

I

Allgemeine Grundlagen und Präanalytik

1

Grundlagen der molekularen Diagnostik

Frank Thiemann

Die molekulare Diagnostik umfasst den Nachweis verschiedenster Krankheitserreger wie Viren und Bakterien oder relevanter Mutationen im menschlichen Genom auf der Ebene der Nucleinsäuren.

Die Nucleinsäuren können mit verschiedenen Verfahren vervielfältigt und nachgewiesen werden. Die verschiedenen Techniken in der molekularen Diagnostik bedienen sich dabei der spezifischen Eigenschaften dieser Moleküle, deren chemische Grundstruktur immer die Gleiche ist.

1.1

Die DNA

Die DNA setzt sich aus vier unterschiedlichen *Nucleotiden* zusammen. Ein Nucleotid besteht aus einem *Zucker*, einem *Phosphat* und aus einer der vier stickstoffhaltigen Basen *Adenin*, *Thymin*, *Cytosin* oder *Guanin*.

Der Zucker ist eine *Ribose*, deren Grundgerüst aus fünf Kohlenstoffatomen (C) aufgebaut ist (Pentose). Diese Kohlenstoffatome sind von 1' bis 5' gekennzeichnet. Am 1'-C-Atom ist jeweils eine Base mit dem Zucker verknüpft. Am 2'-C-Atom der Ribose ist keine OH-Gruppe vorhanden, sondern zwei Wasserstoffmoleküle. Aus diesem Grund bezeichnet man diese Ribose auch als *Desoxyribose* (Desoxyribonucleinsäure, DNS; das A in DNA steht für *acid* = Säure).

Die Nucleotide sind über eine *Phosphodiester-Bindung*, die zwischen der OH-(Hydroxyl)-Gruppe des 3'-C-Atoms und dem Phosphat des 5'-C-Atoms des benachbarten Nucleotids gebildet wird, kovalent miteinander verbunden. Es entsteht somit eine Polynucleotidkette mit einem *Zuckerphosphat-Rückgrat*, deren beiden Enden sich chemisch voneinander unterscheiden. An dem einen Ende ist die *3'-OH-Gruppe*, an dem anderen Ende die *freie Phosphatgruppe* des 5'-C-Atoms lokalisiert. Jeder DNA-Strang hat somit eine chemische Polarität. Konventionell bezeichnet man die beiden Enden als *5'- und 3'-Ende* (Abb. 1.1 und Abb. 1.2).

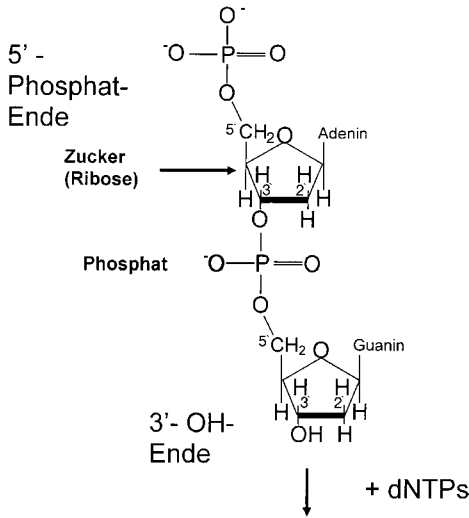


Abb. 1.1 Ausschnitt eines DNA-Einzelstranges. Die DNA besteht aus Zuckermolekülen (Ribose), die jeweils am 5'- und 3'-Kohlenstoff (C) über ein Phosphat (P) miteinander verknüpft sind (Zucker-Phosphat-Rückgrat). Mit dem ersten Kohlenstoffatom des Zuckers sind die einzelnen Basen Guanin, Cytosin, Thymin oder Adenin verbunden. Am 5'-Ende einer DNA steht immer ein Phosphat, am 3'-Ende ist die OH-Gruppe lokalisiert. An dieser 3'-OH-

Gruppe wird von der DNA-Polymerase ein weiteres Nucleotid eingebaut. Am 2'-Kohlenstoff der Ribose sind bei der DNA zwei Wasserstoffatome (H) gebunden. Bei der RNA ist hier eine weitere OH-Gruppe vorhanden. Daher die Bezeichnungen RNA = Ribonukleinsäure und DNA = Desoxyribonukleinsäure (A = Acid: Säure) (Abbildung aus: *MTA-Dialog*, Heft 4, 2000, S. 28, mit freundlicher Genehmigung der Hoppenstedt Bonnier Zeitschriften GmbH, Darmstadt).

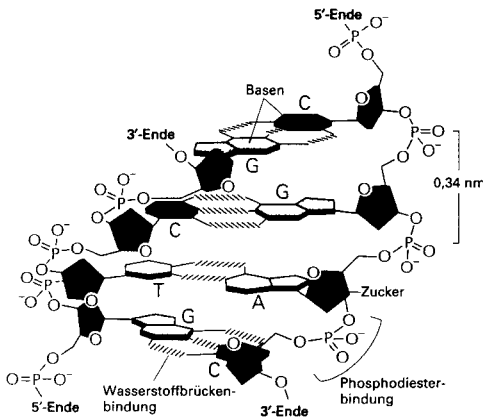


Abb. 1.2 Ausschnitt eines DNA-Doppelstranges. In einem DNA-Doppelstrang sind Guanin und Cytosin über drei, Adenin und Thymin über zwei Wasserstoffbrückenbindungen komplementär miteinander verbunden. Daher sind

DNA-Abschnitte mit einem hohen GC-Gehalt stabiler als Abschnitte, die reich an Adenin und Thymin sind (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

Die dreidimensionale Struktur der DNA wurde 1953 von Watson und Crick mit Hilfe von Röntgenstrukturanalysen, die Rosalind E. Franklin und Maurice Wilkins am Londoner King's College durchführten, aufgeklärt. Sie postulierten anhand der erzielten Ergebnisse die doppelhelikale (schraubenförmig gewundene) Struktur der DNA. In der *DNA-Doppelhelix* winden sich zwei DNA-Stränge unterschiedlicher Polarität um eine senkrecht im Raum stehende Achse. Das Zuckerphosphat-Rückgrat zeigt dabei nach außen, die Basen liegen in Innern der Doppelhelix. Die beiden DNA-Stränge sind über *Wasserstoffbrückenbindungen* der Basen miteinander verbunden. Die chemische Struktur der Basen erlaubt nur zwischen Adenin und Thymin und zwischen Cytosin und Guanin die Ausbildung dieser nicht kovalenten Wasserstoffbindungen. Die Basenabfolge des einen Stranges muss also komplementär zur Basenabfolge des anderen Stranges sein. Man bezeichnet zwei so miteinander verbundene DNA-Stränge als *komplementäre* DNA-Stränge und die entsprechenden Basen als komplementäre Basen oder Basenpaar.

Eine Umwindung in der DNA-Doppelhelix umfasst 10,4 Basenpaare (Abb. 1.3).

In der Abfolge der Basen der DNA ist die genetische Information verschlüsselt. Da die DNA-Stränge in einer Doppelhelix komplementär sind, kann jeder Strang

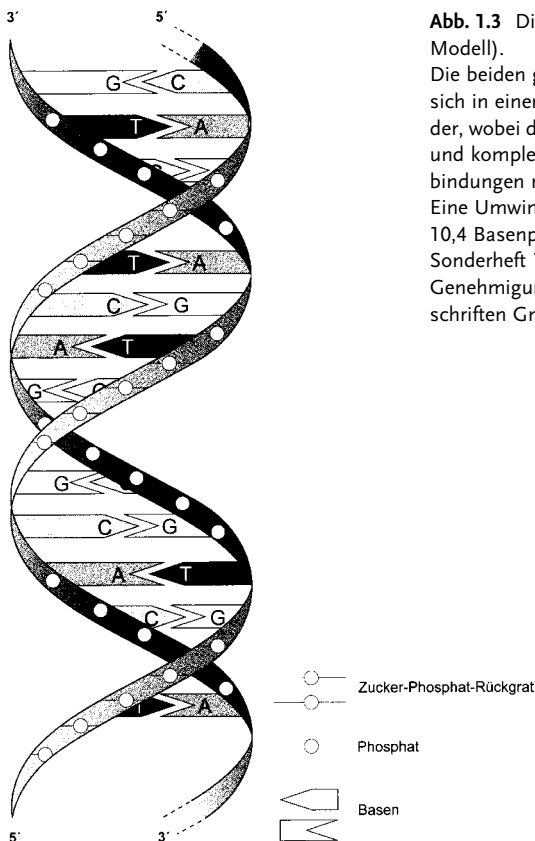


Abb. 1.3 Die DNA-Doppelhelix (Watson-Crick-Modell).

Die beiden gegenläufigen DNA-Stränge winden sich in einer rechtshändigen Spirale umeinander, wobei die Basen ins Innere der Helix weisen und komplementär über Wasserstoffbrückenbindungen miteinander verbunden sind.

Eine Umwindung in der Doppelhelix umfasst 10,4 Basenpaare (Abbildung aus: *MTA-Praxis*, Sonderheft 1, 2002, S. 10, mit freundlicher Genehmigung der Hoppenstedt Bonnier Zeitschriften GmbH, Darmstadt).

als Vorlage (*Matrize*) zur Herstellung eines neuen DNA-Stranges dienen. Werden die Basen der beiden Matrizenstränge komplementär ergänzt, erhält man zwei identische doppelsträngige DNA-Moleküle. Die genetische Information auf der DNA wird so verdoppelt oder *repliziert*. Da bei der *Replikation* jeweils ein DNA-Strang der DNA-Doppelhelix erhalten bleibt und als Vorlage dient, spricht man auch von der *semikonservativen Replikation* der DNA. In der Zelle (*in vivo*) läuft dieser Vorgang während jeder Zellteilung ab, damit gewährleistet wird, dass jede Tochterzelle die gleiche genetische Information erhält (s. Abschnitt 1.3). In Kapitel 4 wird gezeigt, wie man sich diese Eigenschaft der DNA zunutze macht, um diese *in vitro* zu vervielfältigen (*amplifizieren*) und nachzuweisen.

1.2

Die RNA

Von ihrer chemischen Zusammensetzung unterscheidet sich die RNA nur in zwei wesentlichen Punkten von der DNA (Abb. 1.4).

Sie enthält

1. eine Ribose anstatt einer Desoxyribose und
2. Uracil anstelle von Thymin als Base.

Die Nukleotide der RNA sind, wie in der DNA, über eine Phosphodiesterbindung miteinander verknüpft.

Die DNA liegt in der Zelle als doppelsträngiges Molekül vor. Die RNA hingegen ist ein *einzelsträngiges* Molekül. Bedingt durch interne komplementäre Abschnitte innerhalb eines RNA-Einzelstranges, die sich über Wasserstoffbrücken miteinander verbinden, kann die RNA verschiedene Formen und Strukturen annehmen.

Man unterscheidet hauptsächlich drei verschiedene RNAs, die unterschiedliche Aufgaben in der Zelle wahrnehmen:

1. mRNA: *messenger-RNA*
2. tRNA: *Transfer-RNA*
3. rRNA: *ribosomale RNA*

Die Funktionen der einzelnen RNA-Moleküle werden in Abschnitt 1.7 näher erläutert.

1.3

DNA-Replikation

Während der Zellteilung wird die genetische Information der Elternzelle auf die zwei neu entstehenden Tochterzellen verteilt. Damit jede Tochterzelle die identi-

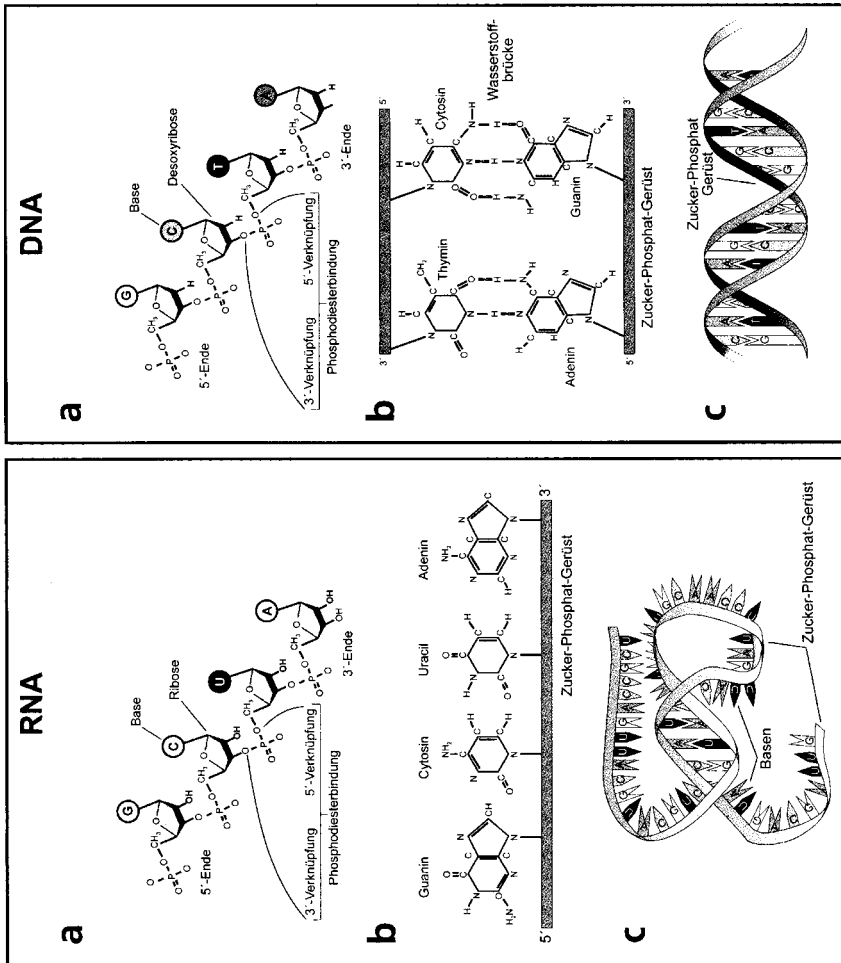


Abb. 1.4 Vergleich von RNA und DNA.

(A) Das Zucker-Phosphat-Rückgrat:

Die DNA enthält im Zucker-Phosphat-Gerüst eine Desoxyribose. Die RNA enthält eine Ribose.

Desoxyribonukleinsäure und Ribonukleinsäure (Säure = **Acid**)

In beiden Molekülen sind die Zucker über eine Phosphodiester-Bindung miteinander verknüpft.

(B) Die Basen:

Beide Nukleinsäuren beinhalten 4 Basen, von denen sich 2 jeweils komplementär miteinander paaren. In der DNA paart sich Adenin mit Thymin und Guanin mit Cytosin. Die RNA

enthält statt Thymin die Base *Uracil*. Die Basenpaarungen erfolgen in beiden Molekülen über spezifische, nichtkovalente Wasserstoffbrückenbindungen.

(C) Die Struktur:

Die DNA liegt in der Zelle als doppelsträngiges Molekül vor.

Die RNA ist ein einzelsträngiges Molekül. Kurze Abschnitte können sich aber komplexitär miteinander paaren und Sekundärstrukturen ausbilden.

(Abbildung aus: *MTA-Praxis*, Sonderheft 1, 2002, S. 14, mit freundlicher Genehmigung der Hoppenstedt Bonnier Zeitschriften GmbH, Darmstadt.)

sche genetische Ausstattung erhält, muss die DNA vorher verdoppelt werden. Dieser Vorgang wird als *DNA-Replikation* bezeichnet (Abb. 1.5).

Die DNA-Replikation ist ein enzymatisch gesteuerter Prozess. Die DNA-Doppelhelix wird zunächst durch das Enzym *Helikase* aufgetrennt. Die beiden nun vorliegenden DNA-Einzelstränge der *Replikationsgabel* werden durch Einzelstrangbindende Proteine (*single stranded binding proteins*, SSB) stabilisiert. Jeder Einzelstrang dient als Vorlage (*Matrize*, *Template*) für den neu zu synthetisierenden Strang.

Das Enzym *DNA-Polymerase* benötigt für diese komplementäre Ergänzung der Matrize ein Startermolekül (*Primer*) das eine freie 3'-OH-Gruppe für die chemische Anknüpfung des komplementären Nukleotids zur Verfügung stellt. Der Primer ist ein kurzes RNA-Molekül, das von einer *Primase* synthetisiert wird.

Die DNA-Polymerase erkennt das Startermolekül und beginnt mit der komplementären Strangsynthese in 5'→3'-Richtung, indem sie die vorliegenden Desoxyribonukleotid-Triphosphate unter Abspaltung zweier Phosphat-Gruppen (PP_i = *Pyrophosphat*) über eine *Phosphodiester-Bindung* miteinander verknüpft. Durch die Abspaltung des Pyrophosphats wird die Energie für diese Verbindung zur Verfügung gestellt.

Eine kontinuierliche Synthese in 5'→3'-Richtung der DNA ist aber nur an einem Strang möglich (*Leit-Strang-Synthese*, *leading strand*). Die Ergänzung des anderen Stranges (*Folge-Strang-Synthese*, *lagging strand*) erfolgt diskontinuierlich. Die Pri-

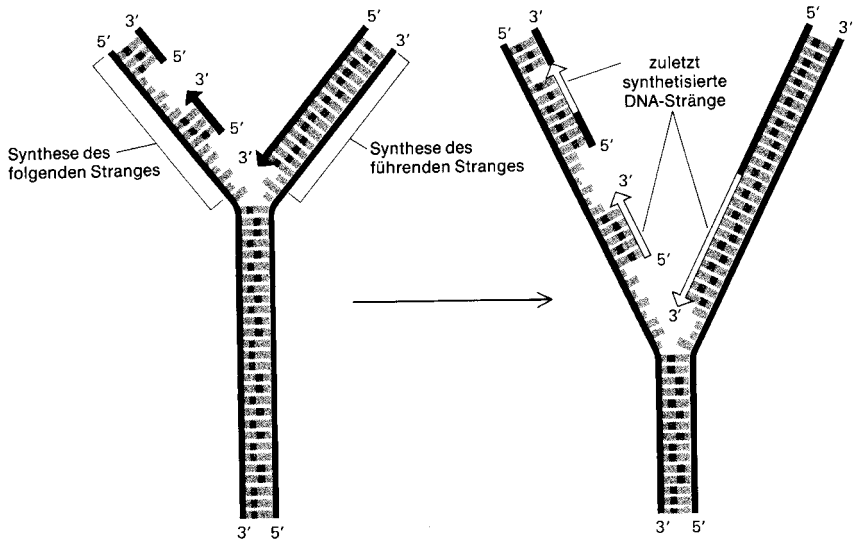


Abb. 1.5 DNA-Replikation.

Die beiden DNA-Stränge der DNA-Helix haben gegensätzliche Polaritäten. Die DNA-Polymerase kann aber einen DNA-Strang nur in 5'→3'-Richtung synthetisieren. Daher verläuft die DNA-Replikation des führenden oder Leit (*leading*)-Stranges kontinuierlich und die des

folgenden oder Folge (*lagging*)-Stranges diskontinuierlich. Die Replikationsgabel ist *asymmetrisch*. Während der Replikation bleibt ein parental DNA-Strang erhalten (*semi-konservative Replikation*) (Abbildung aus: Alberts, B. et al., 2004, *Molekularbiologie der Zelle*, 4. Auflage, Wiley-VCH, Weinheim).

mase synthetisiert hier im Abstand von 100–200 Nukleotiden *RNA-Primer*, die von der DNA-Polymerase verlängert werden. Die so entstandenen DNA-Fragmente nennt man *Okazaki-Fragmente*. Die Synthese eines Okazaki-Fragmentes endet, wenn die DNA-Polymerase auf einen vor ihr liegenden RNA-Primer stößt.

Die RNA-Primer werden anschließend von einem DNA-Reparatursystem entfernt und durch DNA ersetzt. Eine *DNA-Ligase* knüpft abschließend die Phosphodiester-Bindung zwischen den benachbarten Okazaki-Fragmenten.

Da bei der enzymatischen Verdopplung immer ein "alter" DNA-Strang erhalten bleibt, bezeichnet man diesen Prozess als *semikonservative DNA-Replikation*.

Die DNA-Replikation läuft in Pro- und Eukaryonten ähnlich ab. Sie startet immer an einer definierten Stelle im Genom, dem *Replikationsursprung* (*origin of replication*, ORI). Prokaryonten haben pro Genom meist nur einen ORI, während Eukaryonten aufgrund ihres größeren Genoms die Replikation an mehreren ORIs gleichzeitig initiieren müssen.

Die an der DNA-Verdopplung beteiligten Enzyme liegen in der Zelle in einem Multi-Enzym-Komplex vor, um diesen Prozess effizient ablaufen zu lassen. Eukaryonten replizieren ihre DNA mit einer Geschwindigkeit von etwa 50 Nukleotiden/Sekunde, Prokaryonten mit 500–1000 Nukleotiden/Sekunde.

1.4

Das Gen

Proteine sind Polypeptide, die jeweils eine bestimmte Anzahl von spezifischen Aminosäuren enthalten. Es kann entweder aus einem Polypeptid und/oder verschiedenen Polypeptiden bestehen. Jedes Polypeptid wird von einem definierten Abschnitt auf der DNA codiert. Diesen Abschnitt bezeichnet man als *Gen*, die Gesamtheit der Gene eines Organismus als *Genom*.

Die Genome bei Pro- und Eukaryonten sind unterschiedlich organisiert (s. Abschnitt 1.5 und 1.6). Aufgrund dieser Unterschiede ergeben sich auch andere Abläufe bei der *Proteinbiosynthese*.

1.5

Genomorganisation bei Prokaryonten

Zu den *Prokaryonten* zählt man die *Eubacteria* und die *Archaea*. Diese einzelligen Mikroorganismen besitzen keinen abgegrenzten, membranumschlossenen Zellkern.

Das Genom besteht aus einem großen, *ringförmigen doppelsträngigen* DNA-Molekül das *superhelikal* (überspiralisiert) aufgewunden ist. Es wird auch als *Kernäquivalent* oder *Nukleoid* bezeichnet, da es sich in elektronenmikroskopischen Abbildungen deutlich von dem umgebenen Cytoplasma abhebt.

Bakterien besitzen außerdem kleinere ringförmige doppelsträngige DNA-Moleküle, die *Plasmide*. Plasmide enthalten immer einen Replikationsursprung und ein bis zwei Gene, die z. B. für Virulenz-, Konjugationsfaktoren oder Antibiotikaresistenzen codieren. Sie replizieren unabhängig vom Bakteriengenom und schaffen dem Bakterium manchmal einen Selektionsvorteil. Plasmide, die sich in das Wirtsgenom integrieren können, bezeichnet man als *Episomen*.

Die Genomgröße bei Prokaryonten variiert zwischen 1×10^6 – 5×10^6 Basenpaaren (bp). 85% des prokaryotischen Genoms, das entspricht circa 4000–5000 Genen, codiert für Proteine oder RNA, der Rest sind regulatorische DNA-Sequenzen und nur sehr wenig nicht codierende Sequenzen.

1.6

Genomorganisation bei Eukaryonten

Tiere, Pflanzen, Pilze und Protisten (Algen, Protozoen, Schleimpilze) werden zu den *Eukaryonten* gezählt.

Eukaryonten besitzen im Gegensatz zu den Prokaryonten einen membranumschlossenen *Zellkern (Nukleus)* und weitere durch Membranen voneinander abgegrenzte *Zellkompartimente* oder Organellen wie Mitochondrien, Golgi-Vesikel, Endoplasmatisches Retikulum, usw.

Im Nukleus ist die DNA einer Zelle lokalisiert: Fast alle zellulären Prozesse werden von hier aus koordiniert. Mitochondrien und Chloroplasten enthalten neben dem Zellkern noch eigene DNA, die für einen Teil ihrer eigenen Stoffwechsellzyme codieren. Mitochondriale DNA (*mtDNA*) wird maternal vererbt. Diese Tatsache macht man sich bei der Untersuchung evolutionärer Verwandtschaftsverhältnisse zunutze. Stammbäume werden anhand von Sequenzunterschieden der mtDNA verwandter Populationen aufgestellt.

Das menschliche Genom, das hier als Beispiel eines typischen eukaryotischen Genoms dienen soll, besteht aus ca. $3,2 \times 10^9$ Basenpaaren (bp), die für ca. 30.000 Gene codieren. Die DNA einer menschlichen Zelle wäre als Faden abgewickelt etwa 1,80 m lang. Sie liegt im Zellkern komplexiert mit basischen Proteinen, den *Histonen*, vor, um die sie in definierten Abständen gewickelt ist. Durch andere, sog. *Nicht-Histon-Proteine*, wird die DNA weiter verpackt. Während der Zellteilung wird die DNA so kompakt verdichtet, dass sie im Lichtmikroskop in Form der *Chromosomen* sichtbar wird (Chromos = Farbe, Soma = Körper) (Abb. 1.6).

Der Mensch hat 46 Chromosomen, davon zwei geschlechtsbestimmende *Gonosomen* (XX oder XY) und 44 *Autosomen* (diploider Chromosomensatz).

70% der DNA ($2,1 \times 10^9$ bp) entfallen auf Bereiche, die keine Gene enthalten. Nur etwa 30% ($0,9 \times 10^9$ bp) der DNA lässt sich bestimmten Genen zuordnen, wobei hiervon nur 3% ($0,09 \times 10^9$ bp) direkt codierenden Sequenzen entsprechen. Dieses muss bedeuten, dass die Gene Sequenzen enthalten, *auf denen keine*

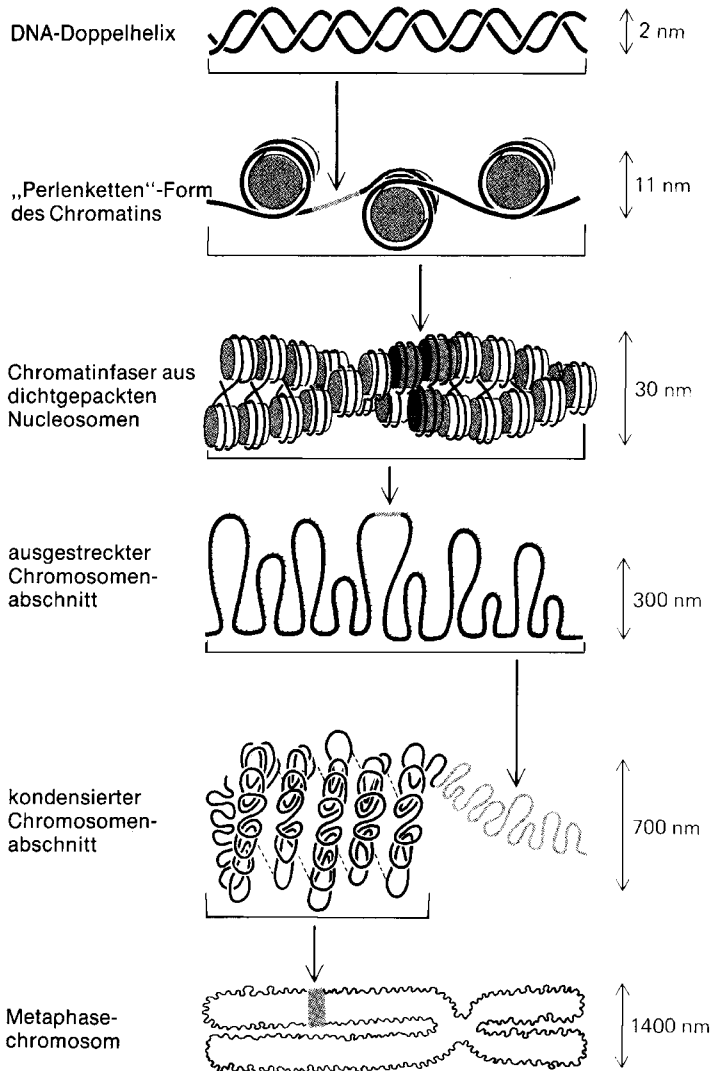


Abb. 1.6 DNA-Verpackung.

Übersicht über die Verpackung der DNA (von oben nach unten):

Die DNA ist in regelmäßigen Abständen um basische Proteinkomplexe, den Histonen, gewunden. Diese sog. Nucleosomen werden durch Nicht-Histon-Proteine weiter zur 30 nm-

Faser gepackt. Durch Faltung der 30 nm-Faser entstehen Schleifendomänen, die weiter kondensieren und das Chromosom in der Metaphase sichtbar werden lassen (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

Information codiert ist. Tatsächlich sind solche Abschnitte identifiziert worden. Sie werden als *Introns* bezeichnet. Die codierenden Abschnitte eines Gens nennt man *Exons*.

Introns werden nach der Transkription aus der *prä-mRNA* herausgeschnitten (s. Abschnitt 1.7.1.3).

1.7

Die Proteinbiosynthese

trallala

Wird ein Protein in der Zelle benötigt, muss die in dem entsprechenden Gen enthaltene Information abgelesen und weiter prozessiert werden, damit das Protein synthetisiert werden kann. Die Nukleotidsequenz eines Gens wird zuerst abgeschrieben (*transkribiert*) und mittels eines Boten (*messenger*) zu dem Ort der Proteinsynthese, den *Ribosomen*, transportiert. Anschließend wird diese in die entsprechende Aminosäuresequenz übersetzt (*translatiert*) und das Genprodukt synthetisiert.

1.7.1

Die Transkription

Die Transkription ist ein Prozess, der hauptsächlich von den *RNA-Polymerasen*, katalysiert wird. (Abb. 1.7). Prokaryonten besitzen nur eine, Eukaryonten hingegen besitzen drei verschiedene RNA-Polymerasen.

Die RNA-Polymerase I transkribiert die meisten Gene, die für die ribosomale RNA (rRNA) codieren (s. Abschnitt 1.7.2.3), die RNA-Polymerase III transkribiert hauptsächlich die Gene für die Transfer-RNAs (tRNA) (s. Abschnitt 1.7.2.2). Für die Transkription der proteincodierenden Gene in Eukaryonten ist die *RNA-Polymerase II* zuständig.

Generell bestehen diese Enzyme aus mehreren Untereinheiten, die als *Transkriptionsfaktoren* bezeichnet werden.

Bakterielle RNA-Polymerasen benötigen den *Sigma* (σ)-Faktor zur Transkription.

Die RNA-Polymerase II der Eukaryonten benötigt dagegen mehrere, als *allgemeine Transkriptionsfaktoren* (*TFIIA*, *TFIIB* usw.) bezeichnete Untereinheiten.

Sie erkennen und binden an spezifische DNA-Sequenzen im Genom, den *Promotoren*. Ein Promotor liegt oberhalb (*upstream*) am 5'-Ende eines Gens. Die DNA-Sequenz verschiedener bakterieller Promotoren ist sehr heterogen. Diese Heterogenität ist dafür verantwortlich, wie häufig ein Gen transkribiert wird, denn RNA-Polymerasen haben zu bestimmten Sequenzen eine größere Affinität. Gene mit "starken" Promotoren werden demnach häufiger abgelesen. Sie codieren für Genprodukte, die in der Zelle häufiger oder in größerer Anzahl benötigt werden.

Ein Promotor enthält zwei kurze DNA-Abschnitte, die im Laufe der Evolution konserviert wurden (*Konsensus-Sequenzen*). Eine Konsensus-Sequenz (TTGACA) liegt ca. 35 Basenpaare, die andere (TATATT; sog. TATA-Box) ca. 10–25 Basenpaare *upstream* vom Startpunkt des jeweiligen Gens entfernt.

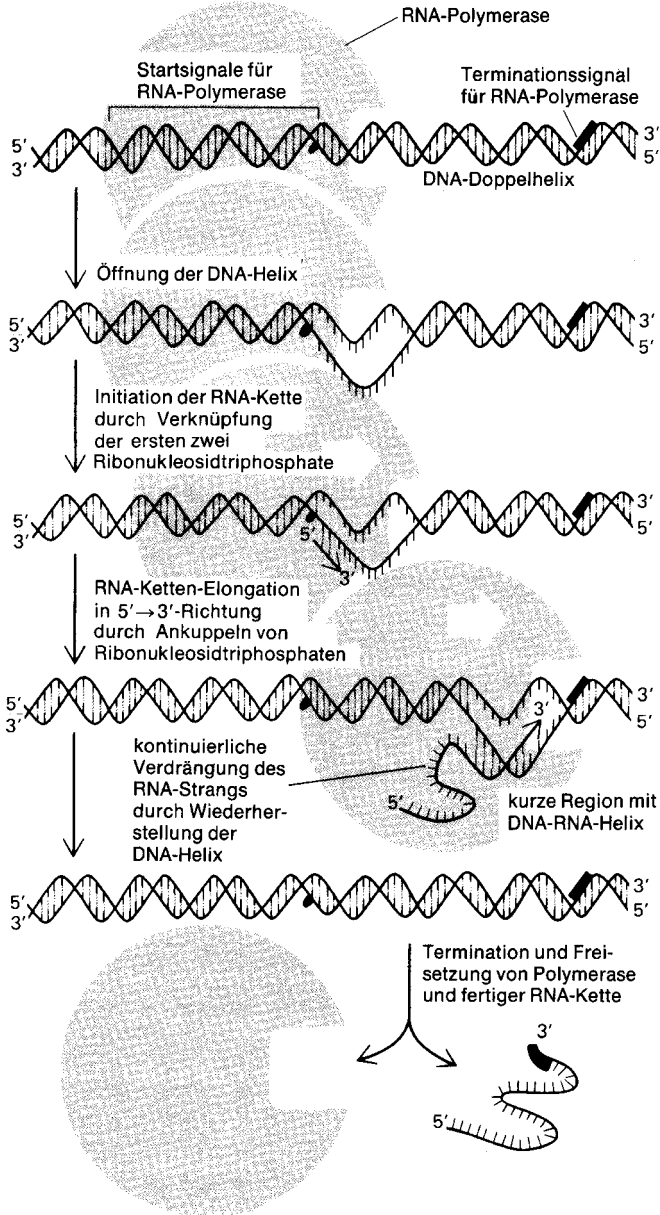


Abb. 1.7 Transkription.

Die RNA-Polymerase bindet an den Promotor eines Gens und beginnt mit der Transkription. Die DNA-Helix wird aufgewunden und die mRNA synthetisiert. Die Ribonukleosidtriphosphate werden komplementär zum abgelesenen DNA-Strang eingebaut (codogener Strang). Die

DNA-Helix schließt sich hinter der RNA-Polymerase und das Transkript wird verdrängt. Erreicht das Enzym den Terminator, endet die Synthese. Polymerase und mRNA werden freigesetzt (Abbildung in Anlehnung an: Alberts, B. et al., 2004, *Molekularbiologie der Zelle*, 4. Auflage, Wiley-VCH, Weinheim).

Promotoren bei Eukaryonten enthalten noch weitere Konsensus-Sequenzen (BRE-, INR- oder DPE-Element), die für die jeweilige "Stärke" eines Promotors verantwortlich sind. Weitere regulatorische DNA-Abschnitte können sehr weit vom Startpunkt entfernt liegen.

Hat die RNA-Polymerase mit Hilfe der Transkriptionsfaktoren an den Promotor gebunden, wird die DNA-Doppelhelix geöffnet und entwunden. Dieser Prozess benötigt keine Energie. Die Lage und Richtung des Promotors entscheidet darüber, welcher der nun vorliegenden DNA-Einzelstränge als Vorlage (Matrize) für die RNA-Polymerase dient. Das Enzym synthetisiert aus den zukommenden Nukleotidtriphosphaten ATP, GTP, UTP und CTP ein RNA-Polymer, die *messenger-RNA* (*mRNA*). Die Nukleotide werden komplementär zum Matrizenstrang in 5'→3'-Richtung eingebaut. Sie werden enzymatisch durch eine Phosphodiester-Bindung kovalent miteinander verbunden. Die Energie für diese Reaktion wird durch die Abspaltung des Pyrophosphats (PPi) von den Nukleotidtriphosphaten bereitgestellt. Der DNA-Doppelstrang schließt sich hinter der RNA-Polymerase wieder.

Das Enzym synthetisiert so lange, bis es auf eine weitere spezielle DNA-Sequenz, den *Terminator* trifft. Diese DNA-Sequenz veranlasst die Polymerase sich von der DNA zu lösen. Die synthetisierte mRNA und der Matrizenstrang werden freigesetzt.

Die mRNA der Prokaryonten wird direkt anschließend translatiert.

Bei Eukaryonten wird sie weiter prozessiert bevor sie vom Zellkern in das Cytoplasma zu den Ribosomen transportiert wird. Man bezeichnet diese mRNA daher auch als *prä-mRNA* (Abb. 1.8 a und Abb. 1.8 b).

1.7.1.1 Die 5'-Cap-Struktur (Capping)

Der erste Schritt hin zu einer "reifen" eukaryotischen mRNA ist das sog. *Capping*. Dabei wird ein modifiziertes Guaninnukleotid an das 5'-Ende der prä-mRNA gebunden (Abb. 1.9). Dieser enzymatisch gesteuerte Prozess beginnt, wenn die RNA-Polymerase ca. 25 Nukleotide der prä-mRNA synthetisiert hat und lässt sich in drei Schritte unterteilen:

1. Das 5'-Phosphat der prä-mRNA wird von einer *Phosphatase* entfernt.
2. Das Enzym *Guanylyltransferase* überträgt ein Guanosin-Nukleotid.
3. Eine *Methyltransferase* methyliert das Guanosin (*7-Methyl-Guanosin*).

Die 5'-Kappe markiert die mRNA und kann so in der Zelle von anderen RNAs unterschieden werden, denn die von den anderen RNA-Polymerasen synthetisierten RNAs werden nicht auf diese Art modifiziert (s. Abschnitt 1.7.1).

Dieser chemisch modifizierte Teil der mRNA wird von dem *Cap-Binding-Komplex* gebunden. Dieser Komplex spielt eine wichtige Rolle beim Export der RNA aus dem Zellkern in das Cytoplasma, zum Ort der *Translation*. Die 5'-Kappe dient dann als Bindungsstelle der Ribosomen (s. Abschnitt 1.7.2.4).

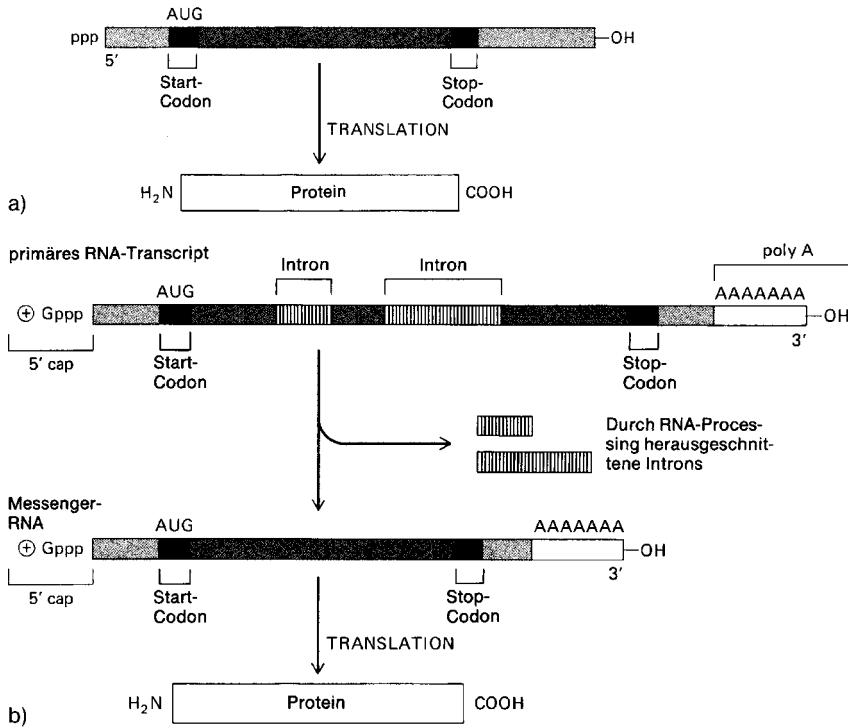


Abb. 1.8 (a) Die mRNA der Prokaryoten. Die mRNA der Prokaryoten enthält keine Introns und wird nicht weiter prozessiert, sondern direkt translatiert. (b) Die mRNA der Eukaryoten. Die prä-mRNA der Eukaryoten enthält Introns, die von dem Spleißosom heraus

geschnitten werden. Erst dann wird das Transkript translatiert.

Außerdem erhält sie eine 5'-Kappe (Cap) und einen poly(A)-Schwanz (Abbildung in Anlehnung an: Alberts, B. et al., 2004, *Molekularbiologie der Zelle*, 4. Auflage, Wiley-VCH, Weinheim).

1.7.1.2 Der Poly(A)-Schwanz

Das 3'-Ende einer eukaryotischen prä-mRNA wird ebenfalls modifiziert. Am Ende eines transkribierten DNA-Abschnitts befinden sich spezifische Sequenzen. Diese Signalsequenzen werden von mit der RNA-Polymerase II assoziierten Proteinen auf der RNA erkannt. Die mRNA wird am 3'-Ende spezifisch geschnitten und das Enzym *Poly(A)-Polymerase* synthetisiert ca. 200–250 Adenin-Nukleotide an, wobei die Länge von den *Poly(A)-Bindungsproteinen* festgelegt wird.

Poly(A)-Polymerasen benötigen keine Matrize. *Der Poly(A)-Schwanz der mRNA ist also nicht im Genom codiert.*

Der Poly(A)-Schwanz wird von Poly(A)-Bindeproteinen (PABP) besetzt und ist notwendig für die Initiation der Translation. Er schützt die RNA außerdem vor RNA-abbauenden Enzymen (*RNAsen*).

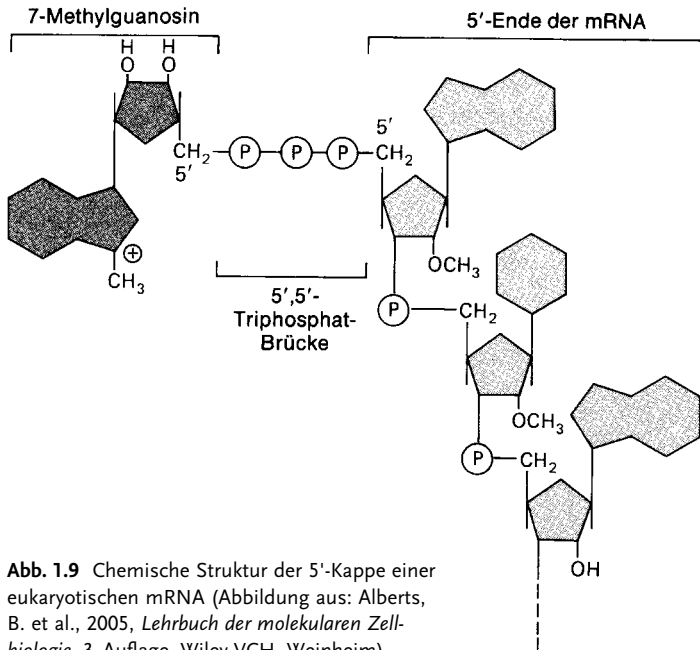


Abb. 1.9 Chemische Struktur der 5'-Kappe einer eukaryotischen mRNA (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

1.7.1.3 Spleißen von RNA

Die weitere Prozessierung der während der Transkription synthetisierten prä-mRNA wird als *Spleißen* bezeichnet. Die in der prä-mRNA enthaltenen *Introns* werden herausgeschnitten und die *Exons* direkt miteinander verbunden (Abb. 1.10).

An diesem Prozess sind hauptsächlich kurze, mit Proteinen assoziierte RNA-Moleküle beteiligt, die sog. *small nuclear ribonuclearproteins (snRNP)*. Sie bilden mit weiteren Proteinen das *Spleißosom*.

Das Spleißosom erkennt die Sequenzen auf der prä-mRNA, die als *Spleißsignale* dienen. Die beiden Enden des Introns werden durch das Spleißosom in räumliche Nähe zueinander gebracht. Nach zwei *Phosphoryltransfer-Reaktionen* und mannigfaltigen Umlagerungen der RNPs werden die Exons exakt miteinander verknüpft. Die ausgeschnittene Intronsequenz liegt am Ende dieses Prozesses in Form eines "Lassos" oder "Lariats" vor. Sie wird im Zellkern abgebaut.

Die so prozessierte RNA wird jetzt als *mRNA* bezeichnet und kann im Cytoplasma der eukaryotischen Zelle translatiert werden.

Ein eukaryotisches Gen kann *alternativ gespleißt* werden. Die Anzahl der herausgespleißten Introns kann dabei variieren. Aus einem *Primärtranskript* können so unterschiedliche mRNAs für verschiedene Proteine entstehen.

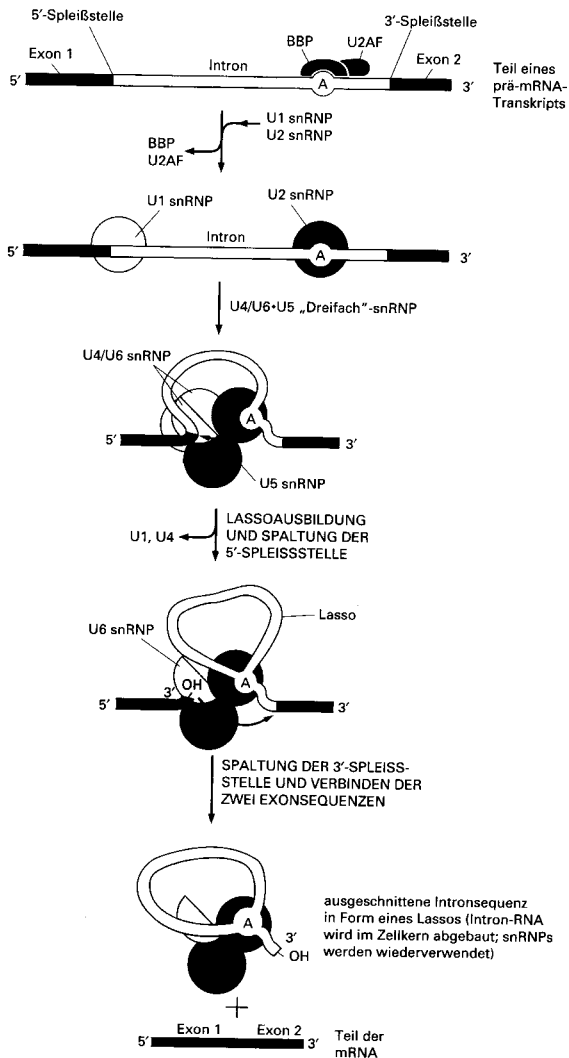


Abb. 1.10 Der Mechanismus des RNA-Spleißens. Das Spleißen von RNA wird durch die snRNPs (*small nuclear ribonuclearproteins*) katalysiert, die mit anderen, nicht dargestellten Proteinen, das Spleißosom bilden. Das Spleißosom erkennt die Spleißsignale auf der prä-mRNA und führt die beiden Enden des Introns zusammen. Durch mehrfache Umlagerungen der in den RNPs enthaltenen RNA wird das aktive Zentrum innerhalb des Spleißosoms

geschaffen und die jeweils passenden Abschnitte der prä-mRNA zusammen geführt. An der Reaktion sind insgesamt über 50 Proteine beteiligt, von denen einige ATP hydrolysieren, um die Umlagerungen zwischen den RNAs anzutreiben.

Die ausgeschnittene Intronsequenz wird im Zellkern abgebaut. Die snRNPs werden wieder verwendet (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

Der genetische Code

| Position 1 (5'-Ende) ↓ | Position 2 | | | | Position 3 (3'-Ende) ↓ |
|------------------------------|--------------------------|--------------------------|----------------------------|---------------------------|------------------------------|
| | U | C | A | G | |
| U | Phe Phe Leu Leu | Ser Ser Ser Ser | Tyr Tyr STOP STOP | Cys Cys STOP Trp | U C A G |
| C | Leu Leu Leu Leu | Pro Pro Pro Pro | His His Gln Gln | Arg Arg Arg Arg | U C A G |
| A | Ile Ile Ile Met | Thr Thr Thr Thr | Asn Asn Lys Lys | Ser Ser Arg Arg | U C A G |
| G | Val Val Val Val | Ala Ala Ala Ala | Asp Asp Glu Glu | Gly Gly Gly Gly | U C A G |

Aminosäuren und ihre Symbole

Codons

| | Aminosäure | Protein | Codon 1 | Codon 2 | Codon 3 | Codon 4 |
|---|------------|----------------|---------|---------|---------|-------------|
| A | Ala | Alanin | GCA | GCC | GCG | GCU |
| C | Cys | Cystein | UGC | UGU | | |
| D | Asp | Asparaginsäure | GAC | GAU | | |
| E | Glu | Glutaminsäure | GAA | GAG | | |
| F | Phe | Phenylalanin | UUC | UUU | | |
| G | Gly | Glycin | GGA | GGC | GGG | GGU |
| H | His | Histidin | CAC | CAU | | |
| I | Ile | Isoleucin | AUA | AUC | AUU | |
| K | Lys | Lysin | AAA | AAG | | |
| L | Leu | Leucin | UUA | UUG | CUA | CUC CUG CUU |
| M | Met | Methionin | AUG | | | |
| N | Asn | Asparagin | AAC | AAU | | |
| P | Pro | Prolin | CCA | CCC | CCG | CCU |
| Q | Gln | Glutamin | CAA | CAG | | |
| R | Arg | Arginin | AGA | AGG | CGA | CGC CGG CGU |
| S | Ser | Serin | AGC | AGU | UCA | UCC UCG UCU |
| T | Thr | Threonin | ACA | ACC | ACG | ACU |
| V | Val | Valin | GUA | GUC | GUG | GUU |
| W | Trp | Tryptophan | UGG | | | |
| Y | Tyr | Tyrosin | UAC | UAU | | |

Abb. 1.11 Der genetische Code und die Aminosäuren.

Ein Codon besteht aus drei Nukleotiden (Triplet). Jedes Codon codiert für eine spezifische Aminosäure. Einige Aminosäuren werden allerdings von mehreren Codons determiniert. Dabei ist immer die dritte Base

eines Triplets variabel. Dargestellt ist hier die Sequenz der mRNA.

Die Aminosäuren werden konventionell durch einen oder drei Buchstaben abgekürzt wiedergegeben (Abbildung aus: Alberts, B. et al., 2004, *Molekularbiologie der Zelle*, 4. Auflage, Wiley-VCH, Weinheim).

1.7.2

Die Translation

Die auf der DNA codierte genetische Information ist von der RNA-Polymerase abgeschrieben (*transkribiert*) worden und muss jetzt übersetzt (*translatiert*) werden, um ein funktionelles Protein zu erhalten.

1.7.2.1 **Der genetische Code**

Die DNA ist der Speicher der genetischen Information. Diese Information ist in der Abfolge der Basen, der Basen- oder Nukleotidsequenz, verschlüsselt oder *codiert*. Wie wird dieser Code in Proteine übersetzt?

Proteine bestehen aus einer bestimmten Abfolge von Aminosäuren. Der Code für die Abfolge der Aminosäuren ist durch die Nukleotidsequenz der DNA festgelegt. Eine spezifische Sequenz von drei Nukleotiden (*Triplet-Codon*) entspricht jeweils einer Aminosäure. Jedes Codon ist durch die drei Basen genau definiert und kann auf eine bestimmte Aminosäure zurückgeführt werden. Allerdings kann eine Aminosäure durch mehr als ein Codon codiert sein. Die größte Variabilität weisen die Triplets dabei an ihrer dritten Position auf. Diese Variabilität bezeichnet man als *Redundanz*. Geringfügig verschiedene Nucleotidsequenzen können demnach für dieselben Aminosäuresequenzen codieren (Abb. 1.11).

Spezifische Triplets werden als *Start- oder Stop-Codons* bezeichnet. Diese sind spezifische Codons, an denen die Proteinbiosynthese startet bzw. stoppt. Sie definieren den *offenen Leserahmen (ORF)* einer DNA-Sequenz (Abb. 1.12).

Der genetische Code ist universal, d.h. alle Organismen benutzen, bis auf einige Ausnahmen, die gleichen Codons für dieselben Aminosäuren.

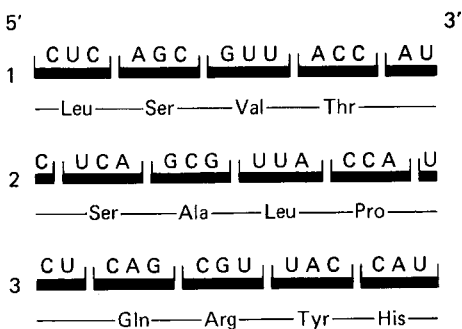


Abb. 1.12 Leseraster (Open Reading Frame).

Das Leseraster einer DNA ist durch die Triplets, die jeweils für eine spezifische Aminosäure codieren, festgelegt. Die Verschiebung des Leserasters um nur ein

Nukleotid ergibt in diesem Beispiel eine andere Abfolge von Aminosäuren (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

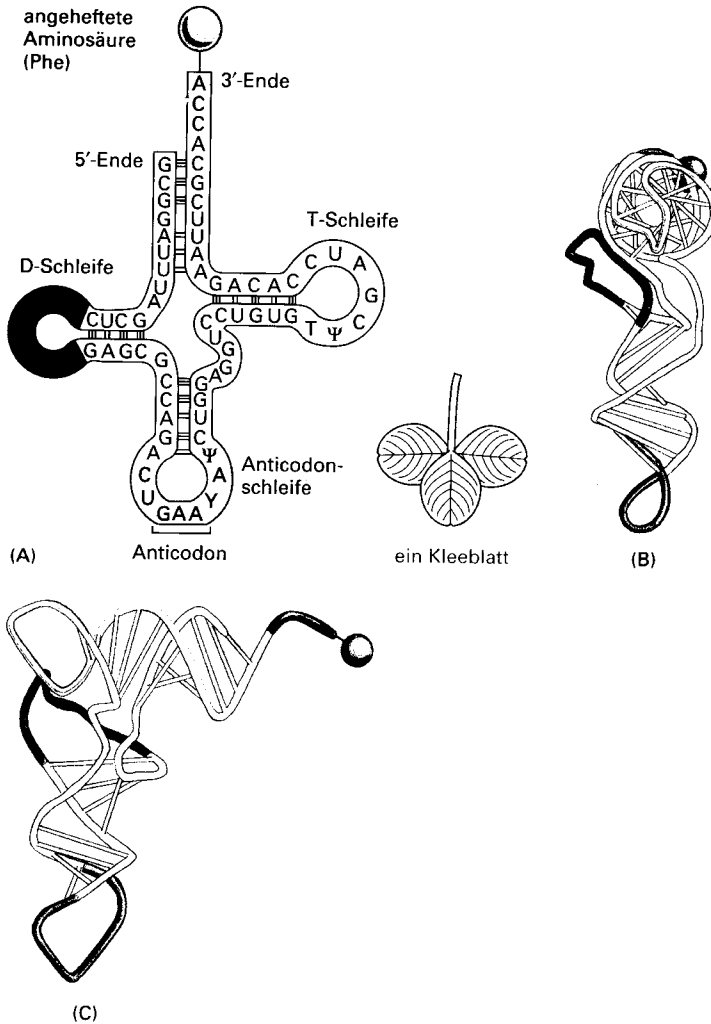


Abb. 1.13 Die tRNA.

In der Abb. ist die spezifische tRNA für Phenylalanin dargestellt.

(A) Die Kleeblattstruktur: In der zweidimensionalen Kleeblattstruktur werden die komplementären Basenpaarungen innerhalb eines tRNA Moleküls sichtbar, die die doppelhelikalen Strukturen des Moleküls ausmachen. Das Anticodon besteht aus drei Nucleotiden. Dieser Bereich bindet über komplementäre Basenpaarungen während der Translation mit dem Codon der mRNA. An das 3'-Akzeptor-

Ende bindet die Aminoacyl-Transferase die spezifische Aminosäure. Die nicht gepaarten Bereiche bilden Schleifen aus (Anticodon-, D- und T-Schleife). tRNAs enthalten ungewöhnliche Basen, die nach der tRNA-Synthese durch chemische Modifikationen entstehen (ψ = Pseudouridin, D = Dihydrouridin). (B) und (C) In der dreidimensionalen Ansicht wird die L-förmige Struktur der tRNA sichtbar (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

1.7.2.2 Transfer-RNA

Transfer-RNAs oder tRNAs sind RNA-Moleküle, die die Aminosäuren zum Ort der Proteinsynthese transportieren. Sie haben eine durchschnittliche Länge von ca. 80 Nukleotiden und zeichnen sich durch ihre besondere Struktur aus. Diese kommt durch interne komplementäre Basenpaarungen zustande. Dadurch entsteht ein Molekül mit doppelhelikalen und einzelsträngigen Bereichen. Zweidimensional erscheint die tRNA als *kleblattförmige Struktur*, die durch weitere Auffaltungen eine *dreidimensionale L-förmige Struktur* annimmt (Abb. 1.13.).

Transfer-RNAs enthalten Introns, die wie bei einer prä-mRNA herausgespleißt werden müssen. Der Mechanismus des tRNA-Spleißens kommt bei Pro- und Eukaryonten vor. Der Spleißvorgang unterscheidet sich aber von dem des mRNA-Spleißens.

Transfer-RNAs enthalten außerdem *chemisch modifizierte* RNA-Nukleotide, die entweder zur Konformationsstabilisierung oder zur besseren Erkennung des mRNA-Codons dienen.

Die wichtigsten funktionellen Bereiche einer t-RNA sind das *3'-Akzeptorende*, an der sie mit der für sie spezifischen Aminosäure beladen wird und das *Anticodon*, das komplementär zum Codon der mRNA ist.

Die Beladung der tRNA mit der spezifischen Aminosäure erfolgt durch Enzyme, den *Aminoacyl-tRNA-Synthetasen*. Eine beladene tRNA wird allgemein als *Aminoacyl-tRNA* bezeichnet. Die jeweilige Aminosäure beziehungsweise das Anticodon legt die spezifische Bezeichnung fest. Eine tRNA, die die Aminosäure *Lysin* transportiert, wird als *Lysyl-tRNA* bezeichnet.

1.7.2.3 Die Ribosomen

Die eigentliche Proteinsynthese findet im Cytoplasma der Zelle, an den Ribosomen statt. Ribosomen enthalten mehr als *50 verschiedene Proteine* die mit den *ribosomalen RNAs (rRNA)* assoziiert sind.

Ribosomen der Pro- und Eukaryonten sind sehr ähnlich und bestehen aus einer *großen* und einer *kleinen Untereinheit* (Abb. 1.14). Die einzelnen Untereinheiten als auch das gesamte Ribosom werden oft mit der *S-Einheit* versehen (z. B. 50S und 30S-Untereinheit beim Prokaryonten-Ribosom). Die S-Einheit gibt die Sedimentationsgeschwindigkeit der Ribosomen während einer Ultrazentrifugation an.

Die Untereinheiten lagern sich erst unmittelbar vor Beginn der Proteinsynthese in der Nähe des 5'-Bereichs der mRNA zusammen. Ein funktionelles Ribosom hat eine *Bindungsstelle für die mRNA* und drei Bindungsstellen (*A-, P- und E-Bindungsstelle*) für die Aminoacyl-tRNAs.

1.7.2.4 Der Beginn der Proteinsynthese

Für den Translationsstart wird sowohl in Pro- als auch in Eukaryonten eine *Initiator-tRNA* benötigt. Die spezielle tRNA ist immer mit der Aminosäure *Methionin* beladen. In Prokaryonten ist diese Aminosäure allerdings modifiziert (*N-Formylmethionin*). Das entsprechende *Start-Codon* auf der mRNA ist AUG.

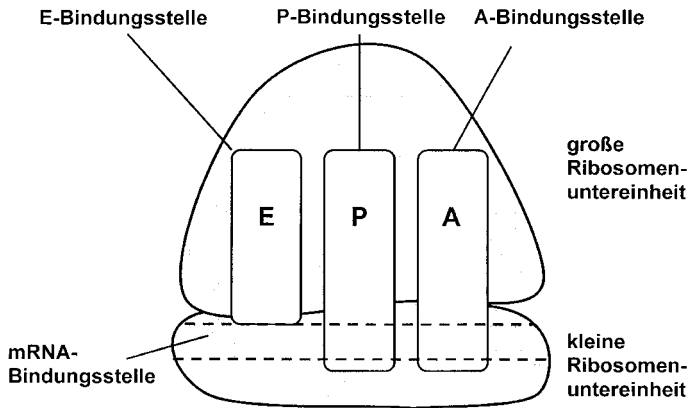


Abb. 1.14 Das Ribosom mit seinen RNA-Bindungsstellen. Jedes Ribosom hat drei Bindungsstellen für die tRNA: A-Bindungsstelle (Aminoacyl-tRNA), P-Bindungsstelle (Peptidyl-tRNA) und Ausgang (Exit). Die mRNA bindet an die kleine Untereinheit.

Die mRNA der Prokaryonten enthält eine Konsensus-Sequenz, die als *Bindungsstelle* für die kleine 16S-Einheit des Ribosoms fungiert (*Shine-Dalgarno-Sequenz*). Ist die mRNA an die 16S-Untereinheit gebunden, kommt die große Untereinheit dazu.

Die Initiator-tRNA der Eukaryonten bindet zuerst mit *eukaryotischen Initiationsfaktoren (eIFs)* an die kleine ribosomale Untereinheit. Die mRNA wird mittels der 5'-Cap-Struktur (s. Abschnitt 1.7.1.1) erkannt und ebenfalls gebunden. Die kleine Untereinheit wandert nun an der mRNA bis zum ersten AUG-Codon, wo sich die große Untereinheit anlagert.

Die Translation der mRNA kann beginnen.

1.7.2.5 Das Protein entsteht

Die mit Methionin beladene Initiator-tRNA befindet sich in der *P-Bindungsstelle* des Ribosoms. Eine mit ihrer spezifischen Aminosäure beladene Aminoacyl-tRNA bindet in der *A-Bindungsstelle*. Welche tRNA binden kann, wird durch die komplementäre Basenpaarung zwischen der mRNA und dem Anticodon der tRNA festgelegt. Zwischen der Aminosäure der zuvor gebundenen tRNA in der P-Stelle und der Aminosäure der zuletzt gebundenen tRNA in der A-Stelle wird durch die *Peptidyltransferase-Aktivität* der großen Ribosomen-Untereinheit die *Peptidbindung* geknüpft. Durch Konformationsänderung bewegt sich mRNA genau drei Nukleotide, also ein Codon, weiter. Die verbrauchte tRNA wird aus der *E-Stelle* entlassen und kann neu beladen werden. Die mit der wachsenden Polypeptidkette beladene tRNA befindet sich nun in der P-Bindungsstelle. Die A-Bindungsstelle ist wieder frei für die nächste Aminoacyl-tRNA (Abb. 1.15).

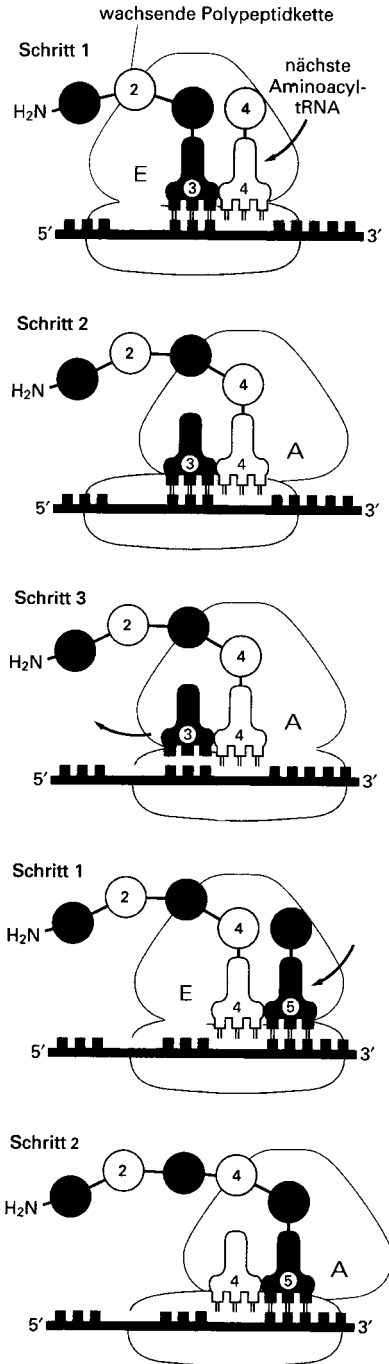


Abb. 1.15 Die Translation.

Die Proteinsynthese läuft zyklisch in drei immer wiederholenden Schritten ab:

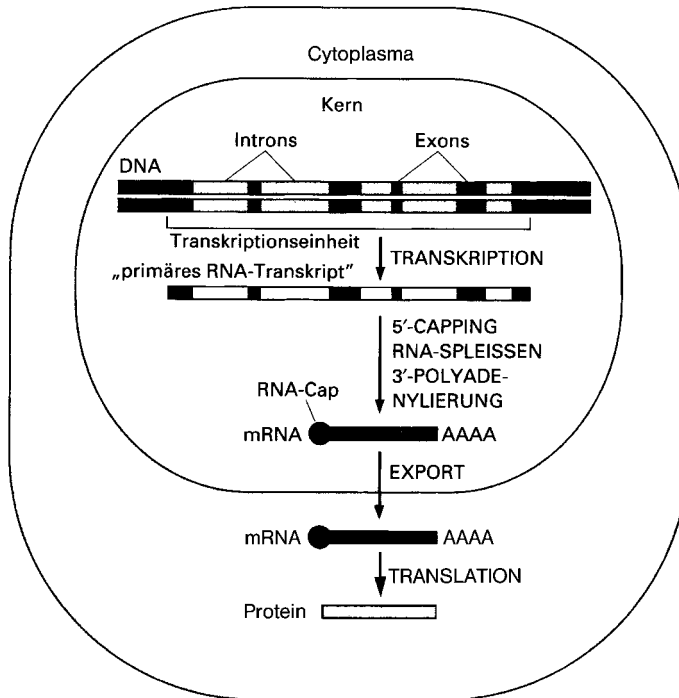
Schritt 1: Die Aminoacyl-tRNA bindet an die freie A-Bindungsstelle des Ribosoms. Durch die komplementäre Basenpaarung zwischen dem Codon der mRNA und der Aminoacyl-tRNA wird die Aminosäure bestimmt, die an das Ende der Polypeptidkette angehängt wird.

Schritt 2: Die Peptidyltransferase-Aktivität der großen Ribosomen-Untereinheit knüpft die Peptidbindung zwischen dem Carboxy-Ende der in der P-Bindungsstelle lokalisierten Aminosäure und dem Amino-Ende der Aminosäure in der A-Bindungsstelle. Das Carboxy-Ende der Polypeptidkette wird dabei von der in der P-Bindungsstelle liegenden tRNA abgekoppelt. Das Ribosom verändert dabei seine Konformation, wobei die beiden tRNAs in die E- und P-Bindungsstelle gezwungen werden.

Schritt 3: Die mRNA wird durch weitere Konformationsänderungen des Ribosoms genau drei Nucleotide weiter bewegt. Die nicht mehr beladene tRNA wird aus der E-Bindungsstelle entlassen und die ursprüngliche Konformation des Ribosoms wird wieder hergestellt.

Die Polypeptidkette wächst so von ihrem Amino- zu ihrem Carboxy-Ende. Der Zyklus endet, wenn ein Stop-Codon erreicht wird (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

(A) EUKARYONTEN



(B) PROKARYONTEN

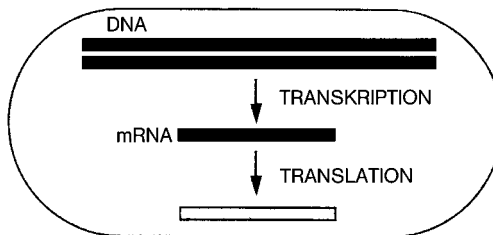


Abb. 1.16 Von der DNA zum Protein.
 (A) Eukaryonten: Die in dem Zellkern lokalisierte DNA enthält codierende (*Exons*) und nicht codierende (*Introns*) Abschnitte. Alle Abschnitte eines Gens werden transkribiert und das primäre Transkript, die prä-mRNA synthetisiert. Die prä-mRNA wird weiter prozessiert: die *Introns* werden durch das Spleißosom herausgeschnitten. Die mRNA erhält an ihrem 5'-Ende eine chemische Kappe, an ihrem 3'-Ende einen poly(A)-Schwanz.

Sie wird in das Cytoplasma der Zelle transportiert und an den Ribosomen translatiert.
 (B) Prokaryonten besitzen keinen Zellkern. Die Transkription und Translation der mRNA findet daher in einem gemeinsamen Zellkompartiment statt. Noch während der Transkription kann die mRNA translatiert werden (Abbildung aus: Alberts, B. et al., 2005, *Lehrbuch der molekularen Zellbiologie*, 3. Auflage, Wiley-VCH, Weinheim).

1.7.2.6 Das Ende der Translation

Drei spezifische Codons markieren das Ende der Translation. Diese *Stop-Codons* UAA, UGA oder UAG bewirken die Bindung von Freisetzungsfaktoren in der A-Bindungsstelle des Ribosoms. Die Peptidyltransferase transferiert daraufhin ein Wassermolekül anstatt eine Aminosäure auf das *Carboxy-terminale Ende* der Polypeptidkette, was wiederum zur Folge hat, dass die Bindung zwischen tRNA und Polypeptidkette sofort gelöst und das fertig synthetisierte Protein in das Cytoplasma entlassen wird.

Dort wird es weiter prozessiert und in seine funktionelle Form gebracht. Die *Sekundär- und Tertiärstruktur* ist dabei von der Aminosäurezusammensetzung abhängig. Hydrophobe Bereiche werden sich in das Innere des Moleküls falten, hydrophile nach außen. Hilfsproteine (*Chaperone*) helfen dabei, die Proteine in die richtige Struktur zu falten.

Signalsequenzen zeigen an, ob die Proteine in bestimmte Organellen transportiert, oder sezerniert werden sollen.

Abbildung 1.16 zeigt schematisch die Schritte der Transkription und Translation vergleichend bei Eukaryonten (A) und Prokaryonten (B).

