*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

Gerhard Krauss

# Biochemistry of Signal Transduction and Regulation

## Second Edition

WILEY-VCH

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

Gerhard Krauss

# Biochemistry of Signal Transduction and Regulation

## Second Edition

Translated by Nancy Schönbrunner
and Julia Cooper

WILEY-VCH

Weinheim · New York · Chichester · Brisbane · Singapore · Toronto

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

Prof. Dr. Gerhard Krauss
Laboratorium für Biochemie
Universität Bayreuth
D-95440 Bayreuth
Gemany
e-mail: Gerhard.Krauss@uni-bayreuth.de



1st English edition 1999
2nd English edition 2001

Die Deutsche Bibliothek – CIP-Cataloguing-in-Publication-data
A catalogue record for this publication is available from
Die Deutsche Bibliothek



Printed on acid-free paper



Composition: Mitterweger & Partner Kommunikationsgesellschaft mbH, D-68723 Plankstadt
Printing: betz-druck GmbH, D-64291 Darmstadt
Bookbinding: Wilh. Osswald & Co., D-67433 Neustadt/Weinstr.

Printed in the Federal Republic of Germany

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

*For Silvia, Julia, Hannes, and Enno*

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Related Titles from Wiley-VCH

U. Diederichsen / T. K. Lindhorst / B. Westermann / L. A. Wessjohann (eds.)
**Bioorganic Chemistry – Highlights and New Aspects**
1999. XIV, 416 pages with 155 figures and 28 tables. Softcover.
ISBN 3-527-29665-4

H. Waldmann / J. Mulzer (eds.)
**Organic Synthesis Highlights III**
1998. XV, 444 pages with 309 figures and 10 tables. Softcover.
ISBN 3-527-29500-3

G. C. Ferreira / J. J. G. Moura / R. Franco (eds.)
**Iron Metabolism – Inorganic Biochemistry and Regulatory Mechanism**
1999. XXXII. 379 pages with 344 figures. Hardcover.
ISBN 3-527-29653-0

F. Marks / G. Fürstenberger (ed.)
**Protein Phosphorylation**
1996. XXIII. 404 pages with 67 figures, 8 in color, and 13 tables. Hard cover.
ISBN 3-527-29241-1

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Preface

This book is based on lectures on regulation and signal transduction that are offered to students of biochemistry, biology and chemistry at the University of Bayreuth. During the preparation of these lectures I realized that it is extremely difficult to achieve an overview of the area of signal transduction and regulation. Our knowledge of signal transduction processes has exploded in the past ten years and this fast progress has been reflected only slowly in the major textbooks. Furthermore, our progress in understanding signal transduction processes has increased to a point where – in contrast to the situation a decade ago – the basic priciples of intra- and intercellular signaling are quite well established. Importantly, signaling processes can be described nowadays more and more on a molecular level. The great increase in structural and biochemical information on signaling processes provides us now the rational chemical and biochemical basis that is required for understanding the interplay between signaling molecules and the biological function of signaling pathways.

It is the aim of the present book to describe the structural and biochemical properties of signaling molecules and their regulation, the interaction of signaling proteins at the various levels of signal transduction and to work out the basic principles of cellular communication. As far as possible molecular aspects have been included. Starting from regulation at the level of genes and of enzymes the book concentrates on the major intracellular signaling molecules and signaling pathways and then describes the interplay and cooperation of various signaling pathways in central cellular processes like cell cycle regulation, tumorigensis and apoptosis.

Signaling and regulation processes influence all aspects of cellular function and a book on this topic necessarily must confine on the exemplary aspects. Numerous studies in very diverse systems have revealed that the basic principles of signaling and regulation are similar in all higher organims. Therefore the book concentrates on the best studied reactions and components of selected signaling pathways and does not try to describe distinct signaling pathways (e.g. the vision process) in a complete way. Furthermore results from very different eucaryotic organisms and tissues have been included. Due to the huge number of publications on the topic, the references cited had to be highly selected for and it may be forgiven that mostly reviews are cited and that original articles have been selected on a more or less subjective basis.

Cellular signaling in higher organisms is a major topic in modern medical and pharmacological research and is of central importance in biomolecular sciences. Accordingly, the book concentrates on signaling and regulation in animal systems and in man. Plant systems could not be considered and results from lower eucaryotes and procaryotes are only cited if they are of exemplary character.

The present book is based on a german edition which appeared in 1997. Where necessary the book has been updated citing data from up to 1998. The rapid progress in some areas made it necessary to rewrite some chapters as e.g. on apoptosis completely.

I am grateful to all people who have encouraged me to write the book and who have supported me with many helpful comments and corrections. In first place I want to thank my colleague Mathias Sprinzl and my former coworkers Carl Christian Gallert and Oliver Hobert. I am also grateful to Ralph Schubert, Joachim Reischl and Hannes Krauss for the figures and structure presentations.

Bayreuth, October 1999 Gerhard Krauss

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Overview of Chapters

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Contents

## Chapter 1
## The Regulation of Gene Expression

## Chapter 2
## The Regulation of Enzyme Activity

## Chapter 3
## Function and Stucture of Signaling Pathways

## Chapter 4
## Signaling by Nuclear Receptors



## Chapter 5
## G-protein Coupled Signal Transmission Pathways

## Chapter 6
## Intracellular Messenger Substances: “Second Messengers“

## Chapter 7
## Ser/Thr-specific Protein Kinases and Protein Phosphatases

## Chapter 8
## Signal Transmission via Transmembrane Receptors with Tyrosine-specific Protein Kinase Activity

## Chapter 9
## Signal Transmission via Ras Proteins



## Chapter 10
## Intracellular Signal Transduction: the Protein Cascades of the MAP Kinase Pathways

## Chapter 11
## Membrane Receptors with Associated Tyrosine Kinase Activity



## Chapter 12
## Other Receptor Classes



## Chapter 13
## Regulation of the Cell Cycle

## Chapter 14
## Malfunction of Signaling Pathways and Tumorigenesis: Oncogenes and Tumor Suppressor Genes

## Chapter 15
## Apoptosis



## Chapter 16
## Ion Channels and Signal Transduction

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Subject Index

## G

## R

## S

## S



## T

## U

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Chapter 1

# The Regulation of Gene Expression

## 1.1 Regulation of Gene Expression: How and Where? A Schematic Overview

The transfer of genetic information from the level of the nucleic acid sequence of a gene to the level of the amino acid sequence of a protein or to the nucleotide sequence of RNA is termed gene expression. The entire process of gene expression in eucaryotes includes the following steps:

- *transcription*: formation of a primary transcript, the pre-mRNA
- *conversion of the pre-mRNA into the mature mRNA which includes*: Processing, splicing, transport from the nucleus to the cytosol
- *translation*: synthesis of the protein on the ribosome.

The expression of genes follows a tissue and cell-specific pattern, which determines the function and morphology of a cell. In addition, all development and differentiation events are also characterized by a variable pattern of gene expression. The regulation of gene expression thus plays a central role in the development and function of an organism. Due to the multitude of individual processes which are involved in gene expression, there are many potential regulatory sites (Fig. 1.1).

### Regulation of Transcription

At the level of transcription it can be determined if a gene is transcribed *at all* at a given time point.

The chromatin structure plays an important role in this decision. Certain chromatin structures can effectively inhibit transcription and totally shut down a gene. This „silencing“ of genes is often observed in development and differentiation processes. The methylation of DNA at cytidine residues is involved in the silencing of genes. The activation of silenced genes requires a reorganization of the chromatin. This little understood process fulfills the prerequisites for transcription initiation and, furthermore, represents a further possibility to regulate gene expression at the level of transcription. Efficient transcription initiation requires the formation of a transcription initiation complex at the starting point of transcription. Involved in this event are, aside from the RNA polymerase, further proteins (transcription factors) which can inf-

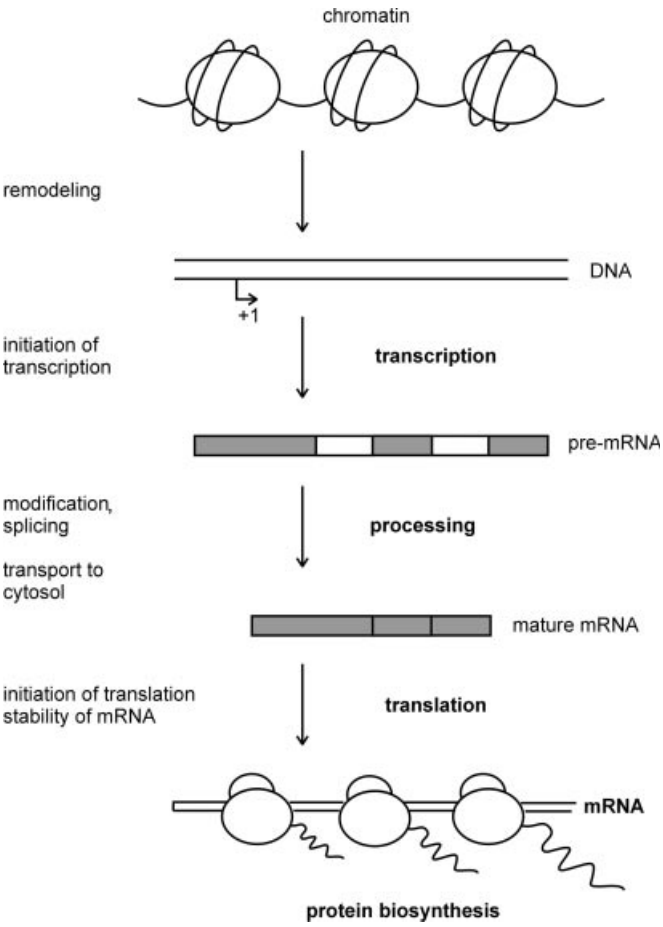



**Fig. 1.1.** Levels of regulation of eucaryotic gene expression

luence transcription in a specific or unspecific manner. The formation of a functional initiation complex is often the rate limiting step in transcription and is subject to a variety of regulation mechanisms.

### Conversion of the pre-mRNA into the mature mRNA

Transcription of genes in mammals often initially produces a pre-mRNA, whose information content can be modulated by subsequent polyadenylation or splicing. Various final mRNAs coding for proteins with varying function and localization can be produced in this manner starting from a single primary transcript.

**Regulation at the Translation Level**

The use of a particular mature mRNA for protein biosynthesis is also highly regulated. The regulation can occur via the accessibility of the mRNA for the ribosome or via the initiation of protein biosynthesis on the ribosome. In this manner a given level of mature mRNA can specifically determine when and how much a protein is synthesized on the ribosome.

**Nature of the Regulatory Signals**

Regulation always implies that signals are received, processed and translated into a resulting action. The nature of the signals employed in the course of the regulation of gene expression, which are finally translated into a change in protein concentration, can vary dramatically. Regulatory molecules can be small molecular metabolites, hormones, proteins or ions. The signals can be of external origin or can be produced internally. External signals can be environmental in nature, such as light, warmth, pressure or electrical signals, or can originate from other tissues or cells of the organism. The external signals are transferred across the cell membrane into the interior of the cell where they are transduced to the level of transcription or translation. Complex signal chains are often involved in the transduction.

## 1.2 Protein-Nucleic Acid Interactions as a Basis for Specific Gene Regulation

A recurring motif on the pathway of information transfer from gene to protein is the binding of proteins to nucleic acid. Specific interactions between proteins and nucleic acids are found not only at the level of DNA, but also at the RNA level. At the DNA level, specific DNA-binding proteins aid in the identification of genes for regulation via transcriptional activation or inhibition. At the RNA level, specific RNAs are recognized in a sequence-specific manner to attain a controlled transfer of genetic information further on to the mature protein.

The basis of all specific regulation processes at the nucleic acid level is the recognition of nucleotide sequences by binding proteins. A binding protein usually recognizes a certain DNA or RNA sequence, termed the *recognition sequence* or *DNA-binding element*. Due to the enormous complexity of the genome, the specificity of this recognition plays a significant role. The binding protein must be capable of specifically picking out the recognition sequence in a background of a multitude of other sequences and binding to it. The binding protein must be able to discriminate against related sequences which differ from the actual recognition element at only one or more positions.

An understanding of the mechanism by which the highly specific and selective recognition of a nucleotide sequence is achieved is only possible with knowledge of the structural details of specific protein-nucleic acid complexes. For the regulation of gene activity the binding of proteins to double-stranded DNA is of central importance. We

will therefore limit our following discussion to specific complexes between double-stranded DNA and protein.

The current structural information on specific protein-DNA complexes allow the first answers to the following basic questions:

– which structural elements of the protein participate in the recognition?
– which interactions impart the specific contact between protein and DNA?
– what role is played by sequence and conformation of the DNA?

### 1.2.1 Structural Motifs of DNA-Binding Proteins

DNA-binding proteins contact their recognition sequences via defined structural elements, termed DNA-binding motifs (overview: Pabo & Sauer, 1992; Burley, 1994). DNA-binding motifs are often found in structural elements of the protein which can fold independently from the rest of the protein and therefore represent separate DNA-binding domains. They can, however, also occur within sequence elements which can not independently fold, but whose folding depends on the tertiary structure of the rest of the protein.

The region of the binding protein which interacts with the recognition sequence often displays a characteristic small structural element which is stabilized through the help of other structural elements and is thereby brought into a defined position relative to the DNA. These structural elements, the „DNA binding sites“, contain short α-helical and β-sheet structures. Contact of the binding site with the DNA sequence usually occurs within the major groove; there are, however, examples for interactions with the minor groove of the double helix (TATA-Box binding protein, see 1.2.3.2 and Fig. 1.16). The dimensions of the major groove of the DNA make it well suited to accept an α-helix. Accordingly, α-helices are often utilized as recognition elements. There are examples of other DNA-binding proteins in which flexible structures are involved in contact to the DNA.

Altogether the variety of participating structural elements is much greater than originally assumed. A number of other structural elements have joined the originally described helix-turn-helix motif of bacterial repressors, to demonstrate the wide variety of mechanisms proteins employ to contact specific DNA sequences, and how the recognition motif can be integrated into the overall structure of the DNA-binding protein. The numerous sequential and structural information available on DNA-binding proteins allow them to be classified into various classes of DNA-binding motifs. The classification of a newly identified protein is often performed on the basis of sequence comparison alone, although, strictly speaking, one should await the analysis of crystal data. Following is an introduction to the most common and well-characterized DNA-binding motifs:

### 1.2.1.1 Helix-Turn-Helix Motif

The helix-turn-helix motif (HTH motif) is – historically seen – the first DNA-binding motif whose structure could be solved in a complex with DNA. It is often found in bacterial repressors. Many eucaryotic DNA-binding proteins also utilize the helix-turn-helix motif for specific binding on the DNA. An example is the homeodomain binding protein „engrailed“ from Drosophila (review: Wolberger, 1996). Characteristic for the helix-turn-helix motif is the positioning of an α-helix in the major groove of DNA (Fig. 1.2a&b). The recognition helix is connected by a turn to another helix, whereby the position of the recognition helix is fixed. The two helices occur at a 120° angle to one other. The binding motif is usually stabilized by further helices of the same or another subunit. The detailed arrangement can differ significantly among the various helix-turn-helix motifs.

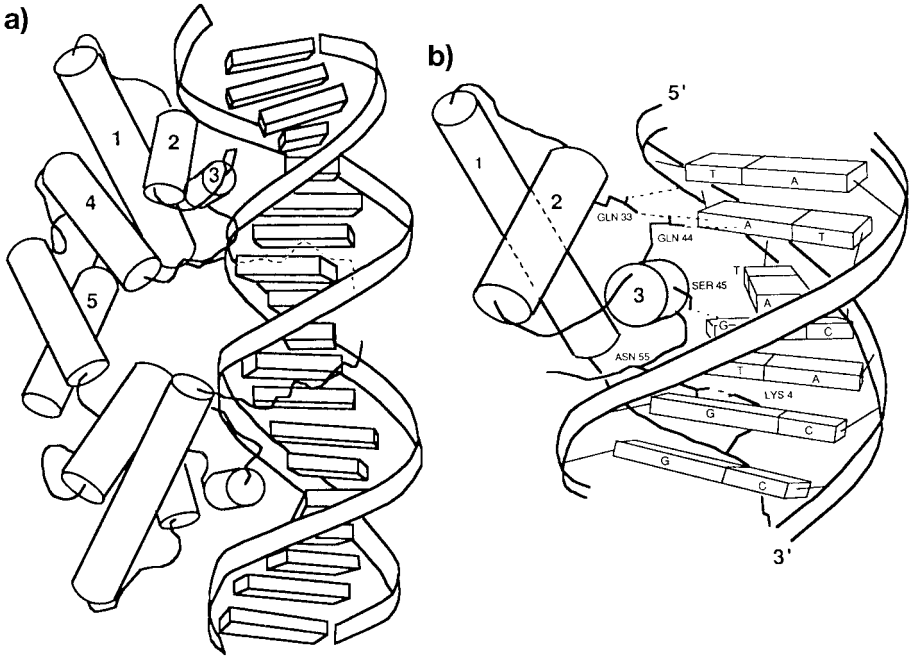



**Fig. 1.2.** The helix-turn-helix motif in complex with DNA. a) side view of the λ-repressor in complex with DNA. The α-helices are drawn as cylinders. In the upper subunit the α-helices are numbered. Helices 2 and 3 form the classic helix-turn-helix motif. b) detailed side view of the binding of a monomer of the λ-repressor to the recognition half site emphasizing the most important protein-DNA contacts. This view displays the embedding of helix 3 in the major groove of DNA. After Pabo and Sauer (1992), with permission. c) side view of the complex of the eucaryotic „homeodomain“ binding protein „engrailed“ with the cognate TAATX binding element. d) the DNA-binding domain of the repressor of the 434 phage in complex with the recognition sequence ACAA. After Harrison (1991), with permission.

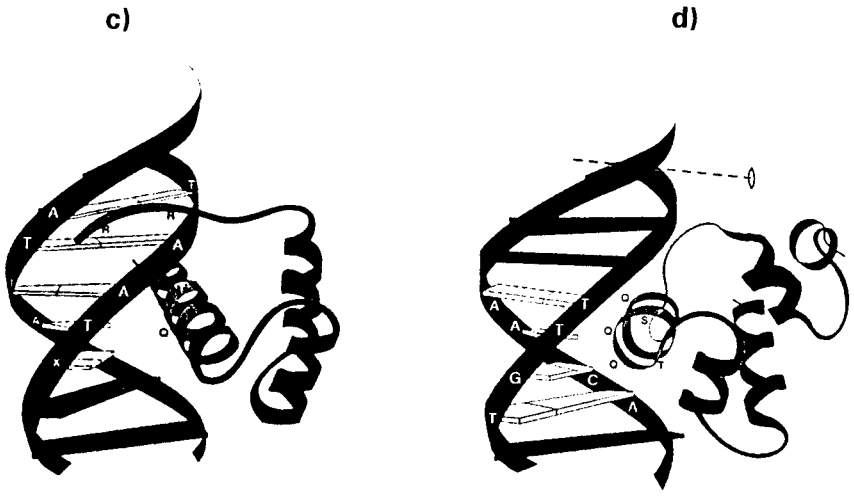



**Fig. 1.2.** continued

#### 1.2.1.2 Binding Motifs with Zinc Ions

The zinc binding motifs contain $Zn^{2+}$ complexed by four ligating Cys and/or His residues (review: Berg 1993). Based on the stoichiometry of the complex, zinc fingers of the type Zinc-$Cys_2His_2$, Zinc-$Cys_4$ and $Zinc_2$-$Cys_6$ can be distinguished (Fig. 1.3).

**Classical Zinc Fingers**

The first zinc binding motif discovered was that of the eucaryotic transcription factor TFIIIA of *Xenopus laevis* which contains 9 copies of a $Cys_2His_2$-Zinc motif. The structure of the binding motif is shown in Fig. 1.4. The central zinc ion serves to pack an α-helix against a β-sheet and thereby position the α-helix. The recognition of the DNA sequence occurs via this α-helix.

A very similar zinc finger is found in Zif268, a regulatory DNA-binding protein of mice (Pavletich and Pabo, 1991). The structure of the Zif268-DNA complex is shown in Fig. 1.5. In Zif268, three of the zinc-fingers are arranged along the coil of the DNA. The DNA-binding element contains three repeats of the recognition sequence. This results in a modular construction of the protein, so that the periodicity of the DNA is reflected in the protein structure.

The zinc binding element plays, above all, a structuring role by ensuring that the recognition helix is correctly oriented and stabilized. The zinc ion does not contact the DNA directly. In Zif268 the zinc motif participates directly in the DNA-binding via formation of a H-bond between the His residue of the zinc complex and the N7 of a G:C base pair of the DNA.

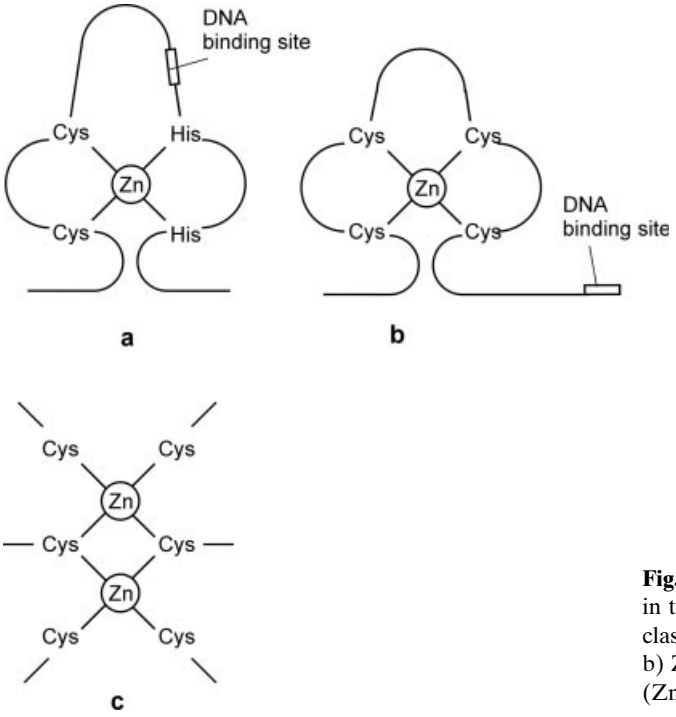



**Fig. 1.3.** Complexation of $Zn^{2+}$ in the Zn-binding motif. a) classical $Zn^{2+}$ $Cys_2His_2$ finger; b) $Zn^{2+}$ $Cys_4$ binding motif; c) $(Zn^{2+})_2$ $Cys_6$ binding motif.

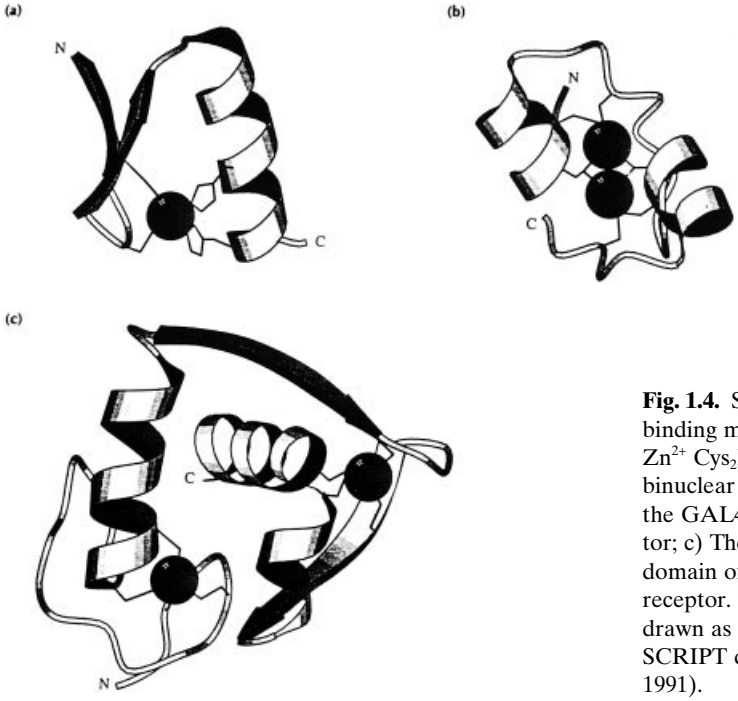



**Fig. 1.4.** Structures of $Zn^{2+}$ binding motifs. a) TFIIA-like $Zn^{2+}$ $Cys_2His_2$ finger; b) the binuclear $(Zn^{2+})_2$ $Cys_6$ motif of the GAL4 transcription activator; c) The DNA-binding domain of the gluccocorticoid receptor. The $Zn^{2+}$ ions are drawn as spheres. MOLSCRIPT drawing (Kraulis, 1991).

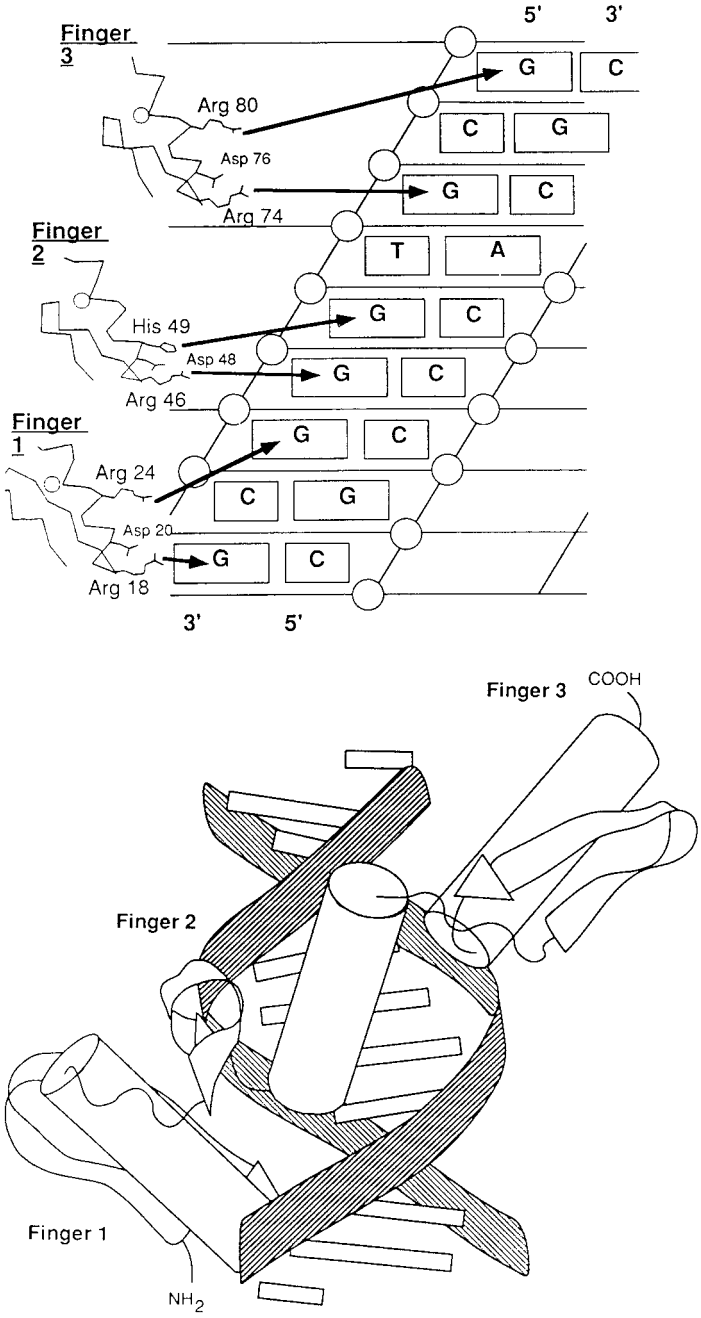



**Fig. 1.5.** Zif268 in complex with DNA. a) specific H-bonds between amino acid side chains of fingers 1–3 of Zif268 and bases of the recognition sequence. The DNA is drawn as cylinders. The arrows emphasize contact with the major groove. b) periodic arrangement of fingers in the major groove of the DNA. According Pabo and Sauer (1992), with permission.

**Zinc Finger of the Steroid Hormone Receptor**

The DNA-binding of the steroid hormone receptor occurs via an approx. 60 amino acid DNA-binding domain with two zinc-$Cys_4$-motifs (see also 4.3.2). The structure displays two so-called helix-loop-helix elements, each with a bound $Zn^{2+}$ ion (Fig. 1.6).

Both zinc ions are each complexed by 4 Cys residues, whereby a non-equivalent arrangement of the two $Zn^{2+}$ ions is observed. The binding specificity is accomplished by amino acid residues near the N-terminus of the first helix in a helix-loop-helix element.

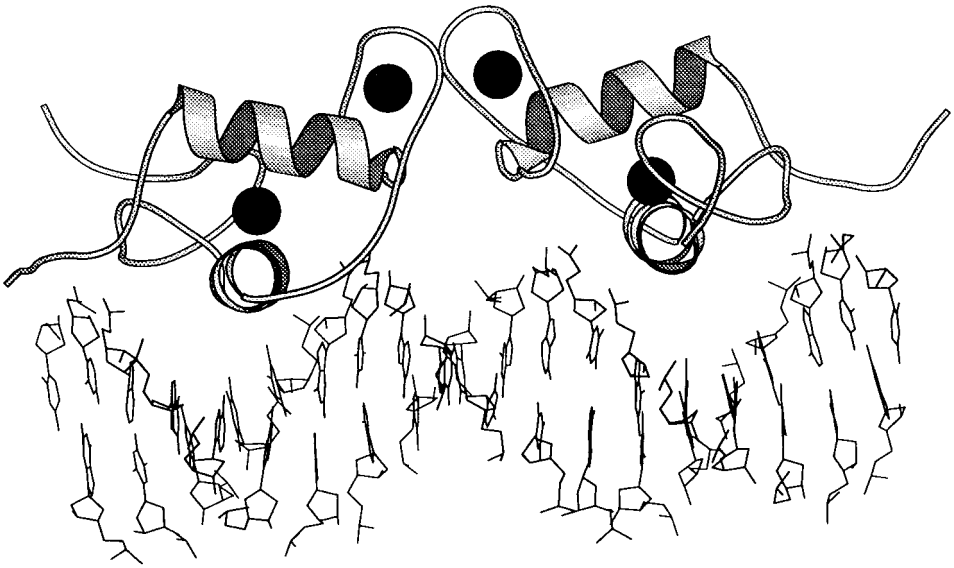


**Fig. 1.6.** The Zn binding motif of the glucocorticoid receptor in complex with DNA. Shown is the complex of the dimeric DNA-binding domain of the glucocorticoid receptor with the cognate DNA element (Luisi et al., 1991). The $Zn^{2+}$ ions are shown as spheres. The two $Zn^{2+}$ ions are clearly non-equivalent. While one of the $Zn^{2+}$ ions aids in the fixation of the recognition helix in the major groove, the other correctly positions a structural element for the dimerization of the monomers. MOLSCRIPT drawing (Kraulis, 1991).

**Transcriptional activator GAL4 of Yeast**

Here two zinc ions are complexed by 6 Cys residues, whereby two of the Cys residues bind to both $Zn^{2+}$ ligands (see Fig. 1.3 & Fig. 1.4). The structure of GAL4 complexed with its recognition sequence indicates that the $Zn^{2+}$ ions mainly act to stabilize the small globular GAL4 protein and to orient the recognition helix correctly within the major groove.

Overall, the zinc-binding motifs display a great variety of structural diversity. The occurrence of a zinc binding motif can often be predicted based solely on a characteristic series of Cys and His residues in a protein sequence. The complexation of a $Zn^{2+}$ by His and Cys residues serves to bring the recognition element of the protein into a stable and unambiguous position relative to the DNA, thereby enabling specific contacts with the recognition sequence.

### 1.2.1.3 Basic Leucine Zipper and Helix-Loop-Helix Motifs

This group of binding motifs displays as characteristic structural element an extended bundle of two α-helices. The two α-helices are wound around each other in the form of a „coiled-coil". At their end is a basic region which mediates the DNA-binding (review: Ellenberger, 1994).

**Basic Leucine Zipper**

An example for the structure of a basic leucine zipper in complex with DNA is shown by the transcription factor GCN4 from yeast in Fig. 1.7a. *The leucine zipper takes its name from the regular occurrence of leucine residues (or other hydrophobic residues) in an α-helix.* A leucine or other hydrophobic amino acid is found at every seventh position of the helix (Fig. 1.8). This sequential arrangement brings the hydrophobic residues all along one face of the helix, and the hydrophobic residues of two helices can interlock via hydrophobic interaction in a zipper-like manner. The leucine zipper is, above all, a tool to associate proteins in higher dimensions, whereby homodimers as well as heterodimers can be formed. The oligomerization of DNA-binding proteins is usually a prerequisite for strong binding to the cognate DNA element.

The leucine zipper itself does not participate in the recognition; it is only utilized for dimerization of the proteins. The N-terminal end of the basic leucine zipper motif is relatively unstructured in the absence of DNA. A helical structure is induced upon binding to DNA allowing specific contacts to the recognition sequence. Dimer formation is a prerequisite for the exact positioning of the N-terminal basic end in the major groove of the DNA. Analogous to the dimeric structure of the protein, the DNA sequence displays 2-fold symmetry (see 1.2.4).

**The Helix-Loop-Helix Motif**

One example of the basic helix-loop-helix motif (HLH-motif) is found in the eucaryotic transcription factor Max (Fig. 1.7b and 15.3.2). The DNA-binding occurs by a paral-

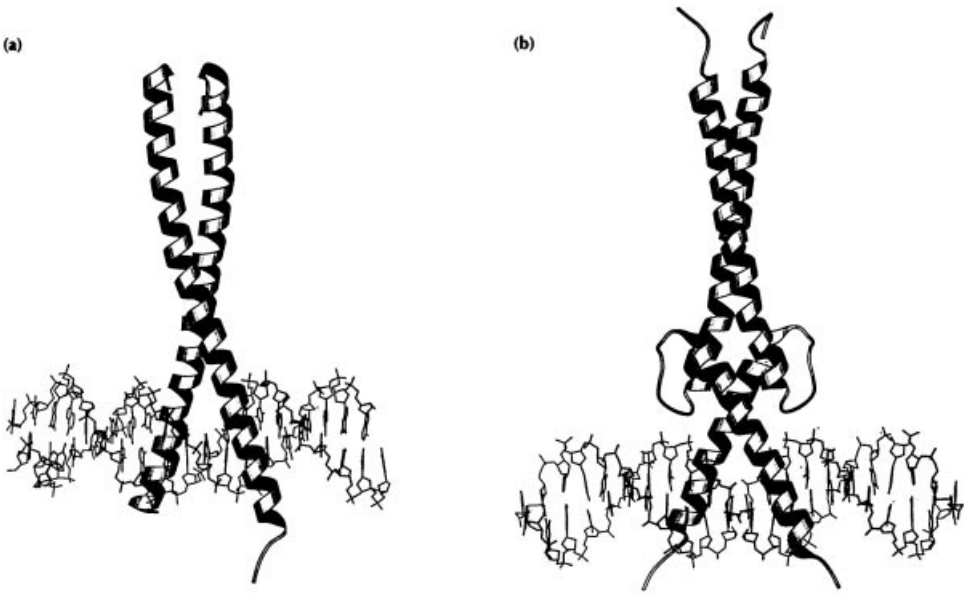



**Fig. 1.7.** Basic leucine zipper and helix-loop-helix motif in complex with DNA. A) The basic leucine zipper of the transcription activator GCN4 of yeast consists of two slightly curved α-helices, which dimerize with the help of the leucine zipper motif. The sequence specific binding of DNA occurs via the basic ends of the two helices. They insert themselves into the major groove of the DNA. B) The helix-loop-helix motif of the eucaryotic transcription factor Max complexed with DNA. Molscript drawing (Kraulis 1991).

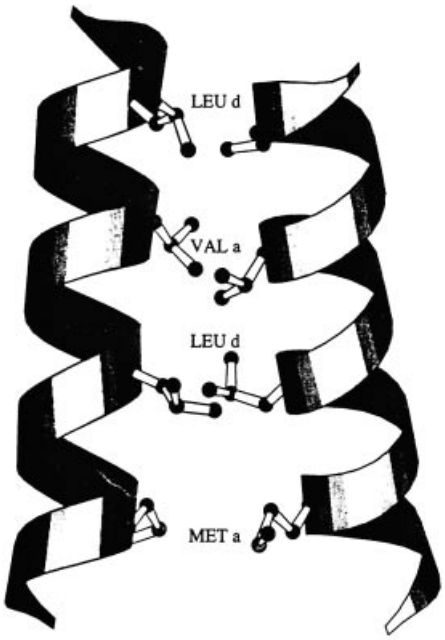



**Fig. 1.8.** Packing of the amino acids in the interior of a leucine zipper, after Ellenberger (1994), with permission.

lel bundle of 4 helices with two basic ends. As with the basic leucine zipper motif, the basic ends only attain a defined structure upon binding the DNA. The 4 helix bundle forms via dimerization of two subunits of the Max protein. A structural element similar to that of the leucine zipper is responsible for the dimerization and stretches from the helix-loop-helix structure in the direction of the C-terminus.

#### 1.2.1.4 DNA-binding via β-Sheet Structures

β-sheet structures as DNA-binding motifs are found in pro- and eucaryotic DNA-binding proteins. As an example, the structure of the MetJ repressor from E. coli is shown in Fig. 1.9. The DNA is contacted in the major groove by the protruding β-strands.

The eukaryotic transcription factor NFκB also binds DNA via β-sheet structure (Fig. 1.10). Noteworthy is the enshrouding of the DNA by the β-sheets of NFκB. The recognition of the DNA elements is also achieved by interaction with the major groove of the DNA.

#### 1.2.1.5 Flexible Structures in DNA-binding Proteins

A series of DNA-binding proteins utilize additional flexible structures aside from defined structural DNA-binding motifs in order to increase the stability and specificity of the complex. The λ repressor grabs around the DNA helix with the flexible N-terminal arm of the protein to contact the back side of the helix. The basic region of the leucine zipper and HLH binding protein is a further example for the importance of protein flexibility in DNA-binding. In the absence of DNA the basic portion of this binding motif is poorly structured, and only following DNA-binding is an α-helix formed in the basic region. The α-helix induced upon binding lies in the major groove of the DNA and establishes specific interactions with the recognition sequence.

**Fig. 1.9.** DNA-binding via β-pleated sheets. The repressor MetJ (E. coli) complexed with the half-site of its operator sequence. The binding occurs via two parallel β-sheets in the major groove of the DNA. Molscript drawing (Kraulis 1991).

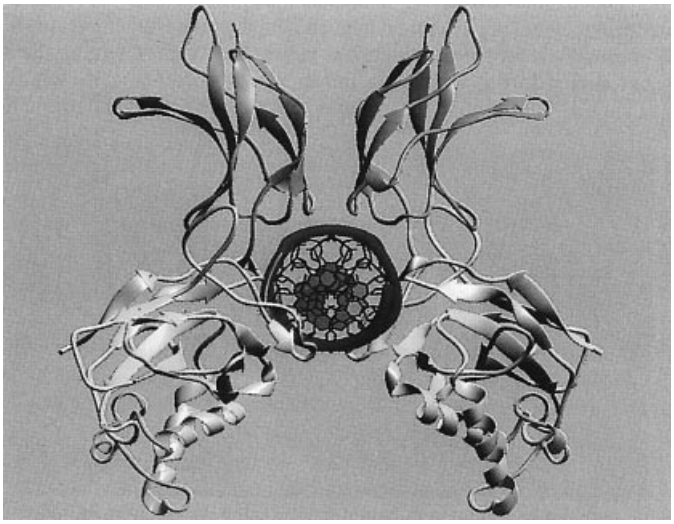


**Fig. 1.10.** The eucaryotic transcription factor NFκB in complex with DNA. Shown is the structure of a fragment of the p50 subunit of NFκB complexed with the recognition sequence. p50 of NFκB binds DNA as a dimer. Each of the subunits contains a bundle of β-sheets which envelops the DNA so that only the minor groove is exposed. After Ghosh et al. (1995), with permission.

## 1.2.2 The Nature of the specific Interactions in Protein-Nucleic Acid Complexes

The binding of a protein to nucleic acid is accomplished by weak, non-covalent interactions. The interactions are the same as those involved in the formation of the tertiary structure of a protein:

- Hydrogen bonds (H-bonds)
- Electrostatic interactions
- Van der Waals interactions
- Hydrophobic interactions

### 1.2.2.1 H-bonds in Protein-Nucleic Acid Complexes

Of central importance for the formation of a specific protein-DNA complex are hydrogen bonds. The H-bonds are clearly identifiable in high resolution structures. H-bonds occur where a H-bond donor and acceptor lie with 0.27–0.31 nm of each other. Energetically most favorable is the linear arrangement of the H-bond, with deviations from linearity leading to a reduction in energy. This characteristic is responsible for the stereospecific orientation of H-bond acceptors and donors. The H-bond thus contributes significantly to the spatial orientation between protein and nucleic acid.

There are many different H-bond donors as well as acceptors in proteins and nucleic acids which contribute to the specific recognition. Important H-bond donors and acceptors in proteins are Asn, Gln, Ser, Thr, Tyr, Glu, Asp, Arg, Lys, Cys and His. The peptide bonds of the backbone often participate, as well.

The heteroatoms and exocyclic functional groups of the bases within the nucleic acid can form H-bonds to residues of a binding protein, in addition to base pairing. Also, the oxygen of the ribose or deoxyribose and the phosphate moiety of DNA can be used as H-bond acceptors.

The various base pairs, e.g. A:T vs G:C, can be individually distinguished based on their pattern of H-bond donors and H-bond acceptors, as viewed from the major groove (Fig. 1.11).

The available structural information on protein-DNA complexes shows that mother nature uses the spectrum of possible H-bond interaction in a flexible manner. Originally it was assumed that, similar to the genetic code, a specific code for contacting a base pair by amino acids existed.

This idea has been refuted by the available structural information. There are many possibilities for an amino acid to contact a base pair, and this repertoire is put touse. Examples for the variety of H-bond interactions are shown in Fig. 1.12.

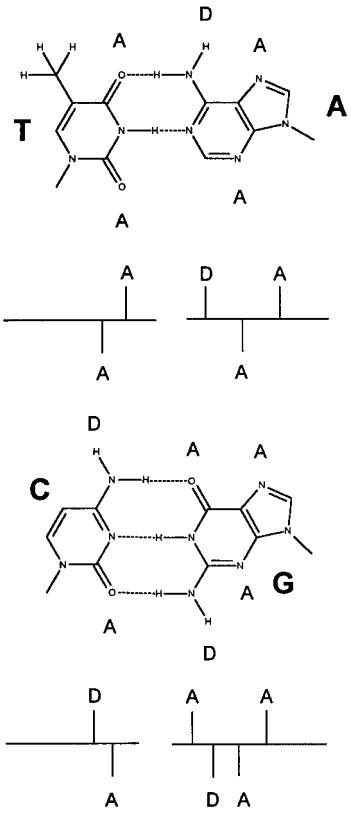



**Fig. 1.11.** H-bond donors (D) and H-bond acceptors (A) in A:T and G:C base pairs. Schematic display of the differing pattern of H-bond acceptors and donors in the Watson-Crick base pairs. The groups above the base pairs (above the line) are accessible in the major groove, and those below the line are accessible from the minor groove.

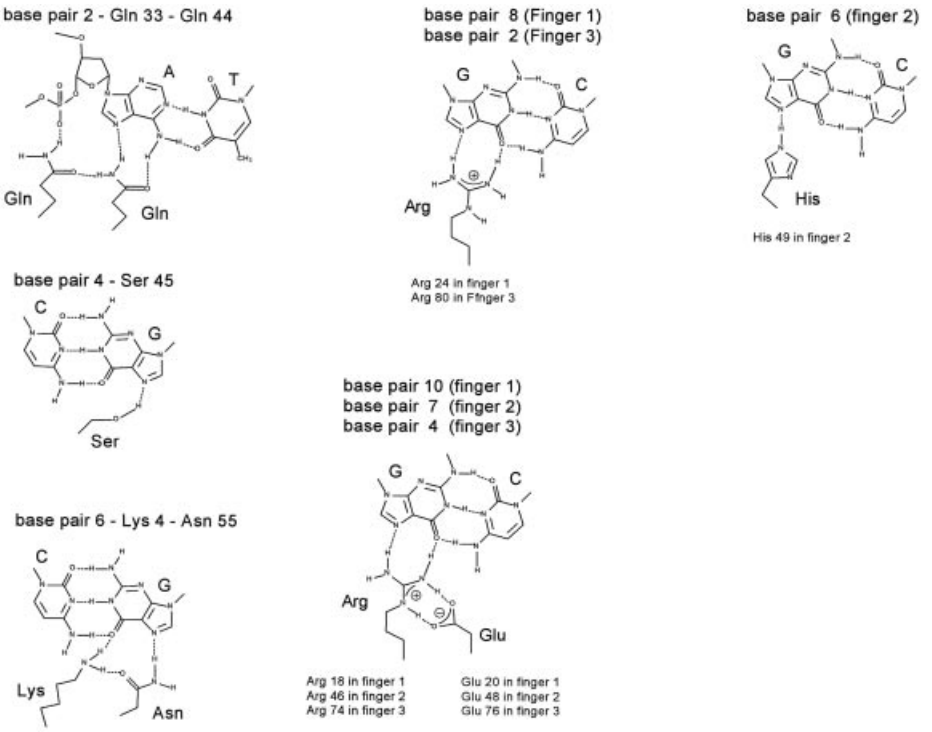



**Fig. 1.12.** Examples for the H-bonds in protein-nucleic acid complexes. A) H-bond contacts of the λ-repressor in complex with its operator sequence. After Jordan & Pabo, (1988). B) H-bonds in the complex between the Zinc fingers of Zif268 with the cognate recognition helix. Zif268 contacts the DNA with three Zn-fingers (finger 1–3 in Fig. 1.5). Shown are the H-bond contacts formed between the fingers and the base pairs of the recognition sequence. After Pavletich & Pabo, (1991).

The following points are noteworthy:

A base can be contacted by more than one amino acid residue. Furthermore, there are many examples of one amino acid residue, e.g. Arg, contacting two sequential bases. This type of interaction functions as a clip and maintains a spatially defined arrangement.

The contact between protein and DNA can also be transmitted via bound water molecules. In the crystal structure of the complex of the bacterial Trp-repressor and the cognate operator sequence are found only a few direct H-bonds between the amino acid residues of the protein and the bases of the recognition sequence. Rather, the contacts between protein and nucleic acid are frequently established indirectly by a chain of well-defined bound water molecules which contact the protein and the bases, and thereby function as transmitter between the protein and DNA.

There are always numerous H-bond contacts formed between the recognition sequence and the binding protein. The pattern of H-bond donors and H-bond acceptors is determined by the sequence and conformation of the DNA as well as by the specific structure of the protein. Both together lay the foundation for a specific recognition of the DNA by the protein.

**The Role of the Peptide Backbone**

An important factor in the structure of protein-DNA complexes can be the peptide backbone. The amide bond can function as an H-bond acceptor as well an H-bond donor. Due to the reduced flexibility of the backbone vs. side chain (resonance stabilization of the peptide bond), H-bonds to the peptide backbone lead to a rigid and tight arrangement in the complex and contribute extensively to the exact fit between protein and nucleic acid.

### 1.2.2.2 Ionic Interactions

Ionic interactions result from the electrostatic attraction or repulsion between charged groups. As opposed to H-bonds, ionic interactions are not directed and are effective over greater distances.

DNA presents itself to a binding protein as a negatively charged, anionic substrate. Accordingly, the protein displays a complementary positive potential, resulting from an accumulation of basic amino acid residues. The electrostatic interaction between the two oppositely charged binding surfaces of DNA and protein make a significant energetic contribution to the formation of a stable complex.

The ionic interactions are, however, less suitable to distinguish between various base pairs since only the phosphates of the backbone from the DNA are involved in the interaction. Together with the specific H-bonds, the non-specific ionic interactions contribute significantly to the formation of a stable complex. The positively charged surface of DNA-binding proteins is also the reason for the ability of many such proteins to bind DNA nonspecifically.

The compensation of the negative charges of DNA can also have a further effect. It has been shown that the neutralization of the negative phosphate charge on one side to the DNA helix can lead to bending of the DNA (Fig. 1.13). A charge neutralization by a binding protein can, in this manner, favor DNA bending (Strauss and Maher, 1994).

### 1.2.2.3 Van der Waals Contacts

The van der Waals' contacts are a type of electrostatic interaction and arise from an interaction between permanent and/or induced dipoles in the bond pair. They are typically effective over a much shorter range than ionic interactions. The contribution of van der Waals contacts to the binding of a protein to a DNA sequence is difficult to

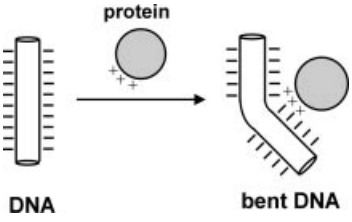



**Fig. 1.13.** Bending of DNA as a result of charge neutralization by a DNA-binding protein. The negatively charged DNA bends upon binding the positively charged protein surface. On the side of the DNA facing away from the protein excess negative charges build up and repel each other. After Strauss & Maher (1994).

estimate, since many small contributions must be considered. An example for a contact surface with many van der Waals interactions can be found in the complex of the TATA box binding protein with the TATA box (see Fig. 1.16). In this case there are extensive van der Waals contacts between the sugar residues of the DNA backbone and the hydrophobic surface of the protein. Furthermore, phenylalanine residues stack between the bases and are thus fixed via hydrophobic interactions (Kim et al., 1993).

## 1.2.3 The Role of the DNA Conformation in Protein-DNA Interactions

The double helix of the DNA can only to a first approximation be considered a linear, rod-like structure with the typical coordinates of B-DNA. Actually DNA possesses considerable flexibility and conformational variability. The flexibility and structural polymorphism of DNA are prerequisites for many of the regulatory processes on the DNA level (review: Harrington, 1994; Alleman and Egli, 1997). Local deviation from the classical B-structure of DNA, as well as bending of the DNA, are observed in many protein-DNA complexes.

### 1.2.3.1 Local Conformational Changes of DNA

In recent years an astonishing structural variety has been uncovered for DNA. Crystal structures have shown that, apart from the structural motifs of the A-, B- and Z-forms of DNA, other, sequence-dependent structural variations exist which are observed when smaller sequence fragments are examined in detail.

The structural variations can affect the width of the major groove, the extent of base stacking, as well as the tilt of the basepairs to each other. The local conformational changes are sequence dependent and can be intrinsic properties and thus permanent occurrences; they can, however, also be induced by protein binding. The DNA sequence can thus serve a double purpose for the recognition between DNA and protein.

*Direct recognition:* The order of bases can determine the pattern of weak interaction and the specificity of the complex formation. In this case there is a direct recognition of the sequence by the protein.

*Indirect recognition:* The DNA sequence can predetermine a particular conformation, which is a prerequisite for specific protein binding. Alternative DNA conformations will not be bound and recognized. We speak here of an indirect recognition of a sequence:

DNA sequence → DNA conformation → recognition

The detailed analysis of DNA structure in the region of contact with the binding protein often displays distinct divergence from the parameters of classical B-DNA structure. The specific sequence-determined conformation of the DNA is often a prerequisite for a specific recognition. This recognition mechanism is, for example, realized with the Trp-repressor, where the sequence determines a certain spatial arrangement of the

sugar-phosphate backbone. Only this arrangement is complementary to the binding surface of the repressor and enables a strong binding (Otwinowski et al., 1988).

### 1.2.3.2 Bending of DNA

If one traces a longer stretch of a DNA molecule in solution, a clear divergence from linearity becomes evident. Thermally induced structural fluctuations allow a bending of DNA, which is why long DNA molecules are described as a random coil. This bending of the DNA occurs in molecules with a length of more than approx. 200 bp.

The bending of shorter DNA fragments can be attained via specific sequence with intrinsic bending ability as well as by binding of proteins.

#### Sequence Determined Bending of DNA

There are DNA sequence motifs which induce an intrinsic bending of the DNA. For both natural and synthetic DNA it has been show that the periodic occurrence of short dA:dT sequences causes bending of the DNA (Fig. 1.14). Such a short dA-repeat (e.g. $d_{A5}$) leads to an intrinsic bending of the DNA by ca. 18°. If the dA-repeats in the sequence are properly arranged, then a definite bending of the DNA results. The intrinsic bending of DNA is easily detectable by gel electrophoresis: a bent DNA migrates in a native electrophoresis slower than a linear DNA of the same length.

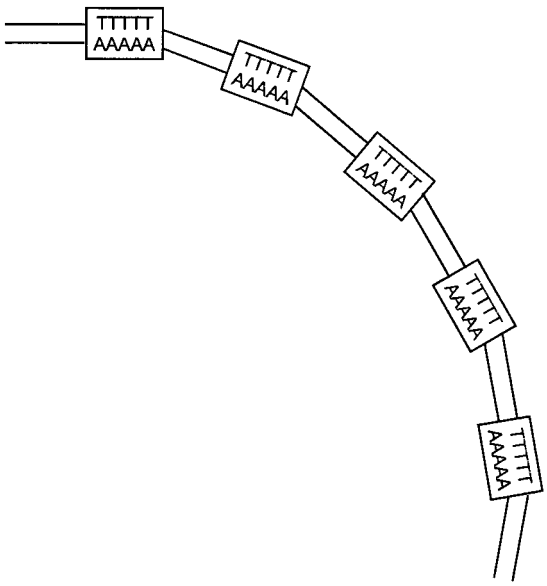



**Fig. 1.14.** Intrinsic bending of DNA via periodic repeat of $(dA)_{5-6}$ sequences. An intrinsic bending of DNA of ca. 18° is induced per $(dA)_{5-6}$ sequence. Poly-dA repeats in 10 bp steps (the rise of the DNA) result in a strong bending of the DNA, since in this configuration the axis of bending lies on the same side of the DNA.

**Protein Induced Bending of DNA**

There are numerous examples for a protein-induced bending of DNA. The bending of a short segment of DNA (150–200 bp) leads to a loss of stacking interactions of the π-electron system of neighboring bases and is energetically unfavorable. Stacking interactions arise from interactions of the π-electron systems of bases atop one another and contribute extensively to the stability of the double helix. An active bending of a short piece of DNA is therefore only possible if the energy loss is compensated for by other favorable interactions. For protein-induced bending of DNA, the energy is provided by the complex formation with the protein. A portion of the favorable interaction energy (H-bonds, hydrophobic interaction, etc.) compensates for the energy required to bend the DNA. An important contribution to the entire binding energy derives from the neutralization of the negative charge of the DNA. A neutralization of the negative charges on only one side of the phosphate backbone by a positively charged protein surface can lead to a bending of the DNA (see Fig. 1.13).

The divergence of the DNA conformation from a rod-like structure is observed to a variable extent. The DNA can be slightly curved or abruptly kinked.

If the DNA is only slightly bent, as observed for the nucleosome-bound DNA, then the required deformation energy is distributed over many base pairs. The energy requirement per base pair is small and can easily be provided by the interaction energy with the protein. Furthermore, such bending displays little sequence specificty.

Kinking of the DNA is observed, for example, in the DNA complex of the CAP protein, as well as for the TATA-box binding protein. In the complex of the CAP protein there are two successive kinks in the DNA, each of which lead to a bending of ca. 40°, resulting in a net bend of 80°-90° (Fig. 1.15).

The TATA-box binding protein causes a kinking of the bound DNA at an angle of ca. 100° (Fig. 1.16). The flexibility of the alternating purine-pyrimidine sequences of the binding site favor a prominent deformation of the DNA with little energy requirement. Thus, in the region of the kink, the minor groove is obviously widened and the DNA strands partially separated. The widening of the minor groove allows numerous van der Waals contacts with the protein (see 1.2.2.3).

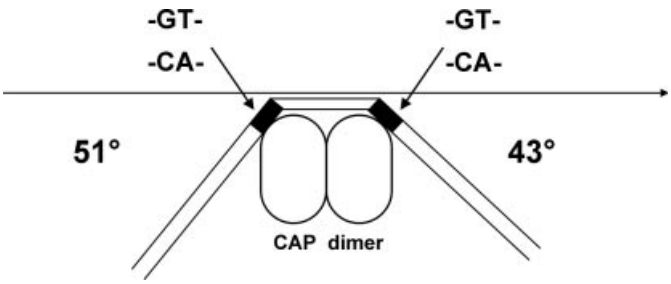



**Fig. 1.15.** Bending of the DNA in the CAP protein-DNA complex. The CAP protein (*E. coli*) binds as a dimer to the two-fold symmetric operator sequence. The DNA is bent nearly 90deg in the complex. The turns are centered around two GT sequences (shown in black) of the recognition element.

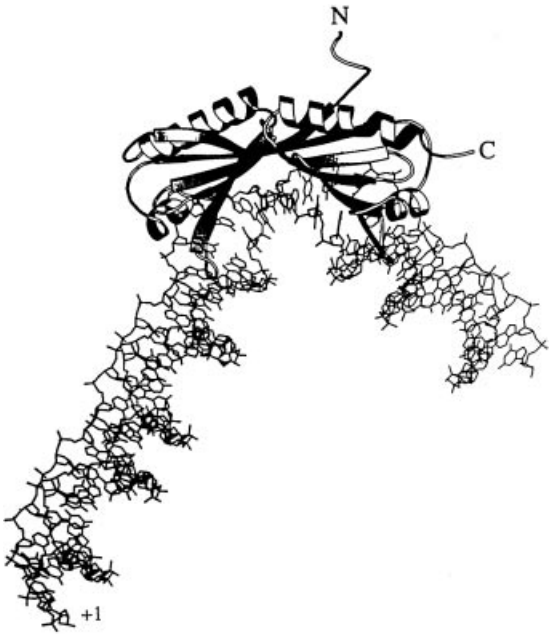



**Fig. 1.16.** Bending of DNA in the TATA box. The DNA is kinked in the complex of the TATA box binding protein (yeast) with the 8 base pair TATA box (Kim et al., 1993). The DNA is deformed in the region near the kink: the minor groove, which faces the protein, is clearly widened. Molscript drawing (Kraulis, 1991).

**What Purpose does the Bending of DNA Serve for Regulatory Processes?**

Regulatory processes at the protein-DNA level require, above all, communication between various DNA-bound proteins, which may not be bound to neighboring sequences. An important role of the actively induced or intrinsic DNA bending is to bring linearly separated DNA sequences, and hence their bound protein, together. (Fig. 1.17). Only by bending DNA is an effective interaction between DNA-binding proteins bound to distant DNA-binding elements possible.

An active bending of the DNA induced by regulatory proteins is of particular importance when a defined arrangement of the DNA in a small volume within a larger nucleoprotein complex is required.

Bending plays an important role in the initiation of transcription. Studies in eucaryotic systems have shown that a highly ordered DNA multi-protein complex is formed at the starting point of transcription. The DNA does not exist as a linear cylinder, but rather is spatially bent due to its bound proteins. The TATA box binding protein functions initially to induce bending of the bound DNA. This bending creates defined binding sites for other components of the transcription initiation complex. Furthermore, other sites are brought within proximity of each other which are separated in their linear sequence. The bending of DNA therefore plays an essential function in gene activation.

A further important function of the bending and deformation of DNA in a protein complex can be to partially melt the sequences, thereby making them accessible for recognition in transcription or DNA replication processes.

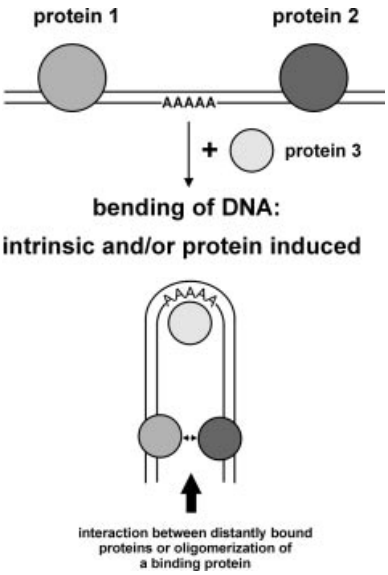



**Fig. 1.17.** The significance of the bending of DNA for protein-protein interactions. DNA-bound proteins, which would not interact if associated with linear DNA, can be brought together through intrinsic or protein-induced bending of the intervening sequences. The bending of the DNA creates a high local concentration of the two proteins and thus enables their effective interaction.

### 1.2.4 Structure of the Recognition Sequence and Quarternary Structure of DNA-binding Proteins

The recognition sequences for specific DNA-binding proteins usually include only 3–8 base pairs, arranged either palindromically or in direct repeats (Fig. 1.18). The symmetry of the sequence in the DNA element is often reflected in the subunit structure of the binding protein. Less common is the occurrence of a singular recognition sequence.

#### Palindromic Arrangement

Palindromic sequences with 2-fold symmetry are usually bound by dimeric proteins in which each subunit of the protein contacts one half-site of the DNA element. The use of 2-fold symmetry in the binding sequence and the protein dimers is an economical approach to achieve high affinity binding. The DNA-binding motif of one subunit often contacts only a few base pairs of the recognition sequence when in a complex. This is generally not sufficient to ensure tight binding of a subunit. The recognition sequence of the E2-protein of papillomavirus is composed e.g. of only three base pairs (Hedge et al., 1992). Due to the repetition of the recognition sequence in a DNA element, binding by the two subunits of a dimeric binding protein occurs in a *cooperative manner*: if one subunit of a protein contacts one half of the recognition sequence, then binding by the other subunit to the other half is strongly favored. Both subunits bind cooperatively and a high affinity binding results. The 2-fold symmetry in the DNA sequence and binding protein plays an important role in the specific binding process. If, for example, a mutation inactivates one half of the recognition sequence, the other intact site often no longer suffices to provide for a tight binding. The protein can then only bind weakly and the mutated DNA element is often inactive in the *in vivo* situation.

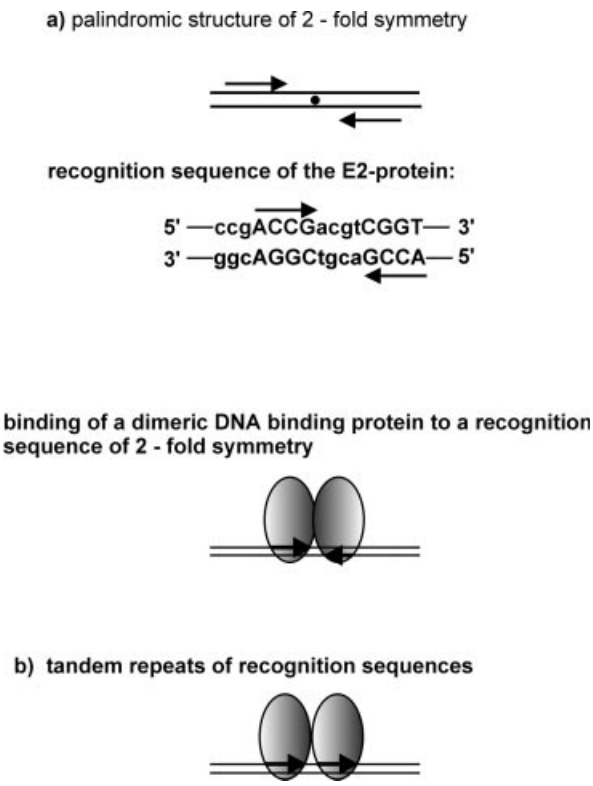



**Fig. 1.18.** Structure and symmetry of DNA recognition elements and the oligomeric structure of DNA-binding proteins. sequence and binding protein plays an important role in the specific binding process. If, for example, a mutation inactivates one half of the recognition sequence, the other intact site often no longer suffices to provide for a tight binding. The protein can then only bind weakly and the mutated DNA element is often inactive in the *in vivo* situation.

### Direct Repeats of the Recognition Sequence

Direct repeat of the recognition sequence requires a nonsymmetrical spatial arrangement of the bound protein subunits (see chapter 4, Nuclear Receptors). The protein-DNA complex has, in this case, a polar character and the protein bound on each of the two halves of the recognition element can carry out different functions. Direct 2-fold repeats are commonly observed for the DNA-binding elements of the steroid hormone receptors (see chapter 4).

The promoter region of procaryotes and eucaryotes often contains multiple repeats of a DNA element. In this case there can be a tandem-like arrangement of the multimers of the DNA-binding protein.

An example for such an arrangement is the MetJ repressor of E. coli. The palindromic recognition sequence of the MetJ repressor occurs in 2–5 copies on the DNA. The repressor itself binds as a dimer on one copy of the recognition sequence. Protein-protein interactions mediate cooperative binding of the repressor dimers to the adjacent copies of the recognition sequence.

The occurrence of tandem-like repeats of the DNA elements, in conjunction with the oligomerization of the cognate DNA-binding protein, allow specific structures to be created which are vital for further regulatory processes. This functional principle is

demonstrated poignantly by the *E. coli* Lac repressor/operator system. Three Lac repressor binding sites are found within a 500 bp stretch in the Lac operon of *E. coli*. Each of the three binding sites has a two-fold palindromic structure on which the Lac repressor binds as a dimer (Lewis et al., 1996). The Lac repressor however exercises its full repressive function as a tetramer. It is therefore assumed that dimers bound to adjacent binding sites associate into tetramers (Fig. 1.19). The intervening sequence of 93 resp. 401 bp forms a loop, also termed the „repression loop". The repressor acts in this arrangement as a clip to bring together widely separated DNA sequences. It is assumed that the specific arrangement of DNA in the loop has decisive consequences for the ensuing transcription activity: the binding of RNA polymerase is *hindered* by the DNA loop, while the loop creates the structural framework for further regulatory proteins to bind, e.g. the Cap protein.

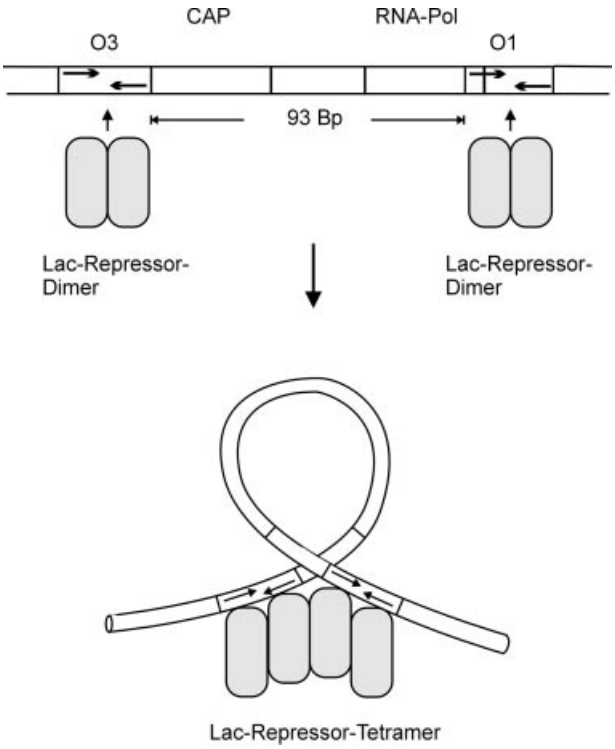



**Fig. 1.19.** Tetramerization of the Lac repressor and loop formation of the DNA. The Lac repressor from *E. coli* binds as a dimer to the two-fold symmetric operator sequence, whereby each of the monomers contacts a half-site of a recognition sequence. The Lac operon of *E. coli* possesses three operator sequences O1, O2 and O3, all three of which are required for complete repression. O1 and O3 are separated by 93 bp, and only these two sequences are displayed in the figure above. Between O1 and O3 is a binding site for the CAP protein and the contact surface for the RNA polymerase. The Lac repressor acts as a tetramer. It is therefore assumed that two dimers of the repressor associate to form the active tetramer, whereby one of the two dimers is bound to O3, the other dimer binds to O1. The intervening DNA forms a so-called repression loop. After Lewis et al., 1996.

The distance and nature of the bases between both the palindromic as well as the direct repeats of the recognition sequence plays an important role. It is evident that a dimeric protein would bind optimally to a 2-fold symmetric sequence only if the distance between the recognition elements matches the distance as determined by the protein structure. If one increases the distance by a few base pairs, a loss in cooperative binding capacity of the dimerization motifs of the protein to the rigid intervening DNA may result. The distance between the contacting sequence elements is particularly important for the DNA-binding elements of the nuclear receptors (see chapter 4). The DNA-binding element of the estrogen receptor differs from, e.g. that of the $T_3$-receptor, only with respect to the number of bases between the two half-sites of the recognition sequence. In this case the distance between the half sites decides which of the two receptors will bind and act as gene regulators.

A further aspect of the occurrence of multimeric recognition elements is the possibility for the formation of *heterodimers* (see 1.4.4.3). There exist related classes of DNA-binding proteins which recognize similar DNA-binding motifs and possess a common dimerization motif. Among these both homodimers as well as heterodimers can be formed which bind to DNA with slightly different specificities. The possibility for formation of heterodimers and homodimers of related DNA-binding proteins represents an important strategy to expand the specificity of the regulatory process. A notable example is the nuclear receptors (chapter 4).

# 1.3 The Principles of Transcription Regulation

## 1.3.1 General Mechanism

### 1.3.1.1 Elements of Transcription Regulation

Transcription represents the most important point of attack for the regulatory processes which control the flow of genetic information from DNA to mature protein. Primarily it is the initiation of transcription that is regulated, since this represents the rate-limiting step. The essential elements of such a regulation are (1) the *cis-acting DNA sequences*, which usually represent specific protein binding sites, and (2) *trans-acting DNA-binding proteins*, which specifically bind these DNA sequences to thereby influence the transcription process.

Cis-acting DNA elements can lie near the start site of transcription or be quite distanced from it. Furthermore, there are examples among eucaryotes in which the cis element is found within the transcribed region. If the cis element is located far from the site of action and its effect is also orientation-independent, then it is termed an *enhancer*. Furthermore, one frequently observes in eucaryotes so called composite control regions which contain various cis elements. In this case, several transcription factors act cooperatively in the initiation of transcription. Examples for such cooperative effects are observed among the genes controlled by nuclear receptors.

DNA-binding proteins can exercise a negative or positive influence on transcription upon binding to their cognate cis element (Fig. 1.20).

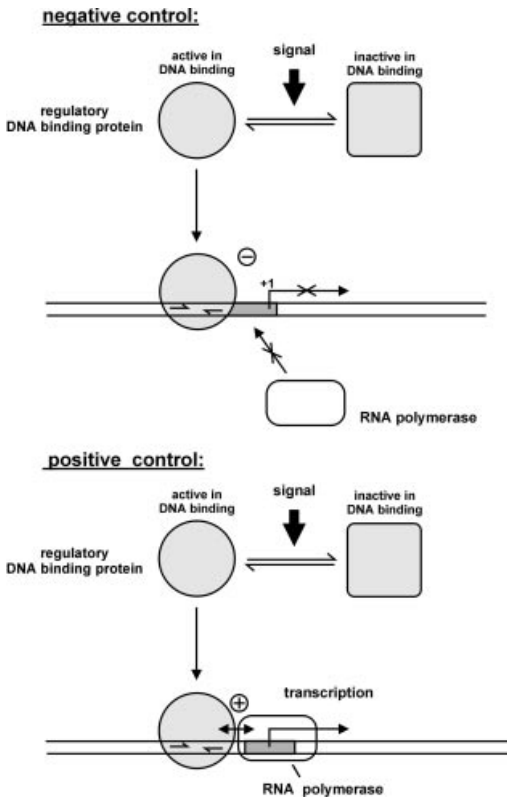



**Fig. 1.20.** Positive and negative control of gene activity. Negative control: the DNA-bound protein inhibits transcription. For bacterial repression, the binding sites for the RNA polymerase and the repressor partially overlap, and the bound repressor tends to inhibit the association of the RNA polymerase to the promoter. Positive control: the bound protein stimulates transcription.

#### 1.3.1.2 Negative Regulation of Transcription

Negative regulation of transcription implies that the binding of a regulatory protein leads to inhibition of transcription. Such proteins are described as transcriptional *repressors*. Negative regulation among procaryotes is often accomplished by the bound repressor blocking the access of the RNA polymerase to the promoter. This occurs if, for example, the binding sequence of the repressor and promoter sequence partially overlap. Bound repressor proteins can also cause a change in the conformation and topology of the DNA, which can indirectly inhibit transcription. Another mechanism for negative control involves binding of the regulatory protein to another protein whose function is essential for transcription; such binding then interferes with the function of the latter.

#### 1.3.1.3 Positive Regulation of Transcription

Positive regulation implies that the bound protein stimulates transcription. Such proteins are termed transcriptional activators. Transcriptional activation plays a central role in eukaryotes (see 1.4.3). There are various mechanisms of transcriptional activation.

Usually protein-protein interactions between the transcriptional activator and components of the transcription apparatus are involved.

#### 1.3.1.4 Functional Requirements for Repressors and Transcriptional activators

Regulatory DNA-binding proteins are multi-functional. Aside from their DNA-binding property, they also have the ability to register regulatory signals and transmit the signals on to the transcription apparatus. (Fig. 1.21).

**Specific DNA-binding**

Regulatory DNA-binding proteins generally display specific and selective DNA-binding capacity. In this way only those genes which possess a copy of a particular DNA-binding element are subjected to regulation by the corresponding binding protein.

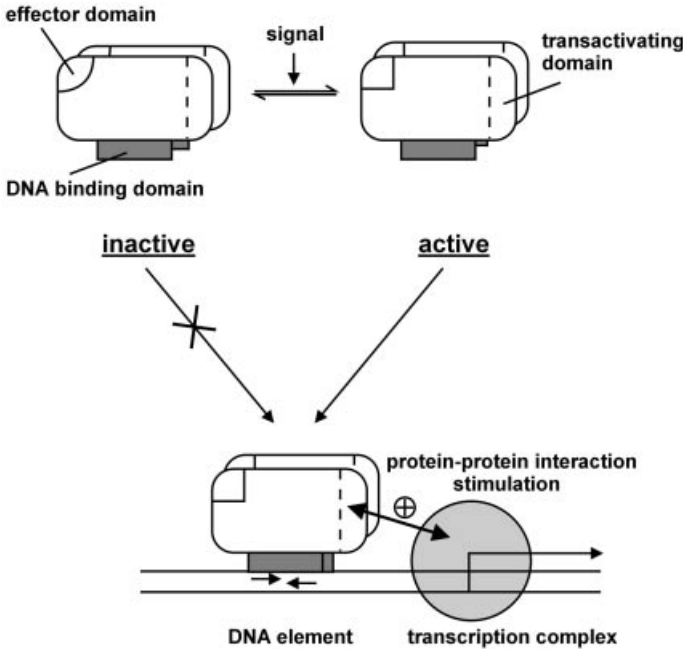



**Fig. 1.21.** Structural and functional principles of transcription activators. Typical transcription activators of eucaryotes possess a DNA-binding domain, an effector domain and a transactivating domain. An incoming signal is registered by the effector domain and transformed into a change in affinity for DNA. In the active state, the transcription activator is capable of binding to its cognate DNA-binding element. Protein-protein interactions with the transcription apparatus bound to the promoter mediate a stimulation of transcription initiation.

**Registering a Regulatory Signal: Activation and Inactivation**

A regulatory DNA-binding protein possesses structural elements for the registration of incoming signal, which leads to a change in concentration of the active binding protein. The activation (or inactivation) of the binding protein can be connected with a change in the ability to bind DNA, or can influence the capacity of the protein to interact with the transcription apparatus.

**Communication with the Transcription Apparatus**

The DNA-binding protein must be capable of transmitting signals to the transcription apparatus via protein-protein interactions. DNA-binding alone can be ascribed the function of increasing the effective concentration of the transcription regulator at the site of the transcription apparatus.

**Turning off the Transduction of Signal**

Regulatory signals should only be effective for a limited period of time and under certain external conditions. This also holds, of course, for regulation at the transcription level. It is therefore necessary to turn off the transduction of signal by the DNA-binding protein after the mediated demands have been fulfilled. Cells use common mechanisms for both the activation and inactivation of signal pathways. These are summarized below.

## 1.3.2 Mechanisms for the Control of the Activity of DNA-binding Proteins

The binding ability of regulatory DNA-binding proteins can be controlled by the following mechanisms:

### 1.3.2.1 Binding of Effector Molecules

Low molecular weight effectors are commonly emloyed in bacteria to change the DNA-binding activity of repressors or transcriptional activators and to control the amount of active DNA-binding proteins. This type of regulatory mechanism is frequently used for metabolic pathways, as in, for example, the biosynthesis and degradation of amino acids. The effector molecules represent components arising from the particular metabolic pathway. The goal of this regulation is to adjust the transcription rate to the current demand of the gene product.

The binding of small molecular weight effectors to regulatory DNA-binding protein can lead to an increase or decrease in the affinity of the protein for its recognition sequence.

Increase in the binding affinity upon binding of the effector:

| | |
|---|---|
| Trp repressor *(E. coli)* | Effector: tryptophan |
| CAP protein *(E. coli)* | Effector: cAMP |

Decrease in the binding affinity upon binding of the effector:

| | |
|---|---|
| Lac repressor *(E. coli)* | Effector: allolactose |

The strategies and mechanisms of effector molecules on regulatory DNA-binding proteins can be elucidated on the example of the Trp repressor of *E. coli*.

The Trp repressor controls the transcription of a total of five enzymes required for the biosynthesis of tryptophan (Fig. 1.22a). The genes for the five enzymes are encoded in a single operon, whereby the binding site for the Trp repressor overlaps with the promoter. The bound repressor blocks the RNA polymerase's access to the promoter, thereby inhibiting transcription.

The enzymes of Trp-biosynthesis are only required if too little tryptophan is available to the bacteria from the growth medium. In such a case the Trp requirement is fulfilled by the cell's own Trp biosynthesis. If however, there is enough Trp supplied by the medium, then it is prudent to shut down the Trp operon. The sensor is the Trp concentration. The Trp repressor registers the current Trp concentration with the help of its own Trp binding site. If a great deal of Trp is present, then the Trp binding site of the repressor is occupied by Trp. The Trp repressor binds Trp with high affinity ($K_D$=$10^{-9}$-$10^{-10}$ M), upon which transcription of the operon is then blocked.

At low Trp concentration the Trp repressor is mainly in the unbound, inactive form. The free form of the Trp repressor binds with a ca. $10^4$-fold lower affinity to the recognition sequence than the Trp-bound form. The promoter remains free under these conditions and transcription of the genes for Trp biosynthesis can occur. The shutting on and off of the Trp operon is based on the disparate DNA affinities of the free and Trp-bound repressor .

The Trp repressor is representative of many other DNA-binding proteins which occur in a binding and non-binding form as regulated by effector molecules. The binding of the effector molecule determines whether the protein is in either the active or the inactive form. Active and inactive forms differ by a factor of $10^4$-$10^5$ in their affinity for their cognate sequence. The affinity of the inactive form for the recognition sequence usually lies in the same range as its affinity for random, non-specific DNA sequence. The inactive binding protein is incapable of selectively binding the specific DNA element. The structural basis for the affinity differences are changes in the protein structure induced upon binding the effector molecule.

### Molecular Basis for the Control of Binding Activity of a Repressor by Effector Molecules

The comparison of the structure of a binding protein in the inactive form and in the active form bound to DNA gives an impression of the conformational changes correlated with binding of effector molecules. The Trp repressor is, next to the Lac repressor from E. coli, one of the few examples in which the structural basis for the difference in DNA-binding affinity of the inactive vs. active form is understood (Fig. 1.22b).

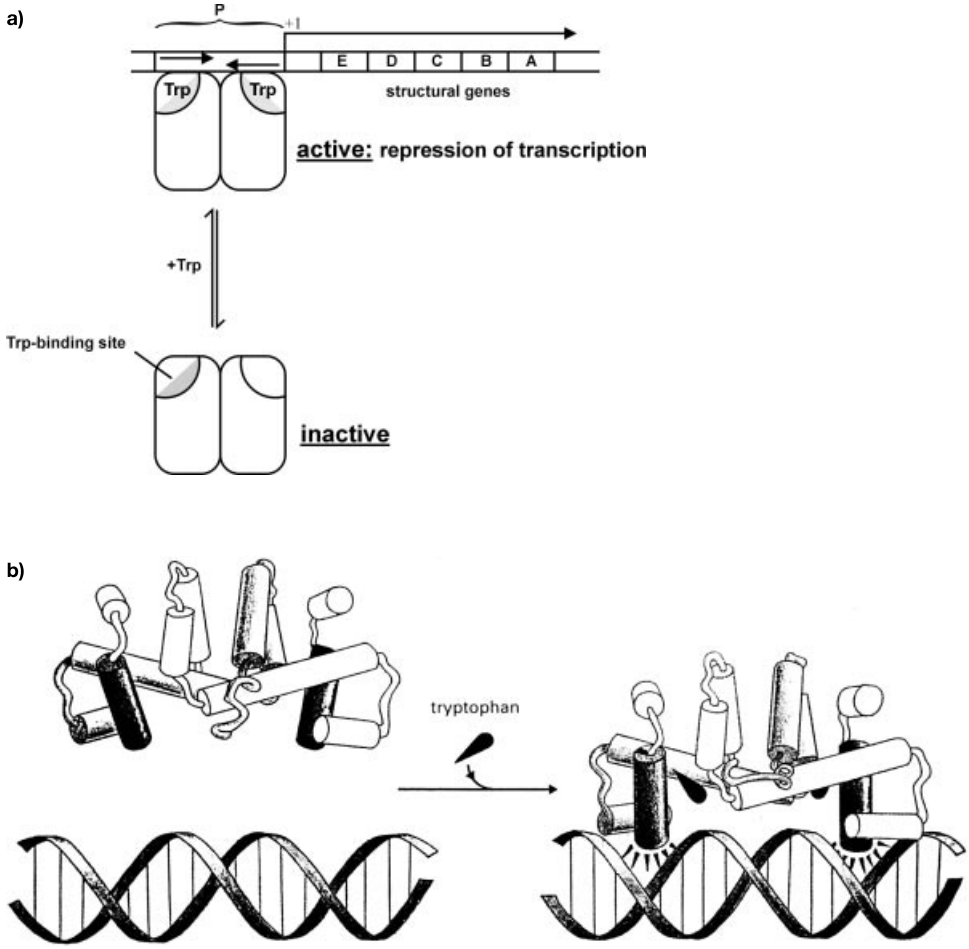



**Fig. 1.22.** Regulation of the Trp operon in E. coli. A) The Trp repressor requires Trp in order to bind its affiliated DNA-binding element. In the absence of tryptophan, the Trp repressor can not bind to the regulatory sequence and is therefor inactive. Upon an increase in the tryptophan concentration, tryptophan binds to the Trp repressor and transforms it into a binding-proficient form. The DNA bound Trp repressor prevents the transcription of the structural genes, and the biosynthesis of tryptophan is halted. B) structural basis for the activation of the DNA-binding of the Trp repressor by tryptohan molecules. The dimeric Trp-repressor recognizes the two-fold symmetric recognitions sequence with the help of a helix-turn-helix motif. In the absence of tryptophan the spacing between the two recognition helices is too small to allow entry into and binding of the major groove. Upon binding tryptophan, the flexible recognition helices are pushed apart and oriented for optimal contact to the half sites of the DNA elements. Reproduction with permission from Alberts et al., 1994, p. 418.

In the Trp-bound, binding-competent form, the helix-turn-helix motif of the repressor is found in a position favorable for contacting the recognition sequence, and the recognition helix can interact with the major groove of the DNA. The effector molecule tryptophan binds near the helix-turn-helix motif and performs several tasks: first, it orients and fixes the recognition helix in such a way that the specific interactions with the DNA recognition element can be formed. Furthermore, the bound tryptophan is indirectly involved in interactions with the DNA, in that it supports the formation of H-bonds to the DNA by certain amino acid residues. In the Trp-free form the prerequisite for a strong cooperative binding of the repressor dimer are not fulfilled since the recognition helices are not postioned optimally for binding to the recognition sequence in the major groove of the DNA.

#### 1.3.2.2 Metal Ions as Effector Molecules

Metal ions can serve as effector molecules as well as control the DNA-binding activity of regulatory proteins. An example is the regulation of the metallothionein gene in eucaryotes (Fig. 1.23). The metallothioneins are small, cysteine rich proteins which can specifically bind metal ions like $Cu^{+}$ or $Zn^{2+}$. The complexation of metal ions functions to sequester the ions in a form that is not damaging to the cell.

The gene for metallothionein is induced by metal ions. The goal of the regulation is to provide enough metallothionein for the complexation of metal ions and thus to maintain the concentration of free metal ions at a tolerable level for the cell.

The transcription of genes for metallothionein is under positive regulation by the DNA-binding protein ACE, which binds an ACE-specific DNA element in the metal-

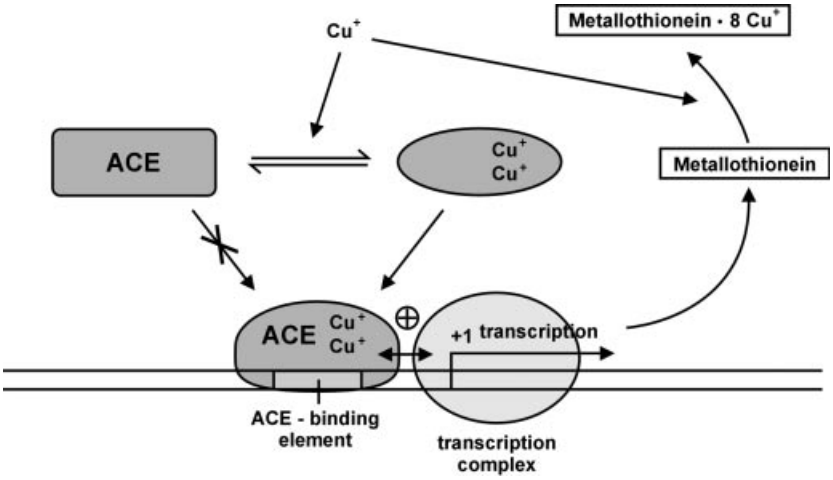



**Fig. 1.23.** Regulation of the transcription of metallothionein by $Cu^{+}$. The transcription of metallothionein is controlled by the transcription activator ACE, whose specific DNA-binding capacity is regulated by $Cu^{+}$which functions as an effector of the ACE protein by binding to the protein and regulating its DNA-binding ability.

lothionein promoter. The DNA-binding activity of the ACE protein is furthermore controlled by the level of $Cu^+$. The ACE protein requires $Cu^+$ in order to be able to specifically bind its recognition element. It possesses a total of eight $Cu^+$ binding sites and its active conformation is induced upon binding of $Cu^+$. In the $Cu^+$-bound form the ACE protein acts as a transcriptional activator. In the absence of $Cu^+$ a specific binding of the ACE recognition sequence is not possible, and the expression of the metallothionein gene is inhibited.
A recent example for the regulation of DNA-binding proteins by metal ions is the transcriptional repressor DREAM that binds to the cognate DNA element only in the absence of $Ca^{2+}$ (Carrion et al., 1999). An increase of $Ca^{2+}$ in the form of a $Ca^{2+}$ signal (see Chapter 6) leads to reduced affinity of the repressor for its DNA element and to an increased expression of the target gene.

#### 1.3.2.3 Binding of Inhibitory Proteins

Specific DNA-binding proteins can be constrained in their ability to function as gene regulators by complex formation with inhibitor proteins. Examples are the steroid hormone receptors which, in the cytosol, are bound in their inactive form to the proteins hsp90, hsp56 and p23 (see chapter 4). Also the GAL4 transcriptional activator of yeast occurs in the cytosol as a complex with the inhibitory protein GAL80, and in this form is regulatory inactive (Fig. 1.24).

In response to an incoming signal, e.g. a low molecular weight effector molecule, the DNA-binding protein is released from the inactive complex. The steroid hormone receptors are a noteworthy example, which are found in the cytosol in inactive form bound to an inhibitory protein. The concentration of steroid hormone serves as a signal. The binding of steroid hormones to the receptor enables the dissociation of the inhibitory protein and the subsequent transport into the nucleus where the receptor can function as a gene regulator (see 4.4.1).

Protein phosphorylation also serves as a tool to release the DNA-binding protein from an inhibitory complex in the cytosol. The free DNA-binding protein can then, for example, be transported into the nucleus where it is then available for the regulation of gene expression. In this manner the activity of the transcription factor NFκB is regulated (Thanos and Maniatis, 1995; see 2.7).

#### 1.3.2.4 Modification of Regulatory Proteins

Post-translational covalent modification of DNA-binding proteins is a mechanism commonly employed among eucaryotes to control the activity of DNA-binding proteins.

Of particular importance is the *phosphorylation* of eucaryotic transcription factors. Functional and mechanistic consequences of the phosphorylation of transcription factors will be discussed in more detail in the section on the regulation of eucaryotic transcription (see 1.4.3.2). Specific or non-specific protein phosphatases (see 7.5) can remove the phosphate residues and terminate the phosphorylation signal.

A singular example of how covalent modification can be used to achieve transcription control is found in the regulation of the adaptive repair of DNA damage in proca-

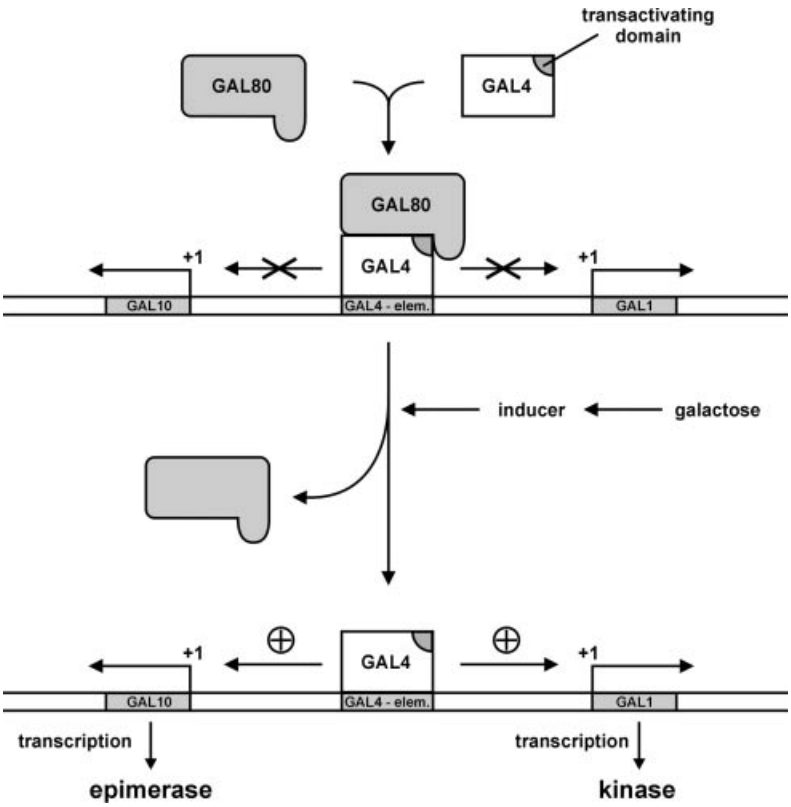



**Fig. 1.24.** Regulation of trancription of the enzymes of galactose metabolism by the inhibitory protein GAL80 and the transcription activator GAL4. The trancription activator GAL4 controls the transcription of the gene for the protein GAL1 (galactose-1-kinase) and GAL10 (galactose epimerase). In the absence of galactose, GAL4 is inactive due to complexation with the inhibitor protein GAL80. In this case, Gal4 can bind the associated GAL4 element, but no transactivation occurs. If galactose is available, then a currently unidentified product of the galactose metabolic pathway leads to dissociation of GAL80 from the inhibitory complex and the inhibition of transcription is uplifted.

ryotes. Here, the DNA-binding activity of a transcriptional activator is controlled by its *methylation* (Fig. 1.25).

In adaptive repair, primarily alkylation damage of DNA is repaired. Alkylating agents such as N-methyl-N-nitroso-urea methylate DNA, whereby various O- or N-methylated base adducts are produced. A methlyation of the phospate residue is also possible. The base adducts produced, for example $O^6$-methly-dG, can lead to mispairing during DNA replication and are potentially mutagenic. The cell repairs such DNA damage with the help of S-alkyl transferases. The S-alkyl transferases accept the methyl group from the DNA and are thereby methylated themselves. The methylation occurs on an essential Cys-SH of the S-alkyl transferase, and irreveribly inactivates the enzyme. Thus, S-alkyl transferase is used in an amount stochiometric to the damaged DNA and commits „suicide“ upon repair (suicide enzyme).

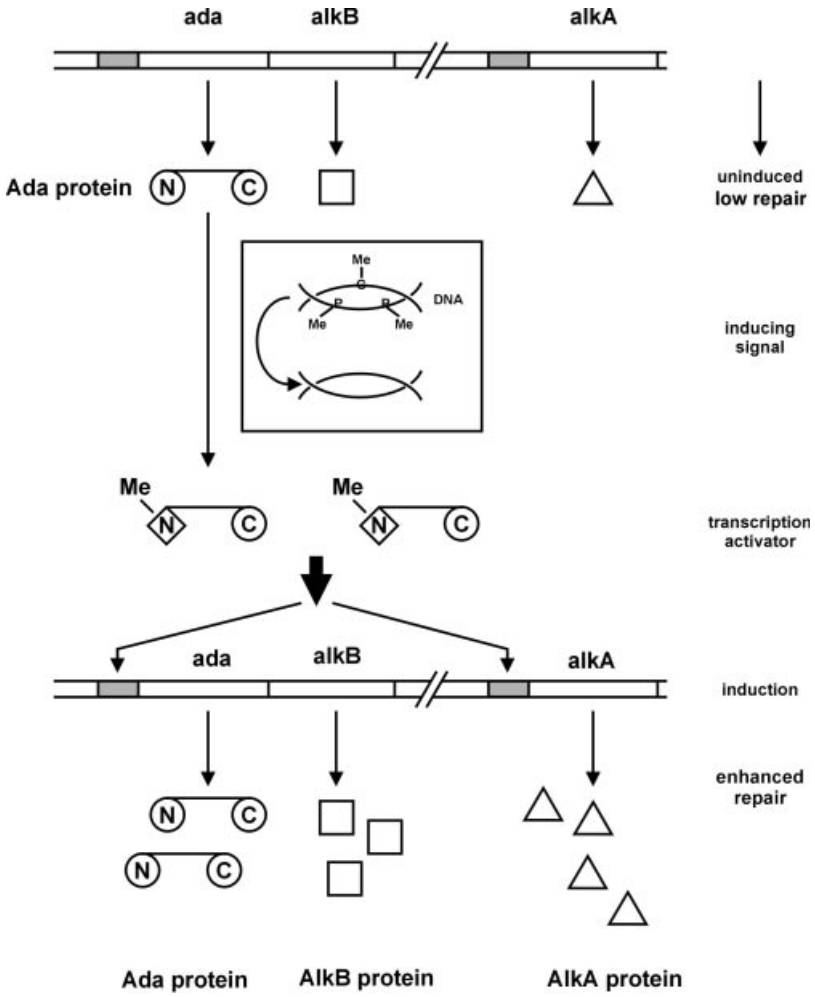



**Fig. 1.25.** Regulation of alkylation repair in E. coli by methylation of the Ada protein. The effect of methylating agents, such as N-nitroso-N-methyl urea lead to the formation of methyl phosphotriesters (P-Me) of DNA, as well as various base adducts. The Ada protein possesses an N-terminal and a C-terminal domain. In one of the first steps of alkylation repair the methyl groups of the phosphotriester is transferred to the Ada protein. The Ada protein is methylated on a Cys residue at its N-terminal domain and thereby transformed into an active transcription activator. In its methylated form the Ada protein binds to the control region of various genes to stimulate their transcription. Among the genes under the control of the Ada protein are its own gene, as well others required for DNA repair (alkB, alkA). After Lindahl et al., 1988.

In *E. coli*, repair is induced by alkylation damage. The repair of methylated DNA is performed by, among others, the Ada protein which possesses S-alkyl transferase and transcription activation activity. The Ada protein is methylated during the first repair process. In the methylated form the Ada protein can function as a transcriptional activator. It binds to the corresponding DNA element of an operon which encodes for

Ada and other proteins. Upon the binding of the methylated Ada protein to its cognate DNA element, the transcription of the genes is stimulated such that more repair proteins are available.

Further covalent modifications of regulatory DNA-binding proteins include acetylation (e.g. of Lys residues in the transcription factor GATA-1 (Boyes et al., 1998)) and ADP-ribosylation (see 5.5.2), to name a few.

#### 1.3.2.5 Changes in the Concentration of Regulatory DNA-binding Proteins

The amount of available DNA-binding proteins is, in many situations, a critical factor for the extent of transcription regulation. The concentration of regulatory DNA-binding proteins can be regulated within the framework of the following processes in eucaryotes:

- transcription
- splicing, transport
- translation
- compartmentalization
- protein stability

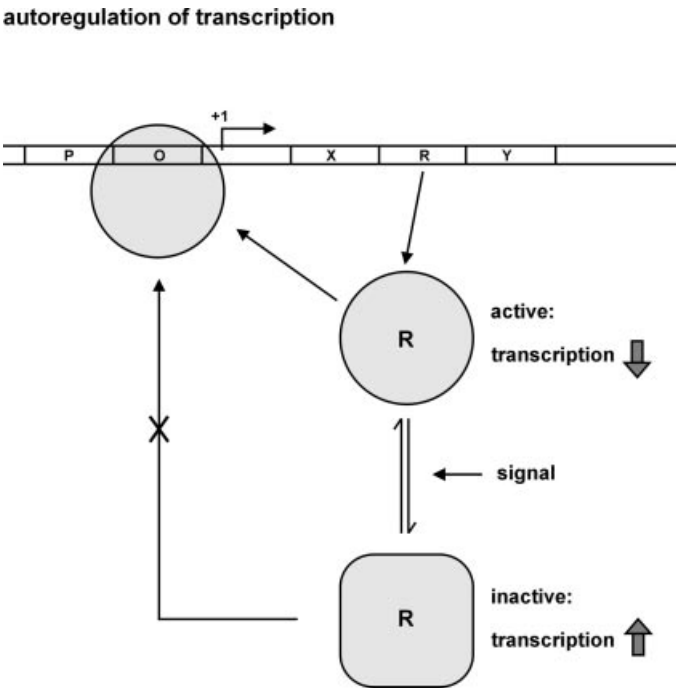



**Fig. 1.26.** Principles of autoregulation of transcription. In autoregulation, a repressor controls the transcription of its own gene as well as the transcription of other genes (X,Y).

The above points will be discussed in more detail in the following section (1.4) in the context of eucaryotic gene regulation.

Only *autoregulation* will be introduced as an example for the regulation of DNA-binding proteins at the level of transcription.

Autoregulation implies that a repressor regulates the transcription of its own gene (Fig. 1.26). In the operator region for the genes of the repressor is found a binding site for the repressor itself so that it can function as its own negative regulator. If little repressor is available, then the associated DNA element remains unoccupied and the transcription of the repressor gene is no longer blocked. Increasing concentration of the repressor leads to increased occupation of the repressor binding site and to an inhibition of the transcription of the repressor. Usually, binding sites also exist for the repressor in other operons. The extent of occupation of the various operons is determined by the affinity of the repressor to the various operator binding sites.

An example for autoregulation is found in the hut-operon of E. coli and in the regulation of the SOS response in bacteria via the lexA repressor. There are examples of autoregulation at the level of translation as well (see 1.5.6).

With the aid of autoregulatory processes it is possible for the cell to maintain a minimal concentration of repressor.

## 1.4 Regulation of Transcription

Procaryotes and eucaryotes differ decisively in the structure of the transcription start site and the complexity of the transcription appartus. For a better understanding we want to briefly summarize procaryotic transcription and then contrast it to eucaryotic transcription (review: Eick and Heumann, 1994).

### 1.4.1 Overview of Transcription Initiation in Procaryotes

Transcription initiation in procaryotes is controlled via promoters and regulatory DNA sequences located near the promoter. The role of the promoter is to provide a defined association site for the RNA polymerase and to correctly orient it. The binding of the RNA polymerase to its promoter is controlled by the *sigma* factor, a component of the RNA polymerase holoenzyme. The sigma factor selects which genes are to be transcribed by specifically recognizing the promoter sequence and structure and by allowing the RNA polymerase to form a *transcription-competent complex* at the transcription start site.

#### Mechanism of Promoter Recognition

A transcription-competent complex must be present at the initiation site, with partial melting of the DNA, for the RNA polymerase to be able to add ribonucleotides complementary to the DNA template.

The formation of a transcription-competent complex can be described according to a two step mechanism (fig 1.27). The initial binding of the RNA polymerase to the pro-

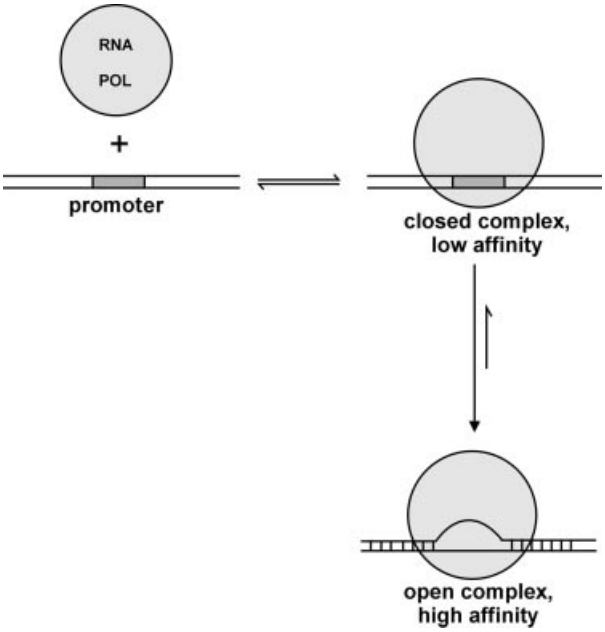



**Fig. 1.27.** Two-step mechanism of transcription initiation. The binding of a procaryotic RNA polymerase to its promoter can be subdivided into two steps. In the first step the RNA polymerase binds to the closed promotor with low affinity. The closed complex isomerizes in a second step to an open complex in which the promoter is partially unwound. Detailed consideration reveals that further steps can be distinguished. These are not shown here for simplicity reasons.

motor leads to the formation of a *closed complex* in which the RNA polymerase is only weakly bound. Isomerization of the closed complex transforms it into a transcription-competent open state. In the *open complex* the RNA polymerase is tightly bound and the DNA is partially unwound at the transcription start site.

The RNA polymerase of E. coli possesses with its subunit construction ($\alpha_2\beta\beta'\sigma$) a simple structure in comparison to eucaryotic RNA polymerases. The sigma factor is only required for the recognition of the promoter and the subsequent formation of a tight complex. After the incorporation of the first 8–10 nucleotides into the transcript, the sigma factor dissociates from the holoenzyme, and the remaining core enzyme carries out the rest of the elongation.

There are several sigma factors in E. coli ($\sigma^{70}$, $\sigma^{54}$, $\sigma^{32}$, $\sigma^{28}$) which can associate with the core enzyme to form the holoenzyme. The various sigma factors differ significantly with respect to their cognate promoter sequences. The overwhelming majority of the promoters in E. coli are recognized and activated by $\sigma^{70}$-containing holoenzymes.

A comparison of the activation of $\sigma^{70}$ and $\sigma^{54}$-dependent promoters helps us understand some of the basic points of transcription activation, which also play an important role in eucaryotic transcription.

#### 1.4.1.2 $\sigma^{70}$-Dependent Transcription

Key elements of $\sigma^{70}$-dependent promoters are the TATA box with the consensus sequence TATAAT 10 bp upstream from the transcription initiation site (pos. -10), and the sequence TTGACA at the position -35 (Fig. 1.28). Both sequences are necessary

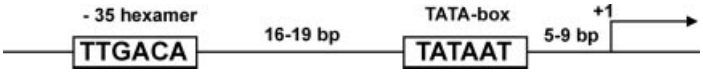



**Fig. 1.28.** Promoter structure of $\sigma^{70}$-dependent genes in *E. coli.*

for the recognition of the promoter by $\sigma^{70}$. The intervening sequences, as well as other upstream sequences, can also influence the efficiency of transcription initiation. It is not possible to define consensus sequences at these positions. An optimal $\sigma^{70}$-dependent promoter can be defined as a sequence with the -35 hexamer as well as the -10 hexamer 17 bp away. The latter lies 7 bp upstream from the transcription initiation site.

The more than 1000 currently known promoters in E. coli display a significant graduation in efficiency of transcription initiation. There are strong and weak promoters depending upon their deviation from the optimal promoter. In bacteria part of the variability in the efficiency of transcription initiation can be explained at the level of the different promoter sequences. Some promoters do not possess any of the above mentioned structural characteristics.

An important aspect of $\sigma^{70}$-dependent promoters is the fact that the holoenzyme is capable of initiating a weak transcription even without accessory proteins. In this way, the $\sigma^{70}$ containing holoenzyme can independently carry out all necessary steps, e.g. melting of the DNA, so that a constitutive transcription without the participation of regulatory proteins is possible. In this case, the extent of transcription depends on the affinity of the holoenzyme for the promoter and, thus, depends indirectly on the promoter sequence.

The recognition sequences of regulatory proteins may overlap not only the promotor site, but can also be found in the immediate vicinity of the $\sigma^{70}$ promoter. The sequence elements are relatively simple and often include only one binding site for regulatory proteins.

Transcriptional activity is controlled mainly according to two mechanisms:

*Repressors* turn off transcription by, for example, competing with the holoenzyme for binding to the promoter. RNA polymerases and repressors compete if the repressor binding site overlaps with the promoter sequence (see 1.3).

*Transcriptional* activators turn a gene on by increasing the efficiency of transcription initiation above the basal level. Transcription activation plays a particularly important role for the promoters classified as weak based on their sequence. It is assumed that the transcription activation occurs via protein-protein interactions between the DNA-bound transcriptional activator and the holoenzyme. This form of activation demands a close and defined contact. A change in the distance between the transactivator binding site and the promoter may lead to a loss in the stimulatory effect if direct communication between the two proteins is no longer possible. On the other hand, a shift in the binding site of the transcriptional activators upstream in the direction of the promoter can lead to inhibition of transcription. Characteristic for the regulation of $\sigma^{70}$-dependent promoters by DNA-binding proteins is the tightly defined region within which the regulatory protein must bind relative to the binding site of the RNA polymerase holoenzyme.

**Mechanism of Transcription Activation**

Structural data on the interaction of transcriptional activators with the RNA polymerase holoenzyme are not yet available. We therefore rely on models to explain the mechanism of transcriptional activation. A plausible and experimentally supported model of transcriptional activation for $\sigma^{70}$ promoters assumes that the transcriptional activators recruit the RNA polymerase holoenzyme to the promoter. According to this model, the DNA-bound transcriptional activator interacts with the holoenyzme, itself either free or DNA-bound, thus enabling the RNA polymerase to form a tight complex with the promoter. The recruitment increases the life span of the holoenyzme-DNA complex and/or eases the transition to the open, high affinity complex, from which transcription initiation occurs (review: Ptashne and Gann, 1997). Without transcriptional activators present, the holoenzyme binds only weakly to the promoter and the transition to the open complex occurs only at a low frequency. Components of the holoenzyme that may be involved include the α-subunit or the σ-factor. Multiple, synergistically acting contacts can be formed between the transcriptional activator and the holoenzyme. This concept does not predict a conformational change in the holoenzyme, but does assume that the interactions between the holoenzyme and transcriptional activator are rather non-specific in nature. As already proven experimentally, heterologous binding surfaces can also carry out this function.

### 1.4.1.3 $\sigma^{54}$-dependent Promoters

This type of promoter displays markedly different characteristics compared to the $\sigma^{70}$-dependent promoter. The $\sigma^{54}$-containing holoenzyme binds tightly to the promoter in the absence of transcriptional activators. In this closed state, however, it is not capable of initiating transcription. The transcriptional activators are required in this case to activate the promoter-bound holoenzyme for initiation, i.e. to transform it into the open complex (see Fig. 1.29). Activation is mediated via protein-protein interactions between the transcriptional activator and the RNA polymerase holoenzyme, and is accompanied by ATP hydrolysis. The binding site for the transcriptional activator is found at a distance of ca.110 bp upstream form the start site and can be shifted further upstream without loss of stimulatory effect. Direct interaction of the holoenzyme with the bound transcriptional activator is possible due to *loop formation* of the intervening DNA. The strict dependency on transcriptional activators for transcription initiation indicates that the DNA-bound holoenzyme alone is not capable of isomerizing to the transcription-competent open complex. The transition to the open complex requires interactions with the transcriptional activator, an event which occurs with ATP hydrolysis.

A detailed, structural based description of the mechanism of transcription stimulation in procaryotes is currently not possible since such data is not yet available.

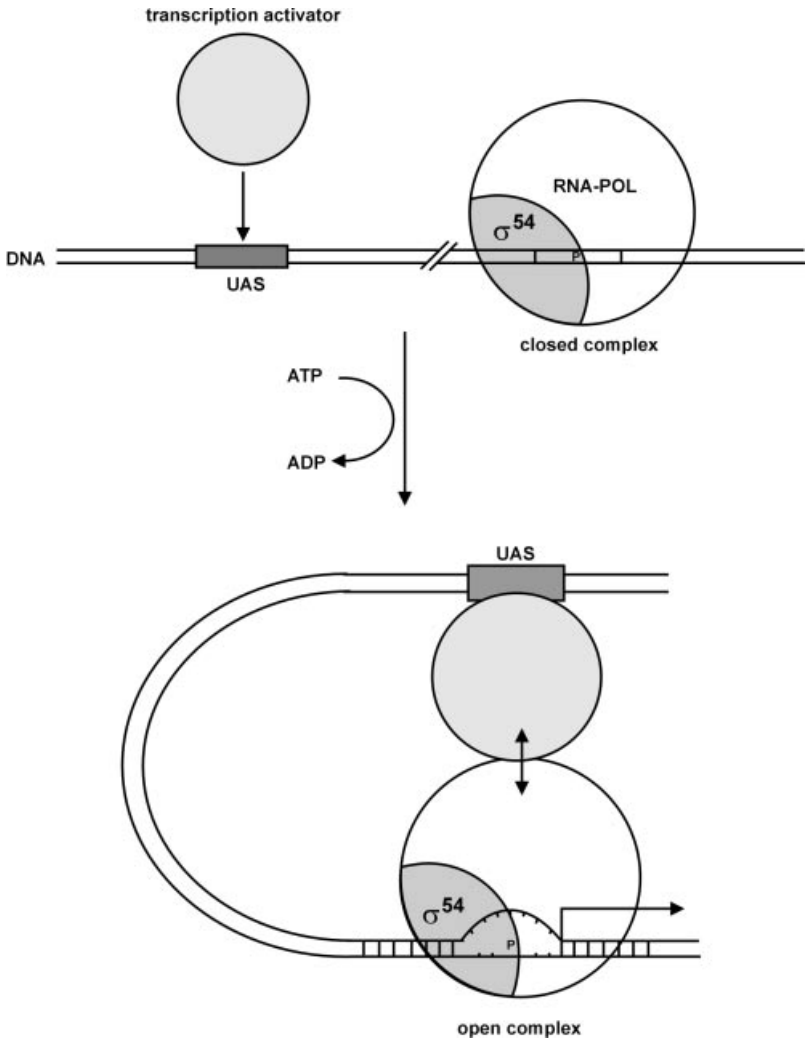



**Fig. 1.29.** Mechanism of promoter activation of $\sigma^{54}$-dependent genes in procaryotes. The formation of an open, initiation-competent transcription complex for $\sigma^{54}$-dependent genes requires the assistance of transcription activators, which bind to their cognate UAS element. Upon loop formation of the intervening DNA sequences, the transcription activator interacts with the $\sigma^{54}$-containg RNA polymerase bound to the promoter. The activation is accompanied by ATP hydrolysis and leads to the formation of an open complex.

## 1.4.2 Structure of the Eucaryotic Transcription Apparatus

Three types of RNA polymerases exist for the transcription of eucaryotic genes, each of which transcribes a certain class of genes. All three enzymes are characterized by a complex subunit structure.

RNA polymerase I is responsible for the transcription of the ribosomal RNA genes (class I genes), RNA polymerase II transcribes the genes encoding proteins (class II genes), and RNA polymerase III transcribes the genes for the tRNAs and the 5S ribosomal RNA (class III genes). Below we will limit the discussion to RNA polymerase II and the genes transcribed by it since it plays the most important role for regulatory processes and signal transduction. Aside from this, many characteristics of the transcription of the genes of class II are also valid for genes of the class I and III.

### 1.4.2.1 Structure of the Transcription Start Site and Regulatory Sequences

The promoter activity of class II genes is guided primarily by three structural elements (Fig. 1.30):

The TATA box and/or an initiation sequence are structural elements which define a minimal promoter from which in vitro transcription can be initiated. A classical TATA box is often, though not always, ca. 30bp from the transcription start site. The initiation sequence includes sequences in the immediate vicinity of the transcription start site. The TATA box and initiation sequence are sufficient for the formation of a basal transcription apparatus composed of general initiation factors for transcription and RNA polymerase II (see Fig. 1.31).

Not every eucaryotic promoters possesses a TATA box. For promoters devoid of a TATA box, the initiation sequence is determining for promoter selection and formation of the pre-initiation complex.

An increase in the basal transcription activity originating from the promoter requires regulatory DNA sequences. These sequences can be proximally or distally located and serve as binding sites for transcriptional activators. The *cis-acting DNA elements* can be found near the promoter in either orientation. They are often termed „*upstream activating sequences*“ (UAS). Regulatory sequences can also be located far from the promoter. Their effect is independent of their orientation and they are known as *enhancers.*

A combination of several cis-elements, and thus several transcriptional activators, are often involved in the regulation of eucaryotic transcription. Transcription activation, in these cases, results from the complex concerted action of various specific DNA-binding proteins.


**Fig. 1.30.** Structure of a typical eucaryotic transcription start site. Enhancer elements and UAS elements (UAS: upstream activating sequences) are binding sites for positive and negative regulatory DNA-binding proteins. The TATA box is the binding site for the TATA box binding protein (TBP) and serves to position the RNA polymerase holoenzyme on the promoter. For promoters that do not possess a TATA box, this function is fulfilled by an initiator region.

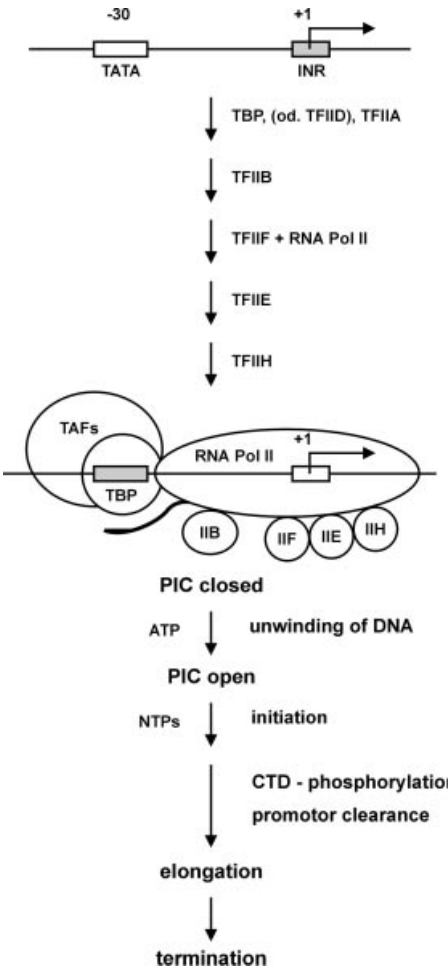



**Fig. 1.31.** Model for the formation of a pre-initiation complex and further steps for transcription initiation in eucaryotes. A functional pre-initiation complex (PIC) can be reconstructed *in vitro* from purified general transcription factors and RNA polymerase II. Successful reconstruction requires the addition of general transcription factors in the indicated order, whereby each of the given steps is supported by the experimentally confirmed existence of the implied intermediates. In the case of a TATA box-containing promoter, the first step in the reconstruction consists of the binding of the TATA box binding protein (TBP). It is assumed that under more physiological conditions the TFIID complex (together with the associated TBP) binds to the promoter in the first step. TFIIA can join the pre-initiation complex after the binding of TBP at any of the following steps, but is not essential for the *in vitro* formation of a functional pre-initiation complex. The pre-initiation complex is stabilized and favored by interactions of the general transcription factors with the TAFs (TAF: TATA box binding protein associated factors). After the formation of the closed pre-initiation complex and under ATP requirement, melting of the DNA and transition to the open complex occurs, in which the first nucleotide is added in the actual initiation step. The subsequent phosphorylation of the carboxyl terminal domain (CTD) of the large subunit of RNA polymerase II causes the dissociation of the RNA polymerase from the promoter and transition to the elongation phase. After termination the RNA polymerase is dephosphorylated and is ready for the next round of transcription (not shown in the above figure). INR: initiation region.

### 1.4.2.2 Elementary Steps of Eucaryotic Transcription

As in procaryotes, the elementary steps of initiation, elongation and termination can be distinguished in eucaryotic transcription. Aside from the specific RNA polymerases, transcription in eucaryotes requires the action of numerous other proteins which are collectively known as transcription factors. Transcription factors are required at the level of initiation, elongation, and termination and are accordingly known as initiation factors, elongation factors and termination factors of transcription.

Transcription in eucaryotes can, as shown schematically in Fig. 1.31, be subdivided in the following steps (Review: Roeder, 1996, Nikolov and Burley, 1997):

1) Formation of a Pre-Initiation Complex:
– Promoter selection, binding of basal initiation factors
– binding of the RNA polymerase II
– formation of a basal transcription apparatus

2) Activation of the Pre-Initiation Complex:
– melting of the DNA in the vicinty of the start site

3) Initiation:
– incorporation of the first nucleotide

4) Transition from Initiation to Elongation:
– processive RNA synthesis by RNA polymerase

6) Termination:
– End of synthesis at defined sequence elements

Transcription is regulated to a great deal at the start of transcription, i.e. steps (1–4). These steps are thus the center point of the following discussion. One must keep in mind that *in vivo* transcription does not occur on naked DNA, but rather on chromatin, i.e. nucleosome-coated DNA. Activation of transcription requires an active remodeling of the chromatin structure at the transcription start sites and within the transcribed regions. Every discussion of the individual steps of transcription must consider this fact. However, only incomplete in vitro systems are available for the study of transcription of chromatin-associated DNA, so that in the following discussion only the data on the transcription of naked DNA is presented.

#### 1.4.2.3 Formation of a Basal Transcription Apparatus from General Initiation Factors and RNA Polymerase

In contrast to the procaryotes, where the $\sigma^{70}$-holoenzyme of the RNA polymerase can initiate transcription without the aid of accessory factors, the eucaryotic RNA polymerase requires the help of numerous proteins to begin transcription. These proteins are termed *basal* or *general initiation factors of transcription*. Together with RNA polymerase II, they participate in the basal transcription apparatus. The various components must associate in a defined order for the formation of a transcription-competent complex, from which a low level of transcription is possible. An increase in the basal transcriptional level requires the effect of specific transcriptional activators, which bind cognate DNA sequences at a variable distance from the promoter. The transcriptional activators themselves require the aid of further protein factors, known as coactivators (see 1.4.3.2), in order to attain full stimulatory activity.

The purification and structural and functional characterization of the general initiation factors has proven extremely difficult. The specific function of the various factors, as well as their structural role in the entire complex, remains poorly resolved. According to the current model, the general transcription initiation factors, with which an exact start of transcription is possible in vitro, are required for the formation of a basal

transcription apparatus (summarized in table 1.1; review: Koleske and Young, 1995; Roeder, 1996).

A transcription-competent pre-initiation complex consisting of general transcription initiation factors and RNA polymerase II, can be reconstituted in the test tube from the individual components. As outlined in Fig. 1.31, efficient reconstitution requires a defined order for the addition of the individual components.

The following points are noteworthy:

The transcription factor TFIID is a multi-protein complex that binds specifically to the promoter region. It consists of the TATA-Box binding protein (TBP) and TATA-Box binding protein associated factors (TAFs). It is assumed that the binding of TFIID to the TATA box represents an important regulatory step in the recognition and selection of the promoter in vivo. TBP is the only sequence specific DNA-binding protein of the basal transcription apparatus and its binding to the TATA-Box leads to a distinct bending of the DNA (see Fig. 1.16). In this manner a particular topology of the DNA is created that serves as a prerequisite for the defined binding of further basal transcription factors, such as TFIIA and TFIIB.

**Table 1.1.** General initiation factors of transcription by RNA polymerase II
After Roeder (1996). TAF: TATA box binding protein associated factor; TBP: TATA box binding protein; RNA Pol II: RNA polymerase II

| protein | number of subunits | subunit size (kDa) | function |
|---|---|---|---|
| TFIID: | | | |
| TBP | 1 | 38 | sequence specific binding to TATA box, recruitment of TFIIB |
| TAFs | 12 | 15–250 | promoter recognition, regulation, chromatin modification |
| TFIIA | 3 | 12,19,35 | stabilization of TBP-DNA binding; antirepression |
| TFIIB | 1 | 35 | recruitment of RNA Pol II – TFIIF; selection of start site by RNA Pol II |
| TFIIF | 2 | 30,74 | assists in promoter binding by RNA Pol II |
| RNA polymerase II | 12 | 10–220 | enzymatic activity of RNA synthesis, binding of TFIIF |
| TFIIE | 2 | 34,57 | binding of TFIIH, modulation of activities of TFIIH |
| TFIIH | 9 | 35–89 | helicase, protein kinase and ATPase-activity; promoter unwinding, promoter clearance (?) |

The TAFs fulfill numerous functions (Review: Burley and Roeder, 1996; Struhl and Moqtaderi, 1998) . On the one hand they are ascribed a structure promoting function. Some of the TAFs display a high degree of homology to the histones H2A, H3 and H4, and it is speculated that they help to create a nucleosome-like structure at the promotor. Furthermore, the TAFs are targets for protein-protein interactions with transcriptional activators. TAFs also posses enzymatic activity. TAFII250 has both a histone acetylase activity and a protein kinase activity. While the former presumable plays a role in the reorganization of the nucleosome, the latter can lead to phosphorylation of TFIIF.

There are also indications that the composition of TFIID is not fixed, but may vary depending on the detailed structure of the promoter. This idea is corroborated by the isolation of a transcription-competent TFIID that supports transcription without requiring TBP binding (Apone & Green, 1998).

TFIIA and TFIIB support TFIID in the formation of a stable complex with the promotor. TFIIB is necessary for the downstream selection of the start site for RNA polymerase II. Interactions with TFIIB ensure correct positioning of the RNA polymerase II on the promoter. Crystal structures have been solved for several of the intermediates of the pre-initiation complex (review: Sokolev and Burley, 1997), showing, for example, that TBP affects a predominant kink in the DNA (see Fig. 1.16). TFIIB binds to the TBP-DNA complex, contacting both TBP and the DNA.

TFIIF is found in a pre-formed complex with RNA polymerase II and suppresses the non-specific binding of RNA polymerase to DNA. TFIIF supports the association of RNA polymerase with the promoter bound complex of TFIIA, TFIIB and TFIID.

TFIIE binds TFIIH to assist the latter with the melting of the promoter.

The binding of TFIIH completes the formation of the pre-initiation complex. TFIIH is a multi-protein complex with a variable composition (see 1.4.2.5) and which possesses protein kinase, ATPase and helicase activities. The helicase activity of TFIIH is required for the melting of the promoter.

The general transcription initiation factors can be assigned the role fulfilled by a single protein in procaryotes-namely the σ-factor. This role includes the correct positioning of the RNA polymerase on the promoter and the preparation for the incorporation of the first nucleotide.

The addition of ATP to the pre-initiation complex leads to a rapid melting of the promoter, initiation of RNA synthesis and dissociation of the RNA polymerase from the promoter.

**Holoenzyme Forms of RNA Polymerase II**

RNA polymerase II represents a multi-enzyme complex of at least 12 proteins, but whose exact composition is difficult to determine. This is due to the instability of the holo-complex, which makes the purification and characterization of the enzyme difficult. Furthermore, it is likely that multiple forms of RNA polymerase II exist, each of slightly different composition and performing different functions.

Studies in yeast have shown that *pre-formed RNA polymerase II holoenzyme complexes* can be found in the cell associated with some of the general transcription factors.

General transcription initiation factors TFIIB, TFIIE, TFIIF and TFIIH have been identified as components of the RNA polymerase II holoenzyme of yeast. Various forms of the yeast holoenzyme contain further proteins, known as *mediators* or *SRB proteins* (SRB, suppressor of RNA polymerase B). The mediators function as coactivators (see 1.4.3.2). The holoenzyme is difficult to define structurally because the proteins accessory to the core enzyme (see table 1) may not be permanently associated with RNA polymerase II.

The use of a pre-formed holoenzyme complex of the RNA polymerase II appears to be an economical mechanism for the formation of an initiation-competent transcription complex. There are two basic processes involved in the formation of a pre-initiation complex in the cell:

– promoter recognition and promoter binding by TFIID (TBP and TAFs)
– binding of the RNA polymerase holoenzyme to the promoter-bound TFIID

Transcriptional activators can intervene as regulators at various steps in the initiation of transcription. They can interact with components of TFIID, as well as with components of RNA polymerase II, to stimulate transcription. Regulated transcription generally requires the aid of further protein components, which are commonly termed *coactivators* (see 1.4.3.2). An understanding of the details of coactivator function is only just emerging.

Some of the coactivators are ascribed a mediator function between the DNA-bound transcriptional activators and the pre-initiation complex, while others are attributed an active role in the restructuring of the chromatin.

#### 1.4.2.4 Phosphorylation of RNA Polymerase II and the Onset of Transcription

The large subunit of RNA polymerase II plays an important role at the beginning of the transcription process. The large subunit of the mammalian enzyme contains 52 copies of the heptamer sequence YSPTSPS in the C-terminal domain (CTD) at which phosphorylation occurs. Phosphorylation occurs extensively on the Ser-residues of the CTD, to a lesser degree at the Thr-residues, and, very rarely, at the Tyr-residues. Two forms of RNA polymerase II can be isolated from cellular extracts: a underphosphorylated form and a hyper-phosphorylated form. The isoforms fulfill different functions: RNA polymerase found in the initiation complex tends to display little or no phosphorylation at the C-terminus of the large subunit, while RNA polymerase II active in elongation is hyperphosphorylated in this region of the protein.

Based on the above observation, phosphorylation at the C-terminus is believed to serve as a trigger for disosciation from the inititation complex (Fig. 1.32). It is suspected that the unphosphorylated C-terminus of the large subunit is able to form contacts to the TATA Box binding protein, as well as with the mediator complex (see below). The high density of negative charges at the C-terminus resulting from phosphorylation disrupts these interactions, thereby releasing RNA polymerase into the elongation process.

The phosphorylation of the CTD plays an additional role in the maturation of the pre-mRNA and in the formation of the mRNA processing factory. For this reason,

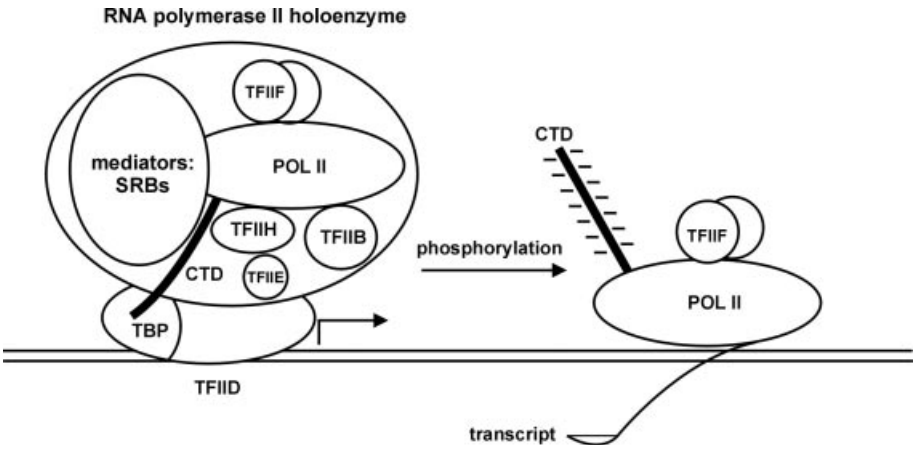



**Fig. 1.32.** Phosphorylation of the C-terminal domain of RNA polymerase II and the beginning of transcription. The transition from the initiation complex to actual begin of transcription is regulated via phosphorylation of the C-terminal domain (CTD) of RNA polymerase II. In the above model it is assumed that initially a complex is formed between TFIID and a holoenzyme of RNA polymerase consisting of RNA polymerase II and associated factors (mediators, SRB proteins) and the basal transcription factors. Phosphorylation of the C-terminal domain effects the dissociation of the RNA polymerase from the initation complex and the transition to the elongation phase. A protein kinase, which is part of TFIIH, is responsible for the phosphorylation. The nature of the signal that induces phosphorylation of RNA polymerase II remains unknown. SRB: suppressor of RNA polymerase B. After Koleske and Young (1995).

the CTD is associated with the splicing apparatus and with the proteins involved in polyA processing.

The protein kinase responsible for phosphorylation at the C-terminus is localized in TFIIH. It has been shown (Shiekhattar et al., 1995) that TFIIH contains a Ser/Thr-specific protein kinase termed CDK7 (or MO15). CDK7 belongs to the family of cyclin-dependent protein kinases (see 14.2.1). The cognate cyclin, cyclin H, is also found in TFIIH.

CDK7/cyclin H were shown to be identical to the CDK-activating protein kinase CAK. CAK is ascribed an important role in the regulation of the cell cycle (see 14.2.1). The functional importance of this identity is not yet fully understood.

#### 1.4.2.5 TFIIH-A Pivotal Regulatory Protein Complex?

TFIIH is a multi-protein complex consisting of at least eight different subunits (review: Drapkin and Reinberg, 1994; Svjestrup et al, 1996).

A more in-depth characterization of the subunits brought surprising results. Some of the polypeptides of TFIIH could be identified as proteins shown in other experiments to participate in repair of DNA damage, as well as in the regulation of the cell cycle (CDK7/Cyclin H, see above). TFIIH itself, or other individual components of TFIIH, thus participate in the following fundamental processes in the cell:

– transcription
– excision repair of DNA lesions
– regulation of the cell cycle

For the various functions there are different forms of TFIIH which differ in protein composition and activity (Svjestrup et al., 1995; Svjestrup et al., 1996). There are transcription-competent forms which are effective as basal transcription factors. A repair-competent form is known as the *repairosome*. It has long been known that DNA repair and transcription can be coupled. An increased repair of DNA damage is observed while a gene is being transcribed. The association of a „core TFIIH“ with further, repair-specific proteins (in yeast: Rad1, among others) provides the obvious explanation for this coupling phenomenon.

The picture of the structure and function of TFIIH is a varied and complex one. The mechanistic details are long not understood.

## 1.4.3 Regulation of Eucaryotic Transcription by DNA-binding Proteins

Primary controlling elements of the transcriptional activity in eucaryotes are specific DNA-binding proteins. They bind cis-acting DNA elements and have a specific influence on the initiation of transcription.

As in procaryotes, there are transcription activating proteins, the transcriptional activators, as well as proteins which inhibit transcription, the transcriptional repressors. Of the two classes, the activators are the most extensively studied and characterized. The mechanism of eucaryotic repressors has only been clarified in a few cases.

### 1.4.3.1 The Structure of Eucaryotic Transcriptional activators

It is generally true that that eucaryotic transcriptional activators act via direct or indirect protein-protein contacts with the transcription apparatus. In this context the definition of the transcription apparatus is rather broad: it includes the basal apparatus as well as proteins which act as co-activators or mediators (see 1.4.3.2).

For a DNA-binding protein to engage in regulation of activation of transcription it must possess the following functions:

– specific DNA-binding
– communication with the transcription apparatus
– ability to be regulated by effectors

In many cases the various functions are located on independently folding protein domains, resulting in modularly constructed transcriptional activators. The function of a DNA-binding protein can thus often be deduced from the primary sequence information and from homology with other eucaryotic transcriptional activators. Furthermore, the modular structure of transcriptional activators is the prerequisite for the

swapping of individual activities using recombinant DNA technology to create hybrid transcriptional activators (see below).

### DNA-binding

Transcriptional activators bind specifically to cognate DNA elements variably located relative to the promoter (see 1.4.2.2) and can interact directly or indirectly with the transcription apparatus. Transcriptional activators depend on the occurrence of regulatory DNA elements for their action and perform their function on specific genes. They are thus termed *specific transcriptional activators*, to distinguish them from proteins that activate transcription independent of specific DNA elements (see 1.4.3.2).

The DNA-binding serves to bring the binding protein in proper orientation to the pre-initiation complex. Binding to cis-elements creates a high effective concentration of binding proteins close to the transcription appartus to enable productive interactions. Because protein-induced conformation changes of the DNA aid in orienting the transcription complex on the promoter, the conformation and bending of DNA also plays an important role (see chapter 1.2.3). The structural motifs of eucaryotic transcriptional activators for the recognition of specific DNA sequences are varied and include all the DNA-binding motifs summarized in 1.2.1.

In many cases the specific DNA-binding serves only to create a high concentration of the transcription factor in the vicinity of the site of transcription initiation. This property can be used in domain swapping experiments to alter the specificity of regulation by design. Experiments, such as the one described in Fig. 1.33, have contributed significantly to our understanding of the structure and function of eucaryotic transcriptional activators. A prerequisite for successful *domain swapping experiments* is that the DNA-binding domain not influence the function of the transactivating domain. This is often, but not always, the case.

### Communication with the Transcription Apparatus

The first data concerning the structural requirements for communication with the transcription apparatus came from domain swapping experiments with the GAL4 protein of yeast.

One structural domain in the GAL4 protein could be shown to mediate a stimulation of transcription and was thus termed the *trans-activating domain*. The essential structural element of the trans-activating domain of GAL4 is an amphipathic α-helix with negatively charged and nonpolar amino acid side chains on opposite sides of the helix. The amount of negative charge correlates with the extent of the trans-activating effect and also determines the distance over which an effect on the transcription complex is still possible.

Examples for other trans-activating domains are the glutamine rich domains of the transcription factor Sp1 and the proline rich domain of the transcription factor CTF/NF1, which contains 20 % proline residues.

It is assumed that trans-activating domains are structural elements that can adapt to become complementary to a surface of the transcription apparatus in a flexible and rather unspecific manner. Information on the structure of the trans-activating domain,

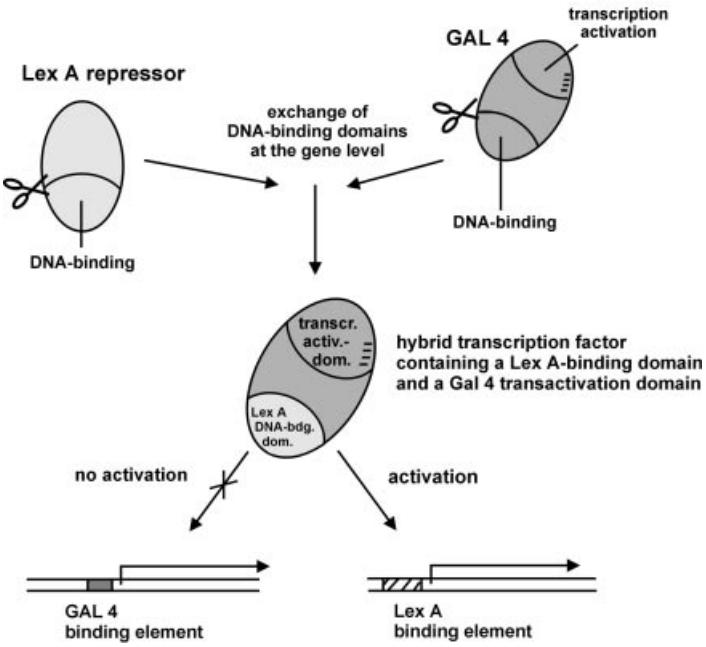



**Fig. 1.33.** „Domain Swapping" experiment for the creation of hybrid transcription factors. The functional domain of transcription factors often behave as independent structural units that can be exchanged among various proteins by genetic means. In the above example of a domain swap experiment, the DNA-binding domain of the bacterial LexA repressor is transferred to the GAL4 protein of yeast. GAL4 is a transcription activator that positively regulates the transcription of the enzymes of galactose metabolism. The genetically engineered hybrid transcription factor has the DNA-binding specificity of the LexA repressor and the transactivating specificity of the GAL4 protein. The hybrid transcription factor activates promoters with the LexA binding element, but not promoters with the GAL binding element.

as well as to the overall protein structure, is currently available for the steroid hormone receptors, among others (see chapter 4).

### 1.4.3.2 Concerted Action of Transcriptional activators and Co-activators in the Regulation of Transcription

The formation of an active, regulation-competent initiation complex for transcription in eucaryotes demands the concerted action of a large number of proteins. It is estimated that more than 50 different proteins participate in the initiation of transcription in eucaryotes. The basal transcription complex, consisting of the general initiation factors, as well as RNA polymerase II, allows only for a slow transcription rate. For a regulated acceleration of this low transcription rate it is necessary to have – apart from the regulatory DNA-binding proteins – mediation by further *co-activator proteins.*

The regulated activation of transcription thus requires at least two types of proteins:

**Specific Transcriptional activators**

The specific transcriptional activators (see 1.4.3.1), represented by the GAL4 protein of yeast, are sequence-specific DNA-binding proteins. They possess both a DNA-binding domain and a trans-activating domain to allow them to interact directly with the transcription apparatus.

**Coactivators**

Extensive studies on the reconstitution of a regulated transcription initiation in *in vitro* systems have shown that most of the specific transcription factors are not capable of stimulating transcription above the basal level without the assistance of further proteins. Further coactivators are required for this task. The coactivators can be subdivided into three classes (Fig. 1.34).

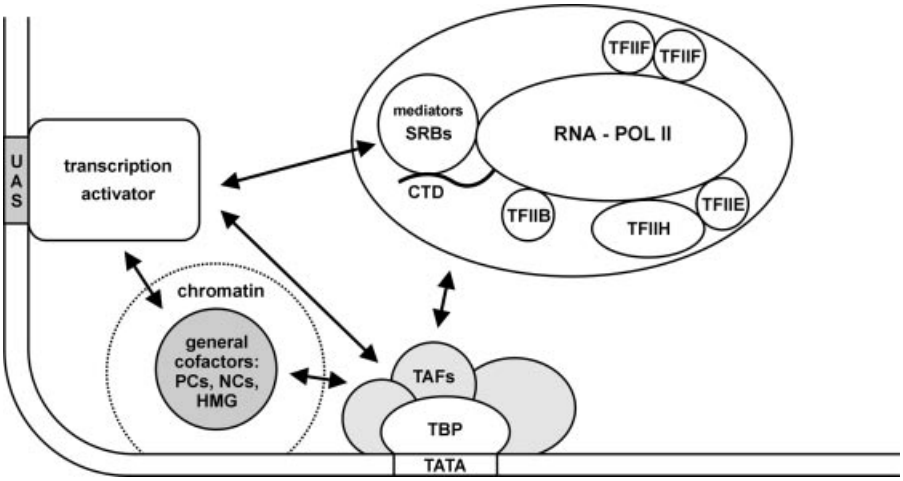



**Fig. 1.34.** Activators and coactivators of transcription intiation. The figure shows the function of three groups of proteins that function as coactivators. The general cofactors mediate the interactions between the specific transcription activators and the TFIID complex as well as with various forms of the RNA polymerase II holoenzyme. The TAFs are components of the TFIID complex and serve as contact points for specific transcription activators. The mediators are components of various forms of holoenzymes of RNA polymerase II. SRB proteins belong to the class of mediators, which, among other things, interacts with the CTD of RNA polymerase. The simplified diagram does not show the interactions with chromatin.
TBP: TATA box binding protein, TAF: TATA box binding protein associated factor, SRB: suppressor of RNA polymerase B, CTD: C-terminal domain of RNA polymerase II, PC: positive cofactor, NC: negative cofactor, HMG: high mobility group proteins, UAS: upstream activating sequence.

**1) TATA Box Binding Protein Associated Factors (TAFs)**

The TAFs are components of TFIID (see table 1.1) and are required for a regulated transcription (review: Verriijzer and Tijan; 1996, Burley and Roeder, 1996;). Thus, the stimulation of transcription by the transcriptional activators Sp1 and NTF-1 depends upon the presence of specific TAFs in the TFIID complex. The TAFs mediate interactions between the transcriptional activators and the TFIID complex; in many cases direct protein-protein interactions could be demonstrated between the activators and TAFs. Some of the TAFs possess additional enzymatic activities which allow them to participate in the regulation of transcription. By this token, the histone acetylase and protein kinase activity of $TAF_{II}250$ is ascribed a regulatory function in the remodeling of chromatin and in the control of the activity of the basal transcription factors.

Overall, the physiological function of individual TAFs is still incompletely understood. The situation is complicated by the fact that a subset of the TAFs, including the histone-like TAFs, have been identified as components of a large histone acetylase complex, termed SAGA in yeast and PCAF (p300/CBP associated factor) in humans (see 1.4.6).

**2) Mediators**

Mediators include proteins which are components of various forms of the holoenzyme of RNA polymerase II and interact with the C-terminal domain (CTD; review: Bjorklund and Kim, 1996). Mediators include, among others, the SRB protein; see 1.4.2.3).

**3) General Cofactors**

The general cofactors perform their stimulating function within the context of the entire basal transcription apparatus and are not associated specifically with any general transcription factor. Among the general cofactors are proteins which have a positive or a negative influence on the stimulating effect of specific transcriptional activators (review: Kaiser and Meisterernst, 1996). They are classified accordingly as *positive cofactors* (PCs) and *negative cofactors* (NCs).

Among the general cofactors are several proteins which bind DNA and associate with the chromatin. Some possess enzymatic activity. Examples for positive cofactors are DNA-topoisomerase I, poly-ADP-ribose polymerase, and HMG1 and HMG2 proteins (HMG: „high mobility group“), both associated with the chromatin. Of particular interest for regulation is the positive cofactor PC4, whose DNA-binding capacity is controlled via protein phosphorylation.

An example of a negative cofactor is the NC2 complex, which can repress the basal transcription level. The NC2 complex consists of two subunits, both displaying homology to the histone proteins. The repressive function of NC2 is due to its competition with TFIIB and TFIIA for the promoter binding site, thus blocking formation of the pre-initiation complex.

Among the general cofactors are also included proteins with histone acetylase activity (HAT) or histone deacetylase activity (HDAC). The acetylation/deacetylation of histones plays a significant role in transcription regulation of chromatin-coated DNA.

Examples for coactivators with histone acetylase activity are the CBP protein, the p300 protein and the GCN5 protein (see 1.4.6).

The existence of various classes of coactivators make the cooperative formation of a regulation-competent transcription initiation complex likely. All three classes of co-activators appear to be necessary for an efficient and regulated transcription event. The presence of one particular transcriptional activator and its cis DNA element is not sufficient for transcription activation. Generally, several coactivators must be present simultaneously, such as the TAF proteins, the mediators, as well as the general cofactors, in order for the regulatory signal to be transmitted to the basal transcription apparatus. The coactivators appear to be present as large multiprotein complexes with variable composition and they serve as a link between the DNA-bound transcriptional activator and the transcription apparatus, as well as for the restructuring of the chromatin and/or nucleosomes. The availability of the co-activators can, as with the transcriptional activators, be controlled in a tissue-specific manner either via gene expression or via signal transduction chains. This opens up a realm of possibilities for tissue-specific transcription activation.

#### 1.4.3.3 Interactions with the Transcription Apparatus

The activating domain of transcriptional activators requires specific binding partners within the transcription apparatus. Unequivocal identification of these binding partners has proven difficult due to the large number of proteins that participate in the formation of the pre-initiation complex. The targets can either belong to the group of basal transcription factors, such as TFIIB, or to the group of co-activators, such as the TAF protein of the TFIID complex. Using techniques to detect protein-protein interactions, such as co-immunoprecipitation or affinity chromatography, specific interaction between the acidic domain of the viral transcriptional activator VP16 and the 40kDa and/or 60kDa TAF protein from the TFIID complex were identified (further examples are given in Verrijzer and Tijan, 1996). Furthermore, *in vitro* and *in vivo* experiments have shown that the negatively charged domain of VP16 can interact with the positively charged region of TFIIB. The complex formation between VP16 and TFIIB is accompanied by a conformation change in TFIIB, as evidenced by increased sensitivity of TFIIB to protease (Roberts and Green, 1994).

The pathways by which specific transcriptional activators and co-activators influence transcription initiation remain little understood and, for the most part, an item of speculation. There are two main pathways in discussion which possibly act cooperatively:

In one model it is assumed that transcriptional activators and coactivators increase the efficiency of formation of the pre-initiation complex. This function includes a restructuring of chromatin at the transcription start site. In this context the formation of the TFIID complex at the promoter plays an important role.

The other model views the activators and coactivators as responsible for the stable and defined spatial arrangement of proteins in the holo-complex and for the induction of topology in the DNA which promotes initiation.

It can be assumed that the extent to which either pathway is used depends on the structure of the specific gene, as well as on the structure of the chromatin.

## 1.4.4 Regulation of the Activity of Transcriptional activators

The repertoire of mechanisms for control of the activity of eucaryotic transcriptional activators (and also of coactivators) is varied and allows a spatially and temporally coordinated regulation of transcription.

The principle means by which the activity of sequence-specific DNA-binding proteins is controlled have already been presented in section 1.2. The importance of these mechanisms for regulation in eucaryotes will be discussed below. Altogether, the demands on eucaryotic organisms with regard to the regulation of transcription activity are much more complex than for procaryotes. This tenet holds for the structure of the transcription apparatus as well as for the mechanism of transcription regulation.

### 1.4.4.1 The Principal Pathways for the Regulation of Transcriptional activators

The principal mechanisms for the control of the DNA-binding activity of DNA-binding proteins have been presented in section 1.3.2.

Fig. 1.35 gives an overview of the most important mechanisms by which the transcription regulating activity of specific DNA-binding processes in eucaryotes can be controlled. They include de *novo synthesis*, as well as the *modification* and *availability of pre-existing proteins*. The specific expression of transcriptional activators is of great importance during development and differentiation of organisms where long-term changes in gene expression are required. *Concentration gradients* of diffusable regulatory proteins for the specific control of gene expression are used in development, as shown for the bicoid protein of Drosophila (Driever and Nuesslein-Vollhard, 1989).

For the coordination of metabolic processes or for the regulation of cell division the cell relies primarily on post-translational modification of pre-existing regulatory proteins. The activity or availability of preexisting regulatory proteins can be adjusted rapidly and effectively by post-translational modifications so that an immediate reaction within the framework of intercellular communication is possible. The modification of pre-existing regulatory proteins thus plays a more important role amongst eucaryotes than in bacteria. A further significant regulation mechanism is the *binding of effector molecules* to the DNA-binding proteins (compare 1.3.2). The influence of steroid hormones on the DNA-binding activity of steroid hormone receptors is an example of this regulatory mechanism.

Chromatin structure plays a distinctive role in the regulation of gene expression in eucaryotes. The nuclear eucaryotic DNA is associated with histones and other proteins, whereby the DNA is tightly wound and extremely compact. The packaging of DNA into chromatin can affect the accessibility of genes for regulatory proteins and can inhibit gene expression. Modifying DNA, e.g. formation of 5-methylcytidine ($m^5C$), is another means by which the cell controls gene expression.

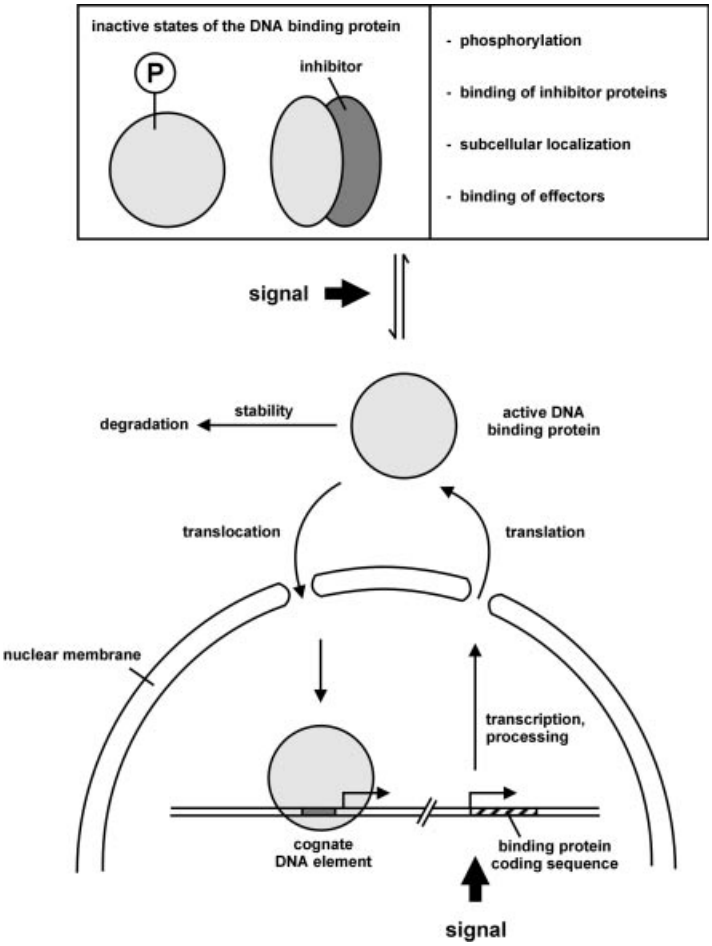



**Fig. 1.35.** Mechanism for the control of the activity of transcription factors. Regulatory DNA-binding proteins can occur in binding active and binding inactive forms. The transition between the two forms is primarily controlled by the mechanism shown schematically above. Activation or inactivation of transcription factors is determined by signals that can be of an internal or external origin. Furthermore, the amount of available transcription factor can also be regulated via its degradation rate or rate of expression.

### 1.4.4.2 Phosphorylation of Transcriptional activators

The phosphorylation of proteins on Ser, Thr or Tyr residues is a basic tool for the regulation of protein activity (see 7.1). Many eucaryotic transcriptional activators are isolated as phosphorylated proteins. The phosphorylation occurrs mainly on the Ser and Thr residues, but can also be observed on the Tyr residues. The extent of phosphorylation is regulated via specific protein kinases and protein phosphatases, each components of signal transduction pathways (see ch. 7). The phosphorylation of transcriptio-

nal activators often represents the final event of a signal transduction chain targeted for a change in gene expression.

An example for how protein phosphorylation can influence the transcription process is the transition from the initiation to the elongation process for RNA polymerase II (see 1.4.2.4).

Phosphorylation of transcriptional activators can influence the transcription activity according to the following mechanisms:

**Regulation of the Nuclear Localization by Phosphorylation**

Proteins which act in the nucleus require specific sequences, known as nuclear localization sequences, to direct their transport from the cytoplasm to the nucleus. The *nuclear localization sequences* are generally found at the C-terminus of a protein and often comprise basic amino acids.

Phosphorylation in the nuclear localization sequence of transcriptional activators can decide whether transport into the nucleus, and subsequent transcription activation, occurs or not.

An example for this regulation mechanism is the SWI5 protein of yeast (Fig. 1.36): SWI5 is a transcriptional activator which up-regulates the expression of the HO endonuclease in yeast. SWI5 occurs in two different forms during the cell cycle:

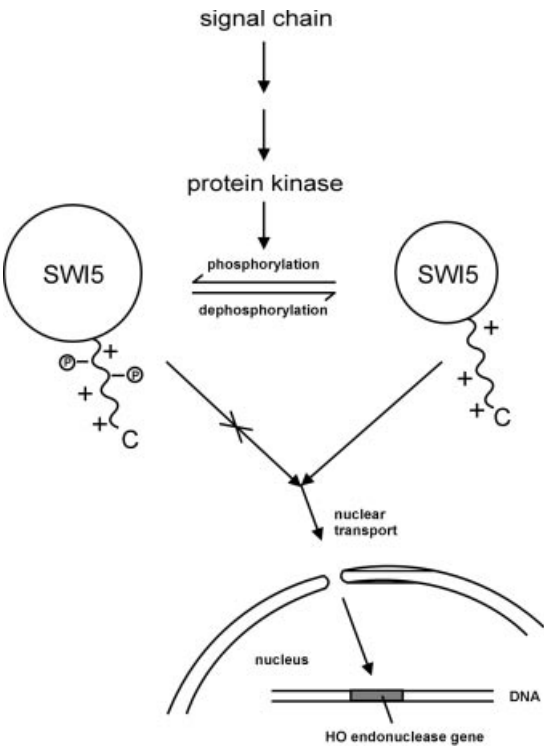



**Fig. 1.36.** Regulation of the subcellular localization of the transcription factor SWI5 in yeast by phosphorylation. The subcellular localization of the SWI5 protein is regulated by phosphorylation/dephosphorylation. In the phosphorylated state, SWI5 is found in the cytoplasm, while in the underphosphorylated state it is localized in the nucleus. Phosphorylation and dephosphorylation are catalyzed by either protein kinases or protein phosphatases and can be controlled via signal transduction chains.

In the G1-phase SWI5 is localized in the nucleus and induces the gene for the HO endonuclease.

In the S-, G2- and M-phases SWI5 is localized in the cytoplasm and can thus not be active as a transcriptional activator. The reason for the change in subcellular localization of SWI5 is phosphorylation in the region of the nuclear localization sequence. SWI5 possesses three sequences in the nuclear localization signal for phosphorylation at Ser and Thr residues. Cytoplasmically localized SWI5 is phosphorylated at these positions, thus blocking transport into the nucleus. The protein thus remains in the cytosol.

SWI5 is dephosphorylated at the beginning of anaphase, whereupon transport into the nucleus, binding to the cognate DNA element and stimulation of transcription becomes possible. The significance of Ser-phosphorylation for the function of SWI5 has been well documented experimentally. Mutation of the specific Ser residues to nonphosphorylatable Ala leads to a constitutive nuclear localization of the mutated protein and permanent activation of the SWI5-target genes.

The phosphorylation state of the transcription factor NF-AT has a different effect on translocation. The phosphorylated form of this protein is localized in the cytosol and requires dephosphorylation by the protein phosphatase calcineurin in order to be translocated to the nucleus (see also 7.5.2). Other examples for phosphorylation-dependent nuclear translocation include the STAT-proteins (see 11.1.3.2) and the SMAD-proteins (see 12.1.2).

In these cases however, phosphorylation of the transcription factors is required before translocation into the nucleus can occur.

The biochemical basis for the phosphorylation-dependent cytoplasmic localization of transcription factors appears to be a specific interaction of the phosphorylated forms with the nuclear export machinery, allowing the specific export into the cytoplasm of the phosphorylated form only (Kaffman et al., 1998). The preferential export of the phosphorylated form will lead to an increased cytoplasmic localisation of the protein.

**Phosphorylation of the DNA-binding Domain**

There are many examples for the specific phosphorylation of gene regulating proteins within their DNA-binding domain. Phosphorylation can influence either positively or negatively the abiltiy to specifically bind DNA. The situation becomes complicated by the fact that many transcriptional activators possess multiple phosphorylation sites, whose phosphorylation can have different effects on DNA-binding.

How phosphorylation interferes with DNA-binding is not well understood. Several mechanisms are conceivable:

- direct interference with the DNA-binding due to electrostatic effects
- inhibition of the dimerization of the transcriptional activators
- induction of conformational changes of the protein which cause inhibition or enhancement of DNA-binding

**Phosphorylation of the Trans-Activating Domain**

The trans-activating domains of transcriptional activators are also common substrates for phosphorylation by protein kinases. The details by which the phosphorylation

affects the interactions with the basal transcription apparatus is known in very few cases. The reason for this is the difficulty of identifying the specific interaction partner in the complex transcription apparatus.

Exemplary is the regulation of the CREB protein of higher eucaryotes, displayed in Fig. 1.37.

The CREB protein is a transcriptional activator for genes with cis-regulatory, cAMP-sensitive DNA elements (cAMP responsive elements, CREs).

CREs are DNA sequences which mediate cAMP-regulated transcription. An increase in the cAMP concentration due to hormonal stimulation (see chapter 5,6) activates protein kinases, which can then lead, either directly or indirectly, to phosphorylation

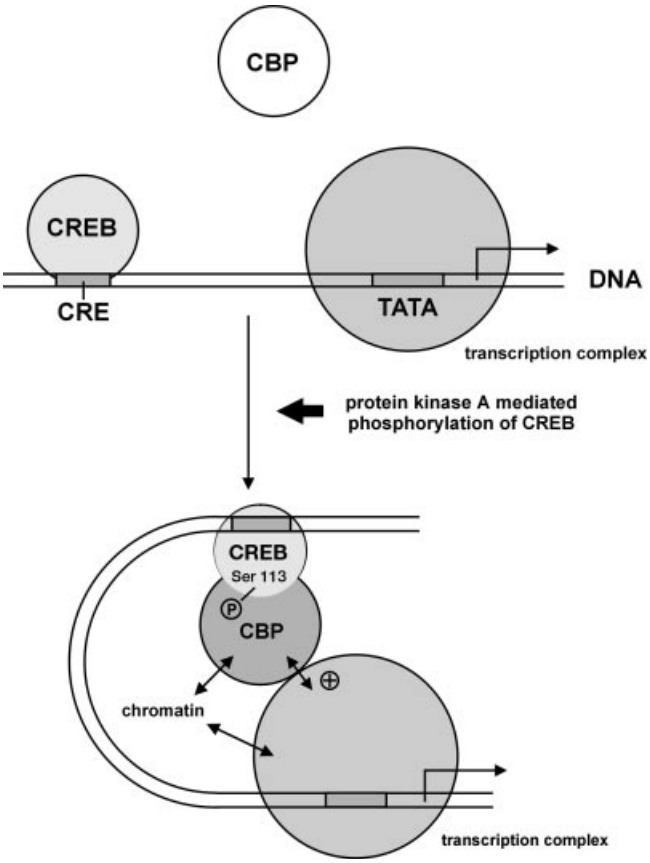



**Fig. 1.37.** Regulation of the activity of a transcription factor by phosphorylation. The CREB protein is a transcription factor that binds to CREs (CRE: cAMP responsive element) tand thereby activates the cognate genes. The CREB protein (CREB: CRE-binding protein) requires the assistance of CBP (CBP: CREB binding protein) for efficient transcription activation. CBP can only act as a coactivator if the CREB protein is phosphorylated on Ser113. The phosphorylation of CREB is controlled by a signaling pathway involving cAMP as an intracellular messenger. The DNA element is termed CRE because the cognate gene is regulated by a cAMP-dependent signaling pathway.

and regulation of transcriptional activators. The transcription stimulation of the cognate genes requires the binding of CREB to the CREs and the phosphorylation of CREB at Ser133. This phosphorylation event is mediated by a cAMP-dependent signal transduction pathway. A second protein was shown to participate in the transactivation via CREB. This protein, termed CBP (CREB binding protein), binds specifically to CREB and mediates the interactions between CREB and TFIIB of the basal transcription apparatus. CBP has also histone acetylase activity and is part of the multiprotein complex PCAF (see 1.4.6). The binding of CBP to CREB depends upon whether it is phosphorylated at Ser133: only if Ser133 of CREB is phosphorylated can CREB and CBP interact, and only then can CBP fulfill its function as a mediator between the basal transcription apparatus and the transcriptional activator.

#### 1.4.4.3 Heterotypic Dimerization

Most transcription activators bind to DNA as a dimer or higher multimer (see 2.4). The dimerization relies on structural motifs which commonly occur in many different proteins. Examples for dimerization motifs are the *helix-loop-helix motif* and the *leucine zipper*. The dimerization motifs permit the formation of DNA-bound *homodimers* or *heterodimers*, depending upon whether the same or different proteins interact with each other (Fig. 1.38). The different dimers have different requirements for the

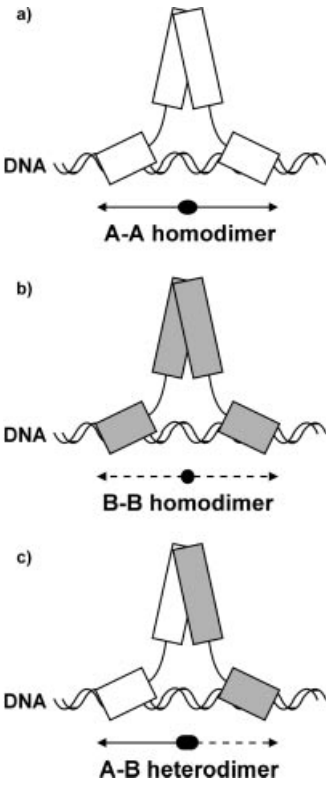



**Fig. 1.38.** Formation of homo- and heterodimeric transcription factors and the specificy of DNA-binding. Shown are two different helix-loop-helix proteins, which bind as homodimers (a, b) to the each of their cognate palindromic DNA elements (drawn as arrows). The two homodimers display different DNA-binding specificity. The heterodimerization (c) of the two proteins creates a complex that recognizes a hybrid DNA element.

sequence of the DNA-binding elements and they can influence transcription activity in very different ways. As shown in Fig. 1.39, families of interacting transcriptional activators can be distinguished (review: Lamb and Knight, 1992).

The heterotypic dimerization significantly expands the repertoire for tissue-specific regulation of transcription activity. The tissue-specific expression of a particular pattern of transcriptional activators can be used to select only certain DNA-binding elements out of a series of similar elements, and thus to specifically induce certain genes. This strategy is extensively used by the receptors for retinoic acid (see chapter 4).

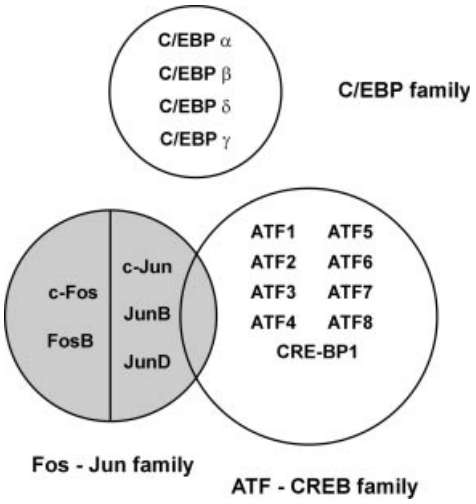



**Fig. 1.39.** Examples for families of interacting transcription factors. The circles indicate groups of eucaryotic transcription factors that can form homo- and heterodimers amongst themselves. The intersection of the circle of the ATF family with the circle of the Jun family indicates possible heterodimerization between the two families. The members of the Jun family can form complexes with members of the Fos family and with the members of the ATF family. The Fos family is unique in that its members can not form homodimers, but must heterodimerize with members of the Jun family. C/EBP: CCAAT/enhancer binding protein; ATF: activating transcription factor; CRE-BP: cAMP responsive element binding protein. After Lamb and McKnight (1992).

#### 1.4.4.4 Regulation by Binding of Effector Molecules

The activity of eucaryotic transcriptional activators can be regulated by the binding of low molecular weight effectors, as well as by the binding of inhibitor proteins (see 1.3.2.3). The most significant example for transcriptional activators regulated by low molecular weight effectors are the nuclear receptors, which will be discussed in more detail in chapter 4.

In this system, cognate hormones act as positive regulating effectors. The transcription regulating activity of nuclear receptors can also be negatively influenced by specific inhibitor proteins. These proteins are characterized as repressors.

### 1.4.5 Specific Repression of Transcription

A further significant mechanism of transcription control is the repression of gene expression (review: Cowell, 1994; Johnson, 1995). There are two types of gene repression to be distinguished in eucaryotes. On the one hand the chromatin structure can cause an unspecific repression of gene expression (see 1.4.6). On the other hand, analogous to the specific transcriptional activators, there are specific repressors of transcription. Their effect, in contrast to that of unspecific repressors, is sequence-dependent and thus suitable for selective repression (Fig. 1.40). DNA sequences that mediate repression of transcription factors are termed *silencers*.

Specific repressors can exert their influence directly or indirectly. *Indirect repression* is, for example, if a repressor protein competes with transcriptional activators for the binding site on the promoter. The extent of repression is then determined by the relative affinity of both proteins to the DNA element and their concentration ratios. The DNA-bound repressor in this case does not act as a transcriptional activator.

A further possibility for indirect repression results from heterodimerization (see 1.4.4.3). Heterodimers between two transcription factors, in which one of the partners possesses a DNA-binding domain with weak affinity, can inactivate a transcriptional activator in a heterodimer complex. Since a strong binding to the DNA element usually requires both subunits of a DNA-binding protein, transcription activation by this type of heterodimer is not possible.

*Direct repressors* interact with the basal components of the transcription apparatus or with transcriptional activators to inhibit their activity. Specific repressors, analogous to transcriptional activators, are constructed modularly, with a DNA-binding domain and a repressor domain. The repressive character of such domains has been proven in domain swapping experiments. The mechanism of specific repression remains speculative. The following mechanisms are, however, conceivable:

- Direct inhibition of the formation of a pre-initiation complex: complexation of basal transcription factors, such as TFIID or TFIIB, or competition with TFIIB for binding to the promoter. An example for this type of repression is the negative cofactor NC2 (see 1.4.3.2). Transcription repression can also result from phosphorylation of the basal transcription factors. By this token, the repression of transcription observed during mitosis is attributed to the hyperphosphorylation of TBP and TAFs.
- Inactivation of the trans-activating protein by specific complex formation.
- Inhibition of the transition from the initiation phase to the elongation phase.
- Induction of a chromatin structure that does not allow the efficient formation of the pre-initiation complex; e.g. by deacetylation of histones ( see.1.4.6).

An interesting and functionally important aspect of transcriptional activators is that one and the same protein can act as both an activator and a repressor. The alternative functionality is determined by the sequence environment, the presence of other transcriptional activators (steroid receptors, see ch. 4), by specific repressors or by low molecular weight effectors. Examples are the receptors for vitamin A acid, which, in the absence of its ligand, represses the genes with cognate DNA elements. The repression is exerted in the DNA-bound form. In the presence of its ligand, vitamin A acid, the same receptor acts as a transcriptional activator (see ch. 4).

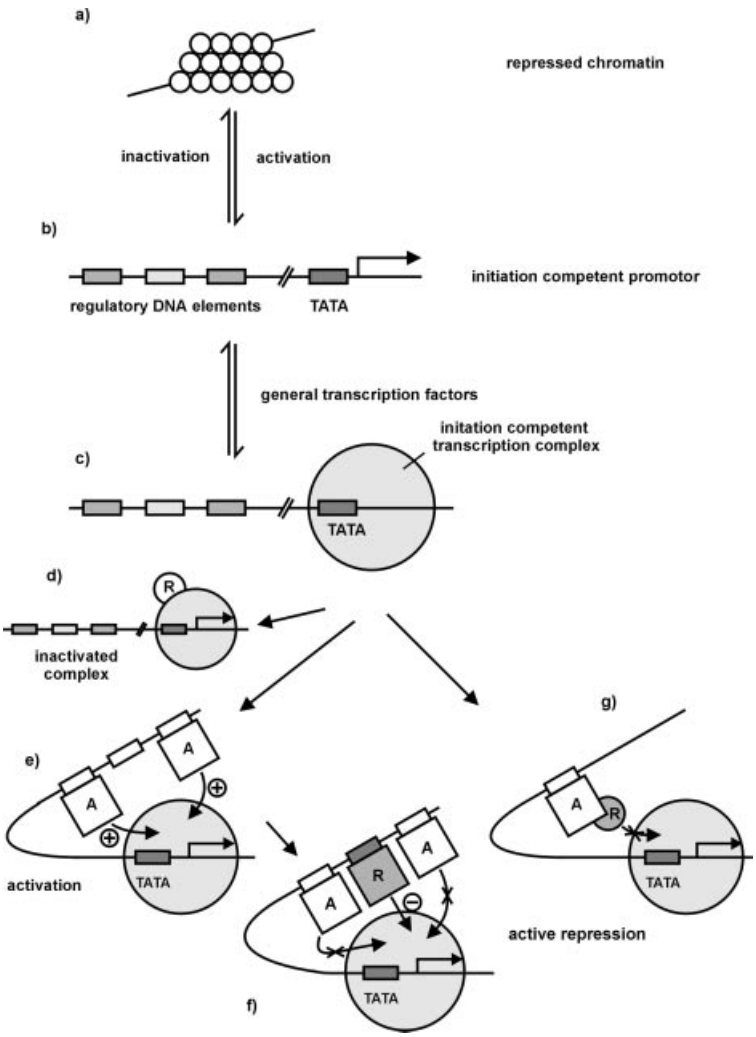



**Fig. 1.40.** Model of repression and activation of transcription. The figure illustrates various mechanisms of repression of transcription. a) genes are in a generally repressed states in inactive chromatin. In a first phase of activation the chromatin is restructured. b) The promoter is now accessible for the binding of the basal transcription factors and for RNA polymerase II. c) An initiation complex is formed that contains the central components of the transcription apparatus, but which enables transcription only at a low rate. d) the binding of repressors to the transcription initiation complex can prevent further activation of transcription at this step. e) the binding of transcription activators to their DNA elements leads to activation of transcription. f) an active repression is affected by proteins that bind sequence specifically to DNA elements and in their DNA-bound form inhibit the transcritption preventing interactions with the transcription apparatus.

The mechanism by which these receptors repress transcription in the absence of their ligands remains unclear. Specific proteins, termed nuclear receptor co-repressors, also participate in the repression. They are an integral part of a larger protein complex

that binds the nuclear receptor and possibly weakens the transcription of the gene via the activation or recruitment of a histone deacetylase activity (see 1.4.6). It is assumed that the repressor protein dissociates from the receptor in the presence of the cognate hormone to enable transcription activation.

### 1.4.6 Chromatin Structure and Transcription Activation

In the previous considerations the function of chromatin structure on transcription activation has been ignored. Generally it holds true that chromatin structure is decisive for gene activity and certain configurations of chromatin are associated with transcription repression.

The basic structural unit of chromatin is the nucleosome, in which the DNA is wrapped 1.65 turns around the histone octamer $(H2A, H2B, H3, H4)_2$. The chromatin is further condensed to a so-called solenoid with the aid of histone H1. The following observations are relevant to the discussions about the role of chromatin structure in gene regulation (Lewin, 1994, Workman and Kingston, 1997, Kadonaga, 1998):

- Actively transcribed chromatin displays a change in nucleosome structure
- The arrangement of nucleosomes on the DNA plays an important role in the activation of transcription and can be actively changed
- There are chromatin structures which generally inhibit transcription

In view of the size of the nucleosome, it appears very difficult to carry out transcription initiation and elongation on nucleosome-covered DNA. The special role played by the nucleosome in the total transcription process is confirmed by the following experimental findings:

- During elongation the RNA polymerase is capable of displacing the nucleosome. The histone octamer does not dissociate completely from the DNA. The nucleosome is thus a mobile entity that plays a dynamic role in the transcription process.
- During transcription activation, a *defined rearrangement of nucleosomes* can be observed. The movement of the histones is coordinated with the shift of the DNA sequence from the nucleosome surface into the linker region between nucleosomes. A shift of the octomer, which has a height unequal to the 10.3 bp rise of DNA, causes DNA sequences to change their position relative to the surface of DNA. Sequences which point toward the core of histone octamers are not available for recognition by DNA-binding proteins.
- Nucleosome-bound DNA is accessible for the binding of transcriptional activators. This means that the DNA sequences associated with the nucleosome surface can still specifically be recognized by DNA-binding proteins. For example, DNA footprinting studies demonstrated the specific binding of a steroid hormone receptor to the cognate DNA-binding element, even if this sequence lies in a nucleosome-bound region. A requirement for sequence-specific binding of nucleosome-bound DNA is that the DNA-binding element be oriented outward from the nucleosome. This also requires an exact positioning of the DNA. The winding of the DNA aro-

und the nucleosome has the additional effect that DNA sequences separated by great distances in one dimensional space are brought close together in three dimensional space (Fig. 1.41).

- Proteins of the yeast SWI/SNF family can initiate a modification of nucleosome structure to enable the interaction of the transcription factors with the nucleosome-bound DNA. Proteins of this class can function as anti-repressors by opposing the general repression of chromatin structure. The SWI/SNF proteins are contained in a large protein complex. They can modify nucleosome structure under ATP hydrolysis in such a way as to strengthen the binding of transcription factors, such as GAL4 or the TATA box binding protein, to nucleosome-bound DNA (Cote et al., 1994,).

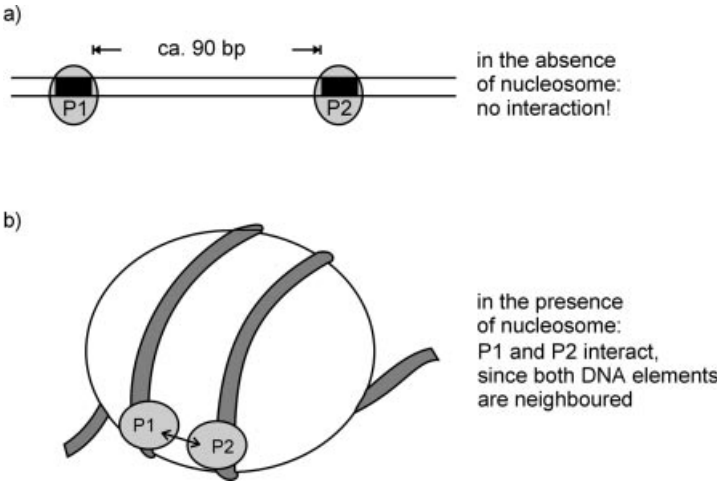



**Fig. 1.41.** The influence of the nucleosomes on the positioning of DNA-binding proteins. Example of a control region in which two regulatory DNA elements are separated by 60–90 bp but are brought near each other in 3D space via nucleosome formation. The super-helical arrangement of the DNA in the nucleosome brings the the two DNA elements close together. The DNA element-bound proteins P1 and P2 are brought into closer contact with each other in this configuration than in a linear arrangement.

It is becoming increasingly evident that multiple linkages exist between transcription and chromatin. A large part of the function of transcription factors is dependent on and directed to the structure of chromatin. A diagram of the network of linkages between chromatin and transcription is shown in Fig. 1.42A. Within this network, interactions between transcriptional activators (or repressors) and the following multiprotein complexes have been shown to occur (review: Kadonaga, 1998):

- RNA polymerase II holoenzyme
- chromatin remodeling machines
- chromatin assembly factors
- Histone acetylase and Histone deacetylase complexes

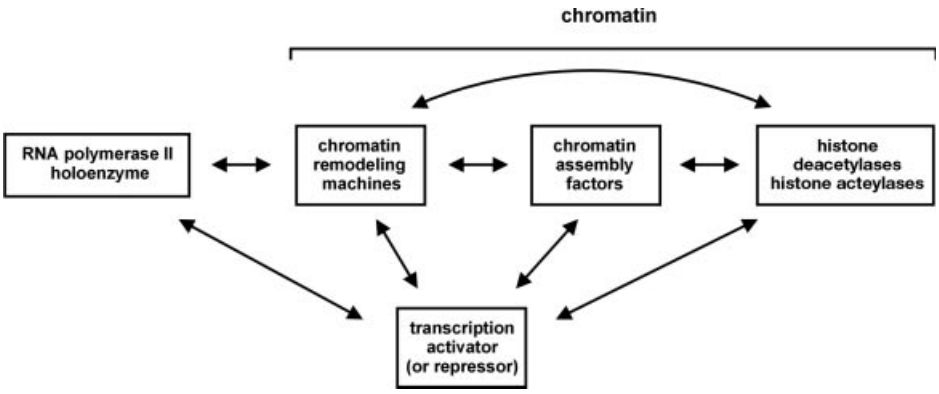



**Fig. 1.42A.** Schematic diagram of potential linkages between chromatin and transcription. The figure shows a simplified representation of some of the linkages between chromatin and transcription proposed on the basis of experimental evidence (Kadonaga, 1998) .

Although there is not universal agreement with regard to the hierarchy and the temporal order of the linkages depicted in Fig.1.42A, the action of transcriptional activators (or repressors) must be discussed in the context of this network. The large number of possible – and already experimentally shown – interactions between transcriptional activators and multiprotein complexes other than the RNA polymerase II holoenzyme makes it very difficult to adress specific questions on the detailed function of transcriptional activators in the context of chromatin covered DNA. The following dicussion therefore concentrates on the best studied aspect of chromatin modification during transcription activation, namely the acetylation and deacetylation of histones.

### 14.6.1 Transcriptional Activity and Histone Acetylation

Recent studies have shown that the acetylation or deacetylation of the histones of the nucleosome plays an important role in the regulation of transcriptional activity. Acetylation of the histones ( review: Hassig and Schreiber, 1997) is a postranslational modification which is usually performed on lysine residues at the N-terminus and requires specific enzymes,the *histone acetyl transferases (HATs).* Removal of the acetyl group also requires specific enzymes, the *histone deacetylases (HDAC).* Most importantly, the acetylation of histones is accompanied by a loss of postive charges which is thought to have a profound influence on the nucleosome structure and on the strength of DNA-binding.

#### Histone Acetylases

In many, although not all, cases, the histone acetylation is correlated with a stimulation of transcriptional activity. HAT activity is found in proteins associated with the transcription apparatus or identified as coactivators. Examples are, among others, members

of the TAF proteins, including $TAF_{II}250$, The GCN5 protein (yeast and human) and the coactivator CBP and the related protein, p300. The histone acetylases are part of large protein aggregates of complex and varying composition (review: Struhl and Moqtaderi, 1998). Two types of protein complexes have been characterized, the SAGA complex from yeast and the human PCAF. Both complexes contain several histone-like TAFs and HAT enzymes, GCN5 in the yeast SAGA complex and a human GCN5 homogue in the PCAF complex.

The mechanism of transcriptional activation via histone acetylation remains speculative. Two models, which are not mutually exclusive, are currently discussed:

- alteration of nucleosome structure: the neutralization of positive charges of the histones by acetylation leads to a modification, or loosening, of the nucleosome structure (fig 1.42B). The loss of positive charges will weaken the interaction of the nucleosomes with the DNA: Related to this can be an increased mobility of the nucleosome on the DNA as well as improved accessibility of the DNA for transcriptional activators.
- Binding of additional accessory proteins: the acetylation pattern can serve as a signal for the association of further protein components wit the genetic element, whereby transcription of the gene is modified.

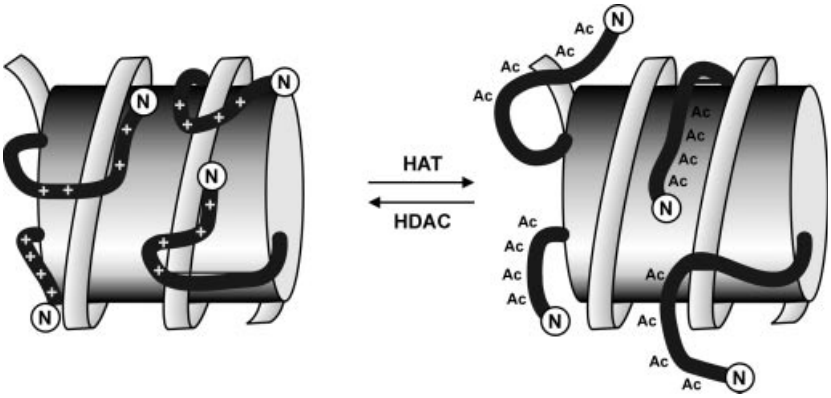



**Fig. 1.42B.** Model of the influence of histone acetylation and deacetylation on the nucleosome structure.

The amino terminal tails of the four core histones contain lysines that are acetylated by HATs and deacetylated by HDACs. The histone octamer $(H2A, H2B, H3, H4)_2$ is represented as a cylinder wrapped by DNA. It is thought that neutralization of the positive charges on the histone tails results in alterations of the nucleosome structure that may lead to a higher mobility of the nucleosome and/or an improved accessibility of the bound DNA, with accompanying changes in chromatin structure, chromatin hierarchy and transcription. In most, but not all cases, deacetylation correlates with the repressed state and acetylation correlates with the transcriptionally active state.

The amino terminii of the histones are indicated by a circled N; charged lysines are represented by +; and acetylated lysines are indicated by Ac. Only the hyper- and hypoacetylated states are depicted.

**Histone Deacetylases**

The histone deacetylases are found in large protein complexes, often together with repressive transcription factors. By this token, interactions of the repressive heterodimeric transcription factor Mad-Max and a complex with the histone deacetylase HDAC I and the mSin3A protein have been demonstrated. A complex of HDAC I and the nuclear receptor-corepressor (see chapter 4) binds to unliganded nuclear receptors and is believed to exercise a repressive effect. A further example is the tumor suppressor protein pRb (see chapters 13, 14), which can occur as a transcription repressor in the hypo-phosphorylated form and transcriptionally activating in the hyperphosphorylated form. The repressive form of the pRb protein recruits the histone deacetylase HDAC I to the DNA and thereby initiates an active repression of the gene (see 13.3.2).

Overall, histone acetylation and deacetylation represents an important tool with which transcription can be positively or negatively influenced. The nucleosomes and, in a further sense, chromatin structure assume a central role in the regulation of transcription. Nucleosome structure and nucleosome position can decisively contribute to the accessibility of DNA elements for transcription factors. The nucleosomes function as a framework that determines the spatial arrangement of a region of the DNA. The nucleosome constellation must be modified during transcription initiation, whereby the post-translational modification of histones in the form of acetylation or deacetylation plays a significant role. The participation of other non-histone proteins remains an open issue and it is also unclear how a constitutive and permanent inactivation of a section of DNA can be accomplished via the chromatin structure.

## 1.4.7 Methylation of DNA

Two states can be distinguished in chromatin: heterochromatin and euchromatin. The two states describe for every cell type a characteristic difference in degree of condensation and transcription activity of DNA. Genes located in the condensed heterochromatin can not be transcribed, whereby genes located in euchromatin are accessible for transcription.

An essential instrument for the suppression of transcription activity in heterochromatin, as well as for the differential regulation in euchromatin, is the *methylation of DNA* on the C5 atom of cytidine in the CpG sequence (Fig. 1.43). CpG sequences occur unevenly distributed in the genome. They may be concentrated in CpG islands. Higher eucaryotes possess a characteristic distribution pattern of 5-methyl cytidine ($m^5C$), which remains intact upon cell division. Mechanisms must therefore exist to ensure that the methylation pattern is precisely retained in the daughter cells following cell division. A methyl transferase that carries out *hemi-methylation* in the CpG sequences (Fig. 1.43) is responsible for the inheritance of the methylation pattern. The methyl group is derived from S-adenosyl methionine. The preferential substrates for the hemi-methylation are DNA sequences in which the complementary strand is already methylated. Such a hemi-methylation occurs, for example, shortly after replication of the sequence.

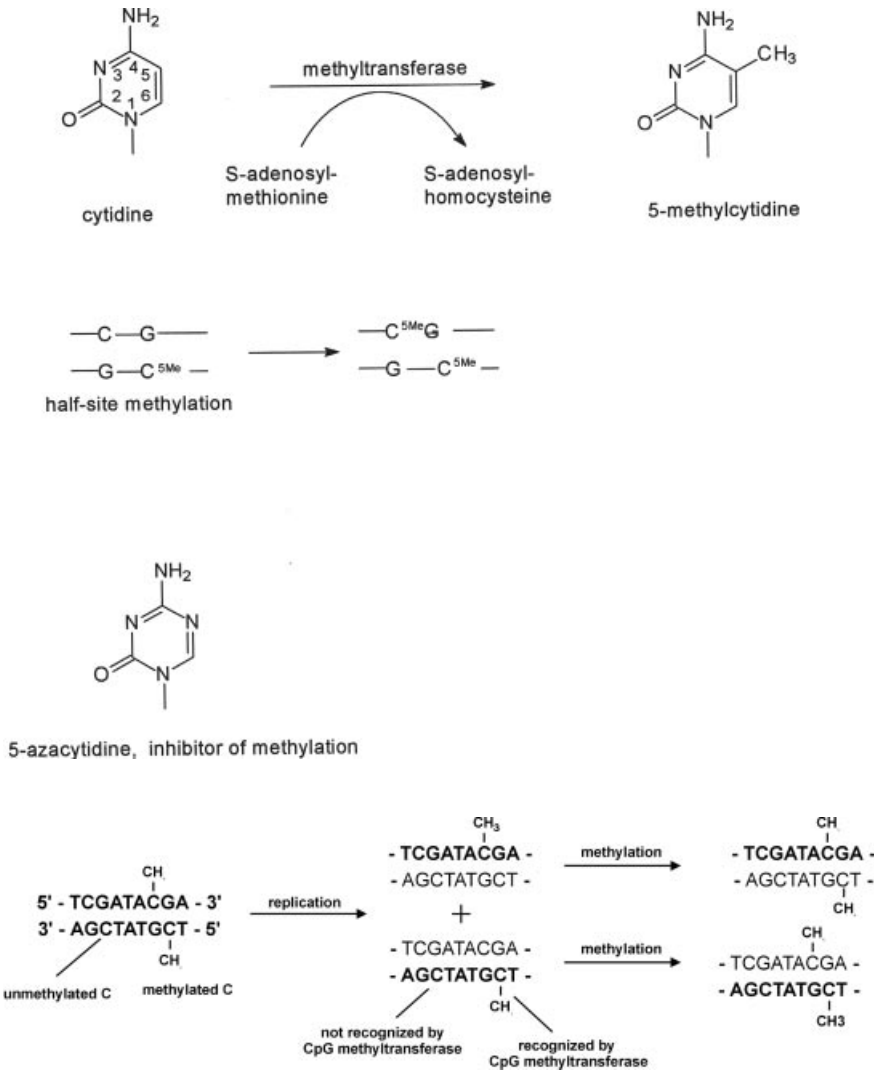



**Fig. 1.43.** The methylation of DNA: 5-methyl-cytidine and maintenance methylation. a) The methylation of cytidine residues on DNA is catalyzed by a methyl transferase that employs S-adenosine methionine as a methyl group donor. The peferable substrate for the methyl transferase are hemi-methylated CpG sequences. 5-aza-cytidine is a specific inhibitor of methyl transferses. b) The methylation pattern of DNA remains intact upon DNA replication and is passed on to the daughter cells. The newly synthesized strands are unmethylated immediately after DNA replication. The methyltransferase uses the previously methylated parent strand as a matrix to methylate the CpG sequences of the newly synthesized strand.

Based on the following observations, methylation can be viewed as a mechanism of *long-term repression*:

- Constitutively inactive genes display a high density of methylation; active genes tend to be under methylated.
- Inhibition of methyl transferase by 5-azacytidine (Fig. 1.42) leads to a change in the methylation pattern and to a reversal of differentiation of a cell culture.
- Methylation of foreign DNA, such as viral DNA, is used as a tool for transcription inactivation.
- There are proteins that specifically bind to methylated CpG sequences. They are termed MeCP1 and MeCP2 and function as transcriptional repressors. MeCp2 is a member of the H1 family of histones and has been shown to direct a histone-deacetylase to chromatin (Nan et al., 1999). Possibly the histone-deacetylase provides a mechanistic link between DNA methylation and transcription repression.
- Transcriptional activity can be modulated by the methylation of CpG sequences near promoters. It is still unclear exactly how methylation in the promoter region influences gene activity.
- DNA methylation participates in genetic imprinting. The term „genetic imprinting" describes a situation where genes are expressed unequally depending upon whether they were maternally or paternally inherited. Normally both copies of the parental genes are equally transcribed in a diploid chromosome. However, with imprinting, a gene inherited from either the mother or father is selectively inactivated. Methylation is obviously involved in such an inactivation. The inactive copy is more strongly methylated than the active copy.

In summary, methylation of CpG sequences reveals itself to be a tool to modulate the activity of promoters as well as a means of long-term repression of genes. Recent studies (review: Razin, 1998) suggest that methylation plays an important role in establishing an inactive state of a gene by targeting a histone deacetylase to the chromatin and rendering the chromatin structure inaccessible to the transcription machinery.

## 1.5 Post-Transcriptional Regulation of Gene Expression

Transcription and translation are spatially separated events in eucaryotes. The product of nuclear transcription is *pre-mRNA*. In order to enable translation, the information contained within the pre-mRNA must be transported out of the nucleus and into the cytosol. The quantity of processed mRNA available for translation decides to a high degree how much protein is formed by de novo synthesis.

From the primary transcript to translated protein there are many possible points for regulatory processes. The most important regulatory points are:

– modification of the 3'-end of the pre-mRNA
– splicing of the pre-mRNA
– transport of the pre-mRNA

– initiation of translation
– stability of the mRNA

The transport from nucleus to cytoplasm is accompanied by modification at the 5'- and 3'-end of the pre-RNA, as well as by processing (splicing) of the primary transcript. The 3'-end modifications and splicing decide which information contained in the primary transcript is made available for protein biosynthesis. The information content of the processed mRNA can be specifically influenced by these processes. This has an important impact on the tissue- and cell-specific protein expression. 3'-modification and splicing are tightly coupled to extranuclear transport. Interventions in the transport process are another possibility for a regulation at the post-transcriptional level.

The translation of the correctly modified mature mRNA by the ribosome is also subject to regulation. The regulatory site of translation is mainly at the initiation of translation. Further regulatory elements include the availability of mRNA for ribosomal protein biosynthesis, as well as the concentration of mRNA. The availability of mRNA can be controlled by, for example, sequence-specific protein binding to the mRNA. The concentration of a specific mRNA is determined by a balance between its rate of synthesis (i.e. transcription) and its rate of degradation by RNases. The stability of a mRNA against nucleolytic degradation is thus a further factor that can determine the extent of biosynthesis of a protein.

## 1.5.1 Modifications at the 5'- and 3'-Ends of the Pre-mRNA

Modifications at the 5'- and 3'-end include the processes of capping and polyadenylation (Fig. 1.44).

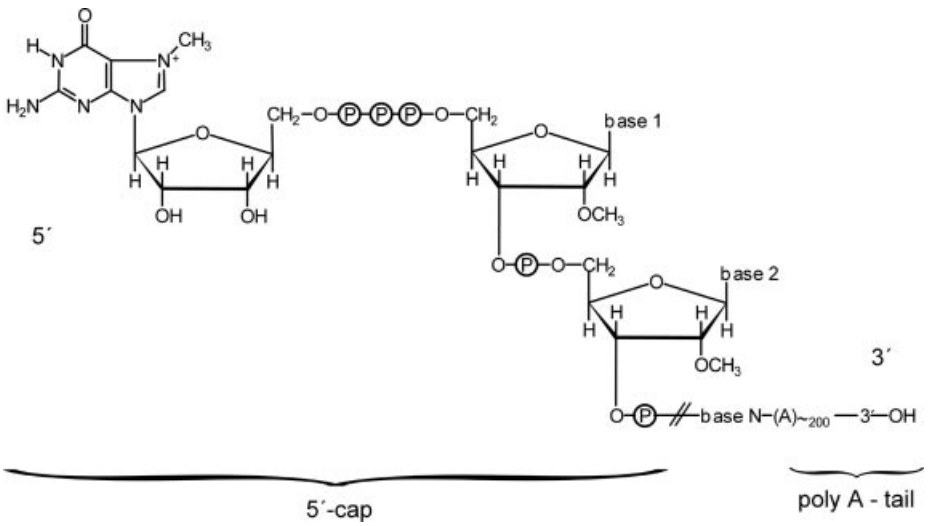



**Fig. 1.44.** Modifications at the 5' and 3' ends of eucaryotic mRNA. Eucaryotic mRNAs possess a „cap structure" at their 5' ends and a 100–200 base long poly-A tail at their 3' ends.

*Capping* at the 5'-end of the pre-mRNA occurs immediately after incorporation of about 30 nucleotides in the primary transcript. The 5' cap structure is required for the binding of the mRNA to the 40S subunit of the ribosome during the initiation of translation. Capping is also ascribed a stabilizing function for mRNA.

In higher eucaryotes the 3'-end of mature mRNA is not produced as a result of termination of transcription. Rather, the 3'-end of the primary transcript is cut at a specific site and a poly-A sequence is appended. Polyadenylation precedes the splicing of the primary transcript.

During *polyadenylation* the primary transcript is shortened in an endonucleolytic step and appended with ca. 200 A-residues. The endonucleolytic incision requires two signal sequences on the pre-mRNA. A highly conserved AAUAAA sequence 10–30 nucleotides upstream from the hydrolysis site serves as one signal. Another signal in the form of a less well conserved GU- or U-rich element upstream of the hydrolysis site. Both together constitute the *polyadenylation signal* (Fig. 1.45). Polyadenylation occurs in a multiprotein complex, whose composition is not yet explained in all details.

*In vitro* experiments show that correct modification of the 3'-end requires at least three protein factors: the CPSF protein, the poly-A polymerase and the poly-A binding protein. The CPSF protein (CPSF: cleavage and polyadenylation specificity factor) binds to the AAUAA signal and brings the poly-A polymerase to the polyadenylation site. The poly-A polymerase is supported by the poly-A binding protein. The latter binds to the poly-A sequence and is required for the transition from the phase of synthesis of short poly-A sequences to the formation of mature poly-A sequences (ca. 200 A-residues).

## 1.5.2 Formation of Alternative mRNA by Alternative Polyadenylation

Some pre-mRNAs carry several polyadenylation signals on their 3'-end. Depending upon which signal is used for the polyadenylation, various mRNAs can be formed from a single primary transcript (Fig. 1.46). The mechanism of alternative polyadenylation offers the possibility to form cell- and tissue-specific mRNAs from the same primary transcript. It is currently unkown how the cell decides which poly-A signal to use. It is speculated that the composition of the polyadenylation complex regulates the efficiency of the various competing pol-A signals (Bienroth et al., 1993). This implicates the tissue- and cell-specific availability of the protein responsible for the polyadenylation as the key regulatory element for the choice of the pol-A tail.

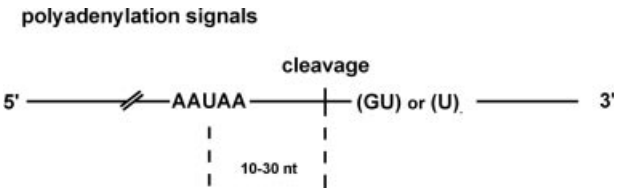



**Fig. 1.45.** Sequence signals for polyadenylation.

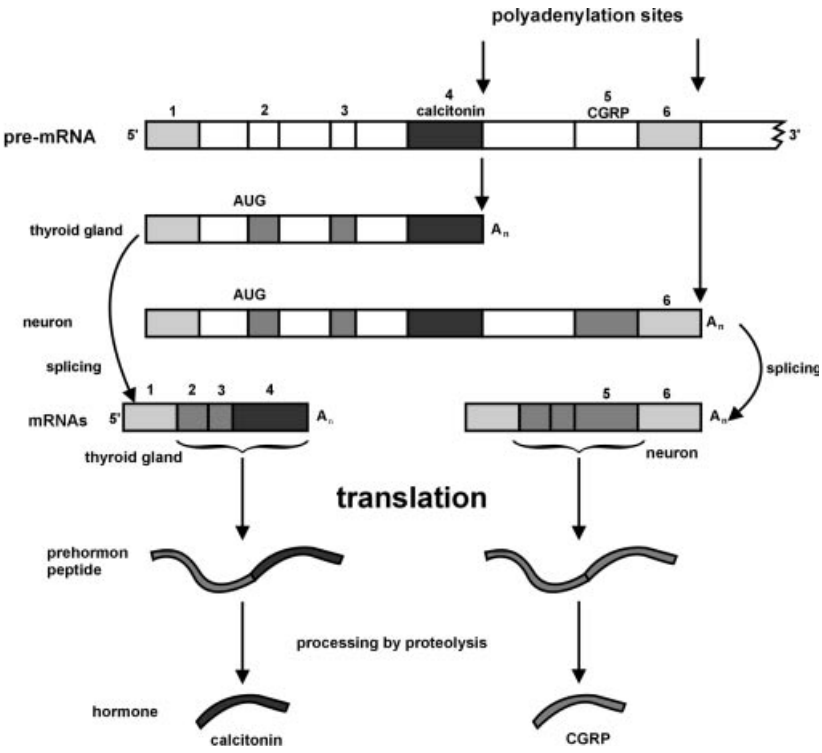



**Fig. 1.46.** Alternative polyadenylation in the expression of calcitonin genes of rat. The primary transcript of the calcitonin gene possesses two polyadenylation sites. One site is used in the processing of RNA in the thyroid, another site in the brain, and yet another in nerve tissue. The translation of the two mRNAs creates two pre-hormones, from which two different polypeptide hormones (calcitonin and the „calcitonin-related peptide“, or CGRP) are created via proteolysis.

## 1.5.3 Alternative Splicing

The genetic information encoding a protein in higher eucaryotes is usually not found in a continuous sequence. Usually the genetic information is found in pieces of coding sequences, or exons, interrupted by non-coding sequences, the introns. For the formation of the mature mRNA, the introns must be excised and the exons rejoined in the proper order. This process is termed splicing. The number of introns in eucaryotic genes can be very large; there are 50 introns in the human dystrophin gene.

Splicing occurs in a large protein-nucleic acid complex, termed the spliceosome. Components of the spliceosome are, apart from the pre-mRNA, a number of proteins and small RNAs, termed the U1, U2, U4, U5 and U6. The RNAs found in the spliceosome are bound to specific proteins. The complexes are termed snRNPs (small nuclear ribonucleoprotein). Depending upon the type of RNA bound, there are U1, U2, U5 and U4/U6 snRNPs.

Based on the observation of self splicing by the 23S RNA of Tetrahymena, it is assumed that the cleavage and rejoining of the phosphodiester bond is catalyzed by the RNA components of the spliceosome. The proteins of the spliceosome are believed to be important for the recognition of the 5' and 3' splice sites and for the formation of a defined structure in the spliceosome. Thus, the proteins of the spliceosome play a decisive role in the choice of the splice site and the effeciency of splicing.

The signal sequences that determine the location of splicing are shown in Fig. 1.47. Decisive are sequences at the border between exon and intron, found as conserved sequences at the 5' (donor site) and 3' (acceptor site) of the splice site. The intron is removed in the form of a lasso. Donor and acceptor sites are then precisely joined together.

The occurrence of several exons and intron in a gene opens the possibility of *alternative splicing*.

Starting with a single pre-mRNA several alternative mRNAs via the rejoining of various exons can be formed, each coding for proteins with different activities and functions. The creation of different proteins due to alternative splicing is a commonly used tool for cell- and tissue-specific expression of proteins. By the formation of variable spliced proteins, a tissue-specific constitutive protein pattern can be created. An example for this is shown in Fig. 1.48. Alternative splicing can also be used to rapidly and flexibly adjust the expression of specific proteins in response to a transduced external signal.

Alternative splicing can create proteins with varying enzymatic activity, cellular localization and regulatory activity.

The splicing of pre-mRNAs often allows several possible combinations of exons. Despite this fact, only a few of the theoretically possible combinations are realized. This accentuates that the choice of alternative splice sites must be strictly controlled. The exact mechanism by which a splice site is chosen is not understood. The following control mechanisms are conceivable (review: Lopez, 1998):

1. The *composition of the spliceosome* determines the pattern of chosen splice sites. Several proteins have been identified which act as antagonists in the selection of splice sites. The SF2 protein (and related proteins) belong to the family of SR proteins (SR: Ser- and Arg-rich) and supports the use of 5'-splice sites. Another protein, the hnRNP A1 protein supports the use of 3'-splice sites.

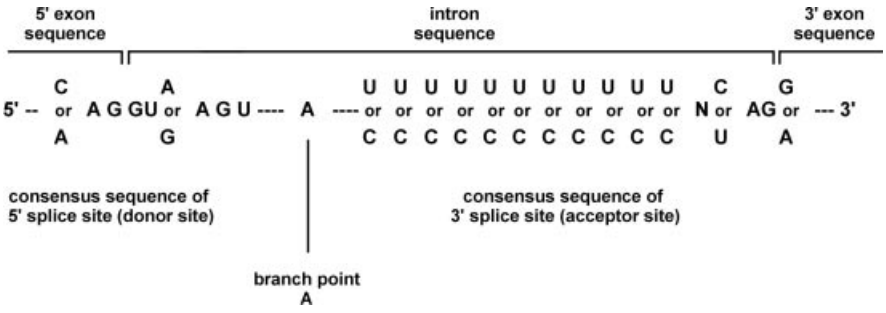



**Fig. 1.47.** Signal sequences for splicing. A splice site is determined by sequences on the 5' side of an intron (5' splice site) and on the 3' side of an intron (3' splice site).

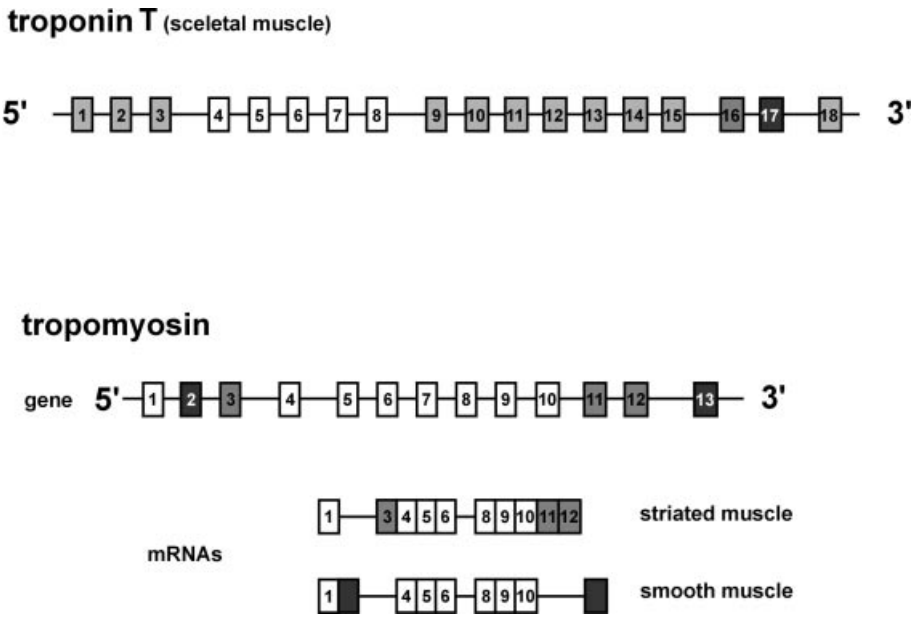



**Fig. 1.48.** Differential splicing in muscle proteins. a) The troponin gene of rats possess 18 exons that encode 258 amino acids. Different subtypes of troponin are found in various types of muscle tissue.The exons 1–3 and 9–15 are found in all subtypes, while the exons 4–8 appear in various combinations, allowing 32 possible combinations. Exon 16 or 17 are found in every subtype. Altogether, 64 different mRNAs can be formed from the troponin pre-mRNA. After Breitbart and Nadal-Grinard (1986). b) Tropomyosin is a muscle protein that can be alternatively spliced in different muscle tissue. Shown in the figure are the predominant subtypes for striated and smooth muscle. After Wieczorek et al., 1989.

2. *Positive and negative regulatory proteins* can determine the selection of the splice site. A positive acting protein would accordingly stimulate the use of a specific splice site, a negative acting protein would inhibit the use of a splice site (Fig. 1.49). Both processes allow for timely control of splicing and can affect a short term change in gene expression in response to external signals.

## 1.5.4 Regulation via Transport and Splicing of pre-mRNA

Splicing and transport of the transcript from the nucleus to the cytosol are tightly coupled. Unspliced pre-mRNA usually does not leave the nucleus, so that only correctly spliced mRNA reaches the cytosol. The human immunodeficiency virus (HIV) is an example of how manipulation of the nucleocytoplasmic transport can be used to create different mRNAs from one pre-mRNA (Cullen and Malim, 1991).

Following integration into the host genome as a provirus, the genome of a retrovirus is similar in many respects to a typical eucaryotic cellular gene. The transcription of the HIV coding DNA, modification and processing of the HIV transcript is performed by

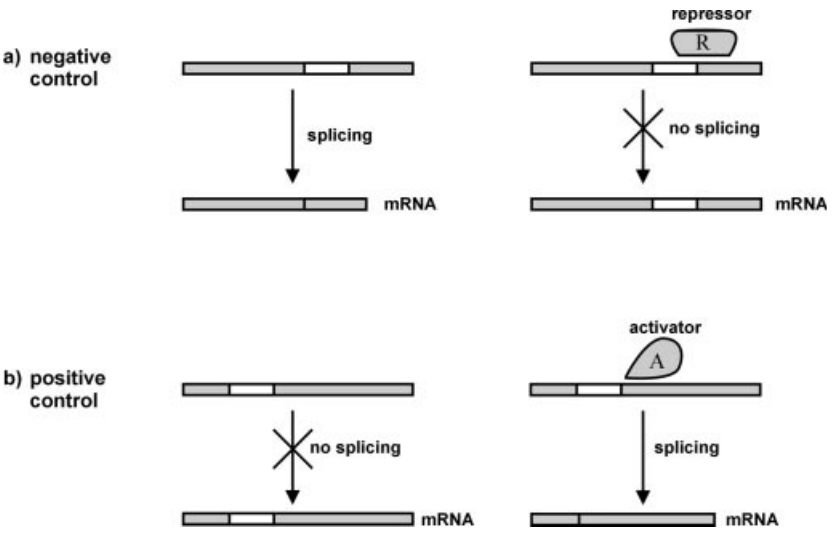



**Fig. 1.49.** Model for the negative and positive control of splicing. a) In negative control, the repressor protein binds to the primary transcript and prevents the excision of an intron. b) In positive control, an activator protein is required to excise an intron.

most of the same protein complexes which process cellular mRNAs. Among the proteins encoded by the HIV genome, some are required at the beginning of the replication cycle, while others are required at a much later phase. The regulatory proteins Tat and Rev are required at the beginning. Only when sufficient levels of HIV RNA have been produced are structural proteins, such as Gag, Pol and Env, required. The primary transcript of the HIV genome is a 9kb long pre-mRNA, out of which further RNAs arise by differential splicing (Fig. 1.50). The remarkable thing about the splicing of HIV RNA is the fact that the splicing and transport of the pre-mRNA is regulated by a viral protein.

A key element of regulation is the Rev protein. In the absence of the Rev protein, one finds a multiply spliced class of 2S RNA in the cytoplasm that encodes the regulatory proteins Tat, Rev and Nef. At this phase the mRNAs which encode structural protein are transported to the cytosol only at a low level. They are furthermore unstable and are only partially loaded with polysomes.

In the presence of the Rev protein, however, mainly longer, unspliced 9S mRNAs and partially spliced viral RNAs are found in the cytosol, from which the expression of the structural proteins occurs. Normally incompletely spliced or unspliced mRNAs are retained in the nucleus and can not be transported to the cytosol.

At the beginning of the replication cycle of HIV, before the Rev protein is present, primarily regulatory proteins are formed. If, in the framework of this expression pattern, enough Rev protein is present, then unspliced or incompletely spliced viral mRNAs appear in the cytosol and structural proteins are formed.

The critical point for the Rev protein is the processing of the 9S and 4S RNA. In the presence of Rev the splicing of long mRNAs to the short 2S RNA is repressed and the unspliced mRNA forms can be transported from the nucleus to the cytosol. The Rev

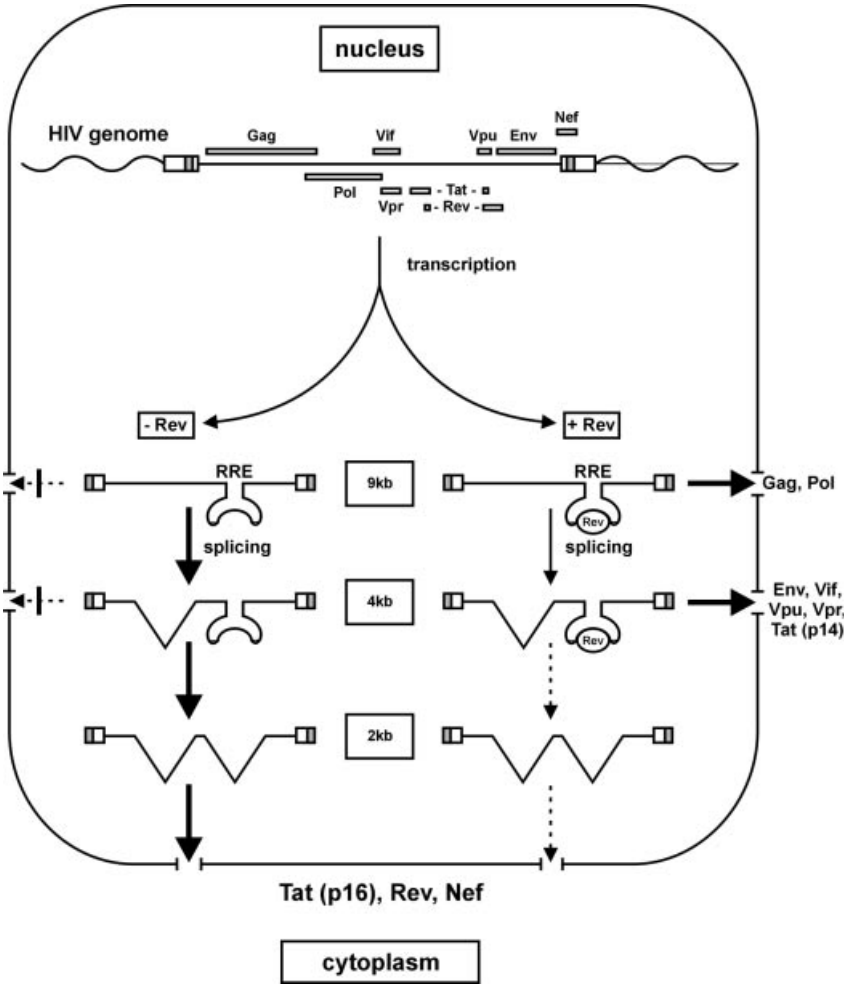



**Fig. 1.50.** The function of the Rev protein in the processing of the transcripts of the Human Immunodeficincy Virus (HIV). The Rev protein is a regulatory protein required for the processing of the primary transcripts of HIV. In the early stadium of viral replication, before Rev protein is available, spliced mRNA of ca. 2kb are transported to the cytosol. The spliced mRNAs are created by multiple splicing events of the primary transcript and encode the regulatory proteins Tat, Rev and Nef. Once a critical level of Rev protein is formed, unspliced (9kb) or only singly spliced (4kb) forms of the viral mRNA appear in the cytosol. These encode for structural proteins and reverse transcriptase. The Rev protein binds to a particular sequence of the viral RNA, the Rev responsive element (RRE). The binding of Rev to the RRE enables the transport of unspliced viral transcript into cytosol.

protein binds to a ca. 230 long RNA element, termed the „Rev Responsive Element; RRE, Fig. 1.51), that includes sequences on the 9S and 4S RNA. The binding of Rev to the RRE obviously allows the transport of the unspliced or partially spliced protein into the cytosol.

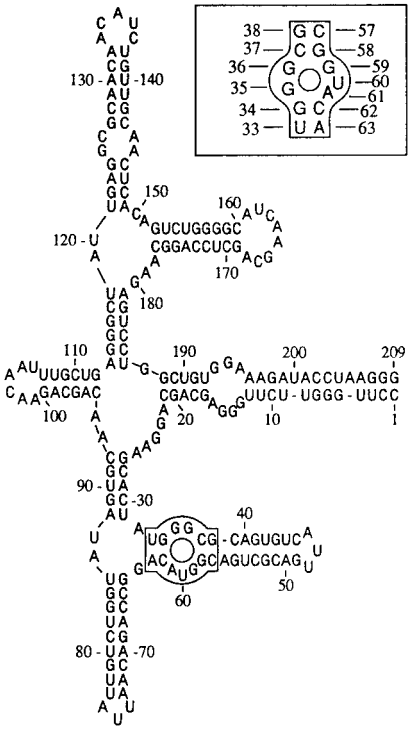


**Fig. 1.51.** Sequence and presumed secondary structure of the „Rev responsive element" of HIV RNA. The framed region is postulated to be the binding site of the Rev protein based on mutation studies. After Heaphy et al., 1991.

The exact mechanism by which Rev interferes with the transport and splicing process is largely unknown. One possible scenario is that Rev interacts with components of the spliceosome, which leads to the release of splice factors that allow cytosolic transport without actual splicing.

## 1.5.5 Stability of the mRNA

The concentration of mRNA available for translation is determined by the *de novo* synthesis, as well as by the degradation rate. The specific degradation of mRNAs plays an important role in cell- and tissue-specific gene expression (review: Sachs, 1993). The stability of various mRNAs can vary significantly: from 20 min to 24 hour half-life in the same cell.

The stability of a mRNA is determined from cis-acting elements, as well as by trans-acting protein factors. Destabilizing sequence elements were identified by site-specific changes in the mRNA sequences. They can be found at the 5' end, in the coding region, or at the 3' end. It can be generally noted that the degradation of mRNA by nucleases is coupled with translation. It is suspected that the proteins required for degradation are brought to the RNA by the translation apparatus. The nature of these proteins is not yet clear.

The following discussion gives examples of control elements important for the stability and degradation of mRNA.

### Destabilizing Sequences at the 3'-Non-Translated End

Tubulin is a well-studied example of the interference of gene products with the stability of mRNA. Processed tubulin binds as a dimer to the growing tubulin chain on the ribosome. The binding of the tubulin facilitates the attack by endonucleases on the ribosome-bound tubulin mRNA and thereby initiates the degradation of the mRNA (Fig. 1.52). The goal of this regulation process is to prevent the formation of excess tubulin. If tubulin is in excess, then the degradation of its own mRNA is induced and the synthesis of more tubulin is prevented.

### Regulation of mRNA Stability by Iron

Two proteins are important for iron metabolism in mammalian cells: the *transferrin receptor* (TFR) and *ferritin*. Ferritin is a protein for the storage of iron. The production and its level is increased when more iron is available.

TFR is instrumental in the uptake of iron under conditions of low iron concentration. The concentration of cellular TFR is inversely correlated with the level of iron: if little iron is present, then the TFR concentration is increased; if high levels of iron are present, then the TFR concentration is decreased.

The regulation of the TFR and ferritin concentrations occurs at the mRNA level for both proteins (review: Klausner et al., 1993; Hentze and Kühn, 1996). The key element for the regulation of the TFR concentration is a region at the 3'- non-translated end of

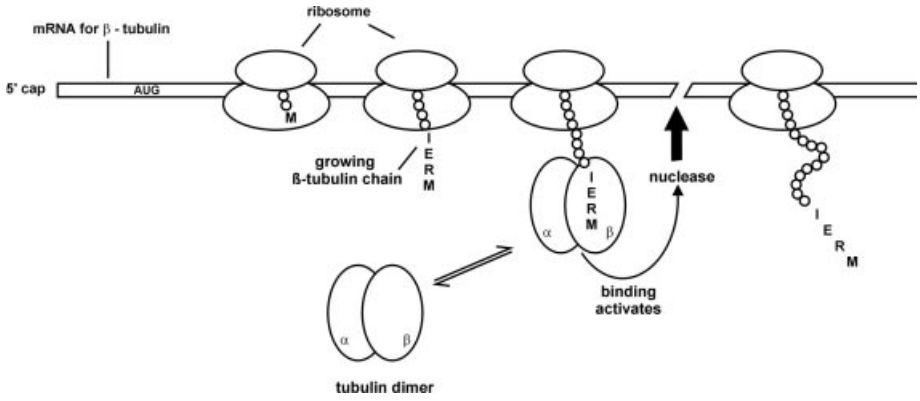



**Fig. 1.52.** Model for the control of translation by tubulin. The amount of tubulin in animal cells is determined partially by the stability of β-tubulin mRNA, whereby tubulin itself acts as the regulating signal. Starting from the 5' cap, various stages of the translation of β-tubulin mRNA, represented as a chain of small circles, is illustrated in the figure. As soon as the N-terminus of the growing β-chain emerges from the ribosome, the α- and β- subunits of tubulin bind to the terminal MREI sequence, upon which an endonuclease becomes activated by a presently unknown mechanism. The degradation of the β-tubulin mRNA then proceeds.

the TFR mRNA. Here are found several hairpin structures essential for the stability of the mRNA. The hairpin structures offer binding sites for a specific RNA binding protein, the iron responsive element binding protein (IRE-BP). Binding of IRE-BP to the hairpin structure protects the mRNA from nucleolytic degradation (Fig. 1.53).

The binding capacity of IRE-BP to the hairpin structures is controlled by the amount of iron. Low iron concentrations favor the formation of a binding competent form, high iron concentrations favor the formation of a binding incompetent form of IRE-BP. Both forms of the IRE-BP differ from each other in terms of their content of a 4Fe-4S cluster. Iron favors the insertion of the 4Fe-4S cluster into the protein and thereby transmits it into the binding incompetent state. In the presence of high levels of iron, the hairpin structures are not occupied, the mRNA can be degraded, and the level of TFR drops.

The regulation of the ferritin concentration is also related to the iron concentration. The vulnerable point is not the stability of the mRNA, but rather of the initiation of translation. The mRNA for ferritin possesses a hairpin structure in the 5'- non-coding

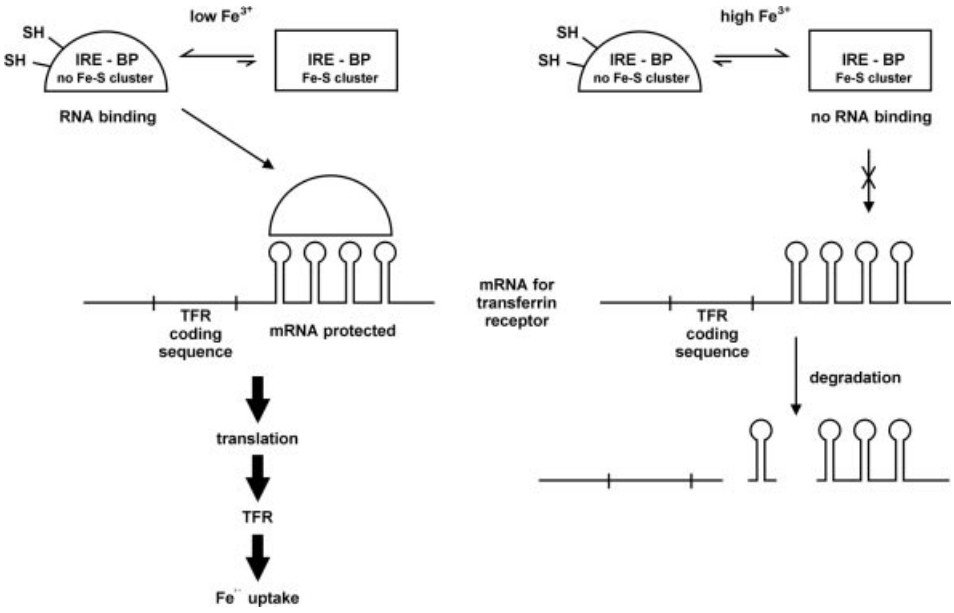



**Fig. 1.53.** Regulation of the stability of the mRNA of the transferrin receptor by $Fe^{3+}$. The translation of the mRNA of the transferrin receptor (TFR) is subject to regulation by the Fe concentration. Fe exerts its regulatory effect via the IRE binding protein (IRE-BP). The IRE-BP binds to a control segment at the 3' terminal region of the TFR mRNA, known as the iron responsive element (IRE). Binding of the IRE-BP to a hairpin structure of the IRE stabilizes the mRNA of the transferrin receptor and protects it agains nuclease attack. If high levels of Fe are present, then the IRE-BP is in its binding inactive form, the IRE is free and nucleases can attack the non-translated 3' end of the mRNA. Upon depletion of Fe, the IRE-BP is found in its binding compatible form and it protects the mRNA of the transferrin receptor from degradation. Binding-active and inactive forms of the IRE-BP differ in the content of Fe-S clusters.

region similar to that found in the TFR receptor at the 3'-end. The IRE-BP binds in its active form to the hairpin structure of the ferritin mRNA and blocks accessibility of the ribosome to the coding sequence. The translation of ferritin mRNA is halted under this mechanism upon low levels of iron, and the amount of ferritin decreases. At high iron concentrations the IRE-BP is found in its binding incompetent form, the 5' regions is free to be translated and de novo synthesis of ferritin for the storage of iron is possible.

## 1.5.6 Regulation at the Level of Translation

In order to allow a better understanding of regulation at the level of translation, some of the specific features of eucaryotic translation will be summarized briefly.

As opposed to the case in procaryotes, eucaryotic translation does not require a specific sequence for the binding of the ribosome. Procaryotes rely on the *Shine-Dalgarno sequence*, which is complementary to sequences of the 16S RNA of the 30S subunit. The Shine-Dalgarno sequence mediates the binding of mRNA to the 30S ribosome and ensures the correct positioning of the AUG initiation codon.

The mRNA of eucaryotes does not possess specific initiation sequences. Rather, the AUG start codon is identified by scanning the eucaryotic mRNA: the 40S subunit of the ribosome threads the 5' non-translated end of the mRNA and uses the first AUG codon encountered to initiate translation. Whether a AUG codon is used as an initiator depends, additionally, upon the sequence context. If the sequence environment is unfavorable for initiation, then the scanning is continued and initiation occurs at one of the next AUG. With the help of this „leaky scanning" strategy it is possible to produce proteins with different N-termini from the same mRNA. Since there are often signal sequences found at the N-terminus, this mechanism may lead to alternative compartmentalization of a protein.

Eucaryotic translation is controlled, analogously to transcription, primarily via initiation. Regulation by attenuation, an elementary regulation mechanism in procaryotes, is unknown to eucaryotes. Regulation by attenuation demands a tight coupling of transcription and translation. In eucaryotes such a coupling does not exist, since transcription and translation are spatially separated. As discussed above, however, eucaryotes posses an extensive repertoire of regulation possibilities in the framework of transport and processing of mRNA.

### 1.5.6.1 Regulation by Binding of Protein to the 5'-End of the mRNA

As shown in the example of ferritin mRNA , the binding of the ribosome to the 5'-end of the mRNA can be blocked by RNA binding proteins at the same region. This case is exemplary for a negative regulation of translation by sequence-specific RNA binding proteins. Such proteins are categorized as translation repressors (see Fig. 1.54). The repressor proteins can either be directly or indirectly related to the protein encoded by the regulated mRNA, thus allowing adaptation of the translation process to the current needs by a feedback mechanism. Often the gene product acts as a repressor of

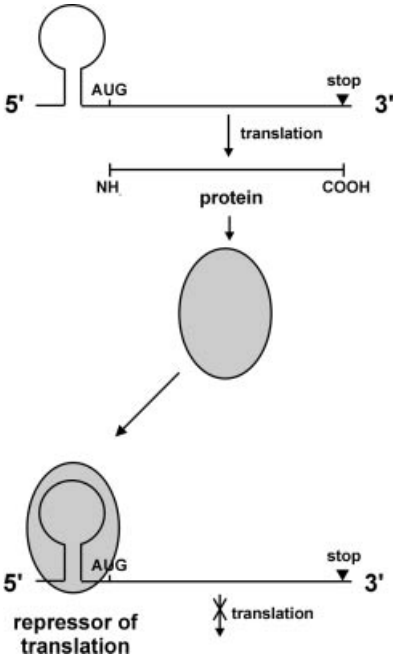



**Fig. 1.54.** Principle of negative control of translation initiation by protein binding to mRNA. Proteins can negatively effect translation by binding to the sequences in the 5' non-translated region of their own or other mRNAs. The participating proteins are sequence-specific RNA binding proteins and recognize RNA sequences in hairpin structures or other secondary structures of RNA. The protein binding interferes with the scanning of ribosomes and prevents the translation of mRNA.

its own translation. This mechanism is utilized by procaryotes to regulate the synthesis of ribosomal proteins: excess ribosomal proteins block the translation of its own mRNA.

An interesting case of regulation at the translation level was discovered on the example of the homeodomain protein *bicoid* (bcd), which is important in *Drosophila* differentiation (Dubnau and Struhl, 1996). The bicoid protein is a transcriptional activator that binds a cognate DNA element and stimulates the transcription of the neighboring genes. Apart from its specific DNA-binding capability, the bicoid protein binds to 3'-, non-translating sequences of the mRNA of another homeodomain protein (caudal protein) to inhibit its translation.

### 1.5.6.2 Regulation by Modification of Initiation Factors

The de novo synthesis of proteins can be varied in response to external stimuli, such as hormones or heat stress. The regulation of protein biosynthesis occurs primarily via phosphorylation of translation initiation factors. The regulatory points in eucaryotes are, above all, the translation factors eIF-2 and eIF-4.

The function of eIF-2 is illustrated schematically in Fig. 1.55. eIF-2 belongs to the superfamily of regulatory GTPases (see ch. 5). eIF-2 fulfills the task of bringing the methionyl-initiator-tRNA to the 40S subunit of the ribosome. The active eIF-2 *GTP form binds the methionyl-initiator-tRNA, associates with the cap structure of the mRNA, then commences to scan along the mRNA. Once an AUG codon is encountered, the bound GTP is hydrolyzed to GDP, resulting in the dissociation of the

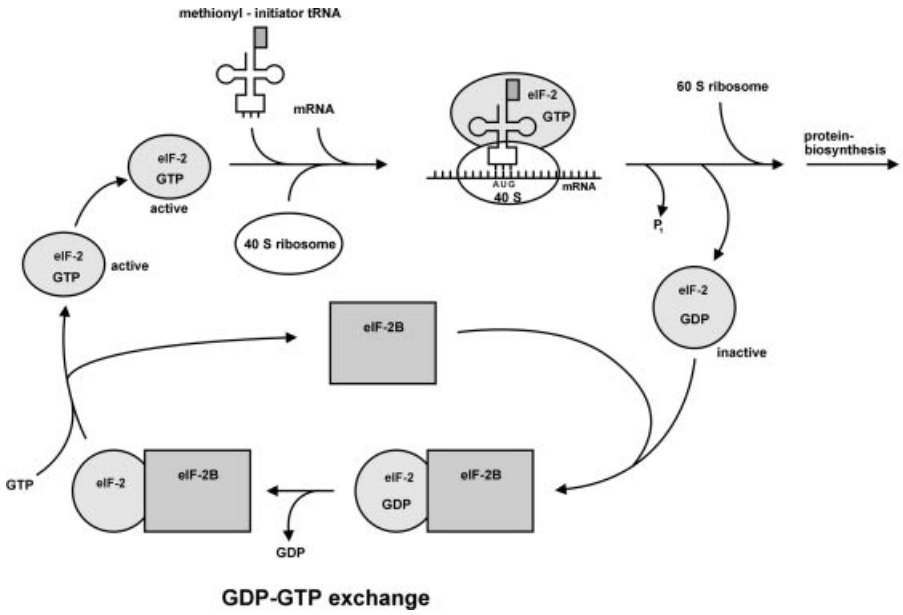



**Fig. 1.55.** The function of eIF-2 in eucaryotic translation. eIF-2, the initiator protein for the translation is a regulatory GTPase that occurs in an active GTP-form and in an inactive GDP form (see ch. 5). The active eIF-2*GTP forms a complex with the initiator-tRNA, fMet-tRNA$^{fmet}$ and the 40S subunit of the ribosome. This complex binds to the cap structure of mRNA to initiate the scanning process. eIF-2 undergoes an activation cycle typical for regulatory GTPases: the inactive eIF-2*GDP fom is activated with the assistance of the eIF-2B protein into the active eIF-2*GTP form. eIF-2B acts as a G-nucleotide exchange factor in the cycle (see ch. 5).

eIF-2*GDP from the 40S ribosome. The transition from the inactive eIF-2*GDP form into the active eIF-2*GTP form requires a G-nucleotide exchange factor, termed the eIF-2B.

The transition of the inactive eIF2*GDP into the active GTP form is subject to regulation by phosphorylation of the α-subunit of eIF-2. Phosphorylation of the α-subunit of eIF-2 increases its affinity for the nucleotide exchange factor eIF-2B, without inducing nucleotide exchange. eIF-2 is found in excess of eIF-2B in the cell, so that phosphorylated eIF-2 binds the entire reservoir of eIF-2B. As a consequence, no further eIF-2B is available for nucleotide exchange, and protein biosynthesis is halted.

The translation of certain mRNAs, including that encoding the transcription factor GCN4 of yeast, is stimulated rather than inhibited according to the above mechanism. In this case there exists a complex interaction between several initiation sites.

eIF-2 is phosphorylated by the specific protein kinase eIF-2α (review: Samuel, 1993; Wek, 1994). The eIF-2α kinase is activated by hormonal signal transduction pathways or via external stimuli, such as lack of nutrients, heat shock or viral infection.

A well studied example for control at the level of eIF-2 is the regulation of protein biosynthesis in erythroid cells (review: Chen and London, 1995). A decrease in the heme concentration in reticulocytes leads to inhibition of globin synthesis at the level

of translation. This regulation mechanism ensures that only so much globin is produced as is heme available. If the level of heme drops, then the heme-regulated kinase (HRI) becomes activated. The activated HRI phosphorylates the eIF-2α subunit, which in turn shuts off protein biosynthesis (Fig. 1.56).

The mechanism of regulation of HRI kinase by heme is not well understood. Inactive and active forms of the HRI kinase differ in the content of intermolecular disulfide bridges of the HRI kinase dimer. However it is still unclear how heme targets the kinase and whether the disulfide bridges directly influence the kinase activity.

Mammalian cells contain at least one additional eIF-2-specific kinase, known as the RNA-specific eIF-2 kinase (abbreviated PKR). PKR is induced by interferon and can be activated by minimal concentrations of double-stranded RNA. Activated PKR phosphorylates eIF-2, which, according to the above mechanism, switches off protein biosynthesis.

The above mentioned examples illustrate the variety of regulatory mechanisms under the control of eIF-2. Together, they play a central role in the control of initiation of translation in eucaryotes.

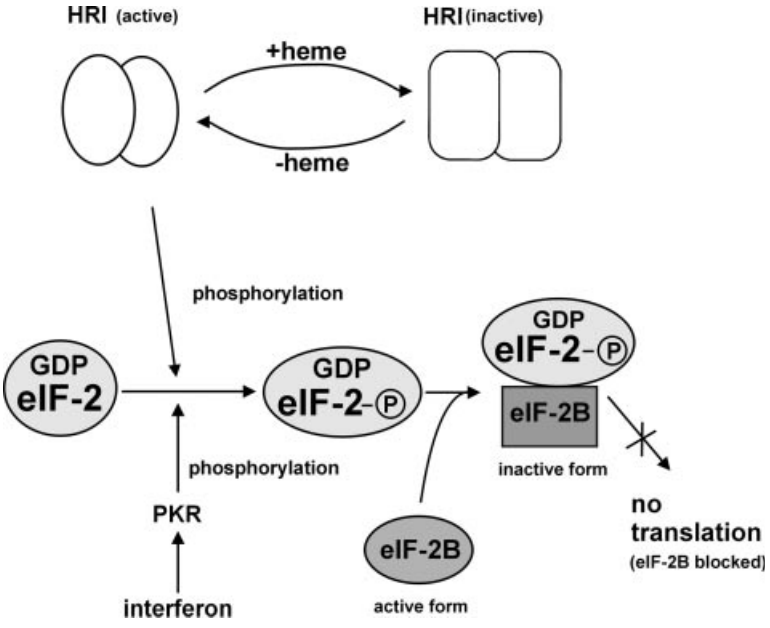



**Fig. 1.56.** Control of eIF-2 by phosphorylation. Phosphorylated eIF-2*GDP binds strongly to eIF-2B without nucleotide exchange occurring. Initiation of protein biosynthesis is not possible in this case.In reticulocytes, eIF-2 is subject to phosphorylation by the heme-regulated eIF-2-kinase (HRI). The activity of the dimeric HRI is regulated via the heme concentration. Another protein kinase that can phosphorylate and regulate eIF-2 is the RNA-dependent eIF2α-kinase (PKR). The latter is induced by interferons and activated by double stranded RNA.

### 1.5.6.3 Regulation of Translation via Insulin

It has long been known that insulin ( and other hormones and growth factors) can stimulate protein biosynthesis. The signal transduction pathway linking insulin to the translation apparatus was, until recently, unclear. However, insight is now being gained into this mechanism (Pause et al, 1994, review: Proud and Denton, 1997).

The vulnerable point for insulin-mediated regulation of translation is the initiation factor eIF-4E. This factor binds specifically to the 5'-cap structure of mRNA and is part of a larger complex, termed eIF-4F. A further component of eIF-4F is the eIF-4A protein, which possesses helicase activity. The binding of eIF-4E to the cap structure is necessary for the association of the 40S subunit with the 5'-end of the mRNA and for

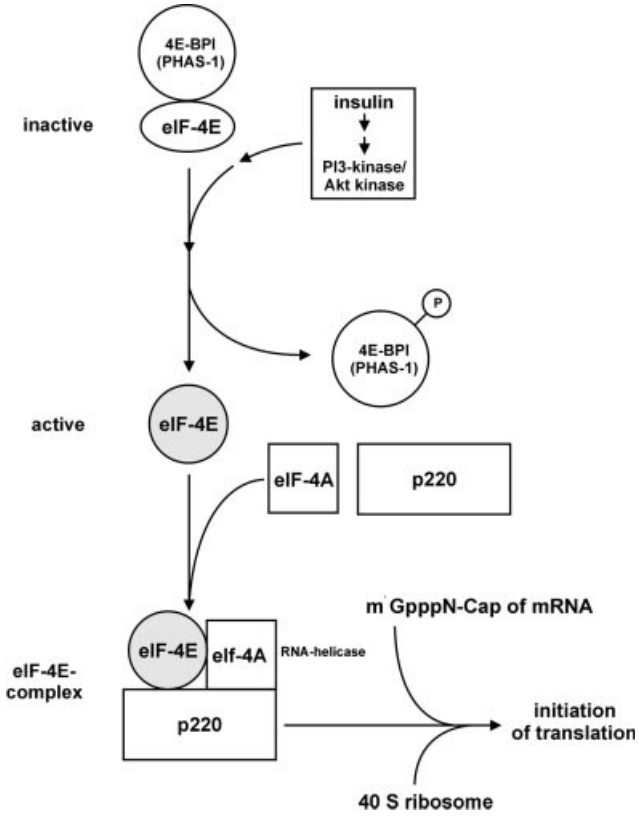



**Fig. 1.57.** Model of the regulation of translation by insulin. Insulin ( and other growth factors) activates the Akt kinase pathway (see ch. 10), whose final result is the phosphorylation of 4E-BP1, a regulatory protein of translation initiation. The 4E-BP1 protein inactivates the initation factor eIF-4E by complex formation. eIF-4E is required, together with the proteins eIF-4A and p220, for the binding of the 40S subunit of the ribosome to the cap structure of the mRNA. If the 4E-BP1 protein becomes phosphorylated as a result of insulin-mediated activation of the PI3 kinase/Akt kinase cascade, then eIF-4E is liberated from the inactive eIF-4E*4E-BP1 complex and protein biosynthesis can begin.

the subsequent scanning by the 40S subunit. It is assumed that the RNA helicase activity of eIF-4A is required to melt the secondary structure at the 5'-end of the mRNA. The activity of eIF-4E is controlled by a specific binding protein, 4E-BP1. The 4E-BP1 protein binds and inactivates eIF-4E, so that the latter is no longer available to bind the 5' cap structure and translation of mRNA is inhibited.

Under the influence of insulin a signaling chain is activated that results in the phosphorylation of 4E-BP1. The responsible protein kinase is most probably the serine/threonine kinase Akt/PKB (see 6.6.2).

It is assumed that the phosphorylation of 4E-BP1 represents the signal for the release of eIF-4E from the complex. The free eIF4E can then associate with eIF-4A and the 5'cap of the mRNA to initiate translation (Fig. 1.57).

The regulation of translation is accomplished in this system via a specific inhibitory protein and an initiation factor of translation. The binding activity of the inhibitor protein is regulated by protein phosphorylation, and thus, by protein kinases. The activity of protein kinases can be regulated in a multitude of ways. A signal initiating from insulin, for example, can activate the PI3-kinase and the Akt kinase pathway (see 6.6.2), resulting in phosphorylation of 4E-BP1.

A further susceptible point for insulin-regulated signaling pathways is the ribosomal protein S6. Under the influence of insulin, S6 is phosphorylated by a specific protein kinase, the $p70^{S6}$ kinase, resulting in increased levels of translation of certain mRNAs. Several pathways including the MAPK/ERK pathway (see chapter 10) and the Akt kinase pathway can contribute to the activation of the $p70^{S6}$ kinase.

# References Chapter 1

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. 'Molecular Biology of the Cell' (1994) 3rd edition, Garland Publish. Inc. pp 419

Alleman,R.K. and Egli, M. 'DNA recognition and bending' (1997) Chem. Biol. *4*, 643–650

Berg, J.M. 'Zinc-finger proteins' (1993) Curr. Op. Struct. Biol. *3*, 11–16

Bienroth, S., Keller, W. and Wahle, E. 'Assembly of a processive messenger RNA polyadenylation complex' (1993) EMBO J. *12*, 585–94

Bird,A. 'The essentials of DNA methylation' (1992) Cell *70*, 5–8

Björklund, S and Kim, Y.-J. 'Mediator of transcriptional regulation' (1996) Trends Biochem. Sci. *21*, 335–337

Burley, S.K. 'DNA-binding motifs from eukaryotic transcription factors' (1994) Curr. Op. Struct. Biol. *4*, 3–11

Burley, S.K. and Roeder, R.G. 'Biochemistry and structural biology of transcription factor IID (TFIID)' (1996) Annu Rev Biochem. *65*, 769–799.

Carrion, A.M., Link, W.A., Ledo, F., Mellstrom, B. and Naranjo, J.R. 'DREAM is a $Ca^{2+}$-regulated transcriptional repressor' (1999) Nature. *398*; 80–4.

Chen, J.J. and London, I.M. 'Regulation of protein synthesis by heme-regulated eIF-2α kinase' (1995) Trends Biochem. Sci. *20*, 105–108

C–oté, J., Quinn, J. Workman, J.W. and Peterson, C.L. 'Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex' (1994) Science *265*, 53–60

Cowell, I.G. 'Repression versus activation in the control of gene expression' (1994) Trends Biochem. Sci. *19*, 38–42

Cullen, B.R. and Malim, M.H. 'The HIV-1 Rev-protein: prototype of a novel class of eukaryotic post-transcriptional regulators' (1991) Trends Biochem. Sci. *16*, 346–350

Drapkin, R. and Reinberg, D. 'The multifunctional TFIIH complex and transcriptional control' (1994) Trends Biochem. Sci. *19*, 504–508

Driever, W. and Nüsslein-Volhard, C. 'The bicoid protein is a positive regulator of *hunchback* transcription in the early *Drosophila* embryo' (1989) Nature *337*, 138–143

Dubnau,J. and Struhl, G. 'RNA recognition and translational regulation by a homeodomain protein' (1996) Nature *379*, 694–699

Eick, D., Wedel,A. and Heumann, H. ' From initiation to elongation: comparison of transcription by prokaryotic and eukaryotic RNA polymerases' (1994) Trends Gen. *10*, 292–296

Ellenberger, T. 'Getting a grip on DNA recognition: structures of the basic region leucine zipper, and the basic region helix-loop-helix DNA-binding domains' (1994) Curr. Op. Struct. Biol. *4*,12–21

Ghosh, G., Van Duyne, G., Ghosh, S. and Sigler, P.B. 'Structure of the NF-ϰB p50 homodimer bound to a ϰB site' (1995) Nature *373*, 303–310

Guarente, L. 'Transcriptional coactivators in yeast and beyond' (1995) Trends Biochem. Sci. *20*, 517–521

Harrington, R.E. and Winicov, I. ' New concepts in protein-DNA recognition: sequence directed DNA bending and flexibility' (1994) Progr. Nucl. Acid Res. Mol. Biol. *47*, 195–270

Harrison, S.C. 'A structural taxonomy of DNA binding domains' (1991) Nature *353*, 715–719

Hassig, C.A. and Schreiber, S.L. 'Nuclear histone acetylases and deacetylases and transcriptional regulation: HATs off to HDACs' (1997) Curr Opin Chem Biol. 1, 300–308.

Heaphy, S., Finch, J.T., Gait, M.J., Karn, J. and Singh, M. 'Human Immunodeficiency virus type-1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich 'bubble' located within the rev responsive element' (1991) Proc. Natl. Acad. Sci. U.S.A. *88*, 683–688

Hegde, R.S., Grossman, S.R., Laimins, L.A. and Sigler, P.B. 'Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target' (1992) Nature *359*, 505–12

Hentze, M.W. and Kühn, L.C. 'Molecular control of vertebrate iron metabolism: mRNA-based regulatory circuits operated by iron, nitric oxide, and oxidative stress' (1996) Proc Natl Acad Sci U S A. *93*, 8175–82.

Horowitz, D.S. and Krainer, A.R. 'Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing.' (1994) Trends Genet *10*, 100–6

Johnson, A.D. 'The price of repression' (1995) Cell *81*, 655–658

Jordan, S.R. and Pabo, C.O. 'Structure of the lambda complex at 2.5 Å resolution: details of the repressor-operator interactions' (1988) Science *242*, 893–899

Kadonaga, J.T. 'Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines' (1998) Cell *92*, 307–13.

Kaffman A., Rank N.M., O'Neill E.Mm, Huang L.S. and O'Shea E.K. 'The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus' (1998) Nature *396*, 482–6.

Kaiser, K. and Meisterernst, M. ' The human general co-factors' (1996) Trends Biochem. Sci. *21*, 342–345

Kim, Y., Geiger, J.H., Hahn, S. and Sigler, P.B. 'Crystal structure of a yeast TBP/DNA complex' (1993) Nature *365*, 512–520

Kissinger, C.R., Liu, B.S., Martin-Blanco, E. Kornberg, T.B., and Pabo, C.O. 'Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions' (1990) Cell *63*, 579–90

Klausner, R.D., Roualt, T. and Harford,J .B. 'Regulating the fate of mRNA: the control of cellular iron metabolism' (1993) Cell *72*, 19–28

Koleske, A.J. and Young, R.A. 'The RNA polymerase II holoenzyme and its implications for gene control' (1995) Trends Biochem. Sci. *20*, 113–116

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Lamb, P. and McKnight, S.L. 'Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization' (1991) Trends Biochem. Sci. *16*, 417–422

Lewin, B. 'Chromatin and gene expression: constant questions, but changing answers' (1994) Cell *79*, 397–406

Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G. and Lu ,P. 'Crystal structure of the lactose operon repressor and its complexes with DNA and inducer' (1996) Science *271*, 1247–1254

Lindahl, T., Sedgwick, B., Sekiguchi, M. and Nakabeppu, Y. Regulation and expression of the adaptive response to alkylating agents' (1988) Ann. Rev. Biochem. *57*, 133–157

Löwe, J., Stock,D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. 'Crystal structure of the 20S proteasome from the archeon T. acidophilum at 3.4 Å resolution' (1995) Science *268*, 533–539

Lopez A.J. 'Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation' (1998) Annu Rev Genet. *32*, 279–305

Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. and Sigler, P.B. 'Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA' (1991) Nature *352*, 497–505

Nan X, Ng H.H., Johnson C.A., Laherty C.D., Turner B.M., Eisenman R.N. and Bird A. 'Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex' (1998) Nature *393*, 386–389

Nikolov, D.B. and Burley, S.K. 'RNA polymerase II transcription initiation: a structural view' (1997) Proc Natl Acad Sci U S A 94, 15–22.

Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F. and Sigler, P.B. 'Crystal structure of trp repressor/operator complex at atomic resolution' (1988) Nature *335*, 321–329

Pabo, S. and Sauer, R.T. 'Transcription factors: structural families and principles of DNA recognition' (1992) Annu. Rev. Biochem. *61*, 1053–1095

Pause, A., Belsham, G.J., Gingras, A.C., Donzé, O., Lin, T., Lawrence, J.C. and Sonenberg, N. 'Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function' (1994) Nature *371*, 762–767

Pavletich, N.P. and Pabo, S. 'Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1Å' (1991) Science *252*, 809–814

Proud, C.G. and Denton, R.M. ' Molecular mechanism for the control of translation by insulin' (1997) Biochem. J. *328*, 329–341

Ptashne, M. and Gann, A. 'Transcriptional activation by recruitment' (1997) Nature *386*, 569–77

Raumann, B.E., Brown, M.B. and Sauer, R.T. 'Major Groove recognition by β-sheets: the ribbon-helix-helix-family of gene regulatory proteins' (1994) Current Op. Struct. Biol. *4*, 36–43

Razin, A. 'CpG methylation, chromatin structure and gene silencing-a three-way connection' (1998) EMBO J. *17*, 4905–8

Roberts, S.G.E. and Green, M.R. 'Activator-induced conformational change in general transcription factor TFIIB' (1994) Nature *371*, 717–720

Roeder, R.G. ' The role of general initiation factors in transcription by RNA polymerase II' (1996) Trends Biochem. Sci. *21*, 327–335

Rosenfeld, M.G., Amara, S.G. and Evans, R.M. 'Alternative RNA processing events as a critical developmental regulatory strategy in neuroendocrine gene expression' (1984) Biochem. Soc. Symp. *49*, 27–44

Sachs, B. 'Messenger RNA degradation in eukaryotes' (1993) Cell *74*, 413–421

Samuel, C.E. 'The eIF-2α protein kinases, regulators of translation in eukaryotes from yeast to humans' (1993) J. Biol. Chem. *268*, 7603–7606

Schleif, R. 'DNA looping' (1992) Annu. Rev. Biochem. *61,* 199–223

Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O. and Reinberg, D. 'Cdk-activating kinase complex is a component of human transcription factor TFIIH' (1995) Nature *374*, 283–287

Strauss, J.K. and Maher, L.J. 'DNA bending by asymmetric phosphate neutralization' (1994) Science *266*, 1829–1833

Struhl, K. and Moqtaderi, Z. 'The TAFs in the HAT' (1998) Cell *94*, 1–4

Svjestrup, J.Q., Wang, Z., Feaver, W.J., Wu, X., Bushnell, D.A., Donahue, T.F., Friedberg, E.C. and Kornberg, R.D. ' Different forms of TFIIH for transcription and DNA repair: Holo-TFIIH and a nucleotide excision repairosome' (1995) Cell *80*, 21–28

Svjestrup,J.Q., Vichi,P. and Egli,J.-M. 'The multiple roles of transcription/repair factor TFIIH' (1996) Trends Biochem.Sci. 21, 346–350

Thanos, D. and Maniatis, T. ' NF-κB. a lesson in family values' (1995) Cell *80*, 529–532

Verrijzer, C.P. and Tijan, R. 'TAFs mediate transcriptional activation and promoter selectivity' (1996) Trends Biochem. Sci. *21*, 338–342

Wek, R.C. 'eIF-2 kinases: regulators of general and gene-specific translation initiation' (1994) Trends Biochem. Sci. *19*, 491–496

Wolberger, C. ' Homeodomain interactions' (1996) Curr. Op. Struct. Biol. *6*, 62–68

Workman, J.L. and Kingston, R.E. 'Alteration of nucleosome structure as a mechanism of transcriptional regulation' (1998) Annu Rev Biochem. *67*, 545–79.

Wieczorek, D.F., Smith, C.W. and Nadal-Ginard, B. 'The rat alpha-tropomyosin gene generates a minimum of six different mRNAs coding for striated, smooth, and nonmuscle isoforms by alternative splicing' (1988) Mol. Cell Biol. *8*, 679–94

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 2**

# The Regulation of Enzyme Activity

## 2.1 Enzymes as Catalysts

Enzymes function as biocatalysts and, as such, are involved in all metabolic reactions. Characteristic for enzymes is their high *efficiency*, high *specificity* and extreme *stereoselectivity*, as well as their *ability to be regulated.* Analogous to chemical catalysts, enzymes do not alter the equilibrium of a reaction, but only accelerate the establishment of the equilibrium of the reaction.

The mechanism of action of enzymes can best be described with the aid of transition state theory. On the pathway from substrate A to product B in a reaction catalyzed by a chemical catalyst or an enzyme A passes through a transition state $A^{\#}$ which is found at the highest point of the energy diagram (fig. 2.1). The energy difference between the ground state of A and the transition state $A^{\#}$ represents the activation energy. The transition state as such can not be isolated. It is the state of A in which the bonds

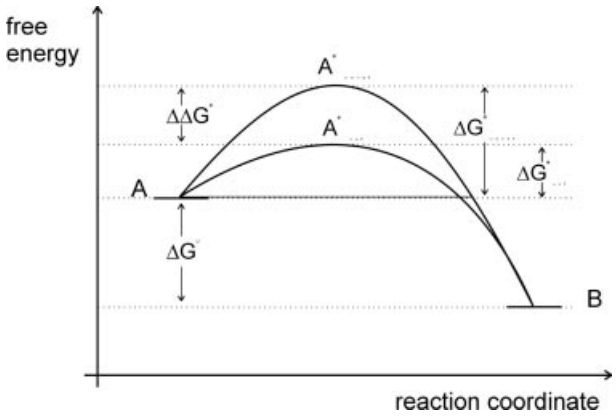



**Fig. 2.1.** The energy profile of a catalyzed and an uncatalyzed reaction. The figure shows the energy diagram for the conversion of A → B for a catalyzed and an uncatalyzed reaction. The binding of A to the catalyst (=enzyme) is left out to simplify the figure. In the uncatalyzed reaction, the energy difference between the ground state A and the transition state $A^{\#}$ is the activation energy $\Delta G^{\#}_{uncat}$. The transition state of the catalyzed reaction is at a lower free energy, so that the activation energy $\Delta G^{\#}_{cat}$ is less and the reaction proceeds at a faster rate. The energy difference $\Delta\Delta G^{\#}$ is a measure of the efficiency of the catalyst and determines how much faster the catalyzed reaction is compared to the uncatalyzed reaction. The equilibrium of the reaction, which is characterized by $\Delta G_0$, remains unchanged upon catalysis.

participating in the reaction are in the process of opening and closing. The transition state is the most unstable state on the path from substrate to product. Enzymes, like chemical catalysts, increase the rate of a reaction by decreasing the activation energy for the transition from A → B.

For the tight binding of the transition state the binding surface of the enzyme must be complementary to the structure of the transition state, so that optimal interactions between the enzyme and the transition state are possible. This demand implies that enzymes display a high affinity to molecules which are chemically similar to the transition state of the reaction. Complexes of such transition state analogues with enzymes are well suited for X-ray structure analysis to elucidate the structural principles of the active site and the catalytic mechanism.

The pathway from enzyme-bound substrate to the transition state involves changes in the electronic configuration and geometry of the substrate. The enzyme itself is also not static. The ability to tightly bind the transition state requires flexibility in the active site. Such flexibility has been experimentally demonstrated in many cases. A corollary to this is that the effectivity of enzyme catalysis can easily be influenced and regulated by conformational changes in the enzyme. An extensive consideration of the mechanisms of enzymes can be found in the works by J. Kraut (1988) and A. Fersht (1998).

The binding of an effector molecule or a covalent modification of the enzyme, such as phosphorylation, can prevent the necessary restructuring of the enzyme that enables strong binding of the transition state, thereby inhibiting the reaction. Effectors and enzyme modifications can also affect the substrate binding site such that binding of the substrate in the ground state is impossible or very weak. On the other hand, activation by effectors and modifications can be achieved by stabilizing a conformation of the enzyme in which substrate binding is favorable and a high complementarity between enzyme and substrate in the transition state is possible.

The ability of proteins to exist in different conformations is termed allostery (see 2.2). Allosteric enzymes can assume various conformations which differ in catalytic activity and/or substrate binding capacity.

## 2.2 Regulation of Enzymes by Effector Molecules

The regulation of the activity of enzymes by the binding of effector molecules is a ubiquitous and general principle for the fine tuning and control of metabolic activity. Effector molecules are often low molecular weight organic compounds. Proteins and metal ions can also exercise the function of effectors. The effector molecules bind specifically to the enzymes and the binding results in inhibition or stimulation of enzymatic activity.

For the regulation of metabolic pathways metabolites are often used which are a product of that pathway. The basic strategy for the regulation is exemplified in the mechanisms employed in the biosynthetic and degradation pathways of amino acids, purines, pyrimidines, as well as in glycolysis. In most cases a metabolite (or similar molecule) of the pathway is utilized as the effector for the activation or inhibition of enzymes in that pathway.

Effector molecules bind to an enzyme of a metabolic pathway and modify its actity in a concentration-dependent manner. This regulation serves to adjust the level of production of metabolites by the pathway to the current needs in the cell. The regulatory signal of such a „*feedback regulation*“ strategy is the concentration of the final product or of an intermediate. In feedback inhibition, if the concentration of a particular product exceeds a certain level, then the product occupies the active site of the enzyme and thereby inhibits it. The enzymes of a metabolic pathway that act as a bottleneck for the total throughput, i.e. are rate limiting for the overall reaction, are generally the primary targets for feedback regulation. In this case the allosteric regulation of one enzyme can control the overall throughput of the metabolic pathway. These enzymes are recognized as *key enzymes* and are usually involved in early steps in the reaction path (fig. 2.2).

The basic elements for the regulation of enzyme activity by effector molecules are *allosteric conformational changes* of the enzyme. Allostery means that the enzyme can exist in various conformations which differ in activity and substrate or ligand binding. The typical titration curve for the binding of a ligand to an allosteric protein is sigmoidal in shape (fig. 2.3), while binding curves for non-allosteric enzymes are hyperbolic.

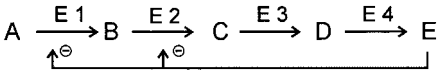



**Fig. 2.2.** Feedback inhibition of a metabolic pathway. In feedback inhibition, the final product, E, of a metabolic path acts as an inhibitor of an early reaction in a cycle, which in the figure leads eventually to the formation of E with the aid of enzymes E1, E2, etc. E acts as an effector on the enzymes E1 and E2 to inhibit their activity.

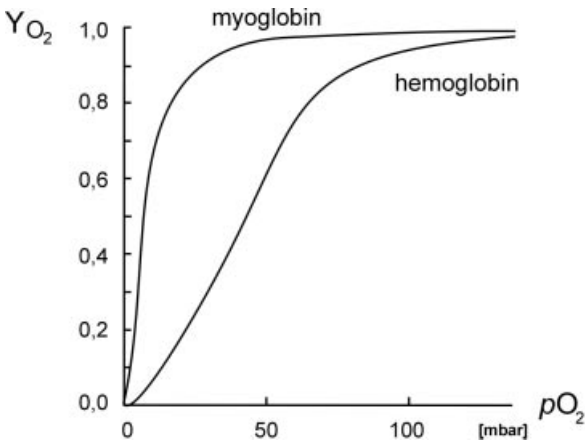



**Fig. 2.3.** Ligand-binding to allosteric and non-allosteric proteins: oxygen bindig to myoglobin and hemoglobin. a) the binding of $O_2$ to myoglobin can be described by a hyperbolic curve. b) the sigmoid form of the $O_2$ binding curve to hemoglobin is characteristic for a ligand-binding to an allosteric protein.Y= degree of binding, ratio of binding sites occupied with $O_2$ to the total O binding sites of hemaglobin; $pO_2$: partial pressure of $O_2$.

Enzymes that are regulated by effector molecules in an allosteric manner possess, apart from the binding site for the substrate, a specific binding site for the effector molecule. The binding of effector molecules to the effector site leads to a shift in equilibrium between the various conformations and thus to a change in activity (see 2.3).

## 2.3 Mechanistic Descriptions of Allosteric Regulation

Various models have been offered to describe the mechanisms of allosteric regulation of enzyme activty, of which the symmetry model (Monod et al., 1965) is the simplest. The symmetry model has proven suitable in many cases to explain the experimentally observed characteristics of allosteric enzyme regulation.

The symmetry model (fig. 2.4) of allostery can describe the cooperative binding of substrate to enzyme (homotropic effect), as well as the influence of effector molecules on the activity of enzymes (heterotropic effect).

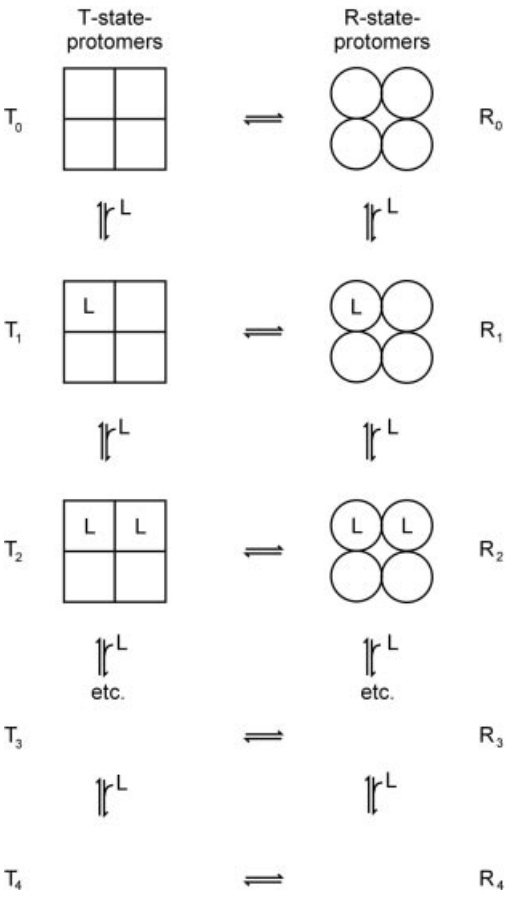



**Fig. 2.4.** The symmetry model of allostery. Shown here is the succesive binding of a ligand L to a protomer of a tetrameric protein with four ligand-binding sites according to the symmetry model. T: tense form, R: relaxed form.

The model assumes the following with regard to the binding of a ligand (=substrate) to the protein:

1. Allosteric proteins are oligomeric and composed of symmetric subunits.
2. The subunits can exist in two forms, termed the T-form and the R-form. The R-form („relaxed“) is the relaxed, active state; the T-form („tense“) is the less active state. T and R forms are in equilibrium with each other, independent of whether the ligand is bound or not.
3) The ligands can bind both to the T and R forms. The two forms differ in their affinity to the ligand.
4) Both R and T forms are symmetrical in structure. The subunits of the oligomers are either all in T or all in R form. T and R foms can not coexist in an oligomer.

The influence of effector molecules is described by the symmetry model in the following manner:

*Activators* bind preferentially to the R form of the enzyme and thereby stabilize it. In the R form the enzyme possesses greater affinity for the substrate and is therefore more active.

*Inhibitors* bind preferentially to the T form. In the presence of an inhibitor more molecules occur in the T form. In this form the enzyme possesses a lower affinity for the substrate and the enzyme is thus less active.

Fig. 2.5 illustrates the influence of an activator and of an inhibitor on the binding of a ligand (and thereby on the activity) by a tetrameric protein. The activator shifts the binding curve to lower ligand concentrations and can, in the extreme case, lead to a hyperbolic binding curve. The inhibitor shifts the binding curve to higher ligand concentrations, so that higher concentrations are required to saturate the ligand binding sites.

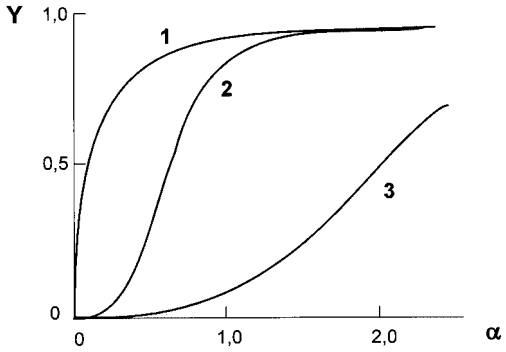



**Fig. 2.5.** The influence of an activator and an inhibitor on the ligand-binding curve of an allosterically regulated enzyme. In the absence of inhibitor or activator, the ligand-binding curve is sigmoidal (curve 2). In the presence of an activator the binding curve is shifted to lower ligand concentrations (curve 1) and approaches hyperbolic form, similar to that observed in Fig. 2.3 for the binding of $O_2$ to myoglobin. An inhibitor shifts the binding curve to higher ligand concentrations (curve 3). Y: degree of binding, ratio of occupied to toal binding sites of the protein; a: L/P0; $L_0$: total ligand concentration; $P_0$: total concentration of binding sites on the protein.

## 2.4 Structural Basis of Allosteric Regulation on the Example of Phosphofructokinase

Based on high resolution crystal structures, detailed understanding of the molecular principles of allosteric regulation for several enzyme systems could be gleamed. One of the best studied examples is that of phosphofructokinase from *Bac. stearothermophilus* (Schirmer and Evans, 1990).

Phosphofructokinase (PFK) is a key enzyme for the substrate turnover in glycolysis. PFK catalyzes the folowing reaction:

Fructose-6-P + ATP ⇄ Fructose-1,6-Bis-P + ADP

PFK displays sigmoidal kinetics for the conversion of fructose-6-P, although not for ATP. The enzyme is *allosterically activated* by ADP and *allosterically inhibited* by phosphoenolpyruvate (fig. 2.6).

Two different enzymatically active forms of PFK could be identified which may be considered the R and T form in the framework of the symmetry model. The R form possesses a high affinity for the substrate fructose-6-P, the T form binds fructose-6-P with lower affinity. Upon binding of the inhibitor phosphoenolpyruvate, PFK converts to the T form. The enzyme is found in the R form upon binding the substrates (ATP or fructose-6-P) or the activator (ADP). There exist high resolution crystal structures of both forms.

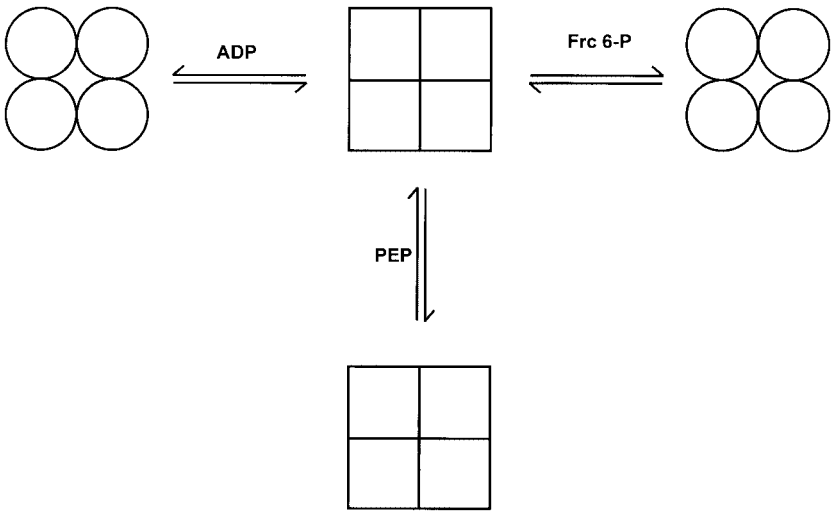



**Fig. 2.6.** Regulation of Phosphofructokinase from *Bacillus stearothermophilus.* The tetrameric phosphofructokinase is allosterically regulated by ADP, Frc-6-phosphate, and phosphoenolpyruvate (PEP). The binding of ADP and Frc-6-phosphate converts the enzyme into the active R state. PEP binds to the T state and inhibits phosphofructokinase. The circles represent the R state, and the squares represent the T state of the enzyme.

PFK from *Bac. stearothermophilus* consists of four identical subunits, each 319 amino acids. Each subunit has two domains, a large domain with the ATP binding site and a smaller domain with the binding site for furctose-6-P. The subunits are packed in the form of two dimers (fig. 2.7, subunits AB and CD) with a large contact surface between the monomers of the dimer and a smaller contact surface between the dimers. The substrate binding site for fructose-6-P is found near the small contact surface between the two dimers. At this surface, prominent differences in the conformation of the high affinity R form and the low affinity T form occur. The effector binding site for the activator, ADP, is nearly identical with that for the inhibitor, phosphoenolpyruvate, and is found near the large contact surface within a dimer.

The *structural* differences between the R- and T-form are strongest at the interface between the dimers and impact heavily the binding site for fructose-6-P. At the binding site for the second substrate, ATP, there are only minor structural differences between the two conformations. Overall, the transition from R- to T-form is accompanied by a rotation of the two dimers around the small interface; the large contact surface within a dimer remains relatively unchanged.

The following discussion is restricted to the fructose-6-P binding site and the 6-F loop since the conformational change between the R and T-form is most noticeable here. Fig. 2.8 shows the binding site for fructose-6-P in the R- and T-form. In the T-form, both dimers are tightly coupled by salt bridges, the cleft between both subunits is constricted, and the binding site for fructose-6-P is blocked. In the R-form the cleft between the dimers is widened and the binding site for fructose-6-P is accessible. Residues from both subunits are involved in the binding of fructose-6-P, whereby the residues Arg162 and Glu161 of the 6-F loop clearly assume different positions in the R- and T-form. In the T-form Glu161 is found near the fructose-6-P binding site so that binding is prevented by electrostatic repulsion. Upon the transition from T->R, Glu161 swings out of its position and is replaced by Arg162. The latter binds, together with Arg243, via a salt bridge to the phosphate moiety of fructose-6-P. This salt bridge, which is only possible in the R-form, contributes significantly to the binding energy of fructose-6-P.

The movement of the residues involved in fructose-6-P binding is mediated by neighboring helices and loops, whose own position is influenced by the effector binding site. Binding of the activator ADP to the effector binding site stabilizes the R conformation. Binding of the inhibitor phosphoenolpyruvate (or 2-P-glycolic acid) to the same site fixes the enzyme in the T-form. The disparate effect of the two effectors can probably be explained by their different size. How the occupation of the effector binding site is coupled to the conformational change in the region of the fructose-6-P binding site is not yet understood.

A central event in the allosteric mechanism is the rotation of the dimers relative to each other. The movement is coupled with the conformational change of all four binding sites for fructose-6-P, since this is found at the dimer interace.

The following points should be noted:

1) changes in the conformation of the subunits are coupled with one another via the inter-subunit contact surface. In this manner, the subunit interface plays a central role in the allosteric mechanism.
2) The binding site for the fructose-6-P substrate is found at the contact surface of the subunits and its accessibility depends on the mutual orientation of the subunits.

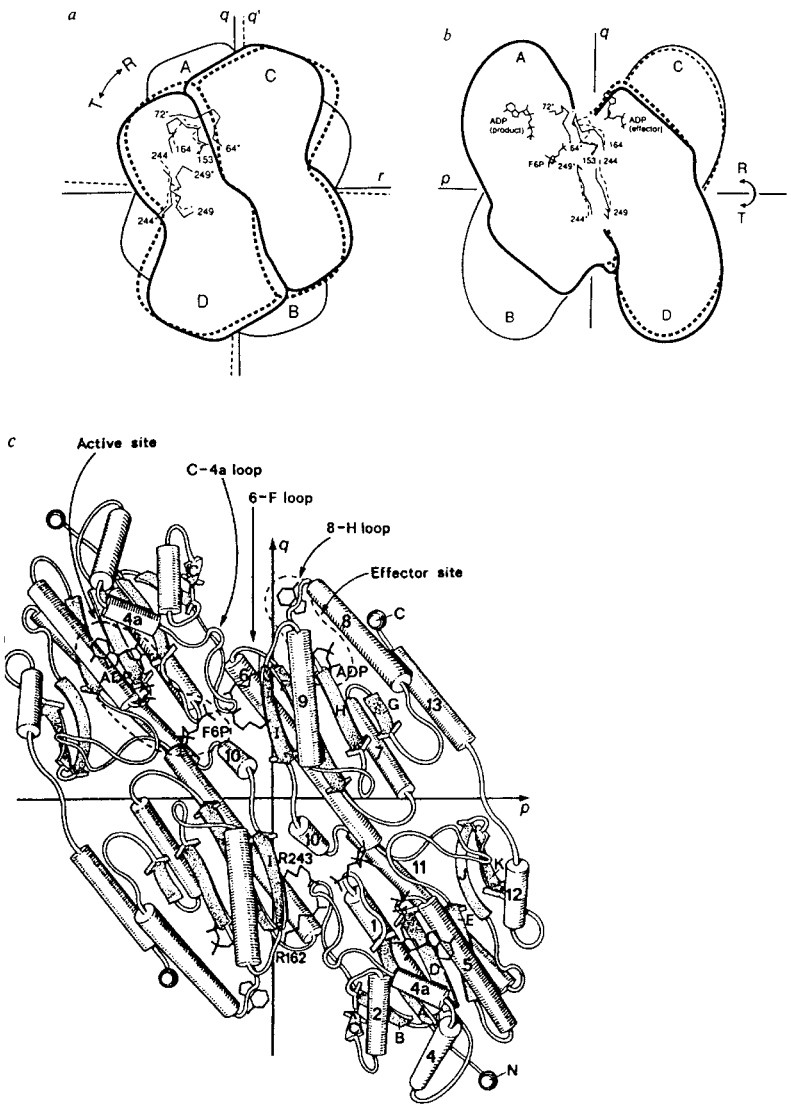



**Fig. 2.7.** Structure of Phosphofructokinase from *Bacillus stearothermophilus* displaying the binding sites for substrate and effector. a, b) schematic representation of the tetrameric phosphofructokinase and the change in quarternary and tertiary strucutre of the enzyme upon T-R interconversion. The subunits are numbered A-D; p, q, and r are the various parallel two-fold symmetry axes. a) view along the p axis. b) view along the r axis highlighting the substrate and effector binding sites. The solid lines indicate the T-state, the dashed lines the R-state. c) structure of an AD dimer showing the binding site for fructose-6-P and for the effectors ADP or Phosphoenolpyruvate. The view is, as in b), along the r axis.The contact surface between the A and D subunit runs approximately along the q axis. The helices are drawn as cylinders and are numbered from 1–13. The β-pleated sheets are drawn as strands and are numbered A-K. After Schirmer and Evans (1990), with permission.

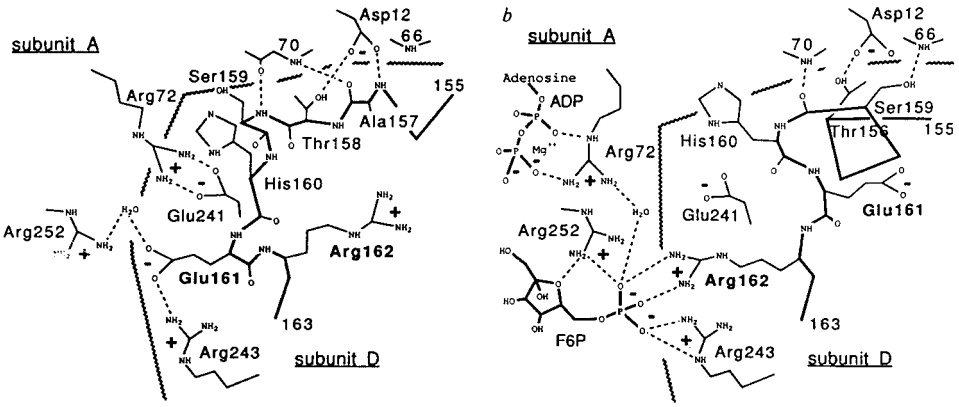



**Fig. 2.8.** Schematic representation of the interactions between the dimer contact surface of phosphofructokinase in the T- and R-state. The jagged lines mark part of the contact surface between an A and D subunit. The view is along the r axis, as in Fig. 2.7b, and thus along the contact surface between the AB and CD dimers. a) T-state; In the T-state the two subunits interact tightly via salt bridges in the indicated region. b) R-state: Fructose-6-phosphate binds in the cleft between the two subunits. The binding of ADP is also shown in the binding site for the ATP substrate. After Schirmer and Evans (1990), with permission.

3) The T- and R- form differ above all in their ability to bind the substrate fructose-6-P. In the T-form the binding site for fructose-6-P is inaccessible due to electrostatic and steric effets. In the R-form the binding site is accessible and, furthermore, the amino acid side chains necessary for a tight binding are optimally situated.
4) The binding of an effector molecule to the effector binding site leads to a coupled conformational change to fixate either a high affinity R-form or a low affinity T-form.

The high resolution X-ray structural data of the R- and T-form help us understand the differences in the conformation of the two states. The crystallographic data however mediates only a static picture of the allosteric mechanism. Structures that occur between the two conformational extreme situations of the R- and T-states are not captured. Furthermore, the quest for the understanding of the mechanism of the conformational coupling, as well as the influence of conformational changes on the enzymatic reaction, remains elusive. Although we still do not understand all the details and the intermediate states of the allosteric process, the structural data offers an indispensable framework with which to understand the principles, which guide allosteric regulation.

## 2.5 Regulation of Enzyme Activity by Binding of Inhibitor and Activator Proteins

Enzyme-specific inhibitor and activator proteins can be considered as a type of effector molecules.

### Inhibitor Proteins

There are numerous examples of inhibitor proteins that specifically bind a particular enzyme and block its activity.

Inhibitor proteins themselves are subject to a variety of regulation mechanisms. The function of an inhibitor protein can be regulated, for example, by protein phosphorylation (see chapter 7, Inhibitors of Protein Kinases), by degradation or by de novo synthesis (see chapter 13, Inhibitors in the Cell Cycle).

The structurally best characterized example is found amongst the proteases: specific inhibitor proteins ensure that the proteases carry out their proteolytic function only in certain tissues and only under certain metabolic conditions. By this token, the protease inhibitors assume an important protective and regulatory function in metabolism.

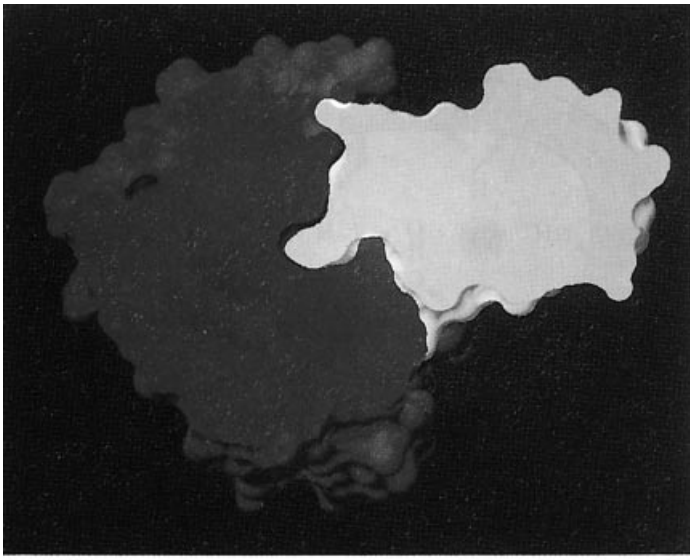
The structural analysis of the trypsin inhibitor from bovine pancreas (BPTI) in complex with trypsin shows that the inhibitor occupies and blocks the substrate binding pocket in a highly complementary manner (fig. 2.9). In the trypsin-BPTI complex, the catalytically essential Ser-OH of trypsin contacts a CO group of the inhibitor in a manner very similar to the tetrahedral transition state of amide or ester bond hydrolysis (see fig. 2.9b). The inhibitor can be likened to a *pseudo-substrate* and, as such, is bound with high affinity. The cleavage of the peptide bond is, however, not possible due to other circumstances, such as the fact that water is prevented from reaching the active site with the inhibitor bound.

A further example of inhibitor proteins is the *protein kinases* and the cognate *protein kinase inhibitors*. The activity of protein kinases in the cell cycle are controlled via the association with inhibitor proteins. The basic mechanism and the goals of this phenomenon will be discussed in detail in chapters 7 and 13.

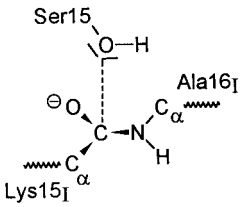
### Activator Proteins

An example for the reversible association of activator proteins with an enzyme is the $Ca^{2+}$-calmodulin dependent enzymes. Calmodulin is a $Ca^{2+}$-binding protein which can activate target enzymes, e.g. phosphorylase kinase (see 6.7.1 and 7.4) in its $Ca^{2+}$-bound form . Another example for activating proteins is the cyclins (see chapter 14). The cyclins are activators of protein kinases that regulate the cell cycle.

Activator proteins themselves can be bound in regulatory networks, as shown in the example of the cyclins (chapter 14). The function of an activator protein can be regulated at the level of gene expression, degradation, or post-translational modification (e.g. protein phosphorylation).



(a)




(b)

**Fig. 2.9.** The complex of trypsin and trypsin inhibitor (BPTI). a) A cross section through the structure of the complex shows how trypsin (dark) binds to the trypsin inhibitor (light). The light-colored tongue represents the side chain of Lys 15 of the inhibitor invading into the substrate binding pocket of trypsin. b) The catalytically active Ser195 of trypsin interacts tightly with the carbonyl-C of Lys15 of the inhibitor (Lys15I). The peptid bond between Lys15 and Ala16 of the inhibitor (Ala16I) assumes a conformation with respect to Ser195 similar to the peptide bond to be hydrolyzed. However, the conditions for peptide bond hydrolysis of the inhibitor are not met for steric reasons: the contact with the Lys15-Ala16 peptide bond is too close. Furthermore, the leaving group can not exit the active site and the hydrolyzing water molecule can not enter. After Voet and Voet (1992) with permission.

### Metal Ions

The availability of metal ions can also be employed for regulation of enzyme activity. Of primary importance is $Ca^{2+}$. An important example in this regard is protein kinase C, which is activated by $Ca^{2+}$ (see chapter 7). The availability of $Ca^{2+}$ is further regulated in various ways by hormone-controlled pathways (see chapter 6).

## 2.6 Regulation of Enzyme Activity by Phosphorylation

The phosphorylation of enzymes by specific protein kinases is a widespread mechanism for the regulation of enzyme activity. It represents a flexible and reversible means of regulation and plays a central role in signal transduction chains in eucaryotes.

Proteins are phosphorylated mainly on *Ser/Thr* residues and on *Tyr* residues. Occasionally Asp or His residues are phosphorylated, the latter especially in procaryotic signal transduction pathways (see chapter 7, chapter 12). For the regulation of enzyme activity the phosphorylation of Ser and Thr residues is most significant. Apart from regulation of Tyr kinases, Tyr phosphorylation serves the function of creating specific attachment sites for proteins. Both of these functions will be discussed in more detail in chapter 8.

Protein phosphorylation is a specific enzymatic reaction in which one protein serves as a substrate for a protein kinase. Protein kinases are phophotransferases. They catalyze the transfer of a phosphate group from ATP to an acceptor amino acid in the substrate protein (fig. 2.10). A detailed discussion of protein kinases can be found in chapter 7.

The response of a protein upon phosphorylation is dictated by the special properties of the phosphate group. The phosphate group has a $pK_a$ of about 6.7 and carries two negative charges at neutral pH. Therefore, two negative charges are generated in a substrate protein upon phosphorylation of an uncharged amino acid side chain. This fact, and the presence of four oxygen atoms, allows the phosphate group to form an extensive network of H-bonds which can link different parts of a polypeptide chain. Similar twofold negatively charged groups do not occur in other structural elements of proteins. Electrostatic interactions and a network of H-bonds are therefore of special importance for the control of protein functions by phosphorylation.

Analysis of protein-protein-interactions in existing structures has shown that phosphate groups most commonly interact with the main-chain nitrogens at the start of a helix, where often glycine is found. In non-helix interactions phosphate groups most commonly contact arginine residues. The guanidinium group of arginine is well suited for interactions with phosphate because of its planar structure and its ability to form multiple hydrogen bonds. The electrostatic interaction between arginine residues and the phosphate group provide tight binding sites that often function as organizers of small-range as well as long-range conformational changes.

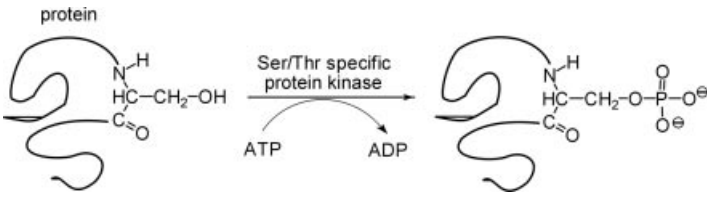



**Fig. 2.10.** Change in charge state of proteins via phosphorylation. The phosphorylation of Ser residues is catalyzed by a Ser/Thr-specific kinase that utilizes ATP as the phosphate group donor. The product of the reaction is a Ser-phosphate ester which carries a net charge of -2 at physiological pH.

The phosphate ester of Ser, Thr, or Tyr residues are quite stable at room temperature and neutral pH; the rate of spontaneous hydrolysis is very low. Therefore, to remove the phosphate residue the cell utilizes specific enzymes termed phosphatases. Based on substrate specificity, these can be classified as Ser-, Thr- or Tyr-specific phosphatases (see chapter 8).

**Structural Consequences of Protein Phosphorylation**

The molecular basis of the control function of protein phosphorylation will be illustrated for two enzymes, glycogen phosphorylase from rabbit muscle and isocitrate dehydrogenase from *E. coli*. Both examples represent very different mechanisms by which the activity of an enzyme is altered by protein phosphorylation. Glycogen phosphorylase displays extensive allosteric conformational changes upon protein phosphorylation. Isocitrate dehydrogenase is phosphorylated directly in the substrate binding site with only minimal conformational changes resulting. For further examples on the structural aspects of the control by phosphorylation see Johnson and O'Reilly (1997).

## 2.6.1 Regulation of Glycogen Phosphorylase by Phosphorylation

Glycogen phosphorylase was the first enzyme shown to be regulated via protein phosphorylation (Krebs, 1959). In recognition of their trail-blazing work, Edwin Krebs and Edmond Fisher were rewarded the Nobel prize for Chemistry in 1992.

Glycogen phosphorylase catalyzes the transfer of a glucose unit from glycogen to an inorganic phosphate to form Glc-1-P. This first step of glycolysis is the target of several regulatory processes. The phosphorylase is subject to allosteric regulation by effector molecules, as well as a reversible regulation via phosphorylation. The regulation of the phophorylase by phosphorylation is the last step in a hormone controlled signal chain, the individual steps of which will be discussed in chapters 5, 6 and 7 (see also fig. 7.18).

In the framework of the symmetry model (see 2.3), a T- and R-form can be formulated for glycogen phosphorylase. In the T-form glycogen phosphorylase binds its substrates and activating effectors with lower affinity, while in the R-form it possesses higher affinity for substrates and activating effectors.

A simplified representation of the various conformational forms of glycogen phosphorylase is given in fig. 2.11. The non-phosphorylated form is termed *phosphorylase b.* In the absence of effectors, phosphorylase b is found in the inactive T-form. Phosphorylation on Ser14 by phosphorylase kinase transforms phophorylase b into *phosphorylase a*, for which equilibrium lies toward the active, or R-form. The transition of phosphorylase b from the T-form to the R-form can also be induced by the binding of the activator AMP. Binding of the inhibitor Glc-6-P transforms the R-form of phosphorylase b back into the T-form. IMP and glucose have also been described as inhibitors that fix the enzyme in the T-state. In the absence of effectors, glycogen phosphorylase b occurs as a dimer. Binding of the activator AMP or phosphorylation leads to a partial aggregation of the dimers to tetramers. For structural discussions, however, it is sufficient to consider only the dimeric form of the enzyme.

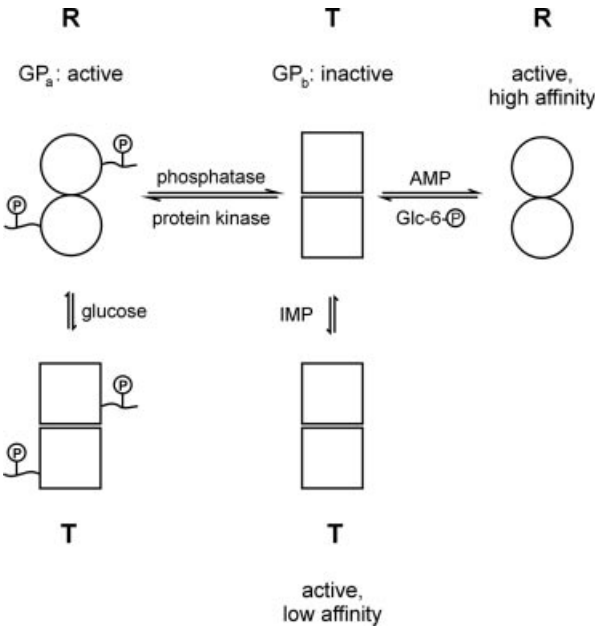



**Fig. 2.11.** Model of allosteric and covalent activation of glycogen phosphorylase of muscle. The R-form of the subunits are represented by circles, the T-form by squares. The active state of glycogen phosphorylase (GP) is characterized by a high affinity, the inactive state by low affinity for the substrate P1.

High resolution X-ray structures of the a- and b- form of rabbit muscle phosphorylase permit a view into some of the structural differences of the various allosteric forms of the enzyme. Furthermore, the data give an impression of the mechanism of binding of effectors and the influence that phosphorylation has on substrate binding and enzyme activity (Barford et al., 1991). The following discussion will be restricted to the observed consequences of phosphorylation.

Phosphorylation of the glycogen phosphorylase of muscle occurs on Ser14 near the N-terminus. In the unphosphorylated T-state the amino acids 10–18 are found in disordered structures or irregular β-sheets and interact with other residues in the same subunit. In the T-form, the side chain of Ser14 is located near Glu501 of the same subunit. Phosphorylation of Ser14 leads to a radical restructuring (fig. 2.12): the N-terminus reorients itself in the opposite direction and interacts with the other subunit. The reorientation brings the negatively charged phosphate moiety closer to positively charged residues, substantially stabilizing the alternative conformation. Coupled to the restructuring of the N-terminus is a reorientation of additional structural elements, which eventually affects the binding site for the phosphate substrate. The final result is an active site conformation more favorable for substrate binding.

Phosphorylation of glycogen phosphorylase is the initiator for the coupled conformational changes, which are communicated over a large distance to the active site. Similar to the allosteric mechanism of phophofructokinase, the inter-subunit contact surfaces play a decisive role for the communication between the phosphorylation site

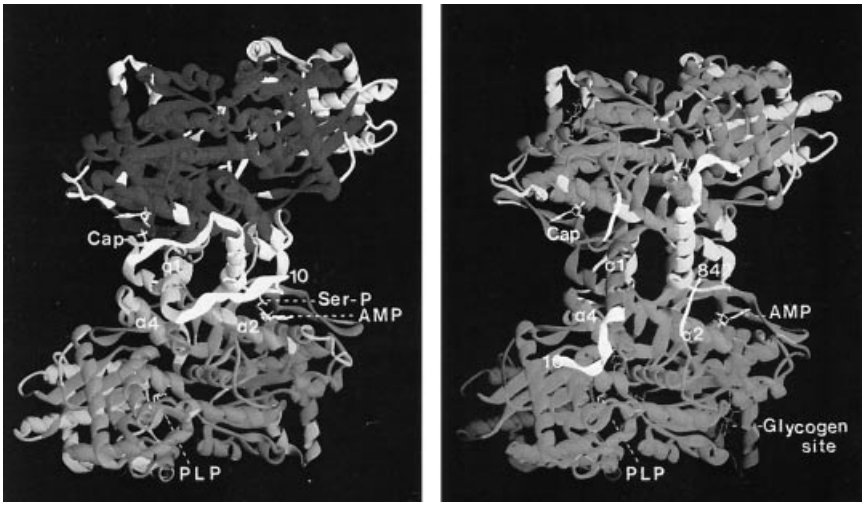



**Fig. 2.12.** Structural changes at the N-terminus of glycogen phosphorylase as a result of phosphorylation. a) R-form of the dimer of glycogen phosphorylase a. b) T form of the dimer of glycogen phosphorylase b. Phosphorylation at Ser13 near the N-terminus transforms the inactive glycogen phosphorylase b into the active glycogen phosphorylase a. The N-terminus rearranges significantly as a result of phosphorylation. In the inactive T-state the N-terminus interacts with the same subunit, while in the R-form it forms interactions with the other subunit. After Barford and Johnson (1991), with permission.

and the substrate binding site. They act as hinges for the transmission of the signal, allow interactions over large distances and provide for a cooperative coupling of conformation changes between the subunits. For a more detailed account one should refer to the original literature (Barford et al., 1991).

### 2.6.2 Regulation of Isocitrate Dehydrogenase (*E. coli*) by Phosphorylation

Isocitrate dehydrogenase catalyzes the NAD-dependent reduction of isocitrate to α-ketoglutarate. The dimeric enzyme is regulated via phosphorylation. Phosphorylation on Ser113 leads to a complete inactivation of the enzyme.

An understanding of the molecular basis for regulation of isocitrate dehydrogenase by phosphorylation was facilitated by X-ray crystallography of the phosphorylated enzyme in complex with isocitrate. The crystal structures of mutants of the enzyme in which Ser113 had been exchanged for aspartate or glutamate were also solved (Hurley et al., 1990). The structure of the enzyme in complex with the substrate isocitrate revealed the phophorylation site to be localized near isocitrate. Ser113 itself binds the substrate directly via a H-bond with the $O^4$ of isocitrate (fig. 2.13).

The comparison of various forms of the enzyme demonstrated that phosphorylation does not always imply extensive conformational changes. The observed small confor-

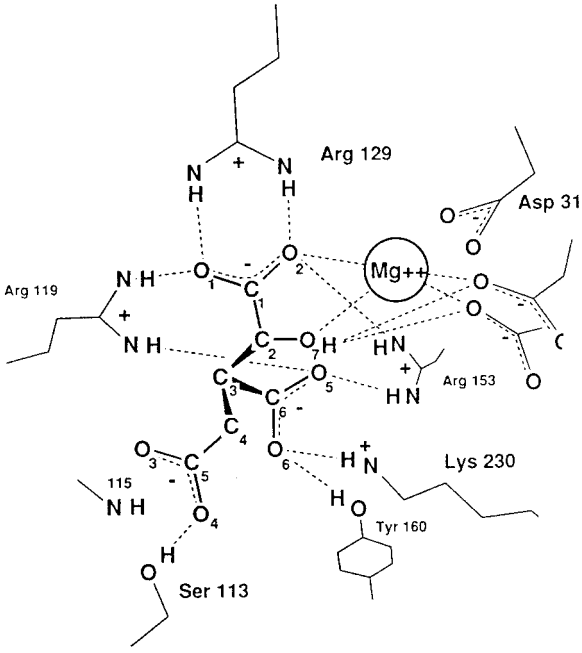



**Fig. 2.13.** Substrate binding site of isocitrate dehydrogenase (*E. coli*). The interactions involved in the binding of $Mg^{2+}$-isocitrate at the active site of isocitrate dehydrogenase are shown. After Hurley et al., (1990), with permission.

mational differences between the phosphorylated and unphosphorylated form may be irrelevant for the regulation mechanism.

Decisive for the inactivation of the enzyme via phosphorylation of Ser113 is the creation of negative charges at the substrate binding site, which aggravates binding of the negatively charged isocitrate.

As a consequence of electrostatic repulsion, the isocitrate substrate can only bind with low affinity to the phosphorylated enzyme. This interpretation is supported by structural data on mutant forms of isocitrate dehydrogenase. If one replaces Ser113 with amino acids that possess negatively charged side chains, then isocitrate can still bind, albeit with markedly reduced affinity, and the binding can only occur with an unfavorable geometry. Theoretical calculations of the change in free energy for the binding of isocitrate upon phosphorylation corroborates the conclusion that introduction of a negative charge to Ser113 is sufficient to explain the experimentally observed decrease in affinity for isocitrate.

## 2.7 Regulation of Enzyme Activity by Proteolysis

Proteolysis is a versatile tool in the cell for the targeted change in structure, activity, function and subcellular distribution of proteins.

Examples for directed proteolysis are:

- removal of the N-terminal methionine residue from de novo synthesized proteins
- removal of the signal peptide upon protein translocation across the membrane
- hydrolysis of the translation product of monocistronic mRNA in viruses
- maturation of proteins: transformation of inactive proteins into active proteins
- targeted degradation of proteins

Most important, in terms of regulation, is the maturation of proteins and targeted protein degradation.

## 2.7.1 Maturation of Proteins via Proteolysis

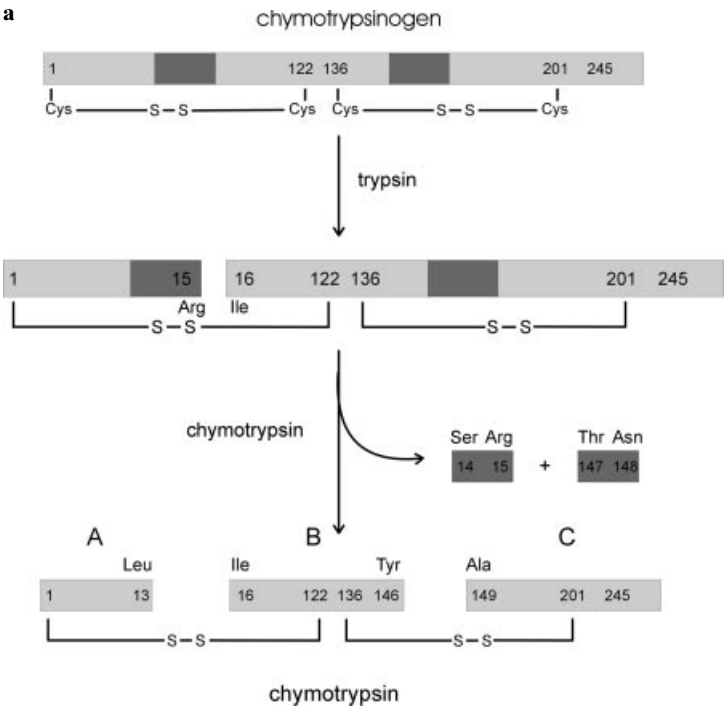
Many proteins are formed as inactive precursors and become activated by proteolysis. The inactive precursors are termed proenzymes, zymogens or – for hormones like e.g. insulin – prehormones. Processing to the active form occurs in a cell- and tissue-specific way and usually requires a specific protease. Activation can also occur intramolecularly by autoproteolysis. In most cases, short sequences of the protease substrate serve as a recognition signal for the attack of the processing protease. Of the numerous examples of proteolytic processing of proteases only the digestive proteases will be discussed in more detail.

**Activation of Digestive Enzymes**

Many digestive enzymes in the pancreas or in the stomach are formed as zymogens and are activated by specific proteolysis (refer to text books).

The activation of the zymogen chymotrypsinogen to chymotrypsin is shown schematically in fig. 2.14a. Chymotrypsinogen is a protein of 245 amino acids that contains 5 disulfide bridges and is enzymatically inactive. The activation to π-chymotrypsin is performed by trypsin, which cleaves the peptide bond between arginine 15 and isoleucine 16. The π-chymotrypsin formed catalyzes the cleavage of further π-chymotrypsin molecules to mature chymotrypsin which consists of three chains linked by disulfide bridges. The key step in activation is the cleavage of the peptide bond between Arg15 and Ile16 leading to formation of the enzymatically fully active π-chymotrypsin. The newly formed N-terminus (Ile16) of the B-chain plays an important role in the activation process:

a) The positively charged amino terminus of Isoleucine 16 rotates to the interior of the chymotrypsin molecule and interacts electrostatically with Asp 194, which further coordinates the Ser residue of the catalytic triad of chymotrypsin (Ser195, His57, Asp102).
b) As a consequence of the interactions between Ile16 and Asp194 further conformational changes occur which contribute to the formation of the specific substrate binding site. This site is not fully formed in the zymogen.
c) The conformational changes create an orientation of the peptide backbone that is particularly suitable for the stabilization of the transition state. In the postulated



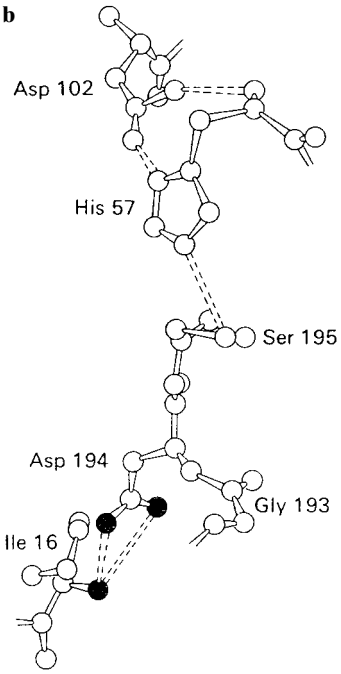




**Fig. 2.14.** The activation of chymotrypsin via proteolytic cleavage. a) Chymotrypsinogen is transformed into the active forms of chymotrypsin π and α by trypsin and autoproteolysis. b) The N-terminal isoleucine residue Ile6 is particularly important for the activity of chymotrypsin. The positively charge NH2 group of Ile16 interacts electrostatically with Asp194 and stabilizes an active conformation of the catalytic center. After Stryer „Biochemistry“, with permission.

transition state of the reaction a negative partial charge is formed on a peptide oxygen. The oxyanion is stabilized by H-bonds with NH-groups of the peptide backbone. The binding site for the oxyanion ('oxyanion pocket') is not completely formed in the zymogen, which explains why the transition state can only be partially stabilized.

d) In other regions of the molecule there are only minor conformational changes observed. The activation of chymotrypsinogen to chymotrypsin is thus attributed to the cleavage of a single peptide bond.

Trypsin is responsible for the activation of chymotrypsin, as well as for a range of other digestive enzymes synthesized in the pancreas. Trypsin itself is formed from its zymogen via digestion by the enzyme enteropeptidase. Enteropeptidase is secreted from intestinal cells and cleaves trypsinogen to trypsin as soon it travels from the pancreas to the intestine.

The activation of pepsin from its zymogen pepsinogen occurs by a different mechanism. In this case, the pH of the environment plays a decisive role. In the strongly acidic milieu of the stomach cleavage of a 44 amino acid peptide occurs from the inactive precursor pepsinogen. The activation is intramolecular and depends on the pH of the solution.

#### Formation of Hormones from Precursors

Many other protein hormones are also synthesized as inactive precursors. Examples are insulin, which is formed in a two-step proteolytic process from the precursor pre-pro-insulin. Another noteworthy example is that of pre-pro-opiomelanocortin, which is the precursor for eight peptide hormones and neuropeptides in the epiphyse.

#### Specific Proteolysis in Blood Clotting

During blood clot formation a complex cascade of zymogen activation occurs, whereby each currently activated zymogen activates the subsequent zymogen. For details one is referred to text books.

### 2.7.2 Specific Degradation of Proteins in the „Ubiquitin-Proteasome" Pathway

The function of most proteins in the cell is tightly restricted both locally and temporally. To what extent an enzyme can participate in metabolism depends not only on the rate of biosynthesis and on the extent of covalent modification, such as phosphorylation, but also on the rate of degradation of a protein. The comparison of the *life span* of various proteins (table 2.1) shows that the stability of proteins can differ significantly. Among the most unstable, short-lived proteins are found many regulatory proteins and enzymes with key positions in a metabolic pathway. The targeted degradation of proteins involved in the cell cycle (see chapter 13) is an essential element of cell cycle regulation.

**Tabelle 2.1.** Lebensdauer ausgewählter Proteine.

| Life span (h) | Nucleus | Cytosol | ER and plasma membrane |
|---|---|---|---|
| <2 | Transcription factors Fos, Myc, Myb, p53, Hsp70 | Hsp70, proteinkinase C Tyrosine aminotransferase | HMGCoA reductase |
| 2–8 | Cyclins | Proteinkinase A Tryptophan oxygenase | γ-Glutamyl transferase |
| 9–40 | Ubiquitin | Calmodulin, glucokinase, Ubiquitin | EGF rezeptor, epoxide hydrolase |
| 41–200 | Histone H1, HMG1, HMG2 | Alcohol dehydrogenase, catalase Myoglobin, Phospholipase A2 | Cytochrome b5, Dipeptidyl-dipeptidase |
| >200 | Histone H2A, H2B, H3, H4 | Phosphoglycerate kinase Glycogen phosphorylase | NAD glycohydrolase, Acteylcholine rezeptor |

Attempts to explain the variable life span of proteins showed that, apart from the non-specific degradation of proteins, there also exist specific degradation mechanisms. Such specific mechanisms allow the function of a protein to be temporally restricted and specifically modified. There are two main pathways for the degradation of proteins in mammalian cells. In the *lysosomal* path proteins that enter the cell via endocytosis are degraded. The degradation of proteins in the lysosome is rather unspecific and used mainly to eliminate foreign proteins. More recent findings indicate that a specific degradation pathway also exists in the lysosome (see below).

The *non-lysosomal* degradation pathways allows for the selective degradation of proteins under normal cellular conditions. These degradation pathways are also responsible for the degradation of cellular proteins under conditions of stress. The most significant and well characterized non-lysosomal degradation pathway is that of the ubiquitin-proteasome pathway in which proteins are degraded in a 26S proteasome after they have been conjugated by one or more ubiquitin molecules. The ubiquitin-proteasome system (review: Hershko and Ciechanover, 1998) is a tool for the selective proteolysis of proteins and thus plays an important regulatory role in the cell.

#### 2.7.2.1 Components of the Ubiquitin System

Ubiquitin is a 76 residue protein found in nearly every eucaryote. It occurs either in free form or bound to other proteins. All known functions of ubiquitin are transmitted via its covalent linkage with other proteins. This serves the purpose, among others, of marking the proteins for proteolytic degradation.

The ubiquitinylation of proteins is a complex process which involves several specific enzymatic reactions. Three sequential steps can be distinguished (fig. 2.15A):

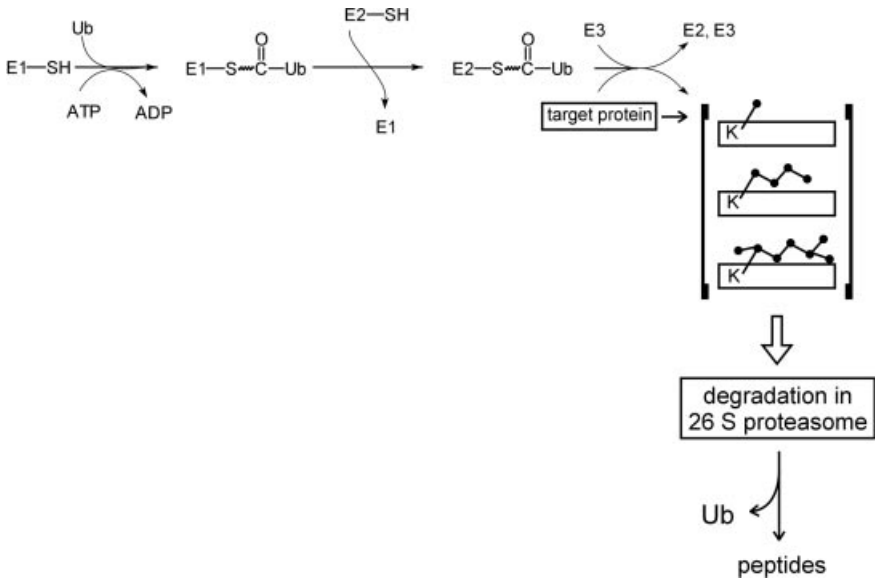



**Fig. 2.15A.** Pattern of ubiquitinylation of proteins and degradation in the proteosome. Ubiquitin (Ub)is initially activated by an enzyme E1, whereby the C-terminal carboxyl group of ubiquitin becomes attached to a SH group of E1 via a thioester bond. Ubiquitin is then transferred from E1-Ub to E2, from which it is transferred with the help of E3 to the target protein. Several ubiquitin molecules can attach to the target protein in a linear or in a crosslinked fashion. The mono- or polyubiquitinylated protein is degraded to peptides in the 26S proteosome. In the above diagram the filled circles represent the ubiquitin residues attached to the target protein. K: lysine residues of the target protein.

### Activation of Ubiquitin, Formation of E1-Ub

In an initial reaction ubiquitin is activated by forming a reactive thiolester with an SH-group of the ubiquitin-activating enzyme E1. This step requires ATP and consists of an intermediate formation of ubiquitin adenylate followed by the binding of ubiquitin to a Cys residue of E1 in a thiolester linkage, with the release of $PP_i$ and AMP.

### Transacylation to E2

In a transacylation reaction the ubiquitin moiety is transferred from E1-Ub to the SH-group of the ubiquitin-carrier protein E2 to form E2-Ub.

### Transfer to the Target Protein with the participation of E3

The third step of ubiquitinylation, the transfer of ubiquitin to the target protein, is catalyzed by a ubiquitin-protein-ligase, or E3 enzyme. In this reaction ubiquitin is linked by its C-terminal glycine in an amide isopeptide linkage to an ε-$NH_2$-group of the substrate proteins' Lys residues.

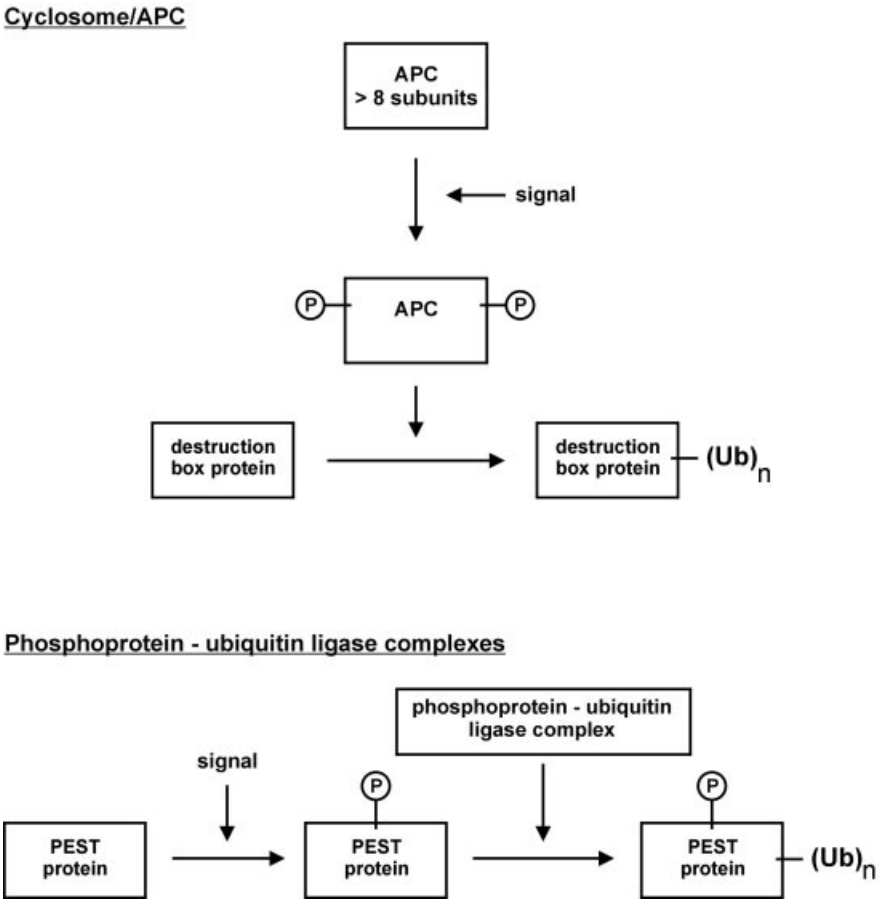



**Fig. 2.15B.** Examples of E3 enzyme complexes.
Two of the four different E3 enzymes mentioned in the text are shown. The cyclosome/anaphase promoting complex (APC) ligates ubiquitin to regulatory and structural cell cycle proteins containing a destruction box as recognition signal (see also 13.3.2). The activity of the APC is thought to be controlled by phosphorylation. For simplicity, the subunits of APC are not shown.
The phosphoprotein-ubiquitin ligases ligate ubiquitin exclusively to phosphorylated proteins. In this system, ubiquitination and degradation are controlled by the phosphorylation status of the target proteins, which is in turn dependent on the regulated activity of protein kinases (or protein phosphatases). Phosphorylation of the target proteins often occurs in sequence elements rich in the aminoacids P,E,S, and T (PEST sequences). For target proteins and the subunit structure, see 13.3.1.

While there is usually only one E1 enzyme, many species of E2 proteins and multiple families of E3 enzymes or E3 multiprotein complexes exist. Selection of substrates for ubiquitin-ligation occurs mainly by specific E3 enzymes which target substrate proteins that contain specific recognition signals (fig. 2.15B). E3 enzymes also can bind indirectly to the substrate, via an adaptor protein.

Ubiquitin can be transferred to the substrate protein by two different mechanisms:

- In some families of E3 enzymes, ubiquitin is first transferred from the E2 carrier to an active site cysteine of the E3 enzyme and subsequently to the ε-$NH_2$-group of an acceptor lysine on the substrate protein.
- In other families of E3 enzymes, no intermediate E3-ubiquitin linkage can be demonstrated. In this case ubiquitin is transferred directly from E2 to the substrate protein. The E3 enzymes are nevertheless required for ubiquitinylation since the E3 enzymes are responsible for substrate selection and are found in tight complexes with the cognate E2 proteins.

#### 2.7.2.2 Degradation in the Proteasome

The degradation of protein-ubiquitin conjugates occurs in an ATP-dependent reaction within a large protease complex, the *26S proteasome* (review: Baumeister et al., 1998). The substrate protein is degraded to peptides in the 26S proteasome, while the ubiquitin is released and again available to form protein conjugates.

The 26S proteasome is composed of two protein aggregates, a 19S and a 20S particle. The main proteolytic component of the 26S proteasome is the 20S particle, the structure of which from an archaebacterial system and a yeast system has been solved (Löwe et al., 1995, Groll et al., 1997).

The structure of the 20S proteasome (fig. 2.16) from *Thermoplasma acidophilum* displays four rings stacked upon each other surrounding a central cavity in which pro-

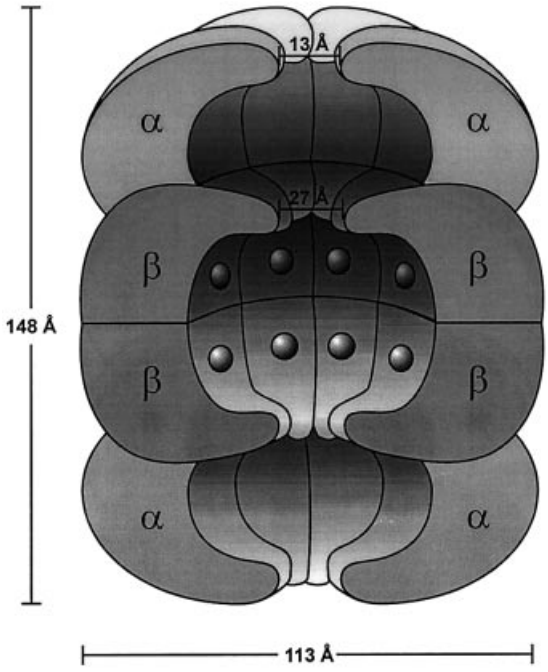



**Fig. 2.16.** Structure of the 20S proteosome of *Thermoplasma acidophilum*. The figure shows the schematic structure of the 20S proteosome (Loewe et al., 1995). Four stacked rings can be identified in the 20S proteosome, each consisting of 7 protomers. The two external rings contain 7 copies of the 26kDa α-subunit of the proteosome, while the inner rings are composed of 7 copies of the 22kDa β-subunit. The rings form a central channel with three chambers. The catalytic centers of proteolytic cleavage are localized on the β-subunits of the inner chamber and are represented in the figure as spheres.

teolysis takes place. A N-terminal threonine has been identified as an essential active site residue of the protease center. The OH-group of the threonine functions as a nucleophile during hydrolysis of the petide bond. A similar mechanism of hydrolysis has been shown for other hydrolases that, due to this property, are now included in the family of N-terminal nucleophile hydrolases. For some β-subunits of eucaryotes the N-terminal threonine is generated by autoproteolysis of a N-terminal prosequence.

The presence of the protease center in the central cavity ensures that the proteolysis is compartmentalized and shielded from the surrounding media. The substrate proteins are accessible to the proteolytic center only via a ring-shaped opening at the end of the 20S proteasome and require the assistance of the 19S particle. The structure of the 20S proteasome also indicates that proteins are accessible to the catalytic center only in the unfolded state.

The 20S proteasome from yeast has a similar overall structure as the archaebacterial proteasome. However, the proteasome from yeast and from higher eucaryotes is composed of seven different α- and β-subunits that occupy specific positions in the proteasome. By exchange of β-subunits different 20S complexes can form which function in specific degradation reactions. The proteasome from yeast and of higher eucaryotes clearly is of a more complicated structure, however it is functionally more versatile than that of archaebacteria and distinct variants of it can be used in specific degradation reactions.

The binding of ubiquitinylated substrates requires the 19S complex of the proteasome, which possesses a ubiquitin binding site and several ATPase sites. It is assumed that the recognition and ATP-dependent unfolding of the substrate protein occurs within the 19S complex.

#### 2.7.2.3 Recognition of the Substrate in the Ubiquitin-Proteasome Degradation Pathway

By far not all cellular proteins are subject to ubiquitinylation. It is therefore evident that the target proteins must exhibit unique features in order to ensure ubiquitin conjugation. It is now generally assumed that the E3 ubiquitin ligases are responsible for the selection of proteins for ubiquitinylation and subsequent degradation. The search for recognition signals in target proteins has led to the classification of ubiquitin ligases into four categories which differ strongly in the complexity of the poteins involved, in the nature of the target sequences recognized and in the regulation of their function (Hershko and Ciechanover,1998).

- **N-end rule E3 enzymes**

This class of E3 enzymes recognizes the nature of the N-terminus of target proteins. As shown by Varshavsky (1992), yeast proteins may be selected for ubiquitinylation according to the 'N-end rule'. Proteins containing basic or bulky-hydrophobic amino acids at the N-terminus are recognized by distinct N-end rule E3 enzymes of which the E3α enzyme (UBR1 protein in yeast) is best characterized.

- **Hect domain E3 enzymes**

The discovery of this family of E3 enzymes started from the studies on the targeted degradation of the p53 tumor suppressor protein. Ubiquitinylation and degradation of p53 can be mediated by the papillomavirus E6 oncoprotein (see below) in collaboration with a further protein, E6-AP (E6 associated protein). E6-AP was the first member of a large family of E3 enzymes, the Hect (homologous to E6-AP C-terminus ) domain family. These proteins contain an essential active site Cys residue near the C-terminus and one or several WW domains ( see Chapter 8.2.6).

- **Cyclosome, Anaphase promoting complex (APC)**

The cyclosome (or APC) is a high molecular weight complex that degrades proteins containing a specific recognition sequence, the destruction box (see chapter 13.2.4). Substrates are cell cycle regulators as e.g. cyclins, kinase inhibitors and spindle-associated proteins. Importantly, some forms of the cyclosome require phosphorylation in order to be active (fig. 2.15B). It is still unclear which of the many subunits carries the E3 enzyme activity.

- **Phosphoprotein-ubiquitin ligase complexes**

Another type of multisubunit ubiquitin ligase is also involved in the degradation of cell cycle regulators, such as the CDK inhibitor Sic1 and G1 cyclins (see 13.2.4). It is characteristic for this class of ubiquitin ligases that the substrates must be phosphorylated in order to be recognized as a substrate for the ubiquitin ligase complex (fig. 2.15B). Protein degradation in this system can be highly regulated since signal pathway mediated phosphorylation can convert a substrate into an active form that is susceptible to ubiquitinylation by the ligase complex. Very often the regulatory phosphorylation events are found in sequence elements called PEST sequences (see 13.2.4).

#### 2.7.2.4 Regulatory Function of Ubiquitin Conjugation and the Targeted Degradation of Proteins

The normal functioning and growth of a cell requires that the life span, and thus the activity, of regulatory proteins and key enzymes be limited. To this end, ubiquitin conjugation and the subsequent proteolysis in the proteasome represents an important tool.

Ubiquitin conjugation has been correlated with the following cellular functions:

– degradation of proteins under stress situations
– degradation of denatured and damaged proteins
– targeted degradation of regulatory proteins: oncoproteins, tumor suppressor proteins
– transmembrane receptors, mitotic cyclins, transcription activating proteins

– modulation of the activity of cell surface receptors
– protein import into cellular organelles
– repair of DNA
– processing and presentation of antigens
– assembly of ribosomes

Of particular interest for regulatory processes are mechanisms by which the activity of growth regulating proteins and central transcription factors are controlled via ubiquitinylation. Often the cell uses signal pathway mediated protein phosphorylation in order to induce the regulated degradation of a signal protein. Examples are the G1 cyclins, the tumor suppressor p53 and the inhibitor IκB.

Two examples will be discussed in more detail.

**Tumor Suppressor Protein p53**

The p53 protein assumes an important role in the control of growth of higher organisms. It functions as a tumor suppressor, that is, it suppresses the growth of tumors (see chapter 14). An inactive, mutant form of p53 is found in nearly half of all human tumors.

p53 is degraded via the ubiquitin-proteasome pathway. The signals which induce p53 degradation are only partially understood (see 14.4.4.4). An interesting aspect of the p53 degradation is the finding that the oncoprotein E6 of human papilloma virus can mediate the ubiquitinylation of p53. The ubiquitinylation of p53 then leads to its degradation, resulting in the loss of an important growth control in the cell (Scheffner et al., 1993). Recognition and conjugation in the ubiquitin system occurs in a complex between the viral E6 protein, a protein associated with E6, E6-AP and p53 (see fig. 2.17). In this reaction, p53 is first bound by the E6 protein. E6-AP functions as the E3 enzyme; it recognizes the p53-bound E6 protein and transfers the ubiquitin in collaboration with the appropriate E2 enzyme to an acceptor lysine on p53. The ubiquitinylation, initiated by the E6 protein, and the ensuing degradation of p53 results in a loss of p53 function, thus offering an explanation for the tumor causing effect of the papilloma virus.

**NFκB**

The transcription activator NFκB regulates a variety of genes involved in the immune response and the inflamatory process. NFκB is required for the expression of genes for the light x-chain of immunoglobulins, interleukin 2 and 6, as well as for interferon b (see chapter 11).

The function and regulation of NFκB is shown schematically in fig. 2.18. The active form of NFκB is a heterodimer consisting of one p50 and one p65 subunit. In the cytosol NFκB is found in an inactive complex bound to the inhibitor JκB. IκB masks the nuclear translocation signal of the heterodimer, thus preventing its transport into the nucleus. The activity of NFκB is highly regulated.

NFκB is activated upon extracellular stimuli, e.g. the action of growth factors, cytokines (see chapter 11) or the exposure to UV light. The signal pathway that leads to

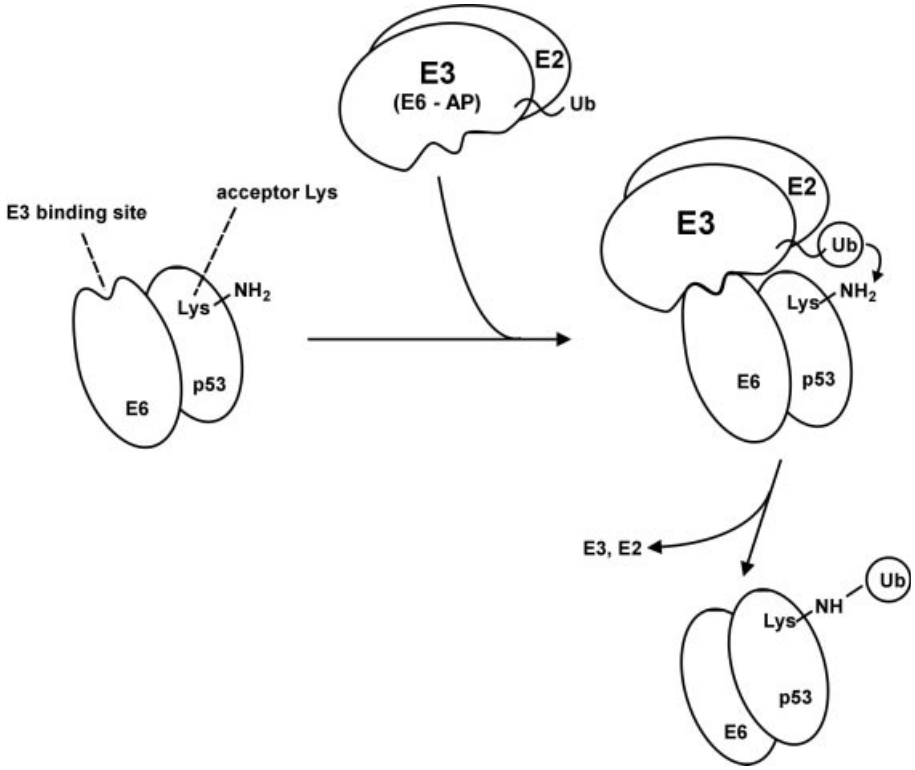



**Fig. 2.17.** Degradation of the tumor suppressor protein p53 by the ubiquitin-proteosome system. The oncoprotein E6 of the human papilloma virus (HPV) forms a specific complex with the p53 protein and can thus induce the degradation of p53. The E6-p53 complex is recognized by E6-AP, a E3 enzyme of the ubiquitin pathway, as a target protein, whereby a ubiquitin residue is transferred to a lysine residue of p53. In this process, the E6 protein serves as the recognition element for ubiquitin ligation of p53.

phosphorylation and subsequent degradation of $I_{\varkappa}kB$ has been well characterized for the cytokines IL-1 and for Tumor Necrosis Factor (review: Maniatis, 1997). Following binding of a cytokine to its transmembrane receptor, a family of specific protein kinases, including a high molecular mass $I_{\varkappa}kB$ kinase complex, is activated to phosphorylate the inhibitor $I_{\varkappa}kB$. This phosphorylation is the signal for ubiquitinylation and degradation of $I_{\varkappa}kB$ . $NF_{\varkappa}kB$ is thus released from its inhibited state to translocate in the nucleus and activate transcription of target genes. .

The ubiquitin-proteasome pathway participates in the regulation of $NF_{\varkappa}kB$ at two points:

- The p50 subunit of $NF_{\varkappa}kB$ results from the proteolytic processing of a 105 kDa precursor protein (p105) in the cytosol. The processing requires the poly-ubiquitinylation of p105 mediated by the 26S proteasome.
- The degradation of the inhibitor protein $I_{\varkappa}kB$ involves the ubiquitin-proteasome pathway.

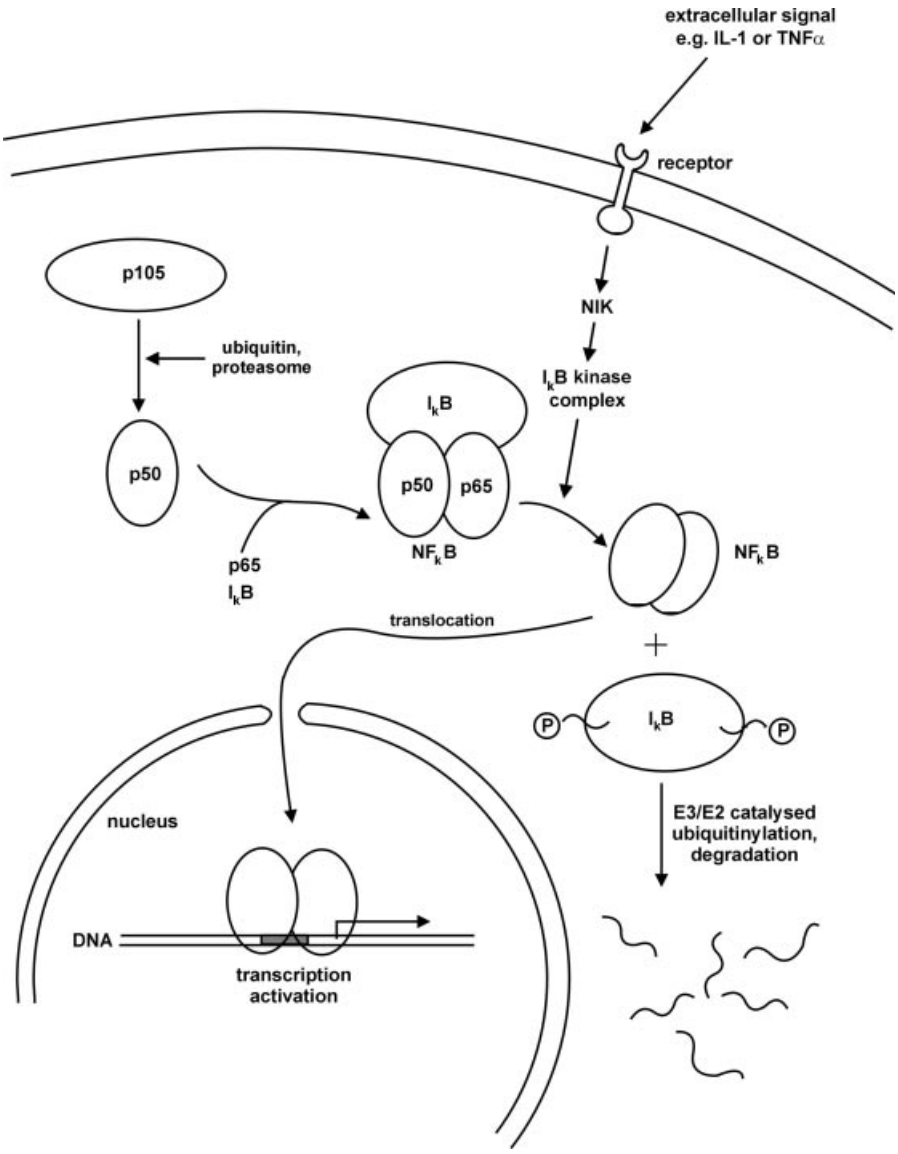



**Fig. 2.18.** Regulation and ubiquitinylation of NFκB and its inhibitor IκB.
The ubiquitin-proteasome path participates by two ways in the regulation of the transcription factor NFκB. The 50 kDa subunit of the heterodimeric NFκB is generated by ubiquitin-dependent proteolysis from a 105 kDa precursor. NFκB is found in the cytosol in a latent, inactive state, bound to the inhibitor IκB. External signalling molecules ( e.g. the cytokine IL-1 or TNFα) activate a signal path that finally leads to the phosphorylation of IκB. At least two different protein kinases, the NIK (a MAPKKK, see chapter 10) and the IκB kinase complex are involved in this signalling. The phosphorylated IκB is recognized by a E3/E2 complex and is ubiqutinylated and targeted for proteolysis in the proteasome. NFκB is now released from the inactive state and can be translocated into the nucleus where target genes (e.g. immunoglobulin genes) are activated.

The ubiquitin-proteasome system thus has significance for $NF_{\varkappa}kB$ in two ways. On the one hand, it participates in the specific processing of the p105 precursor protein to the small subunit of $NF_{\varkappa}kB$. On the other hand, $NF_{\varkappa}kB$ is activated due to the degradation of $I_{\varkappa}kB$.

This example illustrates nicely how extracellular signals can induce the ubiquitinylation and degradation of specific proteins. As shown by the processing of the p105 precursor, ubiquitinylation can be also used for partial proteolysis and for specific activation of a regulatory protein.

# References Chapter 2

Barford, D., Hu, S.H. and Johnson. L.N. 'Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP' (1991) J. Mol. Biol. *218*, 233–260

Baumeister, W., Walz, J., Zuhl, F. and Seemuller, E. 'The proteasome: paradigm of a self-compartmentalizing protease' (1998) Cell *92*, 367–380

Ciechanover, A. 'The Ubiquitin-Proteasom-Pathway' (1994) Cell *79*, 13–21

Fersht, A. 'Enzyme Structure and Mechanism' (1998) Freeman, New York

Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H.D. and Huber, R. 'Structure of 20S proteasome from yeast at 2.4 A resolution' (1997) Nature *386*, 463–471

Hicke, L. and Riezman, H. 'Ubiquitination of a yeast plasma membrane receptor signals its ligand-stimulated endocytosis' (1996) Cell *84,* 277–287

Hurley, J.H., Dean, A.M., Sohl, J.L., Koshland, D.E. and Stroud, R.M. 'Regulation of an enzyme by phosphorylation at the active site' (1990) Science *249*, 1012–1016

Hershko A. and Ciechanover A. 'The ubiquitin system' (1998) Annu Rev Biochem. *67*, 425–479

Johnson L.N. and O'Reilly M. 'Control by phosphorylation' (1996) Curr Opin Struct Biol. *6*, 762–769

Kraut,J. 'How do enzymes work?' (1988) Science 242, 533–540

Krebs, E.G., Graves,D.J. and Fischer, E.H. 'Factors affecting the activity of muscle phosphorylase kinase' (1959) J. Biol. Chem. *234,* 2867–2873

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. 'Crystal structure of the 20S proteasome from the archeon T. acidophilum at 3.4 Å resolution' (1995) Science *268*, 533–539

Maniatis,T. ' Catalysis by a multiprotein IκB kinase complex' (1997) Science *278*, 818–819

Monod, J., Wyman, J. and Changeux, J.-P. 'On the nature of allosteric transitions: a plausible model' (1965) J. Mol. Biol: *12*, 88–118

Palombella, V.J., Rando, O.J., Goldberg, A.L. and Maniatis, T. 'The Ubiquitin-proteasom pathway is required for processing the $NF_{\varkappa}B1$ precurser protein and the activation of $NF_{\varkappa}B$' (1994) Cell *78*, 773–785

Scheffner, M., Huibregste, M. Vierstra, R.D. and Howley, P.M. 'The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53' (1993) Cell *75*, 495–505

Schirmer, T., Evans, R.P. 'Structural basis of the allosteric behaviour of phosphofructokinase' (1990) Nature *343*,140–145

Stryer, L. ' Biochemie' (1991) Spektrum Akad. Verlag, Heidelberg

Varshavsky, A. 'The N-end rule' (1992) Cell *69*, 725–735

Voet, D. and Voet, J.G. 'Biochemie' (1992) Verlag Chemie, Weinheim, pp. 373

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 3**

# Function and Structure of Signaling Pathways

## 3.1 General Function of Signaling Pathways

The enormous structural variety and functional capacity of multicellular organisms is due to their ability to coordinate the biochemical reactions of the various cells of the total organism. The basis for this coordination is the intercellular communication, which allows a single cell to influence the behavior of other cells in a specific manner.

We currently know of various forms of communication between cells (fig. 3.1):

a) *Chemical Messengers*: cells send out signals in the form of specific chemical messengers that the target cell transmits into a biochemical reaction. Signaling cells can simultaneously influence many cells via chemical messengers so as to enable a temporally coordinated reaction in an organism.
b) *Gap Junctions*: Communication between bordering cells is possible via direct contact in the form of „gap junctions". Gap junctions are channels that connect two neighboring cells to allow a direct exchange of metabolites and signaling molecules between the cells.
c) *Cell-cell interaction via cell surface proteins*: Another form of direct communication between cells occurs with the help of surface proteins. In this process a cell surface protein of one cell binds a specific complementary protein on another cell. As a consequence of the complex formation, an intracellular signal chain is activated which initiates specific biochemical reactions in the participating cells.

A further intercellular communication mechanism relies on electrical processes. The conduction of electrical impulses by nerve cells is based on changes in the membrane potential. The nerve cell uses these changes to communicate with other cells at specialized nerve endings, the synapses (see chapter 16). It is central to this type of intercellular communication that electrical signals can be transformed into chemical signals (and vice versa, see chapter 16).

Intercellular signaling influences nearly every physiological reaction. It ensures that all cells of a particular type receive and transform a signal. In this manner, cells of the same type react synchronously to a signal. A further function of signaling pathways is the coordination of metabolite fluxes between cells of various tissues.

In higher organisms intercellular signaling pathways have the important task of *coordinating and regulating cell division*. The pathways ensure that cells divide synchronously and, if necessary, arrest cell division and enter a resting state.

Cellular communication assumes great importance in *the differentiation and development* of an organism. The development of an organism is based on genetic pro-

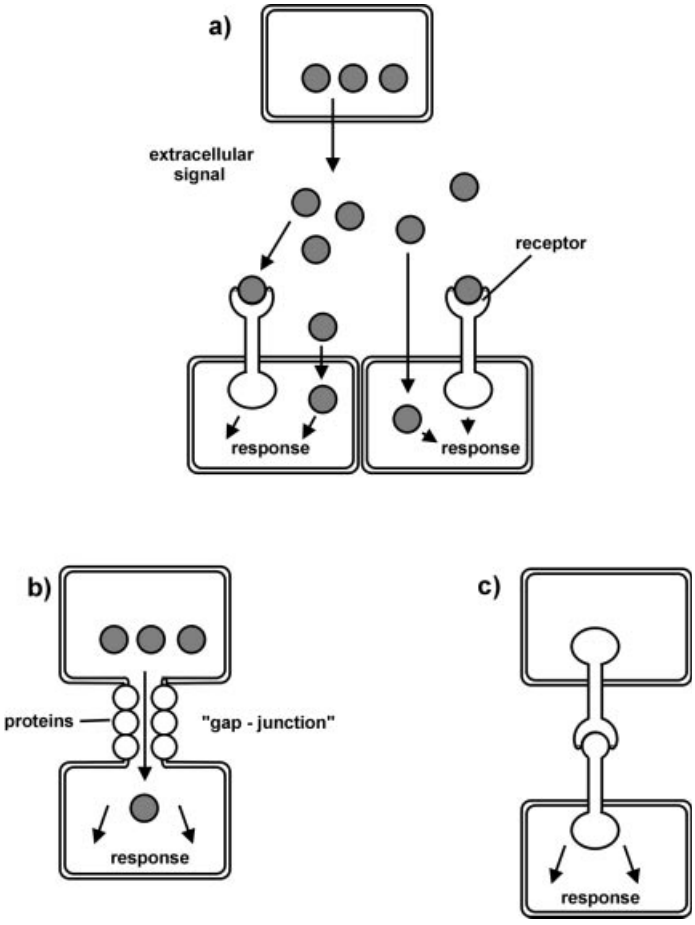



**Fig. 3.1.** Principal mechanisms of intercellular communiciation. a) communication via extracellular messengers and receptor systems. b) communication via gap junctions. Gap junctions are direct connections between cells. They are coated by proteins (drawn as circles) that can have a regulatory influence on the transport. c) communication via cell surface proteins.

grams that always utilize inter- and intracellular signaling pathways. Signal molecules produced by one cell influence and change the function and morphology of other cells in the organism.

Signaling pathways are also critical for the processing of *sensory information*. External stimuli, such as optical and acoustic signals, stress, gradients of nutrients, and so on, are registered in sensory cells and are transmitted to other cells of the organism via signaling pathways.

## 3.2 Structure of Signaling Pathways

Intercellular communication relies on the creation of specific signals by a signaling cell. The signals are registered by a target cell and are thereupon transmitted and processed further with the help of *intracellular* signal chains.

### 3.2.1 The Principle Mechanisms of Intercellular Communication

In the communication between cells of an organism the signals (chemical messengers or electrical signals) are produced in specialized cells. The signal-producing function of these cells is itself regulated, so that the signal is only produced upon a particular stimulus. In this way signaling pathways can be coupled to one another and coordinated.

The following steps are involved in intercellular communication (fig. 3.2):

1) Formation of a signal in the signal-producing cell as a result of an external trigger
2) transport of the signal to the target cell
3) registration of the signal in the target cell
4) further transmission of the signal into the target cell
5) transformation of the signal into a biochemical or electrical reaction in the target cell
6) termination of the signal

A target cell that receives a signal within the framework of intercellular communication transmits the signal in intracellular pathways. These signaling pathways are characterized by the following parameters:

1) the nature of the triggering, external signal
2) mechanism of the registration of the signal
3) mechanism of the transmission and termination of the signal
4) nature of the biochemical reaction induced in the target cell

The sum of these reactions determines the response of the target cell.

#### Nature of the External Signal

Cells can receive and process signals in the form of chemical messengers, and electrical, optical and other stimuli.

#### Reception of Signals by Receptors

Specialized proteins, termed receptors, are utilized for the reception of signals. The reception of the signals by the receptor is equivalent to the binding of a chemical messenger on the receptor or the transmission of physical stimuli into a structural change in the receptor.

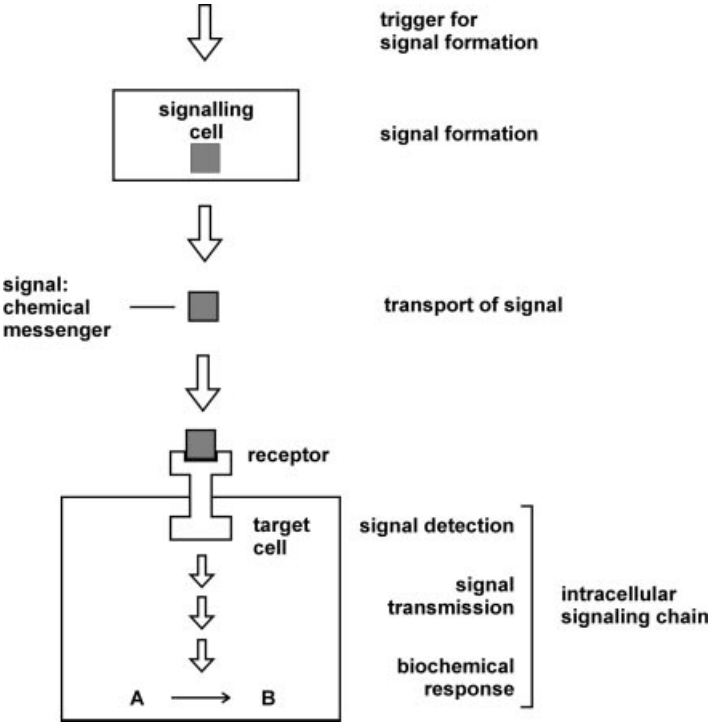



**Fig. 3.2.** The individual steps of intercellular communication. Upon reception of a triggering stimulus, the signal is transformed into a chemical messenger within the signaling cell. The messenger is secreted and transported to the target cell, where the signal is registered, transmitted further, and finally converted into a biochemical reaction. Not shown are processes of termination or regulation of communication which can act at any of the above steps.

There are two principal ways by which target cells can process incoming signals.

- Cell surface receptors receive the signal (e.g. a chemical messenger) at the outside of the cell, become activated and initiate a signaling chain in the interior of the cell. In such signaling pathways the membrane bound receptor transduces the signal at the cell membrane so that it is not necessary for the signal to actually enter the cell.
- The chemical messenger enters into the target cell and binds and activates the receptor localized in the cytosol or nucleus.

### Transmission of the Signal

Upon receiving a signal a receptor becomes activated to transmit the signal further. The *activated receptor* passes the signal onto components, usually proteins, further downstream in the signaling pathway which then become activated themselves for further signal transmission. A chain of serially operating, intracellular signal transduction processes results. Finally, a specific biochemical process is triggered in the cell, which represents the endpoint of the signaling pathway.

**Regulation and Termination**

Signaling pathways always possess multiple mechanisms to regulate the intracellular signal transduction. This allows a specific attenuation or termination of the signal. Very often different signaling pathways communicate with one another. This communication is called crosstalk.

## 3.2.2 Components of the Intracellular Signal Transduction

In the following we will deal with the basic components and principles of intracellular signal transduction and signal processing. The specific reactions and levels of signal transduction will be dealt with in detail in later chapters.

Components of the intracellular signal transduction are proteins and small molecule messengers (fig. 3.3). An incoming signal is passed on from the receptor to down-

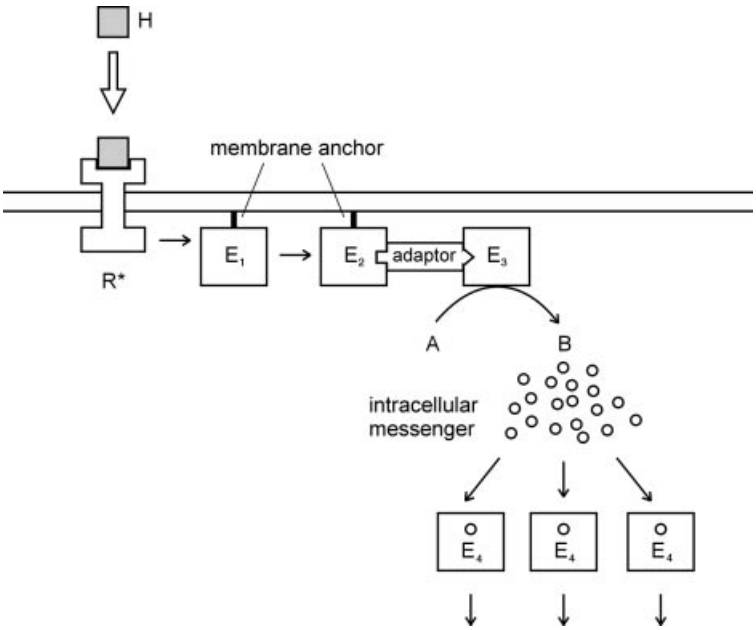



**Fig. 3.3.** Components of intracellular signal transduction. The reception of an extracellular signal by a membrane receptor, shown here as binding of a hormone H to its receptor, activates the receptor for further signal transduction. The activated receptor R* passes the signal onto downstream effector proteins E. Often adaptor proteins mediate and organize the communication between effector proteins. The transduction of signal from the receptor to its downstream effector is frquently a membrane-associated process. The example shown in the diagram above is only to be construed as an example for the composition of a generic signaling pathway. The structure of the intracellular signaling pathways of a cell are highly variable. There are signal transduction pathways that are simpler than the one represented in the figure above, and others that involve many more components and are much more complicated.

stream proteins, which themselves have other proteins as the next partner in the signal cascade. In this manner, further signaling proteins are recruited to act in the signaling chain. The participating signaling proteins can be *enzymes* or they can act as connectors to recruit the other proteins in the signaling pathway. The latter type of proteins are termed *adapters*. The activation of signaling pathway enzymes often leads to the formation of diffusible chemical messengers which transmit the signal further.

### Diffusible Intracellular Messengers

The intracellular activation of enzymes in a signaling chain can lead to the formation of diffusible chemical signaling molecules in the cell. These intracellular signaling molecules are also termed *second messengers*. The second messenger molecules activate and recruit cognate enzymes for the further signal transduction.

### Proteins as Element of Signal Transduction

The most important components of intracellular signal transduction are the protein kinases, protein phosphatases, regulatory GTPases and adapter proteins:

### Protein Phosphorylation

A central tool for signal transmission in a cell is phosphorylation of proteins via protein kinases. Proteins can be reversibly activated or inactivated via phosphorylation. The phosphorylation status of a protein is controlled by the activity of both protein kinases and protein phosphatases (see chapter 7). Both classes of enzymes are elementary components of signaling pathways and their activity is subject to manifold regulation.

### Regulatory GTPases

The regulatory GTPases function as switches that can exist in an active or inactive form. In the active form the GTPases can transmit signals to downstream components in the signaling chain. In the inactive form signal transmission in repressed.

### Adaptor Proteins

Adaptor proteins mediate the signal transmission between proteins of a signaling chain by bringing these proteins together. They function as clamps to co-localize proteins for an effective and specific signaling. Furthermore, adaptor proteins help to target signaling proteins to specific subcellular sites and to recruit signaling molecules into multiprotein signaling complexes. In the latter case, the adaptor proteins may function as a scaffold or docking site for organizing different signaling molecules at distinct sites. The proteins are then also termed docking or scaffolding proteins.

## 3.3 Extracellular Signaling Molecules

The signal-producing cells release their chemical signaling molecules either by exocytosis or passive diffusion into the extracellular space. The messengers reach their target cells via the circulatory system. In special cases, as in the communication between cells of the immune system, membrane bound proteins are also used as signaling molecules. Communication is then only possible upon direct contact between the target cell with the surface of the signaling molecules (fig. 3.1).

Signaling molecules for the communication between cells are known as *hormones.* Hormones that are proteins and regulate cell proliferation are known as *growth factors.*

### 3.3.1 The Chemical Nature of Hormones

The chemical nature of hormones is extremely variable. Hormones can be:

- proteins
- peptides
- amino acids and amino acid derivatives
- derivatives of fatty acids
- nucleotides
- steroids
- retinoids
- small inorganic molecules, such as NO

Table 3.1 shows a selection of hormones from mammals and man.

**Table 3.1.**

**a) Examples for hormones that bind to nuclear receptors**

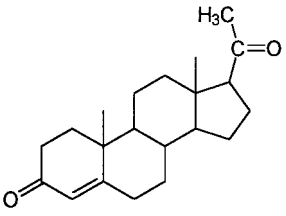
| **Steroids** | **biochemical and/or physiological funktion** |
|---|---|
| Progesterone  | Preparation of the uterus for implantation of the embryo, maintenance of early pregnancy |

**Table 3.1.** continued

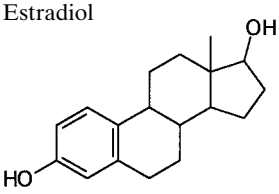
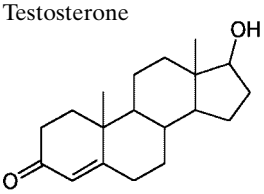
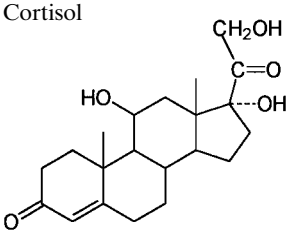
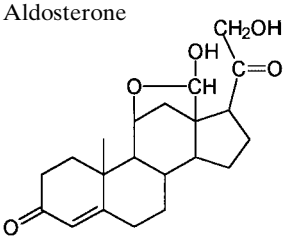
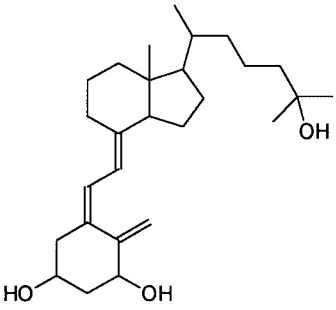
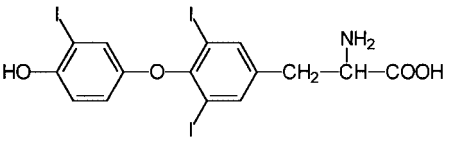
| | | |
|---|---|---|
| Estradiol |  | Preparation of the uterus to receive the blastocyst, control of uterine constraction, generation of secretory system of breasts during pregnancy |
| Testosterone |  | Differentiation and growth of the male reproductive tract, stimulation of male secondary sex characteristics, sceletal muscle growth |
| Cortisol |  | Metabolism of carbohydrates, lipids and proteins, anti-inflammatory, immunsuppressive Induction of Tyr-aminotransferase and of Trp-cyclooxygenase |
| Aldosterone |  | Water and ion balance, backresorption of ions in the kidney |
| **Steroid-related hormones** | | |
| 1, 25-Dihydroxychol-ecalciferol (from vitamine $D_3$) |  | Metabolism of $Ca^{2+}$- and phosphate, bone mineralization, resorption of $Ca^{2+}$ und phosphate in the intestine |

**Table 3.1.** continued

**Other hormones**

3,5,3'-Triiodothyronine, $T_3$-Hormon

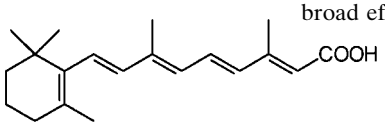
Increased oxygen consumption and increased heat formation, stimulation of glycolysis and of protein biosynthesis,



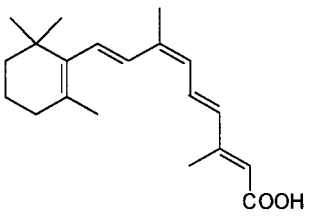

**Retinoids**

All-trans-retinoic acid

Formed from all-trans-retinal, broad effect on differentiation and morphogenesis




9-cis-VitA-Säure




**b) Examples of hormones, that bind to transmembrane receptors**

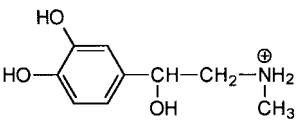
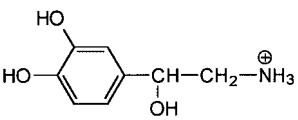
| **Hormone** | **Function and biochemical action** |
|---|---|
| Epinephrine  | Raise of blood pressure, contraction of smooth muscles, glycogen breakdown in liver, lipid breakdown in adipose tissue |
| Norepinephrine  | Contraction of arteria |

**Table 3.1.** continued

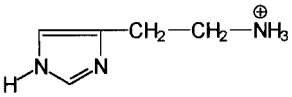
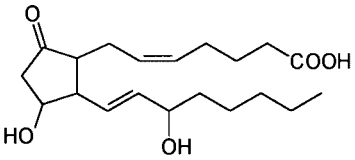
| | |
|---|---|
| Histamine  | Relaxation of blood vessels |
| Derivatives of arachidonic acid | |
| Prostaglandin $E_2$  | Contraction of smooth muscles |
| **Peptide hormones and protein hormones** | |
| Glucagon<br>polypeptide, 29 aa | Glycogenolysis in liver, release of fatty acids from triglycerides in adipose tissue |
| Insulin<br>polypeptid, A-chain, 21 aa,<br>B-Kette 30 aa, | Stimulation of:<br>glucose uptake in muscle and adipose tissue, catabolism of carbohydrates,<br>storage of triglycerides in adipose tissue, protein synthesis, cell proliferation<br>Inhibition of glycogenolysis, |
| Gastrin<br>polypeptidw, 17 aa, | Secretion of HCl und pepsin in stomach |
| Secretin<br>polypeptide, 27 as, | Stimulation of sekretion of pancreatic proteases |
| Adrenocorticotropin<br>polypeptide, 39 aa, | Biosynthesis in anterior pituitary, stimulation of formation of corticosteroids in adrenal cortex,<br>release of fatty acids from adipose tissue, |
| Follicle stimulating hormone (FSH)<br>polypeptide,<br>α-chain 92 As,<br>ß-chain 118 As | Stimulation of growth of oocytes and follicle |
| Thyrotropic hormone<br>Hormon (TSH)<br>polypeptide,<br>α-chain, 92 aa,<br>ß-chain, 112 aa | Release of thyroxine ($T_4$ hormone) and of $T_3$ in thyroid gland |

**Table 3.1.** continued

| | |
|---|---|
| TSH-Releasing-Hormon<br>peptide 3 aa, | Formation in hypothalamus,<br>stimulates synthesis and release of TSH in anterior pituitary |
| Vasopressin<br>peptide, 9 aa, | Formation in posterior pituitary,<br>backresorption of water in the kidney, contraction of small blood vessels |
| Parathyroid hormone<br>polypeptide, 84 aa | Formation in parathyroid gland,<br>increase of $Ca^{2+}$ in the blood, mobilization of $Ca^{2+}$ from the bone |

### 3.3.2 Hormone Analogs: Agonists and Antagonists

The modification of hormones can lead to compounds that are known as agonists or antagonists.

*Antagonists* are hormone derivatives that bind to a receptor but do not initiate signal transduction. Antagonists block the receptor and thus terminate signal transduction. Hormone antagonists find broad pharmaceutical and medical application since they specifically interfere with certain signal transduction pathways in the case of hormonal disregulation. Antagonists with a much higher affinity for a receptor than the unmodified hormone are medically very interesting. Such high affinity antagonists require very low dosages in therapeutic applications. A few important antagonists and agonists of adrenaline are shown in fig. 3.4. Propranolol is an example for a medically important hormone antagonist. Propranolol binds with an affinity three orders of magnitude greater than its physiological counterpart, adrenaline, on the β-adrenergic receptor. In this manner a very effective blockage of the adrenaline receptor is possible.

Hormone analogs that bind specifically to a receptor and initiate the signal transduction pathway in the same manner as the genuine hormone are termed *agonists*. Application in research and medicine is found especially for those agonists which posses a higher affinity for a receptor than the underivatized hormone.

### 3.3.3 Endocrine, Paracrine and Autocrine Signaling

Various forms of intercellular communication can be discerned based on the range of the signal transmission (fig. 3.5).

#### Endocrine Signaling

In endocrine signaling the hormone is synthesized in specific signaling, or endocrine, cells and exported via exocytosis into the extracellular medium (e.g. blood or lymphatic fluid in animals). The hormone is then distributed throughout the entire body via the circulatory system so that remote regions of an organism can be reached.

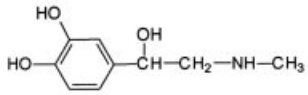
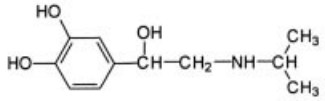
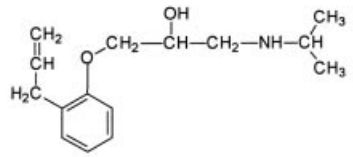
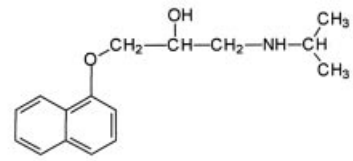
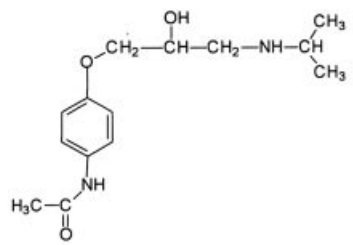
| Structure | compound | $K_D$ for binding to the receptor from frog erythrocytes |
|---|---|---|
|  | adrenaline | $5 * 10^{-6}$ M |
| agonist: | | |
|  | isoproterenol | $0{,}4 * 10^{-6}$ M |
| antagonist: | | |
|  | alprenolol | $0{,}0034 * 10^{-6}$ M |
|  | propranolol | $0{,}0046 * 10^{-6}$ M |
|  | practolol | $21 * 10^{-6}$ M |

**Fig. 3.4.** Structure of important agonists and antagonists of adrenalin and their affinity for the β-adrenergic receptor (source: Lefkowity et al., 1976).

## Paracrine Signaling

Paracrine signal transduction occurs over medium range. The hormone reaches the target cells from the hormone-producing cell by passive diffusion. The producing cell must be found in the vicinity of the receiving cells for this type of communication. The signaling is rather local and the participating signaling molecules are termed *tissue hor-*

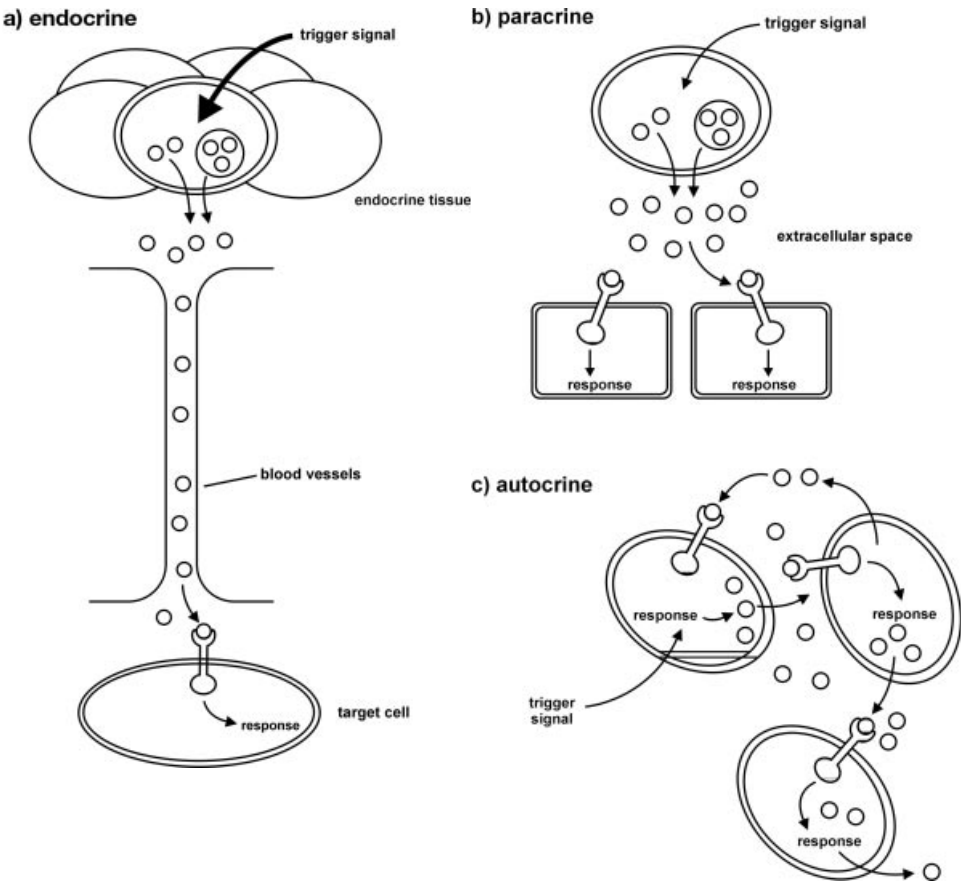



**Fig. 3.5.** Endocrine, paracrine and autocrine signal transduction. a) endocrine signal transduction: the hormone is formed in the specialized endocrine tissue, released into the extracellular medium and transported via the circulatory system to the target cells. b) paracrine signal transduction: the hormone reaches the target cell, which is found in close juxtaposition to the hormone producing cell, via diffusion. c) autocrine signal transduction: the hormone acts on the same cell type as the one in which it is produced.

mones or *local mediators.* A special case of paracrine signal transduction is the synaptic neurotransmission in which a nerve cell communicates with either another nerve cell or with a muscle cell (see chapter 16).

### Autocrine Signaling

In autocrine signaling, cells of the same type communicate with one other. The hormone produced by the signaling cell affects a cell of the same type by binding to recep-

tors on these cells and initiating an intracellular signal cascade. If an autocrine hormone is secreted simultaneously by many cells then a strong response occurs in the cells. Autocrine mechanisms are of particular importance in the immune response (see chapter 11).

### 3.3.4 Direct Modification of Protein by Signaling Molecules

A special case of signal transduction is represented by a class of small, reactive signaling molecules, such as NO (see chapter 6.10). NO is synthesized in a cell in response to an external signal and is delivered to the extracellular fluid. Either by diffusion or in a protein-bound form, the NO reaches neighboring cells and modification of target enzymes ensues, resulting in a change in the activity of these enzymes. NO is characterized as a mediator that lacks a receptor in the classical sense.

## 3.4 Hormone Receptors

### 3.4.1 Recognition of Hormones by Receptors

Hormones are usually produced by specialized cells and initiate a reaction in only a certain cell type. Only those cells that possess a cognate protein, the receptor of the hormone, can act as target cells. Receptors specifically recognize and bind the cognate hormone based on their chemical nature. The binding of the hormone to the receptor in the target cell induces an intracellular cascade of reactions at whose end lies a defined biochemical response. The pathway from receptor bound signaling molecule to final biochemical response is complex and occurs under the participation of many proteins.

The receptors of the target cell can be divided into two classes: the membrane bound receptors and the soluble cytoplasmic or nuclear localized receptors (fig. 3.6).

*Membrane bound receptors* are actually transmembrane proteins; they display an extracellular domain linked to an intracellular domain by a transmembrane domain. Binding of a hormone to the extracellular side induces a specific reaction on the cytosolic side, which then triggers further reactions in the target cell. The mechanism of signal transmission over the membrane will be discussed in more detail in chapters 5, 8 and 11. Characteristic for signal transduction via membrane bound receptors is that the signaling molecule does not need to penetrate the target cell to activate the intracellular signal chain.

In the case of *intracellularly localized receptors* the hormone must enter the cell in order to be able to interact with the receptor. The hormone usually penetrates the target cell by passive diffusion. The nuclear receptors can be classified as ligand-controlled transcription activators. The hormone acts as the activating ligand; the activated receptor stimulates the transcriptional activity of genes which carry DNA elements specific for the receptor.

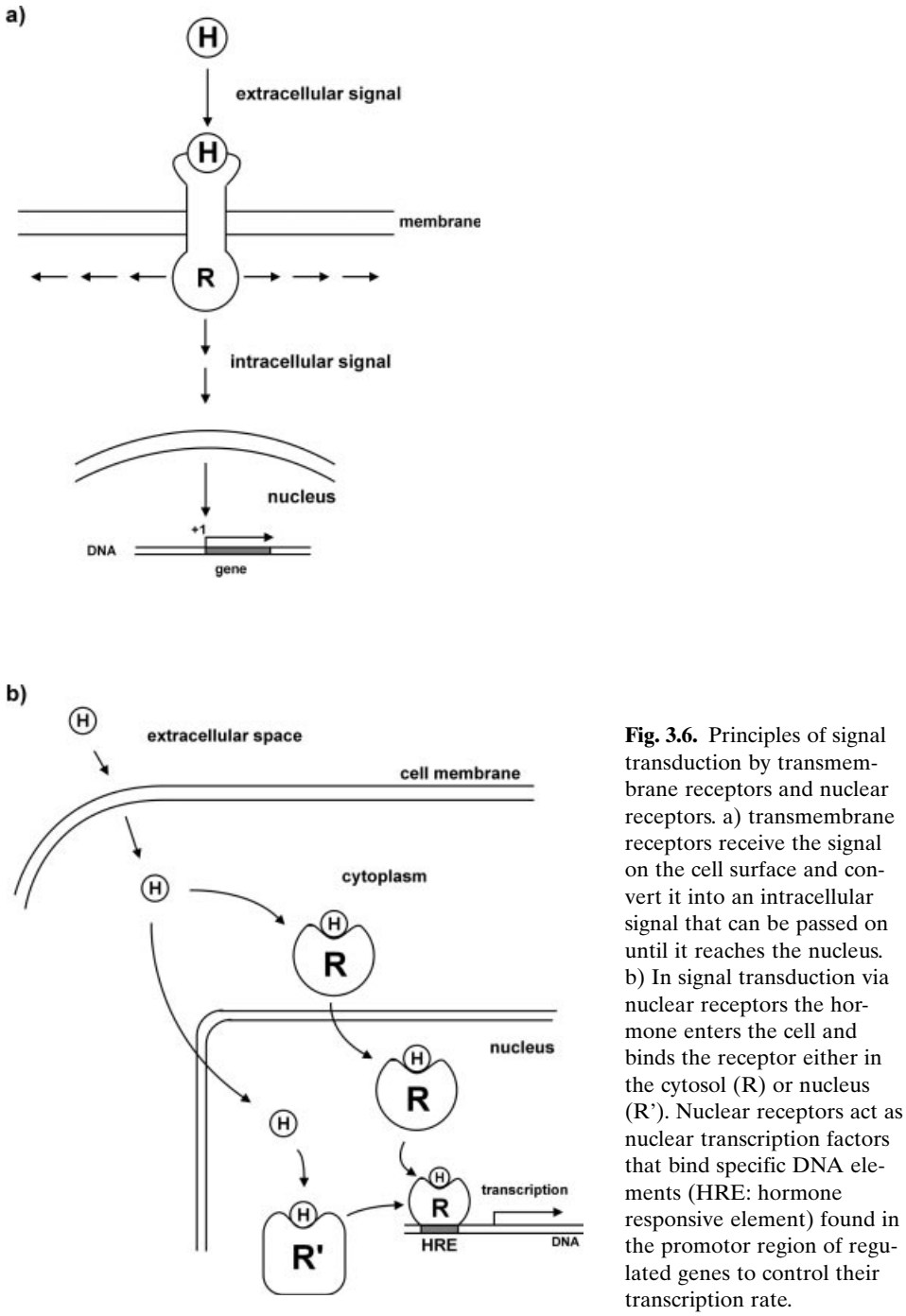



**Fig. 3.6.** Principles of signal transduction by transmembrane receptors and nuclear receptors. a) transmembrane receptors receive the signal on the cell surface and convert it into an intracellular signal that can be passed on until it reaches the nucleus. b) In signal transduction via nuclear receptors the hormone enters the cell and binds the receptor either in the cytosol (R) or nucleus (R'). Nuclear receptors act as nuclear transcription factors that bind specific DNA elements (HRE: hormone responsive element) found in the promotor region of regulated genes to control their transcription rate.

### 3.4.2 The Interaction between Hormone and Receptor

Receptors are the specific binding partners for signaling molecules; the former are able to recognize and specifically bind the latter based on their chemical structure. The binding and recognition are governed by the same principles and the same non-covalent interactions as for the binding of a substrate to an enzyme, namely H-bonds, electrostatic interactions (including dipole-dipole interactions), Van der Waals interactions and hydrophobic interactions. Signaling molecules bind their cognate receptors with an affinity greater than usually observed for an enzyme and substrate.

The binding of a hormone to a receptor can in most cases be described by the simple reaction scheme:

$$[H] + [R] \rightleftarrows [HR], \text{ and } K_D = \frac{[H] \cdot [R]}{[HR]}$$

where [H] is the concentration of free hormone, [R] is the concentration of the free receptor, and [HR] is the concentration of hormone-receptor complex. The value for the equilibrium constant, $K_D$, usually lies in the range of $10^{-6}$ to $10^{-12}$ M. The binding curve for hormones and receptor are, according to the above simple equilibrium, hyberbolic in form (compare binding curve for $O_2$ to myoglobin, fig. 2.3).

Decisive for the intensity of the signal transmission is the concentration of the hormone-receptor complex, since the activation of the signal pathway requires that this complex be formed. The concentration of the hormone-receptor complex depends on the concentration of the available hormone, the affinity of the hormone for the receptor, as well as the concentration of the receptor. All three parameters represent, at least in principle, control points for signal transduction pathways. The variable signal, whose change is registered to thereby activate a signal transmission, is in most cases the concentration of the freely circulating hormone.

Synthesis and release of the hormone in endocrinic tissue is triggered in many hormonally regulated processes by an external signal. As a consequence, the concentration of the circulating hormone is increased, which implies an increased concentration of the hormone-receptor complex at the surface or within the target cell, and thus an increased activation of the downstream components of the signal pathway. For an efficient signal transduction, it is important that the change in concentration of the circulating hormone is approximately proportional to a change in the degree of occupation of the hormone receptor. This condition is fulfilled when the concentration of the hormone is in the range of the equilibrium constant, $K_D$.

The switch for the activation of an intracellular signaling pathway is in most cases an increase in the concentration of the freely circulating hormone. This leads to an increase in the concentration of the hormone-receptor complex, which results in an increased activation of subsequent reactions in the cell. The *concentration of the circulating hormone* is thus the main regulatory parameter in cellular communication. The relation between hormone concentration, binding of the hormone to the receptor, and subsequent reaction in the cell is illustrated in fig. 3.7 for the case of adrenaline and the β-adrenergic receptor.

An increase in the hormone concentration is the main regulating element when a rapid activation of a signaling pathway is required. Hormones can be stored in the

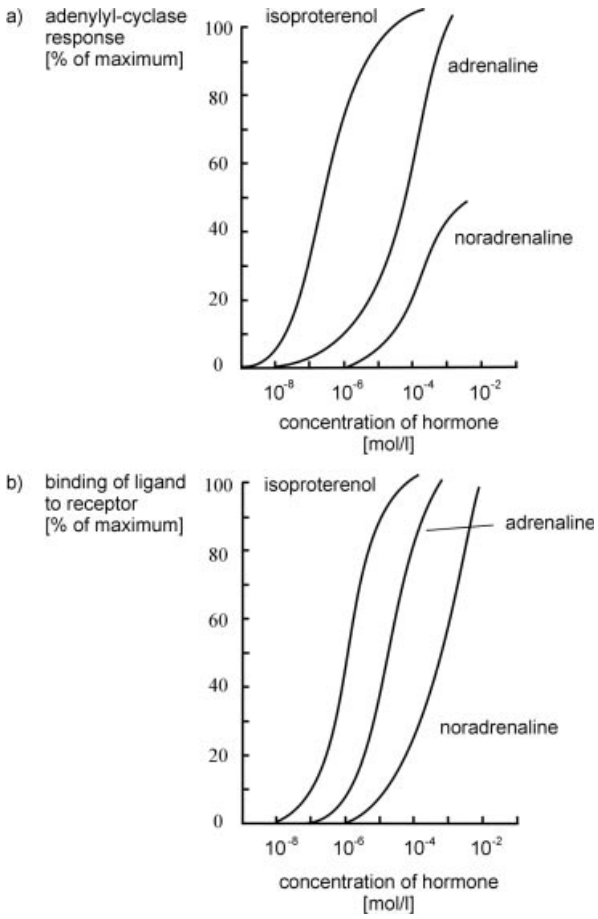



**Fig. 3.7.** Receptor binding and stimulation of the synthesis of cAMP by adrenaline, noradrenaline and isoproterenol. Adrenaline is the first member of a signal transduction chain that begins with the binding of of adrenaline to the β-adrenergic receptor. The signal is transmitted to adenylyl cyclase via a G-protein. The activated adenylyl cyclase forms cAMP, which serves as a diffusible intracellular messenger (see ch. 5& 6). The activation of cAMP synthesis and the binding to the β-adrenergic receptor as a function of hormone concentration in frog erythrocytes is shown in the figure above. a) A cell suspension is incubated with different concentrations of the three hormones, the cells are lysed and the adenylyl cyclase activity determined. b) The degree of binding of adrenaline, noradrenaline and isoproterenol to the β-adrenergic receptor of frog erythrocytes as a function of hormone concentration. The curves show that the ability of a hormone to bind the β-receptor in a particular concentration range is highly correlated with the stimulation of adenylyl cyclase.The concentration of hormone required for half-maximal receptor binding is approximately the same as required for half-maximal adenylyl cyclase stimulation. Of the three ligands, isoproterenol binds the receptor the tightest and noradrenalin the weakest. Accordingly, stimulation of adenylyl cyclase is already observed at a very low isoproterenol concentration, while for the weaker binding noradrenaline much higher concentrations are required to stimulate the adenylyl cyclase. After Lefkowitz et al. (1976).

signal producing cell in specialized organelles from which they can be quickly released when the appropriate external stimulus is received.

A modulation and regulation of signaling is also possible at the level of the receptor by varying the affinity of the receptor for the hormone or by specifically altering the concentration of the receptor. A change in the affinity of the receptor can be achieved by, for instance, phosphorylation of the receptor protein. The concentration of receptor available on the cell surface can be decreased either by degradation of the receptor as well as by internalization of the receptor (see chapter 5). Both processes affect the intensity of the signal transduction on a long time scale.

### 3.4.3 Variability of the Receptor and Signal Response in the Target Cell

Hormone-receptor systems and the corresponding signaling chains display great variability and flexibility (fig. 3.8). There are often many different receptors for a given hormone with similar binding specificity that activate different signaling pathways. Furthermore, there exist subtypes of receptors, which differ in their affinity to the hor-

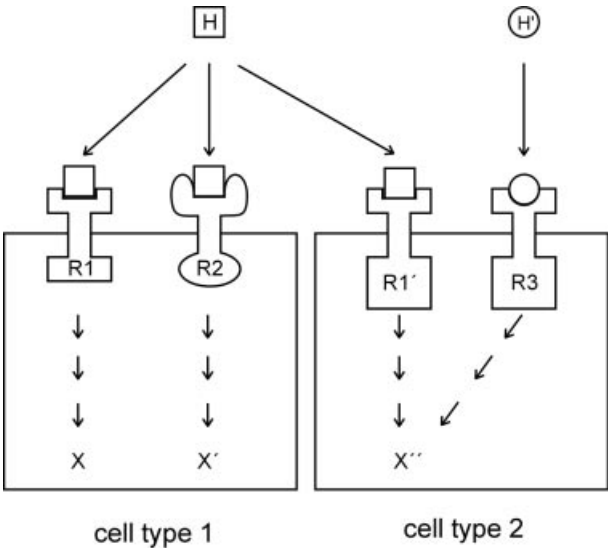



**Fig. 3.8.** Variability of receptor systems and signal pathways. a) For one receptor of a given binding specificity (binding to hormone H) there can be different subtypes in the same cell (R1, R2) or in other cell types (R1'). b) The hormone H can induce different reactions (X, X') upon binding the different receptor types (R1, R2). The receptor types R1 and R2 can be found simultaneous in one cell. c) the binding of two different hormones (H, H') to different receptors (R1', R3) can induce the same intracellular reaction. The characteristics a) and b) contribute to a high degree to the diversity and variability of hormonal signal transduction. Point c) illustrates the principle that important cellular metabolites or reactions can be controlled by different signal transduction pathways.

mone, in the nature and intensity of the reaction triggered in the cell, as well as in their capacity for regulation. Thus, the same hormone can trigger very different reactions in different tissues. An example for such a phenomenon is adrenaline, which on the one hand can initiate a cAMP-mediated signal transduction and, on the other hand, an inositol triphosphate-mediated reaction (see chapter 6). The variability of the receptor system is also illustrated by the fact that in the same cell different receptors for the same signaling molecule can exist which trigger different secondary reactions. An external signaling molecule can thus induce a broad spectrum of biochemical reactions in a cell and a pleiotropy of biological responses.

The highly variable nature of signaling pathways is also expressed by the fact that different receptors and signaling pathways can induce the same biochemical reaction in a cell. This is exemplified by the release of $Ca^{2+}$, which can be regulated via different signaling pathways (see chapters 5–7).

## 3.5 Signal Amplification

Signal pathways commonly amplify the initial signal received by the receptor during the course of the signal transduction (fig. 3.9). In many cases only a few molecules of a hormone are sufficient to initiate an enzymatic reaction in a cell, in which many substrate molecules are turned over.

The extent of amplification, or amplification factor, varies greatly at the different levels of the signal transmission.

An initial amplification often occurs at the level of the hormone-receptor complex. An activated receptor is capable of activating many downstream effector proteins.

The signal amplification at the level of the hormone-receptor complex depends upon many factors:

#### Life Span of the Hormone-Receptor Complex

The life span of the hormone-receptor complex is controlled primarily by the dissociation rate of the bound hormone.

#### Frequency of the Reaction with the Effector Protein

An activated receptor can only transmit the signal further if it encounters an effector molecule. The frequency with which this occurs depends on the concentration and rate of diffusion of both components.

#### Deactivation of the Hormone-Receptor Complex

The signal transmission by the hormone-receptor complex can be actively inhibited via covalent modifications (e.g. protein phosphorylation) which deactivate the hormone-receptor complex. Another mechanism for termination of signaling pathways is the

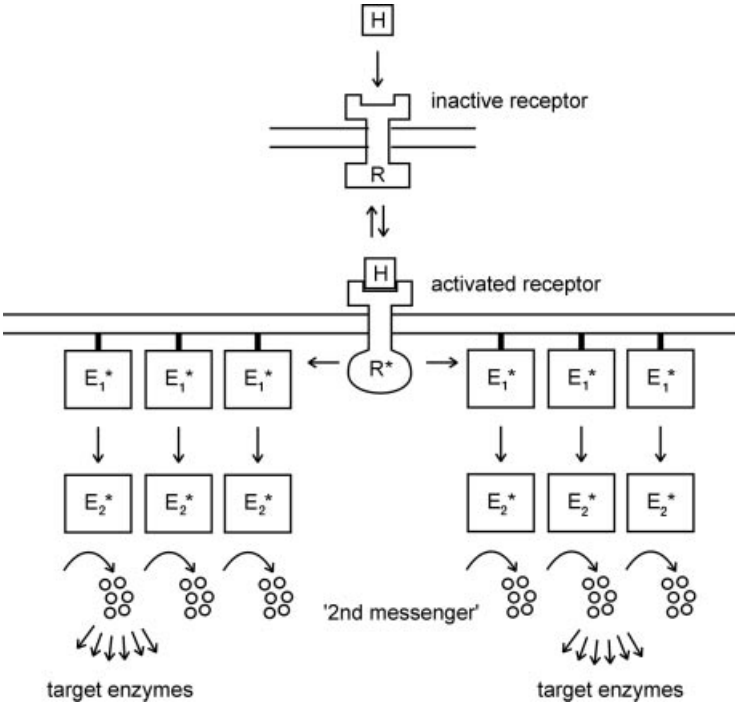



**Fig. 3.9.** Mechanisms of amplification in signal transduction pathways. Starting from an activated receptor R* many downstream effector molecules can be activated. E1* passes the signal on to E2, which is then converted to E2*. If E2 possesses enzymatic activity, and forms a diffusible messenger in its activated form, then a further amplification can occur at this step. The figure also emphasizes that, apart from the steps with a high amplification factor, there are also steps that are not accompanied by amplification (here E1* → E2*).

internalization of the hormone-receptor complex. During internalization a section of the membrane, together with the proteins bound to it, is pinched off and transported into the interior of the cell. There the receptor can be returned to the cell membrane or be degraded. The internalization can affect the free receptor as well as the hormone-receptor complex.

### Amplification of Signaling during the Visual Process

One of the few examples for which the amplification factor at the level of the activated receptor could be determined is for the visual signal transduction pathway (review: Lamb, 1996).

In the visual process a light signal is received by the photoreceptors, rhodopsin, of the rod cells which are then converted to the activated state, R*. The activated rhodopsin passes the signal on to the cognate G-protein, transducin, which in turn activates the next effector molecule, a cGMP phosphodiesterase. The phophodiesterase hydro-

lyzes cGMP to GMP. In the first step of the signal transduction, from the activated photoreceptor R* to transducin, there is a high amplification factor: *one activated rhodopsin molecule can activate 1000–2000 molecules of transducin per second.* No amplification occurs for the signal transmission from transducin to cGMP phosphodiesterase, since each activated transducin molecule only activates one phophodiesterase molecule. A further increase of the signal is found at the level of the activated phosphodiesterase, which rapidly hydrolyzes cGMP to GMP ($k_{cat}$ ca. 4000 sec-1).

The *life time* of the activated state of a signaling protein is an important regulatory point in the signal cascade. As will be discussed later in more detail for the regulatory GTPases (see chapters 5,9), an extension or shortening of the activated state of a signaling protein can lead to an enhancement or attenuation of the signal transmission.

An exact determination and analysis of the amplification factor of signaling cascades *in vivo* is rarely possible. To determine the amplification factor, the life span of the activated state of a signaling protein, the concentration of the signaling protein and its cognate effector molecule, as well as the extent of deactivating processes in the cell must be known. These parameters are very difficult to determine experimentally. The concentration of the participating proteins is an especially elusive parameter. A further complicating factor is the membrane association of many signaling pathways which hinders an accurate concentration determination.

## 3.6 Regulation of Inter- and Intracellular Signaling

The result of communication between the signaling and receiving cells is a defined biochemical reaction in the target cell. The nature and extent of this reaction depends on many individual processes that participate either directly or indirectly in signal transduction.

Beginning with the hormone-producing cell, the following processes are all contributing factors for hormonal signal transduction in higher organisms (fig. 3.10):

1) biosynthesis of the hormone
2) storage, secretion of the hormone
3) transport of the hormone to the target cell
4) reception of the signal by the hormone receptor
5) transmission and amplification of the signal, biochemical reaction in the target cell
6) degradation and excretion of the hormone

All of the above steps are subject to regulation. Thus, the effective concentration of a hormone at the target cell can be adjusted to a wide range according to numerous mechanisms. The biosynthesis of a hormone can, for example, be controlled by other signal transduction pathways. There are signals to trigger the secretion of stored hormones. The distribution of a hormone in an organism contributes to the accessibility of that hormone at a particular location. In addition, degradation and secretion of the hormone also plays an essential role in the effective concentration of the hormone in the cell.

The amount, activity and specificity of receptors at the target cell influences the extent of the final biochemical reaction. The induced reaction cascade in the target cell

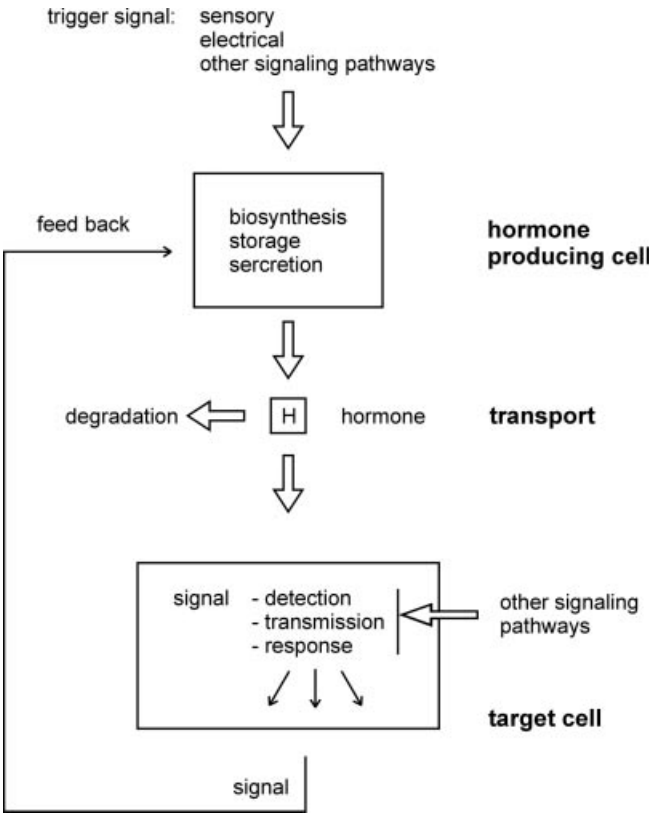



**Fig. 3.10.** Schematic representation of processes that may influence hormonal action in a cell. To note is the possibility for feedback in the framework of intercellular communication. A signal released in the target cell can regulate the hormone producing cell by, for example inhibiting the synthesis or secretion of the hormone. Furthermore, the possibility of a hierarchical structure and the mutual influence of different signaling pathways should also be noted.

can be modulated at many positions by, for example, phosphorylation and subsequent changes in the activity of central proteins.

A signal transduction chain can not be viewed as an isolated event within an organism, but should rather be interpreted in the context of other signaling pathways. The cell possesses a large repertoire of mechanisms by which the extent of signal transduction can be regulated and by which different signaling pathways communicate. Many of these mechanisms will be dealt with in detail in later chapters.

Every individual cell of a multicellular organism is programmed to react to the many external signals in a characteristic and specific manner. The reaction pattern of a cell type depends on the unique pattern of receptors and the corresponding coupled reaction pathways. This furthermore determines the function and morphology of the cell to a large extent. This pattern of regulation and networking of pathways is not constant during the course of development of an organism, but rather is subject to a genetically determined variation.

## 3.7 Membrane Anchoring and Signal Transduction

Transmission of signals over the cell membrane requires cooperation of the signaling proteins, each of which either exist as transmembrane proteins or are associated with the membrane. Extracellular signals are initially transmitted across the membrane with the aid of transmembrane proteins. In many cases the further signal transduction is tightly coupled to the membrane. This signal transduction, localized at the cytoplasmic side of the membrane, occurs with proteins whose function is tightly coupled with an association with the cell membrane.

An example for such a signal transduction can be found in the pathway leading to the production of cAMP, in which the signal is transmitted from a G-protein-coupled receptor to the cytoplasmic side of the membrane (see chapter 5) where membrane-anchored heterotrimeric G-proteins become activated and transmit the signal to downstream transmembrane adenylyl cyclases. The entire activation process occurs in tight association with the cell membrane. A tight coupling between the membrane and signal transduction process is also observed for enzyme activation initiated by signaling processes where the substrate is localized in the membrane. Examples for such processes are found in phospholipases and PI3 kinases (see 5.6).

Often the cell accomplishes the association of signal proteins with the membrane by post-translationally affixed lipid anchors composed of hydrophobic residues, such as fatty acids, isoprenoids or complex glycolipids (see fig. 3.11). These lipid moieties of lipidated proteins favor membrane association by inserting themselves into the phospholipid bilayer.

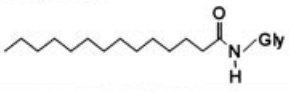
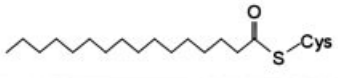
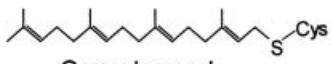
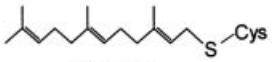
| lipid | examples of modified proteins | site of modification |
|---|---|---|
| N-myristoyl   | heterotrimeric G-proteins (α-subunit), see chapter 5<br><br>cytoplasmic tyrosine kinases, see chapter 8 | N-terminus |
| S-palmitoyl (S-Acyl)   | heterotrimeric G-proteins (α-subunit),<br>G-protein-coupled receptors, see chapter 5<br>Ras proteins, see chapter 9 | internal, no distinct consensus sequence |
| S-prenyl   Geranylgeranyl   Farnesyl | heterotrimeric G-proteins (γ-subunit)<br><br>Ras proteins<br>rhodopsin kinase<br>see chapter 5 | C-terminus |

**Fig. 3.11.** Structure of lipid anchors and representative examples for lipid-modified signal proetins.

The main function of lipidation is to promote membrane association of signaling proteins. Lipid anchors target proteins to the membrane, as is the case for the cytoplasmic protein tyrosine kinases, so that they can participate in membrane-associated signaling pathways. Furthermore, protein lipidation is thought to mediate protein-protein association and/or stabilize protein conformations (review: Casey, 1995).

**Switch Function of Lipid Anchors**

Lipid anchors can be also used as a switch in signaling pathways. Depending on the conformational state of the signaling protein, lipid anchors may be buried in the hydrophobic interior of the protein or they may be exposed on the protein surface and accessible for membrane insertion. The transition between the two states may be controlled by specific ligands in a signal pathway controlled manner (fig. 3.12). Examples are the $Ca^{2+}$-myristoyl switch of recoverin (see 6.7.3) and the GTP-myristoyl-switch of the ARF GTPase (Goldberg, 1998). In both cases ligand-induced conformational changes of the signaling protein are coupled to membrane binding.

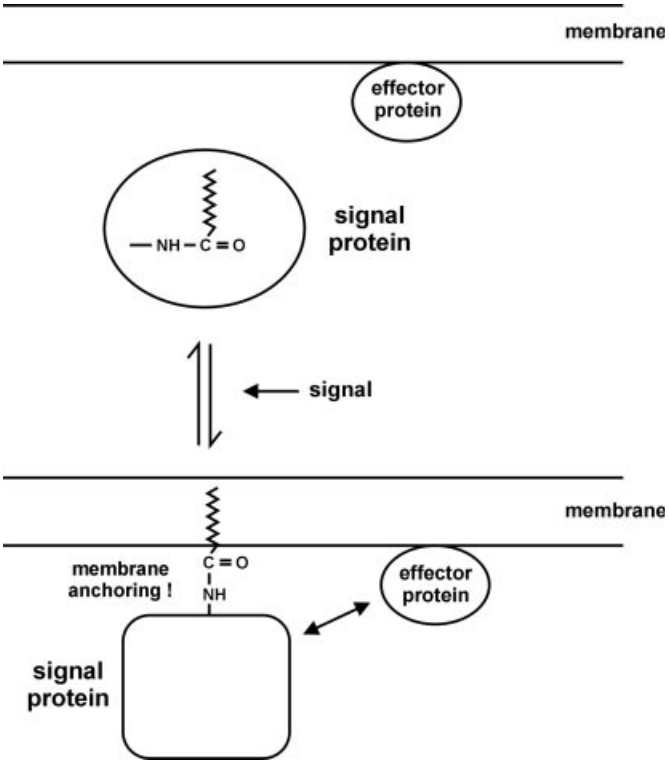



**Fig. 3.12.** Model of the switch function of the myristoyl anchor in signal proteins.
The myristoyl anchor of a signal protein can exist in a state accessible for membrane insertion or in a state buried in the interior of the protein. The transition between the two states may be controlled by specific cellular signals (e.g. $Ca^{2+}$, GDP/GTP exchange). In the membrane-associated form, interactions with membrane-bound effector proteins become possible and the signal can be transduced further.

### 3.7.1 Myristoylation

Myristoylated proteins contain a saturated acyl group of fourteen carbons, myristoic acid (n-tetradecanoic acid) added cotranslationally via an amide bond to the amino group of the $NH_2$-terminal glycine residue. The consensus sequence for myristoylation is Gly-X-X-X-Ser (X:any amino acid), where the residue following the glycine is often a cysteine (review: Resh, 1994). A clumping of basic amino acids at the N-terminus can serve as additional signal for myristoylation (fig. 3.13). An example for this is c-Src (see chapter 8), which displays a net charge of +5 at the N-terminal region. Clusters of basic residues at the C-terminus also aid to anchor the protein to the interface of the membrane, as observed in the example of the K-ras protein (see chapter 9 and fig. 3.12). The basic amino acids help to anchor the protein to the cell membrane via electrostatic interactions with the negatively charged head groups of the phospholipids.

Myristoylation is generally considered a constitutive process and a permanent modification. As shown above the myristoic anchor may function as a switch during regulated membrane anchoring. Examples for myristoylated proteins are the cytoplasmic protein tyrosine kinases (family of the Src-kinases, chapter 8), as well as the α-subunit of the heterotrimeric G-proteins (chapter 5).

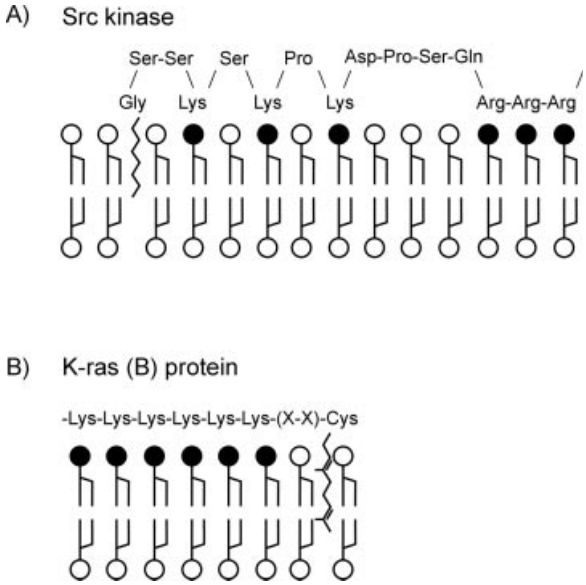



**Fig. 3.13.** Lipid anchors and basic regions as elements of the membrane association of proteins. Examples for proteins which exhibit basic residues near a lipid anchor. a) Src kinase (see ch. 8) possess a myristoyl anchor at the N-terminus as well as a stretch of basic residues. b) In Ki-Ras proteins (see ch. 9) there is a farnesyl residue at the C-terminus that serves as a lipid anchor, as well as a stretch of Lys residues. Negatively charged head groups of phospholipids are shown as filled circles. X: any amino acid.

### 3.7.2 Palmitoylation

Palmitoylated proteins contain a long-chain fatty acid, such as palmitoic acid (n-hexadecanoic acid), or other long chain fatty acids connected to the protein via a labile thioester bond to cysteine residues. The thioester bond of palmitoylated proteins is less stable than the amide bonds of the myristin anchor. The lability conveys a reversible character to the modification and thus permits regulation of the membrane anchor. The reversible palmitoylation of signal proteins is thus a potential instrument for the modulation or regulation of signaling pathways. In the activation of G-protein-mediated signal pathways (see ch. 5), the palmitoic acid anchor on the Gα-subunit is exchanged (review: Mumby, 1997). The distribution of the signal protein between the membrane and cytosol is possibly regulated via a cyclic acylation and deacylation. The acylation and deacylation enzymes have only been partially characterized.

Palmitoylation is, after myristoylation, the most common modification of the α-subunit of the heterotrimeric G-proteins (see chapter 5). The α-subunit of G-proteins can be lipidated in a two-fold manner, with a myristoic acid and a palmitoic acid anchor at the N-terminus. It appears in this case that two lipid anchors are necessary to mediate a stable association of the protein with the membrane. The lipidation of cytoplasmic protein tyrosine kinase also includes both myristoylation and palmitoylation. H-Ras protein also requires, apart from C-terminal farnesylation (see below), a palmitoyl modification in order to bind to the plasma membrane. In all mentioned examples the fatty acid anchors play an essential role in the signal transduction.

### 3.7.3 Farnesylation and Geranylation

Proteins with an isoprenoid modification possess either a C15-farnesyl residue or a C20-geranyl-geranyl residue. Both residues are bound via a thioester linkage to a cysteine residue. As with myristoylation, these are constitutive, stable modification performed by farnesyl or geranyl transferases.

The isoprenylation occurs at the Cys-residue of the consensus sequence Cys-A-A-X-COOH, whereby the nature of the C-terminal X-residue determines if farnesylation or geranylation occurs (fig. 3.13+14). After the isoprenoid residue is appended the three C-terminal residues are removed and the new COOH-group of the Cys-residue is methylated to increases the hydrophobicity of the C-terminus. A two-fold geranylation is found on two Cys residues of the Rab protein (see chapter 9).

The isoprenoid modification can be found, among others, on the Ras protein and other members of the Ras superfamily (see chapter 9), as well as with the α-subunit of G-protein (see chapter 5). The βγ-complex of G-proteins is also associated with the membrane via geranylation.

### 3.7.4 The Glycosyl-Phosphatidyl-Inositol Anchor (GPI Anchor)

GPI is an *extracellular* anchor for proteins. The GPI anchor consists of a phospholipid with an appended glycosyl and ethanolamine residue in a complicated arrangement (fig. 3.15). It is the most commonly employed anchor for the surface proteins of *Trypa-*

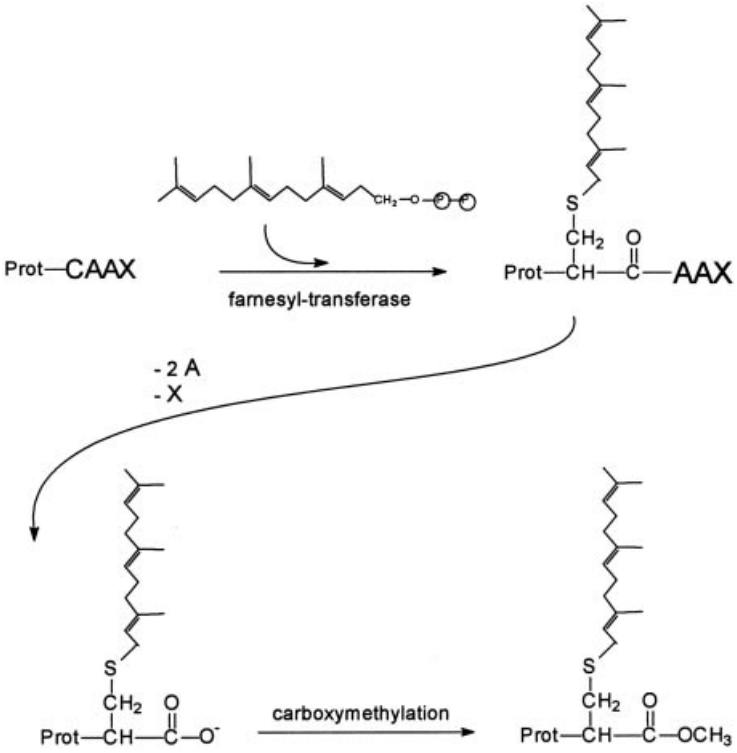



**Fig. 3.14.** Farnesylation at the C-terminus. The signal sequence for farnesylation is the C-terminal sequence CAAX. In the first step a farnesyl moiety is transferred to the cystein in the CAAX sequence. The farnesyl donor is farnesyl pyrophosphate and the responsible enzyme is farnesyl transferase. Subsequently, the three C-terminal amino acids are cleaved (A: alanine, X: any amino acid) and the carboxyl group of the N-terminal Cys-residue becomes methylated.

*nosoma brucei* and prevalent in yeast (review: Takeda and Kinoshita,1995). The function of the GPI anchor in mammals is not yet understood in detail. Cell surface proteins associated with the membrane via a GPI anchor participate in the uptake of nutrients, cell adhesion, and cell-cell interactions in the immune system. In T-lymphocytes, GPI-anchored proteins participate in signal transduction processes which lead to the activation of T-lymphocytes.

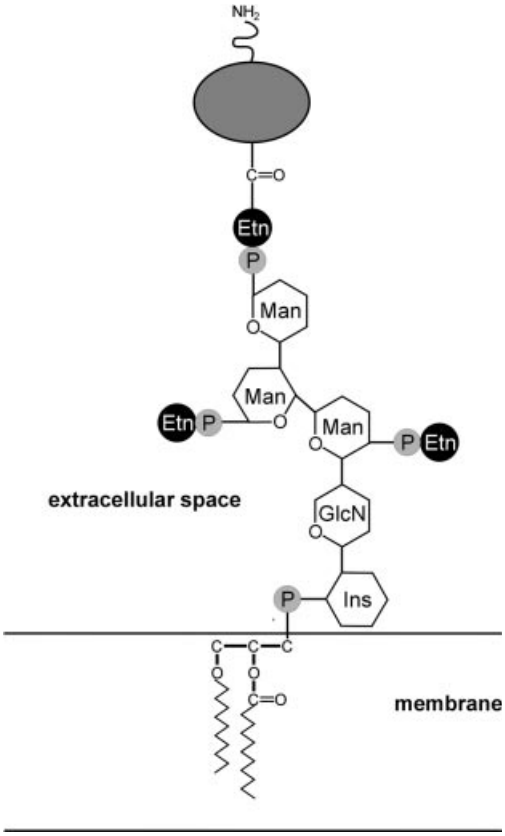



**Fig. 3.15.** Typical strucutre of a glycosyl phosphatidyl inositol (GPI) anchor. Ins: inositol; GlcN: 2'-amino, 2'-deoxy-glucose; Man: mannose; Etn: Ethanolamine, P: phosphate.

# References Chapter 3

Ames, J.B., Ishima, R., Tanaka, T., Gordon, J.I., Stryer, L. and Ikura, M. 'Molecular mechanics of calcium-myristoyl switches' (1997) Nature *389*, 198–202

Casey, P.J. 'Protein lipidation in cell signalling' (1995) Science *268*, 221–225

Goldberg J 'Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching' (1998) Cell *95*, 237–248

Lamb, T.D. 'Gain and kinetics of activation in the G-protein cascade of phototransduction' (1996) Proc. Natl. Acad. Sci. U.S.A. *93*, 566–570

Lefkowitz, R.J.; Limbird, L.E.; Mukherjee, J. and Caron, M.G. 'The ß-adrenergic receptor and adenylate cyclase' (1976) Biochem. Biophys. Acta *457*, 1–55

Mumby, S.M. 'Reversible palmitoylation of signaling proteins' (1997) Curr. Op. Cell Biol. *9*, 148–154

Resh, M.D. 'Myristylation and palmitylation of src family members: the fats of matter' (1994) Cell *76*, 411–413

Takeda, J. and Kinoshita, T. 'GPI-anchor biosynthesis' (1995) Trends Biochem. Sci. *20*, 367–370

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

# Chapter 4
# Signaling by Nuclear Receptors

Nuclear receptors regulate gene expression in response to binding small lipophilic molecules and are thereby involved in the control of a diversity of cellular processes. These proteins are ligand-activated transcription factors that are localized in the cytoplasm and/or in the nucleus. The hormone ligands pass the cell membrane by simple diffusion and bind to the cognate receptors in the cytoplasm or in the nucleus. By binding to cognate DNA elements the ligand-bound receptor activates target genes and thus transmits hormonal signals into a change of gene expression.

## 4.1 Ligands of Nuclear Receptors

The naturally occurring ligands of nuclear receptors are lipophilic hormones, among which the steroid hormones, the thyroid hormone T3, as well as derivatives of vitamin A and D have been long known as central regulators. These hormones play a significant role in metabolic regulation, organ function and in development and differentiation processes. The most important natural ligands of the nuclear receptors are shown in fig. 4.1; the cognate receptors and their DNA elements are summarized in table 4.1 (see Mangelsdorf, 1995).

The primary regulating factor for signal transduction via nuclear receptors is the hormone concentration. The amount of a hormone available for intracellular binding to the receptor is controlled by a number of processes:

### Formation and Secretion of Hormones in Endocrine Tissue

The biosynthesis and secretion of a hormone is regulated by a number of mechanisms and they themselves are involved in complex regulatory schemes. An informative example is the hierarchically ordered hypothalamus-hypophysis system which regulates the biosynthesis and secretion of many hormones of the nuclear receptors.

The hypothalamus is a top-level regulatory switch which registers and integrates electrical and chemical signals. The signals are transmitted further in the form of *releasing factors* to the downstream hormone-producing organ, the hypophysis, or pituitary gland where *glandotropic hormones* are synthesized and secreted which reach the endocrine organs, such as the adrenal cortex or gonads, by way of the bloodstream. In the target organs the production and secretion of specific hormones is induced. The hormones secreted by the endocrine organs are transported in the bloodstream to the target cells where they bind and activate cognate receptors. In the target cells intracellular signal pathways are activated which determine the biochemical response of the

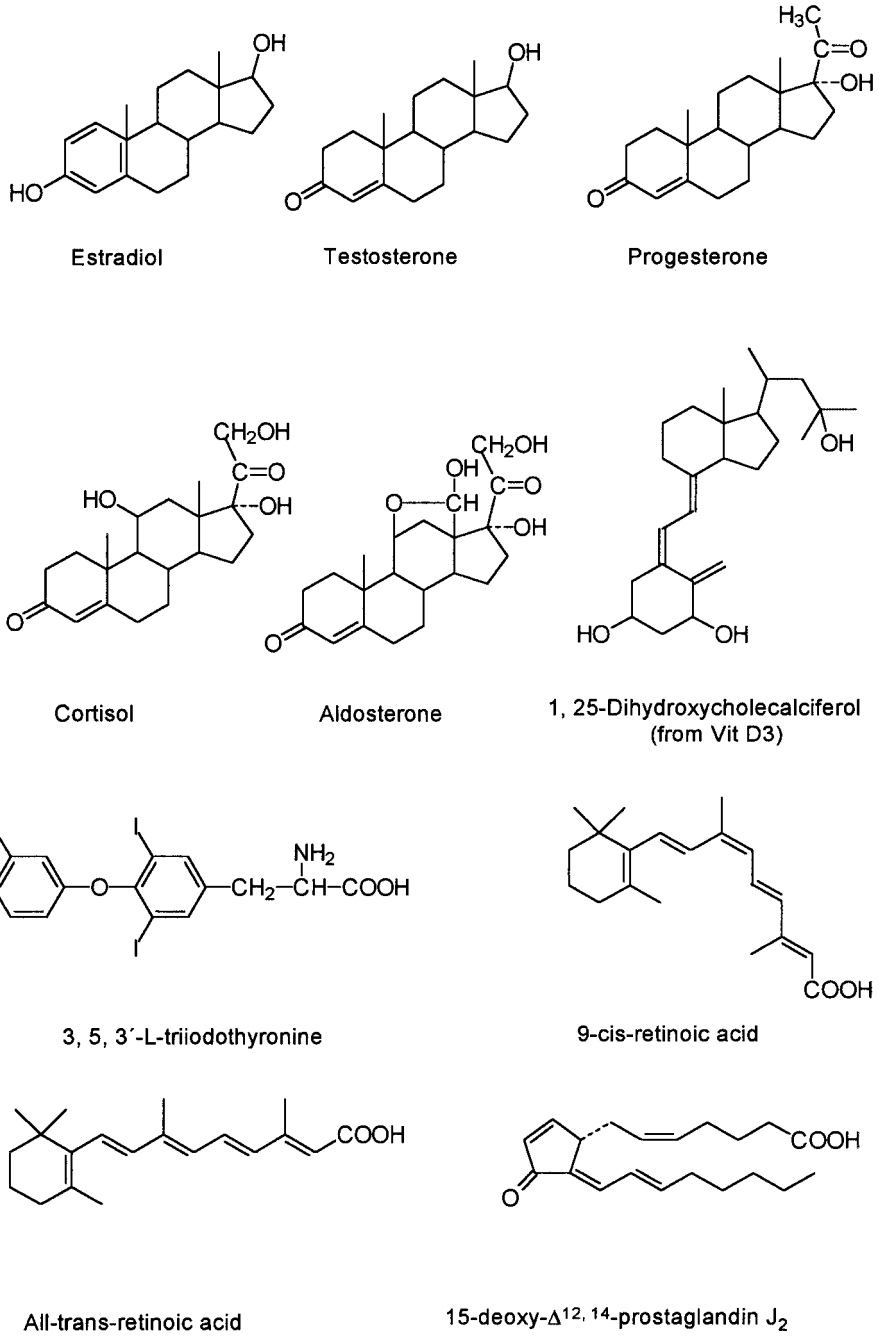



**Fig. 4.1.** Natural ligands of nuclear receptors.

**Tabelle 4.1.** Ligands and structure of HREs of selected nuclear receptors from mammals.

| Receptor | Hormone | Half site sequence | configuration of HRE |
|---|---|---|---|
| Glucocorticoid receptor, GR | Cortisol | AGAACA | IR-3 |
| Mineralcorticoid receptor, MR | Aldosterone | AGAACA | IR-3 |
| Progesterone receptor, PR | Progesterone | AGAACA | IR-3 |
| Androgen receptor, AR | Testosteron | AGAACA | IR-3 |
| Estrogen receptor, ER | Estrogen | RGGTCA | IR-3 |
| Farnesoid receptor, FXR | Farnesoids | AGGTCA | IR-1, DR-5 |
| Thyroid hormone receptor, T3R | 3,5,3'-L-triiodothyronine, $T_3$ hormone | RGGTCA | IR-0, DR-4, ER-6,8 |
| Receptor for vitamine $D_3$, VDR | 1,25-Dihydroxy vitamine $D_3$ | RGKTCA | DR-3 |
| Receptor für 9-cis retinoic acid, RXR | 9-cis retinoic acid, Terpenoids | AGGTCA | DR-1 |
| Receptor for all-trans retinoic acid, RAR | all-trans retinoic acid | AGTTCA | IR-0, DR-2,5; ER-8 |
| Peroxisome proliferator-activted receptor, PPAR | 15-Deoxy-Δ12,14-prostaglandin $J_2$ | AGGTCA | DR-1 |
| COUP-TF(α, β, γ) | ? | RGGTCA | DRs, IRs |
| NGFI-B,( α, β, γ) | ? | AAAGGTCA | NR, DR-5, IR-O |
| ROR (α, β, γ) | ? | YYCYRGGTCA | NR |

**Abbreviations:** IR: "inverted repeat"; DR: "direct repeat"; ER: "everted repeat"; NR: "no repeat". Numbers of the HREs give the number of pairs separating the half-sites. R: purine, Y: pyrimidine. NGFI: Nerve growth factor induced receptor; COUP-TF: Chicken ovalbumin upstream promoter transcription factor; ROR: Retinoic acid related orphan receptor. α, β und γ are receptor subtypes coded by distinct genes. After Mangelsdorf et al., 1995.

target cell to the hormonal stimulation. The communication within the various levels of this control system occurs via hormone receptor systems.

The system is constructed hierarchically which allows an amplification of the signals registered in the hypothalamus. The hypothalamus-pituitary path can uptake nerve impulses and translate them into specific hormone production. A further characteristic of the system are regulatory mechanisms with whose help the formation and secretion of hormones can be controlled. The circulating hormone exerts a *feedback inhibition* at various levels thus ensuring that an overproduction of hormones is avoided.

The hypothalamus-pituitary system is of particular importance for steroid hormones. The formation of steroid hormones in the adrenal cortex is controlled by the pituitary hormone corticotrophin (ACTH), whose formation in turn is regulated by the corresponding releasing factor of the hypothalamus, corticotrophin releasing hormone (CRH).

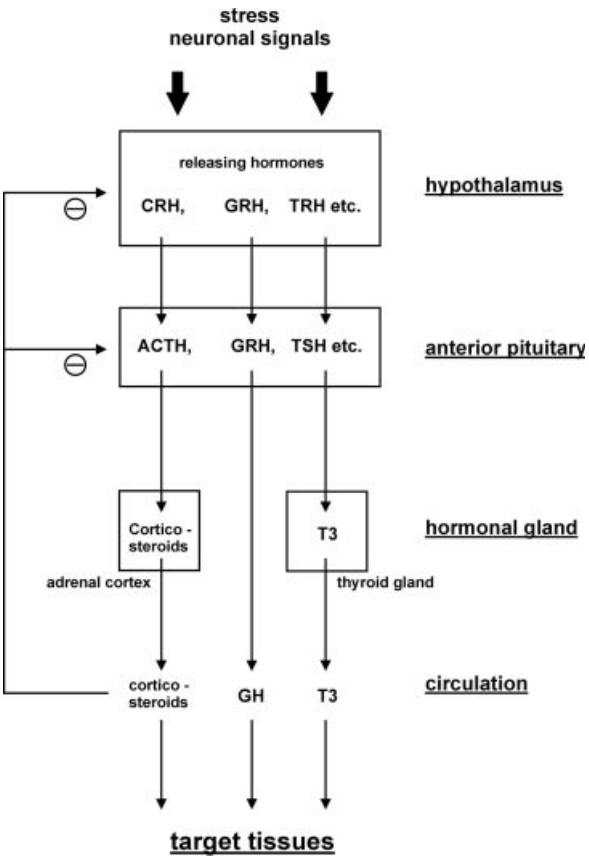



**Fig. 4.2.** Hierarchy of hormone producing organs as observed for the hypothalamus-pituitary system. The hormone cascades of the the hypothalamus-pituitary system are strictly regulated. The concentration of circulating hormone is registered by the superior organs of the system and an overproduction of hormone is prevented via feedback regulation. CRH: corticotropin releasing hormone; GRH: growth hormone releasing hormone; TRH: thyrotropin releasing hormone; ACTH: corticotropin; GH: growth hormone; TSH: thyrotropic hormone.

In addition to biosynthesis and secretion the degradation and excretion of hormones also influences their concentration in the body.

### Availability of Hormones in the Cytosol

The intracellular availability of a hormone can be modulated by its binding to a specific intracellular storage proteins. For the binding of all-trans-retinol and all-trans retinoic acid there exist specific binding proteins which participate in the storage and concentration of both retinoids. The cytoplasmic storage forms can have a regulatory influence on the metabolization and availability of a hormone in the cytosol. Furthermore, there are examples known of hormones produced from a pre-hormone in the target cell (see 4.5.3).

### Modifications of the Hormone in the Target Tissue

After the hormone enters the target cell it can be specifically modified such that the concentration of active, binding-proficient hormones is controlled by the modification.

An example of the importance of enzymatic modification of hormones for the tissue specificity of hormone action is the effect of the mineral corticoid aldosterone in the presence of a large excess of the glucocorticoid cortisol. Aldosterone regulates the $Na^+$-export and $K^+$-retention in the kidney by binding on the aldosterone receptor.

Aldosterone circulates in the blood at much lower concentration than the most important glucocorticoid cortisol. In view of the fact that the aldosterone receptor displays the same affinity to aldosterone as for cortisol, it was long puzzling how aldosterone could exercise a specific regulatory effect on the kidney in the presence of the large excess of cortisol. An explanation was provided when it could be shown that the kidney performs a specific modification of the cortisol via an 11β-OH-steroid dehydrogenase (fig. 4.3). The enzymatic conversion of the 11-OH group of cortisol to an 11-keto group eliminates the binding capacity to the aldosterone receptor (Funder, 1988). The dehydrogenase, however, does not act on the aldosterone because it is protected at the 11 position. The formyl group of aldosterone reacts with the 11-OH-group under the formation of an aldehyde hydrate. The specific occurrence of the 11β-OH-deydrogenase in the kidney cells thus ensures the signal transduction by aldosterone in the presence of excess cortisol. This example demonstrates the important contribution of tissue specific modification toward the tissue specificity of hormone action.

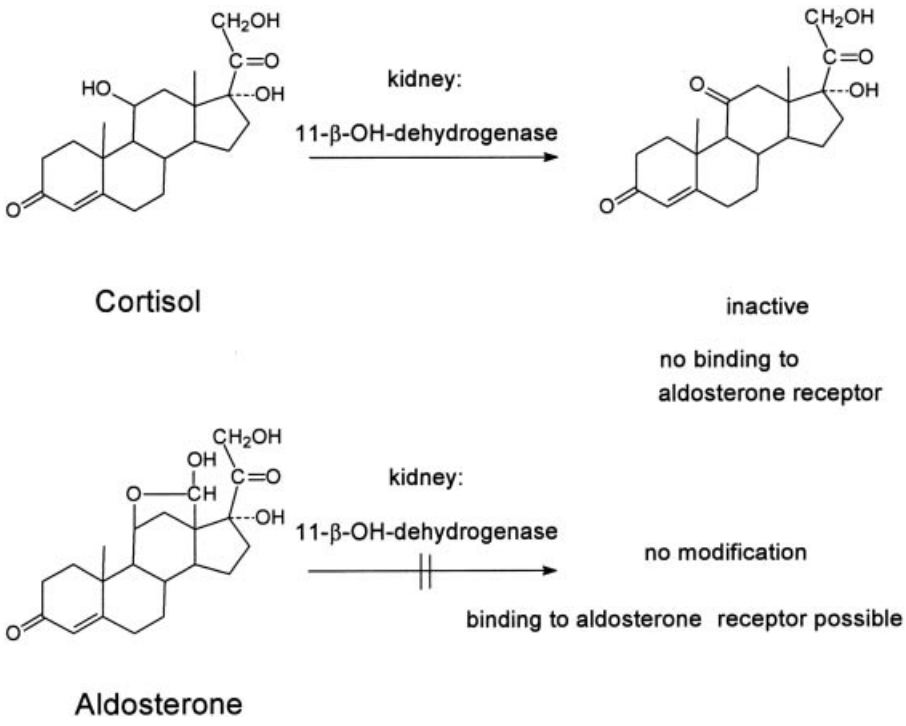



**Fig. 4.3.** Example of a tissue-specific hormone inactivation. Cortisol is enzymatically converted to an keto compound at the C11 position. The modified cortisol compound can no longer bind and activate the aldosterone receptor. Aldosterone can not be turned over by 11-β-dehydrogenase because the OH group at position 11 forms a half-acetal with the formyl group (C18).

## 4.2 Principles of Signaling by Nuclear Receptors

Signal transduction by nuclear receptors is shown schematically in fig. 4.4. The natural ligands of nuclear receptors are lipophilic hormones that can enter the cell in a passive manner. A transmittance of the signal at the cell membrane, is not necessary, as is the case in the signaling via transmembrane receptors. Once inside the cell the hormone

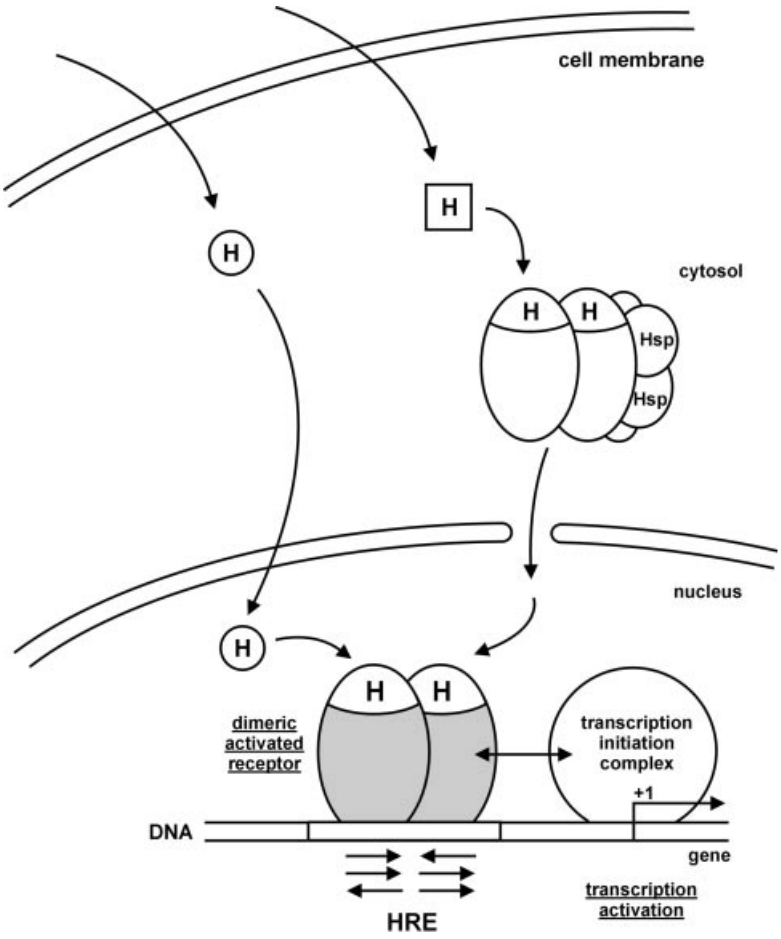



**Fig. 4.4.** The principle of signal transduction by nuclear receptors. Nuclear receptors are ligand-controlled transcription factors that bind cognate DNA sequences, or hormone responsive elements (HRE). The hormone acts as a regulating ligand. Most nuclear receptors bind their cognate HREs, which tend to be symmetrically organized, as homo- or heterodimers. The DNA-bound, activated receptor stimulates transcription initiation via direct or indirect protein-protein interactions with the transcription initiation complex. The arrows demonstrate the different possible configurations of the HRE (see also 4.6). H: hormone; Hsp: heat shock protein.

ligand binds the cognate receptor which is localized in the cytosol and/or in the nucleus. The hormone binding activates the transcription regulation function of the receptor. In the case of the cytosolic receptors, the hormone binding induces translocation into the nucleus where the hormone-receptor complex binds a cognate DNA element and alters the transcription of the target gene.

A detailed review of the structure and function of nuclear receptors is found in Mangelsdorf et al. (1995); Beato et al. (1995); Kastner et al. (1995), Mangelsdorf and Evans (1995), Perlman and Evans, 1997).

The nuclear receptors are localized in the cytosol or nucleus. Due to their lipophilic nature the natural ligands of the nuclear receptors can cross the cell membrane in a passive manner and bind the cognate receptor in the cell interior. Binding by hormones activates the transcription regulatory effect of the receptor.

The nuclear receptors are specific DNA binding proteins which bind particular DNA sequences termed *hormone responsive elements* (HREs). Accessible to regulation by nuclear receptors are genes whose promotors carry cognate HREs. Often the hormones together with nuclear receptors function as *inducers of gene expression*. The response to a hormone signal in this case is an enhanced transcription level of neighboring genes.

In order to distinguish nuclear receptors from membrane-bound receptors, the former are sometimes referred to as *soluble receptors*. They can be found in the *cytoplasm* or *nucleus*. Their regulatory function is always performed, however, in the nucleus.

In comparison to signaling pathways which utilize transmembrane receptors (see chapter 5, 8, 11), signaling via nuclear receptors is of relatively simple structure. The pathways lead directly, with only a few participating protein components, from the extracellular space to the level of the DNA in the nucleus. Most important protein components of the signal pathway are known and well characterized. Nevertheless, we understand very little of the mechanism by which the activated receptors lead to a transcription initiation. This is due to the extreme complexity of transcription initiation in eucaryotes (see 1.2). Both the variety of proteins involved in the formation of a competent initiation complex, as well as the influence of chromatin structure, make it difficult to elucidate the exact function of nuclear receptors in transcription initiation.

Regulation processes which act at the level of transcription are slower than those which affect the activity of already existing enzymes. The latter process is used mainly to induce a rapid reaction in an organism to an external signal or to achieve a rapid communication between cells. Signal transduction by nuclear receptors, on the other hand, is intended to achieve long-term changes in the activity of enzymes, whereby the speed of the regulatory process is not very prominent. Regulation with the aid of nuclear receptors is therefore used mainly to adapt the activity of key metabolic enzymes to modified external conditions or a change in the function of an organism.

Signaling via nuclear receptors is paramount for the development and differentiation of higher organisms, since these processes do not require a rapid response but do require long-lasting functional changes. Accordingly, many hormones of the signal transduction pathways involving nuclear receptors participate in the development and differentiation of organs. Examples are the sexual hormones, the thyroid hormone, the $D_3$ hormone, as well as retinoic acid.

## 4.3 Classification and Structure of Nuclear Receptors

The first receptors to be characterized were those for the glucocorticoids, for estrogen and for progesterone. Many other receptors could be identified based on common structural and functional features. This family *of nuclear receptors*, sometimes termed the family of *steroid receptors* or the family of *intracellular receptors*, carry out similar functions. The nuclear receptors display a high degree of homology at the level of amino acid sequence, which indicates that they operate with similar functional principles. For many receptors the cognate hormone and their function in the cell remain unknown. Such *„orphan receptors*" were usually identified with the help of screening techniques employing DNA probes based on known receptors. The most important representatives of the nuclear receptors are summarized in table 4.1.

Extensive deletion and mutation studies, as well as sequence comparisons, have shown that the nuclear receptors are constructed modularly. At the level of the primary structure the steroid hormone receptors can be divided into five different domains (fig. 4.5), each with specific functions.

A typical nuclear receptor contains a domain responsible for *the DNA binding* (domain C), the *ligand binding and dimerization* (E), and for the *transactivation* and *other protein-protein interactions* (A,B,E,F). Furthermore, there are also *nuclear localization signals* (D).

Especially the domains for DNA binding and ligand binding display characteristics of independently folding structural units. The single domains can therefore be exchanged using recombinant DNA technology and their functionality in other protein environments examined in „domain-swapping" experiments (see 1.4.2, fig. 1.33).

### 4.3.1 DNA Binding Elements of Nuclear Receptors, HREs

The steroid hormone receptors are sequence specific DNA binding proteins whose cognate DNA elements are termed „hormone responsive elements" (HREs). The HREs known to date possess a common structure. They are composed primarily of two copies of a hexamer sequence. In table 4.1 are listed the hexamer sequences of the HREs of important nuclear receptors.

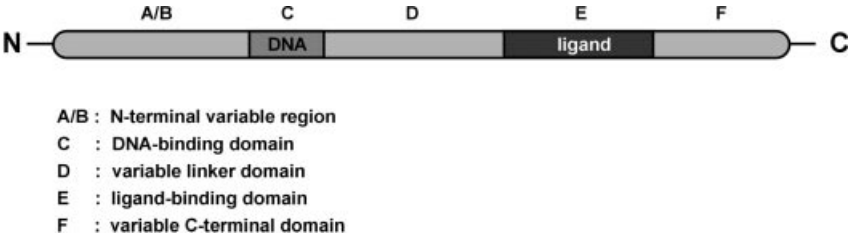



**Fig. 4.5.** Domain structure of the nuclear receptors. Functional domains of nucear receptors are portrayed in a one dimensional, linear fashion.

The identity of a HRE is determined by the sequence, polarity, and distance of the hexamers. Mutation and duplication of an ancestral recognition sequence have allowed the creation of many various DNA elements during the course of evolution, whose sequence, polarity and distance is characteristic for a given hormone receptor or receptor pair (review: Glass, 1994; Mangelsdorf, 1995; Beato, 1995). The *configuration* of a HRE is determined by the polarity and spacing of the hexamers. Fig. 4.6 illustrates on the example of the HREs for the RXR heterodimer (see 4.5.1) the various configurations of an HRE.

The receptors bind to the cognate HRE mainly as dimers, allowing the formation of *homodimers* as well as *heterodimers* between various receptor monomers. We know of very few nuclear receptors whose HRE contains only a single copy of the recognition sequence. These receptors bind as monomers to the cognate HRE.

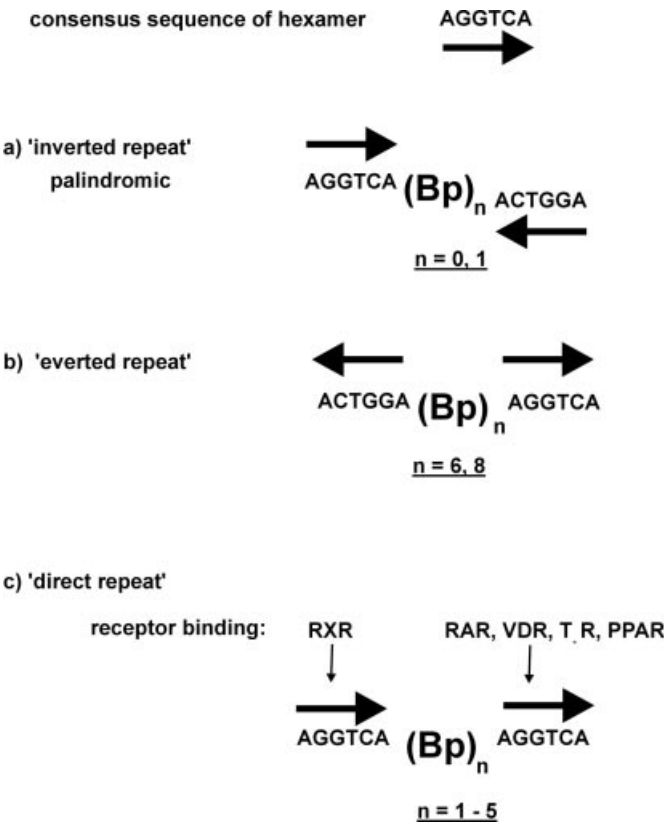



**Fig. 4.6.** HRE structure of the RXR heterodimer. Shown is the consensus sequence of the HREs of the RXR heterodimers (see Fig. 4.7) and the different possible arrangements of the hexameric half-site sequences. The hexamers can be arranged palindromically as inverted repeats (a), as everted repeats (b), or as direct repeats (c). n indicates the number of base pairs that lie between the two hexamers. RXR: receptor for 9-cis retinoic acid; RAR: receptor for all-trans retinoic acid; $T_3R$: receptor for the $T_3$ hormon; PPAR peroxisome proliferator-activated receptor; VDR: receptor for vitamin $D_3$.

Based on the subunit structure of DNA bound receptors and on the structure of the HREs, four classes of nuclear receptors can be distinguished (fig. 4.7).

### Dimers of the Steroid Hormone Receptors

The HREs of the steroid hormone receptors posses a *palindromic structure*, comparable to the DNA binding elements of procaryotic repressors (see fig. 4.7a). The glucocorticoid receptor, for example, binds as a homodimer to the two-fold symmetrical recognition sequence, whereby the receptor is already dimerized in solution. In complex with the DNA each subunit of the dimer contacts one half-site of the HRE. As a consequence of the two-fold repeat of the recognition sequence, a high affinity binding of the receptor dimer results (compare 1.2.4).

### Heterodimers containing RXR

The DNA binding element of the nuclear receptors for all-trans retinoic acid, for 9-cis retinoic acid, for the $T_3$ hormone and for the vitamin $D_3$ hormone usually exhibit a direct repeat of the recognition sequence, resulting in formation of heterodimers on the DNA (fig. 4.7b). One of the partners in the heterodimer is always the receptor for 9-cis retinoic acid, RXR, and which usually occupies the 5' side of the HRE.

Of particular importance for receptor binding on HREs with direct repeats of the recognition sequence is the fact that the hexamers of these HREs are arranged head to tail and thus require a *polar arrangement* of the receptor dimers.

HREs of this type display a further unique characteristic: half-site sequences can be used to create different HREs by varying only the *spacing* between the repeats. The spacing can vary between 1 and 5 bp of any sequence, so that with one given repeat five different HREs can exist ( see also fig. 4.6). In this case the spacing between the recognition sequences determines which hetero- or homodimer can form. Further multiplicity is achieved by combining different half-site sequences in a single HRE.

### Dimeric Orphan Receptors

The „orphan receptors" derive their name from the fact that the cognate hormones for these receptors were originally unknown or little understood. Orphan receptors bind as homodimers to recognition sequences arranged as direct repeats (fig. 4.7c). The receptor for 9-cis vitA acid, which also binds as a homodimer an HRE with two half-sites in direct repeats, is also considered a member of this class of receptors.

### Monomeric „Orphan Receptors"

There are also orphan receptors known which bind as monomers to asymmetric recognition sequences (fig. 4.7d). The function of these receptors is still poorly characterized.

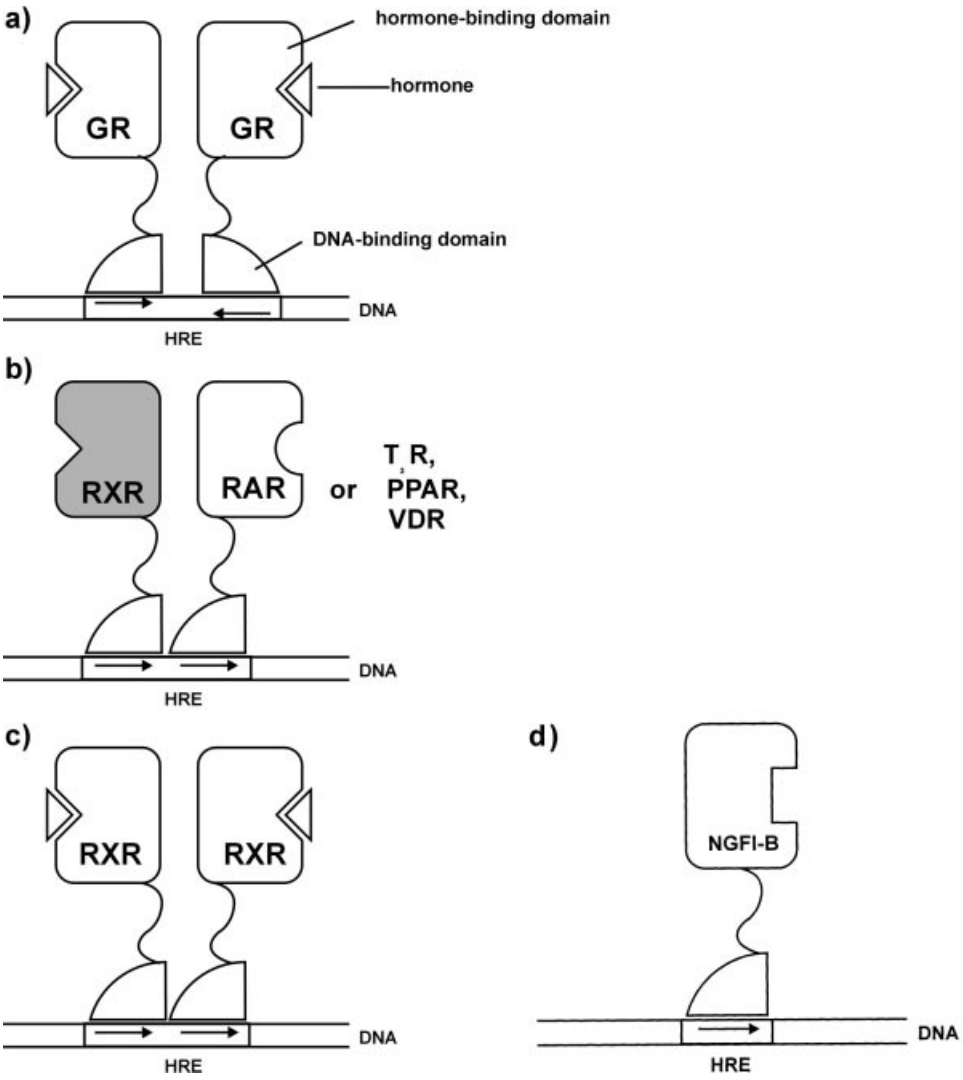



**Fig. 4.7.** Oligomeric structure of nuclear receptors and structure of the HREs. The nuclear receptors can be subdivided into four groups based on structures of the receptors and HREs. Shown above are some representative examples. a) binding of a homodimeric receptor to a two-fold symmetric palindromic DNA element, GR: gluccocorticoid receptor. b) binding of a heterodimeric receptor to a DNA element with direct repeats of the recognition sequence, whereby the 5' side of the HRE is occupied by 9-cis retinoic acid receptor (RXR). RAR: receptor for all-trans retinoic acid, $T_3R$: receptor for the $T_3$ hormone; PPAR: peroxisome proliferator-activated receptor; VDR: receptor for vitamin $D_3$. c) binding of RXR as a homodimer to an HRE with direct repeat of the recognition sequence. d) binding of a monomeric receptor to an asymmetric recognition sequence, NGFI-B: nerve growth factor induced receptor, is involved in the regulation of enzymes of steroid biosynthesis. After Mangelsdorf et al., 1995.

### 4.3.2 The DNA Binding Domain of Nuclear Receptors

The binding of receptors to their cognate HRE occurs via a DNA binding domain, which are largely independent folding domains. Within the family of nuclear receptors, the DNA binding domain is the most conserved structural element and is located in region C of the primary structure (see fig. 4.8a). The DNA binding domain possesses structural elements that mediate the specific recognition of the HRE, as well as for the

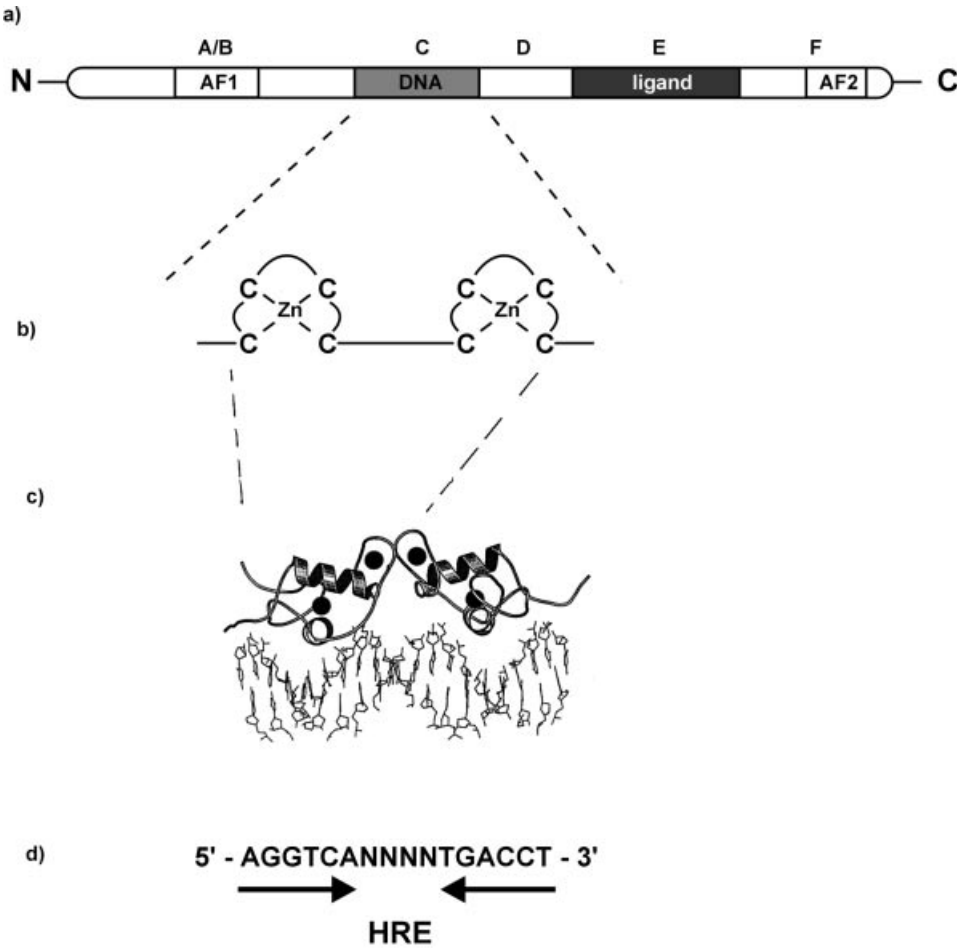



**Fig. 4.8.** Functional domains, DNA-binding and HRE structure of the steroid hormone receptors. a) domain structure of the steroid hormone receptor. AF1, AF2 domains that mediate the stimulation of the transcription. b) schematic representation of the two $Zn^{2+}$-$Cys_4$ binding motifs of the DNA-binding domains. c) Complex formation between the dimeric DNA-binding domains of the gluccocorticoid receptor and the HRE. The black spheres represent $Zn^{2+}$ ions. After Luisi et al., 1991. d) Consensus sequence and configuration of the HRE elements of the steroid hormone receptor.

dimerization of the receptor on the HRE. The core of the DNA binding domain includes a span of 70–80 amino acids, in which all information for the specific recognition of the cognate half-site is contained.

In the core of the DNA binding domain are two $Zn_2Cys_4$-motifs. As shown in fig. 4.8, the Zn-motifs serve to position a recognition helix in the major groove of the DNA. It is not involved in any direct interactions with the bases of the DNA (Luisi et al., 1991). Via the recognition helix specific contacts are formed with the hexamer half-site of the HRE. The Zn-motifs of the nuclear receptor are an integral part of the overall structure of the DNA binding domain and do not represent an independent structural element, as is the case with TFIIIA or ZIF268 (compare 1.2.1 as well as fig. 1.3 and 1.4). The two Zn-motifs assume non-equivalent positions in the DNA binding domain. While the N-terminal Zn-motif participates in the positioning of the recognition helix and the interactions with the sugar-phosphate backbone, the C-terminal Zn-motif serves to impart a certain functionality to the dimerization surface and to contact the phosphate backbone of the DNA.

### 4.3.3 HRE Recognition and Structure of the HRE-Receptor Complex

It is a characteristic feature for the structure of HREs that the recognition sequences of different receptors are actually very similar, and that they differ primarily in the polarity and spacing of the sequences. It is therefore not surprising that the structure of the receptor dimers is accurately adapted to the exact orientation and spacing of the recognition sequences. High resolution structures of DNA-bound homo- and heterodimers have confirmed this aspect of receptor-DNA recognition.

In the case of the palindromically arranged recognition sequences the binding occurs mostly via pre-formed homodimers of the receptor in solution. Sequence and spacing of both recognition sequences in the HRE are highly complementary to the binding surface of the recognition helix, as well as to the spacing between the DNA binding domains of the dimeric receptor. The spacing of the recognition sequences is firmly pre-determined by the structural elements of the receptor participating in the dimerization. The geometry of the dimer determines, furthermore, what spacing of the recognition sequences is tolerated in the formation of a high affinity complex.

For direct repeat HREs, the spacing of the two half-sites is often the decisive, if not only, element based on which the receptor (homodimer or heterodimer) recognizes its own HRE and discriminates against related HREs. The solution of the structure of a DNA-bound receptor-heterodimer composed of a DNA binding domain of RXR and the $T_3$-receptor (Rastinejad et al., 1995) has shown how these receptors can distinguish between highly related HREs (fig. 4.9).

The HRE sequence used for the structure determination (AGGTCA$(N)_4$ AGGTCA) consists of two identical hexamers in direct repeat and separated by 4 bp (D-4 arrangement). The heterodimer RXR-$T_3$R binds in a polar manner on the HRE, with RXR occupying the 5'-side of the HRE. Both hexameric sequences lie on the same side of the DNA double helix and are contacted by an $\alpha$-helix of each of the receptors in a nearly identical manner. In the complex, the DNA binding domain of

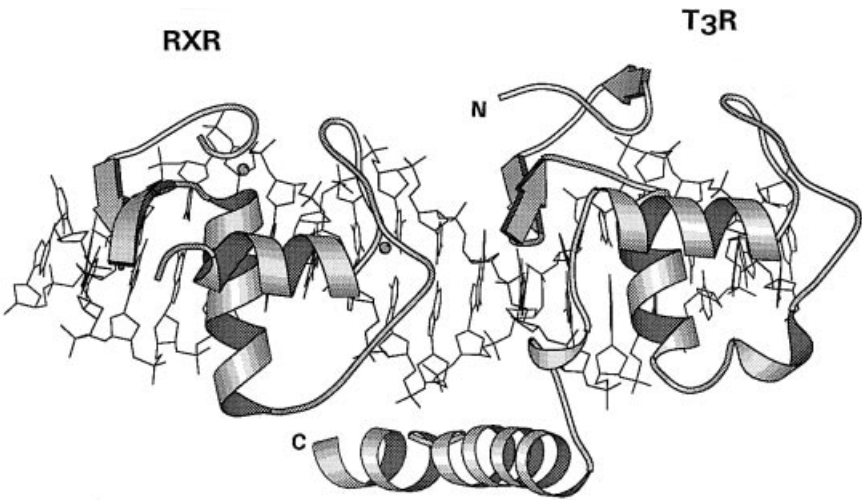



**Fig. 4.9.** Structure of the RXR-$T_3$R heterodimer in complex with DNA. Illustrated is a complex between the DNA-binding domain of the RXR-$T_3$R heterodimer and an HRE with direct repeats of the sequence AGGTCA separated by 4 bp. The two receptor subunits contact the hexameric sequences with a recognition helix in a manner very similar to that of the gluccocorticoid receptor (see Fig. 4.7). The Zn atoms are drawn as spheres. The figure illustrates the polarity of the binding of the two subunits. The interaction between the two subunits is mediated mainly via an extension of the C-terminal DNA-binding domain of the $T_3$R. A greater or smaller distance between the two hexamers of the HRE would act contrary to the interaction between the two subunits as shown. MOLSCRIPT drawing (Kraulis, 1991).

$T_3$R interacts more with the half-site than does RXR. Different structural elements of each of the monomeric receptors are involved in the dimerization process, leading to the polar configuration of the monomers on the DNA.

The structure determination confirmed the importance of the spacing of the two hexamers as a discrimination factor in an impressive manner. A spacing of only 3 nucleotides between the two hexamers would lead to steric overlap of both receptors; a high affinity, cooperative binding would not be possible. With a spacing of more than 4 nucleotides a high affinity complex could also not be formed due to the relative rigidity of the two monomers.

The dimeric structure of the receptors thus plays an important role in the discrimination between closely related hexamers. The structural elements of the receptors which participate in the dimerization ensure that the recognition helices assume a defined mutual spatial arrangement adapted to the spacing of the hexamers of the cognate HREs. Only in this configuration is a high affinity, cooperative binding possible. For a correct binding it is necessary that both recognition helices optimally contact both repeat sequences. If, as a result of an incorrect spacing of the hexamers, only one of the two recognition sequences bind, then a high affinity complex can not be formed.

### 4.3.4 Ligand Binding Domains

The ligand binding domain (section E in fig. 4.5) of the nuclear receptors harbors several functions. Apart from the specific *binding site for the hormone*, one finds further structural elements in this domain which mediates *dimerization* of the receptors as well as structural elements important for the *ligand-mediated transactivation*.

A contribution to the dimerization of the receptors – in addition to that from the DNA binding domain – is provided by a dimerization element in the ligand binding domain. The structure of the ligand binding domain of RXR without bound hormone shows a homodimer with a symmetric dimerization surface, formed essentially from two antiparallel α-helices (Bourguet et al., 1995). The folding of the dimerization surface is supported by the periodic arrangement of hydrophobic amino acids with a 7 residue spacing, similar to the configuration of leucine residues found in leucine zippers (see 1.2). These hydrophobic amino acids, however, do not participate directly in the dimerization.

The structures of the ligand binding domains of two nuclear receptors with bound hormone (RAR, Renaud et al., 1995; $T_3$R, Wagner et al., 1995) show that the hormones are bound in the hydrophobic core of the ligand binding domains. The ligand binding induces a distinct conformation change of an amphipathic α-helix, the essential element of the activation domain AF2 (see 4.3.5). The AF2 domain is also involved directly in ligand binding, which suggests that the activation domain assumes a different conformation in the presence or absence of ligand.

During ligand binding, the hormone and binding pocket of the receptor undergo mutual structural adjustments. An exception is found in the PPARγ-receptor. Structural studies showed that the synthetic ligand thiazoladinedione only partially occupies the binding pocket. There appear to be several alternative possibilities for ligand binding by this receptor (Oberfield et al., 1999).

### 4.3.5 Transactivating Elements of the Nuclear Receptors

A structural element necessary for the ligand-dependent transcriptional activation is found within the C-terminal region of the ligand binding domain. An essential structural feature of this transactivating domain, designated AF-2, is an α-helix with amphipatic character. All known structures to date and extensive experimental data indicates that this helix is involved in communication with components of the transcrptional apparatus. The amphipathic α-helix of the transactivating domain possesses a hydrophilic and a hydrophobic face. Upon ligand binding the helix alters its position and forms a part of the ligand binding pocket with its hydrophobic face (Renaud et al., 1995; Wagner et al., 1995).

A further structural element with transactivating function is localized the N-terminal A/B domain of the receptor. This region, termed AF-1, is structurally and functionally only partially characterized.

## 4.4 The Signaling Pathway of the Steroid Hormone Receptors

The receptors of steroid hormones were the first representatives of the family of nuclear receptors to be characterized. With the characterization of further receptors it became clear that the signaling pathway of the nuclear receptors differ significantly in detail. Based on the receptor activation mechanism the nuclear receptors can be divided into two basic groups:

In the first group, those of the steroid hormone receptors, the receptors can be localized in the nucleus or in the cytoplasm. The receptors of the other group are always localized in the nucleus. Representative ligands of these receptors are the derivatives of retinoic acid, the $T_3$ hormone and $VitD_3$.

Signal transduction by steroid hormones is distinguished by the fact that the receptors can be found either in the cytoplasm or in the nucleus. The steroid hormone receptor receives the hormonal signal in the cytosol, becomes activated by hormone binding, at which point it enters the nucleus to regulate the transcription initiation of cognate genes. Fig. 4.10 shows the most important steps in the signal transduction by steroid hormones.

### 4.4.1 Activation of the Cytoplasmic Apo-Receptor Complexes

The steroid hormones are distributed throughout the entire organism by means of the circulatory system. Transport often occurs in the form of a complex with a specific binding protein. An example for such a binding protein is transcortin, which is responsible for the transport of the corticosteroids. The steroid hormones enter the cell by diffusion and activate the cytosolic receptors.

In the absence of steroid hormones the receptors remain in an inactive complex, designated the *apo-receptor complex* (review Pratt, 1993; Bohen, 1995). In the aporeceptor complex the receptor is bound to proteins belonging to the chaperone class. Chaperones are proteins whose levels are increased as a result of a stress situation, such as a rise in ambient temperature. The chaperones assume a central function in the folding process of proteins in the cell. Chaperones aid proteins in avoiding incorrectly folded states. They participate in the folding of proteins during and after ribosomal protein biosynthesis, during membrane transport of proteins, as well as in the correct assembly of protein complexes.

With the help of co-immunoprecipitation it could be shown that the receptors of steroid hormones interact with at least three chaperones, *Hsp90, Hsp70* and *Hsp56* (fig. 4.10). The term „Hsp“ (Heat shock protein) is derived from the observation that these proteins were produced at higher levels following heat treatment. Furthermore, one finds a 23 kDa acidic protein in the apo-receptor complex whose function is not yet clear.

It is assumed that the receptor, when complexed with Hsp90, Hsp56 and p23, is fixed in an optimal conformation for hormone binding. The activation of the receptor by the hormone requires the presence of Hsp90, which binds the receptor as a dimer. Hsp90

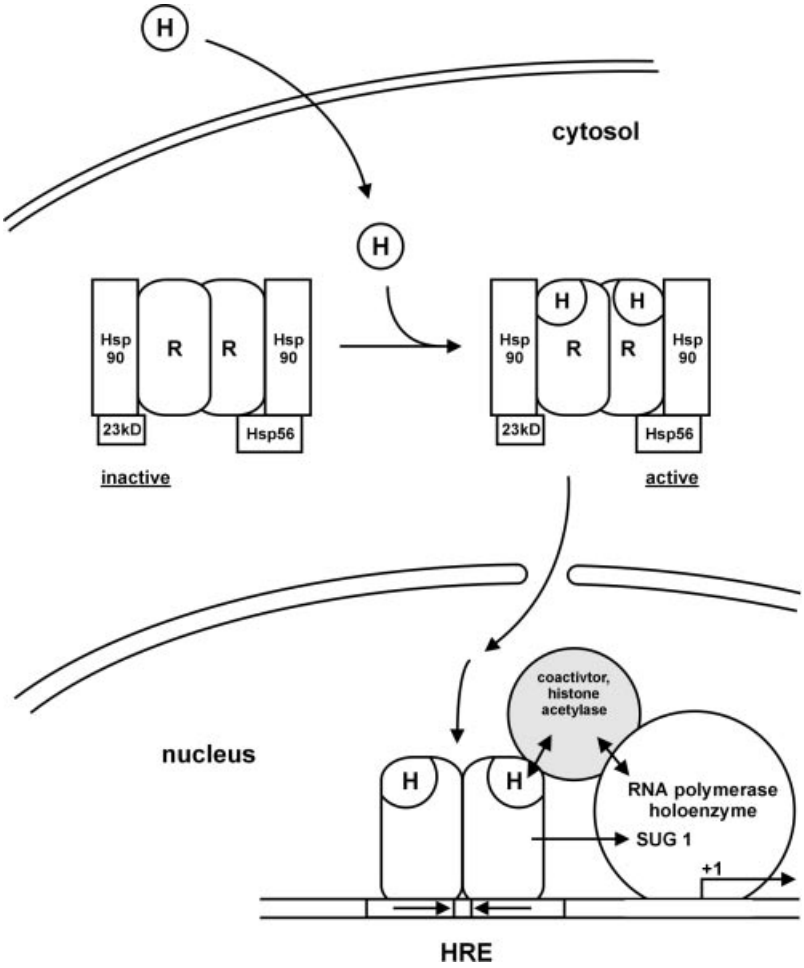



**Fig. 4.10.** Principle of signal transduction by steroid hormone receptors. The steroid hormone receptors in the cytosol are found in the form of an inactive complex with the heat shock proteins Hsp90 and Hsp56 and with protein p23. The binding of the hormone activates the receptor so that it can be transported into the nucleus where it binds to its cognate HRE. It remains unclear in which form the receptor is transported into the nucleus, and to which extent the associated proteins are involved in the transport. One mechanism of activation of transcription initiation involves activation of a histone acetylase and remodeling of chromatin. Furthermore, the receptors interact directly or indirectly with components of the RNA polymerase II holoenzyme (e.g. SUG1).

binds preferentially on partially unfolded proteins and it is believed that Hsp90 holds the receptor in a partially unfolded conformation. It is assumed that Hsp70 facilitates the correct association of Hsp90 with the receptor. In this function, Hsp70 is supported by a further heat shock protein which belongs to the DnaJ family. Hsp70 is possibly also involved in partially unfolding the steroid binding domain, a process which appears to be necessary for hormone binding. The binding of the hormone to the aporecep-

tor complex leads to activation of the receptor and initiates the translocation of the receptor into the nucleus. The activated receptor possesses an accessible nucleus localizing sequence and is furthermore capable of DNA binding and transactivation. The ability to transactivate implies that the transactivating domain is properly positioned, as a result of the hormone binding, to allow stimulatory interactions with the transcription apparatus.

Details of the activation and transport into the nucleus remain unknown. It is assumed that the complex between receptor and hormone and the loosely associated heat shock proteins reaches the nucleus via a protein transport system. After dissociation of the heat shock proteins, the hormone-receptor complex is capable of specific binding on the HRE and of transactivation.

The chaperones are used as tools in this system for regulation of activity of the steroid hormone receptors. The function of the chaperones is obviously to fix the receptor in a conformation which allows high affinity binding to the hormone and the subsequent steps of specific DNA binding and transactivation. For the steroid hormone receptors this means that they must exist in functionally different conformations. It may be a function of the chaperones to stabilize the particular conformation optimal for hormone binding.

### 4.4.2 DNA Binding and Transactivation

The HREs are components of the control region of hormone-regulated promotors. Typically, the binding of the activated, hormone-bound receptor to the HRE stimulates the expression of the cognate genes. The mechanisms of transcription activation still remains speculative. Many proteins have been identified which act as coactivators, i.e. they interact with the activated receptor (review: Glass et al., 1997). The nature of these proteins suggest that at least two, separate but not mutually exclusive pathways of transcription activation must be considered:

- interaction of the transactivating domain with components of the transcription initiation complex
- active remodeling of the chromatin

Due to the complex structure of the initiation complex it remains unclear which interactions are responsible for the first mechanism. The coupling between the transactivating domain and the initiation complex can be direct or indirect. There is evidence which indicates that proteins with co-activator function mediate the interaction between HRE-bound receptors and the transcription initiation apparatus. One such protein is RIP-140, which mediates the transcription activation of the estrogen receptor. The AF2 domain can also directly contact the transcriptional apparatus. One component of the RNA polymerase II holoenzyme, the SUG1 protein, has been identified as a binding partner for the AF2 domain. The SUG1 protein has the function of a co-activator in transcription initiation and is considered a mediator (see 1.4.3.2).

A second mechanism by which transcriptional activation occurs via remodeling of the chromatin. Steroid receptors, in cooperation with other proteins, can induce a

structure of chromatin which promotes transcription. A large protein complex takes part in this remodeling. Some of the proteins in the complex, like the CBP/p300 protein (see also 1.4.6), possess histone acetylase activity. The activated, DNA-bound receptor possibly recruits a histone acetylase to the chromatin. It can thus create the conditions necessary for the formation of a transcription initiation complex by this histone modification.

### 4.4.3 Transcription Repression by Steroid Hormone Receptors

The steroid hormone receptors can not only activate transcription, but can also repress transcription (review Beato, 1995). This function is mechanistically poorly understood. Repression by nuclear receptors has been observed for genes whose control regions have, apart from the HREs, binding sites for other transcription factors, e.g. AP-1 and NFκB. In AP-1 and NFκB controlled genes it has been observed that the steroid hormone receptors have an inhibitory effect on the transcription activation of both transcription factors. The mechanism of this mutual interaction remains unclear.

The transcription factors AP-1 and NFκB are at the end of a signal cascade activated by growth factors (see chapters 9, 10, 11). The ability of steroid hormone receptors in certain situations to nullify the effect other transcription factors demonstrates that two different signaling pathways can converge at the level of transcription.

The observation that steroid hormone receptors act as activators as well as repressors of gene activity suggests that receptors can assume an activating and a repressing conformation. In the latter state, the transcriptional activating domain of the receptors is masked.

### 4.4.4 Regulation of the Receptor Activity by Phosphorylation: Crosstalk

The steroid hormone receptors are phophoproteins which are usually phosphorylated on several positions. The phosphorylation sites are mainly found in the N-terminal region of the receptors. Serine phosphorylation prevails. One rare example of tyrosine phosphorylation is described for the case of estrogen receptors. The consequences of phosphorylation for the receptor proteins are varied. It is conceivable, and in some cases experimentally proven, that it has influence on hormone binding, nuclear transport, DNA binding and transactivation.

Regulation by phosphorylation is well studied in the case of the estrogen receptor. The estrogen receptor is, among others, activated by phosphorylation on Ser111, which is located in the AF-1 domain. As a result of the phosphorylation the hormone-induced activation of the receptor is further enhanced. The kinase function was identified to come from a mitogen-activated protein kinase (MAPK, see chapter 10; Kato et al., 1995). The MAPK pathway is activated by growth factors and includes participation of the Ras protein. The phosphorylation of the estrogen receptor by MAP-kinases allows

mitogens, i.e. cell division signals, to branch into and modulate the signaling pathway of the steroid hormone receptors.

There has also been a report of activation of the estrogen receptor mediated by the neurotransmitter dopamine (Power et al., 1991). This mechanism of activation is independent of that by the hormone.

The modification of nuclear receptors by protein phosphorylation is an example of *„crosstalk"* between the signaling by nuclear receptors and other signaling pathways. Protein phosphorylation is a tool by which other structured signaling pathways can influence the signal transduction by the steroid hormone receptors. At the level of the steroid hormone receptor both the hormonal signals and mitogenic signals arising from growth factors, can be received and processed.

## 4.5 Signaling by Retinoids, Vitamin $D_3$, and the $T_3$-Hormone

Ligands of the RXR-heterodimer group and the „orphan" receptors are chemically more diverse than the ligands of the steroid family. Representative ligands of this group are the retinoids all-trans retinoic acid, 9-cis retinoic acid, the $T_3$ hormone and vitamin $D_3$ (fig. 4.1).

In contrast to signal transduction by the steroid hormone receptors there are multiple pathways by which the ligands of this group are made available for receptor activation (fig. 4.11):

- the hormone ligands can be secreted in the classical endocrinological pathway and transported to the target cell where they bind the receptor.
- the active hormones can be formed intracellularly from inactive precursors. The inactive precursor is transported through the bloodstream to the target cell where it is enzymatically converted to the active hormone. An example for this pathway is that of 9-cis retinoic acid, which is synthesized from the alcohol of vitamin A (vitamin A1; retinol).
- the hormone is synthesized in the same cell in which receptor activation takes place. The ligand is synthesized intracellularly and is not secreted. An example for such an intracellular hormone is prostaglandin $J_2$.

A further, more dramatic difference to the steroid hormone receptors is the localization of the receptors. The receptors for the retinoids (RAR and RXR, see table 4.1), the $T_3$ hormone ($T_3R$) and vitamin $D_3$ (VDR) are mainly localized in the nucleus and their activity is not controlled by the heat shock proteins. The receptors also bind the corresponding HRE in the absence of hormone, in which case they can then act as *repressors* of gene activity. In the presence of the hormone an activation of gene expression is usually observed.

Initial studies on the receptors of all-trans retinoic acid, vitamin $D_3$ and the $T_3$-hormone assumed that these receptors bound their HRE's in homodimeric form. It became clear with the discovery of the receptors for 9-cis retinoic acid that this simple

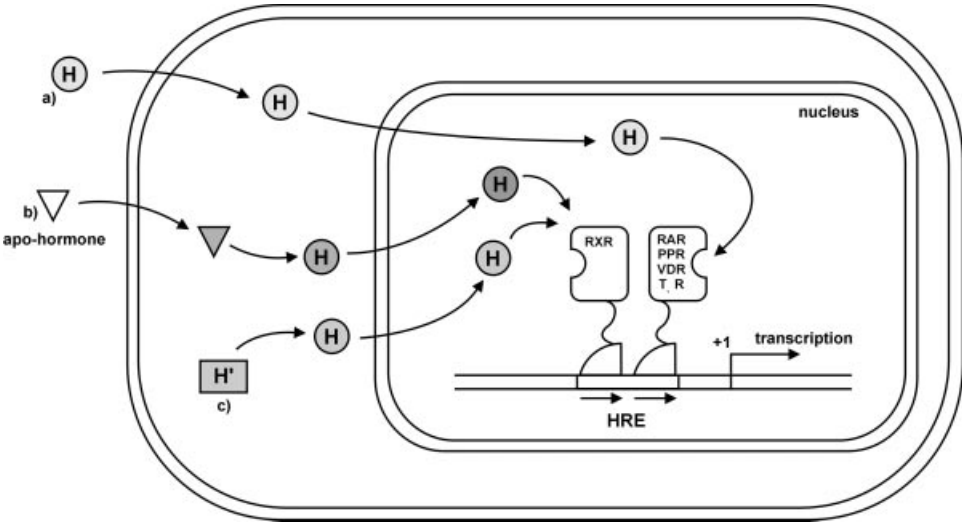



**Fig. 4.11.** Principle of signal transduction by RXR heterodimers. The activated hormone can be made available to the RXR heterodimer in three different ways. a) The hormone (e.g. $T_3$ hormone) is synthesized in endocrinal tissue and reaches the DNA-bound RXR-$T_3$R heterodimer in the nucleus via passive transport. b) The active hormone is formed in the cytosol from an inactive apo–hormone (as for, e.g. 9-cis-retinoic acid). c) The hormone is synthesized intracellularly. In all three cases, the binding of the hormone-RXR-heterodimeric complex is the signal that induces transcription activation of the downstream genes. After Mangelsdorf and Evans, 1995.

picture would have to be replaced by a more complex one. The source of this complexity is the possibility for the formation of heterodimers or of homodimers, as well as the many combination possibilities in the composition of the corresponding HRE's.

The receptors for all-trans retinoic acid, vitamin $D_3$ and the $T_3$-hormone (as well as other receptors of this class, see table 4.1) usually perform their regulatory function as heterodimers. RXR plays a special role in the formation of heterodimers: the receptor for 9-cis retinoic acid is usually one of the binding partners in the heterodimer.

### 4.5.1 The Structure of the HREs of RXR-Heterodimers

The HREs and RXR-heterodimers are usually composed of two identical or nearly identical copies of the hexamer sequence AGGTCA in direct repeat. The apparently simple structure of the HREs leads to the question of how the receptors of this class can distinguish between the various HREs. Studies with artificial HRE constructs, as well as of naturally occurring HREs, indicate a complicated cooperative effect between HRE structure on the one hand and homo- or heterodimer formation of the receptors on the other hand.

The following points were identified as important for the recognition and discrimination of a particular HRE:

- In the case of identical hexamer sequences of a HRE the spacing between the hexamers is a specificity determining element *(n-rule)*. The spacing can be between 1–6 bp. Grounds for the discrimination based on spacing is the structure of the receptor dimer. A given receptor demands a particular spacing of the hexamers in the HRE due to steric requirements. If the distance requirement is not met, the high affinity and cooperative formation of a dimer on the HRE is not possible.
- The receptor for 9-cis retinoic acid (RXR) usually occupies the 5' position in the heterodimer. The RXR serves quasi as a vehicle to bring other receptor monomers to the 3' half-site of the HRE.
- Apart from the spacing, the sequence of the hexamers is also important for the discrimination. In naturally occurring HREs one finds sequences very closely related to the consensus sequence AGGTCA; the nature of the sequence also contributes to the identity and specificity of a HRE as well.
- for heterodimers out of VDR and $T_3R$ it could be shown that the hormone concentration necessary for activation depends on which of the two monomers occupies the 5' and which the 3' position in the HRE.

## 4.5.2 Complexity of the Interaction between HRE, Receptor and Hormone

There is an altogether complex interaction between HRE, receptor and hormone in the group of receptors for retinoids, vitamin $D_3$ and the $T_3$-hormone. The complexity is determined by the following factors:

### Formation of the Homo- or Heterodimer

To what extent hetero- or homodimers are formed depends upon the concentration of the various receptors in the particular cell type. The receptor concentration can furthermore be regulated via the general mechanisms as described in chapters 1 and 2, whereby regulation at the level of expression of the receptors is particularly important.

### Multiplicity of the HREs

The identity of an HRE is determined by sequence, spacing and polarity of the hexamer sequences. There are many possible combinations of these specificity determinants, such that the hormonal activation of a gene decidedly depends on the nature of the cognate HRE. There is an additional, poorly understood influence from the neighboring sequences, and, furthermore, HREs are often arranged tandemly. Another factor which can influence HRE activity is the occurrence of several, different regulatory elements on one promoter. Often times further DNA elements are found in the neighborhood of the HRE which can serve as further transcriptional activators, such as, for example, AP1. In this manner, other transcriptional activators can act cooperatively with or antagonistically against the receptor.

### Multiplicity of the Receptors

Among the receptors for the retinoids (RAR and RXR) there are at least three subtypes characterized, known as RARα, β, and γ (RXR α,β and γ). The various subtypes differ from each other mainly in their amino acid sequence and are encoded in each case by their own gene. Furthermore, one finds isoforms of some of the receptors created by alternative splicing of the primary transcript. Altogether, the repertoire of the various receptor variants is enormously increased through the existence of the subtypes and by alternatively spliced receptors. Because the individual variants differ in their ability to form hetero- and homodimers as well as in their ability to activate and bind ligands, this results in a great variety of functionally different receptors.

### Binding and Activation via Hormones

With regard to ligand binding, heterodimer formation brings a further element of complexity into the picture. The ligand concentration necessary for activation varies strongly depending upon the nature of the heterodimer. For example, there are RXR-heterodimers, such as RXR-$T_3$R, in which the RXR binding site for 9-cis retinoic acid is not accessible. In this case, the RXR is a „silent" partner. In other combinations (e.g. RXR-PPAR) the ligands of both receptors are required for full activation of the heterodimers.

The hormone concentration necessary for the binding and activation of a receptor dimer can be very different for the various dimers. The available hormone concentration in a particular cell type thus plays an important role in receptor activation. The effective intracellular concentration of a hormone is furthermore subject to a diverse regulation mechanism (see 4.1).

The picture sketched above for the function of the receptors of the retinoids and the $T_3$- and Vitamin $D_3$-hormones is in many points still incomplete. The cooperation of the activated receptors with other transcription factors, their interactions with the transcription apparatus and the influence of chromatin structure are still poorly understood.

## 4.5.3 Ligand Binding, Activation and Corepression of the RXR-Heterodimers

The ligands for the RXR-heterodimer group control – as with the steroid hormones for the steroid hormone receptor – the receptor function in a decisive manner. An active conformation of the receptor is induced upon binding of the ligand. Activation results upon DNA binding by the receptor, whereby the hormone is not necessary for DNA binding. It is assumed that a conformational change in the amphipathic helix of the transactivating domain occurs upon binding, resulting in the activation of the receptor. In the activated state, the transactivating domain can interact stimulatory with the transcription apparatus.

The receptors for the $T_3$-hormone ($T_3$R) and for the all-trans-vitaminA acid (RAR) exist in mammalian cells in two states. In the absence of hormone they exercise a

repressive influence on transcription activity and thus act as *transcriptional repressors.* In the hormone-bound state, on the other hand, they stimulate transcription.

The understanding of the biochemical principles of repressive function is still in its initial stages, and thus only models for their mode of action can be currently offered:

**Repression**

In the case of the T3R and RAR specific repressor proteins have been identified which bind to the receptor and mediate an inhibition of transcription (Hoerlein et al., 1995). The repressor proteins are also termed *co-repressors* Among the repressors are found proteins with histone deacetylase activity (see 1.4.6). The RXR heterodimers possibly stabilize the repressed state of chromatin by recruiting a histone deacetylase to the chromatin.

**Activation**

The binding of the ligand to the DNA-bound receptor reverses the repression and leads to an active transcriptional enhancement. One model assumes that the hormone induces dissociation of the repressing proteins so that interactions with the co-activators can occur. This leads to stimulation of transcription initiation.

Activation is a complex process in which various types of co-activators participate. Among these are found histone acetylases (see 1.4.6).

The diversity of nuclear receptor systems is also demonstrated by the fact that orphan receptors exist, which exercise a constitutive repressive or activating function independent of ligand.

# References Chapter 4

Beato, M., Herrlich, P and Schütz, G. 'Steroid hormone receptors: many actors in search for a plot' (1995) Cell *83*, 851–857

Bohen, S.P., Kralli, A. and Yamamoto, K.R. 'Hold'em and fold'em: chaperons and signal transduction' (1995) Science *268*, 1303–1304

Bouruet, W., Ruff, M., Chambon, P. Gronemeyer, H. and Moras, D 'Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-α' (1995) Nature *375*, 377–382

Funder, J.W., Pearce, P.T. Smith, R. and Smith, A.I. 'Mineralcorticoid action: target tissue specificity is enzyme, not receptor, mediated' (1988) Science *242*, 583–585

Glass, C. 'Differential recognition of target genes by nuclear receptor monomers, dimers and heterodimers' (1994) Endocrine Rev. *15*, 391–407

Glass,C.K., Rose,D.W. and Rosenfeld, M.G. (1997) 'Nuclear receptor coactivators' Curr. Op. Cell Biol. 9, 222.232

Hörlein, A.J., Näär, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Söderström, M., Glass, C. and Rosenfeld,M.G. 'Ligand independent repression by the thyroid hormone receptor mediated by a nuclear repressor co-repressor' (1995) Nature *377*, 397–404

Kastner, P., Mark, M. and Chambon, P. 'Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? (1995) Cell *83*, 859–869

Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H. Masushige, S., Gogoh, Y., Nishida, E., Kawashima, H., Metzger,D. and Chambon, P. 'Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase' (1995) Nature *270*, 1491–1494

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. and Sigler, P.B. 'Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA' (1991) Nature *352*, 497–505

Mangelsdorf, D.J. and Evans, R.M. 'The RXR heterodimers and orphan receptors' (1995) Cell *83*, 841–850

Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg,B., Kastner, P., Mark, M., Chambon, P. and Evans, R.M. 'The nuclear receptor superfamily: the second decade' (1995) Cell *83*, 835–839

Oberfield, J.L., Collins, J.L., Holmes, C.P., Goreham, D.M., Cooper, J.P., Cobb, J.E., Lenhard, J.M., Hull-Ryde, E.A., Mohr, C.P., Blanchard, S.G., Parks, D.J., Moore, L.B., Lehmann, J.M., Plunket, K., Miller, A.B., Milburn, M.V., Kliewer, S.A., Willson and T.M., 'A peroxisome proliferator-activated receptor gamma ligand inhibits adipocyte differentiation' (1999) Proc Natl Acad Sci USA S A *25*, 6102–6106

Perlmann, T and Evans, R.M. 'Nuclear receptors in Sicily: all in the famiglia' (1997) Cell *90*, 391–397

Pratt, W. 'The role of heat shock proteins in regulating the function, folding, and trafficking of the glucocorticoid receptor' (1993) J.Biol.Chem. *268*, 21455–21458

Power, R.F., Mani, S.K., Codina, J., Conneely, O.M.and O'Malley, B.W. 'Dopaminergic and Ligand-independent activation of steroid hormone receptors' (1991) Science *254*, 1636–1639

Rastinejad, F., Perlman, T., Evans, R. and Sigler, P.B. 'Structural determinants of nuclear receptor assembly on DNA direct repeats' (1995) Nature 375*,* 203–211

Renaud,J .P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H. and Moras,D., 'Crystal structure of the RAR- ligand-binding domain bound to all-trans retinoic acid' (1995) Nature *378*, 681–689

Wagner, R.L., Apriletti, J.W., McGrath, M.E., West, B.L., Baxter, J.D. and Fletterick, R.J. 'A structural role for hormone in the thyroid hormone receptor' (1995) Nature *378*, 690–697

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 5**

# G-protein Coupled Signal Transmission Pathways

## 5.1 Transmembrane Receptors: General Structure and Classification

During intercellular communication, extracellular signals are registered by the cell and converted into intracellular reactions. Signal transmission into the cell interior takes place by reaction chains, in which several individual reactions generally run in sequence and involve many signal proteins. The nature of the extracellular signal can be very diverse and may include extracellular signal molecules, such as low molecular weight messenger substances or proteins, or sensory signals such as light signals.

The first step in signal transmission is always the reception of the signal by the target cell. Signals can be registered and transmitted by the cell by two principal means. Firstly, the signal molecule may penetrate the target cell and bind to corresponding receptors in the cell interior, as explained in Chapter 4 for the nuclear receptors. However, in a multitude of signaling pathways, a conversion of the extracellular signal takes place with the help of a transmembrane receptor. The signal molecule does not penetrate the target cell, but rather it binds on the extracellular side to the transmembrane receptor. This conducts the signal into the interior of the cell and sets an intracellular *signal chain* in motion, which finally triggers a defined biochemical response of the target cell (Fig. 5.1a).

An example of a transmembrane receptor that registers sensory signals is rhodopsin. Rhodopsin is a sensory receptor that plays a role in vision, by receiving light signals and converting them into intracellular signals.

In addition to reception of hormonal signals by transmembrane receptors or nuclear receptors, the cell also has other means to conduct signals into the cell interior.

One simply designed path of signal transmission is found in neuronal communication (see Chapter 16). Transmembrane receptors are also used for signal transmission here. These have the character of a *ligand-gated ion channel* (Fig. 5.1b). Binding of a ligand (neurotransmitter or neurohormone) to the transmembrane receptor leads to a conformational change of the receptor that enables the flow of ions through the membrane. In this case, the receptor presents itself as an ion channel with an open state controlled by ligand binding to the outer side (or also to the inner side).

Another mechanism uses changes in membrane potential. A change in membrane potential induces the opening of an ion channel and ions cross the membrane (see Chapter 16). In this case, the change of the ion's milieu is the intracellular signal. Ion channels with an open state regulated by changes in membrane potential are known as

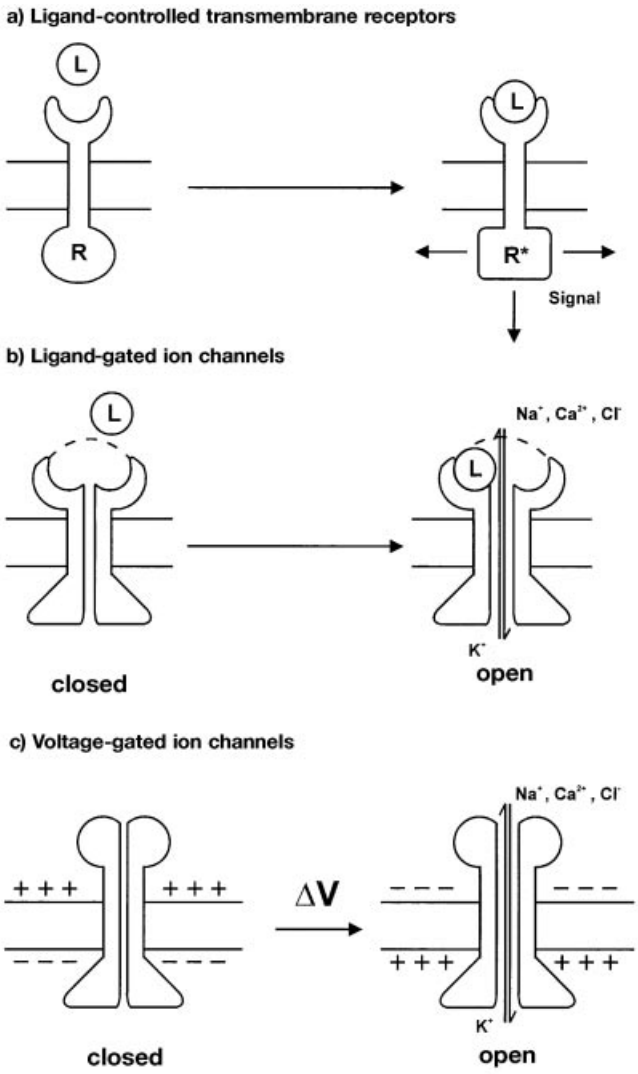



**Fig. 5.1.** Mechanism of signal transduction at membranes. a) Signal transmission via ligand-controlled transmembrane receptors. The ligand L binds to the extracellular domain of a transmembrane receptor, whereby the receptor is activated for signal transmission to the cytosolic side. The cytosolic domain of the activated receptor R* transmits the signal to signal proteins next in sequence. b) Signal transduction via ligand-gated ion channels. The ligand binds to the extracellular side of a receptor that also functions as an ion channel. Ligand binding induces the opening of the ion channel, there is an ion efflux and a change in the membrane potential. c) Signal transduction via voltage-gated ion channels. A change in the membrane potiential $\Delta V$ is registered by an ion channel which transitions from the closed to the open state.

*voltage-gated ion channels* (Fig. 5.1c). The potential-driven passage of ions through ion channels is the basis for stimulation in nerves.

Transmembrane receptors have the following roles in signal transduction:

- Reception of the signal via specific binding of the ligand at the outer side of the membrane and/or registering a change in membrane potential. Reception of the signal is synonymous with activation of the receptor for further transmission of the signal.
- Further transmission of the signal through the membrane. The activated receptor conducts the signal to the inner side of the membrane. Transmission of the signal implies specific communication with the next component of the signal transmission pathway on the inner side of the cell membrane. In the case of volatage- or ligand-gated ion channels, reception of the signal is linked to opening of the channel.

**Intracellular Activation of Receptors**

We also know of receptors for which the reception of the signal and activation take place on the *inner side* of the membrane. The cGMP-dependent ion channels involved in signal conduction in the vision process (see Chapter 16) are ligand-regulated ion channels with an open state controlled by intracellularly created cGMP. Another example are the receptors for inositol triphosphate which are localized in the membrane of $Ca^{2+}$ storage organelles and also have the character of ligand-controlled ion channels. Inositol triphosphate is an intracellular messenger substance that binds to the cytosolic side of the corresponding receptor. Ligand binding leads to opening of the ion channel via a conformational change and thus to influx of $Ca^{2+}$-ions from the storage organelle into the cytosol (see 6.5).

## 5.2 Structural Principles of Transmembrane Receptors

Transmembrane receptors are *integral membrane proteins*, i.e., they possess a structural portion that spans the membrane. An *extracellular domain*, a *transmembrane domain* and an *intracellular* or *cytosolic domain* can be differentiated within the structure (Fig. 5.2a).

### 5.2.1 The Extracellular Domain of Transmembrane Receptors

In many receptors, the extracellular domain contains the ligand binding site. Glycosylation sites, i.e., attachment sites for carbohydrate residues, are also located nearby in the extracellular domain.

The structure of the extracellular domain can be very diverse and is determined by the number of transmembrane sections, as well as the subunit structure of the receptor.

The extracellular localized protein portion may be formed from a continuous protein chain and may include several hundred amino acids. If the receptor crosses the

membrane with several transmembrane segments, the extracellular domain is formed from several loops of the protein chain that may be linked by disulfide bridges.

Transmembrane receptors may show *homotrophic composition* (*identical subunits*) or *heterotrophic* composition (*different subunits* Fig. 5.2b), so that the extracellular domain may be made up of several identical or different structural elements.

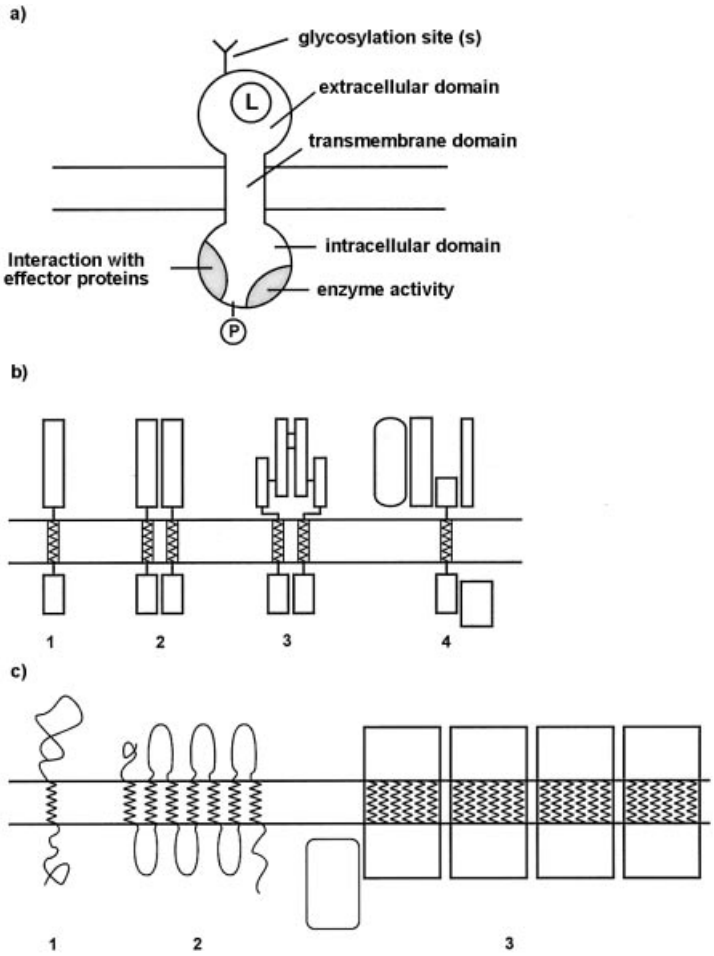



**Fig. 5.2.** Structural principles of transmembrane receptors. a) Representation of the most important functional domains of transmembrane receptors. b) Examples of subunit structures. Transmembrane receptors can exist in a monomeric form (1), dimeric form (2) and as higher oligomers (3,4). Further subunits may associate at the extracellular and cytosolic domains, via disulfide bridges (3) or via non-covalent interactions (4). c) Examples of structures of the transmembrane domains of receptors. The transmembrane domain may be composed of an α-helix (1) or several α-helices linked by loops at the cytosolic and extracellular side (2). The 7-helix transmembrane receptors are a frequently occurring receptor type (see 5.3). Several subunits of a transmembrane protein may associate into an oligomeric structure (3), as is the case for voltage-controlled ion channels (e.g., $K^+$ channel) or for receptors with intrinsic ion channel function (see Chapter 17).

We also know of receptors in which only one subunit spans the membrane, whilst other subunits are bound to this subunit on the extracellular side via protein-protein interactions or via disulfide bridges (Fig. 5.2b and examples in Chapter 11).

## 5.2.2 The Transmembrane Domain

The transmembrane domains have different functions, according to the type of receptor. For ligand-controlled receptors, the function of the transmembrane domain is to pass the signal on to the cytosolic domain of the receptor. For ligand-controlled ion channels, the transmembrane portion forms an ion pore that allows selective passage of ions (see Chapter 16).

The transmembrane receptors span the ca. 5 nm thick phospholipid bilayer of the cell membrane with structural portions known as *transmembrane elements*. The inner of a phospholipid layer is hydrophobic and, correspondingly, the surface of the structural elements that come into contact with the inner of the phospholipid double layer also has hydrophobic character.

The transmembrane domain may be made up of one or many transmembrane elements. Generally, the transmembrane elements include 20–25 mostly hydrophobic amino acids. At the interface with aqueous medium, we often find hydrophilic amino acids in contact with the polar head groups of the phospholipids. In addition, they mediate distinct fixing of the transmembrane section in the phospholipid double layer. A sequence of 20–25 hydrophobic amino acids is seen as characteristic for membrane-spanning elements. This property is used in analysis of protein sequences, to predict possible transmembrane elements in so-called „hydropathy plots".

### Structure of Transmembrane Elements

High resolution structural information about the transmembrane elements of membrane receptors is not currently available, since it is not yet possible to obtain transmembrane receptors in crystalline form for structural analysis. Due to the hydrophobic nature of the transmembrane elements, crystallization is very difficult.

Thus, the important question of the secondary structure of the transmembrane elements can only be addressed with models and by structural comparison with other transmembrane proteins for which the structure has been resolved. Detailed information on the structure of transmembrane elements is available for the photoreaction center of *Rhodopseudomonas viridis* (review: Deisenhofer and Michel, 1989), cytochrome c oxidase (Iwata et al., 1995) and the OmpF porin of *E. coli* (Cowan et al., 1992; Fig. 5.3), amongst others. In addition, high resolution electron microscopic investigations and X-ray studies of bacteriorhodopsin, a light-driven ion pump with seven transmembrane elements, have yielded valuable information on the structure and configuration of membrane-spanning elements (Henderson et al., 1990; Kimura et al., 1997; Pebay-Peyrula et al., 1997; Fig. 5.4). With the successful crystallization of the photoreaction center of *Rhodopseudomonas viridis*, a membrane protein was displayed at atomic resolution for the first time (Deisenhofer et al., 1985). The membrane-

positioned structural elements of the photoreaction center have — in agreement with the prediction — an α-helical structure. For bacteriorhodopsin, the high-resolution electron microscopy and the X-ray investigations confirm earlier predictions that the transmembrane elements have α-helical character.

In contrast, the transmembrane domain of the bacterial OmpF porin is made up of β-elements. The β-elements, in this case, are not mostly made up of hydrophobic amino acids.

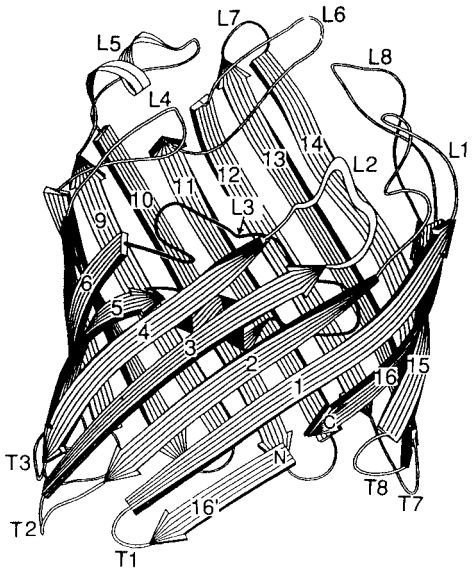



**Fig. 5.3.** Structure of the OmpF porin of *E. coli*. The porin is a bacterial membrane protein with β-sheet structures as transmembrane elements. The structure of a monomer of the OmpF porin is shown. In total, 16 β-bands are configured in the form of a cylinder and form the walls of a pore through which selective passage of ions takes place. L1—L8 are long loops, T1,2,3 and T7,8 are short bends (T turn) that link the β-sheets. According to Cowan et al. (1992), with permission.

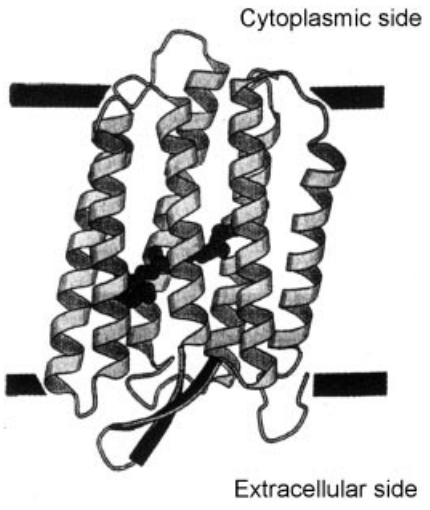



**Fig. 5.4.** Structure of the bacteriorhodopsin from *Halobacterium halobium*. Ribbon diagram of bacteriorhodopsin and retinal as a ball-and-stick model. Bacteriorhodopsin crosses the membrane with seven α-helices that are arranged in a bundle form with the chromophore retinal bound in the interior. According to Kimura et al. (1997), with permission.

At present, it is generally assumed that transmembrane receptors span the cell membrane as α-helices. However, it is not known how often other structural elements occur in the transmembrane domains of receptors. Thus, the presence of β-sheet structures, particularly in the case of receptors with complex structures, cannot be excluded (Hucho et al., 1994).

**Composition of the Transmembrane Domain**

The transmembrane domain may consist of one or several transmembrane elements (see also Fig. 5.2). In the latter case, these are arranged in the form of bundles, as shown in Figure 5.4 for bacteriorhodopsin. In the case of ion channels, in which several subunits are involved in formation of the transmembrane domain (see acetylcholine receptor, Fig.16.12), prediction of the structure of the membrane portion is very difficult. The different transmembrane elements are no longer equivalent in these cases. Part of the element is involved in formation of the inner wall of the pore, other structural elements form the surface to the hydrophobic inner of the phospholipid bilayer. It is evident that the polarity requirements for the amino acid side chains vary according to the position of the transmembrane elements (see Chapter 16).

In these cases, how the different transmembrane elements associate to an ordered structured transmembrane domain can only be discussed based on electron microscopy and crosslinking experiments and affinity marking. Predictions based on these experiments are really more like models.

## 5.2.3 The Intracellular Domain of Membrane Receptors

Two basic mechanisms are available for conduction of the signal to the inner side of the membrane (Fig. 5.5).

- Via specific protein-protein interactions, the next protein component in the signal transmission pathway, the effector protein, is activated. Activation of the effector molecule must be preceeded by activation of the receptor by a signal.
- Arrival of the signal triggers enzyme activity in the cytosolic domain of the receptor that, in turn, pulls other reactions along with it. The enzyme activity of the cytosolic domain is often tyrosine kinase activity; however, there are other examples where tyrosine phosphatase or Ser/Thr-specific protein kinase activity is activated. In all these examples, the cytoplasmic domain carries an enzyme activity regulated by ligand binding. The enzyme activity may be an integral part of the receptor, or it may also be a separate enzyme associated with the receptor on the inner side of the membrane (cf. Chapter 8 and Chapter 12).

Starting from the activated receptor, a large number of reactions can be set in motion (Fig. 5.5). One main route of signal transmission takes place by activation of G-proteins, another via activation of tyrosine-specific protein kinases, and a further route is via activation of ion channels. In the further course of G-protein mediated signal transmis-

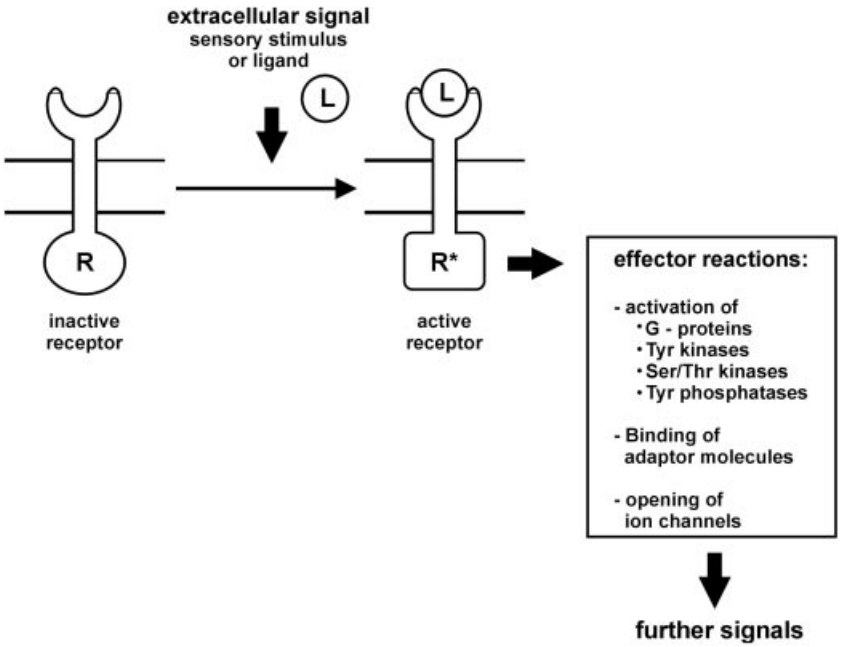



**Fig. 5.5.** General functions of transmembrane receptors. Extracellular signals convert the transmembrane receptor from the inactive form R to the active form R*. The activated receptor transmits the signal to effector proteins next in the reaction sequence. Important effector reactions are the activation of heterotrimeric G-proteins, of protein tyrosine kinases and of protein tyrosine phosphatases. The tyrosine kinases and tyrosine phosphatases may be an intrinsic part of the receptor or they may be associated with the receptor. The activated receptor may also include adaptor proteins in the signaling pathway or it may induce opening of ion channels.

sion, secondary diffusible signals are often formed, the „second messenger" molecules (see Chapter 3 and Chapter 6). These function as effectors and activate further enzyme systems in the sequence, especially protein kinases.

The activated receptor can also associate with adaptor molecules, which serve as coupling elements for further signal proteins.

### 5.2.4 Regulation of Receptor Activity

A physiologically important aspect of signal transmission via transmembrane receptors is its regulation. The cell has various mechanisms available, with the help of which the number and activity of transmembrane receptors can be regulated. The aim of regulation is, for example, to weaken signal transmission via the receptor during conditions of long lasting hormonal stimulation. The structural elements involved in regulation of receptor activity are generally located in the cytosolic domain. These are, above all, protein sequences that permit phosphorylation of the receptor by protein kinases.

Phosphorylation at Ser/Thr or Tyr residues of the cytosolic domain may lead to inactivation or activation of the receptor and thus weaken or strengthen signal transmission. In this way, Ser/Thr-phosphorylation is used in the process of internalization of receptors, in order to remove the receptor from circulation after it has been activated (see 5.3.4). Furthermore, Ser/Thr phosphorylation can be used as a switch for coupling a given receptor to different $G_\alpha$ subunits. Protein kinase A mediated phosphorylation of the β-adrenergic receptor has been shown to switch coupling of the receptor from $G_s$ to $G_i$ and initiate a new set of signaling events (Daaka et al., 1997).

The cytosolic domain thus carries sequences important for short-term or long-term regulation of receptor activity.

## 5.3 G-protein Coupled Receptors

Of the transmembrane receptors that receive signals and conduct them into the cell interior, the G-protein coupled receptors are of central importance. Vertebrates contain ~1000 different G-protein coupled receptors that may be activated by extracellular ligands or sensory signals. The ligands include biogenic amines, such as adrenaline and noradrenaline, histamine, serotonin, retinal derivatives, peptides such as bradykinin and large glycoproteins such as luteinizing hormone and parathormone (see also Tab. 3.1.).

Activation of receptors can also be mediated by proteolytic cleavage of the extracellular domain of the receptor by proteases like thrombin. For these *protease activated receptors*, a proteolytically produced peptide functions as the activating ligand.
In addition, physical stimuli such as light signals are registered and converted into intracellular signals by G-protein coupled receptors; they are also involved in perception of taste and smell.

Ligand binding or reception of a physical signal is linked to activation of the G-protein coupled receptor. As a consequence, the receptor undergoes a conformational change that is transmitted to the inner side of the membrane, whereby the next sequential member of the signal chain, a G-protein, is activated. This conducts the signal further via other reaction pathways (see Fig. 5.14). A review of G-protein coupled receptors is given in Strader et al. (1994).

### 5.3.1 Structure of G-Protein Coupled Receptors

Based on sequence data of a large number of G-protein coupled receptors, a distinct structual homology can be demonstrated. The comparable function of the different receptors is reflected in the appearance of common structural elements.

The occurrence of seven sequence segments, each made up of 20—25 amino acids, is characteristic for the G-protein coupled receptors. From this, it is assumed that they form transmembrane domains, and span the membrane in the form of α-helices. The transmembrane elements are linked by loops of various sizes on the outer and inner side.

Due to the common appearance of 7 transmembrane helices, the family of G-protein coupled receptors is also known as the family of the *7-helix transmembrane receptors.* The G-protein coupled receptors are also sometimes called the *serpentine receptors*, pointing to the serpentine-like configuration of transmembrane helices.

Figure 5.6 shows the primary sequence of the β-receptor for adrenaline with the assumed topology of the seven transmembrane helices. The highest sequence homology of the G-protein coupled receptors is found in the transmembrane elements, whilst the hydrophilic loop regions show stronger divergence between different receptors.

The G-protein coupled receptors are often *glycoproteins.* Glycosylation sites are located in the extracellular region, e.g. in the form of the consensus sequence Asn-X-Ser/Thr for an N-linked glycosylation.

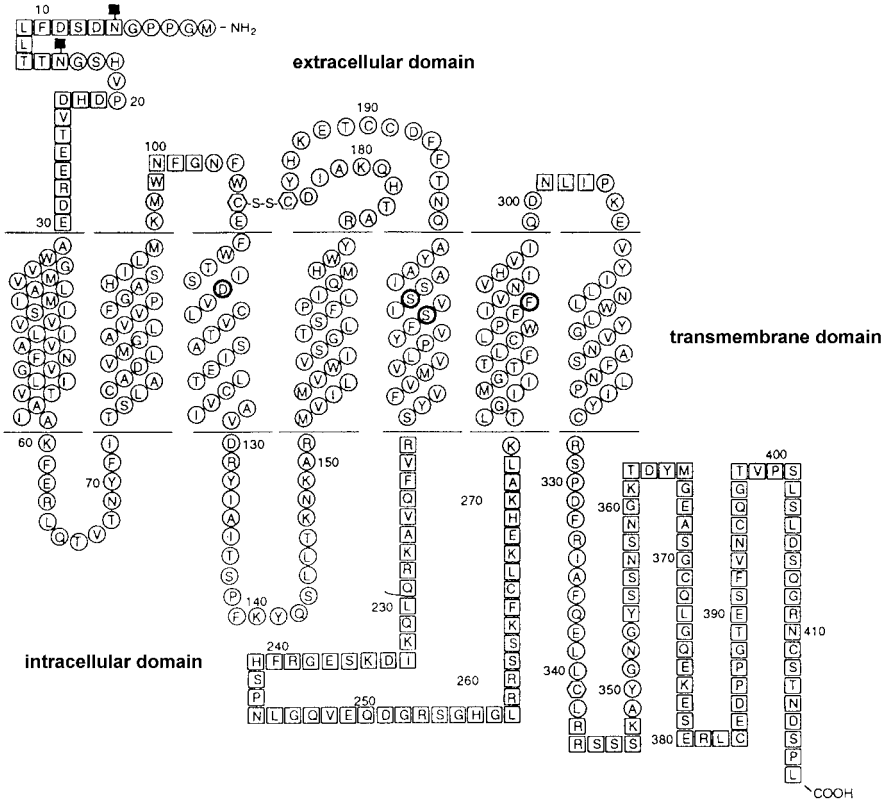



**Fig. 5.6.** Topology of the β-adrenergic receptor of hamster. The primary structure is shown of the β-receptor for adrenaline from hamster, with the assumed topology of the seven transmembrane helices. The extracellular domain is shown at the top of the picture. The interface of the cell membrane is indicated by the dashed line. The filled squares show glycosylation sites. Amino acids not required for ligand binding, according to mutagenesis studies, are shown as open squares. Reprinted with permission of the *American Journal of Respiratory Cell and Molecular Biology* (1989), 1, No.2, p.82.

The extracellular loops contain frequently *conserved Cys residues*. It is assumed that these stabilize the conformation of the extracellular domain, via disulfide bridges.

Posttranslational modification in the form of *palmitoylation* of the cytosolic domain (see 3.7.2) has been demonstrated for the α- and β-adrenergic receptor. The palmitoylation takes place on a Cys-residue localized at the C-terminus on the membrane inner side. The modification may possibly serve to anchor the C-terminus in the membrane.

The currently accepted structural models of the G-protein coupled receptor tend strongly towards the well established structure of bacteriorhodopsin (Fig. 5.4) that is also a 7-helix transmembrane protein. The model assumes that the seven helices are bedded bundle-wise in the membrane. Detailed structural information on the conformation of the extracellular and intracellular structural portions is still lacking.

### 5.3.2 Ligand Binding

The area of ligand binding has been particularly well defined for the receptors of small ligands, with the help of targeted mutagenesis and pharmacological investigations. The receptors, to which small ligands are assigned, bind these in the interior of the transmembrane domain. In agreement with this idea, it has also been shown that the extracellular and intracellular sequence portions of the receptors are not needed for ligand binding in these cases.

For the receptors that have peptides or proteins as ligands, structural portions of the extracellular domain, in addition to areas of the transmembrane domain, are involved in ligand binding.

### 5.3.3 Mechanism of Signal Transmission

The mechanism by which the activated receptor talks to the G-protein is only partially understood. Generally, the switch function of the receptor is considered in terms of allosteric conformational changes of the 7-helix membrane bundle (review: Bourne, 1997). According to this representation, changes in the structure of the transmembrane bundle are passed on to the cytoplasmic loops of the receptor. Communication with the α-subunit of the heterotrimeric G-protein takes place via these loops.

Electron microscopic investigations of 2D crystals of rhodopsin show a bundle-shaped structure of the seven transmembrane helices, in which three helices are vertical and the others are at more of an angle to the membrane. It is assumed that ligand binding is associated with a change of the mutual configuration of the helices, whereby separation of helices is also under discussion. As a consequence, a conformational change takes place of the intracellular loop of the receptor, which creates a binding surface for high affinity binding of the α-subunit of the G-protein.

The heterotrimeric G-protein, which exists as the inactive GDP form, now binds via its α-subunit to the activated receptor and is activated itself. An exchange of GDP for GTP takes place and the βγ-subunit of the G-protein dissociates (see 5.5.3). Once the G-protein is activated, it frees itself from the complex with the receptor, which either returns to its inactive ground state or activates further G-proteins.

### 5.3.4 Switching off and Desensitization of G-Protein Coupled Receptors

A phenomenon often seen in transmembrane receptors in general, and in G-protein coupled receptors in particular, is desensitization (Fig. 5.7). *Desensitization* means a weakening of the signal transmission under conditions of long-lasting stimulation by hormones, neurotransmitters or by sensory signals. Despite the persistent effect of extracellular stimuli, the signal is no longer passed into the cell interior, or only in a weakened form, during desensitizing conditions. This is a mechanism with which both short-term and long-term regulation of receptor activity is possible.

The best investigated is the desensitization of the adrenaline receptor type β2 and of rhodopsin. Rhodopsin has the function of a light receptor in the process of vision. It receives light signals and conducts them to the relevant G-protein, transducin. The *key reaction* in desensitization of both systems is the phosphorylation of the receptor at the cytoplasmic side by specific protein kinases.

Two classes of protein kinases are involved in the phosphorylation and desensitization (review: Freedman and Lefkowitz, 1996):

#### Phosphorylation by cAMP-dependent Protein Kinases

Phosphorylation of the cytoplasmic domain of G-protein coupled receptors can take place via *cAMP-dependent protein kinase*s (*protein kinase A*) or via *protein kinase C* (Chapter 7) (Fig. 5.8). This is a feedback mechanism: the hormonal activation of the receptor leads, via G-proteins and adenylyl cyclase/cAMP, to activation of protein kinases of type A (see 5.6.1, 6.1 and Chapter 7). The activated protein kinases phosphorylate the receptor in the region of the cytoplasmic domain on Ser/Thr residues (see Fig. 5.6).

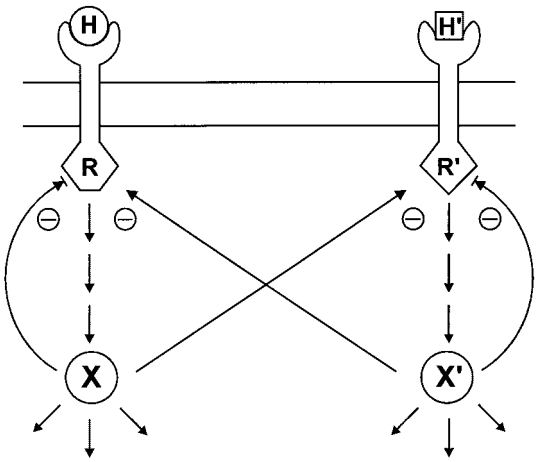



**Fig. 5.7.** General principle of densensitization of G-protein-coupled receptors. Desensitization of a hormone-bound receptors can take place by two principle routes, schematically represented in the figure. A suppressing influence may be exerted on the receptor system via proteins (X) of a signal chain, triggering activation of the signal chain. Receptor systems may also mutually influence one another in that a signal protein X formed in one signal chain mediates the densensitization of another receptor system R*, and vice versa.

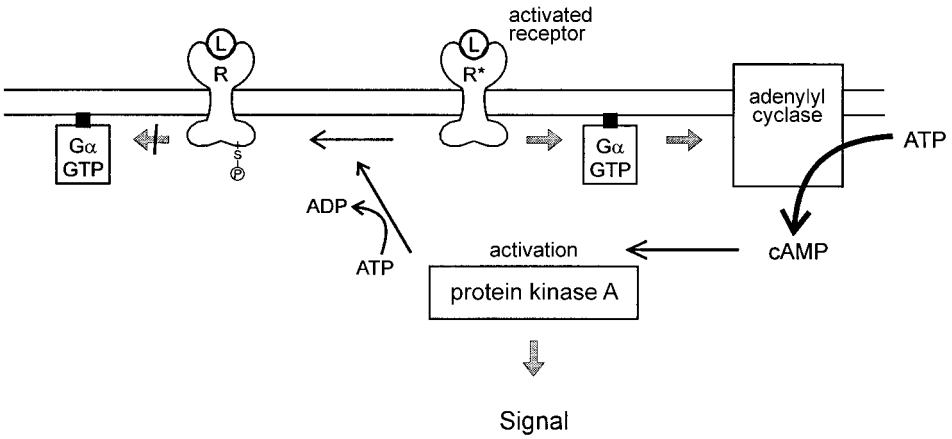



**Fig. 5.8.** Desensitization of G-protein-coupled receptors via cAMP-dependent protein kinases. Starting from an activated receptor R*, the signal is transmitted via the $G_\alpha$-subunit of the G-protein to adenylyl cyclase. The latter is activated and forms cAMP. This activates a protein kinase of type A that passes the signal in the form of a Ser/Thr-specific protein phosphorylation to substrate proteins. One of the substrates is also the receptor that is phosphorylated in the region of the cytoplasmic domain by the activated protein kinase A. The ligand-bound receptor is preferentially phosphorylated. As a consequence of phosphorylation, activation of further G-proteins by the receptor is suppressed.

### Phosphorylation via G-protein Coupled Receptor Protein Kinases (GRK)

Well characterized GRKs (review: Pitcher et al., 1998) are those for rhodopsin, rhodopsin kinase and for the β2 adrenaline receptor, the β-adrenergic receptor kinase (βARK). The GRKs are protein kinases that are not regulated via cAMP. Phosphorylation of the receptor takes place in the C-terminal region and/or in the third cytoplasmic domain. Only the activated, i.e. occupied by an agonist, receptor is phosphorylated. It is assumed that the GRK is transported from the cytosol to the inner side of the membrane to phosphorylate the receptor. During translocation of the GRK to the membrane-localized receptor, the βγ-subunit of the G-protein is thought to play an important role (see Fig. 5.9 and 5.5.7). Furthermore, binding of phosphatidyl-inositol-messengers (see Chapter 6) to the PH-domain (see Chapter 8) of GRKs helps to translocate the GRK to the membrane.

Phosphorylation of the receptor can have two consequences (Fig. 5.9):

### Translocation

The phosphorylated receptor is brought into the cell interior. It is internalized in the membrane-associated form, dephosphorylated and then transported back to the cell membrane. The translocation into the cell interior serves, in particular, to weaken signal transmission during conditions of long-lasting hormonal stimulation.

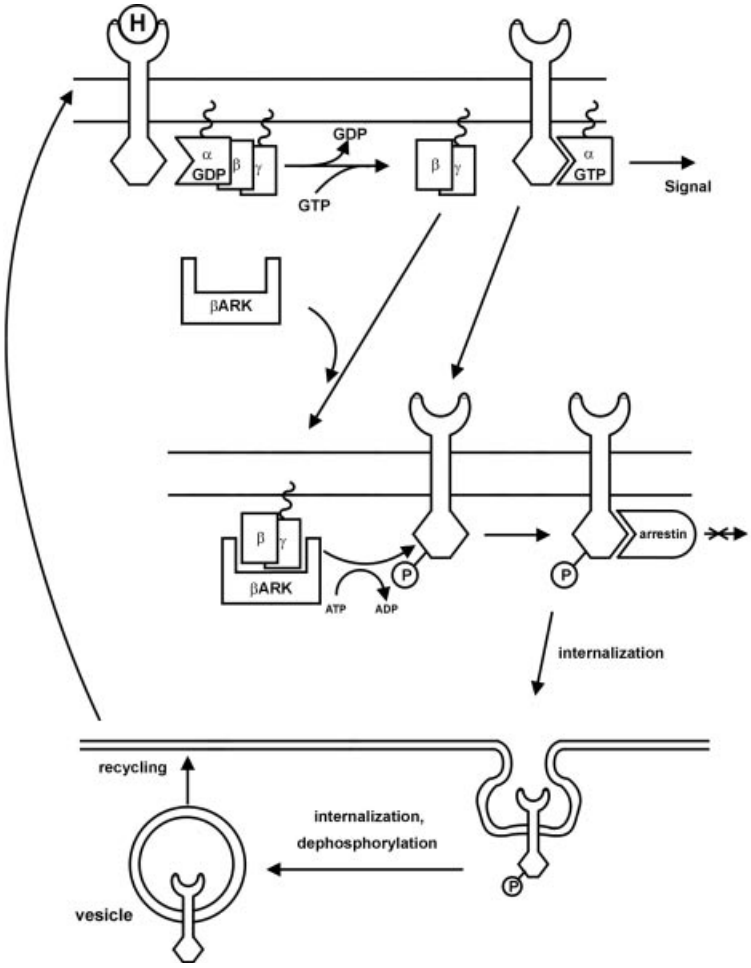



**Fig. 5.9.** Receptor desensitization: translocation and arrestin binding. The βγ-complex released on activation of the G-protein associates with the β-adrenergic receptor kinase (βARK) and recruits this to the membrane. Consequently, the βARK phosphorylates the activated β-receptor and removes it from the signal chain. Arrestin binds to the phosphorylated receptor. In the arrestin-bound form, the signal can no longer be transmitted to the G-protein and signal conduction is disrupted. The phosphorylated receptor is transported in the form of vesicles into the cell interior (internalization) and, after dephosphorylation, is returned to the membrane (recycling).

### Binding of Arrestin

Phosphorylation of the receptor may lead to creation of a specific binding site for proteins known as *arrestins*. Binding of arrestins at the phosphorylated receptor decouples the receptor from the interaction with the G-protein next in the sequence so that signal transmission is suppressed. Arrestin binding serves, e.g. to rapidly weaken signal transmission during the vision process, during conditions of long lasting light stimulus.

## 5.4 Regulatory GTPases

The G-proteins belong to the large family of regulatory GTPases; these bind GTP and hydrolyze it, thereby functioning as a switch in central cellular processes. The family of regulatory GTPases is also called the GTPase superfamily.

### 5.4.1 The GTPase Superfamily: General Functions and Mechanism

Proteins of the GTPase superfamily are found in all plant, bacterial and animal systems. The following examples illustrate the central functions of the regulatory GTPases in the cell.

Regulatory GTPases are involved in:

- Protein biosynthesis on ribosomes
- Signal transduction at membranes
- Visual perception
- Sense of smell and taste
- Control of differentiation and cell division
- Translocation of proteins through membranes
- Transport of vesicles in the cell

The members of the GTPase superfamily show an extensively conserved reaction mechanism. A common trait is a switching function that enables a reaction chain to be switched on or off (review: Bourne et al., 1990).

#### The Switch Function of the GTPases

The regulatory GTPases are involved in reaction chains by functioning as a switch. The switch function is based on a cyclical, unidirectional transition between an *active, GTP bound form* and an *inactive, GDP bound form* (Fig. 5.10). The binding of GTP brings about the transition into the active form. Hydrolysis of the bound GTP by an intrinsic GTPase activity converts the protein into the inactive, GDP bound form. In both inactive and active forms, the proteins of the GTPase superfamily possess a specific affinity to other proteins employed earlier or later in the reaction chain. In the active GTP form, the signal can be passed on to the effector molecule next in sequence. Conversely, if the GTPase is in the inactive, GDP bound form, the reaction chain is disrupted. To reactivate the reaction chain, the GDP must be exchanged for GTP. This takes place by dissociation of GDP from the inactive complex, a process that generally occurs by interaction with upstream protein components of the signaling chain.

#### The GTPase Cycle

The GTPases bring about the transition between the active and inactive states in a *cyclic process* that can only run in one direction, due to the irreversible hydrolysis of GTP.

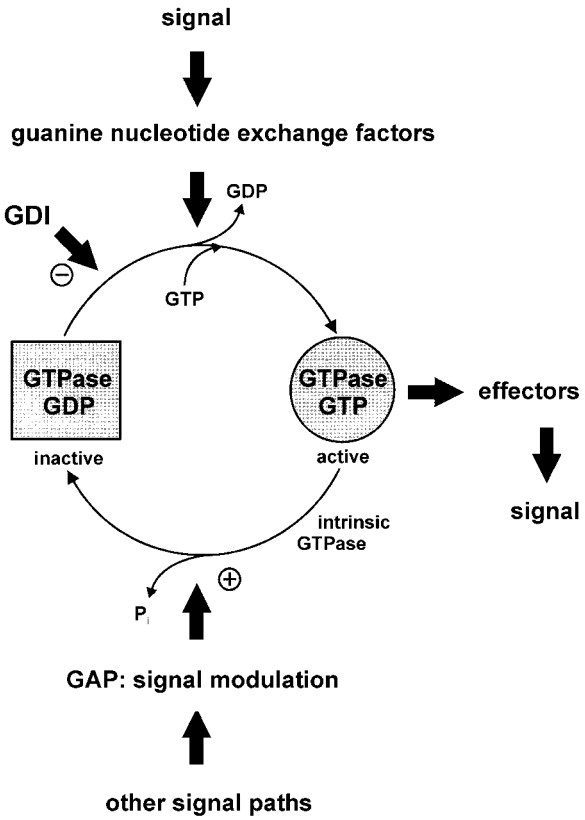



**Fig. 5.10.** The switch function of the regulatory GTPases. The GTP form of the regulatory GTPases represents the „switched on" form of the GTPase, the GDP form, in contrast, the „switched off" form. The switch function of the regulatory GTPases may be controlled by guanine nucleotide exchange factors, by GTPase activating proteins (GAPs) and by G-nucleotide dissociation inhibitors (GDIs). The regulatory GTPases run through a GTPase cycle which signals flow into via GEFs and are conducted further in the form of the GTPase-GTP complex to effector molecules further down the sequence. Hydrolysis of the bound GTP ends the activated state. The rate of GTP hydrolysis is either intrinsically determined or may be accelerated via GAPs.

At least three different GTPase states can be differentiated in the GTPase cycle: the active GTP form, the inactive GDP form and an „empty" form of the GTPase, which is generally a short-lived state.

The switch function of the GTPase is based on the specific ability of the different functional states of the GTPase to interact with the proteins that precede and follow in the signal chain. A particular GTPase is characterized by the proteins with which the active and inactive forms interact. A special characteristic of the active GTP form is that it may activate effector enzymes further on in the reaction chain, e.g., adenylyl cyclase, and thus actively transmits the signal.

**Modulation and Regulation of the Switch Function**

How effectively a signal can be transmitted by a GTPase depends on the relationship of the concentration of the active GTP form to the inactive GDP form. This, in turn, is determined by the relationship of the rate constant for the dissociation of GDP, $k_{diss,GDP}$ to the rate constant of GTP hydrolysis, $k_{cat,GTP}$:

$$\frac{GTPAse \cdot GTP}{GTPAse \cdot GDP} = \frac{k_{diss.GDP}}{k_{cat.GTP}}$$

This relationship is valid if it can be assumed that the GTP concentration is not limited and that GTP binds very rapidly to the „empty" form of the GTPase. A special characteristic of the regulatory GTPases is that both rate constants may be regulated by specific proteins. The proportion of GTPase that exists in the active form can be altered by at least three processes:

1. Acceleration of the dissociation of GDP increases the proportion of the active form. The rate of dissociation of GDP may be increased by specific proteins. These proteins are known as *guanine nucleotide exchange factors* (GEF). For the heterotrimeric G-proteins, the agonist-bound, activated receptor is the exchange factor.
2. Dissociation of GDP may be inhibited by specific proteins known as *guanine nucleotide dissociation inhibitors* (GDI). Proteins with this function are found in all members of the Ras family (see Chapter 9). The GDIs have the function, above all, to provide a cytosolic pool of inactive, GDP-bound proteins (see Chapter 9.1).
3. Increasing of the rate of GTP hydrolysis by *GTPase activating proteins* (GAP) reduces the lifetime of the active, GTP-bound state. The GAP protein class is an important instrument for control of the rate of signal transmission. Activation of the GAPs leads to termination or weakening of signal transmission. Often, the activity of the GAPs is regulated by other signaling pathways. Thus, a regulatory influence on signal transmission via G-proteins can be achieved from another signaling pathways.

The various GTPases may differ to a large extent in the rate of GTP hydrolysis and thereby in the influence of GTPase-activating proteins. The Ras protein and the $G_{\alpha,t}$, known as transducin, involved in the process of vision, are cited as examples.

The Ras protein has low intrinsic GTPase activity. This may be increased ca. $10^5$-fold by the corresponding GTPase-activating protein (see also Chapter 9). In comparison, the intrinsic rate of GTP hydrolysis of transducin is ca. 100-fold higher that that of the Ras protein. The effector molecule next in the reaction chain, the cGMP phosphodiesterase, functions as the GAP here and stimulates GTPase activity of the transducin 100-fold.

## 5.4.2 Inhibition of GTPases by GTP Analogs

*Non-hydrolyzable GTP analogs* are an indispensable tool in identification and structural and functional characterization of GTPases. The GTP analogs shown in Fig. 5.11, GTPγS, β,γ-methylene GTP and β,γ-imino GTP, are either not hydrolzed by GTPases

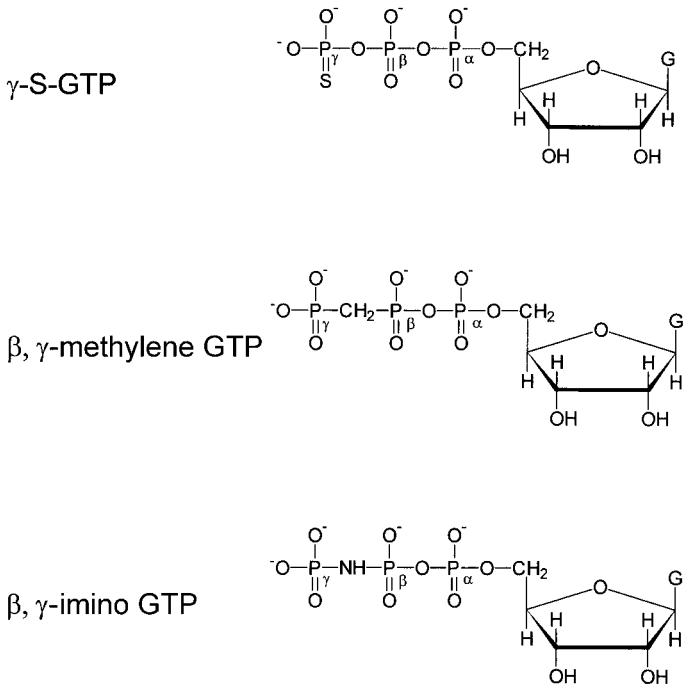



**Fig. 5.11.** Examples of non-hydrolysable GTP analogs.

or only very slowly. Addition of these analogs fixes the G-protein in the active form; it is permanently „switched on". For cellular signal transduction, this means permanent activation of the signal transmission pathway. In many cases, a role of G-proteins in a signal chain was concluded from the observation that non-hydrolyzable GTP analogs bring about a lasting activation of signal transmission. The GTP analogs were equally important for structural determination of the activated form of GTPases. Formation of a stable complex between the non-hydrolyzable GTP analog and different GTPases has enabled crystallization of the complex in its activated form.

### 5.4.3 The G-Domain as Common Structural Element of the GTPases

A common property of the GTPases is the enzymatic activity of GTP hydrolysis. GTP binding and hydrolysis takes place in a domain of the GTPases known as the *G-domain*. Fig. 5.12 shows the G-domain of the bacterial elongation factor EF-Tu. In all GTPase structures known at present, the G-domain has very similar architecture and very similar means of binding the guaninine nucleotide. The sequence element $GX_4GK(S/T)$ is a consensus sequence for guanine nucleotide binding; this sequence is involved in binding the β- and γ-phosphate of GTP and GDP and is also known as the P-loop. Other consensus sequences, such as $RX_2T$ and $DX_2G$ are both involved

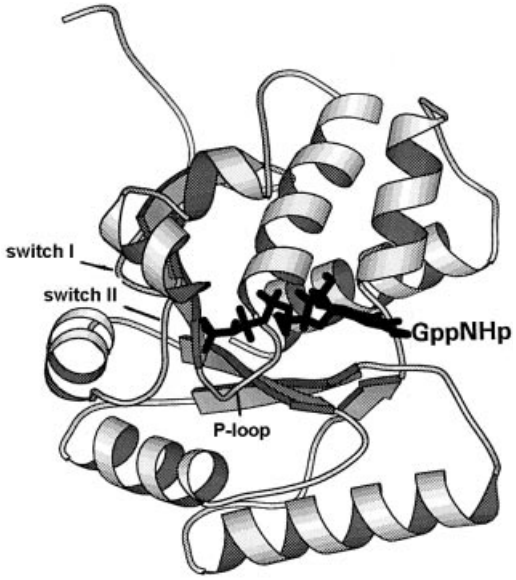



**Fig. 5.12.** Structure of the G-domain of the elongation factor EF-Tu from *T. thermophilus* with bound GppNHp, according to Berchthold et al., (1993). The non-hydrolysable analog GppNHp, the P loop and the switch regions I and II are shown, which play an important role in transition from the inactive GDP form to the active GTP form (see also 5.5.6 and 9.2.1). MOLSKRIPT representation according to Kraulis, (1991).

in binding the γ-phosphate and in the GTPase reaction (X = any amino acid). A further consensus sequence (N/T)(K/Q)XD and SA interacts with the guanosine.

### 5.4.4 The Different GTPase Families

The superfamily of GTPases with their more than hundred members are divided by sequence homologies, molecular weight and subunit structure into further (super)families. These are the families of the *heterotrimeric G-proteins*, the *Ras/GTPase superfamily* and the *family of initiation and elongation factors* (Fig. 5.13).

The heterotrimeric G-proteins are built of three subunits, with the GTPase activity localized on the largest subunit (see 5.5). The members of the Ras/GTPases, in contrast, are monomeric proteins with a molecular weight of ca. 20 kDa (see Chapter 9).

A further functionally diverse class is made up of the proteins involved in protein biosynthesis and membrane transport. GTPases with functions in protein biosynthesis include the elongation factors, termination factors and peptide translocation factors. These are mostly monomeric proteins with molecular weights of 40–50 kDa. GTPases of this class are also found in protein complexes such as the „signal recognition particle“ (SRP) and the corresponding receptor. Both protein complexes are needed during ribosomal protein biosynthesis, for transport of newly synthesized proteins through the endoplasmic reticulum.

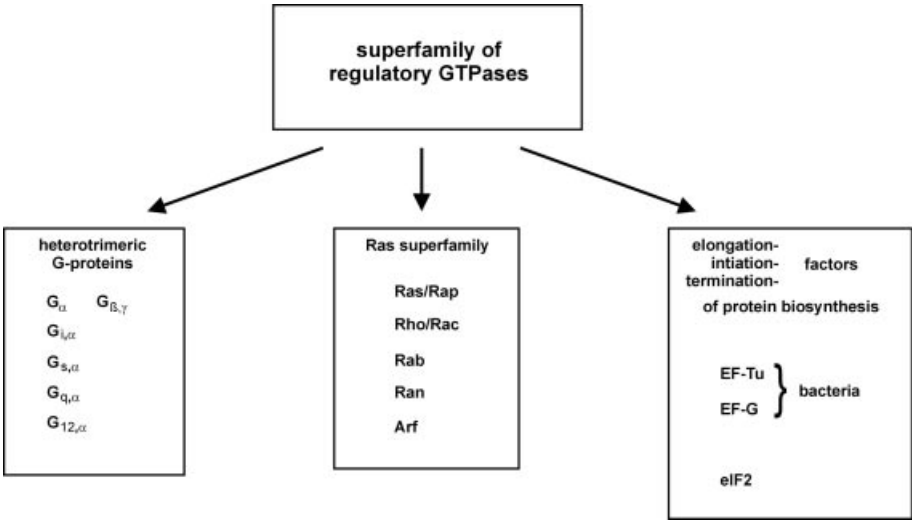



**Fig. 5.13.** The GTPase superfamily.

## 5.5 The Heterotrimeric G-Proteins

The heterotrimeric G-proteins are the specific reaction partners in signal transmission via 7-helix transmembrane receptors, which is why these receptors are also known as G-protein-coupled receptors. From the G-protein, the signal is then passed on to the effector protein next in the sequence (review: Hepler and Gilman, 1992; Neer, 1995).

A common structural feature of the G-proteins is their construction from three subunits (Fig. 5.14), a large α-subunit of 39–46 kDa, a β-subunit of 37 kDa and a γ-subunit of 8 kDa. The α-subunit has a binding site for GTP or GDP and carries the GTPase activity. The β- and γ-subunits exist as a tightly associated complex and are active in this form. All three subunits show great diversity, so that at least 20 different genes for α-subunits, 5 for β-subunits and 12 for γ-subunits are known in mammals. Some G-proteins are ubiquitous, whereas others only occur in specialized tissue.

Specificity of the switch function is mostly determined by the α-subunit: the α-subunit carries out the specific interaction with the receptors preceding in the signal chain and with the subsequent effector molecules. The βγ-complex may also be involved in signal transmission to the effector proteins.

### 5.5.1 Classification of the Heterotrimeric G-Proteins

The important functions of signal transmission by G-proteins are realized by the α-subunit. Since different G-proteins interact with very different partners, there are significant differences in the structure of the α-subunits. Due to the common GTPase

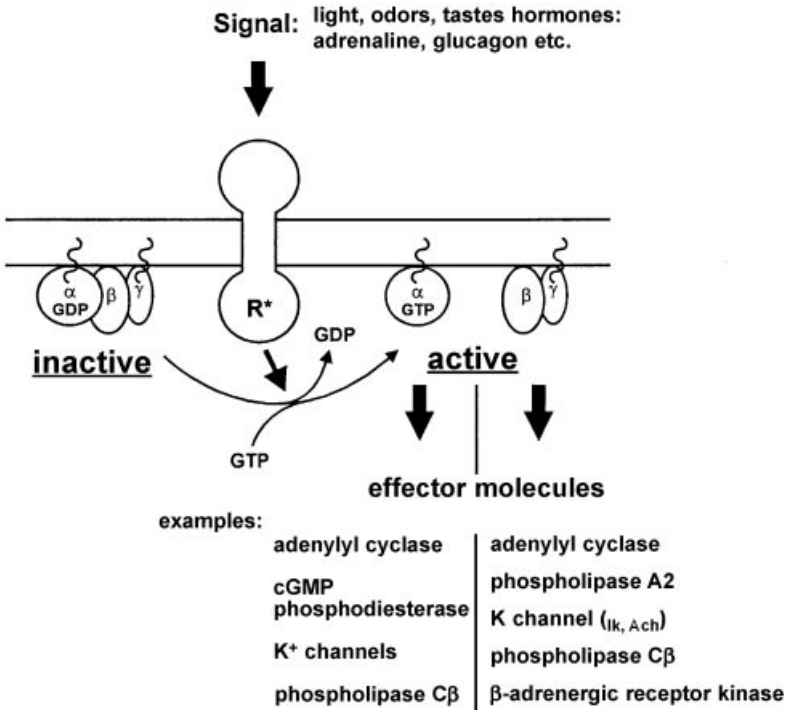



**Fig. 5.14.** Structure and activation of the heterotrimeric G-proteins. Reception of a signal by the receptor activates the G-protein, which leads to exchange of bound GDP for GTP at the α-subunit and to dissociation of the βγ-complex. Further transmission of the signal may take place via Gα-GTP or via the βγ-complex, which interact with corresponding effector molecules. The α- and γ-subunits are associated with the cell membrane via lipid anchors. Signal reception and signal transmission of the heterotrimeric G-proteins take place in close association with the cell membrane. This point is only partially shown in the figure.

domain and the common interaction with the βγ-subunits, however, there are also considerable sequence homologies. Based on the sequence of the α-subunits, the G-proteins may be divided into *four families* (Hepler and Gilman, 1992). These families are summarized in Table 5.1, together with representative members and their characteristic properties. Classification based on homology of amino acid sequences does not give any information about the functional properties of the different α-subunits.

### $G_s$ Subfamily

Members of the $G_s$ subfamily are activated by hormone receptors, by odor receptors and by taste receptors. $G_s$-proteins mediate, e.g., signal transmission by adrenaline receptors of type β and by glucagon receptors. During perception of taste, the taste receptors are activated, which then pass the signal on via the olfactory G-protein $G_{olf}$. Perception of „sweet“ taste is also mediated via a $G_s$-protein. Transmission of the

**Table 5.1.** Classification of the heterotrimeric G-proteins according to the α-subunits

| Subunit | tissue | Examples of receptors | Effector protein, function |
|---|---|---|---|
| **$G_s$** | | | |
| $\alpha_s$ | ubiquitous | ßAR, glucagon receptor | adenylyl cyclase ↑<br>$Ca^{2+}$ channels ↑ |
| $\alpha_{olf}$ | nasal epithelium | olfactory receptor | adenylyl-cyclase ↑ |
| **$G_i$** | | | |
| $\alpha_{i1}$ , $\alpha_{i2}$, $\alpha_{i3}$ | mostly ubiquitous | $\alpha_2$ adrenergic receptor | $K^+$ channels ↑<br>$Ca^{2+}$channels ↓ |
| $\alpha_{oA}$ | brain | $\alpha_2$ adrenergic receptor | $K^+$-channels ↑<br>$Ca^{2+}$-channels ↓ |
| $\alpha_{t1,}$transducin | retina | rhodopsin | cGMP specific phosphodiesterase ↑ |
| $\alpha_g$ | taste buds | | |
| $\alpha_z$ | brain | | adenylyl cyclase ↓ |
| **$G_q$** | | | |
| $\alpha_q$ | ubiquitous | $\alpha_1$ adrenergic receptor | phospholipase Cß ↑ |
| $\alpha_{11,}$ $\alpha_{14,}$ $\alpha_{15,}$ $\alpha_{16,}$ | | | |
| **$G_{12}$** | | | |
| $\alpha_{12,}$ $\alpha_{13,}$ | ubiquitous | | |

signal further involves an adenylyl cyclase in all cases, the activity of which is stimulated by the $G_s$-proteins. Characteristic for the members of the $G_s$ subfamily is that they are inhibited by cholera toxin (see 5.5.2).

### $G_i$ Subfamily

The first members of the $G_i$ subfamily to be discovered displayed an inhibitory effect on adenylyl cyclase, thus the name $G_i$, for inhibitory G-proteins. Further members of the $G_i$ subfamily have phospholipase C as the corresponding effector molecule. Signal transmission via phospholipase C flows into the inositol triphosphate and diacylglycerol pathways (see Chapter 6).

The $G_t$- and $G_g$-proteins are also classed as $G_i$-proteins, based on sequence homologies. The $G_t$- and $G_g$-proteins are involved in transmitting sensory signals. Signal transmission in the vision process is mediated via G-proteins known as *transducins* ($G_t$). The $G_t$-proteins are activated by the photoreceptor rhodopsin and are located in the rods and cones of the retina. The sequential effector molecules of the $G_t$-proteins are cGMP-specific phosphdiesterases (see Fig. 17.9).

Perception of bitter taste can take place via G-proteins; the α-subunit of these G-proteins is know as *gustducin* and is highly homologous with transducin. The corresponding receptors are just beginning to be characterized (Hohn, 1999). A phosphodiesterase with specificity for cyclic nucleotides and a cyclic nucleotide-gated ion channel

have been implicated as downstream components of the signaling cascade. Signal transmission evidently takes place here in a similar way to the vision process.

Apart from a few exceptions ($G_z$), the members of the $G_i$ family are characterized by inhibition by pertussis toxin (see 5.5.2).

#### $G_q$ Subfamily

The members of the $G_q$ subfamily are not modifiable by pertussis toxin or cholera toxin. The signal protein next in the reaction sequence is generally the β-type of phospholipase C.

#### $G_{12}$ Subfamily

Activation by thromboxane and thrombin receptors has been described for the $G_{12}$ subfamily. The effector molecules are not known at present.

The many α-, β- and γ-subunits now known present a large number of possibilities for subunit combination in heterotrimers. With the subunits identified at present, close to 1000 different heterotrimers could theoretically be formed. Although only a fraction of these exist in the cell, this nevertheless emphasizes the enormous complexity and diversity of regulation by G-proteins.

### 5.5.2 Toxins as Tools in Characterization of Heterotrimeric G-proteins

Two bacterial toxins, namely pertussis toxin and cholera toxin, were of great importance in determining the function of G-proteins. Both toxins catalyze ADP ribosylation of proteins. During ADP ribosylation, an ADP-ribose residue is transferred from $NAD^+$ to an amino acid residue of a substrate protein (Fig. 5.15).

Cholera toxin catalyzes the ADP-ribosylation of an arginine residue (Arg174 in $G_{\alpha t}$, Arg201 in $G_{\alpha s}$) in various α-subunits. The Arg174 residue of $G_{\alpha t}$ contacts the phosphate group of the bound GTP and is thus directly involved in GTP binding and possibly also in GTP hydrolysis. Modification of Arg174 by ADP-ribosylation interferes with this function and inactivates the GTPase activity of the G-protein. Consequently, the intrinsic deactivation mechanism of the $G_s$-protein is suspended. The G-protein is constitutively activated; the downstream effector molecules are — without any hormonal stimulation — permanently activated.

Constitutive activation of $G_s$-proteins by cholera toxin is the cause of the devastating effect of the cholera bacterium, *Vibrio cholerae*, on the water content of the intestine. Due to the lack of deactivation of the $G_s$-protein, adenylyl cyclase next in the reaction sequence is constantly activated, so that the level of cAMP in the cells of the intestinal epithelium is greatly increased. This, in turn, leads to increased active transport of ions and an excessive efflux of water and $Na^+$ takes place in the intestine.

Pertussis toxin, formed by *Bortedella pertussis*, the causative organism of whooping cough, carries out an ADP-ribosylation at a cysteine residue close to the C-terminus of

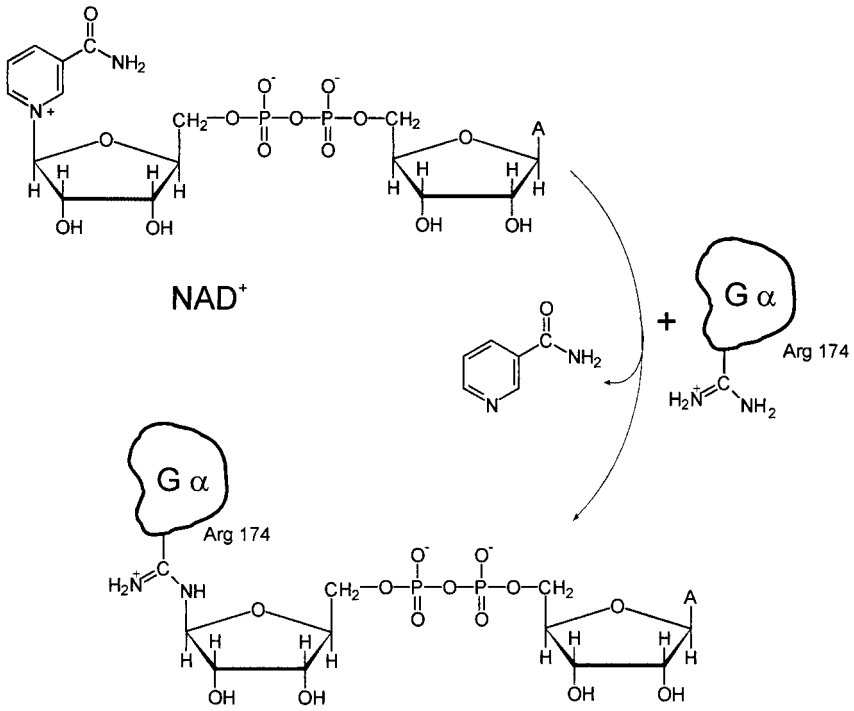



**Fig. 5.15.** ADP-ribosylation of the Gα-subunit of transducin by cholera toxin. Cholera toxin catalyzes the ADP-ribosylation of the α-subunit of the G-protein transducin. During the reaction, the ADP-ribose residue of $NAD^+$ is transferred to Arg174 of $G\alpha_t$, which inactivates the GTPase activity of $G\alpha_t$.

α-subunits. The modification prevents activation of the G-protein by the receptor, whereby the signal transmission is blocked.

## 5.5.3 The Functional Cycle of Heterotrimeric G-Proteins

Signal transmission via G-proteins takes place in close association with the inner side of the cell membrane. Both the α-subunit and the βγ-complex are associated with the membrane via membrane anchors (see 5.5.8).

Like all regulatory GTPases, the heterotrimeric G-proteins run through a cyclical transition between an inactive, GDP-bound form and an active, GTP-bound form. Fig. 5.16 sketches the different functional states and the role of the individual subunits.

### Inactive Ground State

In the inactive ground state, the G-proteins exist as $G\alpha \cdot GDP \cdot (\beta\gamma)$-heterotrimers. The receptor is not occupied by the hormone; the sequential effector molecule is inactive.

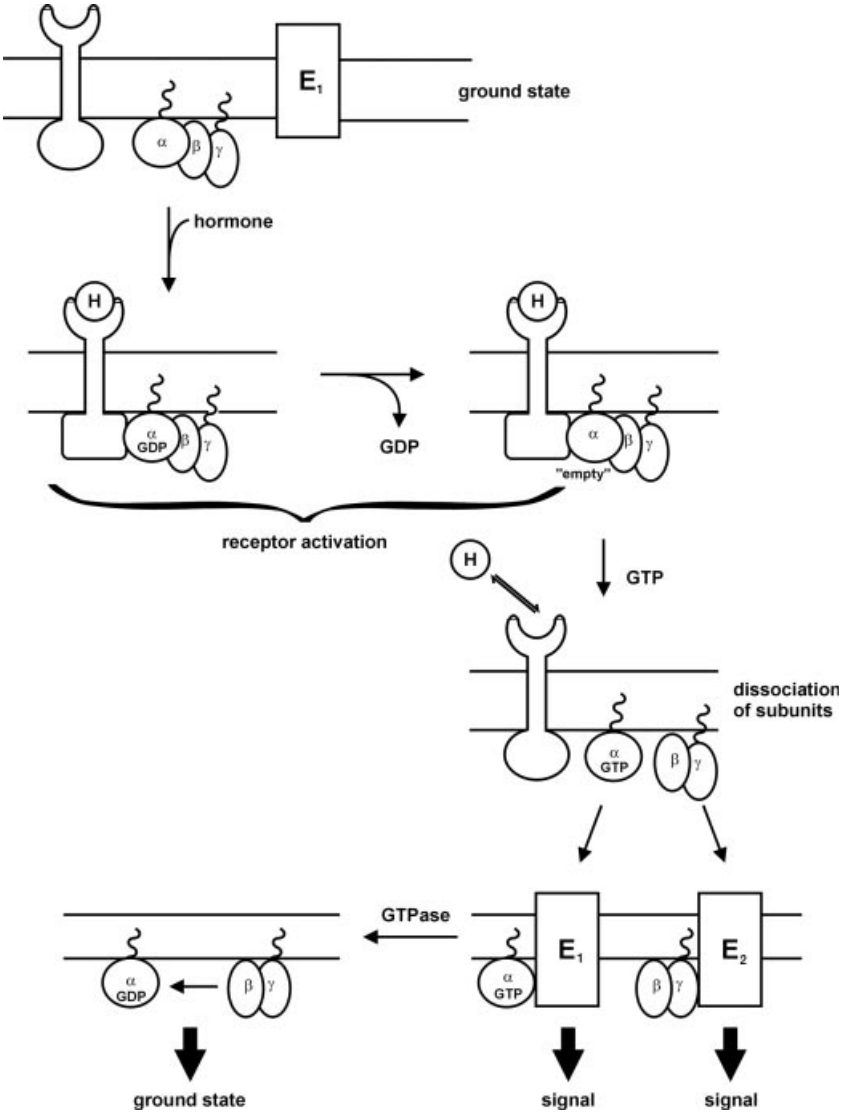



**Fig. 5.16.** Functional cycle of the heterotrimeric G-proteins. a) The G-proteins exist in the ground state as a heterotrimeric complex $(G_\alpha \cdot GDP) \cdot (\beta\gamma)$. b) The activated receptor binds to the inactive heterotrimeric complex of the G-protein and leads to dissociation of the bound GDP and the $\beta\gamma$-complex. c) Binding of GTP to the „empty" $G_\alpha$-subunit transforms the latter into the active $G_\alpha \cdot GTP$ state. $G_\alpha \cdot GTP$ interacts with an effector molecule in the sequence E1 and activates the latter for further signal transmission. The released $\beta\gamma$-complex may also take part in signal conduction by binding to a corresponding effector molecule E2 and activating the latter for further signal conduction. d) Hydrolysis of the bound GTP terminates the signal transduction via the $\alpha$-subunit.

### Activation

Binding of extracellular signal molecules (hormones, neurotransmitters) to the receptor initiates activation of the G-protein. The activated receptor associates with the α-subunit of the heterotrimeric complex $G_\alpha \cdot GDP \cdot (\beta\gamma)$ and induces a conformational change that leads to dissociation of GDP. It is assumed that cytoplasmic structural elements of the receptor interact with the C-terminal tail of $G_\alpha$ and induce an extensive conformational change that propagates to the nucleotide binding site and leads to dissociation of GDP. The heterotrimer is now in an „empty" state in which it possesses high affinity for the activated receptor. The free nucleotide binding site is immediately occupied by GTP since GTP exists in a large excess compared to GDP in the cell and because the $G_\alpha$-subunit binds GTP more strongly than GDP.

### GTP binding

GTP binding has two consequences: firstly, the βγ-complex dissociates and secondly, the binding to the activated receptor is cancelled. The free α-subunit with bound GTP represents the activated $G_\alpha \cdot GTP$ form of the G-protein and transmits the signal further. The receptor released from the complex can activate other G-proteins, enabling amplification of the signal. For a detailed model of Gα activation see Iiri et al., (1998).

### Transmission of the Signal

The interaction of $G_\alpha \cdot GTP$ with the corresponding effector molecule leads to inactivation of the former and thus to initiation of the next step in the signal transmission chain. The βγ-complex released during activation can also perform a signal-mediating function (see 5.5.7).

### Termination of the Signal

Hydrolysis of GTP by the intrinsic GTPase activity of the α-subunit ends signal transmission at the level of the G-proteins. The rate of GTP hydrolysis functions as an inner clock for signal transmission; it determines the lifetime of the activated state and the extent of the reactions next in sequence. At this point, regulatory mechanisms may take effect in which proteins with the character of GTPase-activating proteins (GAPs) specifically stimulate the GTPase activity of the α-subunit. Proteins known as *regulators of G-protein signaling (RGS)* can increase the GTPase activity by more than one order of magnitude (see 5.5.9).

The rate of GTP hydrolysis may also be increased via the downstream effector molecule. Phospholipase C-β1 stimulates the intrinsic GTPase activity of the corresponding $G_{q\text{-}11}$ by close to two orders of magnitude (Bernstein et al., 1992). A further effector molecule, adenylyl cylase, has been shown to function as a GAP for the monomeric $G_\alpha$-GTP state (Scholich et al., 1999).

### 5.5.4 Mechanistic Aspects of the Switch Function of G-Proteins

G-proteins are enzymes that can exist in different conformations, and that can undergo regulatory interactions with different partners. For a detailed understanding of the different functions of G-proteins, it is essential to comprehend the structural and mechanistic basis of their reactions. To this end, we need to answer the following question:

Which mechanism is the basis of GTPase activity and which residues of the α-subunit are involved in catalysis?

What is the structural difference between the active and inactive state of the Gα-subunit?

Answers to these questions are currently founded on the highly resolved structure of transducin $G_{\alpha t}$, a $G_{i,\alpha}$-subunit and on many biochemical experiments, in particular with mutated proteins. Furthermore, comparisons with the highly resolved structures of other regulatory GTPases (Ras protein, bacterial elongation factor EF-Tu) have helped to identify common mechanisms and structural principles for GTPases, even though the different GTPases have significant structural differences in subdomains, due to different effector molecules.

### 5.5.5 Mechanism of GTP Hydrolysis

The $G_{\alpha}$-subunit possesses a slow GTPase activity, which is synonymous with a long lifetime of the activated GTP state. For transducin $G_{\alpha t}$, the $k_{cat}$ value for GTP hydrolysis is in the region of $0.05 sec^{-1}$. This value is, however, close to two orders of magnitude higher than that of the Ras protein (see Chapter 9). The $k_{cat}$ for GTP hydrolysis may be increased by up to two orders of magnitude by GTPase activating proteins (GAP, see below) for some $G_{\alpha}$-proteins. A much larger increase, namely $10^5$-fold, is observed for the GAP belonging to the Ras protein (see Chapter 9).

It is generally assumed that hydrolysis of the γ-phosphate bond proceeds via an $S_N2$ mechanism, as shown in Fig. 5.17. The hydrolysis proceeds by an „in-line“ attack of a water molecule on the γ-phosphate, in which the GDP residue is displaced from the γ-phosphate.

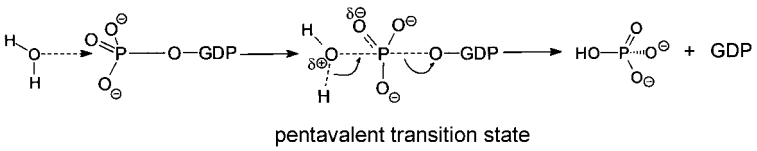



**Fig. 5.17.** „In-line“ attack in GTP hydrolysis. Hydrolysis of GTP takes place via an „in-line“ attack of a water molecule at the γ-phosphate. The reaction passes through a pentavalent transition state in which the ligands of the γ-phosphate adopt a trigonal bipyramidal configuration. The mechanism by which the water molecule is activated for the attack on the γ-phosphate is not shown in the figure. Possible mechanisms are presented in Fig. 5.18.

In the postulated transition state, the γ-phosphorus atom is penta-coordinated, whereby the ligands are configured in the form of a triagonal bipyramid. $Mg^{2+}$ is indispensable for the catalysis; it is needed for binding of substrate and product, as well as for the catalysis itself. Activation of the water molecule for nucleophilic attack at the γ-phosphate requires involvement of side groups of the protein in the sense of a general base catalysis.

For the structural determination of the activated form of the α-subunit, use of $AlF_4^-$ as a ligand was of great importance. $AlF_4^-$ is an activator of GDP-bound α-subunits and due to this characteristic — in addition to the bacterial toxins mentioned above — is often used for detection of G-proteins and for their structural characterization. In the presence of $AlF_4^-$, permanent activation of the G-protein is observed: $G_\alpha \cdot GDP$ is fixed by binding of $AlF_4^-$ in a conformation that permits activation of the effector molecule.

Comparison of the structure of the active conformation of $G_{i\alpha} \cdot GDP \cdot AlF_4^-$ and $G_{i\alpha} \cdot GTP\gamma S$ has brought important insights into the mechanism of GTP hydrolysis and the function of the catalytic residues of the α-subunit (Coleman et al., 1994). Using the same experimental method, it was also possible to postulate a structure of the transition state of GTP hydrolysis for the α-subunit of the transducin, $G_{t,\alpha}$ (Sondek et al., 1994). The same result was obtained for both systems. The structure of the $G_{i\alpha} \cdot GDP \cdot AlF_4^-$ complex shows that the $AlF_4^-$ is bound at the position of the γ-phosphate of $G_{i\alpha}$*GTPgS, whereby there are noticeable differences for two residues of the protein, namely Gln204 and Arg178, between the GTP structure and the $GDP \cdot AlF_4^-$ structure. In the $G_{i\alpha} \cdot GDP \cdot AlF_4^-$structure, Gln204 and Arg178 adopt a position that is very well suited to stabilize the postulated transition state and to nucleophilically activate the hydrolytic water molecule. The X-ray structure indicates that $GDP \cdot AlF_4^-$ has the role of a *transition state analog*, in which the $AlF_4^-$ adopts the position of the γ-phosphate in the supposed transition state of GTP hydrolysis (Fig. 5.18a). It is assumed that the transition state is stabilized by interaction of Arg178 and Gln204 with the oxygen atom of the γ-phosphate and that Gln204 helps to activate the hydrolytic water molecule (Fig. 5.18b). In the ground state of the reaction, represented by the GTPγS structure, both residues do not interact with the hydrolytic water molecule and the γ-phosphate. In the case of $G_{t,\alpha}$, the same amino acids are involved in stabilization of the transition state. In Fig. 5.18c, an alternative mechanism for stabilization of the transition state is presented. According to this mechanism, the oxygen atom of the γ-phosphate has the function of a base that nucleophilically activates the water molecule. The model is supported by theoretical considerations and is increasingly accepted.

In comparison to the Ras protein (see Chapter 9) and bacterial EF-Tu, there are differences in the details of the residues involved. The general hydrolysis mechanism, as formulated in Fig. 5.17, also applies to these proteins, however.

In all, the α-subunits of the G-proteins possess a slow GTPase activity. A reduction of the lifetime of the activated $G_\alpha \cdot GTP$ state and thus weakening of the signal transmission, can be achieved by binding of specific GTPase activating proteins such as the RGS proteins (see 5.5.9) to $G_\alpha \cdot GTP$. The RGS proteins stimulate the GTPase activity of different α-subunits by close to two orders of magnitude. Mechanistically, the GTPase-activating activity of the RGS proteins is explained, in particular, by stabilization of the transition state. It is assumed that the RGS proteins fix the catalytic residue

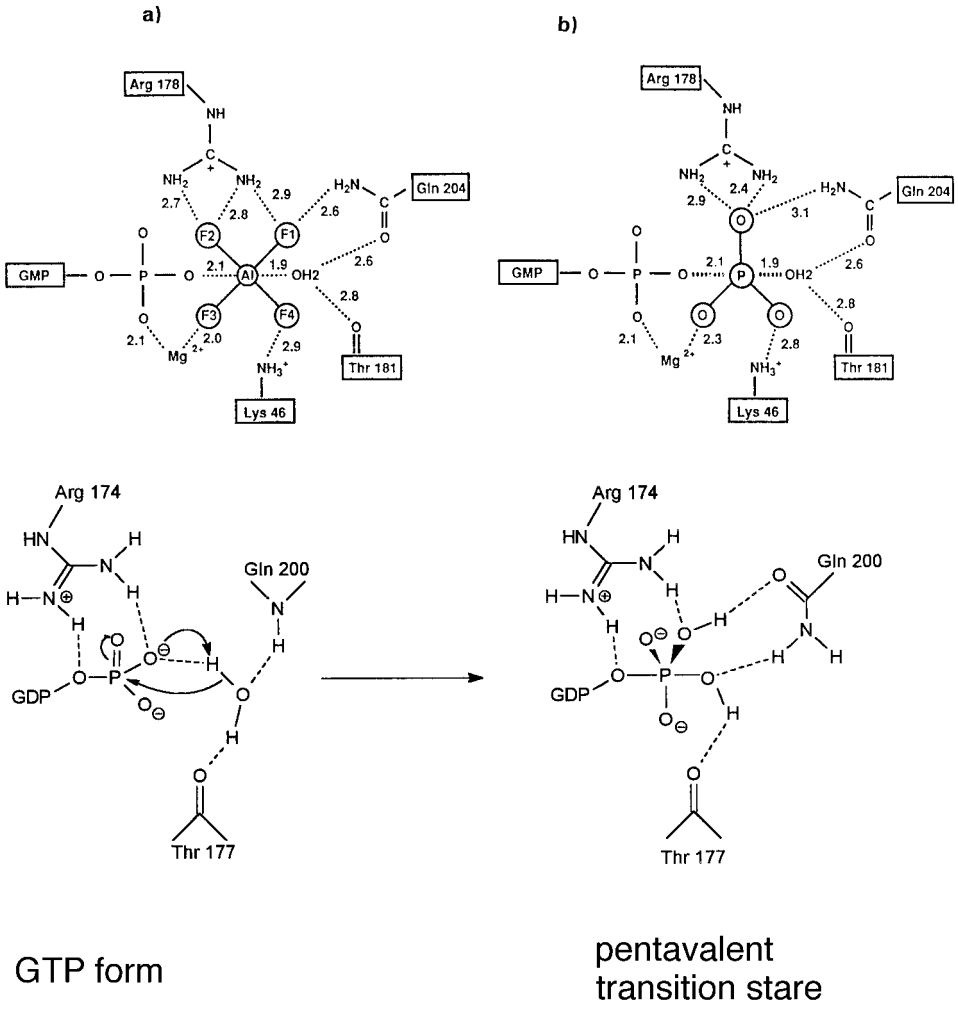



**Fig. 5.18.** Postulated transition state and mechanism of GTP hydrolysis. a) Binding of $GDP \cdot AlF_4^-$ at the active center of $G_{i\alpha}$. The representation is based on the structure of the $G_{i\alpha} \cdot GDP \cdot AlF_4^-$ complex. The coordination sphere and the distances from $AlF_4^-$ (in Å) to Arg178, Gln204, the $Mg^{2+}$ ion and other residues of the GTPase center are shown. The catalytically important residues Gln204 and Thr181 fix a water molecule that is located „in line" to an oxygen atom of the γ-phosphate of GDP. b) Model of the transition state of GTP hydrolysis at the active center of $G_{i\alpha}$. The structure is modeled on the structure of the $G_{i\alpha} \cdot GDP \cdot AlF_4^-$complex. The P atom of an assumed trigonal „$PO_3$ group" (γ-phosphate of GTP) was projected onto the aluminum atom of the $AlF_4^-$ grouping and the levels of the „$PO_3$ group" were placed on the level of the $AlF_4^-$. The water molecule fixed by Gln204 and Thr181 is located at a favorable distance for a nucleophilic attack on the γ-phosphate. c) Mechanism of GTP hydrolysis involving the γ-phosphate as an activating base. A postulated mechanism of GTP hydrolysis is shown for the α-subunit of transducin, $G_{\alpha,t}$, in which the O atom of the γ-phosphate prepares the catalytic base for activation of the water molecule. The water molecule is precisely positioned and fixed by Gln200 (corresponds to Gln204 of $G_{i\alpha}$) and Thr 177 (corresponds to Thr181 of $G_{i\alpha}$).

of the GTPase center and bring it into a position favorable for the hydrolysis. GTPase stimulation of the Ras protein by the corresponding GAP proteins proceeds, in contrast, by another mechanism (see 9.2.2).

### 5.5.6 Structural Basis of the Activation of the α-Subunit

The switch function of the α-subunit of the heterotrimeric G-proteins is founded on the change between an active $G_\alpha$-GTP confirmation and an inactive $G_\alpha$-GDP conformation. The structural difference between the two conformations was explained for the transducin, $G_{t,\alpha}$, by crystallization and structural characterization of the inactive GDP form and the active GTPγS form (Lambright et al., 1994). The structures of both forms of $G_{t,\alpha}$ are shown in Fig. 5.19.

$G_{t,\alpha}$ is made up of two domains, a GTPase domain and a helical domain. The *GTPase* or *G-domain* indicates that $G_{t,\alpha}$ is a member of the superfamily of regulatory GTPases. In addition, $G_{t,\alpha}$ possesses a *helical domain*, which represents a characteristic feature of the heterotrimeric G-proteins. The nucleotide binding site is in a cleft between the two domains. It is assumed that the presence of the helical domain is the reason that bound nucleotide dissociates only very slowly from transducin and that the activated receptor is therefore necessary to initiate the GDP/GTP exchange.

The γ-phosphate group of GTP must be assigned the function of a trigger of activation of $G_{t,\alpha}$. The comparison of the active and inactive conformations gives an insight into this function. In all, the active and inactive forms of $G_{t,\alpha}$ have a very similar structure. Significant conformational changes on transition between the two functional states were found for three structural elements, known as switch I, II and III, that include only 14 % of the amino acids of transducin. The γ-phosphate interacts with three amino acids that move switch I upwards and thus cause a coupled movement of switches II and III (Fig. 5.19).

The GTP binding to $G_{t,\alpha}$ has several consequences:

Firstly, it is assumed that the conformational changes in switch II triggered by GTP binding lead to dissociation of the βγ-complex. The βγ-complex binds to the switch regions I and II of the α-subunit.

A further consequence of the conformational change in the α-subunit induced by the γ-phosphate is the activation of the effector molecule next in sequence. The binding site of the sequential effector molecule adenylyl cyclase includes the switch II (Tesmer et al., 1997). It is therefore assumed that the conformational change of switch II also mediates the binding and activation of the effector molecule. The binding site for the effector and for the βγ-complex partially overlap, so that a binding of the effector is only possible if the βγ-complex has dissociated.

The question of the structural basis and explanation of the dissociation of the activated receptor from the βγ-complex is little understood. Equally open is the question of the structure of the nucleotide-free heterotrimeric state, postulated as a short-lived intermediate form in the activation cycle of the G-protein.

Highly resolved structures are also available for the $G_\alpha \cdot GDP \cdot \beta\gamma$-complex (Wall et al., 1995; Lambright et al., 1996). From these structures, the βγ-complex binds in the region of the switch regions I and II and in the region of the N-terminus of the α-sub-

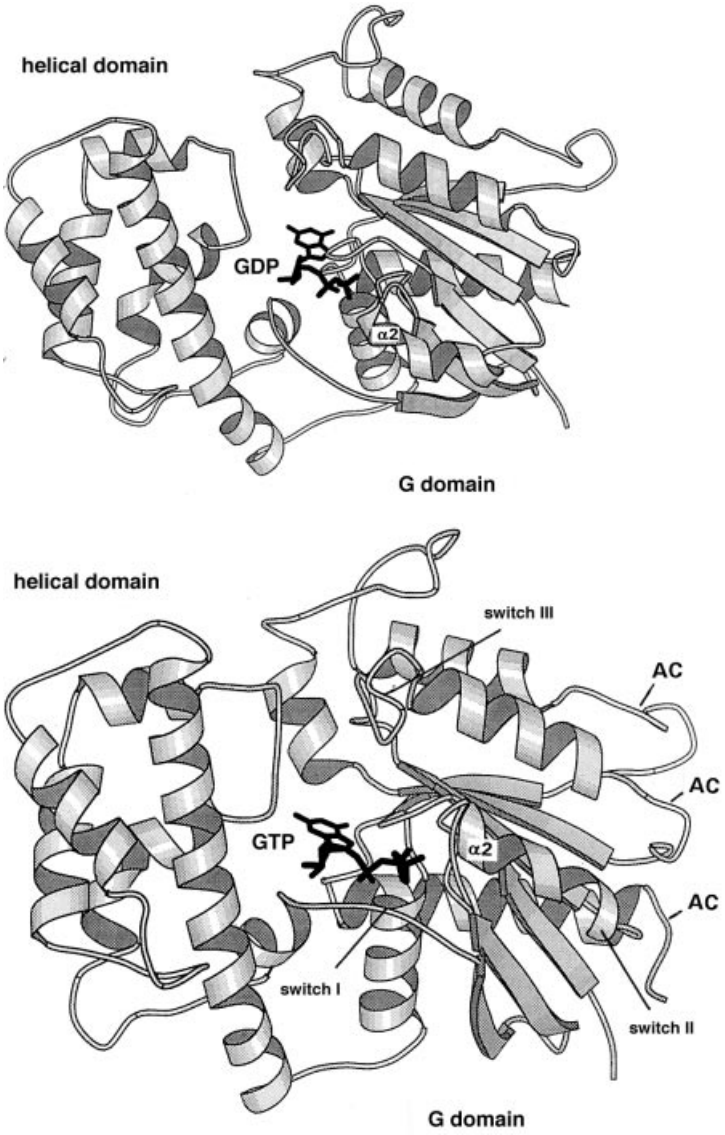



**Fig. 5.19.** GTP and GDP structures of transducin. The $G\alpha_{,t}$ subunit of transducin possesses—in contrast to Ras protein and to other small regulatory GTPases—an α-helical domain that hides and closes the G-nucleotide binding pocket. The conformational changes that accompany the transition from the inactive $G_{\alpha,t} \cdot GDP$ form (a) into the active $G_{\alpha,t} \cdot GTP$ form (b), are restricted to three structural sections that are known as switches I, II and III. Switch I includes the link of the α-helical domain with β2, switch II affects in particular helix α2, and switch III, the β3—α3 loop. Switch III includes a sequence that is characteristic for the α-subunits of the heterotrimeric G-proteins. The conformational changes of switches II and III affect structural sections that are assumed to be binding sites for the effector molecule adenylyl cyclase (AC) and the γ-subunit of cGMP-dependent phosphodiesterase (PDEγ), based on mutation experiments and biochemical investigations. MOLSKRIPT representation according to Kraulis, (1991).

unit. The structure of the βγ-complex has an interesting configuration of 7 β-sheet structures for the β-subunit. These are in the form of a propellor with 7 configured leaves. The γ-subunit is located at the side of the β-subunit and does not itself interact with the α-subunit.

As a consequence of binding of the βγ-complex, significant conformational changes are observed in the α-subunit. These affect, in particular, the switch regions I and II and the N-terminal α-helix.

The βγ-complex does not show any great structural differences in the free and $G_\alpha$-bound forms. Activation of the βγ-complex for the interaction with the corresponding effector molecule (see below) appears to be based only on its release from the inactive $G\alpha \cdot GDP \cdot \beta\gamma$ complex. The $G_\alpha$-subunit has the function of a negative regulator here, that inactivates the βγ-complex by masking the interaction region for signal proteins next in the sequence.

For the Ras protein and the EF-Tu from *Thermus thermophilus*, highly resolved structures of the active GTP form and the inactive GDP form are also available. The analogous switch elements I and II are also present in Ras protein (Fig.9.3) and EF-Tu (Fig. 5.12), even though little sequence homology is visible between the analogous structural elements. The switch element III is, in contrast, characteristic for the heterotrimeric G-proteins. On activation of Ras protein, the conformational changes are limited to the structural elements involved in binding the γ-phosphate (Pai et al., 1990). For EF-Tu, in contrast, the nucleotide exchange leads to a greater structural rearrangement (Berchthold et al., 1993). The conformational changes associated with activation thus differ significantly for these three GTPases.

### 5.5.7 Function of the βγ-Complex

Originally, it was assumed that the βγ-complex only played a passive role in the functional cycle of the G-proteins. It soon became apparent, however, that the βγ-complex, in addition to binding to the α-subunit, also carries out other functions and interacts specifically with corresponding effector molecules (review article: Neer, 1995). The βγ-complex must be assigned its *own regulatory function*; it takes part itself in the propagation and termination of signal transmission.

The first evidence of the special function of the βγ-complex was obtained for adenylyl cyclase. Some of the adenylyl cyclases occurring in the brain are inhibited by the βγ-complex (see below); others, in contrast, are stimulated by the βγ-complex, whereby stimulation only takes place in the presence of the α-subunit.

Other effector molecules of the βγ-complex are specific subtypes of phospholipase C, and $K^+$- and $Ca^{2+}$-specific ion channels. In the case of $Ca^{2+}$ channels, a direct interaction between the βγ-complex and cytoplasmic loops of the α-subunit of the ion channel has been demonstrated (De Waard et al., 1997). Regulation of the activity of ion channels is thus a further important role of the βγ-complex.

Mutational analysis has revealed the structural elements of the βγ-complex that interact with the effectors (Ford et al., 1998). The interaction sites with different effectors overlap partially and are only available for binding in the absence of the α-sub-

unit. This organization of the interaction regions allows the key regulator $G_\alpha$ to control $G_{\beta\lambda}$signal transmission to multiple effectors.

The interaction of the βγ-complex with G-protein coupled receptor kinases (see 5.3.4, β-adrenergic receptor kinase, βARK) appears to be of special regulatory importance. The function of the βγ-complex in this system is shown in Fig. 5.9. The βγ-complex binds specifically to the βARK and translocates this to the cell membrane. The translocation of βARK is necessary to switch off and modulate signal transmission via adrenaline.

The structural element of βARK that interacts specifically with the βγ-complex is localized in the C-terminal third of the βARK sequnce (Inglese et al., 1994). It possesses the characteristics of an independently folding protein domain and is ranked with the *pleckstrin homology domains* (PH domains). The PH domains are protein modules (see Chapter 8) , found in many proteins, that by binding of inositol lipids (see Chapter 6) mediate protein-membrane interactions.

### 5.5.8 Membrane Association of the G-Proteins

Signal transmission via G-proteins is inseparably linked with their membrane association. The preceding reaction partners are transmembrane proteins and the subsquent effector molecules, such as adenylyl cyclase, are either also transmembrane proteins or they are associated with the membrane (Fig. 5.20).

The membrane association of the G-proteins is mediated by membrane anchors that are introduced in the course of a posttranslational modification at the N-terminus of the α-subunit and at the C-terminus of the γ-subunit (cf. Section 3.7).

The α-subunits of $G_i$ and $G_o$ subtypes possess a lipid anchor in the form of a myristoylation at the N-terminal glycine residue. The γ-subunits have a membrane anchor in the form of prenyl residues, in a similar way to Ras protein. In addition, the terminal carboxyl group is esterified with a methyl group which further increases the hydrophobicity of the C-terminus. The length of the appended isoprenoid grouping is variable. Whilst the γ-subunit of the $G_{t,\gamma}$ protein has a farnesyl chain encompassing 15 C atoms, a modification with a $C_{20}$ geranyl-geranyl subunit is to be found in γ-subunits of $G_o$-proteins in the brain.

### 5.5.9 Regulators of G-Proteins: Phosducin and RGS Proteins

Signal transmission via G-proteins and the corresponding receptors is subject to tissue- and cell-specific regulation at different levels. The regulation is mostly of a negative, suppressing character and serves two purposes in particular: Firstly, the cell must try to weaken the cytoplasmic answer under conditions of persistent activation of the receptor. Secondly, the cell needs mechanisms to rapidly terminate the signal. Typically, the rate of GTP hydrolysis of the α-subunit is very slow, about 4 $min^{-1}$. The cell must be able to shorten the associated long lifetime of the activated state in a regulatable way.

The most important regulatory attack points at the level of the G-proteins and their receptors are:

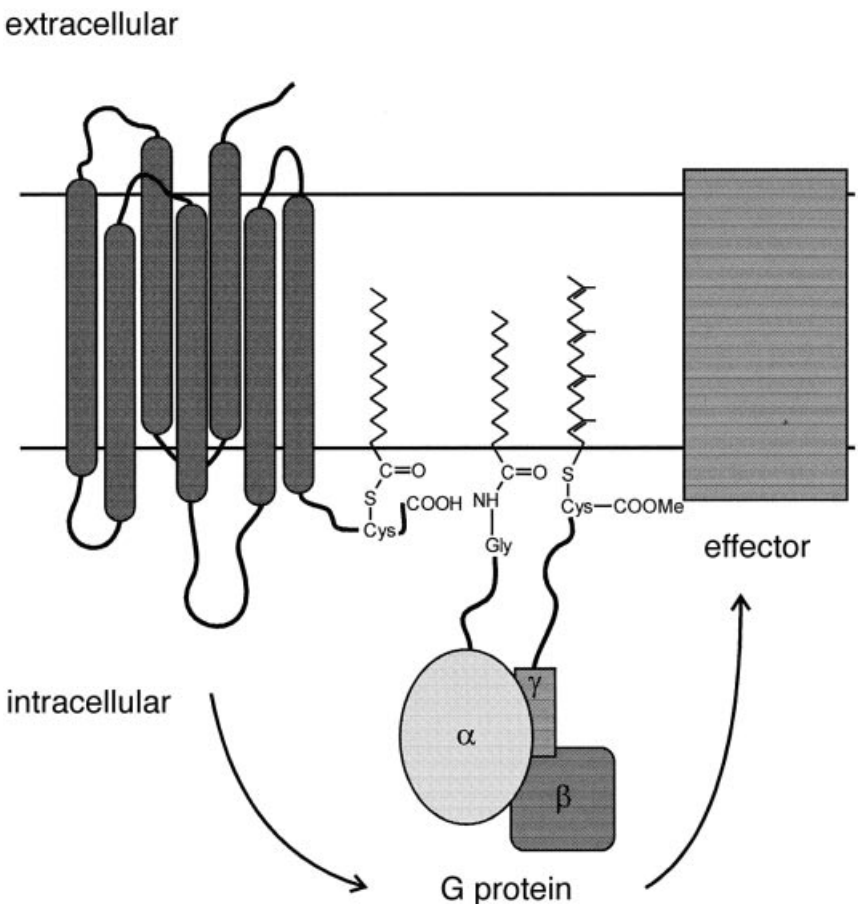



**Fig. 5.20.** Membrane anchor of the heterotrimeric G-proteins. The lipid anchoring in the system of G-protein-coupled receptors and the corresponding G-proteins is shown. In the figure, it is assumed that the lipid anchors are located in the membrane. A possible involvement of the lipid anchor in protein-protein interactions is not shown. The G-protein-coupled receptor carries a palmitoic acid anchor at the C-terminus. The α-subunit of the heterotrimeric G-protein is associated with the membrane via a myristoic acid anchor at the N-terminus, whilst the γ-subunit of the βγ-complex uses a prenyl residue as a membrane anchor.

- Desensitization: phosphorylation of the receptor on the cytoplasmic side (see 5.3.4) as a reaction to persistent stimulation. This is a long-term adaptation.
- Downregulation of the number of receptor molecules: regulation at the levels of expression, stability and internalization of the receptor.
- Inactivation of the βγ-complex: binding of phosducin to the βγ-complex.
- Reduction of the lifetime of the $G_\alpha \cdot GTP$ complex: activation of the GTPase of $G_\alpha \cdot GTP$ by RGS proteins

At the level of the G-proteins, negative regulation by phosducin or RGS proteins stands out in particular.

**Regulation by phosducin**

Phosducin is a protein that binds in a specific manner to the βγ-complex and inhibits the signal-giving function of the latter. Phosducin is assigned an important role as a negative regulator, in the sense of long-term adaptation in signal transmission during the process of vision.

It is likely that phosducins play a role in many G-protein coupled signal transduction pathways. Phosducin-like proteins have been identified in a variety of tissues, e.g., in brain and in the pineal gland.

Binding of phosducin to the βγ-complex leads to its translocation from the membrane into the cytosol. In this way, the number of βγ-complexes available for the G-protein cycle is reduced and signal transmission is weakened. Interestingly, the phosducin function is subject to regulation by phosphorylation. In the Ser-phosphorylated form, binding to the βγ-complex is greatly weakened.

**Regulation by RGS Proteins**

The RGS proteins have the function of GTPase activating proteins (review. Berman & Gilman 1998). They bind specifically to α-subunits and activate their GTPase rate by close to two orders of magnitude. At present, at least 19 different genes are known in mammals for RGS proteins. Specific assignment to particular α-subunits is to be assumed, whereby most of the known RGS proteins act as GAPs towards members of the $G_i$ subfamily. For the most part, it is still unclear which factors regulate the activity of RGS proteins and how exactly they are integrated into G-protein signaling pathways. Possibly they are also involved in crosstalk between signaling pathways.

## 5.6 Effector Molecules of G-Proteins

Activated G-proteins pass the signal on to subsequent effector molecules that have enzyme activity or function as ion channels (see Fig. 5.14). Important effector molecules are adenylyl cyclase, phospolipases, and cGMP-specific phosphodiesterases. The activation of these enzymes leads to concentration changes of diffusable signal molecules, such as cAMP, cGMP, diacylglycerol or inositol triphosphate, and $Ca^{2+}$, that trigger further specific reactions (see Chapter 6 and Fig. 6.1). G-protein-mediated opening of ion channels may lead to changes in membrane potential and to changes in the ion environment (cf. Chapter 16), whereby the changes in $Ca^{2+}$ concentration are of particular importance.

### 5.6.1 Adenylyl Cyclase and cAMP as „Second Messenger“

The adenylyl cyclases catalyze the formation of 3’-5’-cyclic AMP (cAMP) from ATP (Fig. 5.21). cAMP is a widespread signal molecule that primarily functions via activa-

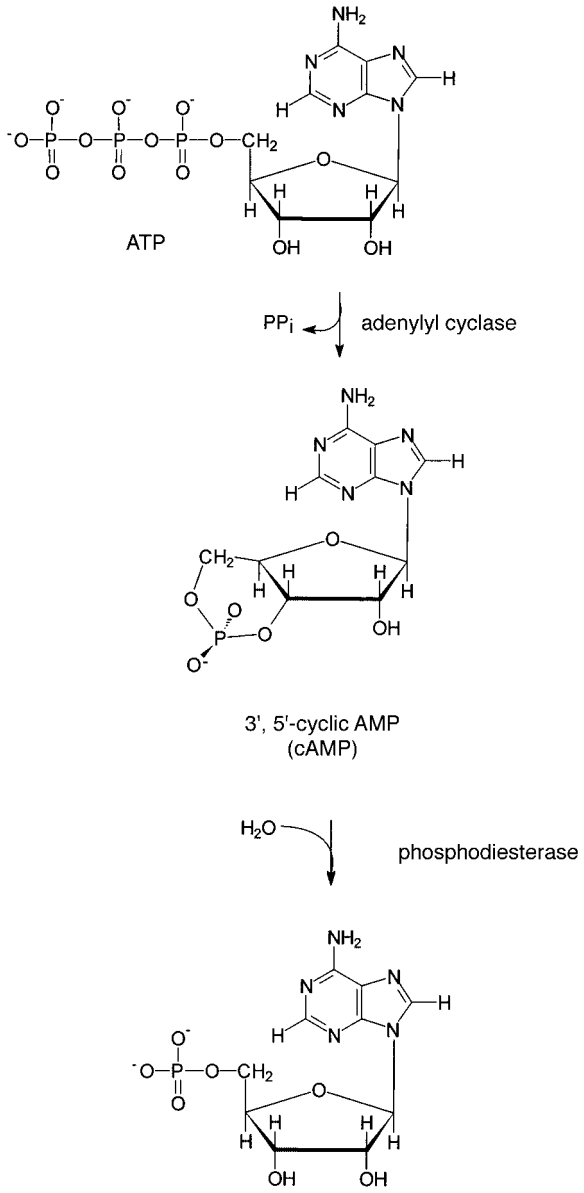



**Fig. 5.21.** Formation and degradation of cAMP.

tion of protein kinases (see 6.1 and Chapter 7). Synthesis of cAMP by adenylyl cyclase is opposed by degradation and inactivation by phosphodiesterases.

### Structure of Adenylyl Cyclase

Despite the central importance of adenylyl cyclase for hormonal signal transduction, its structural and functional characterization is incomplete. In mammals, at least 9 dif-

ferent types of adenylyl cyclase are described; these are known as adenylyl cyclases of type I—IX and show a high degree of sequence homology (ca. 50 %) (review: Tang and Gilman, 1992; Taussig and Gilman, 1995).

The adenylyl cyclases are large transmembrane proteins with a complex transmembrane topology. The assumed topology (Fig. 5.22) shows a short cytoplasmic N-terminal section followed by a transmembrane domain M1 with six transmembrane sections, and a large cytoplasmic domain C1. The structural motif is repeated so that a second transmembrane domain M2 and a second cytoplasmic domain C2 can be differentiated. The complicated structure resembles the structure of some ATP-dependent membrane transport systems such as the P glycoprotein. A transport function has not yet been demonstrated for adenylyl cyclase.

Information on the structure-function relationship of adenylyl cyclase is available, in particular for the cytoplasmic domain. According to this, the important functions of adenylyl cyclase, namely the interaction with the G-protein and the synthesis of cAMP, are localized on the cytoplasmic C1 and C2 domains. The C1 and C2 domains are homologous to a high degree between the different subtypes; the transmembrane domain, in contrast, is little conserved. Structural determination of the complex of $G_{s\alpha}$, $\cdot$ GTP and a C1-C2 dimer indicates that the active center is at the interface of the C1-C2 dimer (Tesmer et al., 1997; Zhang et al., 1997). The ATP binding site and a binding site for the activator forskolin are located there. The binding site for the regulator $G_{s\alpha} \cdot$ GTP is relatively far away from the catalytic center. It is assumed that, during signal transmission by the G-protein, an extensive conformational change is initiated that leads to a productive reorganization of the catalytic center at the C1-C2 interface.

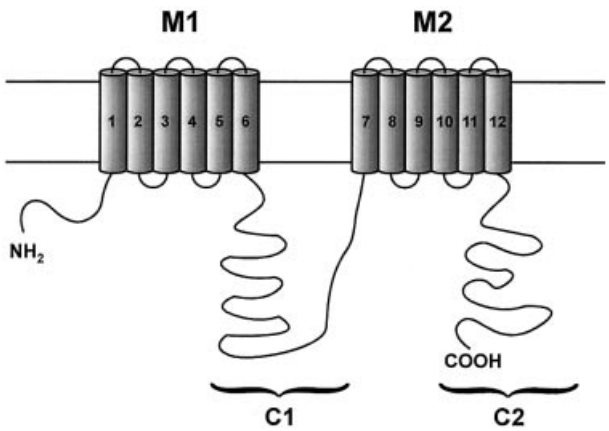



**Fig. 5.22.** Topology of adenylyl cyclase. The adenylyl cyclase of mammals is a transmembrane protein. It is composed of two homologous domains, which each have a transmembrane domain (M1 and M2) and a larger cytoplasmic portion (C1 and C2). Sequence analysis predicts 6 transmembrane helices for each of the domains (numbering from 1–12). The active site is formed by residues from C1 and C2.

## Regulation of Adenylyl Cyclase

A common feature of the different adenylyl cyclases is the stimulation of their enzyme activity by the GTP-bound form of the α-subunit of the $G_s$-protein. Furthermore, all subtypes are stimulated by the diterpene *forskolin*. In addition to the central regulation by the activated α-subunit, there are other stimulatory or inhibitory influences on the different subtypes of adenylyl cyclase, in a manner characteristic for the particular subtype. The various subtypes differ in these regulatory influences, whereby none of the different subtypes have an identical pattern of regulation.

Fig. 5.23 summarizes the stimulatory and inhibitory influences that take effect on the various adenylyl cyclases.

Stimulation of adenylyl cyclase may take place by:

- $G_{s,\alpha} \cdot GTP$
- $Ca^{2+}$/calmodulin
- Protein kinase C
- βγ-subunits of G-proteins

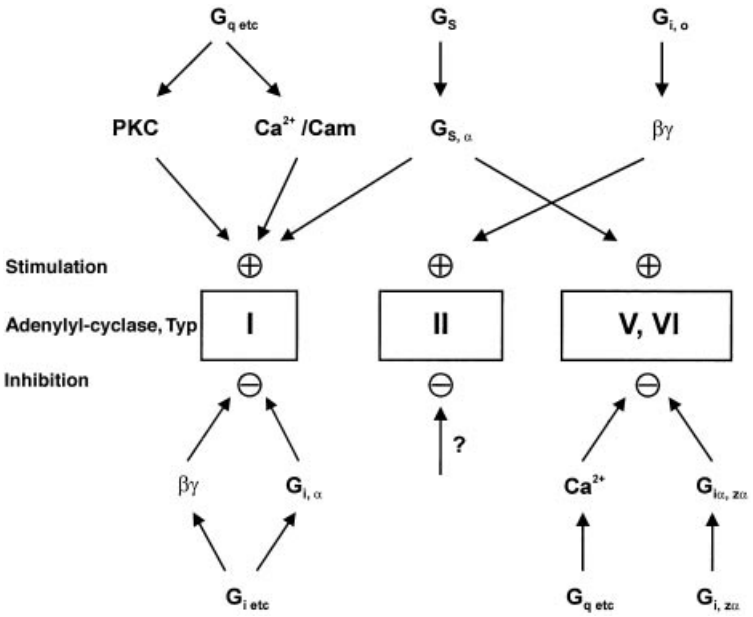



**Fig. 5.23.** Diversity of regulation of adenylyl cyclase. The figure summarizes schematically the regulation of adenylyl cyclases of type I, II, V and VI (after Taussig and Gilman, 1995). The individual subtypes are negatively (-) or positively (+) regulated by heterotrimeric G-proteins by various pathways. Regulation takes place both via $G_\alpha$-subunits and via βγ-complexes. For nomenclature of the G-proteins, see 5.5.1; PKC: protein kinase C; $Ca^{2+}$/CaM: $Ca^{2+}$/calmodulin (see Chapters 6 and 7).

Inhibition of adenylyl cyclase is possible by:

- $G_{i,\alpha} \cdot GTP$
- $Ca^{2+}$
- βγ-subunits of G-proteins

The regulation of the subtypes I, III and VIII by $Ca^{2+}$/calmodulin stands out. All three subtypes are stimulated by $Ca^{2+}$, although in different concentration regions. $Ca^{2+}$ is, as discussed in more detail in Chapter 6, a central intracellular messenger substance, and an increase in the $Ca^{2+}$ concentration is observed on activation of different signal transduction processes.

In the brain, the $Ca^{2+}$/calmodulin regulation of adenylyl cyclase is of particular importance. One finds adenylyl cyclase concentrated in the vicinity of receptors for N-methyl-D-aspartate (see Chapter 16), that represent regulatable entry points for $Ca^{2+}$. Since the entry point for $Ca^{2+}$ and adenylyl cyclase are in the neighborhood of one another, a rapid reaction of the cyclase to changes in $Ca^{2+}$ concentration is ensured.

Regulation of the βγ-complex is another interesting aspect of the function of adenylyl cyclase. The inhibitory βγ-complexes probably do not come from the $G_s$-proteins from which activation of adenylyl cyclase originates. Rather, it is speculated that the inhibitory βγ-complex is released as a consequence of a parallel activation of $G_i$ and $G_o$-proteins, which both exist at much higher concentrations than the $G_s$-proteins.

The cooperation of the different regulatory signals that may have an effect on adenylyl cyclase is not yet understood. Due to the existence of different subtypes of adenylyl cyclase and their various regulation patterns, it is assumed that the individual subtypes have specific roles in different cell compartments, cells and tissues, and are each subject to defined regulatory processes.

The diverse regulation of adenylyl cyclase emphasizes the important role of this enzyme class in the signal processing in a cell. The adenylyl cyclases represent a meeting point at which different regulatory signals arrive and are weighed up against each other. In many aspects, the adenylyl cyclases are like a coincidence detector that is only activated when several signals become effective simultaneously. $Ca^{2+}$/calmodulin-dependent adenylyl cyclases are seen as an important element in learning processes and in memory formation. Both are processes for which a coincidence mechanism is postulated.

### 5.6.2 Phospholipase C

Another large class of effector molecules that are activated by G-proteins are the phospholipases of type C.

Phospholipases are enzymes that cleave phospholipids. Phospholipases of type A1, A2, C and D are differentiated according to the specificity of the attack point on the phospholipid. The bonds cleaved by these phospholipases are shown in Fig. 5.24a.

Cleavage of inositol-containing phospholipids by phospholipase C is of particular regulatory importance. Phospholipase C catalyzes the release of diacyl glycerol and inositol-1,4,5-triphosphate from phosphatidyl inositol-4,5-diphosphate, a phospholipid

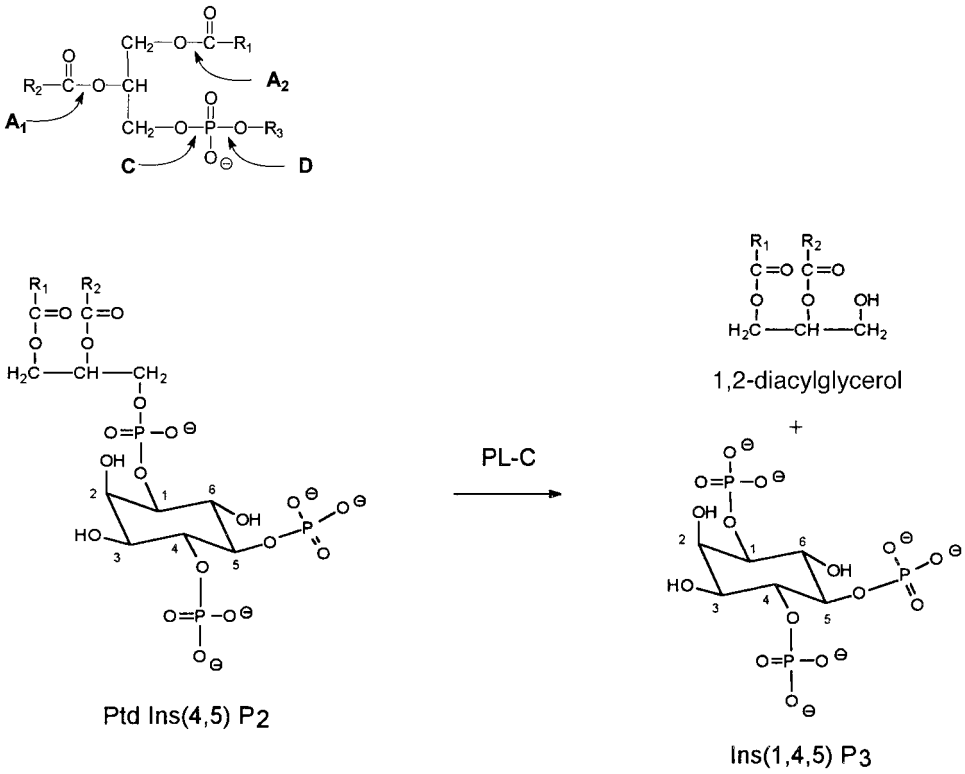



**Fig. 5.24.** Classification of the phospholipases and the reaction of phospholipase C. a) Cleavage specificity of phospholipases A1, A2, C and D. b) Cleavage of inositol-containing phospholipids by phospholipase C. In a reaction of particular importance for signal transduction, phospholipase C (PL-C) catalyzes the cleavage of phosphatidyl inositol-4,5-bisphosphate (PtdIns(4,5)$P_2$) into the messenger substances diacylglycerol and inositol 1,4,5-triphosphate (Ins(1,4,5)$P_3$).

occurring at low concentrations in the membrane (Fig 5.24b). Thus, phospholipase C has a key function in the formation of the intracellular messenger substances diacylglycerol, inositol-1,4,5-triphosphate and $Ca^{2+}$ (see Chapter 6).

Like all other protein components of the G-protein-mediated signal transduction pathways, phospholipase C also shows great heterogeneity. Three subfamilies β, γ and δ are currently differentiated within the family of phospholipase C; further subtypes exist within each of these subfamilies. Common to all phospholipases is the occurrence of pleckstrin homology domains (PH domains). The PH domains are protein modules for which a role in mediation of protein-membrane interactions and protein-protein interactions is assumed (see Chapter 8.2). For the PH domain of phospholipase Cβ, specific binding of phosphatidyl inositol phosphates and also the head group of $InsP_3$ has been demonstrated. This finding suggests that the PH domain has the function to associate the phospholipase with the membrane-localized substrate, the PtdIns(4,5)$P_2$ and to ensure an effective conversion of the substrate.

Phospholipases of type Cβ and Cγ are activated via central signaling pathways:

**Phospholipase Cβ**

Phospholipases of type Cβ are activated by $G_q$ proteins which communicate themselves with various 7-helix transmembrane receptors. The initiating external signals are diverse (see Fig. 5.14) and include hormones, neurohormones and sensory signals such as odorous agents and light (in non-vertebrates).

**Phospholipase Cγ**

Phospholipases of type Cγ are activated by receptor tyrosine kinases (see Chapter 8), and thus phospholipase Cγ is involved in growth factor controlled signal transduction pathways. The receptor tyrosine kinases (see Chapter 8) phosphorylate the enzyme at specific tyrosine residues and initiate activation of the enzyme. Characteristic for the structure of phospholipase Cγ is the occurrence of SH2 and SH3 domains (see Chapter 8). These represent protein modules that serve to attach further partner proteins.

The mechanism of activation of phospholipase Cδ is unknown.

# References Chapter 5

Berchthold, H., Reshetnikova, L., Reiser, C.O.A., Schirmer, N.K., Sprinzl, M. and Hilgenfeld, R. 'Crystal structure of active elongation factor Tu reveals major domain rearrangements' (1993) Nature *365*, 126–132

Berman, D.M. and Gilman, A.G. 'Mammalian RGS proteins: barbarians at the gate' (1998) J. Biol Chem *16*, 1269–1272

Bernstein, G., Blank, J.L., Jhon, D., Exton, J.H., Rhee, S.G. and Ross, E.M. 'Phospholipase C-ß1 is a GTPase-activating protein for $G_{q/11}$, its physiological regulator' (1992) Cell *70*, 411–418

Bourne, H.R., Sanders, D.A. and McCormick, F. 'The GTPase superfamily: a conserved switch for diverse cell functions' (1990) Nature *348*, 125–132

Bourne, H.R ' How receptors talk to trimeric G proteins' (1997) Curr. Op. Cell Biol. *9*, 134–142

Coleman, D.E., Berghuis, A.M., Lee, E., Linder, M.E., Gilman, A.G. and Sprang, S.R. 'Structures of the active conformations of $G_{i\alpha1}$ and the mechanism of GTP hydrolysis' (1994) Science *265*, 1405–1412

Conklin, B.R. and Bourne, H.R. 'Structural elements of $G_\alpha$ subunits that interact with $G_{\beta\gamma}$, receptors and effectors' (1993) Cell *73*, 631–641

Cowan, S.W.; Schirmer, T., Rummel,G., Steiert, M., Ghosh, R., Pauptit, R.A., Jansonius, J.N. and Rosenbusch,J .P. 'Crystal structures explain functional properties of two *E. coli* porins' (1992) Nature *358*, 727–733

Daaka Y., Luttrell, L.M. and Lefkowitz R.J. 'Switching of the coupling of the beta2-adrenergic receptor to different G proteins by protein kinase A' (1997) Nature *390,* 88–91

De Waard, M., Liu, H., Walker, D., Scott, A.E., Gurnett, C.A. and Campbell, K.P. ' Direct binding of G-protein βγ complex to voltage-dependent calcium channels' (1997) Nature 385, 446- 450

Deisenhofer, J. and Michel, H. 'The photosynthetic reaction centre from the purple bacterium *Rhodopseudomonas viridis*' (1989) EMBO J. *8*, 2149–2170

Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. 'Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution' (1985) Nature *318*, 618–624

Ford, C.E., Skiba, N.P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L.R., Rosal, R., Weng, G., Yang, C.S., Iyengar, R., Miller, R.J., Jan, L.Y., Lefkowitz, R.J. and Hamm, H.E. 'Molecular basis for interactions of G protein βγ subunits with effectors' (1998) Science *280*, 1271–1274

Freedman, N.J. and Lefkowitz, R.J. 'Desensitization of G protein-coupled receptors'. (1996) Prog Horm Res *51*, 319–351

Hepler, J.R. and Gilman, A.G. 'G Proteins' (1992) Trends Biochem. Sci. *17*, 383–387

Hoon, M.A., Adler, E., Lindemeier. J., Battey, J.F., Ryba, N.J. and Zuker, C.S., Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity' (1999) Cell *19*, 541–551

Hucho, F., Görne-Tschelnokow, U. and Strecker, A. 'ß-structure in the membrane-spanning part of the nicotinic acetylcholine receptor' (1994) Trends Bioch. Sci. *19*, 383–387

Iwata, S., Ostermeier, C., Ludwig, B and Michel, H. 'Structure at 2.8 A resolution of cytochrome c oxidase from Paracoccus denitrificans' (1995) Nature *376*, 660–669.

Kimura, Y., Vassylyev, D.G., , A, Kidera A., Matsushima, M., Mitsuoka, K., Murata, K., Hirai, T., Fujiyoshi, Y. 'Surface of bacteriorhodopsin revealed by high-resolution electron crystallography' (1997) Nature *389*, 206–211

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Lambright, D.G., Noel, J.P., Hamm, H.E. and Sigler, P.B. 'Structural determinants for the activation of the α-subunit of a heterotrimeric G protein' (1994) Nature *269*, 621–628

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

Iiri, T., Farfel, Z. and Bourne H.R. 'G-protein diseases furnish a model for the turn-on switch' (1998) Nature 394, 35–38

Neer, E.J. 'Heterotrimeric G proteins: organizers of transmembrane signals' (1995) Cell *80*, 249–257

Offermanns, S., Laugwitz, K.L., Spicher, K. and Schultz, G. 'G proteins of the $G_{12}$ family are activated via thromboxane $A_2$ and thrombin receptors in human platelets' (1994) Proc. Natl. Acad. Sci. U.S.A. *91*, 504–508

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

Pebay-Peyroula, E., Rummel, G., Rosenbusch, J.P. and Landau E.M. 'X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases' (1997) Science *277*, 1676–1681

Pitcher, J..A., Freedman, N.J. and Lefkowitz, R.J. 'G protein-coupled receptor kinases' (1998) Annu Rev Biochem *67*, 653–92

Scholich, K., Mullenix, J, Wittpoth, C., Poppleton, H.M., Pierre, S.C., Lindorfer. M.A., Garrison, J.C. and Patel, T.B. 'Facilitation of signal onset and termination by adenylyl cyclase' (1999) Science 283, 1328–1331

Sondek, J., Lambright, D.G., Noel, J.P., Hamm, H.E. and Sigler, P.B. 'GTPase mechanism of G proteins from the 1.7 A crystal structure of transducin $\alpha \cdot GDP \cdot AlF_4^-$ (1994) Nature *372*, 276–279

Spiegel, A.M., Backland, P.S., Butrynski, J.E., Jones, T.L.Z. and Simonds, W.F. 'The G protein connection: molecular basis of membrane association' (1991) Trends Biochem. Sci. *16*, 338–341

Strader, C.D., Fong, T.M., Tota, M.R. and Anderwood, D. 'Structure and function of G-protein coupled receptors' (1994) Ann. Rev. Biochem. *63*, 101–132

Tang, W-J. and Gilman, A.G. 'Adenylyl cyclases' (1992) Cell *70*, 869–872

Taussig, R. and Gilman, A,G. 'Mammalian membrane-boand adenylyl cyclases' (1995) J. Biol. Chem. *270*, 1–4

Tesmer, J.J., Sunahara, R.K., Gilman, A.G. and Sprang, S.R. 'Crystal structure of the catalytic domains of adenylyl cyclase in a complex with $G\alpha \cdot GTP\gamma S$.' (1997) Science *278*, 1907–1916

Wall, M.A., Coleman, D.E., Lee, E., Iniguez-Lluhi, J.A., Posner, B.A., Gilman, A.G. and Sprang, S.R. 'The structure of the G protein heterotrimer $G_{i\alpha}\ \beta_1\gamma_2$' (1995) Cell *83*, 1047–1058

Zhang, G., Liu, Y., Ruoho, A.E. and Hurley, J.H. 'Structure of the adenylyl cyclase catalytic core' (1997) Nature *386*, 247–253.

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 6**

# Intracellular Messenger Substances: "Second Messengers"

## 6.1 General Functions of Intracellular Messenger Substances

Extracellular signals are registered by membrane receptors and conducted into the cell via cascades of coupled reactions. The first steps of signal transmission often take place in close association with the membrane, before the signal is conducted into the cell interior. The cell mainly uses two mechanisms for transmission of signals at the cytosolic side of the membrane and in the cell interior. Signal transmission may be mediated by a protein-protein interaction. The proteins involved may be receptors, proteins with adaptor function alone, or enzymes. Signals may also be transmitted with the help of low molecular weight messenger substances. These are known as „*second messengers*". The intracellular messenger substances are formed or released by specific enzyme reactions during the process of signal transduction, and serve as effectors, with which the activity of proteins further in the sequence is regulated (Fig. 6.1).

The intracellular messengers are *diffusible signal molecules* and reach their target proteins mostly by diffusion. Close spatial proximity of the signal components, as achieved for transmembrane receptors and their effector proteins with the help of membrane anchoring or with specific protein-protein modules (see Chapter 5, Chapter 8), is not necessarily required for this type of signal transduction.

Two types of intracellular messenger substance can be differentiated (see Fig. 6.1):

- Messenger substances with *hydrophobic character* such as diacyl glycerol or the phosphatidyl inositol derivatives are membrane localized. The hydrophobic messengers reach membrane-associated effector proteins by diffusing through the plasma membrane and there regulate their activity.
- *Hydrophilic* messengers with good aqueous solubility are localized in the cytosol and reach their protein substrates in the cytosol.

The most important 'second messengers' are:

- hydrophilic, cytosolic:
  - cAMP, cGMP
  - inositol phosphates
  - $Ca^{2+}$
- hydrophobic, membrane associated:
  - diacylglycerol
  - phosphatidyl inositol phosphates

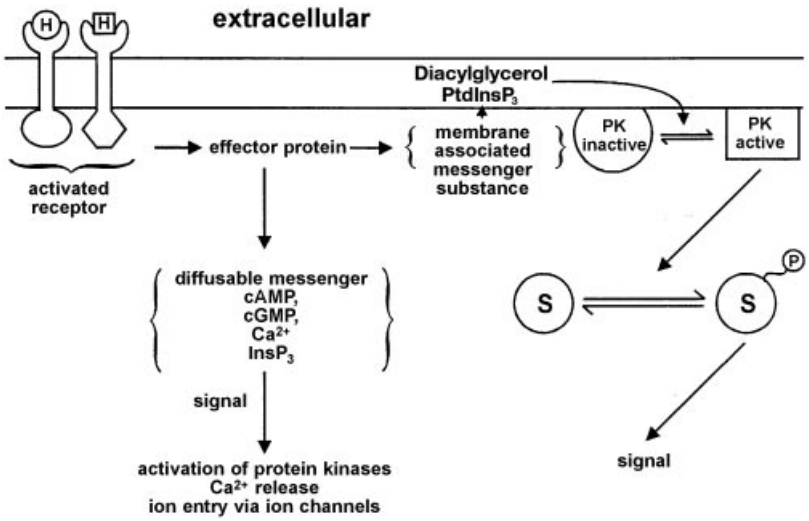



**Fig. 6.1.** Function and formation of intracellular messenger substances in signaling pathways. Starting from the activated receptor, effector proteins next in sequence are activated that create an intracellular signal in the form of diffusible messenger substances. The hydrophilic messenger substances diffuse to target proteins in the cytosol and activate these for signal transmission further. Hydrophobic messenger substances, in contrast, remain in the cell membrane and diffuse at the level of the cell membrane to membrane-localized target proteins. PK: protein kinase; S: substrate of the protein kinase.

The intracellular 'second messengers' are characterized by a series of properties that make them particularly suitable as elements of signal transduction:

- Intracellular messenger substances can be formed and degraded again in specific enzyme reactions. Via enzymatic pathways, large amounts of messenger substances can be rapidly created and inactivated again.
- Messenger substances such as $Ca^{2+}$ may be stored in special storage organelles, from which they can be rapidly released by a signal.
- Messenger substances may be produced in a location-specific manner and they may also be removed or inactivated according to their location. It is therefore possible for the cell to create signals that are spatially and temporally limited.

## 6.2 cAMP

3'-5'-cyclic AMP is a central intracellular 'second messenger' that influences many cellular functions, such as gluconeogenesis, glycolysis, lipogenesis, muscle contraction, membrane secretion, learning processes, ion transport, differentiation, growth control and apoptosis.

cAMP functions mainly in two ways:

### Regulation of Ion Channels

An important function of cAMP is the regulation of ion passage through ion channels. cAMP binds to cytoplasmic structural elements of ion channels and regulates their open state. An example is the cAMP-regulated $Ca^{2+}$ passage through cation channels. cAMP also performs this function during perception of smell in mammals.

### Activation of Protein Kinases

The general role of cAMP formed by adenylyl cyclase (see Fig. 5.21) is to activate protein kinases. Protein kinases regulated by cAMP are also classified as protein kinase A (see 7.5).

The mechanism of activation of protein kinases of type A by cAMP is schematically represented in Fig. 6.2. In the absence of cAMP, protein kinase A exists as a tetramer, composed of two regulatory (R) and two catalytic (C) subunits. In the tetrameric $R_2C_2$ form, protein kinase A is inactive since the catalytic center of the C subunit is blocked by the R subunit.

Regulation of protein kinase A by cAMP takes place by the following mechanism. An increase in cAMP concentration, triggered by activation of adenylyl cyclase, leads to binding of cAMP at specific binding sites on the regulatory subunit. The R subunits dissociate from the tetramer, the catalytic subunits are released from inhibition by the regulatory subunits and can thus phosphorylate substrate proteins.

Concentration of cAMP is controlled primarily by two means, namely via new synthesis by adenylyl cyclase and degradation by phosphodiesterases (review: Houslay & Milligan, 1997). In addition to adenylyl cyclase, the activity of which is subject to diverse regulation (see 5.6.1), the cAMP phosphodiesterases are also an important point of attack for control of the cAMP level. There are phosphodiesterases regulated by $Ca^{2+}$/calmodulin and by protein phosphorylation. More than 10 different isoforms of phosphodiesterase are known, which vary in their cyclic nucleotide specificity and in their regulation.

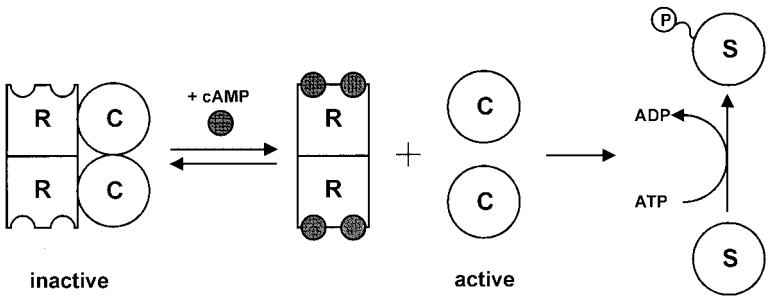



**Fig. 6.2.** Regulation of protein kinase A via cAMP. Protein kinase A is a tetrameric enyzme composed of two catalytic subunits (C) and two regulatory subunits (R). In the $R_2C_2$ form, protein kinase A is inactive. Binding of cAMP to R leads to dissociation of the tetrameric enyzme into the $R_2$ form with bound cAMP and free C subunits. In the free form, C is active and catalyzes the phosphorylation of substrate proteins (S) at Ser/Thr residues.

An increase in cAMP concentration over the threshold required for activation of protein kinase A may be reached via both paths, via stimulation of adenylyl cyclase and/or inhibition of the phosphodiesterase. Another important regulatory aspect of cAMP-mediated signal transduction is the specific localization or compartmentalization of the cAMP reaction, which allows the creation of a spatially and temporally limited signal on the inner side of the cell membrane. The proteins involved in cAMP signal conduction perform their function, without exception, in association with the cell membrane. Protein kinase A may also specifically associate with the cell membrane via specific anchor proteins (see 7.2.2). There is much to suggest that the formation and degradation of cAMP and activation of protein kinase A occurs at spatially restricted sites on the inner side of the cell membrane and a localized reaction is thus initiated. This aspect of signal transduction, known as targeting, is described in more detail in Chapter 7.7.

The nature of the substrate proteins of protein kinase of type A is very diverse; the substrate may be, e.g., other proteins or enzymes of intermediary metabolism (see Chapter 7.2).

**Regulation of a Guanine Nucleotide Exchange Factor (GEF)**

A further second-messenger function of cAMP is the activation of GEFs (see 5.4.1). Increases in the concentration of cAMP have been shown to result in the binding of cAMP to a protein termed Epac, a GEF acting on the GTPase Rap1 (see 9.5.2). This causes a conformational change leading to increased exchange activity towards Rap1 (de Rooij et al., 1998).

## 6.3 cGMP

Like cAMP, 3'-5'-cGMP is widespread as an intracellular messenger substance. Analogous to cAMP, cGMP is formed by catalysis via *guanylyl cyclase* from GTP (review: Lohmann, 1997).

There are several types of guanylyl cyclase. We know of membrane-located guanylyl cyclases that are activated via G-proteins, and we also know of purely cytoplasmically localized guanylyl cyclases. Regulators of activity of cytoplasmic guanylyl cyclase include NO.

Two main functions are attributed to cGMP:

- Regulation of activity of cGMP-dependent protein kinases
- Control of the open state of cation channels

In the vision process, cGMP has the role of regulating $Ca^{2+}$ influx via cation channels.

The cGMP-dependent protein kinases have structural elements similar to those of protein kinase A (review: Franci and Corbin, 1994;), and their activity is also regulated by an inhibitory structural element. In contrast to protein kinase A, the regulatory and catalytic functions are localized on *one* protein chain in cGMP-dependent protein kinases.

The functions of the cGMP-dependent protein kinases are not as well defined as those of protein kinase A. An important function of cGMP and the cGMP-dependent protein kinases is the regulation of the cytosolic $Ca^{2+}$ level. In smooth muscle cells, cGMP brings about a reduction in the $Ca^{2+}$, by a mechanism that has not yet been elucidated.

## 6.4 Metabolism of Inositol Phospholipids and Inositol Phosphate

Inositol-containing phospholipids of the plasma membrane are the starting compounds for formation of the signal substances diacylglycerol and inositol trisphophate. The plasma membrane contains the phospholipid phosphatidyl inositol, in which the phosphate group is esterified with a cyclic alcohol, myo-D-inositol (Fig. 6.3).

Phosphatidyl inositol (PtdIns) is first phosphorylated by specific kinases at the 4' and 5' positions of the inositol residue, leading to formation of phosphatidyl inositol-4,5-bisphosphate (PtdIns(4,5)$P_2$).

From PtdIns(4,5)$P_2$, two paths lead to physiologically important messenger substances. PtdIns(4,5)$P_2$ may be further phosphorylated by PI3-kinase to PtdIns(3,4,5)$P_3$, which functions as a membrane-localized messenger (see 6.6). In a further reaction, PtdIns(3,4,5)$P_2$ may be cleaved by phospholipase C, forming the „second messengers" inositol 1,4,5-triphosphate (Ins(1,4,5)$P_3$) and diacylglycerol (DAG). Both compounds are messengers that can activate further specific reaction chains. Ins(1,4,5)$P_3$ activates the release of $Ca^{2+}$, whilst diacylglycerol acts primarily by stimulation of protein kinase C.

Ins(1,4,5)$P_3$ is subject to rapid degradation to compounds without any regulatory activity, due to the effects of phosphatases. In addition, there are many other inositol phosphate derivatives (see Berridge & Irvine, 1989). Thus, further phosphorylation of Ins(1,4,5)$P_3$ may take place. Of the more highly phosphorylated inositol compounds, Ins(1,4,5,6)$P_4$ in particular is attributed a regulatory function.

A review of structure and metabolism of the inositol compounds important for regulation is given in Berridge and Irvine (1989), and Berridge (1993).

### Regulation of Phospholipase C

Phospholipase C, which occurs in different subtypes in the cell, is a *key enzyme* of phosphatide inositol metabolism (for cleavage specificity, see Fig. 5.24). Two central signaling pathways regulate phospholipase C activity of the cell in a positive way (Fig. 6.4). Phospholipases of type Cβ (PL-Cβ) are activated by G-proteins and are thus linked into signal pathways starting from G-protein-coupled receptors. Phospholipases of type γ (PL-Cγ), in contrast, are activated by transmembrane receptors with intrinsic or associated tyrosine kinase activity (see Chapter 8, Chapter 10). The nature of the extracellular stimuli activated by the two major reaction pathways is very diverse (see Fig 6.4), which is why the phospholipase C activity of the cell is subject to multiple regulation.

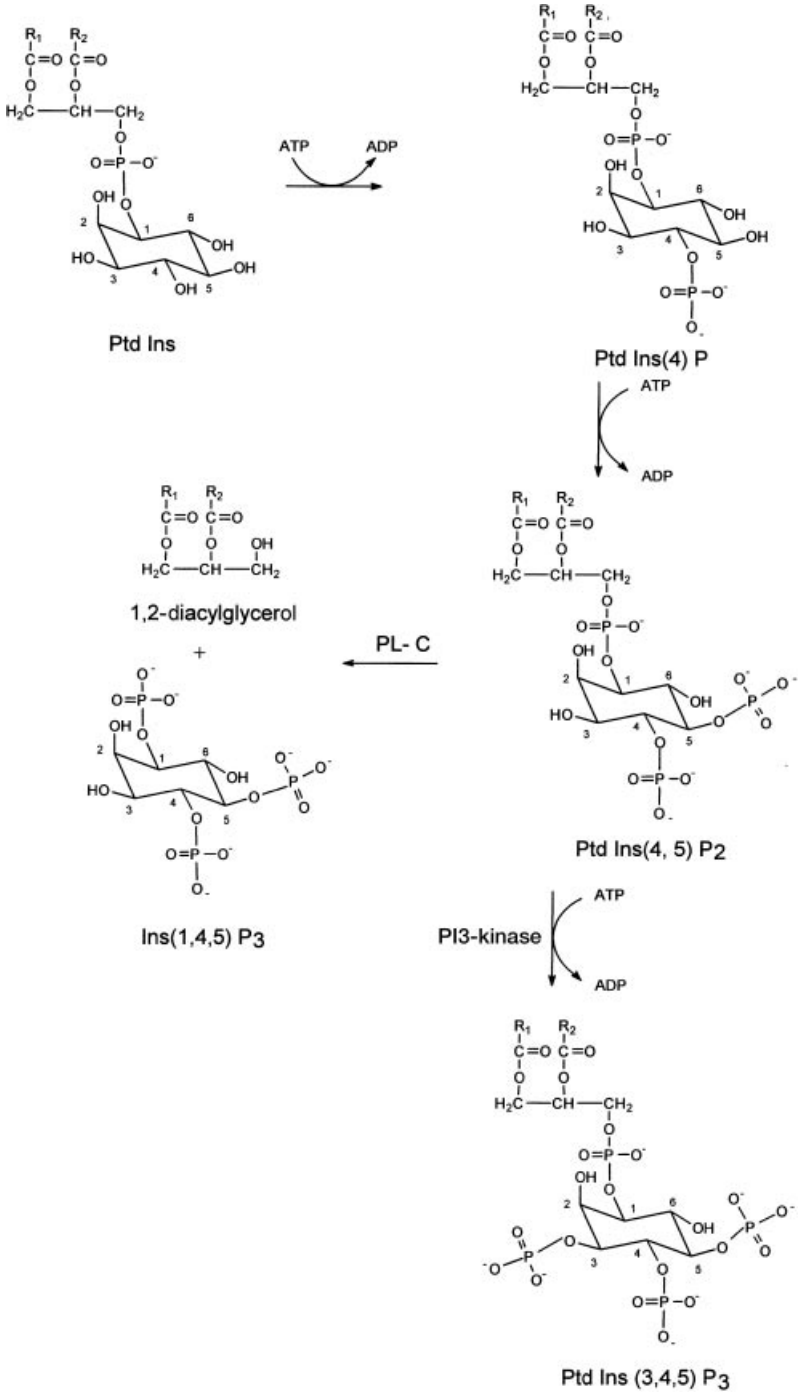



**Fig. 6.3.** Formation of diacylglycerol, Ins(1,4,5)$P_3$ and PtdIns(3,4,5)$P_3$. PL-C: phospholipase of type C: PI3-kinase: phosphatidyl inositol-3'-kinase.

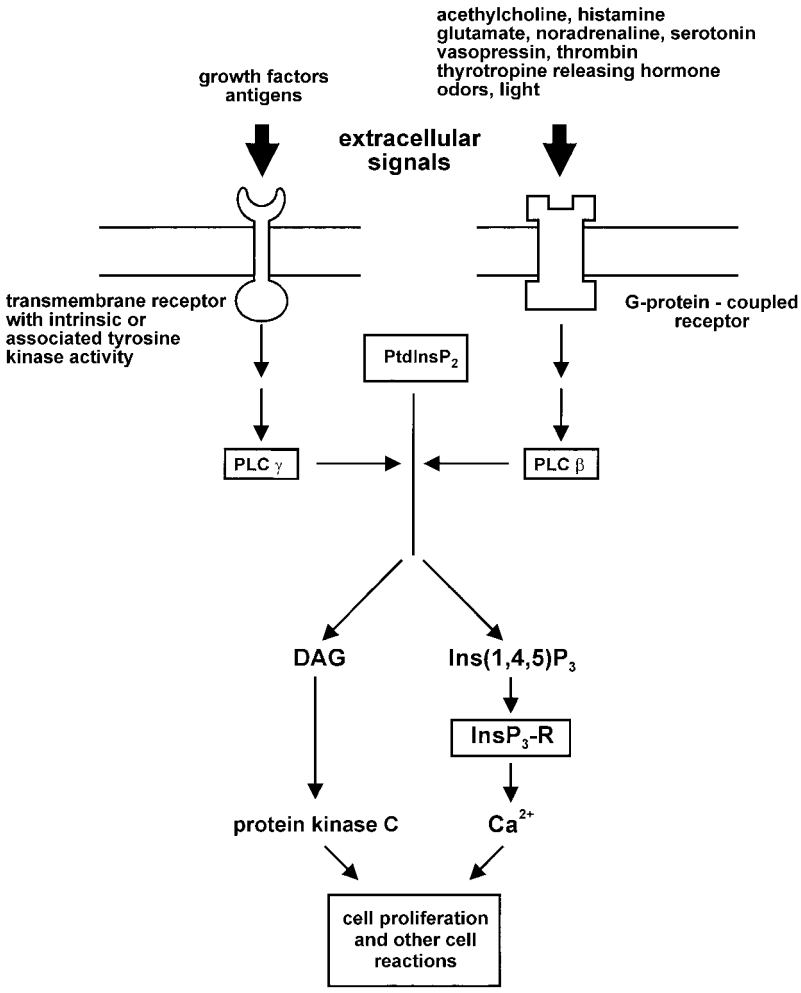



**Fig. 6.4.** Formation and function of diacylglycerol and Ins(1,4,5)$P_3$. Formation of diacylglycerol (DAG) and Ins(1,4,5)$P_3$ is subject to regulation by two central signaling pathways, which start from transmembrane receptors with intrinsic or associated tyrosine kinase activity (see Chapters 8 & 11) or from G-protein-coupled receptors. DAG activates protein kinase C (PKC, see Chapter 7), which has a regulatory effect on cell proliferation, via phosphorylation of substrate proteins. Ins(1,4,5)$P_3$ binds to corresponding receptors ($InsP_3$-R) and induces release of $Ca^{2+}$ from internal stores. The membrane association of DAG, PtdIns(3,4)$P_2$ and PL-C is not shown here, for clarity.

## Metabolic Cycle of Inositol Phosphate

The inositol phosphates are linked into a metabolic cycle (Fig. 6.5) in which they can be degraded and regenerated. Via these pathways, the cell has the ability to replenish stores of inositol phosphate derivatives, according to demand. PtdIns may be regenerated from diacylglycerol via the intermediate levels of phosphatidic acid and CDP-glycerol.

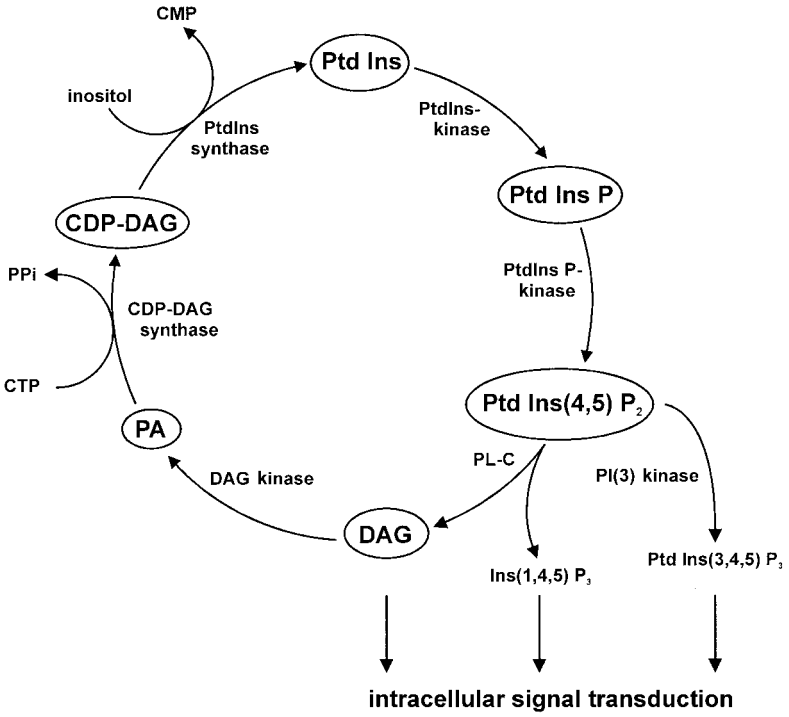



**Fig. 6.5.** Metabolic cycle of regeneration of PtdIns(4,5)$P_2$.

Regeneration of PtdIns in the inositol cycle is of particular importance in the vision process in *Drosophila* (Wu et al., 1995). In *Drosophila*, $InsP_3$ serves as a messenger during perception of light. On incidence of light, $InsP_3$ is formed from $PtdInsP_2$. It has been shown that CDP diacylglycerol synthase, which supplies CDP diacylglycerol for the resynthesis of PtIns (see Fig. 6.5), has an essential role in light perception in *Drosophila*. If regeneration of PtdIns is not possible due to a defective CDP diacylglycerol synthase, severe disorders of the vision process occur.

## 6.5 Inositol 1,4,5-Triphosphate and Release of $Ca^{2+}$

The primary signal function of Ins(1,4,5)$P_3$ is the mobilization of $Ca^{2+}$ from storage organelles.

The concentration of free $Ca^{2+}$ in the cytosol is very low, about $10^{-7}$M. In the extracellular region, in contrast, the $Ca^{2+}$ concentration is in the millimolar region. One reason that the cell tries to keep the free $Ca^{2+}$ concentration low is the ability of these ions to form poorly soluble complexes with inorganic phosphate. The low concentration of free $Ca^{2+}$ is opposed by a large storage capacity for $Ca^{2+}$ in the mitochondria and in the form of special storage organelles of the endoplasmic reticulum. In the storage associated with the endoplasmic reticulum, $Ca^{2+}$ exists in complex with the storage protein

calreticulin. Calreticulin is a low affinity $Ca^{2+}$ binding protein with a high binding capacity. In the protein-bound and compartmentalized form, $Ca^{2+}$ is not freely available but may be released in the process of signal transduction.

In muscle cells, $Ca^{2+}$ is stored in the sarcoplasmic reticulum. The storage takes place particularly by binding to the storage protein calsequestrin. It is released from storage by a neural stimulus (see 6.5.1) and initiates muscle contraction.

In the cell, the free $Ca^{2+}$ concentration is subject to strict regulation, and targeted increase of $Ca^{2+}$ concentration is a universal means of controlling metabolic reactions. Several processes are involved in its regulation (Fig. 6.6): release from $Ca^{2+}$ storage, $Ca^{2+}$ influx from the extracellular space and transport back into the $Ca^{2+}$ storage.

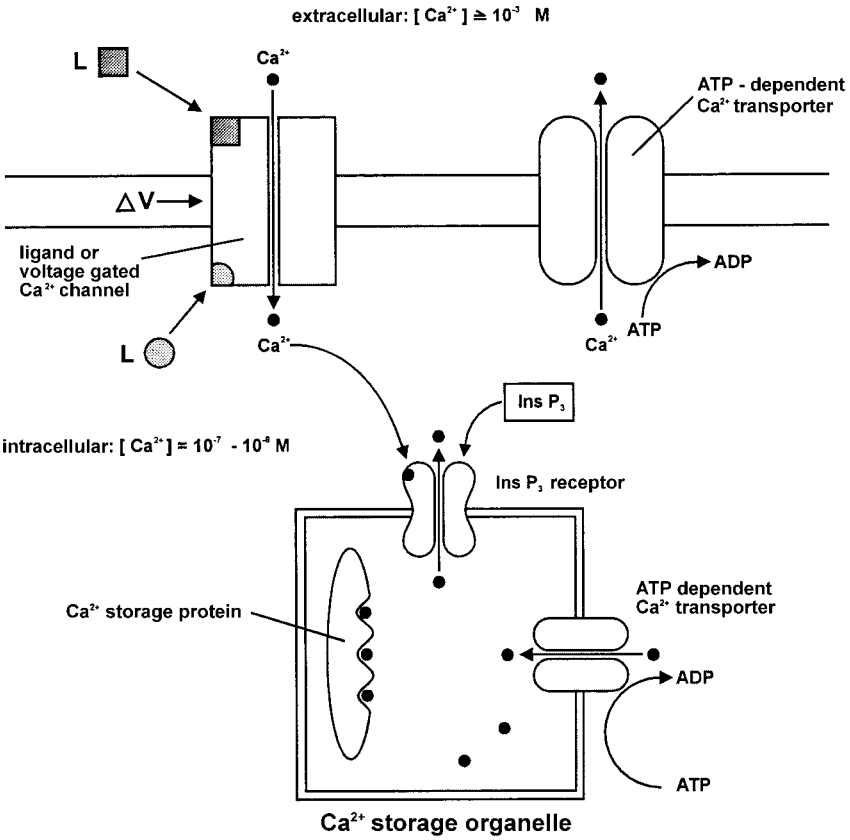



**Fig. 6.6.** Paths for increase and reduction of cytosolic $Ca^{2+}$concentration. Influx of $Ca^{2+}$ from the extracellular space takes place via $Ca^{2+}$channels; the open state of these is controlled by binding of ligand L or by a change in the membrane potential ($\Delta V$). According to the type of ion channel, the ligand may bind from the cytosolic or the extracellular side to the ion channel protein (cf. Chapter 16). The entering $Ca^{2+}$binds to $InsP_3$ receptors on the membrane of $Ca^{2+}$storage organelles and induces, together with $InsP_3$, their opening. $Ca^{2+}$flows out of the storage organelle into the cytosol via the ion channel of the $InsP_3$ receptor. Transport of $Ca^{2+}$ back into the storage organelles takes place with the help of ATP-dependent $Ca^{2+}$ transporters.

### 6.5.1 Release of $Ca^{2+}$ from $Ca^{2+}$ Storage

Mobilization of $Ca^{2+}$ from the $Ca^{2+}$ stores of the endoplasmic reticulum takes place with the help of $Ca^{2+}$ channels, of which two types stand out: the *$InsP_3$ receptors* and the *ryanodin receptors*. Both are ligand-gated $Ca^{2+}$ channels (review: Berridge, 1993; see also Chapter 16), in which receptor and ion channel form a structural unit. The $InsP_3$ receptors and ryanodin receptors are localized in the endoplasmic and sarcoplasmic reticulum, respectively, and may be opened during the process of signal transduction (Fig. 6.7).

#### The $InsP_3$ Receptor

Binding of $InsP_3$ to the $InsP_3$ receptor leads to opening of the receptor channel so that stored $Ca^{2+}$ can flow into the cytosol. The $InsP_3$ receptor is a transmembrane protein, probably with two transmembrane domains in the vicinity of the C terminus. The

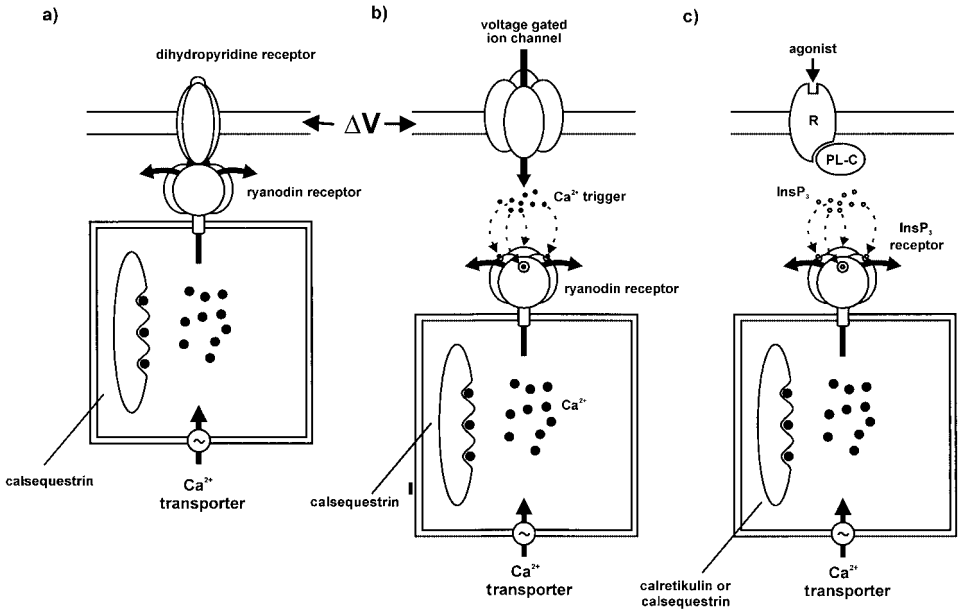



**Fig. 6.7.** Tetrameric $Ca^{2+}$ channels and control of $Ca^{2+}$ release. a) A change in the membrane potential ($\Delta V$) induces a conformational change in the dihydropyridine receptor of skeletal muscle; this is transmitted as a signal to the structurally coupled ryanodin receptor. Opening of the $Ca^{2+}$ channel takes place and efflux of $Ca^{2+}$ from the sarcoplasmic reticulum into the cytosol occurs. b) In cardiac muscle, the release of $Ca^{2+}$ takes place by a $Ca^{2+}$-induced mechanism. A potential change $\Delta V$ induces opening of voltage-gated $Ca^{2+}$ channels. $Ca^{2+}$ passes through, which serves as the trigger for release of $Ca^{2+}$ from $Ca^{2+}$ storage organelles by binding to ryanodin receptors on the surface of the storage organelles. c) Membrane-associated signaling pathways are activated by ligands and lead, via activated receptor and phospholipase C (PL-C) to formation of $InsP_3$ and to release of $Ca^{2+}$ from storage organelles. According to Berridge, (1993).

active receptor is composed of four identical subunits. It is assumed that the $Ca^{2+}$ channel is formed by the C-terminal transmembrane element and that the binding site for $InsP_3$ is localized in the large cytoplasmic region of the receptor. Opening of the $InsP_3$ receptor is subject to complex regulation involving $Ca^{2+}$, $Mg^{2+}$ and ATP, in addition to $InsP_3$.

## Ryanodin Receptor

The ryanodin receptor takes its name from its stimulation by the plant alkaloid ryanodin. In all, it has a similar composition to the $InsP_3$ receptor and is involved in $Ca^{2+}$ signal conduction in many excitatory cells (cells of banded and smooth musculature, neurons, etc.).

The open state of the ryanodin receptor is controlled in part by $Ca^{2+}$ that binds to the receptor and induces its opening. With the $Ca^{2+}$-induced opening of the ryanodin receptor, the cell has a cooperative, self amplifying mechanism that can trigger a rapid increase in the $Ca^{2+}$ concentration. An initial increase in $Ca^{2+}$ concentration, induced by $Ca^{2+}$ influx from the extracellular space due to opening of voltage-gated $Ca^{2+}$ channels, for example, initiates the opening of ryanodin receptors. The additional $Ca^{2+}$ emerging from the membrane compartments now can open more ryanodin receptors, leading to a steep increase in the $Ca^{2+}$ concentration.

In some cell types (including cardiac muscle cells, pancreatic cells), another „second messenger", the cyclic ADP-ribose (Fig. 6.8), is involved in opening the ryanodin receptors (Lee et al., 1994). The cADP-ribose is formed from $NAD^+$ by an enzymatic pathway with the help of an ADP-ribosyl cyclase.

A special coupling between extracellular $Ca^{2+}$ influx and the ryanodin receptor exists in muscle cells. There, a voltage-dependent $Ca^{2+}$ channel, the dihydropyridine receptor, is coupled directly to the cytoplasmic domain of the ryanodin receptor (see Fig. 6.7a). A depolarization of the cell membrane is transmitted in this system via an electromechanical coupling directly to the gating state of the ryanodin receptor.

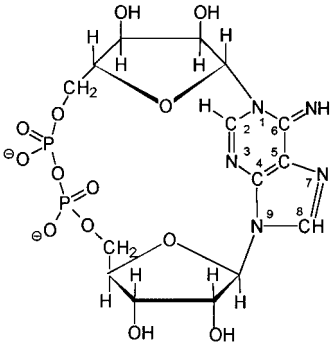



**Fig. 6.8.** Structure of cyclic ADP-ribose based on X-ray structural investigations (Lee et al., 1994).

### 6.5.2 Influx of $Ca^{2+}$ from the Extracellular Region

In the extracellular region, the $Ca^{2+}$ concentration is over $10^{-3}$M, which is very high in comparison to the free cytosolic $Ca^{2+}$ concentration. The cell membrane contains various $Ca^{2+}$ channels that enable $Ca^{2+}$ influx from the extracellular region into the cytosol. One of the primary functions of $Ca^{2+}$ entry is to charge up the internal stores, which can then release an internal $Ca^{2+}$ signal.

The open state of the $Ca^{2+}$ channels of the cell membrane is controlled by different signals. We know of *voltage-gated $Ca^{2+}$ channels* that are opened by a depolarization or change in membrane potential. There are also $Ca^{2+}$ channels that are controlled by G-protein-mediated signal transmission pathways, and *ligand-gated $Ca^{2+}$ channels* (see Chapter 16). It is also reported that $InsP_3$ can activate $Ca^{2+}$ channels in the cell membrane.

### 6.5.3 Removal and Storage of $Ca^{2+}$

The cytosolic $Ca^{2+}$ concentration is generally only temporarily increased during the process of a signal transduction. The cell possesses efficient $Ca^{2+}$ transport systems which can rapidly transport $Ca^{2+}$ back into the extracellular region or into the storage organelles. $Ca^{2+}$-ATPases, in particular, are involved in draining the cytosol of $Ca^{2+}$ back into the extracellular region. The $Ca^{2+}$-ATPases perform active transport of $Ca^{2+}$ against its concentration gradient, using the hydrolysis of ATP as an energy source. Other transport systems in the plasma membrane exchange $Na^+$ ions for $Ca^{2+}$. These *$Na^+$-$Ca^{2+}$ exchange proteins* are located especially in muscle cells and in neurons. $Ca^{2+}$-ATPases, which can fill the empty $Ca^{2+}$ storage, are also located in the membrane of the endoplasmic reticulum.

Opening of $Ca^{2+}$ leads to a local increase in the cytosolic $Ca^{2+}$ concentration from $10^{-7}$ M to $10^{-6}$M. In this concentration region, the $Ca^{2+}$ transport systems mentioned above work very efficiently. However, if an increase in $Ca^{2+}$ concentration over $10^{-5}$M takes place, e.g., due to cell damage, a level critical for the cell is reached. In this case, $Ca^{2+}$ is pumped into the mitochondria with the help of $Ca^{2+}$ transport systems localized in the inner membrane of the mitochondrion.

### 6.5.4 Temporal and Spatial Changes in $Ca^{2+}$ Concentration

Opening of $Ca^{2+}$ channels of the endoplasmic reticulum by $InsP_3$ leads to an initial increase in free $Ca^{2+}$ by emptying the $InsP_3$-sensitive $Ca^{2+}$ stores. As a consequence, a temporally and spatially limited increase in $Ca^{2+}$ is observed that can flow through the cell in the form of a $Ca^{2+}$ wave (review: Berridge, 1993; Cooper et al., 1995).

Several processes contribute to creation and propagation of the $Ca^{2+}$ wave. Firstly, an initial increase in $Ca^{2+}$ concentration can induce the opening of further $Ca^{2+}$ channels. This involves ryanodin receptors, in particular, that are activated by $Ca^{2+}$. Secondly, $Ca^{2+}$ also activates phospholipase C, leading to increased formation of $InsP_3$ that diffuses to $InsP_3$ receptors and can bring about release of more $Ca^{2+}$.

$Ca^{2+}$ release from the storage organelles occurs by a cooperative process, in opposition to transport of $Ca^{2+}$ back via the $Ca^{2+}$ transporter. The sum of both processes is the formation of spatially and temporally limited maxima in free $Ca^{2+}$ in the cytosol. Generally, there is only a short-lived effective increase in the $Ca^{2+}$ concentration. The cell receives a rapidly subsiding signal that can initiate other biochemical reactions in a narrow time window and over a limited area.

$Ca^{2+}$ signals are very versatile signals that can store different information. Like electronic or optical signals in control engineering, the information content of $Ca^{2+}$ signals may be determined by location, frequency, period and amplitude of the $Ca^{2+}$ peak. Thus, the temporal sequence of $Ca^{2+}$ signals has a regulatory function in many physiological processes. For example, in fluid secretion in the salivary glands of insects, an oscillatory $Ca^{2+}$ signal is observed, the frequency of which is determined by the intensity of the external triggering signal.

How the frequency of oscillating $Ca^{2+}$ signal is decoded or integrated and incorporated into specific biochemical reactions is not understood. There is evidence that the CAM kinase II (see Chapter 7.4.2) is involved in decoding repetitive $Ca^{2+}$ signals (De Koninck and Schulmann, 1998).

Amplitude modulation of $Ca^{2+}$ signals is seen in B lymphocytes. These can create $Ca^{2+}$ signals of higher or lower amplitude, initiated by the same extracellular signal, namely binding of an antigen to the B cell receptor. The different $Ca^{2+}$ signals activate different effector molecules and trigger different intracellular reactions. $Ca^{2+}$ signals of higher amplitude bring about activation of the transciption factor NFκB, whilst $Ca^{2+}$ signals of lower amplitude lead to activation of transcription factor NF-AT (review: Berridge, 1997). Whether a $Ca^{2+}$ signal of higher or lower amplitude is created depends on the differentiation state of the B lymphocytes.

## 6.6 Phosphatidyl Inositol Phosphate and PI3-Kinase

Several metabolic pathways lead from phosphatidyl inositol to compounds with „second messenger" character (review: Liscovitch and Cantley, 1994; Divecha and Irvine, 1995). One main pathway, the formation of diacylglycerol and $Ins(1,4,5)P_3$ from $PtdIns(4,5)P_2$, has already been described in 6.4 and Fig. 6.3. Other compounds of regulatory importance can be formed by phosphorylation at the 3' position of the inositol part of PtdIns (Fig. 6.9). The reaction is catalyzed by a class of enzymes known as phosphatidylinositide 3-kinases (PI3-kinases). The PI3-kinases phosphorylate various phosphatidyl inositol compounds at the 3' position. For example, $PtdIns(3,4,5)P_3$, produced by 3' phosphorylation of $PtdIns(4,5)P_2$, has an important function as an intracellular messenger (see 6.6.2).

### 6.6.1 PI3-Kinases

The family of PI3-kinases includes various enzymes that differ in the size of the catalytic subunit, the nature of the associated subunits and regulation (review: Krugmann

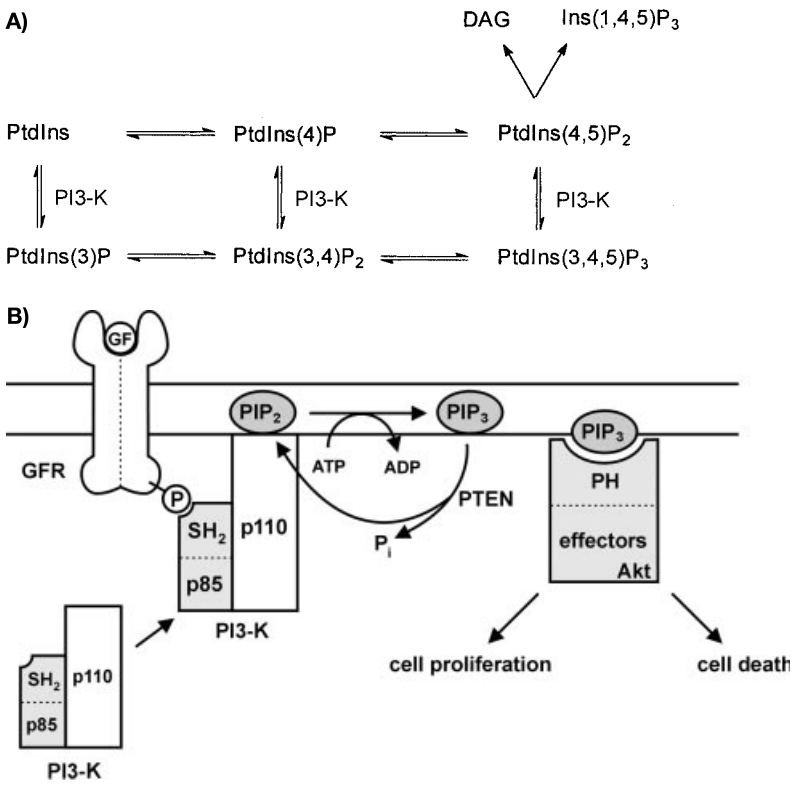



**Fig. 6.9.** A) Metabolism of phosphatidyl inositol lipids. The figure shows different membrane-associated phosphatidyl inositol phosphate compounds and their mutual conversion. PI3-kinase (PI3K) is responsible, in particular, for the formation of PtdIns(3,4,5)$P_3$ from PtdIns(3,4)$P_2$. For the reactions marked with a question mark, involvement of PI(3) kinase is not certain.
B) Signaling by PI3 kinase and PtdIns(3,4,5)$P_3$
The PI3 kinase (PI3-K) is translocated to the membrane by interaction of the SH2 domain of its p85 subunit with phosphotyrosine residues of the activated receptor. There it converts PtdIns(3,4)$P_2$.into PtdIns(3,4,5)$P_3$ which binds to PH domains of various effector molecules and recruits them into the signaling chain. The effector molecules can stimulate cell division or can induce the programmed cell death. The tumor suppressor PTEN hydrolyses phosphates from PtdIns(3,4,5)$P_3$ and thus inhibits the growth promoting effect of the PI3 kinase signaling. An important effector of PI3 kinase is the protein kinase Akt which is also termed protein kinase B (PKB). GF: growth factor; GFR: growth factor receptor.

and Welch, 1998, Anderson et al., 1999). Of the three classes (I—III) of PI3-kinases, only class I will be presented as an example here. Most members of class I are associated with a subunit that functions as an adaptor in signal transduction. The best investigated PI3-kinase, PI3-kinase α, is a heterodimer with adaptor function, made up of a catalytic subunit (p110α) and a regulatory subunit of 85 kDa (p85α). The p85α subunit has a SH3 domain, two SH2 domains and two Pro-rich domains. These domains function as binding modules which the PI3-kinase uses for specific protein-protein interactions in the process of signal transduction and for association with other signal proteins (see Chapter 8).

Other members of class I of the PI3-kinases, such as PI3-kinase of the γ subtype, are stimulated by interaction with βγ-complexes (see Chapter 5.5.7) and have their own regulatory subunit. It is interesting that both a lipid kinase activity and a protein kinase activity have been identified in the catalytic domain of the PI3-kinase γ subtype in brain (Bondeva et al., 1998). Activation of the MAPK pathway (see Chapter 10) may take place via the protein kinase activity, so that this enzyme can produce a bifurcated signal: the lipid kinase activity stimulates the Akt kinase (see below), the protein kinase the MAPK pathway. Proliferation promoting signals are transmitted via both pathways.

Most data are available for the p110α · p85α subtype of PI3-kinase. For brevity, this is referred to as PI3-kinase in the following. The PI3-kinase phosphorylates various PtdIns derivatives at the 3 position (see Fig. 6.9) *in vitro* (review Divecha and Irvine, 1995). A physiologically important substrate is PtdIns(4,5)$P_2$ that is converted to PtdIns(3,4,5)$P_3$ by PI3-kinase. PtdIns(3,4,5)$P_3$ is an intracellular messenger that has a regulatory effect in many elementary functions of the cell, such as growth control, chemotaxis and glycogen synthesis (see 6.6.2).

An important function in growth regulation is attributed to the PI3-kinase. PtdIns(3,4,5)$P_3$ is not detectable in resting cells. On stimulation of the cells with a growth factor, a rapid increase in PtdIns(3,4,5)$P_3$ occurs. An associated translocation of PI3-kinase to the membrane is observed.

Many observations indicate that PI3-kinase functions as a signal protein that receives signals on the cytoplasmic side of the cell membrane and transmits them further, although its primary role is to produce membrane-localized messenger substances.

PI3-kinase is activated by two means:

**Interaction with Activated Receptor Tyrosine Kinases**

The SH2 domain of the p85 subunit mediates an interaction with tyrosine residues on signal proteins involved in transduction of growth-regulating signals. Thus, binding of the PI3-kinase is observed to tyrosine phosphate residues of the activated PDGF receptor (see Chapter 8.1.4). Another binding partner is the insulin receptor substrate (IRS, see 8.5). In both cases, it is assumed that the binding of the SH2 domain of p85 to the tyrosine residue of the signal protein serves to target the PI3-kinase to its membrane-localized substrate. The interaction between PI3-kinase and the insulin receptor substrate links insulin signaling to the PI3-kinase pathways. Accordingly, most of the physiological functions of insulin are mediated by the PI3-kinase /Akt kinase pathway (review: Alessi and Downes, 1998).

**Activation in the Ras Pathway**

The PI3-kinase has also been identified as a part of the Ras signaling pathway (see Chapter 9). Signals originating from transmembrane receptors can be transmitted from the Ras protein to PI3-kinase. In this case, the PI3-kinase acts as the effector molecule of the Ras protein.

### 6.6.2 The Messenger Substance PtdIns(3,4,5)$P_3$

The products of the PI3-kinase reaction are different phosphoinositide derivatives phosphorylated at the 3 position, of which PtdIns(3,4,5)$P_3$ has the greatest regulatory importance. PtIns(3,4,5)$P_3$, like cAMP, has the function of a messenger substance that activates effector molecules in the sequence for further signal conduction. In contrast to cAMP, PtdIns(3,4,5)$P_3$ is localized in the cell membrane and performs its function in close association with processes at the cell membrane.

The central function of PtdIns(3,4,5)$P_3$ is to bind to pleckstrin homology domains (PH domains) of signal proteins. PH domains are found as independent protein modules in many signal proteins (see Chapter 8.2.4) that mediate protein-lipid and possibly also protein-protein interactions. PtdIns(3,4,5)$P_3$ formed by PI3-kinase serves to recruit signal molecules next in sequence to the membrane and to involve them in signal conduction. In addition, PtdIns(3,4,5)$P_3$ can also bring about an allosteric activation of its effector proteins.

An important target protein of PtdIns(3,4,5)$P_3$ is *Akt kinase*, also known as *protein kinase B (PKB)*. The signaling pathway for Akt kinase shown in Fig. 6.9b illustrates the role of PI3-kinase and PtdIns(3,4,5)$P_3$ in growth factor controlled signal pathways that lead from the cell membrane into the cytosol and the nucleus.

In the Akt signaling pathway (review: Downward, 1998), first an extracellular growth factor activates the corresponding transmembrane receptor (e.g., PGDF receptor, see 8.1). Consequently, tyrosine phosphorylation takes place on the cytoplasmic domain of the receptor. The tyrosine residues serve as docking sites for the SH2 domain of the p85 subunit of the PI3-kinase. The associated translocation of PI3-kinase is synonymous with its activation. The PtdIns(3,4,5)$P_3$ formed binds to the PH domain of the signal protein next in sequence, the Akt kinase, which recruits the latter to the membrane.

The membrane-associated Akt kinase is now a substrate for protein kinase PDK1 that phosphorylates a specific Thr and Ser residue of Akt kinase. The double phosphorylation converts Akt kinase to the active form. It is assumed that the Akt kinase now dissociates from the membrane and phosphorylates cytosolic substrates such as glycogen synthase kinase, 6-phosphofructo-2-kinase and ribosomal protein S6 kinase, $p70^{S6K}$. According to this mechanism, Akt kinase regulates central metabolic pathways of the cell. Furthermore, it has a promoting influence on cell division and an inhibitory influence on programmed cell death, apoptosis. A role in apoptosis is suggested by the observation that a component of the apoptotic program, Bad protein (see Chapter 15) has been identified as a substrate of Akt kinase.

The great importance of PtdIns(3,4,5)$P_3$ metabolism for growth regulation is illustrated by the observation that an enzyme of PtdIns(3,4,5)$P_3$ metabolism has been identified as a tumor suppressor protein (Wu et al., 1998). PTEN tumor suppressor protein has lipid phosphatase activity that is specific for hydrolysis of PtdIns(3,4,5)$P_3$. It is assumed that PTEN lipid phosphatase is a negative regulator of the Akt pathway by lowering the concentration of PtdIns(3,4,5)$P_3$ and counteracting stimulation of Akt kinase.

### 6.6.3 Functions of PtIns(4,5)$P_2$

Inositol phosphatides have another role in formation of microfilaments of the cytoskeleton.

Polymerization and depolymerization of actin, the main component of microfilaments, is controlled by a series of proteins, the activity of which is controlled by $Ca^{2+}$ and/or PtdIns(4,5)$P_2$. The $Ca^{2+}$ regulated proteins (see 6.7) are chiefly involved in processes of depolymerization of actin. Many of the proteins involved in the opposite process, actin polymerization, have specific binding sites for PtdIns(4,5)$P_2$ and PtIns(4)P and are regulated by the availability of phosphoinositides. Examples of such proteins are profilin, gelsolin, villin and talin (review: Janmey, 1994).

## 6.7 $Ca^{2+}$ as a Signal Molecule

$Ca^{2+}$ is a central signal molecule of the cell. Following a hormonal or electrical stimulation, an increase in cytosolic $Ca^{2+}$ occurs, leading to initiation of other reactions in the cell. Examples of $Ca^{2+}$-dependent reactions are numerous and affect many important processes of the organism, including:

- Muscle contraction
- Vision process
- Cell proliferation
- Secretion
- Cell motility, formation of the cytoskeleton
- Gene expression
- Reactions of intermediary metabolism

$Ca^{2+}$ signals in the form of temporally and spatially variable changes in $Ca^{2+}$ concentration serve as elements of intracellular signal conduction in many signaling pathways (review: Ghosh and Greenberg, 1995; Bootman and Berridge, 1995; Clapham, 1995).

Three main paths for increase in $Ca^{2+}$ concentration stand out (Table 6.1; Fig. 6.4; Fig. 6.6):

- G-protein-mediated signaling pathways
- Signaling pathways involving receptor tyrosine kinases
- Influx of $Ca^{2+}$ via voltage- or ligand-gated $Ca^{2+}$ channels.

#### What is the Basis of the Function of $Ca^{2+}$ as a Signal Molecule?

There are two principle mechanisms by which $Ca^{2+}$ can perform a regulatory function:

Many proteins have a specific binding site for $Ca^{2+}$ and their activity is directly dependent on $Ca^{2+}$ binding. The available $Ca^{2+}$ concentration thus directly activates the activity of these proteins (see Table 6.2).

**Table 6.1.** Receptors of the plasma membrane that mediate increase of intracellular $Ca^{2+}$.

| mediated via Phospholipase Cβ | mediated viaPhospholipase Cγ | direct |
|---|---|---|
| $\alpha_1$ adrenergic receptor | epidermal growth factor-receptor | nicotinic acteylcholine receptor |
| muscarinic acetylcholin receptors | platelet derived growth factor receptor | glutamate receptors |
| glucagon | fibroblast growth factor receptor | |
| serotonin receptor | T-cell receptor | |
| vasopressin receptor | | |
| ocytocin receptor | | |
| angiotensin II receptor | | |
| thrombin receptor | | |
| bombesin receptor | | |
| bradykinin receptor | | |
| tachykinin receptor | | |
| thromboxan receptor | | |

**Table 6.2.** $Ca^{2+}$ binding proteins

| protein | function |
|---|---|
| troponin C | modulator of muscle contraction |
| caldesmon | modulator of muscle contraction |
| α-actinin | bundling of actin |
| villin | organization of actin filaments |
| calmodulin | modulator of protein kinases and other enzymes |
| calcineurin B | protein phosphatase |
| calpain | protease |
| phospholipase A2 | release of arachidonic acid |
| proteinkinase C | ubiquitous protein kinase |
| $Ca^{2+}$-activiated $K^+$ channel | effektor of hyperpolarization |
| $InsP_3$ receptor | intracellular $Ca^{2+}$ release |
| ryanodin receptor | intracellular $Ca^{2+}$ release |
| $Na^+/Ca^{2+}$ transporter | exchange of $Na^+$ and $Ca^{2+}$ via the cell membrane |
| $Ca^{2+}$ ATPase | transport of $Ca^{2+}$ through cell membrane |
| recoverin | regulation der guanylyl cyclase |
| parvalbin | $Ca^{2+}$ storage |
| calreticulin | $Ca^{2+}$ storage |
| calbindin | $Ca^{2+}$ storage |
| calsequestrin | $Ca^{2+}$ storage |

There are many enzymes that have a specific binding site $Ca^{2+}$ in the active center and for which $Ca^{2+}$ has an essential role in catalysis. An example of a $Ca^{2+}$-dependent enzyme is phospholipase A2. Phospholipase A2 catalyses the hydrolysis of fatty acid esters at the 2' position of phospholipids (see Fig. 5.24), whereby $Ca^{2+}$ plays an essential role. The enzyme has two $Ca^{2+}$ ions bound tightly at the active center. One of the two $Ca^{2+}$ ions is directly involved in catalysis. It binds the substrate in the ground state and also helps to neutralize charge in the transition state of ester hydrolysis. The second $Ca^{2+}$ ion is assigned a role in stabilization of the transition state, in addition to a structural function (White et al., 1990).

Another example of a $Ca^{2+}$ regulated enzyme is protein kinase C (see 7.3).

We also know of many proteins without enzyme activity that have $Ca^{2+}$-regulated functions. Proteins involved in the complex process of polymerization and depolymerization of the cytoskeleton are also often regulated by $Ca^{2+}$ binding (Janmey, 1994). These include the annexins, fimbrin, gelsolin and villin. The latter two are also regulated via $PtInsP_2$. $Ca^{2+}$ and $PtInsP_2$ also have antagonistic effects on the polymerization state of microfilaments. $Ca^{2+}$ promotes depolymerization of microfilaments and $PtInsP_2$ promotes their polymerization.

Another central mechanism of signal transduction via $Ca^{2+}$ is its binding to $Ca^{2+}$ binding proteins also known as *$Ca^{2+}$ receptors*. $Ca^{2+}$ receptors have a high affinity ($K_D$ $<10^{-6}$M) and a high selectivity for $Ca^{2+}$. The receptor proteins function as regulatory proteins that can activate target proteins; their activating function is regulated by $Ca^{2+}$ ions.

We know of many $Ca^{2+}$ receptors that require an increase in $Ca^{2+}$ concentration for activation of a target protein. There are also $Ca^{2+}$ receptors with an activating function inhibited by high $Ca^{2+}$ concentrations, so that they are only active at low $Ca^{2+}$ concentrations.

### 6.7.1 Calmodulin as a $Ca^{2+}$ Receptor

The most widespread $Ca^{2+}$ receptor is *calmodulin*. Calmodulin is a small protein of ca. 150 amino acids (review of structural properties; Weinstein and Mehler, 1994; Ikura, 1996). The structure of the $Ca^{2+}$/calmodulin complex has two globular domains that are separated by a long α-helical section (Fig. 6.10a). Both globular domains have two binding sites for $Ca^{2+}$. $Ca^{2+}$ is bound via a characteristic helix-loop-helix structure, also known as an EF structure. Similar EF structures are found in many, but not all, $Ca^{2+}$ binding proteins. The binding affinity for $Ca^{2+}$ varies widely for different EF structures. The dissociation constants of $Ca^{2+}$ binding lie between $10^{-5}$M and $10^{-9}$M. In calmodulin, the four binding sites are occupied by $Ca^{2+}$ in a positive cooperative way.

With the help of NMR measurement, it has been shown that the $Ca^{2+}$/calmodulin complex has a flexible structure. Flexibility is probably of great importance for the function of $Ca^{2+}$/calmodulin. In the complex with the protein substrate (Fig. 6.10b), $Ca^{2+}$/calmodulin has a collapsed structure in which the two globular domains are much closer together than in free $Ca^{2+}$/calmodulin.

Calmodulin only associates with the target protein in the $Ca^{2+}$-bound form in most cases. There are also examples, however, in which calmodulin constantly socializes

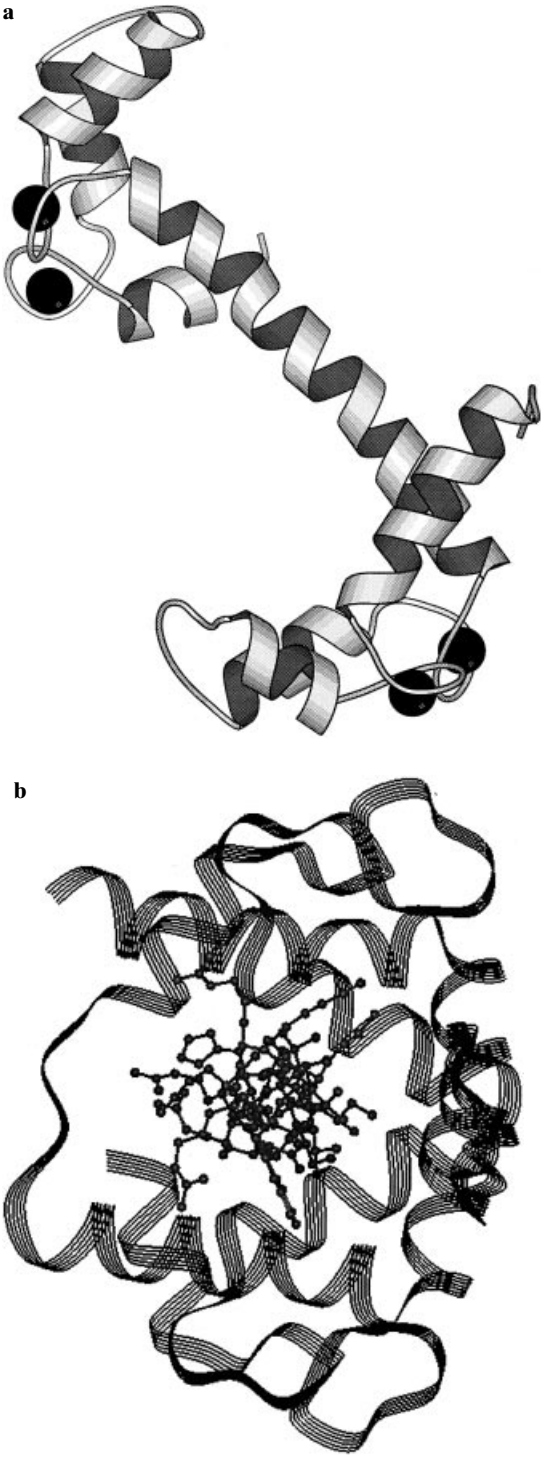



**Fig. 6.10.** Structure of $Ca^{2+}$-calmodulin. a) Structure of free $Ca^{2+}$-calmodulin. Calmodulin is composed of two domains that can each bind two $Ca^{2+}$ ions. Both domains are linked via a flexible structural element. The $Ca^{2+}$ ions are represented as spheres. MOLSKRIPT representation according to Kraulis, (1991). b) $Ca^{2+}$-calmodulin in complex with a peptide that is derived from the $Ca^{2+}$-binding domain of the myosin light chain kinase. $Ca^{2+}$-calmodulin is represented as a band, and the peptide substrate is a „ball and stick" model. The perspective is along the longitudinal axis of the linear peptide substrate.

with its target protein. One such case is the phosphorylase kinase of muscle that is regulated both by phosphorylation and by $Ca^{2+}$ binding to calmodulin (see 7.6). Here, calmodulin is one of the subunits of phosphorylase kinase.

### 6.7.2 Target proteins of $Ca^{2+}$/Calmodulin

The $Ca^{2+}$/calmodulin complex is a signal molecule that is involved in many signal transduction pathways. $Ca^{2+}$/calmodulin is involved in regulation of proliferation, mitosis, and in neuronal signal transduction. Different calmodulin subtypes are known which regulate different target proteins.

The best characterized substrate of $Ca^{2+}$/calmodulin is the *$Ca^{2+}$/calmodulin-dependent protein kinase* (CaM kinase). CaM kinase has an important function in neuronal signal transduction. The mechanism of $Ca^{2+}$/calmodulin activation of CaM kinase is described in more detail in Section 7.4, together with regulation of protein kinases. Another substrate of $Ca^{2+}$/calmodulin is *myosin light chain kinase* (MLCK), involved in contraction of smooth musculature.

### 6.7.3 Other $Ca^{2+}$ Receptors

The cell contains other $Ca^{2+}$ receptors, some of which are related to calmodulin, that occur in specialized tissue and perform specific functions there.

*Troponin C* in muscle is structurally closely related to calmodulin. It has 4 EF structures, of which only two can be occupied by $Ca^{2+}$. Troponin C is a component of the contraction apparatus of muscle. $Ca^{2+}$ binding to troponin C induces a conformational change in the troponin complex that leads to contraction of muscle.

Another regulatory $Ca^{2+}$ receptor is *recoverin*, which performs an important control function in the signal transduction cascade of the vision process, by inhibiting the activity of rhodopsin kinase (see Chapter 5.3.4).

Recoverin is a $Ca^{2+}$ receptor with four EF structures and two $Ca^{2+}$ binding sites; it can exist in the cytosol or associated with the membrane and has an N-terminal myristoyl residue as a lipid anchor. The distribution between free and membrane-associated forms is regulated by $Ca^{2+}$. Binding of $Ca^{2+}$ to recoverin leads to its translocation from the cytosol to the membrane of the rod cells. Structural determination of recoverin in the $Ca^{2+}$ bound and $Ca^{2+}$ free forms (Ames et al., 1997) indicates that membrane association of recoverin is regulated by a *$Ca^{2+}$-myristoyl switch*. The myristoyl residue can adopt two alternative positions in recoverin. In the absence of $Ca^{2+}$, recoverin exists in a conformation in which the myristoyl residue is hidden in the inner of the protein and is not available for membrane association. On $Ca^{2+}$ binding, a conformation change of recoverin takes place; the myristoyl residue moves to the outside and can now associate with the membrane.

Another $Ca^{2+}$ receptor is the *protein phosphatase calcineurin* (see 7.5), which has a specific $Ca^{2+}$ binding protein as a subunit.

## 6.8 Diacylglycerol as a Signal Molecule

During cleavage of $PtInsP_2$ by phospholipase C, two signal molecules are formed, $InsP_3$ and diacylglycerol. Whilst $InsP_3$ acts as a diffusible signal molecule in the cytosol after cleavage, the hydrophobic diacylglycerol remains in the membrane. Diacylglycerol can be produced by different pathways and it has at least two functions (Fig. 6.11). Diacylglycerol is an important source for the release of arachidonic acid, from which biosynthesis of prostaglandins takes place. The glycerine portion of the inositol phosphatide is often esterifed in the 2' position with arachidonic acid; arachidonic acid is cleaved off by the action of phospholipases of type A2.

The second important regulatory function of diacylglycerol is stimulation of protein kinase C (see 7.7). Protein kinase C is a protein kinase occurring in almost all cells, and has a regulating effect on many reactions of the cell. Characteristic for protein kinase C is its stimulation by $Ca^{2+}$, diacylglycerol and phosphatidyl serine.

## 6.9 Other Lipid Messengers

In addition to the membrane associated messenger substances diacylglycerol and $PtdIns(3,4,5)P_3$ mentioned above, other lipophilic compounds have also been identified that are specifically formed in the process of signal transduction and which function as messenger substances (review: Liscovitch and Cantely, 1994). Two such compounds are presented here:

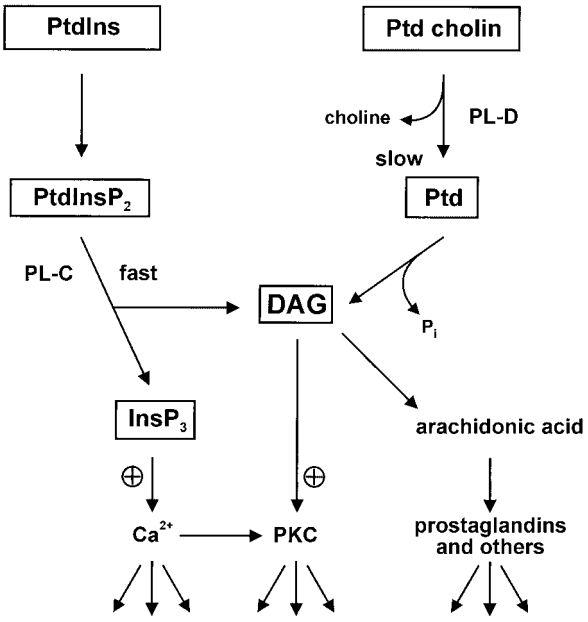



**Fig. 6.11.** Formation and function of diacylglycerol. The figure schematically shows two main pathways for formation of diacylglycerol (DAG). DAG can be formed from $PtdInsP_2$ by the action of phospholipase C (PL-C). Another pathway starts from phosphatidyl choline. Phospholipase D (PL-D) converts phosphatidyl choline to phosphatidic acid (Ptd), and the action of phosphatases results in DAG. Arachidonic acid, the starting point of biosynthesis of prostaglandins and other intracellular and extracellular messenger substances, can be cleaved from DAG. PKC: protein kinase C; PtdIns: phosphatidyl inositol.

## Ceramide

Ceramide is a lipophilic messenger that has a similar function and action to diacylglycerol. Ceramide influences growth and differentiation processes, and derivatives of ceramide have also been identified as inducers of programmed cell death (Chapter 15) (Obeid et al., 1993). The starting point for formation of ceramide is sphingomyelin, which occurs especially in the outer layer of the plasma membrane. Ceramide is produced from sphingomyelin by the action of the enzyme sphingomyelinase (Fig. 6.12). Sphingomyelinase has similar cleavage specificity to phospholipase C, in that it cleaves an alcohol-phosphate bond. Activation of sphingomyelinase is observed in stimulation of signaling pathways that are activated by the extracellular signal protein tumor necrosis factor α (TNFα, Chapter 11). The ceramide produced (or also ceramide-1-phosphate) is a membrane-located messenger substance that regulates, for example, the activity of a specific ceramide-activated protein kinase and a protein phosphatase (review: Kolesnick and Golde, 1994, Igarashi, 1997). Due to this property, ceramide functions as a messenger in signal transduction of tumor necrosis factotor α (TNFα) and of interleukin 1 (IL-1).

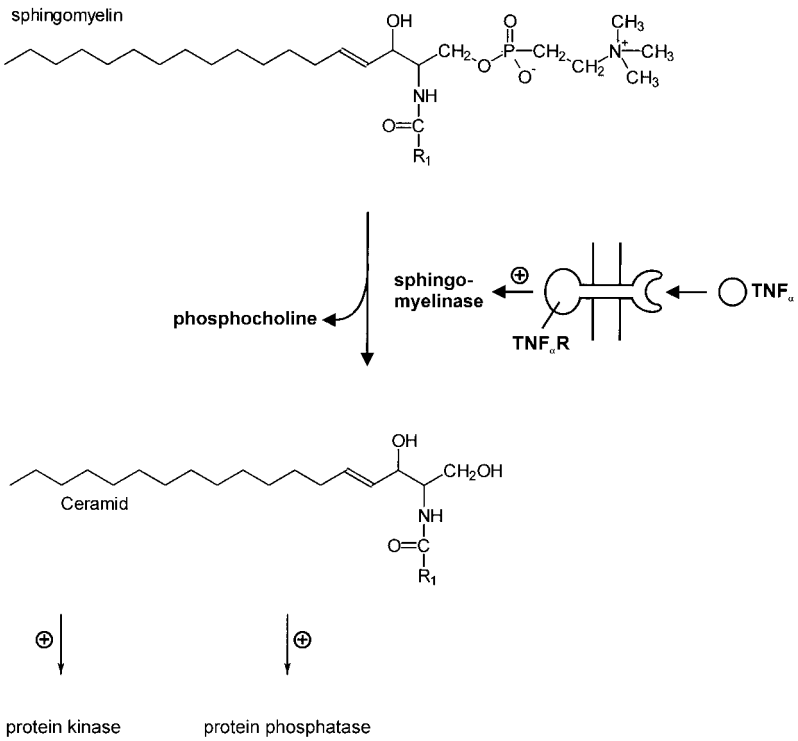



**Fig. 6.12.** Formation and function of the messenger substance ceramide. The starting point for the synthesis of ceramide is sphingomyelin, which is converted to phosphocholine and ceramide by the action of a sphingomyelinase. Sphingomyelinase is activated via a pathway starting from tumor necrosis factor α (TNFα) and its receptor. Ceramide serves as an activator of protein kinases and protein phosphatases. R1: fatty acid side chain.

**Lysophosphatidic Acid, LPA**

Messenger substances derived from phospholipids can also function as hormones and serve for communication between cells. An important *extracellular* messenger substance formed from phospholipids is *lysophosphatidic acid* (1-acyl-sn-glycerine-3-phosphate). Lysophosphatidic acid (lysophosphatidic acid, LPA) is released by platelets and other cells and reaches its target cells via the circulation. As a product of the blood clotting process, LPA is an abundant constituent of serum, where it is found in an albumin-bound form.

LPA binds and activates specific G-protein-coupled receptors found in many cells (review: Moolenaar et al., 1997). The LPA receptor can transmit the signal to $G_q$-, $G_i$- or $G_{12}$-proteins. If $G_q$ is involved, an $InsP_3$ and $Ca^{2+}$ signal is produced in the cell, whereas signal conduction via $G_i$- or $G_{12}$-proteins flows into the Ras pathway or activates the Rho proteins, respectively (see Chapter 9).

## 6.10 The NO Signal Molecule

The biological importance of nitrogen monoxide (NO) as a messenger substance was originally recognized in connection with contraction and relaxation of blood vessels. In the meantime, it has become clear that NO is a universal messenger substance that takes part in diverse forms in intercellular and intracellular communication. Practically every cell in mammals is subject to regulation by NO in one form or another.

NO fulfills many criteria required to qualify as an intracellular and intercellular messenger. NO is formed with the help of specific enzyme systems activated by extracellular and intracellular signals. NO is synthesized intracellularly and reaches its effector molecules, which may be localized in the same cell or in neighboring cells, by diffusion. Thus, it has the character of an autocrine or paracrine hormone, as well as an intracellular messenger.

Classical extracellular messengers, such as the steroid hormones, bind to their corresponding receptor but do not undergo any chemical reaction with the latter, and the binding event is sufficient to activate the receptor. NO can also accumulate at the target protein; however, it can—in a clear difference to classical messengers—bring about a *covalent modification* at the target protein. The modification of the target protein is, for the most part, reversible and the modified target protein can transmit the signal to other effector proteins. Signal transduction via NO mostly takes place by covalent modification of the target protein, in which *redox reactions* are involved in passing on the signal.

### 6.10.1 Reactivity and Stability of NO

NO is a radical that is water soluble and can cross membranes fairly freely by diffusion. Due to its radical nature, NO has only a short lifetime in aqueous solution of ca. 4 sec. Important reaction partners of NO in biological systems are oxygen $O_2$, the $O_2^-$ radical

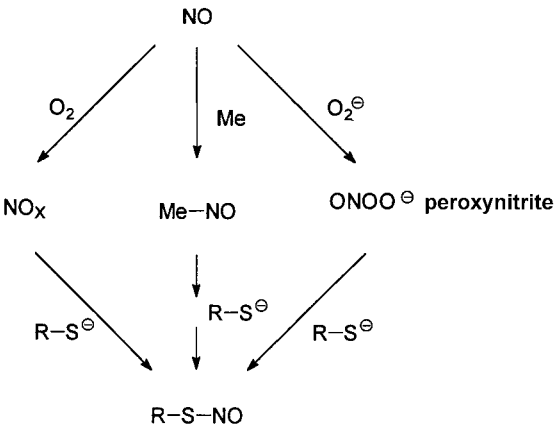



**Fig. 6.13.** Reactions of NO in biological systems. NO reacts in biological systems primarily with $O_2$, with the superoxide anion $O_2^-$ and with transition metals (Me). The products of the reaction, -$NO_x$, metal-NO adducts (Me-NO) and peroxynitrite ($OONO^-$) react further by nitrosylation of nucleophilic centers. In the cell, these are especially –SH (or thiolate-$S^-$) groups of peptides and proteins ($RS^-$).

and transition metals in free or complex form, e.g. $Fe^{2+}$ in heme (review: Stamler et al., 1992, Knowles, 1997). Furthermore, NO readily reacts with nucleophilic centers in peptides and proteins, in particular with the SH groups of Cys residues (Fig. 6.13).

#### Binding to Metal Ions

Physiologically important reactions of NO with metals take place at metal ion centers of enzymes and at Fe ions of heme proteins. Binding of NO can lead to oxidation of the complexed metal ions.

#### Reaction with Thiolene: S-Nitrosylation

Reaction of NO with cysteine residues requires firstly the one-electron oxidation of NO to the $NO^+$ ion. This then accumulates at thiolate groups of peptides or enzymes ($RS^-$), forming S-nitrosyl groups RS-NO. The mechanism of S-nitrosylation of proteins by NO in the cell is not clear. From the S-nitrosyl group, the $NO^+$ can be cleaved off in a heterolytic reaction and transferred to nucleophilic centers (thiol groups) in acceptor proteins, creating a redox signaling chain. A crucial aspect of S-nitrosylation of peptides or proteins is the increased stability of the –S-NO grouping in comparison to free NO. Whilst free NO is only stable for a few seconds, the –S-NO grouping, e.g., in glutathione, has a half life of several hours.

### 6.10.2 Synthesis of NO

NO is formed enzymatically from arginine, with the help of *NO synthase*, producing citrulline (Fig. 6.14). Citrulline and arginine are intermediates of the urea cycle and arginine can be regenerated again from citrulline by urea cycle enzymes.

There are three forms of NO synthase (NOS I, NOS II, and NOS III), and these have differing sensitivity to $Ca^{2+}$ (review: Nathan and Xie, 1994). NO synthases of

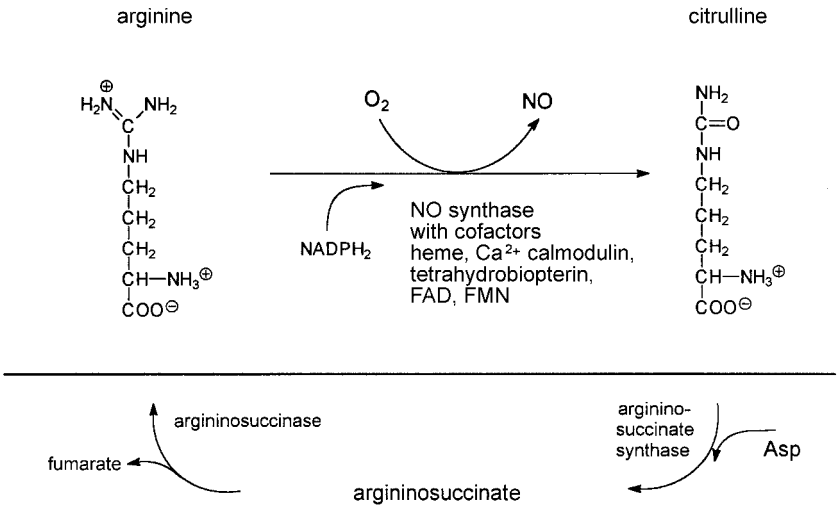



**Fig. 6.14.** Biosynthesis of NO. The starting point of NO synthesis is arginine. Arginine is converted by NO synthase, together with $O_2$ and NADPH, to NO and citrulline. Arginine can be regenerated from citrulline via reactions of the urea cycle.

types I and III require $Ca^{2+}$ for activity and are predominantly consitutively activated. The activity of NOS II, in contrast, is not regulated by $Ca^{2+}$ and the enzyme is inducable at the level of the gene.

The NO synthases are enzymes of complex composition (MW ca. 300 kDa) that are active as dimers but can also exist as inactive monomers. Furthermore, the NO synthases of types I and III undergo complex regulation by $Ca^{2+}$/calmodulin. The following cofactors and substrates are required for reaction of the NO synthase:

- FAD, FMN
- L-arginine
- Tetrahydrobiopterin
- Heme
- $Ca^{2+}$/calmodulin
- NADPH
- $O_2$

### 6.10.3 Physiological Functions and Attack Points of NO

The physiological importance of NO is due to both its regulatory and its toxic functions (review: Stamler, 1994).

### Toxic Action of NO

The toxic action of NO has an important role, especially in nerve systems. It is assumed that during a stroke, for example, excess production of NO leads to death of nerve cells. Direct modification of ligand-gated ion channels, such as the receptor for N-methyl-D-aspartate (NMDA receptor), by NO has a special role in connection with this.

### Regulatory Function of NO

NO-mediated functions are described for the following processes in particular (review: Schmidt and Walter, 1994):

- Endothelium-dependent relaxation of blood vessels
- Neurotransmission, regulation of neurotransmitter release
- Cellular immune response

Due to its high reactivity, NO can interact and react with many effector proteins. Targets are proteins with bound metal ions and specific cysteine residues of proteins. In Table 6.3, some important bioregulatory proteins are summarized, for which direct modification by NO is assumed. Two target proteins should be mentioned in particular:

**Table 6.3.** Regulatory attack points of NO.
Proteins are included for which a direct regulation by NO is assumed (according to Stammler, 1994). Direct evidence of regulatory nitrosylation has only been shown for hemoglobin, however.

| Binding site | subcellular localization | | | |
|---|---|---|---|---|
| | membrane | cytosol (incl. compartments) | nucleus | extracellular |
| thiol | NMDA receptor<br>NADPH oxidase<br>protein kinase C<br><br>adenylyl cyclase (type I) | aldolase<br>GAPDH<br>plasminogen activator<br>aldehyde dehydrogenase | AP-1<br>$NF_{\chi}B$<br>OMDM transferase | glutathione<br>albumine |
| metal | | guanylyl cyclase<br>hemoglobin<br>aconitase/IRE-BP<br>cyclooxygenase<br>cyt P450 | | |

Abbreviations: NMDA: N-methyl-D-aspartate; GAPDH: glycerine aldehyde-3-phosphate dehydrogenase; IRE-BP: iron responsive element binding protein; OMDM transferase: $O^6$-methylguanine-DNA methyltransferase

### NO-sensitive Guanylyl Cyclase

Stimulation of NO synthase leads to activation of a NO-sensitive guanylyl cyclase. The associated increase in the cGMP level has multiple consequences. The cGMP can stimulate cGMP-dependent protein kinases; it can also open cGMP-controlled ion channels. As a consequence, an increase in the intracellular $Ca^{2+}$ concentration takes place and a $Ca^{2+}$ signal is produced. NO can influence both protein phosphorylation and $InsP_3$/diacylglycerol and $Ca^{2+}$ metabolism by this mechanism and activate a broad palette of biochemical reactions in the cell.

### S-nitrosylation of Hemoglobin

Hemoglobin was the first protein for which a regulatory action of S-nitrosylation has been clearly shown (Lia et al., 1996; Fig. 6.15). Hemoglobin (Hb) is a tetramer, composed of two α and two β chains. In man, each chain has a heme system and the β chains

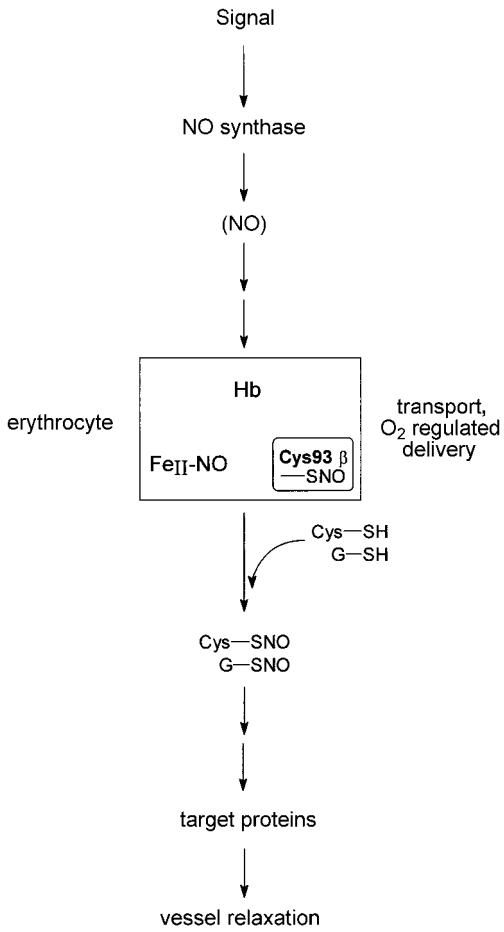



**Fig. 6.15.** Scheme of the function of nitroso-hemoglobin. NO synthase is activated by a stimulatory signal (e.g. a $Ca^{2+}$ signal) and NO is formed. The NO is transferred by direct or indirect means to hemoglobin in the erythrocytes. NO can bind to hemoglobin as a Fe-NO complex with the heme, and it can exist as a S-nitroso derivative of Cys93 of the β subunit (Cys93β). NO is transported in the blood in this form, and transferred to low molecular weight SH compounds such as glutathione (GSH) or free cysteine (Cys). The resulting nitrosyl compounds Cys-SNO and G-SNO can diffuse to target proteins and pass the NO signal on to these. The figure does not show the complex regulation of NO compounds by hemeglobin via oxygen.

have a reactive cysteine group (Cys93). The Hb may bind NO at two sites. Firstly, NO can bind to the Fe(II) of the heme grouping; secondly, NO can accumulate at Cys93 of the β chain by forming an S-nitrosyl.

Nitrosylation of hemoglobin is a dynamic vesicle function. The Hb functions as a means of transport which helps the erythrocytes to store NO and to transport it in the vessel system. The heme- and Cys93-bound NO can be delivered by Hb in a regulated way. The NO binding capacity of hemoglobin is linked to $O_2$ binding to hemoglobin and is allosterically controlled. Binding of $O_2$ to hemoglobin and the associated transition from the T to the R form (see 2.3) facilitates binding as S-nitrosyl, whilst delivery of $O_2$ and the transition from the R to the T form leads to dissociation of NO. The released NO can now be passed on to low molecular SH-containing compounds such as glutathione or cysteine. It is assumed that the S-nitroso derivatives of glutathione or cysteine are a diffusible transport form for NO, which enables NO to reach target proteins in the endothelium of small blood vessels, where it has a vessel widening effect, reducing blood pressure.

# References Chapter 6

Alessi, D.R. and Downes C.P. 'The role of PI3-kinase in insulin action' (1998) Biochim Biophys Acta *1436*, 151–64

Anderson, R.A., Boronenkov, I.V., Doughman, S.D., Kunz, J. and Loijens, J.C. 'Phosphatidylinositol phosphate kinases, a multifaceted family of signaling enzymes'(1999) J Biol Chem *274*, 9907–10

Berridge, M.J. und Irvine, R.F. 'Inositol phosphates and cell signalling' (1989) Nature *341*, 197–205)

Berridge, M. 'Inositol trisphosphate and calcium signalling' (1993) Nature *361*, 315–325

Berridge, M.J. ' The AM and FM of calcium signalling' (1997) Nature *386*, 759–760

Bondeva, T., Pirola, L., Bulgarelli-Leva, G., Rubio, I., Wetzker, R. and Wymann M.P. 'Bifurcation of lipid and protein kinase signals of PI3Kgamma to the protein kinases PKB and MAPK' (1998) Science *282*, 293–296

Bootman, M.D. und Berridge, M.J. 'The elemental principles of calcium signalling' (1995) Cell *83*, 675–678

Clapham, D.E. 'Calcium signalling' (1995) Cell *80*, 259–268

Cooper, D.M.F., Mons, N. und Karpen, J.W. 'Adenylyl cyclase and the interaction between calcium and cAMP signalling' (1995) Nature *374*, 412–424

Divecha, N. und Irvine, R.F. 'Phospholipid signalling' (1995) Cell *80*, 269–278

Downward, J. 'Mechanisms and consequences of activation of protein kinas B/Akt' (1998) Curr. Op. Cell Biol. *10*, 262–267

Francis, S.H. und Corbin, J.D. 'Structure and function of cyclic nucleotide-dependent protein kinases' (1994) Ann. Rev. Physiol. *56*, 237–272

Ghosh, A. und Greenberg, M.E. 'Calcium signalling in neurons: molecular mechanisms and cellular consequences' (1995) Science *268*, 239–247

Houslay, M.D. and Milligan, G. 'Tailoring cAMP-signalling responses through isoform multiplicity' (1997) Trends Biochem Sci. *22*, 217–224

Hunter, T, 'When is a lipid kinase not a lipid kinase? When it is a protein kinase' (1995) Cell *83*, 1–4

Igarashi, Y. 'Functional roles of sphingosine, sphingosine 1-phosphate, and methylsphingosines: in regard to membrane sphingolipid signaling pathways' (1997) J Biochem. (Tokyo) *122*, 1080–7

Ikura, M. 'Calcium binding and conformational response in EF-hand proteins' (1996) Trends Biochem. Sci. *21*, 14–17

Janmey, P.A. 'Phosphoinositides and calcium as regulators of cellular actin assembly and disassembly' (1994) Ann. Rev. Physiol. *56*, 169–191

Knowles RG 'Nitric oxide biochemistry' (1997) Biochem Soc Trans *25*, 895–901

De Koninck, P. and Schulman, H. 'Sensitivity of CaM kinase II to the frequency of Ca2+ oscillations' (1998) Science *279*, 227–230

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Krugmann, S.and Welch, H. 'PI 3-kinase' (1998) Curr Biol. *8*, 828

Lee, H.C., Galione, A. und Walseth, T.F. 'Cyclic ADP-ribose: metabolism and calcium mobilizing function' (1994) Vitam. Horm. *48*, 199–257

Liscovitch, M. und Cantley, L.C. 'Lipid second messengers' (1994) Cell *77*, 329–334

Lohmann, S.M., Vaandrager, A.B., Smolenski, A., Walter, U. and De Jonge, H.R. 'Distinct and specific functions of cGMP-dependent protein kinases' (1997) Trends Biochem Sci. *22*, 307–12

Moolenaar, W.H., Kranenburg, O., Postma, F.R. and Zondag, G.C.M. 'Lysophosphatidic acid: G-protein signalling and cellular responses (1997) Curr. Op. Cell Biol. *9*, 168–173

Nathan, C. und Xie, Q. 'Nitric oxide synthases: roles, tolls, and controls' (1994) Cell *78*, 915–918

de Rooij, J., Zwartkruis, F.J., Verheijen, M.H., Cool, R.H., Nijman, S.M., Wittinghofer, A. and Bos, J.L. 'Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP' (1998) Nature *396*, 474–477

Schmidt, H.H. und Walter, U. 'NO at work' (1994) Cell *78*, 919–925

Stamler, J.S., Singel, J.D. und Loscalzo, J. 'Biochemistry of nitric oxide and its redox-activated forms' (1992) Science *258*, 1898–1902

Stamler, J.S. 'Redox signalling: nitrosylation and related target interactions of nitric oxide' (1994) Cell *78*, 931–936

Weinstein, H. und Mehler, E.L. 'Calcium binding and structural dynamics in the function of calmodulin' (1994) Ann. Rev. Physiol. *56*, 213–236

White, S.P., Scott, D.L., Otwinowski, Z., Gelb, M.H. und Sigler, P.B. 'Crystal structure of Cobra phospholipase $A_2$ in a complex with a transition state analogue' (1990) Science *250*, 1560–1566

Wu, L., Niemeyer, B., Colley, N., Socolich, M. und Zuker. C.S. 'Regulation of PLC-mediated signalling *in vivo* by CDP-diacylglycerol synthase' (1995) Nature *373*, 216–222

Wu, X., Senechal, K., Neshat, M.S., Whang, Y.E. and Sawyers, C.L. 'The PTEN/MMAC1 tumor suppressor phosphatase functions as a negative regulator of the phosphoinositide 3-kinase/Akt pathway' (1998) Proc Natl Acad Sci U S A *95*, 15587–91

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 7**

# Ser/Thr-specific Protein Kinases and Protein Phosphatases

Reversible phosphorylation of amino acid side chains is a widely used principle for regulation of the activity of enzymes and signaling proteins (see Chapter 3). Via this function, protein kinases and protein phosphatases play pivotal roles in regulating aspects of metabolism, gene expression, cell growth, cell division and cell differentiation. Almost all intracellular signaling pathways use protein phosphorylation to create signals and conduct them further. The protein kinases are certainly one of the largest protein families in the cell. Conservative estimates suggest that more than 1000 protein kinases are coded in the mammalian genome. Of the various protein kinases, the Ser/Thr-specific and Tyr-specific enzymes are the best characterized. Tyr-specific protein kinases are dealt with in Chapters 8 and 11. Before going on to the protein family of Ser/Thr-specific protein kinases, a rough classification of protein kinases will be presented.

## 7.1 Classification, Structure and Characteristics of Ser/Thr-specific Protein Kinases

### 7.1.1 General Classification and Function of Protein Kinases

The first protein kinase obtained in a purified form was the Ser/Thr-specific phosphorylase kinase of muscle, in 1959 (Krebs et al., 1959). With the discovery of the Tyr-specific protein kinases (Erikson et al., 1979), the Ser/Thr-specific protein kinases were joined by another extensive class of protein kinases of regulatory importance, to which a central function in growth and differentiation processes was soon attributed. At present, several hundred different protein kinases are known in mammals, most of which are Ser/Thr- or Tyr-specific. In addition, there are some protein kinases that phosphorylate other amino acids (review: Hunter, 1991).

Based on the nature of the acceptor amino acids, four classes of protein kinases can be distinguished (Fig. 7.1):

- *Ser/Thr-specific protein kinases* esterify a phosphate residue with the alcohol group of Ser and Thr residues.
- *Tyr-specific protein kinases* create a phosphate ester with the phenolic OH group of Tyr residues.
- *Histidine-specific protein kinases* form a phosphorus amide with the 1 or 3 position of His. The members of this enzyme family also phosphorylate Lys and Arg residues.

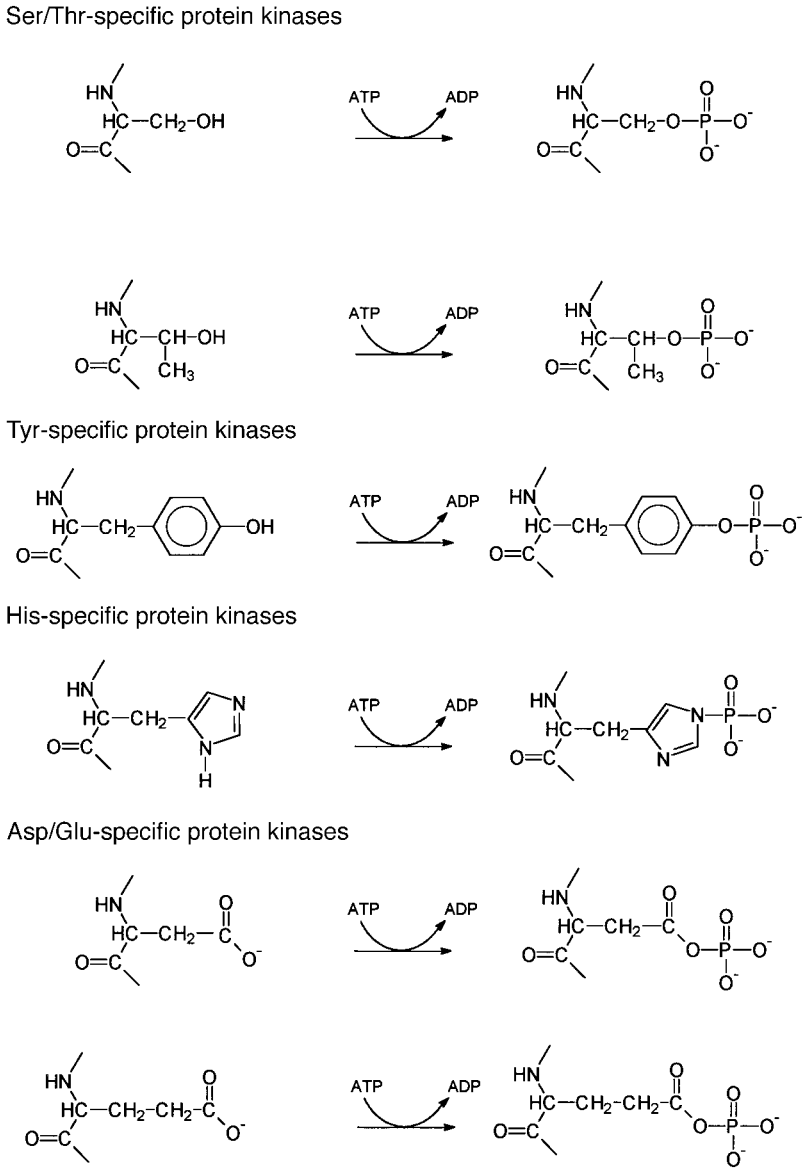



**Fig. 7.1.** Amino acid specificity of protein kinases.

– *Aspartate- or glutamate-specific protein kinases* create a mixed phosphate-carboxylate anhydride.

Whilst comprehensive data are available on structure and function of the Ser/Thr- and Tyr-specific protein kinases, both the other families of protein kinases are incompletely characterized.

In the sequence of protein kinases, a homologous catalytic domain can be identified, that includes ca. 270 amino acids. Based on the sequence of the catalytic domain, it is possible to differentiate between Ser/Thr- and Tyr-specific protein kinases. Furthermore, homology considerations enable identification of subfamilies within both the larger families of Ser/Thr- and Tyr-specific protein kinases.

Protein phosphorylation by protein kinases is a general and versatile tool in the cell, with the function of a switch in signaling pathways. Examples of cellular activities regulated by protein kinases are diverse, affecting practically all the cell's performance. Protein phosphorylation is found in:

- Enzymes: as elements of signal chains (Chapters 7, 8, 10, 13) and in enzymes of intermediary metabolism (Chapter 2)
- Adaptor proteins (Chapter 8)
- Signal proteins (Chapters 5, 7, 8)
- Transcription factors (Chapter 1)
- Ion channels (Chapter 16)
- Transmembrane receptors (Chapters 5, 8, 11, 12)
- Ribosomal proteins (ribosomal protein S6, Chapter 6)
- Structural proteins
- Transport proteins

The switch function of protein phosphorylation is based on different mechanisms that may work alone or in combination. Phosphorylation by protein kinases influences function and activity of the protein substrate, especially in the following ways:

- Induction of conformational changes (Chapters 2, 11, 13)
- Direct interference with binding of substrate or other binding partners, e.g., isocitrate dehydrogenase (Chapter 2)
- Creation of binding sites for effector molecules in the sequence: examples of this are binding of Tyr-P to SH2 and PTB domains (Chapter 8) and binding of Ser-P to 14–3-3 proteins (Chapter 8).

## 7.1.2 Classification of Ser/Thr-specific Protein Kinases

Due to the complexity and number of the subfamilies, only selected subfamilies of the Ser/Thr-specific protein kinases in vertebrates are presented in the following (according to Hunter, 1991).

**Subfamilies of the Ser/Thr-specific Protein Kinases**

– Protein kinases regulated by a cyclic nucleotide
  - cAMP regulated protein kinase, protein kinase A, PKA
  - cGMP regulated protein kinase
– Diacylglycerol regulated protein kinases
  Protein kinase C, with further division into the α, βI, βII, γ, δ, ε, ζ, η, ι, λ, and μ subtypes

- Calcium/calmodulin regulated protein kinases
  - γ subunit of phosphorylase kinase
  - Myosin light chain kinase, MLCK
  - $Ca^{2+}$/calmodulin dependent protein kinase II
- Ribosomal S6 protein kinase
  Kinases that specifically phosphorylate ribosomal protein S6
- Serpentine receptor kinase
  β-adrenergic receptor kinase, βARK
- Casein kinase II (casein kinase gets its name from the observation that casein, milk protein, is a good substrate)
- Glycogen synthase kinase
- CDC2 kinases (representatives of this family are central elements of regulation of the cell cycle, see Chapter 13
- Mitogen activated kinases, MAP kinases (the MAP kinases are involved in transduction of growth promoting signals, see Chapter 10)
- Mos/raf protein kinases (see Chapter 9, the mos/raf protein kinases are also involved in signal transduction of growth factors).

There are many other protein kinases that do not show any close relationship to these subfamilies. These include protein kinases with two-fold specificity, in that they can phosphorylate Ser/Thr and also Tyr residues. An example of a protein kinase with two-fold specificity is the MAP kinase kinase (see Chapter 10).

### 7.1.3 Substrate Specificity of Ser/Thr-specific Protein Kinases

Taking into account the many Ser and Thr residues in proteins, the question arises of which parameters define the phosphorylation site of a substrate protein. With the help of targeted exchange of amino acids in substrate proteins, sequence comparison of phosphorylation sites, and use of defined peptides as substrates, it has been possible to clearly show that the sequence in the neighborhood of a Ser/Thr residue is an important determinant of specificity. The different Ser/Thr-specific protein kinases show different requirements with respect to the neighboring sequence of the Ser or Thr residue to be phosphorylated, so that each subfamily has its own consensus sequence for phosphorylation. Within the family of Ser/Thr kinases, however, the sequences of the phosphorylation sites show great variability and it is not possible to give a consensus sequence for all Ser/Thr-specific protein kinases. Due to the different roles and the very different substrate proteins, this is not surprising. For more information on consensus sequences for protein phosphorylation, the reader is referred to sequence databases and to review articles (Pearson and Kemp, 1991). It should be pointed out that several Ser/Thr residues are found in many phosphorylation sequences, so that multiple and cooperative phosphorylation is possible in a sequence segment. Phosphorylation of the large subunit of the RNA polymerase II (see 1.2) is particularly marked. At the C terminus, this contains 52 copies of the heptamer sequence YSPTSPS as potential phosphorylation sites.

Many protein kinases show indistinct substrate specificity, especially in *in vitro* experiments. Certain phosphorylation sites of the histone H1 can be phosphorylated by various protein kinases. Insight into the specificity requirements of protein kinases was only possible once highly resolved structural information on the binding of model substrates to protein kinases was available (see 7.1.4).

## 7.1.4 The Catalytic Domain of Ser/Thr-specific Protein Kinases

Structural information on Ser/Thr-specific protein kinases (review: Goldsmith and Cobb, 1994, Johnson et al., 1996, Johnson, 1998) indicates a markedly conserved structure of the catalytic domain. In Fig. 7.2a, the structure of the catalytic subunit of protein kinase A is shown in complex with an inhibitor peptide (Knighton et al., 1991). In this case, the inhibitor peptide serves as a model for a phosphorylation substrate. In the form of two Arg residues, it possesses part of a sequence characteristic of phosphorylation sites of protein kinase A, which is defined by two Arg residues in the neighborhood of the Ser residue to be phosphorylated (Fig. 7.3). The inhibitor peptide lacks the

**a**

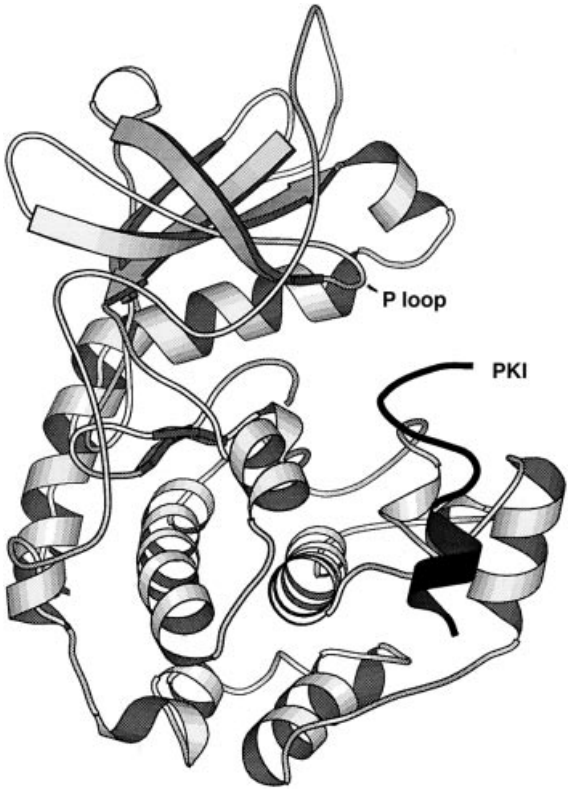



**Fig. 7.2**

**b**

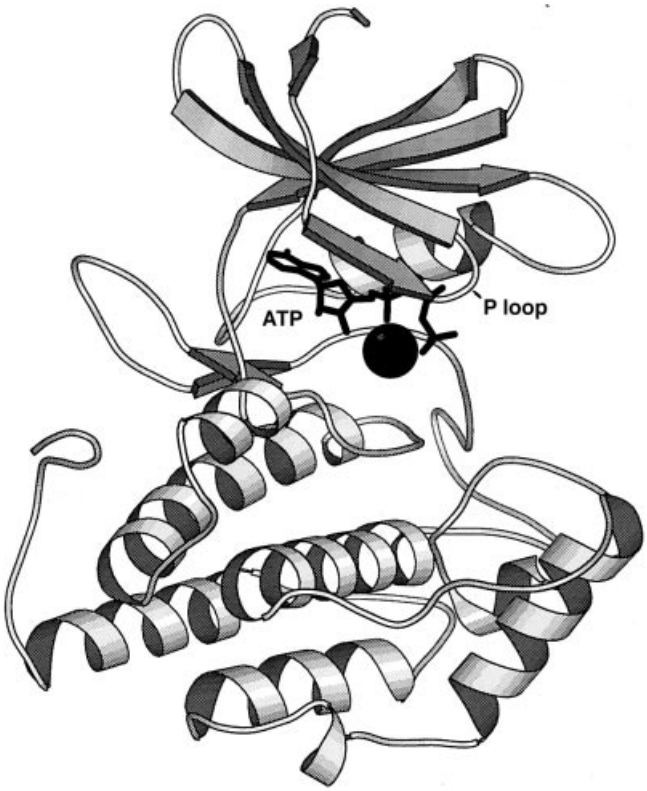



**Fig. 7.2.** Structure and substrate binding sites of Ser/Thr-specific protein kinases. a) Peptide binding site: structure of the catalytic subunit of the cAMP-dependent protein kinase A from mouse, with bound inhibitor peptide PKI (5–22), shown in dark in the figure. PKI (5–22) is a fragment (amino acids 5–22) of the naturally occurring heat-stable protein kinase inhibitor PKI. The inhibitor peptide binds in the region of the substrate binding site between the two lobes of protein kinase A (Knighton et al., 1991). The P-loop is involved in binding the phosphate residue of ATP. b) ATP binding site: structure of casein kinase I with bound Mg-ATP. The $Mg^{2+}$ is shown as a sphere. MOLSKRIPT representation according to Kraulis, (1991).

Ser residue to be phosphorylated; this is replaced by alanine. Due to these characteristics, the inhibitor peptide is bound in a very similar way and with similar affinity to a substrate, but it cannot react and has the property of a pseudosubstrate. The contacts between protein kinase A and a substrate are shown as a model for a peptide known as kemptide, which serves as a phosphorylation substrate.

The catalytic domain of protein kinase A has a two lobe structure, composed of a smaller lobe with a large portion of β-sheet structures and a larger lobe that is mostly α-helical. All Ser/Thr- and Tyr-specific protein kinases structurally characterized to date show a similar domain structure.

The binding site for the protein substrate and the ATP binding site are located in the cleft between the two lobes. The link between the two lobes is flexible and functions as

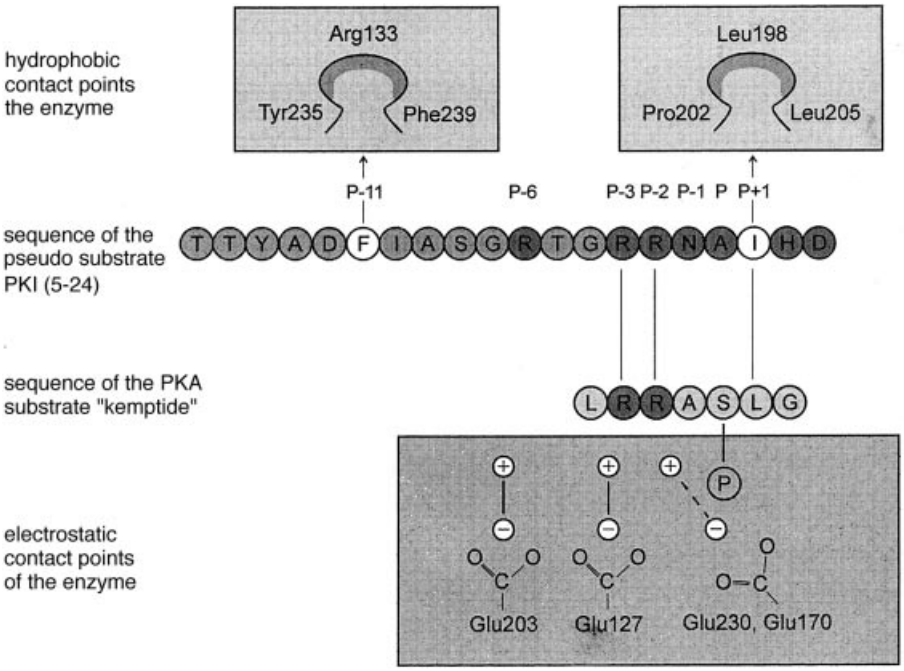



**Fig. 7.3.** Contact points of protein kinase A with inhibitor and peptide substrates. The contact points are shown between the catalytic subunit of protein kinase A, the inhibitor PKI (5—24) and a peptide substrate (kemptide). The inhibitor binds, via two Arg residues (P3, P-2), to the same Glu residue of protein kinase A that also binds to the substrate. The phosphorylation site is labelled as P, the *N*-terminal and *C*-terminal situated amino acids are labelled as (-) and (+) and are numbered starting from the phosphorylation site. According to Kemp et al., (1994).

a hinge. On binding substrate proteins and ATP, the two lobes fold together, bringing the ATP binding site into the interior of the molecule, between the two lobes. A glycine-rich loop is involved at the ATP binding site, with a consensus sequence Y-Gly-X-Gly-(Phe/Tyr)-Gly-X-Val. The loop is flexible and it is attributed an important role in catalysis of phosphate transfer and regulation of the protein kinase, in addition to ATP binding (Bossemeyer, 1994). ATP binding is shown, using casein kinase I as an example, in Fig. 7.2b.

The catalytic center is formed by residues from both lobes. Sequence comparisons, mutation experiments and biochemical studies indicate an essential function in catalysis of phosphate transfer for the conserved amino acids Lys72, Asp166 and Asp184 (numbering of PKA). However, the catalytic mechanism of phosphate transfer is not definitely established. It is generally assumed that Asp166, which is invariant in all protein kinases, serves as a catalytic base for activation of the Ser/Thr hydroxyl and that the reaction takes place by an „in-line“ attack of the Ser-OH at the γ-phosphate.

The inhibitor is bound in an elongated form to the substrate binding surface and forms many contact points to the protein kinase. Hydrophobic residues of the inhibitor find their binding equivalents in hydrophobic pockets of the enzyme. The Arg residues important for specificity of substrate binding are in contact with Glu residues of the

protein via ionic interactions (Fig. 7.3). In total, there are many possible contact points between kinase and inhibitor. The diversity of the contact points explains why certain variations in the neighboring sequence of the Ser residue are tolerated. For tight binding and correct orientation of a substrate protein, it is not necessary that all possible contacts are formed. A substrate that only provides some of the contact points can still be bound in a productive way, assuming that some key amino acids are present. Such key amino acids for protein kinase A are basic residues in the neighborhood of the Ser residues (see Fig. 7.3); for casein kinase II, acidic residues are the determinants of specificity. A particular protein kinase thus does not have any strictly defined specificity elements on the substrate side, which is not the case for specific protein-nucleic acid interactions. Protein kinases do not recognize any fixed sequence, rather they recognize *families of related sequences.*

### 7.1.5 Autoinhibition and Intrasteric Regulation of Ser/Thr-specific Protein Kinases

Protein kinases can exist in active and inactive forms, which is why they are able to perform the function of a switch in signaling pathways. Protein kinases are particularly suitable as switches in signal pathways due to their flexible structure of two domains that can adopt different orientations with respect to one another. Furthermore, in the cleft between the two domains, it is possible to initiate signal-controlled conformational changes of great importance for substrate binding and catalytic activity.

The transition between inactive and active forms of the protein kinase may be controlled by different mechanisms:

- Binding of activator proteins
- Binding of inhibitor proteins
- Ligand-controlled binding of regulatory subunits
- Activation by cofactors
  Example: activation of protein kinase C by diacylglycerol and $Ca^{2+}$
- Covalent modification by phosphorylation at Ser/Thr and/or Tyr residues
- Regulated subcellular localization
- Changes in the oligomerization state (see chapter 8)

Many structure-function investigations have shown that protein kinases can exist in an inhibited state in which structural elements in the active center fold back and inhibit substrate binding and/or catalytic activity. For protein kinases that are regulated by inhibitor proteins or inhibitory elements, the inhibition is also caused by a direct or indirect blocking of the active center by the bound inhibitor. On arrival of a signal (e.g., phosphorylation signal, binding of a messenger substance to the inhibitor), the enzyme is released from the inhibited state. In the process of activation, the inhibiting structural element vacates the active center or the inhibitor protein dissociates (Fig. 7.4).

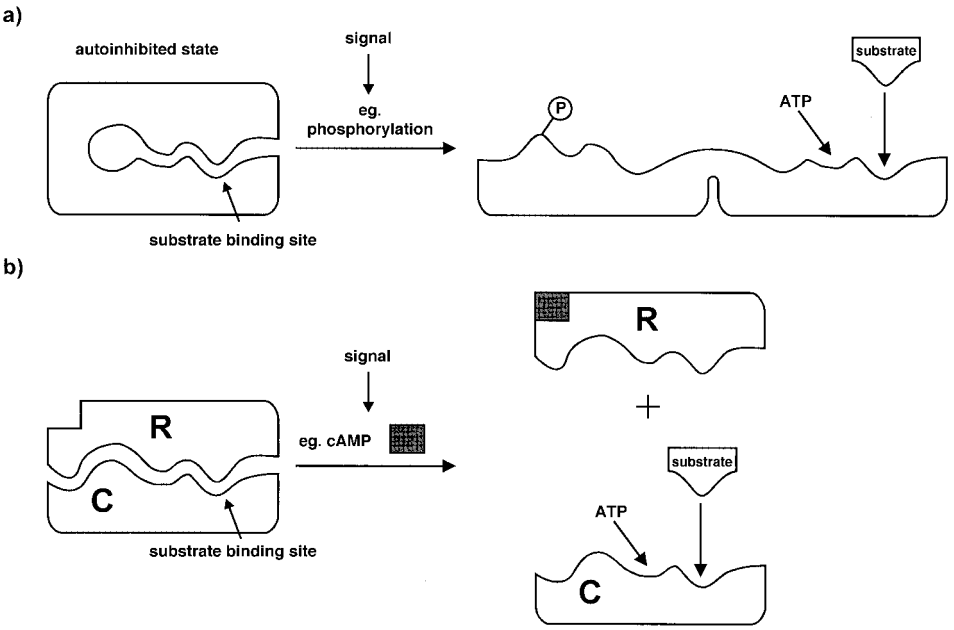



**Fig. 7.4.** Schematic representation of autoinhibition and activation of protein kinases. a) Release of autoinhibition by phosphorylation. b) The catalytic subunit C of a protein kinase, in an inhibited state due to binding of an inhibitory subunit R, can be released by an external signal and thereby activated.

This mechanism is known as *intrasteric regulation* since it involves regulation in the direct vicinity of the active center. Two mechanisms of intrasteric regulation are highlighted: regulation by autoinhibition and regulation by phosphorylation in the activation loop.

## Autoinhibition

If the inhibitory structural element is itself a part of the protein kinase, this is known as *autoinhibition*. The inhibitory structural elements often have the character of a pseudosubstrate. They possess a similarity to the proper substrate and can accumulate in the substrate binding site. Since the pseudosubstrate lacks a phosphate receptor, no phosphorylation takes place (review; Kemp et al., 1994).

Structural information on autoinhibition is available for the twitchin kinase. The twitchin kinase is a Ser/Thr-specific protein kinase of the nematode *Caenorhabditis elegans* and is homologous to the myosin light chain kinase of mammals (see 7.4.1). The crystal structure of a catalytic fragment of twitchin kinase (Hu et al., 1994) has an autoinhibitory element at the C-terminus, which makes specific contact with parts of the active site and the ATP binding site. The active site of twitchin kinase is blocked by the autoinhibitory structural element by:

– Forming contacts with the substrate binding site (pseudosubstrate function)
– Resulting interference with ATP binding
– Contacting essential residues and shielding these.

The autoinhibitory element is highly complementary to the active center and due to this complementarity, performs effective inhibition.

**Phosphorylation in the Activation egment**

Many protein kinases require phosphorylation of Ser/Thr or Tyr residues to reach full activity. The activating phosphorylation often takes place in a part of the structure in the vicinity of the active center, known as the *activation segment* which spans two conserved sequence motifs (DFG to APE, using single-letter code) present im almost all kinases (review: Johnson et al., 1996). Phosphorylation of the activation segment may be catalyzed by other protein kinases or by the protein kinase's own active center. In the latter case, this is generally an autophosphorylation in trans, i.e., between the subunits of an oligomeric protein kinase (see Chapter 8.1.3).

Ser/Thr (e.g. Thr197 of PKA) or Tyr phosphorylation sites (see insulin receptor, Chapter 8) are located at the activation segment. As shown for the phosphorylation of the CDK2-cyclin A complex (see 13.2.4), phosphorylation in the activation segment leads to reorganization of the catalytic center in the sense of an optimal orientation of the catalytic groupings (review: Johnson and O'Reilly, 1996).

## 7.2 Protein Kinase A

Of the protein kinases, protein kinase A is the best investigated and characterized (review: Francis and Corbin, 1994). The functions of protein kinase A are diverse. Protein kinase A is involved in the regulation of metabolism of glycogen, lipids and sugars. Substrates of protein kinase A may be other protein kinases, as well as enzymes of intermediary metabolism. Protein kinase A is also involved in cAMP-stimulated transcription of genes that have a cAMP-responsive element in their control region (review: Montminy, 1997). An increase in cAMP concentration leads to activation of protein kinase A which phosphorylates the transcription factor CREB at Ser 133. CREB only binds to the transcriptional coactivator CBP in the phosphorylated state and stimulates transcription (see Chapter 1.4.4.2).

### 7.2.1 Structure and Substrate Specificity of Protein Kinase A

The activity of protein kinase A is controlled by cAMP. In the absence of cyclic AMP, protein kinase A exists as a tetramer composed of two regulatory R subunits and two catalytic C subunits (see Fig. 6.2). The catalytic activity is masked in the holoenzyme $C_2R_2$, since an inhibitory structural element of the R subunit blocks the entrance to the

active site. Binding of cAMP to the R subunit leads to a reduction in the affinity between R and C by a factor of 10,000–100,000. The holoenzyme dissociates into the dimer of the R subunits and two monomers of C, that now become catalytically active.

In mammals, four isoforms of the R subunit (RIα, RIβ, RIIα and RIIβ) and three subtypes of the C subunit, namely C, Cβ and Cγ, are known.

The composition of the subunits is shown schematically in Fig. 7.5. The R subunit has two cAMP binding sites of differing affinity. In addition, the R subunit has a domain containing an autophosphorylation site which is involved in the autoinhibition of protein kinase A.

The C subunit has a myristinic acid residue at the amino terminus, the function of which is unknown. In addition, the C subunit has specific Ser/Thr phosphorylation sites, namely Thr197 and Ser338. Thr197 is located in the activation loop and is phosphorylated by an autophosphorylation mechanism. This is linked to an increase in the affinity for ATP and to the catalytic efficiency.

The *consensus sequence* for phosphorylation of proteins by protein kinase A is RRXSX. The RII subunit contains such a sequence in the autoinhibitory domain and is therefore subject to phosphorylation by the C subunit in the holoenzyme, but without release of inhibition. Inhibition of the C subunit by the R subunit is based on binding of the autoinhibitory sequence of R at the substrate binding site and at parts of the active center of the C subunit.

## 7.2.2 Regulation of Protein Kinase A

In addition to regulation by cAMP, protein kinase A is also subject to other regulatory influences. Thus, the C subunit may be specifically phosphorylated (see 7.2.1). It is not clear, however, which protein kinase is responsible for phosphorylation of the C subunit.

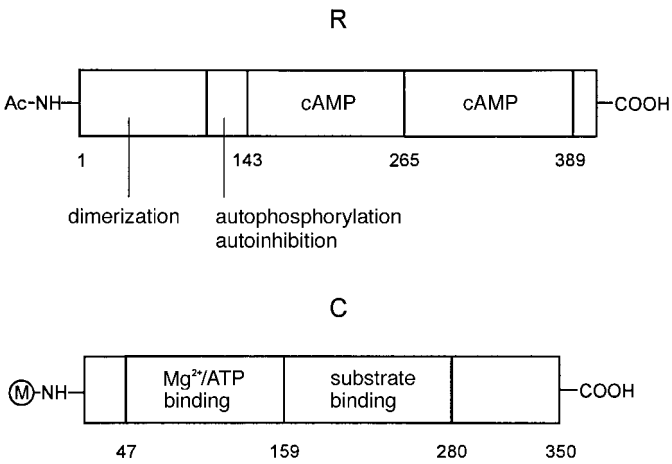



**Fig. 7.5.** Functional domains of protein kinase A. The functional domains of the catalytic (C) and regulatory (R) subunits of protein kinase A (bovine) are shown in a linear configuration. M: myristoylation.

Regulation of protein kinase A may also take place via its subcellular localization. Protein kinase A containing the RII subunit is found associated with the cytoskeleton and with the Golgi apparatus. Anchoring of protein kinase A to the subcellular compartments is mediated by specific proteins known as *protein kinase A anchor proteins* (AKAP: A kinase anchor protein). The RII subunit functions as a specific localization subunit by mediating the interaction of protein kinase A with the anchor proteins.

The changes in concentration of cAMP that lead to activation of protein kinase A in the cell are relatively small. In many tissues, a 2- to 3-fold increase in cAMP concentration is sufficient to bring about the maximum physiological effect. The cell has different mechanisms available that limit the increase in cAMP concentration to a relatively narrow concentration region and contribute to damping of signal transduction via protein kinase A. An example of a mechanism with a damping effect in signal transduction by protein kinase A is a *feedback control* by a cAMP phosphodiesterase. The activated protein kinase A phosphorylates and activates a phosphodiesterase that hydrolyses cAMP to AMP (Fig. 7.6). This mechanism enables protein kinase A to control its own steady state activity. It also ensures that, as the external signal diminishes, the cAMP signal rapidly subsides.

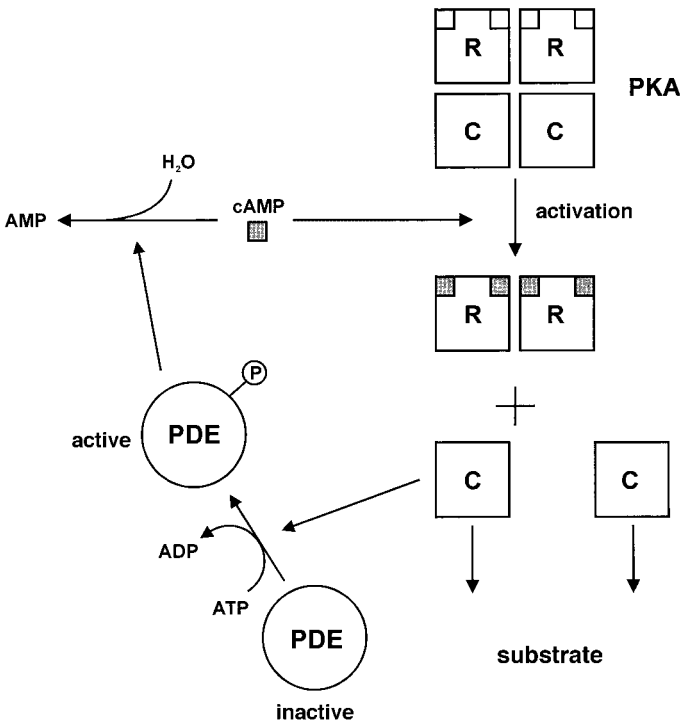



**Fig. 7.6.** Feedback control of protein kinase A by a phosphodiesterase. On activation of protein kinase A, the catalytic C subunits are released, which then phosphorylate a phosphodiesterase, in addition to other substrates. The phosphodiesterase is activated by the phosphorylation and hydrolyzes cAMP to AMP, whereby the signal transduction via protein kinase A is reduced or terminated.

# 7.3 Protein Kinase C

## 7.3.1 Characterization and Classification

The family of protein kinase C enzymes includes Ser/Thr-specific protein kinases that require the following cofactors for activation (review articles: Dekker and Parker, 1994; Newton, 1997, Oancea and Meyer, 1998):

- $Ca^{2+}$
- Phospholipids such as phosphatidylcholine
- Diacylglycerol

Regulation by $Ca^{2+}$ and diacylglycerol identify protein kinase C as a component of signal transduction pathways, in the course of which, phospholipase C is activated and the messenger substances $Ins(3,4,5)P_3/Ca^{2+}$ and diacylglycerol are produced. Activation of protein kinase C may take place via two central pathways:

Signaling pathways starting from receptor tyrosine kinases trigger stimulation of protein kinase C by activating phospholipase C$\gamma$. An activating signal may also be despatched in the direction of protein kinase C – via activation of phospholipase C$\beta$ – from G-protein-coupled membrane receptors (see Fig. 6.4).

### Stimulation by Phorbol Esters

A property of the protein kinase C enzyme family that is highly valuable for their identification and characterization is their activation by tumor promoters such as *phorbol esters* (Fig. 7.7). Protein kinase C binds to the tumor promoter, tetradecanoyl phorbol acetate (TPA), with high affinity and is activated by this binding. The specific activation of protein kinase C by phorbol esters is an important tool to demonstrate their involvement in signal transduction pathways. By external addition of TPA, it is possible to use cellular model systems to test which biological responses of a signal transduction pathway involve, and are controlled by, protein kinase C.

Tumor promoters such as TPA do not themselves initiate tumor formation, but rather they promote triggering of the tumor by carcinogenic substances, e.g., benzo[a]pyrene. At present, the mechanism by which the structurally very heterogeneous tumor promoters bring about their tumor promoting activity is not understood. In the case of TPA, stimulation of protein kinase C forms the basis of the explanation. Since one of the roles of protein kinase C is in regulation of proliferation and of differentiation processes, unregulated activation of protein kinase C could lead to undesired protein phosphorylation and thus bring about misregulation of cell proliferation.

### The Protein Kinase C Family

Like most of the Ser/Thr-specific protein kinase family, the protein kinase C family also shows significant heterogeneity. At the present time, at least 12 different subtypes of protein kinase C have been discovered in mammals, based on different criteria such

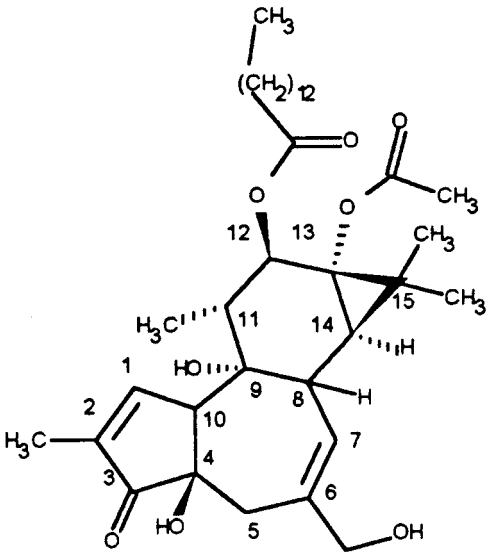


**Fig. 7.7.** Structure of tetradecanoyl phorbol acetate (TPA). Tetradecanoyl phorbol acetate functions as a tumor promoter and is a specific activator of protein kinase C.

as sequence, stimulation and regulation. These are known as α, βI, βII, γ, δ, ε, ζ, η, ι, λ and μ subtypes. Not all members of the protein kinase C family are stimulated by $Ca^{2+}$/diacylglycerol. The subtypes ζ and ι are not dependent on $Ca^{2+}$/diacylglycerol, but are clearly part of the protein kinase C family due to sequence homology.

The existence of the many subtypes of protein kinase C in mammals suggests that the individual subtypes perform specific functions in the organism (Hug and Sarre, 1993; Dekker and Parker, 1994). The different subtypes are distinguished by different cellular localization, different activation by cofactors, and a different pattern of substrate proteins. For example, the α, δ and ζ subtypes are widespread in almost all tissues whereas the other subtypes only occur in specialized tissues.

The members of the protein kinase C family are composed of a polypeptide chain with a molecular weight of 68–83 kDa. The N-terminal regulatory domains C1 and C2 and a C-terminal catalytic domain can de differentiated in the primary structure (Fig. 7.8). In addition, a pseudosubstrate sequence with autoinhibitory function is located at the N-terminus.

The C1 domain includes ca. 50 amino acids and contains a cysteine-rich motif with two bound $Zn^{2+}$ ions. In many isoenzymes, two copies of the C1 domain are present, known as C1A and C1B (or also as Cys1- and Cys2 elements). When expressed in the isolated form, both C1 domains can specifically bind phorbol esters, whereas *in vivo* only one of the two binding sites is occupied by phorbol esters.

In the C2 section, which is not present in all members of the protein kinase C family, the $Ca^{2+}$ binding site and a binding site for acidic phospholipids are found. The conserved C3 and C4 sections form the catalytic domain with the binding sites for ATP and substrate proteins.

Most of the biochemical information is available for protein kinase Cα. In the following, protein kinase C refers to the protein kinase Cα enzyme.

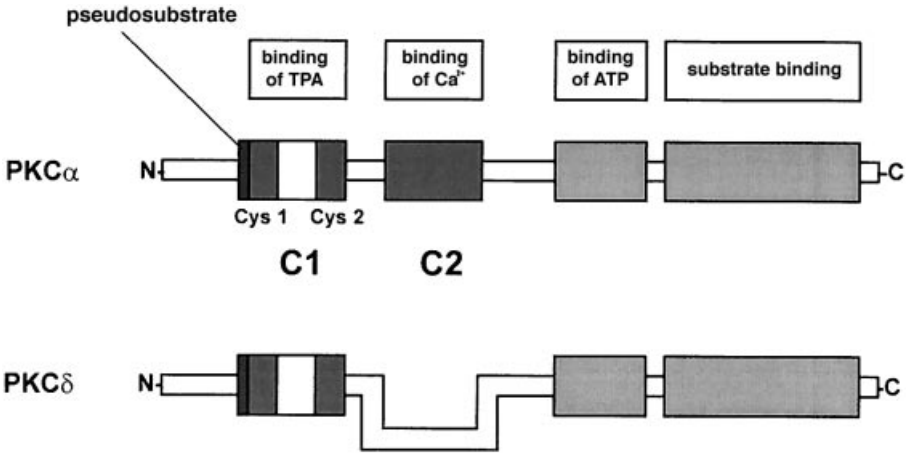



**Fig. 7.8.** Functional domains of protein kinase C. The functional domains of protein kinase Cα and Cδ are shown as a linear representation. The binding site for TPA lies in domain C1. Domain C2 contains the $Ca^{2+}$ binding site. Protein kinase Cδ lacks the C2 elements and thus regulation by $Ca^{2+}$. According to Azzi et al., (1992). Pseudosubstrate: autoinhibitory sequence with pseudosubstrate character.

### 7.3.2 Structure and Activation of Protein Kinase C

In the absence of activating cofactors, the catalytic domain is subject to *autoinhibition* by the regulatory domain (Orr and Newton, 1994). A sequence motif is found in the regulatory domain which serves as a pseudosubstrate. It resembles the consensus sequence for phosphorylation sites of protein kinase C but does not have a Ser or Thr residue for phosphorylation. This sequence motif is found in all protein kinase C family members. It is assumed that the active center is inhibited by occupation by the pseudosubstrate.

Two functions are attributed to the binding of the activating cofactors $Ca^{2+}$, diacylglycerol and phospholipid:

- Stabilization of a structure of protein kinase C in which the active center is accessible for substrate proteins
- Promotion of membrane association

Detailed structural information on protein kinase C is not available at present. The reason is probably the flexibility and membrane association of protein kinase C. The first insight into the mechanism of activation was obtained by structural determination of a C1 domain (Cys2 element) of protein kinase Cδ in complex with phorbol ester (Zang at al., 1995).

From the structure of the Cys2 element with bound phorbol ester (Fig. 7.9), it was concluded that the activating function of the phorbol ester is based, in particular, in promotion of membrane association of protein kinase C. The binding site of the phor-

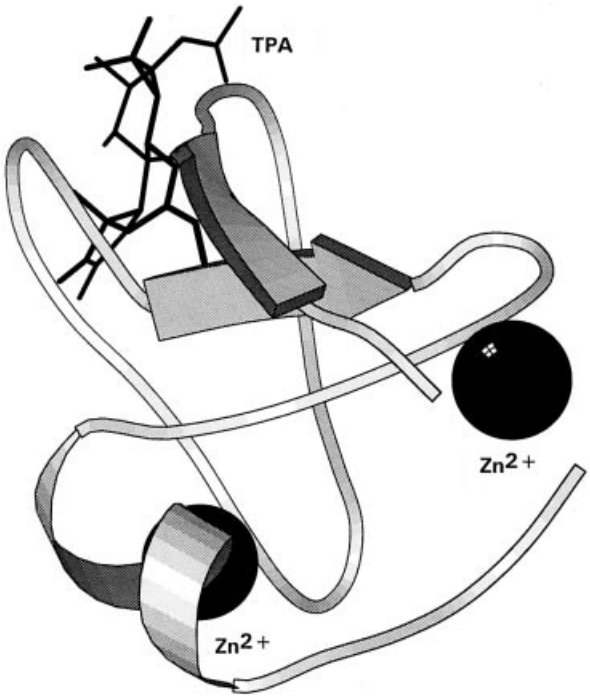



**Fig. 7.9.** Structure of the Cys2 element of protein kinase Cδ with bound phorbol ester. (Zang et al., 1995). MOLSKRIPT representation according to Kraulis, (1991).

bol ester lies in a hydrophobic region of the Cys2 element that is broken by a hydrophilic region. On binding of the phorbol ester, a continuous hydrophobic surface is created in this region of Cys2. It is assumed that the Cys2 section is involved in membrane association and that membrane association is energetically unfavorable in the absence of the phorbol ester. According to this theory, the phorbol ester binding enlarges the hydrophobic surface of the Cys2 element and favors a partial insertion of the Cys2 in the membrane.

Membrane association, and thus activation of protein kinase C, is controlled by the cofactors $Ca^{2+}$, diacylglycerol and phosphatidyl serine. A special role is assigned to the intensity and frequency of the $Ca^{2+}$ and diacylglycerol signals. Binding of $Ca^{2+}$ to the C2 domain leads to an increased membrane association and to activation by release of the catalytic center from interaction with the autoinhibitory structural element. Further activation takes place by binding of diacylglycerol to the C1 domain and binding of phosphatidyl serine to the C2 domain, whereby these ligands serve as an anchor for membrane association. Use of the two membrane-targeting domains C1 and C2 apparently helps to ensure high affinity, specificity and regulation of the membrane interaction.

The importance of coordination of the $Ca^{2+}$ and diacylglycerol signals is very nicely illustrated for protein kinase Cγ of brain (Oancea and Meyer, 1998). For persistent activation of protein kinase Cγ, it is necessary that high frequency $Ca^{2+}$ signals and a diacylglycerol signal are active simultaneously. Low frequency $Ca^{2+}$ signals, in contrast, only lead to low activation in the presence of diacylglycerol. Due to this property, the

function of a molecular device, able to decode $Ca^{2+}$ and diacylglycerol signals, is assigned to protein kinase Cγ.

### 7.3.3 Regulation of Activity of Protein Kinase C

Functions and regulation of protein kinase C are shown schematically in Fig. 7.10.

#### Regulation by Membrane Association

Many functions of protein kinase C in signaling pathways are closely linked with the membrane association of the enzyme. Activation of protein kinase C, initiated by addi-

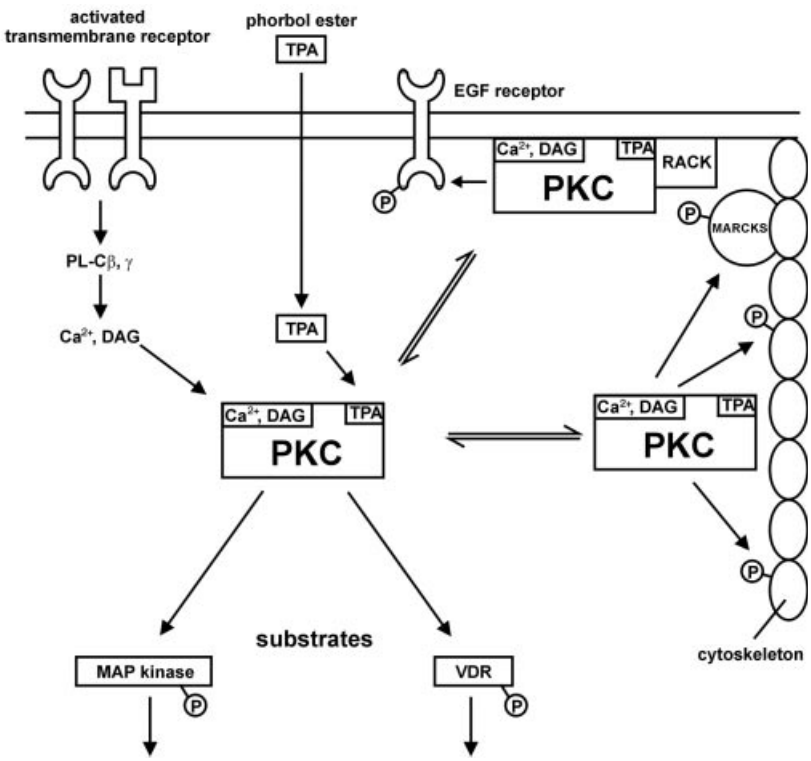



**Fig. 7.10.** Functions and regulation of protein kinase C. Receptor-controlled signal pathways lead to formation of the intracellular messenger substances $Ca^{2+}$ and diacylglycerol (DAG), that, like phorbol ester (TPA), activate protein kinase C (PKC). Translocation to the cell membrane is linked with activation of protein kinase C; receptors for protein kinase C, the RACK proteins, are also involved. Substrates of protein kinase C are the MARCKS proteins and other proteins associated with the cytoskeleton. Other substrates are the Raf kinase (see Chapter 10) and the receptor for vitamin $D_3$ (VDR, see Chapter 4).

tion of phorbol esters, for example, is associated with a redistribution of the enzyme from the cytosol to the membrane. An increase in cellular $Ca^{2+}$ is also sufficient for the redistribution, whereby $Ca^{2+}$ is directly involved in translocation of protein kinase C. An equilibrium seems to exist between cytosolic and membrane bound forms of protein kinase C; it can be assumed that this equilibrium can be shifted in the direction of membrane association with the help of the cofactors $Ca^{2+}$, phosphatidyl serine and diacylglycerol.

**Regulation by Localization**

Specific receptor proteins are also involved in the membrane association of protein kinase C; these are known as *receptors for activated protein kinase C* (*RACK proteins*). The RACK proteins are membrane associated proteins that prepare the binding site for protein kinase C at the membrane and thereby mediate specific membrane association of protein kinase C.

**Regulation by Phosphorylation**

Protein kinase C has three phosphorylation sites to which regulatory functions are assigned *in vivo*. One phosphorylation site (Thr500 in PKC βII) is located in the activation loop of the catalytic domain, the other ly in the vicinity of the C terminus. Activating phosphorylation in the catalytic domain is probably performed by another – as yet unknown – protein kinase. The two phosphorylations at the C terminus are autophosphorylations, to which an activating function is also assigned.

Following activation of signal transduction pathways that contain protein kinase C as a regulation element, a longer lasting stimulation of protein kinase C is observed in many cases that consequently results in long lasting biological effects. This *long term activation* of protein kinase C cannot be explained by release of $InsP_3$ and diacylglycerol within the bounds of activation of phospholipase C alone. Activation of protein kinase C is also observed when the phospholipase C signal and the associated $Ca^{2+}$ have died away.

Lasting stimulation of protein kinase C is associated with other pathways for formation of diacylglycerol. Following a transient increase and decrease of the diacylglycerol concentration due to activation of phospholipase C, a renewed, longer lasting increase in diacylglycerol is observed (Fig. 7.11). Other possibilities for production of diacylglycerol must be taken into consideration as the explanation for this increase (Liscovitch, 1992). In particular, phospholipids of the cell membrane may be an important source for release of diacylglycerol, whereby phosphatidyl choline is attributed a special role. Release of diacylglycerol from phosphatidyl choline may take place via phospholipases of type C or phospholipase D (see Fig. 6.11). Phospholipase D converts phosphatidyl choline to phosphatidic acid, from which diacylglycerol may be produced by removal of the phosphate residue. Activation of these phospholipases by $Ca^{2+}$ mobilizing agonists has often been described, although the underlying mechanisms are not explained.

Free fatty acids are also attributed a role in activation of protein kinase C. Free fatty acids, in particular arachidonic aicd, may be released by the activity of phospholipase A2 on phospholipids. Direct activation of protein kinase C by fatty acids has been des-

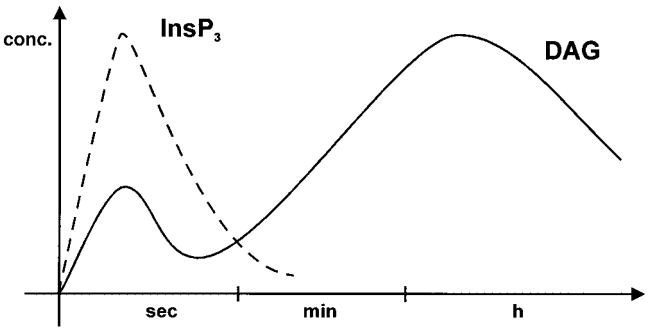



**Fig. 7.11.** Kinetics of formation of Ins(1,4,5)$P_3$ and diacylglycerol. The figure shows a model for the different dynamics of formation of Ins(1,4,5)$P_3$ and of diacylglycerol (DAG), observed as a consequence of hormonal stimulation in an idealized cell. An extracellular stimulus causes activation of the PtdIns specific phospholipase C (PL-Cβ or PL-Cγ) on a sec timescale, and thus formation of Ins(1,4,5)$P_3$ and DAG, and release of $Ca^{2+}$ (not shown). The renewed increase in concentration of DAG is caused by the activation of phosphatidyl choline specific phospholipases of type C and phospholipase of type D. According to Liscovitch, (1992).

cribed many times, as has stimulation of phospholipase A2 following activation of G-protein-coupled signal transduction pathways (review: Liscovitch and Cantely, 1994). The details of this activation are still unknown, however.

### 7.3.4 Functions of Protein Kinase C

The members of the protein kinase C family are central signal proteins and as such, are involved in the regulation of a multitude of cellular processes. A problem in the identification of substrates of protein kinase C is its low substrate specificity which often cannot be differentiated from that of protein kinase A, particularly in *in vitro* experiments. The consensus sequence of the phosphorylation sites in substrate proteins are similar to those of protein kinase A, in that basic amino acids are required in the neighborhood of the Ser/Thr residue to be phosphorylated. The following consensus sequences may be formulated for phosphorylation by protein kinase C: (* = phosphorylation site) S*/T*XK/R; K/RXXS*/T*; K/RXXS*/T*XK/R; K/RXS*/T*; K/RXS*/T*XK/R (Pearson and Kemp, 1991).

Of the many substrates of protein kinase C, the MARCKS proteins are highlighted as very well characterized and specific substrates of protein kinase C (review: Aderem, 1995). The abbreviation MARCKS stands for myristoylated, alanine-rich C-kinase substrate.

The MARCKS proteins are a family of proteins that are involved in physiologically important processes such as cell mobility, secretion, membrane transport and in regulation of the cell cycle. All these processes are associated with changes and restructuring of the actin cytoskeleton. The role of converting extracellular signals into changes in the structure of the actin cytoskeleton is attributed to the MARCKS proteins. A

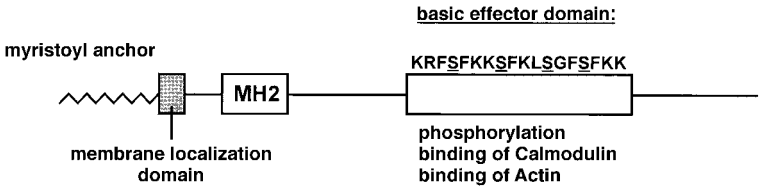



**Fig. 7.12.** Functional domains or the MARCKS proteins. Linear representation of the characteristic domains of the MARCKS proteins. The Ser phosphorylation sites in the effector domain are underlined. The function of the MH2 domain is unknown.

protein kinase C mediated phosphorylation of the MARCKS proteins is involved in this regulation. These are widespread and specific substrates of protein kinase C and their phosphorylation is used as an indicator of activation of protein kinase C *in vivo*.

The MARCKS proteins are acidic proteins with a high content of the amino acids Ala, Gly, Pro and Glu. An N-terminal domain carries a lipid anchor in the form of myristinic acid, from which it is assumed that it mediates the association with the membrane. A basic effector domain is important for regulation of the MARCKS proteins; a binding site for $Ca^{2+}$/calmodulin and the phosphorylation site for protein kinase C are located in this domain (see Fig. 7.12).

In the unphosphorylated form and in the absence of $Ca^{2+}$, the MARCKS proteins bind to actin filaments and bring about crosslinking of the latter. Binding of $Ca^{2+}$/calmodulin or phosphorylation by protein kinase C inhibits the crosslinking activity. The MARCKS proteins can thus modulate the aggregation status of actin filaments and function as effectors for the conversion of extracellular signals that are carried into the cell via G-protein-coupled receptors and/or tyrosine kinase receptors.

Further examples of substrates of protein kinase C are the epidermal growth factor receptor (see Chapter 8), a $Na^+/H^+$ exchanger protein, and Raf kinase (Chapter 9). Activation of protein kinase C may, as the examples show, act on other central signal transduction pathways of the cell; it may have a regulating activity on transcription processes and it is involved in the regulation of transport processes. Many substrates of protein kinase C are membrane proteins and it is evident that membrane association of protein kinase C is of great importance for the phosphorylation of these proteins.

## 7.4 $Ca^{2+}$/calmodulin Dependent Protein Kinases

### 7.4.1 Importance and General Function

The signal-mediating function of $Ca^{2+}$ is performed as a $Ca^{2+}$/calmodulin complex in many signaling pathways. $Ca^{2+}$/calmodulin can bind specifically to effector proteins and modulate their activity. In first place as effector proteins of $Ca^{2+}$/calmodulin are the *$Ca^{2+}$/calmodulin protein kinases* (*CaM kinases*) (review: Braun and Schulman, 1995). The CaM kinases are widespread and are found in practically all cells of mam-

mals. Like other target proteins of intracellular messengers, the CaM kinases also show great heterogeneity caused by the existence of distinct genes and alternative splicing.

A rough categorization of the CaM kinases differentiates between *specialized* CaM kinases and *multifunctional* kinases.

An example of a specialized CaM kinase is myosin light chain kinase (MLCK), the primary function of which is to phosphorylate the light chain of myosin and thus to control the contraction of smooth musculature.

The multifunctional CaM kinases are collectively referred to as CaM kinases of type II, whereby further subtypes α, β, γ and δ are differentiated. The α and β subtypes of CaM kinase II only occur in the brain whereas the other subtypes are also found in other organs. The multifunctional CaM kinases regulate many processes (see Table 7.1) such as glycogen metabolism, activity of transcription factors, microfilament formation, synaptic release of neurotransmitters from storage vesicles, biosynthesis of neurotransmitters and many more. An important cellular function is assigned to CaM kinase II in brain, where it makes up 0.25 % of the total protein.

## 7.4.2 Structure and Autoregulation of CaM Kinase II

From a regulatory point of view, CaM kinase II is of particular interest as it has the characteristic of an enzyme with a built-in „memory switch". The „memory" allows the CaM kinase to conserve a stimulatory signal over a longer period of time and to remain in an activated state, even when the initiating stimulus has died away.

CaM kinase is regulated by both autophosphorylation and by $Ca^{2+}$/calmodulin. An N-terminal catalytic domain, a regulatory domain and an association domain can be

**Table 7.1.** Examples of substrates of $Ca^{2+}$/calmodulin dependent protein kinases (source: Ann. Rev. Physiol. (1995), 57.

| protein | function |
|---|---|
| AcetylCoA carboxylase | Biosynthesis of fatty acids |
| Glycogen synthase | Glycogen synthesis |
| HMGCoA reductase | Biosynthesis of cholesterol |
| NO synthase | Biosynthesis of NO |
| $Ca^{2+}$ channel (N-type) | Presynaptic $Ca^{2+}$ influx |
| $Ca^{2+}$ ATPase (heart) | Storage of $Ca^{2+}$ |
| Synaptogamin | Release of neurotransmitters |
| Ryanodin receptor | Release of $Ca^{2+}$ |
| p56 Lck tyrosine kinase | Activation of T cells |
| EGF receptor | Growth control |
| Cyclic nucleotide phosphodiesterase | cAMP and cGMP metabolism |
| Phospholipase A2 | Hydrolysis of phospholipids |
| Ribosomal protein S6 | Protein biosynthesis |
| CRE binding protein | Transcription control |

differentiated in the structure of CaM kinase. Electron microscopic pictures show that CaM kinase II has an oligomeric structure (Kanaseki et al., 1991) in which 8—12 copies of the monomeric enzyme (α, β, γ or δ subtype) are configured in the form of a cylinder with the catalytic domains orientated outwards (Fig. 7.13). The aggregated form of CaM kinase may occur as the homooligomer or as a heteroligomer composed of different subtypes.

In the absence of $Ca^{2+}$/calmodulin, the catalytic domain of CaM kinase exists in an autoinhibited state. The active center is blocked by occupation by a pseudosubstrate sequence of the regulatory domain. An increase in the $Ca^{2+}$ concentration and the associated $Ca^{2+}$ signal lead to binding of $Ca^{2+}$/calmodulin at the C-terminal end of the regulatory domain, which releases the enzyme from its inhibited state (Fig. 7.14). The enzyme is now activated and can perform an autophosphorylation in the autoinhibitory sequence. The phosphorylation takes place at a conserved Thr residue (Thr286 of the α subtype) and is intermolecular, i.e., neighboring subunits of the holoenzyme mutually phosphorylate one another.

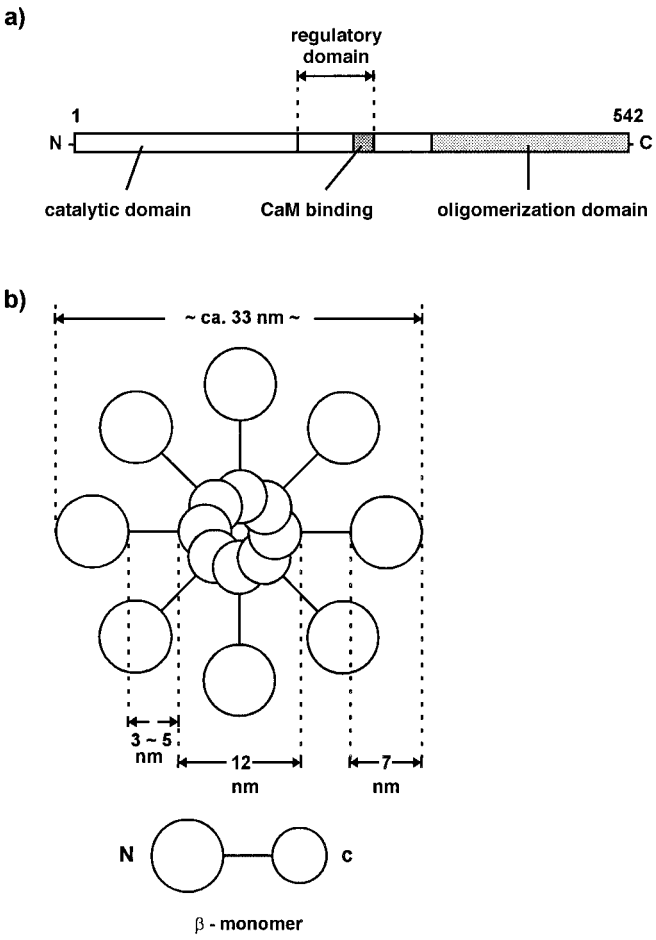



**Fig. 7.13.** Primary structure and oligomeric structure of CaM kinase II of type β. a) Linear representation of the functional domain of CaM kinase IIβ. b) The oligomeric structure shown is proposed for an octamer of type β, based on electron microscopic investigations (Kanaseki et al., 1991). The *N*-terminal catalytic domain is represented as a larger circle, the *C*-terminal oligomerization domain by a smaller circle. CaM: calmodulin.

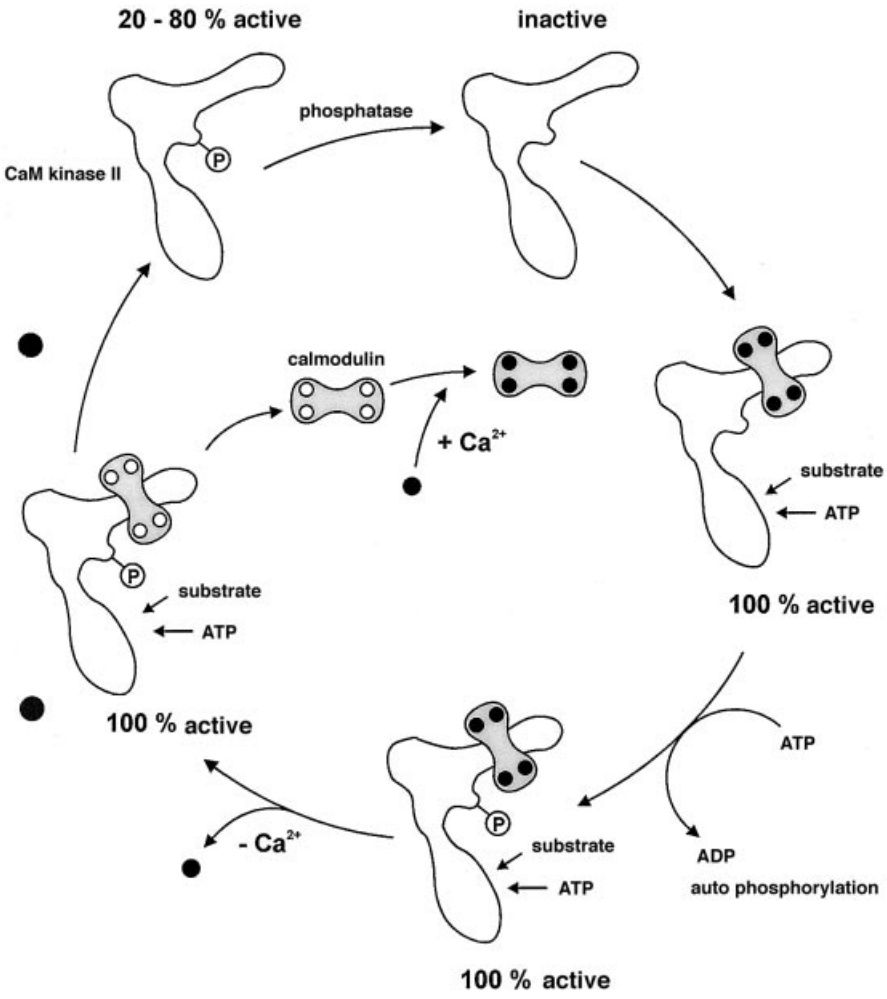



**Fig. 7.14.** Regulation of CaM kinase II. Scheme of regulation of CaM kinase II by $Ca^{2+}$/calmodulin and by autophosphorylation. CaM kinase II is inactive in the unphosphorylated form and in the absence of $Ca^{2+}$/calmodulin. Binding of $Ca^{2+}$/calmodulin activates the kinase for phosphorylation of protein substrates. In the process, autophosphorylation takes place at a conserved Thr residue that stabilizes the active state of the enzyme. In this state, significant residual activity is still present after dissociation of $Ca^{2+}$/calmodulin and the enzyme remains in an active state for a longer time after the $Ca^{2+}$ signal has died away. The active state is only terminated when the activating phosphate residue is cleaved off by a protein phosphatase.

The autophosphorylation has two important consequences:

The affinity for $Ca^{2+}$/calmodulin is increased by close to three orders of magnitude. $Ca^{2+}$/calmodulin only dissociates very slowly from this high affinity complex. The activated state is thus preserved over a longer period of time. Even when the $Ca^{2+}$ signal has died away and the $Ca^{2+}$ concentration has fallen to a level of $10^{-7}$M, the enzyme

remains in the activated state for several more seconds, since calmodulin can remain bound to the enzyme without bound $Ca^{2+}$.

After dissociation of calmodulin, the phosphorylated enzyme still has 20—80 % of the activity of the $Ca^{2+}$/calmodulin bound form. This ensures that significant activity remains after the $Ca^{2+}$/calmodulin signal has died away. In the phosphorylated form, CaM kinase is in an autonomous, $Ca^{2+}$/calmodulin independent state. This is only terminated when phosphatases cleave off the activating phosphate residue and thus lead the enzyme back into the inactive state.

The special feature of regulation of CaM kinase is the memory effect within the activation process. Activation of the enzyme is initiated by a generally transient increase in cellular $Ca^{2+}$. $Ca^{2+}$ activates CaM kinase in the form of the $Ca^{2+}$/calmodulin complex; the kinase remains active even after the $Ca^{2+}$ signal has died away, because the enzyme is converted into an autonomous activated state during autophosphorylation. A special importance is attributed to this property, particularly for detection and differentiation of repetitive $Ca^{2+}$ signals. The interval between the occurrence of staggered $Ca^{2+}$ signals is a determining factor for the intensity of activation. If the $Ca^{2+}$ signals occur with a higher frequency, a long lasting and effective activation is possible since the kinase remains in the activated state between signals, due to the memory effect (De Koninck and Schulman, 1998). Due to this special property, it is assumed that CaM kinase actively participates in memory formation. In agreement with this is the observation that the memory capability of transgenic animals is influenced by expression of CaM kinase ( Mayford et al., 1996).

There are different mechanisms that lead to increased intracellular $Ca^{2+}$ concentration and thus to activation of CaM kinase (see 6.5 and Fig. 6.6). CaM kinases are activated as a consequence of $InsP_3$-mediated release of $Ca^{2+}$ from intracellular storage. Influx of $Ca^{2+}$ from the extracellular region, triggered by opening of various ligand-controlled or voltage-controlled $Ca^{2+}$ channels, also brings about an activation of CaM kinases. Thus, CaM kinase has a key function in several signal transduction pathways. Some important substrates of CaM kinases are shown in Table 7.1.

## 7.5 Ser/Thr-specific Protein Phosphatases

Under physiological conditions, phosphate esters of Ser and Thr residues are stable and only show a low rate of spontaneous hydrolysis. Thus, the cell requires its own tools for regulated cleavage of phosphate residues, to terminate and damp signals mediated by protein phosphorylation. This role is performed by specific protein phosphatases.

### 7.5.1 Structure and Classification of Ser/Thr Protein Phosphatases

We know of Ser/Thr phosphate specific protein phosphatases (referred to in the following as Ser/Thr phosphatases) and Tyr phosphate specific protein phosphatases. The latter are dealt with in Chapter 8.

In mammals, at least four groups of Ser/Thr phosphatases can be differentiated; the members of these groups are known as *protein phosphatases (PP) 1, 2A, 2B and 2C*. PP-1, PP-2A and PP-2B are highly homologous with respect to the sequence of the catalytic domain but they differ in substrate specificity and type of regulation.

Protein phosphatase 2B is also known as *calcineurin*. Calcineurin is made up of a catalytic A subunit and a regulatory B subunit that shows similarity to calmodulin. A binding site for $Ca^{2+}$/calmodulin, required for activation of calcineurin, is found in the catalytic subunit.

The protein phosphatase calcineurin was of particular interest since it mediates the immunosuppressive effect of the pharmaceuticals cyclosporin and FK506, often used in organ and tissue transplantations. The biochemical point of application of both pharmaceuticals was unclear for a long time. In initial experiments, it was found that cyclosporin and FK506 bind specifically to two proteins known as cyclophilin and FK506 binding protein, respectively. Both proteins function as peptidyl prolyl cis/trans isomerases (review: Fischer, 1994).

The immunosuppressive effect of cyclosporin and FK506 could not initially be explained by these observations. Only with the discovery that cyclosporin and FK506 achieve their immunosuppressive effect via inhibition of calcineurin did it become clear that the immunosuppression is mediated by a complex reaction chain involving calcineurin. It was shown that the complexes of cyclosporin/cyclophilin and FK506/FK506 binding protein bind to calcineurin and inhibit the phosphatase activity of the latter.

Calcineurin is part of a signaling pathway that is activated by a rise in intracellular calcium upon ligand binding to a cell membrane receptor. The rise in $Ca^{2+}$ activates calcineurin's phosphatase activity, which dephosphorylates cytoplasmic NF-AT transcription factor family members. Dephosphorylated NF-AT-members enter the nucleus and bind to DNA in cooperation with other transcription factors, e.g. AP-1. By this way many target genes in diverse tissues can be activated.
Transcription of interleukin 2 is inhibited as one of the consequences of inhibition of calcineurin; as an extracellular signal, interleukin 2 can stimulate proliferation of lymphocytes. This discovery made it clear that the protein phosphatase calcineurin has an essential role in signal transduction processes in T lymphocytes (Fig. 7.15).

The Ser/Thr phosphatases exist mostly as heterodimers composed of a catalytic subunit and another subunit to which a specific function for localization of protein phosphatases is often attributed. In the case of protein phosphatase I, this subunit is known as the *targeting subunit*; its function is described in more detail in 7.7.

A crystal structure is available of the catalytic subunit of protein phosphatase I (Goldberg et al., 1995). The enzyme contains two metal ions in the active center, which are probably $Mn^{2+}$. Both metal ions are attributed a function in catalysis and also in substrate binding.

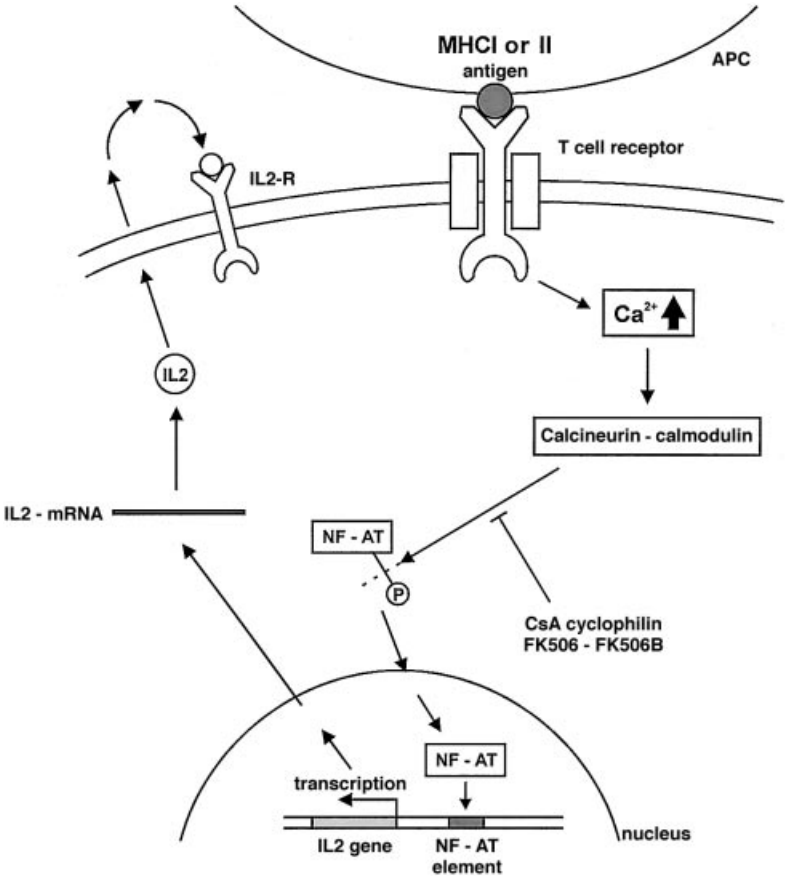



**Fig. 7.15.** Model of the function of calcineurin in T lymphocytes. Antigenic peptides are presented to the T lymphocytes by an antigen-presenting cell (APC) within a cell-cell interaction (see also Chapter 11). Antigen binding activates the T cell receptor that starts a signal chain leading to an increase in cytosolic $Ca^{2+}$ and activation of calcineurin. The activated calcineurin cleaves an inhibitory phosphate residue from the transcription factor NF-AT. Consequently, NF-AT is transported into the nucleus where it stimulates the transcription of corresponding genes. Amongst the genes controlled by NF-AT is the gene for the cytokine interleukin 2 (IL-2). Following secretion into the extracellular space, the IL-2 so formed binds to IL-2 receptors of the same cell or cells of the same type. A proliferation signal is created by the activated IL-2 receptor, leading to proliferation of T lymphocytes. Complexes of the immunosuppressants cyclosporin A (CsA) or FK506 with their binding proteins cyclophilin and FK506 binding protein (FK506B), respectively, inhibit calcineurin and disrupt the signal transmission to NF-AT.

## 7.5.2 Function and Regulation of Ser/Thr-specific Protein Phosphatases

Protein phosphatases are the antagonists of protein kinases. They perform a dual function. They can have a damping effect on protein kinase mediated signal transduction, by diminishing and terminating a signal created by protein phosphorylation. Protein phosphatases can also have a positive, reinforcing effect in signaling pathways. Dephosphorylation of a signal protein by a protein phosphatase can lead to its activation and thus to amplification of the signal (Fig. 7.16).

Due to these functions, the protein phosphatases are an indispensable part of signal transduction processes involving protein phosphorylation. It is therefore not surprising that the protein phosphates are subject to diverse and complex regulation.

Regulation of the Ser/Thr phosphatases takes place predominantly by three mechanisms.

### Targeted Localization

With the help of the localization subunit, a protein phosphatase can be directed to distinct subcellular locations at which the substrates of the protein phosphatase are also localized (see 7.7).

### Specific Inhibitor Proteins

Specific inhibitor proteins for Ser/Thr phosphatases exist which can control the activity of the protein phosphatases. These inhibitors are generally subject to regulation themselves, e.g., by phosphorylation (Fig. 7.17).

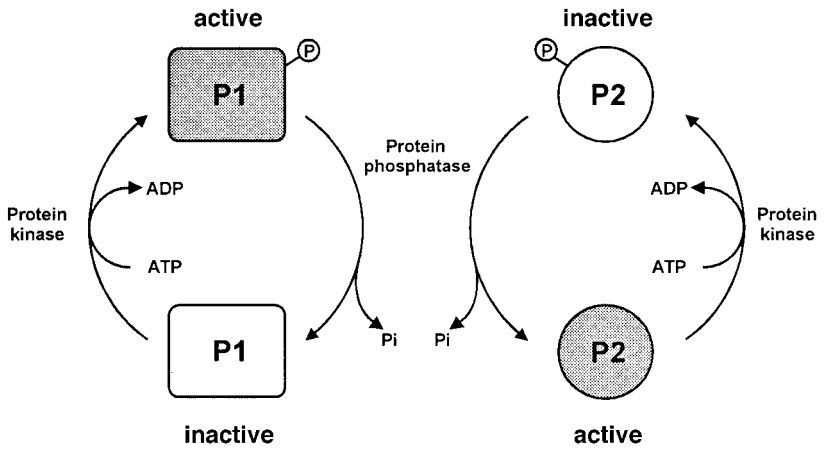



**Fig. 7.16.** The dual function of protein kinases and protein phosphatases. Phosphorylation of proteins (P1, P2) can fix the latter into an active or inactive state. In the case of P1, protein kinases have an activating effect and protein phosphatases are inactivating; the reverse is true for P2.

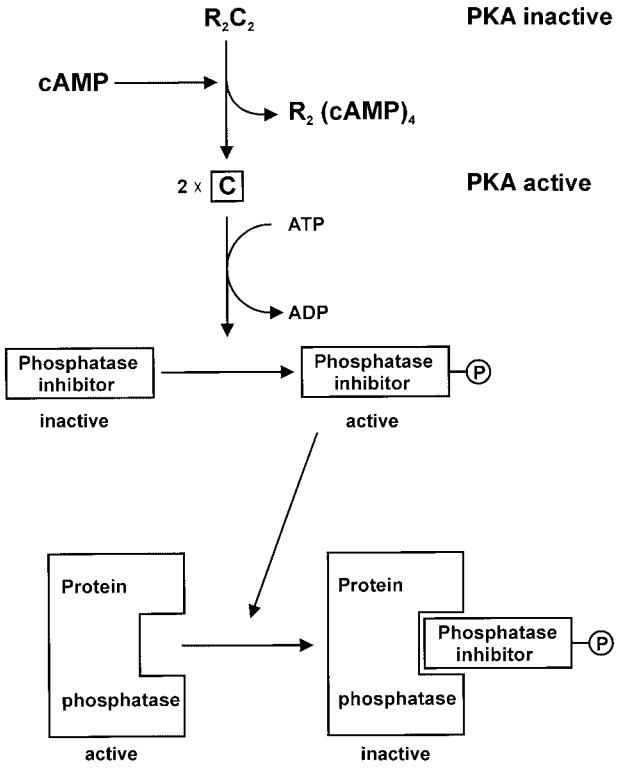



**Fig. 7.17.** Regulation of protein phosphatases by inhibitor proteins. The substrates of protein kinase A include protein phosphatase inhibitors that are phosphorylated by the C subunit of protein kinase A. In the phosphorylated state, the protein phosphatase inhibitors bind to the protein phosphatase and inhibit its enzyme activity

**Phosphorylation**

Protein phosphatase I has a phosphorylation site at the C terminus. Phosphorylation at this site by a cyclin-dependent protein kinase inhibits the phosphatase activity.

## 7.6 Coordinated Action of Protein Kinases and Protein Phosphatases

The current state of Ser/Thr phosphorylation of a protein is determined by the relative activity of Ser/Thr-specific protein kinase and protein phosphatase. It is therefore understandable that the cell has had to develop special mechanisms to balance the two activities with one another, and, when needed, to allow kinase or phosphatase activity to dominate. One of the best investigated examples of coordinated activity of protein kinases and protein phosphatases is the regulation of glycogen metabolism in skeletal muscle. Glycogen metabolism is an example of how two different signals, namely a cAMP signal and a $Ca^{2+}$ signal meet in one metabolic pathway and control the activity of one and the same enzyme.

### 7.6.1 Protein Phosphorylation and Regulation of Glycogen Metabolism

Degradation and synthesis of glycogen is subject to control via $Ca^{2+}$ and via adrenaline (Fig. 7.18). The $Ca^{2+}$ control comes into effect predominantly in contraction of muscle, whereas the hormonal regulation occurs in the resting state of the muscle. The hormonal regulation takes place via a signal transduction pathway, the components of which were described in previous chapters. Following binding of adrenaline to the β-adrenergic receptor, activation of adenylyl cyclase occurs, mediated by a stimulatory G-protein, and thus the cAMP level increases. cAMP activates protein kinase A, which then phosphorylates phosphorylase kinase and activates it. In a further step of this activation cascade, glycogen phosphorylase is phosphorylated and activated by phosphorylase kinase.

Glycogen synthase is regulated via the same pathway. Glycogen synthase is inactive in the phosphorylated form whereas in the dephosphorylated form, it is active. Three key enzymes of glycogen metabolism are thus controlled with the help of reversible protein phosphorylation.

$Ca^{2+}$ brings about its regulatory effect at the level of phosphorylase kinase. Phosphorylase kinase is activated by $Ca^{2+}$. $Ca^{2+}$ influx comes into play particularly in contraction of muscle, a process by which $Ca^{2+}$ is released from storage and activates phosphorylase kinase.

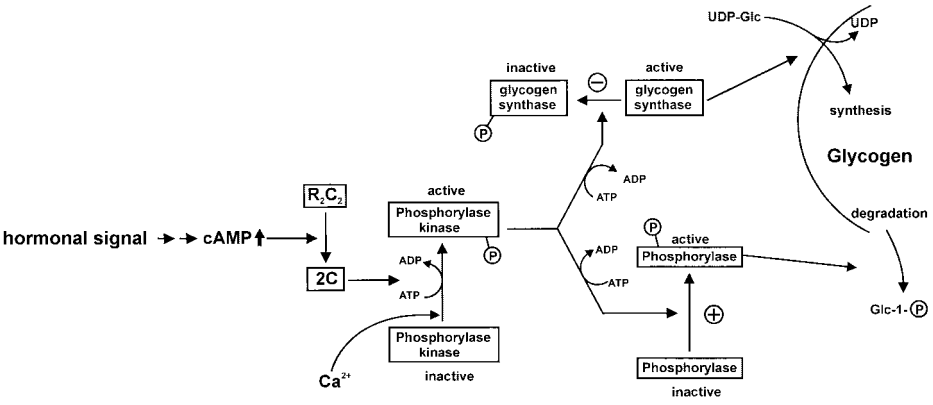



**Fig. 7.18.** Regulation of glycogen metabolism in muscle. Phosphorylase kinase stands at the center of regulation of glycogen metabolism. Phosphorylase kinase may exist in an active, phosphorylated form and an inactive, unphosphorylated form. Phosphorylation of phosphorylase kinase is triggered by hormonal signals (e.g. adrenaline) and takes place via an activation of protein kinase A in the cAMP pathway. In the absence of hormonal stimulation, phosphorylase kinase can also be activated by an increase in cytosolic $Ca^{2+}$. The active phosphorylase kinase stimulates glycogen degradation and inhibits glycogen synthesis, in that, on the one side, it activates glycogen phosphorylase by phosphorylation, and on the other side, it inactivates glycogen synthase by phosphorylation.

At the level of the phosphorylase kinase, two signals meet: a hormonal signal in the form of adrenaline and the $Ca^{2+}$ signal (review: Cohen, 1992). The structural basis for this dual regulation of phosphorylase kinase is its subunit structure (Fig. 7.19). Phosphorylase kinase is composed of four subunits: two regulatory α and β subunits, a catalytic γ subunit and calmodulin as δ subunit. In the dephosphorylated state, phosphorylase kinase is inactive, since the catalytic γ subunit is inhibited by interaction with the other subunits. On increase in $Ca^{2+}$ concentration, the inhibition is partially lifted in that $Ca^{2+}$ binds to calmodulin. A further contribution to release of inhibition is made by the contractile protein troponin C, which enters into an interaction with the inhibitory β subunit. In this state, phosphorylase kinase is active and can phosphorylate glycogen phosphorylase. Glycogen degradation begins and finally ATP is made available for muscle contraction.

In the resting state, i.e., under conditions of low $Ca^{2+}$ concentration, phosphorylase kinase may be activated by a hormonal signal, in that phosphorylation by protein kinase A takes place at the regulatory α and β subunits. In the phosphorylated form, the kinase also needs $Ca^{2+}$; however, the phosphorylated enzyme can be activated by appreciably lower $Ca^{2+}$ concentrations than the unphosphorylated enzyme. Due to the higher sensitivity towards $Ca^{2+}$, the phosphorylated enzyme can still be active at the very low $Ca^{2+}$ concentrations that correspond to the intracellular $Ca^{2+}$ concentration in the resting state of the muscle. Thus, hormonal stimulation of glycogen degradation is possible in the resting state.

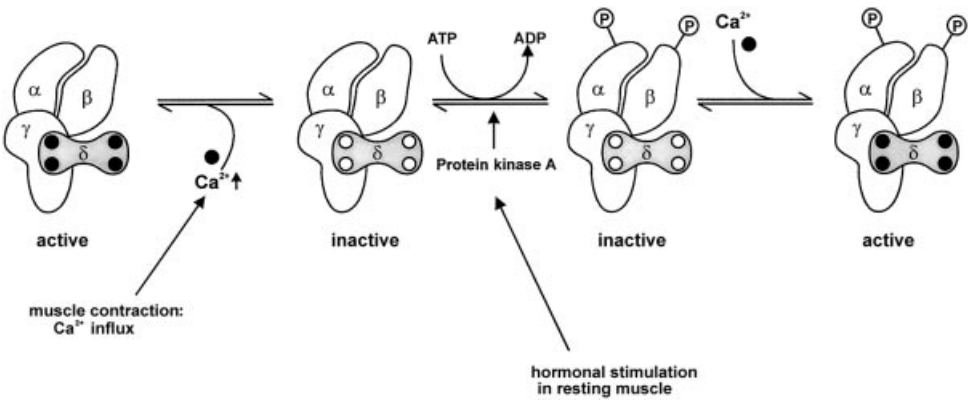



**Fig. 7.19.** Subunit structure and regulation of phosphorylase kinase of muscle. Phosphorylase kinase is – according to the excitation state of the muscle – regulated by two pathways. On nervous stimulation of the muscle, voltage-controlled $Ca^{2+}$ channels are opened, the cytosolic $Ca^{2+}$ concentration increases and $Ca^{2+}$ binds to calmodulin, activating phosphorylase kinase. In resting muscle, activation of phosphorylase kinase is triggered by a hormonal signal. A hormonal signal initiates phosphorylation of the α and β subunits of phosphorylase kinase. In the phosphorylated form, $Ca^{2+}$ binding affinity of the calmodulin subunit (δ) is strongly increased and activation is also possible at low $Ca^{2+}$ concentrations.

## 7.6.2 Protein Phosphatase I and Regulation of Glycogen Metabolism

Protein phosphatase I also has an important role in regulation of glycogen degradation (review: Cohen, 1992; Hubbard and Cohen, 1993). The hormonal signals leading to activation of protein kinase A and phosphorylase kinase simultaneously initiate inactivation of protein phosphatase I. The resultant inhibition of dephosphorylation and stimulation of phosphorylation of the key enzyme of glycogen degradation ensures that rapid mobilization of the glycogen reserves of muscle can follow a hormonal signal.

The basis of regulation of protein phosphatase I is the ability of the catalytic subunit of the enzyme to reversibly associate with corresponding regulatory proteins. As a result, the activity and specificity of the enzyme are modulated.

### Glycogen-associated Proteins, G Subunit of Protein Phosphatase I

The catalytic subunit of protein phosphatase I is associated with a glycogen binding protein, also known as the *G subunit of protein phosphatase I* (Fig. 7.20). The G subunit is tightly bound to glycogen. Association of the catalytic subunit with the G subunit creates a form of protein phosphatase I known as *protein phosphatase I-G*

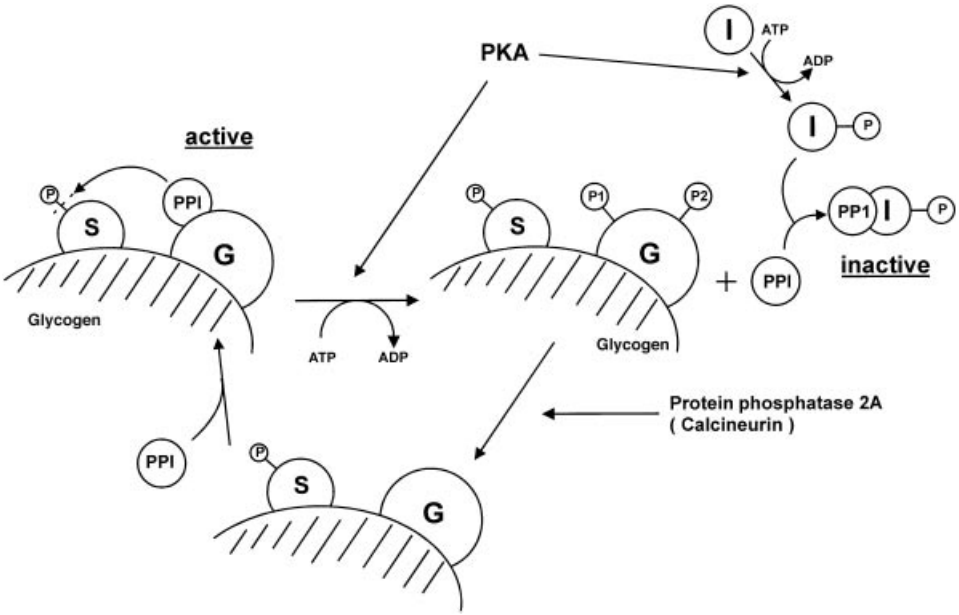



**Fig. 7.20.** Regulation of glycogen-bound protein phosphatase I. Regulation of the activity of protein phosphatase I (PPI) takes place by phosphorylation of the G subunit. The G subunit is phosphorylated at positions P1 and P2, in the process of a signal chain mediated activation of protein kinase A. As a consequence of the phosphorylation, the catalytic subunit dissociates. The phosphatase activity of the free catalytic subunit is inhibited by association with a cytosolic protein phosphatase inhibitor (I), the binding of which is also controlled via a protein kinase A mediated phosphorylation. The phosphorylated G subunit can be dephosphorylated again by protein phosphatase 2A and may bind a catalytic PPI subunit once more.

(*PPIG*). PPIG is a highly active form of the protein phosphatase in which the G subunit functions to mediate a targeted localization of the catalytic subunit to glycogen, so that the enzyme comes into the direct vicinity of its substrate. The substrates are the phosphorylated forms of glycogen phosphorylase and phosphorylase kinase, both of which are associated with glycogen.

Regulation of the activity of protein phosphatase I is also mediated by the G subunit. This may be phosphorylated at two positions P1 and P2 localized in the phosphoregulatory domain of the G subunit. Phosphorylation at P1 modulates the substrate preference of protein phosphatase I, in that the activity towards glycogen synthase and phosphorylase kinase is increased. The P1 phosphorylation thus has a stimulatory effect on dephosphorylation of important substrates of protein phosphatase I. Phosphorylation at P2 has an opposing effect, bringing about a reduction in affinity for the catalytic subunit by a factor of about $10^4$ and thus leading to dissociation of the catalytic subunit from glycogen. Via this mechanism, phosphorylation at P2 brings about an inhibition of protein phosphatase activity at glycogen. The dephosphorylation and inactivation of glycogen phosphorylase is hindered by phosphorylation at the P2 position of the G subunit.

Hormonal stimulation of the cell and the associated activation of protein kinase A lead to phosphorylation at the P1 and P2 sites of protein phosphatase I. Consequently, the enzyme dissociates from the G subunit and can no longer dephosphorylate the glycogen-bound substrate. PPI is released into the cytosol but still retains activity in this form towards substrates not associated with glycogen.

Renewed docking of the catalytic subunit requires the removal of the phosphate residue at the G subunit phosphorylated at the P2 site. This takes place via the protein phosphatases 2A and 2B (calcineurin). Thus, a cascade of protein phosphatases is involved in the regulation of dephosphorylation of key enzymes of glycogen degradation, whereby a phosphatase, namely protein phosphatase I, is indirectly activated by other protein phosphatases. With calcineurin, a $Ca^{2+}$-dependent protein phosphatase is involved and thus it is possible to influence glycogen metabolism via $Ca^{2+}$-mediated signals.

Phosphorylation at P1 also delivers an explanation for the influence of insulin on glycogen metabolism. Insulin stimulates glycogen synthesis and inhibits glycogen degradation by initiating dephosphorylation and activation of glycogen synthase and dephosphorylation and inhibition of glycogen phosphorylase. Both enzymes are dephosphorylated by the protein phosphatase PPIG. Insulin activates protein phosphatase PPIG by causing phosphorylation of the G subunit at the P1 site, but not at the P2 site, via an insulin-stimulated protein kinase (Fig. 7.21).

### Inhibitors of Protein Phosphatase I

Protein phosphatase I that has dissociated from glycogen may be inactivated by association with inhibitor proteins, preventing undesired dephosphorylation of proteins in the cytosol. The activity of inhibitor proteins may in turn be controlled by reversible phosphorylation. A hormone-induced activation of protein kinase A leads to phosphorylation of inhibitor protein I; the phosphorylated form is the active form of the inhibitor. This mechanism ensures that stimulation of protein phosphorylation mediated by cAMP and protein kinase A is not weakened by an opposing dephosphorylation (Fig. 7.17).

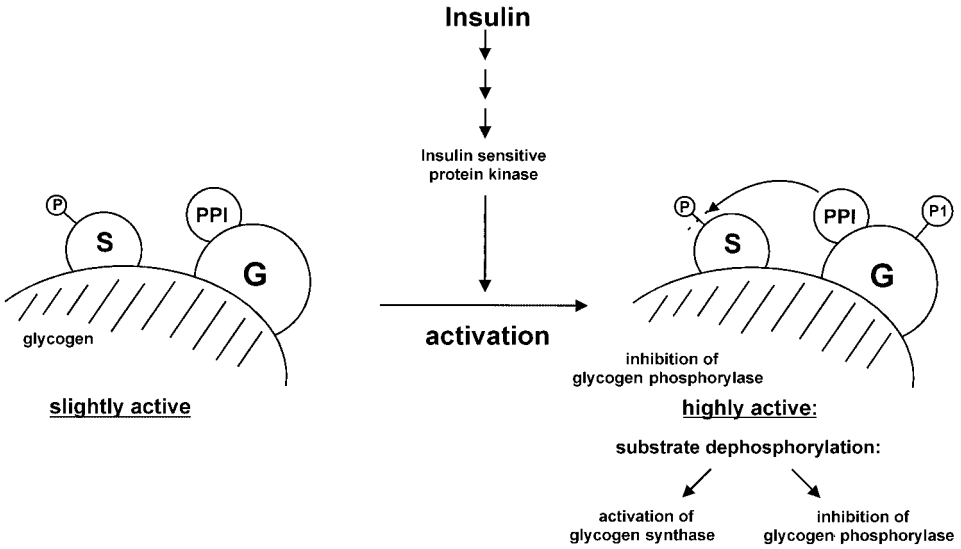



**Fig. 7.21.** Activation of glycogen-bound protein phosphatase I by insulin. Insulin has a stimulating effect on glycogen synthesis by initiating the dephosphorylation and activation of glycogen synthase and the dephosphorylation and inhibition of glycogen phosphorylase. Both enzymes (substrate S in the figure) are dephosphorylated by protein phosphatase PPIG. Insulin mediates the activation of a protein kinase (insulin-sensitive protein kinase) within an insulin-stimulated signal pathway, which phosphorylates and thus activates protein phosphatase PPIG at the P1 site.

## 7.7 Regulation of Protein Phosphorylation by Specific Localization at Subcellular Structures

The extent and specificity of the reactions of protein kinases and protein phosphatases are extremely dependent on the degree to which substrate and enzyme are localized at the same place in the cell. Many substrates of protein kinases occur either as membrane associated or particle associated forms (see 7.6.1, enzymes of glycogen metabolism). For protein kinases or protein phosphatases to perform their physiological function in a signal transduction process, they must be transported to the location of their substrate in many cases (review: Hubbard and Cohen, 1992; Mochly-Rosen, 1995). This is valid both for the Ser/Thr-specific protein kinases as well as for many Tyr-specific protein kinases. In the course of activation of signal transduction pathways, compartmentalization of protein kinases, redistributed to new subcellular locations, is often observed.

Translocation of protein kinases is a specific process, encompassing at least two mechanisms to decide the location in the cell at which the kinase will become active. In one mechanism, sequence sections of the protein kinase are used as leader sequences for compartmentalization.

In another mechanism, an associated subunit of the protein kinase or protein phosphatase determines in which compartment of the cell and at which membrane section

the protein phosphorylation signal will become active. The subunit functions as a *localization subunit*; it determines at which place in the cell the protein kinase gains access to its substrates. With the help of the localization subunit, protein kinase and substrate are brought into the close vicinity of one another and thus rapid conversion of substrate is achieved, due to the high local concentration of the enzyme.

The principle of targeted localization is shown in Fig. 7.22. In addition to the binding site for the corresponding protein kinase (or protein phosphatase), the localization subunit also has a specific binding site for an *anchor protein*, found at a subcellular site in the region where protein phosphorylation should take place. Through the interaction of anchor protein and localization subunit, the catalytic subunit is fixed at the desired location and is able to preferentially convert substrate localized at the same location.

Among the protein kinases already discussed, there are some for which the function is linked with specific subcellular localization to a high degree:

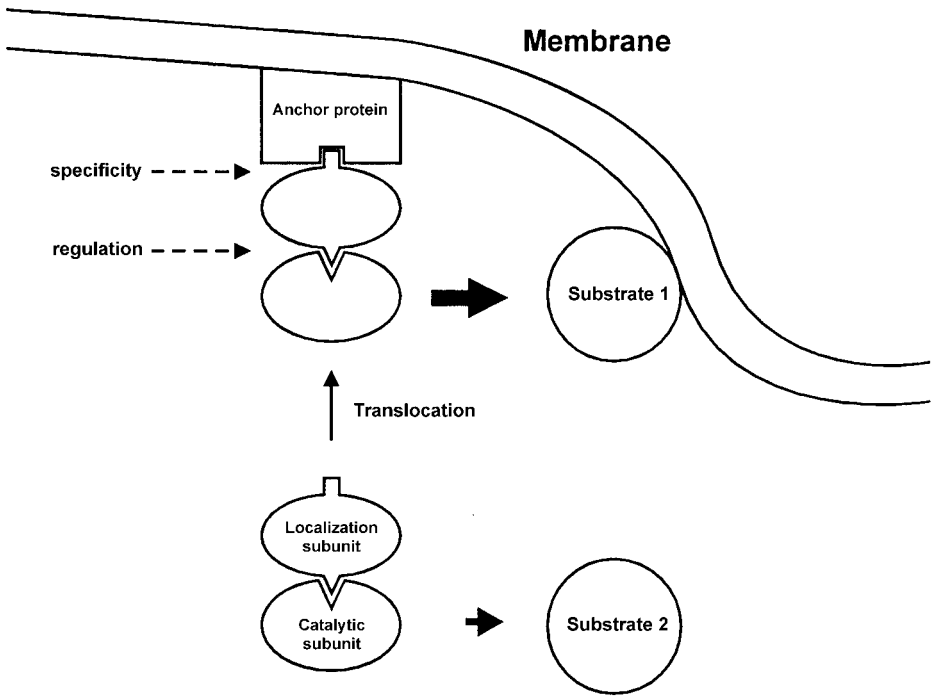



**Fig. 7.22.** The principle of targeted localization of protein kinases and protein phosphatases. The spatial configuration between the catalytic subunit of a protein kinase or protein phosphatase and a membrane-associated substrate is mediated by localization subunits that specifically bind to membrane-localized anchor proteins. The specificity of co-localization is predominantly achieved at the level of binding of the localization subunit to the anchor protein. The co-localization is regulated, in particular, by the interaction of the catalytic subunit with the localization subunit. In the membrane-associated form, the catalytic subunit has increased activity towards membrane-bound substrates.

### Protein Kinase A

An increase in cAMP and activation of protein kinase A are accompanied, in many cases, by a change in the subcellular location of the protein kinase A holoenzyme and the catalytic subunit. In particular, protein kinase A in cells of the nerve system is often found associated with the cytoskeleton. This association is mediated by the RII subunit. Binding of cAMP to the regulatory subunit releases the catalytic subunit that can phosphorylate substrates in the near vicinity. The released catalytic subunit can also be transferred to other compartments of the cell. Parallel to the increase in cAMP, translocation of the catalytic subunit is observed in many cells from the Golgi apparatus to the nucleus via the cytosol, and is accompanied by stimulation of transcription.

The A kinase anchor proteins (AKAP), which are tightly associated with the cytoskeleton, serve as anchors for protein kinase A. In addition to the RII subunit of protein kinase A, the AKAP proteins also bring protein phosphatases and other protein kinases to the cytoskeleton in a targeted fashion. The AKAP79 protein binds protein kinase C and protein phosphatase 2B (calcineurin) as well as the RII subunit. The possibility to bring both a protein kinase and a protein phosphatase to the same place in the cell opens up the prospect of a coordinated and layered regulation of both enzyme activities.

By bringing together various protein kinases and protein phosphatases, anchoring proteins organize signal transduction events and can create localized and efficient signal events at specific subcellular sites.

### β-adrenergic Receptor Kinase (βARK)

βARK is responsible for switching off and desensitizing activated β-adrenergic receptors (see 5.3.4 and Fig. 5.8). On binding of an agonist to the β-adrenergic receptor, translocation of βARK from the cytosol to the membrane is observed. The βγ subunits of the G-protein involved in signal transduction serve as membrane anchors. On receptor activation, the βγ subunits are released and bind specifically to βARK. Since the βγ subunits are associated with the membrane via a lipid anchor (see 5.5.4.7), binding of the βγ complex to βARK is linked to translocation of βARK to the membrane.

### Protein Kinase C

In many cases, stimulation of cells with phorbol esters or with hormones that activate phospholipase Cβ or phospholipase Cγ leads to translocation of protein kinase C (or a subtype of protein kinase C) from the cytoplasm to the cell membrane, cytoskeleton or into the nucleus. The differential localization of the various protein kinase C isoforms seems to be mediated by PKC targeting proteins, of which various classes have been identified (review: Pawson and Scott, 1997). The RACK proteins (see 7.7.4) belong to a class of targeting proteins that mediate anchoring of PKC to the membrane. The RACK proteins bind to a structural section of protein kinase C that is not identical to the catalytic center or the substrate binding site. Protein kinase C can thus be brought to the membrane in a targeted manner, without any major intervention in catalytic activity. Another class of proteins, termed substrate binding proteins (SBP), bind PKC

in the presence of the substrate by forming a ternary complex with the kinase. Phosphorylation of the SBPs by protein kinase C abolishes the targeting interaction and disrupts the ternary complex.

**Protein Phosphatase I**

The G subunit of protein phosphatase I, occurring in the glycogen-bound form, is considered as its localization subunit. The G subunit enables targeted localization of the catalytic subunit of protein phosphatase I to glycogen so that a close spatial orientation of protein phosphatase and its substrates, the enzymes of glycogen metabolism, is created (see 7.6.2).

## 7.8 General Principles of Regulation of Enzymes by Phosphorylation and Dephosphorylation

The examples of phosphorylase kinase and protein phosphatase I illustrate some important principles of regulation of enzyme activity by phosphorylation and dephosphorylation events. They clearly indicate how different signal transduction paths can meet in key reactions of metabolism, how signals can be coordinated with one another and how common components of a regulation network can be activated by different signals. The following principles are highlighted:

Protein phosphorylation may be used to *switch enzyme activities on and off.* The same is true for dephosphorylation of enzymes. In the cell, we often find phosphorylation cascades in which several phosphorylation reactions are connected to one another. Dephosphorylation reactions may also be elements of these cascades. Furthermore, networks of phosphorylation and dephosphorylation of proteins exist which permit the cell to respond to external stimuli in a finely tuned way.

A protein kinase or protein phosphatase may be regulated by different signal transduction pathways. Thus, different external stimuli may influence the phosphorylation status of a protein. This differential stimulation may be mediated by the subunits of the enzyme, for example. For phosphorylase kinase, a $Ca^{2+}$ signal is registered by the $Ca^{2+}$/calmodulin subunit whereas a cAMP protein kinase A signal is picked up in the form of a phosphorylation of the $\alpha$ and $\beta$ subunits. Which of the signals comes into play depends on the current metabolic situation.

Differential signals may also be registered and differentiated by different phosphorylation sites of a protein. Due to the substrate preference of protein kinases, multiple phosphorylation sites present in a substrate protein may be phosphorylated by different protein kinases. The individual phosphorylation sites may have distinct functions in regulation of the enzyme activity so that a specific regulation of the enzyme is possible.

A phosphorylation cascade in which several protein kinases are coupled in sequence can be modulated by various signal transduction pathways, according to this mechanism. Good examples of the importance of multiple phosphorylation sites are the protein kinases that control passage through the cell cycle (see Chapter 13).

The *subcellular localization* of the kinase or phosphatase also plays a crucial role for the activity of protein kinases and protein phosphatases. Many physiological functions of protein kinases and protein phosphatases depend on the enzyme being brought, with the help of specific protein-protein interactions, to certain subcellular locations in the vicinity of its substrate.

# References Chapter 7

Aderem, A. 'The MARCKS family of protein kinase C substrates' (1995) Biochem. Soc. Trans. *23*, 587–591

Bossemeyer, D. 'The glycine-rich sequence of protein kinases: a multifunctional element' (1994) Trends Biochem. Sci. *19*, 201–205

Braun, A.P. and Schulman, H. 'The multifunctional calcium/calmodulin dependent protein kinase: From form to function' (1995) Annu. Rev. Physiol. *57*, 417–445

Cohen, P. 'Signal integration at the level of proteinkinases, protein phosphatases and their substrates' (1992) Trends Bioch. Sci. *17*, 408–413

De Koninck, P. and Schulman, H. 'Sensitivity of CaM kinase II to the frequency of Ca2+ oscillations' (1998) Science *279*, 227–230

Erikson, R.L., Colett, M.S., Erikson, E. and Purchio, A.F. 'Evidence that the avian sarcoma virus transforming gene product is a cyclic AMP-independent protein kinase' (1979) Proc. Natl.Acad.Sci. U.S.A. *76*, 6260–6264

Hubbard, M.J. and Cohen, P. 'On target with a new mechanism for the regulation of protein phosphorylation' (1993) Trends Bioch. Sci. *18*, 172–177

Dekker, L.V. and Parker, P.J. 'Protein kinase C: a question of specificity' (1994) Trends Biochem. Sci. *19*, 73–77

Fischer, G. 'Über Peptidyl-Prolyl-cis/trans-Isomerasen and ihre Effektoren' (1994) Angew.Chem. *106*, 1479–1501

Francis, S.H. and Corbin, J.D. 'Structure and function of cyclic nucleotide-dependent protein kinases' (1994) Ann. Rev. Physiol. *56*, 237–272

Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A. and Kuriyan, J. 'Three-dimensional structure of the catalytic subunit of protein serin/threonine phosphatase-1' (1995) Nature *376*, 745–753

Goldsmith, E.J. and Cobb, M.H. 'Protein kinases' (1994) Curr. Biol. *4*, 833–840

Hu, S.H., Parker, M.W., Lei, J.Y., Wilce, C.J., Benian, G.M. and Kemp, B.E. 'Insights into autoregulation from the crystal structure of twitchin kinase' (1994) Nature *369*, 581–584

Hug, H and Sarre, T.F. 'Proteinkinase C isoenzymes: divergence in signal transduction' (1993) Biochem. J. *291*, 329–343

Hunter, T. 'Protein kinase classification' (1991) Methods in Enzym. *200*, 3–37

Kanaseki, T., Ikeuchi, Y., Sugiura, H. and Yamauchi, T. 'Structural features of $Ca^{2+}$/ Calmodulin-dependent protein kinase II revealed by electron microscopy' (1991) J. Cell Biol. *115*, 1049–1060

Kemp, B.E., Parker, M.W., Hu, S., Tiganis, T. and House, C. 'Substrate and pseudosubstrate interactions with protein kinases: determinants of specificity' (1994) Trends Bioch. Sci. *19*, 441–448

Knighton, D.R., Zheng, J., Ten Eyck, L.F., Ashford, V.A., Xuong, N., Taylor, S.S. and Sowadski,J .M. 'Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase' (1991) Science *253*, 407–413

Kolesnick, R. and Golde, D.W. 'The sphingomyelin pathway in tumor necrosis factor and interleukin-1 signaling' (1994) Cell *77*, 325–328

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Krebs, E.G., Graves, D.J. and Fischer, E.H. 'Factors affecting the activity of muscle phosphorylase kinase' (1959) J. Biol. Chem. *234*,2867–2873

Liscovitch, M. 'Crosstalk among multiple signal-activated phospholipases' (1992) Trends Biochem. Sci. *17*, 393–400

Liscovitch, M. and Cantley, L.C. 'Lipid second messengers' (1994) Cell *77*, 329–334

Johnson L.N. and O'Reilly M. 'Control by phosphorylation' (1996) Curr Opin Struct Biol. *6*, 762–769

Johnson, L.N., Noble, M.E., Owen, D.J. 'Active and inactive protein kinases: structural basis for regulation' (1996) Cell *85*, 149–158

Johnson, L.N., Lowe, E.D., Noble, M.E.and Owen, D.J., The Eleventh Datta Lecture. The structural basis for substrate recognition and control by protein kinases' (1998) FEBS Lett *430*, 1–11

Mayford, M., Bach, M.E., Huang, Y.Y., Wang, L., Hawkins, R.D. and Kandel, E.R. 'Control of memory formation through regulated expression of a CaMKII transgene' (1996) Science *274*, 1678–1683

Mochly-Rosen, D. 'Localization of protein kinases by anchoring proteins: A theme in signal transduction' (1995) Science *268*, 247–251

Montminy, M. 'Transcriptional regulation by cAMP' (1997) Ann. Rev. Biochem. *66*, 807–822

Newton, A.C. ' Regulation of protein kinase C' (1997) Curr. Op. Cell Biol. *9*, 161–167

Nishizuka, Y. 'Intracellular signalling by hydrolysis of phospholipids and activation of protein kinase C' (1992) Science *258*, 607–614

Oancea, E. and Meyer, T. 'Protein kinase C as a molecular machine for decoding calcium and diacylglycerol signals' (1998) Cell *95*, 307–318

Orr, J.W. and Newton, A.C. 'Intrapeptide regulation of protein kinase C' (1994) J. Biol. Chem. *269*, 8383–8387

Pawson, T., Scott, J.D. 'Signaling through scaffold, anchoring, and adaptor proteins' (1997) Science 278, 2075–2080

Pearson, R.B. and Kemp, B.E. 'Protein kinase phosphorylation site sequences and consensus specificity motifs; tabulations' (1991) Methods in Enzymol. *200*, 62–81

Zang, G., Kazanietz, M.G., Blumberg, P.M. and Hurley, J.H. 'Crystal structure of the cys2 activator-binding domain of protein kinase C delta in complex with phorbol ester' (1995) Cell *81*, 917–924

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 8**

# Signal Transmission via Transmembrane Receptors with Tyrosine-specific Protein Kinase Activity

Tyrosine-specific protein kinases, like Ser/Thr-specific protein kinases, are widely employed as tools of intracellular signal conduction. Activation of tyrosine-specific protein kinases is triggered, in particular, by signals that control cell growth and differentiation. Extracellular signals are often protein hormones which – if they have a regulating influence on cell proliferation – are also classed as growth factors. Table 8.1 shows a selection of mammalian growth factors, together with the corresponding receptors that conduct the signal into the cell interior.

Coupling of an extracellular signal to tyrosine phosphorylation in the interior of the cell can take place by two means and involves two different types of receptor (Fig. 8.1):

**Receptors with Intrinsic Tyrosine Kinase Activity**

Some transmembrane receptors possess intrinsic tyrosine kinase activity. These receptors are known as *receptor tyrosine kinases.* Ligand binding to an extracellular domain of the receptor is coupled to the stimulation of tyrosine kinase activity localized on a cytoplasmic receptor domain. The ligand binding domain and the tyrosine kinase domain are part of one and the same protein.

**Transmembrane Receptors with Associated Tyrosine Kinase Activity**

Another type of transmembrane receptor is associated, on the cytoplasmic side, with a tyrosine kinase that is activated when a ligand binds to the extracellular receptor domain (see Chapter 11). The tyrosine kinase and the receptor are not located on the same protein in this case.

## 8.1 Structure and Function of Receptor Tyrosine Kinases

Receptor tyrosine kinases possess binding sites at the surface of the cell membrane that are specific for extracellular ligands. Ligand binding to the receptor activates a tyrosine-specific protein kinase activity of the receptor, located on the cytoplasmic domain. Consequently, tyrosine phosphorylation is initiated, at the receptor itself and also on substrate proteins; these in turn trigger the biological response of the cell by

**Tabelle 8.1.** Selected growth factors and growth factor receptors from mammals.

| Growth factor | Characteristics | Receptors |
|---|---|---|
| Platelet derived growth factor, PDGF, types AA, AB and BB | Dimers, A (17 kD )- and B (16kD) chains, B chain is product of *c-sis* proto-oncogene | 2 types of receptor tyrosine kinases, PDGF-Rα (170kD), PDGF-Rβ (180kD) |
| Epidermal growth factor, EGF<br>Transforming growth factor-α, TGF-α | ca 6 kD, EGF und TGF-α are up to 40 % identical | Receptor tyrosine kinase. EGF-R is a produkt of the c-erbB proto-oncogene |
| Transforming growth factor-β, TGF-β1,-β2, –β3 | Homodimer of 25 kD | TGFβ receptor I und II, contains Ser/Thr specific protein kinase activity |
| Insulin-like growth factor, IGF-1 and IGF-2 | 7kD, related to proinsulin | Receptor tyrosine kinase, IGF-R |
| Fibroblast growth factor, FGF-1, FGF-2, FGF-3, FGF-4. | related proteins of 16–32 kD | Receptor tyrosine kinase |
| Granulocyte colony stimulating factor, G-CSF | 24 kD | 150 kD, G-CSF-R, receptor with associated tyrosine kinase |
| Granulocyte macrophage-colony stimulating factor, GM-CSF | 14 kD | 51 kD, GM-CSF-R, receptor with associated tyrosine kinase |
| Interleukins 1–7, IL-1 – IL-7, Interleukin 9, 12, 15, IL-9, IL-12, IL-15 | | IL-R-1 – I IL-R-7, IL-R-9, IL-R-12, IL-R-15, receptors with associated tyrosine kinase |
| Interleukin 8, IL-8 | | IL-R-8, G-protein coupled receptor |
| Erythropoietin | | Epo-R, receptor with associated tyrosine kinase |
| Tumor necrosis factor, TNF | | TNF-R, receptor with associated tyrosine kinase |
| Leukemia inhibitory factor, LIF | | LIF-R, receptor with associated tyrosine kinase |
| Interferon α, β, γ | | INF-Rα, INF-Rβ, INF-Rγ, receptors with associated tyrosine kinase |

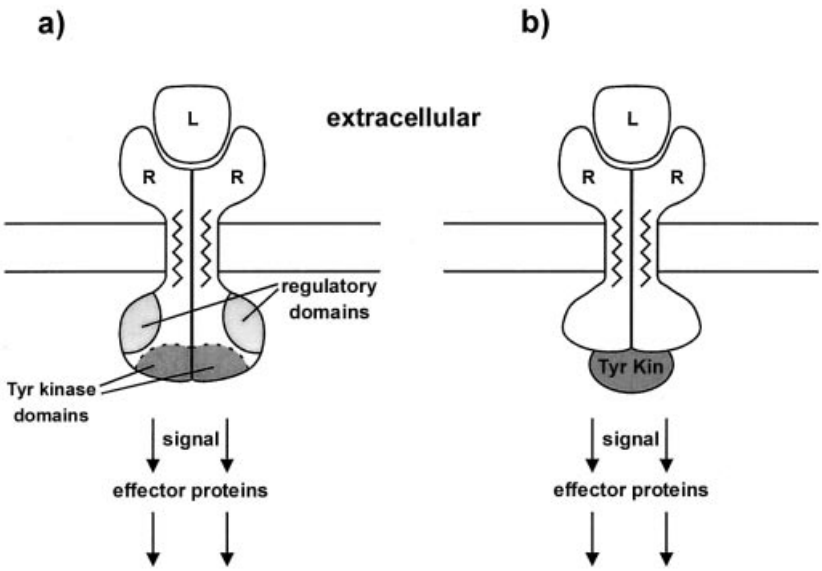



**Fig. 8.1.** Scheme of signal transmission by receptors with intrinsic and associated tyrosine kinase activity. a) Tyrosine kinase receptors possess a tyrosine kinase domain in the cytoplasmic region. Binding of a ligand L to the extracellular domain of the receptor produces a signal on the cytoplasmic side by activating the tyrosine kinase. Regulatory sequence segments are also located on the cytosolic side. b) Receptors with associated tyrosine activity pass the signal on to a tyrosine kinase that is not an intrinsic part of the receptor, but is permanently or transiently associated with the cytoplasmic receptor domain. The receptor shown has been simplified as a dimer.

switching on a further chain of reactions. The response can reach as far as the cell nucleus, where transcription of particular genes is activated. It can also affect the reorganization of the cytoskeleton, cell-cell interactions and reactions of intermediary metabolism. In particular, the receptor tyrosine kinases regulate cell division activity, differentiation and cell morphogenesis by this mechanism.

### 8.1.1 General Structure and Classification

Receptor tyrosine kinases are integral membrane proteins that have a ligand-binding domain on the extracellular side and a tyrosine kinase domain on the cytosolic side (see Fig. 8.1). The transmembrane portion is made up of just one structural element; thus it is assumed that it crosses the membrane in an α-helical form. On the cytoplasmic side, in addition to the *conserved tyrosine kinase domain*, there are also further regulatory sequence portions at which autophosphorylation, and phosphorylation and dephosphorylation by other protein kinases and by protein phosphatases, can take place.

The more than 100-strong family of mammalian receptor tyrosine kinases can be divided into different subfamilies, which are named according to their naturally occurring ligands (Fantl et al., 1993). The subfamilies are classified according to the structure

of the extracellular ligand-binding domains, in which different sequence portions can be differentiated (Ullrich and Schlessinger, 1990; Fig. 8.2). In the extracellular domain, for example, there are Cys-rich sequences that occur as multiple repeats, and sequences with an immunoglobulin-like structure. Other subunits may be attached at the extracellular side of the transmembrane protein chain of the receptor, and receptor chains may also be connected by disulfide bridges.

## 8.1.2 Ligand Binding and Activation

The tyrosine kinase activity localized on the cytosolic side of the receptor is stimulated by extracellular ligand binding. Activation by extracellular ligands is based on a ligand-mediated *change of oligomer status* of the receptor. Stimulation of the tyrosine kinase activity can have a two-fold effect: Firstly, an autophosphorylation of the receptor may take place in trans, i.e., between different protomers of the receptor (Fig. 8.3). Secondly, substrate proteins may be phosphorylated on Tyr residues.

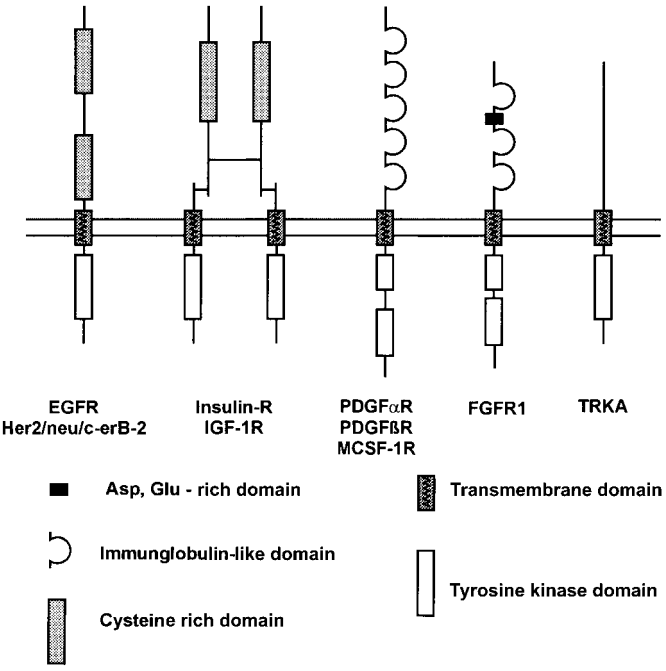



**Fig. 8.2.** Examples of domain structures of receptor tyrosine kinases in mammals. Linear representation of the domain structure of some receptor tyrosine kinases according to the classification of Ullrich and Schlessinger (1990). EGFR: epidermal growth factor receptor; Her2, new, erbB2: alternative name for the 185 kDa „new receptor tyrosine kinase"; InsulinR: insulin receptor; IGF-1R: insulin related growth factor-1 receptor; PDGF-αR, PDGF-βR: platelet-derived growth factor receptor, type α and β; MCSF-1R: macrophage colony stimulating factor-1 receptor; FGFR1: fibroblast growth factor receptor type 1; TrkA: receptor for neutrophins, such as nerve growth factor.

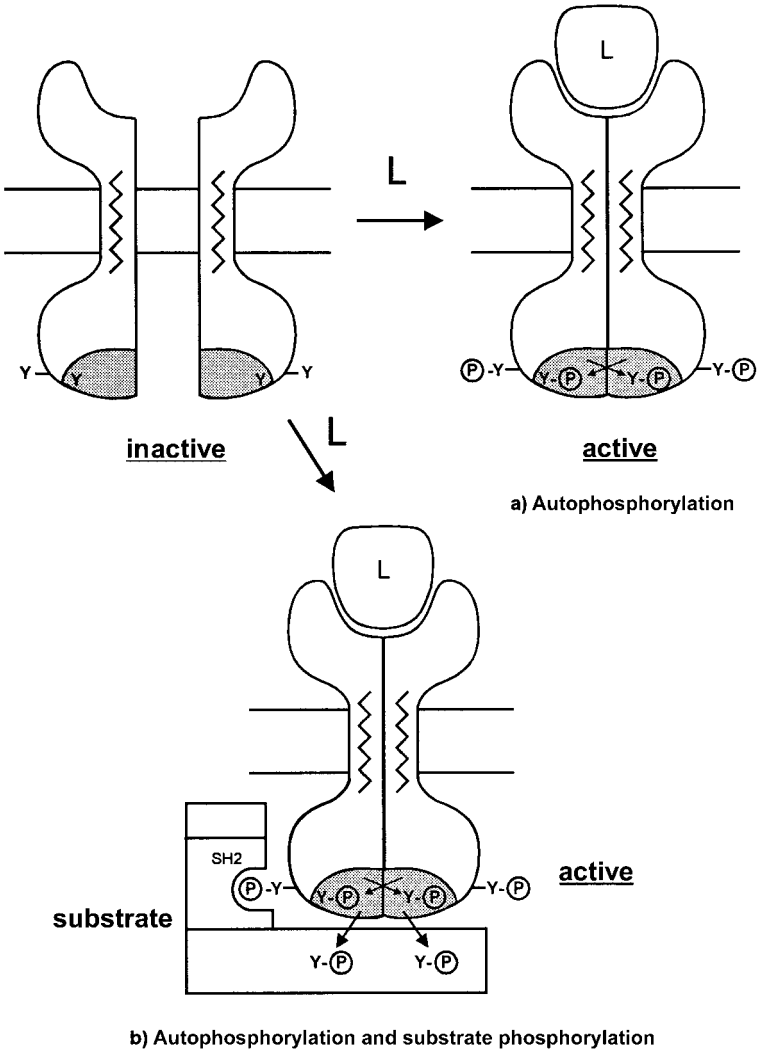



**Fig. 8.3.** Ligand-induced autophosphorylation and substrate phosphorylation of receptor tyrosine kinases. The tyrosine kinase domain of the receptor tyrosine kinase is activated by ligand binding. Consequently, autophosphorylation and/or phosphorylation of substrate proteins takes place. The substrate proteins possess specific phosphotyrosine binding domains (SH2 in the figure or PTP domains, see 8.2), which bind to phosphate residues formed in the process of autophosphorylation.

Ligand-mediated change of the oligomeric structure of the receptor is a general method of transmission of signals through the membrane into the cell interior, with the help of receptor tyrosine kinases (Review: Lemmon and Schlessinger, 1994; Heldin, 1995). It is assumed that signal transmission via receptors with associated tyrosine kinases takes place by a similar mechanism. However, the situation is more complica-

ted here because receptors with associated tyrosine kinase activity are often composed of many subunits (see Chapter 11).

There are at least two mechanisms by which ligand binding can change the oligomer status of the receptor (Fig. 8.4a). In the first case, the ligand has two binding sites for the receptor molecule and brings about a *dimerization* of the receptor. In the absence of the ligand, the receptor exists in a monomeric form. Alternatively, the bound ligand may *stabilize* a *pre-assembled dimeric form* of the receptor. The insulin receptor, for example, is a dimeric protein composed of two $\alpha\beta$-units linked by disulfide bridges. It is assumed that in this case, the activation takes place by an *allosteric mechanism*. The extracellular ligand binding brings about a change in the relative configuration of the two tyrosine kinase domains, in such a way that mutual Tyr phosphorylation is enabled.

### Ligand Structure and Receptor Oligomerization

The ligands of the receptor tyrosine kinases are generally multivalent. The multiple binding sites of the ligand enable it to orient two (or more) of the receptor subunits so that their cytosolic tyrosine kinase activity is stimulated.

In some cases, the ligand itself has a dimeric structure and induces formation of active receptor dimers on binding to the receptor. One example is PDGF, which exists as a disulfide bridge-linked dimeric protein.

Other ligands, such as EGF, hGF and FGF, exist as monomeric proteins in solution. How monomeric ligands can induce a change in the oligomer structure of the receptor is illustrated by the example of hGH; a crystal structure is available of the complex of this ligand with the ligand-binding domain of the hGH receptor. The structure proves that the monomeric hGH protein is able to bind two molecules of the ligand binding domain of the receptor (Fig. 8.4c). Accordingly, the monomeric hGH protein is divalent with respect to binding of its receptor.

### Formation of Heterodimers

An aspect of ligand-induced oligomerization of receptors of regulatory importance is the possibility to form heterodimers. Protein families composed of closely related members can be identified for a number of growth factors and corresponding receptors.

Analogous to heterotypical dimerization of transcription factors (see Chapter 1.2), heterologous dimerization is observed within different members of a receptor family. A certain growth hormone can thus bind to and activate different dimeric combinations of the members of a receptor family. Fig. 8.5. shows the possibilities for heterodimerization of receptors, using the PDGF receptor as an example (review: Lemmon and Schlessinger, 1994).

Heterodimerization of receptor molecules is a mechanism that can increase and modulate the diversity and regulation of signal transduction pathways. Since the various members of a receptor family differ in the exact structure of the autophosphorylation sites and the other regulatory sequences, it is assumed that activity and regulation are different for the various combinations of receptor subtypes. Tissue-specific expression of receptor subtypes enables the organism to process growth hormone signals in a differential way.

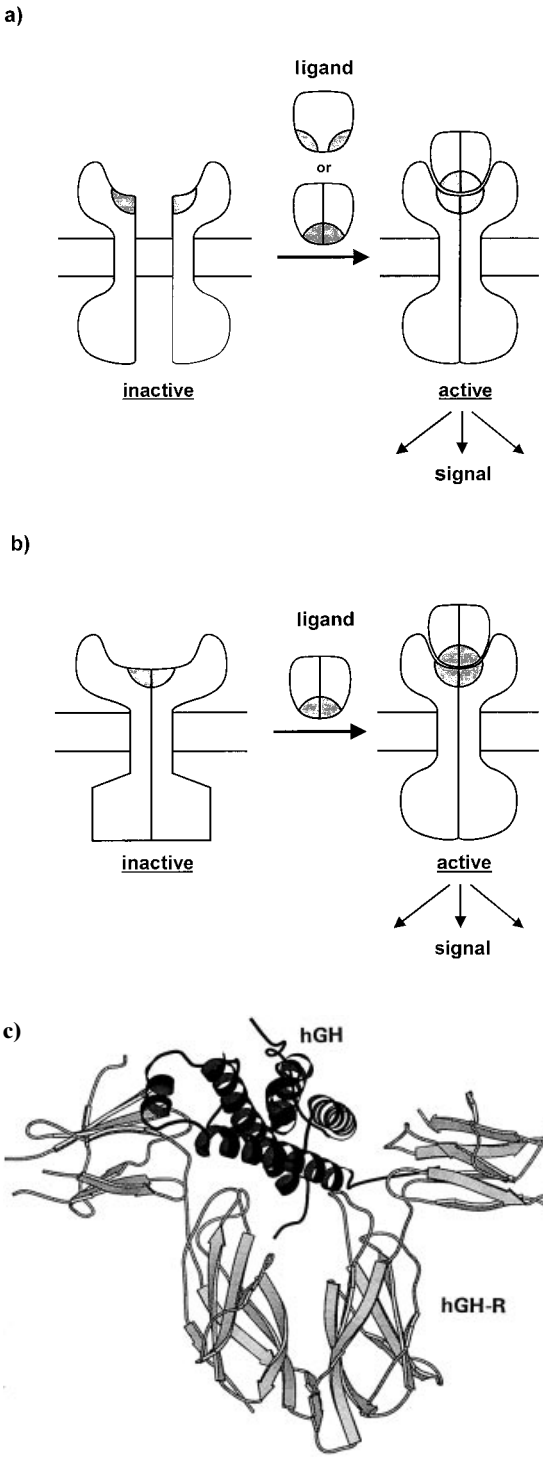



**Fig. 8.4.** Mechanism of activation of receptor tyrosine kinases by ligand binding. Activation of receptor tyrosine kinases is based on a ligand-induced oligomerization and/or conformational change of the receptor. An example is shown of a dimeric receptor; however, activation can also occur in a higher receptor oligomer. a) A bivalent ligand (monomer or dimer) induces a dimerization of a receptor which exists in a monomeric form without the ligand. b) A dimeric receptor is activated via an allosteric mechanism by ligand binding. c) Complex of the human growth hormone (hGH) with the ligand binding domain of the correspoding receptor (hGH-R). The 4 $\alpha$-helices, which arrange themselves in a bundle form, are characteristic for the structure of hGH. The monomeric hGH (dark bands) binds in a non-equivalent manner to two molecules of the ligand binding domain of the hGH receptor.

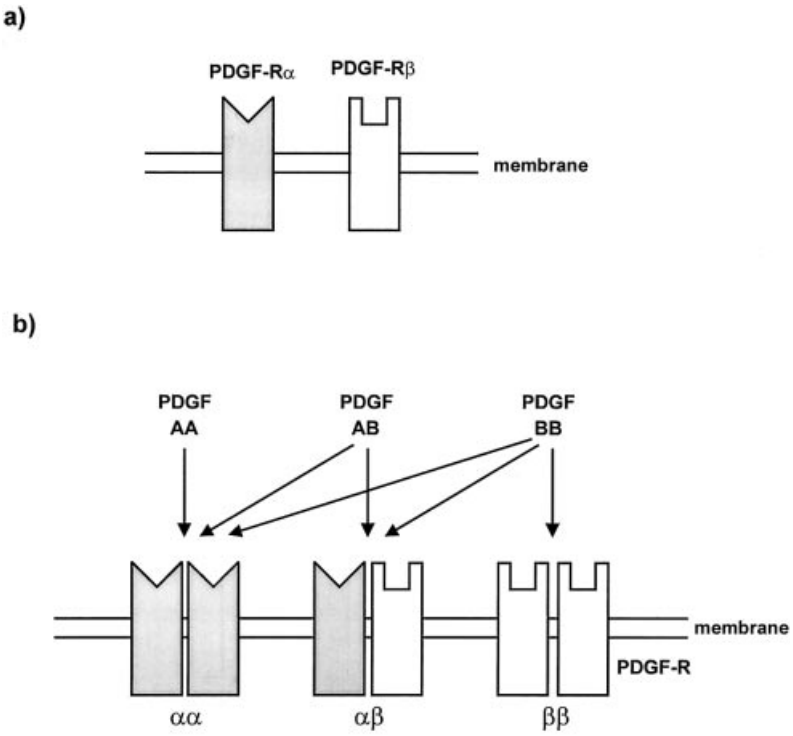



**Fig. 8.5.** Heterodimerization of PDGF receptors. a) There are α and β subtypes of platelet derived growth factor receptor (PDGF-R); these are induced by ligand binding to form homodimers and heterodimers. b) Platelet derived growth factor (PDGF) is a dimeric growth factor, composed of chains A and/or B. The protein may exist as a homodimer (AA, BB) or heterodimer (AB). The AA homodimer of PDGF binds to the αα dimer of PDGF-R, AB binds to the αα and αβ types, BB binds all three combinations. According to Lemmon and Schlessinger, (1994).

### 8.1.3 Structure and Activation of the Tyrosine Kinase Domain

Stimulation of the cytoplasmic localized tyrosine kinase activity is observed as a consequence of binding of the protein ligand to the extracellular side of the receptor tyrosine kinase.

Ligand-induced tyrosine phosphorylation can take place both at the receptor tyrosine kinase itself as well as at specific protein substrates. In the first case, autophosphorylation occurs; in the second case, the signal is passed on to corresponding effector proteins. The effector proteins may carry enzyme activity themselves and be activated by tyrosine phosphorylation. They may also function as adaptor molecules, functioning to pass the signal on to other components of the signaling pathway. With the help of adaptor molecules, other signal proteins are directed to the activated receptor and to the cell membrane and are thus incorporated into signal transduction.

*Autophosphorylation* and *phosphorylation of substrate proteins* are essential elements of signal transduction via receptor tyrosine kinases. Autophosphorylation fre-

quently occurs at several tyrosine residues in the cytoplasmic domain of the receptor tyrosine kinase. Tyrosine phosphorylation sites are found in the Tyr kinase domain itself or close by. It is generally assumed that autophosphorylation takes place by a *trans* mechanism. Accordingly, two neighboring Tyr kinase domains in the receptor oligomer perform a mutual phosphorylation (see Fig. 8.3).

Autophosphorylation may be attributed two functions: first, *activation* of the own Tyr kinase activity by cancelling autoinhibition; second, *creation* of binding sites for corresponding effector proteins, in that Tyr phosphate binds to the SH2 domains or PTB domains (see 8.2) of effector molecules (Fig. 8.6).

### Structure of the Tyr Kinase Domain of the Insulin Receptor

The first insight into the mechanism by which autophosphorylation controls the activity of the Tyr kinase was possible via the crystal structure of the *Tyr kinase domain of the human insulin receptor* (Hubbard et al., 1994).

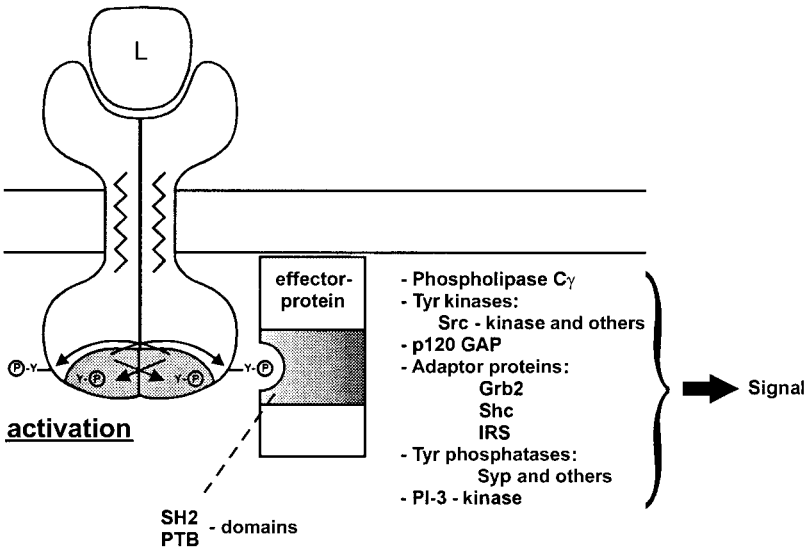



**Fig. 8.6.** Functions of autophosphorylation of receptor tyrosine kinases. Autophosphorylation of receptor tyrosine kinases takes place in trans, i.e., between neighboring protomers of the receptor. The catalytic domain of the receptor is shown as a shaded segment. As a consequence of autophosphorylation, the intrinsic tyrosine kinase activity of the receptor is stimulated. Effector proteins can also bind to the activated receptor. Binding takes place with specific phosphotyrosine binding domains (SH2 or PTB domains) at phosphotyrosine residues of the activated receptor. A critical factor for further signal transduction is the membrane association of the effector proteins that enter into binding with the activated receptor. Details of the effector proteins can be found as follows: phospholipase Cγ: 5.6.2; Src kinase: 8.3.2; p120 GAP: 9.4; Grb2, Shc, IRS: 8.5; PI3-kinase: 6.6.1; Syp tyrosine phosphatase: 8.4.

The insulin receptor has a $\alpha_2\beta_2$ structure (see Fig. 8.2). The $\alpha$ subunit is completely extracellular and is bound to the $\beta$ chain via disulfide bridges. The $\beta$ chain has a transmembrane portion and the tyrosine kinase domain is on the cytosolic side. On binding insulin on the extracellular side, the Tyr kinase activity of the $\beta$ chain is activated and phosphorylation of a total of seven Tyr residues takes place in the *cytoplasmic domain*. Two of these are located in the vicinity of the transmembrane element, three are in the Tyr kinase domain and a further two are in the region of the C terminus. Autophosphorylation activates the Tyr kinase of the insulin receptor. Consequently, an effector protein, *insulin receptor substrate* (*IRS*), is phosphorylated at tyrosine residues that act as docking sites for the SH2 domains of other assigned proteins (see 8.1.4, 8.5). Targeted mutations in the region of the catalytic center of the insulin receptor have shown that the Tyr kinase activity is an essential function in signal transduction via insulin.

The crystal structure of a 306 amino acid fragment of the $\beta$ chain of the insulin receptor indicates that the Tyr kinase domain has a similar construction to that of the catalytic domain of the Ser/Thr-specific protein kinases (see Chapter 7.1.4). Like the latter, the Tyr kinase domain is composed of two lobes: a smaller N-terminal lobe and a larger C-terminal lobe. An essential Arg residue is found in the larger lobe to which the function of the catalytic base during phosphate transfer is attributed. The activation segment located on a loop (see Chapter 7.1.4) is of great importance for regulation of the Tyr kinase activity by phosphorylation. In the activation segment, a Tyr residue (Tyr1162 in Fig. 8.7) is located, which undergoes autophosphorylation on insulin binding to the receptor; this autophosphorylation brings about activation of the receptor. In the inactive state of the Tyr kinase domain, the activation segment is oriented with Tyr1162 directly in the active center and blocking the ATP binding site. Autophosphorylation is not possible under these conditions since ATP cannot be bound.

It is assumed that the catalytic domain of the receptor can exist in two conformations. In the inactive conformation, the activation segment lies in the catalytic center, whereas in the active conformation, it swings out of the active center and both the ATP binding site and the binding site for protein substrate are free and accessible. According to this model, the equilibirum of the two conformations lies on the side of the inactive form in the non-phosphorylated state; phosphorylation at Tyr1162 shifts it towards the active form.

It is postulated that a change in the quarternary structure of the receptor takes place on insulin binding, bringing about a change in the mutual configuration of the $\beta$ chains on the cytosolic side. Consequently, the phosphorylation sites and the active centers of two $\beta$ chains are orientated so that a mutual phosphorylation is possible. The inhibitory Tyr1162 is removed from the active center during this process and the ATP binding site becomes accessible. The conformational change permits a *trans phosphorylation of both* $\beta$ *chains* that also includes a Tyr phosphorylation in the activation segment. As a result, the equilibrium of the receptor conformations is shifted to the side of the active form. According to this theory, phosphorylation of Tyr1162 of the activation segment has a key function in receptor activation. This interpretation is supported by the finding that all known receptor tyrosine kinases have a Tyr residue in an equivalent position.

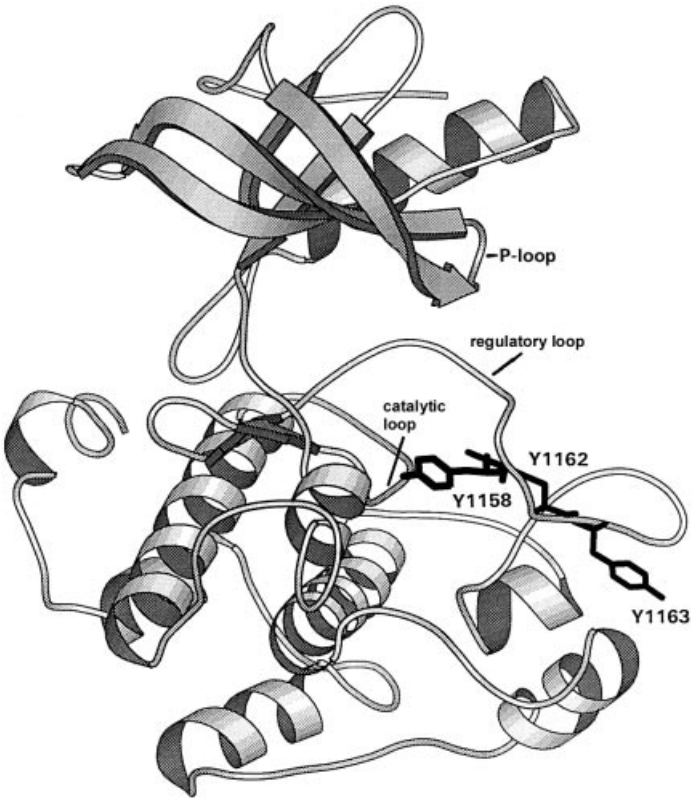



**Fig. 8.7.** Structure of the catalytic domain of the insulin receptor. The crystal structure of the tyrosine kinase domain of the insulin receptor (Hubbard et al., 1994) has a two-lobe structure that is very similar to the structure of the Ser/Thr-specific protein kinases. Structural elements of catalytic and regulatory importance are shown. The P loop mediates binding of the phosphate residue of ATP; the catalytic loop contains a catalytically essential Asp and Asn residue, found in equivalent positions as conserved residues in many Ser/Thr-specific and Tyr-specific protein kinases. Access to the active center is blocked by a regulatory loop containing three Tyr residues (Tyr1158, Tyr1162 and Tyr1163). Tyr1162 undergoes autophosphorylation in the course of activation of the insulin receptor. MOLSKRIPT representation according to Kraulis, (1991).

### 8.1.4 Effector Proteins of the Receptor Tyrosine Kinases

Autophosphorylation of receptor tyrosine kinases has a double effect: The tyrosine kinase activity undergoes autoactivation by phosphorylation of Tyr residues localized in or close to the active center (see 8.1.3). In addition, Tyr residues that lie outside the active center are phosphorylated. The phosphotyrosine residues thereby created serve as *binding sites* for effector molecules next in the sequence of the signal transduction pathway (see Fig. 8.6).

The phosphotyrosine residues of the activated receptors are attachment points for effector proteins that possess a phosphotyrosine binding domain, such as the SH2

domain or the PTB domain (see 8.2). The effector protein is involved in the signal transduction pathway via interaction of its phosphotyrosine binding domain with the phosphotyrosine residue (and neighboring sequence portions) of the activated receptor. Effector proteins may be enzymes or proteins that only possess adaptor function and link further proteins to the activated receptor via protein-protein interactions.

At the level of coupling of receptors and effector proteins, significant *branching* and *variability* of signal transduction is possible. Which effector protein is bound to the activated receptor tyrosine kinases depends on the nature of the SH2 domain of the effector protein and on the sequence environment of the phosphotyrosine residue. The receptor tyrosine kinase frequently has several autophosphorylation sites with different neighbouring sequences, whereby every phosphotyrosine residue of the receptor tyrosine kinase may serve as the binding site for another effector molecule. Fig. 8.8 illustrates the diversity of effector proteins that can interact with a receptor type, using the PDGF receptor as an example.

Important effector molecules of receptor tyrosine kinases are

- The p85 subunit of phosphoinositide-3-kinase (PI3-kinase) (see Chapter 6.6.1)
- Phospholipase C-γ (see 5.5.2)
- Nonreceptor tyrosine kinases of the Src family (see 8.3)
- p120 GAP, a GTPase activating enzyme of Ras signal transduction (see Chapter 9)
- The adaptor protein Grb2 of Ras signal transduction (see 8.5 and Chapter 9)
- The Tyr-specific protein phosphatase SH-PTP2 (see 8.4)

These examples illustrate the diversity of receptor tyrosine kinase-initiated signal transduction pathways:

Starting from receptor tyrosine kinases, signal pathways are activated that involve the intracellular signal molecules $Ins(1,4,5)P_3$, $PtdinsP_3$, $Ca^{2+}$ and diacylglycerol (see Fig. 6.4). The receptor tyrosine kinases are also main components of signal transduction pathways that control growth and differentiation of the cell. The proteins listed above are key components of central signal transduction pathways and, starting from very different external stimuli, they are activated via receptor tyrosine kinases and bound into signal transduction.

The specific interaction between the phosphotyrosine residues of the activated receptor and the SH2 domain of the effector protein is the basis for the activation of effector proteins for further signal transduction. Like the nature of the effector proteins, the mechanism of their activation is also very variable. Important mechanisms are:

- Phosphorylation of the effector molecule at Tyr residues
  Example: phospholipase C-γ

- Induction of a conformational change in the effector molecule
  Example: PI3-kinase

- Translocation of the effector molecule to the plasma membrane
  Example: Grb2-Sos, Shc-Grb2

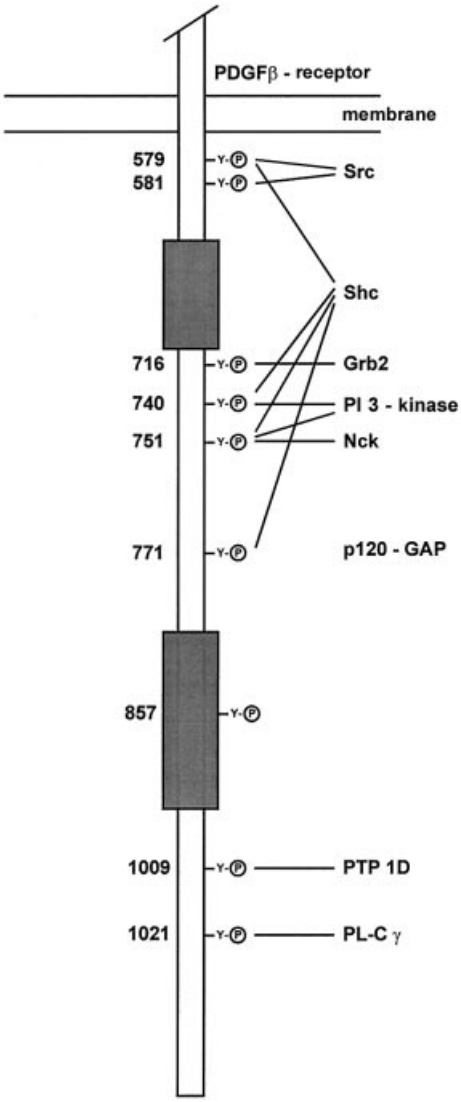



**Fig. 8.8.** Phosphotyrosine residues in the PDGF receptor and specificity of binding of SH2-containing signal proteins. The figure illustrates the diversity of the different effector proteins that can interact with an activated receptor. The tyrosine residues of platelet derived growth factor receptor (PDGF-R), for which autophosphorylation has been demonstrated, are designated according to their position in the receptor sequence. PDGF-R has at least nine different tyrosine phosphorylation sites in the cytoplasmic domain. The phosphotyrosine residues are found in different sequence environments and are recognized by the SH2 domains of the assigned effector proteins. The filled rectangles indicate the two-part tyrosine kinase domain of PDGF-R. Src: members of the Src tyrosine kinase family; Sh2, Grb2, Nck: adaptor proteins; PI(3) kinase: phosphatidyl inositol-3 kinase; GAP: GTPase activating protein; PTP-1D: protein tyrosine phosphatase 1D; PL-Cγ: phospholipase C-γ. According to Heldin, (1995).

## 8.2 Protein Modules as Coupling Elements of Signal Proteins

Starting from an activated receptor tyrosine kinase, further conduction of the signal takes place with the help of specific protein-protein interactions between the activated receptor and one or more effector proteins next in the sequence. In many cases, the effector molecules pass the signal on to other proteins of the signaling pathway, forming chains of signal proteins in sequence. Specific protein-protein interactions are the

basis of the coupling of signal proteins.These function mainly to bring about close spatial configuration of signal-carrying enzymes with their substrates, by leading a substrate protein to the catalytic center or by targeting an enzyme usually located in the cytosol to the cell membrane, where it has direct access to its substrates.

The cell uses defined structural elements for communication between different proteins of a signal transduction pathway; these are found in the form of self-folding protein domains in many signal transmitting proteins. The protein domains are modules of 60–100 amino acids that – with a common basic structure – occur in very different proteins in slightly modified forms. The protein modules mediate protein-protein interactions or protein-membrane interactions in signaling pathways. They are used to associate proteins of a signaling pathway or to target them to the vicinity of the cell membrane (review: Pawson, 1995; Cohen et al., 1995, Pawson and Scott, 1997, Sudol, 1998).

Two points are of particular importance for the coupling function of the protein modules. The variability of the protein modules: For a particular basic motif of a module, there are generally many *variants* that have a slightly different binding specificity and are thus assigned to different structural motifs in the target protein. This results in great variability and diversity of coupling. Another functionally important aspect of coupling of signal proteins is the occurrence of *several* protein modules in a protein (multivalence of signal proteins, Figs. 8.9, 8.10). If a signal protein has multiple binding valence for different effector proteins, networks of interacting proteins can be created that contribute greatly to specificity and diversity of signal transduction and permit linking of different signaling pathways.

At present, a number of structural motifs have been described for signal proteins, to which specific coupling functions in signal transduction are attributed. The most important are:

- SH2 domains and PTP domains
- SH3 domains
- PH domains
- PDZ domains
- WW domains.

Fig. 8.10 illustrates the diversity and variability of the occurrence of some of these domains in signal proteins.

### 8.2.1 SH2 Domains

The SH2 domains are the best investigated protein modules with a regulatory function. They were first discovered as a sequence motif showing homology with a sequence of the Src tyrosine kinase, thus the name SH, from Src homolog. Src tyrosine kinase was the first known tyrosine kinase. The enzyme was discovered in avian retroviruses that induce sarcomas (see 15.3). Later it became apparent that Src tyrosine kinase also belongs to the normal enzyme repertoire of mammalian cells and is involved in regulation of cell growth and differentiation (see 8.3).

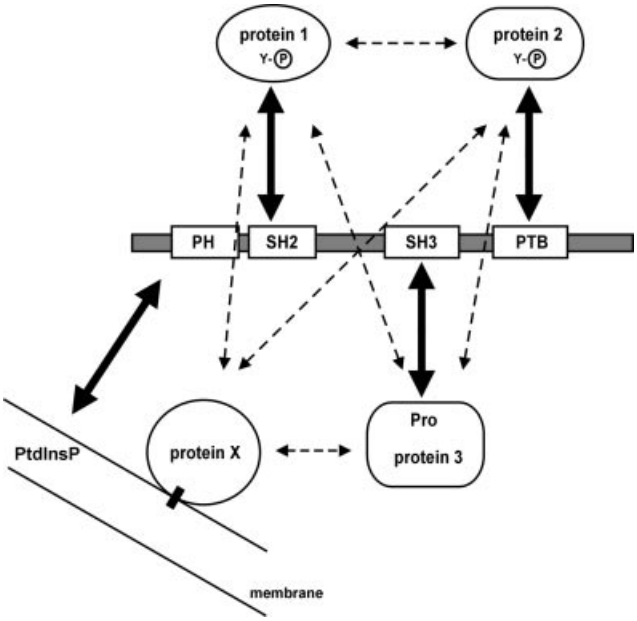



**Fig. 8.9.** Crosslinking of signal proteins with the help of protein modules. A hypothetical protein is shown which contains SH2, SH3, PTB and PH domains. Recognition of phosphotyrosine residues occurs with the help of SH2 or PTB domains; SH3 domains bind to proline-rich sequences (Pro in Protein 3) whilst the pleckstrin homology domains (PH domains) mediate binding to phosphatidyl-inositol-phosphates (PtdInsP) in the membrane. In an idealized scheme, the modular protein can associate several proteins (Protein 1 – Protein 3) and mediate interactions between these proteins (shown as broken arrows). The PH domain helps to recruit the complex to the cell membrane favoring interactions with other membrane-associated proteins (Protein X).

#### 8.2.1.1 Binding Specificity and Structure of SH2 Domains

The SH2 domain functions to specifically recognize phosphotyrosine residues in target proteins and to bind these. The presence of a Tyr phosphate grouping is obligatory for binding of an SH2 domain to a target protein or to a model peptide. In addition, the neighboring sequence is crucial. The sequence environment of the phosphotyrosine residue defines the binding substrate of a particular SH2 domain and differentiates the binding preference of different SH2 domains. For structure-function investigations of SH2 domains, it was very important that phosphotyrosine-containing peptides bind with high specificity to the SH2 domains and thus can serve as models for the SH2 substrate interaction. The great variability of SH2 domains and their substrates is emphasized by the observation that two SH2 domains occur in many signal proteins. These mostly have different substrate binding preferences.

The current structural information on SH2-substrate complexes, the results of binding experiments using peptide libraries and the systematic investigation of binding sites for different SH2 domains showed that the SH2 domains can be divided into at least five classes (1A, 1B, 2, 3, 4), differing in the sequence requirements of the substrate

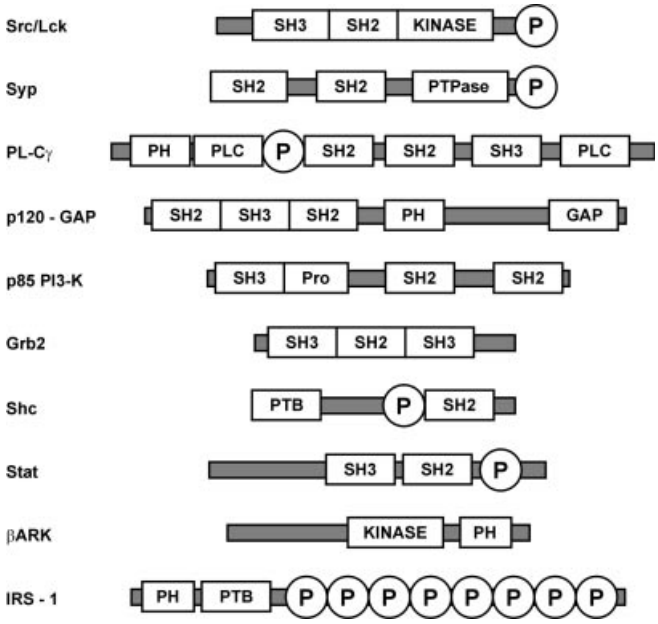



**Fig. 8.10.** Protein modules in signal proteins. Modular structure of important signal proteins in linear representation. Signal proteins: Src/Lck: family of the Src protein tyrosine kinases; Syp: Syp protein tyrosine phosphatases; Stat: Stat transcription factor (see 11.1.5.2); p85/PI3K: 85 kDa regulatory subunit of the phosphatidyl inositol-3 kinase; βARK: β-adrenergic receptor kinase; IRS-1: insulin receptor substrate 1; adaptor proteins: IRS-1, Shc, Grb2, protein domains: SH2, SH3: Src homology regions 2 and 3; Pro: proline-rich SH3 binding site; P: phosphotyrosine-containing binding site for SH2 domains; Kinase: catalytic domain of the protein kinase; PH: pleckstrin homology domain; PTB: phosphotyrosine binding domain; PLC: catalytic domain of phospholipase; GAP: GTPase activating domain.

(Songyang and Cantely, 1995). Furthermore, another phosphotyrosine binding motif was found, known as the *phosphotyrosine binding domain* (*PTB*) (review: Van Geer and Pawson, 1995, Sudol, 1998). In the following, the principles of the SH2-phosphotyrosine interaction will be described, based on the SH2 domains of class 1A and class 3.

### SH2 Domains of Class 1A

The crystal structures of the SH2 domains of Src tyrosine kinase and Lck tyrosine kinase in complex with Tyr phosphorylated peptides have enabled important insight to be obtained into recognition of the phosphotyrosine residue and the neighboring amino acids in class 1A of SH2 domains. The phosphate residue is bound in a deep pocket of the SH2 domain, at the end of which an invariant Arg residue (Arg βB5) is located which contacts the negatively charged phosphate by a two-pronged interaction. It can be estimated that a phosphoserine or phosphothreonine residue would be too short to enter into a similar interaction with the Arg residue.

The neighboring sequences of the phosphotyrosine residue are decisive for binding specificity of a SH2 domain. At this point, the structure shows that particularly the isoleucine residue at position +3 relative to phosphotyrosine is bound in a very specific manner in a pocket of the SH2 domain (Fig. 8.11). Binding of the peptide to the SH2 domain in Src kinase has therefore been compared to binding of a two-pole plug in a complementary socket, where one of the poles is phosphotyrosine and the other is the amino acid at position +3.

**SH2 Domains of Class 3**

Phospholipase Cγ and protein tyrosine phosphatase Syp possess an SH2 domain of class 3. Their substrate binding site has mostly hydrophobic character. The substrate is bound in a stretched form in a flat pit where contacts are formed to a hydrophobic sequence section of the substrate, including 5—6 amino acids on the C-terminal side of the phosphotyrosine residue.

### 8.2.1.2 Function of the SH2 Domain

The interaction between the SH2 domain of a signal protein and a phosphotyrosine-containing protein ligand serves to pass the signal on in a signal chain (see Pawson, 1995; Cohen, 1995). The phosphotyrosine-containing protein ligands are often receptors, which are activated and autophosphorylated on binding an extracellular ligand.

Phosphotyrosine-SH2 interactions are involved in passing on the signal for at least three different types of transmembrane receptor:

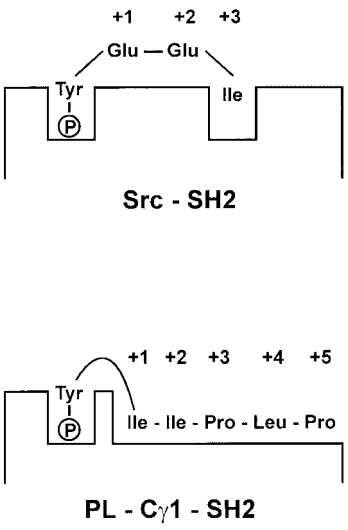



**Fig. 8.11.** Recognition of phosphotyrosine-containing substrate peptides by SH2 domains of Src kinase and phospholipase Cγ1. Binding of phosphotyrosine-containing peptides to SH2 is shown schematically, based on crystal structures of the complexes. The SH2 domain of Src kinase has a basic binding pocket for the phosphotyrosine residue and a hydrophobic pocket for the isoleucine residues at position +3 of the peptide substrate. The SH2 domain of PL-Cγ1 has a hydrophobic binding surface to which the C-terminal part of the peptide P-Tyr-Ile-Ile-Pro-Leu-Pro-Asp binds. According to Cohen, (1995).

### Receptor Tyrosine Kinases

Receptor tyrosine kinases pass signals on in that specific effector proteins are brought into the vicinity of the autophosphorylated receptor, via phosphotyrosine-SH2 interactions, and are functionally activated. The diversity and branching of signal transmission at this point is shown by the example of the receptor for platelet derived growth factor (PDGF-R). The β subtype of PDGF-R has several autophosphorylation sites with different sequence environments and which are therefore assigned to different SH2-containing effector molecules (see Fig. 8.8). At least eight different effector molecules can bind to the different phosphotyrosine sequences on the cytoplasmic side of PDGF-R. The same receptor can be involved in very different signaling pathways, as shown by this example. Which pathway is used will depend on the availability and activity of the different effector proteins, a situation regulated in a cell-specific and tissue-specific manner.

### T cell Receptors

The T cell receptors (see Chapter 11) are an example of receptors composed of several subunits. Ligand binding to the receptor activates an associated tyrosine kinase which phosphorylates Tyr residues in the cytoplasmic region of the receptor and thus creates binding sites for SH2 domains of effector molecules downstream.

### Cytokine Receptors

The receptors for cytokines and interferons are the starting point for signal transduction chains that bring about an activation of transcription factors. The signaling pathway involves the Janus protein kinase and Stat transcription factors (see 11.1.4). Phosphotyrosine-SH2 interactions are also involved in several steps of signal transduction here.

The SH2 domains also have diverse linking functions in signal reactions sequential to activated receptors. The adaptor proteins should be mentioned in particular here (see 8.5); their SH2 groups are used to interlink signal proteins.

Coupling of signal proteins by phosphotyrosine-SH2 interactions can serve different purposes:

### Membrane Localization

Via binding of an SH2-containing signal protein to an activated receptor tyrosine kinase, the signal protein is brought to the membrane and into the vicinity of the corresponding target protein or substrate. Examples are phospholipase Cγ and PI3-kinase, which have substrates in the phospholipid membrane. The same is true for the GTPase activating protein of the Ras pathway and for the Grb2-Sos complex, which both have the Ras protein as target protein (see Chapter 9). In these cases, the target proteins are localized in or at the membrane, and activation of the signaling pathway initiates translocation of the cytosolic signal protein to the membrane where it is in the immediate vicinity of the substrate or target protein.

### Phosphorylation of SH2-containing Proteins

Many SH2-containing signal proteins are brought, via interaction of their SH2 group with phosphotyrosine residues, into the neighborhood of the catalytic center of the tyrosine kinase and are themselves substrates for tyrosine phosphorylation. By this mechanism, new attachment sites can be generated for other SH2-containing proteins within SH2-containing signal proteins. In this way, several components of a signaling pathway can be sequentially linked. An example is the SH2-containing Syp phosphatase which is phosphorylated at Tyr residues by the PDGF receptor, creating a binding site for the adaptor protein Grb2, which also contains a SH2 domain (see also Chapter 9).

### Allosteric Activation of an SH2-containing Protein

Several cases are described (Cohen et al., 1995) in which binding of an SH2-containing enzyme to an activated receptor tyrosine kinase leads to increased catalytic activity of the enzyme. Examples are the PI3-kinase and phospholipase Cγ. The mechanism of activation is not clear. It is possible, however, that the basis is an allosteric mechanism, as is assumed for activation of Src tyrosine kinase. The Src kinase can be phosphoryla-

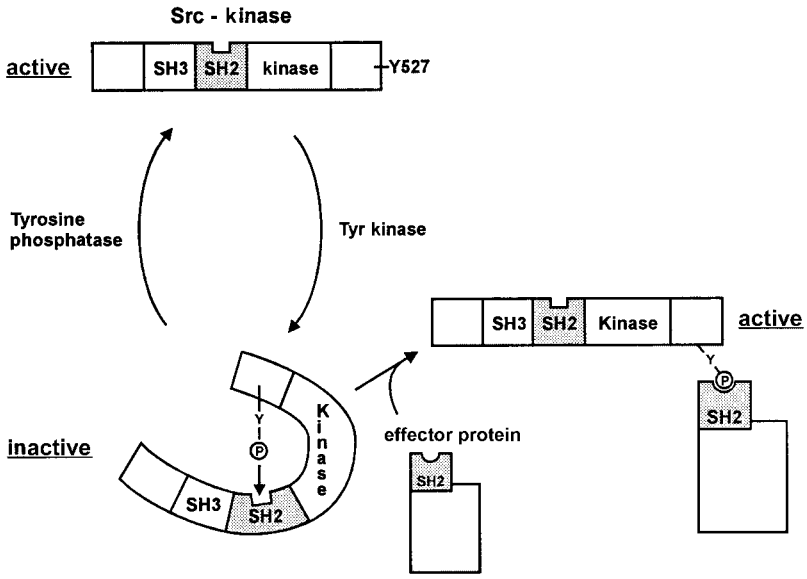



**Fig. 8.12A.** Model of regulation of Src kinase by phosphorylation. At the C-terminal end of the Src tyrosine kinase, a Tyr phosphorylation site (Tyr527) is located, which, when phosphorylated by a tyrosine kinase leads to inactivation of the Src kinase. P-Tyr527 binds intramolecularly to the SH2 domain, blocking the kinase activity. Removal of the Tyr phosphate P-Tyr527 by a tyrosine phosphatase converts the Src kinase into the active state again. Activation of the Src kinase can also be brought about by a SH2-containing effector protein; the SH2 domain of this effector protein competes with the SH2 domain of Src for binding to P-Tyr527. Alternatively, P-Tyr527 may bind to the SH2 domain of another signal protein (not shown in the figure). The src kinase is also regulated by phosphorylation.

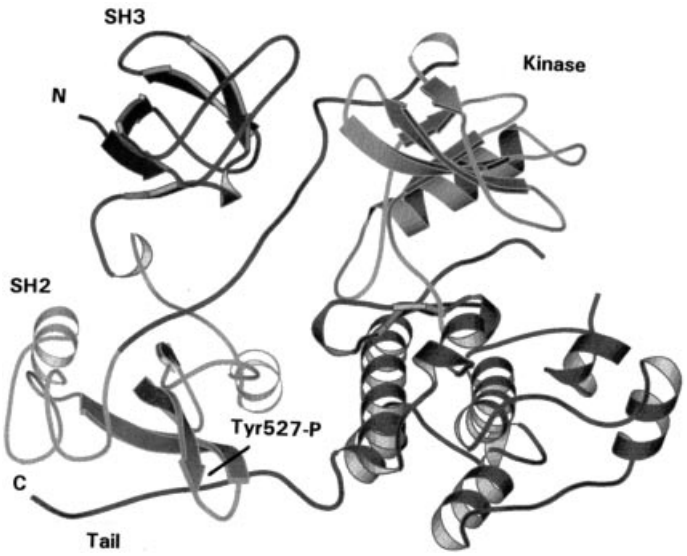



**Fig. 8.12B.** Structure of c-Src kinase phosphorylated at Tyrosine 527. Ribbon diagram showing the structure and organization of the „closed conformation" of c-Src kinase. Two aspects of the structure are important for the regulation of c-Src kinase: i) The phosphorylated Tyr 527 of the C-terminal tail is engaged in an intramolecular interaction with the SH2 domain. ii) The SH3 domain binds to the linker between the SH2 domain and the kinase domain. Both interactions are assumed to fix an inactive state of the kinase.

ted at Tyr residues and it contains a SH2 domain (see 8.3.2). A C-terminal phosphotyrosine residue (pY527) plays a special role in regulation of the activation of Src kinase, controlling the kinase activity in a negative manner. pY527 binds intramolecularly to the SH2 domain of Src kinase, bringing about inhibition of the tyrosine kinase activity. If the SH2 domain binds a phosphotyrosine residue of an exogenous ligand, autoinhibition of Src kinase is lifted (Fig. 8.12).

## 8.2.2 Phosphotyrosine Binding Domain, PTB Domain

The phosphotyrosine-binding (PTB) domains harbor – like the SH2 domains – specific binding sites for phosphotyrosine-containing peptides. PTB domains are found particularly in proteins that have a docking or adaptor function by recruiting additional signaling proteins to the vicinity of an activated receptor. Examples are the adaptor molecule Shc (see 8.5), which has an SH2 domain of class 3 in addition to the PTB domain, and the insulin receptor substrate IRS1. PTB domains recognize phosphotyrosine residues in context with sequence sections towards the N terminus (Songyang and Pawson, 1995). Despite their common preference for binding to phosphotyrosine residues, the PTB and SH2 domains are very different in structure. In contrast, PTB domains have a very similar folding topology to PH domains (see 8.2.4).

## 8.2.3 SH3 Domains

SH3 domains occur in signal proteins that are involved in Tyr kinase signaling pathways (see Cohen et al., 1995; Pawson, 1995). They are also found in proteins of the cytoskeleton and in a subunit of the neutrophilic cytochrome oxidase. Ligand binding at SH3 domains takes place via *Pro-rich sequences* of ca. 10 amino acids, and Pro-rich peptides are very good binding substrates.

### 8.2.3.1 SH3 Structure and Ligand Binding

The structural determination of SH3 domains with bound Pro-rich peptides has shown that the Pro-rich section of the ligand is bound as a left-handed polyproline Type II helix with three amino acid residues per turn. The polyproline Type II helix was described for polyproline some time ago.

The sequence X-P-p-X-P is a consensus sequence for SH3 ligands, in which the two proline residues P are invariant, X is usually an aliphatic residue and p is often a Pro residue. The two invariant proline residues are each bound in a hydrophobic pocket of the SH3 domain. Peptide ligands can be bound in the C→N and also in the N→C direction. Like the SH2 domains, there are many different SH3 domains. The different SH3 domains demonstrate differing binding preferences for Pro-rich sequences, the specificity being determined by the neighboring residues of the invariant proline.

### 8.2.3.2 Functions of the SH3 Domain

A general function of the SH3 domains is the binding of Pro-rich sequences in target proteins of a signaling pathway. SH3 domains are thus coupling modules in signaling pathways. The biological importance of the SH3 domains is emphasized by the observation that deletion of the SH3 domains of the cytoplasmic tyrosine kinases Abl and Src leads to a significant increase in the tumor-transforming potential of both tyrosine kinases. Furthermore, a number of cellular processes have been identified in which SH3 domains mediate functionally important protein-protein interactions. SH3 domains can be attributed the following principal functions:

#### Mediation of Specific Subcellular Localization

A primary function of the SH3 domains is to form functional oligomeric complexes at defined subcellular sites, frequently in cooperation with other modular domains. SH3 domains are found in many proteins associated with the cytoskeleton or with the plasma membrane. Examples are the actin binding protein α-spectrin and myosin Ib. Furthermore, SH3 interactions are involved in signal transduction in the Ras pathway (see Chapter 9).

### Regulation of Enzyme Activity

An example of regulation of enzyme activity via SH3 domains is the negative regulation of the activity of Src tyrosine kinase by SH3-mediated interactions (review: Cohen et al., 1995).

### Contribution to Substrate Specificity of Tyrosine Kinases

Tyrosine kinases demonstrate intrinsically low substrate specificity and it is therefore difficult to formulate consensus sequences for phosphorylation sites of the substrates of tyrosine kinases (Songyang and Cantley, 1995). One means of increasing the specificity and selectivity of Tyr phosphorylation seems to be the use of SH3 domains for specific coupling of substrates to Tyr kinases (Fig. 8.13). The adaptor proteins Crk, Grb2 and Nck (review: Cohen et al., 1995) are specifically phosphorylated by the Abl tyrosine kinase at Tyr residues. All three adaptor proteins possess SH3 domains that can bind to Pro-rich sequences of the Abl tyrosine kinase. This interaction mediates tight binding of the substrate to the tyrosine kinase and enables an effective tyrosine phosphorylation. Other possible substrates that have no SH3 domain are excluded from Tyr phosphorylation since they cannot enter into tight binding with the tyrosine kinase.

The specific function of the SH3 domain is based on *increased substrate specificity* of tyrosine kinases in this case (Shokat, 1995). In classical enzymes, the substrate binding site and the catalytic center are close together and the substrate binding site is generally highly specific for a particular substrate. The situation is different for tyrosine kinases. Here, the substrate binding site near the catalytic center shows moderate selectivity. The specificity of the reaction is increased, however, by mediation of asso-

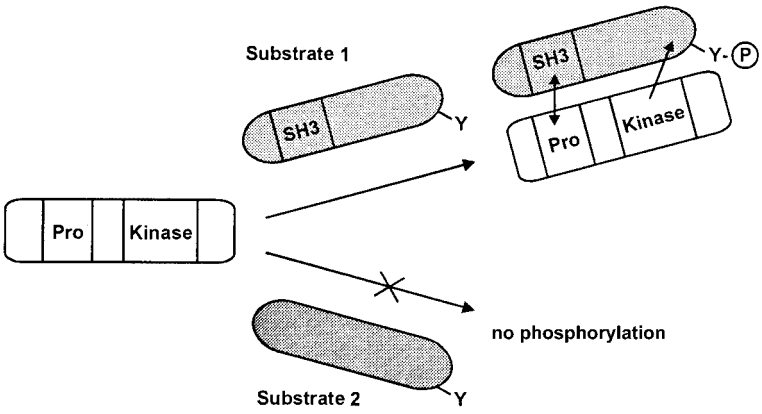



**Fig. 8.13.** SH3 domains and specificity of tyrosine kinases. Binding of the SH3 domain of the substrate protein S1 to the Pro-rich sequence of the tyrosine kinase increases affinity of the tyrosine kinase for its substrate and favors its phosphorylation. Proteins (S2) that possess Tyr phosphorylation sites but do not have SH3 domains are not converted by the tyrosine kinase, or only to a much lower extent.

ciation with the tyrosine kinase via additional structural elements of the substrate, namely its SH3 domain. The interaction between SH3 domain and Pro-rich sequences of the tyrosine kinase increases the specificity of the reaction and contributes to substrate recognition, bringing high specificity to an otherwise unspecific catalytic reaction.

### 8.2.4 Pleckstrin Homology Domains

The pleckstrin homology (PH) domain is a structural motif of ca. 100 amino acids found in many signal molecules such as Ser/Thr-specific protein kinases, tyrosine kinases, isoforms of phospholipase C (PL-Cβ, γ and δ), in G nucleotide exchange factors, adaptor proteins, and in proteins of the cytoskeleton (see also Fig. 8.10). Originally, the PH domain was found in the 47 kDa pleckstrin protein, which is the main substrate of protein kinase C in platelets.

PH domains bind phosphatidyl inositol derivatives and, due to this property, are able to mediate membrane association of signal proteins. The PH domain of PL-Cδ1 binds to phospholipids such as Ptd(Ins)$P_2$ with high affinity and specificity. The crystal structure of the PH domain of PL-Cδ1 with bound Ptd(Ins)$P_2$ surprisingly has a very similar folding topology to the PTB domain that specifically binds phosphotyrosine-containing peptides (see 8.2.3; review: Lemmon et al., 1996). The importance of this similarity is not understood.

In total, the PH domains can be attributed specific anchoring functions at membranes (review: Cohen, 1995; Pawson, 1995, Pawson and Scott, 1997) with phospholipids of the membrane serving as binding substrates. Signal-induced availability of phosphatidyl inositol lipids such as PtdIns$P_3$ thus permits regulated membrane anchoring of PH-containing signal proteins (see 6.6). Due to the great structural similarity between the PH and PTB domains, it has been speculated that the PH domains also mediate specific protein-protein interactions.

### 8.2.5 PDZ Domains

PDZ domains were first identified in proteins of postsynaptic cells and their designation comes from their occurrence in the proteins PSD-95, DlgA and ZO-1 (see Saras and Heldin, 1996). In the meantime, PDZ domains have been found in many other proteins, particularly in proteins that form structures in the cell membrane (e.g. in ion channels) and in signal proteins (review: Craven and Brett, 1997). PDZ domains recognize short peptide sequences with a C-terminal hydrophobic residue and a free carboxyl group, such as the E(S/T)DV motif at the C terminus of certain subunits of ion channels.

An important function of the PDZ domains lies in the formation of macromolecular associates at the cell membrane (review: Pawson and Scott, 1997). PDZ proteins can also provide a framework for clustering of proteins, such as ion channels, at the cell membrane and they may help to recruit proteins into membrane-bound macromolecular complexes.

Many proteins contain multiple PDZ domains with various sequences that may show different binding specificities. In this way, a protein with multiple PDZ domains can help to organize different proteins in supramolecular complexes. The InaD protein of *Drosophila* is an example, which is composed exclusively of 5 PDZ domains with different binding specificity and to which different target proteins are assigned (Tsunoda et al., 1997). The target proteins are three proteins involved in processing of light signals in the eye of *Drosophila*. During phototransduction, InaD associates via its distinct PDZ domains with phospholipase C-β, which is the target of rhodopsin-activated $G_{q,\alpha}$, with the calcium channel TRP and with protein kinase C. The signaling complex formed allows efficient activation of the TRP channel by phospholipase C, in response to stimulation of rhodopsin. Furthermore, the presence of protein kinase C in the signaling complex provides for efficient deactivation by phosphorylation of the TRP channel. It is assumed that the InaD protein functions as an adaptor or scaffolding protein, which organizes light-induced signaling events into supramolecular complexes.

### 8.2.6 WW Domains

The WW domains are small protein modules of 35—40 amino acids that mediate protein-protein interactions in ubiquitin-dependent protein degradation and in recognition of protein substrates by protein kinases. The binding specificity of WW domains is not yet clear. Specific binding to pro-rich sequences has been reported. However, it has been also shown that the WW domain functions as a phosphoserine- or phosphothreonine- binding module (Lu et al., 1999). The WW domain of the mitotic peptidyl prolyl cis-trans isomerase Pin1 binds to its physiological substrates in a phosphorylation dependent manner and is required for Pin1 to perform its essential function during mitosis *in vivo* (Yaffe et al., 1997).

## 8.3 Nonreceptor Tyrosine-specific Protein Kinases

In addition to receptor tyrosine kinases, the cell also contains a number of tyrosine-specific protein kinases that are not an integral component of transmembrane receptors. These „nonreceptor“ tyrosine kinases are localized in the cytoplasm at least occasionally or they are associated with transmembrane receptors on the cytoplasmic side of the cell membrane. They are therefore also known as *cytoplasmic tyrosine kinases*. The nonreceptor tyrosine kinases perform essential functions in signal transduction via cytokine receptors (see Chapter 11) and T cell receptors, and in other signaling pathways.

### 8.3.1 Structure and General Function of Nonreceptor Tyrosine Kinases

Permanent or transient association with subcellular structures, and variable subcellular distribution, are characteristic for the cytoplasmic tyrosine kinases. The nonreceptor tyrosine kinases are intracellular effector molecules that can associate with specific substrates during the process of signal transduction and activate these by tyrosine phosphorylation, to pass on the signal. Many of the functions of the nonreceptor tyrosine kinases are performed in the immediate vicinity of the cell membrane, whether a signal is received from an activated membrane receptor or a signal is passed on to a membrane-associated protein.

In Fig. 8.14, the structures of some nonreceptor tyrosine kinases are shown schematically (review: Taniguchi, 1995). The nonreceptor tyrosine kinases can be divided into different families based on sequence homology; the Src family is the best characterized. In addition to the catalytic domain, the nonreceptor tyrosine kinases often have SH2 and SH3 domains responsible for specific association with substrate proteins. Nonreceptor tyrosine kinases that contain both SH2 and SH3 domains may be involved in signal transduction in larger protein complexes.

### 8.3.2 Src Tyrosine Kinase and Abl Tyrosine Kinase

Two of the nonreceptor tyrosine kinases are highlighted, Src kinase and Abl kinase.

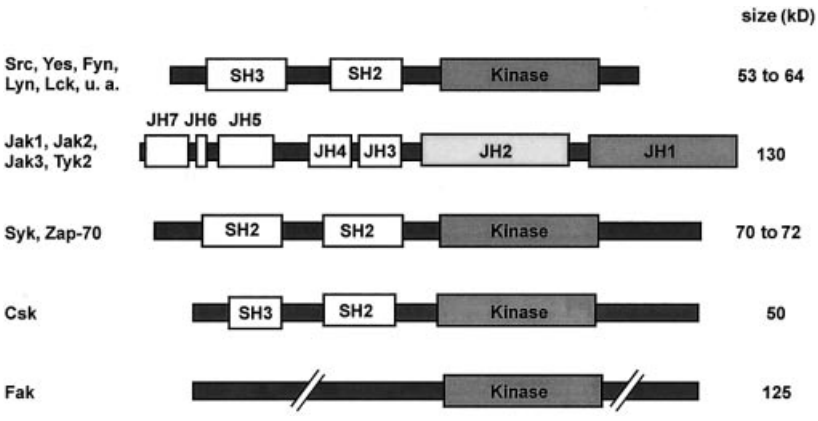



**Fig. 8.14.** Domain structure of cytoplasmic tyrosine kinases. Linear representation of the domain structure of selected cytoplasmic tyrosine kinases. Details of the cytoplasmic tyrosine kinases: Src: 8.3.2; Jak, Tyk: 11.1.5; Zap-70: 11.2.2; Fak: 11.3; JH: Jak homology region. According to Taniguchi, (1995).

### Src Kinase

Src kinase belongs to a family of closely related tyrosine-specific protein kinases involved in regulation of cell division, cell differentiation and cell aggregation. In the meantime, at least 9 different protein kinases are numbered amongst the family of Src kinases. Src kinase is involved, for example, in signal transduction via growth factor receptors and via integrins. Functional interactions have been described with Tyr-P residues of PDGF receptor, EGF receptor and focal adhesion kinase, FAK (see Chapter 11) (review: Parson and Parson, 1997).

Src kinase was discovered during the search for the tumor-causing principle of retroviruses. The viral oncogene product of these viruses, v-Src kinase, was the first tyrosine kinase to be identified. In comparison to its cellular counterpart, c-Src kinase, v-Src lacks the autoinhibitory structural element that controls protein kinase activity. As a consequence of this loss, v-Src kinase is constitutively active and is a potent transforming protein.

These observations and mutation studies have shown that the activity of c-Src kinase is subject to strict regulation. The domain structure of c-Src kinase is shown in Fig. 8.12. Src kinase carries a myristinic acid residue as a membrane anchor and is regulated in a complex manner by SH2 and SH3 domains and via Tyr phosphorylation (see Fig. 8.12).

The SH2 and SH3 domains, the phosphorylation at Tyr 527, which is catalyzed by a c-Src specific protein kinase (Csk), and phosphorylation at Tyr416 are key elements of regulation. The phosphorylation at Tyr527 inhibits the Src kinase; phosphorylation at Tyr416, which is found in the activation segment, in contrast has an activating influence. Furthermore, activity of Src kinase is modulated by its SH2 and SH3 domains. The first insight into the mechanism of regulation of Src kinase was obtained from structural determination of an inactive form of Src kinase (Fig. 8.12B), in which Tyr527 is phosphorylated (Xu et al., 1997, review: Sicheri and Kuriyan, 1997). The crystal structure clearly shows the domain architecture of Src kinase and the interaction between the domains. As already predicted from mutation experiments, Tyr527-P enters into intramolecular binding with the SH2 domain. Furthermore, the SH3 domain is involved in an interaction with the linker between the catalytic domain and the SH2 domain and it helps to fix the SH2 domain in a conformation favorable for interaction with Tyr527. The interaction of the domains and the Tyr-P-SH2 interaction stabilize a conformation of Src kinase in which the active site is disrupted due to distortion of the C helix. The inactive conformation of the catalytic domain is very similar to that of CDK2 (see Chapter 13).

The crystal structure clearly shows that Src kinase is held by intramolecular interactions in an inactive conformation, in which the binding surfaces of the SH2 and SH3 domains are sequestered. Several paths for conversion to the active, open form are under discussion. This could take place by activation via dephosphorylation of the C-terminal tail. It could also take place by the apposition of a high affinity ligand for the SH2/SH3 domains. In addition, phosphorylation of Tyr416 is also needed for full activation.

### Abl Tyrosine Kinase

Abl tyrosine kinase also carries a myristinic acid residue as a membrane anchor. The complex structure is a distinctive feature of Abl tyrosine kinase (Fig. 8.15). The enzyme

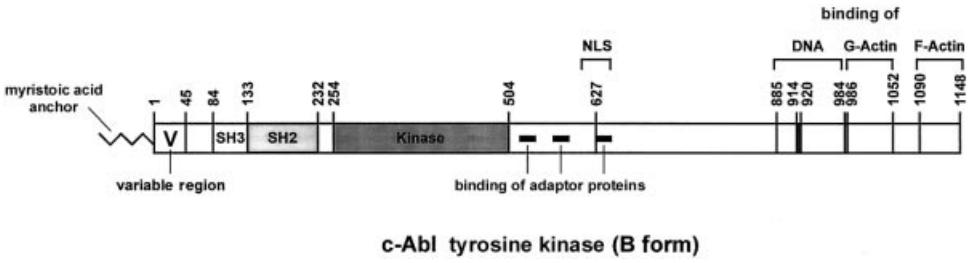



**Fig. 8.15.** Domain structure of Abl tyrosine kinase. The functionally characterized domains of Abl tyrosine kinase are shown in linear configuration. NLS, nuclear localization signal.

possesses a Tyr kinase domain, a SH2 and a SH3 domain, a nuclear localization signal, a DNA binding domain and binding domains for microfilament proteins, namely G-actin and F-actin (review: Feller, 1994). The largest fraction of the Abl kinase activitiy is found in the cell nucleus and it is thought, due to the occurrence of a DNA binding domain, that Abl kinase is active on DNA and chromatin. The C-terminal region of RNA polymerase II (see Chapter 1.2) and the retinoblastoma protein pRb (see Chapter 14) have been identified as Abl kinase substrates located in the nucleus. Abl kinase is also found in association with the plasma membrane and with the cytoskeleton. The complex structure and the variable cellular localization suggest that Abl kinase is involved in several signaling pathways.

Like many other nonreceptor tyrosine kinases, Abl tyrosine kinase may be converted by mutations into a dominant oncoprotein and thus contribute to tumor formation. Abl tyrosine kinase was first discovered as the oncogene of murine Abelson leukemia virus. Chronic myelogenic leukemia in humans is cause by a chromosome translocation in which a fusion protein is created of Abl tyrosine kinase and a Bcr protein (cf. Chapter 14). The result is a greatly increased tyrosine kinase activity, to which a causal role in occurrence of this leukemia is attributed.

## 8.4 Protein Tyrosine Phosphatases

The extent of tyrosine phosphorylation of signal proteins is determined both by the activity of the tyrosine kinases and also the activity of *tyrosine-specific protein phosphatases*. If the total activity of both enzymes in the cell is considered, it is found that there is a preponderance of protein tyrosine phosphatase activity compared to tyrosine kinase activity. In contrast, the activities of the Ser/Thr-specific protein kinases and protein phosphatases are approximately balanced. It is estimated that the activity of the protein tyrosine phosphatases is about 3–4 orders of magnitude higher than the activity of the protein tyrosine kinases. With this relationship between the activities, it is not surprising that the net level of tyrosine phosphorylation in the cell is very low and that tyrosine phosphorylation is often only transient. Consequently, it took a relatively long time until the importance of tyrosine phosphorylation for signal transduction was assessed correctly.

The protein tyrosine phosphatases are, like the protein tyrosine kinases, elementary components of many signal transduction pathways and, as such, are involved in control of cell-cell interactions, in signal transduction via growth factors and in regulation of the cell cycle (review: San and Tonks, 1994; Hunter, 1995, Neel and Tonks, 1997). The biological importance of the protein tyrosine phosphatases is underlined by the observation that defects in the activity of protein tyrosine phosphatases can lead to phenotypically demonstrable errors in function in higher eucaryotes. One example is the „moth-eaten" mutation in mouse, which is due to a defect in protein tyrosine phosphatase PTP1C. An error in splicing of the gene for the PTP1C protein heads to immunodeficiency and autoimmune disease in the mouse.

A medically important protein tyrosine phosphatase is found in the bacterium *Yersinia pestis*, the causative organism of plague. *Yersinia pestis* possesses a highly active protein tyrosine phosphatase (structure: Stuckey et al., 1994) which makes an important contribution to the pathenogenicity of this bacterium. The pathogen brings a protein tyrosine phosphatase into the host organism, which changes the steady state level of tyrosine phosphorylation and leads to extensive disturbance of cellular functions.

## 8.4.1 Structure and Classification of Protein Tyrosine Phosphatases

The protein tyrosine phosphatases (review: Neel and Tonks, 1997) can be roughly divided into two groups: receptor protein tyrosine phosphatases (also called receptor-like protein tyrosine phosphatases) and cytoplasmic protein tyrosine phosphatases. Both groups catalyze hydrolysis of Tyr phosphate by a common mechanism and correspondingly, both groups have a homologous catalytic domain. In addition to these two large groups, of which $> 75$ representatives have currently been identified, there are also tyrosine phosphatases with double specificity that hydrolyze both Tyr-P and also Ser-P. One example is CDC25 phosphatase (see Chapter 13). This group of enzymes is only related to the monospecific Tyr phosphatases to a low degree.

### Receptor Protein Tyrosine Phosphatases

The receptor-like protein tyrosine phosphatases have a transmembrane and, in some cases, a large extracellular domain with a very variable structure (Fig. 8.16). Many, but not all, membrane protein tyrosine phosphatases have two catalytic domains in the cytoplasmic region. The complete structure is very similar to the structure of transmembrane receptors. Understanding of their function is far from complete. Both the natural ligands and the substrate proteins following in the sequence are incompletely characterized. Several studies have demonstrated a role for receptor-like PTPs in neuronal cell adhesion signaling pathways. In cells of the neural tissue, a surface protein, contactin has been identified as ligand for the extracellular domain of a protein tyrosine phosphatase (Peles et al., 1995).

Interesting information on the mechanism of regulation by extracellular ligands was obtained from the crystal structure of the catalytic domain of PTPα, a receptor-like PTP. PTPα crystallizes as an inactive dimer in which structural elements of one subunit

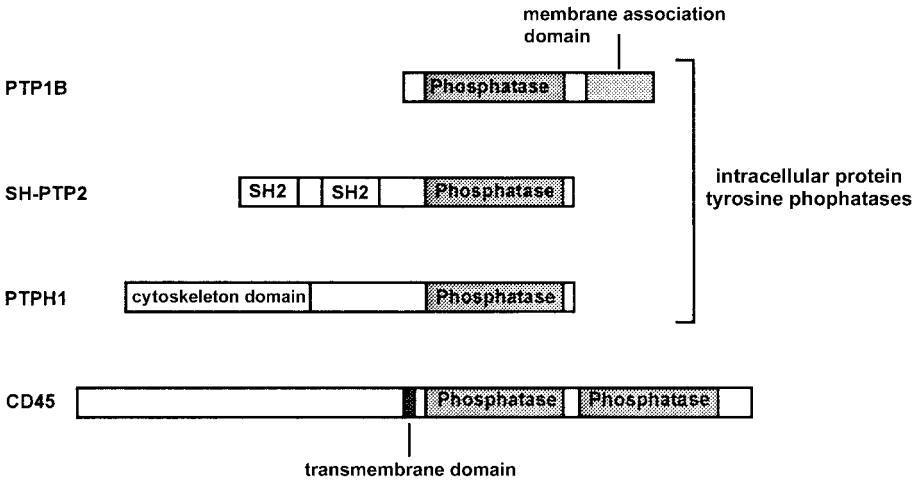



**Fig. 8.16.** Domain structure of protein tyrosine phosphatases. Linear representation of functional domains of the transmembrane tyrosine phosphatase CD45 and some cytoplasmic tyrosine phosphatases.

are located in the catalytic site of the other subunit and block the latter (Majeti et al., 1998). The structure suggests that binding of an extracellular ligand should promote oligomerization and inactivation of the receptor-like PTP. This still speculative inhibitory effect of ligand binding is in contrast to the activating effect of ligands on receptor tyrosine kinases.

### Cytoplasmic Protein Tyrosine Phosphatases

Cytoplasmically localized protein tyrosine phosphatases have a catalytic domain and other structural elements that specify the subcellular localization and association with effector molecules. These structural elements contain sequence signals for nuclear localization, for membrane association and for association with the cytoskeleton (see Fig. 8.16). The presence of SH2 domains suggests that these molecules might interact with signaling pathways involving growth hormones and receptor tyrosine kinases.

### Mechanism of Dephosphorylation

The catalytic center of the protein tyrosine phosphatases includes ca. 230 amino acids and contains the conserved sequence motif H/V-C-$(X)_5$-R-S/T-G/A/P (X is any amino acid) which is involved in phosphate binding and in catalysis and is part of a loop known as the P loop. The available structural data on the catalytic domains of protein tyrosine phosphatases indicate that the mechanism shown schematically in Fig. 8.17 is likely (see Tainer and Russell, 1994). The invariant Cys and Arg residues of the P loop have a central function in binding and cleavage of the phosphate residue.

The Cys residue exists as a thiolate, which carries out a nucleophilic attack on the phosphate of the phosphotyrosine residue. The thiolate is stabilized by the positively

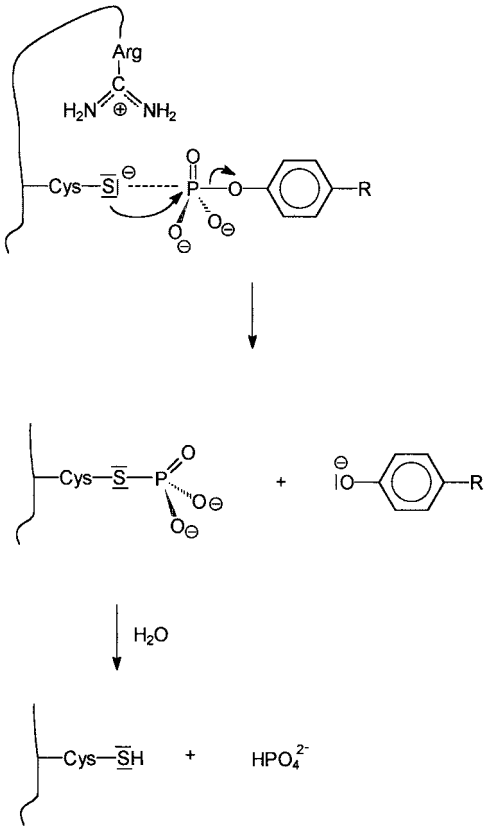



**Fig. 8.17.** Mechanism of hydrolysis of phosphotyrosine residues by tyrosine phosphatases. Cleavage of phosphate from phosphotyrosine residues takes place by an „in-line" attack of a nucleophilic cysteine thiolate of the tyrosine phosphatase at the phosphate of the phosphotyrosine residue. The negative charge on the thiolate is stabilized by the positive charge of a conserved Arg residue. In the course of the reaction, an enzyme-Cys-phosphate intermediate is formed, which is hydrolytically cleaved to phosphate and enzyme-Cys-SH. The figure shows selected interactions. Other interactions in the active center involved in substrate binding and catalysis are not shown. According to Tainer and Russel, (1994). R: substrate protein.

charged arginine residue. The Tyr residue is displaced by the thiolate via an „in-line" attack and an enzyme-bound Cys-phosphate is formed. Discharge of tyrosine from Tyr phosphate is facilitated by a catalytically essential Asp residue donating a proton to Tyr phenolate. The phosphate anion in Cys phosphate is stabilized by the Arg residue and by the positive dipole end of a neighboring α-helix. The phosphate is then released from the intermediate Cys phosphate by nucleophilic attack of a water molecule.

### 8.4.2 Cooperation of Protein Tyrosine Phosphatases and Protein Tyrosine Kinases

The cellular functions of protein tyrosine phosphatases are closely associated with signal transduction via protein tyrosine kinases. The growth and differentiation promoting signals mediated by protein tyrosine kinases include autophosphorylations and phosphorylation of effector proteins (see 8.1.4). According to current ideas, the activity of protein tyrosine phosphatases may have a negative or positive influence on signal transduction via protein tyrosine kinases. On the one hand, protein tyrosine

phosphatases may have an *antagonistic* effect on the activity of protein tyrosine kinases; on the other hand, they may *positively cooperate* with signal transduction via protein tyrosine kinases (review: San and Tonks, 1994; Hunter, 1995).

### Antagonism of Protein Tyrosine Phosphatases and Protein Tyrosine Kinases

A schematic representation of how protein tyrosine phosphatases influence signal transduction via protein tyrosine kinases in a negative way is shown in Fig. 8.18a.

A damping effect of protein tyrosine phosphatases on signal transduction may occur, for example, via cleavage of a phosphate residue from an activated receptor tyrosine kinase that has undergone autophosphorylation following ligand binding. Numerous experiments suggest a damping effect of PTP1 on the Janus kinase Jak1 in signal transduction via cytokine receptors. Furthermore, phosphorylated effector molecules of receptor tyrosine kinases, e.g., the insulin receptor substrate IRS, are considered as substrates for dephosphorylation. The activated insulin receptor phosphorylates IRS protein at several Tyr residues. The phosphotyrosine residues so created are of great importance for transmission of the signal further, as they serve as attachment points for SH2-containing signal molecules such as PI3-kinase or the adaptor protein Grb2. It is easy to imagine, although not yet proven, that the action of protein tyrosine phosphatases effects a damping influence in this signal transduction and limits the duration of the signal.

There are numerous examples showing that misregulation of tyrosine kinases plays a decisive role in tumor formation, and that mutations of the genes of tyrosine kinases can convert these into oncogenes. It has therefore always been hypothesized that protein tyrosine phosphatases may play the role of tumor suppressors (see Chapter 14). Loss of the damping function of protein tyrosine phosphatases in signal transduction can bring about an uncoordinated increase in tyrosine phospharylation and thus uncontrolled growth. An inhibitory activity of protein tyrosine phosphatases on formation of tumors has been described many times. For example, protein tyrosine phosphatase PTP1 inhibits tumor transformation by the *Neu* oncogene. Protein tyrosine phosphatase PTP1C is also attributed a function in control of cell proliferation and differentiation (see San and Tonks, 1994).

### Positive Cooperation between Protein Tyrosine Phosphatases and Protein Tyrosine Kinases

Protein tyrosine phosphatases may also carry out a positive regulatory function in a signal transduction by activating protein tyrosine kinases. An example for this regulating mechanism is Src tyrosine kinase. As already explained above (8.2.1), phosphorylation of Tyr527 of Src kinase is linked with inhibition of the kinase activity. The SH2 domain of Src kinase binds in an intramolecular reaction to the Tyr phosphate at the C terminus, leading to blocking of the active center. Activation of Src kinase may be brought about by cleaving off the inhibitory phosphate residue.

Signal transduction via the receptor-like CD45 protein tyrosine phosphatase in cells of the blood forming system requires its intracellularly localized phosphatase activity. The cytoplasmic tyrosine kinases $p56^{lck}$ and $p59^{fyn}$ are thought to be cellular substrates

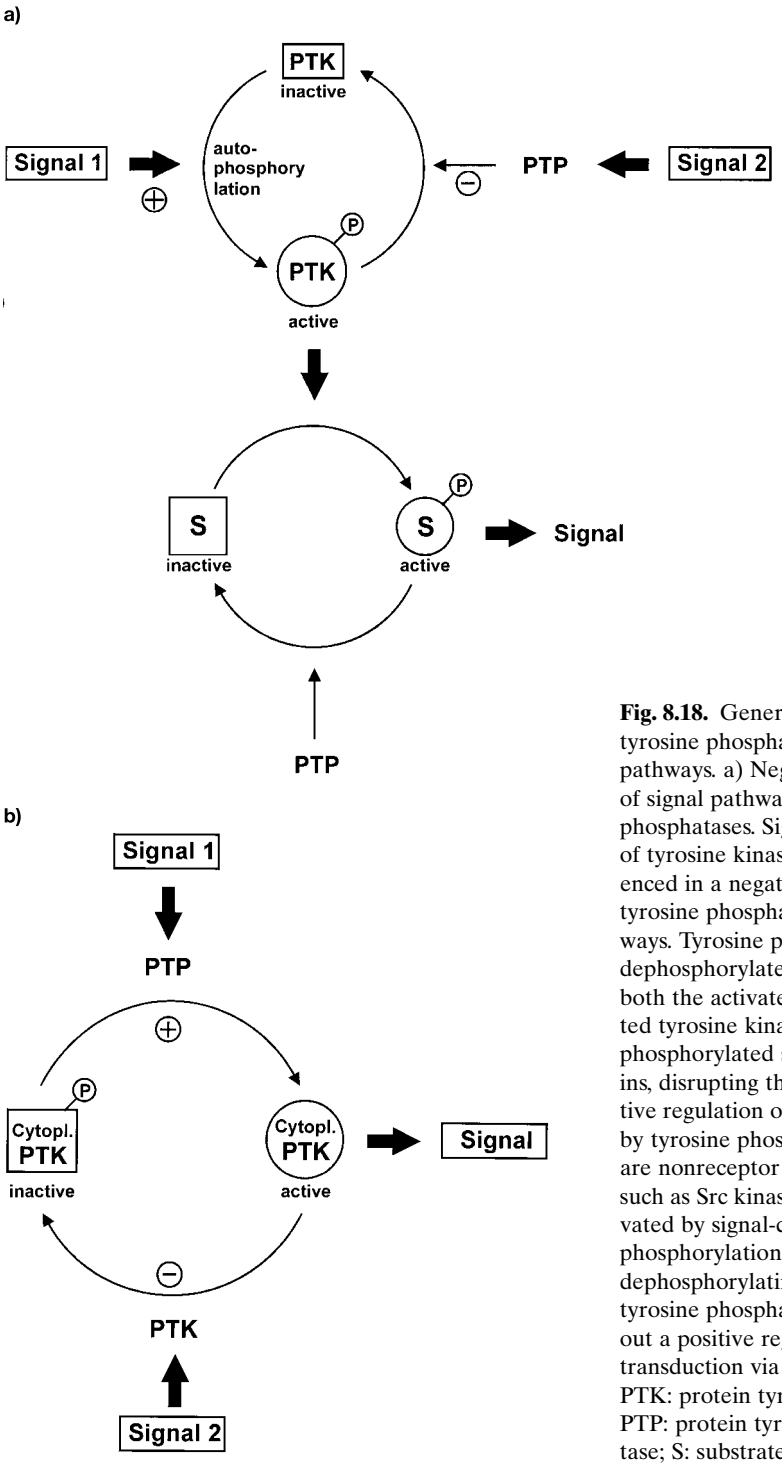



**Fig. 8.18.** General functions of tyrosine phosphatases in signal pathways. a) Negative regulation of signal pathways by tyrosine phosphatases. Signal transduction of tyrosine kinases may be influenced in a negative manner by tyrosine phosphatases in two ways. Tyrosine phosphatases may dephosphorylate and inactivate both the activated, phosphorylated tyrosine kinase and also the phosphorylated substrate proteins, disrupting the signal. b) Positive regulation of signal pathways by tyrosine phosphatases. There are nonreceptor tyrosine kinases, such as Src kinase, that are inactivated by signal-controlled Tyr phosphorylation. In this case, the dephosphorylating activity of tyrosine phosphatases can carry out a positive regulation of signal transduction via tyrosine kinases. PTK: protein tyrosine kinase; PTP: protein tyrosine phosphatase; S: substrate protein.

of CD45 phosphatase. It is assumed that both are activated by CD45 phosphatase. The synergistic action of Tyr phosphatases and Tyr kinases is shown schematically in Fig. 8.18b.

## 8.4.3 Regulation of Protein Tyrosine Phosphatases

The fact that cellular activity of protein tyrosine phosphatases by far exceeds that of protein tyrosine kinases suggests that there is strict control of the dephosphorylation rate in a cell. The mechanisms are those already highlighted in previous chapters as central elements of regulation of activity of signal molecules.

### Subcellular Localization

The subcellular localization of protein tyrosine phosphatases is an important aspect of their function. The sequences of cytoplasmic protein tyrosine phosphatases frequently demonstrate sequence signals specifying a particular subcellular localization. This ensures that protein tyrosine phosphatases are only active at defined subcellular sites. The presence of SH2 domains in cytoplasmic protein tyrosine phosphatases also shows that these are coupled, via SH2-phosphotyrosine interactions, to specific substrates, where they then perform their actual function.

### Ser/Thr Phosphorylation

Another mechanism of regulation of protein tyrosine phosphatases is via Ser/Thr phosphorylation. Specific phosphorylation of protein tyrosine phosphatases by Ser/Thr-specific protein kinases of types A and C has been reported (see Neel and Tonks, 1997). This observation indicates the possibility that signal transductions via Ser/Thr kinases and via Tyr kinases/phosphatases may cooperate and that different signal pathways may be crosslinked in this way.

### Tyr Phosphorylation

Protein tyrosine phosphatases may themselves be the substrate of protein tyrosine kinases and may be phosphorylated at specific Tyr residues. For example, Syp tyrosine phosphatase is phosphorylated at Tyr542 on binding to the activated, autophosphorylated PDGF receptor. This phosphorylation site is a consensus binding sequence for the SH2 group of the adaptor protein Grb2. It may contact other signal proteins via a SH3 domain and pass the signal onwards (Fig. 8.19). Syp phosphatase thus has two distinctly different possibilities to transmit the signal further. On the one hand, the enzyme can involve adaptor molecules in signal transduction; on the other hand, specific substrates can be dephosphorylated.

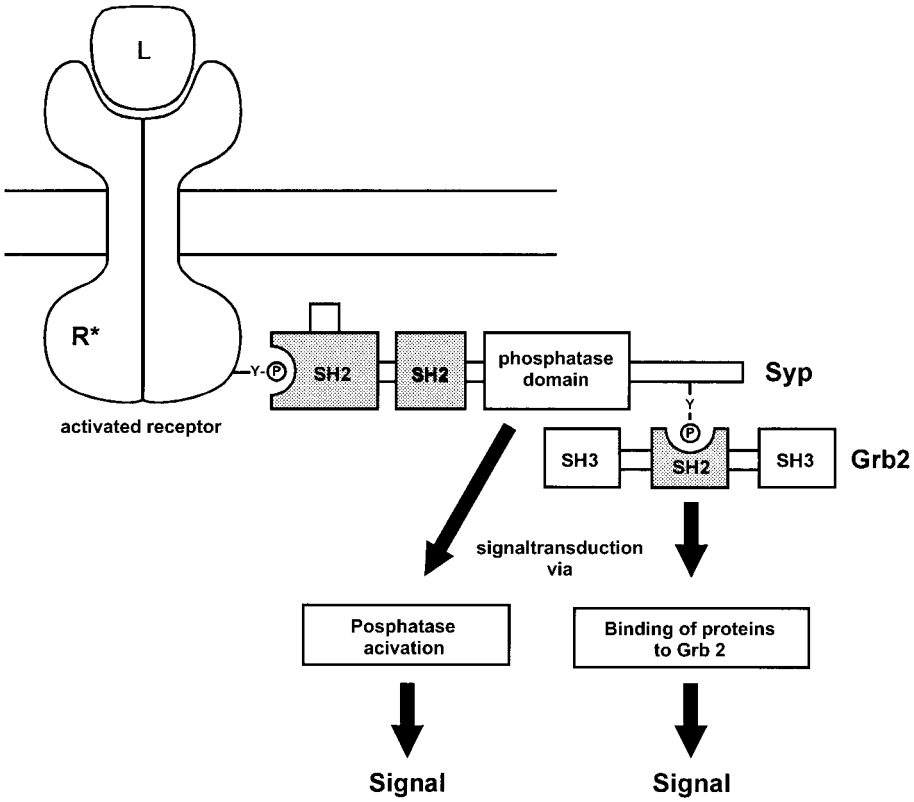



**Fig. 8.19.** Model of signal transduction via Syp phosphatase. Syp phosphatase has two SH2 domains that mediate binding to the activated, autophosphorylated PDGF receptor. Association of Syp phosphatase with the activated receptor has a two-fold effect. The tyrosine kinase activity of the receptor phosphorylates Syp phosphatase at Tyr542. The SH2 domain of the adaptor protein Grb2 binds to the P-Tyr542; Grb2 can bring other signal proteins into the signal transduction process. Furthermore, Syp phosphatase is activated by association with the receptor and can dephosphorylate phosphorylated signal proteins and pass the signal on in this way. According to Hunter, (1995).

## 8.5 Adaptor Molecules of Intracellular Signal Transduction

Signal transduction processes starting from activated transmembrane receptors often have the aim of bringing together particular protein components of the signaling pathway to a defined site in the cell, e.g., to bring an enzyme into the immediate vicinity of its substrate. The cell uses so-called adaptor proteins to bring signal molecules together in a targeted fashion; these adaptor molecules help to decide where and when a certain enzyme, such as a protein kinase, will become active. The adaptor proteins do not have any enzymatic function themselves, but rather they function as a connecting link bet-

ween different signal proteins, mediating specific spatial neighborhood in signal conduction (for review see: Pawson and Scott, 1997). Furthermore, adaptor proteins serve in targeted subcellular localization of signal events. They are an organizational element in signal conduction in that they help to assemble multiprotein complexes of signal conduction at specific subcellular sites, enabling spatially concentrated, and thus site-specific, signals to be created. The specificity and regulation of signal conduction are increased since only certain signal proteins can associate with the adaptor protein.

The occurrence of multiple protein modules is characteristic for adaptor proteins; these mediate different protein-protein interactions and can thus bring about crosslinking of signal proteins.

The function of many adaptor proteins is closely linked with the cell membrane or with the cytoskeleton. The occurrence of PH domains and myristoyl modifications suggests that adaptor proteins are involved in particular in coordination and assembly of signal complexes on the inner side of the cell membrane.

Adaptor proteins are, above all, important elements for controlling the subcellular organization of Tyr and Ser phosphorylation events. Thus, many adaptor proteins contain PTB or SH2 domains that direct specific interactions with autophosphorylation sites on an activated receptor.

Fig. 8.20 shows the schematic composition of some important adaptor molecules. Examples of well-characterized adaptor proteins are the Grb2 protein and Crk protein. Both proteins are predominantly composed of SH2 and SH3 domains and in both, enzyme activity has not been detected. Grb2 protein has two SH3 groups and a SH2 group. Two subtypes of Crk protein are known (see Fig. 8.20), with one or two SH3 groups and a SH2 group (review: Feller et al., 1994). Since the two proteins have two different binding valences, due to the presence of both SH2 and SH3 groups, they can couple and bring together proteins that possess phosphotyrosine residues and Pro-rich sequences.

The *Grb2 protein* (*Grb: growth factor receptor binding protein*) has been identified as a component of signal transduction of growth factors and the Ras signaling pathway (see Chapter 9). The adaptor protein Shc, the EGF receptor, the PDGF receptor and Syp phosphatase have been described as binding partners of the SH2 domain of Grb2

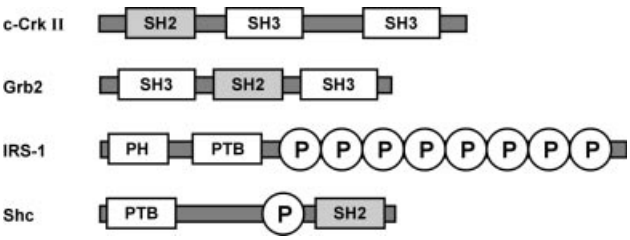



**Fig. 8.20.** Modular composition of adaptor proteins. Adaptor proteins do not show any enzyme activity of their own, but rather they contain protein modules which help to bind signal proteins into signal pathways. IRS-1: insulin receptor substrate 1; PTB: phosphotyrosine binding domain; PH: pleckstrin homology domain; P: phosphotyrosine-containing binding site for SH2 or PTB domains; HLH: helix-loop-helix DNA binding motif.

protein. Grb2 protein is tightly bound via its SH3 domain to the Pro-rich domain of the GTP-GDP exchange factor Sos, which can pass the signal by nucleotide exchange to the Ras protein (Fig. 8.19, see also Chapter 9). In the form of the Grb2-Sos complex, Grb2 protein functions to generate a coupling between the activated receptor tyrosine kinase and the Ras protein. The membrane association of the Sos protein is necessary for its function as a nucleotide exchange factor (GEF) in the Ras signaling pathway (see Chapter 9).

The *Crk protein* was first discovered as the transforming principle of the retroviruses CT10 and ASV-1. Abl tyrosine kinase is under discussion as a binding partner of the SH3 domain of Crk (Feller et al., 1994). Possible binding partners of the SH2 domain have been described but their physiological function is unclear.

The *insulin receptor substrate IRS* couples the insulin receptor to sequential effector molecules (review: Ogawa et al., 1998). On binding of insulin to the insulin receptor, the tyrosine kinase activity of the receptor is stimulated. The IRS protein is phosphorylated at several Tyr residues, which then serve as attachment points for sequential effector molecules as e.g. the Grb2-mSos complex, the PI3-kinase and the protein tyrosine phosphatase SHP-2. The IRS protein also has a phosphotyrosine binding domain and a PH domain. Both modules are required for signal transduction *in vivo*. It is assumed that the PTB domain binds to autophosphorylation sites of the insulin receptor and that the PH domain is involved in membrane association of IRS.

The protein *PSD-95* is an example of a PDZ-containing protein (review: Craven and Bredt, 1998). PSD-95 is found in postsynaptic cells where, via its PDZ domains, it mediates interactions with intracellular domains of receptors such as the NMDA receptor (see 16.4.2.1). The *InaD protein* which is composed solely of PDZ domains has an adaptor function in the vision process in *Drosophila* (see 8.2.5).

# References Chapter 8

Cohen, G.B., Ren, R. and Baltimore,D. 'Modular binding domains in signal transduction proteins' (1995) Cell *80*, 237–248

Craven, S.E., and Bredt, D.S. 'PDZ proteins organize synaptic signaling pathways' (1998) Cell *93*, 495–498

Fantl, W.D., Johnson, D.E. and Williams, L.T. 'Signalling by receptor tyrosine kinases' (1993) Ann. Rev. Biochem. *62*, 453–481

Feller, S.M. Ren, R., Hanafusa, H. and Baltimore, D. 'SH2 and SH3 domains as molecular adhesives: the interactions of Crk and Abl' (1994) Trends Bioch. Sci. *19*, 453–458

Fischer, E.H., Charbonneau, H. and Tonks, N.K. 'Protein tyrosine phosphatases: a diverse familiy of intracellular and transmembrane enzymes' Science (1991) *253*, 401–406

Heldin, C.H. 'Dimerization of cell surface receptors in signal transduction' (1995) Cell *80*, 213–233

Hubbard, S.R., Wei, L., Ellis, L. and Hendrickson, W.A. 'Crystal structure of the tyrosine kinase domain of the human insulin receptor' (1994) Nature *372*, 746–754

Hunter, T. 'Protein tyrosine phosphatases: the yin and yang of protein phosphorylation and signalling' (1995) Cell *80*, 225–236

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Lemmon, A. and Schlessinger, J. 'Regulation of signal transduction and signal diversity by receptor oligomerization' (1994) Trends Bioch. Sci. *19*, 459- 463

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

Majeti, R., Bilwes, A.M., Noel, J.P., Hunter, T. and Weiss, A. 'Dimerization-induced inhibition of receptor protein tyrosine phosphatase function through an inhibitory wedge' (1998) Science *279*, 88–91

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

Neel, B.G. and tonks, N.K. 'Protein tyrosine phosphatases in signal transduction' (1997) Curr. Op. Cell Biol. *9*, 193–204

Ogawa, W., Matozaki, T. and Kasuga, M. 'Role of binding proteins to IRS-1 in insulin signalling' (1998) Mol Cell Biochem *182*, 13–22

Parson, J.T., and Parson, S.J. 'Src family protein tyrosine kinases: cooperating with growth factor and adhesion signaling pathways' (1997) Curr. Op. Cell Biol. *9,* 187–192

Pawson, T. 'Proteine modules and signalling networks' (1995) Nature *373*, 573–579

Pawson, T., Scott, J.D. 'Signaling through scaffold, anchoring, and adaptor proteins' (1997) Science 278, 2075–2080

Peles, E., Nativ, M., Campbell, P.L., Sakurai, T., Martinez, R., Lev, S., Clary, D.O., Schilling, J., Barnea, G. and Plowman, G.D. 'The carbonic acid anhydrase domain of receptor tyrosine phosphatase beta is a functional ligand for the axonal cell recognition molecule contactin' (1995) Cell *82*, 251–260

San, H. and Tonks, N. ' The coordinated action of protein tyrosine phosphatases and kinases in cell signalling' (1994) Trends Bioch. Sci. *19*, 480–485

Shokat, K.M. 'Tyrosine kinases: modular signaling enzymes with tunable specificities' (1995) Chem.Biol. *2*, 509–514

Sicheri, F. and Kuriyan,J. ' Structure of Src-family tyrosine kinases' (1997) Curr. Op. Struct. Biol. *9*, 777–785

Songyang, Z. and Cantley, L.C. 'Recognition and specificity in protein tyrosine kinase-mediated signalling' (1995) Trends Bioch. Sci. *20*, 470–475

Stuckey, J-A., Schubert, H.L., Baumann, E.B., Zhang, Z., Dixon, J.E. and Saper, M.A. 'Crystal structure of Yersinia protein tyrosine phosphatase at 2.5 Å and the complex with tungstate' (1994) Nature *370*, 571–575

Sudol, M. ' From Src homology domains to other signaling modules: proposal of a protein recognition code' (1998) Oncogene, *17*, 1469–1476

Tainer, J. and Russell, P. 'Cracking tyrosine phosphatases' (1994) Nature *370*, 506–507

Taniguchi, T. 'Cytokine signaling through nonreceptor protein tyrosine kinases' (1995) Science *268*, 251–255

Tsunoda, S., Sierralta, J., Sun, Y., Bodner, R., Suzuki, E., Becker, A., Socolich, M. and Zuker, C.S. 'A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade' (1997) Nature *388*, 243–249

Ullrich, A. and Schlessinger, J. 'Signal transduction by receptors with tyrosine kinase activity' (1990) Cell *61*, 203–212

Van der Geer, P and Pawson, T 'The PTB domain: a new protein module implicated on signal transduction' (1995) Trends Bioch. Sci. *20*, 277–280

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

Yaffe, M.B., Schutkowski, M., Shen, M., Zhou, X.Z., Stukenberg, P.T., Rahfeld, J.U., Xu, J., Kuang, J., Kirschner, M.W., Fischer, G., Cantley, L.C. and Lu, K.P. 'Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism' (1997) Science *278*, 1957–1960

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 9**

# Signal Transmission via Ras Proteins

## 9.1 General Importance and Classification of Ras Proteins

Intracellular signal transduction employs central switching stations that receive, modulate and transmit signals further. The Ras proteins (also known as $p21^{ras}$ proteins) make up a switching station of particular importance for growth and differentiation processes. The Ras proteins process signals received by receptor tyrosine kinases, by receptors with associated tyrosine kinase activity and by G-protein-coupled receptors, and transmit these into the cell interior (Fig. 9.1).

The general importance of Ras proteins in growth regulation was recognized at the beginning of the 1980s, when it was demonstrated that close to 30 % of all solid tumors in humans show a mutation in the Ras gene.

Interest in the structure and function of Ras proteins was kindled especially because it was established that certain positions in the Ras protein are particularly sensitive to oncogenic mutations. Replacement of Gly12 in the Ras protein with any of the other natural amino acids (except Pro) leads to an increase in the tumor-transforming potential of Ras protein. Its small size and apparently simple function made the Ras protein an ideal subject for investigation of structure-function relationships of a central signal protein. Ras protein is currently the best-characterized signal protein. The extensive structural information available on Ras protein has helped to understand molecular mechanisms of signal transduction and the cause for tumor-initiating misregulation at the molecular level.

The Ras gene, and the 21 kDa gene product for which it codes, the Ras protein or $p21^{ras}$, were first found in retroviruses that trigger sarcoma-type tumors in rats (*Ras* = *rat sarcoma*). Mammals have at least four different Ras genes: ***H****(arvey)*-***ras, Ki****-(rsten)*-***rasA, Ki-rasB*** and the ***N****(euroblastoma)*-***ras*** gene. Each of these genes can be activated to an oncogene by a mutation. The four genes are closely related and the proteins coded for represent the Ras protein in a narrow sense. Most of the structural and biochemical data are available for the H-Ras protein. This is referred to in the following as *„the Ras protein“* for simplicity. It is assumed that the basic knowledge of structure and function of the H-Ras protein also apply to the other Ras proteins.

Following the discovery of the Ras protein, it was quickly established that Ras proteins are a family within a large superfamily, known today as the *Ras superfamily of monomeric GTPases*. The members of the superfamily of Ras proteins are regulatory GTPases of 16–25 kDa, which are active as monomers.

Within the Ras superfamily, at least five subfamilies can be differentiated (Boguski and McCormick, 1993 and Table 9.1). These are the *Ras/Rap*, the *Rho/Rac*, the *Rab, Ran* and *Arf* subfamilies.

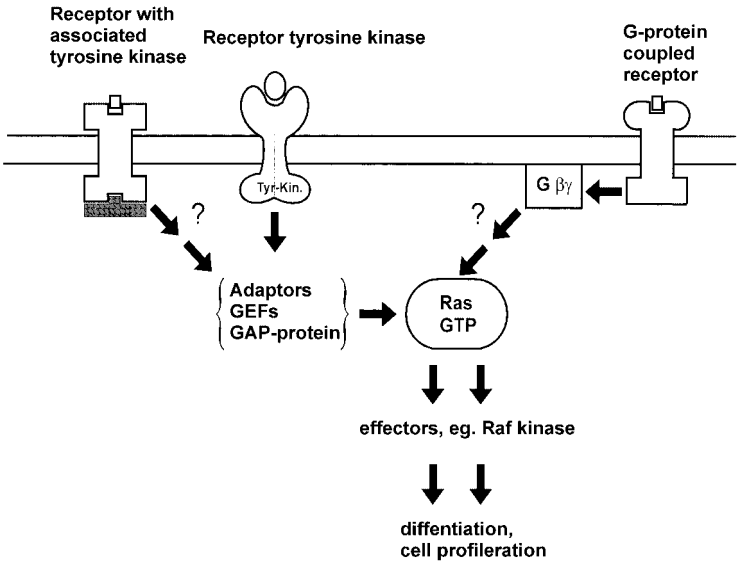



**Fig. 9.1.** The Ras protein as a central switching station of signaling pathways. A main pathway for Ras activation is via receptor tyrosine kinases, which pass the signal on via adaptor proteins and guanine nucleotide exchange factors to the Ras protein. Activation of Ras protein can also be initiated via G-protein-coupled receptors and via transmembrane receptors with associated tyrosine kinase activity. The membrane association of the Ras protein (see Fig. 9.6) is not shown for clarity. In addition, not all signaling pathways that contribute to activation of the Ras protein are shown, nor are all effector reactions. Gβγ: βγ-complex of the heterotrimeric G proteins; GAP: GTPase activating protein; GEF: guanine nucleotide exchange factor.

The members of the Ras superfamily are regulatory GTPases that can exist in an inactive GDP form and an active GTP form. In the GTP-bound state, Ras proteins interact with downstream targets, the Ras effectors. These communicate, in turn, with other signal proteins localized downstream in the signal chain. Effectors are distinguished in that they bind more strongly to the Ras · GTP-form than to the GDP-form. Binding of effectors to the Ras protein is terminated by hydrolysis of the protein-bound GTP.

The Ras proteins run through the unidirectional GTPase cycle as regulatory GTPases (Fig. 9.2). The signal-transmitting function of Ras proteins and Ras-related proteins can be regulated by three mechanisms, which were already presented in 5.4.1 together with the regulatory GTPases:

### GTPase Activating Proteins

The lifetime of the active GTP-bound state may be reduced by regulatory *GTPase activating proteins.* The primary function of the GTPase activating proteins (GAP) is to negatively regulate the Ras proteins and Ras-related proteins.

The GTPase activating proteins show specificity for a particular subfamily within the Ras superfamily. Furthermore, there are several different GTPase activating proteins

**Table 9.1.** Regulatory GTPases and effector proteins of the Ras superfamily of mammals. (according to Boguski and McCormick, 1993).

| Ras family | | GEF, gene or protein name | GAP, gene or protein name |
|---|---|---|---|
| Ras | H-Ras | Ras-GEF, mSos | Ras-GAP; neurofibromin, p120-GAP |
| | N-Ras | | |
| | Ki-Ras A | | |
| | Ki-Ras B | | |
| | Rap subfamily: | | |
| | Rap1A, 1B, 2A, 2B, | | Rap1-GAP |
| | RalB | Ral-GEF | |
| | TC21 (= k-Rev1) | | |
| | R-Ras | | |
| Rho/Rac | Rho A, B, C | Dbl | Bcr, reports of at least 8 others, (Lamarche, 1994) |
| | Rac1, Rac2 | | |
| | TC10 | | |
| Rab | at least 24 different Rab proteins | MSS4 | |
| Ran | Ran,TC4 | RCC1 | |
| ARF (ADP-ribosylation factor) | ARF1–6 | | |

for the individual Ras proteins or Ras-related proteins. Thus, in mammals, there are at least nine different GTPase activating proteins, which show specificity for the subfamily of Rho proteins. It will be interesting to see how the activity of the different GTPase activating proteins is regulated and balanced in the cell, and which specific functions are assigned to individual representatives.

### Guanine Nucleotide Exchange Factors

The transition from inactive GDP state to active GTP state may be accelerated by proteins that cause the bound GDP to dissociate. The *guanine nucleotide exchange factors* (GEF) play an essential role in signal transduction via Ras proteins. Loss of exchange activity by mutation of the exchange proteins has the same effect in lower organisms as loss of the Ras gene.

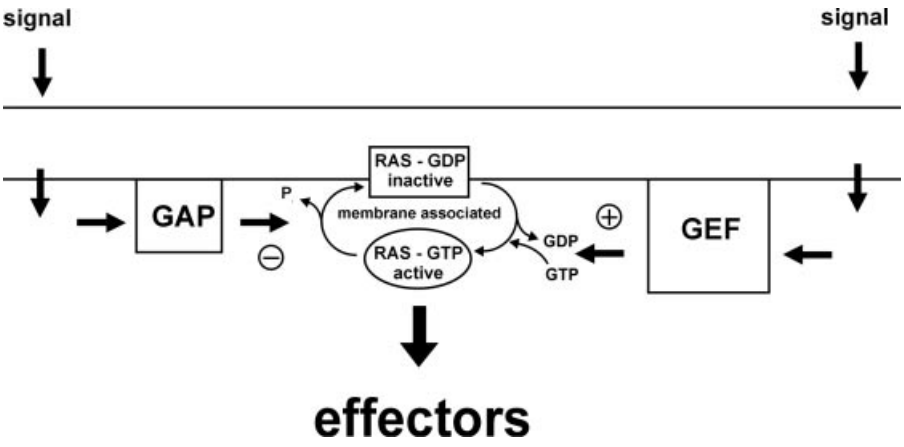



**Fig. 9.2.** The GTPase cycle of the Ras protein. Conversion of the inactive Ras · GDP complex into the active Ras · GTP complex is brought about by guanine nucleotide exchange factors (GEFs). The activated state of the Ras protein is terminated by hydrolysis of the bound GTP. The help of a GTPase actvating protein (GAP) is required, due to the intrinsically slow GTPase activity of the Ras protein. Ras protein performs all its functions in close association with the cell membrane. It carries a membrane anchor and the effector proteins preceding and following in sequence are also associated with the membrane.

**Inhibitors of Guanine Nucleotide Dissociation**

For the Rab and Rho/Rac families, proteins have been described that bind the GDP form of GTPase and prevent dissociation of GDP. Proteins with these characteristics are known as *guanine nucleotide dissociation inhibitors* (*GDI*). The GDIs are localized in the cytosol and bind to Ras-like proteins modified with a prenyl residue. Their function is thought to be to extract the Rab and Rho/Rac proteins from the membrane and to prepare a cytosolic pool of these proteins. In this way, GDIs may prevent early dissociation of the bound GDP and premature nucleotide exchange during intracellular translocation of GTPases.

## 9.2 Structure and Biochemical Properties of Ras Protein

The Ras protein is a monomeric GTPase of ca. 21 kDa. The GTP-bound form represents the active, switched-on state; the GDP-bound form is the inactive, switched-off state (review: Wittinghofer et al., 1993; Schweins and Wittinghofer, 1994; Wiesmüller and Wittinghofer, 1994). The transition between the active and inactive forms occurs in a unidirectional cycle (see Fig. 9.2). Highly resolved crystal structures exist for both forms of Ras protein and for oncogenic mutants of Ras protein. Furthermore, structural information is available on binding of a Ras-related protein, the Rap protein, with the effector Raf kinase (Nassar et al., 1995) and on binding of a GAP protein (Scheff-

zek et al., 1997). 3D-structures are also available for other members of the Ras superfamily (review: Geyer and Wittinghofer, 1997).

The *lifetime* of the active, GTP-bound state of the Ras protein is of great importance for the signal transducing function of Ras protein. Only in the GTP state can the signal be transmitted to the effector molecule next in sequence. The time window available for signal transduction is determined by the rate of GTP hydrolysis. A low rate of GTP hydrolysis and consequently longer occupancy of the active GTP state is associated with a high intensity of signal transduction. Reduction of the time window by stimulation of GTPase leads to weakened signal transduction.

Considered in isolation, the Ras protein is a very inefficient enzyme. On the one hand, the rate of GTP hydrolysis is very low; on the other hand, the complex of Ras protein and GDP is very stable and only dissociates very slowly. The rate constants of both processes are in the region of $10^{-4}$ $sec^{-1}$. Both reactions may be accelerated in the process of signal transduction, however, and have a decisive influence on signal transduction via the Ras protein.

**Activation of GTPase**

Under the influence of GTPase-activating proteins, the rate of GTP hydrolysis of the Ras protein may be increased up to $10^5$-fold. The GTPase-activating proteins control the activity state of Ras protein by drastically reducing the lifetime of the active GTP state. Due to this property, they function as negative regulators of the Ras protein.

**Acceleration of Nucleotide Exchange**

The rate of GDP dissociation is subject to control by *G-nucleotide exchange proteins* (*GEFs*), which promote dissociation of bound GDP, and thus function as positive regulators of the Ras protein.

An important step forward in the understanding of the molecular basis of the Ras function was achieved when structural information and biochemical and genetic data became available for the following questions:

- How does the active GTP state differ from the inactive GDP state?
- What is the mechanism for GTP hydrolysis? Which residues of the Ras protein are involved in the GTPase reaction?
- What is the mechanism of activation of effector molecules such as the Raf kinase?
- How does the occurrence of oncogenic mutations favored at particular positions fit into the structural picture of the Ras function?

### 9.2.1 Structure of the GTP- and GDP-bound Forms of Ras Protein

The structure of the GTP-bound form of the Ras protein is shown in Fig. 9.3.

The Ras protein, as a regulatory GTPase, shows the G domain typical for the superfamily of regulatory GTPases (see Fig. 5.12). The sequence motives characteristic for regulatory GTPases (cf. 5.3.3) are involved in binding the nucleotide and $Mg^{2+}$. Three

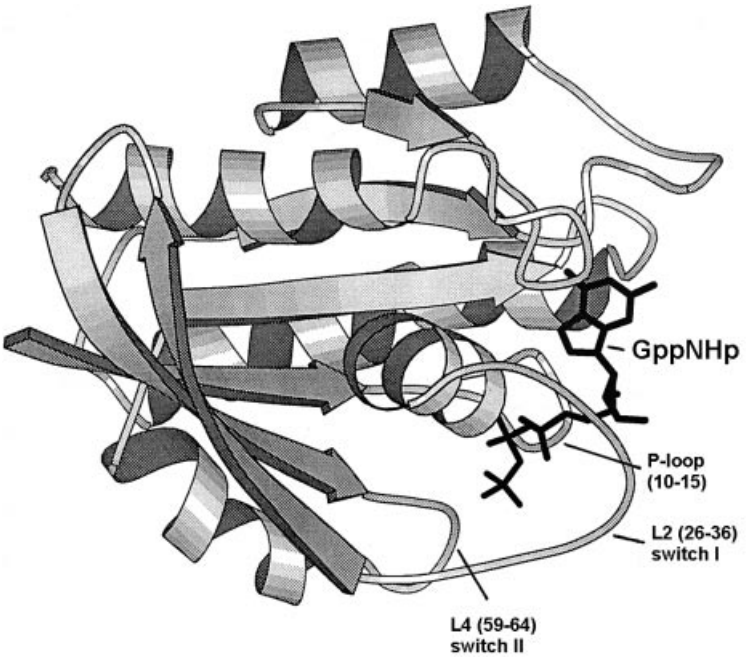



**Fig. 9.3.** Structure of the GTP form of the Ras protein. Crystal structure of the Ras protein in complex with the GTP analog βγ-imino-GTP(GppNHp). The figure shows the P loop, in which Gly12 is located, and the L2 and L4 loops, which have a switch function in GTP hydrolysis. The numbers give the sequence positions of amino acids in the loops. According to Pai et al., (1989). MOLSKRIPT representation according to Kraulis, (1991).

structural elements are of particular importance for the switch function of Ras protein: the L1, L2 and L4 loops. All three loops contact the γ-phosphate of GTP (Pai et al., 1989; Wittinghofer et al., 1993).

The L1 loop winds around the β and γ phosphates; it is also know as the P loop. Gly12 is located in the L1 loop; this amino acid is frequently mutated in oncogenic mutants of Ras protein (mutation „hotspot“).

The L2 loop is of particular importance for the biological activity of the Ras protein. L2 contains the amino acids 32–36, which were recognized in mutation experiments as important for the interaction with GTPase-activating proteins. Thr35 of the L2 loop also participates in H-bridge contacts to the γ-phosphate and to $Mg^{2+}$, and thus helps to correctly position the GTP-$Mg^{2+}$ complex in the active center.

The L4 loop also forms a H-bridge via Gly60 to the γ-phosphate. L4 corresponds to *switch II* (residues 59—76). It contains the catalytically essential Gln61 residue and is also involved in the interaction with GTPase-activating proteins.

The L2 loop (residues 32–37) corresponds to *switch I* of the Gα-subunits (see 5.5.6); it is known as the *effector loop*. It is an important part of the effector domain of the Ras protein, and signals are received and passed on via this domain.

Comparison of the structure of the active Ras · GTP form with the inactive Ras · GDP form indicates signicant conformational changes in the loops. As a cons-

quence of GTP hydrolysis, the contacts of all three loops to the γ-phosphate are lost. The structural difference between the active GTP form and the inactive GDP form can be described by *position changes of switches I and II* (Fig. 9.4). Switch I experiences a drastic reorientation since the coordination to the γ-phosphate and to $Mg^{2+}$ are lost. The interaction of the conserved Gly60 in switch II with the γ-phosphate is also lost in the process of GTP hydrolysis.

It is not surprising that residues corresponding to switch I and switch II, which define the conformational differences between the inactive GDP form and the active GTP state of Ras, are involved in recognition of the Ras effectors, the immediate downstream components in the Ras signaling pathway (see 9.6 and 9.7). Residues 32–40 comprise the core Ras effector domain, which is essential for all effector interactions.

## 9.2.2 GTP Hydrolysis: Mechanism and Stimulation by GAP Proteins

In the high resolution crystal structure of the GTP form of Ras protein, a tightly bound water molecule is visible located in an optimal position for nucleophilic attack on the γ-phosphate (Wittinghofer et al., 1993). The water molecule is fixed in a defined position by H-bridges with Gln61 and Thr35. As described in 5.4.4 for the α-subunits of the heterotrimeric G-proteins, GTP hydrolysis takes place by an „in-line" attack of the nucleophilic water molecule on the γ-phosphate, for which a pentagonal, bipyramidal transition state is postulated.

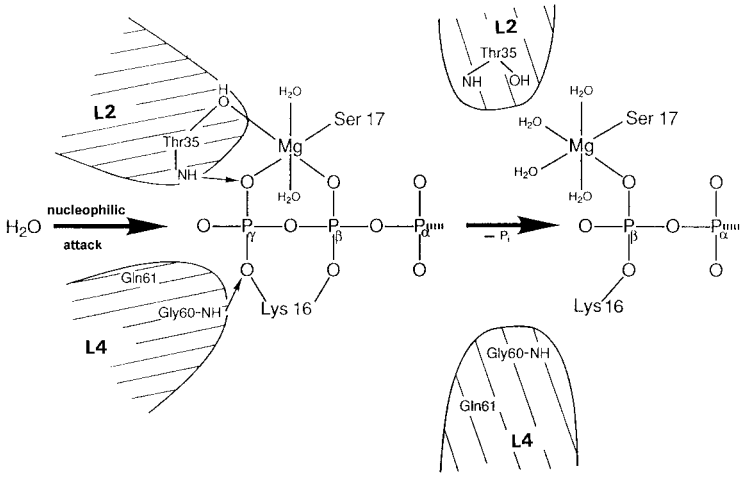
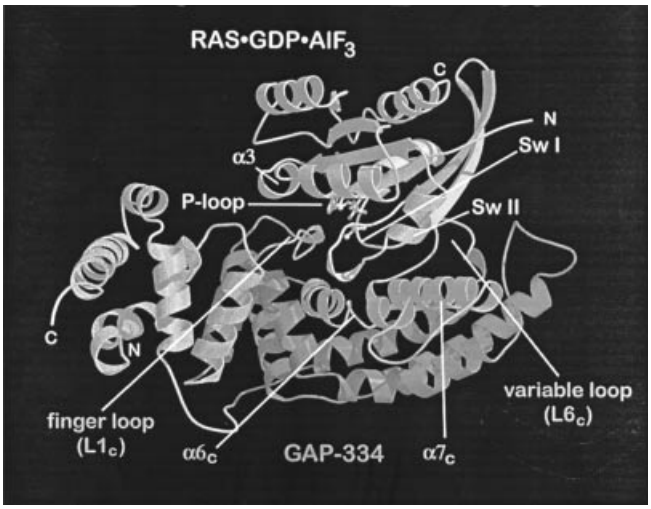
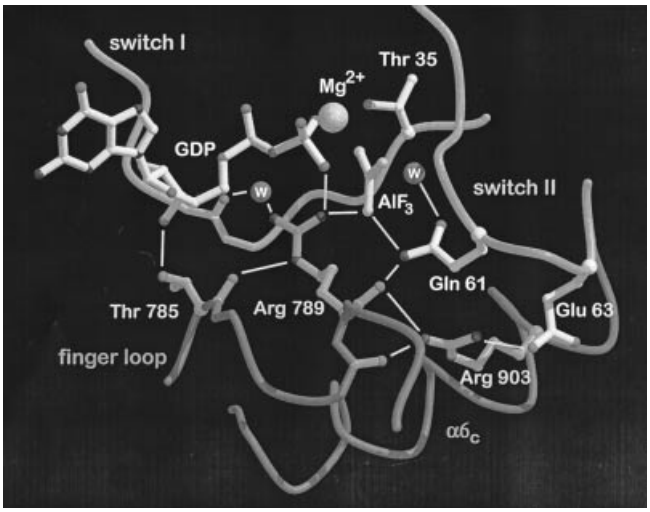



**Fig. 9.4.** Structural changes of the Ras protein on transition from the active to the inactive form. Important structural changes of the Ras protein during GTP hydrolysis are shown in schematic form. The structural changes are derived from comparison of the active Ras · GppNHp structure with the Ras · GDP structure. A defined bound water molecule nucleophilically attacks the γ-P of GTP. The γ-phosphate group of GTP is tightly bound via interactions to Thr35 (L2 loop), Lys16 and Gly60 (L4 loop) and via coordination with a $Mg^{2+}$ ion. In the GDP state, the L2 loop and the L4 loop adopt distinctly different positions. According to Wittinghofer et al., (1993) with permission.

The rate of GTP hydrolysis in the Ras-GTP complex is very low, but is increased by several orders of magnitude by the corresponding GAP protein. The molecular basis of this stimulation was explained by structural determination of the Ras-GAP complex (Scheffzek et al., 1997). The crystal structure of the complex of a fragment of p120-GAP (GAP-334 in Fig. 9.5) and $Ras \cdot GDP \cdot AlF_3$ ended a long discussion on the mechanism of GTPase activation (Fig. 9.5).




**9.5.A**




**9.5.B**

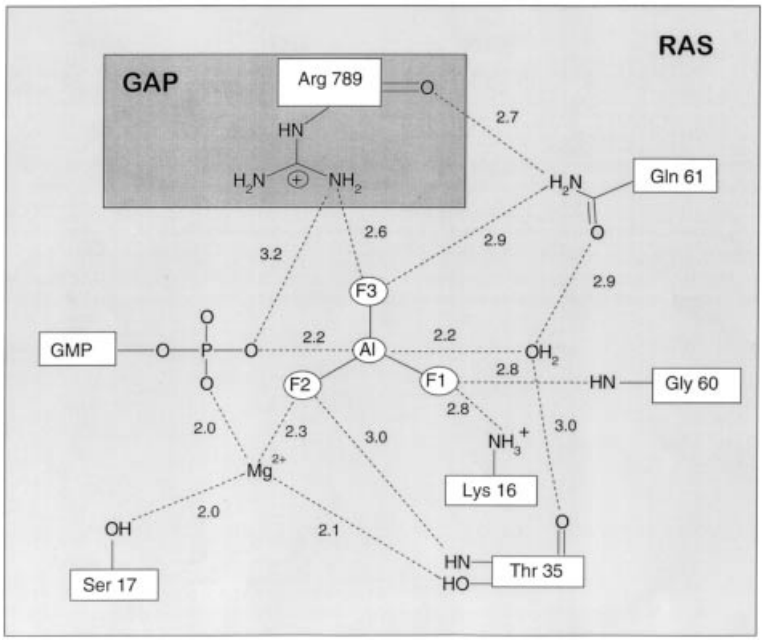



**Fig. 9.5.C** Mechanism of GAP-stimulated GTP-hydrolysis by the Ras protein.
A) Ribbon diagram of the complex of GAP-334 and the Ras protein.
B) Structural view of the active site, with the important elements of catalysis.
C) Schematic view of the active site

The structural data show that the GAP protein actively participates in catalysis by making an Arg residue available, which helps to stabilize the transition state of GTP hydrolysis. In a structural element of the GAP protein known as a *finger loop*, an invariant Arg residue (R789) is located that interacts with $AlF_3$; the latter adopts the position of the γ-phosphate in the transition state of GTP hydrolysis (see 5.5.5). Next, Arg789 has the role of neutralizing the charge of the γ-phosphate developed in the transition state. Furthermore, Arg789 helps to stabilize the L4 loop of the Ras protein, which is a part of the switch II region.

In this process, Gln61 of the L4 loop is attributed a central function in GTP hydrolysis since it is located in an ideal position for exact alignment of the water molecule and for stabilization of the transition state of GTP hydrolysis (Fig. 9.5B,C). The extent to which Gln61 also serves as a base, removing a proton from the water molecule and nucleophilically activating the latter, is still under discussion. A model with increasing acceptance assumes that the negatively charged O-atom of the γ-phosphate performs the function of a nucleophilically activating base in the GTPase mechanism (Schweins et al., 1995).

The observation that position 61 – after position 12 – is the second most frequent site of oncogenic mutations in solid tumors is in agreement with the central importance of Gln61 for GTP hydrolysis. Gln61 is a highly conserved amino acid within the superfamily of GTPases; a Gln residue in an equivalent position is also found in the α-subunits of heterotrimeric GTPases (see 5.5.6 and Fig. 5.18). Exceptions include the bacterial

elongation factor EF-Tu and the Rap GTPase, which have His and Thr/Ile, respectively, at the equivalent position to Gln61. In the case of EF-Tu, it is assumed that the His residue is responsible for activation of the water molecule (Berchthold et al., 1993).

The GTPase of the α-subunits of heterotrimeric G-proteins also uses an Arg residue (Arg178 in Fig. 5.18a,b) for stabilization of the transition state of hydrolysis. In contrast to the Ras protein, this is localized in the cis configuration on the α-subunit itself and is found in the linker between the helical domain and the G-domain.

### 9.2.3 Structure and Biochemical Properties of Transforming Mutants of Ras Protein

Comparison of the biochemical properties of mutated Ras proteins with the wild type Ras protein shows that increased lifetime of the GTP form correlates with *oncogenic activity.*

Oncogenic mutations of Ras protein are found in particular at positions 12, 13 and 61. Position 12 is especially sensitive to amino acid substitutions. Replacement of Gly12 with any amino acid other than proline leads to oncogenic activation of Ras protein. The rate of GTP hydrolysis of oncogenic mutants is about 90 % lower than the wild type. It is important that the low rate of GTP hydrolysis cannot be increased by GTPase-activating proteins, in contrast to the wild type protein (see 9.2). With respect to the lifetime of the activated GTP state, there is therefore a large difference between the wild type Ras protein and the oncogenic mutated Ras proteins. In the presence of GTPase-activating proteins, the oncogenic mutants of Ras protein spend a ca. $10^5$-fold longer period in the activated state than the wildtype Ras protein and can transmit a dominant signal in the direction of cell proliferation, favoring tumor transformation.

Comparison of the wild type structure with the structure of oncogenic Ras proteins, in which Gly12 is replaced by other amino acids, shows only small structural changes in the active center. This is not surprising since the intrinsic GTPase activity of the Ras protein is only slightly changed in the oncogenic mutants. Only when the structure of the Ras-GAP complex was obtained did the means become clear by which oncogenic G12 mutants influence Ras signal transduction. The G12 of the P-loop is located very close to the main chain of the Arg finger of the GAP protein and to the Gln61 of the Ras protein. Replacement of glycine by other amino acids would lead to Van der Waals repulsion and thus to displacement of the Arg finger and of Gln61. In the oncogenic G12 mutant of the Ras protein, an active role of the Arg finger in GTP hydrolysis is, according to this model, no longer possible.

The effect of oncogenic mutations at position 61 can also be explained using the Ras-GAP complex. Gln61 has a central function in GTP hydrolysis in that it contacts and coordinates the hydrolytic water molecule and the O-atom of γ-phosphate of GTP and thus stabilizes the transition state. Amino acids with other side chains apparently cannot fulfil this function, as shown by the oncogenic effect of Gln61 mutants in which Gln61 is replaced by other amino acids (other than Glu).

## 9.3 Membrane Localization of Ras Protein

The function of the Ras protein in cellular signal transduction is inseparably bound with the plasma membrane. The Ras proteins associate with the inner side of the cell membrane with the help of lipid anchors, such as farnesyl residues and palmitoyl residues (see 3.7).

Farnesylation of the Ras protein occurs at the C-terminal CAAX sequence (A: aliphatic amino acid, X: Ser or Thr). The farnesyl residue is attached, with the help of a farnesyl protein transferase, via a thioether bond to the Cys residue of the CAAX sequence. Next, the last three amino acids are cleaved off by proteases and the carboxyl group of the C-terminal cysteine residue undergoes a methylesterification (Fig. 9.6). In addition, the Ras proteins have a palmitinic acid anchor at different Cys residues in the vicinity of the C terminus. The membrane localization of the Ki-Ras protein is also supported by a polybasic sequence close to the C terminus (see 3.7 and Fig. 3.12).

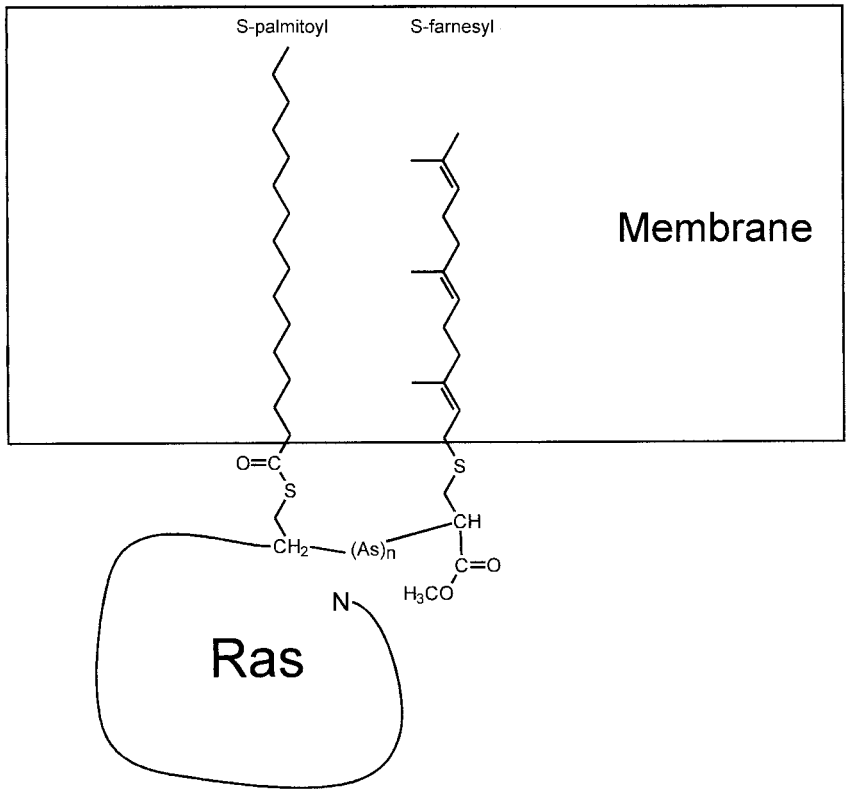



**Fig. 9.6.** Lipid anchor of the Ras protein. Membrane association of the Ras protein is mediated via a palmitoyl and a farnesyl anchor (see also 3.7).

The C-terminal modification is absolutely necessary for the function of the Ras proteins. The lipid anchors, in contrast, have no influence on the catalytic activity of Ras GTPase. Rather, the membrane anchoring of the Ras protein has the role of bringing the latter to the membrane inner side, into the neighborhood of its downstream effector protein, the Raf kinase (see 9.6). Deletion of the sequence signals necessary for lipid modification in Ras protein leads to loss of the signal molecule function. Due to the essential function of lipid modification for signal transduction via Ras proteins, great efforts are being made to develop inhibitors of farnesylation and to employ these therapeutically as inhibitors of signal transduction via Ras protein in tumors.

## 9.4 GTPase-activating Protein (GAP) in Ras Signal Transduction

The Ras proteins possess a very slow intrinsic GTPase activity, and alone, they remain in the activated GTP state for a long time. Inactivation of Ras protein is controlled *in vivo* by specific GAP proteins, which bring about a significant increase in the rate of the GTPase reaction. The GAP proteins specific for the Ras family are known as *Ras-GAP proteins*. There are also specific GAP proteins for the other members of the Ras superfamily (see Lamarche and Hall, 1994), characterized by their own conserved sequence elements. The importance of Ras-GAP proteins for signal transduction via Ras proteins is shown by oncogenic mutants with amino acid substitutions at positions 12, 13 and 61 which are resistant to the influence of GAP and show constitutive activation of Ras signal transduction.

### 9.4.1 Structure of Ras-GAP Protein

Five Ras-GAP proteins are known in mammals, of which the p120-GAP protein and neurofibromin (also abbreviated as NF1) are the best characterized (review: Boguski and McCormick, 1993). The domain structure of p120-GAP is shown in Fig. 9.7. p120-GAP has a hydrophobic amino terminus, two SH2 domains, a SH3 domain, a pleckstrin homology domain and a domain that is homologous to the calcium binding domain of phospholipase A2. The catalytic domain for GAP activity is found in a 250 amino acid section close to the C terminus; three other highly conserved sequence elements are also found in this region.

The gene for the GAP protein neurofibromin is deleted in *Recklinghausen neurofibromatosis type I* disease. The protein shows a high degree of homology with the yeast GAP proteins IRA1 and IRA2, and it may complement loss of the IRA functions in yeast. The homology with p120-GAP is much lower and is limited to the catalytic domain. The stimulating influence of neurofibromin on GTPase activity of Ras protein is comparable to that of p120-GAP. The importance of neurofibromin for growth regulation is emphasised by the observation that mutations leading to loss of catalytic activity have been found in the neurofibromin gene in neurofibromatosis patients.

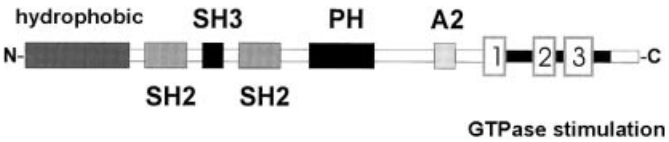



**Fig. 9.7.** Domain structure of p120-GAP. The functional domains of p120-GAP are shown in linear form. PH; pleckstrin homology domain; SH: Sarc homology domain; A2: possibly $Ca^{2+}$-dependent phospholipid binding motif; 1,2,3, conserved sequences characteristic for GTPase-stimulating activity.

The GAP proteins stimulate GTPase activity of the corresponding Ras protein by an active role in catalysis. This mechanism of GTPase stimulation by GAP protein has largely been explained, thanks to the crystal structure of the Ras · GAP transition state analog complex (see 9.2.2). The crystal structure of the transition state complex of a Ras-related protein, the RhoA protein, with the corresponding GAP protein has also shown a similar mechanism of GTPase stimulation.

### 9.4.2 Function of Ras-GAP Protein

The GAP proteins are generally attributed the function of negative regulation of Ras signal transduction. GAP proteins may control the intensity of signal transduction via Ras proteins by reducing the lifetime of the active state of the Ras protein and thus reducing the number of Ras proteins in the GTP state. In yeast, this function has been clearly demonstrated for the GAP proteins IRA1 and IRA2. Due to these properties, GAP proteins may themselves function as signal elements with a negative, damping influence on the Ras pathway.

The presence of SH2 and SH3 domains in p120-GAP indicates a role in signaling pathways starting from receptor tyrosine kinases. In fact, the SH2 domains of p120-GAP mediate specific binding to phosphotyrosine 771 of the β type of PDGF receptor (see also Fig. 8.8). In the process of this interaction, tyrosine phosphorylation of p120-GAP is also observed; the physiological importance of this is unknown. Due to this finding, it is plausible that – starting from activated receptor tyrosine kinases – regulation of GAP activity takes place with the aim of modulating the GTPase activity of Ras protein. The conditions for this control, and whether other signal proteins are involved in control of GAP protein activity, is unclear.

## 9.5 Guanine Nucleotide Exchange Factors (GEFs) in Signal Transduction via Ras Proteins

Ras protein is a central switching station in intracellular signal transduction, which receives, modulates and passes signals on. The Ras protein receives, in particular, signals promoting growth and differentiation, which start from activated receptor tyro-

sine kinases and are transmitted through the cell membrane to the Ras protein. Signal transduction between the activated receptor and the Ras protein takes place via *guanine nucleotide exchange factors, GEFs* (review: Boguski and McCormick, 1993). The GEFs, together with adaptor proteins, form the link between activated receptor tyrosine kinases and Ras protein. The role of the GEFs is to pass the signal to the Ras protein, converting the latter from its inactive GDP form to the active GTP form (see Fig. 9.1).

In addition to growth hormone signals, other signals such as $Ca^{2+}$ signals are also processed by the Ras switching station.

### 9.5.1 Importance of GEFs

The importance of GEFs for activation of Ras protein was first demonstrated in the yeast *S. cerevisiae*. It was possible to show that the *CDC25 protein* in *S. cerevisiae* functions as an exchange factor for yeast Ras protein. The CDC25 protein immediately preceeds the Ras protein in the signal sequence in *S. cerevisiae*. Recombinant CDC25 protein specifically accelerates nucleotide exchange at the Ras protein ca. 1000-fold, a property clearly indicating that CDC25 protein is an activator of the Ras protein. The function of the CDC25 protein is therefore comparable with the function of G-protein-coupled receptors that catalyze nucleotide exchange at heterotrimeric G-proteins (see 5.2).

The question of which signal protein preceeds the Ras protein in higher organisms was unclear for a long time. Genetic investigations in *Drosophila* and in the nematode *Caenorhabditis elegans* showed that at least two types of protein are involved in establishing bonding between the activated receptor tyrosine kinase and the switching station of the Ras protein. These are SH2/SH3-containing adaptor proteins (see 8.5) and nucleotide exchange factors. Corresponding proteins were then identified in mammals. For mammals, the adaptor protein in question is the *Grb2 protein* (Grb=growth factor receptor binding protein). The Grb2 protein exists in the cell in a tight complex with the nucleotide exchange factor known as Sos protein, due to its homology with the exchange factor of *Drosophila* (Sos: son of sevenless, due to the role of this protein in signal transduction of the *sevenless* gene in *Drosophila*).

In addition to Sos protein, there are also other Ras-specific GEFs in mammals (see Table 9.1). However, the specific roles of the different exchange factors are unknown.


**Fig. 9.8.** Domain structure of mSos. The mSos-1 protein of mammals possesses a pleckstrin homology domain (PH), a Pro-rich domain for interaction with Grb2 and a catalytic domain with three sequence motifs (1,2,3) characteristic for Ras GEFs.

Exchange activity with Ras protein has been demonstrated for recombinant Sos protein *in vitro*; however, this is much lower than the exchange activity of CDC25 protein in *S. cerevisiae.*

### 9.5.2 Structure and Activation of GEFs

The structure of the mSos protein (m=mammalian) is shown in Fig. 9.8. Sequence comparison with known Ras-specific GEFs has identified a common domain of ca. 200 amino acids, to which nucleotide exchange activity has been assigned. Within this domain, three highly conserved sequence elements can be differentiated, separated by more variable sections. Other structural elements include a PH domain and a Pro-rich binding domain. The Pro-rich sequence functions as an attachment site for the SH3 group of Grb2 protein.

The complex of Grb2 and mSos proteins forms a link between Ras protein and activated receptor tyrosine kinases. The Grb2 protein has two SH2 domains (see 8.5) and a SH3 domain.

The Sos-Grb2 complex can participate in Ras signal transduction by two pathways. In one pathway, the SH2 domain of Grb2 binds to the phosphotyrosine of the activated receptor, whereby the Grb2-mSos complex, which is predominantly localized in the cytoplasm, is brought to the receptor and thus to the cell membrane (Fig. 9.9).

In the other pathway, an additional adaptor protein, the Shc protein (see 8.5), is involved in the signal transduction. The Shc protein has a phosphotyrosine binding domain (PTB domain) and specifically binds via this domain to autophosphorylated receptors such as the PDGF receptor and the EGF receptor. The Shc protein is phosphorylated itself in the process. The phosphotyrosine residues may also serve as attachment points for the SH2 domain of Grb2 protein, whereby the Grb2-Sos complex is attached to the membrane.

In the membrane-localized form, Sos protein interacts with Ras protein, which is also membrane associated, and induces nucleotide exchange in the latter. It is assumed that relocation of the Grb2-Sos complex from the cytosol to the membrane is the decisive step that establishes binding between the activated receptor and Ras protein. The membrane association of both proteins is sufficient for activation of signal transduction and to „switch on“ the Ras protein, according to this assumption.

The exact mechanism of nucleotide exchange has been explained in terms of a broad outline. The CDC25 protein in yeast binds to the GDP form of the Ras protein and also forms a complex with the nucleotide-free form of Ras protein. The binary Ras-CDC25 complex may be dissociated by GDP or GTP. Since there is a large excess of GTP compared to GDP in the cell, the nucleotide-free Ras protein is preferentially converted to the GTP-bound, active state. The driving force for nucleotide exchange is thus the high GTP concentration in the cell. By stabilizing the nucleotide-free state of Ras protein, GEFs enable binding of the G-nucleotide occurring at the highest concentration in the cell, namely GTP.

The mechanistic basis of nucleotide exchange by GEFs can be inferred from high resolution structures between GEFs and members of the Ras-GTPase superfamily. The data suggest that GEFs trigger nucleotide exchange by inserting residues into the

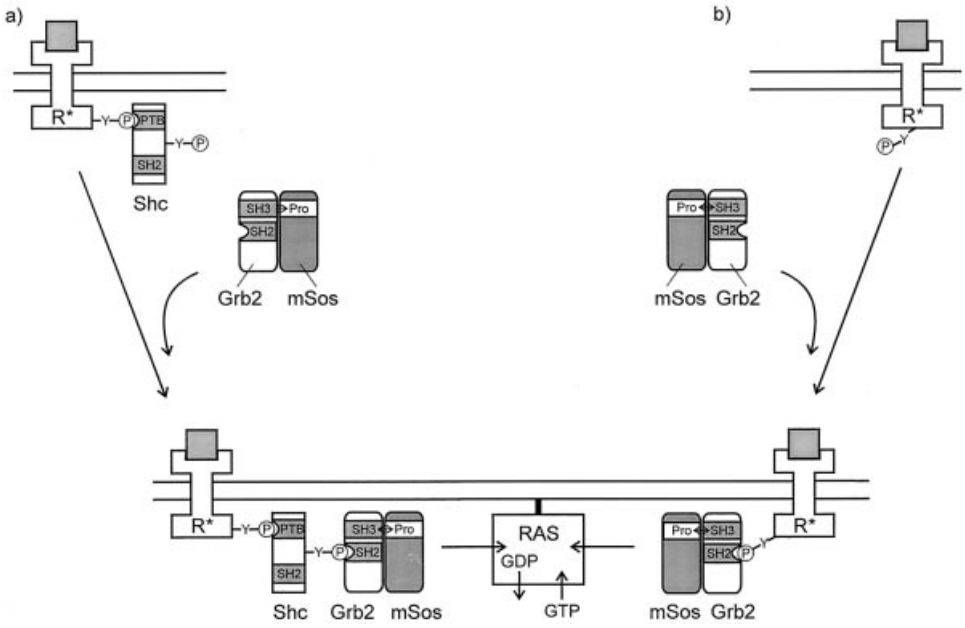



**Fig. 9.9.** Model of the function of Grb2-mSos and the adaptor protein Shc in the Ras pathway. The figure shows a highly simplified version of the two known pathways of involvement of the Grb2-mSos complex in signal transduction via the Ras protein. Phosphotyrosine residues of an activated, autophosphorylated receptor R* may serve as attachment points for the PTB domains of the Shc adaptor protein (a) or for the SH2 domain of the Grb2-mSos complex (b). In case (a), Tyr phosphorylation of Shc is performed by the activated receptor. The Grb2-mSos complex binds to the newly created phosphotyrosine residues and is drawn into the signal pathway. In case (b), the Grb2-Sos complex acts directly between the receptor and the Ras protein. In both situations, the Grb2-mSos complex is targeted to the membrane and from there, it can catalyze nucleotide exchange at the Ras protein.

nucleotide binding site of the Ras protein so as to sterically and electrostatically expel the nucleotide. The GEFs engage the switch II into an interaction and cause the displacement of switch I to open up the nucleotide binding site (Goldberg, 1998). Thus the structural elements of Ras protein recognized by GEFs are identical in part to the sections to which other regulatory proteins bind, such as GAP protein and Raf kinase.

Regulation of the GEFs is very diverse and does not only include adaptor-mediated interactions with the activated receptor. There are other Ras-GEFs that are controlled by distinctly different mechanisms, namely via second messengers. Ras-GEFs have been identified in brain that are subject to control by diacylglycerol and $Ca^{2+}$ (see 9.7). In addition, an exchange factor activated by cAMP has been described for Rap1 protein (see 9.6.2) (Roij et al., 1998).

Activated signals may also be transmitted along the Ras pathway from G-protein-coupled receptors. In this way, stimulation of Ras-GEFs may be mediated by the $\beta\gamma$-complex (Mattingly and Macara, 1996).

The GEFs for the Rho-family of small GTPases contain a characteristic sequence motif, the Dbl homology domain, which encodes the nucleotide exchange activity

towards specific Rho-family members. The Dbl domain has been termed after the oncoprotein Dbl that contains a domain of approximately 180 amino acids for which homologs have been found later in a growing family of oncogenes. The domain has been designated the Dbl (DH) homology domain and proteins containg the DH domain are now included in the Dbl protein family which is comprised of >20 members (review: Whitehead et al., 1997). Important Dbl proteins are the Vav oncoprotein and the Bcr protein (see 14.4).

## 9.6 Raf Kinase as an Effector of Signal Transduction by Ras Proteins

Which signal proteins are next in sequence after the Ras protein? The first insight into this central point of the Ras function was achieved by investigations of the structure and function of the viral oncogene *v-raf*. The *v-raf* gene codes for the Ser/Thr-specific protein kinase Raf (review: Daum et al., 1994; Avruch et al., 1994, Morrison and Cutler, 1997). The *v-raf* gene differs from its cellular counterpart, the *c-raf* gene, by deletions at the N terminus of the coded protein, affecting 200–300 amino acids. This section has an autoregulatory function for protein kinase activity of Raf protein. Loss of the autoregulatory domain is responsible for the transforming property of the viral protein.

Extensive genetic and biochemical investigations rapidly showed that Raf kinase has an effector function in the Ras pathway, immediately downstream from the Ras protein. Via Raf kinase, a linkage between Ras signal transduction and MAP kinase pathways is created (see Chapter 10); one of the functions of this linkage is to regulate transcription of genes of great importance for control and running of the cell cycle. However, Raf kinase is not the only effector molecule of Ras protein. On the contrary, there are other effector molecules that interact with the GTP form of Ras protein in a specific manner and bring about linkages to other signaling pathways (see 9.7).

### 9.6.1 Structure of Raf Kinase

Mammals have at least three different genes for Raf kinases, namely the genes A-*raf*, B-*raf* and c-*raf*1. The structure of c-Raf1 kinase is shown in Figure 9.10. The three Raf kinases have three common conserved domains. Two of the domains, CR1 and CR2, are towards the N terminus and have a regulatory function on Raf activity. Specific mutations in these regions activate the transforming potential of Raf kinase. Particular importance for the Raf function is attributed to the CR1 domain. The CR1 domain has a Cys-rich section with the ability to bind Zn. A similar structural element is found in the N-terminal region of protein kinase C (see 7.3.2). Mutations and deletions of the N-terminal sequences may bring about constitutive activation of Raf kinase. The CR1 element is thus attributed an *autoregulatory function of Raf kinase activity* and loss of this function may lead to oncogenic activation of Raf kinase. The CR2 domain contains

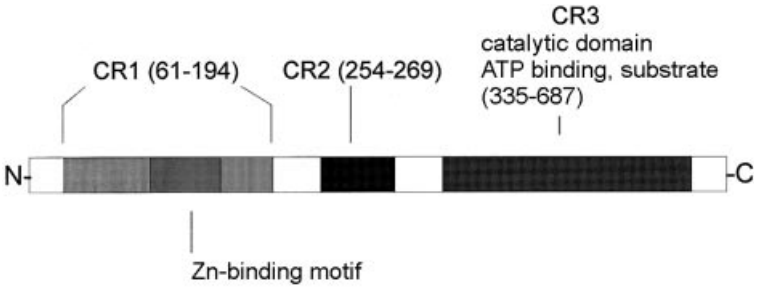



**Fig. 9.10.** Domain structure of Raf kinase. Linear representation of the functional domains of c-Raf1 kinase. CR: control region.

Ser and Thr residues that serve as regulatory phosphorylation sites. Mutations are also described for these regions leading to oncogenic activation of Raf kinase. The protein kinase activity is found in the CR3 domain.

## 9.6.2 Interaction of Raf Kinase with Ras Protein

Genetic and biochemical investigations point to a function as a *central effector of signal transduction via Ras proteins* for the Raf kinase. Raf kinase is immediately downstream from Ras protein in the signal chain. The activated GTP form of Ras protein interacts in a specific manner with Raf kinase and thus mediates membrane localization of Raf kinase. Consequently, the protein kinase activity of Raf kinase is stimulated and the signal is transmitted further via the protein kinase cascade of the MAP kinase pathway.

Signal transduction between Ras protein and Raf kinase is based on a specific interaction of the two proteins, which can only be performed by the activated, GTP-bound form of Ras protein. The inactive, GDP-form of Ras protein shows significantly weaker binding to Raf kinase (Herrmann et al., 1995). Complex formation is not linked to stimulation of GTPase activity of Ras protein, and thus it is assumed that termination of signal transduction only occurs on dissociation of Raf kinase or hydrolysis of the bound GTP. The lifetime of the GTP state is determined by the GTPase activity of the Ras protein, which itself is subject to regulation by GAP proteins.

The structural element involved in complex formation on the c-Raf1 kinase side is the Ras binding domain, found in the CR1 domain. On the Ras protein side, the L2 and L4 loops of the effector region are involved in complex formation. Both structural elements of Ras protein undergo a distinct reorientation on transition from the inactive GDP form to the activated GTP state of Ras protein, and are thus very suitable as signal elements.

Insight into the molecular interaction between Raf kinase and Ras protein was obtained from the crystal structure of the complex between the Ras binding domain of c-Raf1 kinase and a Ras-like protein, Rap1A protein (Nassar et al., 1995, review: Wittinghofer and Nassar, 1996).

Rap1A protein is closely related to Ras protein; it is up to 50 % identical at the amino acid level. High homology is observed particularly in the effector region, where amino acids 32 to 44 are completely identical. The cellular function of Rap1A protein is not completely understood. The Rap1A protein is localized on intracellular membranes. It binds the same effector proteins as Ras protein and binds Raf kinase without any associated translocation of Raf kinase to the cell membrane. Nevertheless it is assumed that Rap1A functions in a pathway distinct from the Ras signaling pathway. GEFs specific for Rap1A have been identified and Rap1A may also interact with unique effectors (review: Bos, 1998).

The structure of the complex indicates a specific interaction between β-sheet structures of Raf kinase and structural elements of Rap1A protein belonging to the switch I region, which are thus part of the Rap1A protein effector domain. Since Rap1A protein has a very similar structure to Ras protein, it is assumed that Ras protein also interacts with Raf kinase via its switch I region.

### 9.6.3 Mechanism of Activation and Regulation of Raf Kinase

Our understanding of the mechanism of activation of Raf kinase by Ras protein is very incomplete. Preliminary information was obtained in experiments using a fusion protein consisting of the membrane localization signal CAAX of Ki-Ras protein (see 9.3) linked to the C terminus of Raf kinase. The presence of the membrane localization sequence of the Ras protein in Raf kinase leads to its constitutive activation, and the activity of Raf kinase in this construct is independent of the Ras protein. In the Ki-Ras protein, the CAAX sequence represents a signal for farnesylation which, together with a basic sequence (see Fig. 3.12), is sufficient to bring about membrane localization of the Ki-Ras protein. The experiment suggests that an important function of the activated Ras protein is to tranport Raf kinase to the membrane in a regulated fashion (Fig. 9.11). It is still unclear how membrane translocation of Raf kinase is linked to its activation. It is assumed that the Ras-mediated membrane localization of Raf kinase is the first step which must be followed by other events, such as protein phosphorylation, oligomerization and interaction with other cofactors, to bring about complete activation.

The complexity of regulation of Raf kinase was shown by the discovery of a further three proteins which can specifically associate with Raf kinase (review: Morrison and Cutler, 1997, Sternberg and Alberola-Ila, 1998). Members of the family of *14-3-3 proteins* are found associated with Raf kinase. The 14–3-3 proteins recognize and bind phosphoserine residues in a particular sequence environment. It is assumed that the 14–3-3 proteins bind to Ser-phosphate residues of Raf kinase and thereby fix it in an inactive conformation. Other Raf-interacting proteins include the molecular chaperones hsp90 and p50. These proteins appear to be important for maintaining protein stability and for the proper localization of Raf kinase within the cell.

The activity of Raf kinase is also regulated by protein phosphorylation (review: Daum et al. 1994; Avruch et al., 1994). Raf kinase is phosphorylated on Ser/Thr and Tyr residues. A negative regulation of Raf kinase is observed by protein kinase A mediated phosphorylation. In addition, protein kinase C activates Raf kinase by phosphorylation. Furthermore, phosphorylation of Tyr residues of Raf kinase is seen in the

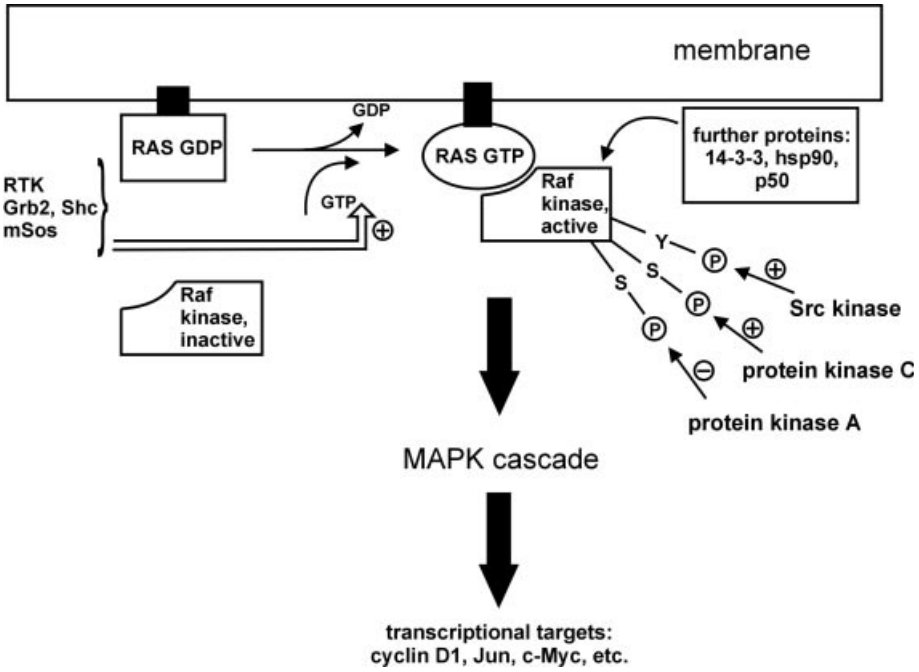



**Fig. 9.11.** Model of regulation and activation of Raf kinase. The active Ras.GTP complex binds and activates Raf kinase, which passes the signal on to the MAP kinase pathway. Various proteins including the 14–3-3 proteins and the molecular chaperons hsp 90 and p50 are thought to be involved in the regulation of the Raf kinase signaling function. In addition, Raf kinase is regulated by phosphorylation. Tyr phosphorylation (possibly via Src kinase) and Ser phosphorylation via protein kinase C have a stimulatory effect. In contrast, Ser phosphorylation via protein kinase A has an inhibitory effect. RTK: receptor tyrosine kinase.

process of activation of the Ras-Raf pathway. Src kinase is involved in the activating Tyr phosphorylation. How the different phosphorylation events are arranged in the overall pathway of Ras/Raf signal transduction is not well understood.

The path of signal transduction leads from the activated Raf kinase to the protein kinase cascade of the MAP kinase pathway (Chapter 10).

## 9.7 Reception and Transmission of Multiple Signals by Ras Protein

The Ras protein is a *multifunctional* signal protein that can be activated by various signaling pathways and transmits signals via different effector proteins. In addition to the well characterized signaling pathways described above leading to the Raf kinase via growth factor receptors, other signal chains have been found that use Ras protein (or the different Ras subtypes) as a central switching station (see Fig. 9.12). These

signaling pathways often show tissue specificity and the details are not well characterized. The original simple picture of Ras signal transmission must therefore be replaced by a more complex picture, in which the Ras protein receives many signals and transmits these to different effector proteins (review: Vojtek and Der, 1998). It is not known how the various activities in this Ras network are coordinated.

**Multiple Input Signals of Ras Protein**

The Ras protein may be activated by different signaling pathways:

- Binding of growth factors to their receptors

This well characterized pathway of Ras signal transmission was the first to be discovered (see above) and involves adaptor proteins (Grb2, Shc) and GEFs (e.g. mSos).

- $Ca^{2+}$, diacylglycerol signals

Changes in the concentration of diacylglycerol and $Ca^{2+}$ lead to activation of the Ras protein in brain. This effect is possibly mediated via specific GEFs. Ras-specific GEFs are found in the brain, which are regulated by $Ca^{2+}$. Examples are the Ras guanyl nucleotide releasing protein (RasGRP), which contains a $Ca^{2+}$-binding motif and a diacylglycerol-binding motif, and the Ras guanyl nucleotide releasing factor 1 (RasGRF1), which is activated by $Ca^{2+}$/calmodulin (Ebinu et al., 1998).

- NO signals

Stimulation of N-methyl-D-aspartate (NMDA) receptors (see Chapter 16) in the nervous system is linked to activation of NO synthase and creation of an intracellular NO signal (Yun et al., 1998). NO can directly activate Ras protein; redox modification of Ras protein is assumed to take place in this process.

- Signals from G-protein signaling pathways

There is evidence that signals starting from G-protein-coupled receptors run into the Ras switch station (Van Biesen et al., 1995). $\beta\gamma$-subunits of G-proteins are under discussion as the link between G-protein-coupled signal transduction and the Ras pathway; these subunits could influence the activity of Ras protein and the subsequent MAP kinase pathway by a presently unknown mechanism.

**Multiple Effector Molecules of Ras Proteins**

In addition to Raf kinase, a number of other signal proteins have been identified to which an effector function in Ras signal conduction has been attributed (review: 1995; Pawson, 1995; White et al., 1995, Katz and McCormick, 1997, Vojtek and Der, 1998). These effector candidates include a very diverse collection of structurally and functionally distinct proteins, which all show preferential affinity for the active Ras form. The

residues of Ras protein involved in binding these effectors lie in the region of switch I (residues 30–37) and switch II (residues 59–76); in addition, an intact Ras effector domain (residues 32–40) is required for this interaction.

- **MEK Kinases**

In addition to Raf kinase activation, Ras protein also mediates stimulation of other protein kinases, known as MEK kinases. These are signal proteins in the MAPK pathway (see Chapter 10) and transmit signals at the level of gene expression.

- **PI3-kinase**

The GTP form of Ras protein specifically binds to the catalytic 110 kDa subunit of phosphatidyl inositol-3-kinase (PI3-kinase; see 6.6) (Rodriguez-Vicinia et al., 1994). Binding to PI3-kinase takes place via the effector domain of Ras protein and is dependent on the Ras protein being in the active GTP form. There is clear experimental evidence that PI3-kinase is downstream from the Ras protein and that its activity may be controlled by Ras protein. Activation of PI3-kinase leads to formation of the membrane-localized messenger substance Ptd-Ins(3,4)$P_3$, which binds to the PH domains of signal proteins and can lead these to the membrane and activate them (see 6.6). The messenger substance Ptd-Ins(3,4)$P_3$ (and related compounds) influence many functions of the cell, e.g., reorganization of the cytoskeleton and suppression of apoptosis.

The Ras · GTP-mediated activation of PI3-kinase links the Ras pathway with functions of the Rho/Rac proteins. Members of this protein family within the Ras superfamily control formation of the cytoskeleton. The exact nature of the linkage with the Ras/PI3-kinase signal conduction to the Rac proteins is unknown. There is evidence that the product of the PI3-kinase, Ptd-Ins(3,4,5)$P_3$, binds to the PH domain of the *Vav protein* and activates the latter. The Vav protein functions as a nucleotide exchange factor for the Rac GTPase (Han et al., 1998). The observation that activation of the Ras pathway is accompanied by reorganization of the cytoskeleton is in agreement with these findings.

Ras-mediated activation of PI3-kinase also links the Ras protein to Akt kinase (see 6.6.1), which mediates antiapoptotic signals.

- **GEFs of Ral GTPase**

Proteins that function as GEFs for the Ral GTPase have also been identified as effectors of Ras proteins. The Ral protein is a Ras-related small GTPase of mostly unknown function. The GEFs are proteins known as Ral-GDS, RLF and Rgl (Feig et al., 1996, Wolthuis et al., 1997;).

The Ral-specific GEFs are activated by the interaction with Ras protein. It is assumed that activation is primarily caused by membrane association coupled to Ras binding.

- **p120 GAP**

The GTPase-activating protein p120 GAP, in addition to negatively regulating the Ras function, also specifically associates with the protein p190, which is a GAP for the Rho family of GTPases. It is assumed that activation of the Rho family of GTPase contributes significantly to the Ras-transformed phenotype.

Other less well-characterized effectors of Ras proteins are the proteins Rin1 and AF6 (review: Vojtek and Der, 1998) and the product of the *ksr gene*. The KSR protein (KSR = kinase suppressor of Ras) has a protein kinase domain which is 31 % identical to the kinase domain of Raf kinase (review: Downward, 1995). It is not clear how KSR kinase participates in Ras signal transduction.

### Ras Protein as a Central Switching Station

Identification of multiple input signals and several effector proteins underlines the high complexity of signal transduction via the Ras protein. The Ras pathway cannot be seen as a linear ordering of signal elements, by which information is conducted verti-

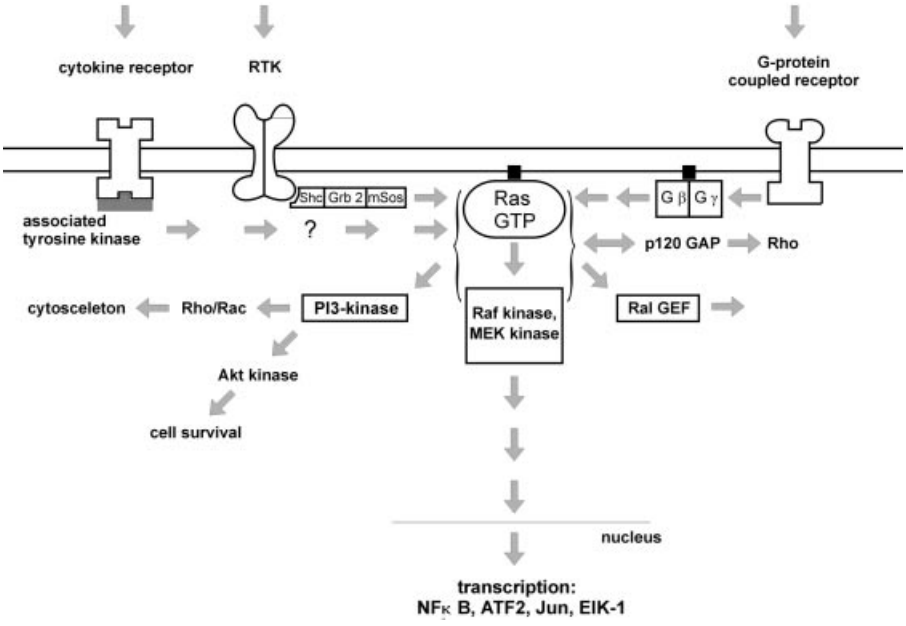



**Fig. 9.12.** Overview of the Ras signaling pathway. Signals from at least three major signaling pathways meet at the Ras protein. Activation of the Ras protein may be initiated by receptor tyrosine kinases, by G-protein-coupled receptors and by receptors with associated tyrosine kinases. The nature of the communication between the Ras protein and receptors with associated tyrosine kinase or G-protein-coupled receptors is mostly unknown. From the activated Ras protein, the signal is passed to various effector molecules including members of the MEK kinases, PI3-kinase, p120 GAP and Ral-GEFs. The best understood is the effector function of Raf kinase, which passes a signal to the transcription level via the MAP kinase pathway.

cally from the cell membrane to the cell interior. Rather, the Ras protein is at the center of a network of different signal chains (see Fig. 9.12). In this network, it functions as a central switching station at which signals are registered, integrated and passed on.

Different signals meet at the Ras switching station. Signals starting from activated receptor tyrosine kinases are registered. These signals may be conducted via GEFs and/or GAP proteins to the Ras protein. $Ca^{2+}$ signals, redox signals in the form of NO, and signals from G-protein signaling pathways are also received at the Ras switching station.

Starting from the activated Ras protein, compounds are produced for various signaling pathways. A main pathway leads via MAP kinase to the level of gene expression, creating proliferation promoting or also inhibiting signals. It is assumed that the transforming effect of oncogenic Ras mutants is mediated by this pathway in particular. A function promoting cell survival is also mediated via linkage to the PI3-Akt pathway.

Surprisingly, a growth inhibiting and pro-apoptotic function has been demonstrated for oncogenic Ras mutants. In primary cell cultures, activation of the Ras pathway is linked to an increase in the concentration of the tumor suppressor proteins p53 and p19ARF (Serrano, 1997), which both promote programmed cell death, or apoptosis (see Chapter 15). This example shows that, according to the cellular context, the Ras protein can promote both cell death and cell survival via interactions with distinct effector proteins.

There are also links to other members of the Ras superfamily, such as the Ral protein and the Rho/Rac proteins. The latter are involved in reorganization of the actin cytoskeleton. Transformation of cells with oncogenic Ras mutants is associated with reorganization of the actin cytoskeleton and it is assumed that this effect is due to coupling of the Ras pathways with the function of Rho/Rac proteins.

# References Chapter 9

Avruch, J., Zhang,X. and Kyriakis, J.M. 'Raf meets ras: completing the framework of a signal transduction pathway' (1994) Trends Bioch. Sci. *19*, 279–283

Berchthold, H., Reshetnikova, L., Reiser, C.O.A., Schirmer, N.K., Sprinzl, M. and Hilgenfeld,R. 'Crystal structure of active elongation factor Tu reveals major domain rearrangements' (1993) Nature *365*, 126–132

Boguski, M.S. and McCormick, F. 'Proteins regulating Ras and its relatives' (1993) Nature *366*, 643–654

Bos, J.L. 'All in the family? New insights and questions regarding interconnectivity of Ras, Rap1 and Ral' (1998) Embo J. *17*, 6776–6782

Han, J., Luby-Phelps, K., Das, B., Shu, X., Xia, Y., Mosteller, R.D., Krishna, U.M., Falck, J.R., White, M.A. and Broek D 'Role of substrates and products of PI 3-kinase in regulating activation of Rac-related guanosine triphosphatases by Vav' (1998) Science 279, 558–560

Daum, G., Eisenmann-Tappe, I., Fries, H.W., Troppmair, J. and Rapp, U.R. 'The ins and outs of Raf kinases' (1994) Trends Bioch. Sci. *19*, 474–480

Downward, J. 'KSR: a novel player in the ras-pathway' (1995) Cell *83*,831–834

Ebinu, J.O., Bottorf, D.A., Chan, E.Y.W., Stang, S.L., Dunn, R.J. and Stone, J.C. 'RasGRP, a ras guanyl nucleotide-releasing protein with calcium- and diacylglycerol-binding motifs' (1998) Science *280*, 1082–1086

Geyer, M., and Wittinghofer, A. 'GEFs, GAPs, GDIs and effectors: taking a closer (3D) look at the regulation of Ras-related GTP-binding proteins' (1997) Curr Opin Struct Biol *7*, 786–792

Goldberg, J. 'Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching' (1998) Cell *95*, 237–248

Herrmann, C. Martin, G.A. and Wittinghofer, A. 'Quantitative analysis of the complex between $p21^{ras}$ and the Ras-binding domain of the human Raf-1 protein kinase' (1995) J. Biol. Chem. *270*, 2901–2905

Katz, M.E. and McCormick, F. 'Signal transduction from multiple Ras effectors' (1997) Curr. Op. Gen. Dev. *7*, 75–79

Kraulis, P,J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Lamarche, N. and Hall, A. 'GAPs for rho-related GTPases' (1994) Trends Gen. *12*, 436–440

Morrison, D.M. and Cutler, R.E. 'The complexity of Raf-1 regulation' (1997) Curr. Op. Cell Biol. *9*, 174–179

Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F. and Wittinghofer, F. 'The 2.2.A crystal structure of the ras-binding domain of the Ser/Thr kinase c-raf1 in the complex with Rap1A and a GTP analogue' (1995) Nature *375*, 554–560

Pai, E.F., Kabsch, W., Krengel, U., Holmes, K.C., John,J. and Wittinghofer, A. 'Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation' (1989) Nature *341*, 209–214

Palmero, I., Pantoja, C. and Serrano, M. 'p19ARF links the tumour suppressor p53 to Ras' (1998) Nature *395*, 125–126

Pawson, T. 'Proteine modules and signalling networks' (1995) Nature *373*, 573–579

Rodriguez-Vicinia, P., Warne, P.H., Dhand, R., Vanhaesebroeck, B., Gout, I., Fry, M.J., Waterfield, M.D. and Downward,J. 'Phosphatidylinositol-3-OH Kinase as a direct target of ras' (1994) Nature *370*, 527–532

de Rooij, J., Zwartkruis, F.J., Verheijen, M.H., Cool, R.H., Nijman, S.M., Wittinghofer, A. and Bos, J.L. 'Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP' (1998) Nature *396*, 474–477

Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F. and Wittinghofer, A. 'The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants' (1997) Science *277*, 333–338

Schweins, T. and Wittinghofer, A. 'GTP binding proteins. Structures, interactions and relationships' (1994) Curr. Biol. *4*, 547–550

Schweins,T., Geyer,M., Scheffzek,K., Warshel,A.,mm Kalbitzer,H.R. and Wittinghofer,A. 'Substrate assisted catalysis as a mechanism for GTP hydrolysis of $p21^{ras}$ and other GTP-binding proteins' (1995) Struct.Biol. 2, 36–44

Sternberg, P.W. and Alberola-Ila, J. 'Conspiracy theory: Ras and Raf do not act alone' (1998) Cell *95*, 447–450

Van Biesen, T., Hawes, B.E., Luttrell, D.K., Krueger, K.M., Touhara, K., Porfiri, E., Sakaue, M., Luttrell, L.M. and Lefkowitz, R.L. 'Receptor-tyrosine-kinase- and Gßγ-mediated MAP kinase activation by a common signalling pathway' (1995) Nature *376*, 781–784

Vojtek, A.B. and Der, C.J. ' Increasing complexity of the Ras signaling pathway' (1998) J. Biol. Chem. *273*, 19925–19928

White, M.A., Nicolette, C., Minden, A., Polverino, A., Van Aelst, L., Karin, M. and Wigler, M.H. 'Multiple ras functions can contribute to mammalian cell transformaton' (1995) Cell *80*, 533–541

Whitehead, I.P., Campbell, S., Rossman, K.L and Der, C.J. 'Dbl family proteins' (1997) Biochim Biophys Acta *1332*, F1–23

Wiesmüller, L. and Wittinghofer, F. 'Signal transduction pathways involving ras' (1994) Cell. Signalling *6*, 247–267

Wittinghofer, A., Pai, E.F. and Goody, R.S. 'Structural and mechanistic aspects of the GTPase reaction of H-ras p21' (1993) Handbook of Exp. Pharmacology *108/I*, 196–211

Wolthuis, R.M., de Ruiter, N.D., Cool, R.H. and Bos, J.L. 'Stimulation of gene induction and cell growth by the Ras effector Rlf' (1997) EMBO J. *16*, 6748–6761

Yun. H.Y., Gonzalez-Zulueta, M., Dawson, V.L. and Dawson, T.M. 'Nitric oxide mediates N-methyl-D-aspartate receptor-induced activation of p21ras' (1998) Proc Natl Acad Sci U S A *95*, 5773–8

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

## Chapter 10

# Intracellular Signal Transduction: the Protein Cascades of the MAP Kinase Pathways

Intracellular signal conduction takes place predominantly by two pathways starting from activated transmembrane receptors. In one pathway, activation of transmembrane receptors initiates formation of diffusible messenger substances that bind effector proteins and activate these for further signal transduction. In this signaling pathway, signals may be carried as far as the cell nucleus and temporally and spatially variable reactions may be triggered.

A second pathway, which is particularly important for regulation of growth and differentiation reactions, takes place via a cascade of sequential protein kinases. In this pathway, an extracellular signal is registered by a transmembrane receptor, the receptor is activated and transmits the signal to downstream effector molecules. The Ras protein or other members of the Ras protein superfamily are often involved in this pathway. Up to this point, all the reactions involved are predominantly membrane associated. From the Ras protein (or other regulatory proteins), the signal is conducted, with the help of protein kinases, in the form of a sequential *cascade*, into the cell interior, possibly into the cell nucleus (Fig. 10.1). Since these signaling pathways are activated by mitogenic (mitogenic = promoting cell division activities) signals such as growth hormone signals, they are known as mitogenic activated protein kinase pathways (MAPK pathways). As a result of stimulation of the MAPK pathways, phosphorylation and activation of gene regulating proteins or enzymes catalyzing key reactions of metabolism are observed (review: Cano and Mahadevan, 1995; Marshall, 1995). Signal transduction via *sequential protein kinase reactions* is a very flexible and efficient principle for amplification, diversification and regulation of signals. Protein kinases, as explained in Chapter 7, are open to a range of regulatory influences. At every level of a protein kinase cascade, positive or negative regulation is possible and the intensity of a signal can be modulated within broad boundaries.

### Organization of MAPK Pathways in MAPK Modules

The mitogenic activated protein kinase (MAPK) downstream from the Ras protein is organized in modules containing three types of protein kinases, which are successively activated by sequential phosphorylation events. The cell contains different MAPK modules, which differ in the nature of the triggering stimuli and the nature and specificity of the protein kinase components. The signal transducing function of a MAPK pathway is thus determined by the nature of the MAPK module involved; this, in turn, depends on the properties of the protein kinases it contains, which differ in regulation and substrate specificity. The exact composition of the MAPK module is not fixed; rather, different subtypes of protein kinase may be recruited to a module in a variable

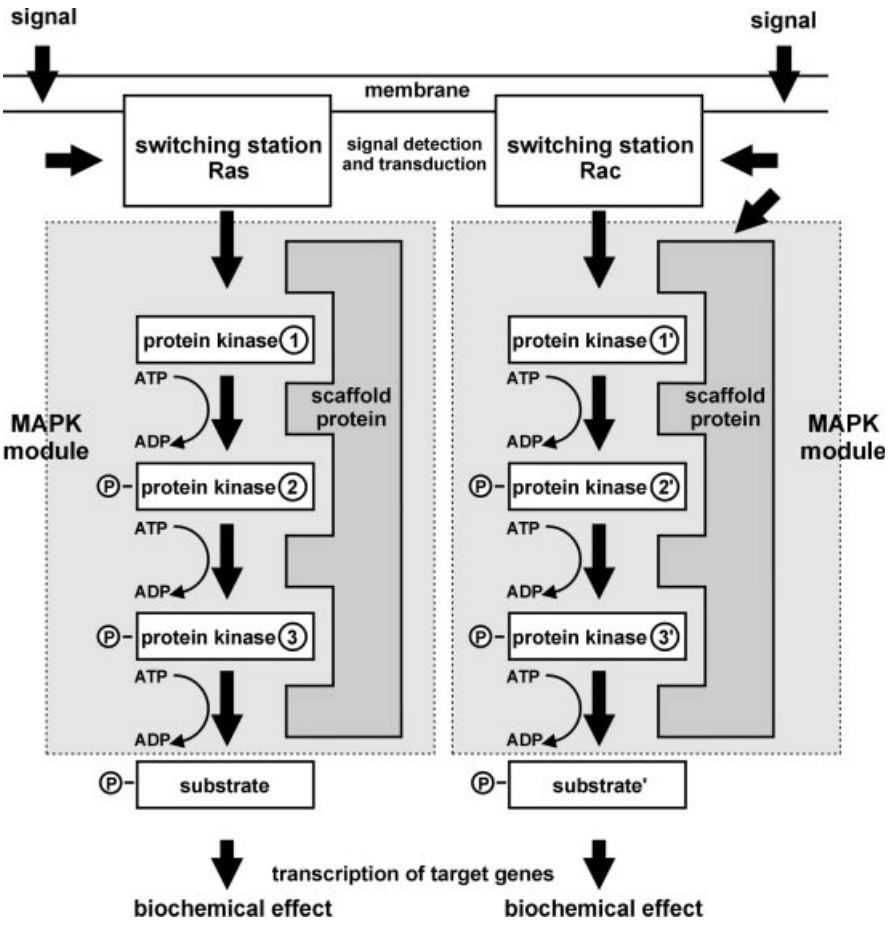



**Fig. 10.1.** Principle of signal transduction through intracellular protein kinase cascades. The intracellular protein kinase cascades are organized in modules composed in most cases of three proteinkinases and a scaffold protein. The modules process signals that are registered, integrated and passed on at the inner side of the cell membrane by central switching stations such as the Ras protein or the Rac protein. In the case of the MAP kinase pathway, the cascade includes at least three different protein kinases. Specific regulatory processes may take effect at every level of the cascade; in addition, signals may be passed from the different protein kinases to other signaling pathways.

manner. Furthermore, the modules can share a common protein kinase but regulate different biological processes. The signal is passed on by the last member in the phosphorylation cascade in the form of a phosphorylation of substrate proteins. In many cases, this process is linked to translocation of the protein kinase into the nucleus, where nuclear localized substrates, particularly transcription factors, are phosphorylated.

The various modules are not strictly independent; rather they mutually influence one another. In mammals in particular, this results in high complexity of function and regulation of MAPK pathways.

## 10.1 Components of the MAPK Pathway

Cytoplasmic protein kinases at the lower end of the MAPK pathways are key elements in intracellular signal transduction starting from Ras protein (or other members of the Ras superfamily). In mammals, they are known, due to their regulation by extracellular, frequently mitogenic, ligands, as *mitogen activated protein kinases* (= *MAP kinase, MAPK*) or as *extracellular regulated kinases* (= *ERK*). Both terms are used for the same set of protein kinases. One of the first MAPK pathways to be characterized leads from mitogens, via the Ras protein, to activation of a protein kinase known as ERK. This pathway is known as the *ERK pathway*. It should be noted that there are different MAP kinases which are active in other MAPK pathways or MAPK modules; these are different to the ERK pathway (see below). In the following, the ERK pathway starting from the activated Ras protein is used to represent other MAPK pathways with similar structure.

The MAPK/ERK proteins are at the lower end of signal transduction within a MAPK module and are generally preceded by two other protein kinases (Fig. 10.2). The MAPK/ERK proteins receive the signal in the form of an activating phosphorylation by a preceding protein kinase known as *MAP/ERK kinase* (*MEK*) or also *MAP kinase kinase* (*MAPKK*).

The MEK proteins are themselves substrates for another type of protein kinase further upstream, the *MEK kinases* (*MEKKs*, also known as *MAPKK kinases, MAPKKK*). The MEK kinases include the various Raf kinases activated by Ras protein, Mos kinase and the protein kinases MEKK1, MEKK2 and MEKK3.

The position of the various protein kinases within a MAPK module is determined by their substrate specificity.

### MEK Kinases

The MEK kinases (MAPKK kinases) are Ser/Thr-specific protein kinases and are the entry point for signal transduction in a MAPK module. The best characterized representative, Raf-1 kinase, is activated by Ras protein in its GTP-bound form. Raf kinase phosphorylates downstream MEK proteins at two Ser residues, which are separated by three other amino acids. All known MEK proteins have a similar phosphorylation site in the conserved sequence LID/NSXANS/T (X: any amino acid). Other representatives of the MEK kinase group are Mos kinase and the protein kinases MEKK1—3.

Activation of MEK kinases occurs particularly via proteins of the Ras superfamily (p21-Ras, Rho/Rac proteins). However, other pathways for activation of MEK kinases, e.g., via other protein kinases such as protein kinase C or PAK (p21 activated kinase) (review: Marshall, 1995; Robinson and Cobb, 1997) have also been reported.

### MAPK Kinases, MEKs

The MEK proteins are a special class of protein kinases since they have two-fold specificity with respect to the nature of the acceptor amino acid at the phosphorylation site of the protein substrate (review: Dhanasekaran and Reddy, 1998). The MEKs activate

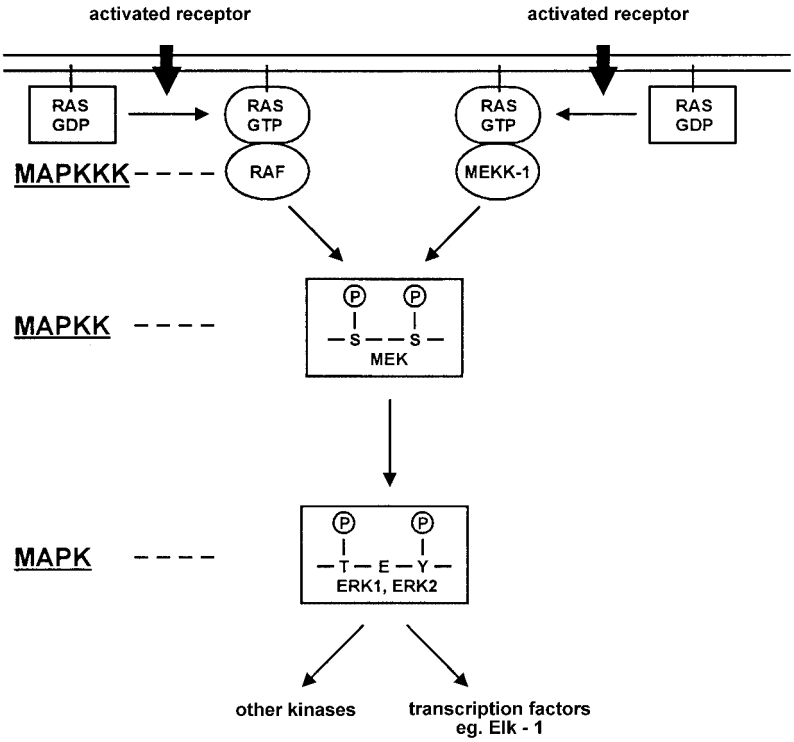



**Fig. 10.2.** Components and activation of the ERK pathway. Ordering and specificity of protein kinases in the ERK pathway. Extracellular signals are registered via receptor tyrosine kinases and passed on to the Ras protein. Ras · GTP activates protein kinases belonging to the group of MAPKK kinases (Raf kinases and MEEKs). The MAPKK kinases phosphorylate the downstream group of protein kinases, the MAPKKs at two Ser residues. The MAPKKs phosphorylate the MAPKs (ERK1 and ERK2) at a Tyr and a Thr residue, and thus are classified as dual specificity kinases. MAPK: mitogenic activated protein kinase; ERK: extracellularly regulated kinase; MEK: MAP/ERK kinase; MAPKK: MAPK kinase; MAPKKK: MAPKK kinase; MEKK: MEK kinase.

the MAP/ERK kinases next in sequence by phosphorylation at a Tyr and a Thr residue in the sequence –TXY-. The MEKs are *dual specificity protein kinases* and, due to this property, differ significantly from the other protein kinases described so far, which are either Tyr or Ser/Thr-specific.

### MAP Kinases, ERKs

The MAP kinases are divided, based on their sequence homology, the nature of the preceding MEK and the input signal, into different subgroups characteristic for the particular MAPK module. Within these subgroups, further diversification is possible by alternative splicing. Most of the known MAP kinases contain a TXY sequence; phosphorylation of this by MEKs is essential for conduction of the signal further to

downstream substrate proteins. The ERK, JNK/SAP and p38 proteins are well-characterized MAP kinases and the preceding protein kinases of the MAPK module have mostly been identified for these proteins.

### Organization in Multiprotein Complexes

There is increasing experimental evidence that coupling of the activities of the MAPK cascade is achieved with the help of organization in multiprotein complexes. In the yeast *S. cerevisiae*, it has been shown that certain proteins, such as the Ste5 protein and Pbs2p protein serve as a kind of scaffold for organization of the protein kinases in a multiprotein complex. MAPK scaffold proteins have also been described in mammals (review: Whitmarsh and Davis, 1998). Co-localization of MAP kinases with the help of scaffold proteins has two important advantages: (1) it favors the rapid transmission of the signal through the cascade, and (2) it ensures specificity of signaling by preventing unwanted crosstalk with other MAPK pathways.

## 10.2 Input Signals and Substrates of the MAPK Pathways

The existence of various MAPK modules is accompanied by a diversity of extracellular stimuli that may activate MAPK pathways. Activation of MAPK pathways is observed on treating cells with growth factors, tumor necrosis factor (TNF), interleukin 1 and upon stress exposure (heat, UV). In addition, the effect of bacterial endotoxins such as lipopolysaccharides may lead to activation of MAPK pathways. The diversity of the extracellular signals shows that very different signaling pathways are involved in activation of MAPK pathways (Fig. 10.3). In higher organisms, signal proteins preceding the MAPK modules have been identified as the Ras protein and members of the Rho family of small regulatory GTPases (Cdc42, Rac, Rho). Signals may also be passed to MAPK modules from G-proteins. In this case, the signal is passed via the $\beta\gamma$-complex to another protein kinase, the PAK kinase, and then to a MAPK module.

The *substrates* of the MAP kinase pathway are very diverse and include both cytosolic and nuclear localized proteins. Phospholipase A2 and transcription factors of the *Ets family* are well characterized substrates of the ERK pathway. Phosphorylation of a Ser residue of phospholipase A2 by ERK proteins leads to activation of the lipase activity. Consequently, there is an increase in release of arachidonic acid and of lysophospholipids, which can act immediately as diffusible signal molecules or may represent first stages in the formation of second messenger molecules.

A central function of the MAP kinase pathway is the activation of gene expression, mediated via phosphorylation of transcription factors. To achieve this, MAP kinases must be active in the nucleus. Important effectors of the MAPK pathway are transcription factors of the Ets family (review: Wasylyk et al., 1998). One member of the Ets family is the *transcription factor Elk-1*, which is positively regulated via the ERK pathway. Elk-1 binds, together with another protein, the *serum response factor*, to the *serum response element*, which is found as a regulatory sequence preceding various genes in higher eucaryotes. The genes regulated by Elk-1 include the gene for the

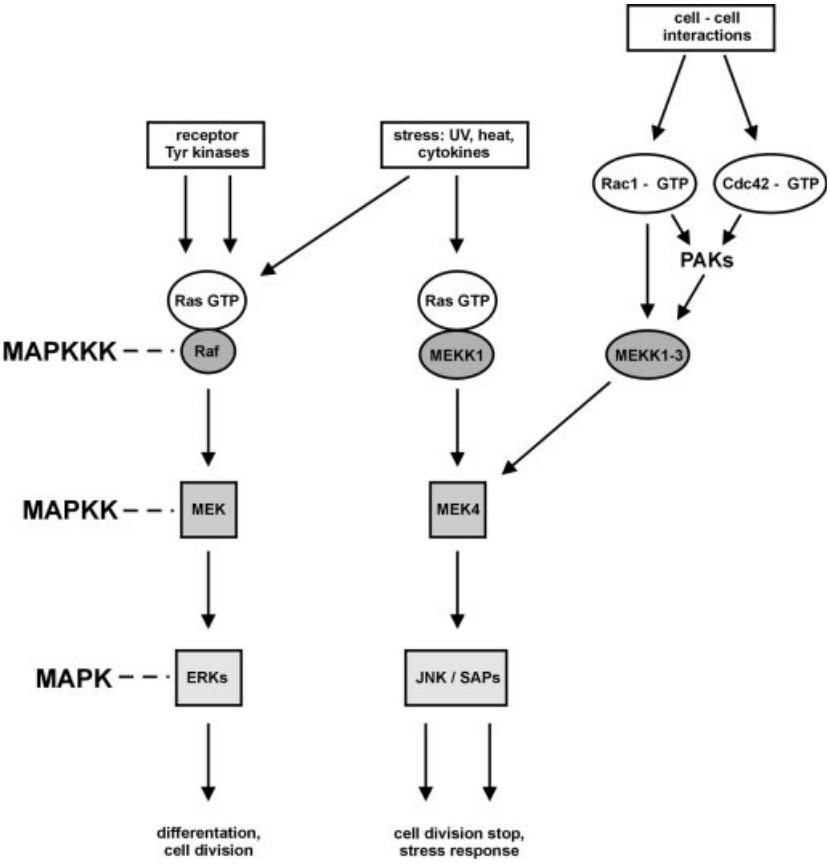



**Fig. 10.3.** MAPK pathways in mammals. In addition to the Ras/Raf/MEK/ERK pathway, which are activated by receptor tyrosine kinases, there are also MAPK pathways activated by other signals, e.g., UV irradiation, heat stress and cell-cell interactions. How these other signals are registered and passed to the small regulatory GTPases is largely unknown. The protein kinases of the JNK/SAP family may trigger a halt in growth or another stress response. Rac-1 and Cdc42 belong to the Rho family of small regulatory GTPases. The figure shows only selected MAPK pathways, for other MAPK pathways see Robinson and Cobb, (1997). JNK: c-Jun $NH_2$ terminal kinases; SAPK: stress activated protein kinases, PAK: p21 activated kinase.

transcription factor c-fos, which forms transcription factor AP1 together with the c-Jun protein (see Chapter 1). Elk-1 is phosphorylated *in vitro* by MAP kinases specifically at the sites essential for transcription activation. Several signals meet at the level of Elk-1 since activation of Elk-1 is mediated by different MAPK proteins, which in turn are activated by different MAP kinase pathways (see Fig. 10.3).

The MAP kinases must enter the nucleus to be able to phosphorylate transcription factors. For ERK2, it has been shown that phosphorylation by the preceding MEK is associated with dimerization and translocation of ERK2 into the nucleus (Khokhlatchev et al., 1998).

## 10.3 The JNK Signaling Cascade

An important subgroup of MAP kinases has the *transcription factor c-Jun* as substrate. These kinases are known as *c-Jun $NH_2$ terminal kinases* (*JNK*) or, due to their activation by stress signals, as *stress activated protein kinases* (*SAPK*). The JNK/SAPK proteins are part of their own protein kinase module that conducts stress signals further at the transcription level, and this signaling pathway is therefore known as the JNK/SAPK pathway.

The JNK/SAPK proteins bind to the N-terminal transactivation domain of c-Jun protein and phosphorylate the residues Ser63 and Ser73. Consequently, increased transcription activity is observed of genes controlled by c-Jun.

External signals leading to activation of c-Jun include the effects of cytokines and genotoxic stress such as UV, ionizing radiation and alkylating agents. Furthermore, the JNK/SAPK pathway is also involved in performing the apoptotic program (see Chapter 15). Activation of the JNK/SAPK proteins involves the small regulatory GTPases Rac1 and Cdc42, which both belong to the Ras superfamily. Signals may be conducted to the JNK pathway independent of the Ras pathway via Rac1 and Cdc42 (Olson et al., 1995). In addition, JNK activity may be regulated via the Ras pathway, so that different signals meet at the level of the JNK proteins.

Overall, the picture of the MAPK pathways is complex and incomplete at many points. Parallel pathways exist which may be activated by very different extracellular signals and use different cascade modules. The superfamily of Ras proteins represent a central switching station in MAPK pathways, which integrates extracellular signals and passes them to MAPK modules. Effector molecules of the small regulatory GTPases are protein kinases that are a part of a MAPK module in which the signal is sequentially conducted further to specific substrates via two other groups of protein kinases. It must be assumed that the different MAPK pathways are crosslinked and that branching may take place at all levels. In addition, the same substrates may be phosphorylated and activated by various MAPK pathways.

## References Chapter 10

Cano,E. & Mahadevan,L.C. 'Parallel signal processing among mammalian MAPKs' (1995) Trends Bioch..Sci. *20*, 117–122

Davis,R.J. 'MAPKs: new JNK expands the group' (1994) Trends Bioch.Sci. *19*, 470–473

Dhanasekaran, N. and Reddy, E.P. ' Signaling by dual specificity kinases' (1998) Oncogene *17*, 1447–1456

Marshall,C.J. 'Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation' (1995) Cell *80*, 179–185

Olson, M.F., Ashworth,A. & Hall,A. 'An essential role for Rho,Rac and Cdc42 GTPases in cell cycle progression through G1' (1995) Science 269, 1270–1272

Khokhlatchev, A.V., Canagarajah, B., Wilsbacher, J., Robinson, M., Atkinson, M., Goldsmith, E. and Cobb, M.H. 'Phosphorylation of the MAP kinase ERK2 promotes its homodimerization and nuclear translocation' (1998) Cell *93*, 605–615

Robinson, M.J. and Cobb, M.H. ' Mitogen-activated protein kinase pathways' (1997) Curr. Op. Cell. Biol. *9*, 180–186

Whitmarsh, A.J. and Davis, R.J. 'Structural organization of MAP-kinase signaling modules by scaffold proteins in yeast and mammals' (1998) Trends Biochem Sci *23*, 481–485

Wasylyk, B., Hagman, J. and Gutierrez-Hartmann, A. 'Ets transcription factors: nuclear effectors of the Ras-MAP-kinase signaling pathway' (1998) Trends Biochem Sci *23*, 213–216

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 11**

# Membrane Receptors with Associated Tyrosine Kinase Activity

Coupling of extracellular signals to tyrosine phosphorylation in the intracellular region may occur by two mechanisms and involves two different receptor types:

Ligand binding on the extracellular side is linked to stimulation of tyrosine kinase activity in the cytoplasmic receptor domain for receptors with intrinsic tyrosine kinase activity, the receptor tyrosine kinases (see 8.1). The ligand binding site and the tyrosine kinase are part of one and the same protein.

The receptor tyrosine kinases contrast with a group of transmembrane receptors that have no tyrosine kinase activity in the cytoplasmic domain. On ligand binding, this receptor type activates an *associated tyrosine kinase* so that a signal is created in the form of an intracellular tyrosine phosphorylation. The tyrosine kinase and the receptor are not localized on the same protein in this case. The tyrosine kinase may be permanently associated with the receptor and be activated as a consequence of ligand binding; alternatively it may be located in the cytosol and only bind to the receptor and become activated following ligand binding. Stimulation of the associated tyrosine kinase is then the starting point for transduction of the signal into the interior of the cell. In many cases, mechanisms described in previous chapters are used for the further signal transmission.

## 11.1 Cytokines and Cytokine Receptors

Cytokines are proteins that serve as signal molecules in cell-cell communication, and as such, perform a central and very diverse function in growth and differentiation of an organism. Representatives of cytokines control proliferation, differentiation and function of cells of the immune system and of cells of the blood-forming system. Furthermore, they are involved in processes of inflammation and in the neuronal, hemapoetic and embryonal development of the organism. Known cytokines include the interleukins (IL), erythropoietin, growth hormone, interferons (INF) and tumor necrosis factor (TNF) (see Table 8.1). A review of cytokines and cytokine receptors is to be found in Hill and Treisman, (1995); Taniguchi et al., (1995) and Moutoussamy et al., (1998).

The cytokines are of considerable medical importance due to their essential function in controlling the immune system, in defense reactions and for processes of inflammation. Great efforts are therefore being made to elucidate the structure and mechanism of activity of the cytokines and their corresponding receptors and to characterize the components of the signals triggered by cytokines. Many of the cytokines have the cha-

racter of autocrine hormones, i.e., they only act locally and their targets are cells of the same or similar type as the cytokine-producing cell.

A characteristic that significantly differentiates some of the cytokines from other hormones is the coupling of their activity to cell-cell interactions. The function of some cytokines such as the interleukins IL-4, IL-5, IL-6 and IL-10 is closely associated with the interaction between B and T lymphocytes.

B lymphocytes may be activated by a cell-cell interaction with T lymphocytes, i.e., they are stimulated to proliferate and produce antibodies. Stimulation of B lymphocytes takes place in a complex with T lymphocytes and this complex formation is mediated by a number of protein-protein interactions of membrane proteins from both cell types. The proteins involved are receptor systems with corresponding ligands on the partner cell. The ligands are either secreted proteins or membrane proteins that specifically bind to receptors on the surface of the partner cell, which in this case is a B or T lymphocyte.

Ligand binding activates signal chains that bring about a change in the rate of cell division or in gene expression of the two interacting cells. The receptor ligands are secreted in soluble form by the emitting cell in many cases, and for some ligands, such as IL-4, secretion is assumed to be directional and spatially limited (review: Paul and Ceder, 1994). Since the ligands are secreted in the narrow spatial region between the two interacting cells, their effect on the contact surface of both cells is limited. This is referred to as an *immunological synapse* (Fig. 11.1). In the limited contact region of the two cells, the ligands only have a short diffusion pathway, and high local concentrations are achieved so that specific binding to the receptors on the other side of the synapse is enabled. The receptors involved are cytokine receptors and B and T cell receptors.

### 11.1.1 Structure and Function of Cytokine Receptors

At least four types of cytokine receptors can be differentiated on the basis of sequence homology (Fig. 11.2). Many members of the cytokine receptors of type 1 regulate growth and transmit mitogenic signals to the cell nucleus. The cytokine receptors of type 2 include the receptors for the interferons $\alpha$ and $\beta$. Type 3 includes the receptors for tumor necrosis factor TNF and for CD40 and Fas protein, which are found on T lymphocytes.

In the extracellular region, cytokine receptors have characteristic sequence sections that specify the particular receptor type. Cys-rich domains, fibronectin type III-like domains and immunoglobulin-like domains can be differentiated. Cytokine receptors of type I often have a conserved *WSXWS motif* in the extracellular region.

The subunit structure of the cytokine receptors is very variable. Amongst the cytokine receptors, there are receptors composed of one polypeptide and receptors made up of two or three different polypeptide chains (Fig. 11.3), which therefore have a hetero-oligomeric structure. The receptors have one transmembrane element per polypeptide chain. The $NH_2$ terminus is located in the extracellular region whilst the COOH terminus is intracellular.

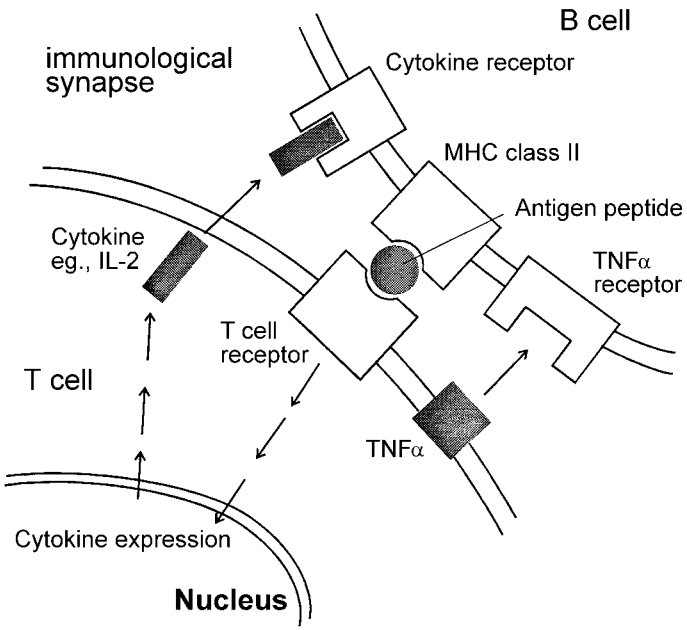



**Fig. 11.1.** Principle of an immunological synapse. Possibilities for communication between B and T cells during an immune response. Antigenic peptides are presented by the MHC complex class II at the surface of the B cell. The antigens are recognized and bound by T cell receptors of the T cell. The T cell receptor is activated and sets a signal chain in motion that leads to activation of the expression of cytokines, such as IL-2. The cytokine is secreted, and binds and activates a cytokine receptor on the B cell. TNFα is shown as another example of a ligand-receptor system. TNFα communicates, as a membrane-bound ligand, with a corresponding receptor on the surface of the B cell. The interactions shown take place in a narrow spatial region between B and T cells, which is why this system is referred to as an immunological synapse. TNF: tumor necrosis factor; MHC: major histocompatibility complex; IL-2: interleukin 2.

A feature of the hetero-oligomeric receptors is that different receptors can use the same receptor subunit (see Fig. 11.3). The γ subunit of the IL-2 receptor also occurs in the receptors for IL-4, IL-7 and IL-9. Other common receptor subunits are the gp130 chain and the β chain. In some receptors, such as the IL-6 receptor, ligand binding is mediated by a polypeptide chain, which only serves for ligand binding and is not able to conduct the signal into the cell interior by itself. The help of common subunits (gp130 for the IL-6 receptor) is required here (Fig. 11.4). *Two separate polypeptide chains* are needed for ligand binding and communication with the cytosolic side. One serves for specific ligand binding, and the other serves to pass the signal into the cell interior. Generally, all the subunits of a hetero-oligomeric receptor are required for effective conduction of the signal through the cell membrane.

No tyrosine kinase activity or other enzyme activity has been found in the intracellular sequence sections of the cytokine receptors. Since activation of the cytokine receptors is linked to an increase in intracellular Tyr phosphorylation, it is assumed that ligand binding is linked to association with a tyrosine kinase, which then conducts the

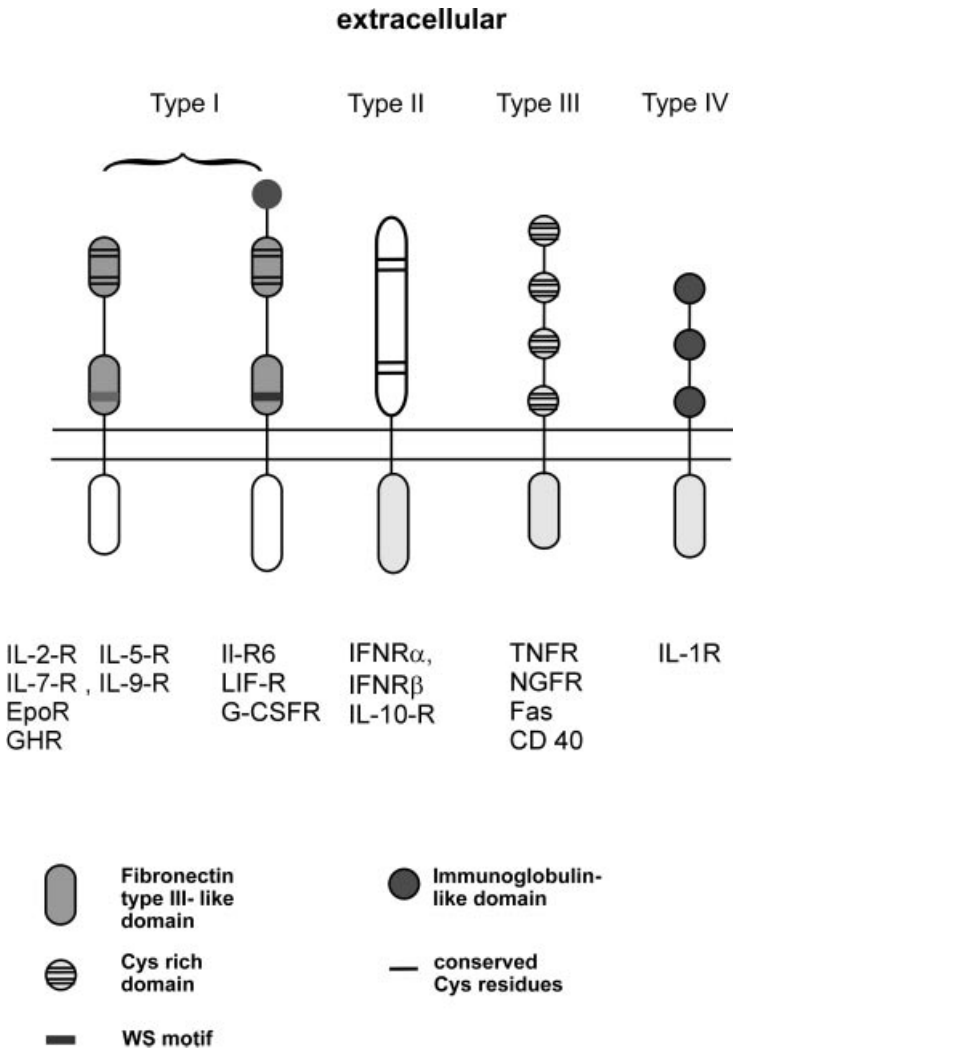



**Fig. 11.2.** Domain structure of cytokine receptors. Schematic representation of the domain structure of selected cytokine receptors. WS motif: conserved WSXWS sequence (W: tryptophan; S: serine; X: non-conserved amino acid); IL: interleukin; EpoR: receptor for erythropoietin; GHR: growth hormone receptor; LIF-R: leukemia inhibitory factor receptor; G-CSFR: granulocyte colony stimulating factor receptor; IFNR: interferon receptor; TNFR: tumor necrosis factor receptor; NGFR: nerve growth factor receptor; Fas, CD40: transmembrane receptors of lymphocytes.

signal further without the activated receptor actually performing any enzyme function (see 11.1.2). The cytoplasmic region of the cytokine receptors is required for association and activation of the protein tyrosine kinase. Furthermore, tyrosine phosphorylation is often observed in the cytoplasmic domain of the receptors, which is explained by the associated tyrosine kinase.

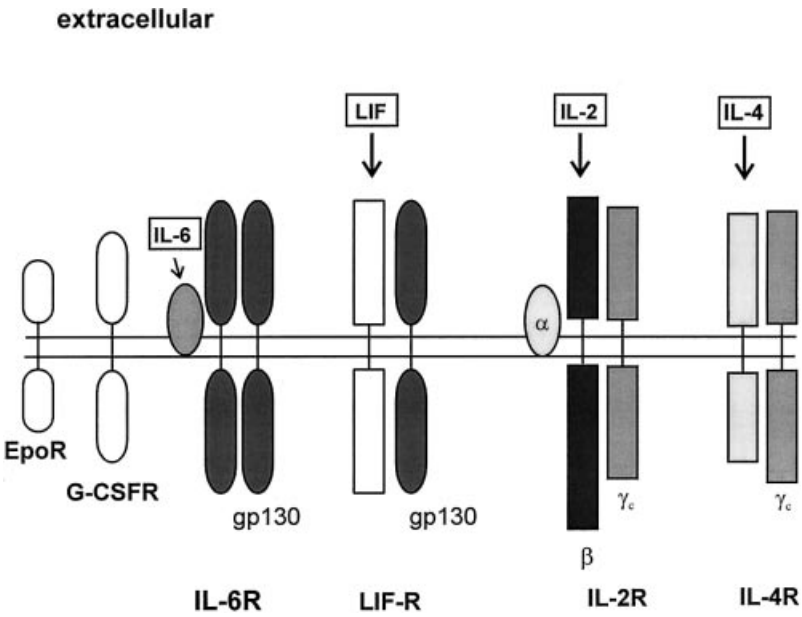



**Fig. 11.3.** Subunit structures of cytokine receptors. EpoR and G-CSFR have a homo-oligomeric structure. The other receptors shown are composed of different subunits, some of which occur in several receptors. The gp130 subunit is common to IL-6R and LIF-R; the $\gamma_c$ subunit is found in IL-2R and IL-4R. The subunits, at which ligand binding occurs, are indicated by arrows. LIF-R: leukemia inhibitory factor receptor.

A *ligand-induced formation of homodimers or higher homo-oligomers* and a *hetero-oligomerization* are assumed to be the mechanism for activation of the receptors. The details of activation and the associated change in oligomer status are unknown. A crystal structure is available for the ligand-binding domains of the receptor for growth hormone (hGH receptor, human growth hormone receptor) with bound growth hormone (see Fig. 8.4c; De Vos et al., 1992). The growth hormone binds as a monomer to the hGH receptor and complexes two molecules of the ligand-binding domain of the receptor. This observation supports the assumption that the receptor is activated by ligand-induced dimerization.

## 11.1.2 Activation of Cytoplasmic Tyrosine Kinases

As a consequence of ligand binding to cytokine receptors, activation of a tyrosine kinase activity, which is not part of the receptor protein, is observed. The *coupling* between cytokine receptor and tyrosine kinase can occur by two means:

Ligand binding *induces* the association of the cytoplasmic tyrosine kinase with the receptor. The extracellular signal leads the tyrosine kinase to make contact with the activated receptor in this case.

The tyrosine kinase may also be *permanently* associated with the receptor in a non-covalent manner and is activated on ligand binding to the receptor.

In both cases, it is assumed that ligand binding leads to restructuring of the intracellular region of the oligomeric receptor, so that binding surfaces are created for productive association of the tyrosine kinase.

Examples of receptor-associated tyrosine kinases are given in Table 8.1. Most of the associated protein tyrosine kinases belong to the family of Src kinases (see 8.3) and the Jak kinases (see 11.1.3).

The first associated tyrosine kinase to be identified was the Lck kinase ($p56^{lck}$) which belongs to the family of Src protein tyrosine kinases. The Lck kinase is activated by the interleukin 2 receptor (review: Taniguchi, 1995). The IL-2 receptor has a heterotrimeric structure (Fig. 11.4). Binding of IL-2 takes place via the β and γ chains whilst the α chain has an amplifying effect on ligand binding. The Lck kinase is tightly associated

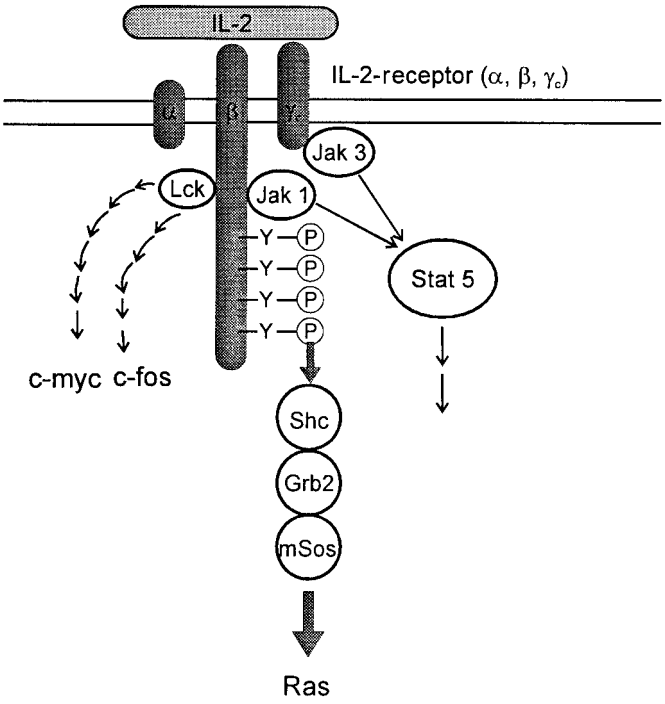



**Fig. 11.4.** Model of signal transduction via the IL-2 receptor. Binding of IL-2 to the IL-2 receptor initiates activation of the Janus kinases Jak1 and Jak3. These phosphorylate tyrosine residues in the β-chain of the IL-2 receptor and in the transcription factor Stat5. SH2 domains or PTB domains of adaptor proteins can bind to the Tyr phosphate residues of the β-chain and, as shown in the figure for the Shc/Grb2/Sos complex, can transmit a signal in the direction of the Ras pathway. The phosphorylated transcription factor Stat5 is translocated into the nucleus and activates the transcription of corresponding gene sections. Another signaling pathway starting from the activated IL-2 receptor involves the Lck and Syk tyrosine kinases (see Chapter 8). The pathway leads to induction of genes for transcription factors such as c-Myc and c-Fos.

with the β chain of the IL-2 receptor and is activated on binding of IL-2. The $NH_2$ terminal region of the catalytic domain of Lck kinase and an acidic region in the cytoplasmic part of the β chain are involved in association of both proteins.

In addition to the Lck kinase, other tyrosine kinases are activated on ligand binding to the IL-2 receptor. These are Fyn kinase and the Janus kinases Jak1 and Jak3.

Starting from the activated tyrosine kinase, the signal may be conducted along different signaling pathways depending on the receptor type.

A first essential step in signal conduction is the Tyr phosphorylation of intracellular structural elements of the cytokine receptor, catalyzed by the activated kinase. The phosphotyrosine residues serve as attachment points for SH2 domains of downstream signal proteins, which form links to central signaling pathways. The PI3-kinase and the adaptor protein Shc have been identified as such linkers. The latter enables coupling to the Ras pathway and the MAP kinases. Activation of the Ras pathway following ligand binding to cytokine receptors has been reported many times and it is assumed that the link to the Ras protein is formed by Shc-Grb2-mSos interaction (see 9.5.2).

Cross-phosphorylation of the EGF receptor has also been observed during the process of activation of cytokine receptors, whereby crosstalk with receptor tyrosine kinases is possible.

The following signaling pathways are activated on ligand binding to cytokine receptors:

- Jak-Stat pathway
- Ras pathway
- MAP kinase pathway
- Protein kinase C
- PI3-kinase pathway

Cytokine signals may be conducted into the nucleus at the transcription level via these central signaling pathways. As a consequence, activation takes place of transcription factors with growth regulating and morphogenetic functions.

## 11.1.3 The Jak-Stat Pathway

The Jak-Stat pathway is a signaling pathway, starting from cytokine receptors, that enables a very direct signal transduction from the membrane to the cell nucleus using only a few coupling elements. Many cytokines use this pathway to bring about a rapid change in the transcription activity of specific gene sequences (review: Darnell et al., 1994; Ihle et al., 1995; Hill and Treisman, 1995; Taniguchi, 1995, Moutoussamy et al., 1998).

### 11.1.3.1 The Janus Kinases

Another family of protein kinases involved in signal transduction via cytokines includes the *Janus kinases* (*Jak kinases*). At least four different Jak kinases are known in mammals (Jak1, Jak2, Jak3 and Jak4). A characteristic feature of the structure of Jak kinases is the occurrence of two tyrosine kinase domains (Fig. 11.5). However, only

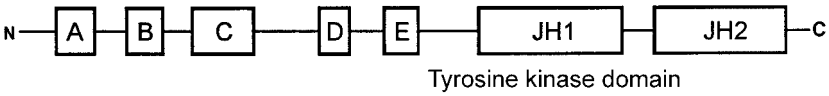



**Fig. 11.5.** Domain structure of the Jak kinases. JH1 is the catalytic tyrosine kinase domain. JH2 shows similarity to a tyrosine kinase domain. The domains A—E are homologous elements of the Jak kinase family. JH: Janus kinase homology region.

one of the two domains possesses all of the structural features considered as necessary for a functioning kinase activity. The importance of the occurrence of two tyrosine kinase domains is not clear.

It is assumed that the Jak kinases bind to a cytoplasmic section of the receptor, which is in the vicinity of the membrane and contains two conserved sequence elements, Box 1 and Box 2. Binding of the Jak kinases leads to their activation, a process linked to mutual phosphorylation of the Jak kinases (Fig. 11.6). Activation of the Jak kinases may take place in a homodimeric receptor complex or it may also occur in hetero-oligomeric complexes. The circumstances are complicated in that two (or more) Jak kinases may associate at an activated receptor. Two different Jak kinases are required for signal transduction via interferon receptors (see Fig. 11.8). Furthermore, the different Jak kinases are specific for the corresponding receptors.

#### 11.1.3.2 The Stat Proteins

Starting from the activated Jak kinases, a signaling pathway leads directly to transcription factors that are phosphorylated by the Jak kinases on tyrosine residues and activated for stimulation of transcription (review: Horvath and Darnell, 1997). These transcription factors belong to a class of proteins known as Stat proteins (Stat = signal transducer and activator of transcription). At least seven different Stat proteins are known (Stat1–4, Stat5a, Stat5b, Stat6). The first Stat proteins, Stat1 and Stat2, were found in association with signal transduction via interferon γ.

The Stat proteins are found in a latent form in the cytosol and are activated by cytokine receptors and their associated kinases. On binding of the cytokine to the receptor and activation of the Jak kinase, the Stat proteins are recruited, via their SH2 domains, to the receptor-kinase complex and are then phosphorylated by the Jak kinase on a conserved Tyr residue at the C-terminus.

The phosphorylated Stat proteins form homodimeric or heterodimeric complexes and are transported as such into the nucleus (Fig. 11.7). In the nucleus, they bind to corresponding DNA elements in the promoter region of cytokine responsive genes. Stimulation of transcription takes place in cooperation with other proteins such as p300, CBP (see 1.4.6), glucocorticoid receptors and c-jun.

The Stat proteins have SH2 and SH3 domains, a DNA binding domain and a C-terminal domain required for transcription activation. The activating phosphorylation takes place for Stat1 on Tyr701 in the vicinity of the C-terminus. In the unphosphorylated form, the Stat proteins exist as monomers, whereas in the phosphorylated form, they are dimers.

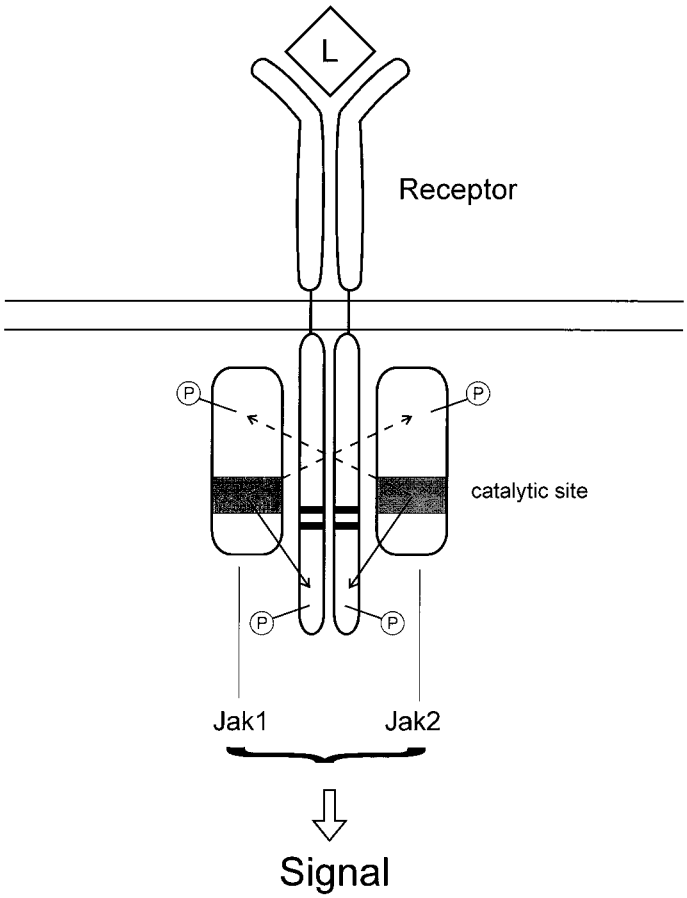



**Fig. 11.6.** Model of activation of Jak kinases. The Jak kinases (Jak1 and Jak2 are shown as examples here) are attributed a two-fold function in signal transduction via cytokine receptors. On binding to the activated cytokine receptor, the Jak kinases are activated and phosphorylation of the Jak kinases takes place, probably by a trans mechanism (dashed arrow). The Jak kinases also catalyze Tyr phosphorylation of the cytoplasmic domain of the receptor (solid arrow). The phosphotyrosine residues serve as attachment points for adaptor proteins or other effector proteins.

Dimerization is mediated by the phosphotyrosine residue and the SH2 domain. Highly resolved structural investigation show that the phosphotyrosine residue of one Stat protein binds to the SH2 domain of the partner and vice versa, so that the phosphotyrosine-SH2 bonds function as a double clasp (structure in complex with DNA: Becker et al., 1998; Chen et al., 1998). The binding to DNA is in the form of a dimer, with the Stat-DNA complex showing a large similarity to the structure of the NFκB-DNA complex (see Fig. 1.10).

The Jak-Stat pathway for the interferons α, β and γ has been well investigated. The interferon γ pathway uses Stat1 protein, which binds as a homodimer to the correspon-

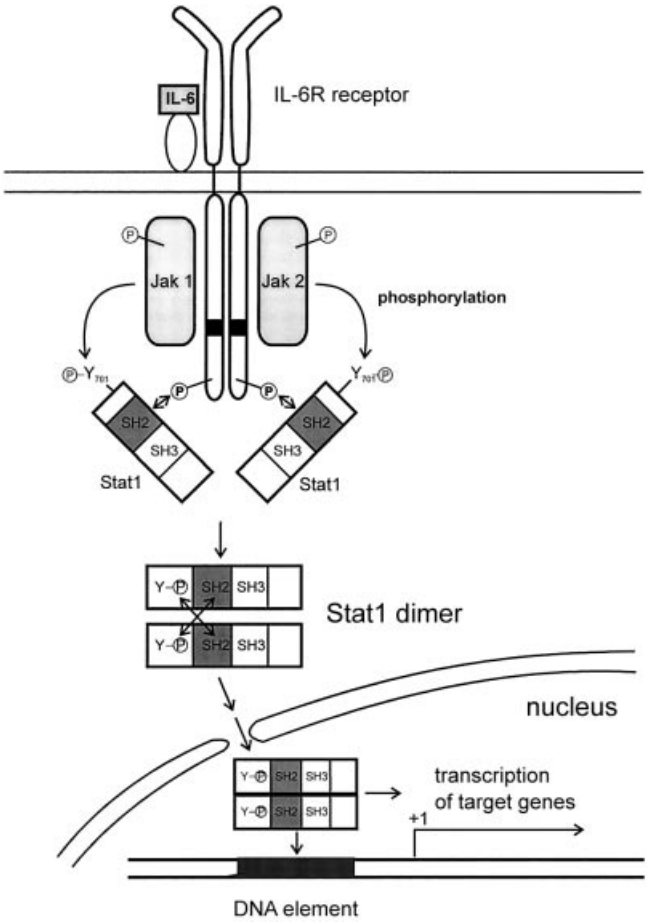



**Fig. 11.7.** Model of activation of Stat proteins. The Stat proteins are phosphorylated (at Tyr701 for Stat1) as a consequence of binding to the receptor-Jak complex, and Stat dimers are formed. The dimerization is mediated by phosphotyrosine-SH2 interactions. In the dimeric form, the Stat proteins are transported into the nucleus, bind to corresponding DNA elements, and activate the transcription of neighboring gene sections. In the figure, activation of Stat proteins is shown using the IL-6 receptor as an example (according to Taniguchi, 1995). Other Jak kinases and Stat proteins may also take part in signal conduction via IL-6, in addition to the Jak kinases and Stat1 shown.

ding GAS element (GAS: interferon γ activation site). The GAS element includes the consensus sequence TTNCNNNAA and is found in a modified form in many cytokine-controlled gene sections. The signaling pathway of the interferons α and β uses the Stat1 and Stat2 proteins in particular; binding to the corresponding DNA binding element also requires the help of a 48 kDa protein (Fig. 11.8).

The Jak-Stat signal transduction is an example of a signaling pathway in which a signal is coupled, in the form of a tyrosine phosphorylation, directly to activation of a transcription factor. In contrast to other signaling pathways that also regulate trans-

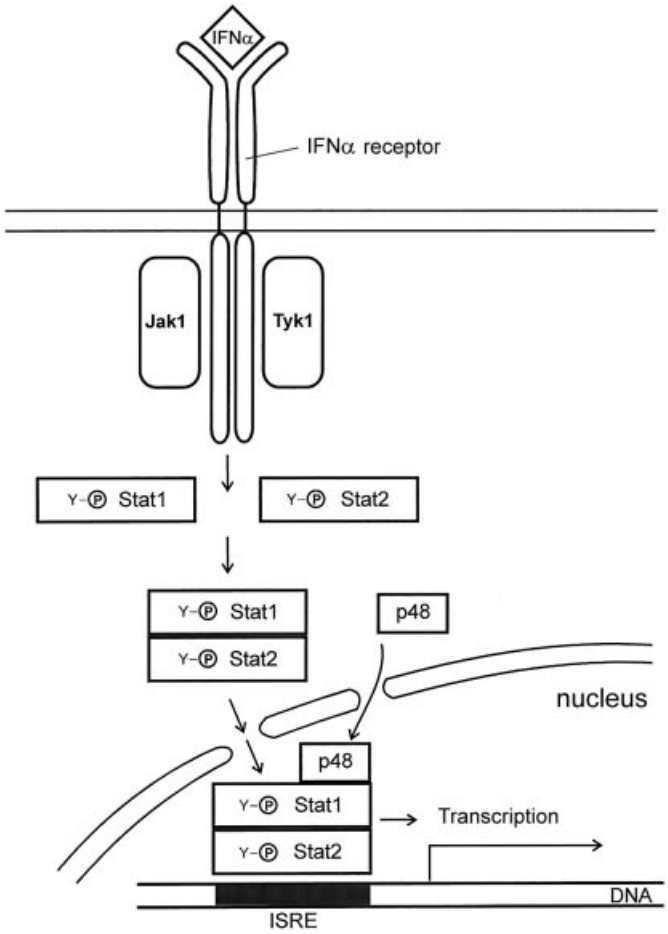



**Fig. 11.8.** Scheme of signal transduction via interferon α. The receptor for interferon α (IFNα) binds and activates the Jak kinases Jak1 and Tyk1. These phosphorylate the Stat factors Stat1 and Stat2, leading to formation of Stat1-Stat2 heterodimers. The heterodimers are transported into the nucleus and bind to a corresponding DNA element known as ISRE (interferon stimulated response element). Another protein, p48, is also involved in transcription activation of the interferon regulated gene.

cription processes, e.g., the Ras/MAPK pathway, the Jak-Stat pathway is impressive in its simple concept and the small number of components involved.

Several factors determine the specificity and diversification of signal transduction in the Jak-Stat pathway (see Darnell et al., 1994):

### Receptor Tyrosine Kinase Combination

At the level of the receptor tyrosine kinase complex, many different combinations are possible since the structure of the cytokine receptors is very variable and there are various (at least 4) Jak kinases.

**Diversity of the Stat Proteins**

There are at least six different Stat proteins, from which it can be assumed that these are phosphorylated with differing efficacy by the various receptor tyrosine kinases. In addition, expression of Stat proteins is cell and tissue specific.

**DNA Level**

At the DNA level, variability of the DNA elements specific for Stat proteins and accessibility of the promoter sites are further determining factors for specificity of signal transduction.

## 11.2 T and B cell Antigen Receptors

At the surface of T and B lymphocytes, specific receptors are found that bind antigens and set intracellular signal chains in motion (review: Weiss and Littman, 1994; Paul and Seder, 1994; Qian and Weiss, 1997). These may lead to increased cell division, programmed cell death or a functional recoining of lymphocytes.

The receptors of the B lymphocytes recognize antigens in the form of foreign proteins, which exist in soluble, particle-bound or cell-bound forms.

The receptors of the T lymphocytes, in contrast, recognize antigens only in the course of a cell-cell interaction between the T lymphocyte and an antigen-presenting cell. The antigen-presenting cell presents the processed (i.e., proteolytically digested to peptides) foreign protein as a peptide. The peptide is bound to the MHC complex (MHC: major histocompatibility complex) of the antigen-presenting cell and is recognized in this form by the receptor of the T lymphocyte (see also Fig. 11.1). The MHC proteins are transmembrane proteins that exist as heterodimers and possess a binding site in the extracellular region for the antigen to be presented (= processed peptide). B and T cell antigen receptors bind the antigen under very different conditions. Despite this, the binding event triggers very similar signal chains and initiates similar reactions.

### 11.2.1 Receptor Structure

The T and B cell antigen receptors are generally composed of several subunits, whereby the functions of ligand binding and conduction of the signal are localized on separate subunits (review: Weiss and Littman, 1994; Wilson and Garcia, 1997).

The T cell antigen receptor may be divided into two functional complexes, in which one mediates ligand binding and the other performs the signal transduction. Antigen binding takes place via the Ti-$\alpha$ and $\beta$ subunits, which only have very short cytoplasmic structural portions and are not directly involved in conduction of the signal on the cytosolic side. The function of signal conduction is performed by other polypeptide chains, namely the CD3$\gamma$, CD3$\delta$ and CD3$\varepsilon$ chains and the $\zeta$ chain (Fig. 11.9). These

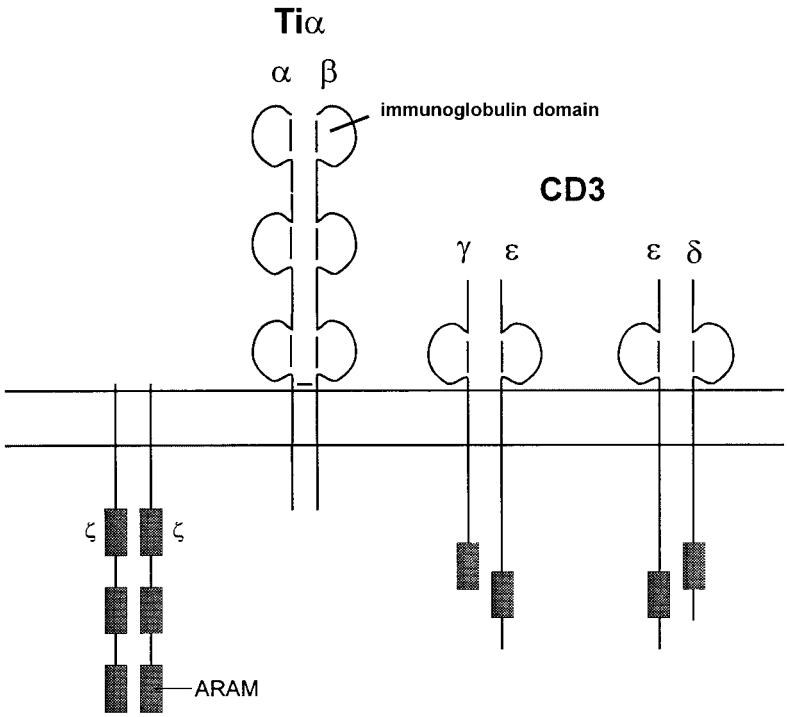



**Fig. 11.9.** Subunit structure of the T cell receptors. The figure shows the different subunits of T cell receptors in a highly simplified representation. The stoichiometry of the subunits in the complete receptor is not clear. The αβ chains are also known as the Tiαβ complex; the γε and δε chains together form the CD3 complex. ARAM: antigen recognition activation motif.

have characteristic sequence motives on their cytoplasmic side, which are required for conduction of the signal further to downstream protein tyrosine kinases. The sequence motives include two pairs of Tyr and Leu residues in the consensus motif (D/E)XXYXXL$(X)_{6-8}$YXXL and are known as the *antigen recognition activation motif* (ARAM) or also as the *immunoreceptor tyrosine activation motif* (ITAM). ARAM motives are found in the CD3γ, CD3δ and CD3ε chains and in the ζ chain. ARAM motives also occur in B cell receptors. Following TCR stimulation, the tyrosine residues within the ITAMs (ARAMs) become phosphorylated, creating binding sites for SH2 domain containing proteins. By providing a link to the various downstream signaling pathways, binding of these proteins is a critical step in the signaling function of these receptors.

Cooperation with other receptors that may help in a synergistic manner to trigger a signal is a particular feature of signal transduction via T and B cell antigen receptors. These other receptors are known as *coreceptors*. Examples are the CD4 and CD8 proteins, which are involved in activation of T cell antigen receptors. The coreceptors are essential for signal transduction and are involved in binding and activation of downstream tyrosine kinases such as Src kinase. Furthermore, they have an amplifying effect on the sensitivity and specificity of antigen binding.

### 11.2.2 Intracellular Signal Molecules of the T and B Cell Antigen Receptors

The T and B cell antigen receptors do not have any intrinsic protein tyrosine kinase activity. Rather, antigen binding leads to recruitment and activation of protein tyrosine kinases on the cytoplasmic side of the receptor. Probably, coupling of the tyrosine kinase takes place via the ARAM motif directly.

Four classes of protein tyrosine kinases are involved in a complex manner in signal transduction via T and B cell antigen receptors. These are the families of the Src (Lck, Fyn), Csk, Tsk and Syk (Syk, ZAP70) kinases (review: Weiss and Littman, 1994; Qian and Weiss, 1997). As a consequence of antigen binding, phosphorylation of the two Tyr residues of the ARAM motif is observed, mediated preferentially by Lck kinase. ZAP70 kinase (ZAP = ζ associated protein 70) binds with the help of its SH2 domains to the thus-created phosphotyrosine residues of the ARAM motif (Hatada et al., 1995). As a consequence of association with the ζ chain, ZAP kinase is phosphorylated and its tyrosine phosphates form attachment points for other signal proteins such as the protein kinase Lck, the Abelson kinase, the guanine nucleotide exchange factor Vav, the GTPase activating protein Ras-GAP and the protein tyrosine phosphatase SHP-1.

In this reaction chain, ZAP kinase performs an essential function. The phosphorylated ARAM motives recruit ZAP70 kinase to the T cell receptor complex and mediate high affinity binding to the kinase. The two SH2 domains of ZAP70 kinase bind in tandem fashion to the twice-phosphorylated ARAM motif. ZAP70 kinase shows a complex pattern of Tyr phosphorylation, including phosphorylations with an inhibitory effect.

An inhibitory phosphorylation of the tyrosine kinases Lck and Fyn is also important for signal transduction, and this is performed by Csk kinase. This phosphorylation takes place at the C-terminus of Lck and inhibits the kinase activity in a similar way to that already structurally illustrated for Src kinase (see 8.3.2).

The further course of signal transduction is variable. The following signaling pathways are activated following stimulation of T cell receptors:

- Ras pathways with MAPK cascades
- Hydrolysis of PtIns phosphates, via phospholipase Cγ
- $Ca^{2+}$ signaling pathways: via phosphorylation of the $InsP_3$ receptor and via PLCγ.

## 11.3 Signal Transduction via Integrins

Structure and function of the cell formations of higher organisms are highly dependent on adhesive interactions based on direct cell-cell contact and on interactions of cells with the extracellular matrix. Adhesion between cells and with the extracellular matrix has a regulatory influence on migratory behavior, proliferation and differentiation of an individual cell within the cell formation. The adhesion processes thus do not only serve to simply hold cells together in the formation. They also have a regulatory effect

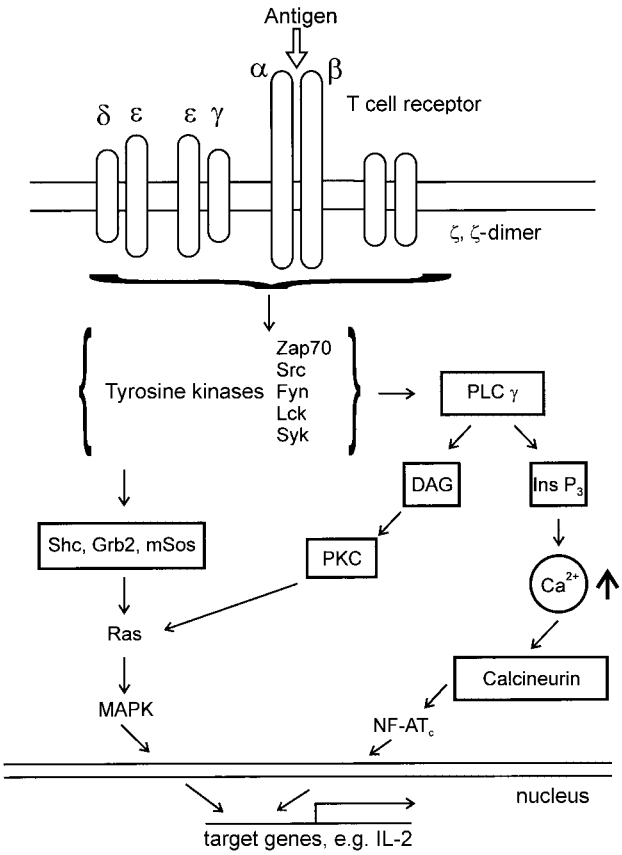



**Fig. 11.10.** Overview of signaling pathways associated with activation of lymphocytes. The triggering signal for activation of T lymphocytes is generally antigen binding to the T cell receptor (see also Fig. 11.1). The activated receptor passes the signal on to different associated tyrosine kinases. The tyrosine kinases phosphorylate and activate downstream effector proteins such as phospholipase Cγ. A $Ca^{2+}$ signal is thereby created, leading to activation of the protein phosphatase calcineurin. Calcineurin dephosphorylates the transcription factor $NF\text{-}AT_c$ which is then transported into the cell nucleus and activates target genes such as the IL-2 gene (see also Fig. 7.16). Activation of the tyrosine kinase also leads to formation of binding sites for adaptor proteins, in the form of phosphotyrosine residues, enabling a link to the Ras MAPK pathway. PLCg: phospholipase Cγ; DAG: diacylglycerol; PKC: protein kinase C; IL-2: interleukin 2.

on central physiological functions of a cell. Surface receptors that can specifically bind to a neighboring cell or to the extracellular matrix serve as mediators of adhesion processes, and as a consequence, intracellular signaling pathways are activated. The protein family of the integrins are one such group of surface receptors which are of central importance (review: Clark and Brugge, 1995; Howe et al., 1998).

The integrins define attachment points for the extracellular matrix and for contact with neighboring cells, and they are involved in signal transduction into the cell interior. With these functions, the integrins are involved in regulation of embryonal

growth, tumor formation, programmed cell death, tissue homeostasis and many other processes in the cell. The integrin receptors (usually just called integrins) have a special position amongst the transmembrane receptors. The integrins form a transmembrane bridge between extracellular scaffold structures and the intracellular cell scaffold. They also convert cell-cell contacts and interactions with the extracellular matrix (ECM) into intracellular signal processes.

The integrin receptors are made up of $\alpha$ and $\beta$ chains which each have a transmembrane element. At least 16 different $\alpha$ chains and 8 different $\beta$ chains are known at present. The integrin receptors form heterodimers, which is why there is such a large structural and functional diversity of integrins. The ligands of the integrins are mostly components of the extracellular matrix such as fibronectin and collagen. These are multivalent ligands immobilized on fibrillar structures. Extracellular ligands may, however, also be soluble proteins or surface proteins of neighboring cells.

The ligands of the integrin receptors are generally multivalent and their binding leads to crosslinking and clustering of integrin receptors. By this process, the integrin receptors are activated for further signal transduction.

Signal transduction by integrins can be roughly divided into two descriptive categories:

- Direct signal transduction: here ligand binding and clustering are the only extracellular signals and direct activation takes place of downstream signaling events, such as activation of nonreceptor tyrosine kinases and other proteins. Targets here are especially the cytoskeleton and MAPK pathways.
- Collaborative signaling: integrin-mediated cell adhesion modulates signal transduction by other receptors (e.g., receptor tyrosine kinases).

**Direct Signal Transduction**

An important function of integrins is the formation of bridges between the extracellular matrix and the intracellular actin cytoskeleton. This bonding often takes place in structures known as *focal adhesion points*. These are multiprotein complexes in which specific contacts are formed on the cytosolic side between the integrin receptors, proteins of the intracellular matrix and actin filaments (Fig. 11.11). Focal adhesion points form linkages between morphological structures of the cell and signal transduction pathways. These multiprotein complexes have an attachment point for the actin cytoskeleton and binding sites for signal proteins. In the mean time, many proteins have been described that bind to the intracellular side of integrins, including the $Ca^{2+}$ binding protein calreticulin.

As a consequence of ligand binding and clustering of integrins, activation of different signaling pathways is initiated:

- Activation of nonreceptor tyrosine kinases
- Activation of Ser/Thr-specific proteins kinases such as the integrin-linked kinase (ILK)
- Increase in $Ca^{2+}$ concentration
- Activation of the MAPK cascade
- Increased formation of Ptd-Ins messenger substances

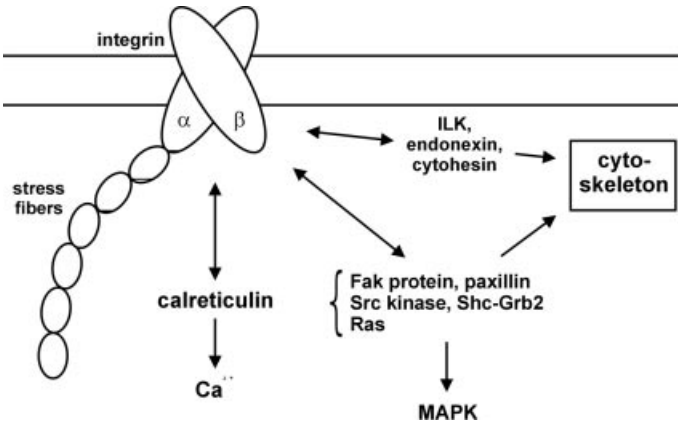



**Fig. 11.11.** Model of signal transduction via integrins. Activated integrins transmit signals from the extracellular matrix to the cytoskeleton and activate various intracellular signaling pathways. The diagram lists some of the signal proteins that have been shown to be involved in integrin signaling. The signal conduction involves – among others – the Ras/MAPK pathway and $Ca^{2+}$-signaling pathways, setting in motion a broad spectrum of subsequent reactions. At least two protein kinases participate in signal transduction, the integrin linked kinase (ILK) and the focal adhesion kinase (FAK).
Furthermore proteins involved in reorganization of the cytoskeleton (paxillin, cytohesin, endonexin) and in formation of stress fibers associate with activated integrins. Association of the $Ca^{2+}$-binding protein calreticulin with integrins is thought to link integrin signaling with $Ca^{2+}$-signaling. Signal conduction via integrins occurs in multiprotein complexes. It is therefore only possible to show selected protein components of these complexes.

The integrins do not have any enzyme activity in their own cytoplasmic domain, but on ligand binding, stimulation of tyrosine phosphorylation is observed on the cytoplasmic side of many cells, such as fibroblasts and platelets. The exact configuration of protein-protein interactions on the cytosolic side of the integrins is not clear and the mechanism of stimulation of protein tyrosine kinases is unknown. Some components of the focal adhesion points, such as the structural protein tensin, have SH2 and SH3 domains that may serve as specific attachment points for tyrosine kinases and other signal proteins.

Of the protein tyrosine kinases, the focal adhesion kinase (FAK, $p125^{FAK}$) plays an important role in integrin signal transduction. The FAK protein undergoes autophosphorylation on integrin activation. It is assumed that the phosphotyrosine residues of FAK serve as binding sites for other SH2-containing signal molecules (Fig. 11.11). The FAK protein also has a specific binding domain for the structural protein paxillin and is found in a defined complex together with paxillin, the Src kinase and another protein. Paxillin is a component of cytoskeleton structures, so recruitment of FAK to the cytoskeleton in the region of focal adhesion points seems possible via a paxillin-FAK interaction.

The integrin-mediated activation of MAPK pathways seem particularly important for integrin function since this has an influence on transcription processes. The model of the mechanism of this linkage usually employs the Ras protein as a central switching

station, although it is not clear how the link to the Ras protein is formed. Furthermore, there are also links between the integrins and the Rho/Rac GTPases. Here, it is interesting to note that the integrin-MAPK linkage can trigger the same biological events as growth factors that bind to transmembrane receptors.

### Collaborative Signaling

In addition to directly generated signals, integrins can also modulate signaling responses to external signaling molecules such as growth factors. The integrins influence efficiency of signal transduction via receptor tyrosine kinases and the subsequent activation of the MAPK pathway. Integrin-mediated activation of the PDGF receptor has also been observed, which is independent of ligand binding at this receptor.

The link between cell anchoring and integrin activation seems to be of great biological importance. A link has been established between cell anchoring and regulation of the cell cycle machinery, although the exact nature of this link is not known. Integrin-mediated cell anchoring is also known to regulate a set of events impacting on apoptotic pathways. Along these lines, activation of PI3-kinase, Akt kinase (see Chapters 6 and 9) and influence on the Bcl-2 proteins and caspases (Chapter 15) have been established.

## References Chapter 11

Becker, S., Groner, B. and Muller, C.W. 'Three-dimensional structure of the Stat3beta homodimer boand to DNA' (1998) Nature *394*, 145–151

Chen, X., Vinkemeier, U.,. Zhao, Y., Jeruzalmi, D., Darnell, J.E. Jr and Kuriyan, J. 'Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA' (1998) Cell *93*, 827–839

Clark, E.A. and Brugge, J.S. 'Integrins and signal transduction pathways: the road taken' (1995) Science *268*, 233–239

Darnell, J.E., Kerr, I.M. und Stark, G.R. 'Jak-Stat pathways and transcriptional activation in response to IFNs and other extracellular signaling pathways' (1994) Science *264*, 1415–1420

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

Hatada, M.H., Lu, X., Laird, E.R., Green, J., Morgenstern, J.P., Lou, M., Marr, C.S., Philipps, T.B., Ram, M.K., Theriault, K., Zoller, M.J. und Karas, J.L. 'Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor' (1995) Nature *377*, 32–38

Hill, C.S. and Treisman, R. 'Transcriptional regulation by extracellular signals: mechanisms and specificity' (1995) Cell *80*, 199–211

Horvath, C.M. and Darnell, J.E. 'The state of the STATs: recent developments in the study of signal transduction to the nucleus' (1997) Curr. Op. Cell Biol. *9*, 233–239

Howe, A., Aplin. A.E., Alahari, S.K. and Juliano, R.L. 'Integrin signaling and cell growth control' (1998) Curr Opin Cell Biol;10, *220*-231

Ihle, J.N. und Kerr,I .M. ' Jaks and Stats in signalling by the cytokine receptor superfamily' (1995) Trends Gen. *2*, 69–74

Miyamoto, S.; Akiymam, S.K. and Yamada, K.M. 'Synergistic roles for receptor occupancy and aggregation in integrin transmembrane function' (1995) Science *267*, 883–885

Moutoussamy, S., Kelly, P.A. and Finidori, J. 'Growth-hormone-receptor and cytokine-receptor-family signaling' (1998) Eur J Biochem *255*, 1–11

Paul, W.E. und Seder, R.A. 'Lymphocyte responses and cytokines' (1994) Cell *76*, 241–251

Qian, D. and Weiss, A. 'T cell antigen receptor signal transduction' (1997) Curr. Op. Cell Biol. *9*, 205- 212

Taniguchi,T. 'Cytokine signalling through nonreceptor protein tyrosine kinases' (1995) Science *268*, 251–255

Weiss, A. and Littman, D.R. 'Signal transduction by lymphocyte antigen receptors' (1994) Cell *76*, 263–274

Wilson, I.A. and Garcia, K.C. 'T-cell receptor structure and TCR complexes' (1997) Curr. Op. Struct. Biol. *7*, 839–848

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 12**

# Other Receptor Classes

The cell has other transmembrane receptors and signaling pathways that do not fit into the „classical" receptor types and signal mechanisms described in Chapters 5–11. The following signaling pathways certainly do not complete the list of intercellular and intracellular communication mechanisms in mammals, and it is to be expected that other classes of signaling pathways will be described in the future.

## 12.1 Receptors with Intrinsic Ser/Thr Kinase Activity: the TGFβ Receptor and the Smad Proteins

In addition to transmembrane receptors with intrinsic tyrosine kinase activity, the cell also contains transmembrane receptors with intrinsic Ser/Thr kinase activity. An example of such a receptor type is the receptor for *transforming growth factor* β *(TGF*β*)*. TGFβ belongs to a class of extracellular signal proteins that includes the *cytokines* (see also 11.1). TGFβ and related proteins regulate cell growth and morphogenesis of various cells. On binding of TGFβ to the TGFβ receptor, a signal chain is activated that leads signals to the transcription level and thus influences progress of the cell cycle. Activation of the TGFβ receptor creates antimitogenic (i.e., inhibiting cell division) signals, which are manifested as increased production of inhibitors of cell cycle specific protein kinases (see Chapter 13).

### 12.1.1 TGFβ Receptor

The extracellular signal molecule TGFβ is the activating ligand for an oligomeric transmembrane receptor with subunits containing a single transmembrane element and a cytoplasmic Ser/Thr kinase domain. The TGFβ receptor has a hetero-tetrameric structure in its activated form; it contains two copies of two different subunits known as TGFβ receptor I and TGFβ receptor II. Both subunits carry a Ser/Thr kinase domain and the activity of TGFβ receptor I is controlled by ligand binding.

The mechanism of activation of the TGFβ receptor is shown schematically in Fig. 12.1 (review: Wrana et al., 1994). The dimeric ligand TGFβ first binds to TGFβ receptor II and is recognized in this form by TGFβ receptor I. TGFβ receptor I binds to the complex of TGFβ and TGFβ receptor II, forming a hetero-oligomeric, ligand-bound receptor complex in which phosphorylation of receptor I by receptor II takes place. The phosphorylation of receptor I occurs in a Gly/Ser-rich domain in the vicinity of the

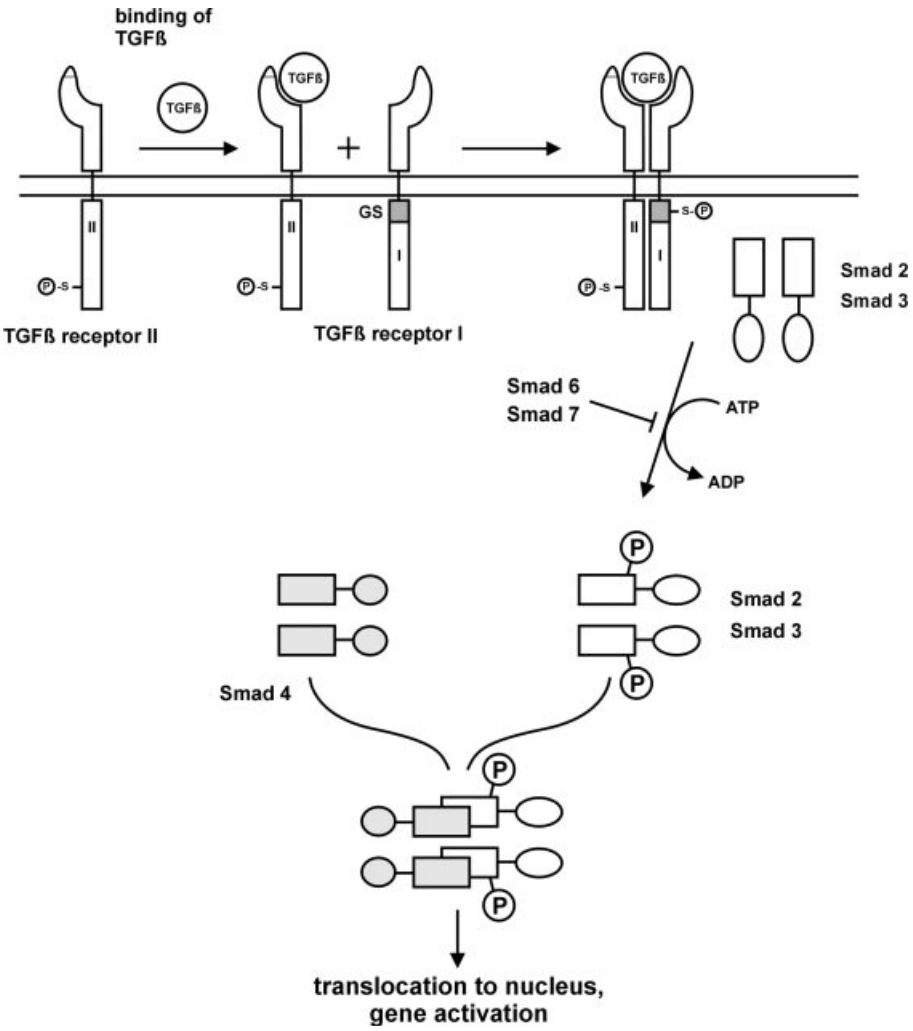



**Fig. 12.1.** Model of signal transduction via the TGFβ receptor.
Signal transduction via TGFβ requires two TGFβ receptors, the TGFβ receptor I and the TGFβ receptor II. TGFβ first binds to TGFβ receptor II, which has a constitutive Ser/Thr-specific protein kinase activity. Binding of TGFβ to the TGFβ receptor II induces association of TGFβ receptor I with TGFβ receptor II; this phosphorylates the TGFβ receptor I on a glycine-serine rich domain (GS) and activates it for conduction of the signal further. The figure does not adress the correct oligomeric state of the activated receptor.
The activated TGFβ receptor phosphorylates the pathway-restricted Smad proteins (Smad2 and Smad 3) which then associate with the common mediator Smad protein (Smad 4). The complex between Smad4 and phosphorylated Smad2/Smad3 translocates into the nucleus where it binds to specific DNA elements of target genes and activates transcription. The inhibitory Smads (Smad 6, Smad 7) can also bind to the TGFβ receptor and may prevent phosphorylation of Smad 2/Smad3 by the TGFβ receptor (According to Wrana et al., (1994) and Heldin et al., 1997).
TGFβ: transforming growth factor type β

cytoplasmic side of the transmembrane element. This converts the receptor into the active state from which the signal is transferred to effector molecules. This receptor type uses two subunits for signal transduction: type II of the TGFβ receptor recognizes the signal, and type I determines specificity for further signal conduction, in that it gives the signal to the downstream effector molecules.

Activation of TGFβ is accompanied by multiple phosphorylation on cytoplasmic parts of the receptor; this includes Tyr phosphorylation in addition to Ser/Thr phosphorylation. It is assumed that these autophosphorylations occur in trans, i.e., between the protomers of the receptor. From the activated TGFβ receptor, signals are directed to the transcription level. A set of proteins, known as *Smad proteins*, are involved in this signal conduction (see Fig. 12; review: Heldin et al., 1997, Massague, 1998). Besides Smad proteins, G-proteins and the MAPK cascade are also involved in the downstream signaling of TGFβ family members.

### 12.1.2 Smad Proteins

At least nine different Smad proteins have been identified in higher organisms, and these may be divided into three functional classes:

- **Pathway-restricted Smad proteins** (Smad 1, 2, 3): These Smad proteins are effector molecules directly downstream of type I TGFβ receptor. They bind to TGFβ receptor I and are phosphorylated and activated for further signal conduction by the receptor. Phosphorylation takes place on a C-terminal SSXS motif.
- **Common mediator Smads:** The mode of action of Smad 4 clearly differs from that of the other members of the Smad family. Smad 4 binds to phosphorylated pathway-restricted Smads and forms hetero-oligomeric complexes. These translocate to the nucleus where they bind to related DNA elements and activate the transcription of target genes, e.g., the genes for the CDK inhibitors p15 and p21 (see Chapter 13) (Fig. new). Activation of transcription often occurs in cooperation with other coactivators (see 1.4.3.2).
- **Inhibitory Smads:** Smad 6 and Smad 7 function as inhibitors of TGFβ signaling. They bind to type I receptors and interfere with phosphorylation of the pathway-restricted Smads. These Smad proteins may be used to modulate and weaken the TGFβ receptor. Such transmodulation has been described for the interferon γ pathway. Activation of the IFNγ pathway leads to increased transcription of Smad 7, which diminishes signal conduction in the TGFβ pathway (Ulloa et al., 1999).

Structure-function investigations have shown that the Smads have a sequence-specific DNA binding domain and two regulatory domains known as the MH1 and MH2 domains. It is assumed that the pathway-restricted Smads are kept in an inactive form by intramolecular interactions between the MH1 and MH2 domains. The inactive conformation is lifted by phosphorylation at the C-terminus upon activation of the TGFβ receptor, and an interaction with the common mediator Smad 4 becomes possible. According to these observations, the MH1 domain functions as a negative regulator of the MH2 domain.

DNA binding of Smad proteins takes place via a DNA binding motif, which differs substantially from classical DNA binding motives. The DNA binding domain of the Smad proteins contains a β-sheet motif, which positions itself in the large groove of the DNA (Shi et al., 1998).

In total, TGFβ-Smad signal conduction has distinct similarities to signal conduction in the Jak-Stat pathway (see 11.1.3). In both pathways, cytosolic transcription factors are activated by phosphorylation and are translocated in oligomeric complexes to the nucleus and the DNA. Common to both pathways is the short distance from the extracellular signal to the transcription level.

## 12.2 Notch: Signaling with Protease Participation

The Notch signaling pathway is of great importance for neuronal differentiation processes. The Notch signaling pathway mediates cell-cell contacts used to control the cell fate of neighboring cells. The Notch proteins are ligand-activated transmembrane receptors with the transmembrane proteins of neighboring cells as ligands. On binding a ligand during a cell-cell interaction, a direct signal is transmitted to the transcription level and various target cells are activated.

The mechanism of Notch signal conduction was a mystery for a long time until it was established that specific proteolysis is an essential step (review: Chan and Yan, 1998, Artavanis-Tsanokas et al., 1999). The Notch receptor is a hetero-oligomeric protein that spans the membrane with a transmembrane element and contains a large cytosolic domain. The intracellular Notch domain is proteolytically cleaved off on ligand binding to the Notch receptor (Fig. 12.2). This ligand-dependent proteolytic release of the intracellular domain is sufficient for signal conduction further. The intracellular domain migrates to the nucleus and activates transcription as part of a complex with a family of transcription factors known as CSL. CSL proteins (e.g. CBF1 in mammals) can bind to specific sites on the DNA activating the transcription of basic helix-loop-helix proteins (see 1.2.1.3) which in turn affect the transcription of downstream target genes.

Proteolytic cleavage of the Notch receptor occurs in the vicinity of its transmembrane section. It is unclear which protease is involved in this processing and how the protease is activated by ligand binding.

The Notch signaling pathway is thus a further example of signal conduction in which the signal is directed by the shortest pathway from the cell membrane into the nucleus.

## 12.3 Signal Transduction via the Two-component Pathway

The two-component pathway was originally discovered in bacteria. It was only recently recognized that this kind of signal transduction is also used in eucaryotes. Bacteria possess signal systems which they use to react to changes in N availability, osmolarity and to chemotactic substances. The signaling pathway responsible for this regulation is

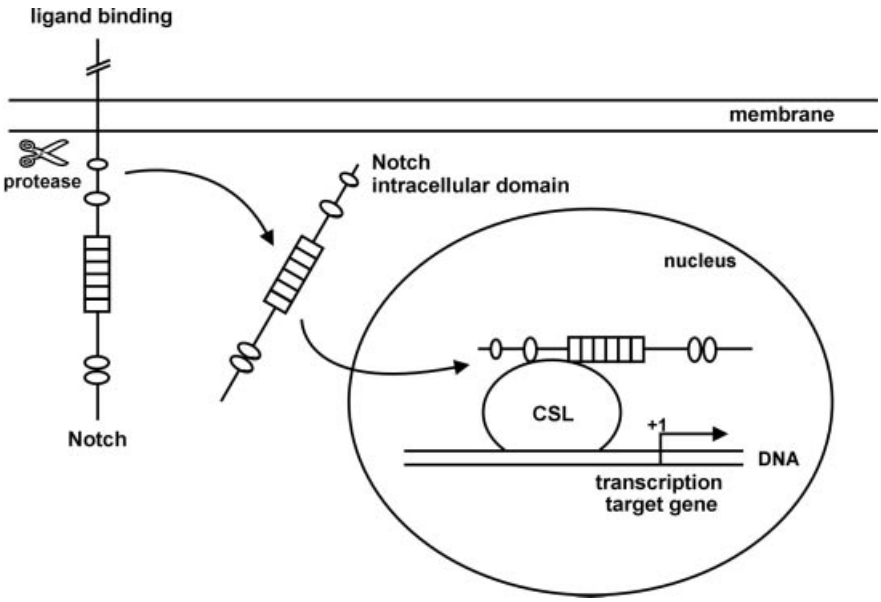



**Fig. 12.2.** Model of Notch signalling. The Notch protein is a ligand activated transmembrane receptor which is subject to proteolysis of the intracellular domain upon ligand binding. The nature of the protease involved is still a matter of debate. The proteolytically released intrancellular domain of Notch (NICD) translocates into the nucleus where it interacts with a family of transcription factors, the CSL proteins, resulting in a change in transcription of target genes.

known as the *two-component pathway*, due to the involvement of two conserved functional and structural domains.

The two-component pathway is characterized by two functional elements. A *histidine-specific protein kinase* functions as a „*sensor*" that registers an external signal and passes this on to a downstream *response regulator*. The latter is activated by phosphorylation during the process of signal transduction, triggering other reactions in the cell (Fig. 12.3).

The composition of the two-component pathway is very variable. The nature of the external signal and the reactions triggered in the cell may be very diverse. The sensor kinase may be a part of the receptor that registers the signal, or it may be on a polypeptide chain separate from the receptor. Furthermore, there are different mechanisms of coupling of the main functions of the two-component pathway. The sensor and reaction regulator may be on a single polypeptide or they may be on separate proteins. In addition, the proteins involved may be membrane proteins.

In the two-component pathway, external signals are registered by a His-specific protein kinase. The His-specific protein kinase is activated as a reaction to reception of an external signal (e.g., binding of a ligand) and is followed by autophosphorylation of a conserved *His residue* using ATP as a phosphate donor. The presence of a His kinase domain in a protein of the two-component pathway can often be diagnosed on the basis of the primary sequence, due to the comprehensive sequence information available (more than 50 bacterial systems are known). The signal is passed on to the *regula-*

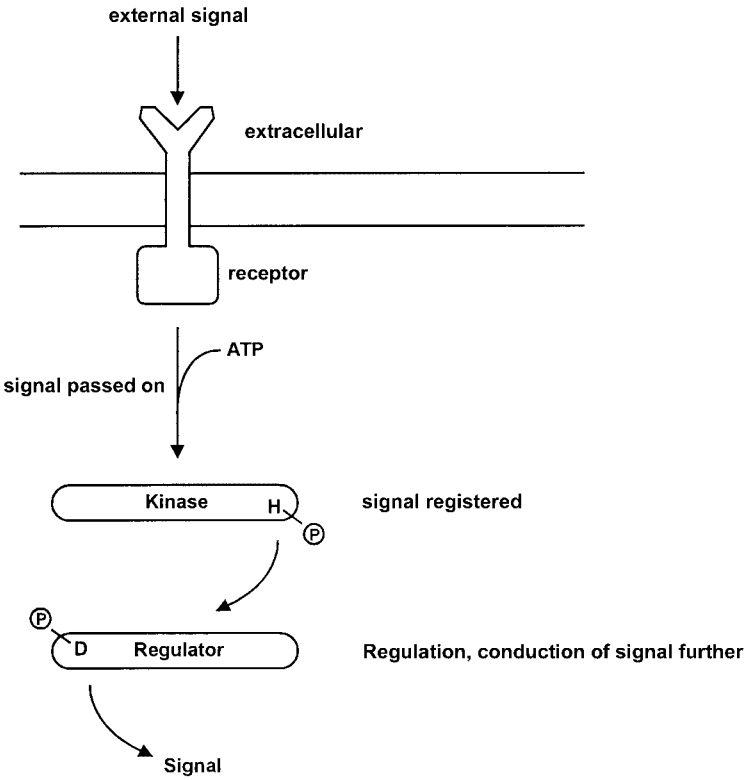



**Fig. 12.3.** Principle of the two-component pathway. The figure shows the principal steps of the two-component pathway in bacterial systems. An extracellular signal (change in osmolarity, N availability, etc.) is registered by a receptor. An interaction takes place with the first component, the „sensor kinase", which undergoes autophosphorylation at a His residue (H). The phosphate residue is transferred to the carboxyl side chain of an Asp residue (D) of the reaction regulator. Phosphorylation of the second component activates this for further signal conduction. The „sensor kinase" may also be localized in the cytoplasmic domain of the receptor.

tor from the phosphohistidine of the activated kinase, with the phosphate residue being transferred to a conserved *Asp residue* in the regulator. The Asp phosphorylation of the regulator controls its activity in signal transduction. Downstream proteins may be activated or the signal may be directly converted into regulation of transcription. A DNA binding domain is often found in the regulator so that DNA binding capacity of the regulator is controlled by Asp phosphorylation.

The two-component system is a signaling pathway of great importance in bacteria. Similar proteins and signaling pathways can also be identified in plant cells and in yeast, based on sequence homology with the bacterial proteins (review: Swanson et al., 1994). It is expected that signaling pathways using the principle of the two-component system will also be found in mammals.

Fig 12.4 shows a model for signal transduction in the two-component system of yeast. The SLN1 protein of yeast is a composite sensor regulator system in which the

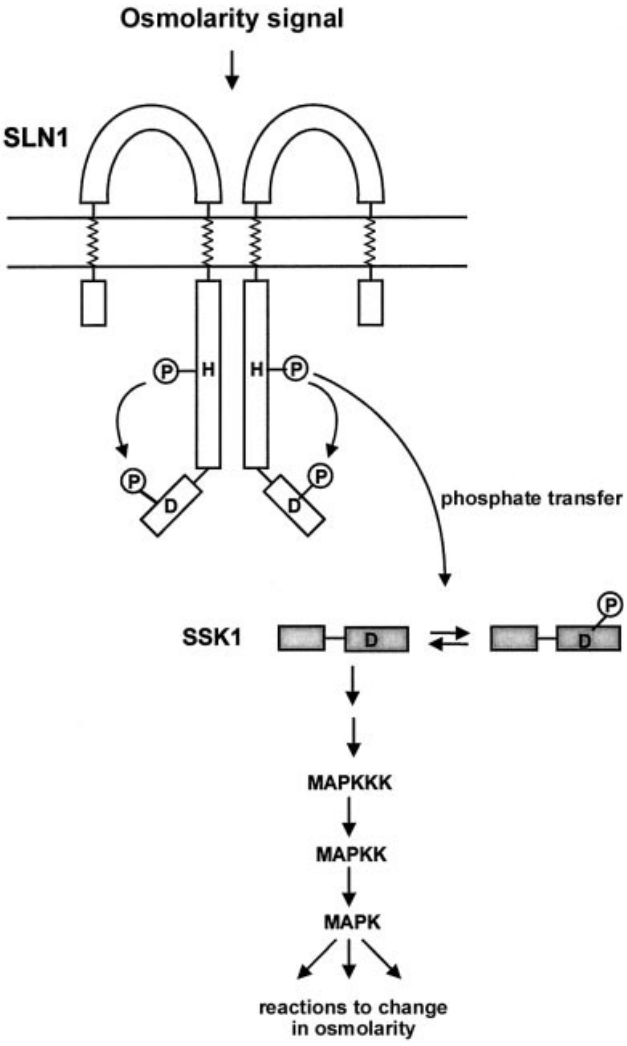



**Fig. 12.4.** Example of a two-component pathway in *S. cerevisiae*. Model of signal transduction via the SLN1 protein. The SLN1 protein is a transmembrane protein with two transmembrane elements, which is assumed to exist as a dimer. The sensor domain and the regulator domain are localized on the same protein chain in the SLN1 protein. The SLN1 protein is activated by an extracellular signal (e.g., decrease in osmolarity). Autophosphorylation takes place on His (H) in the sensor domain and on Asp (D) in the regulator domain. A phosphate transfer takes place from the phosphohistidine to the effector protein SSK1. In the unphosphorylated form, SSK1 activates a MAPK pathway, which contains the protein kinase HOG1 as a MAPK element. Various cellular reactions are triggered by HOG1. If SSK1 is phosphorylated in the course of activation of the two-component pathway, stimulation of the MAPK pathway is stopped. According to Swanson et al., (1994).

His kinase domain and the regulator domain are located on the same protein. Following reception of an extracellular signal (e.g., low osmolarity), autophosphorylation takes place on His and Asp residues of the SLN1 protein. In the process of signal transduction, a phosphate residue is transferred from a His phosphate of the SLN1 protein to an Asp residue on a downstream signal molecule SSK1. In the unphosphorylated form, SSK1 directs a signal to the MAP kinase pathway of yeast. Asp phosphorylation of SSK1 stops this signal conduction. The system is an example of a two-component system in which Asp phosphorylation takes place on both the response regulator and on the substrate protein.

A particular feature of the two-component system is the use of phosphohistidine and phosphoaspartate groupings (see Fig. 7.1) as elements of signal transduction. The

phosphonamide grouping in phosphohistidine is unstable in aqueous medium and only has a short half-life. The duration of the signal transduction is thus limited to a large degree by the chemical instability of phosphohistidine.

The switch for weakening and activation of the signal is already built in to the system. If there is no regulator protein available to receive the signal, the signal dies away due to the chemical instability of His-phosphate. This is similar for the Asp phosphate grouping of the regulator. Here, there are specific phosphatases that can cleave the phosphate off again but the intrinsic instability of the aspartyl phosphate grouping can be large and is determined by the microenvironment of the protein. Thus, Asp phosphate in the bacterial Che-Y protein only has a half-life of a few seconds. In contrast, the OmpR protein has an Asp phosphate with a half-life of ca. 1 h.

The high intrinsic instability of the groupings involved in this chemical signal transduction makes the two-component system particularly suitable for rapid, repeated signal transductions. It is a relatively simple system in which a signal may be rapidly conducted further and rapidly switched off.

# References Chapter 12

Artavanis-Tsakonas, S., Rand, M.D. and Lake, R.J. 'Notch signaling: cell fate control and signal integration in development' (1999) Science *284*, 770–776

Chan, Y. and Jan, Y.N. 'Roles for proteolysis and trafficking in Notch maturation and signal transduction' (1998) Cell *94*, 423–426

Heldin, C.H., Miyazono, K. and ten Dijke, P. 'TGF-beta signalling from cell membrane to nucleus through SMAD proteins' (1997) Nature *390*, 465–671

Massague, J. 'TGF-beta signal transduction' (1998) Ann. Rev. Biochem. *67*, 753–791

Shi, Y., Wang, Y.F., Jayaraman, L., Yang, H., Massague, J. and Pavletich, N.P. 'Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling' (1998) Cell *94*, 585–594

Swanson, R.V., Alex,L.A. und Simon, M.I. 'Histidine and aspartate phosphorylation: two-component systems and the limits of homology' (1994) Trends Bioch. Sci. *19*, 485–490

Ulloa, L., Doody, J. and Massague, J. 'Inhibition of transforming growth factor-beta/SMAD signalling by the interferon-gamma/STAT pathway' (1999) Nature *397*, 710–713

Wrana, J.L., Attisano, L., Wieser, R., Ventura, F. und Massague, J. 'Mechanism of activation of the TGF-ß receptor' (1994) Nature *370*, 341–347

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 13**

# Regulation of the Cell Cycle

## 13.1 Overview of the Cell Cycle

Eucaryotic cells execute their reproduction in a cyclic process, in which at least two phases, a *S phase* and a *M phase*, can be differentiated on the basis of biochemical and morphological features. The biochemical characteristic of the S (synthesis) phase is the replication of nuclear DNA and thus doubling of the genetic information. In M (mitosis) phase, division of the chromosomes between the daughter cells is prepared and carried out.

In most cell types, two further phases can be distinguished, $G_1$ and $G_2$ phase. $G_1$ phase includes the period between M phase and S phase; $G_2$ phase covers the period between S phase and M phase. From $G_1$ phase, the cell may transfer into a quiescent state known as $G_o$ phase. Appropriate signals (e.g., addition of growth factors) can induce the cell to return from $G_o$ into $G_1$ phase and proceed with the cell cycle.

The cyclical sequence of $G_1$, S, $G_2$ and M phases describes a standard cell cycle (Fig. 13.1). Rapidly dividing cells in mammals require 12–24 h for completion of a cell cycle.

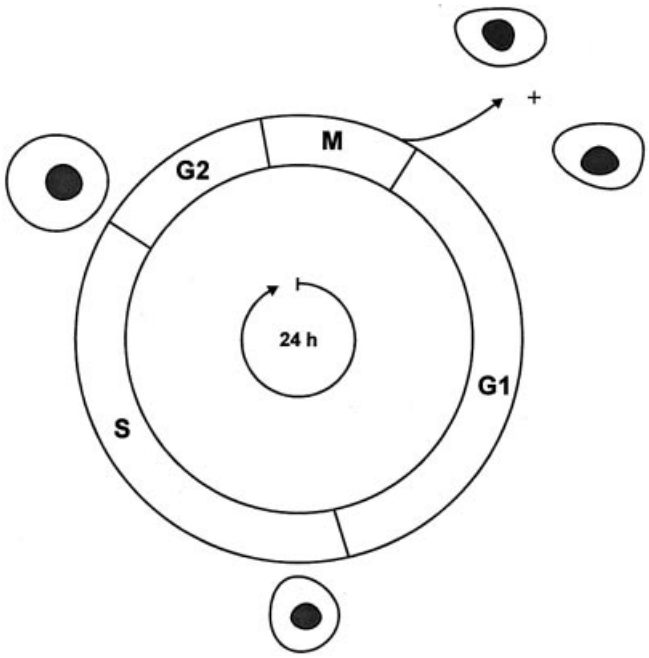



**Fig. 13.1.** The four phases of a typical cell cycle of a eucaryotic cell. $G_1$-, S- and $G_2$ phases form the interphase, whilst the cell grows continuously. Cell division occurs in M phase. New synthesis of DNA is limited to S phase. $G_1$ phase includes the period between M phase and S phase; $G_2$ phase lies between S phase and M phase. The duration shown of 24 h is only approximate. There are cells with cell cycles of shorter or longer duration.

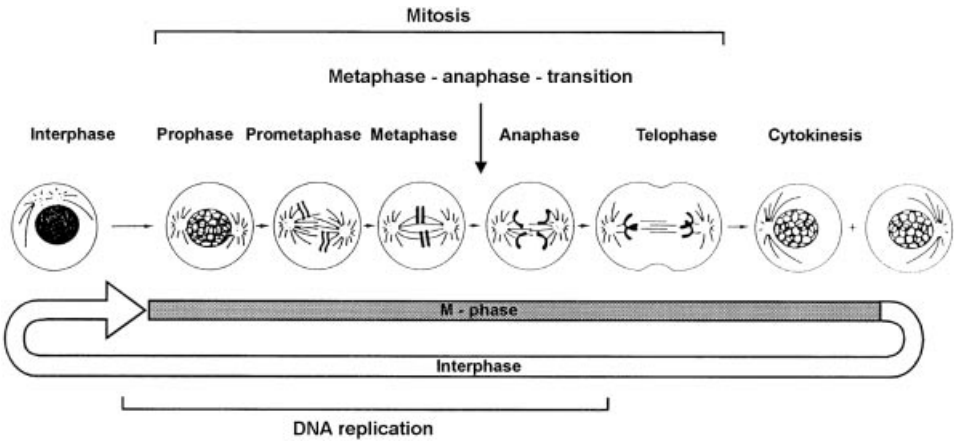



**Fig. 13.2.** Cytologic features of M phase. M phase is divided into the mitotic phases shown, based on characteristic cytologic features. The transition from metaphase to anaphase is an important control point. Cells may stop and pause before this control point. If the control point is crossed, M phase is concluded with cell division.

In some cell types, such as early embryonal cells, the period between the S and M phases is reduced to the extent that discrete $G_1$ and $G_2$ phases cannot be identified. The duration of the cell cycle is then only 8–60 min.

Morphologically, cell division is only visible in M phase. Under the light microscope, condensation, alignment and segregation of the chromosomes and cell division itself may be observed during M phase. In addition, different mitotic phases can be distinguished, as shown in Fig. 13.2.

### 13.1.1 Principles of Cell Cycle Control

The different phases of the cell cycle include a number of highly ordered processes that ultimately lead to duplication of the cell. The various cell cycle events are highly coordinated to occur in a defined order and with an exact timing, requiring precise control mechanisms.

The ordered sequence of cell cycle events is ensured by different control loops that have an inhibitory or promoting effect on the progress(ion?) of the cell cycle. These are monitoring mechanisms that register the completion of important cell cycle events (e.g., complete DNA synthesis) and allow the transition to the next event (e.g., entry into mitosis) to occur.

The control systems of the cell cycle ensure that the various phases are executed completely and in the correct sequence. Entry into a new phase can only take place when the preceding phase has been completed. In addition, the system allows coupling of processes that are not adjacent in the cell cycle sequence. Thus, there are close mutual regulation mechanisms between M and S phase. Examples of the various control mechanisms are shown in Fig. 13.3.

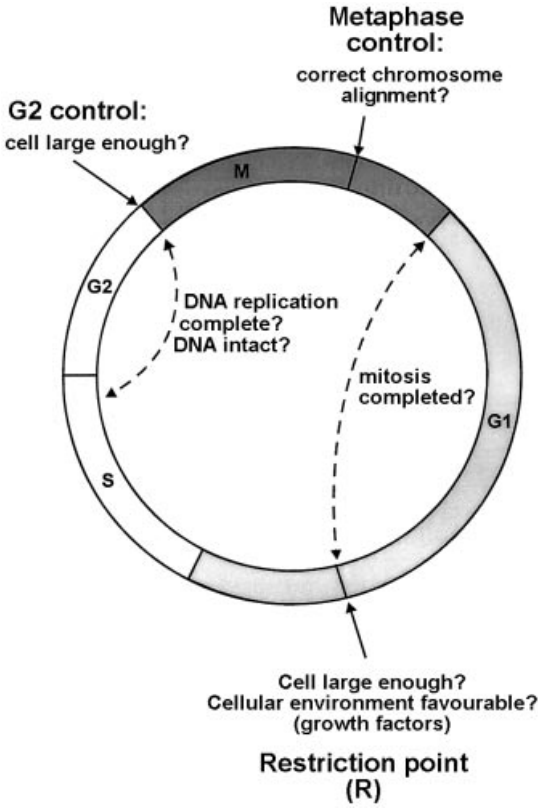



**Fig. 13.3.** Control points of the cell cycle: external and internal control mechanisms. Important control points of the cell cycle lie at the end of $G_2$ phase ($G_2$/M transition), in mitosis (metaphase/anaphase transition) and in $G_1$ phase (restriction point). The internal controls are shown as broken arrows and the external controls are shown as solid arrows.

Many of the control mechanisms of the cell cycle are of *intrinsic* nature and are constitutive, i.e., they are operational in every cell cycle and ensure the ordering of the individual steps. However, other control mechanisms exist that are not active in every cell cycle; these are only induced when defects are detected in central cell cycle events. These control mechanisms are known as *checkpoints*. An example of a checkpoint that is only activated when required is the DNA damage checkpoint (see 13.7). This is a biochemical pathway that detects DNA damage and creates a signal that arrests cells in the G1, S or G2 phase of the cell cycle.

In addition to the built-in protection and control mechanisms, the cell is also subject to a number of *external* controls, which ensure that cell division occurs in balance with the overall development of the organism and with external growth conditions. This is a kind of social control of cell division that regulates the progress of the cell cycle, with the help of circulating signal molecules or via cell-cell interactions.

A biochemical system is at the center of the cell cycle, of which the most important players are Ser/Thr-specific protein kinases and regulatory proteins associated with these. The activity of this central cell cycle apparatus regulates processes downstream that help to carry out the many phase-specific biochemical reactions of the cell cycle in a defined order.

Controlling influences on the progress of the cell cycle are effected from various levels that are linked to one another (Fig. 13.3).

## 13.1.2 Intrinsic Control Mechanisms

Intrinsic control mechanisms ensure that the cycle is executed completely, so that, following cell division, both daughter cells are equipped with the same genetic information as far as possible. Of the many control mechanisms, the following are highlighted:

- **Coupling of mitosis to a completed S phase**

Mitosis is only initiated when the DNA has been completely replicated during S phase. Mechanisms must exist that register completion of S phase and couple this to entry into M phase.

- **Coupling of S phase and mitosis**

Another control mechanism ensures that entry into S phase is only possible if preceded by mitosis. If the cell was able, during $G_2$ phase, to enter a new S phase without mitosis taking place, this would lead to unprogrammed multiplication of the chromosome set and thus to polyploidy. For S phase control, see 13.5.

- **Coupling of cell size and progress in $G_1$ phase**

A further control mechanism, which is also intrinsic, tests whether the cells in $G_1$ phase are large enough to initiate another round of cell division. The daughter cells produced by cell division must reach a critical size in the course of $G_1$ phase before S phase can commence.

- **DNA damage and the course of the cell cycle**

The cellular genome is continually subject to damaging influences that may originate within the cell or externally (see also 14.1.4). DNA damage must be repaired with the help of repair enzymes. Non-repaired DNA damage leads to miscoding in DNA replication and thus to mutations. For this reason, the cell has control mechanisms that register DNA damage and may possibly halt the cell cycle. Thus, time is gained for repair of DNA damage. However, the cell may also enter a resting state and possibly initiate programmed cell death.

## 13.1.3 External Control Mechanisms

**Growth Conditions**

Cell division activities are controlled to a high degree by externally determined growth conditions such as nutrient supply. A cell may stop cell division if the physiological conditions are unfavorable.

### Mitogenic Signals during Cell-cell Communication

Within the bounds of intercellular communication, mitogenic signals in the form of growth factors are produced in the organism. These bind to specific receptors on the target cell and initiate signal chains that influence the progress of the cell cycle (Fig. 13.4). These mitogenic signals control cell division activity and adjust it to the overall development of the organism.

### Antimitogenic Signals during Cell-cell Communication

In addition to growth-promoting signals, growth-inhibiting antimitogenic signals may also take effect in the organism. These lead to a halt in the cell cycle and may lead to transition of the cell into $G_o$ phase. Lack of mitogenic signals can have the same effect on the progress of the cell cycle.

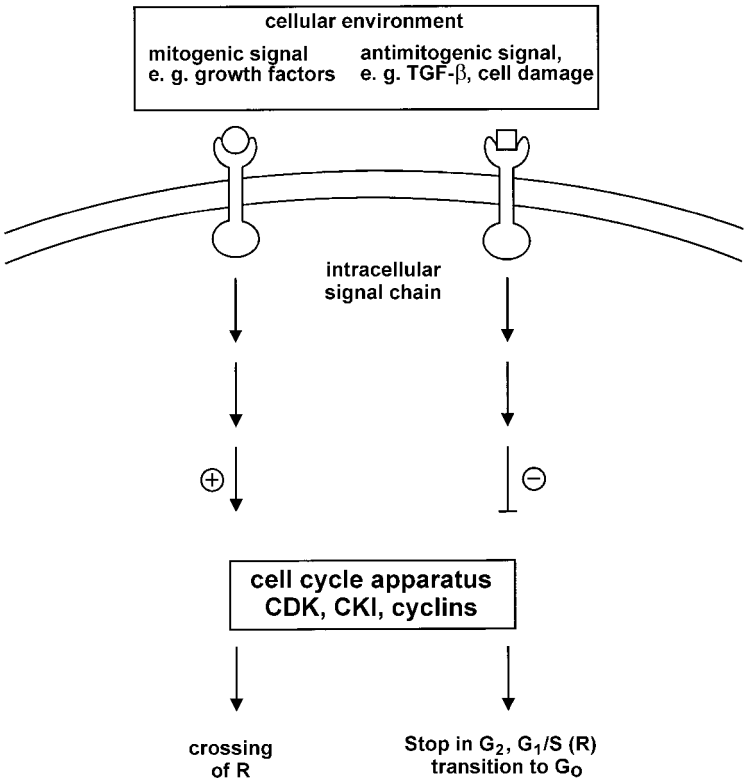



**Fig. 13.4.** Mitogenic and antimitogenic signals in control of the cell cycle. The cellular environment may emit mitogenic or antimitogenic signals. Mitogenic signals (e.g., growth factors) promote passage through the cell cycle; antimitogenic signals (e.g., TGFβ) lead to a halt in the cell cycle. In both cases, the extracellular signal is registered by transmembrane receptors and is passed on to the cell cycle apparatus via an intracellular signal chain. TGFβ: transforming growth factor β; CDK: cyclin-dependent protein kinase; CKI: inhibitor of CDK; R: restriction point.

### 13.1.4 Critical Cell Cycle Events and Cell Cycle Transitions

The cell cycle contains critical events where the cell switches from one state of biochemical activity to another state, in an irreversible manner. These events are called cell cycle transitions. Often, activating and inhibitory signals are received and transmitted at these cell cycle transitions (review: Nurse, 1994; Heichman, 1994; Elledge, 1996)(Fig. 13.3).

An example of an important cell cycle transition is the restriction point R, which occurs in late G1 phase. Crossing the restriction point is an important decision for further progression in the cell cycle and for entry into S phase. At this point, the cell switches from a growth-factor-dependent state to a growth-factor-independent state.

The restriction point is crossed, for example, when the cell is large enough for division and when enough external activating signals (growth factor signals) are present during G1 phase, so that the cell is sufficiently prepared for the events following in S phase. On crossing the restriction point, the cell continues in the cell cycle automatically and there is no need for further activating signals for entry into S phase and continued progress in the cell cycle. If the conditions for crossing the restriction point are not fulfilled, progress through the cell cycle slows down or the cell stops in the cycle until the requirements are fulfilled.

Other important cell cycle transitions are entry into S phase and the G2/M transition. At the G2/M transition, it is registered whether S phase has been completely executed, and the integrity of the DNA is examined at a DNA damage checkpoint. There are other important cell cycle transitions in M phase between metaphase and anaphase. At this point, an important and irreversible decision is made for progress of mitosis: if the spindle apparatus is correctly formed and the sister chromatids are correctly aligned, the cell cycle may proceed.

## 13.2 Key elements of the Cell Cycle Apparatus

Investigations of the cell cycle of yeast have helped to identify the key elements of cell cycle regulation and to understand the underlying principles. The results obtained are mostly applicable to higher eucaryotes; here too, a plausible picture has been obtained of how the cell cycle is maintained and regulated. However, many aspects of this picture are incomplete and we are a long way from being able to understand and biochemically describe the coordinated progress and regulatory aspects of the cell cycle.

Two processes are central to cell cycle regulation:

- Oscillating changes in the activity of the cell cycle machinery, with protein kinases as the most important component
- Specific proteolysis of cell cycle regulators (see 13.3).

The activity of the cell cycle machinery is controlled by the following proteins in particular:

– Cyclin-dependent protein kinases (CDKs)
– Cyclins
– Inhibitors of cyclin-dependent protein kinases (CKIs)

An oscillating system is formed by the interplay of the three protein classes and the activity of this system makes up the specific biochemical functions of the individual phases of the cycle. The activity of the cyclin-dependent protein kinases (CDKs) is central to the oscillating system. These create a signal that initiates downstream biochemical processes and thus determines the individual phases of the cycle. CDK activity is also the starting point for intrinsic and external control mechanisms.

## 13.2.1 Cyclin-dependent Protein Kinases, CDKs

The CDKs are proteins of 34–40 kDa with protein kinase activity. The CDKs must associate with the corresponding cyclin (see 13.2.3) to be active (review: Morgan, 1995; Pines, 1995).

Active cyclin-dependent protein kinases are thus heterodimers in which the CDK subunit carries the catalytic activity and the other subunit, the cyclin, performs an activating and specificity-determining function.

In the fission yeast *Schizosaccharomyces pombe*, the oscillator function is performed by only one CDK subunit, the CDC2-CDK (also known as $p^{34cdc2}$); in the budding yeast *Saccharomyces cerevisiae*, this is CDC28-CDK ($p^{34cdc28}$). In mammals, there are at least seven different catalytic subunits, known as CDC2 kinase (or CDK1) and CDK2 – CDK7. The best characterized are the CDC2-CDK and CDC28-CDK in yeast, and CDK2 in mammals.

The catalytic center of the CDKs includes a core of ca. 300 amino acids in which a high degree of homology between the different members of this protein kinase family may be detected. Within this region, there is a conserved sequence of 16 amino acids known as the *PSTAIRE region*. The PSTAIRE region is involved in binding of the corresponding cyclins and helps to determine specificity of cyclin-CDK binding (see 13.2.5).

## 13.2.2 Activation and Inactivation of CDKs by Phosphorylation

The CDKs may exist in inactive and active states. The transition between the two states is controlled in diverse ways (Fig. 13.5). Specific phosphorylation and dephosphorylation events are of central importance in control of CDK activity. The CDKs possess several phosphorylation sites for protein kinases and these may have an activating or inactivating effect. Phosphorylation at Thr160 of CDC2 kinase, or the equivalent positions Thr161 of CDK2 and Thr172 of CDK4, is *activating*. Phosphorylation at Thr14 and Tyr15 is *inhibiting*. A requirement for this regulation to take effect is the association of the CDK with the corresponding cyclin.

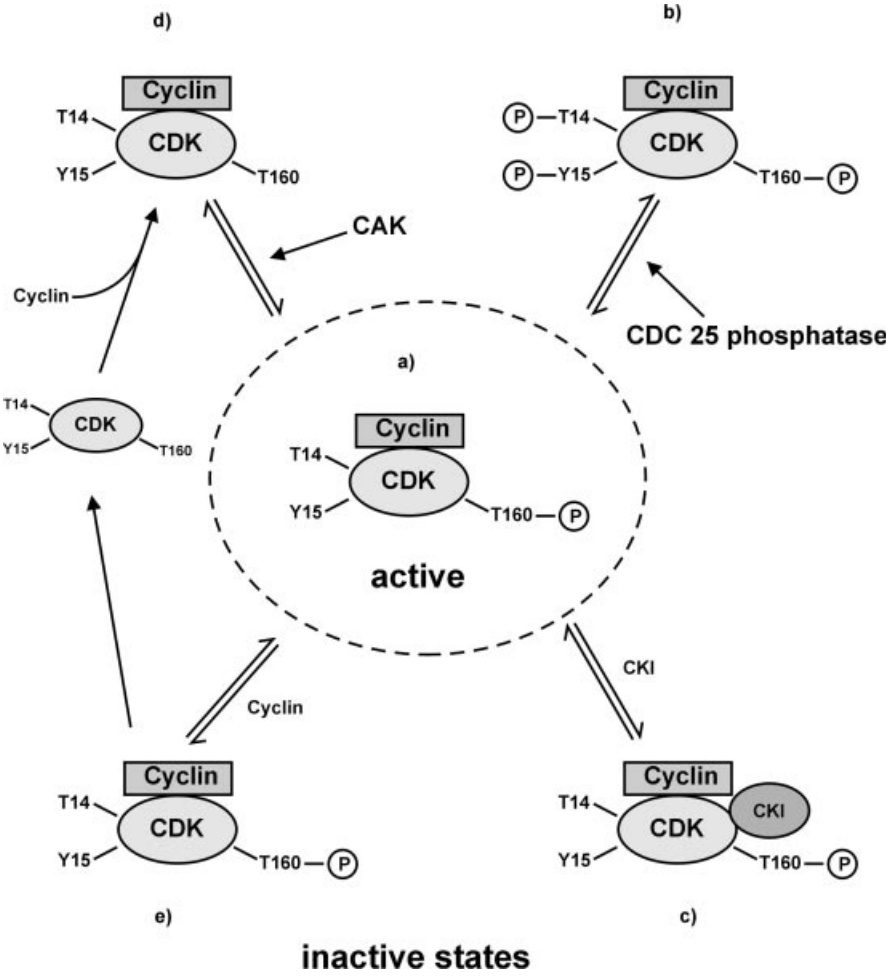



**Fig. 13.5.** Principles of regulation of cyclin-dependent protein kinases. The figure shows the principles of CDK regulation, using the CDC2 kinase (here simply referred to as CDK) as an example. The active form of CDK (a) is associated with the corresponding cyclin; Thr160 of CDK (or equivalent positions in other CDKs) is phosphorylated, and Thr14 and Tyr15 are unphosphorylated. Inactivation may take place by phosphorylation of Thr14 and Tyr15 (b) or by binding of a CKI (c). Other inactive forms of CDKs are the CDK-cyclin complex, in which Thr160 of the CDK is not phosphorylated (d). In addition, the cyclin-free forms of CDK are inactive (e). CDK: cyclin-dependent protein kinase; CKI: inhibitor of CDK; CAK: CDC2 activating kinase. According to Morgan, (1995).

### Phosphorylation at Thr160 (161, 172): Activation

Phosphorylation of the CycA-CDC2 complex at Thr160 leads to a near 300-fold increase in protein kinase activity. Thr160 of CDK2 lies in the activation segment (also known as the T loop) that blocks the access to the substrate binding site in the inactive

form (see Fig. 13.8; Fig 13.9). Phosphorylation of the T loop causes—in cooperation with cyclin binding—the T loop to be fixed in a position that enables optimal access of the substrate to the substrate binding site.

The protein kinase responsible for phosphorylation at Thr160 itself belongs to the family of CDKs and is known as *CAK* (*CDC2 activating kinase*) (review: Solomon, 1994). Its catalytic subunit is known as CDK7; it was previously also known as MO15. CDK7 is related to CDC2 kinase and, like the latter, requires the corresponding cyclin, *cyclin H*, for full activation. CAK demonstrates fairly constant activity during the cell cycle and therefore cannot be the regulating factor for phosphorylation of Thr160 of the CDK. It is assumed that the CAK-catalyzed phosphorylation at Thr160 requires binding of the cyclin to CDK and is regulated by the cyclin concentration.

**Phosphorylation at Thr14 and Tyr15: Inactivation**

Phosphorylation at Thr14 and Tyr15 leads to inactivation of the CDKs. In the fission yeast, the *wee1 kinase* is responsible for this phosphorylation and in mammals, there are enzymes homologous to wee-1 kinase. It is not clear whether this kinase performs both phosphorylations. Phosphorylation at Thr14 and Tyr15 is of particular importance for regulation of CDK activity in mitosis. The CDC2-cyclin B complex is maintained in an inactive state until the end of $G_2$ phase by the phosphorylation of Thr14 and Tyr15. At the $G_2$/M transition, the inactive state is ended by the action of CDC25 phosphatase, which cleaves off the inhibitory phosphate residues.

**Dephosphorylation at Thr14 and Tyr15: Activation**

The inactivating phosphorylation at Thr14 and Tyr15 can be reversed by specific phosphatases in a regulated manner. The dephosphorylation is performed by CDC25 phosphatase. This enzyme, first described for *S. pombe*, is a protein phosphatase with twofold specificity that can cleave phosphate residues from phosphoserine and phosphotyrosine residues of CDKs. CDC25 phosphatases have also been observed in higher eucaryotes where they have a similar function.

In *S. pombe*, activation of the CDC2 cyclin B complex takes place at the start of mitosis by a self-amplifying feedback mechanism. CDC25 phosphatase plays an important role in this mechanism (Fig. 13.6). The CDC25 phosphatase of *S. pombe* is subject to regulation by phosphorylation at Ser/Thr residues. The phosphorylations can have an activating or inactivating effect on the function of the phosphatase. Activation is observed upon phosphorylation at C-terminal Ser and Thr residues. It is assumed (King et al., 1994) that this phosphorylation is performed by the activated cyclinB-CDC2 complex. This results in a positive, self-amplifying feedback leading to rapid further activation of the cyclinB-CDC2 complex.

Furthermore, there is evidence that the protein phosphatases PP1 and PP2A dephosphorylate the activating Ser-phosphates of CDC25 phosphatase and thus have a negative, suppressing effect on the function of CDC25 phosphatase.

The activity of CDC25 phosphatase is also central to the DNA damage and DNA replication checkpoints of the cell cycle. Upon activation of these checkpoints by DNA damage or the presence of unreplicated DNA, the CDC25 phosphatase is inactivated

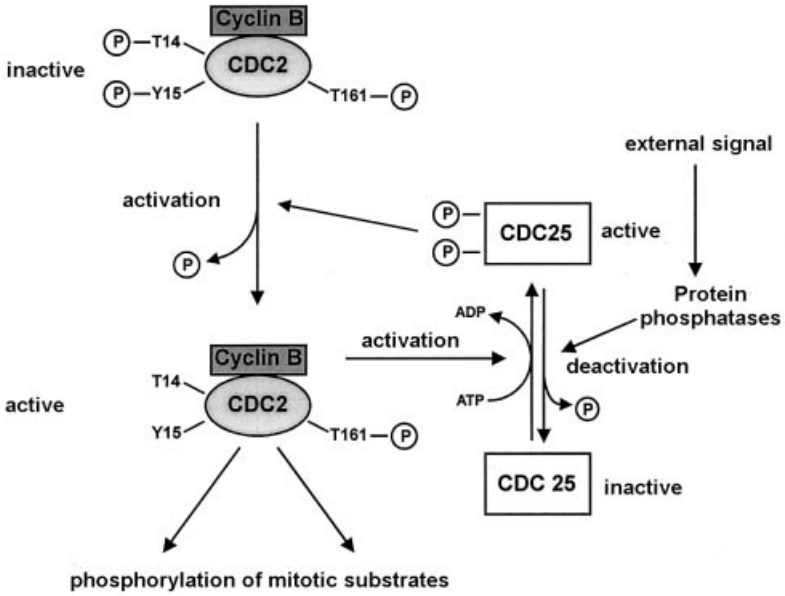



**Fig. 13.6.** Model of function and regulation of the CDC25 phosphatase in *S. pombe*. The CDC2-cyclin B complex is inactive when residues Thr14 and Tyr15 are phosphorylated. CDC25 phosphatase dephosphorylates both residues and activates the CDC2-cyclin B complex. The activity of CDC25 phosphatase is itself regulated by phosphorylation-dephosphorylation. The activated CDC2-cyclin B complex catalyzes – in addition to phosphorylation of mitotic substrate proteins – phosphorylation of the CDC25 phosphatase. The latter is activated by the phosphorylation. A cooperative amplification of CDC2-cyclin B and a rapid transition from $G_2$ to S-phase take place. Suppression of the activating function of CDC25 phosphatase may originate from protein phosphatases that can be activated by external signals and that dephosphorylate and inactivate CDC25 phosphatase.

and the cell cycle is stopped. Inactivation of CDC25 phosphatase is brought about by phosphorylation at specific Ser residues (Ser99, 192, and 359). These phosphoserine residues serve as attachment points for 14–3-3 proteins that sequester the CDC25 phosphatase in an inactive state (see 13.6).

### 13.2.3 Cyclins

The cyclins were originally defined as proteins that show cyclic concentration variations during the cell cycle (Fig. 13.7). A classifying feature of the cyclins today is the *cyclin box*, a conserved domain of ca. 100 amino acids. Binding to the corresponding CDK takes place via the cyclin box.

The cyclins are a protein family whose members become active in different phases of the cell cycle (review: Nurse, 1995; Morgan, 1995; Pines, 1995; Sherr, 1995). The cyclins are mostly localized in the nucleus. Exceptions are the cyclins B1 and B2 that accumu-

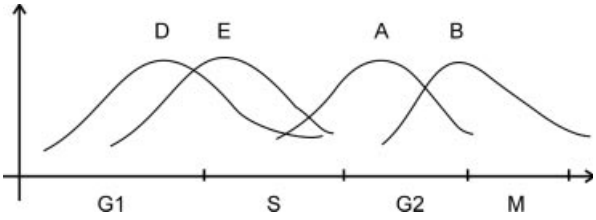



**Fig. 13.7.** Concentration changes in cyclins during the cell cycle. Cyclins of types D, A, E, and B show characteristic concentration changes in the course of the cell cycle. The figure only depicts the course of concentration changes in the cell cycle, and does not indicate the relative cyclin concentrations.

late in the cytosol during $G_2$ phase and are only translocated to the nucleus at the start of M phase.

The different cyclins preferentially form complexes with certain types of CDKs, which is why specific combinations of cyclins and CDKs are observed in the cell cycle (see Table 13.1).

The cyclins in mammals can be roughly divided according to their activity in the different phases of the cell cycle:

The *$G_1$/S cyclins* include the D and E type cyclins; the *M phase specific cyclins* include the B type cyclins. Cyclins of type A are active in S, $G_2$ and M phases.

The equivalent cyclins in yeast have a different classification and are shown in Table 13.1.

**Table 13.1** CDKs, cyclins and CKIs in mammals and in the yeast *S. cerevisiae*.

| Cyclin | CDK | CKI | Phase | |
|---|---|---|---|---|
| **Mammals:** | | | | |
| A1 | CDC2, CDK2 | p21 ? | Meiosis | |
| A2 | CDC2, CDK2 | p21 ? | S, G2, M | |
| B1, B2, B3 | CDC2 | p21? | M | |
| C | CDK8 | ? | ? | |
| D1, D2, D3 | CDK2,4,5,6 | p15, p16, p21, p27 | G1, restriction point | |
| E | CDK2 | p15, p21, p27 | G1/S | |
| F | ? | ? | G2 ? | ca. 80 kDa |
| H | CDK7 | | | |
| ***S.cerevisiae*:** | | | | |
| Cln1, Cln2, Cln3 | Cdc28 | ? | G1 | |
| Clb1, Clb2 | Cdc28 | ? | M | |
| Clb3, Clb4 | Cdc28 | ? | G2 | |
| Clb5, Clb6 | Cdc28 | Sic1 | S | |

The role of the cyclins is to convert CDKs into an active state. With the exception of CDK4 and CDK6, the CDKs exist as an inactive pool of constant concentration and are generally found in excess of the corresponding cyclin. The amount of activatable CDK is therefore limited by the cyclin concentration and for this reason, concentration changes in cyclins may be used as a regulatory element of the cell cycle.

## 13.2.4 Stability of Cyclins

In the cell cycle, the different cyclins show characteristic concentration changes in which temporally defined maxima in cyclin concentration are observed (see Fig. 13.7). The changes in cyclin concentration during the cell cycle are initiated by regulated new synthesis and targeted proteolysis.

The cyclins may be divided into two classes with respect to stability. There are cyclins that are *unstable during the whole cell cycle* and their concentration is determined in particular by the rate of transcription. In addition, there are cyclins that are only *unstable during certain phases of the cell cycle.*

### Degradation of $G_1$ Cyclins

Cyclins D and E in mammals are short-lived proteins with a lifetime of only ca. 20 min. Their instability is due to the occurrence of certain sequences in the C-terminal region; deletion of these sequences is associated with stabilization of the cyclins (see 13.3.1). These sequence elements are known as PEST sequences, based on their composition. Due to the short half-life, the D type cyclins require a constant stimulus at the transcription level to achieve the concentration in $G_1$ phase necessary for activation of CDK4/6 and to initiate crossing of the restriction point.

### Degradation of $G_2$/M Cyclins

The $G_2$/M specific cyclins A and B are stable during interphase and are specifically degraded during and after mitosis by proteolysis. This proteolysis is mediated by a degradation signal („*destruction box*“) located at the N terminus of the cyclin. The degradation signal (consensus sequence: R-ALGVI-N for A type cyclins, R-ALGN/D/EI-N for B type cyclins) is a sequence necessary to mark the cyclins for degradation by the anaphase-promoting complex APC (cyclosome) (see 13.3.2).

## 13.2.5 Structural Basis for CDK Activation

Binding of the corresponding cyclin and activating phosphorylation are required for full activation of the CDKs. Without the cyclin, the CDKs are inactive; the CDK-cyclin complex possesses a basal protein kinase activity that is considerably increased by phosphorylation at Thr in position 160 (or equivalent position). The structural changes

that are the basis of the different states of activity have been worked out for mammalian CDK2. Crystal structures are available for mammalian CDK2 in the inactive form without bound cyclin, in an active form with bound cyclin and in the phosphorylated CDK2-cyclin A complex (Jeffrey et al., 1995; Russo et al., 1996; review: Johnson et al., 1996).

### CDK2-cyclin A Complex

CDK2 has a folding pattern similar to that of the other structurally characterized Ser-specific protein kinases (Goldsmith and Cobb, 1994). Comparison of the inactive form of CDK2 (Fig. 13.8) with the cyclin A-bound form and with the active form of protein kinase A shows that the cause of the inactivity of CDK2 without bound cyclin is mainly due to two reasons: In the inactive form, the binding site for the protein substrate is blocked by a loop known as the T loop, which is equivalent to the activation segment described in 7.1.5. The Thr160 important for regulation is found in this T loop; phosphorylation of Thr160 is required for full activation of CDK2. Secondly, the residues of the active center involved in ATP cleavage are oriented in the inactive state so that cleavage of bound ATP is impossible. Release from the inactive state requires a change in the orientation of the T loop and a conformational change in the active center itself.

Binding of cyclin A has the following consequences for CDK2 structure:

- The active center is reorganized so that ATP may be bound in a similar manner to in the active form of protein kinase A.
- The T loop is partially removed from the catalytic cleft and the Thr160 important for regulation is now accessible for phosphorylation by CAK.

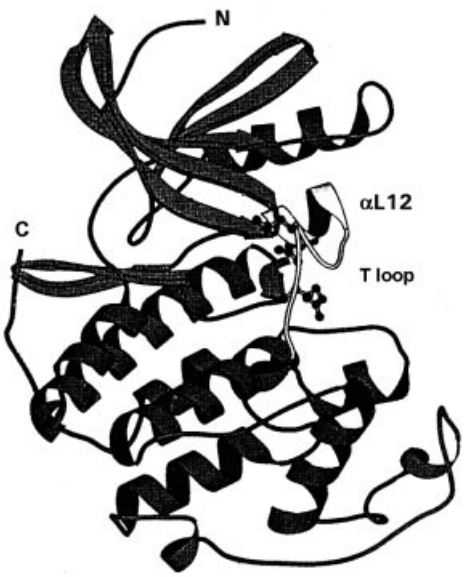



**Fig. 13.8.** Structure of the inactive form of CDK2 in humans. The crystal structure of the human CDK2 apoenzyme shows a very similar folding topology to protein kinase A and other protein kinases (see Chapter 7). The occurrence of the αL12 element is specific for CDK2; this interferes with binding of ATP and substrate in the active form.

- The cyclin box of cyclin A undergoes an interaction with the PSTAIRE helix of CDK2, which dissolves an inhibitory helix in CDK2. The conformational change leads to Glu51, which is involved in catalysis, being brought into the active center.

**Thr160 Phosphorylated CDK2-cyclin A Complex**

Phosphorylation of Thr160 is accompanied by a ca. 300-fold increase in protein kinase activity. The structure of the phosphorylated CDK2-cyclin A complex (Fig. 13.9A) differs considerably from the unphosphorylated form at certain critical points. These differences affect the C-terminal lobes of CDK2, the CDK-cyclin A interface and the T loop. The Thr160-phosphate serves as the organizing center that contacts different structural elements of the complex and structurally reorganizes them. Thanks to its polyvalent coordination sphere, the Thr160-phosphate couples parts of the T loop, the catalytic loop with essential Asp127, the PSTAIRE helix (Fig. 13.9B) and residues of cyclin A. Contacts are formed to Arg50 of the PSTAIRE helix, to Arg150 of the T loop and to Arg126; the latter lies in the vicinity of the essential Asp127. Furthermore, Van der Waals bonds are formed to an Ile270 of cyclin A. The conformational changes induced by these contacts affect the putative substrate binding site and the CDK2-cyclin A interface, in particular.

It is assumed that the activity increase is mainly due to better accessibility of the binding site for the protein substrate. In the unphosphorylated form, the T loop blocks access to the substrate binding site whereas in the phosphorylated form, this site is exposed.

## 13.2.6 Inhibitors of CDKs, the CKIs

Negative control of CDK activity in the cell cycle is performed by specific inhibitor proteins known as *cyclin-dependent kinase inhibitors, CKIs* (review: Morgan, 1995; Pines, 1995; Peter, 1997). These are a heterogenous family of proteins that may associate with a CDK or with a CDK-cyclin complex in a reversible manner, inhibiting CDK activity.

The CDKs may be divided into two groups based on sequence homology:

| | |
|---|---|
| CIP/KIP family | - $p21^{CIP1}$ (also known as CIP1, Waf1) |
| | –$p27^{KIP1}$ (KIP1) |
| | –$p57^{KIP2}$ |
| INK family | $p15^{ink4b}$ |
| | $p16^{ink4a}$ |
| | $p18^{ink4c}$ |
| | $p19^{ink4d}$ |

The members of the CIP/KIP family inhibit all G1 phase CDK-cyclin complexes (Fig. 13.10). The inhibitors of the INK family prefentially bind to CDK4 and CDK6 complexes.

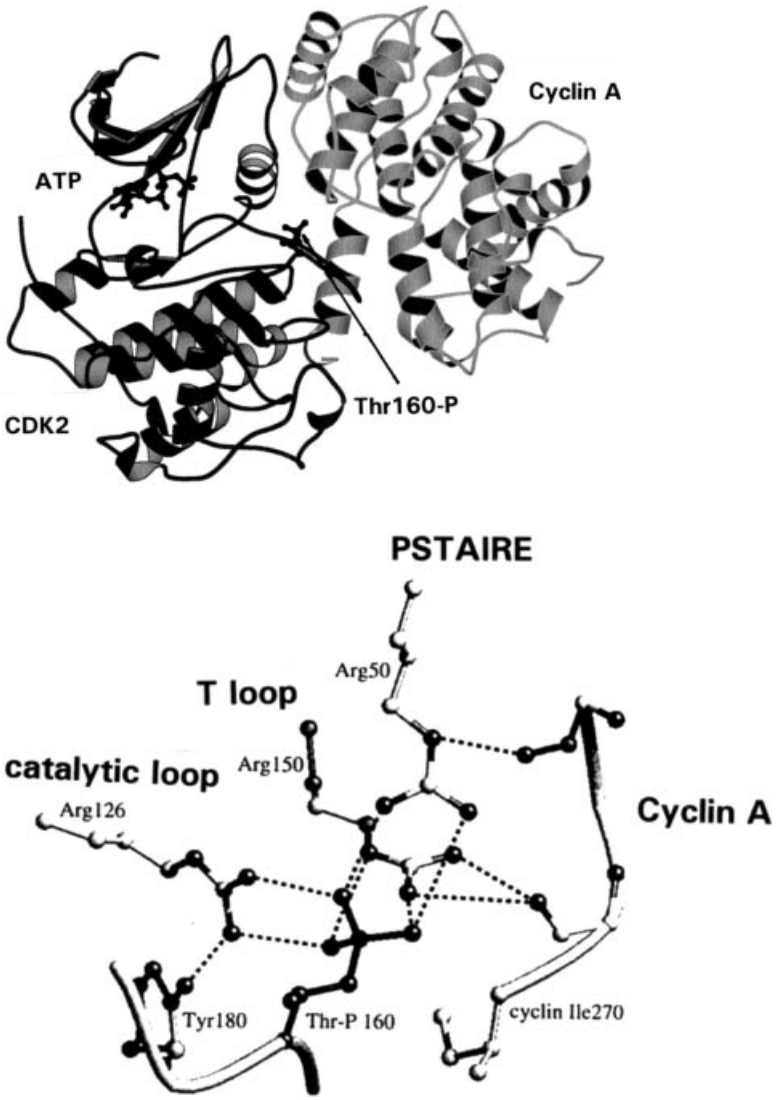



**Fig. 13.9.** Structure of the Thr160 phosphorylated CDK2-CyclinA complex.
A) Ribbon diagram of the complex showing the Thr160 phosphate (Thr-P 160) and bound ATP. Cyclin A is in light grey, CDK2 in dark grey.
B) Diagram of the multivalent interactions of Thr-P 160 at the CDK2-Cyclin A interface. Thr-P 160 forms contacts to Arg50 of the PSTAIRE helix, to Arg150 of the T-loop, to Arg126 which is close to the catalytic Asp127 and to Ile270 of Cyclin A (From Johnson and Reilly (1996), with permission).

### Mechanism of Inhibition

There are concrete data on the mechanism of inhibition for the inhibitor $p27^{KIP1}$, for which a crystal structure of a ternary complex, composed of cyclin A, CDK2 and $p27^{KIP1}$, is available (Russo et al., 1996). The crystal structure shows that the inhibitor interacts with both cyclin A and with CDK2. Inhibition of kinase activity is explained by alignment of structural elements of $p27^{KIP1}$ in the ATP binding site of CDK2. This breaks up the glycine-rich phosphate-binding loop. In addition, the ATP binding site is completely filled by residues of the inhibitor so that ATP binding is no longer possible. A similar functional principle is likely for the related inhibitor $p21^{CIP1}$.

A somewhat different mechanism forms the basis of inhibition by the $p16^{ink4a}$ protein. $p16^{ink4a}$ binds to CDK6 at a site opposite the cyclin binding site and blocks binding of cyclin D, and thus activation of CDK, by an indirect conformational change. Furthermore, the ATP binding site of CDK is deformed by the bound inhibitor (Russo et al, 1998).

### Regulation and Function of Cell Cycle Inhibitors

The CKIs are important entry points for signals that result in a halt in the cell cycle (see Fig. 13.10). The inhibitory effect of CKIs on progress of the cell cycle is highly dependent on the CKI/cyclin concentration ratio and thus the regulation of CKIs is of particular interest.

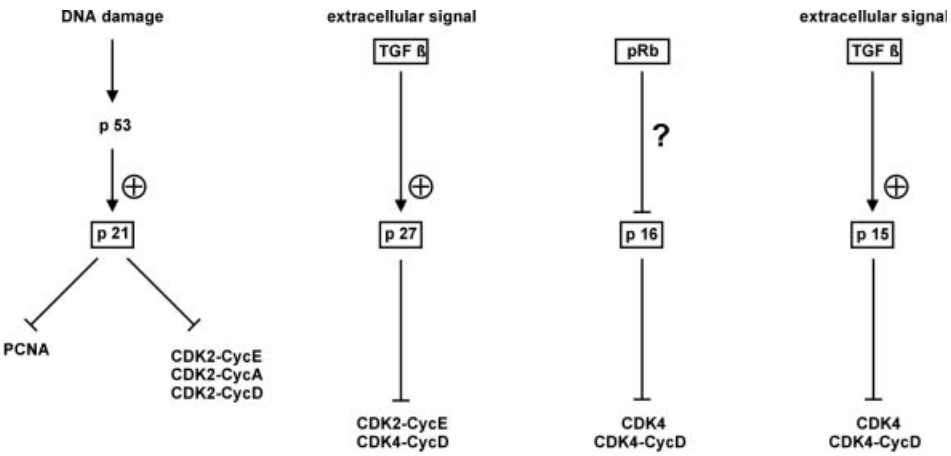



**Fig. 13.10.** Regulation and attack points of inhibitors of cyclin-dependent protein kinases (CKIs) in mammals. The various CKIs are activated and regulated by different signals. Inhibitory activity of the CKIs is specific for the different CDKs and CDK-cyclin combinations. p53: tumor suppressor protein p53; TGFβ: transforming growth factor β; pRb: tumor suppressor protein pRb; p21: $p21^{CIP1}$; p27: $p27^{KIP1}$; p16: $p16^{ink4A}$; CDK: cyclin-dependent protein kinase; Cyc: cyclin; E2F: transcription factor E2F; PCNA: proliferating cell nuclear antigen.

The inhibitor $p21^{CIP1}$ has a two-fold inhibitory function. On the one hand, $p21^{CIP1}$ binds to the complex of CDK2 with cyclin A, D1 and E and leads to their inactivation. On the other hand, $p21^{CIP1}$ has a binding site for the replication-accessory protein PCNA (Luo et al., 1995). PCNA (proliferating cell nuclear antigen) is required for nuclear DNA synthesis and functions in clamping DNA polymerase δ to the DNA thereby increasing the processivity of DNA synthesis. Binding of $p21^{CIP1}$ to PCNA inhibits replication by DNA polymerase δ in *in vitro* systems.

The concentration of $p21^{CIP1}$ is regulated at the transcription level in particular. The tumor suppressor protein p53 (see Chapter 15) is a main component of regulation of $p21^{CIP1}$. p53 protein induces $p21^{CIP1}$ transcription by binding to a DNA element in the $p21^{CIP1}$ promoter. Activation of p53 protein is observed especially on DNA damage. The primary aim of regulation is to halt the cell cycle if DNA damage occurs, to gain time for repair of the damage to the DNA or to induce the programmed cell death, apoptosis.

Regulation at the transcription level has also been described for the inhibitor $p15^{ink4b}$. The cytokine TGFβ is involved in this regulation (see Chapter 12). Binding of TGFβ to its receptor initiates a signal chain that culminates in activation of transcription of the gene for $p15^{ink4b}$ and may lead to a halt in the cell cycle.

The inhibitor $p16^{ink4a}$ is attributed properties as a tumor suppressor since the gene for $p16^{ink4a}$ is mutated in many tumor cell lines. There is evidence that the pRb protein (see 13.3.2) performs a regulating function on transcription of $p16^{ink4a}$.

The inhibitor $p27^{KIP1}$ is regulated at the post-translational level. $p27^{KIP1}$ exists in an inactive, masked form in proliferating cells. It may be converted into the active form by an as yet unknown mechanism so that the cell cycle can be halted. Activation of $p27^{KIP1}$ may be triggered by treatment of cells with TGFβ, by cell-cell contact and by an increase in the cAMP concentration. Furthermore, $p27^{KIP1}$ is subject to specific, ubiquitin-mediated proteolysis (see below).

## 13.2.7 Substrates of CDKs

Due to the problems in identification of cellular substrates of protein kinases, as described in Chapter 7, it has been a difficult and lengthy process to determine the functionally relevant substrates. Fig. 13.11 gives an overview of the cell-cycle-specific activation of CDKs and some important substrates. Comparatively sparse information is available on the $G_1$ and S phase substrates of the CDKs. In contrast, many proteins have been described that undergo specific phosphorylation in $G_2$/M phase. The sequence (K/R)-S/T-P-X-K (X: any amino acid) has been identified as a consensus sequence for phosphorylation by CDKs.

### Substrates in $G_1$/S Phase

Important substrates in $G_1$/S phase are the transcription factor E2F, the tumor suppressor protein pRb and pRb-related proteins. E2F represents a family of heterodimeric transcription factors with activity strictly regulated in the cell cycle. The members of the E2F family regulate transcription of genes with products that are required for the

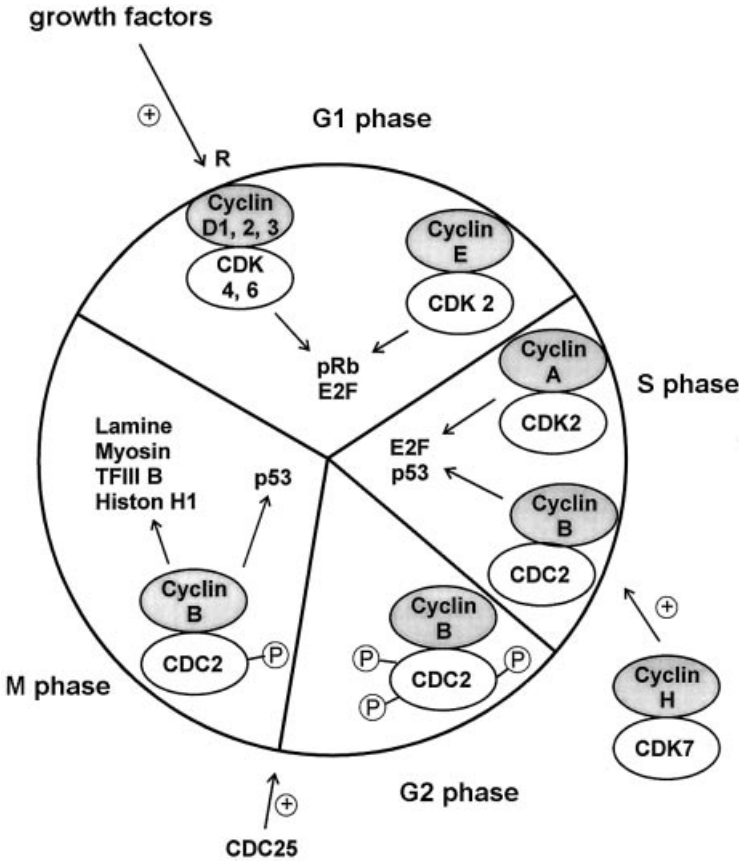



**Fig. 13.11.** Substrates and phase-specific activation of CDKs in the cell cycle. An overview is shown of the phase-specific activation of the most important CDK-cyclin complexes and of selected substrates. The arrows indicate activation and phosphorylation. CDK: cyclin-dependent protein kinase; p53: tumor suppressor p53; pRb: tumor suppressor pRb; CDC25: CDC25 phosphatase; TFIIIB: transcription factor TFIIIB; R: restriction point.

$G_1$/S transition and for progess through S phase. The activity of E2F may be regulated by CDK-mediated phosphorylation and by association with pRb. The tumor suppressor protein pRb is itself a substrate for phosphorylation by the CDK-cyclin D1 complex. Both proteins are described in detail in Chapter 14.

### Substrates in $G_2$/M Phase

In M phase, new phosphorylation of many proteins is observed that starts, in particular, from the CDC2-cyclin B complex. The phosphorylation mostly affects proteins involved in the reorganization of the cytoskeleton, the nuclear membrane and the formation of the spindle apparatus. As a consequence of phosphorylation events, inhibition of vesicular transport and general inhibition of transcription occur.

Examples of proteins that are specifically phosphorylated during the cell cycle are the lamins. Hyperphosphorylation of the lamins leads to disintegration of the nuclear lamina. Myosin in actin-myosin filaments is also specifically phosphorylated during mitosis. Other M-phase-specific phosphorylations occur at transcription factor TFIIIB, leading to inhibition of transcription by RNA polymerase III. Phosphorylation of TAF proteins (see 1.4.2.3) is also involved in general inhibition of transcription.

### 13.2.8 Multiple Regulation of CDKs

Activity of the regulatory components of the cell cycle varies extremely during the course of the cycle, and it is directed by external signals and internal control mechanisms. The CDKs are the central tool for control of the cell cycle. Their activity is regulated by various positively and negatively acting signals that are registered in the cell cycle and can bring about a halt in the cell cycle at various points. In summary, activity may be regulated by the following processes:

The following have a *positive* effect on activity of CDKs:

- Increase in cyclin concentration: by activation of transcription or inhibition of proteolytic degradation
- Phosphorylation of CDKs at Thr160 or equivalent positions
- Dephosphorylation of CDKs at Thr14/Tyr15
- Increase in concentration of CDK4 and CDK6
- Decrease in concentration of CKIs, at the transcription level or by proteolysis

The following have a *negative* effect:

- Decrease in concentration of cyclins: by reduced transcription or by activation of proteolysis
- Phosphorylation at Thr14/Tyr15
- Increase in concentration of CKIs

These regulation mechanisms cannot be considered in isolation. Rather, it must be assumed that the individual mechanisms cooperate, and demonstrate mutual regulation and that feedback mechanisms are built in. All control elements can be activated, in principle, by external signals, resulting in a complex network of cell cycle control with many entry and exit points. The following sections are thus highly incomplete and only describe the elements that have been experimentally well proven.

## 13.3 Regulation of the Cell Cycle by Proteolysis

The ordered course of the cell cycle is ensured by two processes in particular:

- Regulated and temporally coordinated activity changes in CDKs
- Targeted ubiquitin-dependent proteolysis

Both processes, the protein kinase regulatory network of CDKs and targeted proteolysis, are linked to one another and work in mutual dependence. For a review on the control of the cell cycle by targeted proteolysis see King et al., (1996) and Hersko (1997).

Selection of a protein for Ub-dependent proteolysis occurs particularly via the E3 enzymes of the ubiquitin pathway (see 2.7.2) that catalyzes ligation of the target protein with ubiquitin. The specificity of Ub-protein ligation is determined by the nature of the E3 enzymes, which may be composed of several subunits. The composition of this complex is very variable and can be changed according to requirements.

Use of specific proteolysis as a tool for control of the cell cycle has various advantages:

- Proteolysis allows simultaneous and complete inactivation of all functions of a multifunctional cell cycle protein such as the cyclins.
- Proteolysis enables subunit-selective reorganization of hetero-oligomeric protein complexes. An example is the targeted degradation of a CDK inhibitor.
- The total substrate pool of regulatory enzymes of the cell cycle may be inactivated by proteolysis.
- The regulatory system of a cell cycle section can be reset to the ground state by proteolysis.

Two types of E2/E3 complexes are of particular importance for cell cycle control. One complex, the SCF complex, is of outstanding importance for the G1/S transition. The other, the anaphase-promoting complex or the cyclosome, is especially important for the course and control of mitosis. Common to both complexes is the variable collaboration with different proteins to mediate the ubiquitinylation of different substrates.

### 13.3.1 Targeted Proteolysis at G1/S

The transition from G1 into S phase is associated with targeted proteolysis of important regulatory cell cycle proteins such as the G1 or D type cyclins, and degradation of specific CKIs. The ubiquitin-protein ligases involved act on their specific substrates only after phosphorylation of the substrate protein. Thus, the regulator for this degradation pathway is the phosphorylation of the substrate by a regulatory protein kinase and the ubiquitin ligase system may be constitutively active.

The ligases involved in this type of ubiquitin ligation are organized in multiprotein complexes called SCF complexes (Skp1, cullin, F box protein). In *S. cerevisiae*, the complex is composed of the proteins Cdc24 (an E2 enzyme), Cdc53 and Skp1, which form the core of the SCF. This core associates in a variable manner with a further type of protein, which function as specific substrate recognition factors. A common feature of these proteins is a sequence element known as the F box (review: Peters, 1998). Consequently, multiple forms of SCF complexes exist, due to this variable association.

Substrate proteins are selected for Ub ligation based on a C-terminal target sequence. These sequences which, due to the occurrence of common amino acids, are known as PEST sequences, are targets for phosphorylation. In the phosphorylated form, they are recognized by the ubiquitin ligase complex and marked for degradation.

Examples of Ub-mediated degradation of cell cycle control proteins are:

| | |
|---|---|
| Yeast: | G1 cyclins: Cln2, Cln3 |
| | Sic1, an inhibitor of B type cyclin/CDK1 complexes |
| Mammals: | the inhibitor $p27^{KIP1}$ |
| | cyclin E: degradation requires phosphorylation at Thr 380 |
| | cyclin D1 |

For the G1 cyclins in yeast, it is assumed that they are phosphorylated in an autocatalytic process by the activated G1-cyclin-CDK complex and are thus marked for degradation. According to this mechanism, the G1 cyclins are subject to continual degradation, which would explain their short half-life.

Phosphorylation of the inhibitor $p27^{KIP1}$apparently has a twofold function during Ub-mediated degradation (Tomoda et al., 1999). To be broken down, $p27^{KIP1}$ must be transported out of the nucleus which requires its phosphorylation. Furthermore, phosphorylation is needed for recognition by the ubiquitin-conjugating system. Export from the nucleus and proteasome-mediated degradation are both controlled by phosphorylation in this case.

### 13.3.2 Proteolysis during Mitosis: the Anaphase-promoting Complex/Cyclosome

During mitosis, a specific ubiquitin-ligase complex is activated that initiates proteolysis of various mitosis regulators. This complex, also known as the *anaphase-promoting complex* (*APC*) or the *cyclosome*, represents a second type of E2/E3 complex that is active in the cell cycle. The activity of the APC shows considerable changes during the cell cycle, with higher activity at the anaphase/metaphase transition until the end of G1 phase, and a decrease in activity in S and G2 phase. These changes in activity are due to regulatory modification of the APC. The APC differs significantly in this respect from the constitutively active SCF complexes, which degrade G1 cyclins and require phosphorylation of substrate proteins. The APC is a large protein complex composed of at least 8 different proteins. The substrate-specific, activating subunits, which – as for SCF – associate in a variable and regulated manner with APC, are important for regulation of APC activity (review: Peters, 1998). In *S. cerevisiae*, these are the proteins Cdc20 and Hct1 (also called Cdh1*). There are multiple forms of APC that function in a cell-cycle-regulated manner.* For Hct1, it has been shown that association with APC is inhibited by CDK-mediated phosphorylation (Zachariae et al., 1998). It appears that cell-cycle-specific phosphorylation of components of the APC is a major controlling factor of its ubiquitinylation activity.

The APC possesses a ubiquitin-ligase activity with specificity for substrate proteins that contain a particular sequence, the destruction box (see 13.2.4). During and after mitosis, APC-mediated specific proteolysis is observed of the mitotic cyclins B and A and of a number of other regulatory proteins. All these proteins contain one or two copies of the destruction box.

Proteolysis of mitotic cyclins is activated at the metaphase/anaphase transition and is only switched off at the start of S phase. The mitotic cyclin A is degraded before cyclin B. Lasting activity of the cyclosome during G1 phase is thought to be responsible for the lack of detection of mitotic cyclins in G1 phase. Renewed accumulation of mitotic cyclins is only possible again when APC activity is switched off at the start of S phase.

In addition to the mitotic cyclins, other different regulator proteins are marked by the APC for degradation. These include proteins such as the Cut2 protein, which is thought to function as an anaphase inhibitor.

## 13.4 The $G_1$/S phase Transition

$G_1$ phase has a special regulatory function in the cell cycle; here, the decision is made to enter S phase and thus a new round of cell division, or to enter a resting, quiescent state.

When mitosis has been completed, the cell requires signals in the form of growth factors to direct towards a new round of division. The signals become effective in the first two-thirds of $G_1$ phase. In this time window, the cell is programmed to begin a new cell cycle or to enter $G_o$ phase. After a particular point, the restriction point R, no further signals are needed to continue the cell cycle. The cell cycle apparatus is self-contained from this point onwards. S, $G_2$ and M phase occur without external control. The cell cycle may still be halted after crossing the restriction point, however, if the cell detects, via internal control mechanisms or checkpoints, that defects have occurred in the correct course of the phases.

### 13.4.1 Function of the D Type Cyclins

Progress of the cycle in $G_1$ phase is controlled in particular by the *cyclins of type D and E* and by *inhibitors of the CDKs* (review: Morgan, 1995; Pines, 1995; Sherr, 1995).

#### Cyclin D

Cyclins of type D are of particular importance in control of $G_1$ phase. Of the three cyclins of type D (D1, D2, D3), two (D2 and D3) do not occur in all cell types, whilst cyclin D1 has a central function in regulation of $G_1$ phase in all cell types.

A basic function of the *D type cyclins* is to integrate external signals into the cell cycle. Mitogenic signals, such as growth factors, activate the transcription of the gene for cyclin D1. The activating signals are active particularly during the first two-thirds of $G_1$ phase. In addition, mitogenic signals promote the association of D type cyclins with the corresponding CDK.

*Binding partners* of the D type cyclins include CDK4 and CDK6 in particular, and CDK4 activation is attributed a key role. For full activation of CDK4, activating phos-

phorylation at Thr172 of CDK4 is also necessary, in addition to binding of cyclin D1. This step is catalyzed by CAK. The activated CDK4/6-cyclin D1 complexes are now able to perform phosphorylation of the critical substrate, the pRb protein (Fig. 13.12). The CDC2-cyclin E complex is also thought to be involved in this phosphorylation. The activity of both complexes prepares the cell to run through the restriction point.

### Cyclin E

*Cyclin E* also performs its function in $G_1$/S phase. It demonstrates a periodic concentration change with a maximal value at the start of S phase. Afterwards, its concentration falls off sharply within S phase. The gene for cyclin E is also induced by transcription factor E2F which explains the increase in cyclin E at the $G_1$/S transition. Cyclin E binds and activates CDK2. The activated CDK2 complex is also involved in phosphorylation of the pRb protein. As a consequence, a signal is transmitted, with cooperation of cyclin D, in the direction of the transcription of genes that are essential for the continuation of the cell cycle.

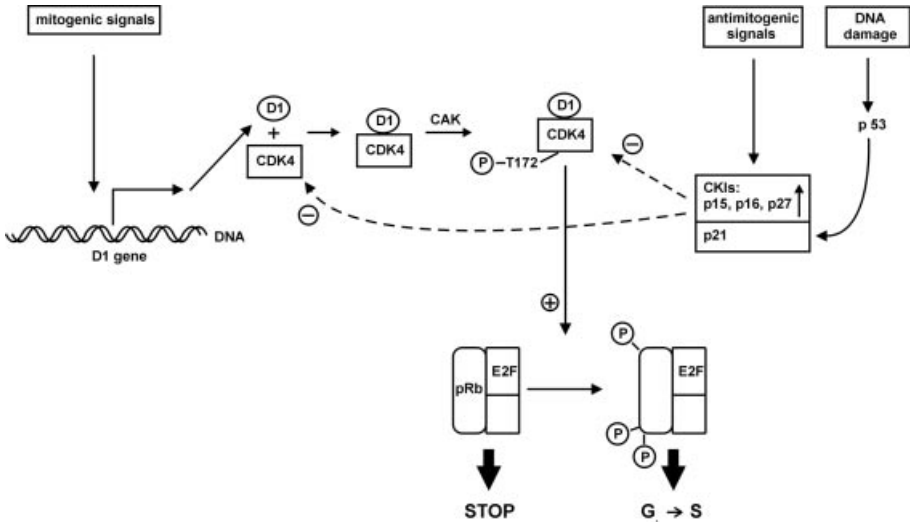



**Fig. 13.12.** Model of control of the $G_1$/S transition via D type cyclins. External signals may induce the cell to halt at the $G_1$/S transition or bring about entry into the S phase. D type cyclins (represented by cyclin D1 here) are involved in both processes. Mitogenic signals such as growth factors stimulate expression of cyclin D1, which associates with CDK4 and is activated with the help of CAK. The activated CDK4-cyclin D1 complex phosphorylates the pRb protein in the pRB-E2F complex and thus prepares the way for entry into the S phase. Antimitogenic signals or DNA damage activate the CKIs, which associate with the corresponding CDK-cyclin complex and inhibit its activity. Under these conditions, the pRb-E2F complex remains in an unphosphorylated state and passage of the cell through the cell cycle is halted at the $G_1$/S transition. CDK: cyclin-dependent protein kinase; E2F: transcription factor E2F; CAK: CDK activating protein kinase; CKI: inhibitor of cyclin-dependent protein kinase; p21: $p21^{CIP1}$; p27: $p27^{KIP1}$; p15; $p15^{ink4B}$; p16: $p16^{ink4A}$.

**Negative Regulation of the G1/S Transition**

In addition to mitogenic signals, antimitogenic signals are also processed during $G_1$ phase. These can lead to a halt in the cell cycle during $G_1$ phase and bring the cell into a resting state. An antimitogenic signal originates, for example, from TGFβ, from cAMP and from certain cell-cell contacts.

Negative regulation of the cell cycle in $G_1$ phase is performed in particular by the inhibitors $p21^{CIP1}$, $p27^{KIP1}$ and $p15^{ink4}$, which are activated by external signals (see above, Fig. 13.10). The inhibitor $p16^{ink4}$ also regulates by binding free D type cyclin, which leads to destabilization of cyclin D.

The balance between activated CDK4-cyclin D and the various inhibitors controls progress through $G_1$ phase. The concentration ratio of the CKIs to the $G_1$-specific CDK-cyclin complexes is attributed an important role in this process. If the CDK4-cyclin D concentration exceeds the concentration of the inhibitors, this presents a positive signal for progress of the cell cycle. If the concentration of one of the inhibitors is increased by an antimitogenic signal, so that the concentration of the $G_1$-specific CDK-cyclin complexes is exceeded, the cell stops in $G_1$ phase. Further reaction and coordination of the positive and negative signals is mediated by a regulation system in which the product of the retinoblastoma gene, the pRb protein, plays a central role (review: Planas-Silva and Weinberg, 1997, Adams and Kaelin, 1998)

## 13.4.2 Function of pRb in the Cell Cycle

The pRb protein is a nuclear phosphoprotein of ca. 100 kDa. The domain structure of pRb is shown in Fig. 13.13 (review; Riley et al., 1994; Weinberg, 1995). The pRb proteins possesses numerous Ser/Thr phosphorylation sites, binding sites for the transcription factor E2F, for the viral oncoproteins TAg, E1A and E7 (see Chapter 14) and a non-specific DNA binding site. Furthermore, an N-terminal sequence section has been identified in pRb protein that is required for oligomerization.

The pRb protein has the characteristics of a *tumor suppressor protein*. Loss of its function is associated with deregulation of cell division and favors tumor formation (see Chapter 14).

A central function of pRb in control of the cell cycle (review: Sherr, 1996) is illustrated by the following:

- At the end of mitosis until the restriction point R, pRb exists in an *underphosphorylated* form. In the underphosphorylated form, pRb has a growth-inhibiting function in that it blocks the activity of activity of transcription factors that control expression of S phase genes.
- During or after crossing the restriction point, pRb exists in a *hyperphosphorylated* form and it remains in this form until the end of mitosis. In the hyperphosphorylated form, pRb has a growth-promoting function.

The restriction point marks the transition of the cell from a growth-factor-dependent state to a mostly growth-factor-independent state. The fact that changes in the degree

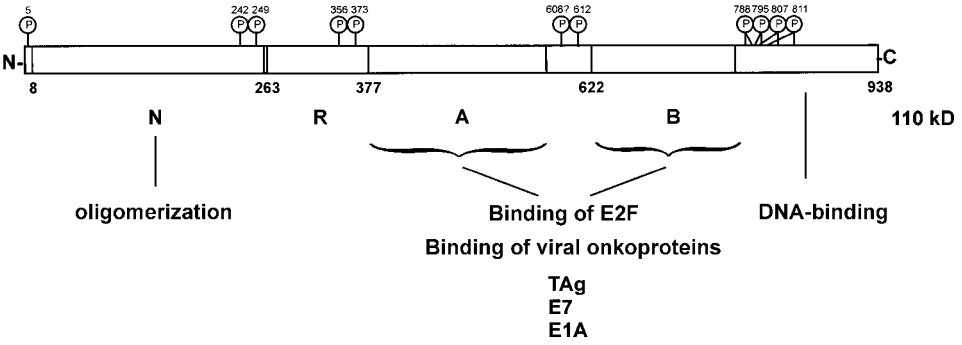



**Fig. 13.13.** Domain structure of the retinoblastoma protein pRb. The phosphorylation sites (P) of pRb and the localization of the sequence sections necessary for interaction with viral oncoproteins and with the transcription factor E2F are shown. In addition, an oligomerization domain and a DNA binding domain can be identified.

of phosphorylation of pRb coincide with the time window of $G_1$ phase, in which external mitogenic or antimitogenic signals become effective, suggests a central role for pRb in processing these signals. In general, the pRb protein can be assigned the function of a guardian for exit from $G_1$ phase. At the pRb protein, signals meet with which it is tested whether the requirements to cross the restriction point are fulfilled. In this guardian function, the pRb protein integrates growth-promoting and growth-inhibiting signals.

## 13.4.3 Model of pRb Function

The experimental data on the function of pRb allow a model to be described that unites the signal elements involved in $G_1$ control in a very logical manner (according to Weinberg, 1995).

### Phosphorylation of pRb

The crucial control element of pRb function is its phosphorylation status, which can be considered as a switch. At the start of $G_1$ phase, pRb exists in an underphosphorylated form. In this form, it functions as a brake on the progress of the cell cycle. The brake is lifted when pRb is phosphorylated by the central components of the cell cycle apparatus. The protein complexes involved in phosphorylation of pRb are CDK4/cyclinD, CDK2/cyclin E and CDK2/cyclin A.

Overall, pRb is phosphorylated by the activated CDKs at many (ca. 11) Ser/Thr residues. The different CDKs preferentially phosphorylate pRb at distinct sites. There is evidence that the various phosphorylation events have different effects on pRb function.

**Effector Function of pRb: Control of E2F**

The effector function of pRb is its control of the function of transcription factors that belong to the E2F family. pRb binds to proteins of the E2F family and thus controls their transcription-activating function.

The E2F proteins are DNA binding proteins, which, in addition to the DNA binding domain, also have a binding site for pRb protein and a transactivating domain (review: Bernards, 1997). In some cell types, the E2F proteins exist as heterodimers in complex with another DNA binding protein, known as DP-1. At least five different E2F proteins are known (E2F-1 – E2F-5), of which three (E2F-1, E2F-2, E2F-3) are under direct control of pRb. The consensus sequence of DNA elements assigned to E2F is TTTCGCGC. This binding element is found in promoter regions of genes that perform an essential function in S phase.

For example, E2F binding sites are found in genes for:

- Thymidine kinase
- Dihydrofolate reductase
- DNA polymerase α
- Cyclin A, Cyclin E
- Transcription factor c-myc
- E2F-1
- pRb

Transcription factor E2F thus controls the expression of proteins that are required for further progress of the cell cycle. Overall, the transcription-controlling activity of E2F can be assigned a central function in progress of $G_1$ phase and also S phase.

A scheme of the control function of pRb and E2F is shown in Fig. 13.14. pRb controls the function of E2F by entering into a complex with the latter. In the underphosphorylated form, the pRb-E2F complex actively represses transcription of DNA. The hypophosphorylated pRb protein represents the active form of pRb since this form mediates repression of the E2F controlled genes. If pRb exists in the hyperphosphorylated form, transcription repression is removed and E2F can stimulate transcription of the target genes.

The mechanism of transcription repression is not completely understood. Initial information was obtained by the observation that pRb interacts in the underphosphorylated form with a histone deacetylase (see 1.4.6). Possibly, the unphosphorylated pRb-E2F complex recruits the histone deacetylase to the chromatin and thus actively initiates a chromatin reorganization that represses transcription (Brehm et al., 1998; Magnaghi-Jaulin et al., 1998).

In addition to negative regulation of the E2F function, the pRb protein also has a regulating influence on the activity of eucaryotic RNA polymerases (Cavanaugh et al., 1995; White et al., 1996). The pRb protein represses transcription activity of all three eucaryotic RNA polymerases and thus performs a general negative control of protein biosynthesis. It is assumed that the growth-inhibiting influence of the pRb protein is also based on this property.

These findings, and the regulating activity on E2F function performed in the promoter-bound state, indicate that pRb protein is able to interact with components of the

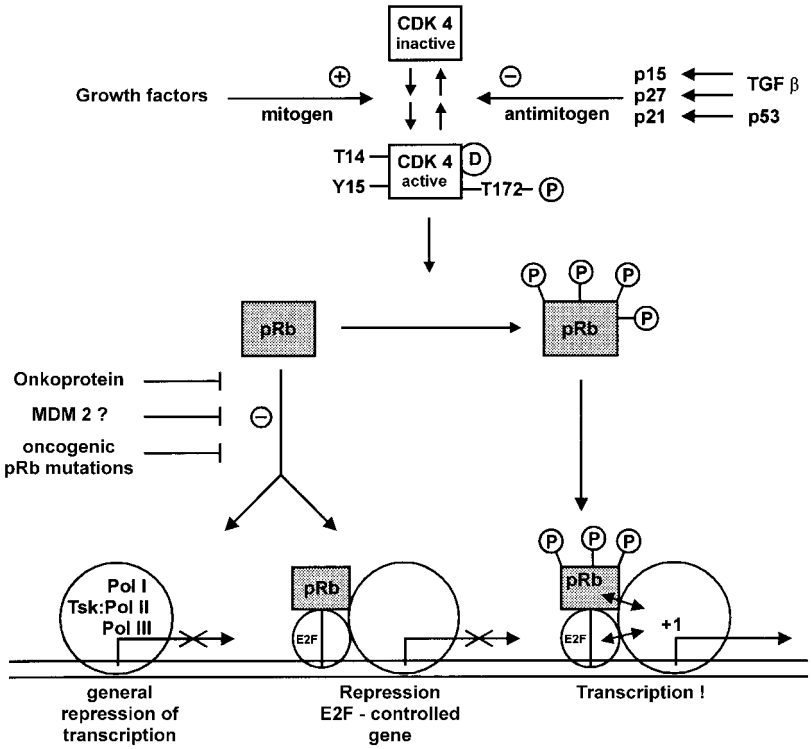



**Fig. 13.14.** Model of pRb-E2F function. The figure schematically summarizes the monitoring function of pRb for the $G_1$/S transition and the importance of pRb phosphorylation. Many aspects of the exact control of $G_1$/S transition are not understood. For this reason, the figure should be considered as highly simplified and it only shows some features of $G_1$/S regulation.

basal transcription apparatus in a positive or negative fashion, depending on its phosphorylation status.

There are two further proteins that are related to pRb which also perform a growth-regulating function but for which the exact function is unclear. These are the proteins *p170 and p130.*

The *MDM2 protein* has been identified as a further control element that can influence pRb-E2F function. The MDM2 protein was discovered as an oncoprotein activated by overexpression. It binds to the p53 protein (see Chapter 15) and to pRb protein. By binding, the growth-controlling function of both proteins is stopped. Furthermore, the MDM2 protein also binds to E2F and stimulates its transcription-activating function (Martin et al., 1995). Overall, the MDM2 protein therefore has a growth-promoting function.

The pRb protein is also involved in differentiation processes. It is therefore assumed that pRb controls other effector molecules in addition to E2F.

### pRb as Integrator of Positive and Negative Signals

Within the pRb function, there is a meeting of *positive* signals, i.e., promoting progress of the cell cycle, and *negative* signals that halt the cell cycle.

### Positive Signals

The extent of phosphorylation, and thus the activation of pRb, is closely associated with the concentration of cyclins, especially cyclin D1. If the concentration of cyclin D1 increases over a certain threshold, due to division-promoting signals, the phosphorylation of pRb, and thus transcription of E2F-controlled genes, is promoted.

A further influence on pRb function may be performed by viral oncoproteins. These bind to the underphosphorylated form of pRb and compete with E2F for pRb binding (see Chapter 14). E2F is released from the inhibiting influence of pRb in this situation and can activate its target genes.

### Negative Signals

The pRb protein also serves indirectly as an integrator of negative signals emitted in the form of a mobilization of inhibitors of CDKs. The antimitogenic TGFβ, for example, strongly increases the concentration of the inhibitor $p15^{ink4b}$ via a transcription-stimulating effect. $p15^{ink4b}$ binds to CDK4 and CDK6 and competes with cyclin D for binding of the CDKs. This reduces the concentration of active CDK complex and hinders phosphorylation of pRb. Furthermore, the TGFβ protein brings about an increase in activity of the inhibitor $p27^{KIP1}$ and this has a repressing effect on the transcription of the CDK4 gene.

The tumor suppressor protein *p53* (see Chapter 14) also indirectly controls the pRb function. p53 induces the transcription of the inhibitor p21, which inhibits the activity of CDK4/CDK6 kinase and CDK2 kinase. Due to this inhibition, activation of p53 ensures that pRb remains underphosphorylated and that the cell cycle is halted in $G_1$ phase. Activation of p53 is observed, for example, on damage of DNA.

*Cyclin A* is assigned a special role in progess of S phase and transition into $G_2$ phase. Cyclin A binds and activates CDK2. The CDK2-cyclin A complex binds to the transcription factor E2F-1 and phosphorylates its DP-1 subunit. As a consequence, the DNA binding capacity of the transcription factor is reduced and the transcription-activating function is inhibited. This function is a control element for progress of S phase. A defect in the interaction of cyclin A-CDK2 with E2F-1 leads to the cell halting in S phase (Krek et al., 1995).

## 13.5 Cell Cycle Control of DNA Replication

Replication of DNA in S phase is subject to strict control in the cell cycle (review: Stillman, 1996), resulting in the following observations:

- DNA replication is restricted to S phase
- DNA is only replicated once in a cycle
- The time sequence of DNA replication during S phase and mitosis is strictly maintained
- If DNA damage is present, DNA replication can be stopped (DNA damage checkpoint)

Control of DNA replication occurs at two levels in particular:

**Control at the Initiation Level**

A main control of DNA replication occurs at the level of initiation of replication. The replication of a DNA sequence starts at specific sequence sections of the DNA, known as replication origins. In yeast, the initiation sites of DNA replication have been very well defined and characterized at the sequence level. In higher eucaryotes, in contrast, initiation has a broad initiation zone and it has not been possible to identify a defined initiation sequence. The size of the genome in eucaryotes necessitates the use of many replication origins, which can be activated in a defined time sequence and position-specific order.

The most important components in cell cycle control of origin activity have been identified, although the questions of the spatial coordination and time sequence of origin activity are still unanswered.

Control of origin activity occurs via specific protein complexes that are bound at certain times of the cell cycle to a replication origin. For replication initiation, two states of this protein complex are important, known as the pre-replication complex and the post-replication complex.

The *pre-replication complex* (pre-RC) is formed during anaphase and is inherited by the sister chromatids. Upon entry into S phase, the pre-RC must be disrupted for initiation to occur. If initiation begins, the pre-RC changes to the post-replication complex state, which does not permit further initiation.

At least four protein complexes are involved in formation of the initiation replication complexes:

- Origin recognition complex (ORC), composed of 6 different proteins. The ORC binds in a sequence-specific manner to the origin DNA and remains constitutively bound there during the whole cell cycle.
- MCM proteins, including 6 different proteins, MCM2—MCM7. The MCM proteins bind the ORC-DNA complex, but dissociate in the process of initiation during S phase.
- Cdc6 protein: The Cdc6 protein is an essential component of the pre-RC. It is synthesized during G2 phase and is available for formation of the per-RC during mitosis.
- „Licensing complexes“: These are other poorly characterized proteins that are necessary for formation of the pre-RC.

The function of these protein complexes is quite well understood for *S. cerevisiae* and is illustrated in the following model (Fig. 13.15).

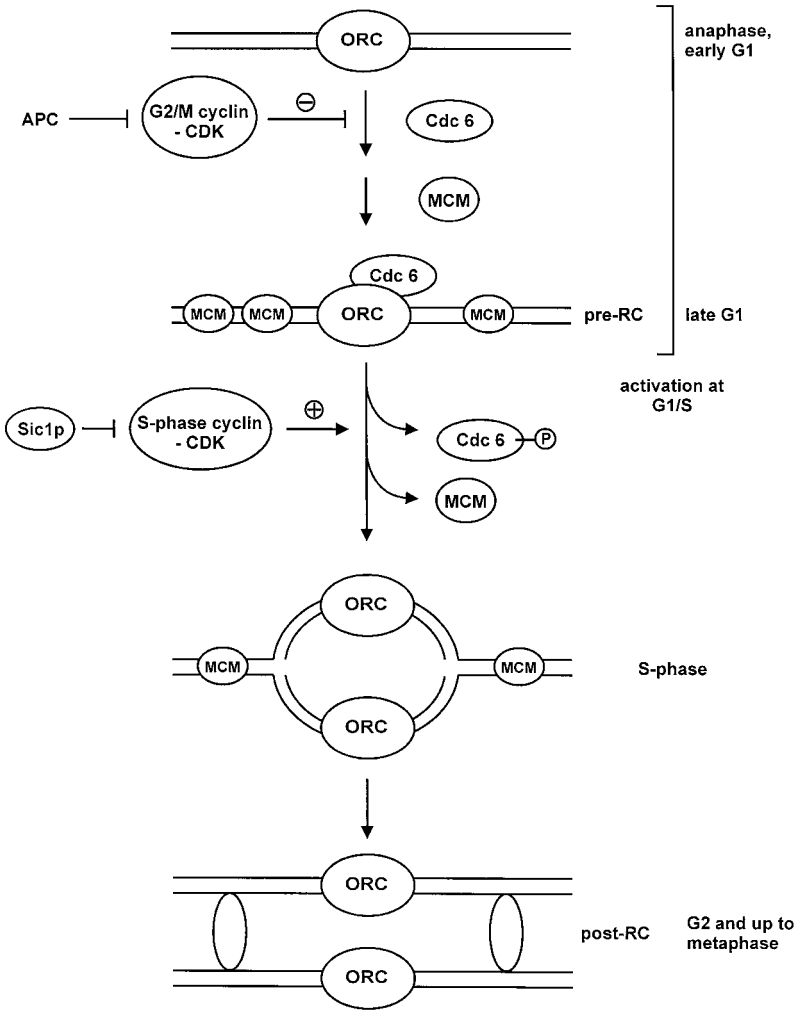



**Fig. 13.15.** Model of the cell cycle control of chromosomal replication in eucaryotes
In anaphase and during G1, a prereplication complex (pre-RC) is formed at replication initiation sites which contains the Cdc6 protein and the MCM proteins, in addition to the constitutively bound origin recognition complex (ORC). Formation of pre-RC is negatively regulated by the G2/M cyclin – CDK complex. The Cdc6 protein and the MDM proteins are removed at the G1-S transition due to phosphorylation by the S-phase cyclin – CDK complex which allows the onset of DNA replication. The activity of the cyclin-CDK complexes is controlled by several means, including degradation (anaphase promoting complex, APC) and by inhibitors (Sic1 in yeast). Upon completion of S-phase, a post-replication complex (post-RC) is established. Pre-RC and post-RC cannot occur simultaneously in the cell, and this prevents re-replication of DNA in G2- and in M-phase. The cell can enter a new S-phase only after the chromatin has been licensed for a new round of DNA replication by the binding of the Cdc6 protein and MDM proteins. The diagram does not include the participation of other protein factors that have been shown to be involved in the licensing process.

During cell division, a replication-competent state is established at the replication start sites, the pre-RC. This contains the ORC, the MCM proteins and the cdc6 protein. Formation of the pre-RC in late M phase and in G1 phase „licenses“ the chromatin for DNA replication. With entry into S phase, the MCM proteins and the Cdc6 protein dissociate from the start site. Their phosphorylation by an active S phase cyclin-CDK complex is responsible for the dissociation.

The activity of the S phase cyclin-CDK complex is, in turn, controlled by an inhibitor, the Sic1 protein. At the start of G1 phase, the S phase cyclin-CDK complex is inactivated by the Sic-1 protein. Only when the inhibitor is degraded by ubiquitin-mediated proteolysis when the start point is crossed is an active S phase cyclin-CDK complex available. This removes the MCM proteins and Cdc6 protein from the origin by phosphorylating them and thus enables replication. A second protein kinase, Cdc7/Dbf4p complex, is also involved in this phosphorylation. Once replication has taken place, the renewed formation of a pre-RC is hindered by the activity of the G2/M cyclin-CDK complex.

#### Availability of the Replication Components

At the start and during S phase, all proteins required for replication and the dNTPs must be available in sufficient quantities. An important control function is performed here by the transcription factor E2F, which induces the different enzymes needed for replication.

In certain situations, such as DNA damage, DNA replication may be halted via inactivation of an important replicatory protein. An example is the binding of the replication protein PCNA to the p53 protein. If DNA damage is present, the p53 protein is activated and binds to PCNA so that it is no longer available for replication and replication is halted.

## 13.6 The $G_2$/M Transition and Cdc25 Phosphatase

Important decisions for the course of mitosis are the $G_2$/M transition and within mitosis, the transition from metaphase to anaphase (review: Draetta and Eckstein, 1997).

Entry into and the course of mitosis are primarily determined by the activity of the CDC2 kinase. The CDC2 kinase in the active form exists as a complex with cyclin B and, together with the cyclin, forms the mitosis promoting factor, MPF. The activity of MPF oscillates in the cell cycle and is the triggering factor for entry of the cell into M phase.

A crucial regulatory element of mitosis is the concentration of cyclin B. The concentration of cyclin B increases with entry into S phase to a threshold at which it is sufficient to trigger mitosis. Mitosis does not begin yet, however, because the CDK2-cyclin B complex requires targeted activation by phosphorylation and dephosphorylation. The phosphorylation occurs on Thr161, Thr14 and Tyr15 (Fig. 13.5). In this three-fold phosphorylated form, the CDC2-cyclin B complex is inactive and remains as such until the end of $G_2$ phase. Targeted activation occurs at the G2/M transition by dephospho-

rylation at Tyr15 and Thr14. This reaction is performed by CDC25 phosphatase, the activity of which is also controlled by phosphorylation (see Fig 13.6). An activated CDC2-cyclin B complex can phosphorylate and activate the CDC25 protein by a positive regulation mechanism. The extent of phosphorylation and the activity of the CDC25 phosphatase are subject to both positive and negative control. Overall, this regulation system ensures a rapid increase in activity of MPF at the $G_2$/M phase transition.

## 13.7 The DNA Damage Checkpoint

The cell cycle contains built-in control mechanisms that register defects in the course of the cell cycle and bring about a halt in the cell cycle to enable the fault to be repaired or to lead the cell to programmed cell death. These control mechanisms are also known as checkpoints. These are biochemical pathways that are activated when a fault occurs and can influence other critical steps of the cell cycle. Of particular importance for the cell cycle is the DNA damage checkpoint. Another important checkpoint is the spindle assembly checkpoint, which is not well biochemically characterized, however.

As a reaction to DNA damage or a replication block, the cell induces a series of different physiological responses that enable DNA repair. These responses include:

- Halt in the cell cycle in G1, S or G2 phases
- Slowing of DNA replication
- Increased transcription of repair genes
- Induction of programmed cell death, apotosis

The signaling pathways that lead from the appearance of DNA damage to a halt in the cell cycle include a number of components of which the function has only partially been characterized. So far, it has been established that the DNA damage checkpoints of different organisms have a common homologous component, namely a protein kinase, which belongs to the superfamily of PI3-kinases (see 6.6).

For the fission yeast *S. pombe*, it has been possible to put together a very plausible model that leads to a halt in the cell cycle at the $G_2$/M transition on damage of DNA and prevents entry into mitosis (Fig. 13.16).

The target of this inhibitory signaling pathway is the CDC25 phosphatase that activates the G2/M transition by dephosphorylation of the mitotic cyclin-CDC2 complex. The starting point of this pathway is damage of the DNA, which is registered by the protein kinase Rad3. The Rad3 kinase phosphorylates and activates a sequentially following protein kinase, the Chk1 kinase, which has the CDC25 phosphatase as a substrate. The latter is phosphorylated by Chk1 kinase on the Ser residues 33, 192 and 359. The Ser residues serve as binding sites for 14–3-3 proteins. In complex with the 14–3-3 proteins, the CDC25 phosphatase is translocated from the nucleus into the cytosol. It is therefore no longer available for activation of the mitotic cyclin/CDC2 complex and the cell cycle is halted (Lope-Girona et al., 1999).

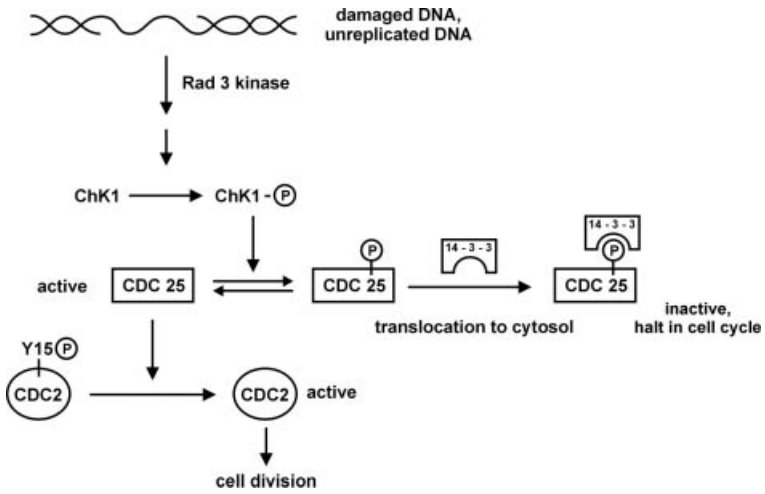



**Fig. 13.16.** The DNA replication and DNA damage checkpoint in fission yeast.
The presence of damaged or unreplicated DNA during mitosis activates a checkpoint of which the CDC25 phosphatase ia an essential element. Upon DNA damage, a signal is transmitted to the proteinkinase Chk1 which becomes phosphorylated and activated. Activated Chk1 then phosphorylates CDC 25 at several Ser residues. The Ser phosphates formed serve as attachment points for binding of 14–3-3- proteins. In the 14–3-3-bound state, the CDC 25 phosphatase is no longer available for dephosphorylation and activation of the CDC2 kinase. Consequently, the cell cycle comes to a halt.

In mammals, there are similar signaling pathways that lead to a halt in the cell cycle when DNA damage is present or when DNA replication is incomplete (Zeng et al., 1998). Important components of these signaling pathways are the ATM kinase and the p53 protein (see 14.4.4.5).

# References Chapter 13

Adams, P.D. and Kaelin, W.G. 'Negative control elements of the cell cycle in human tumors' (1998) Curr. Op. Cell Biol. *10*, 791–797

Bernards, R. 'E2F: a nodal point in cell cycle regulation' (1997) Biochem. Biophys. Acta *1333*, M33-M40

Brehm, A., Miska, E.A., McCance, D.J., Reid, J.L., Bannister, A.J. and Kouzarides, T. 'Retinoblastoma protein recruits histone deacetylase to repress transcription' (1998) Nature *391*, 597–601

Draetta, G. and Eckstein, J. ' Cdc25 protein phosphatases in cell proliferation' (1997) Biochim Biophys Acta *1332*, M53–63

Cavanaugh, A.H., Hempel, W.M., Taylor, L.J., Rogalsky, V., Todorov, G. and Rothblum, L.I. 'Activity of RNA polymerase I transcription factor UBF locked by Rb gene product' (1995) Nature *374*,177–180

Elledge, S.J. 'Cell cyle checkpoints: preventing an identity crisis' (1996) Science *274*, 1664–1672

Goldsmith, E.J. and Cobb, M.H. 'Protein kinases' (1994) Curr. Biol. *4*, 833–840

Heichman, K.A. and Roberts, J.M. 'Rules to replicate by' (1994) Cell *79*, 557–562

Hersko, A. 'Roles of ubiquitin-mediated proteolysis in cell cycle control' (1997) Curr. Op. Cell Biol. *9*, 788–799

Jeffrey, P.D., Russo, A.A., Polyak, K.; Gibbs, E.; Hurwitz, J.; Massague, J. and Pavletich, N.P. 'Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex' (1995) Nature *376,* 313–320

Johnson, L.N., Noble, M.E., Owen, D.J. 'Active and inactive protein kinases: structural basis for regulation' (1996) Cell *85*, 149–158

King, R.W., Deshaies, R.J., Peters, J.M. and Kirschner, M.W. 'How proteolysis drives the cell cycle' (1996) Science 274, 1652–1659

Krek, W., Xu, G. and Livingston, D.M. 'CyclinA-Kinase regulation of E2F-1 DNA binding function underlies suppression of an S-phase checkpoint' (1995) Cell *83*, 1149–1158

Lopez-Girona, A., Furnari, B., Mondesert, O. and Russell, P. 'Nuclear localization of Cdc25 is regulated by DNA damage and a 14–3-3 protein' (1999) Nature *397*, 172–175

Luo, Y., Hurwitz, J. and Massague, J. 'Cell-cycle inhibition by independent CDK and PCNA binding domains in $p21^{cip1}$' (1995) Nature *275*, 159–161

Magnaghi-Jaulin, L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le Villain, .J.P, Troalen, F., Trouche, D., Harel-Bellan, A. 'Retinoblastoma protein represses transcription by recruiting a histone deacetylase' (1998) Nature *391*, 601–605

Martin, K., Trouche, D., Hagemeier, C., Sorensen, T.S., La Thangue, N.B. and Kouzarides, T. 'Stimulation of E2F1/DP1 transcriptional activity by MDM2 oncoprotein' (1995) Nature *375*, 691–698

Morgan, D.O. 'Principles of CDK regulation' (1995) Nature *374*, 131–134

Nurse, P. 'Ordering S phase and M phase in the cell cycle' (1994) Cell *79*, 547–550

Peter, M. 'The regulation of cyclin-dependent kinase inhibitors (CKIs)' (1997) Prog Cell Cycle Res *3*, 99–108

Peters, J.M. ' SCF and APC: the yin and yang of cell cycle regulated proteolysis' (1998) Curr. Op. Cell Biol. *10,* 759–768

Pines, J. 'Cyclins and cyclin-dependent kinases: a biochemical view' (1995) Biochem. J. *308*, 697–711

Pines, J. 'Confirmational change' (1995) Nature *376*, 294

Planas-Silva, M. and Weinberg, R.A. 'The restriction point and control of cell proliferation' (1997) Curr.Op. Cell. Biol. *9*, 768–772

Riley, D.J.R., Lee, E.Y.H.P and Lee, W.H; 'The retinoblastoma protein: more than a tumor suppressor' (1994) Annu. Rev. Cell Biol. 10, 1–19

Russo,A.A., Jeffrey,P.D., Patten,A.K., Massague,J. and Pavletich,N.P. 'Crystal structure of the $p27^{Kip1}$ cyclin-dependent-kinase inhibitor bound to the cyclinA-Cdk2 complex' (1996) Nature 382, 325–331

Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D. and Pavletich, N.P. 'Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a' (1998) Nature *395*, 237–243

Sherr, C.J. 'G1 phase progression: cycling on cue' (1994) Cell *79*, 551–559

Sherr, C.J. 'D-type cyclins' (1995) Trends Bioch. Sci. *20*, 187–190

Sherr, C.J. 'Cancer cell cycles' (1996) Science *274*, 1672–1677

Solomon, M.J. 'The function(s) of CAK, the $p34^{cdc2}$-activating kinase' (1994) Trends Bioch. Sci. *19*, 496–500

Stillman, B. 'Cell cycle control of DNA replication' (1996) Science *274*, 1659–1664

Tomoda, K., Kubota, Y. and Kato, J. 'Degradation of the cyclin-dependent-kinase inhibitor p27Kip1 is instigated by Jab1' (1999) Nature *398*, 160–165

Weinberg, R.A. 'The retinoblastoma protein and cell cycle control' (1995) Cell *81*, 323–330

Weintraub, S.J., Chow, K.N.B., Luo, R.X., Zhang, S.H., He, S. and Dean, D.C. 'Mechanism of active transcriptional repression by the retinoblastoma protein' (1995) Nature *375*, 812–815

White, R.J., Trouche, D., Martin, K., Jackson, S.P. and Kouzarides, T. 'Repression of RNA polymerase III transcription by the retinoblastoma protein' (1996) Nature *382*, 88–90

Zachariae, W., Schwab, M., Nasmyth, K and Seufert, W. 'Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex' (1998) Science *282,* 1721–1724

Zeng, Y., Forbes, K.C., Wu, Z., Moreno, -S, Piwnica-Worms, H. and Enoch, T. 'Replication checkpoint requires phosphorylation of the phosphatase Cdc25 by Cds1 or Chk1' (1998) Nature *395*, 507–510

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 14**

# Malfunction of Signaling Pathways and Tumorigenesis: Oncogenes and Tumor Suppressor Genes

## 14.1 General Comments on Tumor Formation

### 14.1.1 Characteristics of Tumor Cells

Tumor cells have special features compared to normal cells. The phenotype of a tumor cell is characterized by the following characteristics:

- Increased rate of cell division, loss of normal growth control
- Loss of ability to differentiate
- Loss of contact inhibition
- Increased capability for invasion of neighboring tissue, metastasis

The cells of a fully-grown, aggressive tumor have gained these properties in a slow, multi-step process with the characteristics of cellular evolution. This development is associated with a selection process, in the course of which, cells that have lost their growth-regulating mechanisms predominate. The transition of a normal cell to a tumor cell can often be well characterized both morphologically and physiologically. However, only a few of the underlying biochemical changes are understood. It is also unclear to a large extent how the different stages of tumor formation interact and interdepend (Vogelstein and Kinzler, 1993).

### 14.1.2 Genetic Changes in Tumor Cells

A general phenomenon in tumor formation is the *stepwise accumulation* of changes in genetic information (mutations). The genetic changes are passed on from the mother cell to the daughter cell during cell division. The mutations observed in tumors include a broad spectrum of reorganizations and changes in genetic information. The extent of the changes can be very varied. Smaller mutations are observed such as:

- Simple base substitutions
- Insertion, deletions of bases
- Inversions and duplications of DNA sequences

The mutations can lead to changes in the alignment of genes with products involved in signal transduction pathways and in growth regulation. It is assumed that these mutations are particularly important in the early phase of tumor formation.

In the later phases of tumor formation, an increasing *genetic instability* occurs which is observed at the level of the chromosomes, particularly in the form of a change in the normal chromosome number (review: Lengauer et al., 1998).

The following are observed:

- Loss or duplication of whole chromosomes
- Multiplication of the chromosome set
- Chromosome translocations: deletion, addition or exchange of individual chromosomes
- Amplification of DNA sequences

Changes in the chromosome structure are often observed in tumors of the blood-forming system, the leukemias and lymphomas. They are almost always found in the later phases of aggressive solid tumors. These extensive reorganizations have wide-reaching consequences for growth behavior and functional performance.

### 14.1.3 Changes in Methylation Pattern

In addition to the direct changes in the information content of the DNA, changes in the DNA methylation pattern are also observed during the course of tumor formation (review: Counts and Goodman, 1995). The pattern of $^{me}$CpG sequences is an important tool in regulating transcription in eucaryotes (see 1.4.7). Changes in the methylation pattern also bring about changes in the expression patter of specific genes. It is therefore important that the methylation pattern is passed on to the daughter cells during cell division.

### 14.1.4 Causes of Oncogenic Mutations

Mutations that trigger and promote tumor formation may be caused by a number of processes. Briefly summarized, the following factors are involved:

#### Intrinsic Changes in the Genetic Information

There are many endogenous processes that change the information content of the DNA. These processes are inseparable from the performance and characteristics of an organism and they are only influenced by external factors to a small extent.

**DNA Replication Errors**

Replication of DNA does not occur with complete precision, but rather has an intrinsic inaccuracy. The error rate for incorporation of nucleotides in DNA replication is of the order of one error per $10^6 - 10^8$ correctly incorporated nucleotides.

**Spontaneous DNA Damage**

DNA has only a limited stability in the temperature and pH conditions of an organism. Spontaneous changes in the DNA structure may occur and cleavage of purine bases is assigned an important role. The apurinic sites resulting from depurination may give rise to mutations if not repaired.

**Metabolism-related Damage of DNA**

Reactive metabolic products are an important cause of endogenous damage to DNA. In first place is reactive oxygen, which can lead to DNA damage, especially in the form of the superoxide anion $O_2^-$ and in the form of OH● radicals. Dietary components also include many chemical compounds with which DNA can react. Furthermore, dietary components can be converted by metabolic activitation into compounds with a high potential for damage of DNA.

**2) External Damage of DNA**

A number of external factors can lead to DNA damage and thus to mutations. These include especially the effect of cancerogenic chemicals, and UV, X-ray and other high energy radiation (e.g.radioactivity).

**Viruses**

There are a large number of RNA- or DNA-viruses that are directly associated with tumor formation. Three principle mechanisms are responsible (see 14.3):

- Introduction of viral oncogenes into the host genome
- Interaction with virus-specific proteins and signal proteins of the host cell
- Control of expression of proto-oncogenes by a viral promoter

## 14.1.5 DNA Repair and Tumor Formation

Due to the many damaging influences on the DNA, it is essential for the cell that DNA damage can be repaired. The repair systems involved, such as the enzyme apparatus of excision repair, the alkyl transferases and the mispairing repair system, operate almost without error and can repair the vast majority of DNA damage. A deficit in the repair

capacity of a cell, for example, due to inactivating mutation of a repair component, favors the accumulation of mutations and gives tumor formation a boost. Proteins that are involved in the avoidance of DNA damage, either directly or indirectly, make an important contribution to suppression of tumors. Thus, mutated tumor suppressor proteins are often found in association with a hereditary predisposition towards tumor formation (hereditary tumor syndrome), and many proteins involved in repair of DNA damage are tumor suppressor proteins (see 14.4.2).

These observations support the theory that, in an early stage of tumor formation, a mutation occurs in a repair system needed to maintain the integrity of the genome (Loeb, 1991, Loeb, 1998). Loss of the function of the repair system leads to a *mutator phenotype*: a missing or ineffective DNA repair favors further accumulation of mutations and leads to an intrinsic instability of the genome.

Genetic instability due to breakdown of a repair component is only seen in a small fraction of tumors, however. In the majority of tumors, instability at the level of the chromosomes can be established, which is manifested in loss or gain of whole chromosomes or large parts thereof. This instability is associated with failure at cell cycle checkpoints (see 13.7, 14.4.2).

### 14.1.6 Cell Division and Tumor Formation

Cell division itself is also attributed a central function in tumor formation. A pre-cancerous damage to the DNA, such as formation of an adduct of a base and a cancerogenic compound, can only then lead to mutations if the cell undergoes a round of division with its damaged DNA. On replication of the damaged DNA, there is an increased probability of changes in the DNA sequence, unless the error is repaired beforehand. These changes will then be passed on to subsequent generations as mutations.

Furthermore, each cell division round has the potential danger of rearrangement of chromosome sections during mitosis and thus chromosome aberrations. A thorough theoretical and experimental analysis of the dose-effect relation of various cancerogenic substances has shown that an increased cell division activity is an important risk factor for the creation of tumors (Cohen and Ellwein, 1990). All processes that lead to an increase in the rate of cell division will increase the probability of tumor formation, according to these investigations.

## 14.2 Cell Division Activity, Errors in Function of Signal Proteins and Tumor Formation

DNA damage and resulting mutations are generally randomly distributed over the genome and thus affect all genes of an organism with equal probability. Within the confines of the cellular evolution process leading from the normal cell to a tumor cell, cells with normal growth control removed by mutation are favored. Mutations of genes involved in growth control of a cell are therefore of particular importance in tumor creation.

The preceding chapters have shown that growth and cell division are controlled by a complicated process involving many protein complexes and in which intrinsic and external control mechanisms are effective. The process includes mitogenic signals, such as growth factors, and antimitogenic signals, such as TGFβ (see 12.1; 13.1.3; 13.2.6). The external signals serve to adapt cell division activity of a cell to the function of the organism.

## 14.2.1 The Fate of a Cell: Division, Non-division or Death

A cell can enter various stages depending on the intensity and balance of mitogenic and antimitogenic signals (Fig. 14.1):

### Continuous Cell Division

Continuous cell division is found, for example, in stem cells, which serve as precursors for other cells, and in tissue in which dying cells must be replaced. This requires the continuous effects of mitogenic signals. In this case, cell division must ensure homeostasis of the cell structure or the tissue. An increase in the cell number through cell division compensates for loss of cells that die as a part of normal cell turnover or that are eliminated by programmed cell death, or apoptosis.

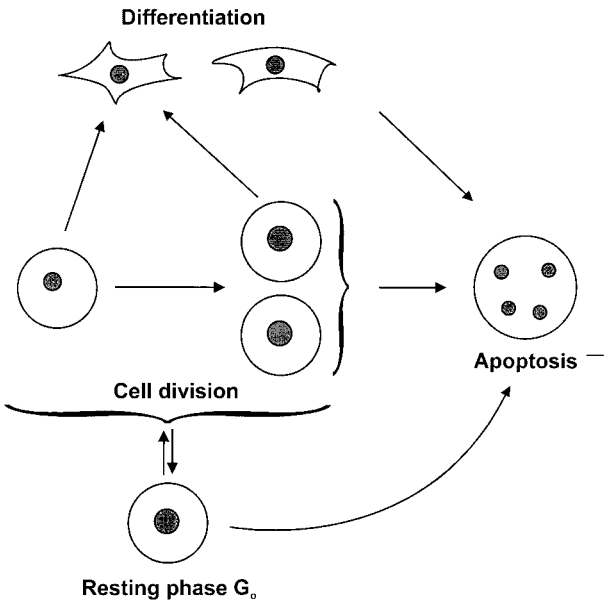



**Fig. 14.1.** Division activity and fate of a cell. Cells can move from a division-active state to a differentiated state or into the resting phase Go. The transition into Go phase is generally reversible, whereas differentiated cells often cannot return to the division-active state. The cell can be directed from all three stadia into programmed cell death, or apoptosis.

**Transition into a Resting Phase**

In the absence of external mitogenic signals or in the presence of a majority of antimitogenic signals, cell division activity may be stopped. The cell enters the resting phase ($G_o$ phase). From $G_o$ phase, cell division may be resumed when mitogenic signals reappear.

**Differentiation of Cells**

During the development of an organism, many cells enter a state of terminal differentiation, in which they perform a specialized function. Differentiated cells originate from dividing stem cells and in the process, partially or completely lose the ability to divide. They are then no longer able to receive and act upon mitogenic signals.

**Apoptosis**

Another route that a cell can take is the pathway of programmed cell death, or apoptosis (see Chapter 15).

Many different factors may be involved in triggering apoptosis, some of which are also components of mitogenic or antimitogenic signaling pathways.

### 14.2.2 Definition and General Function of Oncogenes and Tumor Suppressor Genes

The mitogenic signals that the organism uses to control division activity of individual cells are registered by corresponding receptors and converted into intracellular signal chains. The signals finally meet at the control system of the cell cycle. Here, it is decided whether cell division will take place or not. It is therefore to be expected that mutations, accumulating in a multistep process in the transition of a normal cell to a tumor cell, will be particularly visible in genes involved in control of cell growth. As outlined in Chapter 13, many proteins are involved in the conversion of external signals at the levels of the central cell cycle apparatus (CDKs, cyclins, CKIs).

The growth behavior of a normal cell is determined by a balanced relationship between positive, proliferation-promoting influences and negative, proliferation-inhibiting influences, and this is adapted to the function of the cell within the organism. Defects in cell division activity may arise by mistakes in positive signals and also in negative signals. If the division-promoting signals are excessively active, the affected cell gains a growth advantage in comparison to normal cells and formation of a tumor is favored. The same effect is achieved if the negative, suppressing elements of the mitogenic signaling pathway no longer function. A deficit in suppressing signals that negatively control growth leads to an excess of positive signals and thus promotes tumor growth.

Due to the large number of proteins involved in the network of growth control, it is not surprising that many genes have been identified which, when mutated or not fun-

ctioning normally, lead to a fault in growth control and correlate with tumor formation. The mutated genes are roughly divided into oncogenes and tumor suppressor genes. This division is based on the effect of a mutation on function.

### Cellular and Viral Oncogenes, Proto-oncogenes

The term oncogene was coined in association with the search for the tumor-causing principle in retroviruses. Retroviruses contain RNA as the genetic material and can transcribe RNA into DNA with the help of the virus's own enzyme reverse transcriptase. The DNA form of retroviruses can integrate into the DNA of the host cell and, during cell division, is passed on to the daughter cells as a provirus. From the provirus, viral RNA and complete virus particles may be formed.

Some representatives of the retroviruses cause tumors in animals such as mice or chickens. The discovery of oncogenes initiated from the *src gene* of *Rous sarcoma virus*, which could be identified as the tumor causing principle of this retrovirus. The *src* gene codes for the Src tyrosine kinase (see 8.3). The gene sections of retroviruses responsible for tumor formation were designated *oncogenes.*

Soon after discovery of the oncogene, it was demonstrated that the viral oncogenes are mutated forms of the genes of cellular proteins that are also active in normal cells. The cellular variants of viral oncogenes were named *proto-oncogenes.*

The oncogenes of retroviruses are prefixed with a *v* (e.g., *v-src, v-sis*) whereas the corresponding proto-oncogenes are prefixed with a *c* (e.g., *c-src, c-sis*).

Further investigations revealed that proto-oncogenes can also be converted, via activating mutations, into oncogenes without involvement of viruses. The use of the term „oncogene" was thus extended and its definition was made more general.

Oncogenes are genes that can result in a transforming or immortalizing phenotype on experimental transformation in cellular model systems (see 14.3). Oncogenes arise by activating mutation of their precursors, the proto-oncogenes. Proto-oncogenes are often directly involved in growth regulation of normal cells. Oncogenes generally have *dominant* character. The mutation of a proto-oncogene to an oncogene is phenotypically visible when only one of the two copies of the gene in a diploid chromosome set is affected by the mutation. The dominant mutation is accompanied by a *'gain of function'*, it typically amplifies or increases the yield of a function in growth regulation.

### Tumor Suppressor Genes

Tumor suppressor genes are genes that have a negative, suppressing effect on tumor creation and thus help to prevent formation of tumors. Mutations of tumor suppressor genes are often recessive. On mutation of one allele, the remaining intact allele on the other chromosome continues to perform the growth-suppressing function. Only when both alleles are inactivated does the tumor-suppressing function cease to work. By this property, the inactivation of neighbouring marker genes during tumorigenesis (loss of heterozygosity, LOH) has often helped to identify tumor suppressor genes. However, tumor suppressor genes are known for which mutation of one allele already promotes tumor formation. Halving the gene dose is apparently sufficient in these cases to lift tumor suppression. An example is the cell cycle inhibitor $p27^{KIP1}$ (Fero et al., 1998).

The terms oncogene and tumor suppressor gene do not give any information about the *actual* function of the gene or the protein for which they code. Both are coupled in a complex way to different positions in the network of proliferation regulation and in cell-cell interactions, as shown in the examples below.

### 14.2.3 Cellular Systems for Investigation of the Function of Oncogenes and Tumor Suppressor Genes

The function of oncogenes and tumor suppressor genes is generally investigated in defined cell lines. These serve as model systems in which individual aspects of tumor formation can be investigated.

The cell lines employed differ in many aspects from normal body cells. Human cells taken from normal tissue and maintained in primary culture with the addition of suitable nutrient media only have a limited capacity for division. Cells of primary cultures of human fibroblasts, for example, halt cell division activity after 50—60 cell divisions and enter a state known as *cellular senescence*. In contrast, cells taken from a tumor often have an unlimited ability to divide in cell culture.

The transformation of a normal cell into a tumor cell is marked by the appearance of certain characteristics in the cellular model system, which depend on the cell type and may appear alone or in combination.

The following features are characteristic for cells on the way from a normal state to the status of a tumor cell:

- Immortalization
- Reduced need for growth factors
- Changed morphology
- Formation of cell clumps (foci)
- Loss of contact inhibition

The transformation of cells of a primary culture into a state where they no longer show cellular senscence and are immortalized may take place by treatment with transforming agents (e.g., the SV40 virus) or by genetic manipulation (e.g., introduction of oncogenes). On addition of growth factors into the culture medium, the cells can then be kept in culture for an unlimited duration. Another feature of changed regulation of cell division activity is a reduced need for growth factors. Other parameters that characterize the transformation of a normal cell to a tumor cell are changed morphology, the ability to grow in the form of cell clumps (foci) and loss of contact inhibition. Normal cells generally grow in cell culture in the form of one-cell layers and the cells attach to the walls of the culture dish. Deregulation of growth behavior may be manifested by the cells growing on top of one another and forming foci. The ability to form foci is an important marker that distinguishes a tumor cell from a normal cell.

Immortalization, reduced requirement for growth factors and changed growth behavior are different characteristics of cells undergoing transformation from a normal state to a tumor cell. All three characteristics are important parameters which are used to characterize the function of oncogenes and tumor suppressor genes.

Both primary cultures and immortalized, permanent cell lines are used for functional investigation of oncogenes and tumor suppressor genes. An immortalizing effect may be demonstrated on primary cultures and a transforming effect of a gene may be demonstrated on permanent cell lines. The importance of tumor suppressor genes and oncogenes for growth control has been tested and demonstrated in cellular model systems in many cases. The value of the cellular systems is highlighted by the observation that many of the transforming or immortalizing genes in cellular model systems are also found in a similarly mutated form in tumors in man or mice.

Tumor formation is a *multi-step process* in which activation of proto-oncogenes to oncogenes, deactivation of tumor suppressor genes and extensive rearrangement of DNA occur in a sequential process. Which oncogene and which tumor suppressor gene are involved in a particular step depends on the type of tumor. The number and order of the necessary genetic changes can be very different in different tumor types. It is not clear whether a defined order of the observed genetic changes is necessary for complete formation of a tumor. Equally, the cooperation of genetic changes during tumor formation is not understood. Thus the observations in cellular model system can only offer a small and limited insight into the mechanism by which oncogenes and tumor suppressor genes cooperate and become active in tumors in animals or man.

## 14.3 Oncogenes and Proto-oncogenes

The majority of proto-oncogenes code for proteins involved in transduction and operation of mitogenic signals in a direct or indirect way. Activation of a proto-oncogene to an oncogene leads to an increase in the growth-promoting function of the coded protein. In the cellular model system, activation may be manifested as immortalization, reduced need for growth factors or as loss of contact inhibition. A review of proto-oncogenes and oncogenes is given in Hartwell and Kastan (1994) and Baserga (1994).

### 14.3.1 Mechanisms of Activation of Proto-oncogenes

The activation of a proto-oncogene to an oncogene is based on mutations that can change the function and regulation of the affected protein by various mechanisms. Two pathways of activation can be roughly differentiated (Fig. 14.2). On the one hand, the structure of the coded protein may be affected; on the other hand, activation may lead to a concentration increase in the protein.

#### 14.3.1.1 Activation by Structural Changes

A frequent cause of activation of proto-oncogenes is a change in the structure of the coded protein, affecting the regulation and function. Via the oncogenic activation, there is no creation of completely new functions, but rather the normal function of a proto-oncogene product is modified and/or released from cellular regulation.

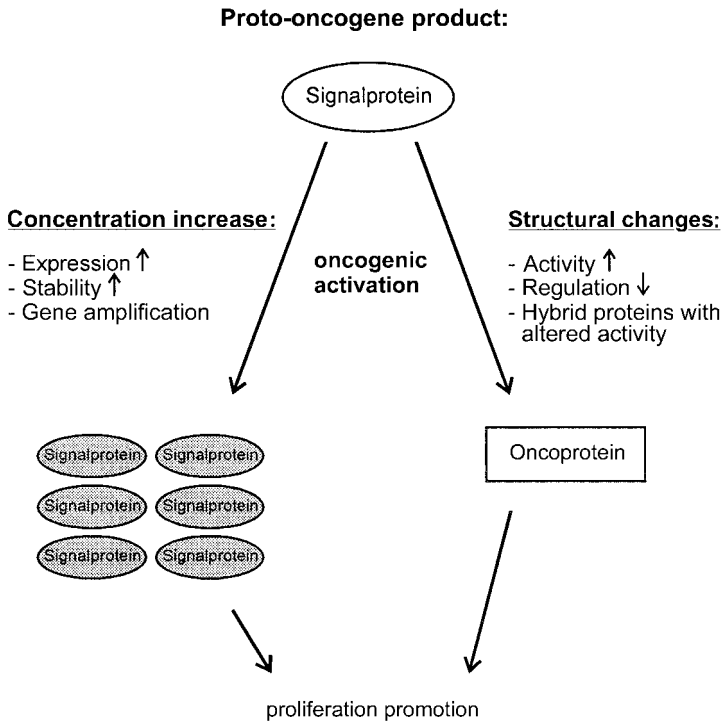



**Fig. 14.2.** Mechanism of activation of proto-oncogenes to oncogenes. Proto-oncogenes may be converted into oncogenes via the concentration increase pathway or the structural change pathway. In the case of the concentration increase, there is an excessive and unprogrammed function of the signal protein coded by the proto-oncogene. In the case of structural change, the proliferation-promoting activity of the oncoprotein results from changed activity, altered regulation or formation of a hybrid protein.

The spectrum of structural mutations that can convert a proto-oncogene into an oncogene is very diverse. Both simple amino acids changes in the coded protein and larger structural changes are observed. In particular, viral oncoproteins demonstrate multiple mutations compared to their cellular counterparts, linked to important and far-reaching structural and functional changes.

### Increase in Activity of Signal Proteins

Activating point mutations may directly affect the enzyme activity of a signal protein. The mechanism of activating oncogenic point mutations is best documented for the Ras protein. Many of the point mutations at positions 12 and 13 of the Ras protein, which are also observed in solid tumors, bring about a reduction in rate of GTP hydrolysis. This can no longer be increased by the GAP protein so that the active GTP state exists for longer duration (cf. 9.2.3). It is assumed that this is the cause of unprogrammed stimulation of the Ras-MAPK pathway, which is manifested as increased cell division activity.

**Change in Regulation of Signal Proteins**

Mutations may lead to loss of cellular control over the activity of a proto-oncogene. Frequently, this brings about constitutive activation of the signal protein. Thus, in the transforming *v-raf* gene, the N-terminal sequence section of Raf kinase is missing, on which both the autoinhibitory function and the phosphorylation sites of Raf kinase are localized (see 9.6).

**Formation of Hybrid Proteins**

In many tumors, a reciprocal exchange is observed of DNA sections on different chromosomes. During this translocation of chromosomes, gene fusions may occur, leading to formation of chimeric proteins (see 14.4). Within the chimeric proteins, there are often structural portions that originate from signal proteins. The function of the signal protein portion is removed from normal regulation in the chimeric protein and can have a tumor-promoting effect. Tyrosine kinases and transcription factors are often affected by gene fusions. The chimeric proteins arising from chromosome translocation often represent a characteristic of a particular tumor type. A review of gene fusions observed as a consequence of chromosome translocations in tumors is given in Rabbitts (1994) and Look (1997).

#### 14.3.1.2 Activation by Concentration Increase

A change in the gene expression or stability of a proto-oncogene product may lead to an increase in the cellular concentration of the protein. Due to the increased concentration, a mitogenic signal mediated by a proto-oncogene product may be amplified.

**Overexpression of Proto-oncogenes**

The overexpression of proliferation-regulating proteins may lead to immortalization and/or transformation in cellular model systems. Overexpression of signal proteins is observed in many tumors and it is assumed that the *overexpression is associated with pathogenesis of the tumor*. The mechanisms leading to overexpression are diverse and theoretically include all processes involved in expression regulation. Of note are translocations of a proto-oncogene into the vicinity of a strong promoter, causing excessive or deregulated expression of the corresponding gene. Stabilization of the mRNA of a proto-oncogene by deletions at the 3' end can also lead to overexpression, as has been shown for cyclin D1 (review: Hunter and Pines, 1994).

Activation of proto-oncogenes by unprogrammed expression is often associated with chromosome translocations in leukemias and in lymphomas. In Burkitt's lymphoma, different translocations of the *c-myc* gene are found, which codes for the c-Myc transcription factor. The translocations bring about movement of the *c-myc* gene into the vicinity of immunoglobulin genes. Consequently, constitutive expression of the c-Myc protein occurs, disturbing the normal regulatory network into which the c-Myc protein is bound (see 14.3.2).

Another tumor-promoting mechanism based on deregulation of transcription of growth factors is the formation of *autocrine loops*. In the course of tumor formation, unprogrammed expression of growth factors may occur in cells in which there would normally be no or little expression of these proteins. If these cells express the appropriate growth factor receptors, the growth factors may bind to these and create a stimulus of division. The cell is no longer dependent on the supply of an external growth factor. The cell then produces its own growth factor and division stimulus (see Fig. 3.5 and 11.1), in a similar manner to antigenic stimulation of cells of the immune system (see Baserga, 1994). An autocrine loop is set in motion, leading to uncontrolled growth of the cell (Fig. 14.3).

### Amplification of Genes for Signal Proteins

Another factor closely associated with tumor formation is an increase in number of copies of a gene for a signal protein (*gene dose effect*). In a number of tumors, amplification is observed of genes with products that have a central position in growth regulation. The tumor-promoting effect of amplification of genes coding for signal proteins can also be observed in cellular model systems.

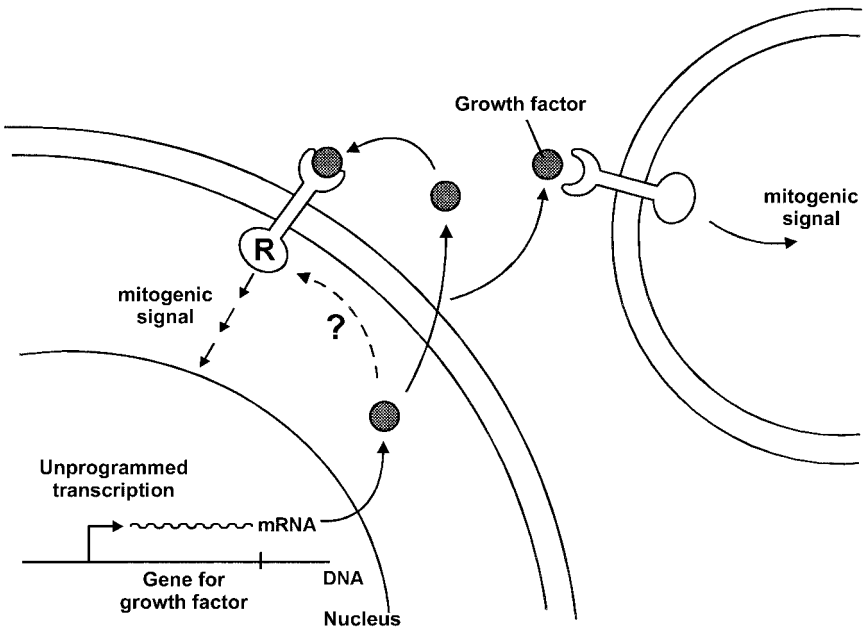



**Fig. 14.3.** Autocrine loops in tumor formation. Due to an error in control of transcription, growth factors may be produced and secreted in the cell which would normally only be formed in low concentrations or not at all. If the cell also possesses the receptors corresponding to the growth factor, the growth factor can then bind and activate a mitogenic signal chain. In this situation, the cell creates the mitogenic signal itself. There is evidence that the growth factors can become active intracellularly. The mechanism behind this is unknown.

### 14.3.2 Examples of Functions of Proto-oncogenes and Oncogenes

If the path of a mitogenic signal is traced from the exterior to the interior of a cell at the level of the cell cycle and transcription, the proto-oncogenes may be categorized into the following groups (according to Hunter, 1991):

#### Group 1: Growth Factors

An example of an oncogene derived from a growth factor is the *sis oncogene*. The viral oncogene *v-sis* codes for a protein with ca. 90 % homology to the B chain of platelet derived growth factor (PGDF-β, see also Chapter 8). The transforming effect of growth factors can also be seen in cellular model systems; for example, overexpression of the PDGF-β protein can lead to transformation in such a system. It is assumed that the oncogenic activation takes place via an autocrine or paracrine loop (see Fig. 14.3).

#### Group 2: Receptor Tyrosine Kinases and Cytoplasmic Tyrosine Kinases

#### Receptor Tyrosine Kinases

There are many examples of oncogenes that code for mutated forms of receptor tyrosine kinases. Examples include the *v-erbB* oncogene, which represents a truncated variant of the receptor for epidermal growth factor (EGF receptor). Another example is the *v-fms* oncogene derived from the receptor for colony stimulating factor 1 (CSF-1 receptor). The transforming properties of mutated receptor tyrosine kinases can often be explained by their *constitutive activation*. The effect of oncogenic mutations of receptor tyrosine kinases may be the permanent activation of the tyrosine kinase activity, which is no longer under control of extracellular ligands.

#### Nonreceptor Tyrosine Kinases

Many of the nonreceptor tyrosine kinases were discovered because the mutated form of the protein is the product of a viral oncogene. The most prominent examples are the Src tyrosine kinase and the Abl tyrosine kinase (see 8.3). The relation of the Abl tyrosine kinase with the Philadelphia chromosome translocation in lymphocytes has been especially well investigated (see Rabbitts, 1994). The *Philadelphia translocation* is a chromosome translocation affecting the *c-abl gene* of chromosome 9 and the *bcr gene* of chromosome 22. The translocation leads to formation of a hybrid gene composed of the *bcr* gene, which codes for a Ser/Thr-specific protein kinase, and the *c-abl* gene. Consequently, the two alternative fusion proteins $p210^{BCR\text{-}ABL}$ and $p180^{BCR\text{-}ABL}$ are created, which are characteristic for various leukemias.

During the translocation, a part of the *c-abl* gene is fused to the first exon of the *bcr* gene (Fig. 14.4). The $p180^{BCR\text{-}ABL}$ hybrid protein demonstrates increased tyrosine kinase activity and it has a changed subcellular location in that it is predominantly bound to the cytoskeleton. Furthermore, the BCR portion of the hybrid protein mediates binding of the adaptor protein Grb2, in that a SH2 domain of the Grb2 protein

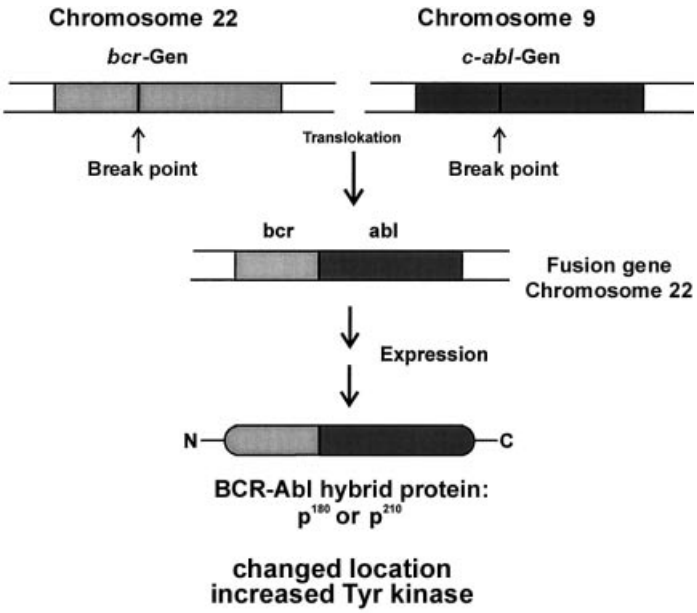



**Fig. 14.4.** Formation of a hybrid oncoprotein, illustrated by translocation of the Abl tyrosine kinase. The gene for the Ser-specific protein kinase BCR is fused with a part of the *c-abl* gene in the process of the Philadelphia chromosome translocation. Fusion genes are produced on chromosome 22, coding for various fusion proteins. The most important fusion proteins are the $p^{180}$- and $p^{210}$-BCR-Abl hybrid proteins, which have increased Tyr kinase activity and an altered subcellular location.

binds to a phosphotyrosine residue of the BCR portion. The Grb2 protein is involved in transduction of mitogenic signals in the Ras pathway (see 9.5). It has been speculated that unprogrammed binding of the Grb2 protein to the hybrid protein brings about deregulation of mitogenic signaling pathways (Pendergast et al., 1993).

### Group 3: Regulatory GTPases

Oncogenic activation of small regulatory GTPases has been documented many times for the example of the Ras proteins (see 9.2.3).

Oncogenic activation of heterotrimeric G-proteins, in contrast, is seldom observed. In some tumors of endocrine organs (thyroid glands, pituitary gland), mutated $G_{s,\alpha}$ subunits occur which have a strongly reduced GTPase activity (Landis et al., 1989). Consequently, there is a constitutive activation of cAMP, which sets in motion uncontrolled cell division in the affected cell types. The mutations of the α-subunits affect the positions Arg201 and Gln227. Arg201 is at the site of ADP ribosylation by cholera toxin (see 5.5.2). The Gln227 is equivalent to Gln204 of the $G_i$ subunits. It is directly involved in the GTPase reaction (see 5.5.4).

### Group 4: Cytoplasmic Ser/Thr Kinases and their Regulatory Subunits

The activity of Ser/Thr-specific protein kinases is often controlled by autoinhibitory sequences (see 7.1.5). Loss or lack of function of autoregulatory sequences due to an oncogenic mutation can remove Ser/Thr kinase activity bound into mitogenic signaling pathways from normal control and thereby promote tumors. An example is the Raf kinase (see 9.6). Viral v-Raf oncoproteins are characterized by a deletion of the $NH_2$-terminal regulatory sequences.

### Cyclins

The cyclins are regulatory subunits of protein kinases of the cell cycle (see Chapter 14).

The D type cyclins are of particular importance for regulation of the $G_1$/S transition and thus for regulation of the whole cell cycle. It is easy to imagine that errors in regulation of D type cyclins, e.g., by an unprogrammed concentration increase, have a proliferation-increasing effect and can promote oncogenesis. In fact, the gene for cyclin D1 has been identified as a proto-oncogene in various tumors. In benign tumors in the parathyroid, a translocation is found of the gene for cyclin D1 into the vicinity of the promoter for the parathyroid hormone. Consequently, cyclin D1 is overexpressed and uncontrolled cell division occurs (review: Hunter and Pines, 1994, Sherr, 1996).

Overexpression of cyclin D1 is also observed in some forms of bowel and breast cancer. The overexpression is due to amplification of the gene for cyclin D1.

### Group 5: Adaptor Proteins of Signal Transduction

Corresponding viral oncoproteins have been described for the adaptor proteins Crk and Shc (see 8.5). The mechanism by which the viral adaptor proteins lead to errors in regulation of cell division activity is unknown. It is possible that they cause unprogrammed and no longer regulated linking of signal proteins in mitogenic signal transduction.

### Group 6: Transcription Factors

A large number of proto-oncogenes code for transcription factors required for progression of the cell cycle and/or for the differentiation of the cell. The best known and investigated examples of oncogenic mutated transcription factors involve the *jun, fos* and *myc* genes and the genes for the T3 receptor and the vitamin A acid receptor.

Oncogenic activation of transcription factors due to chromosome translocation is often observed in leukemias. The aberrant activation by translocation of genes encoding transcription factors occurs mainly by two mechanisms (Look, 1997):

- Transcription factor proto-oncogenes that are silent or expressed at low levels in the progenitor cell are activated when placed under control of potent enhancer elements of genes, which are normally highly expressed.

- Chromosomal breakpoints occur within the coding regions of the transcription factors, producing a fusion gene that encodes a chimeric transcription factor with altered function.

The *c-myc* gene is presented as an example of overexpression of a transcription factor due to translocation.

Chromosome translocations are observed in Burkitt's lymphoma, which bring the *c-myc* gene into the vicinity of immunoglobulin genes (review: Rabbitts, 1994). The translocation has the consequence of increasing expression of the *c-myc* gene in comparison to the normal situation. Furthermore, during the course of tumor formation, mutations also occur in the coding region of the *c-myc* gene. Overexpression of the *c-myc* gene is seen as a dominating factor in the oncogenic activation of the Myc protein. The overexpression interrupts a complex network of interacting transcription factors (Fig. 14.5). The Myc protein functions as a transcription factor. It belongs to the helix-loop-helix proteins (see 1.2.1.3) and can bind via its leucine zipper dimerization motif to other transcription factors. The Myc protein can dimerize with itself and it can form heterodimers with other helix-loop-helix motif transcription factors such as the MAX protein, the MAD protein or the Mxi-1 protein. The various heterodimers have diffe-

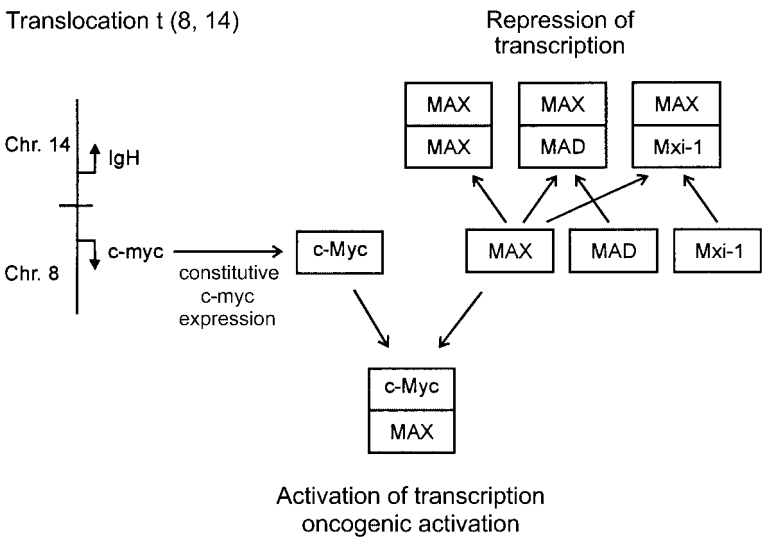



**Fig. 14.5.** The network of the c-Myc-Max transcription factors in Burkitt lymphoma. Burkitt lymphoma is characterized by chromosome translocations in which immunoglobulin gene sections (a gene for IgH in our example) are translocated into the region of the *c-myc* gene. As a consequence of the translocation t(8,14), constitutive expression of the *c-myc* gene occurs. The *c-myc* gene codes for the transcription factor c-Myc, which can form homodimers or may associate with the related transcription factors MAX, MAD and Mxi-1 to form heterodimers. Constitutive expression of c-Myc shifts the homo-heterodimer equilibrium towards the c-Myc-MAX heterodimers. Unprogrammed activation of target genes of the cMyc-MAX heterodimers then occurs. According to Rabbitts, (1994).

rent effects on transcription activity of the corresponding genes. The c-Myc-MAX dimer activates transcription while the MAX-MAD or MAX-Mxi-1 dimers do not show transcription-activating activity. In a normal cell, there is a balanced equilibrium between the different dimers. This equilibrium is disturbed by the chromosomal translocation and the associated overexpression of the Myc protein. It is assumed that there is an excessive transcription of target genes due to overexpression of the Myc protein, leading to tumor formation.

An example of an oncogene derived from a nuclear receptor is the *v-erbA* gene. The cellular counterpart of the *v-erbA* gene is the gene for the receptor for $T_3$ hormone ($T_3$ receptor, $T_3R$). The viral mutations of the $T_3$ receptor are characterized by loss of the binding domain for $T_3$ hormone. The intact $T_3$ receptor obtains its transcription-activating function on binding the ligand, the $T_3$ hormone. In the absence of the ligand, the $T_3$ receptor represses gene expression (see 4.5). Due to the loss of the ligand binding domain, the mutated oncoprotein has a dominant repressing effect on the expression of genes regulated by the $T_3$ hormone.

Possibly, the constitutive repression of target genes of the $T_3$ receptor brings about a switch from a differentiation pathway to unprogrammed cell division.

## 14.4 Tumor Suppressor Genes

### 14.4.1 General Functions of Tumor Suppressor Genes

Tumor suppressor genes are genes that, by their inactivation due to mutations or deletion, promote tumor formation. The proteins for which they code are known as *tumor suppressor proteins.* Many of the known tumor suppressor proteins have a suppressing and negatively regulating effect on processes that are either directly associated with regulation of cell division or influence this in an indirect way. Other, equally important functions of tumor suppressor proteins are in the areas of DNA repair and cell adhesion.

Inactivation of tumor suppressor genes can have various consequences:

#### Loss of Negative, Suppressing Signals in Cell Division

Mitogenic signal transduction pathways and the cell cycle machinery contain a range of negative regulation elements which help to reduce or terminate a mitogenic signal (see Chapter 14). Inactivation of the suppressing function is equivalent to stimulation of cell division in many cases. Negative control of cell division activity is performed in particular by the repression of genes with products that are required for progression in the cell cycle. The retinoblastoma protein pRb is of special importance here.

#### Promotion of Accumulation of Further Mutations

A cell that has suffered DNA damage can bring about a halt in the cell cycle with the help of intrinsic control mechanisms (DNA damage checkpoints, see 13.7). The aim is

to gain time for repair of the DNA damage and to prevent cells with the DNA damage progressing in the cell cycle. Failure of the regulatory functions that couple progression in the cell cycle to integrity of the DNA will favor the establishment of DNA damage as a mutation. The mechanism by which coupling is achieved between DNA damage and progression of the cell cycle is only just starting to be understood. The tumor suppressor protein p53 has a central function at DNA damage checkpoints (see 14.4.5.4).

#### Failure of Apoptotic Signals

Apoptosis or programmed cell death is a program that brings about the death of the cell in a targeted manner in the presence of DNA damage (see Chapter 15). Apoptosis is a protection against formation of tumor cells. If a cell is affected by DNA damage, apoptosis can help to initiate cell death before further mutations accumulate that would favor transition to the tumor state. The tumor suppressor protein p53 also plays an important role here.

Based on the functions in tumor formation, the tumor suppressor proteins can be roughly divided into two classes (Kinzler and Vogelstein, 1998):

#### „Gatekeeper“

This type of tumor suppressor protein directly prevents growth of tumors by inhibiting cell growth or promoting apoptosis. Each cell type has only a few gatekeepers and inactivation of these genes leads directly to neoplastic growth. These genes normally function as gatekeepers to prevent uncoordinated growth. Important gatekeepers are the pRb protein, the p53 protein and the cell cycle inhibitor $p16^{ink4a}$. Inactivation of these genes is rate determining for tumor formation and both the maternal and paternal copies of the gene must be inactivated for tumor formation.

#### „Caretaker“

Genes and proteins known as caretakers have an indirect influence on tumor formation. These are susceptibility genes that indirectly suppress tumor formation. An important class of caretakers includes repair proteins. Inactivation of a caretaker gene of this class leads to a sharp increase in mutation rate and is therefore equivalent to constant exposure to mutagens.

### 14.4.2 DNA Repair, DNA Integrity and Tumor Suppression

Each cell has a range of protection mechanisms to avoid DNA damage and to ensure integrity and stability of the genome (see also 14.1.5). The following caretaker functions are associated with tumor creation and tumor progression.

**Mismatch Repair, hMSH2**

For inherited forms of a certain form of bowel cancer (*hereditary nonpolyposis cancer*, HNPPC), it has been observed that there is an error in function of the repair system for DNA mismatches. Patients with HNPPC have inherited a defect in the hMSH2 gene in their germ cells and their tumor cells have a further mutation in the hMSH2 gene. The hMSH2 gene is a homolog of the MutS gene in *E. coli*, which is involved in repair of DNA mismatches in bacteria. The defect in the mismatch repair is responsible for the genetic instability of this type of tumor cell.

**BRCA1 and BRCA2**

The BRCA1 and BRCA2 genes mediate a hereditary susceptibility for breast cancer. Both gene products bind to the human Rad51 protein which is involved in DNA recombination and DNA repair. The BRCA1 gene has been shown to be involved in the transcriptional control of genes that participate directly or indirectly in the repair of oxidative DNA damage (Gowen et al., 1998).

**Defects in Metaphase-anaphase Checkpoint**

The metaphase-anaphase checkpoint controls the formation of the spindle apparatus and the correct alignment of chromosomes, and may initiate metaphase arrest (see Chapter 13). On failure of the control system, the occurrence of abnormal chromosomes is favored. In various cancer cells, chromosomal instability is associated with loss of function of the BUB1 gene, which is part of the metaphase-anaphase checkpoint (Cahill et al., 1998).

It is becoming increasingly obvious that errors in function of cell cycle checkpoints correlate with increased chromosomal instability of tumor cells (Lengauer et al., 1998).

Other proteins that have a caretaker function in tumor progression are the proteins of nucleotide excision repair and the ATM kinase (see 14.4.5.4).

## 14.4.3 The Retinoblastoma Protein pRb as a Tumor Suppressor Protein

The retinoblastoma gene was the first tumor suppressor gene to be identified and characterized in man. Human genetic investigations of patients suffering from the rare retinoblastoma eye tumor showed a defect in a gene sequence known as the retinoblastoma gene. In the inherited form of this tumor, which appears in childhood, a defect is inherited in the retinoblastoma gene via the germ line. Inactivation of the second allele of the retinoblastoma gene by mutation or deletion leads to complete failure of the function of the gene and thus to tumor formation. Inactivation of the retinoblastoma gene is also observed in many other more frequently occurring tumors.

The product of the retinoblastoma gene, the retinoblastoma protein pRb, functions as a switch for conversion of mitogenic and antimitogenic signals at the transcription

level in the cell cycle (see 14.4.3). The switch function is performed with the help of the phosphorylation status of pRb (see Fig. 13.14). In the hypophosphorylated state, and in cooperation with the transcription factor E2F, pRb represses the expression of genes with products needed for progression through the cell cycle. In the hyperphosphorylated state, in contrast, the pRb protein has an activating effect on these genes. Switching between the two states coincides with crossing of the restriction point, and from this point, further progress in the cell cycle is independent of the effect of growth factors.

The regulatory network, of which the pRb protein is a part, can be changed by various mechanisms in the sense of stimulation of proliferation (Fig. 14.6). First, an error in the function of proteins involved in phosphorylation of pRb may initiate a dominating activation of the pRb function. Second, the pRb protein itself may be activated in an unprogrammed manner by mutation or by binding of viral proteins.

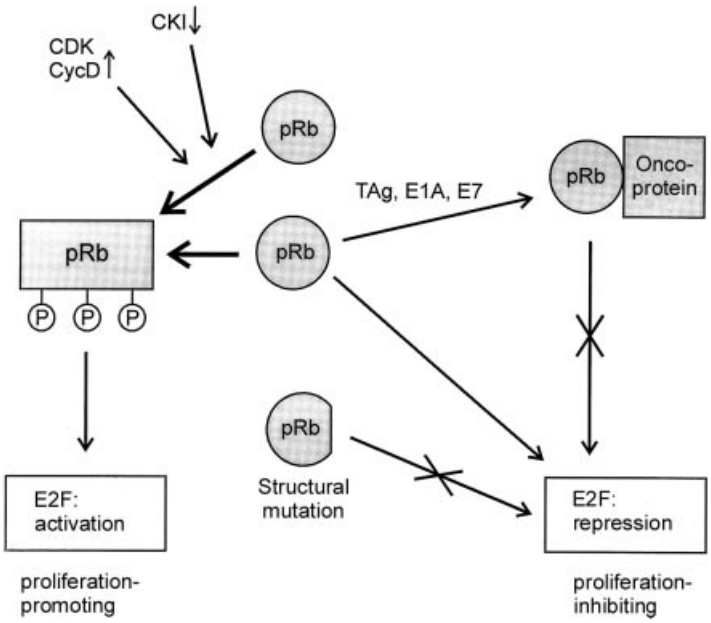



**Fig. 14.6.** Errors in regulation of the tumor suppressor protein pRb. The figure shows a simplified version of well characterized mechanisms by which errors in regulations of the pRb function can occur. The proliferation-promoting activity of pRb is amplified if unprogrammed high activity of the CDK4-cyclin D1 complex is present or the inhibitory activity of CKI is absent. Lack of proliferation-inhibiting activity of the pRb protein can be due to complexation of the pRb protein with viral oncoproteins or due to structural mutation of the pRb protein. The transcription factor E2F no longer has a repressing activity on transcription of cell cycle genes in this situation (see 13.4.3) and the division-inhibiting activity of the pRb protein is absent.

### Errors in Regulation of the Phosphorylation Status of pRb

Phosphorylation of pRb is mainly controlled by the system of G1-specific CDKs, the D and E type cyclins and the CKIs, with the cyclin/CKI relationship being of particular importance (see Chapter 14). An error in function of the proteins involved can override the growth-controlling effect of the pRb control system and promote uncontrolled cell division.

### Cyclins and Cyclin-dependent Protein Kinases (CDKs)

An uncontrolled increase in the concentration and activity of the cyclins D1 and CDK4 can lead to unprogrammed phosphorylation of the pRb protein. Overexpression of cyclin D1 or CDK4 has been found in various tumors of the bladder, breast, pancreas and the lungs (Weintraub, 1995).

### Inactivation of Inhibitors of CDKs (CKIs)

CKIs inhibit the phosphorylation of pRb and bring about a halt in the G1 phase. Of the various CKIs, mutations in the inhibitor $p27^{Kip1}$ and, in particular, the $p16^{ink4a}$ protein (see 14.4.4), are associated with tumor formation. Both proteins are tumor suppressor genes.

### Binding of Viral Oncoproteins

DNA viruses that can trigger tumors are found in the classes of the polyomaviruses, the adenoviruses and the papilloma viruses. The polyoma viruses with the SV40 virus as a well studied representative, adenoma virus and human papilloma virus (HPV) are associated with formation of tumors in humans and have genes coding for proteins with the properties of oncoproteins. The oncoproteins of all three viruses interfere with the pRb function by lifting its inhibition of transcription factor E2F. It is assumed that the tumor-promoting activity of the proteins is due, in particular, to this property.

The oncoproteins are the T antigen (TAg) of SV40 virus, E1A protein of adenovirus and E7 protein of HPV. The three proteins have in common the ability to bind to the hypophosphorylated form of pRb. In all three cases, binding takes place in the same region of pRb referred to as the „pocket“. The transcription factor E2F also binds in the region of the „pocket“ of pRb protein so that competition occurs between the viral oncoproteins and transcription factor E2F for the pocket binding site of pRb. Under these conditions, transcription factor E2F can be released from the inhibition of pRb and can activate its target genes. It is assumed that this mechanism withdraws the transcription-regulating activity of E2F from cellular control so that there is unprogrammed and unregulated expression of important cell cycle genes.

### Genetic Inactivation of pRb

Genetic inactivation of pRb is observed in many tumors. The gene defect may affect the promoter region of the pRb gene, leading to reduced pRb expression, or it may

affect the structure of pRb, for example, by a mutation of the binding site for E2F. The mutations observed in tumors are generally extensive structural changes in the pRb gene.

### 14.4.4 The $p16^{ink4a}$ Gene Locus and Tumor Suppression

The CKIs are important control elements that enable the cell to bring about a halt in the cell cycle in the presence of DNA damage, in the absence of mitogenic signals or during an excess of antimitogenic signals. If the inhibitor function fails, mitogenic signals predominate and cell division takes place in an uncontrolled manner. The *in vivo* importance of CKIs for tumor formation is not easy to estimate. This is illustrated by the example of the $p21^{CIP1}$ inhibitor. Mice with an inactivated $p21^{CIP1}$ inhibitor develop fairly normally. Apparently, the safety mechanisms of the cell are redundant and if one inhibitor fails, its function is taken over by other inhibitors.

Of the various CKIs, $p16^{ink4a}$ has a central role in regulation of the phosphorylation status of pRb and in the G1/S transition (review: Sherr, 1996). Investigations in tumor cells indicate that proteins coded by the p16 gene locus function as tumor suppressors and are of great importance for tumor development (review: Chin et al., 1998).

In many tumors, such as lung carcinomas, inactivation of the gene for the $p16^{ink4a}$ inhibitors has been observed, based on mutations or on an aberrant C-methylation. It was then shown, surprisingly, that two proteins are coded by the $p16^{ink4a}$ gene locus, the $p16^{ink4a}$ protein and another protein, the $p19^{ARF}$ protein (ARF = alternative reading frame). Both proteins have a growth-inhibiting function, however with different points of attack. Whilst the $p16^{ink4a}$ protein inhibits the cyclin D-CDK complex and brings about a halt in the cell cycle via the pRb protein, the $p19^{ARF}$ protein attacks the function of the p53 protein by specifically interacting with the MDM2 protein and interfering with its binding to the p53 protein (see 14.4.5.4, Fig. 14.10).

The $p19^{ARF}$ protein is not homologous to the $p16^{ink4a}$ protein although both originate from the same gene locus. It arises by alternative splicing and by use of a different reading frame.

### 14.4.5 The Tumor Suppressor Protein p53

The most frequently observed genetic changes in human tumors affect the gene for a nuclear phosphoprotein of 393 amino acids, which is known as the p53 protein, after its molecular weight.

Mutations of the p53 gene are observed in over 50 % of all human tumors. Defects in the p53 gene in the germ line lead to a hereditary tendency to develop various tumors, especially of the connective tissue. In affected families, several members of the family may develop tumors during childhood. The disease is known as Li Fraumeni Syndrome after its discoverer.

The close correlation between tumor formation and mutations of the p53 gene indicates that the p53 protein has a central function in tumor pathogenesis. Thanks to

intensive investigations of the structure and functions of the gene and its coded protein, it is clear that the p53 protein is an important component of a regulatory network in which cell cycle control, integrity of the DNA and programmed cell death play a central role.

The following central functions may be assigned to the p53 protein (review: Ko and Prives, 1996, Agarwal et al., 1998).

- The p53 protein functions as a sequence-specific transcription activator. As such, it is able to bind specifically to DNA elements and to activate the transcription of downstream genes.
- The p53 protein is part of a control mechanism that couples progress of the cell cycle to the integrity of the DNA. In the presence of DNA damage, the p53 protein can bring about a halt in the cell cycle at the G1→S transition and at other points of the cell cycle.
- The p53 protein is involved in initiation of programmed cell death.

#### 14.4.5.1 Structure and Biochemical Properties of the p53 Protein

The p53 protein can be seen as a multi-talent amongst the regulatory proteins. As shown in Fig. 14.7, three domains can be identified in the p53 protein, and defined biochemical functions can be assigned to these.

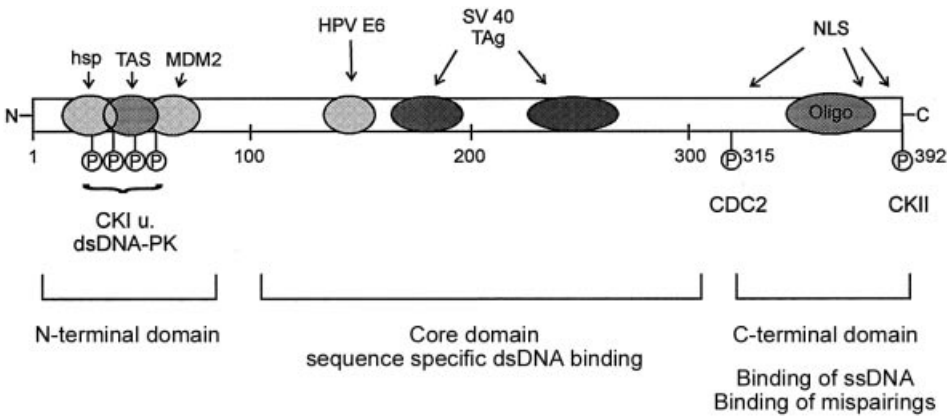



**Fig. 14.7.** Localization of biochemical functions in p53 protein. The sequence sections of the p53 protein, to which a biochemical function has been assigned, are shown in linear form. P: phosphorylation site for protein kinases; CKI: casein kinase I; CKII: casein kinase II (see 7.1.2); dsDNA-PK: double-stranded DNA-dependent Ser/Thr-specific protein kinase; Hsp: binding site for heat shock proteins; TAS: transactivating sequence; MDM2: binding site for the oncoprotein MDM2; HPV E6, SV40 TAg: binding site for the viral oncoprotein E6 of human papilloma virus (HPV) and TAg of SV40 virus; NLS: nuclear localization signal; CDC2: phosphorylation site for the CDC2 kinase; Oligo: oligomerization domain; According to Ko and Prives, (1996).

**1) Transactivation Domain**

In the N-terminal region of p53, there is a transactivation domain which p53 uses to make contact with the transcription apparatus. Different protein binding sites have been identified in this region. These include binding sites for components of the TFIID complex and for coactivators such as the CBP/p300 coactivator (see 1.4.6).

**2) Sequence-specific DNA Binding Domain**

A core domain of 100—300 amino acids includes the binding site for the corresponding DNA element and binding sites for viral oncoproteins such as the large T antigen (TAg) of SV40 virus.

**3) Nonspecific DNA Binding Domain**

The C terminus of p53 contains a domain for nonspecific binding of DNA. In this region, single-stranded and damaged DNA are bound. Furthermore, the C-terminal domain contains several phosphorylation sites, sequence signals for nuclear localization, sequence sections for tetramerization and binding sites for transcription factors. Overall, the C-terminal domain has an important function for activation and regulation of p53. There is experimental evidence that specific DNA binding of the core domain is controlled by phosphorylation of the C-terminal domain.

**Interaction of p53 with Regulatory Proteins**

The p53 protein interacts with proteins that have a regulatory function or are involved in DNA replication or DNA repair. Interactions have been demonstrated with components of the basal transcription apparatus, such as TFIIB and TFIID, with TAF proteins, and with the CBP/p300 coactivators, corresponding to the transcription-regulating function of p53. Furthermore, p53 binds to the single-stranded DNA binding protein RPA and to protein components of the TFIIH complex (see 1.4) which are involved in repair. Other binding partners of p53 are the MDM2 oncoprotein (see 14.4.4.4) and the viral oncoprotein TAg of SV40 virus.

### 14.4.5.2 Sequence-specific DNA Binding of p53

Central to the function of the p53 protein is its ability, as a transcription activator, to specifically bind to corresponding *cis* elements in the promoter region of various genes and to activate their transcription. The importance of sequence-specific DNA binding for the tumor-suppressing function of p53 became clear when the crystal structure of the complex of p53 protein and a corresponding DNA element were resolved and this structure was compared with the spectrum of known mutations of p53 protein occurring in human tumors (Cho et al., 1994).

The p53 protein is a tetrameric protein that binds to DNA elements with the consensus sequence 5'-RRRC(A/T)(T/A)GYYY-3' (R = purine, Y = pyrimidine). The struc-

ture of the complex of the central DNA binding domain of p53 with an oligonucleotide that carries one half-site of the p53 recognition sequence is shown in Fig. 14.8.

The p53 protein contacts the recognition sequence by two means. A loop-helix-loop motif is placed in the large groove of the DNA and makes contact with the bases. In addition, another loop (L3) forms a contact via an Arg side chain (R248) to the minor groove of the DNA.

The large number of known sequences of the p53 gene from tumor patients was particularly valuable for interpretation of the crystal structure since a spectrum could be assembled for p53 mutation in association with tumor formation. The mutation spectrum shown in Fig. 14.9 shows „hotspots", positions at which p53 mutations are seen particularly frequently in tumor patients.

Comparison of the *mutation spectrum* with the structure of the p53-DNA complex indicates that the positions of frequent mutations coincide with the conserved structu-

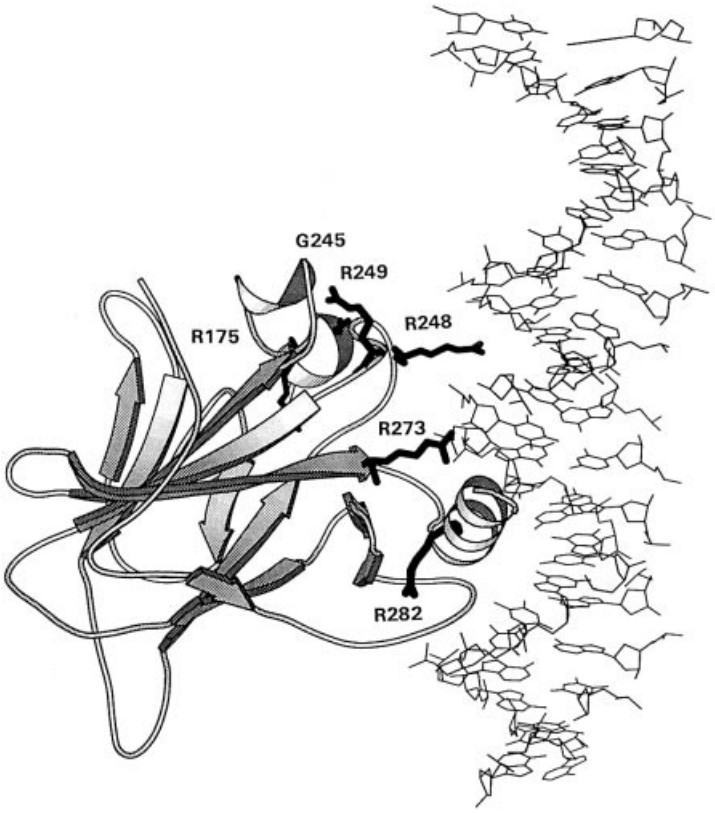



**Fig. 14.8.** DNA binding domain of the tumor suppressor protein p53 in complex with DNA. Crystal structure of the core domain of p53 (amino acids 102–292) in complex with a double-stranded DNA that contains a specific binding site for p53 (Cho et al., 1994). The amino acid positions are highlighted at which frequent oncogenic mutations are observed (see Fig. 14.9). MOLSKRIPT representation according to Kraulis, (1991).

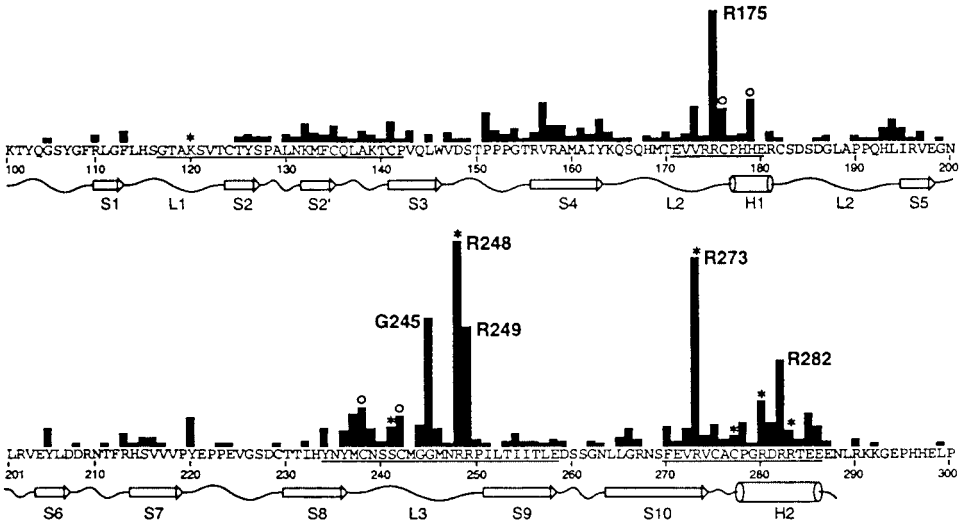



**Fig. 14.9.** Mutation spectrum of the p53 protein in tumors. The linear structure is shown of p53 and the frequency of mutations found in tumors. The black bars indicate the approximate position and the relative frequency of the p53 mutations. The frequency of mutations in the region of the DNA binding domain is of note. The sites of the most frequent mutations coincide with positions of the p53 protein that are directly involved in interactions with the DNA sequence (see Fig. 14.8). According to Cho et al., (1994), with permission.

ral elements of the DNA binding domains. It is particularly noticeable that the most frequently mutated position in tumors, namely Arg248, is also the position at which the p53 protein forms a specific contact to recognition sequences.

The naturally occurring mutations in the DNA binding domain can be divided into two classes:

One class includes mutations of amino acids that enter into direct contact with the DNA element. These include mutations at positions 248 and 175. Both positions show a high mutation frequency in naturally occurring tumors.

The other class includes mutations that are assumed to change the overall structure of the DNA binding domain of p53 so that a specific interaction is no longer possible with the DNA.

Overall, comparison of the structure data with the natural mutation spectrum of p53 shows that specific DNA binding is closely associated with the function as a tumor suppressor. A disruption of specific DNA binding apparently has serious effects on the tumor suppressing function of p53.

#### 14.4.5.3 p53-regulated Genes

The p53 protein functions as a specific transcription activator but it can also bring about a general repression of gene activity.

### 1) Activation of Specific Genes

The p53 protein stimulates transcription of various genes that carry the p53 DNA element in the promoter region. Table 14.1 shows some of the p53-activated genes.

#### Gene for CDK Inhibitor $p21^{CIP1}$

The p53 protein controls the expression of the CDK inhibitor $p21^{CIP1}$ and two copies of the p53 recognition sequence are found in its promoter region. Activation of p53 leads to increased formation of the p21 inhibitor, which brings about a halt in the cell cycle in G1 phase. The p53 protein has a negative control function according to this mechanism. Inactivation of p53 means loss of an important control element of the cell cycle.

#### Bax Gene

Another gene controlled by p53 is the gene for the Bax protein (Miyashita and Reed, 1995). The Bax protein has an activating function in the initiation of apoptosis. Activation of the p53 protein due to DNA damage can initiate the apoptotic program via stimulation of *bax* transcription (see Chapter 15). If the transcription regulating function of p53 is lost due to an oncogenic mutation, this apoptotic path cannot be initiated.

#### Thrombospodin-1 Gene

The thrombospodin-1 gene codes for a protein that inhibits new formation of blood vessels (angiogenesis). The p53 protein activates expression of the thrombospodin-1 gene and can thus suppress angiogenesis (Dameron et al., 1994). If the regulating activity of

**Table 14.1.** Examples of p53-activated genes (according to Ko and Prives, 1996).

| Gene | Function |
|---|---|
| $p21^{Waf1,Cip1}$ | Inhibitor of cyclin-dependent protein kinases (see chapter 13) |
| mdm-2 | Oncoprotein, binds to p53 und promotes its degradation (see chapter 13) |
| GADD45 | binds to the replication factor PCNA and inhibits the entry of cells into S phase |
| Cyclin G | Cyclin, induced on DNA damage, |
| bax | forms heterodimers with the Bcl-2 protein (see Chapter 15), promotes programmed cell death |
| IGF-BP3 | binds to growth factor IGF (Insulin-like growth factor) and inhibits its growth promoting function |
| Thrombospondin I | inhibits new formation of blood vessels |

the p53 protein is inactivated, a situation is created which facilitates new formation of blood vessels, since the inhibitor of angiogenesis is missing. It is assumed that this situation promotes tumor progression especially in the late phase of tumor formation.

**Redox Related Genes**

A detailed analysis of p53-activated genes has shown that this includes many genes that can generate or respond to oxidative stress (Polyak et al., 1997). The link between oxidative stress and p53 may be explained by the apoptotic effect of p53. It is plausible that formation of activated oxygen is involved in triggering of p53-mediated apoptosis.

**2) Transcription Repression**

The p53 protein has – in addition to the activation of specific genes – also a general repressing influence on transcription. The repression is observed for various cellular and viral genes that have no p53 binding site. Examples of genes repressed by p53 are the genes for transcription activators *c-jun* and *c-fos*, the cytokine IL-6, the retinoblastoma protein pRB (see Chapter 13) and the *bcl2* gene (see Chapter 15). It is assumed that the general repressing activity of p53 is based on interference in formation of the preinitiation complex of transcription and that this is associated with the binding of p53 to the TATA box protein TBP and the TFIID complex.

### 14.4.5.4 Activation, Regulation and Modulation of the Function of p53

**1) General Control Functions of p53**

The p53 protein is central to a control function that underlies progress in the cell cycle when DNA damage or other faults in the cell cycle are present. If cells are exposed to damage such as UV irradiation, an *increase in the concentration* of p53 protein is observed and the p53 protein is activated. One of the signals that has been identified as leading to activation of the p53 protein is a DNA *strand break*. Conditions that favor a strand break are the effect of UV irradiation, incomplete repair processes or a pause in DNA replication due to a damaged DNA template.

Depending on the cell type, the extent of the DNA damage and the overall situation of the cell, activation of the p53 protein can bring about two reactions:

**Halt in Cell Cycle**

Activation of p53 can bring about a *halt* in the cell cycle at the important cell cycle transitions. Thus, the p53 protein is involved in the control of the G1/S transition, the mitotic spindle checkpoint and the G2/M transition. There is increasing experimental evidence that the halt is irreversible and that the cell can survive for a very long time in this resting state.

### Initiation of Apoptosis

Another reaction following activation of p53 is the initiation of cell death, apoptosis (see 15.3.6).

### 2) Mechanism of p53 Activation

The mechanism by which DNA damage leads to a concentration increase and thus to activation of the p53 protein is only just starting to be understood. It is generally assumed that regulation of p53 activity takes place at the post-translational level. According to this theory, the p53 protein is in an inactive, latent form in the normal situation of the cell. Effector signals such as DNA damage, arrest of DNA or RNA synthesis or nucleotide depletion lead to activation of p53 protein. There is an increase in the concentration of p53 and a change in the biochemical properties of the p53 protein. At least a part of the concentration change is due to increased stability of the p53 protein. The changes in the biochemical properties are due to post-translational modifications such as phosphorylation, acetylation, to alternative splicing and interactions with other regulatory proteins. The following mechanisms of activation of p53 are under discussion:

### Activation at the DNA Damage Checkpoint: the ATM Protein

The p53 protein has been identified as a component of the DNA damage checkpoint in animals and humans, from which it is assumed to be homologous to the DNA damage checkpoint in the yeast *S. cerevisiae* (see 13.7).

The ATM protein has been identified as an important member of a reaction chain that leads from detection of DNA damage to activation of the p53 protein. Mutations of the ATM protein are causally associated with the disease ataxia telangiectasia, thus the name ATM (ataxia telangiectasia mutated). The ATM protein has protein kinase activity and is counted as a member of the PI3-kinase family, due to sequence homologies (review: Canman et al., 1998). The p53 protein is phosphorylated at Ser15 by ATM kinase (Canman et al., 1998) and it is assumed that this phosphorylation contributes to activation of the p53 protein. The ATM protein is preceded by other protein kinases that are directly or indirectly activated by DNA damage and pass this signal on to the p53 protein via the ATM protein.

### Binding of the Oncoprotein MDM2

The MDM2 protein is recognized as another important control element of p53 function. The MDM protein was first identified as an oncoprotein that negatively regulates p53 function. The regulatory function is performed within a network that includes the inhibitor $p19^{Arf}$ (Fig. 14.10).

The MDM2 protein specifically binds to the p53 protein and thus leads to its inactivation. There is experimental evidence that the MDM2 protein promotes proteolytic degradation of p53 (Haupt and Oren, 1997). There is also a link between the p53 pro-

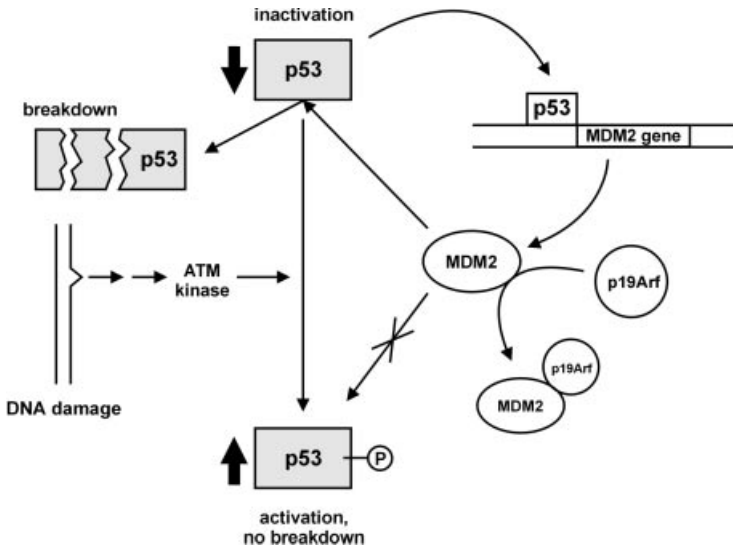



**Fig. 14.10.** Model of the regulation of p53 concentration by DNA damage and by MDM2 and $p19^{ARF}$

The level of p53 is strongly regulated by its interaction with MDM2. Binding of MDM2 to p53 targets p53 for proteolytic degradation thus keeping p53 concentration low. A negative feedback loop exists between p53 and MDM2 since p53 controls the expression of MDM2 at the level of transcription. Another control is exerted by the tumor suppressor $p19^{ARF}$, which binds to MDM2 and sequesters it from the feedback loop thus helping to increase p53 concentration. Upon DNA damage the ATM kinase is activated and phosphorylates p53 at Ser15. The phosphorylated p53 is no longer bound by MDM2 and proteolysis is decreased. The level of p53 now increases and p53 can exert its tumor suppressing effects.

tein and the MDM2 protein at the transcription level, since transcription of the *mdm2* gene is activated by the p53 protein. A negative feedback loop is therefore postulated for the MDM2 protein, which contributes to keeping the p53 concentration low in the normal situation. An important part of this feedback loop is the inhibitor $p19^{Arf}$, which binds to the MDM2 protein and inhibits its control function for the p53 protein.

It is not clear how this network is reprogrammed on activation of the p53 protein. Possibly, the DNA-damage-induced phosphorylation of p53 interferes with its interaction with the MDM2 protein, which leads to attenuation of the negative feedback loop.

The regulatory network of the MDM2-p53 interaction highlights that loss of the p53 function can occur by several pathways:

- Mutation of the p53 protein: direct loss of transcription-activating function of p53
- Inactivation of $p19^{Arf}$: loss of negative control of the MDM2 protein and an associated increased destabilization of p53
- Unprogrammed activation of the MDM2 protein: amplification and overexpression of the MDM2 protein is observed in various tumor types and is often associated with a poor prognosis

It is also important that the MDM2 protein has a regulating influence on the pRb protein (see 14.4.4) so that the MDM2 protein regulates two important cell cycle control proteins.

**Post-translational Modification of p53: Phosphorylation and Acetylation**

The p53 protein has various Ser/Thr phosphorylation sites. These include C-terminal phosphorylation sites for the CDKs of the S and $G_2$/M phase. A controlling effect on DNA binding ability has been described for C-terminal phosphorylation so that a direct connection between the cell cycle machinery and the transcription-activating function of p53 seems possible via this route.

A further post-translational modification of p53 is an acetylation at the C terminus, for which stimulation of specific DNA binding has been reported (Wei and Roeder, 1997). Possibly, this acetylation is mediated by the CBP/p300 proteins (see 1.4.6), with which the p53 protein can specifically interact.

**Activation of DNA Binding**

There is an unspecific binding site for double- and single-stranded DNA in the C-terminal region of the p53 protein (Lee et al., 1995, Jayaraman and Prives, 1995). There is evidence that binding of single-stranded DNA regulates promoter selection of p53. It is also plausible that stabilization of the p53 protein is achieved via DNA binding.

**Binding of Viral Oncoproteins**

The oncoprotein of SV40, TAg (see 14.4.3), binds to the p53 protein and can inactivate the p53 function in a similar way to that assumed for inactivation of the pRb protein.

The functional consequences and the biochemical details of the various regulating influences that meet at the p53 protein have only been partially characterized so far. The nature of the integration of p53 into the regulatory network of cell cycle control, DNA damage, DNA repair and apoptosis is unknown in many aspects. The data available at present do, however, permit a general model to be proposed for the function of the p53 protein in tumor formation (Fig. 14.11).

### 14.4.5.5 Model of p53 Function

The p53 protein represents a control element that links progress in the cell cycle and survival of a cell with the effect of genotoxic stress (especially DNA damage).

Reaction to signals originating from various DNA-damaging influences lead to accumulation, modification and activation of the p53 protein. The mechanisms of activation of p53 are diverse and involve many proteins. An important control element that links DNA damage with p53 activation is the ATM protein. MDM2 protein negatively controls p53 activity: MDM2 protein is part of an autoregulatory loop and promotes degradation of p53. The functions of the p53 protein in this network can be inf-

luenced by various processes such as phosphorylation, binding of cellular regulatory proteins and oncoproteins.

As a reaction to DNA damage, the activated p53 protein may act at several points in the course of the cell cycle and bring this to a halt. The halt at the G1/S phase transition is mediated by the inhibitor $p21^{CIP1}$ in particular. A stop in the cell cycle at the G2/M transition and in mitosis may also be mediated by activated p53.

The second important function of the p53 protein is initiation of apoptosis. This pathway involves induction of the Bax protein, which is a part of the apoptotic program of the cell.

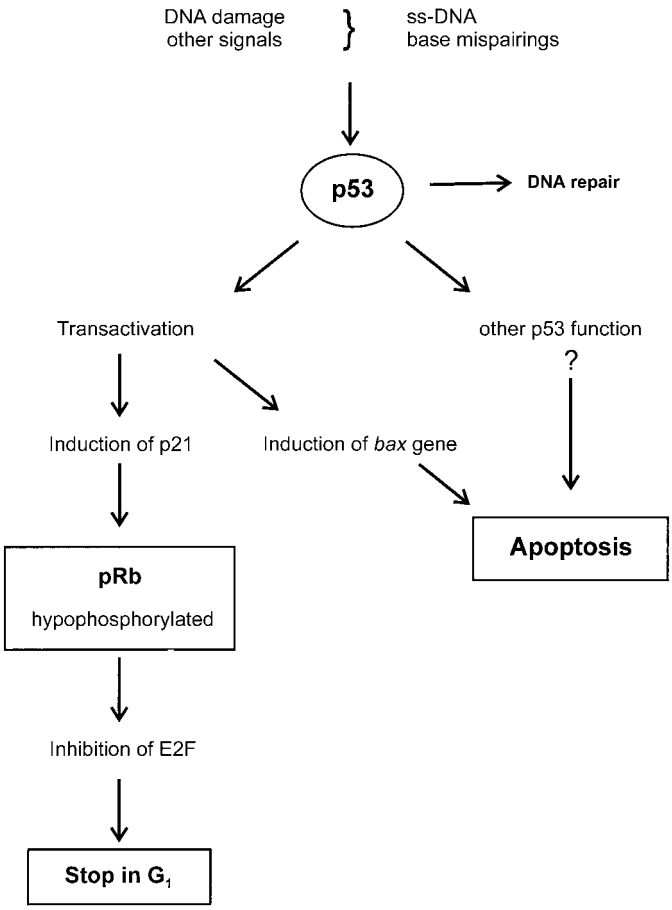



**Fig. 14.11.** Model of the function of the p53 protein. The figure summarizes in a very simplified manner two important functions of the p53 protein, which are assumed to be of importance for the tumor-suppressive activity of p53. The p53 protein is activated by DNA damage and other signals and can either bring about a halt in the cell cycle or initiate apoptosis of the cell. Activation of apoptosis can also be triggered via pathways other than the *bax* gene. The figure does not take into consideration the many other biochemical functions of p53, which can also be linked to the two pathways shown.

The intact network of p53 regulation helps to maintain the genomic integrity of the cell by eliminating damaged cells. This may take place by permanent arrest in the G1 phase, at the G2/M transition and in mitosis or by initiation of apoptosis.

Inactivation of the p53 network enables the cell to continue in the cell cycle with damaged DNA, yet without any DNA repair taking place. Furthermore, failure of the apoptotic control function permits survival of cells with damaged DNA. Both effects lead to increased susceptibility of the genome for accumulation of further mutations. The cells can also divide under conditions in which serious changes of the genome are present, such as DNA amplification and chromosome rearrangement (see Kinzler and Vogelstein, 1996). Failure of the p53 function cancels a central control element that ensures the integrity of the genome. Therefore, p53 has been ascribed the function of a '*guardian of the genome*'.

### 14.4.6 Other Tumor Suppressor Genes

A selection of other tumor suppressor genes is summarized in Table 14.2. Interestingly, an enzyme of phosphatidyl-inositol metabolism has been also identified as a tumor suppressor. The PTEN tumor suppressor gene codes for a phospholipid phosphatase which specifically cleaves a phosphate from the second messenger phosphatidyl-inositol-3,4,5-trisphosphate ($PtdInsP_3$, see 6.6.2). and thus inactivates the messenger (review: Maehama and Dixon, 1999) .

A number of tumor suppressor genes are known with no direct relationship to the regulation of the cell cycle. Some of the tumor suppressor genes in Table 14.2 are involved in the organization of the cytoskeleton or in cell-cell interactions.

**Table 14.2.** Characteristics of some tumor suppressor proteins.

| Gene | Protein | Function |
|---|---|---|
| DCC<br>'Deleted in Colon Carcinoma' | Transmembraneprotein | Cell adhesion |
| Wt-1<br>Wilms Tumor | Transcription factor with Zn-binding motif | Transcription regulation |
| APC<br>'Adenomatous polyposis coli' | binds to the cell adhesion protein ß-Catenin | Regulation of cell-cell adhesion |
| PTEN | Phosphatidylinositol-phosphate phosphatase | blocking of signaling function of PI3-kinase |
| NF-2<br>'Neurofibromatosis' | NF-2 gene product = 'merlin' | Cytoskeleton association |

# References Chapter 14

Agarwal, M.L., Taylor, W.R., Chernov, M.V., Chernova, O.B. and Stark, G.R. (1998) ‚The p53 network'. J Biol Chem *273*, 1–4

Baserga, R. 'Oncogenes and the strategy of growth factors' (1994) Cell *79*, 027–930

Cahill, D.P., Lengauer ‚C,. Yu, J., Riggins, G.J., Willson, J.K., Markowitz, S.D., Kinzler, K.W. and Vogelstein, B. 'Mutations of mitotic checkpoint genes in human cancers' (1998) Nature *392*, 300–303

Canman, C.E. and Lim, D.S. 'The role of ATM in DNA damage responses and cancer' (1998) Oncogene *17*, 3301–3308

Canman, C.E., Lim, D.S., Cimprich, K.A., Taya, Y., Tamai, K., Sakaguchi, K., Appella, E., Kastan, M.B. and Siliciano, J.D. 'Activation of the ATM kinase by ionizing radiation and phosphorylation of p53' (1998) Science *281*, 1677–1679

Chin, L., Pomerantz, J. and DePinho, R.A. 'The INK4a/ARF tumor suppressor: one gene–two products–two pathways' (1998) Trends Biochem Sci *23*, 291–296

Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. 'Crystal structure of a p53 tumor suppressor-DNA complex: understandung tumorigenic mutations' (1994) Science *265*, 346–355

Cohen, S. and Ellwein, L. 'Cell proliferation in carcinogenesis' (1990) Science *249*, 1007–1011

Counts, J.L. and Goodman, J.I. 'Alterations in DNA methylation may play a variety of roles in carcinogenesis' (1995) Cell *83*, 13–15

Dameron, K.M., Volpert, O.V., Tainsky, M.A. and Bouck, N. 'Control of angiogenesis in fibroblasts by p53 regulation of thrombospodin-1' (1994) Science *265*, 1582–1584

Fero, M.L., Randel, E., Gurley, K.E., Roberts, J,M. and Kemp, C.J. 'The murine gene p27Kip1 is haplo-insufficient for tumour suppression' (1998) Nature 396, 177–180

Gowen, L.C., Avrutskaya, A.V., Latour, A.M., Koller, B.H. and Leadon, S.A. 'BRCA1 required for transcription-coupled repair of oxidative DNA damage' (1998) Science *281*, 1009–1012

Gu, W. and Roeder, R.G. 'Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain' (1997) Cell *90*, 595–606

Haffner, R. and Oren, M. 'Biochemical properties and biological effects of p53' (1995) Curr. Op. Gen. Dev. *5*, 91–96

Hartwell, L.H. and Kastan, M.B. 'Cell cycle control and cancer' (1994) Science *266*, 1821–1828

Haupt, Y., Maya, R., Kazaz, A. and Oren, M. 'Mdm2 promotes the rapid degradation of p53' (1997) Nature *387*, 296–299

Hunter, T. 'Cooperation between oncogenes' (1991) Cell *64*, 249–270

Hunter,T. and Pines, J. 'Cyclins and cancer II: Cyclin D and CDK inhibitors come of age' (1994) Cell 79, 573–582

Jayaraman, L. and Prives, C. 'Activation of p53 sequence-specific DNA binding by short single strands of DNA requires the p53 C-terminus' (1995) Cell *81*, 1021–1029

Kinzler, K.W. and Vogelstein, B. 'Life (and death) in a malignant tumor' (1996) Nature *379*, 19–20

Kinzler, K.W. and Vogelstein, B. 'Landscaping the cancer terrain' (1998) Science *280*, 1036–1037

Ko, L.J. and Prives, C. 'p53: puzzle and paradigm' (1996) Genes Devel. *10*, 1054–1072

Kraulis, P.J. 'MOLSKRIPT: A program to produce both detailed and schematic plots of protein structures' (1991) J. Appl. Crystallogr. *24*, 946–950

Landis, C.A., Masters, S.B., Spada, A., Pace, A.M., Bourne, H.R. and Vallar, L. 'GTPase inhibiting mutations activate the $\alpha$ chain of $G_s$ and stimulate adenylyl cyclase in human pituitary tumours' (1989) Nature *340*, 292–696

Lengauer, C., Kinzler, K.W. and Vogelstein, B. 'Genetic instabilities in human cancers' (1998) Nature 396, 643–649

Loeb, L.A. 'Mutator phenotype may be required for multistage carcinogenesis' (1991) Cancer Res *51*, 3075–3079

Loeb, L.A. 'Cancer cells exhibit a mutator phenotype' (1998) Adv Cancer Res *72*, 25–56

Look, A.T. 'Oncogenic transcription factors in the human acute leukemias' (1997) Science *278*, 1059–1064

Maehama, T. and Dixon, J.E. 'PTEN: a tumour suppressor that functions as a phospholipid phosphatase' (1999) Trends Cell Biol *9*, 125–128

Malkin, D. 'Germline p53 mutations and heritable cancer' (1994) Annu. Rev. Gen. *28*, 443–465

Marshal, C.J. 'Tumor suppressor genes' (1991) Cell *64*, 313–326

Miyashita, T. and Reed, J.C. 'Tumor suppressor p53 is a direct transcriptional activator of the human bax gene' (1995) Cell *80*, 293–299

Lee,S., Elenbaas,B., Levine, A. and Griffith, J. 'p53 and its 14 kDa C-terminal domain recognize primary DNA damage in the form of insertion/deletion mismatches' (1995) Cell *81*, 1013–1020

Pendergast, A.M., Quilliam, L.A., Cripe, L.D., Bassing, C.H., Dai, Z., Li, N., Batzer, A., Rabun, K.M., Ger, C.J., Schlessinger, J. and Gishizky, M. 'BCR-Abl-induced oncogenesis is mediated by direct interaction with the SH2 domain of the GRB-2 adaptor protein' (1993) Cell *75*, 175–185

Polyak, K., Xia, Y., Zweier, J.L., Kinzler, K.W. and Vogelstein, B. 'A model for p53-induced apoptosis' (1997) Nature *389*, 300–305

Rabbitts, T.H. 'Chromosomal translocations in human cancer' (1994) Nature *372*, 143–149

Vogelstein, B. and Kinzler, K.W. 'The multistep nature of cancer' (1993) Trends Gen. *9*, 138–141

Wang. X.W., Yeh, H., Schaeffer, L., Roy, R., Moncollin, V., Egly, J.M., Wang, Z., Friedberg, E.C., Evans, M.K., Taffe, B.G., Bohr, V.A., Weeda, G., Hoeijmakers, J.H.J., Forrester, K. and Harris, C.C. 'p53 modulation of TFIIH-associated nucleotide excision repair activity' (1995) Nature Gen. *10*, 188–194

Weinberg, R.A. 'The retinoblastoma protein and cell cycle control' (1995) Cell *81*, 323–330

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

## Chapter 15

# Apoptosis

Apoptosis is a naturally occurring process by which a cell is directed to programmed death. The name apoptosis was coined following investigations of the nematode *C. elegans* and is of Greek origin, describing falling of leaves. The course of apoptosis is accompanied by characteristic changes in cell morphology, shown schematically in Fig. 15.1. Condensation of the chromatin, degradation of DNA, cell shrinkage, fragmentation of the cell nucleus and disassembly into membrane-enclosed apoptotic vesicles are characteristics that clearly distinguish apoptosis from another form of cell death known as necrotic cell death.

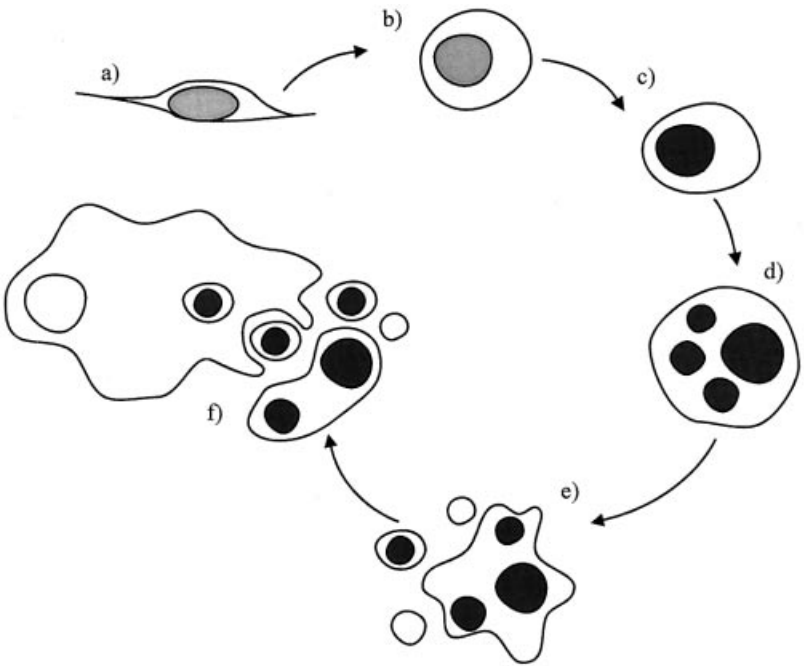



**Fig. 15.1.** Schematic representation of morphologic changes in a cell during apoptosis. On reception of an apoptotic signal, an adhesive cell (a) begins to become rounded (b) and the nuclear DNA condenses (c). The DNA is fragmented and the nucleus begins to break down into discrete chromatin bodies (d). Finally, the cell disintegrates into several vesicles (apoptotic bodies) (e), which are phagocytozed by neighboring cells (f).

## 15.1 Basic Functions of Apoptosis

Apoptosis is based on a genetic program that is an indispensable part of the development and function of an organism. It serves to eliminate undesired or superfluous cells in a targeted manner. The conditions under which the apoptotic program is activated are very diverse.

- **Tissue Homeostasis**

Apoptosis is attributed a central importance in homeostasis of tissues: in an organ or a tissue, the cell number must be kept constant within narrow limits. An increase in cells due to cell division is compensated by processes to eliminate cells that are no longer functional or are old. Apoptosis is a process that helps to keep the cell number in a tissue within limits that are suitable for the development and function of the organism. If defects occur in the apoptotic program, the consequence may be a pathologic increase or decrease in cells (Fig. 15.2). Examples of diseases associated with an increased rate of cell survival are cancer and autoimmune diseases. Diseases associated with increased apoptosis include AIDS and neurodegenerative diseases (Thompson, 1995).

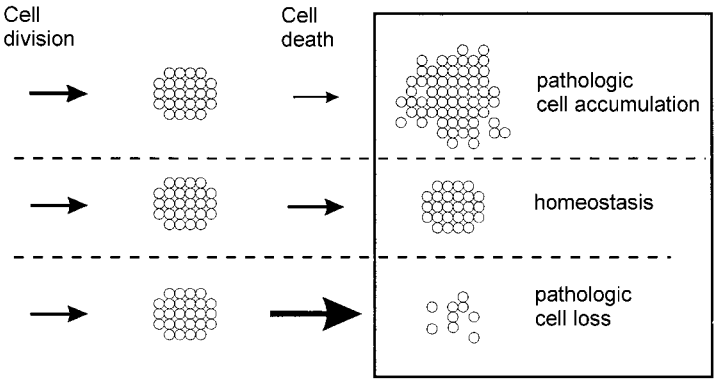



**Fig. 15.2.** Influence of apoptosis on homeostasis of a cell grouping. In a grown organism, the cell number in a tissue is determined by the relation between the rate of cell division and cell death. The rates of both processes are represented in the figure by the size of the arrow. In a normal tissue, the cell number remains constant (homeostasis) since both processes occur at the same rate. If the rate of cell proliferation predominates, diseases occur characterized by increased cell number (e.g., in tumors). In the reverse case, when the rate of cell death predominates, the cell number is reduced in a pathologic fashion. In the absence of compensatory changes in the cell division rate, changes in the extent of apoptosis can lead to either accumulation of cells or loss of cells. According to Thompson, (1995).

- **Development and Differentiation**

Apotosis has an indispensable role in development and differentiation processes (review: Vaux and Korsmeyer, 1999), especially in the embryo. Here, it provides a means to switch off cells no longer needed during embryonal morphogenesis and synaptogenesis.

- **Immune System**

In the immune system, apoptotic programs are activated in various situations. Examples include:

– Elimination of target cells (e.g., virus-infected cells) by cytotoxic T lymphocytes
– Elimination of autoreactive B- or T-lymphocytes, natural selection and elimination of cells in the thymus: 95 % of T cells that migrate to the thymus are eliminated by apoptotis.

- **Cell Damage**

Another function of apoptosis is the destruction of damaged cells. The apoptotic program may be activated in the presence of cell damage or during stress (review: Evan and Littlewood, 1998). Cells with damaged DNA can be eliminated with the help of apoptotic programs before they have the chance to accumulate mutations and possibly degenerate to a tumor cell.

## 15.2 Apoptosis in the Nematode *Caenorhabditis elegans*

The genetic description of programmed cell death originates from observations of the nematode *Caenorhabditis elegans*. In the course of the development of *C. elegans*, a total of 1090 cells are formed of which 131 are eliminated in a targeted manner by an intrinsic apoptotic program. During apoptosis, at least 14 different genes are involved with specific roles in regulation and performance of apoptosis (review: Steller, 1995; Fig. 15.3).

Two groups of genes have special importance:

The *ced3* and *ced4* genes have a *proapoptotic* effect, i.e. promoting apoptosis. Both genes are necessary for apoptosis in *C. elegans*. Mutations in the *ced3* or *ced4* genes permit survival of cells that would be eliminated in the wild type situation by programmed cell death. The *ced3* gene codes for a protein belonging to a family of Cys proteases known as the caspases (see 15.3.1). The protein coded by *ced4* functions as an adaptor or a cofactor (see 15.3.3) needed for the apoptotic program carried out by the gene product of *ced3*.

The *ced9* gene, in contrast, has an *antiapoptotic* effect, i.e., it protects the cell from apoptosis by antagonizing the function of the *ced3*/*ced4* genes. Inactivation of the *ced9* gene by mutation leads to death of cells that would survive in the normal development program of *C. elegans*. As a consequence, the embryo dies. According to these observations, the *ced9* gene has the role of ensuring survival of cells needed for the functioning of the organism.

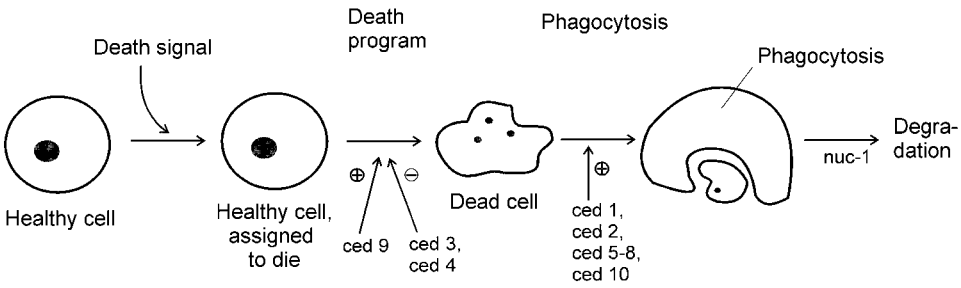



**Fig. 15.3.** Components of the apoptotic program in *C. elegans*. A simplified overview and model of the role of apoptotic genes in programmed cell death in *C. elegans*. A „death signal" is registered by a healthy cell, leading to initiation of the apoptotic program in that cell—the cell is assigned to die. The course of the death program is controlled by genes with negative and positive effects. The *ced9* gene has a positive, promoting effect on the course of apoptosis, whereas the *ced3* and *ced4* genes have a negative, inhibiting effect. Many other genes are needed for phagocytosis and further degradation of the dead cell. According to Steller, (1995).

## 15.3 Components of the Apoptotic Program in Mammals

The identification of components of the apoptotic program in *C. elegans* was the starting point for the discovery of homologous genes or proteins in mammals (review: Thompson, 1995; Green, 1998; Raff, 1998). Homologous proteins in mammals have been identified for the three central proteins of the apoptotic program in *C. elegans* and the first plausible model for the course of apoptosis has been proposed. It was also shown, however, that the apoptotic phenotype is based on a complex and versatile process. Different apoptotic signaling pathways are involved, depending on stimulus and cell type, and these meet in a common effector pathway. Central components of this effector path are a certain class of proteases, the *caspases*. The caspases mediate degradation of a number of key structural and housekeeping proteins and are the components that execute cell death.

*A larger part of the apoptotic program exists in the cell in a latent, inactive form and it only requires an apoptotic stimulus to activate the program and to initiate apoptosis.* Thus, apoptotic processes can also take place without activation of transcription. There are also forms of apoptosis that are dependent on transcription.

The execution of the apoptotic program is strictly controlled at several levels. This creates links to signaling pathways that have a control function in cell division.

### 15.3.1 Caspases: Death by Proteolysis

The link between apoptosis and proteolysis became apparent when a homology was established between the protein coded by the *ced3* gene and a previously known protease in mammals, the ICE protease (review: Martin, 1994; Kumar, 1995). The ICE

protease (interleukin-converting enzyme) is involved in maturation of interleukin 1β (to the interleukins, see 11.1). The mature 15.5 kDa interleukin 1β is formed from a 30 kDa precursor protein by specific proteolysis with the help of the ICE protease.

Based on these observations, proteases related to the ICE protease were rapidly identified in mammals and these are known today collectively as the family of caspases (review: Salvesen and Dixit, 1997; Thornberry, 1998; Earnshaw et al., 1999). The name caspase is based on particular properties of these proteases: caspases use a Cys residue as a nucleophile and cleave the substrate after an Asp residue.

### Structure and Mechanism

Like many other proteases, the caspases are formed as inactive proenzymes of 30—50 kDa and are activated by proteolytic processing. The proenzymes have a prodomain and two cleavage sites for processing, which are consensus sites for the caspases, enabling activating by autoproteolysis.

Structural studies have shown that caspases can form tetramers with two active sites. The catalytically active subunit of a caspase is made up of a large (17—12 kDa) and a small (10—13 kDa) subunit which form a heterodimer with an active site comprised of residues from both large and small subunits. Two heterodimers then align to form a tetramer with two catalytic centers (Fig. 15.4).

The cleavage mechanism of the caspases is shown schematically in Fig. 15.5. They use a typical protease mechanism with a catalytic diad for cleavage of the peptide bond. The nucleophilic thiol of an essential Cys residue forms a covalent thioacyl bond to the substrate during the catalysis. The imidazole ring of an essential histidine is also involved in catalysis and this facilitates hydrolysis of the amide bond in the sense of an acid/base catalysis.

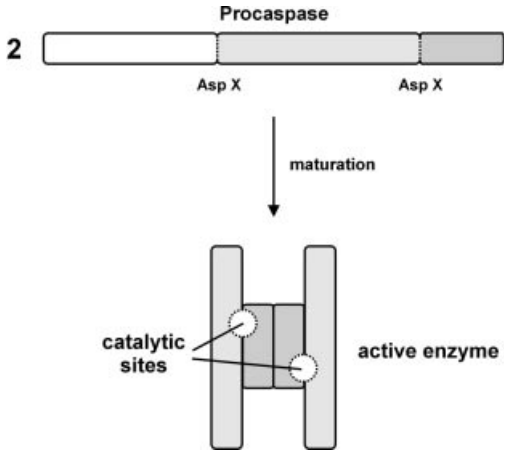



**Fig. 15.4.** Processing and subunit structure of caspases
Proteolytic processing of inactive procaspases by autoproteolysis or by other caspases includes cleavage at two caspase consensus sites (AspX). The N-terminal pro-domain is discarded and the other two fragments assemble into the active caspase which is a tetramer with two active sites.

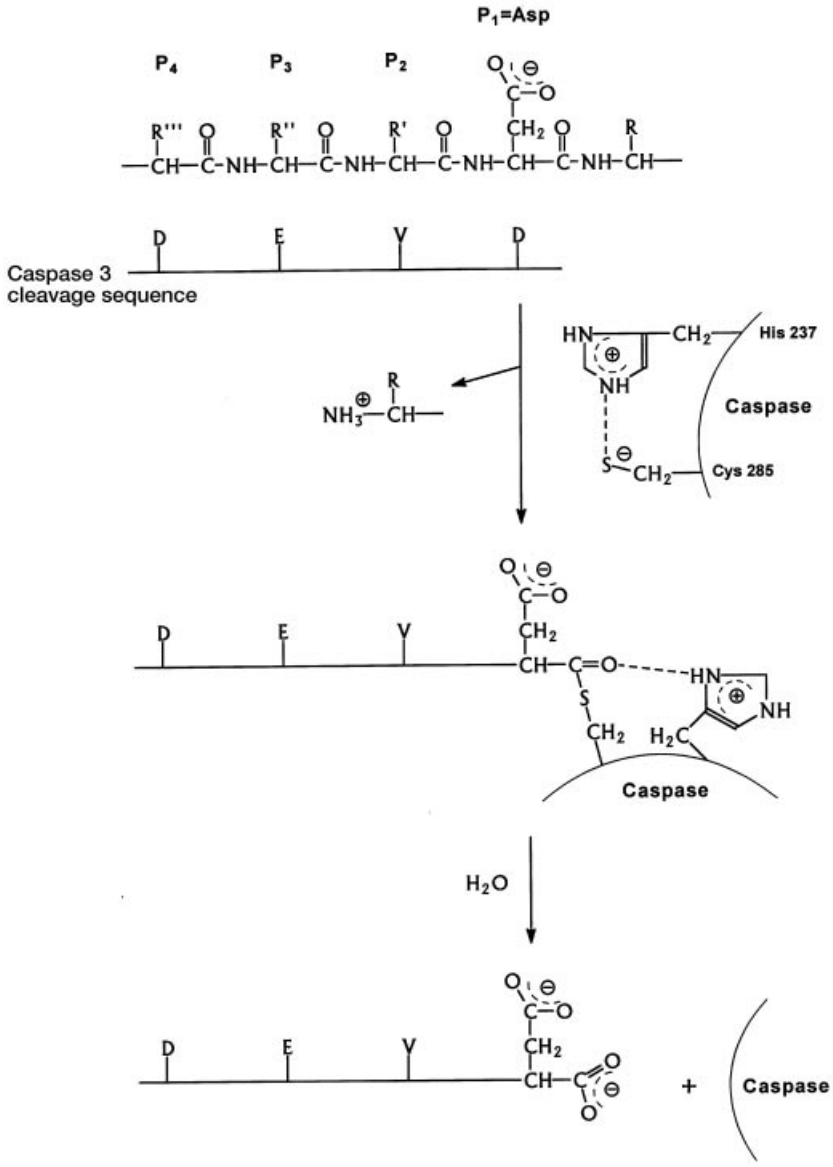



**Fig. 15.5.** Substrate recognition and postulated cleavage mechanism of caspases.
For caspase 3, four specific residues N-terminal to the cleavage site are required for cleavage in addition to the essential Asp residue. In the first step of the reaction, a covalent thioacyl-intermediate is formed between the N-terminal part of the substrate and the caspase, this is hydrolytically cleaved in the second step.

The special feature of the caspases is their high cleavage specificity. The cleavage occurs on the C terminal side of an Asp residue of the substrate protein and requires another recognition element comprised of a certain sequence of at least 3 amino acids

on the N-terminal side of the cleavage site. The various caspases differ significantly in their cleavage specificity and thus have very different proteins as substrates. At present, at least 13 different caspases have been identified, named caspase 1 to caspase 13, which participate in apoptosis or in inflammatory processes. The ICE protease is caspase 1; caspase 3 is also known as apopain or CPP32.

**Substrates**

A large number of caspase substrates have been identified, some of which have a direct relationship to survival of the cell. The following examples illustrate important cellular processes and substrates that are targeted for proteolysis during apoptosis:

Inactivation of proteins that protect from apoptosis:

- DNAase inhibitor ICAD (inhibitor of caspase-activated DNAase): Inhibitor of a DNAase responsible for DNA fragmentation (Enari et al, 1998).
- Bcl-2 protein: has a central antiapoptotic function (see 15.3.2).

Disruption of cell structures

- Lamins
- Cytoskeleton deregulation: gelsolin, focal adhesion kinase (FAK), p21-activated kinase (PAK)

Inactivation of DNA repair and DNA synthesis

- Poly ADP-ribose polymerase
- Replication factor C (RF-C)
- DNA-dependent protein kinase

**Activation and Regulation**

Unprogrammed activation of the caspases has serious consequences for the cell. Therefore, activation of caspases is strictly controlled. In the normal state of the cell, the caspases are maintained in an inactive state but can be rapidly and extensively activated by a small inducing signal. At present, not all proteins involved in the activation are known. Possibly, the activation occurs by a complex mechanism, as in blood coagulation.

Based on their function in apoptosis, the caspases are divided into two classes (see Fig. 15.6): the *initiator caspases* (caspases 8 and 9) and the *effector caspases* (caspases 3, 6 and 7).

The initiator caspases receive proapoptotic signals and initiate the activation of a caspase cascade. They are activated by an interaction with a transmembrane receptor or by cytotoxic influences. A complex is thus formed known as the *apoptosome* (see 15.4). The effector caspases are activated by an upstream caspase via a cascade mechanism. They are the component that executes apoptosis, initiates degradation of central proteins and directs the cell to death.

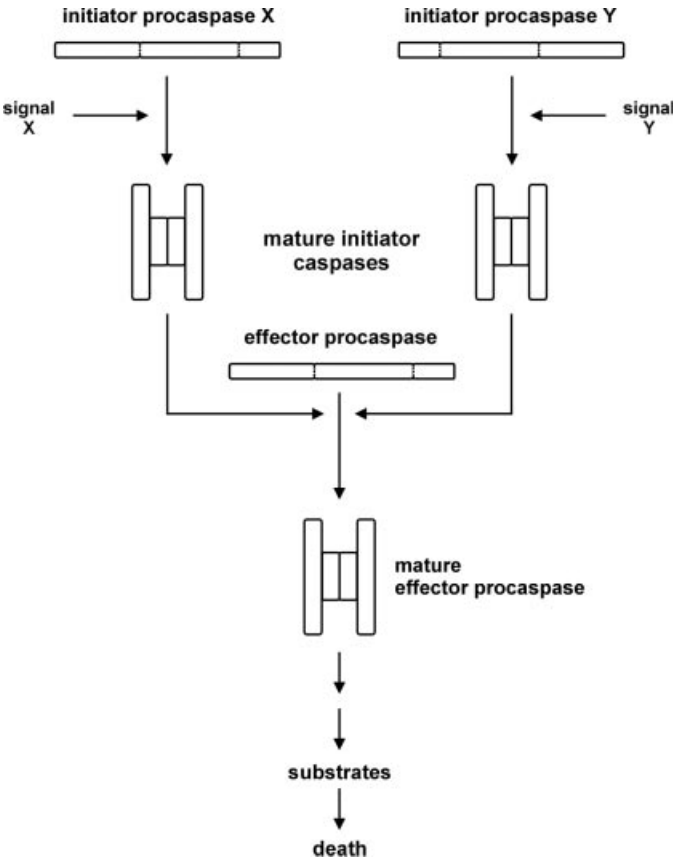



**Fig. 15.6.** Initiator and effector caspases in apoptosis.
An essential part of the apoptotic program is a caspase cascade. Apoptosis is initiated by proteolytic processing of intiator-procapsases under the influence of a variety of signals. The mature initiator caspase catalyzes the processing of a effector-procaspase to the active enzyme, which degrades specific substrates and/or activates further procaspases. In this way, caspases can be activated sequentially in a protease cascade.

Based on the triggering stimulus and the nature of the components involved, at least two apoptotic signaling pathways can be differentiated that lead to activation of the effector caspases. On the one hand, receptor systems may be involved; on the other hand, activation may be triggered by cytotoxic stress. The two pathways differ in the mechanism of activation of the initiator caspase but use the same effector caspase at least partially.

### Mechanism of Caspase Activation

The cell uses two mechanisms for activation of caspases:

- Proteolytic cleavage in a caspase cascade: A procaspase is proteolytically activated by a caspase already activated upstream in the apoptotic signaling pathway. This mechanism is used particularly in the effector part of the apoptotic pathway.
- Proximity-induced activation of an initiator procaspase: Many observations support a model of caspase activation in which an initiator procaspase with little activity is activated by stimulus-induced aggregation (review: Hengartner, 1998) The procaspases have low protease activity corresponding to only 1–2 % of that of the mature caspase. If a procaspase is recruited into a larger aggregate, this activity is sufficient for mutual processing of the procaspase molecules to mature caspases. Aggregation of the procaspases is supported by cofactors such as FADD (see 16.3.3) or Apaf1/Cyt c, which are activated by external signals (ligand binding to death receptors; mitochondrial damage). The basis of the aggregation is protein-protein interaction between specific domains of the caspases and the cofactors. Two general types of interaction domains have been identified. Procaspase 8 contains two tandem *death effector domains (DED)* whereas the procaspases 1, 2, 4, 5 and 9 contain a *caspase recruitment domain (CARD)*. The domains occur in similar forms in the procaspases and in the cofactors and enable formation of larger aggregates. According to this model, the cofactors function as apoptotic chaperones by bringing about the formation of particular aggregates in which activation of the initiator caspase occurs.

## 15.3.2 The Family of Bcl-2 Proteins

The Bcl-2 protein was first identified as an oncoprotein coded by a gene affected by translocations of chromosomes 14 and 18 in B cell lymphomas. It was soon shown, however, that the Bcl-2 protein is not involved in regulation of the cell cycle, in contrast to many other oncoproteins, and thus does not fit into the classical oncogene picture. Furthermore, homology was established with the Ced9 protein of *C. elegans*, which has an antiapoptotic function in this organism.

The Bcl-2 protein is a member of a protein family involved in regulation of the apoptotic program in mammalian cells (review: Reed, 1997; Adams and Cory, 1998). At present, at least 15 members are known of the *Bcl-2 family*, which can have a negative or a positive effect on initiation of the apoptotic program. All Bcl-2 family members have at least one copy of a so-called BH motif (BH, Bcl-2 homolog), of which there are four types (BH1 – BH4).

### Antiapoptotic Bcl-2 Family Members

The antiapoptotic members of the Bcl-2 family (Bcl-2, BclX, BclW, A1, Mcl-1) inhibit apoptosis by various cytotoxic effects. At a minimum, they contain the motives BH1 and BH2. Bcl-2 protein contains all four BH motives. For the mechanism of the antiapoptotic function, see 15.3.4 and 15.4.

The antiapoptotic function of the Bcl-2 protein has been clearly shown experimentally. Its overexpression can prevent initiation of the apoptotic program in various cell types. The oncogenic function of Bcl-2 protein, observed in association with its overex-

pression, can be explained by its antiapoptotic effect: the high level of Bcl-2 protein suppresses initiation of the apoptotic program and an important requirement for further tumor progression is fulfilled. In this situation, damaged cells can survive, which would have been eliminated by apoptosis in the normal situation.

**Proapoptotic Bcl-2 Family Members**

There are two proapoptotic families of Bcl-2 proteins.

One family includes the Bax, Bak and Bok proteins that have a similar structure to the Bcl-2 protein. There is a link via the Bax protein to the tumor suppressor protein p53 (see 15.6).

The proteins of the other family are also known as BH3 proteins, since they only contain the BH3 domain. An important member of the BH3 family is the Bad protein which is part of the PI3-kinase/Akt signaling pathway (see 15.6).

The biochemical basis of the various activities of the Bcl-2 family members is only partially understood. The antiapoptotic Bcl-2 proteins may function by directly inhibiting the activation of the caspases (see 15.3.3). It is assumed that proapoptotic proteins interact with antiapoptotic proteins and halt their inhibition of apoptosis.

Various members of the Bcl-2 family can interact with another via the BH domains and form hetero-oligomeric complexes in which the different activities neutralize one another. The relationship of proapoptotic and antiapoptotic Bcl-2 family proteins thus helps to determine a cell's susceptibility to apoptosis.

## 15.3.3 Cofactors of Caspase Activation

Activation of the caspases requires the help of a number of *cofactors* that are also known as *activators* or *adaptors*. Different cofactors are involved depending on the trigger mechanism of caspase activation. A central function of the cofactors is to bring about aggregation and thus activation of the procaspases. This occurs by specific protein-protein interactions with the help of common structural motives. Examples of such motives are the *death domains (DD), death effector domains (DED)* and the *caspase recruitment domains (CARD)*, which all have a similar structure of six α-helices.

**Cofactors of Death-receptor-mediated Apoptosis:**

In the case of the death receptor Fas (see 15.4), activation of the caspase involves a protein that interacts with the cytoplasmic part of the receptor and is known as FADD protein (Fas-associated death domain). The FADD protein has distinct structural motives that mediate specific interactions with other proteins. It interacts via the death domain with the receptor and via the death effector domain with the corresponding caspase (here caspase 8).

#### Apaf1 and Cytochrome c as Cofactors

The Apaf1 protein has been identified as a central cofactor in cytotoxically-triggered apoptosis. The Apaf1 protein is homologous to Ced4 protein of *C. elegans*. It binds to the initiator caspase with the help of a CARD motif. A CARD motif is found in the Apaf1 protein and in various caspases (caspases 1, 2, 4, 5 and 9).

A further cofactor in activation in this system is cytochrome c. The ATP-dependent activation occurs in a complex of caspase, Apaf1 protein and Cyt c, which is also known as the *apoptosome*.

### 15.3.4 Intracellular Regulation

The ability of the cofactors to activate caspases is regulated by a number of other proteins that appear to interact directly with the cofactors.

Examples of these regulators are:

#### Flip Proteins

These proteins interact with the cofactor FADD in death-receptor-mediated apoptosis (see 15.4) and block its activating effect.

#### Bcl-2 Family Members

The antiapoptotic Bcl-2 family members control apoptosis by various mechanisms without directly binding to the caspases. Bcl-2 proteins can interact with cofactors and inhibit their activity. They can also act antiapoptotically by binding to mitochondria and interfering with release of cytochrome c. Furthermore, they can interact with other proapoptotic proteins, e.g. with propapoptotic members of their own family.

#### Inhibitors of Apoptosis

These proteins inhibit apoptosis but their site of action is uncertain. Some of them apparently bind to caspases and inhibit these directly.

## 15.4 Stress-mediated Apoptosis: the Cytochrome c/Apaf1 Pathway

In addition to receptor-mediated apoptosis (see 15.5), there is another main pathway activated by various forms of cellular stress. Examples of stress effects that can induce apoptosis are $\gamma$- and UV-radiation, treatment with cytotoxic drugs such as actinomycin D and removal of cytokines. As a consequence of stress, procaspase 8 is activated by a

complex mechanism that is not completely understood. A model for activation is shown in Fig. 15.7.

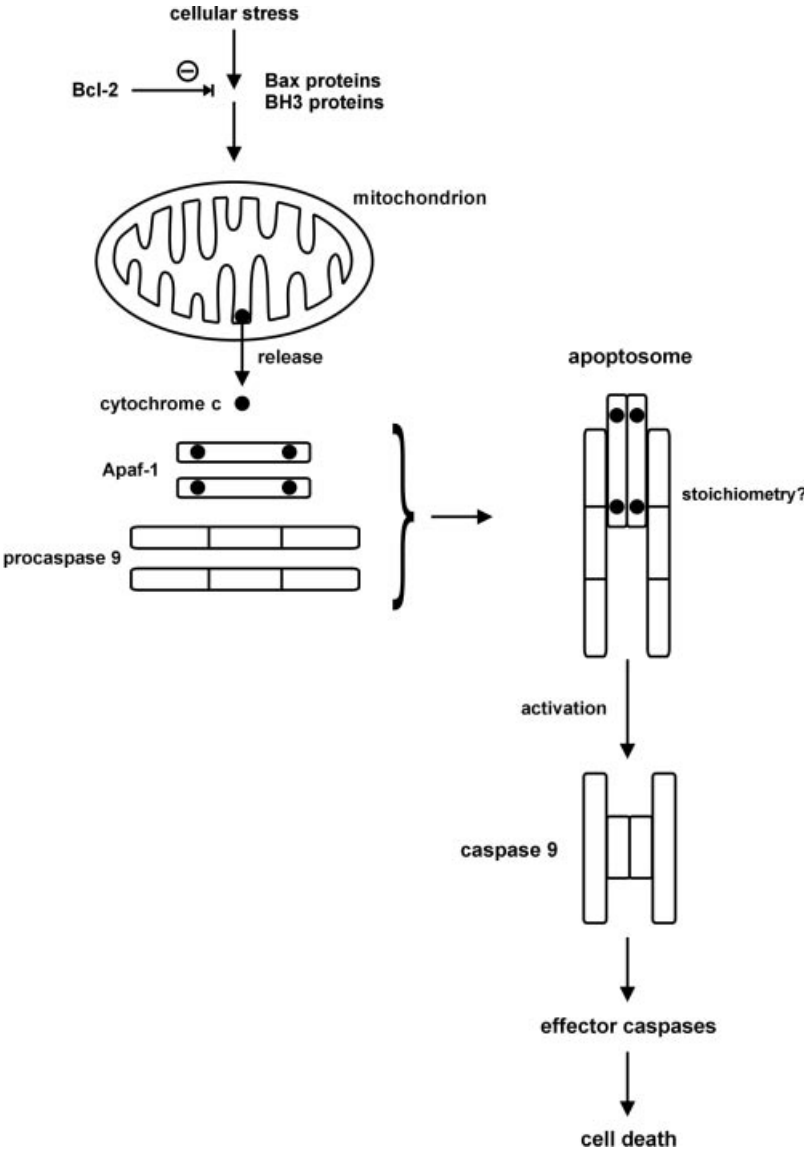



**Fig. 15.7.** Apoptotic signalling during stress-mediated apoptosis
In this model, cellular stress mediates the release of cytochrome C from the mitochondrion. The proapoptotic proteins Bax and BH3 proteins support the release of cytochrome C, while the antiapoptotic Bcl2 protein has an inhibitory effect. Cytosolic cytochrome C binds to the cofactor Apaf 1, which then associates via its CARD motif with procaspase 9 to a complex termed apoptosome. In this complex, procaspase 9 is processed to the mature caspase which subsequently activates downstream effector caspases and commits the cell to death.

In this model, the mitochondrion has an important role in initiation of apoptosis (review: Green and Reed, 1998). According to this, cytochrome c is first released from the mitochondria into the cytosol as a consequence of stress. Cytochrome c functions as a cofactor for the activation of the initiator caspase 8 by binding to a complex of procaspase 8 and the cofactor Apaf1. The hetero-oligomeric complex of procaspase 8, Apaf1 and cytochrome c forms the apoptosome. There is also a requirement for ATP for this type of activation mediated by Apaf1, which contains an ATP/dATP binding site. The function of the ATP binding is not understood; possibly, ATP-dependent conformational changes in the apoptosome are involved. In the apoptosome, autocatalytic activation of procaspase 8 to the mature caspase takes place, which then activates the caspase cascade of the effector caspases 3, 6 and 7.

It is not clear how cytochrome c release from the mitochondrion is initiated. There is evidence that proteins of the Bcl-2 family are involved. The antiapoptotic Bcl-2 protein is itself localized on the outer mitochondrial membrane and is thought to have an inhibiting effect on cytochrome c release. The propapototic Bax protein can shuttle between the cytosol and the mitochondrion. For some members of the Bcl-2 family, including Bax, formation of a pore in the mitochondrial membrane has been postulated based on structural investigations. Thus, the Bax protein can form a pore structure in membranes which permits passage of ions. Possibly, a pore is formed on activation of the Bax protein, permitting passage of cytochrome c and thus enabling activation of the initiator caspase.

Overall, regulation of caspase activation is little understood, since other protein factors of unknown function are also involved in activation. It can be assumed that there are other apoptotic pathways in addition to the two main pathways described here. In addition, „crosstalk“ between the pathways may occur.

## 15.5 Death-receptor-triggered Apoptosis

Apoptosis in the immune system often involves receptor systems known as „death receptors“ (review: Yuan, 1997; Askenazi and Dixit, 1998). The death receptors belong to the superfamily of tumor necrosis factor (TNF) receptors, which are characterized by a Cys-rich extracellular domain and a homologous intracellular domain known as the „death domain“. The best characterized death receptors are Fas (also known as CD95) and tumor necrosis factor receptor 1 (TNFR1). The ligand for the Fas receptor (Fas ligand) is a homotrimeric protein that causes oligomerization of its receptor on binding. Associated with this is a clustering of the death domains and binding of cofactor FADD. The FADD protein binds via its DED motif to a homologous motif in procaspase 8. A model of the apoptotic Fas signaling pathway is shown in Fig. 15.8. The cofactor function of FADD can, in turn, be blocked by interaction with the regulator FLIP.

Procaspase 8 functions as an initiator caspase in this system, since its activation is the signal for activation of the downstream caspase cascade. The DED motif of caspase 8 is localized in its large prodomain. Similar motives are found in other caspases with large prodomains (caspases 2, 8 and 9).

Upon recruitment by FADD, caspase 8 oligomerizes, triggering activation by self-cleavage. Caspase 8 then proteolytically activates downstream effector caspases, such as caspase 9, and commits the cell to death.

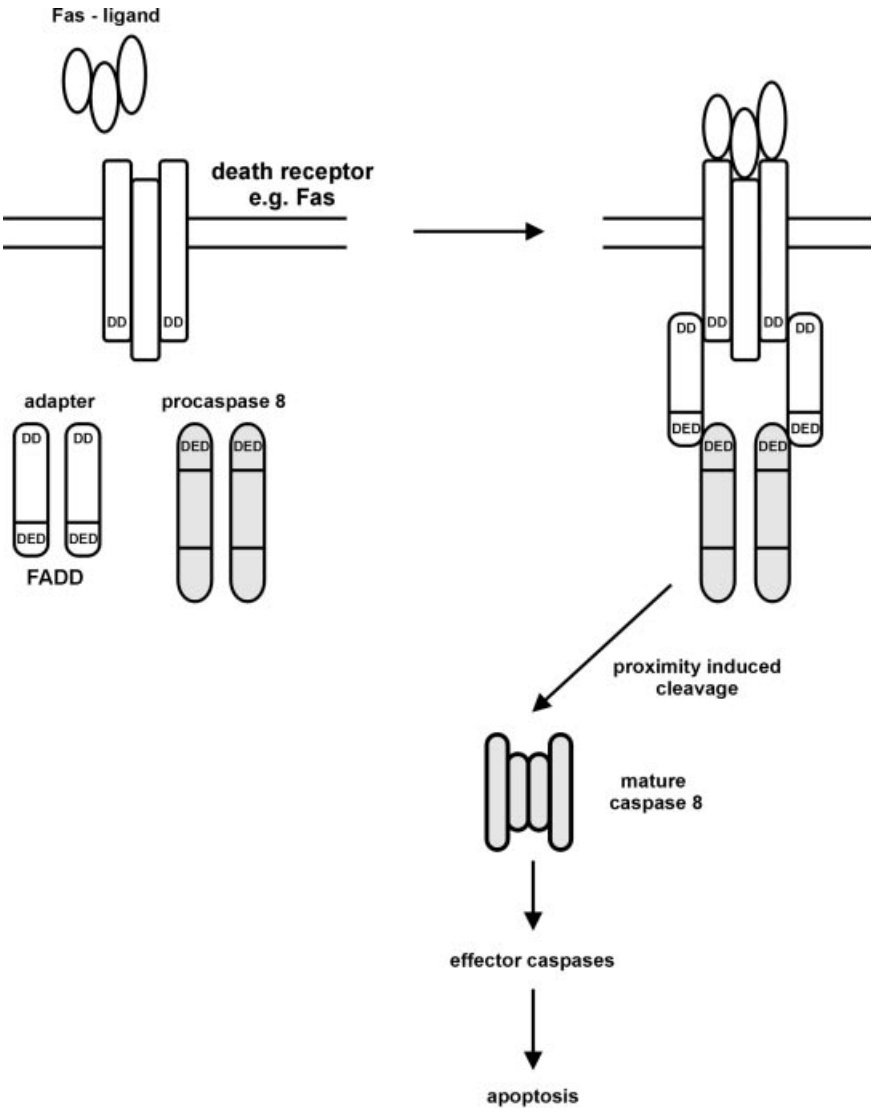



**Fig. 15.8.** Model of Fas receptor-triggered apoptosis
Binding of the ligand of the Fas receptor triggers clustering of the receptor and association of the cofactor FADD (fas-associated protein with death domain) which interacts with the receptor via its death domain (DD). Procaspase 8 binds to FADD via a common DED (death effector domain) motif and is thereby also recruited into the Fas-receptor associated complex. Due to the clustering of the proteins, proximity-induced cleavage of procaspase 8 to the mature initiator caspase 8 takes place. This activates the effector caspases and triggers cell death.

Signal transduction by the Fas-related TNF receptor is of a more complex nature and can mediate both proapoptotic and antiapoptotic signals. In the normal situation, binding of TNF to its receptor initiates a signal chain that activates a MAPK pathway and creates a signal for c-Jun expression. In addition, the IκB-NFκB pathway may be activated. Both NFκB and c-Jun have an antiapoptotic and proliferation-promoting effect.

In certain situations, e.g., inhibition of protein biosynthesis, TNF receptor activation has a proapoptotic effect, however, in that caspases are activated and cell death is initiated.

## 15.6 Apoptosis and Cellular Signaling Pathways

Like most functions in animal cells, the apoptotic program is regulated by signals from other cells, which can activate or suppress. In addition to these extracellular controls, the apoptotic program is also controlled by intracellular signaling pathways. At different levels of the apoptotic program, there are links to cell-cell interactions, to growth-factor-controlled signaling pathways, to the cell cycle and the DNA damage checkpoint system. Overall, our knowledge of links to intracellular and extracellular signaling pathways is very incomplete and is limited to a few examples.

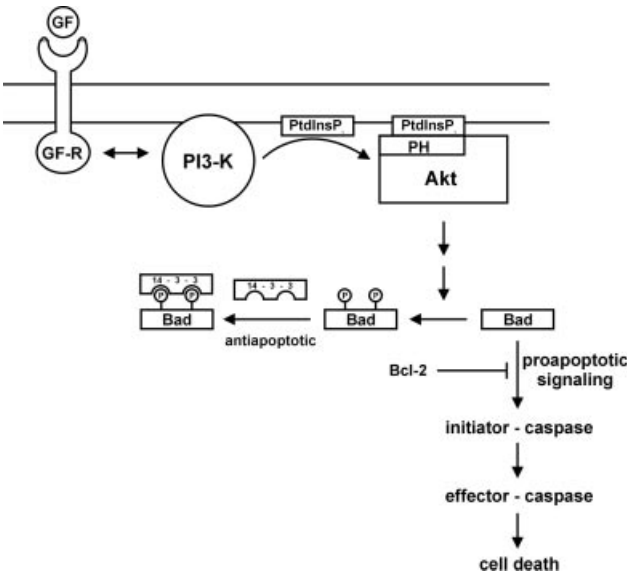



**Fig. 15.9.** Antiapoptotic signalling by the PI3-kinase/Akt kinase pathway
The PI3 kinase/Akt kinase pathway influences apoptosis via phosphorylation of the Bad protein, which is a member of the family of Bcl-2 proteins. Activation of the PI3-kinase pathway leads to Akt-kinase-catalyzed phosphorylation of Bad protein. Bad protein in its unphosphorylated form participates in activation of initiator caspases and thus has a proapoptotic effect. Phosphorylation of Bad protein by Akt kinase (or related kinases) has an antiapoptotic effect since phosphorylated Bad protein is a binding substrate of 14–3-3 proteins. Bad is thus sequestered in an inactive state and is not available for triggering of apoptosis.

Two of these examples are highlighted:

### PI3-kinase and Apoptosis

PI3-kinase (see 6.6) can mediate antiapoptotic signals, in addition to growth-promoting signals (Fig. 14.9). The antiapoptotic signal conduction starts at PI3-kinase to Akt kinase, which is activated by the messenger substance $PtdInsP_3$ formed by PI3-kinase. The Bad protein has been identified as a substrate of Akt kinase. The Bad protein is a proapoptotic member of the Bcl-2 family. It is phosphorylated by Akt kinase at several Ser residues and its proapoptotic effect is thus inhibited (Datta et al., 1997). Experimental evidence exists that the 14–3-3 proteins are involved in this inhibition; these bind to phosphoserine residues of Bad protein and thus inactivate its proapoptotic function.

### p53 and Apoptosis

The p53 protein is at the center of apoptotic signaling pathways initiated by DNA damage and defects in the course of the cell cycle. Depending on cell type, p53-induced apoptosis either requires transcriptional activation (Polyak et al., 1997) or occurs without new RNA and protein synthesis (see Fig. 14.10).

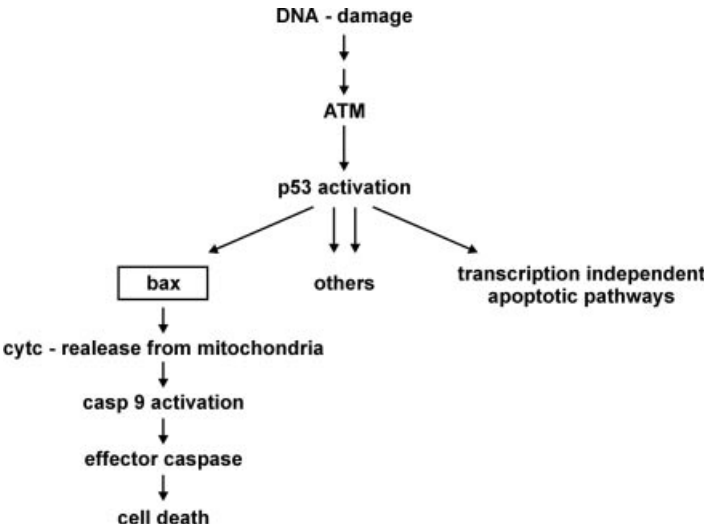



**Fig. 15.10.** Pathways of DNA damage-mediated and p53-mediated apoptosis
The presence of DNA lesions activates the ATM kinase and leads to an increase in p53 concentration. In a transcription-dependent pathway, p53 functions as a transcription activator of the *bax* gene. The increase in Bax protein facilitates release of cytochrome C from mitochondria and this serves as a trigger for activation of initiator and effector caspases. p53 also influences apoptosis by less well characterized ways, some of which are transcription independent.

An important transcriptional target of the p53 protein that can induce apoptosis is the *bax* gene. The Bax protein belongs to the family of Bcl-2 proteins (see 15.3.2) and has a proapototic effect. There is speculation that the p53-induced increase in Bax concentration leads to formation of ion pores in mitochondria and that cytochrome c is released into the cytosol via these pores. Cytochrome c then functions as a cofactor which, together with Apaf1 protein, activates procaspase 8 and initiates the apoptotic program.

Transactivation of proapoptotic genes is not the only way that p53 protein can activate the apoptotic program. There is evidence that variants of p53, which are independent of Bax protein and do not operate at the transcription level, can also result in apoptosis (see Haffner and Oren, 1995; Ko and Prives, 1996). Thus, a p53-regulated redistribution of the Fas death receptor from the cytosol to the cell membrane has been demonstrated (Bennet et al., 1998). Overall, these mechanisms are poorly understood.

# References Chapter 15

Adams, J.M. and Cory, S. 'The BCl-2 protein family: Arbiters of cell survival' (1998) Science *281*, 1322–1326

Askenazi, A. and Dixit, V.M. 'Death receptors: signaling and modulation' (1998) Science *281*, 1305–1308

Bennett, M., Macdonald, K., Chan, S.W., Luzio, J.P., Simari, R. and Weissberg, P., Cell surface trafficking of Fas: a rapid mechanism of p53-mediated apoptosis (1998) Science 282, 290–293

Datta, S.R., Dudek, H., Tao, X., Masters, S., Fu, H., Gotoh, Y. and Greenberg, M.E. 'Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery' (1997) Cell *91*, 231–241

Earnshaw, W.C., Martins, L.M. and Kaufmann, S.H. 'Mammalian Caspases: Structure, Activation, Substrates and Functions during Apoptosis' (1999) Ann. Rev. Biochem. *68*, 383–424.

Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A. and Nagata, S. 'A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD' (1998) Nature *391*, 43–50

Evan, G. and Littlewood, T. 'A matter of life and death' (1998) Science *281*, 1317–1322

Fisher, D.E. 'Apoptosis in cancer therapy' (1994) Cell *78,* 539–542

Green, D.R. ' Apoptotic pathways: the roads to ruin' (1998) Cell *94*, 695–698

Green, D.R. and Reed, J.C. 'Mitochondria and apoptosis' (1998) Science *281*, 1309–1312

Hengartner, M. ' death by crowd control' (1998) Science *281*, 1298–1299

Kinzler, K.W. und Vogelstein,B. 'Life (and death) in a malignant tumor' (1996) Nature *379*, 19–20

Kumar, S. 'ICE-like proteases in apoptosis' (1995) Trends Bioch. Sci. *20*, 198–202

Martin, S.J., Green, D.R. und Cotter, T.G. 'Dicing with death: dissecting the components of the apoptosis machinery' (1994) Trends Bioch. Sci *19*, 26–30

Miyashita, T. und Reed, .C. 'Tumor suppressor p53 is a direct transcriptional activator of the human bax gene' (1995) Cell *80*, 293–299

Nicholson, D.W., Ali, A., Thornberry, N.A., Vaillantcourt, J.P., Ding, C.K., Gallant, M., Gareau, Y., Griffin, P.R., Labelle, M., Lazebnik, Y.A., Munday, N.A., Raju, S.M., Smulson, M.E., Yamin, T., Yu, V.L. und Miller, D.K. 'Identification and inhibition of the ICE/CED-3 protease necessary for mammalian apoptosis' (1995) Nature *376*, 37–43

Polyak, K., Xia, Y., Zweier, J.L., Kinzler, K.W. and Vogelstein, B. ' A model for p53-induced apoptosis' (1997) Nature *389*, 300–305

Raff, M. ' Cell suicide for beginners' (1998) Nature *396*, 119–122

Reed, J.C. 'Double identity for proteins of the Bcl-2 family' (1997) Nature *387*, 773–776

Yuan, J. 'Transducing signals of life and death' (1997) Curr. Op. Cell Biol. *9*, 247–251

Salvesen, G.S. and Dixit, V.M. 'Caspases: intracellular signaling by proteolysis' (1997) Cell *91*, 443–446)

Sherr, C.J. 'Cancer cell cycles' (1996) Science *274*, 1672–1677

Thompson, C.B. 'Apoptosis in the pathogenesis and treatment of disease' (1995) Science *264*, 1456–1462

Thornberry,N.A. 'Caspases: key mediators of apoptosis' (1998) Chem. Biol. *5*, R97-R103

Vaux,D.L. and Korsmeyer, S.J. 'Cell death in development' (1999) Cell **96**, 245–254

*Biochemistry of Signal Transduction and Regulation, Second Edition.* Gerhard Krauss

ISBNs: 3-527-30378-2 (Softcover); 3-527-60005-1 (Electronic)

**Chapter 16**

# Ion Channels and Signal Transduction

## 16.1 Principles of Neuronal Communication

Processing of stimuli in the nerve system includes intracellular communication, in the form of stimulus conduction in the nerve, and intercellular communication that takes place at the synapses. Synapses are specialized cell contacts used by nerve cells to communicate with one another or with other cells, e.g. muscle or sensory cells.

Neuronal communication takes place predominantly by two mechanisms:

### Electrical Communication

In electrical communication, *changes in membrane potential* are used to conduct a stimulus within a nerve cell. Changes in membrane potential can also be used for intercellular communication. In this case, communication between the cells takes place via electrical synapses at which the potential change can be directly passed on to neighboring cell. Central components of electrical communication are voltage-dependent ion channels with open states regulated by changes in the membrane potential.

### Chemical Communication at Synapses

The chemical synapse uses chemical messenger substances to communicate between a presynaptic and a postsynaptic cell. The messengers used for signal transduction at chemical synapses are known as neurotransmitters. The neurotransmitters bind to corresponding receptors at the postsynaptic cell, enabling selective passage of ions through the membrane of this cell. These receptors are ligand-controlled ion channels, with the neurotransmitters serving as ligands. Fig. 16.1 summarizes some important neurotransmitters.

Composition and structure of a chemical synapse are shown in a simplified form in Fig.16.2.

In the presynaptic cell, the neurotransmitters are stored in vesicles. On arrival of an electrical signal (action potential, see 16.2), an influx of $Ca^{2+}$ takes place into the presynaptic cell as voltage-gated $Ca^{2+}$ channels are opened. The increase in $Ca^{2+}$ concentration leads to fusion of the vesicles with the membrane of the postsynaptic cell. The neurotransmitters are released into the synaptic cleft and diffuse to a corresponding receptor on the surface of the postsynaptic cell. Binding of the neurotransmitter to the receptor induces opening of an ion channel that is a component of the receptor. The type of ions that can enter depends on the selectivity of the ion channel. There are $Na^+$, $K^+$, $Ca^{2+}$ and $Cl^-$ specific ion channels. The ion flux creates an electric signal in the post-

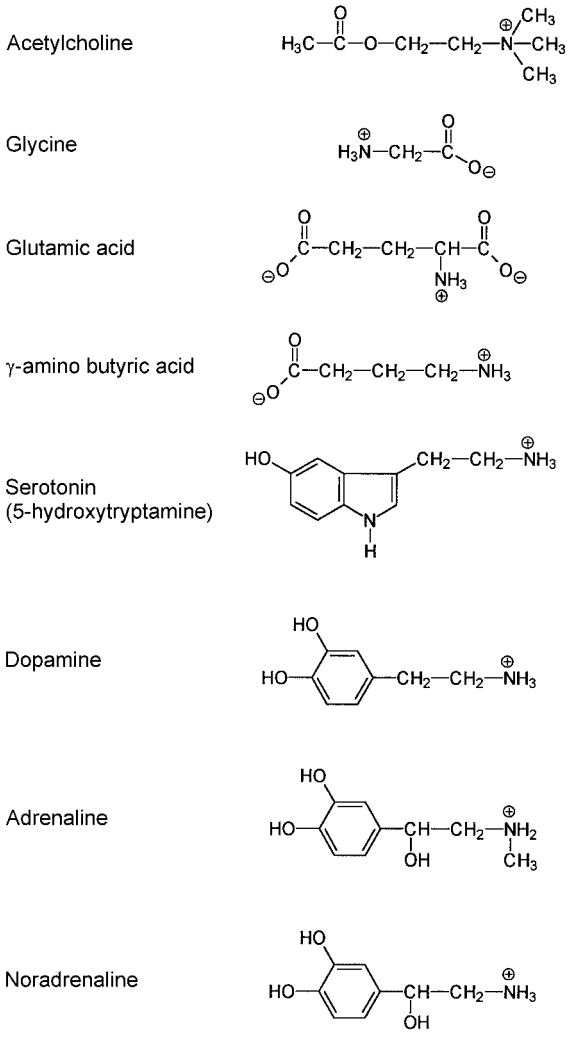



**Fig. 16.1.** Examples of neurotransmitters. Structures are shown for important neurotransmitters in vertebrates.

synaptic cell, and this propagates itself in the postsynaptic cell membrane. Ion flux can also trigger secondary reactions. Thus, the electrical signal may lead to opening of voltage-gated $Ca^{2+}$ channels. $Ca^{2+}$ entering through the opened channels may also be used as a „second messenger“ for further signal conduction.

## 16.2 Membrane Potential and Electrical Communication

In higher organisms, there is a noticeable imbalance in the equilibrium of ion distribution between the cell interior and the extracellular space (see Table 16.1). The concentration of $Na^+$ and $Cl^-$ ions is ca. 10-fold higher in the extracellular space than in the

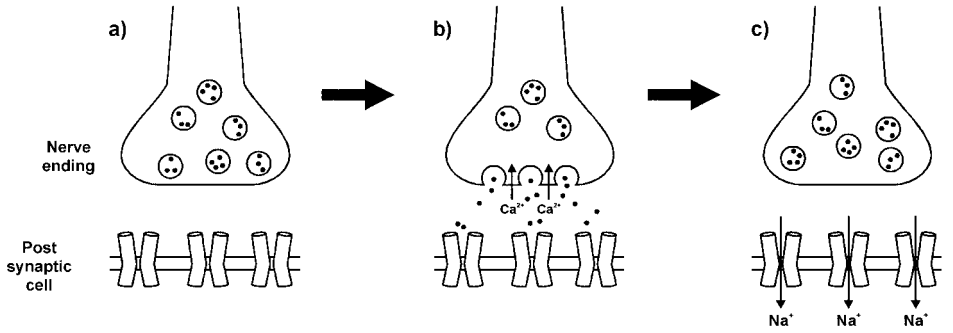



**Fig. 16.2.** The elementary processes at a chemical synapse. a) In the resting state, the neurotransmitter is stored in vesicles in the presynaptic cell. b) An arriving action potential leads to influx of $Ca^{2+}$ into the presynaptic cell. Consequently, the vesicles fuse with the presynaptic membrane and the neurotransmitter is released into the synaptic cleft. c) The neurotransmitter diffuses across the synaptic cleft and binds to receptors at the surface of the postsynaptic cell. Ion channel and receptor form a structural unit. The ion channel opens and there is an influx of $Na^+$ ions into the postsynaptic cell. Recycling takes place in the presynaptic cell and the vesicles are reloaded with neurotransmitter.

cytosol. For $K^+$, the situation is reversed: here, the concentration is higher in the cytosol. The imbalance in ion distribution and the possibility of selective ion passage through the membrane are the reasons for formation of a membrane potential. The membrane potential is –30 to –50 mV in the resting state of the cell and the inner side of the membrane is negatively charged.

The membrane potential of a cell in the resting state is predominantly determined by two processes:

ATP-dependent $Na^+/K^+$ ATPases transport $Na^+$ ions outwards and $K^+$ ions inwards against the concentration gradient.

$K^+$ leak channels permit entry of $K^+$ ions along their concentration gradient from the interior to the exterior.

The activity of both processes determines the resting potential of the cell. Changes in the resting potential are used as signals in intracellular and intercellular communication. In nerve cells, changes in the membrane potential in the form of action potentials are used to conduct a stimulus further along the nerve cell and to communicate with other cells. Formation and propagation of an action potential are associated with spatially and temporally limited changes in the membrane potential. The potential changes are initiated by opening of ion channels that permit selective passage of ions.

**Table 16.1.** Intra- and extracellular ion concentration in typical neuronal tissues.

| Cytosol | Extracellular region |
|---|---|
| 140 mM $K^+$ | 4 mM $K^+$ |
| 12 mM $Na^+$ | 150 mM $Na^+$ |
| 4 mM $Cl^-$ | 120 mM $Cl^-$ |
| 148 mM $A^-$ (negatively charged side chains of proteins) | 34 mM $A^-$ |

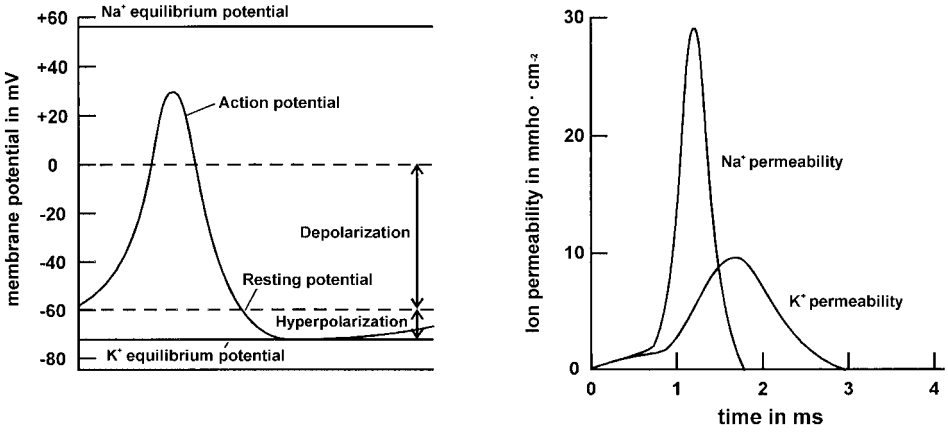



**Fig. 16.3.** Course of an action potential and permeability changes for $Na^+$ and $K^+$. a) The resting potential of the cell is about –30 to –50 mV. On formation of an action potential, depolarization takes place, and in the course of this depolarization, the membrane potential increases by ca. 70 mV. The potential then decreases again within a few msec, passing through a hyperpolarized state in which the membrane potential is more negative than the resting potential. After a few msec, the cell returns to the resting state. b) Changes in ion permeability. Formation of the action potential is associated with increased permeability of the membrane for $Na^+$ ions, due to opening of voltage-controlled $Na^+$ channels. The depolarization induces delayed opening of $K^+$ channels and thus increases $K^+$ permeability. There is an efflux of $K^+$, and $Na^+$ channels close. As the action potential dies away, the $K^+$ channels close and $K^+$ permeability decreases again.

The potential change depends on the nature of the ions crossing the membrane. If there is a $Na^+$ influx, the membrane potential decreases and there is a depolarization. The membrane often gains a positive value by this process. In contrast, a $K^+$ efflux or a $Cl^-$ influx leads to hyperpolarization. The membrane potential becomes more negative than the resting potential.

Stimulus conduction in nerves is associated with sequential opening of $Na^+$ and $K^+$ channels (Fig. 16.3). The arrival of an action potential triggers opening of $K^+$ channels. The $Na^+$ influx leads to a local depolarization, which is manifested as a rise in membrane potential to positive values (up to +30 mV). This depolarization brings about opening of $K^+$ channels delayed relative to $Na^+$ channel opening. There is now an efflux of $K^+$ and the membrane potential falls to negative values below the equilibrium potential. From this hyperpolarized state, there is a slow transition into the resting state.

## 16.3 Structure and Function of Voltage-gated Ion Channels

### 16.3.1 Principles of Regulation of Ion Channels

The cell uses various mechanisms to trigger and regulate the flow of ions through ion channels. Ion flow is subject to the following gating mechanisms, according to the nature of the ion channel:

- For voltage-gated ion channels, the ion flow is regulated by changes in the membrane potential.
- Ligand-gated ion channels are closed (or opened) on binding of specific ligands to extracellular (or intracellular) structural elements of the channel protein.
- Ion channels can also be effector molecules in receptor-regulated signaling pathways (review; Jan and Jan, 1997). The ion channel may be opened, for example, by direct interaction with a $G_{i\alpha}$ subunit in the process of activation of G-protein-coupled receptors.
- There are ion channels which are controlled both by membrane potential and ligand binding. An example is the NMDA receptor (see 16.4.2.1).

In addition to these trigger mechanisms, which permit rapid switching between different states of the ion channel, there is also the possibility to modulate and regulate activity of ion channels on a long-term basis. This regulation is usually due to phosphorylation in the cytoplasmic part (see 16.4.2.1).

Highly resolved structural information is currently available only for a constitutively open $K^+$ channel from *Streptomyces lividans* (see 16.3.3).

## 16.3.2 Characteristics of Voltage-gated Ion Channels

Voltage-gated ion channels are responsible for creation and conduction of electrical signals in cell membranes. The voltage-gated ion channels are transmembrane proteins, and these can exist in an open state or a closed state. Depolarization of the membrane leads to opening of the ion channels and—according to the ion selectivity of the channel—to a change in the permeability of $Na^+$, $K^+$ or $Ca^{2+}$.

Two phases of permeability changes may be differentiated: depolarization initially leads to an increase in the permeability for $Na^+$, $K^+$ or $Ca^{2+}$. This takes place on a timescale of 0.5 to several hundred milliseconds. On a timescale of 2 msec to several seconds, the permeability then falls back to the original state.

The function of the ion channels is determined by the processes of activation and deactivation.

The process of *activation* determines duration, size and voltage dependence of the increase in permeability.

The process of *inactivation* determines the rate and voltage dependence of the return to the resting state.

There are thus three states for an ion channel, namely the resting state, activated state and inactivated state. Only the activated state is conducting; in the resting state and in the inactivated state, the ion channel is closed. For renewed opening, the ion channel must move from the inactivated state to the resting state.

### Ion Selectivity

The voltage-gated $Na^+$, $K^+$ and $Ca^{2+}$ channels differ in their ion selectivity. $Na^+$ channels have a selectivity for $Na^+$ up to 12-fold higher than for other cations. $Ca^{2+}$ channels may have a selectivity of up to 1000-fold for $Ca^{2+}$.

### 16.3.3 Structure of Voltage-gated Ion Channels

The voltage-gated ion channels for $Na^+$, $K^+$ and $Ca^{2+}$ known at present all have a very similar overall structure (review: Catterall, 1995). This permits division of the channels into three classes (Fig. 16.4). Each class includes various members that differ in structural details, in the characteristics of the pore and in regulation. Most information is available for the voltage-gated $Na^+$ channel from the electric organ of the electric eel *Elektrophorus electricus*, for a $K^+$ channel („shaker") from *Drosophila* and for a $Ca^{2+}$ channel in skeletal muscle. These ion channels will be discussed in the following sections.

Generally, voltage-gated ion channels are formed of several subunits. The ion pore is formed from the large subunit whereas the other subunits have a modulating or amplifying effect on the pore function of the large subunit. There are potential transmembrane domains in most subunits (Fig. 16.5). The extracellular part of the subunits is often glycosylated and the intracellular structural parts possess regulatory phosphorylation sites.

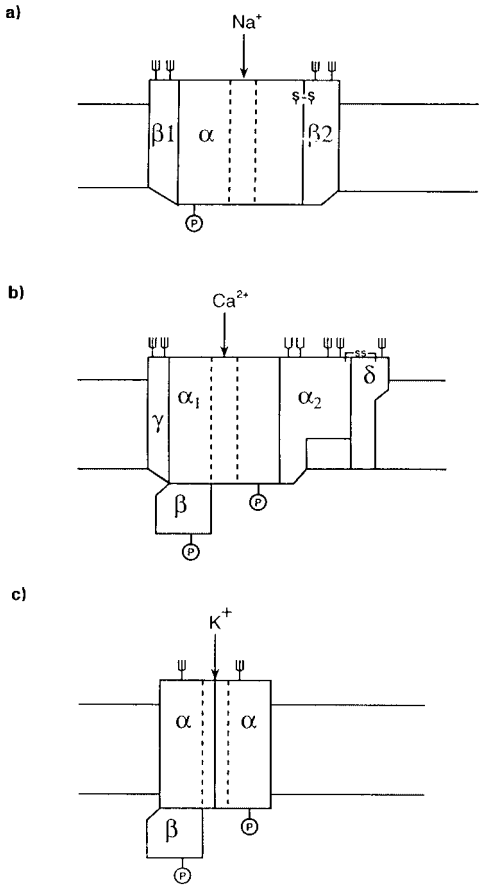



**Fig. 16.4.** Subunit structure of voltage-controlled ion channels. The subunit structure of various voltage-controlled ion channels is shown in schematic form with nomenclature of the different subunits. Phosphorylation sites (P) in the cytoplasmic part are also shown, as well as glycosylation sites (Y) in the extracellular part of the ion channels. The dotted line indicates the location of the ion pore. a) $Na^+$ channel; b) $Ca^{2+}$channel; c) $K^+$ channel. According to Catterall, (1995) with permission.

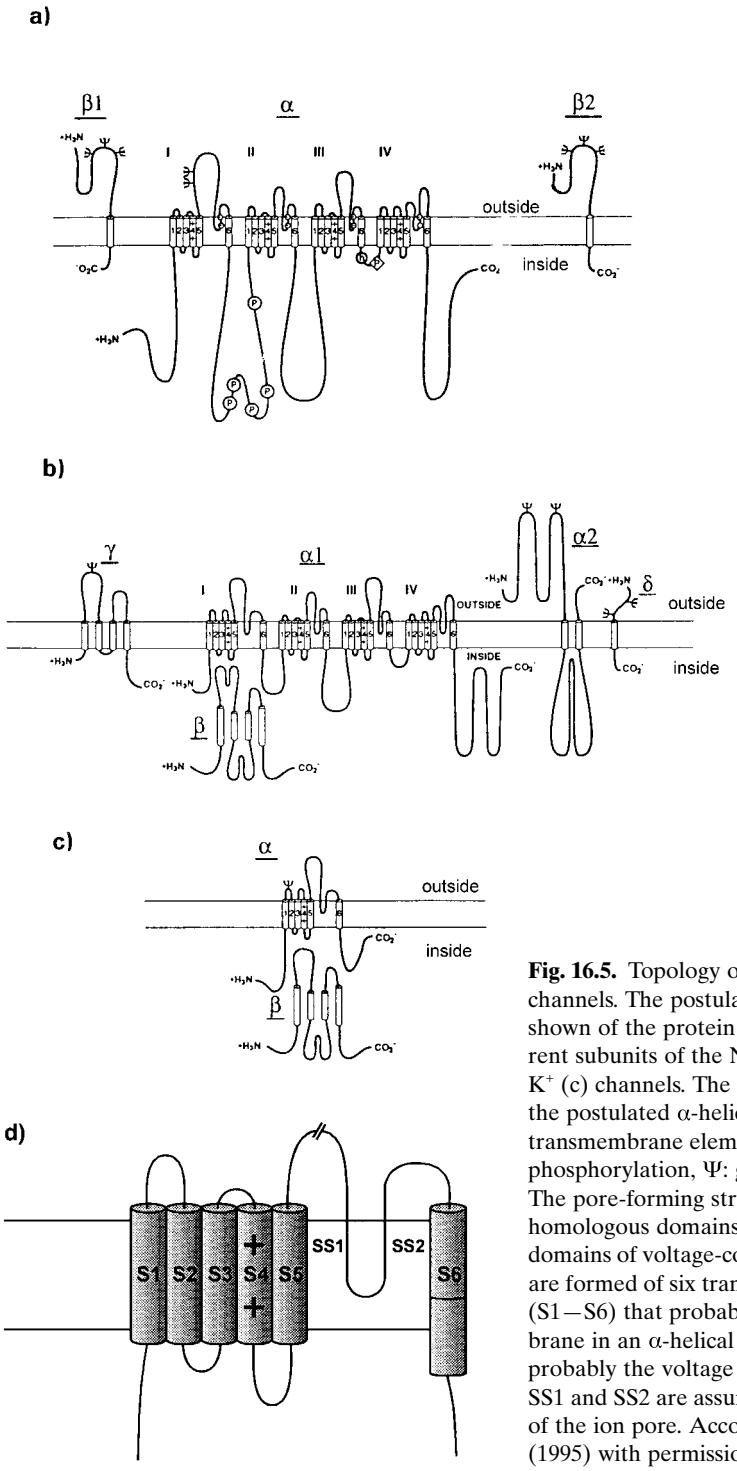



**Fig. 16.5.** Topology of voltage-gated ion channels. The postulated tolopogy is shown of the protein chains of the different subunits of the $Na^+$ (a), $Ca^{2+}$ (b) and $K^+$ (c) channels. The cylinder represents the postulated α-helical structure of the transmembrane element. P: proven sites of phosphorylation, Ψ: glycosylation site. d) The pore-forming structural elements of homologous domains. The homologous domains of voltage-controlled ion channels are formed of six transmembrane elements (S1—S6) that probably cross the membrane in an α-helical form. The S4 helix is probably the voltage sensor; the segments SS1 and SS2 are assumed to form the walls of the ion pore. According to Catterall, (1995) with permission.

The large subunit is generally capable of forming an ion pore alone. In the large subunits of the $Na^+$ and $Ca^{2+}$ channels, four homologous domains can be identified that each have six potential transmembrane helices. The large subunit of the $K^+$ channel has only one of these domains. Despite this, it fits into the structural principle of the $Na^+$ and $Ca^{2+}$ channels since the large subunit is present as a tetramer.

The associated subunits often have a stabilizing effect on the pore formed by the large subunit and they determine the details of activation and deactivation of the pore to a large extent. Furthermore, the associated subunits have phosphorylation sites and it is assumed that these can have a regulating effect on pore function.

### 16.3.4 Structural Basis of Ion Channel Function

The successful crystallization and structural determination of the constitutively open $K^+$ channel of *Streptomyces lividans* (Doyle et al., 1998) represented a large step forward in the understanding of function and selectivity of ion channels.

The $K^+$ channel is a tetramer in which each subunit has two α-helical transmembrane elements. One helix of each subunit forms the inner coating of the pore whilst the other helices form contacts to the phospholipid bilayer, via hydrophobic residues. The helices are in the form of an „inverted teepee" with the broad opening oriented to the outside, into the extracellular region (Fig. 16.6). The loops between the helix pairs are on the extracellular side and are oriented inwards to the pore. These form the selectivity filter that discriminates between $K^+$ and $Na^+$. The narrow selectivity filter, which is only 12 Å long, joins a large hydrophobic cavity and the inner pore, which is hydrophilically coated.

The following points are important for selectivity and ion passage:

- Two $K^+$ ions are bound in a dehydrated form in the selectivity filter. Repulsion between the two ions promotes ion conduction.
- Oxygen atoms of the main chain of the selectivity filter coordinate the bound $K^+$. Exact coordination of the smaller $Na^+$ is not possible, due to the high rigidity of the pore.
- A large water-filled cavity and the macrodipole of the α-helices are positioned so that electrostatic destabilization of the ion in the middle of the phospholipid double layer is reduced.
- The walls of the cavity and the inner pore are hydrophobic, ensuring rapid diffusion of the ion.

It is to be expected that these basic principles of ion selectivity and ion conduction for the $K^+$ channel of *Streptomyces lividans* also apply to other types of ion channels.

### 16.3.5 Voltage-dependent Activation

It is generally agreed that activation of ion channels is caused by a voltage-induced conformational change that opens a transmembrane pore. It is assumed that the ion channel contains a structural element that can register changes in the electrical field.

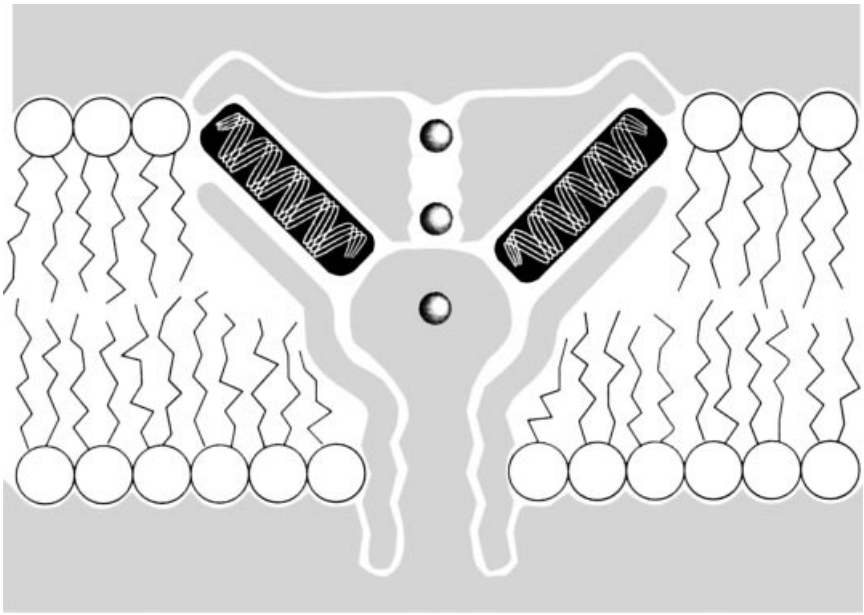


**Fig. 16.6.** Schematic view of the $K^+$ channel from *Streptomyces lividans*
The diagram depicts the overall form of the $K^+$ channel and illustrates the mechanisms by which the channel stabilizes a cation in the middle of the membrane. Two ions are bound in the selectivity filter and repulsion between them promotes ion passage. A large aqueous cavity stabilizes a cation in the otherwise hydrophobic interior. A further contribution to ion stabilization comes from the macrodipoles of oriented α-helices whose negative ends point to the cavity where a cation is located (According to Doyle et al. (1998), with permission).

The structural element functions as a *voltage sensor*, which experiences a force due to depolarization that changes its conformation and/or position relative to the membrane. The voltage sensor contains charged amino acids. The charges of these amino acids move in the electrical field of the membrane during depolarization and cause a *gating current*. Such a gating current can be demonstrated experimentally in the form of an outward movement of positive charges occurring before passage of ions.

Primary structure analysis and targeted exchange of amino acids led to the prediction that the transmembrane helix S4 forms the voltage sensor. The S4 helix contains a repeating motif of a positively charged amino acid and two hydrophobic amino acids (Fig. 16.7). The positive charges are seen as the carrier of the outward gating current. It is assumed that the positive charges are fixed by negative charges in various ways in the closed and open state.

The exact mechanism by which the depolarization leads to movement of the voltage sensor is not known. A simple model is under discussion, in which the S4 helix turns outwards by one helix turn during opening and thus leads to outward transport of 1—2 charges. A more complex model assumes a conformational change of the S4 helix in which the outward transport of charges is associated with conversion of a α-helix into a β-sheet structure.

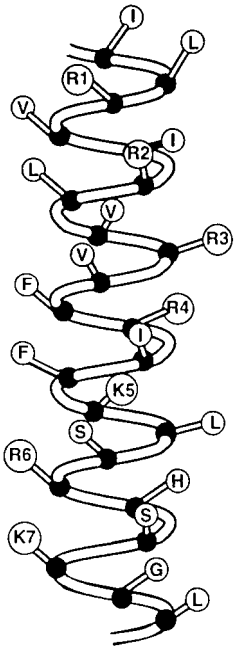



**Fig. 16.7.** Structure of the postulated voltage sensor in the S4 helix. The S4 segment of the $K^+$ channel is shown in an assumed helical configuration with the individual amino acids given as one-letter codes. The regular sequence of a basic amino acid and two hydrophobic amino acids is characteristic for the S4 segment. The basic amino acids are numbered from the *N* to the *C* terminus. According to Catterall, (1995) with permission.

### 16.3.6 Ion Passage and Pore Walls

All structural models of the voltage-gated ion channels assume that the pore walls are formed by structural sections of the four homologous domains. Information on the sequence sections of the domains that coat the pore walls is of great interest with respect to the mechanism of ion passage and ion selectivity. Due to the lack of direct structural information, all binding studies so far have used ion channel blocking agents to identify the amino acids involved in formation of the pore walls.

Blocking of the $Na^+$ channel from the extracellular side can be performed with tetrodoxin, a fish poison, or with tetraethylammonium ions. Binding studies with these agents have helped to identify two structural elements known as the SS1 and SS2 segments (Fig. 16.5d). Amino acids have been identified in the SS1 and SS2 segments that play an important role in both ion passage and ion selectivity. Binding studies with pharmaceuticals that bind to the ion channel from the intracellular side have also identified regions of the SS1 and SS2 segments as part of the pore walls. In addition, amino acid residues of the S6 helix are involved in formation of the pore opening on the cytosolic side.

### 16.3.7 Inactivation of Voltage-gated Ion Channels

After activation and opening, the ion channel enters an inactive state. For the duration of inactivation, which can last some msec to several seconds, the ion channel is closed and cannot be opened again.

Two related models describing the rapid inactivation process of the $Na^+$ channel or $K^+$ channel are generally accepted and are backed up experimentally. Both models are based on conformational changes in which flexible structural elements of the central subunit align themselves in the entrance to the pore from the intracellular side and temporarily block it.

**Inactivation of the $Na^+$ Channel**

The model of inactivation of the $Na^+$ channel assigns the inactivating function to a 15—20 amino acid segment of a loop that binds domains 3 and 4 with one another. It is assumed that this segment aligns itself in the pore entrance from the intracellular side, blocking the pore with hydrophobic residues of the loop (Fig. 16.8). The inactivating segment also contains several positively charged residues and it is assumed that these bind to negatively charged residues of the pore entrance during inactivation. According to this model, the intracellular pore entrance contains a specific binding site for the inactivating segment.

**Inactivation of the $K^+$ Channel**

An N-terminal sequence section with an autoinhibitory function is involved in inactivation of the $K^+$ channel. The model assumes that the N-terminal sequence is aligned in the intracellular pore entrance for inactivation of the $K^+$ channel. The model assigns the structure of a globular domain to the N terminus, linked to the channel via a flexible chain of ca. 200 amino acids. According to this hypothesis, the N terminus has the structure of a „ball“ linked to the pore entrance by a „chain“. During inactivation, the „ball“ aligns itself in the entrance of the pore and closes it. Like the $Na^+$ channel, the inactivating structural element of the $K^+$ channel has hydrophobic and positively charged residues that are important for inactivation. Here, a specific binding site has also been identified for the inactivating structural element on the side of the pore entrance.

The function of the ion channels is also regulated by phosphorylation of intracellular structural elements. The biochemical basis of this regulation is poorly understood.

## 16.4 Ligand-gated Ion Channels

### 16.4.1 Neurotransmitters and Mechanisms of Ligand-gated Opening of Ion Channels

Ligand-gated ion channels are at the center of signal transduction at chemical synapses. The ligands involved are known as neurotransmitters (see Fig. 16.1). These bind as chemical messengers to corresponding receptors on the surface of a target cell and bring about opening of an ion channel at that site (see Fig. 16.2).

The neurotransmitters are comparable to hormones in their mode of action. Both are extracellular ligands for corresponding surface receptors. The characteristic diffe-

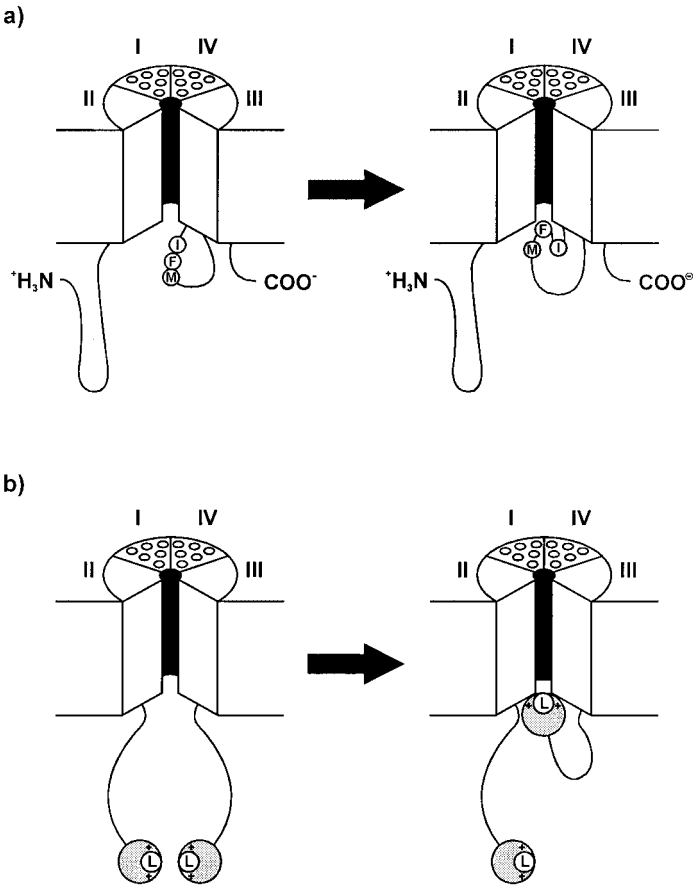



**Fig. 16.8.** Model of inactivation of voltage-gated $Na^+$ and $K^+$ channels. a) Inactivation of the $Na^+$ channel. On inactivation of the $Na^+$channel, the loop, which binds domain III and domain IV of the α-subunit, positions itself in the cytoplasmic entrance of the pore and closes it. The indicated hydrophobic amino acids of the connecting loop are involved in the inactivation. b) Inactivation of the $K^+$ channel. The model assumes that a compact structural part of the C terminus of the β subunit is aligned in the pore and transiently closes it. The inactivating structural part is linked to the pore via a flexible structural element and contains a functionally important leucine residue and a lot of positive charges. According to Catterall, (1995).

rences are in the spatial relationship between ligand release and the receptor. In contrast to hormones, neurotransmitters are released directly in the vicinity of the receptor and are only locally active. In the presynaptic cell, the neurotransmitters are stored in vesicles and are released on arrival of an electrical signal. Release of the neurotransmitter takes place by a complex mechanism involving many protein components. Following diffusion across the synaptic cleft, the neurotransmitter binds to the receptor on the postsynaptic cell. Binding to the receptor opens the ion channel and specific ion passage occurs, which is registered as a signal and processed further (review: Jessel and Kandel, 1993; Kelly, 1993; Unwin, 1993).

The processes at the receptor of a postsynaptic cell are only one example of regulation of ion channels by specific ligands. The cell also has other means of controlling the open state of ion channels, with the help of specific messenger substances (Fig. 16.9).

### a) Binding of Messengers to Receptors with Intrinsic Ion Channel Function

Binding of extracellular messengers to the receptor opens an ion channel, which, in this case, is a part of the receptor. The receptor and the ion channel form one structural unit (see 16.4.3, *Acetylcholine receptor*). This permits very fast translation of the signal „ligand binding“ into the signal „ion passage“.

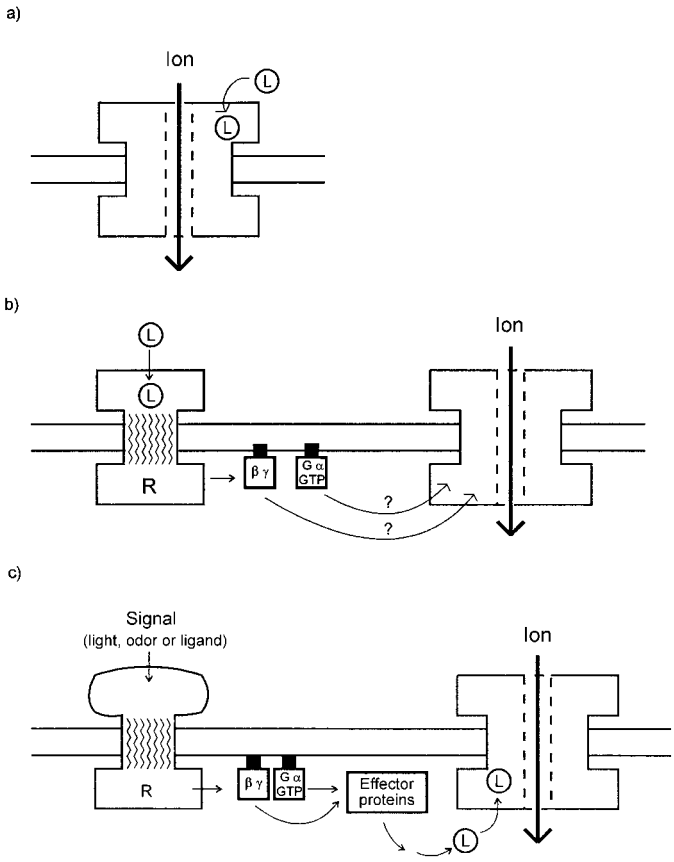



**Fig. 16.9.** Control of ion channels via messenger substances. a) Ligand-gated opening of an ion channel by a messenger substance: the ion channel contains a specific binding site for a ligand. Ligand binding induces opening of the ion channel. b) Direct coupling between ion channel and a G-protein. Activation of G-protein-coupled receptors and the corresponding heterotrimeric G-proteins is directly linked to opening of the ion channel in this case. c) Control of the open state by intracellular messengers. A G-protein-mediated signaling pathway leads to activation of effector proteins that form diffusible messenger substances. The messengers bind to binding sites on the cytoplasmic side of the ion channel and induce its opening.

**b) Direct Coupling between G-proteins and Ion Channels**

Ion channels can also be opened by a direct effect of an activated G-protein. In this case, the extracellular messenger binds to a G-protein-coupled transmembrane receptor and activates a heterotrimeric G-protein. This acts directly on an ion channel and regulates its open state.

The ion channel and the receptor are separate signal elements here, with functional states coupled via the heterotrimeric G-protein. The mechanism of opening of ion channels by G-proteins is unknown.

An example of a receptor coupled to an ion channel is the muscarinic acetylcholine receptor, which brings about opening of a $K^+$ channel on agonist binding. There is evidence that the $\beta\gamma$ subunits of the corresponding heterotrimeric G-protein are involved in opening of the ion channel.

**c) Intracellular Opening of Ion Channels by Diffusible „Second Messenger" Molecules**

Another mechanism for ion channel opening by ligands was discovered in the signal transduction cascade of light and odor perception in mammals. Processing of the sensory signals includes a process leading to opening of ligand-controlled ion channels from the intracellular side (review: Yau, 1994). During light perception, the „second messenger" cGMP serves as a ligand; during odor perception, cAMP is used.

The course of signal transduction during light perception is shown in a simplified form in Fig. 16.10.

The light signal is firstly registered by the photoreceptor (= rhodopsin). The receptor is activated and conducts the signal via the corresponding G-protein (transducin) to the downstream effector protein. The effector molecules are cGMP phosphodiesterases during light perception; during odor perception, these are cAMP phosphodiesterases. As a consequence of activation of the phosphodiesterase, there is a reduction in the concentration of cGMP. cGMP is the ligand of an ion channel that permits entry of $Ca^{2+}$ and $Na^+$ into the cell interior. Binding of cGMP to the ion channel induces its opening. If the concentration of cGMP falls as a consequence of light perception, there is a sharp drop in the intracellular $Na^+$ and $Ca^{2+}$ concentration and thus a hyperpolarization, which is passed on as a signal.

Further examples of intracellularly activated ion channels are the $InsP_3$ receptors and the ryanodin receptors (see Chapter 6).

## 16.4.2 Neurotransmitter-controlled Receptors with Intrinsic Ion Channel Function

At least two classes of neurotransmitter-controlled receptors with intrinsic ion channel function can be differentiated. One class includes receptors with the specific ligands acetylcholine, $\gamma$-aminobutyric acid (GABA), glycine and serotonin.

The other class is formed by the receptors for glutamate. These are divided on the basis of their activation by specific ligands into NMDA (N-methyl-D-aspartate), kainate and AMPA ($\alpha$-amino-3-hydroxy-5-methylisoxazole-4-propionic acid) receptors.

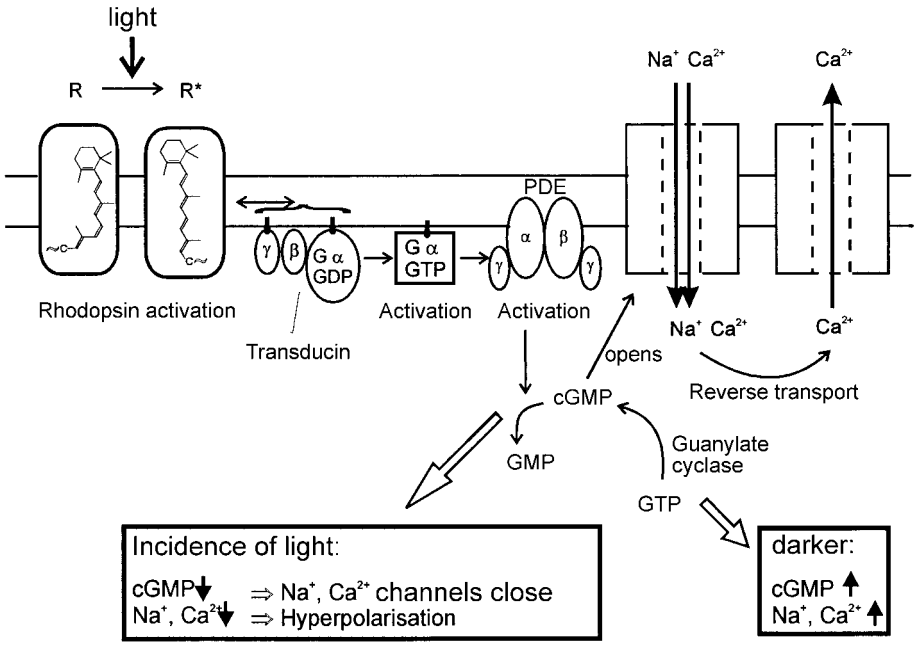



**Fig. 16.10.** Signal transduction during light perception in vertebrates. Important steps in signal transduction during light perception. The figure depicts the central function of an intracellularly activated ion channel. Rhodopsin (R) is a photoreceptor of the eye that is activated by light. The basis of the activation of rhodopsin is the light-induced *cis*/*trans* isomerization of rhodopsin-bound retinal. The activated rhodopsin (R*) passes the signal on to the heterotrimeric G-protein transducin. The effector protein of the activated $G_{t,\alpha} \cdot GTP$ subunit is a cGMP-specific phosphodiesterase (PDE) that cleaves cGMP to GMP. The PDE exists as a heterotetrameric complex, composed of α, β and two γ subunits. As a consequence of the activation of PDE, there is a decrease in the cGMP concentration. cGMP is the intracellular ligand of a ligand-controlled ion channel that permits passage of $Na^+$ and $Ca^{2+}$ ions through the membrane. The decrease in the cGMP concentration is associated with closing of the ion channel, decrease in local $Na^+$ and $Ca^{2+}$ concentration and with hyperpolarization. The hyperpolarization is registered as a light stimulus and is processed further. In darkness, the cGMP concentration increases due to new synthesis of cGMP by a guanylate cyclase. cGMP binds to the ion channel, induces its opening and an influx of $Na^+$ and $Ca^{2+}$ takes place. $Ca^{2+}$ is transported back into the extracellular region via specific $Ca^{2+}$ transport systems.

### 16.4.2.1 The NMDA Receptor

The NMDA receptor belongs to the glutamate-controlled receptors (review: Hollmann and Heinemann, 1994). Glutamate is the most important excitatory neurotransmitter in the brain of mammals. By binding to specific receptors, glutamate mediates ion passage through postsynaptic membranes. Of the glutamate receptors, the NMDA receptor has won special attention because it functions both in synaptic transmission and in regulation of synaptic plasticity. The NMDA receptors are involved in activity-

dependent remodelling of the synapses during development of the brain. The NMDA receptors also have a central function in performance of the mature nervous system, since they are needed for a form of activity-dependent amplification of synaptic transmission known as „long-term potentiation".

Binding of glutamate opens the ion channel of the receptor and permits passage of ions, especially $Ca^{2+}$. The receptor is named after its specific activation by the (non-physiological) ligand N-methyl-D-aspartate.

The NMDA receptor can be characterized by the following properties as a receptor that is controlled by ligand binding and by potential changes.

The most important characteristics of the NMDA receptor are:

- Activation by glutamate binding *in vitro* and in *vivo*
- Glutamate activation requires binding of glycine at a specific binding site.
- Potential-dependent blocking by $Mg^{2+}$: in the presence of $Mg^{2+}$, the ion flux is blocked at potentials more negative than –20 mV. The receptor can also be blocked by $Zn^{2+}$.
- High permeability for $Ca^{2+}$
- Regulation by Tyr phosphorylation in cytosolic structural elements

A functional NMDA receptor is a hetero-oligomer, composed of the subunit NR1 plus one of the subunit types NR2A, NR2B, NR2C or NR2D. It is assumed that each subunit contains 4 transmembrane elements and that two of each subunit align to form a heterotetramer. In addition, further proteins are associated on the cytoplasmic side and these are needed for the physiologic function of the NMDA receptor (Fig. 16.11). One of these proteins is the PSD-95 protein, which contains three PDZ modules (see Chapter 8). The C-terminal cytoplasmic tail of NR2 binds via the sequence –ESVD specifically to the PDZ module 1 or 2 of PSD-95.

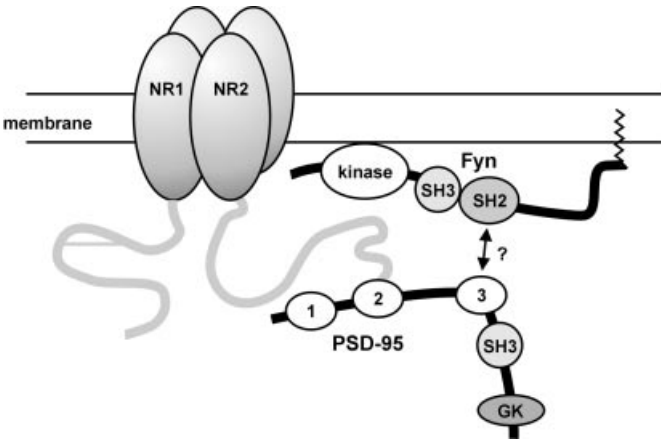



**Fig. 16.11.** Model of the association of Fyn kinase with the NMDA receptor
The NMDA receptor is shown as a tetramer of NR1 and NR2 subunits. The C-terminal tail of NR2 interacts with PDZ2 of PSD-95. The protein tyrosine kinase Fyn is assumed to bind to PDZ3 of PSD-95 via its SH2 domain. Fyn also is anchored to the cell membrane via its myristoylated N-terminus. GK: guanylate kinase domain of PSD-95. According to Sala and Sheng (1999), with permission.

It is assumed that the PSD-95 protein functions as an organizer of large protein complexes in the postsynaptic membrane, since other signal proteins such as a NO synthase and a Ras-GAP protein can also specifically bind to the PDZ domain or other domains of PSD-95. In this way, the NMDA receptor is coupled via PSD-95 to other signal proteins in the cell, whereby the NMDA receptor is directly linked to other signal conduction.

Like many other ligand-gated ion channels, the channel gating properties of the NMDA receptor are regulated by phosphorylation. The NR2 subunit of the receptor contains many Tyr phosphorylation sites in the cytoplasmic region and it has been shown that receptor activity is stimulated by Tyr phosphorylation. Cytoplasmic nonreceptor tyrosine kinases of the Src family are responsible for the Tyr phosphorylation. Thus, a specific association with PSD-95, and thereby mediation of phosphorylation of NR2, has been demonstrated for the Fyn kinase, a member of the Src kinase family (Tezuka et al., 1999). Thus, the nonreceptor tyrosine kinase has an important role in regulation of synaptic activity and plasticity. A model for Fyn association with the NMDA receptor/PSD95 complex is shown in Fig. 16.11.

#### 16.4.2.2 The Nicotinic Acetylcholine Receptor

The *receptors for acetylcholine, GABA, glycine and serotonin* have a homologous structure and form a *superfamily*. The best-investigated representative of this superfamily is the nicotinic acetylcholine receptor, for which extensive biochemical and structural data are available. The nicotinic acetylcholine receptor can be treated as a representative of the other receptors of the superfamily since it can be assumed that the structure-function principles of this receptor apply to the others.

There are two large groups of acetylcholine receptors, which differ in their inhibition properties. The *nicotinic acetylcholine receptors* are inhibited by nicotine whereas the members of the other group are inhibited by the fungal alkaloid muscarine and are therefore known as *muscarinic acetylcholine receptors*. The muscarinic acetylcholine receptors belong to the G-protein-coupled receptors with seven transmembrane helices (see Chapter 5).

The nicotinic acetylcholine receptor is a protein of 290 kD that occurs in chemical synapses where communication between nerve cells and muscle cells takes place. Binding of acetylcholine to the receptor induces opening of the ion channel, which is a part of the receptor. Passage of $Na^+$ and $K^+$ ions through the receptor takes place and depolarization of the postsynaptic cell occurs. The depolarization represents a signal that—according to the nature of the postsynaptic cell—is processed in various ways.

The structure of the acetylcholine receptor is shown in Fig. 16.12. The nicotinic acetylcholine receptor is a pentameric protein complex with the subunit structure $\alpha_2\beta\gamma\delta$. Each of the two $\alpha$-subunits possesses a binding site for acetylcholine. Four potential transmembrane helices can be identified in each of the five subunits and these are known as M1—M4.

In addition, each subunit has an amphipathic helix, the MA helix. It is generally assumed that the inner wall of the pore is formed by the M2 helices.

High-resolution electron microscope pictures of the acetylcholine receptor and mutagenesis studies have enabled a detailed view of the overall structure to be obtai-

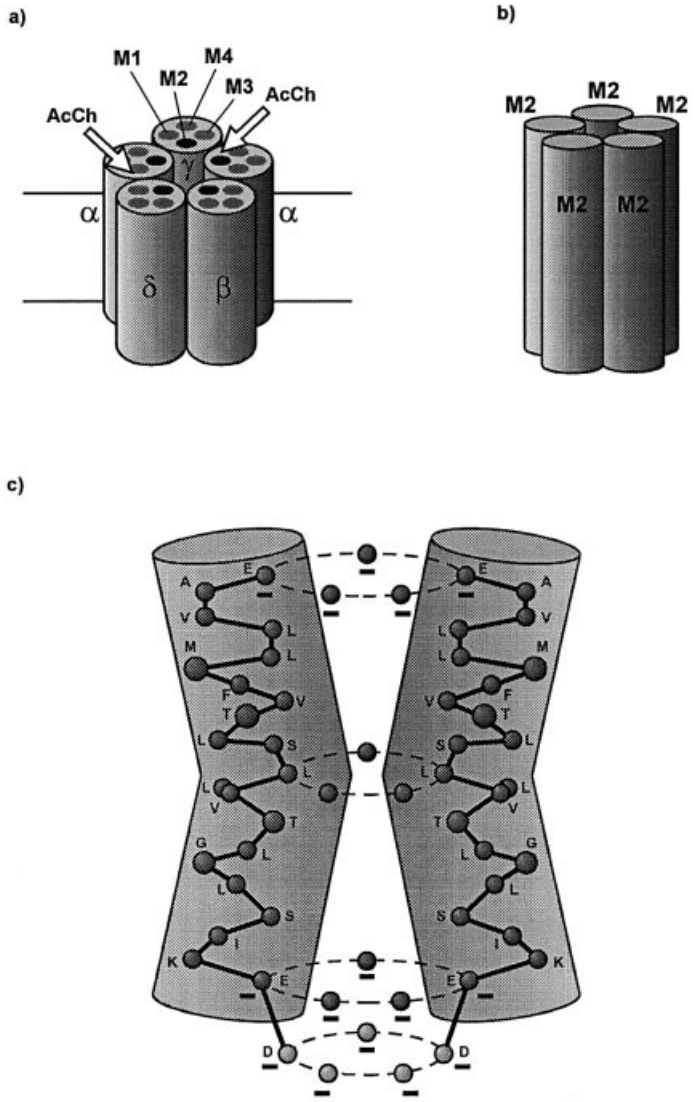



**Fig. 16.12.** Subunit structure of the acetylcholine receptor. a) The acetylcholine receptor has the subunit structure $\alpha_2\beta\gamma\delta$. The four transmembrane elements M1—M4 are shown for the $\gamma$ subunit. The binding sites for acetylcholine (ACh) are located on the $\alpha$-subunits. b) It is assumed that the inner wall of the ion pore is formed by M2 helices of the five subunits. c) Postulated configuration of the M2 helices in the narrowest region of the ion channel. In the closed state, five leucine residues (one per subunit) lie in the ion channel and hinder passage of ions. Above and below the block, there are negatively charged residues that serve as prefilters for ion passage.

ned. The acetylcholine receptor has a funnel-like structure with a relatively large extracellular structural portion and a smaller intracellular region (Fig. 16.13). The funnel narrows from the outer side of the membrane to the narrowest part, which is located

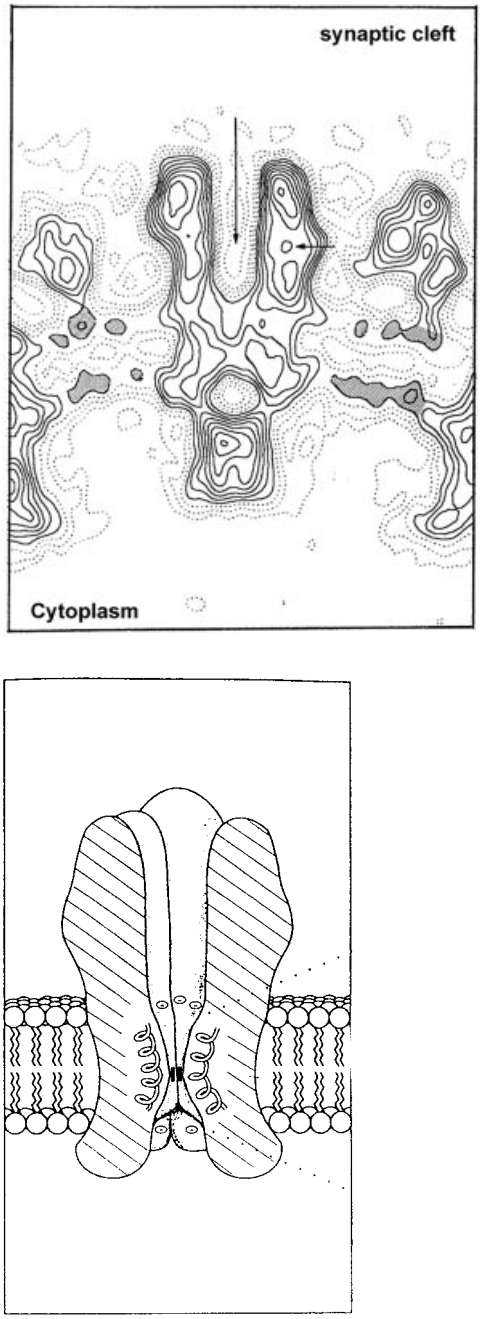



**Fig. 16.13.** Pore structure of the acetylcholine receptor, based on electron microscopy studies. a) Electron density map of the acetylcholine receptor of the postsynaptic membrane of the electric organ of the ray *Torpedo californicus*, based on electron microscopy studies. The receptor has a long funnel-like structure in the extracellular region, which narrows at the center of the pore. A smaller funnel structure is observed in the cytoplasmic region of the receptor. Another protein is situated on the cytoplasmic side. The long arrow indicates the direction of ion passage and the small arrow shows the postulated binding site for acetylcholine. b) Schematic representation of the acetylcholine receptor with the M2 helix as the central block in the ion channel. According to Unwin, (1993).

at the level of the lipid membrane. In the inner wall of the funnel, negatively charged amino acids are frequently found above and below the narrowest site and these are aligned in a ring in the funnel. It is assumed that the negative charges serve as a pre-filter for cation entry. The binding sites for acetylcholine are located on the extracellular side of the receptor and are 2—3 nm away from the narrowest part of the pore.

The electron microscope studies show that the inner wall of the pore is formed by five α-helices. These are probably the M2 helices. The helices are surrounded by a star-shaped structure, which is possibly made up of β-sheet structures.

### Structural Differences in the Open and Closed States

First insight into the structural differences of the acetylcholine receptor in the open and closed states was obtained with the electron microscope (Unwin, 1995). Significant structural changes in the central part of the pore became visible due to binding of acetylcholine (Fig. 16.14).

In the closed state, the M2 helices are crooked with the bend oriented towards to center of the pore. It is assumed that the bulky Leu residues are found at the center of the bend; these are thought to be responsible for closing the ion channel. Above and below the block, the channel expands in a funnel form.

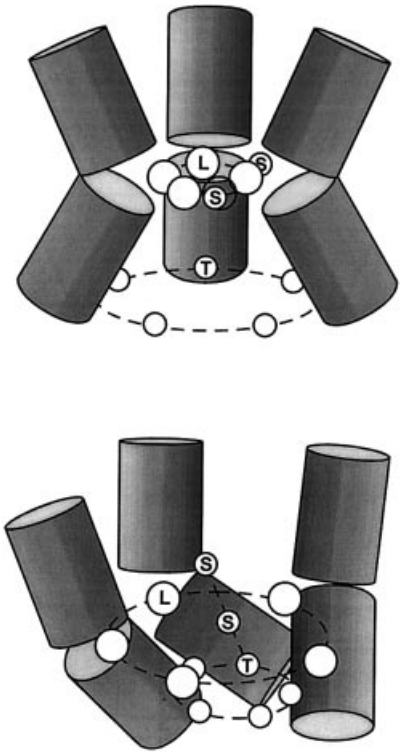



**Fig. 16.14.** Configuration of the M2 helices of the acetylcholine receptor in the closed and open states. The schematic representation is based on a comparison of the electron density map of the acetylcholine receptor in closed and open states. Only three of the five M2 helices are shown. a) Closed state: the M2 helices are bent at the middle. The leucine residues point into the interior of the pore and prevent passage of ions. b) Open state: the M2 helices are turned outwards at a tangent and the bulky leucine residues are removed from the center of the pore. Reorientation of the M2 helices causes a reorientation of polar amino acids that coat the interior of the pore. The polar amino acids (Ser and Thr residues) are oriented closer to the center of the pore and create a hydrophilic coating of the pore inner wall, which facilitates ion passage. According to Unwin, (1995) with permission.

On binding of acetylcholine, small *conformational changes* are observed in the region of the acetylcholine binding sites. These propagate into the center of the pore where larger structural changes are triggered in the M2 helices. The M2 helices rotate and turn at a tangent to the outside. The Leu residues of the block are rotated out of the center of the pore and the channel opens. Through the conformational changes in the M2 helices, polar residues of these helices are exposed to the inner wall of the pore. It is assumed that the polar residues interact with the hydrate shells of the cations and function as a cation-selective filter.

# References Chapter 16

Catterall, W.A. 'Structure and function of voltage-gated ion channels' (1995) Annu. Rev. Biochem. *64*, 493–531

Clapham, D.E. 'Direct G protein activation of ion channels?' (1994) Annu.Rev.Neurosci. *17*, 441–464

Doyle, D.A, Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L. Chait, B.T. and MacKinnon, R. 'The structure of the potassium channel: molecular basis of K+ conduction and selectivity' (1998) Science *280*, 69–77

Hollmann, M. and Heinemann,S. 'Cloned glutamate receptors' (1994) Annu. Rev. Neurosci. *17*, 31 -108

Jan, L.Y. and Jan, Y.N. 'Receptor-regulated ion channels' (1997) Curr. Op. Cell Biol. *9*, 155–160

Jessel, T.M. und Kandel, E.R. 'Synaptic transmission: a bidirectional and self-modifiable form of cell-cell communication' (1993) Cell *72*, 1–30

Kelly, R.B. 'Storage and release of neurotransmitters' (1993) Cell *72*, 43–52

Sala, C. and Sheng, M., 'The fyn art of *N-methyl*-D-aspartate receptor phosphorylation' (1999) Poc. Natl. Acad. Sci. U.S.A. *96*, 335–337

Tezuka, T., Umemori, H., Akiyama, T., Nakanishi, S. and Yamamoto, T., PSD-95 promotes Fyn-mediated tyrosine phosphorylation of the N-methyl-D-aspartate receptor subunit NR2A' (1999) Proc Natl Acad Sci U S A *96*, 435–440

Unwin, N. 'Neurotransmitter action: Opening of ligand-gated ion channels' (1993) Cell *72*, 31–41

Unwin, N. 'Acetylcholin receptor channel imaged in the open state' (1995) Nature *373*, 37–43

Yau, K.W. 'Cyclic nucleotide-gated channels: an expanding new family of ion channels' (1994) Proc. Natl. Acad. Sci. U.S.A. *91*, 3481–3483