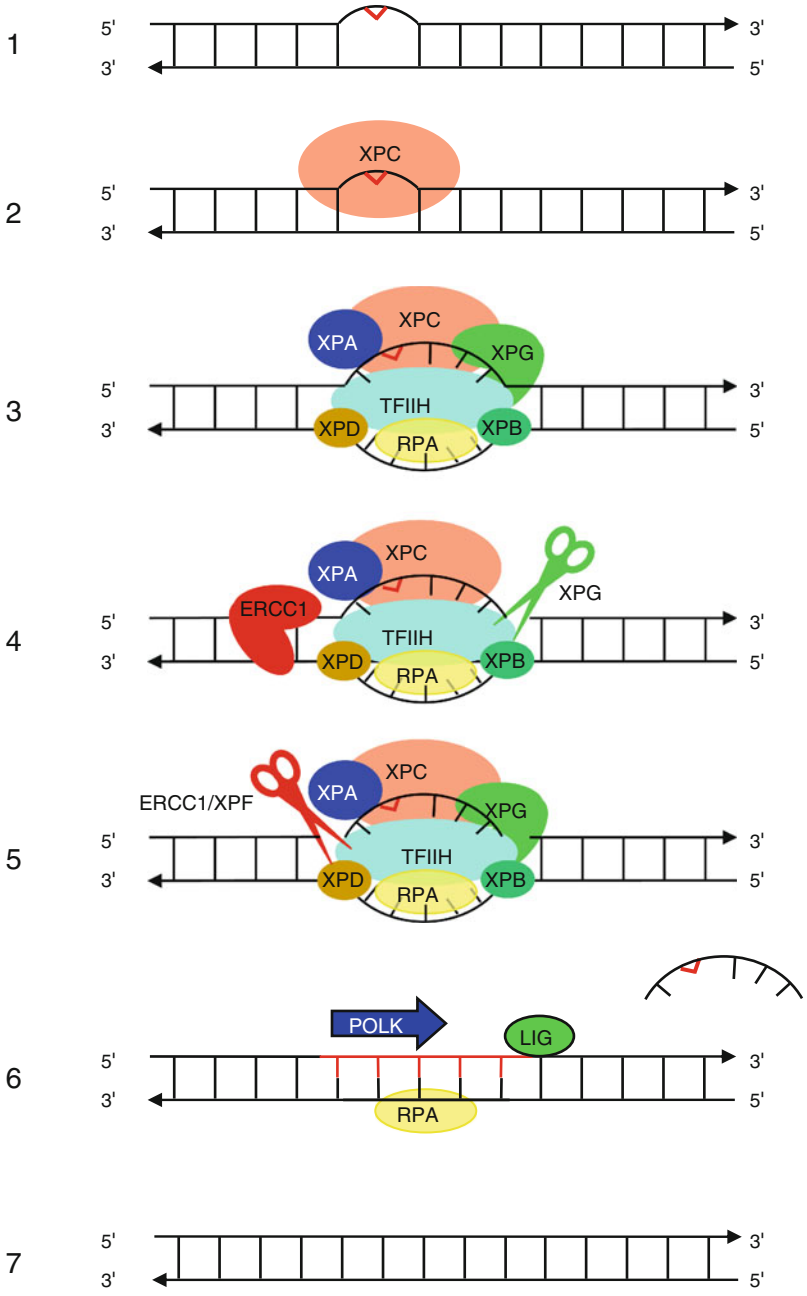


Fig. A.12 Nucleotide excision repair (NER). There are two distinct NER pathways, according to the fact that the lesion is localised within non-transcribed DNA (*global genome NER*) or transcribed DNA regions (*transcription-coupled NER*). They only differ by the mode of detection of the lesion. This detection is ensured by the XPC–RAD23B complex in the case of GG-NER represented here. Once detected, the lesion is made available for repair proteins through double-helix opening by helicases XPB and XPD of the TFIIH complex. Nucleases XPG and XPF–ERCC1 excise part of the damaged strand, then DNA polymerase β ensures the resynthesis of the excised portion. Afterwards, DNA ligase restores DNA continuity

alkylator, trabectedin, which targets the nitrogen atom N^2 of guanine residues, interferes with NER: contrary to other alkylators, it is more active when NER is highly proficient, probably because it stimulates a futile NER activity. Its combination with cisplatin in cancers like ovarian carcinomas could prove synergistic.

Homologous Recombination Repair

This mechanism (HRR, Fig. A.13) is involved in the repair of DNA double-strand breaks, either generated by ionising radiations or resulting from DNA cross-links induced by alkylating agents. It is activated after damage recognition by kinases, which play a role of damage sensor: ATM (*ataxia telangiectasia mutated*) and ATR (*ataxia telangiectasia and Rad-3 related*). These kinases are able to phosphorylate many proteins, directly involved either in DNA repair (histone H2AX) or in cell cycle arrest (CHK1 and CHK2 [*checkpoint kinases*, genes *CHEK1* and *CHEK2*]) (Chap. 17) or in apoptosis (MDM2 [*murine double-minute homologue 2*]). Homologous recombination is a relatively slow process, because it utilises the undamaged chromosome as a template, but it ensures an accurate copy.

A protein complex called MRN, comprised of the three proteins MRE11 (*meiotic recombination homologue 11*), RAD50 (*radiation response yeast homologue 50*) and NBS1 (*Nijmegen breakage syndrome 1*) or nibrin (gene *NBN*), is activated by phosphorylation by ATM and exerts a 3′ exonuclease activity. A series of proteins, namely RAD51, XRCC2, XRCC3, BRCA1, BRCA2 (*breast cancer protein 1 and 2*) as well as PARP1 (*poly*(*ADP*)*ribose polymerase 1*), are involved in the resynthesis of damaged DNA, which is obtained by copying the homologous sequence on the sister chromatid. This resynthesis, starting from the 3′ ends of each strand, extends beyond the lesion, and the restoration of the original DNA fragment is enabled by a specific enzyme called *resolvase*. *BRCA1* and *BRCA2* gene mutations are found in an important proportion of breast and ovarian cancers, especially those occurring with hereditary predisposition.

Inhibitors of homologous recombination have been identified: inhibition of MRE11 can be achieved by a small molecule, mirin, as well as the inhibition of ABL-mediated phosphorylation of RAD51 on a tyrosine residue by imatinib (Chap. 1); the inhibition of CDK1-mediated phosphorylation of BRCA1 by CDK inhibitors is achieved by dinaciclib (Chap. 17). Inhibition of the DNA damage sensor kinases, ATM and ATR, might sensitise cells to anticancer agents inducing DNA double-strand breaks (topoisomerase II inhibitors, ionising radiations). Wortmannin and caffeine have low selectivity; more active agents are being developed for clinical trials.

Non-homologous End Joining

This DNA double-strand break repair mechanism (NHEJ, Fig. A.14) involves a mere reassociation of the ends that were separated by the break. This process is rapid, but it is error-prone because it can introduce point deletions. This is the most common double-strand break repair mechanism in mammals. The split endings are recognised by a protein complex made of KU70 (gene *XRCC6*), KU80 (gene *XRCC5*) and another kinase, DNAPK (*DNA-activated protein kinase*, gene

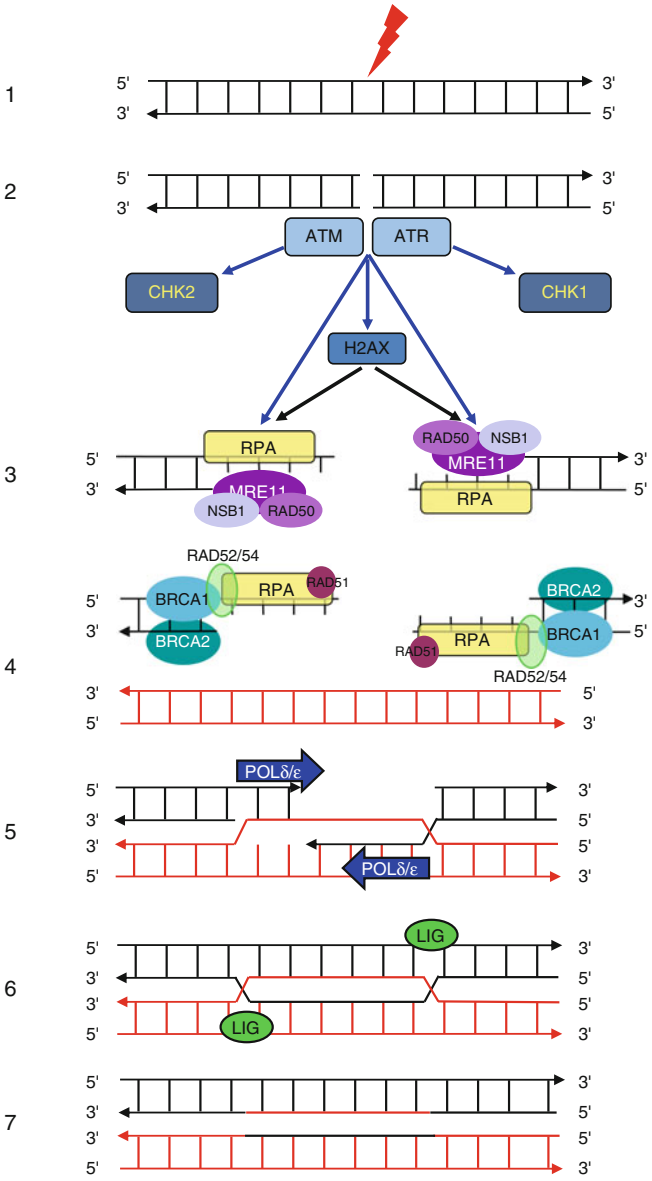


Fig. A.13 Homologous recombination (HRR). Double-strand breaks are detected by sensor proteins belonging to the PIKK family, ATM and ATR, which phosphorylate several effector proteins, among which CHK1 and CHK2 that control cell cycle progression (Chap. 17), histone H2AX that plays a role in chromatin remodelling after DNA damage and the MRN complex (MRE11, RAD50 and NBS1) that exerts 3′ exonuclease activity. There is a resection of DNA endings at the level of the break. A series of proteins, RAD51, XRCC2, XRCC3, BRCA1 and BRCA2 (*breast cancer protein 1* and *2*), are involved in DNA resynthesis, which is obtained by copying the homologous sequence of the damaged DNA fragment on the sister chromatid, and strand invasion allowing new synthesis to occur, using the intact chromosome as a template. This pathway is very secure and not error-prone

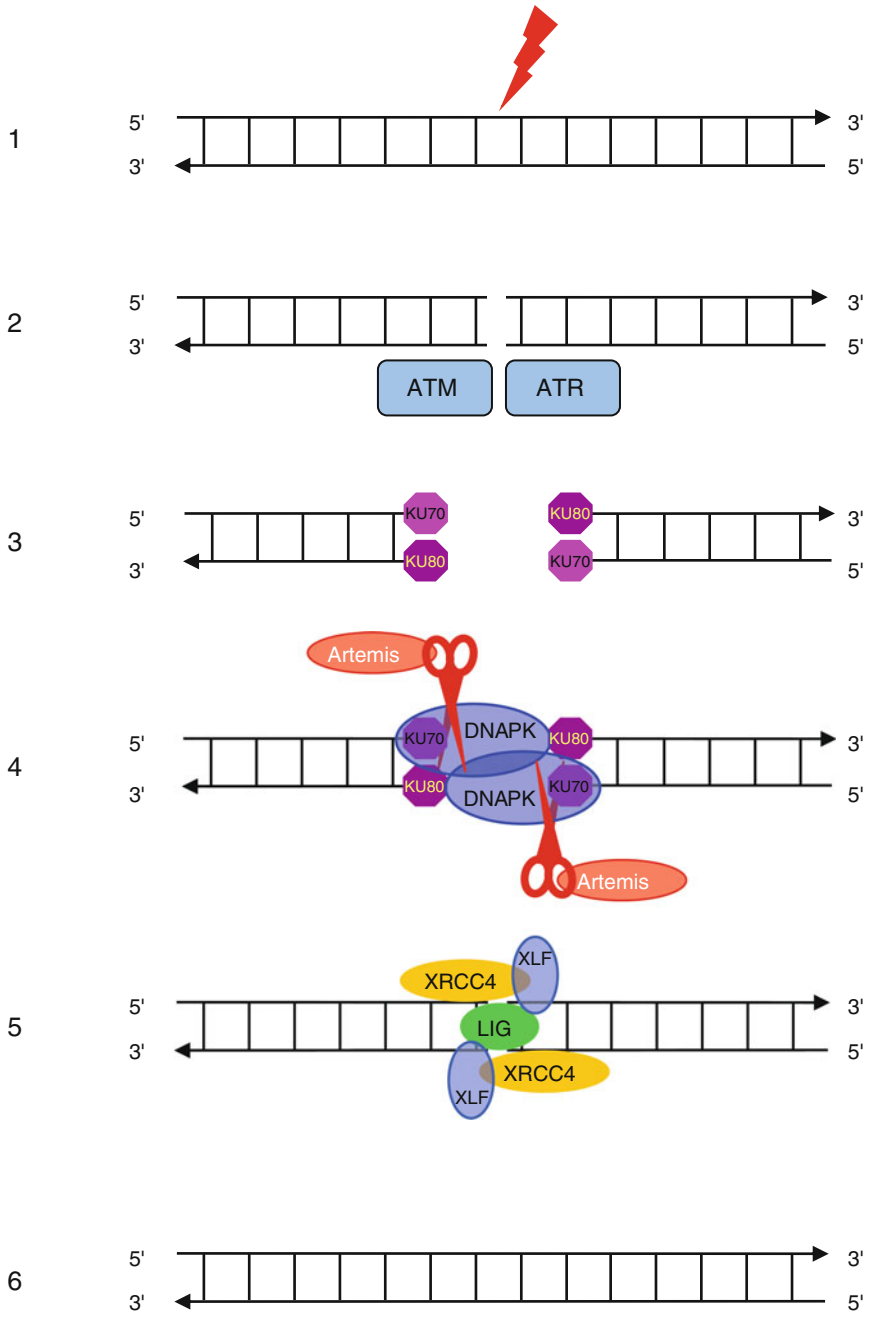


Fig. A.14 Non-homologous end joining (NHEJ). This recombination pathway is the major one in mammals. It consists in the reattachment of the break endings and involves mainly the proteins of the DNAPK complex, KU70 and KU80, associated to the artemis nuclease which prepares the endings to be joined. This is realised by a ligase associated to XRCC4. This pathway is more rapid but more error-prone than HHR, as it generates insertion or deletion of small nucleotide sequences

PRKDC), which activates by phosphorylation a 3′ exonuclease called *artemis* (gene *DCLRE* for *DNA cross-link repair*), which prepares the separated endings to enable ligase 4 (*LIG4*), associated to XRCC4 and XLF (*XRCC4-like factor*) (gene *NHEJ1*, for *non-homologous end-joining factor 1*), to bind the two endings. Defects in NHEJ are likely due to the occurrence of chromosomal abnormalities in cancers and leukaemias (translocations, inversions, etc.)

A.3.3 DNA Damage Tolerance Mechanisms

Prokaryotes and eukaryotes have developed various mechanisms for DNA damage tolerance, a process less expensive in energy than DNA repair, but contributing to genome instability. In mammals, DNA damage tolerance involves especially translesion DNA polymerases, devoid of editing functions, each of them being able to clear a given type of lesion: DNA polymerase η (gene *POLH*) for UV-induced thymine dimers and cisplatin-mediated guanine–guanine cross-links; DNA polymerase ι (gene *POLI*) for oxidative damage; DNA polymerase κ (gene *POLK*) for minor grove adducts such as those formed on guanine N^2 by mitomycin C; DNA polymerase θ (gene *POLQ*) for abasic sites, etc. Some of them are also utilised for post-repair DNA resynthesis. These polymerases are highly mutagenic and translesion synthesis is used by tumours to increase their level of genomic instability. Their expression is often altered in cancers, either decreased (so that DNA repair resynthesis is altered) or increased (favouring thus the accumulation of mutations). Mutations in *POLH* have been identified in some forms of *xeroderma pigmentosum* (XPV). Helicases may also participate to DNA translesion synthesis. High-fidelity DNA polymerases (POLΔ and POLε), when their gene is mutated, can also lose their proofreading function, participate to genome instability and favour the emergence of cancers.

A.3.4 Role of Poly(ADP-Ribose) Polymerases

Poly(ADP-ribose) polymerases (PARP) constitute a family of seven enzymes involved in the regulation of many processes, including transcription, cell cycle progression, cell death and DNA repair. PARPs catalyse the binding, to nuclear proteins glutamic acid residues, of ADP-ribose polymers produced by NAD^+ hydrolysis that they also catalyse (Fig. A.15). Two of them, PARP1 and PARP2, are activated by DNA damage and play a role as damage sensors and in signalling. PARP1 is able to bind the endings of DNA double-strand breaks during homologous or non-homologous DNA repair or during excision of abasic sites by the BER endonucleases. PARP2 interacts with PARP1 by associating with their common partners. Numerous proteins directly involved in DNA repair or in DNA damage response are PARP1 substrates: ATM, p53, KU70 (gene *XRCC6*), KU80 (gene *XRCC5*), DNAPK (gene *PRKDC*), XRCC1, MRE11, POLβ, LIG3, TOP1, etc. PARP1 is also able to bind poly(ADP-ribose) groups to histones, taking thus part in chromatin decondensation, as well as to transcription factors such as NFκB. Protein poly(ADP-ribosylation) constitutes thus a fundamental post-translational covalent modification of proteins.

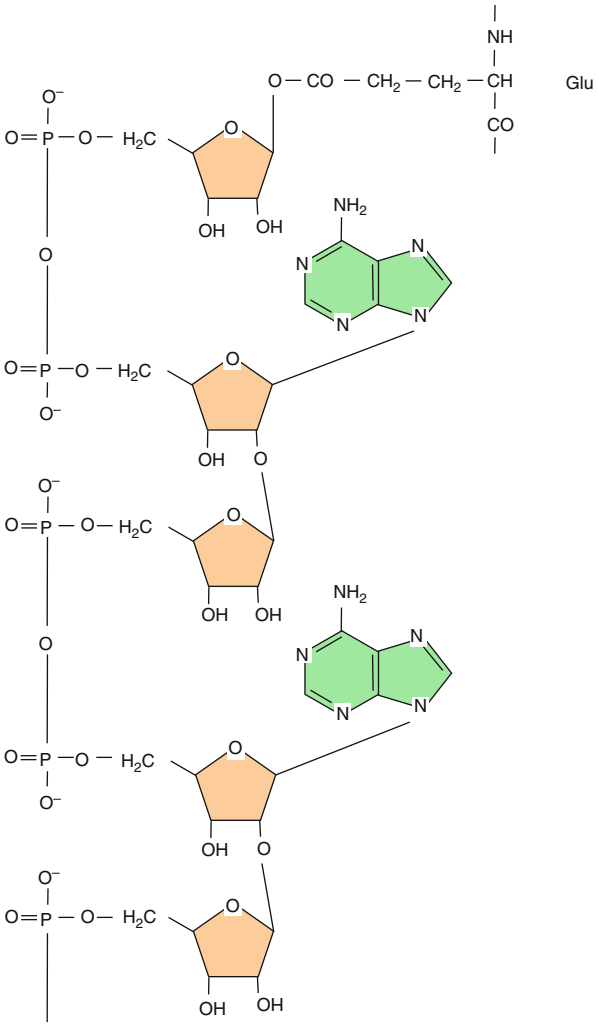


Fig. A.15 Structure of poly(ADP-ribose). Poly(ADP-ribose) consists of chains of ADP molecules bound together through a molecule of ribose. The first ribose is bound to a glutamic acid residue of an acceptor protein by an ester bond

The discovery of the role of PARP1 in BER led to the development of inhibitors aiming at potentiating the effect of alkylating agents (Fig. A.11). In the case of temozolomide, adenine N^3 and guanine N^7 methylations are rapidly repaired by BER with the help of PARP1 and PARP2 and do not contribute to cytotoxicity, only the O^6 guanine methylations being lethal, at least in the absence of MGMT-mediated repair. If PARP1 inhibitors are combined with temozolomide, BER is inhibited and cell death occurs, independently of O^6 guanine methylation. Similarly, these compounds are able to potentiate the effect of ionising radiations, which produce

double-strand breaks, and of topoisomerase I inhibitors, which also generate double-strand breaks during the collision of a replication fork with a cleavage complex. Another potential use of PARP1 inhibitors is for treating tumours presenting a deficiency in double-strand break repair mechanisms: this is the case of breast and ovary cancers displaying a deleterious *BRCA1* or *BRCA2* gene mutation or more generally tumours with BRCA1 or BRCA2 defects, encountered in the phenomenon of *BRCAness*.

3-aminobenzamide was the first molecule identified as a PARP1 inhibitor. More specific and more active compounds have been developed afterwards; they belong to various chemical families: isoquinolines, benzoxazoles, benzimidazoles and quinazolinones. Olaparib, rucaparib, veliparib and some others are presently in clinical trials, either in association with DNA-damaging anticancer drugs (topoisomerase I inhibitors, temozolomide, cisplatin, etc.) or alone in tumours presenting the BRCAness phenotype, in a synthetic lethality approach.

A.4 Protection of Chromosomes Endings

A.4.1 Telomeres and Telomerase

Telomeres are nucleoprotein complexes aimed at protecting chromosome endings from exonucleases and from accidental recombinations that could alter genomic information. At each cell generation, a fragment of telomeric DNA is eliminated, and this progressive chromosome shortening explains why normal cell lineages can realise a limited number of divisions, as if they harboured a 'mitotic clock'. When telomeres become too short, the cell replicative potential is exhausted, what is called the *Hayflick limit* (Fig. A.16). They acquire in a first step (M1) a replicative senescence phenotype then undergo a 'crisis' (M2) accompanied by lethal chromosome fusions generating genomic instability and leading to cell death. However, in germinal tissues as well as during embryogenesis, telomere regeneration occurs after each replication cycle, so that the replicative potential is maintained. This is catalysed by telomerase, an enzyme whose activity and/or expression is null in adult somatic tissues, excepted stem cells.

During the senescence step (M1), genetic or epigenetic alterations may allow some cells to pursue a replicative activity and to enter cell cycle; then, during the crisis (M2), these cells may escape cell death through reactivation of telomerase activity and thus acquire an indefinite replicative potential called 'immortality'. The occurrence of subsequent mutations, facilitated by genomic instability, confers to these cells the phenotype of cancer cells. Telomere maintenance and lineage immortality are characteristic features of cancer cells but are insufficient to determine the full cancerous phenotype. The telomeres of cancer cells are generally shorter than those of the surrounding stromal cells, which does not seem to hinder their function of chromosome protection.

Telomeres adopt a characteristic structure, consisting of several thousands of a six-nucleotide repeat (5′ TTAGGG 3′). They extend over 6–8 kb in somatic cells

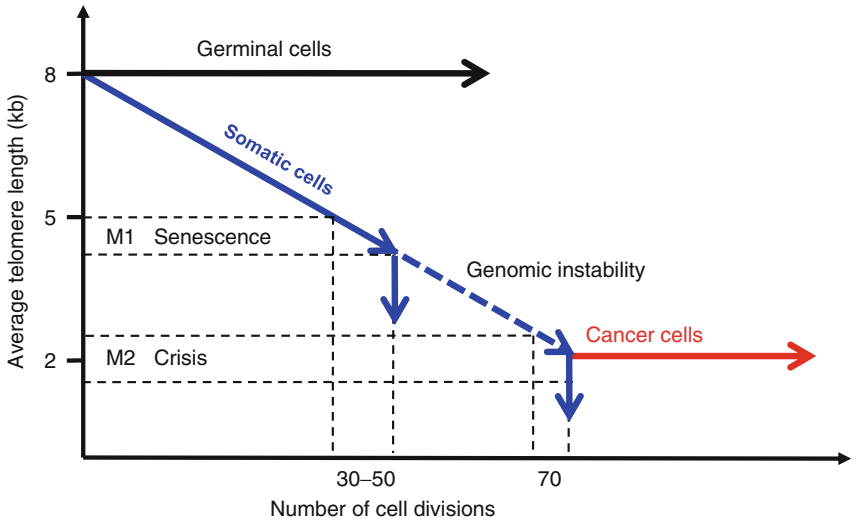


Fig. A.16 Evolution of telomere length as a function of the number of cell divisions. Telomere shortening occurs regularly at the end of each cell cycle; when the M1 threshold is reached (Hayflick limit), cells enter senescence and cannot undergo more divisions unless oncogenic alterations are present. Further telomere shortening leads to the M2 crisis, which generates a high level of genomic instability, generally leading to cell death. If cells survive crisis, they can become immortalised by telomerase reactivation, with short telomeres that are reconstituted at each generation

and 10–20 kb in germline cells. The 200 last nucleotides constitute a single DNA strand folded as a lasso, the *T-loop*, which realises, at its 3′ end, a hybridisation with an upstream sequence to form the *D-loop* (Fig. A.17). Several proteins are associated to these loops: TRF1 and TRF2 (*telomere repeat binding factors*), TIN2 (*TRF1-interacting protein* 2), RAP1 (*transcriptional repressor/activator protein*), POT1 (*protection of telomeres 1*) and TPP1 (*TIN2- and POT1-organising protein*) (Fig. A.17). TRF1, TRF2 and POT1 are bound to double-stranded telomeric structures and are interconnected by other proteins. TRF1 and TRF2 also interact with PARP (*poly(ADP-ribose)polymerase*) and with another enzyme called *tankyrase* (gene *TNKS*), as well as with DNA repair proteins: the MRE11–RAD51–NBS1, the heterodimer KU70–KU80 and the kinase ATM. All together, this constitutes an organelle at the extremity of chromosomes, called the telomeric cap, *telosome* or *shelterin*. The exact role of each protein of the complex is still incompletely understood.

Telosomes enable telomeres to exist under several distinct forms that can be converted from one to the other: a *capped* form, in which DNA is unavailable to exonucleases or to DNA recombination and repair as well as to telomerase, and an open *uncapped* form enabling telomerase to reconstitute the sequence lost during the previous replication cycle. In addition, telomeres can also adopt a third conformation, in which the single-stranded telomere end forms planar guanine tetramers

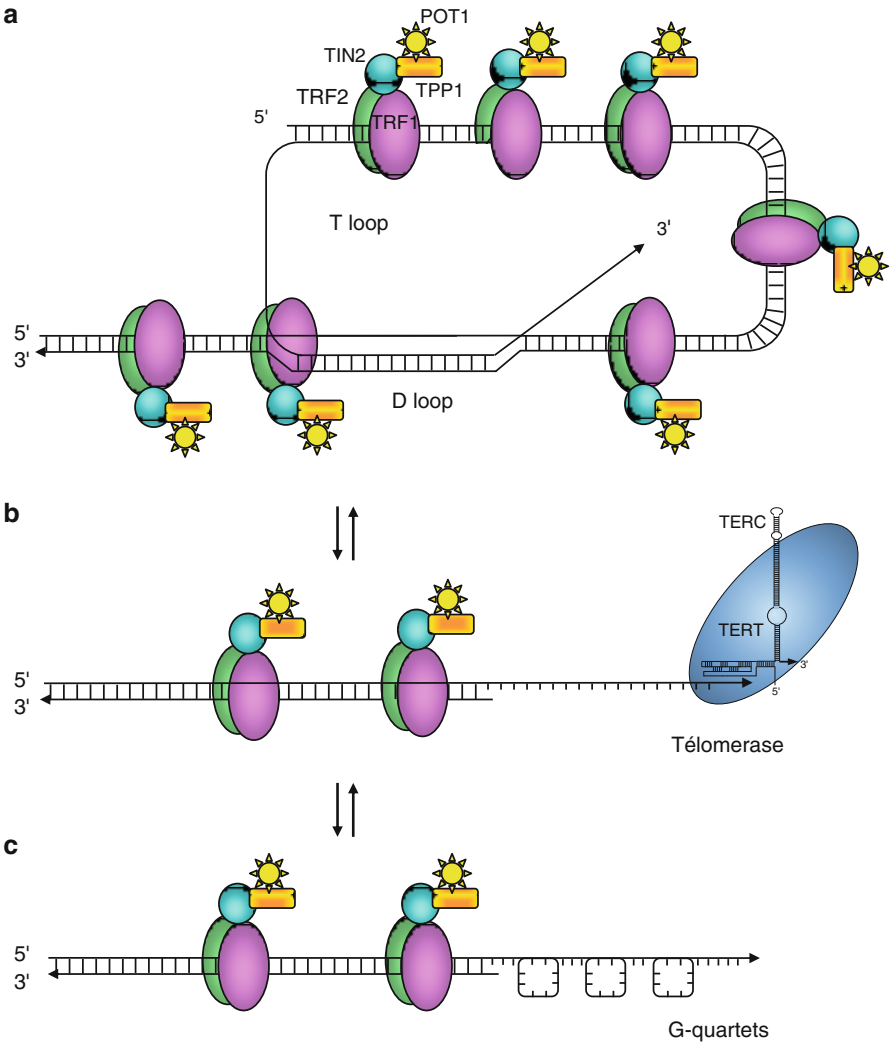


Fig. A.17 Structure of telomeres. Telomeres present a characteristic structure consisting of several thousands of repeats of a 6-nucleotide sequence (5′ TTAGGG 3′). In the *capped* telomere conformation (**a**), the last single-stranded 200 nucleotides are folded to constitute the lasso-like T-loop which hybridises, at its 3′ end, with a complementary segment to form the D-loop. Several proteins are associated to these loops, forming the TRF1 and TRF2 complexes. This capped conformation prevents telomere availability to exonucleases or telomerase. It can be converted to the *uncapped* conformation (**b**), available for telomerase-mediated elongation, in which G-quadruplexes are found (**c**)

called G-quadruplexes. In these structures (Fig. A.18), the consecutive guanine residues are associated by hydrogen bonds different from those involved in the classical Watson-and-Crick pairing. This constitutes exonuclease-resistant compact

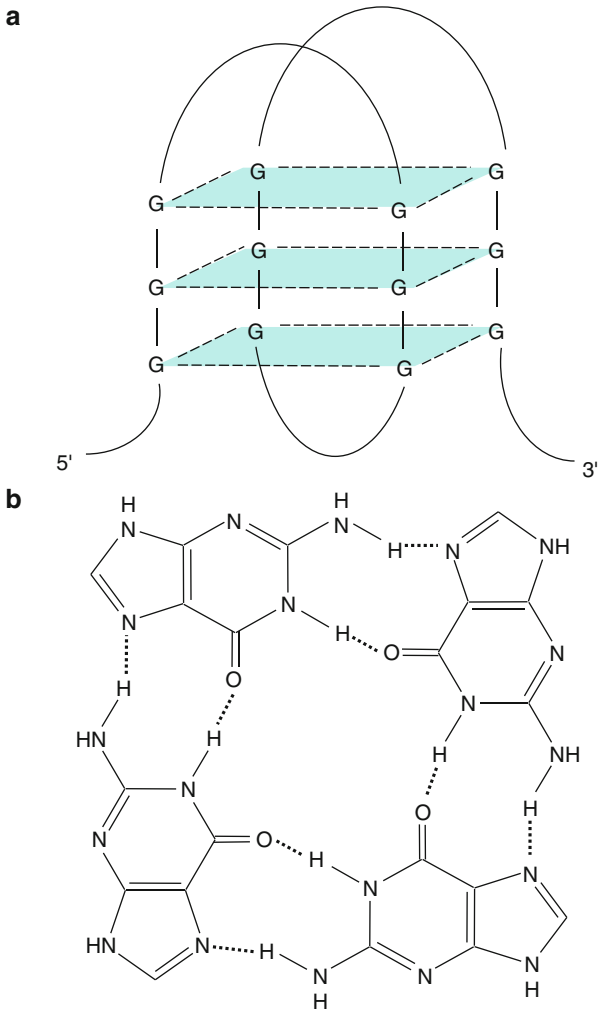


Fig. A.18 G-quadruplex structure. (**a**) Telomere single-stranded ends can form planar guanine tetramers that are stabilised through hydrogen bonds different from those required for the classical Watson–Crick pairing (**b**). These G-quadruplexes constitute a compact structure which is resistant to exonucleases and to telomerase

structures. The mechanism of interconversion between the different telosome conformations is not fully understood. G-quadruplexes are not exclusively present in telomeres; they are also present in other DNA sequences, such as ribosomal DNA and minisatellites.

Telomerase is a dimeric ribonucleoprotein particle containing an RNA chain of 451 nucleotides (TR or TERC, *telomerase RNA component*), in which a 12-nucleotide sequence will be used as a template for telomere resynthesis; a protein of 170 kDa

with retrotranscriptase function (TERT, *telomerase reverse transcriptase*); and an assembly of proteins constituting the *dyskerin* complex (Fig. A.19). TERT function is to catalyse nucleotide polymerisation at chromosome endings to reconstitute the telomeres. This is obtained by copying the RNA template by groups of six nucleotides to elongate the truncated DNA in the 5′ → 3′ direction with a shift of the same number of nucleotides at each step (Fig. A.20). Telomerase activity is regulated during embryonic development thanks to the transcriptional control of TERT, which is the limiting component of the protein complex. TERT is submitted to alternative splicing and a variant called TERTα exerts a negative dominant effect of telomerase activity.

Several proteins are associated to telomerase and contribute to its function: these are small nucleolar ribonucleoproteins (snoRNP), dyskerin (DKC1) and the

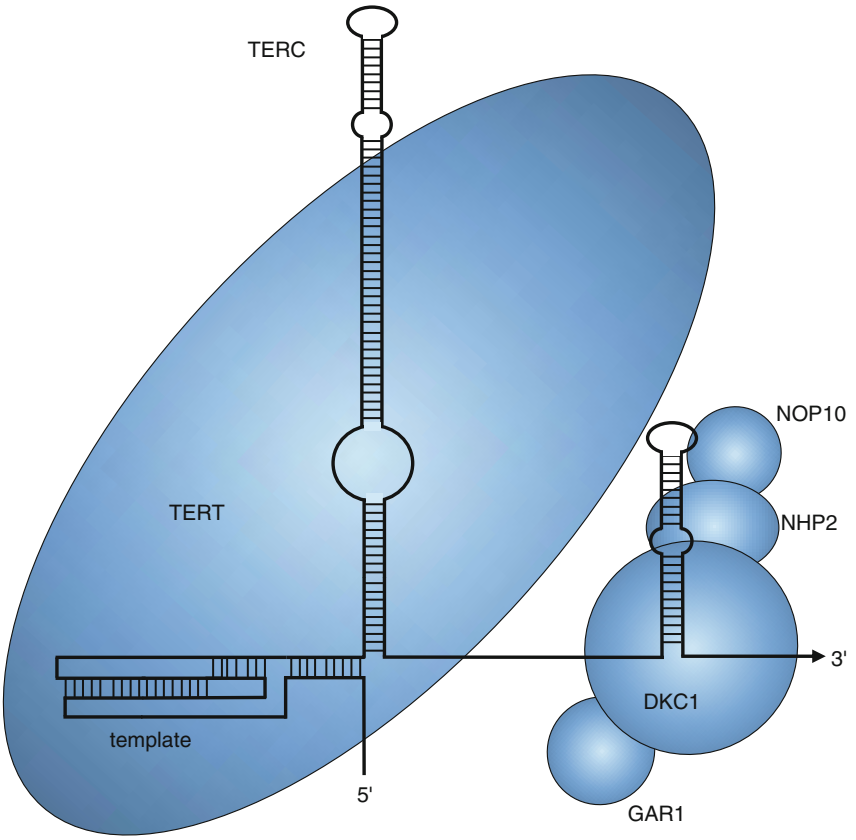


Fig. A.19 Structural organisation of telomerase and associated proteins. In its holoenzyme form, telomerase is a ribonucleoprotein associating: (i) a RNA molecule (TERC), of which a 12-nucleotide sequence is used as a template for telomere synthesis; (ii) a 170-kDa protein with retrotransferase catalytic activity (TERT); and (iii) a non-enzymatic, structural protein called dyskerin (DKC1), associated to the proteins NOLA1 (GAR1), NOLA2 (NHP2) and NOLA3 (NOP10), which are required for stability of telomerase RNA

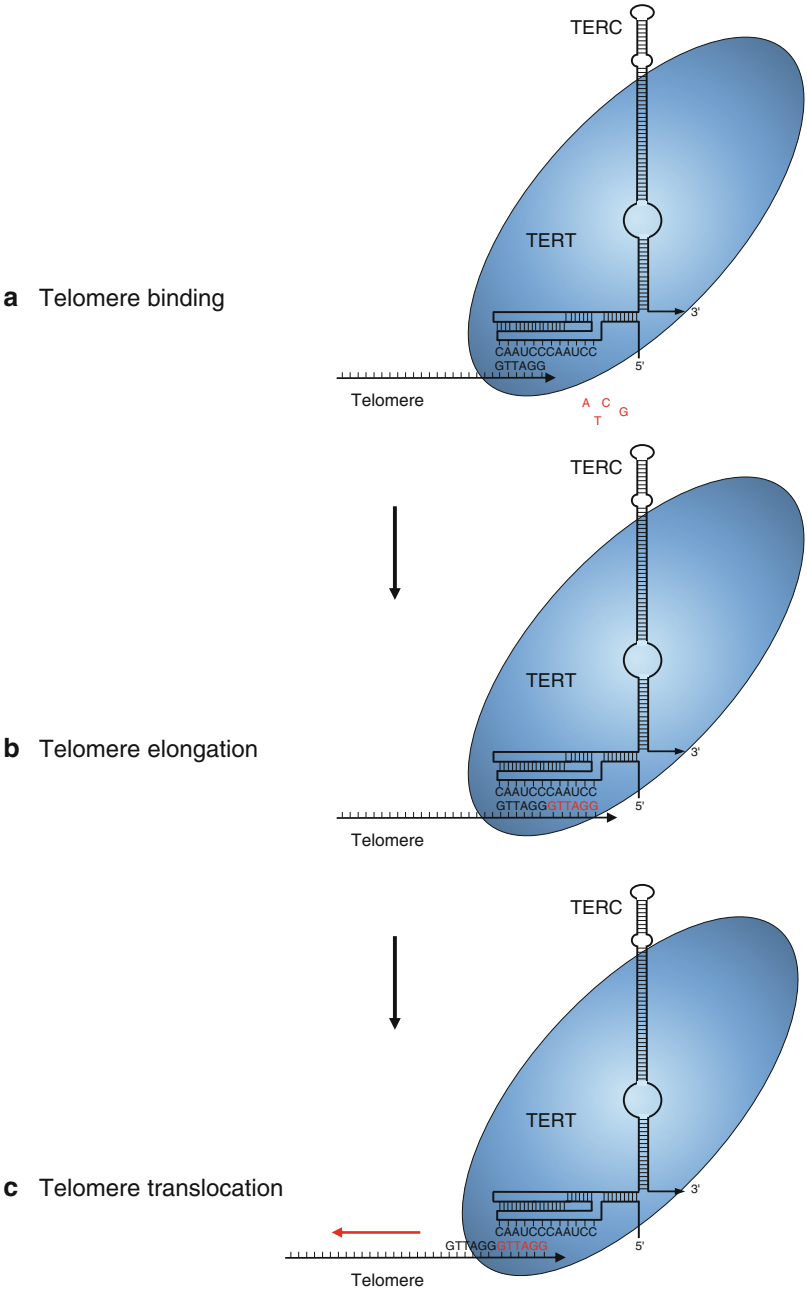


Fig. A.20 Telomerase mechanism of action. Telomerase enzyme subunit catalyses the polymerisation of nucleotides that are incorporated at chromosome endings, by recopying groups of six nucleotides of the RNA template to elongate the truncated DNA, in the $5' \rightarrow 3'$ direction, and with a 6-nucleotide shift at each time

proteins NOLA1 (gene *GAR1*), NOLA2 (gene *NHP2*) and NOLA3 (gene *NOP10*), which are required for the stability and accumulation of telomerase RNA and display distinct nucleolar functions. In addition, two ATPases, pontin and reptin, are required for the assembly of the complex TERT–TERC–dyskerin but do not remain associated to the telomerase functional complex. In addition, telomerase presents functions unrelated to telomere regeneration: it is a transcriptional regulator of the Wnt–β-catenin pathway and acts as a DNA-dependent RNA polymerase for the biosynthesis of small interfering RNAs (siRNAs).

A.4.2 Oncogenic Alterations and Pharmacological Targets

Telomerase activity is absent in somatic cells, with the exception of stem cells and some rapidly proliferating cell types. Reactivation of telomerase activity in cancer cells represents a critical way to ensure lineage immortality, and 85 % of cancer cells present such a reactivation. Another mechanism for telomere maintenance, independent from telomerase, has been described in cancer cells and involves homologous recombination; it is called ALT (*alternative lengthening of telomeres*). TERT reexpression occurring during the crisis of telomere evolution represents an important step of neoplastic transformation. In addition, telomerase inhibition triggers apoptosis or senescence of most cancer cell lines, which occurs rapidly due to the short size of their telomeres. For these reasons, telomerase constitutes a relevant target for inhibiting the proliferation of cancer cells.

Convincing experimental evidence has shown that telomerase reactivation in somatic cells is oncogenic: the introduction of telomerase in primary cultured cells contributes to the malignant phenotype and, conversely, the expression of a TERT negative dominant protein induces the reversal of the cancer phenotype associated with a decrease in telomere length. However, the molecular mechanisms involved in telomerase reactivation in somatic cells remained poorly understood until the recent discovery of a series of *TERT* mutations in various cancer types (central nervous system, bladder, thyroid, melanoma), which are responsible for an increase in TERT expression. In addition, *TERT* genetic polymorphisms, especially the intronic SNP rs2736100, with high allele frequency, have been associated to the risk of cancer and to sensitivity to anticancer drugs.

Several germline mutations of TERT lead to protein inactivation and are associated to various diseases, such as congenital dyskeratosis, which can also be associated to mutations of dyskerin (gene *DKC1*) or other proteins involved in telomere regeneration. However, these diseases are associated to a high risk of myelodysplastic syndromes that frequently evolve to acute leukaemia; this shows that, whereas telomerase is required for indefinite cell proliferation through cell lineage immortality, its loss can also lead to genome instability, which is also potentially oncogenic.

Numerous tracks have been followed for identifying relevant targets for cancer treatment. The most promising targets concern telomerase expression and/or activity and telomere structure. Telomerase catalytic activity can be targeted with nucleosidic drugs similar to those active against viral retrotranscriptases (azidothymidine,

carbovir, acyclovir, ganciclovir, penciclovir, etc.), but these compounds have shown poor anticancer activity. Non-nucleosidic compounds have been sought by high-throughput screening, using a functional test, the *telomere repeat amplification protocol* (TRAP). The long delay required by such drugs to induce antiproliferative effect is a major drawback for their development.

A variety of antisense approaches have been suggested, directed either against the RNA component TERC or the protein component TERT. Imetelstat, a 13-nucleotide thiophosphoroamidate mimicking the telomere sequence TTAGGG, is presently in clinical trials. Immunotherapy ('vaccination') and gene therapy approaches have also been suggested and are under study; this is, for instance, the case of telomelysin, an oncolytic virus preparation, aimed at killing telomerase-overexpressing cells.

Targeting the G-quadruplexes with planar molecules able to intercalate between the guanine plateaus is a strategy developed by numerous laboratories. Porphyrin and acridine derivatives have been especially studied. Pyridostatin, quarfloxin and telomestatin, a natural compound produced by *Streptomyces* species, are undergoing clinical evaluation.

Annex B: Control of Gene Expression

Abstract

This annex aims at deepen the understanding of signalling pathways by presenting the modalities of gene expression regulation, in other words the amount of messenger RNA (mRNA) that is made available for protein synthesis. Most signals received by cells eventually lead to modifications of gene expression; several signalling pathways end in the nucleus by the activation or inhibition of the transcription of specific genes. It appears necessary to describe in a simple way the general rules that govern genome transcription, also because these processes, downstream signal transduction, can harbour oncogenic alterations and pharmacological targets. In addition, we have often used the expression 'cell context' without definition, when describing the opposite effects that may induce the same signalling proteins in different circumstances. The cell context is the state of cell at a given moment and in a given place: level of differentiation, topological localisation in a tissue and in the organism, interactions with other cells, etc. This may be represented by the collection of the genes that the cell expresses at this moment and at this location. There again, we need to know how gene expression regulation proceeds for understanding the 'cell context'.

We briefly present the transcription machinery, as it is described in basic textbooks to which the reader is referred for more details. We then present the basic mechanisms that control transcription and the genetic and epigenetic factors that modify gene products, with special emphasis on cancer-related alterations and on the solutions proposed by pharmacology. Several transcription factors have already been presented: MYC, JUN, ELK1 and ETS in Chap. 2, FOXOs in Chaps. 2 and 3, STATs in Chap. 4, SMADs in Chap. 5, β-catenin in Chap. 7, GLI in Chap. 9, NFκB in Chap. 12, nuclear receptors in Chap. 14, NFAT in Chap. 15, HIF1α in Chap. 16, p53 and E2F in Chap. 17, etc.

B.1 Transcription Machinery

RNA synthesis results from the activity of RNA polymerases. It proceeds in the $5' \rightarrow 3'$ direction, uses as substrates the four ribonucleoside triphosphates (ATP, UTP, CTP and GTP) and generates phosphodiester bonds between the 3′ OH RNA

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8

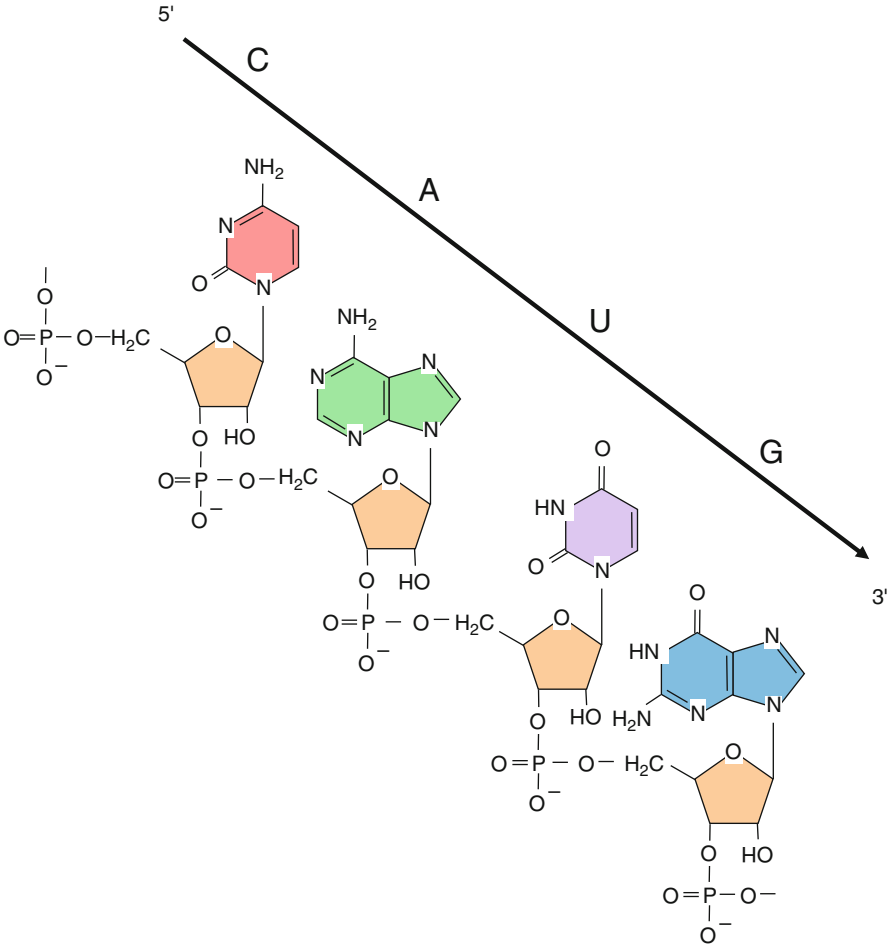


Fig. B.1 Structure of a RNA fragment. As an example, the CAUG sequence is presented here

end in the course of elongation and the 5′ P end of a new nucleotide (Fig. B.1). DNA nucleotide sequence determines RNA nucleotide sequence: RNA synthesis occurs after the opening of the DNA molecule by specialised enzymes and each nucleotide of the RNA strand is complementary to that of the template DNA strand, thymine being replaced by uracil. According to the gene, transcription proceeds on one DNA strand or the other. The copied strand, with a sequence complementary to that of RNA, is the antisense (−) strand; the other one, with a sequence identical to that of the synthesised RNA, is the sense (+) strand (Fig. B.2).

The enzymes responsible for RNA synthesis are the DNA-dependent RNA polymerases. Three distinct enzymes are in charge of RNA synthesis: RNA polymerase I (POLR1) for the synthesis of large ribosomal RNAs; RNA polymerase II (POLR2) for the synthesis of messenger RNAs (mRNA), small nuclear RNAs (snRNAs),

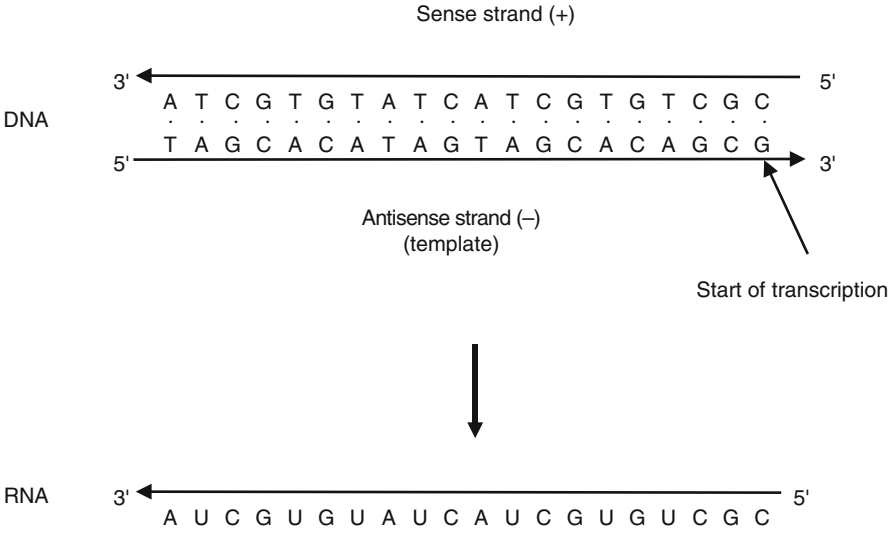


Fig. B.2 RNA synthesis (transcription). RNA synthesis is carried out from 5′ to 3′ and copies the DNA template sequence of the antisense or (–) strand. RNA sequence is thus identical to that of the DNA (+) strand, excepted for the replacement of thymine (T) by uracil (U)

small interfering RNAs (siRNAs) and micro-RNAs (miRNA); and RNA polymerase III (POLR3) for the synthesis of transfer RNAs (tRNAs) and other short stable RNAs. We will consider here essentially POLR2, as it is in charge of the synthesis of mRNAs and consequently of protein production. RNA polymerases are comprised of several subunits encoded by different genes, 15 in the case of POLR2.

As already mentioned, gene sequence is discontinuous in eukaryotes; the coding sequences (exons) are interrupted by non-coding sequences (introns). The transcription of a gene encoding a protein begins upstream (5′) the first exon, comprises all exons and introns, and ends downstream (3′) the last exon. After the transcription step, primary mRNAs are matured: the introns are excised by splicing, so that the mature mRNAs only contain: (1) a coding sequence resulting from exon connecting, which is later translated into proteins, and (2) untranslated (UTR) sequences in 5′ and 3′.

Upstream the transcription start site (TSS) is localised the promoter of the gene, which is the binding site for RNA polymerase, which determines the DNA strand to be copied. In its close neighbourhood are a series of characteristic sequences enabling transcription factor binding, especially a sequence called the *TATA box*, located 25 nucleotides upstream the TSS, which is only present in domestic genes and which is recognised by *TATA box-binding proteins* (TBPs). In non-domestic genes, other specific boxes play the same role as the TATA box.

A series of protein complexes called *general transcription factors*—TFIIA, B, D, E, F and H (*transcription factor IIA*, etc.)—are required for transcription, in addition to the enzyme (POLR2), the DNA template and the substrate nucleotides (Fig. B.3). All together, POLR2 and the general transcription factors (to be

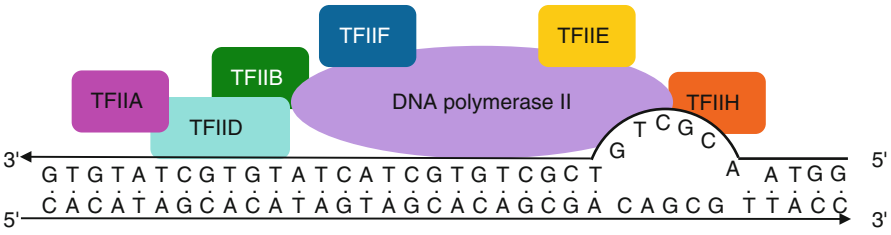


Fig. B.3 Transcription initiation complex. The catalytic activity of RNA polymerase II (POLR2) is possible only in the presence of the general transcription factors TFIIA, B, D, E, F and H, which are protein or protein complexes, some of them bearing a catalytic activity: TFIIH harbours a helicase activity borne by XPD, which allows the unwinding of the double helix to make it available for POLR2. It harbours also kinase and ATPase activities borne by other proteins of the complex

distinguished from the specific transcription factors involved in the control of transcription and described below) constitute the transcription initiation complex.

- TFIIA is a heterodimeric protein associating TF2AL (large, gene *GTF2A1*) and TF2AS (small, gene *GTF2A2*), which facilitates the binding of TBP to the TATA box.
- TFIIB (gene *GTF2B*) is involved in the association of the main POLR2 subunit (RPB1, gene *POLR2A*) with TBP.
- TFIID results from the association of several proteins, including TBP itself and a series of TBP-associated factors (TAFs).
- TFIIE is a heterodimeric protein (genes *GTF2E1* and *GTF2E2*) and is involved in the separation (melting) of the two DNA strands.
- TFIIF is also a dimeric protein (genes *GTF2F1* and *GTF2F2*) that binds primarily to POLR2 and prevents it to bind DNA sequences outside the promoter region.
- TFIIH results from the association of ten subunits, seven of which (XPD [*ERCC2*], XPB [*ERCC3*], p62 [*GTF2H1*], p44 [*GTF2H2*], p34 [*GTF2H3*], p52 [*GTF2H4*] and TTDA [*GTF2H5*]) form the core complex. To the core complex is transiently associated the cyclin-activating kinase complex (CDK7, MAT1 [gene *MNAT1*] and cyclin H [*CCNH*]) already mentioned in Chap. 17.

XPD displays helicase activity enabling the opening or melting of the DNA molecule (i.e. the local separation of the two DNA strands, by forming a loop) in order to make the nucleotides available for copy. CDK7 displays kinase activity for the phosphorylation and activation of RNA polymerase II. Transcription begins 25 nucleotides downstream the TATA box when it exists and is pursued by copying the antisense DNA strand, forming a transient RNA:DNA hybrid. The nascent RNA strand is progressively detached from the DNA template.

Elongation of RNA synthesis does not occur at constant rates and often pauses for various periods of time, aiming, for instance, at coordinating the transcription of

different genes. These pauses are mediated by the NELF (*negative elongation factor*) complex in association with the DISF (*DRB sensitivity-inducing factor*) complex. Elongation also requires positive transcription elongation factors which are organised in the P-TEFb complex containing especially CDK9, which phosphorylates several substrates (POLR2 on a serine residue of its *C*-terminal domain, the two negative elongation factor complexes, DSIF and NELF) and factors required for RNA modifications that occur during RNA synthesis, such as capping which is described below.

Characteristic nucleotide sequences signal the termination of transcription; they are located downstream the gene coding sequence and are recognised by POLR2. Two protein complexes carried by the *C*-terminal domain, CPSF (*cleavage and polyadenylation specificity factor*) and CSTF (*cleavage stimulation factor*), recognise such signals in the transcribed RNA, so that the newly synthesised mRNA is detached from the complex by hydrolysis. POLR2 still continues to move along the template and synthesises a novel RNA molecule which is rapidly cleaved by exonucleases, enabling POLR2 to eventually detach from the elongation complex.

mRNAs undergo a series of modifications during and after synthesis:

- During synthesis, mRNAs are equipped with a cap (Fig. B.4) aimed at protecting their 5′ end and at ensuring their recognition by the translation machinery: the γ phosphate group of the first nucleotide is hydrolysed, and this nucleotide then binds to a molecule of GTP at its 5′ end, generating a 5′ → 5′ bisphosphate bond; a methyl group is then added to GTP N^7, forming thus a MeGTP cap. Methylation of the 2′ OH moieties of initial nucleotides may also occur.
- After synthesis, mRNAs are equipped with a poly(A) tail (Fig. B.5), thanks to poly(A) polymerase (gene *PAPOLA*) which adds, at the 3′ ends, a series of approximately 200 adenylic nucleotides and enables the distinction between the messengers RNAs and the other cell RNAs. The polyadenylation signal in the gene sequence is in fact the termination signal mentioned above.
- mRNAs undergo the deletion of the introns through a process called *splicing*. Splicing is realised by a ribonucleoprotein complex, spliceosome, which recognises the initial nucleotide of the intron $n+1$ (GU, donor site), the terminal nucleotide of intron n (AG, receiver site) and a nucleotide located 40 nucleotides upstream the receiver site (A, branching site). The endonuclease activity of the spliceosome enables the excision of the intron, which is eliminated as a lasso, and the religation of the 3′ end of the upstream exon to the 5′ end of the downstream exon (Fig. B.6).

Once matured, mRNAs can migrate to the cytoplasm. They only contain the sequence to be translated (Annex C). DNA transcription rate, mRNA stability and religation of exons after splicing (alternative splicing) are important features for gene expression control that will be described in this chapter.

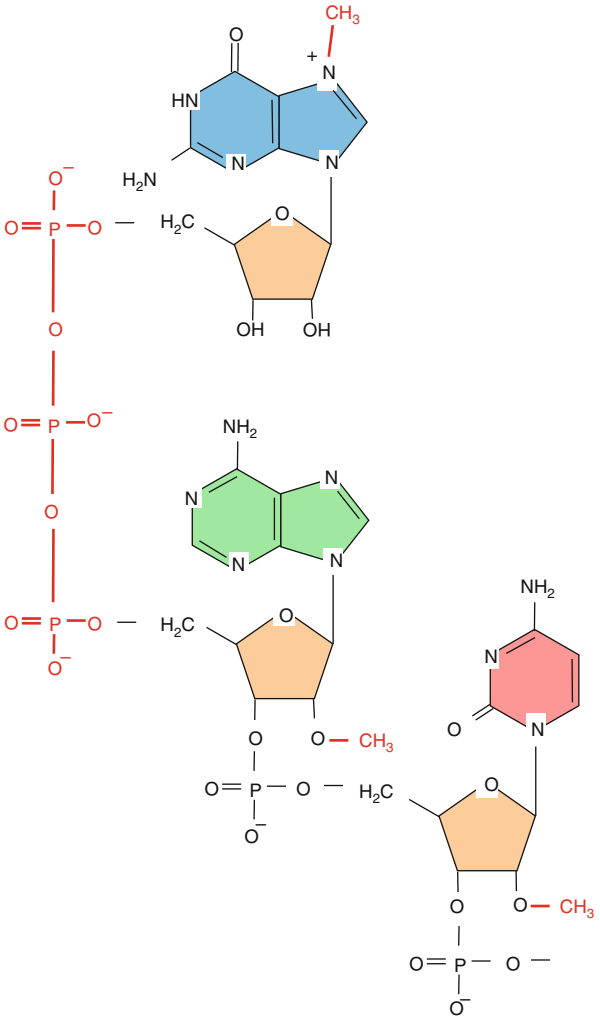


Fig. B.4 Structure of the 5′ end of an mRNA. The characteristic elements of the cap are indicated in *red*: guanine methylation on N^7, 5′ → 5′ phosphate bond, methylation of the first nucleotides in $O^{2'}$

B.2 Transcription Regulation

Gene transcription is controlled by regulatory proteins synthesised from other genes (*trans* regulatory sequences) and recognising short DNA sequences inside or close to the gene (*cis* regulatory sequences). Hundreds of different *cis* regulatory DNA sequences have been identified in the genome, each of them able to be recognised by regulatory proteins synthesised from *trans* regulatory sequences.

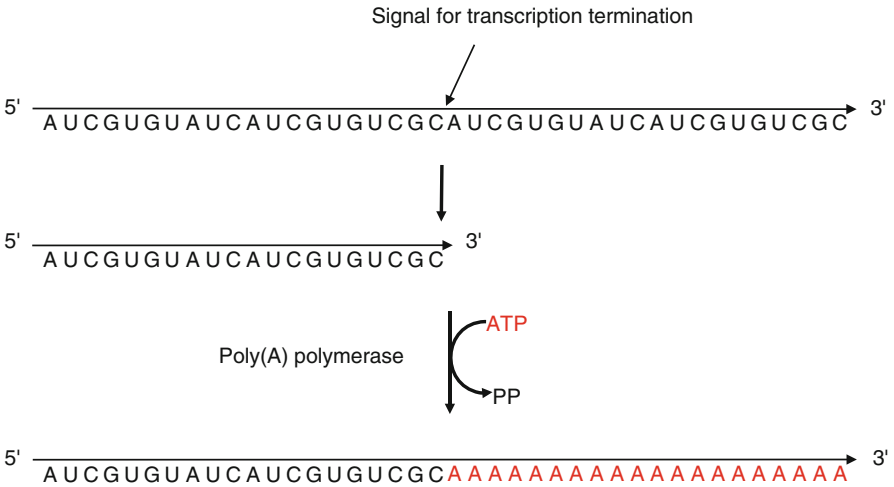


Fig. B.5 Formation of the poly(A) tail at the 3′ end of mRNAs. When the transcription complex meets a characteristic signal sequence, RNA polymerase II stops transcription; a poly(A) polymerase then intervenes to add adenylic nucleotides at the 3′ end of the newly synthesised RNA

- The *cis* regulatory DNA sequences are often designated with the suffix RE (responsive element) such as CRE (*cAMP-responsive element*, Chap. 6) or HRE (*hormone-responsive element*) which recognise ligand-bound nuclear receptors (Chap. 14). Some are responsible for an increase in transcription rate and are called *enhancers*; others are responsible for a decrease in transcription rate and are called *silencers*.
- Regulatory proteins synthesised from *trans* regulatory sequences contain a DNA-binding domain that generally binds DNA at the level of the major groove (where are localised accessible base substituents, not engaged in interstrand complementarity) and an effector domain for transcription regulation. In addition, they may contain a signal-sensing domain able to bind specific ligands that can activate or inhibit their function. These regulatory proteins are generally called *transcription factors* and should be distinguished from the general transcription factors TFII of the transcription machinery. About 2,600 proteins contain at least one DNA-binding domain and most of them are considered to be transcription factors.

Several characteristic DNA recognition motifs are found in transcription factors, the most common being *helix–loop–helix* (HLH), *zinc-finger* and *leucine zipper* motifs. Regulatory proteins often function as dimers, trimers or tetramers. The *helix–loop–helix* motif is comprised of two perpendicular α helices separated by a β turn. One of the helices is inserted in the major groove and establishes contacts with the reactive base atoms. *Zinc fingers* consist in the folding of the peptide chain around a Zn^{2+} ion, bound by coordination to two cysteine residues and two histidine

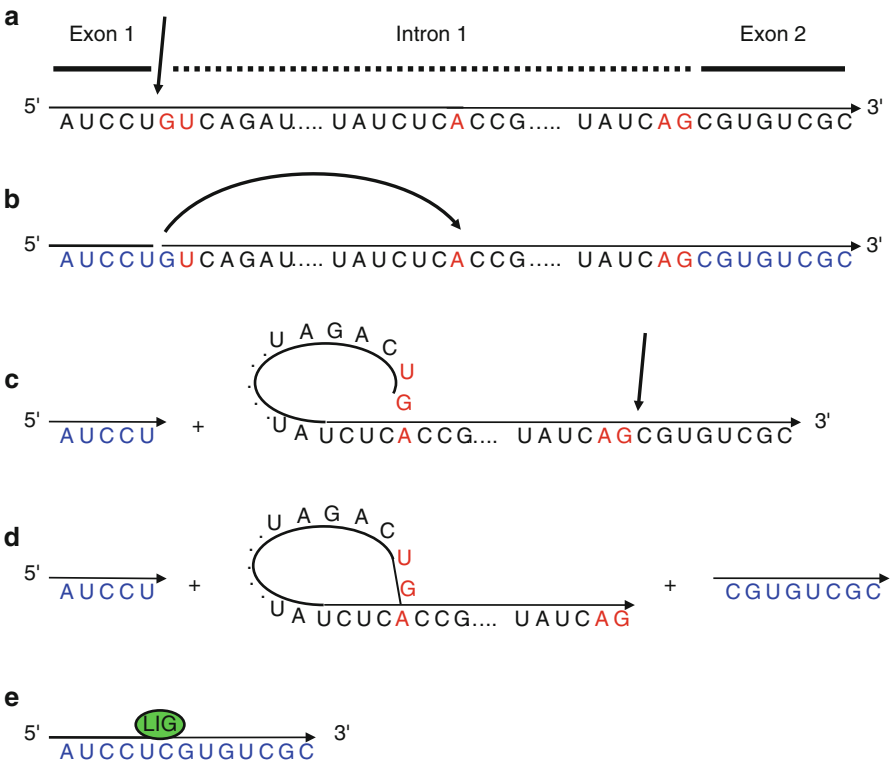


Fig. B.6 Splicing mechanism. Introns are characterised by the presence (1) of a 'donor' site GU at their 5′ end, (2) of a 'receiver' site AG at their 3′ end and (3) by a branching point A, located 40 nucleotides upstream the receiver site. Splicing is carried out by a ribonucleoprotein complex, the *spliceosome*, which recognises these characteristic signals. The events are then as follows: intron excision at its 5′ end (**a**), RNA folding (**b**), reconnection of this end at the branching point and intron excision at its 3′ end (**c**), intron elimination as a lasso (**d**) and reconnection of the 3′ end of the upstream exon to the 5′ end of the downstream exon (**e**). Donor and receiver nucleotides are in *red*, the exonic nucleotides are in *blue* and the intronic nucleotides in *black*

residues (*Cys_2His_2 zinc fingers*). Several consecutive zinc fingers enable the protein to coil up spirally around DNA, allowing contacts between an α helix and the major groove. There is a variant in nuclear receptors, two α helices being separated by a gap where a Zn^{2+} ion is bound by coordination to four cysteine residues (*Cys_4 zinc fingers*). Finally, *leucine zippers* enable contacts between the *C*-terminal hydrophobic, leucine-rich ends of two proteins to form a homo- or a heterodimer, the α helices coiling up one around the other and DNA binding occurring at the *N*-terminal end of the α helix, in the DNA major groove (Fig. B.7).

Transcription factors can interact with the gene promoter, located upstream the transcription initiation site; they can also interact upstream, inside or downstream the transcribed region, interacting with the transcription initiation site thanks to a

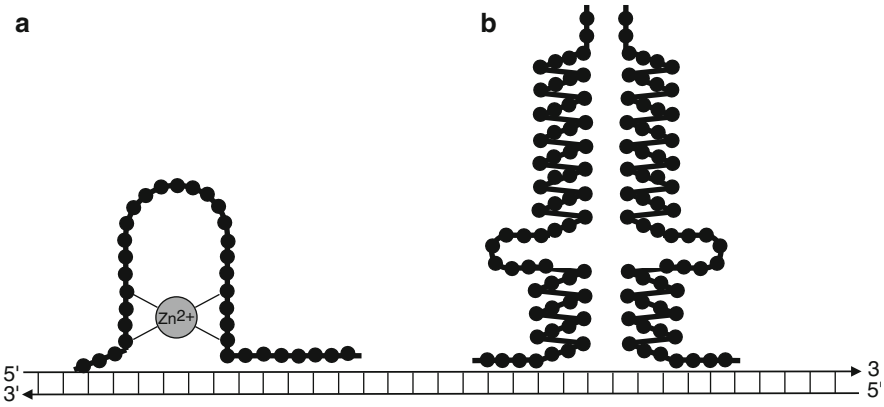


Fig. B.7 Structures of DNA-binding domains. Transcription factors carry precise protein motifs enabling them to recognise and bind specific DNA sequences, especially at the promoter level. (**a**) zinc-finger structure; (**b**) helix–loop–helix structure

folding of the DNA chain. Activator proteins recognising enhancer sequences directly interfere with the transcription initiation site and facilitate transcription through enhancement of TFII factors binding to the promoter, while repressor proteins recognising silencer sequences can interfere either with the binding site of RNA polymerase to prevent its action or with the binding site of activator proteins.

Numerous cofactors (coactivators and corepressors) can interact with transcription factors. These cofactors generally act by recruiting proteins able to remodel chromatin by histone modifications (Sect. B.5), essentially histone deacetylases and histone acetyltransferases. All these regulatory factors are gene products whose transcription is regulated by other factors, and they can undergo post-translational modifications able to activate or deactivate them (Annex C). This emphasises the complexity of transcription regulation and the subtlety of its fine-tuning.

There have been many ways for classifying transcription factors: mechanistic classification takes into account the type of protein–DNA interaction involved; functional classification relies on the ways used for their activation: some transcription factors are constitutively active in cells and others are activated by extracellular messages (nuclear receptors, Chap. 14), or by internal signals, often generated by transduction cascades leading eventually in many cases to their phosphorylation: ETS (Chap. 2), STAT (Chap. 4), SMAD (Chap. 5) and NOTCH (Chap. 8); structural classification takes into account their mode of interaction with DNA (helix–loop–helix, zinc finger, leucine zipper, etc.). There are a number of oncogenic alterations of structure or function of transcription factors, which have been presented in the previous chapters.

Transcription ends up with the synthesis of the messengers that will govern protein synthesis in the cytoplasm. It is possible since the beginning of the 2000s to obtain a snapshot of the relative amount of each transcript present in a cell population or a tissue at a given moment. This has been made possible thanks to the

completion of genome sequencing, the progress in molecular biology techniques and the bioinformatics tools required for high-throughput analysis. Oncology has especially benefited from these expression microarrays, which enable comparisons between tumour and healthy tissue, between sensitive and drug-resistant tumours, etc. From these expression profiles, it has been possible to establish molecular signatures or portraits characteristic of a given situation, which sometimes appear easier to establish than to interpret. Nosological tumour classification, risk of metastatic recurrence and treatment response prediction are the main fields that have been explored, with the aim of helping physicians to refine a diagnosis, to establish a prognosis and to personalise a treatment. Numerous validation clinical trials should be implemented before these approaches become routinely used.

B.3 DNA Methylation

DNA methylation of cytosine residues located in the promoter region of a gene represents an important regulation mechanism of gene expression. The level of DNA methylation allows the distinction between 'active' genes, i.e. transcribed genes, and 'inactive genes', not transcribed. This constitutes a key determinant of cell differentiation: in a tissue where a protein should not be expressed, the promoter of the corresponding gene undergoes a high level of methylation. The gene expression profile of a tissue is, at least in part, the reflect of the methylation level of the gene promoters.

DNA methylation is carried out by DNA methyltransferases on the C^5 of cytosine residues belonging to CG doublets, generally clustered in areas of several dozens of nucleotides called 'CpG islands'. Through a maintenance DNA methyltransferase, DNMT1 (*DNA methyltransferase 1*) using *S*-adenosylmethionine as a substrate, cytosine methylation is transmitted from a cell to its daughter cells as if it was genetic; it is said to be *epigenetic*. DNA replication can only incorporate non-methylated cytosine residues into DNA; this is after DNA replication that DNA methyltransferase adds a methyl group on the cytosine complementary to the guanine residue immediately upstream the methylated cytosine residue of the template strand (Fig. B.8).

The methylcytosine residues of a gene promoter may directly inhibit transcription by preventing the recognition of promoter sites by transcription factors. Several transcription factors, such as AP2, MYC, CREB, E2F or NFκB, recognise CpG island-containing sequences, and their action is inhibited by methylation of these islands. More generally, methylcytosine residues indirectly inhibit transcription, because they are recognised by proteins which induce chromatin remodelling and compaction, through histone modifications. These proteins belong to the MBD (*methyl CpG-binding domain protein*) family. Chromatin compaction reduces promoter recognition by the general transcription factors and highly reduces transcription rate. Conversely, chromatin remodelling influences DNA methyltransferase activity, so that histone deacetylation is accompanied by DNA demethylation,

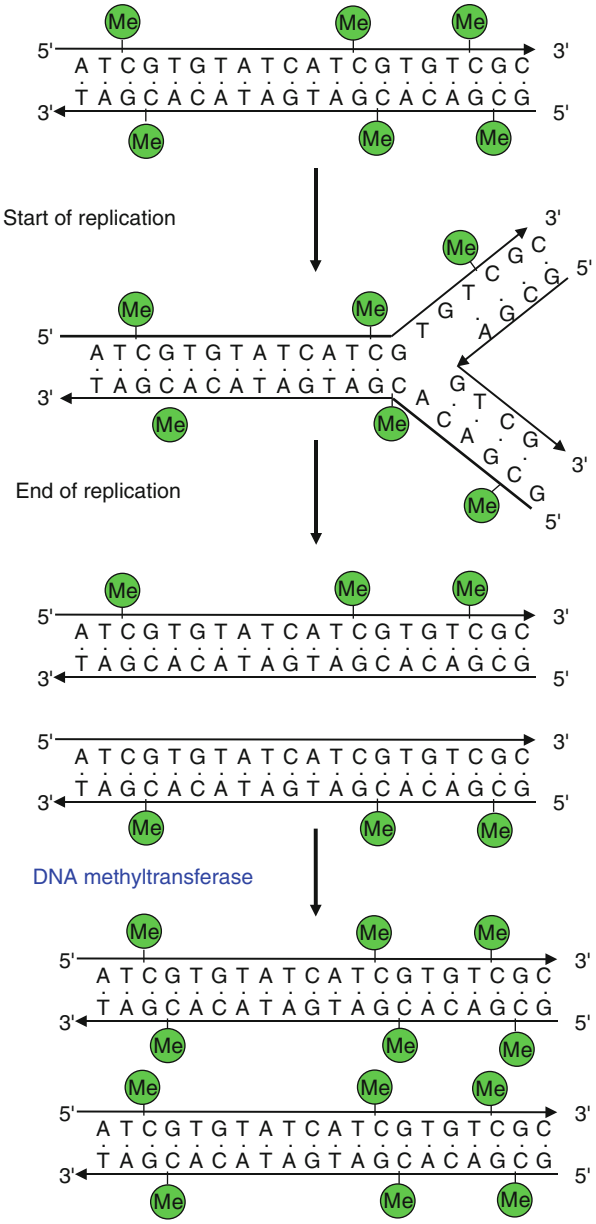


Fig. B.8 DNA methylation. In the promoting regions of numerous genes, the cytosine residues immediately upstream a guanine residue undergo methylation on their C^5. Thanks to a maintenance DNA methyltransferase, cytosine methylation is transmitted from a cell to its daughters as if it were genetic. Replication can only incorporate unmethylated cytosine residues in DNA; after replication, the DNA methyltransferase adds a methyl group on the cytosine residue complementary to a guanine residue next to the methylated cytosine residue of the template strand

demonstrating information exchanges between chromatin and DNA, which concur to gene expression regulation.

In cancer cells, there is overall DNA hypomethylation, especially at the level of repetitive sequences, and hypermethylation localised to certain promoters, especially at the level of tumour suppressor genes. Overall hypomethylation would be a factor of genomic instability, and localised hypermethylation is certainly a process for maintaining tumour suppressor gene expression at low level: it represents, therefore, an important contribution to oncogenic mechanisms. Numerous genes are hypermethylated in cancer, generally in a specific way; among them are genes involved in cell cycle control (*CDKN2A* or $p16^{INK4a}$, *CDKN1A* or $p21^{CIP1}$, *RB1* [*retinoblastoma protein 1*], Chap. 17), DNA repair (*BRCA1* [*breast cancer 1*], *MGMT* [*methylguanine methyltransferase*], *MLH1* [*MutL homologue 1*], Annex A), apoptosis (*DAPK* [*death-associated protein kinase*], Chap. 18), angiogenesis, metastatic spreading (E-cadherin, Chap. 7), signalling pathways (*APC* [*adenomatous polyposis coli*], Chap. 7), etc. The methylation of a DNA repair gene, *MGMT*, is associated to a low repair of the lesions generated by some anticancer drugs such as temozolomide, which potentiates their anticancer activity.

Pharmacological agents that induce DNA demethylation are used in the treatment of malignant diseases such as myelodysplastic syndromes. These are especially nucleoside analogues, 5-azacytidine, decitabine (5-aza-2′-deoxycytidine) and zebularine. These agents are incorporated into DNA during replication and irreversibly block DNMT1, preventing this enzyme to carry out the maintenance genome methylation during replication. The DNA demethylation they induce may however reactivate the oncogenes which were precisely repressed by methylation. Non-nucleosidic inhibitors of DNA methyltransferases and inhibitors of MBD2 are under study. In addition, since DNA methylation can repress transcription through histone deacetylation, histone deacetylase (HDAC) inhibitors can behave as demethylating agents (see next section).

B.4 Histones and Chromatin Structure

Chromatin (defined in Annex A) consists of DNA and structurally and functionally associated proteins, mainly histones, and participates to gene expression regulation. The elementary motifs of chromatin are the nucleosomes, which form linearly chained particles all along DNA, like a string of pearls, with regular spacing of about 180 nucleotide pairs. Nucleosomes are comprised of a cylindrical central core around which DNA is coiled up. This core is a histone octamer: a tetramer $(H3–H4)_2$ surrounded by two H2A–H2B dimers, histone H1 being engaged in internucleosomal interactions (Fig. A.4). In the nucleus, chromatin is present either under condensed form (heterochromatin) in non-transcribed gene regions or under relaxed form (euchromatin) in actively transcribed regions (Fig. B.9).

Histone *N*-terminal extremities that protrude out of the nucleosome can be post-translationally modified by methylation, acetylation and phosphorylation (see Annex C). This gives to each histone molecule a specific mark (*histone code*) generating a high number of possible variants (Fig. B.10). Transcription activation or

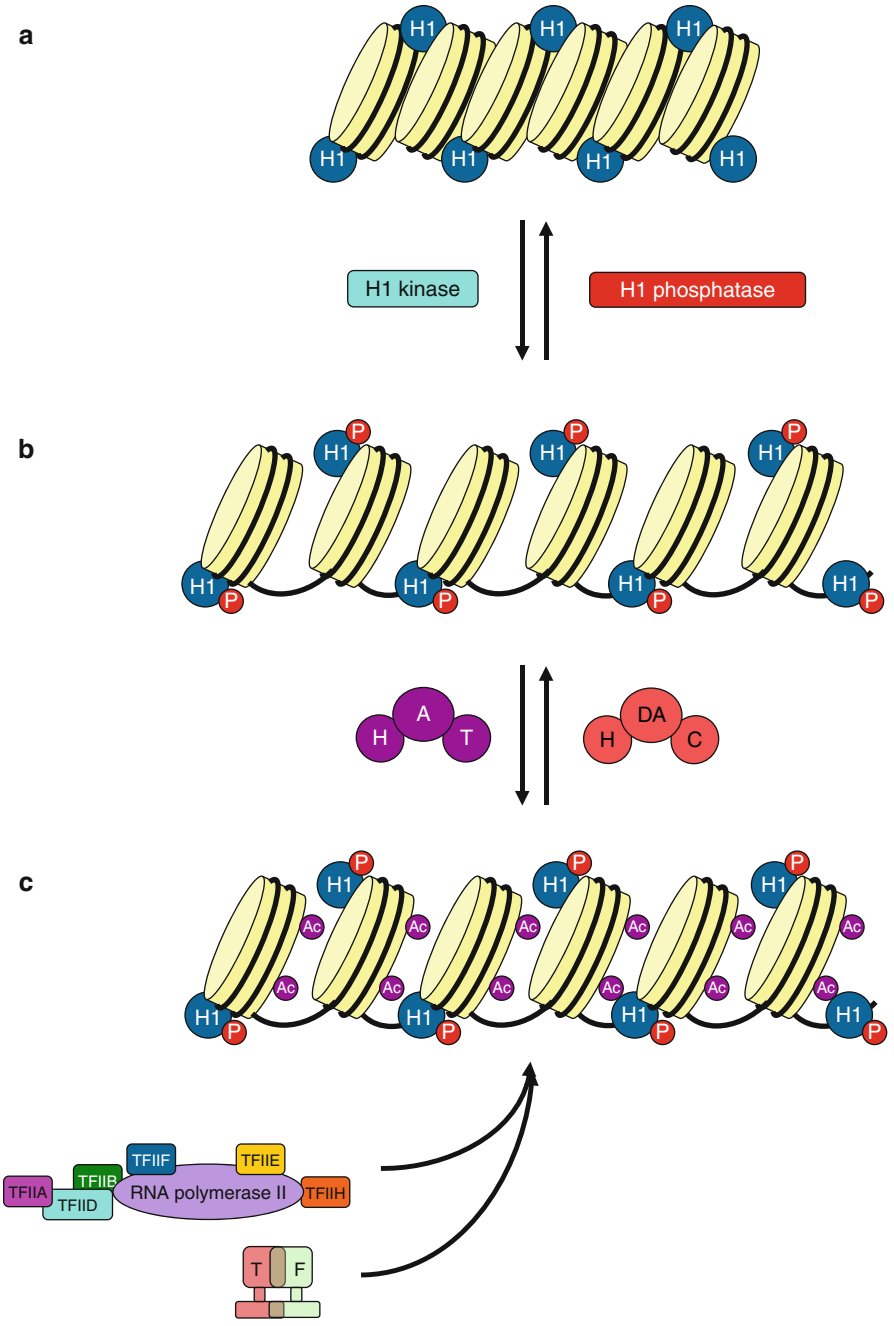


Fig. B.9 Chromatin-directed transcription regulation. (**a**) 'Inactive' chromatin has a compact structure; nucleosomes do not allow the access of the transcription machinery to the DNA template. (**b**) Histone H1 phosphorylation contributes to chromatin relaxation and transcriptional activity. It is carried out by specialised kinases while dephosphorylation is ensured by phosphatases. (**c**) Acetylation of histones 2A, 2B, 3 and 4 also enables transcription, thanks to chromatin decondensation and accessibility of the sequences to be transcribed by the transcription machinery and the transcription control protein factors. Acetylation is carried out by histone acetyltransferases (HAT) and deacetylation by histone deacetylases (HDAC)

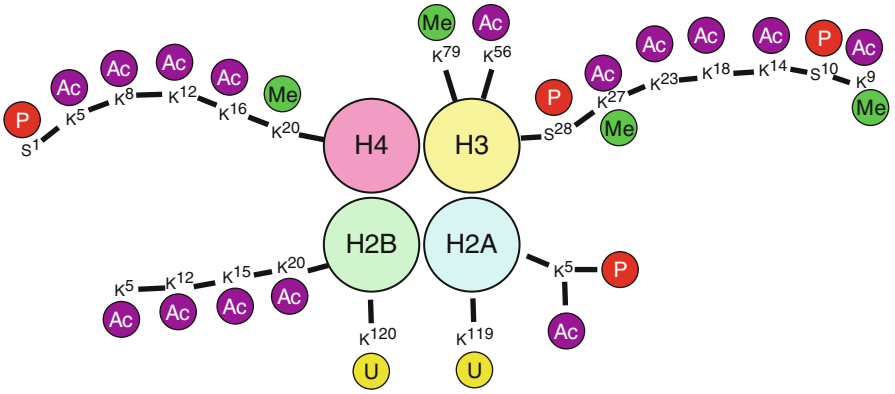


Fig. B.10 Post-translational histone modifications. Histones can undergo numerous post-translational modifications on the *N*-terminal portions, at the level of lysine residues (methylation, acetylation) or serine residues (phosphorylation). They can also undergo ubiquitinylation on several amino acid residues (Annex C)

repression factors recruit the histone-modifying enzymes (methyltransferases and demethylases, acetyltransferases and deacetylases, kinases and phosphatases) at the level of specific genes and thus define histone marks profiles in the gene environment. The histone code is complex, because of the high number of possible sites on each histone molecule and of the combination of the modifications that can be made. This histone marking is reproduced in tissues, from a cell to its daughter cells, and represents, with DNA methylation, another type of epigenetic transmission of information. Specialised proteins, equipped with histone substituents–recognition domains, are in charge of the interpretation of the signals brought by histone marks; in particular, proteins bearing specific domains called *bromodomains* (BET, *bromodomain and extraterminal protein*) can bind histone acetylation marks and participate to chromatin remodelling.

The marking profile of histones is usually abbreviated, with indications of the histone, the amino acid residue and its position in the protein chain and the modification undergone. For instance, H2BK120ac means 'histone 2B, Lys^{120}, acetylation'. Such profiles have been established in some cell types and associated to high, medium or low expression levels of the corresponding gene. For instance, the marks associated to active transcription are histone H3 trimethylation of Lys^{79} (H3K79me3), histone H4 methylation of Lys^{20} (H4K20me1, H4K20me2, H4K20me3) and histone H3 acetylation of Lys^{4}, Lys^{9}, Lys^{18}, Lys^{27} and Lys^{36} (H3K4ac, H3K9ac, H3K18ac, H3K27ac, H3K36ac) and of histone H4 Lys^{5}, Lys^{8} and Lys^{91} (H4K5ac, H4K8ac, H4K91ac). Conversely, the histone H3 trimethylation marks of Lys^{4}, Lys^{9} and Lys^{27} (H3K4me3, H3K9me3, H3K27me3) and of histone H4 Lys^{20} of (H4K20me3) correspond to repression of gene transcription. More generally, a high level of histone acetylation is often associated to the maintenance of an undifferentiated, pluripotent state, while a high level of histone methylation is associated to cell differentiation and the loss of expression of the corresponding genes.

A class of proteins called *polycomb* proteins is involved in post-translational histone modifications. These proteins form transcription repressor complexes that bind specific gene promoter sites (PRE, *polycomb responsive elements*), where they are recruited by transcription factors or by non-coding RNA molecules of about 200 nucleotides called ncRNA. Two polycomb protein complexes are found in mammals: PRC1 (*polycomb repressive complex 1*) and PRC2, which is the support of the enzymatic trimethylation of H3K27. Polycomb proteins induce a permanent repression of the expression of their target genes, which are often involved in tissue development and differentiation, such as the negative regulators of cell cycle of the INK4 and the CIP–KIP families (Chap. 17). The promoters of these genes are frequently hypermethylated; repression of gene expression by DNA methylation and polycomb proteins are likely to be associated in many circumstances, especially for the maintenance of adult tissue differentiated state.

Important alterations of the various histone modifications are found in cancer cells. For instance, the promoter of *CDKN2A* ($p16^{INK4a}$) is hypermethylated on H3K9 and hypomethylated on H3K4 in colorectal cancers. It is impossible to list the specific alterations and generate general features, but some trends can be brought up. A reduction in H4K16ac and H4K20me3 is common during the development of many tumours, especially in repetitive sequences, in association with CpG island DNA methylation. Associations between such histone modifications and cancer evolution and prognosis are actively sought. The origin of such histone marking alterations may lie at the level of polycomb proteins, which are frequently mutated or overexpressed in cancers following gene amplification or translocation. For instance, this is the case in many tumour types for EZH2 (*enhancer of zeste homologue 2*), which is a lysine *N*-methyltransferase.

The level of histone acetylation on lysine residues plays a major role on gene transcription. It is governed by the activity of two enzyme types, histone acetyltransferases (HAT or KAT) and histone deacetylases (HDAC). As a general feature, there exists a trend to HDAC overexpression during oncogenesis and cancer progression. The modulation of histone acetylation by histone deacetylases (HDAC) may represent an interesting pharmacological target, and several natural and synthetic inhibitors have been developed, the first one being valproic acid, now replaced by hydroxamic acid derivatives and cyclic peptides. Several HDAC inhibitors (HDI) are under clinical evaluation and have an indication in the treatment of cutaneous T lymphomas. They induce cell cycle arrest, growth inhibition, cell differentiation and apoptosis in several types of cancer cells. Their mechanism of action involves the alleviation of the downregulation of tumour suppressor genes such as *CDKN2A* ($p16^{INK4a}$), *CDKN1A* ($p21^{CIP1}$) or *RB1*. However, they have pleiotropic cellular actions and it is not proven that histone deacetylation inhibition is their unique, or even principal, mechanism of action. They could potentiate cytotoxic activity of alkylating or platinating agents through facilitating DNA accessibility by chromatin decompaction.

Histone methylation of lysine residues and less often of arginine residues does not play an unequivocal role on transcription, whose stimulation or inhibition depends on the precise amino acid involved. The level of methylation is governed by two

types of enzymes, histone methyltransferases (HMT or KMT) and histone demethylases (HDM or KDM). Multiple alterations of these enzymes may play an oncogenic or a tumour suppressor role and have been identified in various cancer types: mutations, translocations, overexpressions, etc. These enzymes can also represent relevant targets for cancer treatment, and several inhibitors have been identified. The specificity of the role of methyl adducts does not favour their development.

Histone phosphorylation is ensured by multiple kinases, especially those that are involved in proliferation signalling (AKT, JAK2). The activity of the inhibitors of these kinases may depend, at least in part, on their action on histones. Finally, histone ubiquitinylation is catalysed by various E3 ubiquitin ligases, with variable specificity.

B.5 Chromatin Maintenance

In addition to the numerous enzymatic proteins involved in histone post-translational modifications, several protein complexes are able to modulate chromatin structure. These protein complexes can interfere with transcription factors, modulate gene expression and contribute to proliferation and differentiation. Instead of modifying covalently the structure of nucleosomes, as histone modifiers do, they reshape nucleosome organisation, using the energy provided by ATP hydrolysis. One of the most important protein complexes is the SWI/SNF complex. The name of this complex comes from yeast mutants (*switch* and *sucrose non-fermentable*) of genes encoding some subunits of the ortholog proteins. These complexes are connected to the actin cytoskeleton, and β-actin (gene *ACTB*) and actin-like proteins 6 (genes *ACTL6A* and *ACTL6B*) are members of the SWI/SNF complexes.

SWI/SNF complexes are composed of 12–15 subunits and have a large size (2 MDa). In addition to β-actin and actin-like proteins, they contain an ATPase subunit, either BRM (*Brahma*, gene *SMARCA2* for *SWI/SNF-related matrix-associated actin-dependent regulator of chromatin*) or BRG1 (*BRM-related gene 1*, gene *SMARCA4*), and a series of subunits that are either mutually exclusive such as SMARCD1/D2/D3 or unique such as SMARCB1, SMARCC1 and SMARCE1. Two types of complexes are found: the BAF (*BRG1-associated factors*) complexes that are structured around ARID1A (*AT-rich interactive domain-containing protein 1A*) or ARID1B and the PBAF (*polybromo BRG1-associated factors*) complexes that are structured with ARID2 and contain in addition the proteins BDR7 (*bromodomain-containing*) and PBRM1 (*polybromo 1*), absent from the BAF complexes. Due to the various homologous proteins that can be associated, there are hundreds of theoretically possible complexes, so that the exact structure of each complex is variable and tissue dependent.

Several genes encoding the SWI/SNF protein subunits are mutated in cancers, especially *SMARCB1* and *ARID1A*, and these mutations are found in ovary, lung, breast, pancreas, kidney and other cancers. As much as 20 % of cancers present at least one mutation of a SWI/SNF gene. In all cases, it appears that these mutations are invalidating, i.e. that these genes are tumour suppressor genes. The mutations

are sometimes biallelic, but in most cases they concern only one allele, suggesting that their oncogenic effect is related to haploinsufficiency. It has been suggested that these complexes act as DNA protectors and that their alterations generate a mutator phenotype, as is the case for mutations of DNA repair genes. However, the fact that rhabdoid tumours often harbour a unique *SMARCB1* mutation is not in favour of this hypothesis; the oncogenic mechanism associated to their deleterious mutations is rather to be found at the level of the control of specific transcriptional programmes, so that disruption of chromatin structure may directly alter the signalling pathway using such programmes. Several signalling pathways are affected by mutations occurring in SWI/SNF genes: the inactivation of SMARCB1 induces p16INK4a downregulation (Chap. 17), so that it has been proposed to prescribe CDK4 inhibitors to patients with such mutations in a synthetic lethality approach; *BRG1* deletion results in downregulation of WNT receptors and similar approaches have been suggested to target the specific vulnerabilities of SWI/SNF mutant tumours.

B.6 Micro-RNAs

Micro-RNAs (miRNAs or MIRs) are small non-coding RNAs of 19–24 nucleotides that are involved in the regulation of the expression of at least 30 % of the human genome. Their total number is around 1,000 different molecular species. About 70 % of MIR genes are localised within the introns (sometimes the exons) of coding genes and 30 % in intergenic regions. They are transcribed by RNA polymerase II as sequences of several kb, equipped as all mRNAs with a MeGTP cap and a poly(A) tail. These primary precursors (pri-MIRs) are then matured by cleavage carried out by a nuclear RNase called *drosha* (gene *RNASEN*) to generate pre-MIRs of about 65–85 nucleotides. These pre-MIRs have a hairpin-like structure with a double-strand rod and a loop. They are exported to the cytoplasm by *exportin* (gene *XPO5*). Pre-MIRs bind then another RNase called *dicer* (*double-stranded RNA-specific endoribonuclease*, gene *DICER1*) in a complex called RISC (*RNA-induced silencing complex*) which also contains other proteins, in particular TRBP (*transactivation-responsive RNA-binding protein*, gene *TARBP2*) and *argonaute* (genes *AGO1* and *AGO2*). MIRs are generated by the RNase dicer as double strands; the active strand remains attached to the RISC while the antisense strand is hydrolysed (Fig. B.11).

Mature MIRs of 19–25 nucleotides can recognise, through their 2–8 first 5′ nucleotides (the *seed* region), complementary sequences borne by target mRNAs at the level of their 3′ untranslated end (3′UTR). The target mRNA thus binds the RISC, resulting either in mRNA degradation or in the inhibition of its translation into proteins (Fig. B.12). The MIR seed sequence is only partially complementary to the target 3′UTR mRNA sequence, so that the same MIR can inhibit the translation of hundreds of different mRNAs; algorithms allow to predict which messengers are able to recognise a given MIR. Conversely, a given mRNA may be regulated by multiple MIRs. MIRs exert therefore multiple functions and they can participate to the regulation of all cellular activities. Similarly to the mRNA profiles that can be

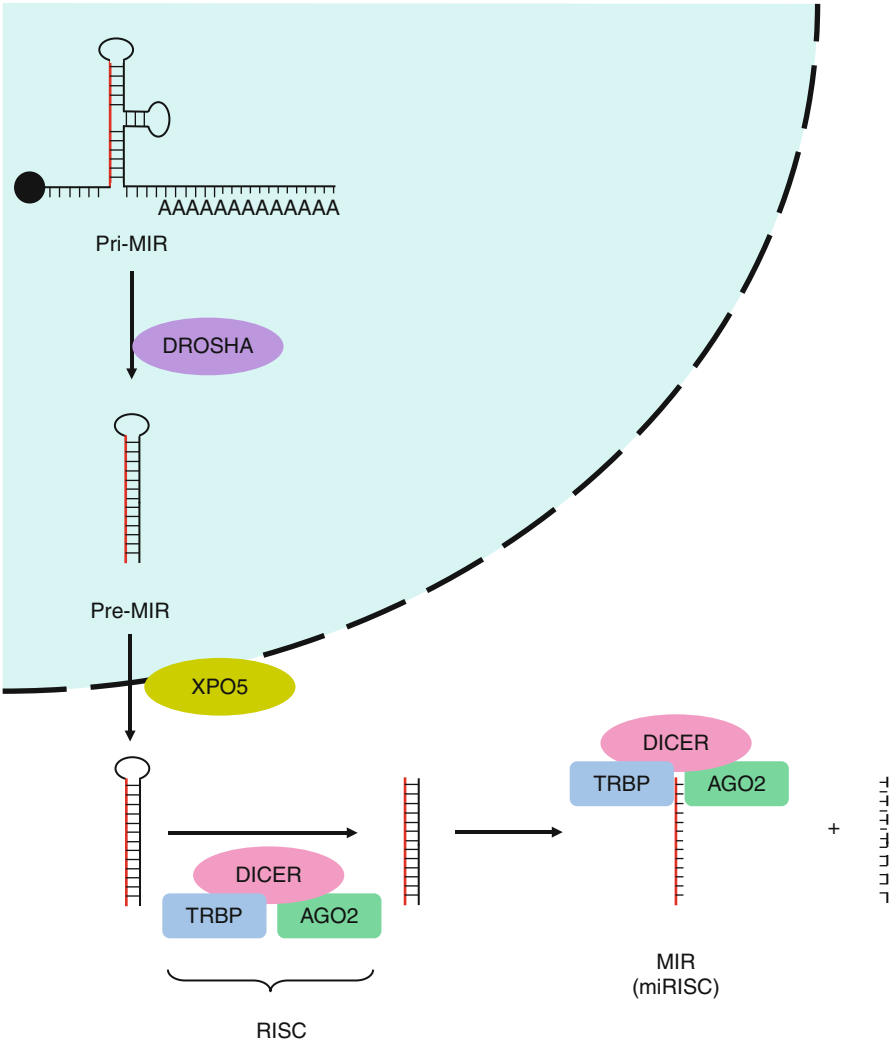


Fig. B.11 Generation of micro-RNAs (MIRs). Transcription of DNA-specific sequences generates special messenger RNAs, called pri-MIRs, with the usual cap and the poly(A) tail. Pri-MIR maturation begins in the nucleus where they are converted into a double-stranded hairpin structure, pre-MIRs, thanks to a nuclease called *drosha*. They are the exported into the cytoplasm by the transporter exportin (*XPO5*) and converted by another nuclease called *dicer*, in a protein complex called RISC (*RNA-induced silencing complex*), which eliminates the antisense strand and remains attached to the MIR

established for a given tissue in a given situation, MIR expression profiles have been established for various cell types. Qualitative and quantitative MIR alterations are able to perturb cell functions.

MIRs involvement in oncogenesis has been recognised soon after their discovery. MIRs can act similarly to oncogene or tumour suppressor genes, depending on their target. MIRs of the LET7 family behave as tumour suppressor genes: they can

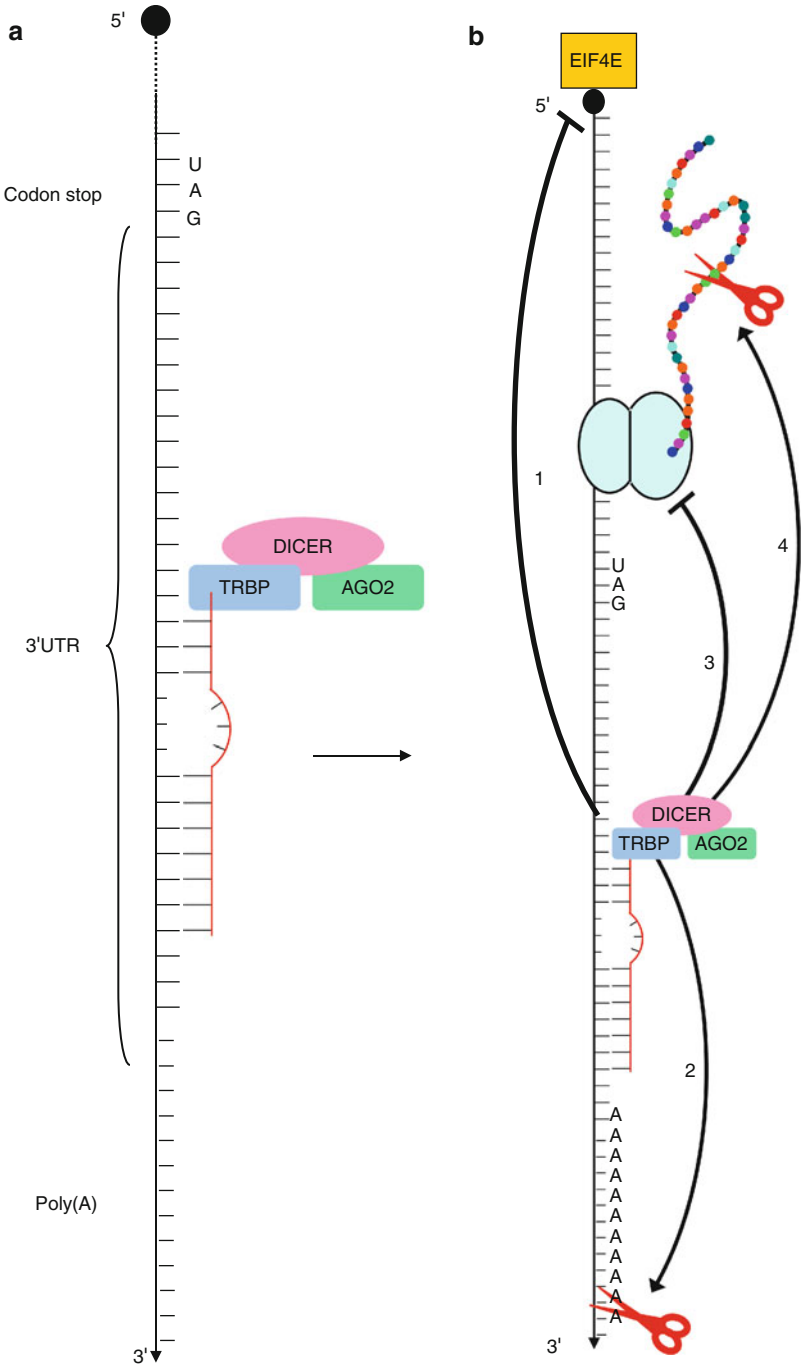


Fig. B.12 Mechanism of action of micro-RNAs (MIRs). (**a**) MIRs imperfectly hybridise mRNA sequences, generally in 3′UTR regions, between the codon stop and the poly(A) tail. (**b**) RISC is able (*1*) to inhibit the translation initiation step, (*2*) to hydrolyse the poly(A) tail of the mRNA, (*3*) to inhibit the translation elongation step and (*4*) to hydrolyse the neosynthesised protein

post-transcriptionally repress *RAS* gene expression (Chap. 2) and can induce cell cycle arrest and cell death. On the contrary, MIR214, which represses *PTEN* expression (Chap. 3), behaves as an oncogene. MIRs intervene in the regulation of most proliferation and cell death signalling pathways. MIR genes undergo, as oncogenes and tumour suppressor genes, qualitative (mutations) and quantitative (expression level) variations that explain their role in oncogenesis. For instance, the transcription factor MYC activates a cluster of MIR genes involved in the regulation of the expression of transcription factors of the E2F family. A complete description of MIR functions in oncogenesis is beyond the limit of this book.

The establishment of MIR expression profiles and the identification of their common sequence variations may lead in the future to their routine use in diagnostic, prognosis and prediction of response of cancers to treatment. MIRs may be targeted by oligonucleotides called *antagomirs* and ways allowing therapeutic applications are actively researched. MIR discovery has enabled the implementation of a general strategy of specific gene expression inhibition, by using short RNA molecules able to interfere with mRNAs that have been called s*mall interfering RNAs* (siRNAs). siRNAs recognise specific mRNA sequences and induce RISC activity to hydrolyse them. This technique is now widely used experimentally in molecular biology to specifically inhibit gene expression. Therapeutic applications in oncology will become available when siRNAs will be formulated as drugs, i.e. when efficient ways for cell entry will be discovered.

B.7 Alternative Splicing

We have already mentioned that splicing was required for the maturation of eukaryote mRNAs, because of the discontinuous gene structure, where exons and introns follow each other. One of the main features is that splicing can intervene on several sites of the mRNA, so that variations exist in the choice of the exons that are maintained when introns are eliminated. These variations concern at least 90 % of genes and are a general rule, not an exception; the ancient dogma—'one gene, one protein'—has been shown to be inexact when alternative splicing was discovered. We do not describe here the mechanisms that govern the choice between 'constitutive exons' and 'alternative exons' during splicing. They involve various nuclear proteins, especially proteins with serine–arginine domains (*SR proteins*) and hnRNP (*heterogeneous nuclear ribonucleoproteins*), which bind activator or repressor exon domains or primarily transcribed intron domains. The activation of such proteins is the outcome of various signalling pathways; alternative splicing is governed by extra- or intracellular signals.

Several elementary splicing modalities exist (Fig. B.13): (1) a given exon (cassette exon) is kept or not; (2) one or the other of two mutually exclusive exons is kept; (3) there is competition between two GU splicing donor sites; (4) there is competition between two AG splicing receiver sites; (5) a usually intronic sequence is kept in mature mRNA. In addition to these alternative splicing events, two processes also generate distinct transcripts of the same gene: the existence of alternative promoters,

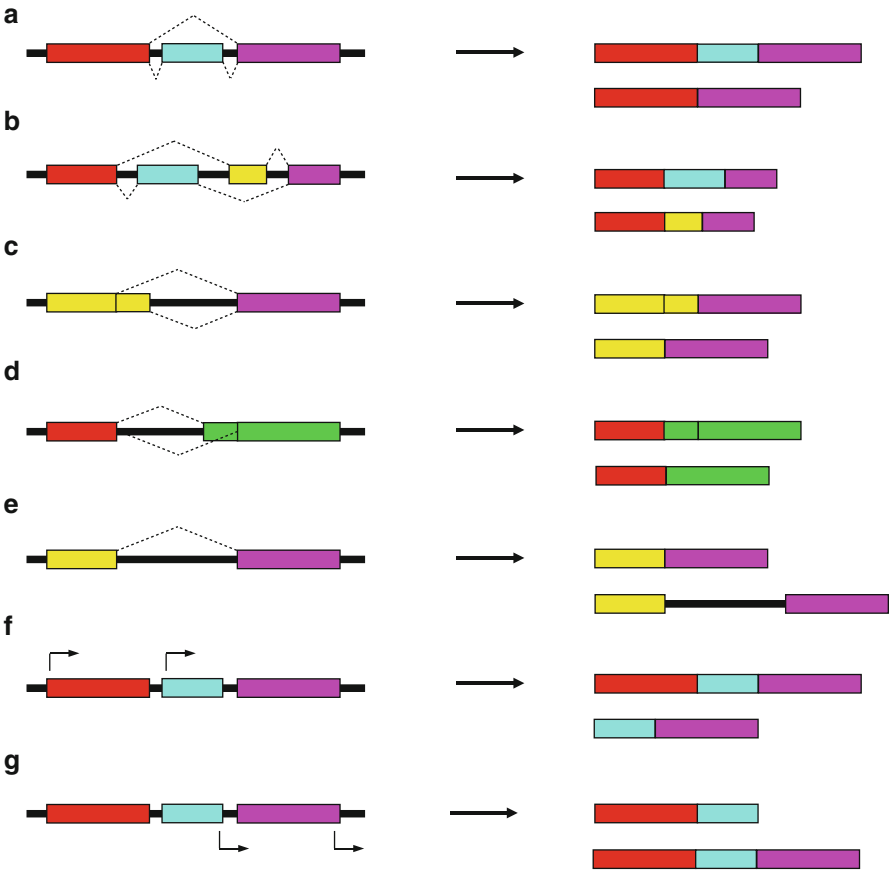


Fig. B.13 Modalities of alternative splicing. Alternative splicing may allow to exclude an exon (**a**), to replace an exon by another (**b**), to use a given donor site or another (**c**), to use a given receiver site or another (**d**), to keep an intronic structure usually spliced (**e**), to use a given promoter or another (**f**), to use a site of polyadenylation or another (**g**)

which alter the *N*-terminal sequence of the protein, and the existence of alternative polyadenylation sites, which modify the *C*-terminal sequence of the protein.

There are multiple alterations of alternative splicing in cancers; some of them are actually oncogenic, others concur to oncogenesis without playing a crucial role. High-throughput approaches have been developed to make a complete directory of splicing alterations in cancer cells, in order to provide splicing profiles analogous to the various profiles of cancer molecular alterations presented in the previous sections. Such alterations may have various origins:

- They may concern the proteins responsible for the splicing operations or their control, such as SR proteins. The overexpression of SFRS1 (*splicing factor, arginine/serine-rich 1*), involved in splicing of signalling proteins of the MAP

kinases (Chap. 2) or the PI3 kinase (Chap. 3) pathways, is associated to oncogenesis, and this protein behaves as a true oncoprotein.

- Mutations in splicing sequences can lead to pathological alternative splicing, with integration or elimination of sequences that should not be integrated or eliminated. Such mutations have been described in the *APC* (*adenomatous polyposis coli*) gene (Chap. 7) or in the *BRCA1* gene, involved in DNA repair (Annex A). Such special oncogenic mutations occur with a non-exceptional frequency.
- Finally, alternative splicing is one of the basic mechanisms allowing the adaptation of a synthesised protein to its function at each step of development. Reactivation of the production of 'normal' protein isoforms outside the right period may also be oncogenic. This is the case for RAC1, a small G-protein found in colorectal cancers under the variant form RAC1b containing 19 supplementary amino acids resulting from an exon that is usually eliminated in adult tissues for the synthesis of the RAC1a isoform.

Alternative splicing can be considered as a potential therapeutic target in oncology. It is possible to interfere with the proteins responsible for splicing regulation, thanks to small molecules selected by high-throughput approaches. It is also possible to target the splicing sites or the splicing regulatory sites on primary mRNAs with antisense approaches. The common problem is always the entry of oligonucleotides into the target cells where they are supposed to act. BCLX (gene *BCL2L1*) is a member of the BCL2 protein family (Chap. 18), which is antiapoptotic in its long isoform ($BCLX_L$) obtained after using a given splicing site and proapoptotic in its short isoform ($BCLX_S$) obtained by the use of another splicing site, located upstream the first one. Blocking the downstream splicing site could contribute to cell death in tumour cells expressing $BCLX_L$. Another example is that of the alternative splicing of the messenger of *FGFR1* (*fibroblastic growth factor receptor 1*, Chap. 1), for which blocking a repressor splicing sequence has been proposed as a therapeutic approach.

Annex C: Control of Protein Activity

Abstract

This third annex describes the events occurring after transcription; it considers the mature mRNAs at the cytoplasmic sites, where they are translated into proteins, and describes afterwards the post-translational protein modifications which allow them to ensure their activities, up to their eventual destruction. Proteins are the effectors of all the orders and programmes contained in the genome; multiple processes of regulation of their activity should be implemented for rendering them able to carry out the orders they receive from inside or outside the cell. Whereas a total of about 23,000 genes is enough for programming the life of a mammal, the total number of proteins that can exist may exceed one million; through alternative splicing, first, as described in Annex B; through post-translational modifications, such as phosphorylations, glycosylations, prenylations and other covalent adducts; and through truncation by a variety of proteinases.

Along the different chapters of this book, we have seen how the regulation processes of protein activity are important for signal transduction; this annex aims at describing the general rules that govern the implementation of protein activity regulation, following them from synthesis to destruction. A variety of excellent molecular biology textbooks can be used to deepen these basic notions. These processes are often altered in cancers, as exemplified by the fact that many enzymes that catalyse phosphorylation, called kinases, are oncogene products.

C.1 Translation Machinery

The existence of four different nucleotides, in DNA as in mRNAs, which should correspond to the 20 amino acids necessary for protein structure, requires the existence of a code of at least three nucleotides per amino acid. A code of two nucleotides would only allow 16 combinations, while a code with three nucleotides allows 64 combinations. Such a three-nucleotide code is actually utilised to allow the correspondence between DNA sequence and protein sequence. The genetic code is degenerated: several 3-nucleotide codons may correspond to the same amino acid. When an amino acid corresponds to several codons, the two first nucleotides are

J. Robert, *Textbook of Cell Signalling in Cancer: An Educational Approach*,
DOI 10.1007/978-3-319-14340-8

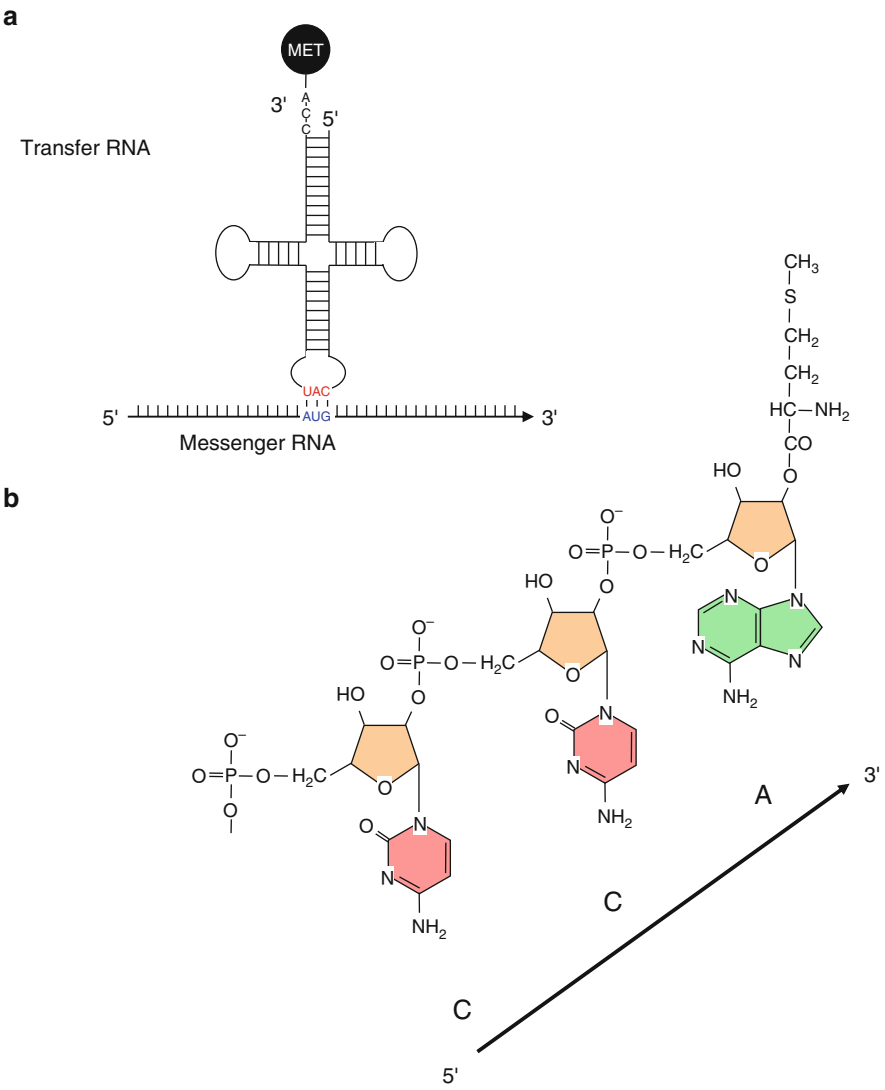


Fig. C.1 Transfer RNA structure. (**a**) tRNAs adopt a folded, cloverleaf-shaped structure, thanks to nucleotide pairings that confer this secondary structure. The three *red* nucleotides represent the anticodon, which recognises the matching codon (*blue*) on the mRNA. At the 3′ end, the three nucleotides consists always of two cytidylic acid residues and one adenylic acid residue, which is bound to the amino acid via an ester bond. The example presented here is the tRNA of methionine ($tRNA^{met}$), characterised by the anticodon CAU which recognises the initiator codon AUG. (**b**) The 3′ terminal sequence CCA and its binding to methionine

generally the same, and the replacement of the third by mutation has no consequence on protein sequence. Three special codons do not specify an amino acid but are used for synthesis termination (*stop* codons, UAA, UAG and UGA). There exist in theory three possible reading frames, but only one is actually used: it is identified by the codon of translation initiation (codon AUG). In addition to the

Table C.1 The genetic code

1st	2nd U		2nd C		2nd A		2nd G		3rd
U	UUU	phe	UCU	ser	UAU	tyr	UGU	cys	U
	UUC	phe	UCC	ser	UAC	tyr	UGC	cys	C
	UUA	leu	UCA	ser	UAA	stop	UGA	stop	A
	UUG	leu	UCG	ser	UAG	stop	UGG	trp	G
C	CUU	leu	CCU	pro	CAU	his	CGU	arg	U
	CUC	leu	CCC	pro	CAC	his	CGC	arg	C
	CUA	leu	CCA	pro	CAA	gln	CGA	arg	A
	CUG	leu	CCG	pro	CAG	gln	CGG	arg	G
A	AUU	ile	ACU	thr	AAU	asn	AGU	ser	U
	AUC	ile	ACC	thr	AAC	asn	AGC	ser	C
	AUA	ile	ACA	thr	AAA	lys	AGA	arg	A
	AUG	met	ACG	thr	AAG	lys	AGG	arg	G
G	GUU	val	GCU	ala	GAU	asp	GGU	gly	U
	GUC	val	GCC	ala	GAC	asp	GGC	gly	C
	GUA	val	GCA	ala	GAA	glu	GGA	gly	A
	GUG	val	GCG	ala	GAG	glu	GGG	gly	G

Two usual abbreviations are used for the amino acids constitutive of protein structure: the 3-letter one used in this table and the one-letter abbreviation explained below

Alanine	ala	A	leucine	leu	L
Aspartic acid	asp	D	lysine	lys	K
Glutamic acid	glu	E	methionine	met	M
Arginine	arg	R	phenylalanine	phe	F
Asparagine	asn	N	proline	pro	P
Cysteine	cys	C	serine	ser	S
Glutamine	gln	Q	threonine	thr	T
Glycine	gly	G	tryptophan	trp	W
Histidine	his	H	tyrosine	tyr	Y
Isoleucine	ile	I	valine	val	V

mRNA template, translation utilises small RNAs, called transfer RNAs (tRNAs), which bring on the one hand the amino acid required for the synthesis of the peptide chain and on the other hand the genetic information, carried out by a nucleotide triplet called the anticodon, enabling it to recognise on mRNA the matching codon. Figure C.1 presents the general structure of a tRNA and Table C.1 the genetic code.

Translation takes place in voluminous ribonucleoprotein complexes called *ribosomes*, which associate ribosomal RNAs (rRNAs) and ribosomal proteins. Ribosomes are comprised of two subunits, the large one (L) and the small one (S). They carry several binding sites for other RNAs: one for mRNA and three for tRNAs. The tRNAs that bring a novel amino acid bind first to the site A; this amino acid is then attached to the peptide chain in the course of elongation, in the P site, and the tRNA that has lost its amino acid is expelled through the E site (Fig. C.2). Translation begins by binding, on the small ribosomal subunit, of a particular initiator tRNA bearing the initiation anticodon, CAU (read from 5′ to 3′) and bound to a methionine. Several

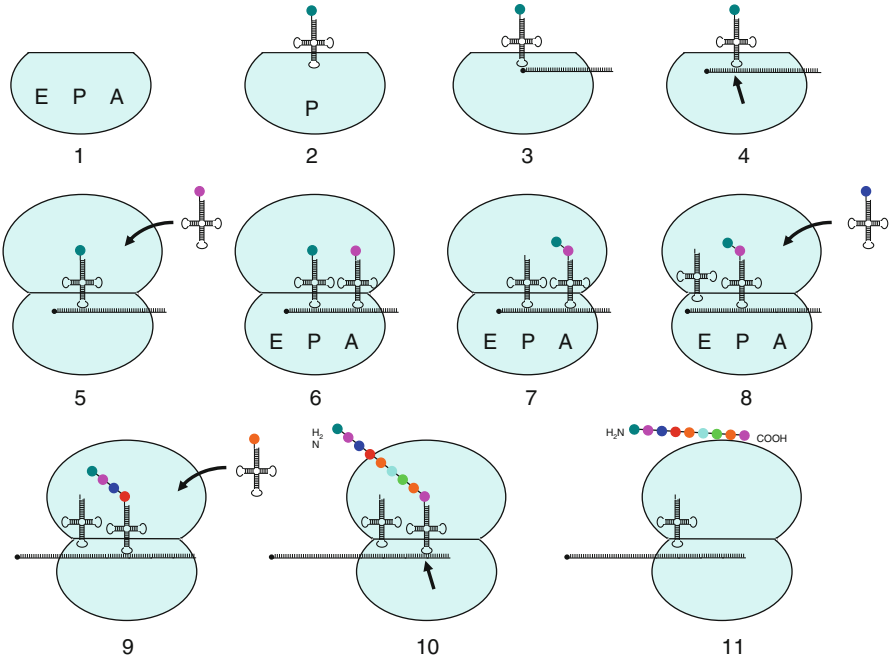


Fig. C.2 Translation steps: from mRNA to protein. Initiation: the small ribosome subunit (*1*) binds the initiator tRNA on site P (*2*) as well as proteic factors EIF (not represented here). The small ribosome subunit binds then to the 5′ end of a mRNA (*3*), which is recognised thanks to its cap, and slides along mRNA up to the first AUG sequence, recognised by the anticodon of the initiator tRNA (*4*, *arrow*). The large ribosome subunit binds to the small one (*5*) to form the complete ribosome. *Elongation*: A tRNA carrying the anticodon that recognised the codon following the initiator codon AUG, and the corresponding amino acid binds to the ribosome site A (*6*). A peptidic bond is formed between the –COOH of the upstream amino acid and the $-NH_2$ of the downstream amino acid (*7*). The binding between the upstream amino acid and its tRNA is disrupted and the tRNA migrates to the ribosome site E, while the tRNA of the following amino acid, carrying now the growing peptide chain, migrates to ribosome site P (*8*). The ribosome then slides by one step, i.e. by three nucleotides, along mRNA; tRNA is removed from site E, a novel tRNA carrying a novel anticodon and a novel amino acid can then bind to site A so that a new cycle starts again (*9*). This is reproduced as many times as the peptide chain will contain amino acid residues. Elongation protein factors EEF (not represented here) are required for these processes. *Termination*: the end of translation occurs when the ribosome meets a stop codon (*arrow*) on the mRNA, not recognised by any tRNA anticodon (UAA, UAG, UGA), which gives a translation termination signal (*10*). Termination protein factors recognise the vacant ribosome site A; the polypeptide chain is detached from the tRNA and the ribosome (*11*). The mRNA is detached from the ribosome, and the two ribosome subunits are dissociated and recycled for further protein synthesis

initiation protein factors EIFs (*eukaryotic initiation factors*) also bind the ribosome: EIF2, comprised of three subunits, α, β and γ, then EIF4E and EIF4G, whose function is regulated by phosphorylation (see Chap. 3). The small ribosomal subunit binds then to the 5′ end of a mRNA, identified thanks to its 7-methylguanosine 5′-trisphosphate cap (Annex B), and slides along mRNA up to the first AUG codon recognised by the initiator tRNA. The mRNA region upstream this AUG codon is not translated and is known as the 5′ UTR. EIFs are then released from the ribosomal small subunit and are replaced by the large ribosomal subunit to constitute the complete ribosome.

A three-step cycle is then undertaken to enable, from the initial methionine, the synthesis of the polypeptide chain. A tRNA bringing the anticodon that recognises the codon immediately downstream the initiator AUG codon binds to the ribosomal site A; a peptide bond is formed between the –COOH of the upstream amino acid (the first one being methionine) and the $–NH_2$ of the downstream amino acid. This bond is catalysed by the peptidyltransferase activity of the large ribosomal subunit. The bond between the upstream amino acid is disrupted and the corresponding tRNA migrates to the ribosomal E site, while the tRNA that brought the new amino acid, now bringing an elongating peptide chain, migrates to the P site. The ribosome then moves forward of one notch, i.e. three nucleotides, along the mRNA; the tRNA of the E site is expelled; a novel tRNA, bringing a novel anticodon and the corresponding amino acid, can then bind to the site A and the cycle can start again. Elongation protein factors EEF (*eukaryotic elongation factor*) watch over the accuracy of the translation process and help the movements of tRNAs from site A to site P and from site P to site E. Figure C.2 presents the elongation cycle of proteins.

The ending of the translation process is ensured by the encounter of a stop codon, which is not recognised by any tRNA, giving thus a translation termination signal. Termination protein factors identify the empty site A; peptidyltransferase cannot transfer any amino acid on the free –COOH extremity of the polypeptide and the chain is detached from the tRNA and the ribosome. mRNA is in turn detached from the ribosome, whose subunits are dissociated to be then reutilised for the synthesis of a new protein. In fact, a given mRNA molecule is simultaneously 'read' by several ribosomal complexes which regularly progress along the polynucleotide chain, giving to the particle the characteristic aspect of a string of pearls called the *polysome*.

Translation control mechanisms are added to transcription control mechanisms (Annex B) to adapt protein synthesis to the needs and possibilities of the cell. Translational repressors binding to the 5′ end of a given mRNA can hinder its reading until the encounter of a ligand that induces their detachment when the cell is required to produce the corresponding protein. Another regulation mechanism is carried out by the phosphorylation of the EIF initiation and EEF elongation factors. EEFs are small G-proteins that can be inhibited by phosphorylation. Finally, a particular regulation occurs at the level of certain mRNAs which carry sequences, most often in the 5′ UTR, able to adopt a secondary structure enabling them to bind to translation initiation factors (EIF4G): these are called IRES (*internal ribosome entry site*) sequences (Fig. C.3).

IRES sequences are used as starting points for protein translation distinct from translation initiated from the first AUG codon met on mRNA, usually called the *cap-dependent translation*. Most of the known IRES have been found in the mRNAs that encode proteins associated to the control of proliferation and cell death. IRES-dependent translation is used for the translation of mRNAs encoding transcription factors MYC, JUN (Chap. 2), SMAD5 (Chap. 5), p53 (Chap. 17) and HIF1α (Chap. 16); apoptosis regulators XIAP, APC and APAF1 (Chap. 18); growth factors FGF2, PDGFA and IGF2 (Chap. 1); and various receptors such as ERα (Chap. 14), IGF1R (Chap. 1) and NOTCH2 (Chap. 8). This modality of translation initiation may constitute a shortcut enabling rapid synthesis of proteins required for mitosis or cell death, especially when the cap-dependent translation initiation factors are absent or inhibited. This process is used by some viruses to enable priority synthesis of their own proteins over the proteins of the infected cell: their mRNAs are equipped with

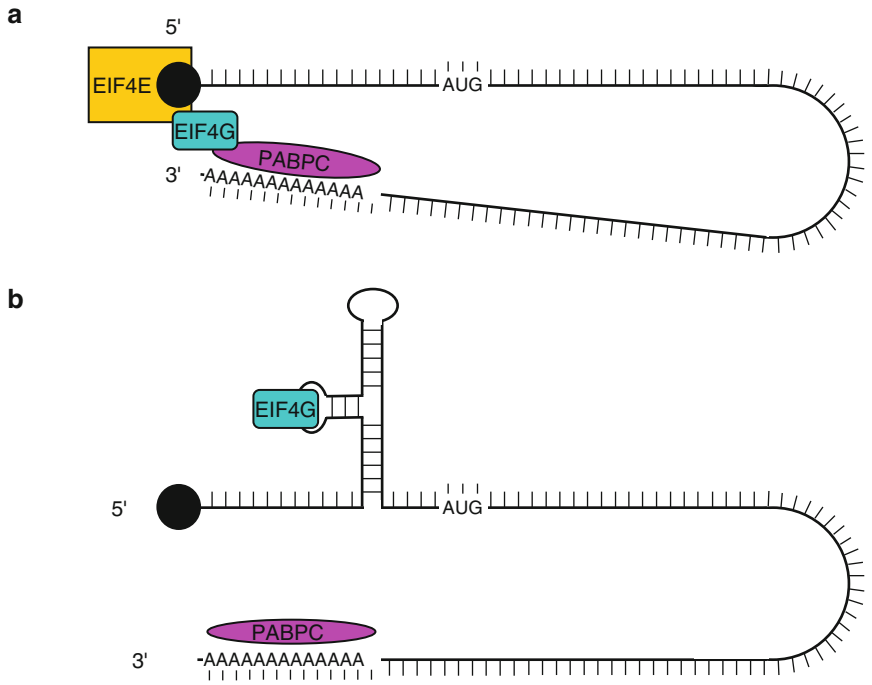


Fig. C.3 Cap-dependent translation and IRES-dependent translation. (**a**) Cap-dependent translation requires the intervention of the initiation factors EIF4, whose availability depends on the intervention of complex regulations involving in part the mTOR pathway (Chap. 3). (**b**) IRES-dependent translation can be more rapidly implemented; IRES are mRNA secondary structures able to activate translation after the simple recognition of an initiation factor EIF4G. PAPBC: *polyA-binding protein*, *cytoplasmic*

an IRES sequence and their synthesis can therefore take place without cap-dependent factors. This explains why alterations of these mechanisms can contribute to oncogenesis.

In multiple myeloma, an increase in MYC expression has been attributed to a mutation of the 5′ UTR sequence of the corresponding mRNA, at the level of an IRES; this mutation would stabilise the IRES and facilitate the depending translation. As a general feature, the research of mutations within the IRES of genes encoding oncoproteins such as p53 may reveal fruitful. In addition, protein factors called ITAF (*internal initiation trans-acting factors*) are required for IRES-dependent translation initiation, in a more or less mRNA-specific way. Qualitative and quantitative alterations of these proteins could well play a role in IRES-dependent translation and be involved in oncogenesis: this is the case for the EIF4G1 factor whose overexpression could explain the overproduction of δ-catenin, a cell adhesion molecule able to bind E-cadherin. Such an overexpression has been observed in a high percentage of metastatic breast cancers and might play a role in invasion processes.

C.2 Protein Conformation and Protein–Protein Interactions

C.2.1 Regulation of Protein Conformation

Proteins are made of linear sequences of amino acids bound together by covalent bonding. The protein amino acid sequence (called its *primary structure*) is determined from the gene DNA nucleotide sequence, via the information transmitted by mRNAs. This protein sequence is sufficient to characterise the protein univocally. However, nothing can be deduced from this sequence about the functions of the protein if its spatial structure is not known. This three-dimensional structure is of course the consequence of the amino acid sequence but involves numerous interaction forces between the lateral chains of the amino acids, so that the most sophisticated softwares are only partially able to predict the spatial structure of the protein from its amino acid sequence. They are able to predict some arrangements of short sequences of amino acids (α helices, β sheets) that constitute *secondary structures*, but the relative arrangement of these domains and their interactions, which represents the *tertiary structure* of the protein, can only be deciphered, thanks to complex physical techniques such as X-ray diffraction of the crystallised protein.

Proteins are synthesised linearly; this synthesis is rapidly followed by protein folding, which involves non-covalent interactions, to which may be added disulphide bridges between two cysteine residues of the polypeptidic chain by a *protein disulphide isomerase* (PDI), which generates an important stabilisation of the 3D structure of the proteins. These non-covalent bonds are:

- Ionic bonds between a positive charge carried by a basic amino acid (lysine, arginine, histidine) and an acidic amino acid (aspartic and glutamic acids)
- Van der Waals interactions, electrostatic in nature, which occur between charge fractions, either permanent or induced by interactions between neighbour atoms
- Hydrogen bonds between amino acids carrying polarised functions that exert attractive or repulsive effects as a function of their electronic properties (serine, threonine, cysteine, proline, etc.)
- 'Hydrophobic' bonds which are due to privileged interactions between aliphatic and aromatic nonpolar amino acid residues that exclude the presence of water molecules (glycine, alanine, valine, leucine, isoleucine, phenylalanine)

Protein folding leads in principle to the adoption of a final minimal energy status; however, chain length and the complexity of the interactions cannot allow the simple computation of a minimal energy status; folding is also a function of the 'history' of the synthesis, i.e. the order of amino acid addition and the interactions that occur during the course of synthesis. It is known for more than a century, for instance, that heating leads to a loss of protein spatial structure: non-covalent bonds are disrupted, and cooling, even if progressive, does not generally allow the recovery of the native protein conformation, which is required for its function. Alterations in the folding of some proteins lead to their aggregation and their insensitivity of proteolytic processes: they are found in some neurodegenerative diseases. In some cases, these protein aggregates display an 'infectious' character, able to catalyse the conversion of a

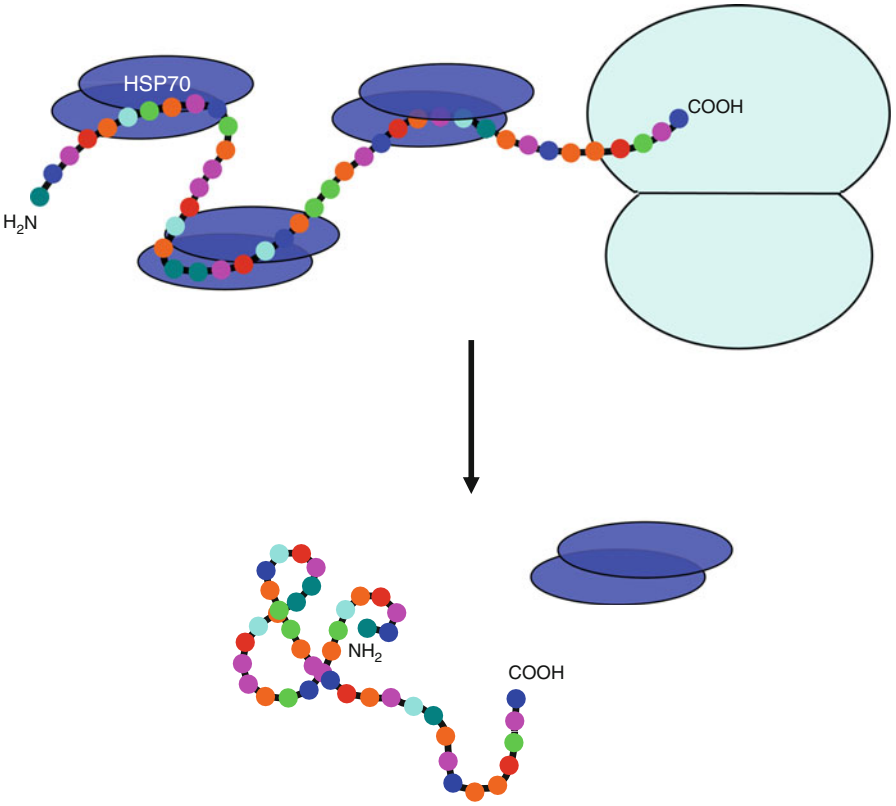


Fig. C.4 Protein folding during synthesis. During protein synthesis, HSP70 chaperone proteins interact with hydrophobic amino acid residues of the growing chain and govern its progressive folding to enable it to acquire its definitive three-dimensional conformation required for its functions

normal tertiary structure to an aberrant conformation: this is the case of the prion protein, which is the support of prion diseases such as Creutzfeldt–Jakob disease.

The optimal folding of a protein (the one allowing it to acquire its functional tertiary structure) can be guided by *chaperone proteins*, whose function is to interact with the protein during its synthesis to help it adopt the eventual spatial conformation it should present (Fig. C.4). These proteins are also involved later for protein protection against sublethal stresses that could alter tertiary structures, such as hypoxia, acidosis, osmotic shock or moderate heating: the misfolded proteins are handled by chaperone proteins that restore its native functional structure (Fig. C.5). When the alterations are irreversible, the chaperones may lead the proteins they protect (their 'clients') to proteasomal degradation. Chaperone proteins belong to an ancestral highly conserved protein family called HSP (*heat-shock proteins*), named after their molecular weight: HSP100, HSP90, HSP70, HSP40, etc. They

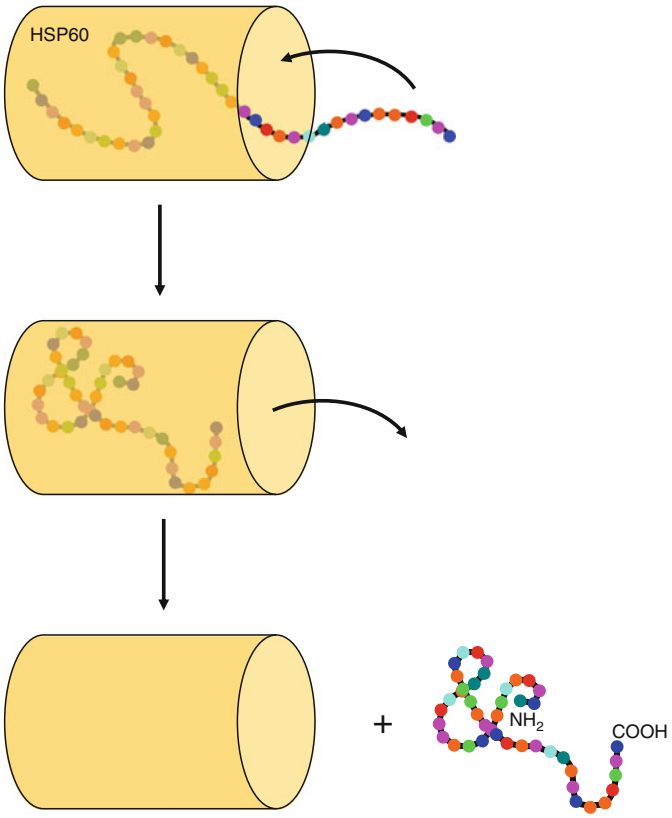


Fig. C.5 Restoration of the conformation of misfolded proteins. When proteins are misfolded, HSP60 chaperone proteins are able to enclose them as within a shell where they can recover their optimal functional conformation

undergo multiple post-translational modifications (phosphorylation, nitrosylation, acetylation). They function together with co-chaperone proteins and adapter molecules to form large-size protein complexes.

HSP genes are transcribed in response to *heat-shock factors* (HSF) which recognise on DNA the activator sequences of heat-shock response elements (HSE). The HSP mRNA transcription and maturation steps are especially rapid as HSP genes are devoid of introns. There are three HSFs in humans (HSF1, 2 and 4). HSF1 is maintained inactive, thanks to binding to HSP90, from which it is detached under stress conditions, allowing it to migrate to the nucleus, get trimerised and induce HSP genes transcription. The implementation of HSPs is relatively complex (Fig. C.6): 'client' proteins are recruited by a complex associating a chaperone HSP70 and a co-chaperone HSP40, bound together by an adapter protein HOP (*HSP organising protein*). The complex is then bound to an ADP-bound HSP90

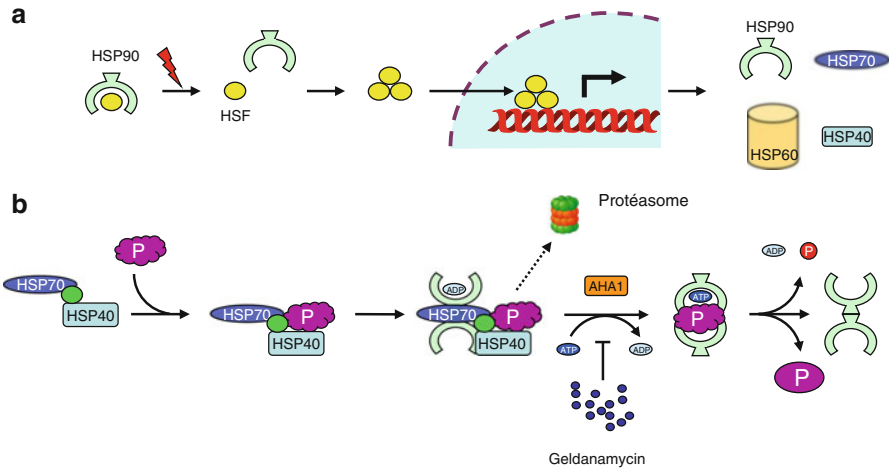


Fig. C.6 Implementation of heat-shock proteins (HSPs). (**a**) HSP genes are transcribed in response to heat-shock factors (HSFs). The transcription factor HSF1 is maintained at an inactive state through binding HSP90, from which it is detached under stress effects. It can thus migrate to the nucleus, where it trimerises and induces HSP transcription. (**b**) Bringing HSP into play occurs through a complex process: the 'client' proteins are recruited by a complex between HSP70 and a co-chaperone HSP40, bound together by an adapter protein HOP. The assembly is then attached to an ADP-bound HSP90 dimer. An ADP–ATP exchange, activated by the protein, enables a conformation change and ensures the protection of the client protein so that it can exert its function. In the absence of ADP–ATP exchange, the client protein is bound to ubiquitin and led to the proteasome. HSP inhibitors such as geldanamycin inhibit this ADP–ATP exchange and favour the destruction of the client protein

dimer. An ADP–ATP exchange, activated by the adapter protein AHA1 (*activator of HSP90 ATPase 1*), enables then conformational change and ensures the protection of the client protein for assuming its function. In the absence of ATP activation, the client protein is bound to ubiquitin and driven to the proteasome.

Numerous proteins involved in cell proliferation and survival are HSP clients: tyrosine kinase receptors such as EGFR, ERBB2, KIT and MET (Chap. 1); intracellular kinases such as BRAF, AKT and CDK4 (Chaps. 2, 3, and 17); transcription factors such as p53, HIF1α and the oestrogen receptor (Chaps. 14, 16, and 17); apoptosis proteins such as BCL2 and APAF1 (Chap. 18); as well as telomerase (Annex A). HSPs enable these proteins to endure the hypoxic and acidic conditions of the tumour environment; they even enable them to tolerate oncogenic mutations that could be lethal. HSPs are overexpressed in a variety of solid tumour types and haematological malignancies, and they participate to oncogenesis, as having a positive effect on cell proliferation and survival. HSF1 knockout in mice is associated to tumour growth inhibition.

Intensive pharmacological research aims at identifying HSP inhibitors, especially at the step of ADP–ATP exchange which allows them to protect client proteins

instead of leading them to the proteasome. The core structure for ATP exchange blocking and HSP inhibition is presented by ansamycins, such as geldanamycin, several analogues of which are being developed. Other targets could be found, for instance, at the level of the interactions between the chaperone or the co-chaperone and its client proteins or at the level of the post-translational HSP modifications. Since chaperones have many clients, inhibiting their function could lead to the interruption of several proliferation pathways and thus constitutes a multi-target approach.

C.2.2 Protein–Protein Interactions

Protein–protein interactions involve the same interaction forces as those that determine protein tertiary structure and folding: ionic bonds, van der Waals forces, hydrogen bonds and hydrophobic interactions. These non-covalent bonds enable proteins to interact with high specificity. The study of protein–protein interactions is a major field in biochemistry and allows the understanding of numerous physiological and pathological processes. Frequently, well-defined selective protein domains are involved in these interactions: SH2 domains enabling proteins to bind phosphotyrosine residues of other proteins (Chaps. 1 and 2), PH domains for the specific binding of the 3-phosphate carried by phosphatidylinositol 3,4,5-trisphosphate by AKT or PDK1 (Chap. 3), etc. Specific techniques allow to isolate the complete set of proteins interacting with a given protein, and this has led to numerous discoveries. In the signalling pathways, several examples of protein–protein interactions which play a crucial role in information transmission have been described: just to name one, such interactions are involved in the recruitment of GRB2 by activated TKRs and of SOS1 by GRB2 in the MAP kinase pathway (Chap. 2).

Mutations of the genes involved in cell proliferation and survival, or in cell adhesion and motility, induce structural modifications able to disturb protein–protein interactions and consequently to stimulate or activate the signalling pathway where the mutated protein intervenes. For the pharmacologist, protein–protein interactions are especially difficult to target, primarily because of the large interaction area between protein partners, involving many non-covalent bonds: small molecules hardly achieve the specific inhibition of these interactions with enough affinity. Peptides that mimic the structure of one of the interacting domains represent the principal pharmacological approach for the inhibition of protein–protein interactions. One can mention, as an example, the development of peptides that disturb the interaction between GRB2 and SOS1 as a mean for interrupting the information process, along the MAP kinase pathway (Chap. 2), between a tyrosine kinase receptor and the RAS protein.

C.3 Covalent Post-translational Protein Alterations

C.3.1 Hydroxylation (Fig. C.7a)

Post-translational hydroxylation of lysine and proline residues is found in collagen, a group of fibre proteins of the extracellular matrix. These hydroxylations allow the

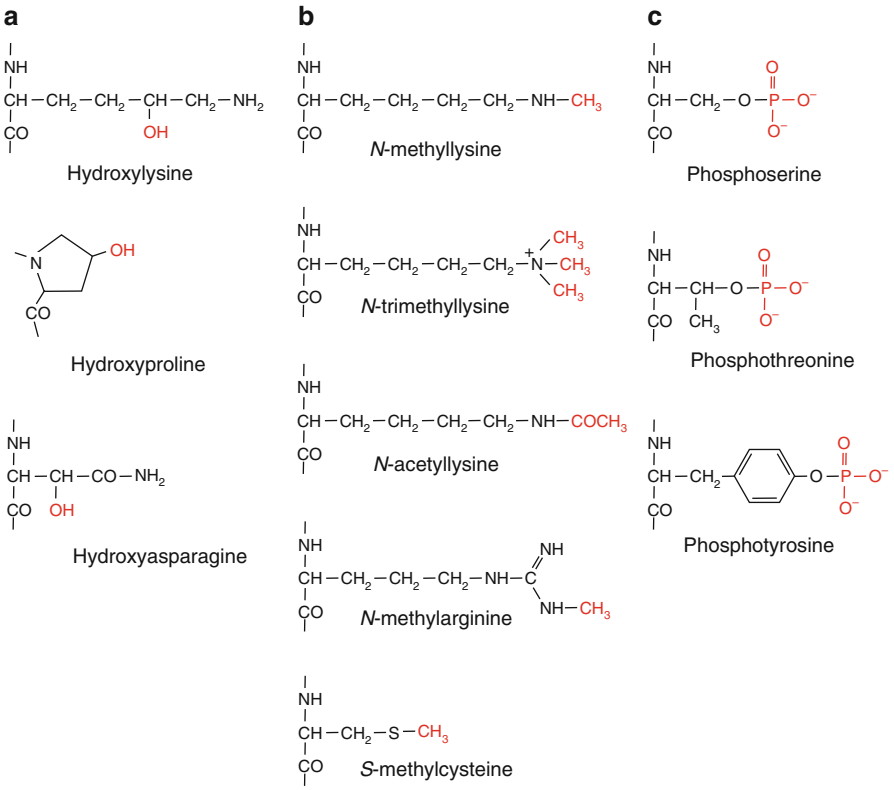


Fig. C.7 Post-translational covalent protein modifications. (**a**) Examples of hydroxylations. Hydroxylysine and hydroxyproline residues are found in collagens. The hydroxylation of two proline residues of HIF1α represents a signal driving it to the proteasome via ubiquitinylation (Chap. 16); the hydroxylation of an asparagine residue of the same transcription factor leads to an inhibition of its transcriptional activity. (**b**) Examples of methylations and acetylations. These modifications are found in particular on histone lysine residues (Annex B). The methylation of arginine residues is found in transcription factors downstream several signalling pathways, such as the JAK–STAT pathway (Chap. 4). The detoxification of methylated DNA bases by MGMT occurs by the binding of the methyl group on a cysteine residue of this enzyme (Annex A). (**c**) Phosphorylation. Kinase-mediated phosphorylation of serine, threonine and tyrosine residues represents a universal mechanism of regulation of protein activity

formation of novel hydrogen bonds between polypeptidic chains and induce the formation of triple-helix structures that confer to collagen its rigid structure.

Some hydroxylations play an important informative role: for instance, the hydroxylation of proline residues of the transcription factor HIF1α (*hypoxia inducing factor*, Chap. 16) enables its binding to an E3 ubiquitin ligase, the VHL (*von Hippel–Lindau disease*) protein, which drives it to proteasomal destruction (see Sect. C.8). These hydroxylations are catalysed by an oxygen sensor, proline hydroxylase. Another oxygen sensor, asparagine hydroxylase, catalyses the hydroxylation

of an asparagine residue of HIF1α, preventing thus its activation. In both cases, HIF1α cannot induce the transcription of the genes required for angiogenesis, such as VEGFA (*vascular endothelial growth factor*). The identification of therapeutic targets for HIF1α gene inactivation could contribute to the inhibition of tumour angiogenesis.

C.3.2 Methylation and Acetylation (Fig. C.7b)

Alkylation and acetylation are frequent post-translational protein modifications, which can intervene on several types of amino acids. Histone modifications by methyl or acetyl groups on lysine residues constitute an important mechanism for regulating their activity and thus for gene expression control, as described in Annex B. The methylation of other proteins, often involved in transcription regulation such as STAT (Chap. 4), can take place on arginine residues, thanks to protein-arginine *N*-methyl-transferases (PRMT). Methylations can also be catalysed by specific methyltransferases on histidine and aspartic or glutamic acid residues. In all cases, the methyl donor is *S*-adenosylmethionine, which is transformed into *S*-adenosylhomocysteine by this reaction. Another important example is that of methylguanine methyltransferase (MGMT), which can transfer, on one of its cysteine residues, a methyl group bound to the O^6 of a DNA guanine residue formed through the action of some alkylating agents (Annex A). This is a direct DNA repair pathway that induces definitive enzyme inactivation.

C.3.3 Phosphorylation (Fig. C.7c)

Proteins can be phosphorylated on serine, threonine and tyrosine residues, which carry a hydroxyl function that can be esterified by a phosphoric acid moiety brought by an ATP molecule whose γ phosphate moiety can be transferred. The enzymes that ensure phosphorylation reactions are called *kinases*; these are either serine/threonine kinases (428 genes have been identified in the human genome) or tyrosine kinases (90 genes). Some serine/threonine kinases are indeed dual kinases and can also phosphorylate tyrosine residues in addition to serine or threonine residues. This is, for instance, the case for MEK1/2 (Chap. 2) and WEE1 (Chap. 17). Some others are lipid kinases and can phosphorylate a phosphoinositide on the 3 hydroxyl group of the inositol moiety (Chap. 3). All kinases together constitute the *kinome* that can be represented under the form of a genealogic tree as a function of their homology. Phosphatases are the enzymes that catalyse the reverse reaction: they hydrolyse the phosphoric ester bond and release a phosphate group. There are serine/threonine phosphatases, tyrosine phosphatases and dual phosphatases on the basis of the amino acid whose phosphate adduct is released.

Phosphorylation of substrate proteins is the leading event for transduction of the signals brought by growth factors and constitutes one of the principal modes of regulation of their activity. The phosphate moiety carried by a protein, sometimes a

lipid, represents the message itself in many signalling pathways. In the 1960s, the second messenger of the action of several hormones such as insulin or growth hormone was actively researched, after the identification of cyclic AMP and trisphosphoinositol (IP3) as second messengers of glucagon and hypophysis hormones. In transduction systems, however, the phosphate group borne by a tyrosine residue of a protein or the inositol moiety of a phosphoinositide is frequently the messenger itself. Numerous kinases have been mentioned in the previous chapters: tyrosine kinase receptors (Chap. 1), serine/threonine kinase receptors (Chap. 5), cytoplasmic tyrosine kinases (Chap. 4) and serine/threonine kinases of the cell cycle (Chap. 17), of the MAP kinase pathway (Chap. 2), of the PI3 kinase pathway (Chap. 3), etc. Similarly, numerous phosphatases have been mentioned: tyrosine phosphatase receptors (Chap. 1), non-receptor tyrosine phosphatases (Chap. 3), CDC25s in the cell cycle (Chap. 17), PTEN in the PI3 kinase pathway (Chap. 3), etc.

Phosphorylation takes place on specific residues of the substrate protein and different amino acid targets may lead to different functional effects. For instance, AKT1 can be phosphorylated on Thr^{308} by PDK1 and on Ser^{473} by TORC2 (Chap. 3), both phosphorylations having activating consequences, while CDK1 can be activated by CDK7-mediated phosphorylation on Thr^{161} and inhibited by WEE1 phosphorylation on Tyr^{15} (Chap. 17). There is generally a strict specificity of the amino acid target, but sometimes several amino acid residues on a protein can be phosphorylated by the same kinase: this is the case for tyrosine kinase receptor autophosphorylation (Chap. 1). The phosphate group bears a negative charge that induces important modifications of the reactivity of the phosphorylated amino acid, which can be accompanied by modifications of protein conformation and of protein–protein interactions. Some mutations are said to be 'phosphomimetic' when they replace a neutral amino acid by an acidic one, carrying a negative charge, close to a hydroxylated amino acid that may undergo phosphorylation; such a phosphomimetic mutation is the V600E mutation of BRAF (Chap. 2).

C.3.4 Glycosylation (Fig. C.8)

Numerous proteins carry glucidic chains, either on the amino group of asparagine residues by *N*-osidic bond (*N*-glycoproteins, Fig. C.8a, c) or on the hydroxyl group of serine, threonine or hydroxylysine residues by *O*-osidic bond (*O*-glycoproteins, Fig. C.8b, d). The glucidic moieties of glycoproteins, called glycans, are generally branched, thanks to the multiple available functions carried by oses and osides. Membrane proteins often carry such glycans that are exposed to the outside of the cell; these glycans play an antigenic role and behave sometimes as a marker of various diseases, in particular cancers. For instance, the determinants of blood and tissue groups are borne by glucidic structures exposed at the surface of circulating blood cells.

In the case of *N*-glycoproteins, the glycosylation site is always comprised of the three amino acids S/T–X–N, and the oside directly bound to asparagine is always *N*-acetylglucosamine. The glycosylation reaction takes place in the lumen of the endoplasmic reticulum (ER): sugars are first successively added in a precise order, with branchings at the level of mannose residues, on a voluminous phospholipid species inserted within the ER membrane, dolichol-pyrophosphate. They are

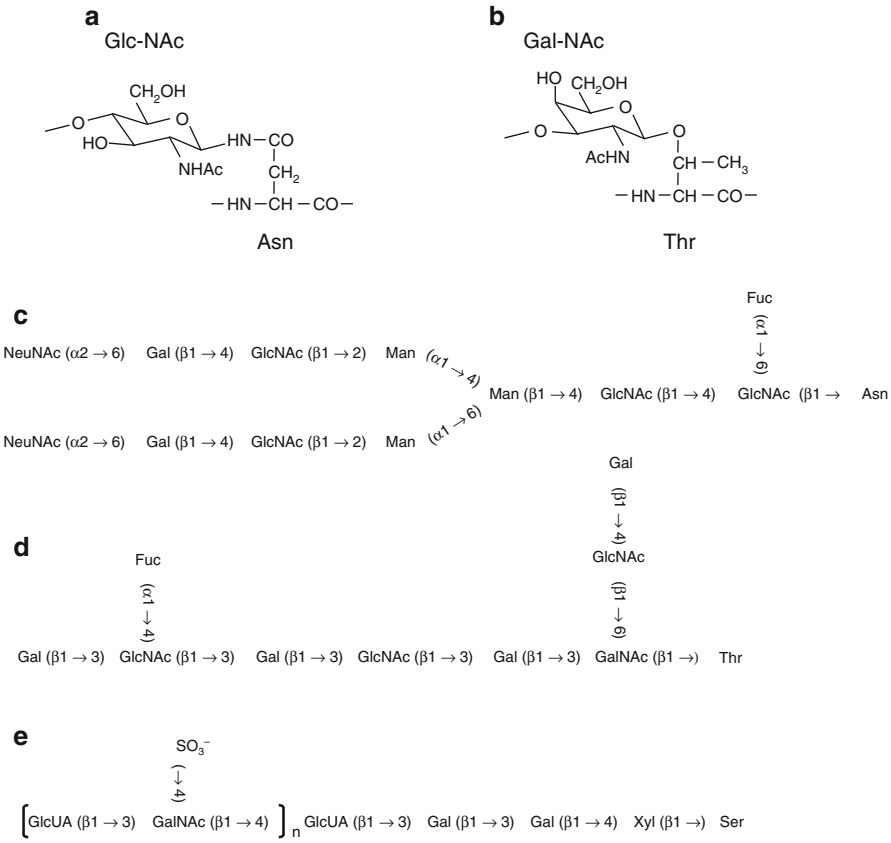


Fig. C.8 Post-translational covalent protein modifications: glycosylations. (**a**) *N*-osidic bond between an asparagine residue and a molecule of β-*N*-acetyl D-glucosamine. (**b**) *O*-osidic bond between a threonine residue and a molecule of β-*N*-acetyl D-galactosamine. (**c**) Enchainment of oses and osides in the glycan of an *N*-glycoprotein, lactotransferrin. (**d**) Enchainment of oses and osides in the glycan of a membrane *O*-glycoprotein, the protein carrying the Lewis[a] antigen marker. (**e**) Structure of a glycosaminoglycan, chondroitin sulphate A, bound through a β D-xylose to a serine residue of a protein chain, all together forming a *proteoglycan*. The characteristic motif, repeated *n* times, is comprised of D-glucuronic acid and β-*N*-acetyl D-galactosamine 4-sulphate

brought under activated forms through binding with a specific nucleotide (UDP, CMP, GDP) and each addition is catalysed by a distinct glycosyltransferase. The glycan chain is then transferred in one piece to a protein carrying a glycosylation site protruding in the ER lumen, by an oligosaccharide transferase. Purely cytosolic proteins are never glycosylated. The glycosidic chain of the protein is then reshaped during the maturation of the glycoprotein and its transfer to the Golgi and then to its final destination; some sugars, such as glucose, are removed while others are added, such as galactose, fucose and *N*-acetylneuraminic acid, which give to the protein its specific properties, especially antigenic. An example of a glycan structure of a *N*-glycoprotein is given Fig. C.8c.

In the case of *O*-glycoproteins (Fig. C.8d), sugars are sequentially directly added on the protein under their activated, nucleotide-bound form, by specific glycosyltransferases, without the intervention of a dolichol intermediate. In the case of mucins, the extracellular part of the transmembrane protein carries multiple glycans, as thorns along a stem. Related to these glycoproteins are *proteoglycans*, which also constitute arrangements between a polypeptide chain and osides that are bound to the protein by *O*-osidic bonds. In that case, the glucidic structure consists of a monotonous linear chain of a uronic acid and an *N*-acetylglucosamine (Fig. C.8e), this type of glycan being called *glycosamino-glucuronoglycan* (GAG). These structures, whose molecular weight can be very high, may bear sulphate groups that confer negative charges (heparan sulphate, chondroitin sulphate, etc.). Whereas glycoproteins are informative molecules, due to the precise sequence of their constitutive oses and osides, proteoglycans mainly play a structural role but may sometimes facilitate the transduction of extracellular messages and act as coreceptors: this is the case of β-glycan in the TGFβ signalling pathway (Chap. 5).

Numerous alterations of protein glycosidic chains are found in tumour cells. Most of them are more likely the consequence of malignant transformation, but some of them may play a crucial role in oncogenesis. The inconstant passage of tumour glycoproteins in the circulation may be used to follow tumour evolution (CA125 [gene *MUC16*] for ovarian cancers, CEA [*carcinoembryonic antigen*] for colon cancers, etc.). Structural alterations of glycoproteins may also serve as therapeutic targets for cancer treatment. A frequent phenomenon seen in cancer cells is the increased complexity of *N*-glycoprotein glycans, due to the hyperactivity of glycosyltransferases such as β1,6-*N*-acetylglucosaminyl transferases and sialyltransferases and to the resulting enrichment in sialic acid and fucose, located at the extremity of glycan chains. The abundance in sialic acids could be a factor favouring metastatic dissemination, through a reduction of cell–cell interactions involving terminal sialic acids. In addition, cancer cells of epithelial origin often present mucin alterations, characterised by glycan truncation leading to new antigenic epitopes. However, there are no obvious constant alterations allowing to separate normal and tumour tissues.

These peculiarities of tumour cells, together with the antigenic character of the glycans exposed at the cell surface, have elicited important research on vaccinal therapies aimed at sensitising the immune system against the altered glycoproteins of tumour cells, especially mucin 1 (MUC1) in prostate cancers. Some antibodies directed against glycan epitopes of malignant cells have also been developed. Finally, it is also possible to target the terminal sialylation of glycans with oligosaccharides used as decoys for sialyltransferases, or inhibitors of sialyltransferases or *N*-acetylglucosaminyltransferases.

C.3.5 Lipidation (Fig. C.9)

Several types of lipid moieties can be covalently bound to proteins. These substituents enable the protein to anchor to the membranes, especially the plasma membrane, thanks to the hydrophobic interactions exerted between the aliphatic chains of membrane phospholipids and those covalently bound to proteins. This anchoring

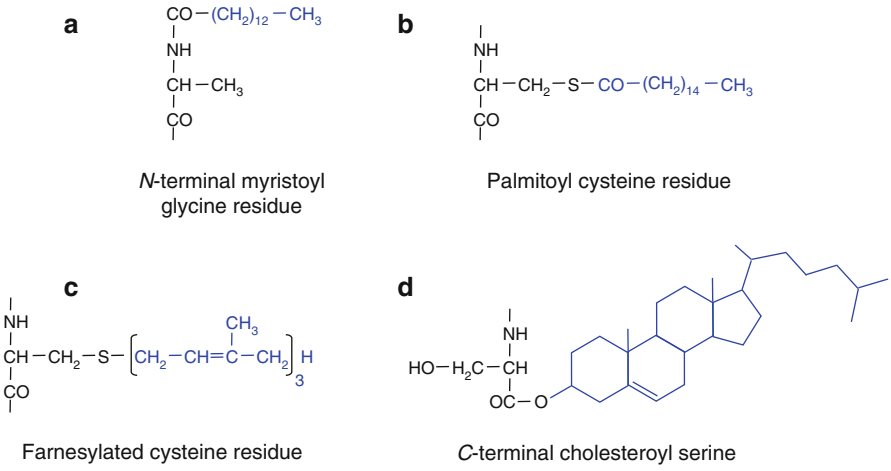


Fig. C.9 Post-translational covalent protein modifications: lipid adducts. (**a**) Myristic acid is bound to the *N*-terminal glycine residue of some proteins, such as the α subunits of large heterotrimeric G-proteins (Chap. 6) through an amide bond. (**b**) Palmitic acid is bound to a cysteine residue of various proteins such as RAS (Chap. 2) or Hedgehog (Chap. 9) through a thioester bond. (**c**) A farnesyl or geranylgeranyl moiety is bound to the cysteine residues of proteins such as RAS by a thioether bond (Chap. 2). (**d**) A molecule of cholesterol is bound to a *C*-terminal serine residue of the Hedgehog proteins (Chap. 9) through an ester bond

may be required for protein activity and may be used to recruit its interactants to the membrane. A classical example in oncology is that of RAS proteins, which are covalently bound to two types of lipid chains (Chap. 2); another is that of the Hedgehog proteins, also covalently bound to two lipid moieties (Chap. 9). Semaphorin 7 is an example of a protein bound to the membrane through a glycosylphosphatidylinositol anchor (Chap. 11).

A frequent substituent is a fatty acyl chain, especially myristoyl (C14:0) and palmitoyl (C16:0) chains, which are bound, respectively, to *N*-terminal glycine residues through amide bonds and to internal cysteine residues through thioester bonds. Other frequent substituents are of polyisoprene structure: farnesyl groups (3 isoprene units, 15 carbon atoms) and geranylgeranyl groups (4 isoprene units, 20 carbon atoms). These groups are provided as pyrophosphates and are bound to cysteine residues through thioester bonds by various farnesyl or geranylgeranyl transferases. A third type of lipid covalent protein adduct is cholesterol, which can be bound to the *C*-terminal amino acid of proteins through an ester bond. Finally, some proteins are attached to the membranes through binding to a glycosylphosphatidylinositol (GPI) moiety; this reaction takes place in ER and replaces the *C*-terminal end of a protein by a glycosylphosphatidylinositol anchor.

A therapeutic approach of prenylation of small G-proteins has been developed, aiming at inhibiting RAS protein farnesylation, which is required for its membrane insertion and its activation of the MAP kinases cascade (Chap. 2). Up to now, this approach has not allowed the discovery of active anticancer drugs able to

specifically inhibit RAS function, but it could well be useful in oncology for the targeting of other small G-proteins of the RHO family.

C.3.6 Poly(ADP-Ribos)ylation

The addition of poly(ADP-ribose) groups, synthesised from NAD coenzymes, on glutamic acid residues of proteins has been described in Annex A. Beyond the role of protein poly(ADP-ribosyl)ation in DNA repair, these covalent modifications of proteins are found in numerous pathways and represent an important modality of regulating protein activity.

C.3.7 Nitrosylation

The addition of the $NO^{\bullet}$ radical on the sulphur atom of cysteine residues (Chap. 16) is one of the mechanisms used by nitric oxide to control protein activity, either activating them (case of RAS proteins) or inhibiting them (case of caspase 3).

C.4 Subcellular Protein Localisation

Proteins are synthesised in the cytoplasm, but they must afterwards migrate to various cell compartments to exert their actions or be secreted. Protein relocalisation or routing represents a crucial step for their function. This relocalisation is enabled by *addressing signals* represented by common amino acid sequences of the protein primary structure. There are addressing signals allowing protein import into mitochondria; long of 20–50 amino acids in the *N*-terminal part, they allow recognition and binding of the protein, guided by a chaperone protein (HSP70), by multiprotein complexes called mitochondrial translocases. One of them, TOM (*translocase of the outer membrane*), enables the protein to reach the intermembrane fluid, while two other proteins, TIMs (*translocases of the inner membrane*), enable it to cross the inner membrane and reach the mitochondrial matrix. Proteins are then freed of the signal sequence by a peptidase and they acquire their definitive conformation thanks to mitochondrial chaperone proteins (Fig. C.10).

Nuclear proteins carry a *nuclear localisation signal* (NLS), rich in basic amino acids. The activation of transcription factors, which can be associated in the cytosol to various signalling proteins, is due to the possibility, in precise conditions such as phosphorylation, to migrate to the nucleus to enable the transcription of their target genes. This is the case of STATs (Chap. 4), SMADs (Chap. 5), GLIs (Chap. 9), of β-catenin (Chap. 7), of the truncated NOTCH receptor (Chap. 8) and of NFκB (Chap. 12). Proteins that should migrate to the nucleus are recognised in the cytoplasm by transporter proteins (*importins*) working as a shuttle between cytoplasm and nucleus (Fig. C.10), thanks to the recognition of nucleoporins, the proteins that constitute the nuclear pore complexes (NPC). Once in the nucleus, the complex protein transported–importin is dissociated thanks to a small G-protein, RAN (*RAS-related nuclear protein*), which binds to importin when associated to GTP; after release of the transported protein in the nucleus, the RAN–importin complex returns to the cytoplasm. RAN then hydrolyses GTP and leaves the transporter protein, which can thus carry on a new transport cycle of an NLS-bearing protein.

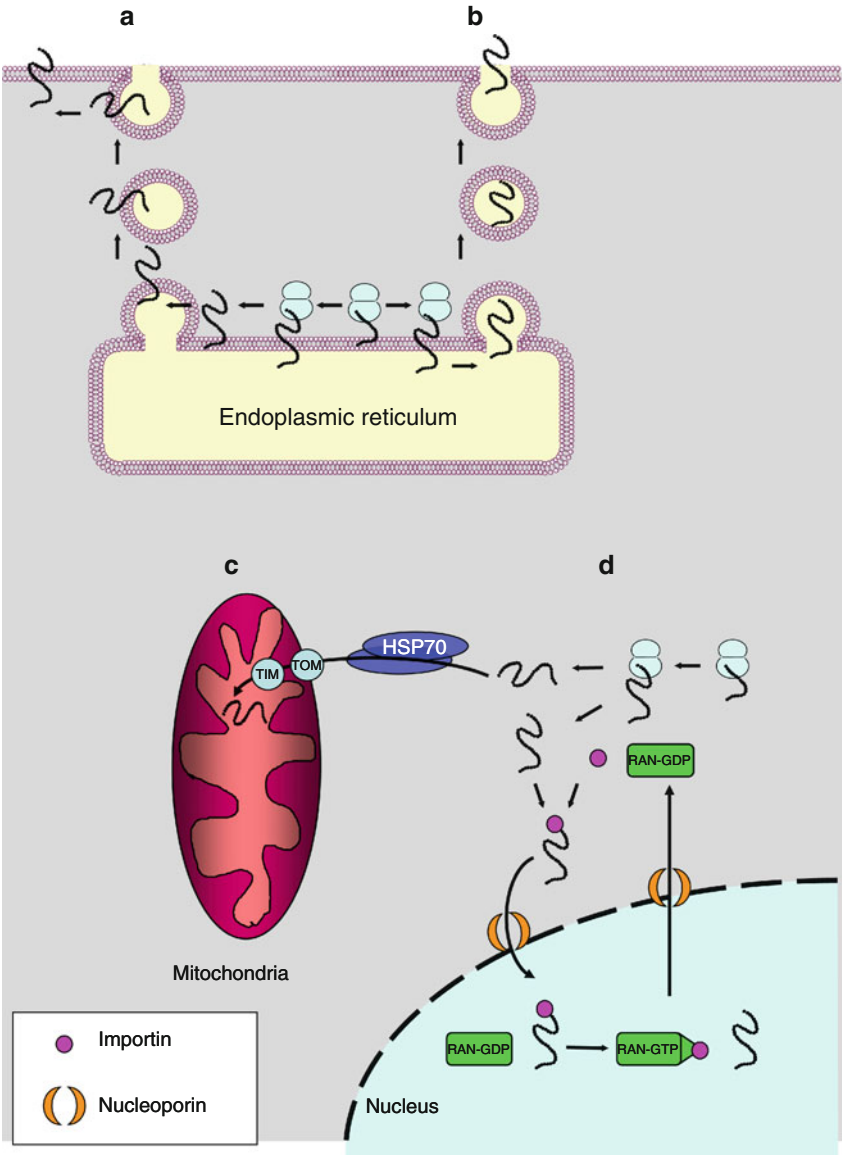


Fig. C.10 Protein translocation. Protein synthesis must be followed by the translocation of the proteins to the cell compartments where they will ensure their functions, thanks to localisation structural domains. (**a**) Protein synthesis at the level of endoplasmic reticulum-associated ribosomes enables their insertion in the plasma membrane after fusion of ER membrane and plasma membrane. (**b**) Such ER-associated synthesis enables also protein secretion to the outside of the cell, when they have no transmembrane domain. (**c**) Protein synthesis in the cytosol enables protein export to the mitochondria, which are reached thanks to specialised transporters localised in the outer membrane (*translocase of the outer membrane*, TOM) or in the inner membrane (*translocases of the inner membrane*, TIM). (**d**) Cytosolic protein synthesis also enables protein export to the nucleus. After coupling to an importin, they can cross the nuclear membrane, thanks to nucleoporins. A small G-protein, RAN, is able to associate importin when bound to GTP; it can thus cross the nuclear membrane, hydrolyse GTP, release importin and return back to the nucleus

Some proteins are directly synthesised in the endoplasmic reticulum, which can bind ribosomal chains (polysomes). They can either entirely get into the ER if they are soluble or remain embedded in the ER membranes if not. The ER serves as an entry point of these proteins in other membrane systems of the cell: Golgi, endosomes and plasma membrane. Polysome binding to the ER is allowed by the recognition of a signal sequence, from the beginning of protein synthesis, by a ribonucleoprotein complex called s*ignal recognition particle* (SRP), which enables the elongating chain to localise straightaway in the ER. This signal sequence is cleaved in the case of soluble proteins, which therefore are entirely located in the ER lumen, where they are taken over by a chaperone protein that ensures their functional folding. Transmembrane proteins preserve one or two α helix sequences that cross the ER membrane and reach their definitive location in this configuration.

Protein localisation in the right cell compartment (the one where they must ensure their function) constitutes an important mode of regulation of their activity. For instance, β-catenin, if non-phosphorylated by the kinase complex, in particular because a mutation of APC (Chap. 7), migrates to the nucleus where it activates the transcription of proliferation genes. Another example is that of p53, of which certain mutations hinder nuclear localisation (Chap. 17). This is the same for $p21^{CIP1}$ or $p27^{KIP1}$ (Chap. 17) whose defective localisation prevents the inhibitory activity of the cyclin–CDK complexes. Conversely, an excessive nuclear import of NFκB (Chap. 12) is an oncogenic event. Therapeutic peptides aimed at targeting the nuclear import of this transcription factor are in development.

C.5 Endoplasmic Reticulum Stress and Unfolded Protein Response

For proteins synthesised in the ER, some environmental features (acidity, hypoxia, glucose deficiency) lead to protein folding abnormalities and to what has been called *unfolded protein response* (UPR). This consists in the generation of adaptation and alarm signals aiming at restoring cell homoeostasis and eliminating misfolded proteins. UPR activates various signalling pathways: the JNK (*JUN N-terminal kinase*) and p38 pathways, for instance, (Chap. 2), and the pathways leading to NFκB (Chap. 12). One of the consequences of UPR is autophagy, a process which enables cells to survive stress but can induce cell death.

UPR starts by the activation of transmembrane proteins whose *C*-terminal extremity is localised in the cytosol and the *N*-terminal extremity in the ER, where they are bound to a special chaperone protein, GRP78 (*glucose-regulated protein*, *78 kDa*, gene *HSPA5* [*heat-shock 70 kDa protein 5*]) (Fig. C.11). Accumulation of misfolded proteins induces the oligomerisation of such proteins, which determines their activation. Among them have been identified IRE1 (*inositol-requiring kinase 1*) (gene *ERN1*, *endoplasmic reticulum to nucleus signalling 1*), PERK (*PRKR-like ER kinase*) (gene *EIF2AK3*, *eukaryotic translation initiation factor 2-alpha kinase 3*) and ATF6 (*activating transcription factor 6*). This activation enables the recruitment of chaperone proteins which take over the misfolded proteins, the arrest of mRNA translation

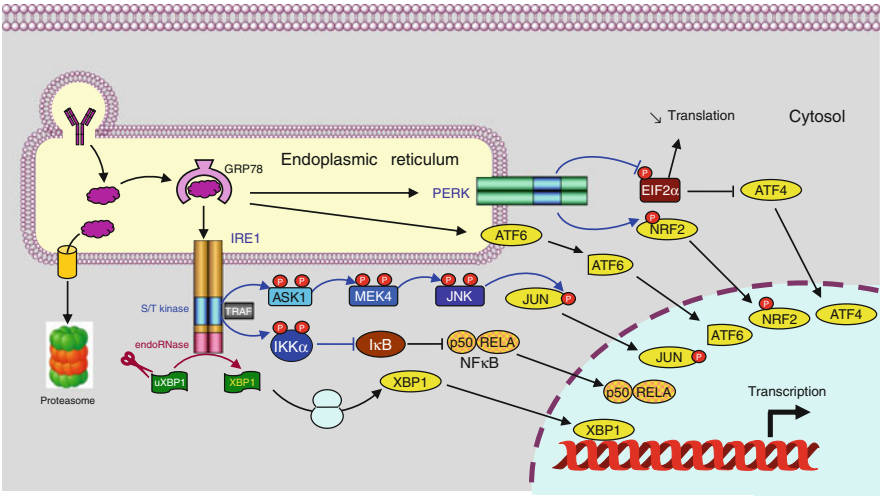


Fig. C.11 Endoplasmic reticulum stress (ERS). The accumulation of misfolded proteins in cancer cells, in particular immunoglobulins in multiple myeloma, may lead them directly to the proteasome; these proteins can also be taken over by GRP78, which activates several signalling pathways initiated by proteins that have their *N*-terminal extremity in ER and their *C*-terminal extremity in the cytosol, like IRE1, PERK and ATF6. IRE1 displays endonucleasic activity and catalyses mRNA splicing of a transcription factor, XBP1, which is thus translated into its functional form and activates transcription of a series of genes. IRE1 also displays serine/threonine kinase activity, which leads, via an adapter protein of the TRAF protein family, to the activation of the MAP3K ASK1 (which activates the JNK and the p38 pathways, Chap. 2) and of the IKK complex (which activates the NFκB pathway, Chap. 12). PERK also displays serine/threonine kinase activity and activates the transcription factor NRF2 as well as, through the elongation factor EIF2α, the transcription factor ATF4. Finally, ATF6 is a transcription factor that is activated by proteolysis in the Golgi. The ERS-activated transcription factors induce the transcription of genes involved in apoptosis, autophagy, protein chaperoning, redox homoeostasis, etc., as well as of the genes encoding GRP78 and XBP1 themselves

and the retrotransport of these proteins to the cytosol where there are ubiquitinylated, then hydrolysed in the proteasome (Sect. C.7) or digested by autophagy.

IRE1 carries both serine/threonine kinase and endoRNase activities, the last one admitting as a substrate the mRNA of XBP1 (*X-box-binding protein 1*) and ensuring its maturation for translation. XBP1 is a transcription factor which, after heterodimerisation with a trimeric transcription factor, NFY, activates the transcription of several genes required for UPR. The kinase activity of IRE1 enables the activation of a special MAP3K, ASK1 (*apoptosis signal-regulating kinase 1*, gene *MAP3K5*, Chap. 2), which initiates the MAP kinase cascade of JNK and p38, initiating the apoptotic process. The proapoptotic proteins BAK and BAX (Chap. 18) directly activate IRE1, demonstrating another connection between apoptosis and UPR.

PERK is also a serine/threonine kinase whose activity is revealed by oligomerisation, enabling autophosphorylation. It phosphorylates and inactivates the translation elongation factor EIF2α, slowing down the translation process and reducing

protein synthesis. However, this phosphorylation of EIF2α has a reverse effect on the mRNA of a transcription factor, ATF4 (*activating transcription factor 4*) whose target genes encode the chaperone proteins GRP78 and GRP94 (gene *HSP90B1*), as well as those involved in protection against oxidative stress, in glutathione synthesis and in autophagy induction, such as the one encoding CHOP (*C/EBP homologous protein*, gene *DDIT3* [*DNA damage-inducible transcript 3*]). This protein is a transcription factor phosphorylated by the MAP kinases JNK and p38 (Chap. 2), involved in apoptosis triggering via the activation of the BH3-only protein BIM and the repression of the antiapoptotic protein BCL2 (Chap. 18).

ATF6 is also a transcription factor which is activated by release from GRP78 and migration to the Golgi, where it is activated by proteolysis. It migrates afterwards to the nucleus to activate transcription of various genes, among which XBP1, whose mRNA is specifically matured by IRE1. Among the target genes of ATF6 are those encoding GRP78, PDI (*protein disulphide isomerase*), CHOP and EDEM1 (*ER degradation-enhancing α-mannosidase-like protein 1*), which increases the ER protection activity toward misfolded proteins or their degradation.

Endoplasmic reticulum stress (ERS) is associated to several diseases such as neurodegenerative diseases, diabetes and cancer. Acidosis, hypoxia and hypoglycaemia, which are often present in the tumour environment, induce such a stress which triggers the processes of UPR protection. It has been observed that GRP78 and XBP1, for instance, are overexpressed in the hypoxic areas of tumours. Pharmacological interventions aimed at preventing tumours from adaptation to hostile environments have been conceived. Inhibition of the serine/threonine kinase activity of IRE1 or PERK is the first approach; inhibition of the recruitment of GRP78 in response to stress is another, and several compounds have been identified. Interleukin 24 (IL24), which binds to and antagonises GRP78, presents antitumour properties. Indirectly, protease or poly(ADP-ribos)ylation inhibitors (Annex A) are able to induce ERS and are developed as potential anticancer drugs.

C.6 Proteolytic Cleavage

There are various kinds of proteolytic cleavages. Some are catabolic and aim at protein destruction, either by effector caspases in the apoptotic processes (Chap. 18) or in specialised organelles such as proteasomes (Sect. C.8) or lysosomes, or else in the extracellular space for the destruction of the extracellular matrix (*matrix metalloproteinases*, MMP). Other types of protein cleavage are in contrast required for protein activation and reveal protein function. They are ensured by a variety of proteases; we give here some examples drawn from oncology.

The initiator caspases (essentially CASP8 and CASP9) are activated by autoproteolysis and elimination of their prodomain (Chap. 18), at the level of dedicated platforms, DISC (*death-inducing signalling complex*) and apoptosome, which enable the contact between the catalytic domain of a caspase and the cleavage site of another. They activate in turn by proteolysis the effector caspases such as CASP3, which are thus cleared of their prodomain. Other caspases, not involved in apoptotic

processes, are able to activate signalling proteins such as interleukin 1 (IL1, Chap. 12), so that CASP1 is also known as *interleukin 1 converting enzyme* (ICE).

Matrix metalloproteinases (MMPs) are also activated by proteolysis: there exists a cascade of proteolytic activities, starting from cathepsin D (CTSD), which cleaves the pro-activator of plasminogen to form the plasminogen activator (PLAU), which cleaves in turn plasminogen (PLG) to form plasmin, which cleaves in turn pro-MMPs to generate collagens, gelatinases and stromelysins. These last proteins are the MMPs responsible for the digestion of the extracellular matrix. They are involved in the metastatic dissemination of tumours, and pharmacological research aims at identifying MMP inhibitors. Some inhibitors were quite efficient in preclinical models but did not reveal anticancer properties during various clinical trials.

ADAM (*a disintegrin and metalloprotease*) proteins, also known as α-secretases, are a group of 21 proteinases with structural homology with the snake venom disintegrins. These are transmembrane proteins and some of them carry, in addition to the metalloproteinase domain, a thrombospondin domain (ADAMTS). They are involved in a large variety of cell processes, inflammation, angiogenesis, cell migration and proliferation and exert an effect on many signalling pathways. They display 'shedding' activity, i.e. they can cleave transmembrane proteins and solubilise the complete ectodomain of their substrates, exerting therefore control on the extracellular side of signalling processes. ADAM17 or TACE (*TNF-alpha converting enzyme*) is especially involved in the activation of EGF and other growth factors by releasing them from their transmembrane precursors (Chap. 1) and in the activation of the Notch pathway (Chap. 8). Overexpression of ADAM proteins has been found in various cancer types and active pharmacological research aims at their inhibition, especially through monoclonal antibodies.

Proprotein convertases constitute a family of nine proteinases, also known as subtilisins, PCSKs (*proprotein convertases subtilisin/kexin*) or PACE (*paired basic amino acid cleaving enzymes*). Some of them, such as furin, have a transmembrane domain. They cleave their substrate proteins at two KR pairs of amino acid residues separated by 0–6 residues. They are involved in the modulation of various signalling pathways; they activate the latent forms of TGFβ (Chap. 5), the NOTCH receptors (Chap. 8), as well as numerous hormone precursors: parathormone, nerve growth factor (NGF), insulin, adrenocorticotrophin (ACTH), renin, enkephalin, dynorphin, somatostatin, gastrin, etc. Some of them are responsible for the proteolytic activation of the gp160 and gp140 of the HIV envelope. They are activated by autocatalytic proteolysis. Oncogenic alterations of these proteins have been described: *PCSK1* gene mutations in carcinoid tumours, *PCSK7* gene rearrangements in malignant lymphomas, etc. Proteinase inhibitors are currently used in AIDS treatment; their possible use in cancer is under consideration.

Gamma secretase is a protein complex to which participate various polypeptides, especially presenilins, nicastrin, cathepsin B and other partners. γ-secretase is involved in the activation of the NOTCH receptors (Chap. 8) and inhibitors are being developed for cancer treatment. These enzymes are responsible for the cleavage of the amyloid protein precursor (APP) and contribute thus to the pathogenesis of Alzheimer disease. γ-secretase components may be involved in oncogenesis: for instance, cathepsin B is overexpressed in oesophagian adenocarcinomas and other tumours.

C.7 Ubiquitinylation and Proteasome

C.7.1 Mechanism of Action

Protein destruction by proteolysis represents an essential mean of regulation of protein life. Systems such as apoptosis, autophagy and lysosome digestion are not protein specific; proteasome also ensures the degradation of most proteins but only after they have been specifically tagged by ubiquitinylation. The ubiquitin–proteasome system is an important component of gene expression control and plays a role in signal transduction, cell cycle progression, DNA repair and apoptosis. In addition, it enables the destruction of misfolded, damaged or mutant proteins, whose accumulation could be deleterious for the cell. Alterations in the ubiquitin–proteasome pathway lead to pathological processes and may constitute a target for cancer therapy.

Ubiquitin is a 76-amino acid protein which is conjugated with the protein to be hydrolysed after a series of three reactions: a first E1 enzyme, *ubiquitin-activating enzyme*, activates ubiquitin through binding the *C*-terminal glycine residue of ubiquitin to a cysteine residue of the enzyme active centre by a thioester bond (Fig. C.12). Ubiquitin is then transferred on a reactive cysteine residue of a second E2 enzyme, *ubiquitin-conjugating enzyme*, by transthiolation. Finally, E3 enzymes, *ubiquitin protein ligases*, transfer ubiquitin on the protein to be destroyed by generating an amide bond between the *C*-terminal ubiquitin glycine residue and the amine group of an internal lysine residue of the substrate protein. There exist two E1 enzymes, two dozens of E2 enzymes and several hundreds of E3 enzymes, which bear the specificity for the protein to be ubiquitinylated. In fact, several ubiquitin moieties are attached together from the beginning by internal lysine residues, generally Lys^{48} (Fig. C.12), so that ubiquitin chains are attached to the target protein substrate. Chain length must be long enough (at least 4 ubiquitin molecules) for the recognition of the complex by the proteasome.

Ubiquitin ligases (E3 enzymes) belong to two major families, those equipped with a HECT (*homologous to E6-associated protein C terminus*) domain and those carrying a RING (*really interesting new gene*) domain, whose mechanisms of action are different. For RING domain-carrying E3 enzymes, the transfer of ubiquitin on the protein substrate is obtained without previous binding of ubiquitin with E3. Ubiquitin ligases may be monomeric or multimeric, with ubiquitin-recognition sites carried by different subunits. The mode of recognition of substrate proteins by E3 enzymes is complex and varies from a protein to another. Phosphorylation is often used for directing ubiquitinylation: this is, for instance, the case for $p27^{KIP1}$ (gene *CDKN1B*) which controls cell cycle progression (Chap. 17) and that of β-catenin (Chap. 7). Other protein substrates with short half-life carry PEST (*proline–glutamic acid–serine–threonine-rich*) domains, which enable their recognition by E3 enzymes.

Proteasomes are corpuscles with similar size as ribosomes (26S); they consist of a multimeric protein complex (PSM proteins), with two main subunits whose assembly requires ATP: the 20S catalytic particle and the 19S regulatory particle (Fig. C.13). The 20S particle sequesters the protease sites inside a cylindrical structure consisting of four heptameric rings: two outer ones made of seven α subunits and two inner ones made of seven β subunits. The β subunits carry the threonine protease activity, which

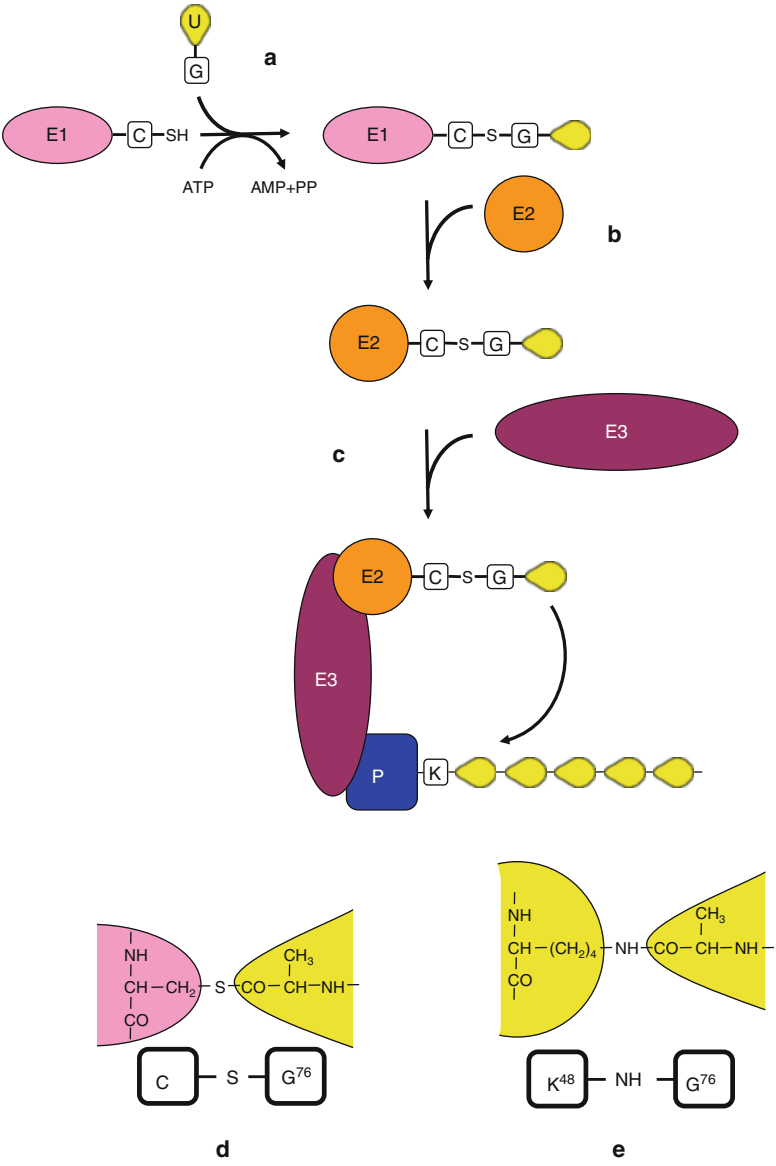


Fig. C.12 Protein ubiquitinylation. (**a**) A first enzyme (*ubiquitin-activating enzyme*, E1) activates ubiquitin through binding the *C*-terminal glycine residue of ubiquitin to a cysteine residue of the enzyme catalytic site through a thioester bond. (**b**) Ubiquitin is then transferred to a cysteine residue of a second enzyme (*ubiquitin-conjugating enzyme*, E2) through transthiolation. (**c**) A third type of enzyme (*ubiquitin protein ligase*, E3) transfers ubiquitin on the substrate protein, thanks to an amide bond between the carboxyl moiety of the *C*-terminal glycine residue of ubiquitin and a lysine internal residue of the substrate protein. (**d**) Thioester bond between the E1 or E2 cysteine residue and the *C*-terminal glycine residue of ubiquitin. (**e**) Peptidic (amide) bond between the Lys^{48} residue of a molecule of ubiquitin and the *C*-terminal glycine residue of the following ubiquitin molecule

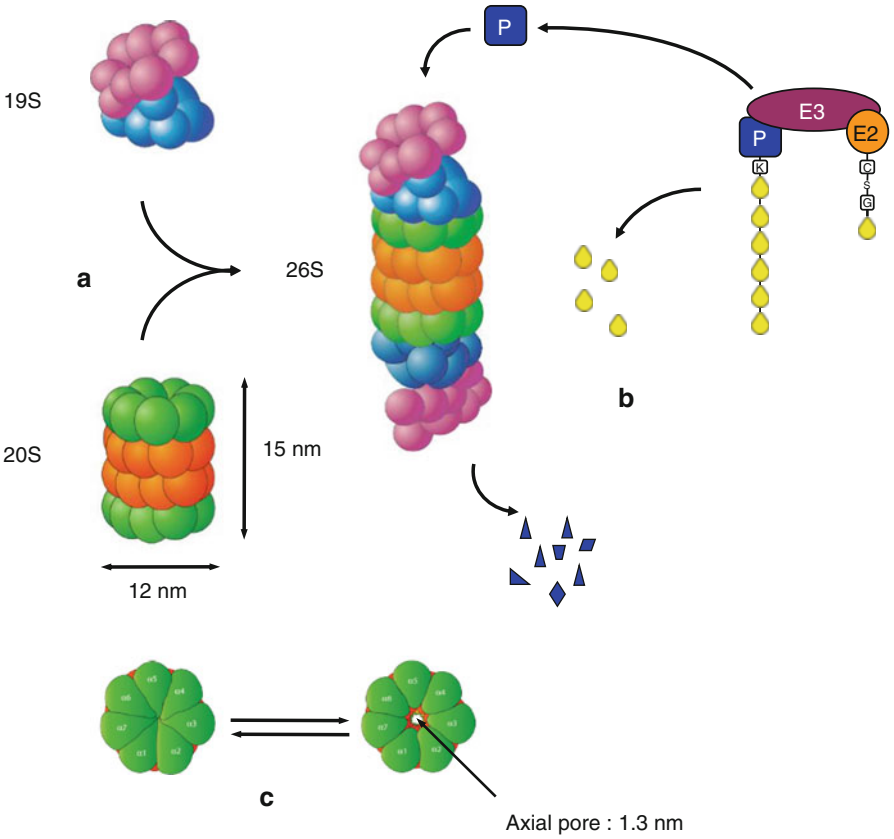


Fig. C.13 Protein destruction in the proteasome. (**a**) The 19S and 20S proteasome subunits are assembled to form the 26S particle. The 19S particle regulates protein entry into the cylinder constituted by the 20S particle. The 20S particle results from the stacking of four heptameric rings, two outer ones (*green*), made of seven α subunits, and two inner ones (*orange*), made of seven β subunits which carry threonine protease activity. (**b**) The ubiquitinylated protein is led by E3 to the proteasome when it is hydrolysed. Ubiquitin molecules are recycled. The protein is fragmented into peptides 6–12 amino acids long. (**c**) The 20S subunit, viewed from above. The outer rings form narrow pores at each extremity of the cylinder, restricting thus the access to the catalytic chamber

can be *chymotrypsin-like* (CTL), *trypsin-like* (TL) or *post-glutamate peptide hydrolase* (PGPH), and enable the degradation of a large variety of proteins into small peptides of 6–12 amino acids. The outer rings comprising the α subunits form thin pores at each extremity of the cylinder, restricting thus the access to the catalytic chamber. This type of architecture isolates the catalytic sites and prevents the degradation of proteins that should not be destroyed. The 19S complex recognises the polyubiquitin chain carrying the protein to be destroyed; hydrolyses the bond between the substrate and the ubiquitin chain, which is recycled in ubiquitin monomers by deubiquitinylases; denatures the substrate; binds to the 20S complex at the level of the outer rings; and opens the catalytic chamber in which the substrate is translocated. This mechanism avoids non-selective protein degradation.

C.7.2 Oncogenic Alterations

A large number of oncogene mutations or tumour suppressor gene inactivations directly affect the ubiquitin–proteasome pathway. Some oncogenes encode ubiquitin ligases. Deleterious mutations occurring in this process prevent the degradation of certain proteins such as transcription factors or angiogenic inducers. Also, gene amplification can lead to the degradation of gene suppressor proteins. Some examples can be mentioned:

- The *VHL* (*von Hippel–Lindau disease*) gene encodes the E3 ligase of a hypoxia-inducible transcription factor, HIF1α (Chap. 16). In response to hypoxia, this factor activates the expression of genes such as *VEGFA* (*vascular endothelial growth factor*). VHL mutations prevent the degradation of HIF1α in normoxic conditions and predispose to the formation of hypervascular lesions and renal tumours.
- The *APC* (*adenomatous polyposis coli*) gene is a tumour suppressor gene which is mutated in 70 % of colorectal cancers. Its protein product binds and regulates β-catenin (Chap. 7). *APC* mutations lead to the formation of truncated proteins unable to bind β-catenin and, therefore, to enable its phosphorylation, then its polyubiquitinylation, enhancing thus its transcriptional stimulatory activity.
- The *MDM2* (*murine double-minute homologue 2*) gene encodes a RING domain-containing E3 ligase which induces p53 ubiquitinylation (Chap. 17). *MDM2* overexpression, resulting especially from gene amplification, is a mechanism responsible for p53 inactivation in several cancer types, in particular soft-tissue sarcomas.
- The *SMURF* (*SMAD ubiquitinylation regulatory factor 2*) gene encodes a HECT domain-containing E3 ligase which regulates SMAD proteins degradation in the TGFβ pathway (Chap. 5). An accelerated degradation of these transcription factors prevents the downregulation of cell proliferation by TGFβ.
- The *CDKN1B* gene encodes the $p27^{KIP1}$ protein, a CDK2 inhibitor with tumour suppressor properties, which regulates the G1 → S transition of cell cycle through interaction with CDK2–cyclin E complexes (Chap. 17). One of the mechanisms responsible for the downregulation or p27 expression in tumours is its ubiquitinylation by an E3 ubiquitin ligase of the SCF (*SKP1/cullin/F-box*) family, whose concentration is especially high in several tumour types.
- Most cyclins are degraded in the proteasome after ubiquitinylation, in order to oblige them to stop their CDK-activating functions after the passage from one phase of the cell cycle to the following one. As for $p27^{KIP1}$, this reaction is catalysed for cyclin E by an E3 ligase of the SCF family, a component of which is the protein CDC4 whose gene (*FBXW7*) is mutated in colorectal and ovarian cancers.
- EGFR (Chap. 1) is the substrate of a RING domain-containing E3 ubiquitin ligase called CBL (*Casitas B-lineage lymphoma*). EGFR ubiquitinylation represents a signal of internalisation in endosomal vesicles, leading to receptor proteolysis after fusion with lysosomes. The EGFR–ERBB2 heterodimers, which are favoured when the *ERBB2* gene is amplified, are less sensitive to CBL action than the EGFR–EGFR homodimers.

C.7.3 Therapeutic Targets

Because of the action of the ubiquitin–proteasome system on multiple proteins involved in cell proliferation and survival processes, pharmacological interventions might contribute to inhibit tumour cell proliferation, induce apoptosis and/or increase the sensitivity of tumour cells to anticancer drugs. The preservation of the activity of damaged proteins that would be driven to the proteasome, such as the CDK inhibitors $p21^{CIP1}$ and $p27^{KIP1}$ or the antiapoptotic agents BAX and p53, might be efficient in cancer therapy. However, these proteins, which are essential for proliferation, are also degraded in the proteasome and their stabilisation could favour tumour growth. The identification of compounds that target the proteasome has enabled the successful development of an inhibitor of chymotrypsin-like activity, bortezomib, which has been marketed for the treatment of multiple myeloma and is presently under evaluation in solid tumours. Other compounds, with structure analogy to bortezomib or not, are also being studied. The initial success has encouraged much research at the level of E3 ligases inhibition, whose activity may well prove more selective than that of bortezomib: for instance, nutlin-3 blocks the p53 binding site of MDM2 and increases p53 stability (Chap. 17).

C.8 Sumoylation

Protein sumoylation consists in binding, on a specific lysine residue of a protein, of a small ubiquitin-related protein of about 100 amino acids, SUMO (*small ubiquitin-like modifier*). This covalent modification is involved in the regulation of protein function and specific interactions. There are four SUMO proteins (SUMO1, 2, 3 and 4), which are activated by proteolytic cleavage of their *C*-terminal extremity to generate a diglycine motif, through a proteinase called SENP (*sentrin-specific protease*). The activated SUMO protein is transferred to a cysteine residue of a *SUMO-activating enzyme* (SAE), the equivalent of the E1 enzyme of the ubiquitinylation process. SUMO is then transferred on an equivalent of the E2 enzyme, called UBC9 (*ubiquitin-conjugating enzyme homologue*) and then on a lysine residue of the target protein, at the level of a consensus sequence, ΨKXD/E (Ψ designating a hydrophobic amino acid). The E3-equivalent enzymes that have been identified are the proteins PIAS (*protein inhibitor of activated STAT*, Chap. 4), the polycomb repressing complex 2 (Annex B) and a RAN-binding protein, RANBP2 (Sect. C.5). Sumoylation is generally a transient process that is reversed by desumoylating proteases. SENP can also detach the SUMO adduct from the target protein and allow its recycling.

The best known sumoylated proteins is the GTPase-activating protein of the nuclear trafficking protein RAN, called RANGAP1. Other sumoylation substrates are expressed in the nervous system and are involved in neurodegenerative diseases: huntingtin (HTT, Huntington disease), ataxin (ATXN, spinocerebellar ataxia), synuclein α (SNCA, Parkinson disease), superoxide dismutase 1 (SOD, amyotrophic sclerosis) and APP (*amyloid beta precursor protein*, Alzheimer disease). In addition, several transcription factors are sumoylated, such as MYB, STAT and HIF1α, as well as nuclear receptors such as the androgen receptor (AR) or the oestrogen receptor (ER, gene *ESR1*) (Chap. 14).

Alterations in protein sumoylation can be associated with oncogenesis. UBC9 overexpression is found in several cancer types and this protein is able experimentally in enhance cell proliferation. Certain E3 proteins such as PIAS3 are also overexpressed in cancer. Sumoylation regulates the activity of several oncogene products and tumour suppressors, such as p53 and its homologues p63 and p73, the retinoblastoma protein RB1 (Chap. 17) and MDM2 (Chap. 17), etc. Finally, the proteinase SENP is overexpressed in prostate and thyroid cancers. The relationships between sumoylation and cancer are still poorly understood but could lead to therapeutic targeting.

Further Reading

Genome Organisation and Replication Machinery

Bell SP, Dutta A. DNA replication in eukaryotic cells. Annu Rev Biochem. 2002;71:333–74.
Davey MJ, O'Donnell M. Mechanisms of DNA replication. Curr Opin Chem Biol. 2000;4:581–6.
Hoffmann JS, Cazaux C. Aberrant expression of alternative DNA polymerases: a source of mutator phenotype as well as replicative stress in cancer. Semin Cancer Biol. 2010;20:312–9.
Johnson A, O'Donnell M. Cellular DNA replicases: components and dynamics at the replication fork. Annu Rev Biochem. 2005;74:283–315.
Lange SS, Takata K, Wood RD. DNA polymerases and cancer. Nat Rev Cancer. 2011;11:96–110.

DNA Repair

Bolderson E, Richard DJ, Zhou BB, Khanna KK. Recent advances in cancer therapy targeting proteins involved in DNA double-strand break repair. Clin Cancer Res. 2009;15:6314–20.
Bouwman P, Jonkers J. The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance. Nat Rev Cancer. 2012;12:587–98.
Branzei D, Foiani M. Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol. 2008;9:297–308.
Bunting SF, Nussenzweig A. End-joining, translocations and cancer. Nat Rev Cancer. 2013;13:443–54.
Cassidy LD, Venkitaraman AR. Genome instability mechanisms and the structure of cancer genomes. Curr Opin Genet Dev. 2012;22:10–3.
Cleaver JE, Lam ET, Revet I. Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. Nat Rev Genet. 2009;10:756–68.
Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. Nat Rev Cancer. 2012;12:801–17.
Fu D, Calvo JA, Samson LD. Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat Rev Cancer. 2012;12:104–20.
Furgason JM, Bahassi EM. Targeting DNA repair mechanisms in cancer. Pharmacol Ther. 2013;137:298–308.
Helleday T, Petermann E, Lundin C, Hodgson B, Sharma RA. DNA repair pathways as targets for cancer therapy. Nat Rev Cancer. 2008;8:193–204.
Hoeijmakers JH. DNA damage, aging, and cancer. N Engl J Med. 2009;361:1475–85.
Lange SS, Takata K, Wood RD. DNA polymerases and cancer. Nat Rev Cancer. 2011;11:96–110.
Lord CJ, Ashworth A. The DNA damage response and cancer therapy. Nature. 2012;481:287–94.
Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability–an evolving hallmark of cancer. Nat Rev Mol Cell Biol. 2010;11:220–8.

O'Connor MJ, Martin NM, Smith GC. Targeted cancer therapies based on the inhibition of DNA strand break repair. Oncogene. 2007;26:7816–24.
Prakash S, Johnson RE, Prakash L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu Rev Biochem. 2005;74:317–53.
Schärer OD. Chemistry and biology of DNA repair. Angew Chem Int Ed. 2003;42:2946–74.
Shrivastav N, Li D, Essigmann JM. Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. Carcinogenesis. 2010;31:59–70.

Poly(ADP-ribosyl)ation

Fong PC, Boss DS, Yap TA, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med. 2009;361:123–34.
Jagtap P, Szabo C. Poly (ADP-ribose) polymerase and the therapeutic effects of its inhibitors. Nat Rev Drug Discov. 2005;4:421–40.
Krishnakumar R, Kraus WL. The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. Mol Cell. 2010;39:8–24.
Rouleau M, Patel A, Hendzel MJ, Kaufmann SH, Poirier GG. PARP inhibition: PARP1 and beyond. Nat Rev Cancer. 2010;10:293–301.

Telomeres and Telomerase

Artandi SE, DePinho RA. Telomeres and telomerase in cancer. Carcinogenesis. 2010;31:9–18.
Blasco MA. Telomeres and human disease: ageing, cancer and beyond. Nat Rev Genet. 2005;6:611–22.
Blasco MA. Telomere length, stem cells and aging. Nat Chem Biol. 2007;3:640–9.
Bochman ML, Paeschke K, Zakian VA. DNA secondary structures: stability and function of G-quadruplex structures. Nat Rev Genet. 2012;13:770–80.
Deng Y, Chan SS, Chang S. Telomere dysfunction and tumour suppression: the senescence connection. Nat Rev Cancer. 2008;8:450–8.
Folini M, Gandellini P, Zaffaroni N. Targeting the telosome: therapeutic implications. Biochim Biophys Acta. 2009;1792:309–16.
Günes C, Rudolph KL. The role of telomeres in stem cells and cancer. Cell. 2013;152:390–3.
Harley CB. Telomerase and cancer therapeutics. Nat Rev Cancer. 2008;8:167–79.
Murnane JP. Telomere loss as a mechanism for chromosome instability in human cancer. Cancer Res. 2010;70:4255–9.
O'Sullivan RJ, Karlseder J. Telomeres: protecting chromosomes against genome instability. Nat Rev Mol Cell Biol. 2010;11:171–81.
Riethman H. Human telomere structure and biology. Annu Rev Genomics Hum Genet. 2008;9:1–19.
Shay JW, Wright WE. Telomerase therapeutics for cancer: challenges and new directions. Nat Rev Drug Discov. 2006;5:577–84.
Shay JW, Wright WE. Role of telomeres and telomerase in cancer. Semin Cancer Biol. 2011;21:349–53.
Stewart SA, Bertuch AA. The role of telomeres and telomerase in cancer research. Cancer Res. 2010;70:7365–71.

Transcription Machinery

Compe E, Egly JM. TFIIH: when transcription met DNA repair. Nat Rev Mol Cell Biol. 2013;2012(13):343–54.
Uptain SM, Kane CM, Chamberlin MJ. Basic mechanisms of transcript elongation and its regulation. Annu Rev Biochem. 1997;66:117–72.

Sawadogo M, Sentenac A. RNA polymerase B (II) and general transcription factors. Annu Rev Biochem. 1990;59:711–54.

Transcription Control

Bywater MJ, Pearson RB, McArthur GA, Hannan RD. Dysregulation of the basal RNA polymerase transcription apparatus in cancer. Nat Rev Cancer. 2013;13:300–14.
Darnell Jr JE. Transcription factors as targets for cancer therapy. Nat Rev Cancer. 2002;2:740–9.
Farnham PJ. Insights from genomic profiling of transcription factors. Nat Rev Genet. 2009;10:605–16.
Neph S, Stergachis AB, Reynolds A, Sandstrom R, Borenstein E, Stamatoyannopoulos JA. Circuitry and dynamics of human transcription factor regulatory networks. Cell. 2012;150:1274–86.
Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. Nat Rev Genet. 2009;10:252–63.

DNA Methylation

Das PM, Singal R. DNA methylation and cancer. J Clin Oncol. 2004;22:4632–42.
Goodman JI, Watson RE. Altered DNA methylation: a secondary mechanism involved in carcinogenesis. Annu Rev Pharmacol Toxicol. 2002;42:501–25.
Issa JP. DNA methylation as a therapeutic target in cancer. Clin Cancer Res. 2007;13:1634–7.
Robertson KD. DNA methylation and human disease. Nat Rev Genetics. 2005;6:597–610.
Szyf M. Targeting DNA, methylation in cancer. Bull Cancer. 2006;93:961–72.

Histones

Batty N, Malouf GG, Issa JP. Histone deacetylase inhibitors as anti-neoplastic agents. Cancer Lett. 2009;280:192–200.
Bracken AP, Helin K. Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nat Rev Cancer. 2009;9:773–84.
Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet. 2007;8:286–98.
Gal-Yam EN, Saito Y, Egger G, Jones PA. Cancer epigenetics: modifications, screening, and therapy. Annu Rev Med. 2008;59:267–80.
Lennartsson A, Ekwall K. Histone modification patterns and epigenetic codes. Biochim Biophys Acta. 2009;1790:863–8.
Lopez J, Percharde M, Coley HM, Webb A, Crook T. The context and potential of epigenetics in oncology. Br J Cancer. 2009;100:571–7.
Minucci S, Pelicci PG. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nat Rev Cancer. 2006;6:38–51.
Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis. 2010;31:27–36.
Yoo CB, Jones PA. Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov. 2006;5:37–50.

Micro RNAs

Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer. 2006;6:857–66.
Gartel AL, Kandel ES. miRNAs: Little known mediators of oncogenesis. Semin Cancer Biol. 2008;18:103–10.

Garzon R, Calin GA, Croce CM. MicroRNAs in cancer. Annu Rev Med. 2009;60:167–79.
Medina PP, Slack FJ. microRNAs and cancer: an overview. Cell Cycle. 2008;7:2485–92.
Nicoloso MS, Spizzo R, Shimizu M, Rossi S, Calin GA. MicroRNAs – the micro steering wheel of tumour metastases. Nat Rev Cancer. 2009;9:293–302.
Schickel R, Boyerinas B, Park SM, Peter ME. MicroRNAs: key players in the immune system, differentiation, tumorigenesis and cell death. Oncogene. 2008;27:5959–74.

Alternative Splicing

Dehm SM. mRNA splicing variants: exploiting modularity to outwit cancer therapy. Cancer Res. 2013;73:5309–14.
Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol. 2013; 13:153–65.
Matlin AJ, Clark F, Smith CW. Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol. 2005;6:386–98.
Shin C, Manley JL. Cell signalling and the control of pre-mRNA splicing. Nat Rev Mol Cell Biol. 2004;5:727–38.
Skotheim RI, Nees M. Alternative splicing in cancer: noise, functional, or systematic? Int J Biochem Cell Biol. 2007;39:1432–49.
Venables JP. Aberrant and alternative splicing in cancer. Cancer Res. 2004;64:7647–54.

Translation Machinery

Kapp LD, Lorsch JR. The molecular mechanics of eukaryotic translation. Annu Rev Biochem. 2004;73:657–704.
Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11:113–27.

Protein Folding, Protein Chaperones

Banerji U. Heat shock protein 90 as a drug target: some like it hot. Clin Cancer Res. 2009;15:9–14.
Daggett V, Fersht A. The present view of the mechanism of protein folding. Nat Rev Mol Cell Biol. 2003;4:497–502.
Jhaveri K, Taldone T, Modi S, Chiosis G. Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers. Biochim Biophys Acta. 2012;1823:742–55.
Khalil AA, Kabapy NF, Deraz SF, Smith C. Heat shock proteins in oncology: diagnostic biomarkers or therapeutic targets? Biochim Biophys Acta. 2011;1816:89–104.
Li Y, Zhang T, Schwartz SJ, Sun D. New developments in Hsp90 inhibitors as anti-cancer therapeutics: mechanisms, clinical perspective and more potential. Drug Resist Updat. 2009;12:17–27.
Neckers L, Workman P. Hsp90 molecular chaperone inhibitors: are we there yet? Clin Cancer Res. 2012;18:64–76.
Scaltriti M, Dawood S, Cortes J. Molecular pathways: targeting hsp90–who benefits and who does not. Clin Cancer Res. 2012;18:4508–13.
Trepel J, Mollapour M, Giaccone G, Neckers L. Targeting the dynamic HSP90 complex in cancer. Nat Rev Cancer. 2010;10:537–49.

Whitesell L, Lin NU. HSP90 as a platform for the assembly of more effective cancer chemotherapy. Biochim Biophys Acta. 2012;1823:756–66.
Whitesell L, Lindquist SL. HSP90 and the chaperoning of cancer. Nat Rev Cancer. 2005;5:761–72.

Post-translational Covalent Protein Modifications

Brooks SA, Carter TM, Royle L, Harvey DJ, Fry SA, Kinch C, Dwek RA, Rudd PM. Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies? Anticancer Agents Med Chem. 2008;8:2–21.
Dube DH, Bertozzi CR. Glycans in cancer and inflammation–potential for therapeutics and diagnostics. Nat Rev Drug Discov. 2005;4:477–88.
Dwek MV, Brooks SA. Harnessing changes in cellular glycosylation in new cancer treatment strategies. Curr Cancer Drug Targets. 2004;4:425–42.
Gorelik E, Galili U, Raz A. On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis. Cancer Metastasis Rev. 2001;20:245–77.
Logsdon CD, Fuentes MK, Huang EH, Arumugam T. RAGE and RAGE ligands in cancer. Curr Mol Med. 2007;7:777–89.
Singh R, Bandyopadhyay D. MUC1: a target molecule for cancer therapy. Cancer Biol Ther. 2007;6:481–6.

Subcellular Protein Localisation

Davis JR, Kakar M, Lim CS. Controlling protein compartmentalization to overcome disease. Pharm Res. 2007;24:17–27.
Dorn GW, Mochly-Rosen D. Intracellular transport mechanisms of signal transducers. Annu Rev Physiol. 2002;64:407–29.
Gorlich D, Kutay U. Transport between the cell nucleus and the cytoplasm. Annu Rev Cell Dev Biol. 1999;15:607–60.
Kau TR, Way JC, Silver PA. Nuclear transport and cancer: from mechanism to intervention. Nat Rev Cancer. 2004;4:106–17.

Endoplasmic Reticulum Stress and Unfolded Protein Response

Hetz C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nat Rev Mol Cell Biol. 2012;13:89–102.
Hetz C, Chevet E, Harding HP. Targeting the unfolded protein response in disease. Nat Rev Drug Discov. 2013;12:703–19.
Kim I, Xu W, Reed JC. Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities. Nat Rev Drug Discov. 2008;7:1013–30.
Kraskiewicz H, FitzGerald U. Interfering with endoplasmic reticulum stress. Trends Pharmacol Sci. 2012;33:53–63.
Li Z, Li Z. Glucose regulated protein 78: a critical link between tumor microenvironment and cancer hallmarks. Biochim Biophys Acta. 2012;1826:13–22.
Luo B, Lee AS. The critical roles of endoplasmic reticulum chaperones and unfolded protein response in tumorigenesis and anticancer therapies. Oncogene. 2013;32:805–18.
Ma Y, Hendershot LM. The role of the unfolded protein response in tumour development: friend or foe? Nat Rev Cancer. 2004;4:966–77.
Moenner M, Pluquet O, Bouchecareilh M, Chevet E. Integrated endoplasmic reticulum stress responses in cancer. Cancer Res. 2007;67:10631–4.

Ron D, Walter P. Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol. 2007;8:519–29.
Rothman JE, Schekman R. Molecular mechanism of protein folding in the cell. Cell. 2011;146:851–4.
Schleicher SM, Moretti L, Varki V, Lu B. Progress in the unraveling of the endoplasmic reticulum stress/autophagy pathway and cancer: implications for future therapeutic approaches. Drug Resist Updat. 2010;13:79–86.
Schönthal AH. Pharmacological targeting of endoplasmic reticulum stress signaling in cancer. Biochem Pharmacol. 2013;85:653–66.

Proteolytic Cleavage

Białas A, Kafarski P. Proteases as anti-cancer targets–molecular and biological basis for development of inhibitor-like drugs against cancer. Anticancer Agents Med Chem. 2009;9:728–62.
Koblinski JE, Ahram M, Sloane BF. Unraveling the role of proteases in cancer. Clin Chim Acta. 2000;291:113–35.
Mohamed MM, Sloane BF. Cysteine cathepsins: multifunctional enzymes in cancer. Nat Rev Cancer. 2006;6:764–75.
Murphy G. The ADAMs: signalling scissors in the tumour microenvironment. Nat Rev Cancer. 2008;8:929–41.

Ubiquitinylation and Proteasome

Adams J. The proteasome: a suitable antineoplastic target. Nat Rev Cancer. 2004;4:349–60.
Bedford L, Lowe J, Dick LR, Mayer RJ, Brownell JE. Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. Nat Rev Drug Discov. 2011;10:29–46.
Ciechanover A. Proteolysis: from the lysosome to ubiquitin and the proteasome. Nat Rev Mol Cell Biol. 2005;6:79–87.
Fraile JM, Quesada V, Rodriguez D, Freije JMP, Lopez-Otin C. Deubiquitinases in cancer: new functions and therapeutic options. Oncogene. 2012;31:2373–88.
Frankland-Searby S, Bhaumik SR. The 26S proteasome complex: an attractive target for cancer therapy. Biochim Biophys Acta. 2012;1825:64–76.
Hoeller D, Dikic I. Targeting the ubiquitin system in cancer therapy. Nature. 2009;458:438–44.
Kirkin V, Dikic I. Ubiquitin networks in cancer. Curr Opin Genet Develop. 2011;21:21–8.
Lipkowitz S, Weissman AM. RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. Nat Rev Cancer. 2011;11:629–43.
Murata S, Yashiroda H, Tanaka K. Molecular mechanisms of proteasome assembly. Nat Rev Mol Cell Biol. 2009;10:104–15.
Nakayama KI, Nakayama K. Ubiquitin ligases: cell-cycle control and cancer. Nat Rev Cancer. 2006;6:369–81.

Sumoylation

Hickey CM, Wilson NR, Hochstrasser M. Function and regulation of SUMO proteases. Nat Rev Mol Cell Biol. 2012;13:755–66.
Kim KI, Baek SH. SUMOylation code in cancer development and metastasis. Mol Cells. 2006;22:247–53.
Kim KI, Baek SH. Small ubiquitin-like modifiers in cellular malignancy and metastasis. Int Rev Cell Mol Biol. 2009;273:265–311.
Sarge KD, Park-Sarge OK. Sumoylation and human disease pathogenesis. Trends Biochem Sci. 2009;34:200–5.