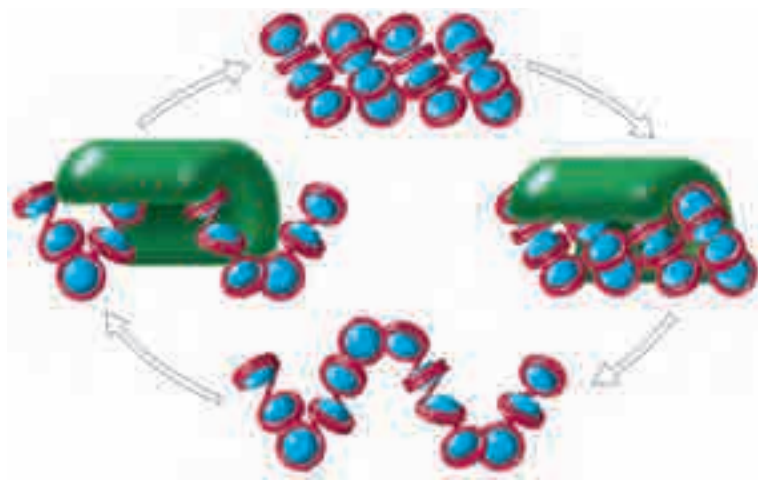


Fritz Höffeler

Bildatlas *Genexpression*



Verlag
Harri
Deutsch

Bildnachweis

Die elektronenmikroskopischen Aufnahmen auf den Seiten 30, 32, 34, 36, 37, 168, 180 und 183 stammen von Herrn Prof. Dr. Gerhard Wanner, Botanisches Institut der Ludwig-Maximilians-Universität München.

Von *eye of science* wurden die kolorierten EM-Aufnahmen der *Drosophila*-Mutanten auf Seite 148 gemacht.

Für die Photos der Viren auf den Seiten 152 und 153 danke ich Herrn Dr. Hans Gelderblom, Robert Koch-Institut, Berlin.

Die rasterkraftmikroskopischen Aufnahmen auf Seite 33 und 153 erstellten Herr Prof. Dr. Hans Oberleithner und seine Mitarbeiter, Westfälische Wilhelms-Universität Münster.

Die Polytäanchrosomen auf Seite 168 photographierte Herr Dr. Christian Laforsch, Department Biologie II, Ludwig-Maximilians-Universität München.

Alle Illustrationen und Graphiken wurden angefertigt durch *Art For Science*, Hamburg.

Autor

Fritz Höffeler, Jahrgang 1964, hat in Münster und Bielefeld Biologie studiert und war für mehrere Jahre im Zentralbereich F & E im Labor für Biotechnologie der Lurgi AG in Frankfurt am Main angestellt. Einige Jahre später studierte er zusätzlich in Hamburg Erziehungswissenschaft mit dem Schwerpunkt Didaktik. 1999 gründete er *Art For Science*, ein Atelier für wissenschaftliche Illustrationen, und arbeitet seitdem als Illustrator und Wissenschaftsvermittler.

Der Autor ist erreichbar unter: Fritz.Hoeffeler@Art4Science.de

Die Webseite zum Buch

<http://www.harri-deutsch.de/1885.html>

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-8171-1885-4

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdrucks und der Vervielfältigung des Buches – oder von Teilen daraus – sind vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder ein anderes Verfahren), auch nicht für Zwecke der Unterrichtsgestaltung, reproduziert oder unter Verwendung elektronischer Systeme verarbeitet werden.

Zu widerhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

Der Inhalt des Werkes wurde sorgfältig erarbeitet. Dennoch übernehmen Autor und Verlag für die Richtigkeit von Angaben, Hinweisen und Ratschlägen sowie für eventuelle Druckfehler keine Haftung.

1. Auflage 2011

© Wissenschaftlicher Verlag Harri Deutsch, Frankfurt am Main, 2011

Umschlaggestaltung: Claudia Holz

Lektorat: Manuela Kupfer

Druck: fgb • freiburger graphische betriebe <www.fgb.de>

Printed in Germany

Vorwort

Ziel eines Repetitoriums ist, größtenteils bereits bekanntes Wissen auf Wesentliches zu komprimieren und in einer übersichtlichen und überschaubaren Form zu präsentieren: Übersichtlichkeit durch Reduktion und Verständnisförderung durch Fokussierung.

Lehrbuch und Repetitorium stellen somit keine alternativen, sondern einander ergänzende Lernsysteme dar.

Der Themenschwerpunkt des vorliegenden Repetitoriums ist die Genexpression. Der Leser und Lerner wird durch aufeinander aufbauende Kapitel mit aufsteigender Komplexität hin zu komplizierten Systemen wie der Embryonalentwicklung von *Drosophila* geführt. Strukturen, Sequenzelemente und Prozesse, die mehreren thematischen Zusammenhängen zuzuordnen sind, werden zusätzlich im Kapitel „Ergänzungen“ intensiver beleuchtet.

Wo immer es sinnvoll und möglich erschien, wurden die Daten kristallographischer Strukturuntersuchungen als Grundlage für symbolische und stilisierte Darstellungen von Komplexen herangezogen.

Hamburg, im Sommer 2011

Fritz Höffeler

Danksagung

Ganz herzlich möchte ich mich bei Herrn Dr. Gelderblom bedanken – sowohl für die Überlassung der Photos als auch für die Beantwortung meiner Fragen. Großer Dank gebührt meiner Lektorin Frau Manuela Kupfer, die mir bereits beim Band *Cytologie* zur Seite stand. Ihr scharfer Verstand und ihre scharfen Augen bewahrten mich vor meiner eigenen Betriebsblindheit und den Tücken der Orthographie- und Interpunktionsregeln. Die Verlagsleitung hatte sehr viel Geduld mit mir, denn die Arbeit an diesem Buch hat enorm viel Zeit in Anspruch genommen. Auch ihr schulde ich Dank!

Gewidmet meinem ehemaligen Biologielehrer Herrn Günter Jarmer,
dessen Lebenskreis sich viel zu früh geschlossen hat.

I. DIE MOLEKÜLKLASSEN

1. Struktur-Funktions-Beziehungen.....	2
2. Proteine.....	4
3. Nucleinsäuren.....	10
4. Interaktionen.....	16

II. SEQUENZELEMENTE

1. Allgemeines.....	18
2. Replikationsursprünge.....	20
3. Centromere.....	21
4. Telomere.....	22
5. Gene.....	23
6. Promotoren.....	24
7. Regulatorregionen.....	25
8. Transponierbare Elemente.....	26
9. Repetitive Elemente.....	27
10. Weitere Elemente.....	28

III. CHROMOSOMALE ELEMENTE

1. Übersicht.....	29
2. Nucleoide.....	30
3. Chromosomen.....	32
4. Nucleoide der Organellen.....	36
5. Virale Chromosomen.....	38
6. Plasmide.....	39
7. Genome.....	40

IV. ZUSAMMENFASSUNG

Strukturelle Grundlagen.....	42
------------------------------	----

V. GENETISCHE MECHANISMEN

1. Übersicht.....	44
2. Replikation.....	46
- Strukturen.....	46
- Prozess.....	48
- Replisom.....	50
- Telomere.....	51
- Chromatinumformung.....	52
- Regulation, Besonderheiten.....	53
3. Transkription.....	54
- Strukturen.....	54
- Prozess.....	56
- Initiationskomplexe.....	58
- Initiation der Pol II.....	59
- Termination.....	60
- Chromatinumformung.....	61
- Transkripte.....	62
4. Processing.....	63
- Übersicht.....	63
- Additionen.....	64
- Modifikationen.....	65
- Editing.....	66
- Spleißen.....	70
- rRNA.....	76
- tRNA.....	77
- mRNA.....	78
- Zusammenfassung.....	79
5. RNA-Export.....	80
6. Translation.....	84
- Genetischer Code.....	84
- Strukturen.....	86
- Prozess.....	90
- Initiation.....	92
- Recodierungen.....	93
7. Zusammenfassung.....	94

VI. EXPRESSIONSREGULATION

1. Übersicht.....	96
2. DNA.....	98
- Amplifikation.....	98
- Transposition.....	99
- Inversion.....	100
- Methylierung.....	101
3. Chromatin.....	104
- Prokaryoten.....	104
- Eukaryoten.....	105
- Dosiskompensation.....	106
4. Transkription.....	108
- Übersicht.....	108
- Prokaryoten.....	109
- Eukaryoten.....	112
5. Processing.....	117
- Poly-A-Stellen.....	117
- Alternatives Spleißen.....	118
- Editing.....	120
6. RNA-Abbau.....	121
- Übersicht.....	121
- Allgemeine Abbauwege.....	122
- RNA-Interferenz.....	124
- Nonsense Mediated Decay.....	126
7. Translation.....	128
- Prokaryoten.....	128
- Eukaryoten.....	129

VII. SIGNALE UND KASKADEN

1. Allgemeines.....	134
2. Lichtinduktion.....	135
3. Immunantwort.....	136
4. Geschlechtsbestimmung.....	138
5. Embryonalentwicklung von <i>Drosophila</i>	144

VIII. VIREN

1. Allgemeines.....	150
2. Bau und Bestandteile.....	152
3. RNA-Viren.....	154
4. Retroviren.....	156
5. DNA-Viren.....	158

IX. ERGÄNZUNGEN

1. RNAs.....	160
2. Histone.....	164
3. Chromatin.....	166
4. Interphase-Chromosomen.....	168
5. DNA-Polymerasen.....	170
6. RNA-Polymerasen.....	172
7. Helicasen.....	175
8. Topoisomerasen.....	176
9. Promotoren.....	178
10. Ribosomen.....	180
11. Translationsfaktoren.....	184
12. Replikation.....	185
13. Regulatorproteine.....	186
14. Geschlechter und Sexualität.....	189
15. Transponierbare Elemente.....	190
16. Mengen und Zahlen.....	194
17. mtDNA.....	195
18. Gene und Genome.....	196
- Gentransfer.....	198

ANHANG

Abkürzungen.....	200
Stichwortregister.....	204
Literaturverzeichnis.....	212

Der Seitenbau ist so konzipiert, dass auf jeder Seite die hierarchische Themenorganisation des konkret behandelten Inhalts ersichtlich ist:

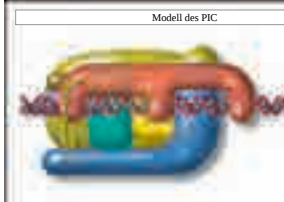
I: Hauptthema und Unterthema

II: Übergeordneter Themenkomplex mit Farbmarkierung

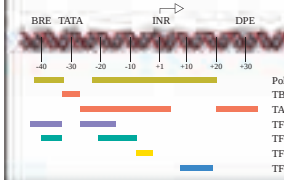
3. TRANSKRIPTION: Initiation der Pol II
V. GENETISCHE MECHANISMEN

INITIATIONSKOMPLEX DER EUKARYOTISCHEN RNA-POLYMERASE II

Modell des PIC



Protein-DNA-Kontaktbereiche



- Pol II
- TFIIA
- TFIIB
- TFIIF
- TFIIH

Aktivierung der Pol II



INITIATIONSKOMPLEX DER POL II

Der **geschlossene Initiationskomplex** (auch als PIC bezeichnet) der eukaryotischen RNA-Polymerase II enthält die geschlossene Promotorregion und den Basalkomplex, der wiederum aus der RNA-Polymerase und sechs basalen Transkriptionsfaktoren besteht. Dieser Basalapparat bildet sich entweder durch sukzessive Aggregation der Einzelkomponenten oder durch Zusammenfügen von zwei Teilkomplexen: der Pol II mit angebundenen TFs sowie der Reinitiationsplattform (s. u.). Die obere Darstellung ist ein vereinfachtes Modell, das nicht berücksichtigt, dass sich die DNA auf Grund der starken Biegung durch TBP/TFIID um diesen Basalkomplex windet. Die Positionierung und Stabilität dieses Komplexes ergibt sich aus einer Vielzahl kooperativer Protein-Protein- und Protein-DNA-Bindungen. Letztere können sich über einen großen DNA-Bereich erstrecken.

AKTIVIERUNG DER POL II

Innerhalb des PIC wird ca. 12 bp stromabwärts der TATA-Box u. a. mit Hilfe der beiden Helicase-UE des TFIIF ein Bereich von 11-15 bp entwunden. In dieser Position kann der codogene Strang leicht zum aktiven Zentrum der Polymerase geführt werden. Dieser Zu-

stand wird als **offener Initiationskomplex** bezeichnet. Es folgt die abortive Synthese mehrerer drei bis zehn Nt langer RNAs; erst ab einer Länge von ca. 30 Nt ist der Zusammenhalt von DNA, RNA und Polymerase so stabilisiert, dass die Polymerase in den Elongationsmodus wechseln kann. Um den Initiationskomplex verlassen zu können, muss die CTD, über die die Polymerase mit TFIIF verbunden ist, phosphoryliert werden. Beim Menschen besteht die CTD aus 52 Heptapeptiden, die die Konsensussequenz Tyr₁-Ser₂-Pro₃-Thr₄-Ser₅-Pro₆-Ser₇ aufweisen. Die Proteinkinase-UE des TFIIF kann jedes Heptapeptid an Ser₅ phosphorylieren. Auf Grund dieser Modifikationen löst sich die CTD vom TFIIF ab und es entstehen an der CTD gleichzeitig Bindestellen für Cappingfaktoren (S. 64). Die Pol II löst sich nun vom Initiationskomplex. Einige Komponenten des Basalapparats bleiben an die Pol II gebunden, einige lösen sich vom Komplex ab und einige können als Reinitiationsplattform am Promotor verbleiben. Diese Plattform ermöglicht einen schnellen Transkriptionsstart, indem sich ein Aggregat aus Pol II und verschiedenen TFs schnell anbinden kann. Dafür jedoch muss die CTD der neuen Pol II in einem unphosphorylierten Zustand vorliegen. Eine Phosphatase dephosphoryliert die Heptapeptide. Die CTD muss sich also für die Initiation in einem unphosphorylierten und für die Elongation in einem phosphorylierten Zustand befinden (→ S. 172, RNA-Polymerasen).

Abbildungen gehören immer zum Text der (Doppel-)seite.

59

III: Konkret behandeltes Thema

Texte enden immer am Ende einer (Doppel-)seite.

REGELN UND KONVENTIONEN

1. Sequenzen

Nucleotid- und Aminosäuresequenzen werden entlang ihrer Syntheserichtung notiert: Einzelsträngige Nucleotidsequenzen in 5' → 3'-Richtung (vom Phosphat- zum Hydroxylterminus) und Peptide in N → C Richtung (vom Amino- zum Carboxylterminus). Aminosäuren werden als Drei- oder Einbuchstaben-Codes wiedergegeben (s. S. 4) und Nucleotide (bzw. die entsprechenden Basen) als Einbuchstaben-Codes (s. S. 11).

Um bei der Notierung von Dinucleotidfolgen und von Basenpaaren Missverständnisse zu vermeiden, werden beide unterschiedlich wiedergegeben: Die Basenfolge Cytosin-Guanin wird als CpG, das Basenpaar Cytosin-Guanin als C/G oder CG formuliert. (Z. B. S. 24: Eukaryotische Promotoren enthalten sowohl GC-Boxen, die aus G/C-Basenpaaren bestehen als auch CpG-Inseln, die mehrere Dinucleotidfolgen CG enthalten.)

Einzelne Aminosäuren und einzelne Nucleotide innerhalb eines Polymers werden mit entsprechenden Positionsziffern versehen, wobei bei Peptiden vom N-Terminus und bei Nucleinsäuren vom 5'-Ende her gezählt wird. Ser³³ bezeichnet demnach die Aminosäure Serin, die sich an der 33. Position des Peptids befindet und A2450 ein Adenin, das in der 2450. Position liegt.

Um relative Positionen von Nucleotidsequenzen zu bezeichnen, werden die Begriffe stromaufwärts und stromabwärts verwendet: In der Sequenzabfolge 5'-X-Y-Z-3' liegt X stromaufwärts von (also vor) Y und Z liegt stromabwärts von (also hinter) Y.

2. Schrift

Bezeichnungen für Gene werden immer kursiv und i. d. R. in Kleinbuchstaben geschrieben.

Proteinnamen werden in Normalschrift aufgeführt, wobei sie mit einem Großbuchstaben beginnen oder vollständig in Großbuchstaben geschrieben werden können.

Falls Strukturen von Menschen und Tieren dieselben Namen tragen, werden die der menschlichen Strukturen in Großbuchstaben, die der tierischen in Kleinbuchstaben notiert.

Die Bezeichnungen von Arten, nicht jedoch von anderen taxonomischen Einheiten, werden kursiv gesetzt (z. B. *Drosophila melanogaster*; *Homo sapiens*; aber: Mammalia, Vertebrata).

3. Nomenklatur

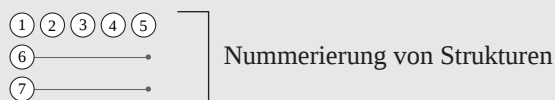
Lange und zusammengesetzte Namen werden – meist in Form eines Akronyms – abgekürzt (z. B. MPF = Maturation Promoting Factor). Gene und die darin codierten Proteine tragen i. d. R. dieselben Bezeichnungen (wie *rec A* und Rec A). Mitglieder von Gen- und Proteinfamilien erhalten zusätzlich zu ihrem Namen eine fortlaufende Ziffern- oder Buchstabenfolge (z. B. Papillomavirus-Gene *E1A*, *E1B*, *E2*, *L1*, *L2* oder Histonproteine H1, H2A, H2B, H3, H4). In einigen Fällen werden Proteine auch dann durchnummeriert, wenn sie zwar nicht verwandt sind, aber einen eindeutigen Funktionszusammenhang aufweisen (z. B. die an der Replikation beteiligten Proteine von *E. coli*: DnaA [Initiatorprotein], DnaB [Helicase], DnaC [Helicaseladeprotein], DnaG [Primase]).

Für die Namensfindung von Genen und Proteinen gibt es keine verbindliche Regelung. Proteine können z. B. nach ihrem Funktionszusammenhang (wie Rec A = Recombination A), nach ihrem Verhalten (wie TBP = TATA-Box Binding Protein), nach ihren Leistungen (wie RNA-Polymerasen) oder nach dem Kontext ihrer Erstentdeckung (wie MPF, der bei der Erforschung der Oocyten-Reifung entdeckt wurde) benannt werden. Enzyme erhalten die Endsilbe „-ase“. Nucleotidsequenzen können nach der tatsächlichen Sequenz benannt werden (z. B. TATA-Box, GC-Box), nach ihrer Funktion (wie Origin und Promotor) oder nach ihrem codierten Inhalt. Proteingene werden häufig nach den Symptomen benannt, die durch ihren Ausfall hervorgerufen werden (z. B. *eyeless* bei *Drosophila*). Komplexe aus mehreren Proteinen erhalten eigene Bezeichnungen (z. B. besteht MPF aus Cyclin B und Cdc2).

4. Mutationen

Die folgende Regelung betrifft besonders die Bakteriengenetik: Mutationen eines Gens tragen hinter dem kursiv geschriebenen Namen des normalen Gens eine Zahl (z. B. *hisC1*, *hisC2*). Die Bezeichnung eines Phänotyps wird normal geschrieben, beginnt mit einem Großbuchstaben und trägt, je nach Leistungsvermögen des genannten Gens, ein hochgestelltes - oder + (z. B. His⁺). Führt z. B. eine *hisC1* genannte Mutation des *hisC*-Gens bei *E. coli* zum Ausfall der Histidinsynthese, so wird der Phänotyp des mutierten *E.-coli*-Stamms mit His⁻ bezeichnet.

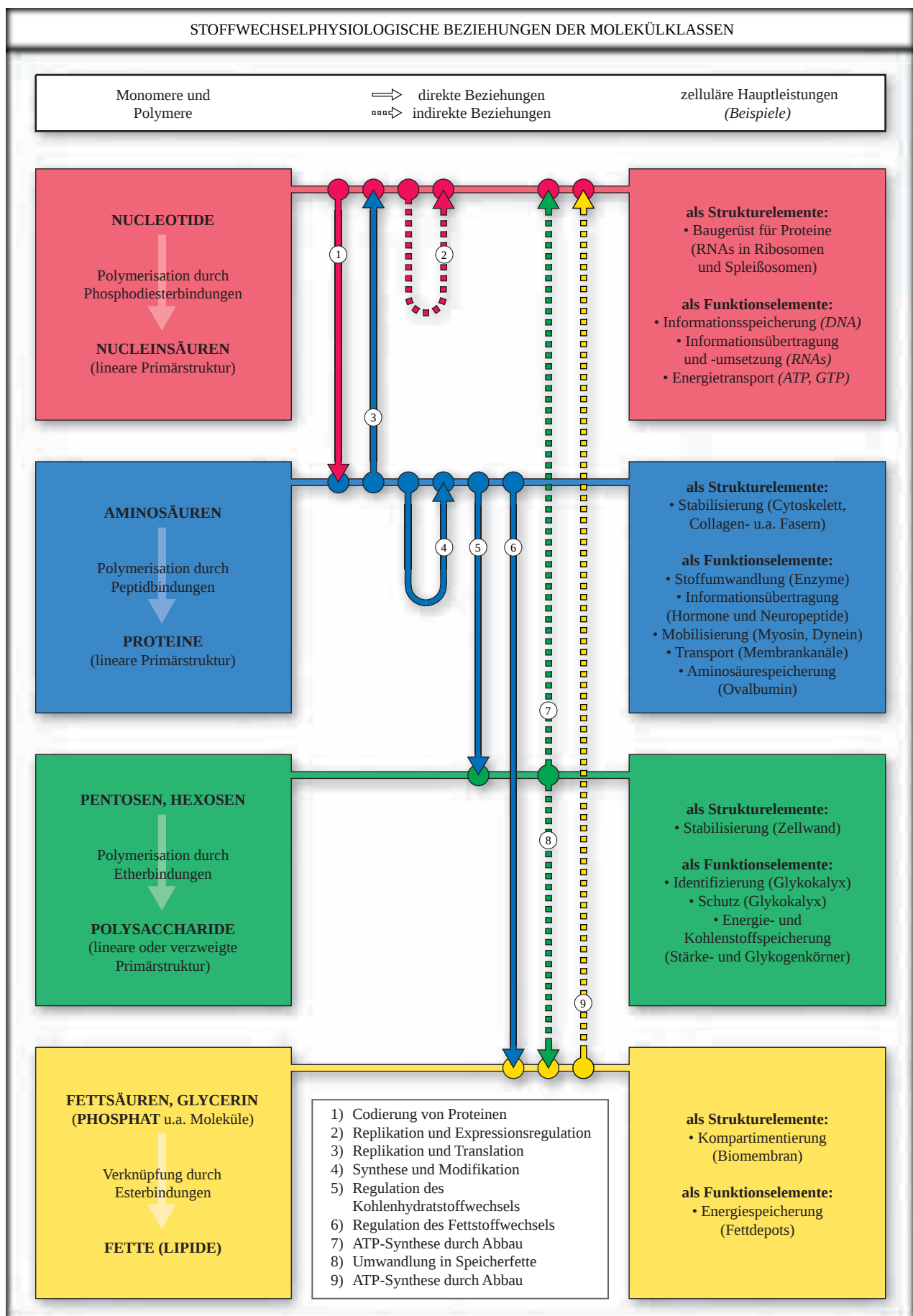
SYMBOLS UND ALLGEMEINE ABKÜRZUNGEN



∅	Durchmesser
<	kleiner als
>	größer als
↪	Querverweis auf zusätzliche Informationen im Kapitel „Ergänzungen“

Abb.	Abbildung
bes.	besonders
ca.	circa
d. h.	das heißt
euk.	eukaryotisch
evtl.	eventuell
Fam.	Familie

hpts.	hauptsächlich
i. d. R.	in der Regel
li	links
Mio	Millionen
Nt	Nucleotid (-e)
NTP	Nucleosidtriphosphat (-e)
o.	oder
prok.	prokaryotisch
re	rechts
S.	Seite
s. a.	siehe auch
sog.	sogenannte (-r, -s)
u. a.	unter anderem, und andere
u. Ä.	und Ähnliche (-s)
UE	Untereinheit (-en)
versch.	verschiedene (-r, -s)
z. B.	zum Beispiel
z. T.	zum Teil



ALLGEMEINES

Die vier wichtigsten Molekülgruppen wirken in z. T. sehr unterschiedlichen physiologischen Bereichen, da sie sich in ihrem chemischen und physikalischen Verhalten deutlich voneinander unterscheiden. Kennzeichen physiologischer Prozesse sind gerichtete Veränderungen von Molekülen (und Atomen) bezüglich des Ortes, der räumlichen Konfiguration und der chemischen Zusammensetzung. Die Grundlage für solche Veränderungen ist die Fähigkeit der Atome, untereinander verschiedene Bindungen eingehen zu können. Innerhalb und zwischen Biomolekülen sind hauptsächlich vier Bindungstypen vertreten: die kovalente Bindung (zwei Atome „teilen“ sich ein oder mehrere Elektronenpaare), die Ionenbindung (zwei entgegengesetzt geladene Ionen ziehen sich an), die Wasserstoffbrückenbindung (zwei Atome „teilen“ sich ein Wasserstoff-Atom) und die Van-der-Waals-Bindung (zwei entgegengesetzt ausgerichtete Dipole ziehen sich an). Die hydrophobe Wechselwirkung beruht nicht auf der gegenseitigen Anziehung hydrophober Moleküle, sondern auf der starken Bindung der Wasser-Moleküle. Im Folgenden werden die allgemeinen Charakteristika der einzelnen Molekülgruppen beschrieben.

NUCLEINSÄUREN

MOLEKÜLE: Es gibt vier verschiedene Desoxyribonucleotide und vier verschiedene Ribonucleotide, die zu linearen, unverzweigten Desoxyribonucleinsäuren (DNAs) bzw. Ribonucleinsäuren (RNAs) polymerisieren. Die Reihenfolge der Nucleotide wird durch enzymregulierte Kopiermechanismen von einem auf das andere Molekül übertragen. DNA-Moleküle liegen als lange, flexible Doppelstränge vor. RNA-Moleküle sind einzelsträngig und können sich durch Basenpaarungen zu komplexen, stabilen Strukturen falten.

FUNKTIONEN: DNA-Moleküle fungieren als Informationsspeicher und codieren die Primärsequenzen aller zellulären Proteine; RNA-Moleküle sind in verschiedenen Bereichen an der Umsetzung und Regulation dieser Informationen beteiligt.

PROTEINE

MOLEKÜLE: Es gibt 20 genetisch codierte Aminosäuren, die zu linearen Polypeptiden polymerisieren. Die Reihenfolge der Aminosäuren eines jeden Proteins wird durch eine entsprechende Nucleotidsequenz der DNA festgelegt. Polypeptide falten sich zu dreidimensionalen Strukturen, durch die letztlich die Funktion des jeweiligen Proteins festgelegt wird.

FUNKTIONEN: Von allen Molekülklassen weisen Proteine das größte Leistungsspektrum auf. Eine der wich-

tigsten Funktionen ist die Katalyse: Als Enzyme sind Proteine an der Umwandlung aller Stoffklassen, auch der eigenen, beteiligt. Da die Replikation der DNA ebenfalls durch Proteine durchgeführt wird und diese wiederum durch die Nucleotidsequenz der DNA codiert werden, sind die Nucleinsäuren indirekt an ihrer eigenen Vervielfältigung beteiligt.

POLYSACCHARIDE (= KOHLENHYDRATE)

MOLEKÜLE: Die Anzahl der in der Natur vorkommenden Monosaccharide ist auf Grund vielfältiger Derivatbildungen nur schwer abzuschätzen. Die Polymerisation der Monomere und somit auch die Gestalt der Polymere wird durch verschiedene Enzyme reguliert. Im Gegensatz zu Nucleotiden und Aminosäuren können Monosaccharide auf unterschiedliche Weisen miteinander verknüpft werden, so dass sowohl lineare als auch stark verzweigte Polymere gebildet werden können.

FUNKTIONEN: Die drei Hauptfunktionen von Polysacchariden sind die Zellstabilisation, die Energiespeicherung und die Zellidentifizierung. Die der Stabilisierung und Speicherung dienenden Polymere bestehen aus nur einem oder wenigen verschiedenen Monomeren; die der Identifizierung dienenden Makromoleküle der Zelloberfläche (= Glykokalyx) sind stark verzweigt und enthalten viele verschiedene Monosaccharide.

Die drei genannten Molekülklassen bilden demnach eine Informationskette mit zunehmender Komplexität der Polymere: DNA → Enzyme → Glykokalyx.

FETTE und LIPIDE

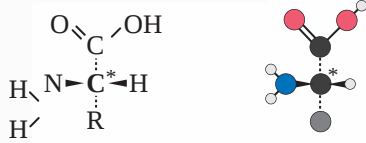
MOLEKÜLE: Fette bestehen aus Glycerin und drei Fettsäuren, Lipide enthalten statt einer Fettsäure eine Phosphorsäure, an die weitere Komponenten gebunden sein können. Fette und Lipide sind daher keine Polymere. Wie Proteine und Polysaccharide werden sie durch Enzyme synthetisiert (und abgebaut).

FUNKTIONEN: Die einfacher gebauten Fette fungieren in der Zelle als Langzeit-Energiespeicher, die komplexeren Lipide sind Bestandteil der Biomembranen und kompartimentieren als solche den Zellinhalt. Ein Kohlenhydratüberschuss kann in Speicherfette umgewandelt werden. Ebenso wie bei den Kohlenhydraten wird auch bei Fetten durch den Abbau der Makromoleküle ATP gebildet, ein energiereiches Molekül, das außerdem Bestandteil der Nucleinsäuren ist.

Wie für alle biologischen Strukturen gilt auch für die am Stoffwechsel beteiligten Moleküle:
Die Struktur determiniert die Funktion!

DIE 22 PROTEINOGENEN AMINOSÄUREN

allgemeine Schreibform einer L-Aminosäure
C* = α -C-Atom, R = Rest (= Seitengruppe)



Aminosäuren unterscheiden sich durch die unten aufgeführten Seitengruppen.

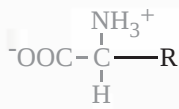
Die Seitengruppen entscheiden über das chemische Verhalten der Aminosäuren.

ungeladene, unpolare, hydrophobe Aminosäuren

ungeladene, polare, hydrophile Aminosäuren

geladene, saure, hydrophile Aminosäuren

geladene, basische, hydrophile Aminosäuren



Bis auf Prolin sind nur die Seitengruppen der jeweiligen Aminosäuren angegeben. Vor den ausgeschriebenen Namen stehen die entsprechenden Drei- und Einbuchstaben-Codes. Hell unterlegt sind die für Menschen essentiellen Aminosäuren. Die seltenen Aminosäuren Selenocystein und Pyrrolysin werden durch Stopp-Codons codiert.

Gly, G, **Glycin** —H

Ala, A, **Alanin** —CH₃

Val, V, **Valin** —CH(CH₃)₂

Leu, L, **Leucin** —CH₂-CH(CH₃)₂

Ile, I, **Isoleucin** —CH(CH₃)-CH₂-CH₃

Pyl, O, **Pyrrolysin** —CH₂-CH₂-CH₂-CH₂-NH-C(=O)-

Met, M, **Methionin** —CH₂-CH₂-S-CH₃

Pro, P, **Prolin** Wegen Ringschluss mit der Seitengruppe besitzt Prolin keine freie Aminogruppe.

Phe, F, **Phenylalanin** —CH₂-

Trp, W, **Tryptophan** —CH₂-

Ser, S, **Serin** —CH₂-OH

Thr, T, **Threonin** —CH(CH₃)-OH

Asn, N, **Asparagin** —CH₂-C(=O)-NH₂

Gln, Q, **Glutamin** —CH₂-CH₂-C(=O)-NH₂

Cys, C, **Cystein** —CH₂-SH

Sec, U, **Selenocystein** —CH₂-SeH

Tyr, Y, **Tyrosin** —CH₂-

Asp, D, **Asparaginsäure** —CH₂-COO⁻

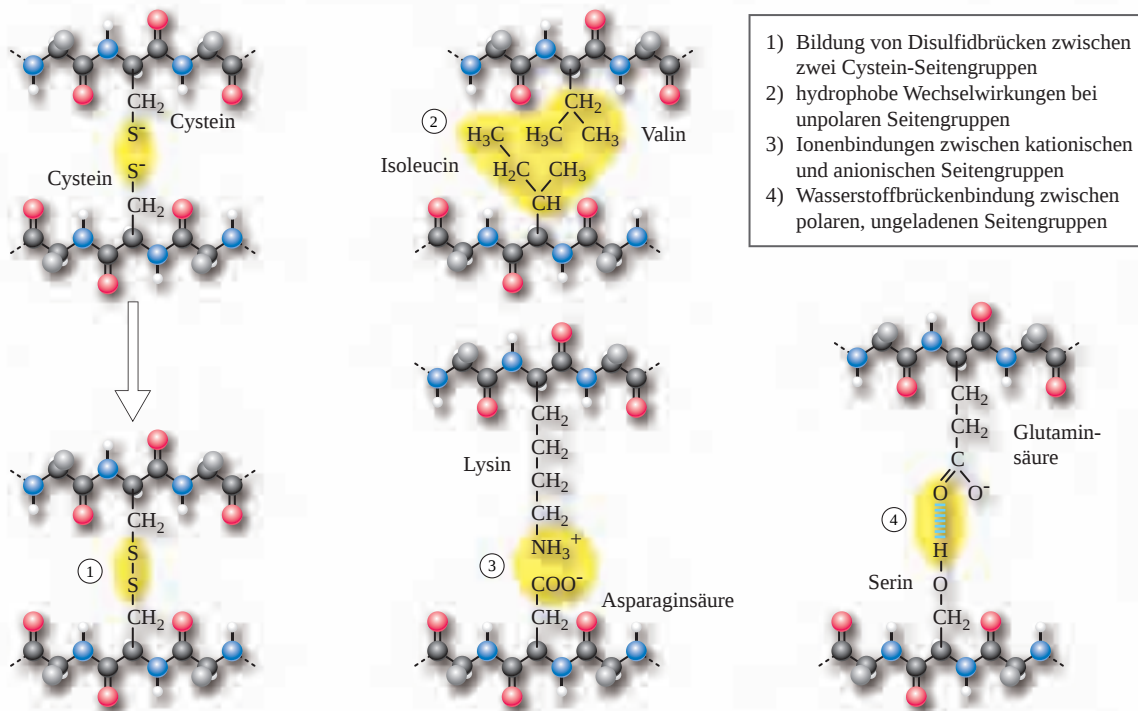
Glu, E, **Glutaminsäure** —CH₂-CH₂-COO⁻

Lys, K, **Lysin** —CH₂-CH₂-CH₂-CH₂-NH₃⁺

Arg, R, **Arginin** —CH₂-CH₂-CH₂-NH-C(=NH₂⁺)₂

His, H, **Histidin** —CH₂-

INTERAKTIONEN DER SEITENGRUPPEN



DEFINITION

Proteine sind unterschiedlich lange, unverzweigte Polymere der L-Aminosäuren. Die funktionsfähige Konformation wird erreicht, indem die Polypeptidketten – in Abhängigkeit ihrer Aminosäuresequenz – gefaltet und geknäuel werden.

LEISTUNGEN

als Strukturelemente:

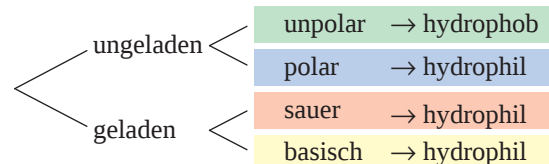
- Festigung und Stabilisierung von anderen Makromolekülen, von Organellen, Zellen, Geweben und Organen (z. B. Keratine, Collagen, Elastin, Histone und Cytoskelettelemente)

als Funktionselemente:

- Katalyse (Enzyme, z. B. DNA-Polymerasen)
- Informationsübertragung (Hormone, z. B. Insulin; Neuropeptide, z. B. Glucagon; Regulatorproteine, z. B. Transkriptionsfaktoren; Rezeptoren der Zellmembran; Proteinkinasen)
- Mobilisierung (z. B. Myosin und Dynein)
- Transport (z. B. Hämoglobin, Fibrinogen, Membrankanäle und -pumpen)
- Schutz (z. B. Antikörper, Toxine, Antibiotika)
- Aminosäurespeicher (z. B. Ovalbumin, Ferritin, Ricin).

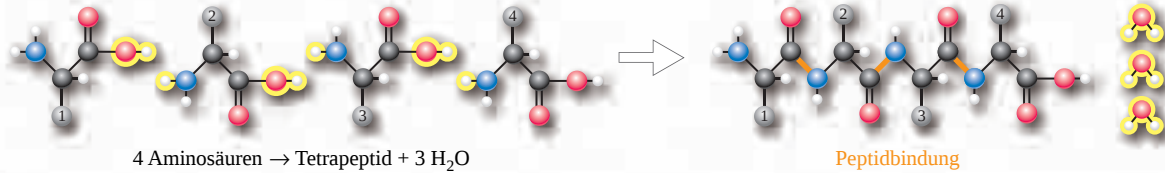
AMINOSÄUREN

Aminosäuren (= Aminocarbonsäuren) besitzen ein asymmetrisches, sog. α -C-Atom, an das ein Wasserstoff-Atom, eine Amino-, eine Carboxyl- und eine für jede Aminosäure charakteristische Seitengruppe gebunden sind. Auf Grund des asymmetrischen C-Atoms existieren stereoisomere **D- und L-Konfigurationen**. Alle in Proteinen vorkommenden (= proteinogenen) Aminosäuren gehören dem L-Typ an. Insgesamt konnten bisher über 260 verschiedene Aminosäuren identifiziert werden, doch kommen in Proteinen i. d. R. nur 22 verschiedene Aminosäuren vor, von denen zwei, Selenocystein und Pyrrolysin, nicht exakt codiert sind. Für den Menschen sind neun Aminosäuren (zusätzlich noch Arginin für Kinder im Wachstumsalter) **essentiell**, d. h. sie können vom Organismus nicht selbst synthetisiert werden. Die einzelnen Aminosäuren unterscheiden sich in Größe, Form und Ladung der Seitengruppen, die maßgeblich die Form und Leistung des Proteins beeinflussen. Anhand dieser Seitengruppen werden die Aminosäuren nach folgendem System unterteilt:



PRIMÄRSTRUKTUR

Die Primärstruktur entspricht der linearen Aminosäuresequenz. Gemäß der Konvention wird die Sequenz entlang der Amino-Carboxyl-Richtung wiedergegeben.

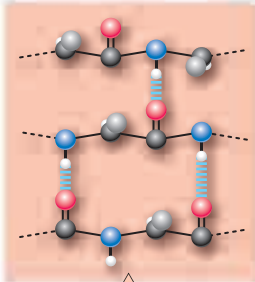
 α -Helix

SEKUNDÄRSTRUKTUR

 β -Faltblatt

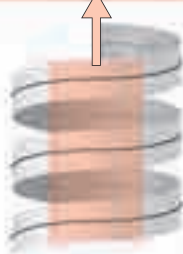
Die Sekundärstruktur beschreibt die durch Wasserstoffbrückenbindungen zwischen CO- und NH-Gruppen hervorgerufene Konfiguration, die als Helix oder Faltblattstruktur dargestellt wird.

Ausschnitt Seitenansicht

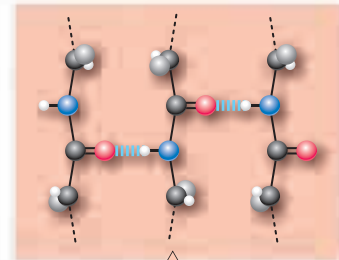


Bei der **α -Helix** ist die Peptidkette um einen imaginären Zylinder gewunden, so dass CO- und NH-Gruppen übereinander liegen und die Seitengruppen aus der Achse nach außen ragen. Brückenbindungen entstehen zwischen der CO-Gruppe einer Aminosäure und der NH-Gruppe der viertnächsten Aminosäure.

Helix-Modell



Ausschnitt Aufsicht



Die **β -Faltblattstruktur** entsteht, indem die Seitengruppen aufeinanderfolgender α -C-Atome alternierend nach oben und unten aus der Ebene ragen. Liegen zwei oder mehrere solcher β -Stränge parallel oder antiparallel nebeneinander, werden sie durch Wasserstoffbrücken zu einem stabilen Faltblatt verbunden.

Faltblatt-Modell



TERTIÄRSTRUKTUR

Die räumliche Konfiguration des Gesamtmoleküls wird als Tertiärstruktur bezeichnet. Diese bestimmt das Leistungsspektrum des Moleküls.



Bei der dreidimensionalen Darstellung von Proteinen werden Helixbereiche als Schrauben oder Zylinder dargestellt und β -Stränge als Pfeile, wobei die Pfeilspitze immer zum Carboxylterminus der Peptidkette weist. Die dazwischenliegenden Abschnitte werden durch dünne Kabel symbolisiert.

links und Mitte: Lysozym
rechts: ein ribosomales Protein
N: Aminoterminus
C: Carboxylterminus

AUFBAU DER PROTEINE

1. PRIMÄRSTRUKTUR

Die Carboxylgruppe einer Aminosäure wird mit der Aminogruppe der benachbarten Aminosäure zu einer **Peptidbindung** verknüpft. Die Länge der Polypeptide ist sehr unterschiedlich und kann mehrere tausend Aminosäuren betragen. Die Polymerisation findet in den Ribosomen statt und verläuft vom Amino-(N-) terminus zum Carboxyl-(C-)terminus, d. h. die erste Aminosäure trägt eine freie Aminogruppe, die letzte eine freie Carboxylgruppe. Die konventionelle Schreibweise einer Aminosäuresequenz erfolgt ebenfalls vom N- zum C-Terminus.

2. SEKUNDÄRSTRUKTUR

Zwischen den Carbonyl- (CO-) und den Amid- (NH-) gruppen verschiedener Aminosäuren können Wasserstoffbrücken gebildet werden, wodurch Bereiche mit geometrischen Raumstrukturen entstehen. Diese Abschnitte werden als α -Helix bzw. β -Strang bezeichnet. Eine **α -Helix** ist eine meist rechtsläufig gewundene Schraube, die i. d. R. 3,6 Aminosäuren pro Windung enthält. Wasserstoffbrücken entstehen zwischen einer CO-Gruppe mit einer NH-Gruppe der viertnächsten Peptidbindung. Die Seitengruppen der an der Helix beteiligten Aminosäuren weisen nach außen. Starke Helixbildner sind die Aminosäuren Ala, Glu, Leu und Met. Helices sind weit verbreitet unter Membranproteinen und Transkriptionsfaktoren. Als **β -Strang** wird ein Sequenzabschnitt bezeichnet, dessen hintereinanderliegende α -C-Atome mitsamt den angebundenen Seitengruppen alternierend nach oben und unten aus der Ebene ragen. Eine Peptidkette faltet sich meist so, dass zwei oder mehrere β -Stränge parallel oder antiparallel nebeneinander liegen und ein stabiles **β -Faltblatt** bilden. Die Wasserstoffbrücken werden zwischen den CO-Gruppen des einen Strangs und den NH-Gruppen des Nachbarstrangs gebildet. Starke Faltblattbildner sind die Aminosäuren Tyr, Val und Ile. Faltblätter sind stabiler als Helices und man findet sie häufig im Innern von Proteinen. Die Seitengruppen der Aminosäuren sind weder an der Ausbildung der Helices noch der Faltblätter beteiligt.

Orientierungsmöglichkeiten der β -Stränge:
parallel (links) und
antiparallel (rechts)



Durchschnittlich 60 % einer Aminosäurekette ordnen sich zu α -Helices und β -Strängen. Diese Bereiche sind durch Haarnadelschleifen oder größere, unregelmäßig geformte Schlaufen miteinander verbunden. Häufig bilden Kombinationen aus Helices und β -Strängen in sich stabile Komplexe, die sich deutlich von der übrigen Proteinstruktur abgrenzen. Diese **Domänen** stellen die Grundeinheiten der nächsten Organisationsebene dar.

3. TERTIÄRSTRUKTUR

Die endgültige dreidimensionale Konformation eines Polypeptids wird durch Einbeziehung der Seitengruppen erreicht. Die Tertiärstruktur, die zugleich die thermodynamisch stabilste Form des Moleküls darstellt, bestimmt letztlich das Leistungsspektrum des Proteins. Durch Biegung und Faltung der Peptidkette können Seitengruppen, die in der Primärsequenz nicht benachbart sind, miteinander interagieren und folgende Wechselwirkungen eingehen:

- Wasserstoffbrücken zwischen polaren Resten
- Ionenbindungen zwischen kationischen und anionischen Resten
- hydrophobe Wechselwirkungen zwischen unpolaren Resten
- Van-der-Waals-Anziehungen
- kovalente Bindung zwischen Cystein-Resten.

Die Hauptkraft der Faltung wird durch die Zusammenlagerung der hydrophoben Seitengruppen ausgeübt, die sich zur Proteinmitte drehen. Die hydrophilen Seitengruppen drehen sich dagegen zur Außenseite des Proteins. SH-Gruppen gegenüberliegender Cystein-Reste können durch das Enzym Proteindisulfidisomerase zu einer Disulfidbrücke verknüpft werden. (Bei Eukaryoten findet dieser Prozess im Endoplasmatischen Retikulum statt.) Da Disulfidbrücken nur in einer oxidativen Umgebung stabil sind, findet man sie nur in extrazellulären Oberflächen- und in Exportproteinen (z. B. bei Immunglobulinen, Fibronectinen und Insulin). In lebenden Zellen wird der Faltungsprozess von Polypeptiden noch während ihrer Synthese von Hilfsproteinen, den sog. Chaperonen unterstützt.

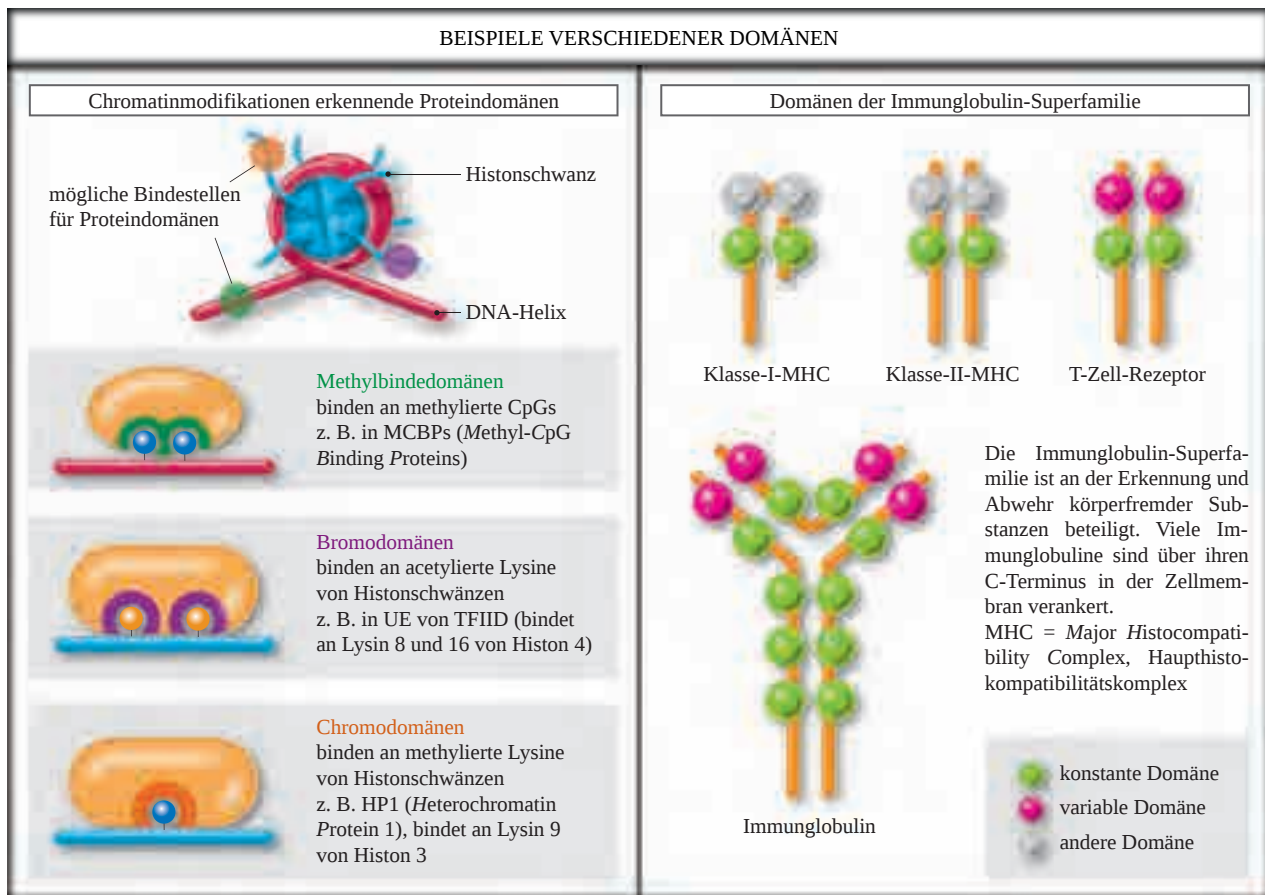
Der Form nach unterscheidet man leicht diffundierbare, wasserlösliche **globuläre Proteine** (mit meist dynamischer Funktion, z. B. Enzyme) und physikalisch widerstandsfähigere, wasserunlösliche **faserförmige Proteine** (mit meist stabilisierender Funktion, z. B. Collagen, Keratin und Intermediärfilamente).

4. QUARTÄRSTRUKTUR

Zelluläre Prozesse und Funktionen werden i. d. R. nicht durch einzelne Proteine, sondern durch Proteinkomplexe ausgeführt. Diese entstehen, indem Proteine über definierte Oberflächenareale nicht-kovalente Bindungen mit anderen Molekülen eingehen. Proteinkomplexe werden auch **oligomere Proteine** genannt und bestehen aus gleichen oder verschiedenen Untereinheiten (UE):

- Hämoglobin 4 UE (2 verschiedene)
- Lac-Repressor 4 UE (identische)
- RNA-Polymerase (prok.) 5 UE (4 verschiedene)
- Histonoctamer 8 UE (4 verschiedene)
- Actinfilament > 1.000 identische UE.

Einige Proteine enthalten nicht-proteinogene Komponenten wie Zucker (z. B. in Glykoproteinen), Fettsäuren (z. B. in Lipoproteinen), Metallionen (z. B. Fe^{2+} in Hämoglobinen) oder andere organische Moleküle (z. B. Porphyrinring in Cytochromen).



DOMÄNEN

Proteindomänen sind strukturell abgrenzbare Einheiten, die von bestimmten Teilen einer Peptidkette gebildet werden; diese 50-350 Aminosäuren langen Einheiten können sich unabhängig von der übrigen Peptidkette zu kompakten, stabilen Strukturen falten. Verschiedene Domänen eines Proteins haben häufig auch verschiedene Funktionen, deren Zusammenwirken die Gesamtleistung des Proteins ausmachen. Die Verbindungsschlaufen zwischen den Domänen fungieren häufig als Bindungsstellen für andere Moleküle.

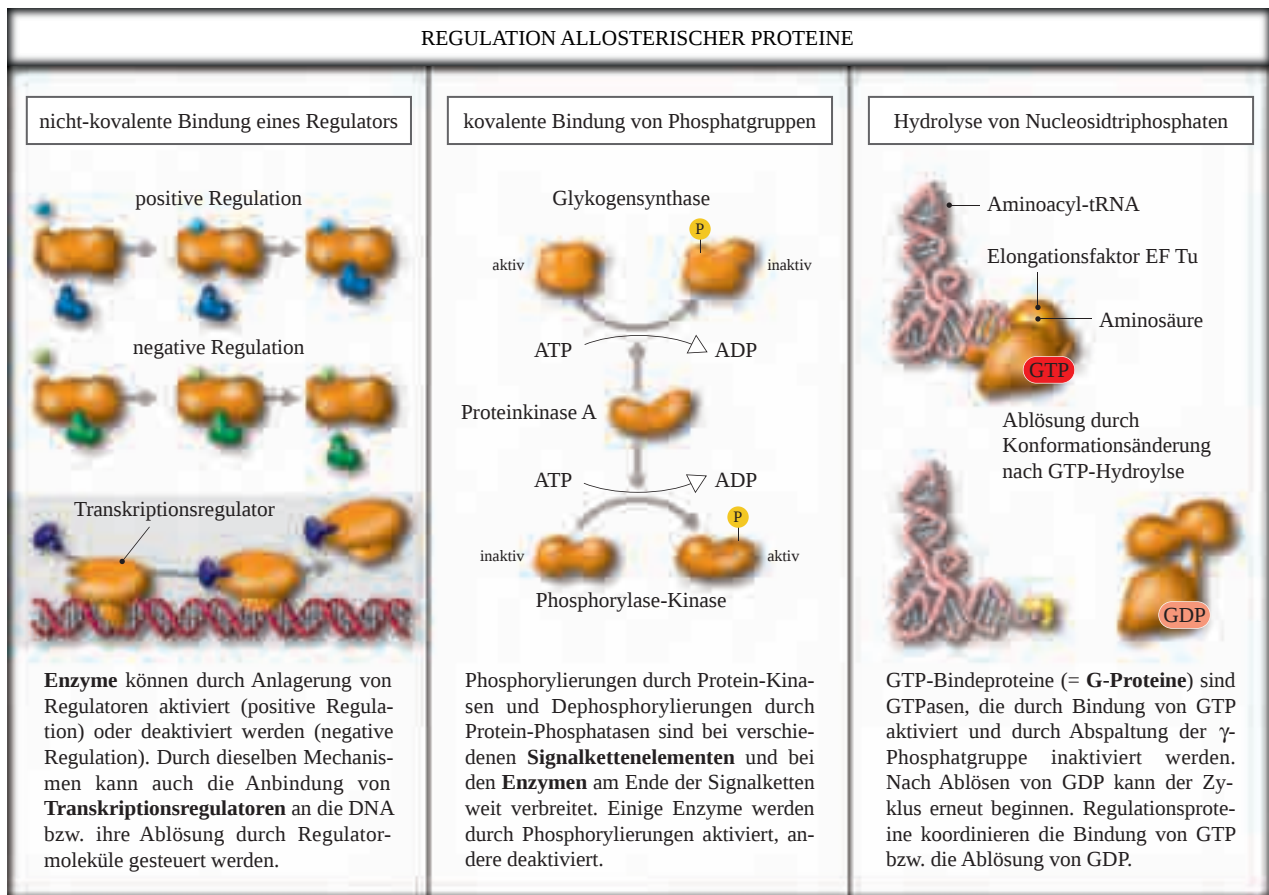
Domänen stellen ein gutes Beispiel für das in der Natur weit verbreitete Baukastenprinzip dar: Aus demselben Domänensatz entstehen durch **Domänenmischung** verschiedene Proteine mit verschiedenen Leistungen. Einige Domänen codierende DNA-Sequenzen sind im Laufe der Evolution innerhalb eines oder zwischen verschiedenen Proteingenen gewandert oder haben sich tandemartig vervielfacht. Viele Proteine entstanden durch Neukombination bereits bestehender Domänen. Beispiele für verschiedene Domänen:

- Zahlreiche regulatorische Proteine besitzen Methylbindedomänen, Bromo- oder Chromodomänen, durch die sie spezifische epigenetische Marker wie Methylierungen und Acetylierungen erkennen.
- Alle Helicasen enthalten eine sog. RecA-ähnliche Domäne.

- Alle Mitglieder der Immunglobulin-Superfamilie besitzen mindestens ein Paar der sog. konstanten Domänen.
 - Die sog. DEAD/H-Box bezeichnet die Abfolge der Aminosäuren Asp-Glu-Ala-Asp/His und kommt bei Proteinen vor, die sehr stabile dsRNA-Doppelbindungen aufbrechen können.
 - dsRBDs (double stranded RNA Binding Domains) kommen u.a. in Proteinen vor, die an der Prozessierung verschiedener RNAs beteiligt sind.
 - Zinkfingerdomänen sind Struktureinheiten, die in vielen Regulatorproteinen zu finden sind.
- ➔ s. a. S. 186, Regulatorproteine.

FUNKTIONSWEISE

Die chemischen Eigenschaften von Proteinen hängen von den physikalischen Wechselwirkungen mit anderen Molekülen ab. Die Art der Wechselwirkung hängt wiederum von den Seitengruppen der Proteinoberfläche ab, denn über diese werden verschiedene nicht-kovalente Bindungen (Wasserstoffbrücken, Ionenbindungen, hydrophobe Wechselwirkungen und Van-der-Waals-Bindungen) mit den **Liganden** aufgebaut. Für eine effektive Wechselwirkung müssen mehrere Bindungen gleichzeitig gebildet werden und dies ist nur möglich, wenn die jeweiligen Oberflächenprofile der Kontakt-



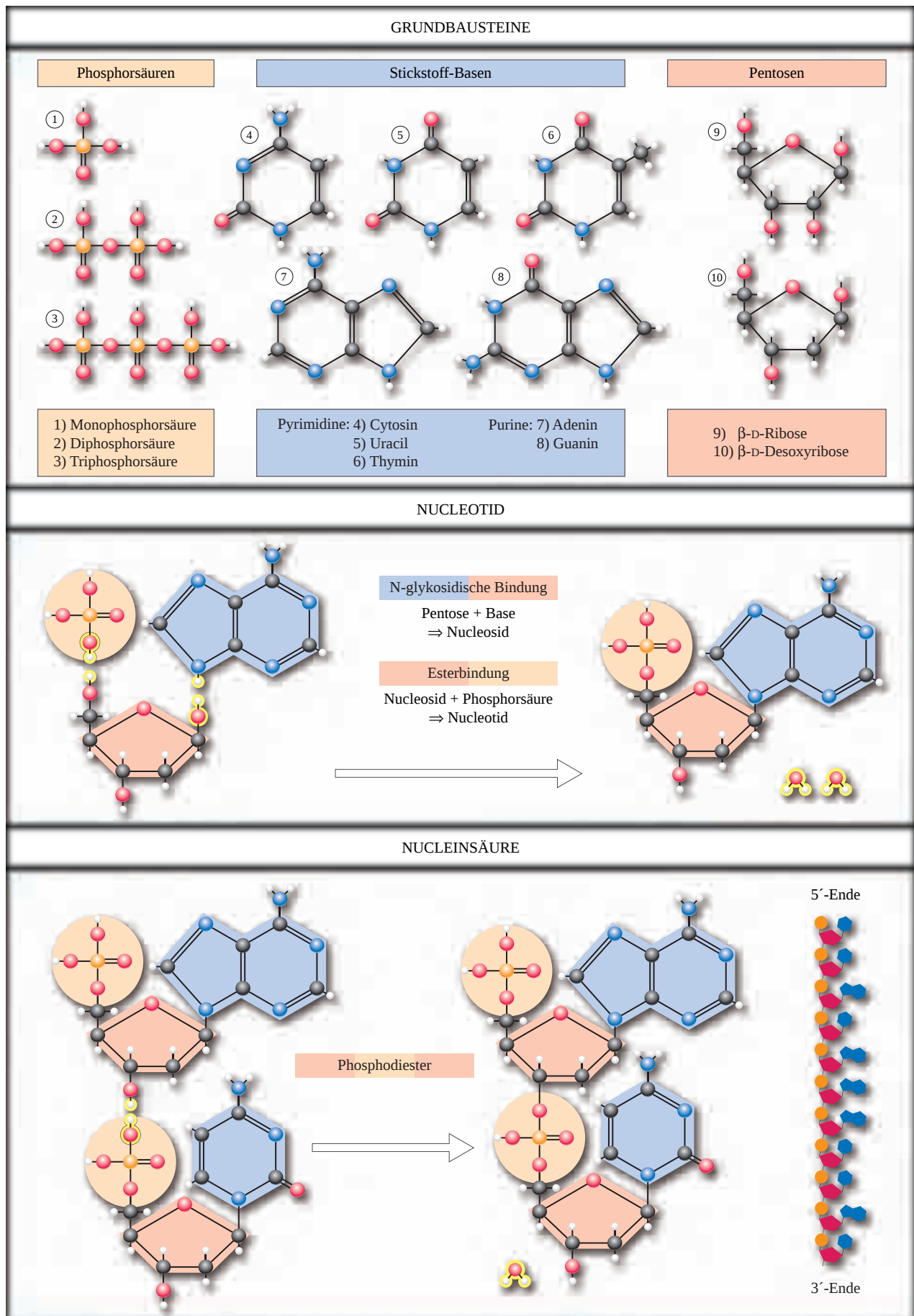
zonen genau zusammenpassen. Auf Grund dieser physikalischen Voraussetzungen haben Proteine die Fähigkeit, selektiv und mit hoher Affinität andere Moleküle zu binden. Die meisten Proteine besitzen mehrere Bindungsstellen für unterschiedliche Liganden. Ein Protein, das seinen Liganden chemisch verändert, wird **Enzym** genannt; der veränderbare Ligand wird als **Substrat** dieses Enzyms bezeichnet. Ein Substratmolekül muss eine Reihe von Zwischenzuständen mit veränderter Geometrie und Elektronenverteilung durchlaufen, ehe es ein stabiles Reaktionsprodukt bildet. Die Energie, die benötigt wird, um den instabilsten Übergangszustand zu erreichen, heißt Aktivierungsenergie; sie ist der Hauptfaktor, der die Reaktionsgeschwindigkeit limitiert. Dadurch, dass Enzyme selektiv die Übergangszustände binden und stabilisieren, beschleunigen sie die Reaktion um den Faktor $10^9 - 10^{23}$. Die Enzyme selbst werden durch die Reaktion nicht verändert, sie fungieren also als Katalysatoren. Eingebundene **Coenzyme** wie FAD und Häm verleihen Enzymen zusätzliche Leistungen. Ein Enzymmolekül kann durchschnittlich 1.000 Substratmoleküle/Sekunde umsetzen! Bei sehr schnell arbeitenden Enzymen ist der limitierende Faktor die Diffusionsgeschwindigkeit der Substrat- und Produktmoleküle. Daher werden komplexe Reaktionsketten in der Zelle nicht durch einzelne Enzyme, sondern durch Enzymaggregate katalysiert, sog. **molekulare Maschinen**.

ALLOSTERISCHE PROTEINE

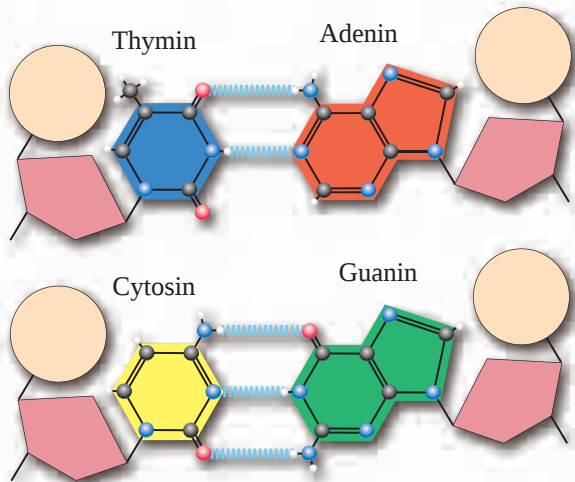
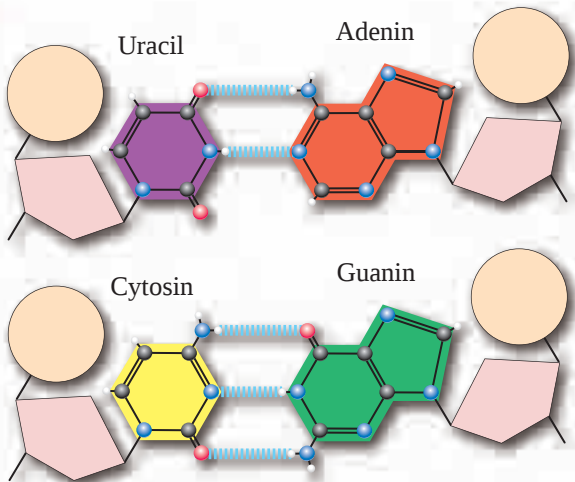
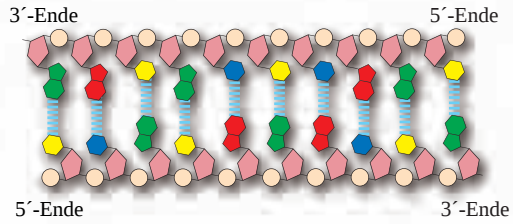
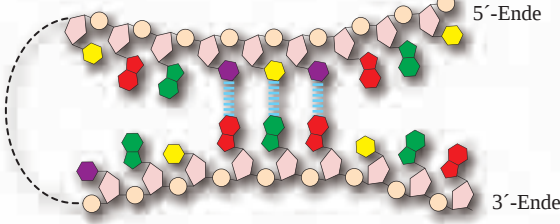
Proteine mit der Fähigkeit zur reversiblen Konformationsänderung werden **allosterische Proteine** genannt. Die Veränderung der Raumstruktur entsteht durch Um- und Neuknüpfung von schwachen Bindungen zwischen den Seitenketten innerhalb des Proteins. Durch die Verschiebung von der einen in die andere Konformation kann die Aktivität des Proteins verändert werden. Ausgelöst werden Konformationsänderungen meist durch einen der drei folgenden Mechanismen:

- nicht-kovalente Bindung eines Regulator-Moleküls (z. B. bei Transkriptionsregulatoren und Enzymen)
- kovalente Bindung einer Phosphatgruppe (z. B. bei Signalkettenelementen und Enzymen)
- Hydrolyse von ATP oder GTP (z. B. bei Motorproteinen, G-Proteinen und Signalkettenelementen)

Allosterische Proteine sind regulierbare Proteine. Sie haben mindestens zwei verschiedene Bindungsstellen: eine für das Kontrollelement und eine für das Substrat. Durch die Bindung des Kontrollelements entsteht eine lokale Konformationsänderung, die sich innerhalb des Proteins fortpflanzt und zu einer Veränderung anderer Bindestellen oder zu einer Bewegung des gesamten Moleküls führt. Bei Proteinkomplexen kann eine koordinierte Kettenreaktion ausgelöst werden. Zu den regulierbaren Proteinen zählen u. a. Enzyme, Rezeptor-, Struktur-, Motor- und Regulatorproteine.



KLASSIFIKATION DER NUCLEOSIDE UND NUCLEOTID-MONOPHOSPHATE			
Basen	Nucleoside	Desoxyribonucleotide (→ DNA)	Ribonucleotide (→ RNA)
<i>Purine</i>			
Adenin (A) 	Adenosin	2'-Desoxyadenosin-5-Monophosphat (dAMP)	Adenosin-5-Monophosphat (AMP)
Guanin (G) 	Guanosin	2'-Desoxyguanosin-5-Monophosphat (dGMP)	Guanosin-5-Monophosphat (GMP)
<i>Pyrimidine</i>			
Thymin (T) 	Thyminidin	2'-Desoxythyminidin-5-Monophosphat (dTMP)	
Uracil (U) 	Uridin		Uridin-5-Monophosphat (UMP)
Cytosin (C) 	Cytidin	2'-Desoxycytidin-5-Monophosphat (dCMP)	Cytidin-5-Monophosphat (CMP)

DNA	BASENPAARUNGEN	RNA
 Pentose: Desoxyribose Basen: Adenin, Guanin; Thymin, Cytosin Paare: Adenin-Thymin (zwei Wasserstoffbrücken) Guanin-Cytosin (drei Wasserstoffbrücken)	 Pentose: Ribose Basen: Adenin, Guanin; Uracil, Cytosin Paare: Adenin-Uracil (zwei Wasserstoffbrücken) Guanin-Cytosin (drei Wasserstoffbrücken)	
 Zwei komplementäre DNA-Einzelstränge bilden einen stabilen Doppelstrang, der zu einer Helix gewunden ist. Eine Trennung der Einzelstränge erfolgt nur kurzfristig und i. d. R. im Rahmen einer Replikation oder Transkription.	 Einzelstränge verbinden sich i. d. R. nicht zu Doppelsträngen; lokale Basenpaarungen innerhalb eines RNA-Moleküls kommen häufig vor (z. B. rRNAs, tRNAs). Komplementäre DNA- und RNA-Stränge können labile Hybrid-Doppelstränge bilden.	

DNA	PRIMÄRSTRUKTUR	RNA
Die Primärstruktur entspricht der linearen Nucleotidsequenz. Gemäß der Konvention wird die Sequenz durch die Anfangsbuchstaben der Nucleotid-Basen entlang der 5'→3'-Richtung wiedergegeben.		
5'-Ende 3'-Ende C T A C G A A T C G T	5'-Ende 3'-Ende U C G U G A G U C A C	
DNA	SEKUNDÄRSTRUKTUR	RNA
Die Sekundärstruktur von Nucleinsäuren beschreibt die durch Basenpaarungen hervorgerufene Konfiguration, die als zweidimensionale Strukturen dargestellt werden.		
5'-Ende 3'-Ende Kreuzstruktur bei Palindromen	Haarnadelschleife Stiel Ausbuchtung Pseudoknoten Drei-Stamm-Verbindung	
Bei DNA-Molekülen bilden i. d. R. zwei gegenläufige, komplementäre Einzelstränge einen linearen Doppelstrang. Sequenzbereiche, die, um 180° gedreht, mit sich selbst übereinstimmen (= Palindrome) können sich vorübergehend zu einem Kreuz umfalten, wenn die Helix unter starker Drehspannung steht.	RNA-Moleküle liegen i. d. R. als Einzelstränge vor, die auf Grund inter- und intramolekularer Basenpaarungen unterschiedliche Formen bilden können. Charakteristische Faltmuster sind Stiel, Haarnadelschleife, Ausbuchtung, Pseudoknoten und Stammverbindungen. Gefaltete RNAs haben meist eine katalytische oder strukturelle Funktion.	
DNA	TERTIÄRSTRUKTUR	RNA
Die dreidimensionale Konformation der Moleküle wird als Tertiärstruktur bezeichnet. Sie wird durch die Sekundärstruktur bestimmt und determiniert das Leistungsspektrum der Moleküle.		
2nm große Furche kleine Furche vollständige Windung 3,4 nm mit ca. zehn Basenpaaren	charakteristische L-Form einer tRNA	RNA-Moleküle falten sich zu unregelmäßigen Formen auf. Neben der durch Basenpaarungen entstandenen Molekülform sind besonders die Position und die Sequenz der ungepaarten Basen verantwortlich für die Interaktionen mit anderen Molekülen und somit für das große Leistungsspektrum der verschiedenen RNA-Klassen.

DEFINITION

Nucleinsäuren sind hochpolymere, unverzweigte Moleküle aus Nucleotiden. Man unterscheidet DNA (Deoxyribonucleic Acid) und RNA (Ribonucleic Acid). DNA liegt meist als Doppelstrang, RNA als Einzelstrang vor.

LEISTUNGEN

als Strukturelemente:

- Baugerüst für Proteine (z. B. RNAs in Ribosomen und Spleißosomen)

als Funktionselemente:

- „Energietransport“ (Nucleotide wie ATP und GTP)
- Informationsspeicherung (DNA, seltener RNA)
- Informationstransport (z. B. mRNAs)
- Informationsumsetzung in Aminosäuresequenzen (z. B. tRNAs, rRNAs und gRNAs)
- Regulation der Informationsumsetzung (z. B. miRNAs, siRNAs)
- Katalyse (z. B. Ribozyme und Ribonuclease P)

NUCLEOTIDE

Nucleinsäuren sind Polymere aus Nucleotiden. Diese bestehen aus drei Komponenten: einer N-haltigen **Base**, einer **Pentose** und einem **Phosphatrest**.

Bei den Basen handelt es sich um die größeren Purine Adenin und Guanin und die kleineren Pyrimidine Cytosin, Uracil und Thymin. Die Base Uracil kommt nur in RNAs vor, Thymin nur in DNAs.

Die Pentosen sind β -D-Desoxyribose oder β -D-Ribose. Ist die Pentose eine Desoxyribose, heißt das Nucleotid Desoxyribonucleotid und das entsprechende Polymer Desoxyribonucleinsäure (oder DNA). Ribosen bilden Ribonucleotide und Ribonucleinsäuren (oder RNAs). Die Base und die Pentose werden über die OH-Gruppe des C₁-Atoms der Pentose N-glykosidisch zu einem Nucleosid verknüpft.

Die OH-Gruppe des C₅-Atoms der Pentose wird mit einer Phosphorsäure zu einem Nucleotid verestert. Je nach Größe der Phosphorsäure entstehen Mono-, Di- oder Triphosphate.

AUFBAU DER NUCLEINSÄUREN

Polymere entstehen ausschließlich aus Ribonucleotiden ($\rightarrow$ RNA) oder ausschließlich aus Desoxyribonucleotiden ($\rightarrow$ DNA). Die Hauptunterschiede zwischen diesen beiden Polymerklassen liegen in kleinen unterschiedlichen Seitengruppen der Nucleotide:

Pentose: OH (RNA) bzw. H (DNA) am C₂-Atom

Base: H bei Uracil (RNA) bzw.
CH₃ bei Thymin (DNA)

1. PRIMÄRSTRUKTUR

Die OH-Gruppe des C₃-Atoms eines Nucleotids und eine OH-Gruppe des Phosphats eines anderen Nucleotids werden verestert. Auf diese Weise entsteht ein langes, unverzweigtes Polymer aus Nucleotiden, das durch das Zucker-Phosphodiesterband gehalten wird. Nach einer allgemeinen Übereinkunft wird die Reihenfolge der Nucleotide in 5' $\rightarrow$ 3'-Richtung aufgezählt, d. h. man beginnt mit dem Nucleotid, das einen freien Phosphatrest am C₅-Atom trägt und endet bei dem Nucleotid mit der freien OH-Gruppe am C₃-Atom.

Die Primärstruktur der Nucleinsäuren entspricht der linearen Abfolge der Nucleotide, die mit ihren Anfangsbuchstaben abgekürzt werden.

2. SEKUNDÄRSTRUKTUR

Die Sekundärstruktur entsteht durch **Basenpaarungen**, die durch **Wasserstoffbrückenbindungen** gebildet werden. Zwischen den Basen Adenin und Thymin sowie zwischen Adenin und Uracil entstehen jeweils zwei und zwischen Guanin und Cytosin drei Wasserstoffbrücken. Es paaren sich also immer eine große (Purin) und eine kleine Base (Pyrimidin).

Bei DNA-Molekülen lagern sich zwei gegenläufige, komplementäre Stränge zusammen und bilden die Form einer Strickleiter. In einem solchen Doppelmolekül ist das Basenverhältnis von Adenin zu Thymin und von Guanin zu Cytosin gleich 1. Das Zahlenverhältnis der Basenpaare A-T/G-C ist jedoch bei verschiedenen Doppelmolekülen unterschiedlich und für jede DNA charakteristisch.

RNA-Moleküle bilden häufig intramolekulare Basenpaarungen aus, wodurch charakteristische Formen entstehen können. Haarnadelschlaufen sind z. B. bei einigen mRNAs zu finden, wo sie als Elemente von Regulationsprozessen fungieren. Da RNAs i. d. R. einzelsträngig vorliegen, sind bei ihnen die A/U und die G/C-Basenverhältnisse nur in Ausnahmefällen gleich 1.

3. TERTIÄRSTRUKTUR

Die Basenpaarungen verursachen eine räumliche Aufaltung der Moleküle, in der sich DNA- und RNA-Moleküle sehr deutlich voneinander unterscheiden.

Das DNA-Doppelmolekül bildet eine i. d. R. rechtsgewundene Doppelhelix mit zwei außen verlaufenden Zucker-Phosphat-Bändern und nach innen gerichteten Basen. Sowohl die durch Wasserstoffbrücken gebildeten Basenpaarungen als auch die hydrophoben Bindungen zwischen den planaren aufgestapelten Basenringen stabilisieren die Helix.

Mit Ausnahme der mRNAs sind einzelsträngige RNA-Moleküle sehr unregelmäßig geformt, können aber, wie z. B. die tRNAs, innerhalb einer Funktionsklasse eine charakteristische Grundform annehmen. Die meisten stark aufgefalteten RNA-Moleküle haben eine strukturelle (z. B. einige rRNAs) oder eine katalytische Funktion (z. B. Ribozyme, snRNAs und wiederum rRNAs).

VORKOMMEN			RIBONUCLEOPROTEINE (RNPs)	
DNA	Einzelstrang (single strand, ss), linear	RNA	RNAs sind als Sekundärstrukturen dargestellt.	
Genom einiger Viren (z. B. Parvoviren)		mRNA, rRNA, tRNA und viele andere kleine RNAs bei allen lebenden Zellen; Genom einiger Viren (z. B. TMV, Picornaviren)		
DNA	Doppelstrang (double strand, ds), linear	RNA	Ribosom	Telomerase
Genom aller Eukaryoten und einiger Viren (z. B. Herpes- und Pockenviren)		Genom einiger Viren (z. B. Reoviren)		
DNA	Einzelstrang (single strand, ss), circular	RNA	SRP	U2 snRNP (Komponente eines Spleißosoms)
Genom einiger Viren (z. B. Microviridae)		Genom der Viroide		
DNA	Doppelstrang (double strand, ds), circular	RNA		
Genom aller Prokaryoten, Mitochondrien, Plastiden; Plasmide; Genom einiger Viren (z. B. Bacteriophagen λ und T4)		bisher unbekannt		

RNP ist ein Sammelbegriff für Strukturen, die aus Proteinen und npc-(non-protein-coding)-RNAs bestehen. Man klassifiziert sie in cytoplasmatische (z. B. Ribosomen und SRPs), nucleäre (z. B. Telomerasen und Komponenten der Spleißosomen) und nucleoläre RNPs (z. B. snoRNPs). Typische Akronyme sind sc (small cytoplasmatic), sn (small nuclear), sno (small nucleolar) und hn (heterogenous nuclear).

TRIPLEX-STRUKTUREN

Einige chromosomalen Elemente enthalten Bereiche mit einem zusätzlichen dritten Strang, der den ursprünglichen Helixstrang verdrängt. Es entsteht eine Triplex-Struktur, die allgemein als **D-Loop** (Displacement-Loop) bezeichnet wird. Handelt es sich bei dem zusätzlichen Strang um ein RNA-Molekül, nennt man den Komplex auch R-(RNA-)Loop.

D-Loops befinden sich z. B. in den Nucleoiden von Mitochondrien (☞ S. 195, mtDNA) und einigen Plastiden, in bestimmten Plasmiden und einigen Viren. Sie sind wesentlicher Bestandteil der Replikationsinitiation (☞ S. 185, Replikation). In Telomeren dienen D-Loops dem Schutz der überhängenden Einzelstrangbereiche, die zurückgebogen und in die Helix eingeflochten werden (s. S. 22). Bei einigen Rekombinationsprozessen treten D-Loops als vorübergehende Erscheinung auf.

Strukturvarianten von D-Loops



Die RNA-Moleküle sind häufig Transkripte, die am DNA-Strang verbleiben; sie sind an Regulationsprozessen beteiligt.

Bei Telomeren wird der überhängende Einzelstrang in die Helix desselben Moleküls eingebunden (s. S. 22).