

## 4 Transkription, Translation und der genetische Code

Die in der Überschrift genannten Begriffe sind unter Molekularbiologen so etwas wie Allerweltswörter: man muss sie kennen und problemlos benutzen können. Das soll mit diesem Kapitel erreicht werden.

Zur Illustration beschreiben wir dabei im wesentlichen die Verhältnisse bei Bakterien – oder genauer bei der Bakterienart *Escherichia coli* (*E. coli*).

Dafür gibt es hauptsächlich zwei Gründe:

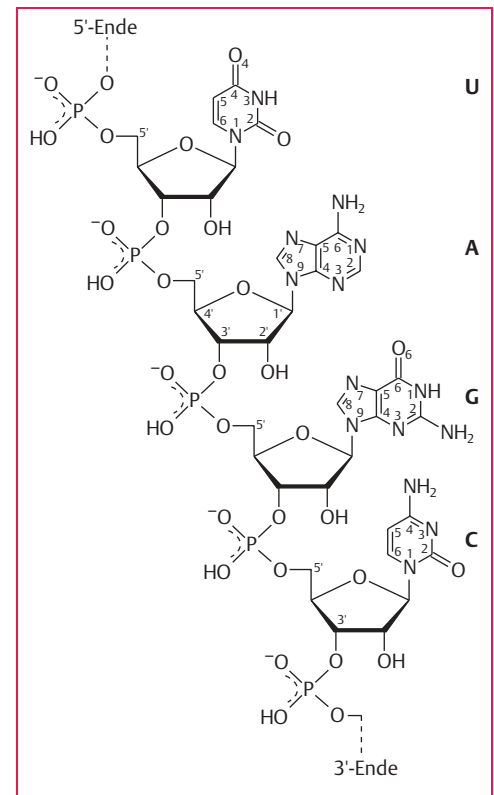
- Forschungsgeschichte. Die Art und Weise, wie genetische Information zur Synthese von Proteinen umgesetzt wird, konnte nach der Entdeckung der DNA-Doppelhelix innerhalb einer recht kurzen Zeitspanne von 10–15 Jahren aufgeklärt werden – zumindest in den Grundzügen. Das lag unter anderem daran, dass die beteiligten Forscher sich auf Untersuchungen an einem besonders einfachen biologischen System konzentrierten, nämlich auf Untersuchungen mit der harmlosen Bakterienart *E. coli*, die sich problemlos und ohne großen Aufwand in Laboratorien kultivieren lässt (S. 88).
- Vorteile für eine einführende Beschreibung. Obwohl Transkription und Translation im Prinzip bei allen Lebewesen gleich ablaufen, sind die Verhältnisse bei Bakterien doch erheblich einfacher als bei Eukaryoten. Dazu kommt noch, dass auch in der heutigen Zeit, in der sich viele Molekularbiologen bevorzugt für Eukaryoten interessieren, eine gründliche Kenntnis der Bakteriengenetik von Nutzen ist, denn Bakterien werden bei den meisten gentechnischen Verfahren eingesetzt (S. 300).

In diesem Kapitel geht es im wesentlichen um die Umsetzung der genetischen Information. Das ist ein Prozess in zwei Schritten:

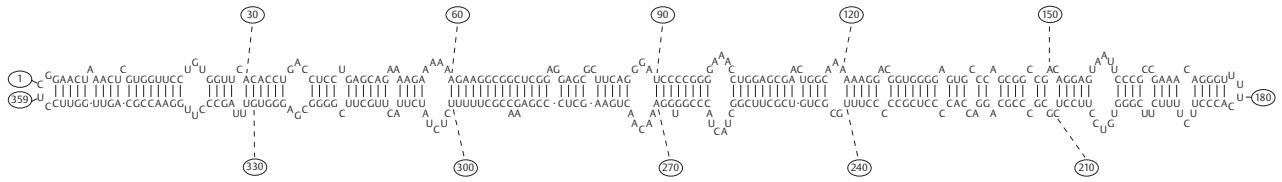
1. **Transkription:** Die Nucleotid-Folgen (Sequenzen) in den Genen der DNA werden „umgeschrieben“ – „transkribiert“ – und zwar in Form der Nucleotid-Folgen (Sequenzen) von RNA-Molekülen. Oder anders gesagt, Transkription ist die Synthese von RNA – und zwar so, dass die entstandenen RNA-Sequenzen komplementär zu einem der beiden DNA-Stränge sind.
2. **Translation:** Die RNA-Sequenzen werden „übersetzt“ – „translatiert“ – und zwar in Aminosäure-Sequenzen von Proteinen. Der genetische Code auf der RNA-Sequenz bestimmt die Reihenfolge der Aminosäuren im Protein.

### Transkription oder die Synthese von RNA

Ribonucleinsäuren, abgekürzt RNA (*ribonucleic acid*), sind Ketten von Nucleotiden, die durch Phosphodiester-Bindungen miteinander verknüpft sind (Abb. 4.1), entsprechend den Bindungen zwischen den Deoxynucleotiden in DNA-Strängen (S. 11). Aber wir registrieren wichtige Unterschiede zwischen den Ribonucleotiden der RNA und den Deoxy-



**Abb. 4.1 Nucleotide in der RNA.** Ein Nucleotid ist zusammengesetzt aus dem Fünf-Kohlenstoff-Zucker Ribose, einer Phosphatgruppe und einer von vier heterozyklischen Basen, nämlich Adenin (A), Guanin (G), Cytosin (C) und Uracil (U).



**Abb. 4.2 Ringförmige RNA.** Struktur des „*Potato spindle tuber viroid*“ (PSTV). Beachte, dass das Molekül aufgrund zahlreicher Basenpaarungen die Form eines Stäbchens einnimmt [nach 9]. Viroide infizieren Kartoffeln, Zitruspflanzen, Kokospalmen u. a. Sie verursachen oft Wachstumshemmung mit Verkrümmung und Vergilbung der Blätter.

ribonucleotiden der DNA: RNA enthält Ribose als Fünf-Kohlenstoff-Zucker sowie Uracil anstelle von Thymin (s. Abb. 2.1).

Durch die Art der Verknüpfung über Phosphatbrücken zwischen der 5'-OH-Gruppe in der Ribose eines Nucleotids und der 3'-OH-Gruppe in der Ribose des benachbarten Nucleotids erhält ein RNA-Molekül eine definierte Richtung mit einem freien 5'-Ende und einem freien 3'-Ende (Abb. 4.1).

Man kann die natürlich vorkommenden RNA-Arten der Zelle (Tab. 4.1) als unterschiedlich lange, unverzweigte und einzelsträngige Ketten von Ribonucleotiden ansehen. Diese Beschreibung ist jedoch nicht umfassend, denn RNAs neigen zur Ausbildung von Doppelstrang-Formen. Diese treten v. a. innerhalb einer RNA-Kette auf, sofern Abschnitte mit komplementären Ribonucleotid-Folgen in einem Molekül vorkommen.

Später werden wir viele teilweise komplexe Sekundärstruktur-Formen von RNA-Molekülen kennenlernen. Wir werden auch sehen, dass die Fähigkeit von RNA-Molekülen zur Schleifenbildung nicht selten wichtige strukturelle und genetische Konsequenzen hat.

Es gibt RNA-Moleküle ohne freie Enden: sie sind ringförmig geschlossen, wie die sogenannten **Viroide**, die zu den Erregern von wichtigen Pflanzenkrankheiten in tropischen und subtropischen Ländern gehören (Abb. 4.2). Auch aus der menschlichen Pathologie kennt man ringförmige RNA, nämlich als Genom des Hepatitis-Delta-Virus, das zusammen mit dem Hepatitis-B-Virus schwere Entzündungen der Leber verursacht. Die ringförmige RNA des **Hepatitis-Delta-Virus** besteht aus etwa 1700 Nucleotiden.

**Tab. 4.1 Einige wichtige RNA-Arten in der Zelle**

	Größe (ungefähre Angaben)	Funktion
transfer-RNA (tRNA)	80–95 Nucleotide	Übertragung von Aminosäuren zum Proteinsynthese-Apparat
ribosomale RNA (rRNA)	drei Arten (bei Bakterien) aus etwa 120, 1540 bzw. 2900 Nucleotiden	Struktur- und Funktionselemente von Ribosomen
messenger RNA (mRNA)	sehr verschieden (von einigen hundert bis zu mehreren tausend Nucleotiden)	die Boten-( <i>messenger</i> )-RNAs enthalten Abschriften der Gene und programmieren den Proteinsynthese-Apparat

Zellen enthalten zusätzlich noch viele andere Arten von RNA. Darüber berichten wir im passenden Zusammenhang in späteren Kapiteln.



Unter diesen Bedingungen kopiert die RNA-Polymerase die Nucleotid-Folge des DNA-Matrizenstranges nach den Regeln der Basenpaarung. Dort, wo in der DNA ein Guanin-Baustein steht, wird in der RNA ein Cytosin-Nucleotid eingebaut, und umgekehrt. Gegenüber einem Thymin- in der DNA gelangt ein Adenin-Nucleotid in die RNA, und gegenüber einem Adenin- ein Uracil-Nucleotid. Uracil nimmt also in der RNA die Stelle des Thymins ein. Die RNA-Polymerase knüpft ein Nucleotid nach dem anderen an das 3'-OH-Ende einer wachsenden RNA-Kette. Dabei werden die endständigen Pyrophosphate von den Ribonucleosid-Triphosphaten abgespalten und Phosphodiester-Bindungen gebildet.

Alle zellulären RNA-Polymerasen sind aus mehreren Untereinheiten aufgebaut: bakterielle RNA-Polymerasen bestehen aus fünf bis sechs Untereinheiten (Tab. 4.2) und eukaryotische RNA-Polymerasen aus mehr als einem Dutzend Untereinheiten (S. 337). Der Aufbau aus mehreren Untereinheiten ist keine notwendige Voraussetzung für die Synthese von RNA. Zum Beispiel sind die RNA-Polymerasen der Bakteriophagen T3 und T7 monomere Proteine mit einem Molekulargewicht von etwa 100 kDa. Dies gilt auch für die RNA-Polymerase in Mitochondrien. Aber diese RNA-Polymerasen sind hoch spezialisiert: Sie erkennen nur die Genanfänge der zugehörigen Genome.

Der komplexe Aufbau zellulärer RNA-Polymerasen hängt mit den vielfältigen Regulationen zusammen, die erforderlich sind, damit die einzelnen Gene einer Zelle spezifisch und genau nach den Bedürfnissen im Zuge der Reaktionen auf Umweltbedingungen aktiviert werden können.

Zurück zur bakteriellen RNA-Polymerase (Tab. 4.2). Für eine einfache RNA-Synthese von der Art der Abb. 4.3 genügt ein Komplex aus je einer  $\beta$ -Untereinheit und einer  $\beta'$ -Untereinheit sowie zwei  $\alpha$ -Untereinheiten, das so genannte **Minimal-** oder **Core-Enzym**. Doch für eine korrekte und effiziente Transkription bakterieller Gene ist ein **Holo-Enzym** notwendig, das außer den Untereinheiten des Minimal-Enzyms noch eine Sigma( $\sigma$ )-Untereinheit besitzt.

Die  **$\beta$ -Untereinheit und die  $\beta'$ -Untereinheit** bilden eine gemeinsame Struktur zur Bindung der DNA-Matrize und der wachsenden RNA-Kette.

**Tab. 4.2 RNA-Polymerase von *E. coli***

Untereinheit	Zahl der Untereinheiten/Enzym	Aminosäuren	Mol. Gew. [kDa]	Gen-Name im <i>E. coli</i> -Genom
$\beta'$	1	1407	155	<i>rpoC</i>
$\beta$	1	1342	150	<i>rpoB</i>
$\alpha$	2	329	36	<i>rpoA</i>
$\omega$ (omega)	1	91	10	<i>rpoZ</i>
$^{70}\sigma$ (sigma)	1	613	70	<i>rpoD</i>

Die Größe einer Untereinheit wird auf zweierlei Weise notiert, nämlich durch Angabe der Zahl der Aminosäuren, aus denen das Protein aufgebaut ist, und durch das Molekulargewicht, ausgedrückt in kDa (Definition auf S. 43).

*E. coli*-Zellen haben mehrere Sigma-Untereinheiten, je eine für die Transkription verschiedener Gen-Gruppen (s. S.104). Die Tabelle enthält als Beispiel die häufigste Sigma-Untereinheit – aufgrund ihres Molekulargewichtes als Sigma-70 oder  $^{70}\sigma$  bezeichnet.

Die Struktur trägt das aktive Zentrum, wo neue Nucleotide an das 3'-OH-Ende der RNA angeheftet werden.

Die  **$\alpha$ -Untereinheiten** übernehmen zwei Funktionen: Die aminoterminale Domäne trägt zum Zusammenbau und zur Stabilität des Gesamt-Enzyms bei; die carboxyterminale Domäne hilft bei der Bindung an die Promotor-DNA und beteiligt sich an der Wechselwirkung mit Regulatoren der Transkription (s. S. 120).

Die **kleinste Untereinheit** ( $\omega$ ) hat eine Funktion bei der Stabilisierung und Aufrechterhaltung der dreidimensionalen Gesamtstruktur der RNA-Polymerase.

Die **Sigma-Untereinheit** hat, wie gesagt, eine spezielle Aufgabe bei der Erkennung von Startstellen der Transkription vor Bakteriengenomen.

## Gen-Anfang: Der Promotor

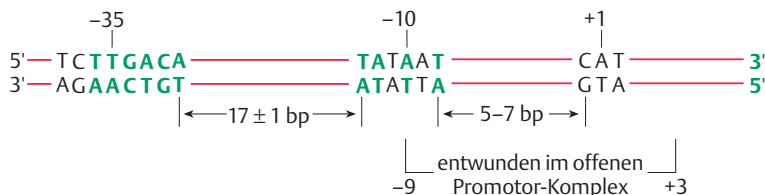
Wie wird die RNA-Polymerase an den Gen-Anfang geleitet? Wie unterscheidet sie den Strang, der transkribiert werden soll, vom komplementären Strang?

Die RNA-Synthese sollte nicht irgendwo auf der DNA beginnen, sondern genau vor einem Gen. Es sollte auch nicht irgendein Strang transkribiert werden, sondern nur der Strang, dessen Transkript die genetische Information trägt, der codogene oder Sinnstrang.

Die RNA-Polymerase (Holo-Enzym) bindet bevorzugt an Stellen auf der DNA, die vor einem Gen-Anfang liegen. Eine solche Erkennungs- und Bindestelle nennt man Promotor (*promoter*).

Man kennt die Nucleotid-Sequenzen der über tausend verschiedenen Promotoren im Genom von *E. coli*. Beim Vergleich dieser Sequenzen fallen einige Regelmäßigkeiten auf (Abb. 4.4):

1. In dem Bereich, der etwa 10 Basenpaare stromaufwärts vor dem Start der RNA-Synthese liegt, kommt oft eine Sequenz von Nucleotiden vor, die eine mehr oder weniger große Ähnlichkeit mit der Folge 5'-TATAAT-3' hat. Diese Sequenz wird gelegentlich nach ihrem Erstbeschreiber als **Pribnow-Box**, meist als **TATA-Box** oder einfach als **-10-Region** bezeichnet (wobei als +1 das Nucleotid am Startpunkt der RNA-Synthese angegeben wird).
2. In dem Bereich, der etwa 35 Nucleotide stromaufwärts vom Start liegt (in der **-35-Region**), gibt es innerhalb eines AT-reichen Abschnitts eine zweite Folge von oft vorkommenden Nucleotiden, im Idealfall 5'-TTGACA-3'.
3. Ein drittes Erkennungselement in vielen bakteriellen Promotoren liegt direkt aufwärts der -35-Region. Daher die Bezeichnung als **UP-Element** (*upstream*). Es enthält viele AT-Basenpaare und ist eine Stelle, wo die  $\alpha$ -Untereinheit der RNA-Polymerase an den Promotor gelangt.



**Abb. 4.4 Ein Musterpromotor des *E. coli*-Genoms.** Der Abstand zwischen dem Transkriptionsstart und dem ersten Nucleotid der -10-Region beträgt 5-7 Basenpaare (bp); der Abschnitt zwischen der -10-Region und -35-Region  $17 \pm 1$  bp. Der untere der beiden DNA-Stränge ist der transkribierte oder „codogene“ Strang, der obere der nicht-transkribierte Strang [nach 8].