

# Molekulargenetik

# 1

## In diesem Kapitel ...

... geht es um diese Themen:

- DNA-Struktur
- Gene
- Der genetische Code
- Von der DNA zum Protein
- DNA-Mutation
- Mutagene und DNA-Reparatur
- Genexpression
- Epigenetik und Chromatinmodifikation

## 1.1 DNA-Struktur

### 1.1.1 Nukleotide

Die Fähigkeit der DNA, die genetische Information zu übertragen, die eine Zelle benötigt, um sich selbst zu vermehren, ist eng mit der Struktur der DNA-Moleküle verbunden. Die DNA ist ein Polymer und besteht aus einer langen Kette von Monomeren, die man **Nukleotide** nennt. Das DNA-Molekül bezeichnet man als Polynukleotid. Jedes Nukleotid besteht aus drei Teilen: einem Zucker, einer stickstoffhaltigen Ringstruktur, die man **Base** nennt, und einer Phosphatgruppe. Der Zucker in der DNA ist eine Pentose, d. h. sie besitzt fünf Kohlenstoffatome. Sie wird 2'-Desoxyribose genannt, da die -OH-Gruppe am Kohlenstoffatom 2 der Ribose durch Wasserstoff ersetzt ist (Abb. 1.1). Die Kohlenstoffatome im Zucker sind von 1 bis 5 durchnummeriert. Den Zahlen wird ein **Strich** (') hinzugefügt, um sie von der Zahl der Atome in der Base zu unterscheiden. Die Nummerierung ist wichtig, da sie angibt, wo andere Komponenten des Nukleotids an den Zucker gebunden sind.

Nukleotide enthalten eine der vier Basen: **Adenin**, **Guanin**, **Thymin** oder **Cytosin** (Abb. 1.2). Dabei handelt es sich um komplexe Moleküle mit Kohlenstoff und Stickstoffringstrukturen. Adenin und Guanin, die zwei Kohlenstoff-Stickstoff-Ringe besitzen, nennt man **Purine**. Cytosin und Thymin besitzen nur einen

2 | 1 Molekulargenetik

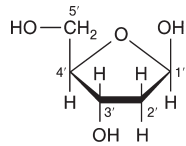
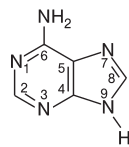
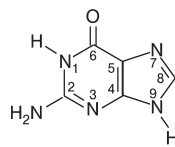


Abb. 1.1 Struktur von 2'-Desoxyribose.

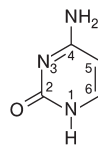
2'-Desoxyribose



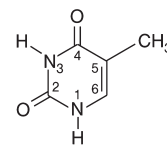
Adenin (A)



Guanin (G)



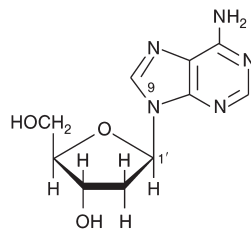
Cytosin (C)



Thymin (T)

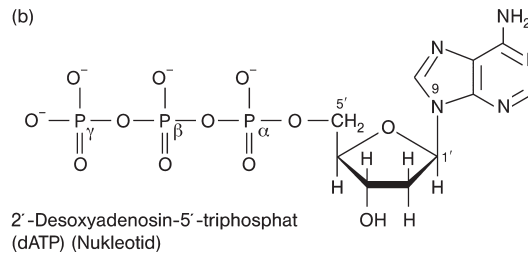
Abb. 1.2 Basen in der DNA.

(a)



2'-Desoxyadenosin (Nukleosid)

(b)



2'-Desoxyadenosin-5'-triphosphat (dATP) (Nukleotid)

Abb. 1.3 Struktur von (a) Nukleosiden und (b) Nucleotiden.

einzelnen Ring und werden als **Pyrimidine** bezeichnet. Die Basen sind mit dem Zucker über eine Bindung zwischen dem 1'-Kohlenstoff des Zuckers und einem Stickstoffatom an Position 9 des Purins oder an Position 1 des Pyrimidins verbunden. Die Kombination aus einem Zucker und einer Base nennt man **Nukleosid** (Abb. 1.3a).

Nucleotide besitzen am 5'-Kohlenstoffatom des Zuckers Phosphatgruppen ( $\text{PO}_4$ ) (Abb. 1.3b). Wenn an das Nucleosid eine Phosphatgruppe angehängt wird, nennt man es Nucleotid. Dieser Anhang kann aus einer, zwei oder drei miteinander verbundenen Phosphatgruppen bestehen. Die Phosphatgruppen bezeichnet man mit  $\alpha$ ,  $\beta$  und  $\gamma$ , wobei  $\alpha$  die direkt mit dem Zucker verbundene Phosphatgruppe ist. Nucleotide können in den Zellen als einzelne Moleküle (Nucleotidtriphosphate spielen in Zellen eine wichtige Rolle als Energieträger beim Antrieb enzymatischer Reaktionen) oder polymerisiert in Form von Nucleinsäuren (DNA oder RNA) vorkommen.

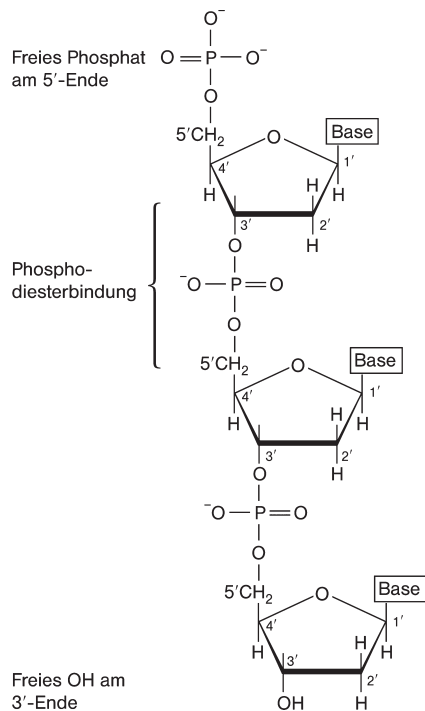


Abb. 1.4 Phosphodiesterbindungen verbinden die Nucleotide in einem DNA-Polynucleotid.

### 1.1.2 DNA-Polynucleotide

Nucleotidtriphosphate werden miteinander verbunden und ergeben Polynucleotide. Für die Synthese von DNA-Polynucleotiden werden vier verschiedene Nucleotidtriphosphate verwendet: 2'-Desoxyadenosin-5'-triphosphat (dATP oder A), 2'-Desoxythymidin-5'-triphosphat (dTTP oder T), 2'-Desoxycytosin-5'-triphosphat (dCTP oder C) und 2'-Desoxyguanosin-5'-triphosphat (dGTP oder G). Die  $\beta$ - und  $\gamma$ -Phosphate gehen während der Polymerisierung verloren, und die Nucleotide werden über die verbleibende Phosphatgruppe miteinander verknüpft. Das 5'-Phosphat des einen Nucleotids geht mit dem 3'-Kohlenstoffatom des nächsten Nucleotids unter Abspaltung der -OH-Gruppe am 3'-Kohlenstoff eine Bindung ein. Man nennt diese Bindung **3'-5'-Phosphodiesterbindung (C-O-P)** (Abb. 1.4). Die Polynucleotidkette besitzt ein freies 5'-Triphosphat am einen Ende (dem 5'-Ende) und eine freie 3'-Hydroxylgruppe am anderen Ende (dem 3'-Ende). Diese Unterscheidung verleiht dem DNA-Polynucleotid Polarität. Ein DNA-Molekül kann somit in 5'→3'- oder 3'→5'-Richtung beschrieben werden.

Die genetische Information ist in der Abfolge der Basen (Basensequenz) im DNA-Polynucleotid verschlüsselt (codiert). Diese Sequenz wird immer in 5'→3'-Richtung angegeben (Polymeraseenzyme kopieren DNA-Moleküle in dieser Richtung). Polynucleotide können extrem lang sein. Es gibt scheinbar keine Obergrenze für die Nucleotidanzahl und keine Einschränkung bezüglich der Nucleo-

#### 4 | 1 Molekulargenetik

tidsequenz. Die maximale Anzahl möglicher Basensequenzen für ein Polynukleotid ist  $4^n$ , wobei  $n$  die Anzahl der Nukleotide ist. Das ist eine gewaltige Zahl. Ein Polynukleotid aus sechs Basen kann zum Beispiel  $4^6 = 4096$  verschiedene Sequenzen haben.

### 1.1.3 Die Doppelhelix

DNA-Moleküle haben eine sehr spezielle und charakteristische dreidimensionale Struktur: die Doppelhelix (Abb. 1.5). Die Struktur der DNA wurde 1953 von Watson und Crick entdeckt. Sie arbeiteten in Cambridge und verwendeten Röntgenbeugungsaufnahmen von Franklin und Wilkins. DNA besteht aus zwei Polynukleotidketten, die umeinander gewunden sind und so eine Doppelhelix bilden. Der Zucker-Phosphat-Teil des Moleküls bildet ein Gerüst oder Rückgrat, das an der Außenseite der Helix liegt. Die Basen, bei denen es sich um flache Moleküle handelt, zeigen zueinander, leicht dezentriert nach innen. Sie sind wie Teller aufeinander gestapelt.

Röntgenbeugungsaufnahmen der Doppelhelix zeigen ein sich wiederholendes Bandenmuster. Es spiegelt die Regelmäßigkeit der DNA-Struktur wider. Für eine Windung benötigt die Doppelhelix 10 Basenpaare. Dies entspricht einer Länge von 34 Å, d. h. der Abstand zwischen zwei Basen beträgt 3,4 Å. Der Durchmesser der Helix beträgt 20 Å. Die Doppelhelix wird als **antiparallel** bezeichnet, d. h. einer der Stränge verläuft in 5' → 3'-Richtung, der andere in 3' → 5'-Richtung. Nur antiparallele Polynukleotide bilden eine stabile Helix. Die Doppelhelix ist nicht absolut regelmäßig. Wenn man sie von außen betrachtet, kann man eine **große Furche** und eine **kleine Furche** erkennen. Das spielt eine wichtige Rolle bei der Wechselwirkung mit Proteinen, bei der DNA-Replikation und bei der Expression der genetischen Information. Die Doppelhelix ist rechtsgängig, d. h. wenn die Doppel-

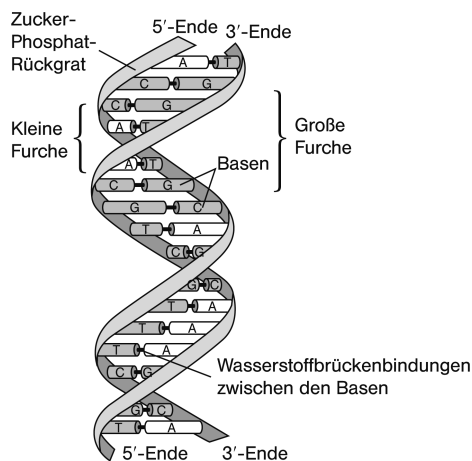


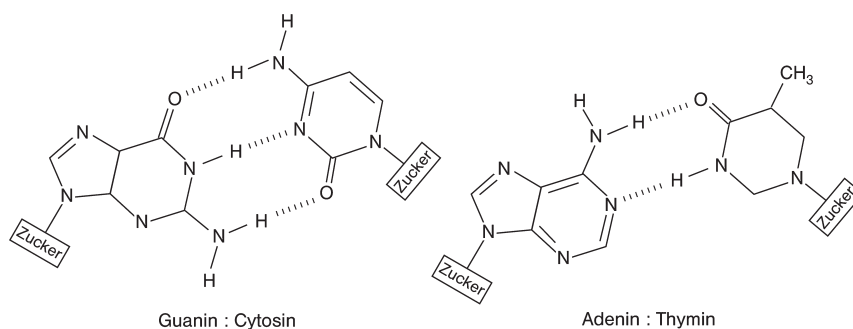
Abb. 1.5 Die Doppelhelix.

helix eine Wendeltreppe wäre, auf der man nach oben ginge, befände sich das Zucker-Phosphat-Rückgrat auf der rechten Seite.

Wenn man unter verschiedenen Bedingungen Kristalle eines DNA-Moleküls herstellt, entstehen viele verschiedene DNA-Formen. Die Form, die in Zellen vorkommt, nennt man **B-Form**. Eine andere Form, die A-Form, hat eine etwas kompaktere Struktur. Außerdem gibt es noch die C-, D-, E- und die Z-Form. Die Z-Form ist bemerkenswert, da es sich um eine linksgängige Helix handelt. Man hat in Chromosomen Bereiche identifiziert, die ungewöhnliche Strukturen wie zum Beispiel Z-DNA enthalten.

#### 1.1.4 Komplementäre Basenpaarung

Die Basen der beiden Polynukleotidketten treten miteinander in Wechselwirkung. Der Abstand zwischen den Polynukleotiden ist derart, dass ein Purin mit Doppelring mit einem Pyrimidin mit Einzelring in Wechselwirkung treten kann. Thymin wechselwirkt immer mit Adenin und Guanin mit Cytosin. Zwischen den Basen bilden sich Wasserstoffbrückenbindungen, die die Wechselwirkung stabilisieren. Zwischen A und T bilden sich zwei Wasserstoffbrückenbindungen, zwischen G und C drei. G-C-Bindungen sind also stärker als A-T-Bindungen. Die Art und Weise, wie die Basen zwischen den beiden DNA-Strängen Paare bilden, bezeichnet man als **komplementäre Basenpaarung** (Abb. 1.6). Sie ist von grundlegender Bedeutung. Andere Kombinationen als G-C und A-T funktionieren nicht, da sie entweder zu groß oder zu klein sind, um in die Helix zu passen. Sie richten sich nicht korrekt aus, sodass sich keine Wasserstoffbrückenbindungen bilden können. Da sich G immer mit C verbrücken muss und A mit T, stehen die Sequenzen der beiden Stränge miteinander in Beziehung. Man nennt sie komplementär. Ist also die Sequenz des einen Strangs bekannt, so kann man die Sequenz des anderen Strangs vorhersagen und bestimmen. Das ist ein lebensnotwendiger Mechanismus zur Erhaltung der genetischen Information und zur Weitergabe der Information an andere Zellen bei der Zellteilung. Komplementäre Basenpaarung ist auch



**Abb. 1.6** Komplementäre Basenpaarung. Wasserstoffbrückenbindungen sind als gestrichelte Linien dargestellt.

## 6 | 1 Molekulargenetik

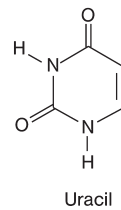
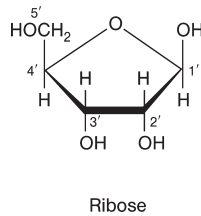


Abb. 1.7 Struktur von Ribose und Uracil.

unerlässlich bei der Expression genetischer Information, bei der Transkription von DNA in mRNA und bei der Translation in Protein.

Die Doppelhelix wird durch Wasserstoffbrückenbindungen zwischen den Basenpaaren stabilisiert. Diese Bindungen können durch Hitze oder einige Chemikalien gelöst werden. Dies führt dazu, dass sich die Doppelhelix in zwei Stränge auftrennt. Man sagt, das Molekül ist **denaturiert**. In Zellen können Enzyme die beiden Stränge der Doppelhelix trennen, um die DNA zu kopieren und genetische Information zu exprimieren.

### 1.1.5 RNA-Struktur

RNA ähnelt in ihrer Struktur der DNA. Es gibt jedoch einige wichtige Unterschiede. In der RNA ist 2'-Desoxyribose durch Ribose ersetzt und statt der Base Thymin wird die Base Uracil eingebaut. Uracil kann, wie Thymin, Basenpaarung mit Adenin eingehen (Abb. 1.7). Außerdem kommen RNA-Moleküle normalerweise als einzelne Polynukleotidstränge vor. Sie bilden keine Doppelhelix. Basenpaarung kann jedoch zwischen komplementären Bereichen desselben RNA-Strangs auftreten. So entstehen kurze doppelsträngige Regionen.

### 1.1.6 DNA-Replikation

DNA wird von Enzymen, den DNA-Polymerasen, kopiert. Diese Enzyme arbeiten an einzelsträngiger DNA. Sie synthetisieren einen neuen Strang, der komplementär zum Originalstrang ist. DNA-Synthese läuft immer in 5'→3'-Richtung ab. Die Replikation wird deshalb als **semikonservativ** bezeichnet (Abb. 1.8). Das bedeutet, dass jedes kopierte DNA-Molekül einen Strang enthält, der vom Ausgangsmolekül stammt, und einen neu synthetisierten Tochterstrang.

Der Mechanismus der DNA-Replikation ist in den meisten Organismen sehr ähnlich. Replikation muss sehr präzise sein, da selbst eine kleine Fehlerrate zum

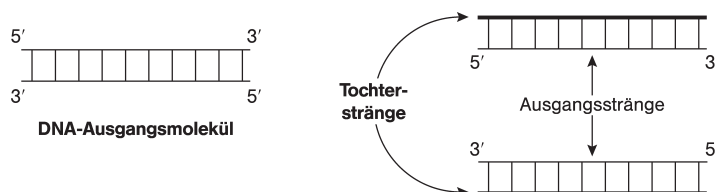


Abb. 1.8 Semikonservative Replikation.

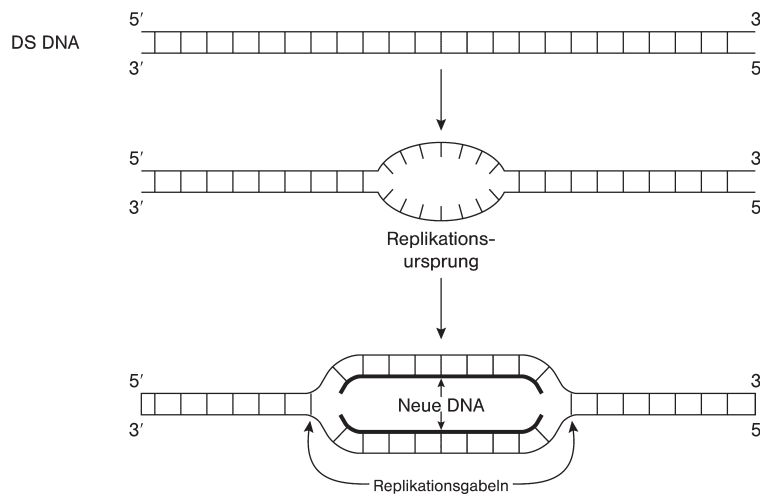


Abb. 1.9 Replikationsursprung und Replikationsgabeln.

Verlust wichtiger genetischer Information führen würde. Genauigkeit wird dadurch erreicht, dass die DNA-Polymerasen überprüfen können, ob in den neu synthetisierten Strang die richtigen Basen eingebaut wurden. Man bezeichnet das als **Korrekturlesen**. Vermutlich wird nur eine Base pro fünf Milliarden falsch eingebaut.

Während der DNA-Replikation wird die Doppelhelix der gesamten Zell-DNA nach und nach entwunden. So entstehen Bereiche einzelsträngiger DNA, die von DNA-Polymerasen kopiert werden können. Das Entwinden der Doppelhelix beginnt an einer bestimmten Stelle, dem **Replikationsursprung**. Von dort aus wird die Doppelhelix schrittweise, entlang des gesamten Moleküls, weiter entwunden. Üblicherweise geschieht dies in beide Richtungen. Ein Replikationsursprung enthält normalerweise A-T-reiche Sequenzen. A-T-Basenpaare besitzen schwächere Bindungen als G-C-Basenpaare. Den Bereich, in dem die Helix entwunden und neue DNA synthetisiert wird, nennt man **Replikationsgabel** (Abb. 1.9). An einer Replikationsgabel laufen mehrere verschiedene Ereignisse ab:

- **Trennung der Doppelhelix.** Dies geschieht mithilfe des Enzyms **Helicase**.
- **Synthese des Leitstrangs und des Folgestrangs.** DNA-Polymerasen synthetisieren DNA nur in 5'→3'-Richtung. Da die beiden Stränge der Doppelhelix in entgegengesetzter Richtung verlaufen (der eine Strang verläuft in 5'→3'-Richtung, der andere in 3'→5'-Richtung), sind etwas unterschiedliche Mechanismen nötig, um sie zu replizieren. Ein Strang, der **Leitstrang**, wird in derselben Richtung kopiert, wie die Helix entwunden wird. Er kann folglich kontinuierlich synthetisiert werden (Abb. 1.10a). Der andere Strang, der **Folgestrang**, wird in entgegengesetzter Richtung synthetisiert und muss diskontinuierlich kopiert werden. Der Folgestrang wird als eine Reihe von Segmenten, den **Okazaki-Fragmenten**, synthetisiert (Abb. 1.10b).

8 | 1 Molekulargenetik

- **Einsatz von Primern.** DNA-Polymerasen benötigen eine kurze doppelsträngige Region, um die DNA-Synthese zu starten. Diese Aufgabe wird von einer RNA-Polymerase, der **Primase**, erledigt. Sie kann die Synthese an einzelsträngiger DNA beginnen. Die Primase synthetisiert eine kurze RNA-Startsequenz (den Primer) an der DNA-Matrize. Auf diese Weise entsteht ein kurzer doppelsträngiger Bereich. Die DNA-Polymerase beginnt dann am RNA-Primer mit der DNA-Synthese. Am Folgestrang stoppt die Synthese wenn der nächste RNA-Primer erreicht ist. An diesem Punkt übernimmt eine andere DNA-Polymerase: Sie entfernt den RNA-Primer und ersetzt ihn durch DNA (Abb. 1.11).
- **Ligation.** Um die Synthese des Folgestrangs abzuschließen, müssen die Okazaki-Fragmente über Phosphodiesterbindungen miteinander verknüpft werden. Dies bewerkstelligt das Enzym **DNA-Ligase** (Abb. 1.11).

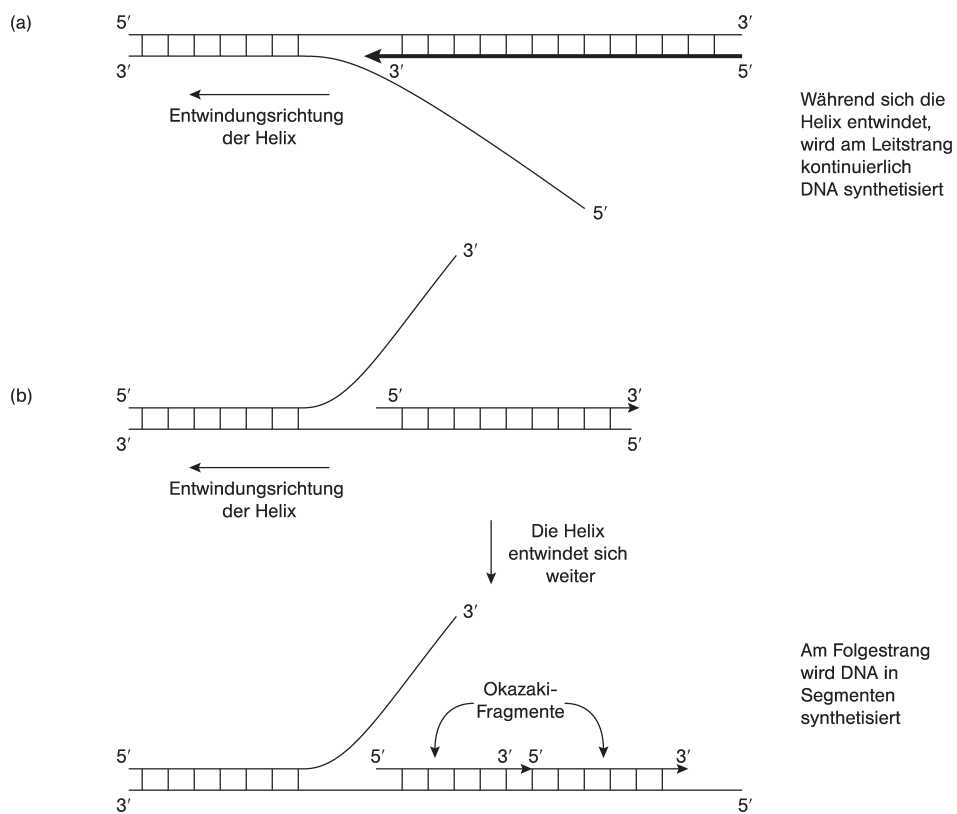


Abb. 1.10 (a) Replikation des Leitstrangs. (b) Replikation des Folgestrangs.



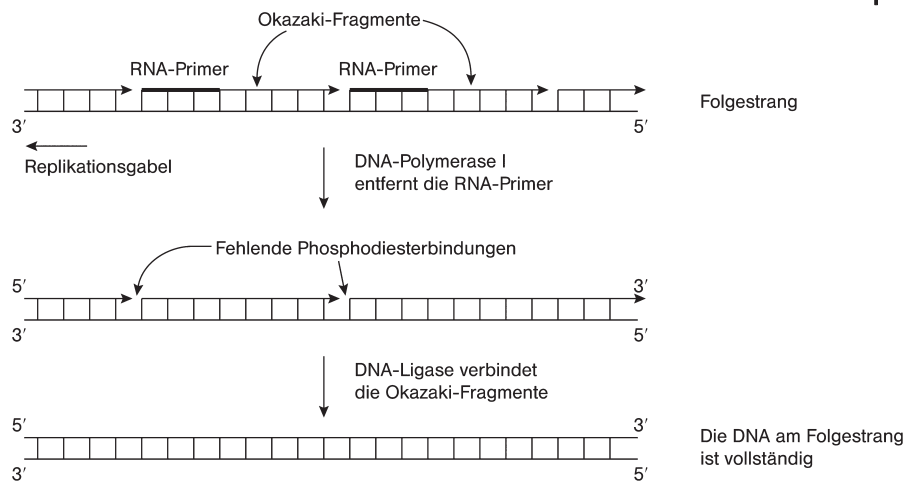


Abb. 1.11 Vervollständigung der Folgestrangsynthese.

## Noch einmal in Kürze

### Nukleotide

Die DNA ist ein Polymer mit Ketten aus Nukleotid-Monomeren. Jedes Nukleotid enthält einen Zucker, eine Base und eine Phosphatgruppe. Bei dem Zucker handelt es sich um 2'-Desoxyribose. Er besitzt fünf Kohlenstoffatome, die mit 1' (eins Strich), 2' etc. durchnummeriert sind. Es gibt vier verschiedene Basen: Adenin und Guanin besitzen zwei Kohlenstoff-Stickstoff-Ringe und gehören zu den Purinen. Thymin und Cytosin besitzen einen einzigen Ring und zählen zu den Pyrimidinen. Die Basen sind an das 1'-Kohlenstoffatom der Desoxyribose gebunden. Die Kombination aus einem Zucker und einer Base nennt man ein Nukleosid. Ein Nukleotid hat eine, zwei oder drei Phosphatgruppen, die an das 5'-Kohlenstoffatom des Zuckers gebunden sind. Nukleotide kommen in Form von einzelnen Molekülen vor oder als Polymer wie in der DNA oder RNA.

### DNA-Polynukleotide

Nukleotidtriphosphate der vier Basen sind zu DNA-Polynukleotidketten verknüpft. Zwei Phosphatgruppen gehen während der Polymerisierung verloren, und die Nukleotide werden über die verbleibenden Phosphatgruppen miteinander verbunden. Zwischen dem 5'-Phosphat des einen Nukleotids und der 3'-Hydroxylgruppe des nächsten Nukleotids bildet sich eine Phosphodiesterbindung. Das Polynukleotid besitzt eine freie 5'-OH-Gruppe am einen Ende (5'-Ende) und eine freie 3'-OH-Gruppe (3'-Ende) am anderen Ende. Die Reihenfolge der Basen codiert für die genetische Information. Sie wird in 5'→3'-Richtung gelesen. Polynukleotide sind extrem lang. Es sind  $4^n$  verschiedene Sequenzen möglich.



### **Die Doppelhelix**

DNA-Moleküle bestehen aus zwei Polynukleotidsträngen, die umeinander herum gewickelt sind und so eine Doppelhelix bilden. Der Zucker-Phosphat-Teil des Moleküls bildet das Rückgrat. Die Basen zeigen nach innen und sind aufeinander gestapelt. Die beiden Polynukleotidketten laufen in entgegengesetzte Richtung (antiparallel). Die Doppelhelix ist rechtsgängig und vollzieht pro 10 Basen eine Windung. Die Helix besitzt eine große Furche, die mit Proteinen in Wechselwirkung tritt, und eine kleine Furche. Es wurden verschiedene DNA-Strukturen identifiziert, darunter die linksgängige Z-DNA.

### **Komplementäre Basenpaarung**

Wasserstoffbrückenbindungen zwischen den Basen der beiden DNA-Stränge stabilisieren die Doppelhelix. Der vorhandene Abstand zwischen den Strängen beschränkt die Art und Weise, wie die Basen miteinander in Wechselwirkung treten können. Ein Purin tritt immer mit einem Pyrimidin in Wechselwirkung. Somit paart sich A immer mit T und G immer mit C. Dies bezeichnet man als komplementäre Basenpaarung. Die Einschränkung der Basenpaarung bedeutet, dass die Sequenz der Basen auf den beiden Strängen voneinander abhängig ist. Die Sequenz des einen Strangs bestimmt die des anderen und macht sie vorhersagbar. Somit kann die genetische Information während der DNA-Replikation und während der Genexpression bewahrt werden. Die Zerstörung der Wasserstoffbrückenbindungen zwischen den Basen durch Hitze, Chemikalien oder durch Enzyme führt dazu, dass sich die Stränge der Doppelhelix voneinander trennen.

### **RNA-Struktur**

In der RNA ist Thymin durch Uracil und 2-Desoxyribose durch Ribose ersetzt. RNA kommt normalerweise in Form eines einzelnen Polynukleotidstrangs vor. In kurzen Bereichen kann es jedoch zwischen den beiden komplementären Sequenzen zur Basenpaarung kommen.

### **DNA-Replikation**

Bei diesem Prozess kopiert eine Zelle ihre DNA, bevor sie sich teilt. Die DNA wird von der DNA-Polymerase in 5'→3'-Richtung kopiert. Einzelsträngige DNA dient dabei als Matrize. Die Replikation ist semikonservativ. Eine Helicase trennt die Doppelhelix. DNA wird am Leitstrang kontinuierlich synthetisiert und am Folgestrang nicht kontinuierlich, in Form von Abschnitten (den Okazaki-Fragmenten). Die DNA-Synthese durch die DNA-Polymerase beginnt an einem kurzen RNA-Primer. Dieser Primer wird von der Primase, einer RNA-Polymerase, synthetisiert. Die Primer-Sequenz wird später durch DNA ersetzt. Die DNA-Ligase verbindet die Okazaki-Fragmente über eine Phosphodiesterbindung.

### Tipp

#### Verwandte Themen:

- (Abschnitt 1.5) DNA-Mutationen
- (Abschnitt 1.4) Von der DNA zum Protein



## 1.2 Gene

### 1.2.1 Genstruktur

Die biologische Information, die ein Organismus für seine Reproduktion benötigt, ist in seiner DNA gespeichert. Die Information wird durch die Basensequenz der DNA codiert und ist in Form vieler Gene organisiert. Jedes dieser Gene enthält die Anleitungen zur Synthese eines Polypeptids. Physikalisch gesehen ist ein Gen ein eigenständiger DNA-Abschnitt mit einer Basensequenz, die für die Aminosäuresequenz eines Polypeptids codiert. Die Größe eines Gens variiert zwischen weniger als 100 Basenpaaren und mehreren Millionen Basenpaaren. In höheren Organismen liegen die Gene in einer Reihe extrem langer DNA-Moleküle, den **Chromosomen**, vor. Beim Menschen sind derzeit 20.563 proteincodierende Gene auf 23 Chromosomen identifiziert. Die Gene sind sehr zerstreut und durch Sequenzen, die scheinbar keine sinnvolle Information enthalten, voneinander getrennt. Diese Sequenzen nennt man **intergene DNA**. Intergene DNA ist sehr lang. Beim Menschen machen die Gensequenzen weniger als 30 % der gesamten DNA aus. Nur einer der beiden Stränge der DNA-Doppelhelix trägt die biologische Information. Man nennt diesen Strang den **Matrizenstrang**. Er wird zur Herstellung eines RNA-Moleküls mit komplementärer Sequenz verwendet, das wiederum die Synthese eines Polypeptids steuert. Der andere Strang wird als **Nichtmatrizenstrang** bezeichnet. Beide Stränge der Doppelhelix haben das Potenzial, als Matrizenstrang zu fungieren. Einzelne Gene können auf verschiedenen Strängen codiert werden. Man kann statt diesen beiden Bezeichnungen auch noch andere Begriffe zur Beschreibung der Doppelhelixstränge verwenden. Dazu gehören die Bezeichnungen „**sense/antisense**“ und „**codierend/nicht codierend**“. Die Begriffe „antisense“ und „nicht codierend“ sind Synonyme für den Matrizenstrang.

DNA-Moleküle besitzen eine enorme Kapazität für die Speicherung von Information. Ein DNA-Molekül mit einer Länge von  $n$  Basen hat  $4^n$  verschiedene Möglichkeiten, die vier Basen zu kombinieren. Sogar bei einem sehr kurzen DNA-Molekül ist die Anzahl der verschiedenen möglichen Sequenzen sehr groß. In der Praxis gibt es Einschränkungen für die Sequenzen mit sinnvoller Information. Die Kapazität für die Codierung von Information bleibt dennoch enorm.

### 1.2.2 Genfamilien

Die meisten Gene sind zufällig über die Chromosomen verteilt. Einige sind jedoch in Gruppen oder Clustern organisiert. Es gibt zwei Arten von Clustern: **Operons** (Abb. 1.12) und **Multi-Genfamilien** (Abb. 1.13).

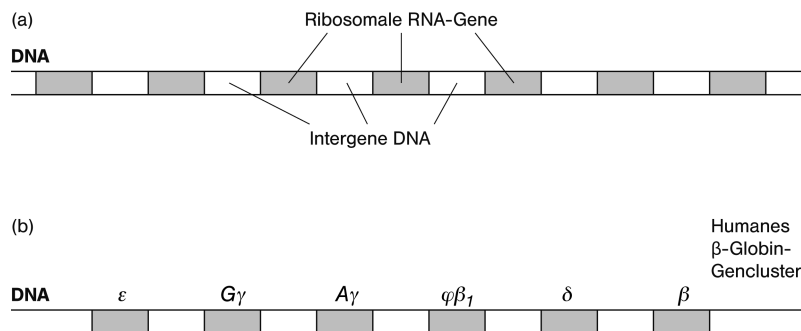
12 | 1 Molekulargenetik



**Abb. 1.12** Das *lac*-Operon. Drei Gene (*lac Z*, *Y* und *A*) sind zusammen angeordnet und werden gemeinsam reguliert.

Operons kommen in Bakterien vor. Sie enthalten Gene, die koordiniert reguliert werden und für Proteine mit eng verwandten Funktionen codieren. Ein Beispiel ist das *lac*-Operon in *E. coli*. Es enthält drei Gene, die für Enzyme codieren, die das Bakterium zum Abbau von Lactose benötigt. Wenn Lactose als Energiequelle vorhanden ist, werden alle vom *lac*-Operon codierten Enzyme zusammen benötigt. Die Clusterbildung von Genen innerhalb des Operons erlaubt das gleichzeitige An- oder Ausschalten der Gene. Der Organismus kann auf diese Weise seine Ressourcen effizient nutzen (Abb. 1.12).

Bei höheren Organismen gibt es keine Operons. Die Gene eines Clusters kommen in Multi-Genfamilien vor. Anders als bei Operons sind die Gene in einer Multi-Genfamilie identisch oder zumindest sehr ähnlich und werden nicht koordiniert reguliert. Der Grund für eine derartige Ansammlung von Genen besteht wahrscheinlich darin, dass viele Kopien dieses Gen nötig sind und dieses Gen daher im Laufe der Evolution vervielfältigt wurde. Einige Multi-Genfamilien existieren als getrennte Cluster auf unterschiedlichen Chromosomen. Diese Trennung geschah wahrscheinlich durch Umordnung der DNA während der Evolution. Dadurch lösten sich die Cluster auf. Multi-Genfamilien können einfach oder komplex sein. Bei **einfachen Multi-Genfamilien** (Abb. 1.13a) sind die Gene identisch. Ein Beispiel dafür ist das Gen für die 5S-rRNA. Beim Menschen gibt es 2000 in Clustern angeordnete Kopien dieses Gens. Das lässt erkennen, welch hohen Bedarf eine Zelle an diesem Genprodukt hat. **Komplexe Multi-Genfamilien** (Abb. 1.13b) enthalten Gene, die sich sehr ähneln, aber nicht identisch sind. Die Globin-Genfamilie ist ein Beispiel für eine solche komplexe Multi-Genfamilie. Die Gene codieren für eine Reihe von Polypeptiden ( $\alpha$ -,  $\beta$ -,  $\gamma$ -,  $\epsilon$ - und  $\zeta$ -Globine), die sich nur durch ein paar Aminosäuren voneinander unterscheiden. Die Globin-Polypeptide



**Abb. 1.13** Multi-Genfamilien. (a) Einfach. (b) Komplex.

bilden miteinander und zusammen mit einem Cofaktor, den man Häm nennt, Komplexe. So entstehen adulte und embryonale Formen des sauerstofftransportierenden Blutproteins Hämoglobin.

### 1.2.3 Genexpression

Die biologische Information eines DNA-Moleküls steckt in seiner Basensequenz. Genexpression ist der Prozess, in dem diese Information für die Zelle nutzbar gemacht wird. Die Nutzung der Information wird durch das ursprünglich von Crick aufgestellte **zentrale Dogma** beschrieben. Es besagt, dass die Information von DNA über RNA an Protein weitergeleitet wird (Abb. 1.14). Während der Genexpression kopieren DNA-Moleküle ihre Information, indem sie die Synthese eines RNA-Moleküls mit komplementärer Sequenz anordnen. Man bezeichnet diesen Prozess als **Transkription**. Die RNA steuert dann die Synthese eines Polypeptids. Die Aminosäuresequenz dieses Polypeptids wird durch die Reihenfolge der Basen der RNA bestimmt. Diesen Prozess nennt man **Translation**. Die Aminosäuresequenz des Proteins bestimmt seine dreidimensionale Struktur. Die Struktur wiederum diktiert die Funktion. Das zentrale Dogma besagt, dass die Übertragung von Information nur in eine Richtung erfolgen kann – von DNA über RNA zum Protein – und nicht umgekehrt. Eine Ausnahme dieser Regel findet man bei den Retroviren. Sie besitzen ein Enzym namens **Reverse Transkriptase**, das RNA zu DNA kopieren kann. Die Arbeitsweise von Zellen und folglich auch von lebenden Organismen hängt von der abgestimmten Aktivität vieler verschiedener Proteine ab. Die biologische Information, die in den Genen gespeichert ist, dient als Anleitung zur Synthese von Proteinen zur richtigen Zeit am richtigen Ort.

### 1.2.4 Genpromotoren

Die Expression der in den Genen steckenden biologischen Information ist sehr stark reguliert. Nicht alle Gene der DNA einer Zelle werden exprimiert. Verschiedene Gene sind in verschiedenen Zelltypen aktiv. Die Gesamtheit der Gene, die aktiv sind, bestimmt die Eigenschaften einer Zelle und ihre Funktion innerhalb des Organismus. Wenn man zum Beispiel die aktiven Gene einer Muskelzelle mit denen einer Blutzelle vergleicht, erkennt man, dass bei jedem Zelltyp viele andere Gene aktiv sind. Die Expression von Genen wird von einem Abschnitt der DNA-Sequenz stromaufwärts der codierenden Sequenz reguliert. Man bezeichnet diese Sequenz als **Promotor**. Konservierte DNA-Sequenzen im Promotor werden von der

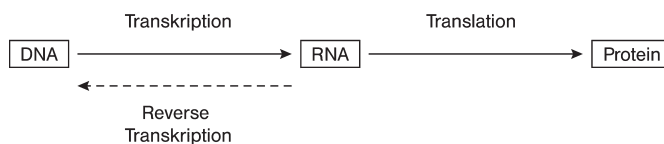


Abb. 1.14 Das zentrale Dogma (oben) und reverse Transkription (unten).

14 | 1 Molekulargenetik

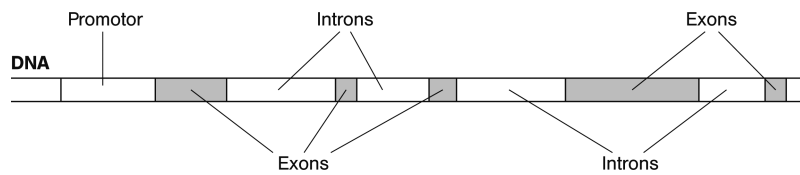


Abb. 1.15 Introns und Exons.

RNA-Polymerase und anderen assoziierten Proteinen – den **Transkriptionsfaktoren** – erkannt und gebunden. Zusammen bewerkstelligen sie die Synthese eines RNA-Transkripts des Gens. Die Expression eines Gens in einer Zelle wird von der Promotorsequenz sowie seiner Fähigkeit, RNA-Polymerase und Transkriptionsfaktoren zu binden, bestimmt.

### 1.2.5 Introns und Exons

Eine der überraschendsten Eigenschaften von Genen besteht darin, dass in höheren Organismen die codierende Information üblicherweise in eine Reihe von DNA-Sequenzsegmenten aufgeteilt ist. Diese Segmente bezeichnet man als **Exons**. Sie sind durch Sequenzen, die keine sinnvolle Information enthalten – die **Introns** –, voneinander getrennt (Abb. 1.15). Die Anzahl der Introns variiert sehr stark. Sie liegt zwischen null und über 50 in einigen Genen. Die Länge der Exons und Introns variiert auch. Die Introns sind jedoch üblicherweise viel länger und machen den größten Teil der Sequenz eines Gens aus. Bevor die biologische Information, die in einem Gen steckt, zur Synthese eines Proteins verwendet werden kann, müssen die Introns aus dem RNA-Molekül entfernt werden. Diesen Vorgang bezeichnet man als **Spleißen**. Nach dem Spleißen sind die Exons und damit die codierende Information durchgehend. Introns kommen nur bei höheren Organismen, nicht jedoch bei Bakterien vor.

### 1.2.6 Pseudogene

Es existieren Gene, die anderen Genen ähneln. Die Untersuchung ihrer Basensequenz ergab jedoch, dass sie Fehler enthalten, die es unmöglich machen, dass diese Gene für sinnvolle biologische Information codieren. Man nennt sie **Pseudogene** (Abschnitt 2.3). Sie stammen von Genen, die im Laufe der Evolution Fehler oder Mutationen in ihrer DNA-Sequenz erworben haben. Ihre biologische Information wurde so zerhackt, dass sie nicht mehr in der Lage sind, die Synthese eines Proteins zu steuern. Pseudogene gelten als evolutionäre Überbleibsel. Im Laufe der Evolution führte der ursprüngliche Basenaustausch erst zum Verlust der biologischen Information und anschließend zu schnelleren Veränderungen, sodass sich die Sequenz des Pseudogens schließlich beträchtlich von der des ursprünglichen Gens unterscheidet. Beispiele dafür sind einige Globin-Pseudogene im Globin-Gencluster.



## Noch einmal in Kürze

### Genstruktur

Ein Gen ist eine Informationseinheit und entspricht einem separaten DNA-Abschnitt, der für die Aminosäuresequenz eines Polypeptids codiert. Humanzellen enthalten etwa 21.000 Gene, die auf 23 Chromosomen angeordnet sind. Die Gene sind zerstreut und durch nicht codierende intergene DNA voneinander getrennt. Die Information ist im Matrizenstrang codiert, der die Synthese eines RNA-Moleküls steuert. Beide DNA-Stränge können als Matrizenstrang fungieren. DNA-Moleküle besitzen eine enorme Kapazität zur Speicherung genetischer Information.

### Genfamilien

Einige Gene sind in Form von Clustern angeordnet. Man bezeichnet diese Cluster als Operons und Multi-Genfamilien. Operons kommen in Bakterien vor und enthalten gemeinsam regulierte Gene mit ähnlicher Funktion. Multi-Genfamilien kommen bei höheren Organismen vor und bestehen aus ähnlichen oder identischen Genen, die nicht koordiniert reguliert werden. Einfache Multi-Genfamilien enthalten identische Gene, deren Produkt in großen Mengen benötigt wird. Komplexe Multi-Genfamilien bestehen aus Genen, die sehr ähnlich sind und für Proteine mit einer ähnlichen Funktion codieren.

### Genexpression

Die biologische Information, die in den Genen verschlüsselt ist, wird durch die Genexpression verfügbar gemacht. Bei diesem Vorgang wird eine RNA-Kopie eines Gens synthetisiert. Diese RNA steuert dann die Synthese eines Proteins. Das zentrale Dogma sagt aus, dass Information immer von DNA über RNA in Protein übertragen wird. Die Arbeitsweise von Zellen hängt von der koordinierten Aktivität vieler Proteine ab. Genexpression stellt sicher, dass Proteine am richtigen Ort zur richtigen Zeit synthetisiert werden.

### Genpromotoren

Die Genexpression ist sehr stark reguliert. Nicht alle Gene einer Zelle sind aktiv. Verschiedene Zellarten exprimieren verschiedene Gene. Die Expression eines Gens wird durch ein DNA-Segment stromaufwärts der codierenden Sequenz reguliert. Man nennt diese Sequenz Promotor. Der Promotor bindet RNA-Polymerase und dazugehörige Transkriptionsfaktorproteine und initiiert die Synthese eines RNA-Moleküls.

### Introns und Exons

Die codierende Sequenz eines Gens ist in eine Reihe von Segmenten aufgespalten. Diese Segmente nennt man Exons. Sie werden durch nicht codierende Sequenzen, die Introns, voneinander getrennt. Introns machen in der Regel den größten Teil der Gensequenz aus. Anzahl und Größe der Introns variiert von Gen zu Gen. Durch einen Prozess, den man Spleißen nennt, werden die

Introns vor der Proteinsynthese aus den RNA-Transkripten entfernt. Bakterielle Gene besitzen normalerweise keine Introns.

### Pseudogene

Von einigen Genen existieren Kopien, die Sequenzfehler enthalten. Diese Fehler haben sie im Laufe der Evolution erworben. Sie verhindern, dass das entsprechende Gen in Protein übersetzt wird. Man nennt diese Gene Pseudogene. Sie stellen evolutionäre Überbleibsel der ursprünglichen Gene dar. Ein Beispiel dafür sind die Globin-Pseudogene.



### Tipp

#### Verwandte Themen:

- (Abschnitt 1.7) Regulation der Genexpression
- (Abschnitt 1.8) Epigenetik und Chromatinmodifikation
- (Abschnitt 2.2) Prokaryotengenome
- (Abschnitt 2.3) Eukaryotengenome

## 1.3 Der genetische Code

### 1.3.1 Genexpression

Die Information, die ein Organismus für seine Reproduktion benötigt, ist in seiner DNA gespeichert. Sie ist in der Basensequenz codiert und in Form von Genen organisiert. Der Ausdruck „Genexpression“ beschreibt den Prozess, in dem Zellen die Information decodieren und sie verwenden, um die Proteine, die eine Zelle für ihr Funktionieren benötigt, zu synthetisieren. Während der Genexpression wird die Information von DNA zu RNA kopiert. Dies geschieht durch die Synthese eines RNA-Moleküls, dessen Basensequenz komplementär zur Sequenz der DNA-Matrize ist. Die RNA steuert dann die Synthese eines Proteins, dessen Aminosäuresequenz durch die Basensequenz der RNA festgelegt ist. Für jedes Gen gilt: Die DNA-Sequenz ist kollinear zur Aminosäuresequenz des Polypeptids, für das sie codiert. Die 5' → 3'-Basensequenz des codierenden Strangs bestimmt die Aminosäuresequenz des codierten Polypeptids vom Aminoende (N-Terminus) zum Carboxylende (C-Terminus).

### 1.3.2 Der genetische Code

Der genetische Code beschreibt, wie Basensequenzen während der Proteinsynthese in Aminosäuresequenzen übersetzt werden. Die DNA-Sequenz eines Gens ist in eine Reihe von Einheiten mit jeweils drei Basen unterteilt. Jede dieser Einheiten nennt man ein Codon; es steht für eine bestimmte Aminosäure. Mit den vier Basen in DNA und RNA können insgesamt  $4^3 = 64$  verschiedene Codons gebildet werden. Diese Codons codieren die 20 Aminosäuren, die in Proteinen vorkommen (Tabelle 1.1). Da die Anzahl der Codons größer ist als die Anzahl der



**Tabelle 1.1** Der genetische Code

Erste Position (5´-Ende)	Zweite Position				Dritte Position (3´-Ende)
	U	C	A	G	
U	Phe UUU	Ser UCU	Tyr UAU	Cys UGU	U
	Phe UUC	Ser UCC	Tyr UAC	Cys UGC	C
	Leu UUA	Ser UCA	Stopp UAA	Stopp UGA	A
	Leu UUG	Ser UCG	Stopp UAG	Trp UGG	G
C	Leu CUU	Pro CCU	His CAU	Arg CGU	U
	Leu CUC	Pro CCC	His CAC	Arg CGC	C
	Leu CUA	Pro CCA	Gln CAA	Arg CGA	A
	Leu CUG	Pro CCG	Gln CAG	Arg CGG	G
A	Ile AUU	Thr AUC	Asn AAU	Ser AGU	U
	Ile AUC	Thr ACC	Asn AAC	Ser AGC	C
	Ile AUA	Thr ACA	Lys AAA	Arg AGA	A
	Met AUG	Thr ACG	Lys AAG	Arg AGG	G
G	Val GUU	Ala GCU	Asp GAU	Gly GGU	U
	Val GUC	Ala GCC	Asp GAC	Gly GGC	C
	Val GUA	Ala GCA	Glu GAA	Gly GGA	A
	Val GUG	Ala GCG	Glu GAG	Gly GGG	G

verschiedenen Aminosäuren, werden alle Aminosäuren außer Methionin und Tryptophan von mehr als einem Codon codiert. Man spricht von der **Degeneration** oder der **Redundanz** des genetischen Codes. Codons, die für die gleiche Aminosäure codieren, bezeichnet man als Synonyme. Sie sind sich in der Regel ähnlich. ACU, ACC, ACA und ACG spezifizieren zum Beispiel alle die Aminosäure Threonin. Abweichungen zwischen den Synonymen treten in der Regel an der dritten Position des Codons auf. Man bezeichnet sie auch als **Wobble-Position**. Die Degeneration des genetischen Codes minimiert die Auswirkungen von Mutationen. Es ist somit unwahrscheinlicher, dass Basensequenzveränderungen zu einem Austausch der codierten Aminosäure führen. Mögliche schädliche Auswirkungen auf die Funktion des Proteins werden vermieden. Von den 64 möglichen Codons codieren 61 für Aminosäuren. Die übrigen drei UAG, UGA und UAA codieren nicht für Aminosäuren, sondern wirken stattdessen als Stoppsignale für die Proteinsynthese. Man bezeichnet sie als **Terminations-Codons** oder **Stopp-Codons**. Das Codon für Methionin, AUG, ist das Signal zum Start der Proteinsynthese. Man nennt es **Start-Codon**. Alle Polypeptide beginnen also mit Methionin. Manchmal wird diese Aminosäure später entfernt.

## 18 | 1 Molekulargenetik

Leseraster 1. 5' -AUG ACU AAG AGA UCC GG -3'  
Met Thr Lys Arg Ser

Leseraster 2. 5' -A UGA CUA AGA GAU CCG G -3'  
Stop Leu Arg Asp Pro

Leseraster 3. 5' -AU GAC UAA GAG AUC CCG -3'  
Asp Stop Glu Ile Arg

**Abb. 1.16** Jede DNA-Sequenz kann in drei verschiedenen Leserastern gelesen werden. Es hängt davon ab, welche Base als Startpunkt des Codons ausgewählt wird.

### 1.3.3 Leseraster

Die Aufgabe des Start-Codons besteht nicht nur darin, den Startpunkt der Proteinsynthese zu identifizieren, es bestimmt auch das Leseraster der RNA-Sequenz. Abhängig davon, welche Base als Start eines Codons ausgewählt wird, gibt es drei mögliche Zusammenstellungen von Codons, die aus einer Sequenz herausgelesen werden können. Das bedeutet, dass es in der Praxis während der Proteinsynthese normalerweise nur ein Leseraster gibt, das sinnvolle Information enthält. Die anderen beiden Leseraster enthalten üblicherweise Stopp-Codons, die verhindern, dass sie zur Steuerung der Proteinsynthese eingesetzt werden (Abb. 1.16). Eine Zusammenstellung von Codons, die kontinuierlich fortläuft und am Beginn durch ein Start-Codon und am Ende durch ein Stopp-Codon begrenzt ist, heißt **offenes Leseraster** (ORF, *open reading frame*). Diese Charakteristik verwendet man, um in Genom-Sequenzierungsprojekten proteincodierende DNA-Sequenzen zu identifizieren.

### 1.3.4 Universalität des Codes

Zunächst nahm man an, der genetische Code würde universell gelten, d. h. alle Organismen erkennen in den verschiedenen Codons dieselben Aminosäuren. Man hat inzwischen jedoch gezeigt, dass es einige, wenn auch seltene Abweichungen im Code gibt. In den tierischen Mitochondrien existieren zum Beispiel kleine DNA-Genome mit etwa 20 Genen, in denen Abweichungen vom genetischen Code auftreten. Veränderungen sind meistens mit Start- oder Stopp-Codons verbunden. UGA zum Beispiel, das normalerweise ein Stopp-Codon darstellt, codiert für Tryptophan, während AGA und AGG, die normalerweise für Arginin codieren, Stopp-Codons sind. Und AUA, das normalerweise für Isoleucin codiert, spezifiziert Methionin. Man vermutet, dass diese Veränderungen nur deshalb funktionsfähig sind, weil es sich bei einem Mitochondrium um ein geschlossenes System handelt. Inzwischen wurden auch außerhalb des mitochondrialen Genoms einige Beispiele für vom Standard-Code abweichende Codons in einzelligen Organismen gefunden.

UAA und UAG, die normalerweise Stopp-Codons darstellen, codieren in einigen Protozoen für Glutaminsäure.

### Noch einmal in Kürze

#### Genexpression

Die genetische Information ist in der Basensequenz der DNA-Moleküle in Form von Genen codiert. Der Ausdruck „Genexpression“ beschreibt, wie Zellen die Information decodieren, um Proteine, die für die Funktion der Zelle benötigt werden, zu synthetisieren. Zur Expression eines Gens gehört die Synthese eines komplementären RNA-Moleküls, dessen Sequenz die Aminosäuresequenz eines Proteins bestimmt. Die DNA-Sequenz des Gens ist kollinear zur Aminosäuresequenz des Polypeptids.

#### Der Genetische Code

Aminosäuren werden von 64 Basentriplets, die man als Codons bezeichnet, codiert. Die Codons codieren die 20 Aminosäuren. Für die meisten Aminosäuren gibt es mehr als ein Codon. Man nennt dies die Degeneration des genetischen Codes. Sie trägt dazu bei, die Auswirkungen von Mutationen zu minimieren. Codons, die für die gleiche Aminosäure codieren, nennt man Synonyme. Sie unterscheiden sich in ihrer dritten Base, der sogenannten „Wobble“-Position. AUG ist ein Start-Codon. Es codiert für Methionin. Es gibt drei Stopp-Codons: UAG, UGA und UAA.

#### Leseraster

Aus einer Sequenz können, abhängig davon, bei welcher Base mit dem Lesen begonnen wird, drei mögliche Codon-Zusammenstellungen gelesen werden. Jede dieser Codon-Zusammenstellungen bezeichnet man als Leseraster. Das Start-Codon bestimmt das Leseraster einer proteincodierenden Sequenz. Andere Leseraster enthalten eher Stopp-Codons und werden nicht für die Proteinsynthese verwendet. Ein offenes Leseraster ist eine Abfolge von Codons, die durch Start- und Stopp-Codons begrenzt ist.

#### Universalität des Codes

Der genetische Code gilt universell bei allen Organismen, die die gleichen Codons für die jeweiligen Aminosäuren verwenden. Es gibt jedoch einige Ausnahmen von dieser Regel in Mitochondrien-Genomen und bei einigen einzelligen Organismen.

#### Tipp

##### Verwandte Themen:

- (Abschnitt 1.4) Von der DNA zum Protein



## 1.4 Von der DNA zum Protein

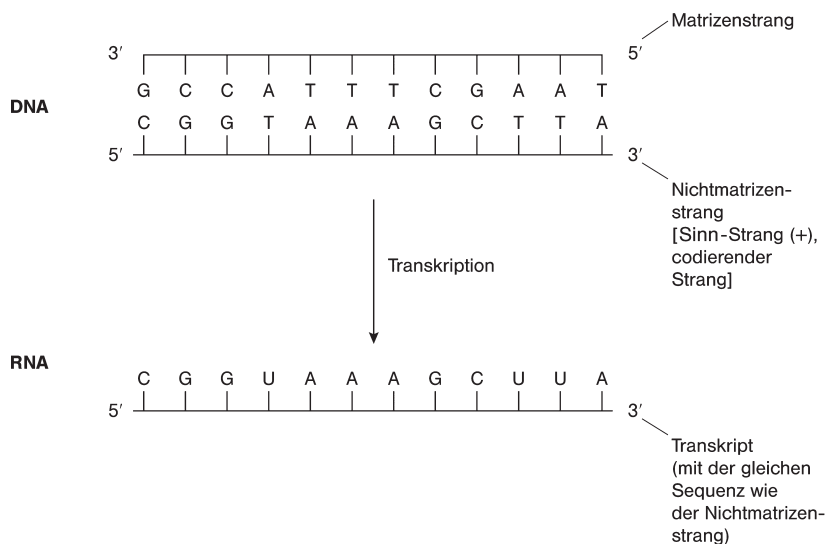
### 1.4.1 Einleitung

Die Expression von Genen, ausgehend von der in der Nukleotidsequenz codierten Information bis hin zur Proteinsynthese, erfolgt in drei Schritten. Bei der **Transkription** wird die Basensequenz der DNA zu RNA kopiert. Die Information bleibt dabei im Nucleinsäure-Code erhalten. Das RNA-Molekül wird anschließend prozessiert und weist schließlich an den Ribosomen die Synthese der Proteine an. An den Ribosomen wird die in der Nucleinsäuresequenz gespeicherte genetische Information in die entsprechende Aminosäuresequenz übersetzt (**Translation**).

### 1.4.2 RNA-Transkription

Mithilfe der DNA als Matrize wird von Enzymen, den **RNA-Polymerasen**, RNA synthetisiert. Die beiden Stränge einer Doppelhelix heißen Matrizenstrang und Nichtmatrizenstrang. RNA wird vom **Matrizenstrang** kopiert. Das so synthetisierte RNA-Molekül hat die gleiche Sequenz wie der **Nichtmatrizenstrang** (Abb. 1.17). Die Gensequenzen beziehen sich üblicherweise auf den Nichtmatrizenstrang. Andere Namen für den Nichtmatrizenstrang sind **Sinn(+)-strang** oder **codierender Strang**. Das synthetisierte RNA-Molekül ist das **Transkript**.

Während der Transkription wird durch Polymerisieren von Ribonucleosidtriphosphat-Einheiten (ATP; UTP, GTP, CTP) RNA synthetisiert. Das 3'-OH des einen Ribonucleotids reagiert mit dem 5'-Phosphat des anderen und bildet eine Phosphodiesterbindung. Die Reihenfolge, in der die Ribonucleotide an die wach-



**Abb. 1.17** Bei der Transkription des DNA-Matrizenstrangs entsteht ein RNA-Transkript, das die gleiche Sequenz besitzt wie der Nichtmatrizenstrang.

sende RNA-Kette angefügt werden, wird durch die Abfolge der Basen in der Matrizen-DNA bestimmt. Während die RNA-Polymerase RNA in 5'→3'-Richtung synthetisiert, werden neue Ribonukleotide am freien 3'-Ende an die wachsende Kette angefügt. Dieser Prozess ähnelt der DNA-Synthese sehr (Abschnitt 1.1).

Die Transkription **startet** gezielt am Anfang eines Gens. Signale für den Beginn der Transkription befinden sich in der **Promotorsequenz**. Diese liegt direkt stromaufwärts von der zu transkribierenden Gensequenz. Der Promotor enthält spezifische DNA-Sequenzen, an denen sich die Doppelhelix teilweise entwindet und einen **offenen Promotorkomplex** bildet. An diesen Komplex lagert sich dann die RNA-Polymerase an.

Die Phase, in der die Transkription fortgesetzt wird, bezeichnet man als **Elongationsphase**. Während der Transkription wird zu jedem Zeitpunkt immer nur ein kleiner Bereich der Doppelhelix entwunden. Der entwundene Bereich enthält die neu synthetisierte RNA in Basenpaarung mit dem Matrizen-DNA-Strang. Er ist zwischen 12 und 17 Basen lang. Das endgültige Transkript enthält nicht nur den codierenden Bereich der DNA, sondern auch die am 5'- und 3'-Ende angrenzenden Regionen. Bei Eukaryoten ist das erste Transkript komplexer. Es enthält sowohl codierende **Exons** als auch nicht codierende **Introns** (Abschnitt 1.2). Spezielle Sequenzen in der DNA signalisieren die Termination der Transkription.

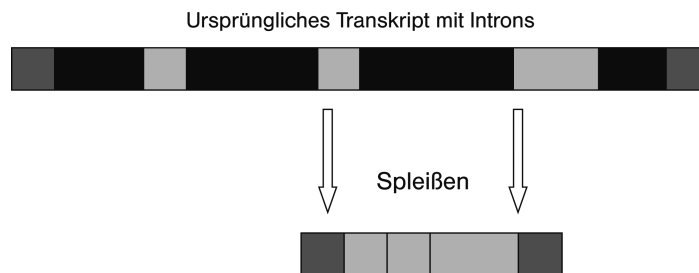
Bei Eukaryoten transkribieren drei verschiedene RNA-Polymerasen verschiedene Gengruppen. Die **RNA-Polymerase II** transkribiert Gene, die für Proteine codieren, und regulatorische mikroRNA-Gene. Die meisten Gene, die für die Nukleinsäurestruktur des Ribosoms, die **ribosomale RNA**, codieren, werden von der **RNA-Polymerase III** transkribiert. **RNA-Polymerase I** transkribiert schließlich die Gene, die für die an der Translation beteiligten tRNAs codieren.

### 1.4.3 RNA-Prozessierung

Bei Eukaryoten ist das ursprüngliche Transkript (prä-mRNA), das von proteincodierenden Genen gebildet wird, groß und besteht größtenteils aus transkribierten Sequenzen nicht codierender Introns. Diese Bereiche werden in einem Prozess, den man als **Spleißen** bezeichnet, entfernt. In Prokaryoten wird mRNA nicht prozessiert. Die Translation beginnt sogar schon bevor die Transkription abgeschlossen ist. Prokaryotische Gene enthalten normalerweise keine Introns. Folglich entfällt das Spleißen.

Das Spleißen findet im Zellkern statt. Introns vom Typ 1 enthalten ein aktives Zentrum in der RNA und spleißen sich selbst heraus. Bei Introns vom Typ 2 werden spezifische Nukleotidsequenzen am 5'- und am 3'-Ende eines jeden Introns von **kleinen nukleären Ribonukleoproteinen (snRNPs, small nuclear ribonucleoproteins)** erkannt. Bei snRNPs handelt es sich um Komplexe aus Protein und spezifischen RNA-Molekülen. Die Kombination aus prä-mRNA und den snRNPs nennt man **Spleißosom**. Es ist dafür verantwortlich, dass sich die prä-mRNA in die richtige Konformation zum Spleißen faltet, wenn die Intron-RNA-Sequenzen herausgeschnitten werden und die benachbarten Exon-Sequenzen zusammengefügt werden (Abb. 1.18). Die snRNPs können in einer Zelle oder in verschiedenen

22 | 1 Molekulargenetik



**Abb. 1.18** Spleißen von prä-mRNA. Das ursprüngliche Transkript enthält Introns (schwarz), Exons (hellgrau) und nicht translatierte 3'- und 5'-Regionen (dunkelgrau). Durch Spleißen des mRNA-Moleküls werden die Introns entfernt. Das Molekül ist nun bereit für das Anfügen einer Kappe (engl. *capping*) am 5'-Ende und eines Poly(A)-Schwanzes am 3'-Ende. Danach wird es zur Translation ins Cytoplasma transportiert.

Zellen variieren und machen die Regulation und Modifikation des Spleißens möglich. In verschiedenen Geweben kann man bei derselben mRNA unterschiedliche Spleißmuster beobachten. Auf diese Weise können von nur einem einzigen „Gen“ viele Varianten eines Proteins gebildet werden. In Tumoren ändern sich diese Muster oftmals (Abschnitt 5.3).

Abgesehen vom Spleißen werden prä-mRNAs auch an ihren 3'- und 5'-Termini verändert. Das 5'-Ende kann durch Aufsetzen einer Kappe (*capping*) modifiziert werden. Dabei wird das modifizierte Nukleotid **7-Methylguanosin** angefügt. An das 3'-Ende von fast allen mRNAs wird ein langer Schwanz aus Adeninnukleotiden angehängt. Das reife mRNA-Molekül kann nun an den Ribosomen im Cytoplasma in Protein übersetzt werden.

#### 1.4.4 Translation

Die Translation erfolgt an den Ribosomen. Transfer-RNA-Moleküle spielen bei diesem Prozess eine wichtige Rolle. Sie liefern an das Ribosom die Aminosäuren in der Reihenfolge, die von der mRNA-Sequenz vorgegeben wird. Das stellt sicher, dass die Aminosäuren in der richtigen Reihenfolge miteinander verbunden werden (Abb. 1.19). Zellen enthalten normalerweise zwischen 31 und 40 verschiedene tRNA-Arten. Jede davon bindet spezifisch eine der 20 Aminosäuren. Folglich gibt es für jede Aminosäure nicht nur eine tRNA. Transfer-RNAs, die die gleiche Aminosäure binden, nennt man **Isoakzeptoren**. Bevor die Translation beginnt, werden die Aminosäuren kovalent mit ihren tRNAs verbunden. Die tRNAs erkennen dann die Codons in der mRNA, die diese Aminosäure codieren. Das Anheften einer Aminosäure an ihre tRNA nennt man **Aminoacylierung** oder **Beladung**. Die Aminosäure wird mithilfe des Enzyms **Aminoacyl-tRNA-Synthetase** kovalent an das Ende des Akzeptorarms der tRNA gebunden. Für jede Aminosäure existiert ein eigenes Enzym. Jedes Enzym kann alle Isoakzeptor-tRNAs für die entsprechende Aminosäure beladen.

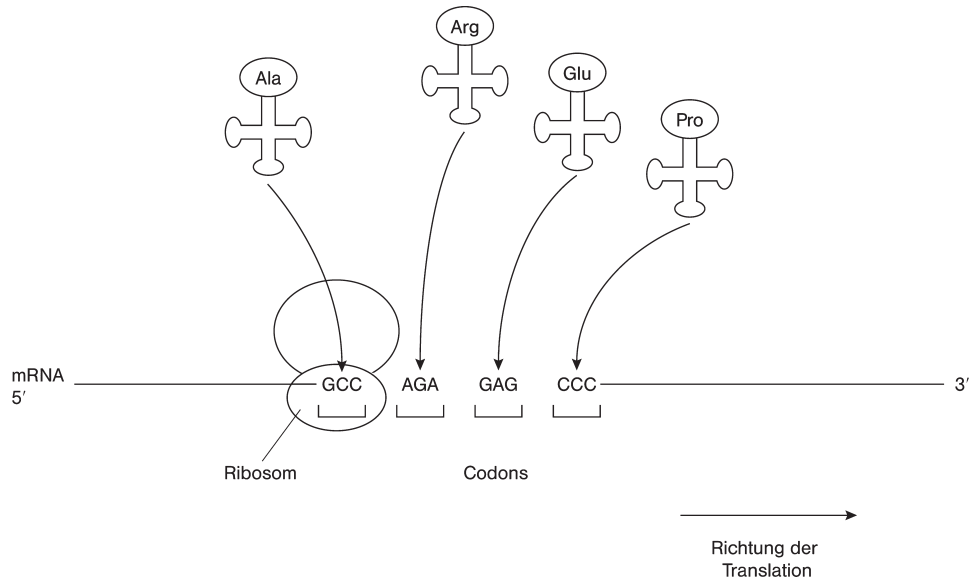


Abb. 1.19 Rolle der tRNAs bei der Translation.

Wenn die tRNA mit der richtigen Aminosäure beladen wurde, erkennt sie das entsprechende Codon in der mRNA und platziert die Aminosäure an der richtigen Position, so wie die Sequenz der mRNA es vorgibt. Auf diese Weise wird sichergestellt, dass die von der mRNA codierte Aminosäuresequenz korrekt übersetzt wird. Die Codonerkenkung erfolgt über die Anticodonschleife der tRNA. Die Anticodonschleife enthält drei Nucleotide, die als **Anticodon** bezeichnet werden. Dieses bindet über komplementäre Basenpaarung an das Codon. Wenn zwei Aminosäuren eng nebeneinander ausgerichtet sind, knüpft das Enzym **Peptidyltransferase** eine Peptidbindung zwischen ihnen. So wächst die entstehende Polypeptidkette immer weiter. Die vier Basen, die in der DNA vorkommen, können zu 64 verschiedenen Codons kombiniert werden. Drei Codons wirken als Signal für die Beendigung der Translation. Die übrigen 61 codieren für die 20 Aminosäuren, die in Proteinen vorkommen (Abschnitt 1.3).

Ribosomen kommen als getrennte große und kleine Untereinheiten vor. Im ersten Schritt der Translation bindet die kleine Ribosomenuntereinheit an die mRNA. Die Translation beginnt üblicherweise an der Sequenz **AUG** (Bakterien verwenden manchmal GUG oder UUG). Diese Sequenz codiert für Methionin und ist das Start-Codon.

Die Translation schreitet an der mRNA entlang und entfernt sich dabei vom Startpunkt. Die Aminosäuren werden über Peptidbindungen miteinander verknüpft. Das an erster Position stehende Methionin wird vom wachsenden Polypeptid entfernt. Eine mRNA kann gleichzeitig an mehreren Ribosomen translatiert werden. Die so entstehenden Strukturen nennt man Polysome. Die Translation stoppt am Terminationscodon (Abschnitt 1.3).

Das Polypeptid wird von Enzymen, die manche Aminosäuren modifizieren können, weiter prozessiert. Dazu gehört das Anheften kleiner chemischer Gruppen wie zum Beispiel Methylierung, Hydroxylierung oder Formylierung. Oft werden Zuckerketten (Polysaccharide) angefügt. Diesen Prozess nennt man Glykosylierung. Im Allgemeinen werden alle Proteine, die aus der Zelle herausragen, glykosyliert. Dies geschieht im Golgi-Apparat. Ein spezielles Beispiel dafür sind die roten Blutkörperchen des Menschen. Blutgruppe A hat an einer Polysaccharidkette, die man die H-Substanz nennt, *N*-Acetylgalactosamin gebunden. Blutgruppe B hat stattdessen Galactose. Zellen der Blutgruppe AB besitzen beides an verschiedenen Ketten, und Blutgruppe O hat nur die nicht modifizierte H-Substanz.



### Noch einmal in Kürze

#### Einleitung

Die Expression von Genen erfolgt in drei Schritten: Transkription, RNA-Prozessierung und Translation. Bei der Translation wird die in Nukleinsäuren codierte genetische Information am Ribosom in Aminosäuresequenzen umgewandelt.

#### RNA-Transkription

RNA-Polymerasen schreiben DNA-Sequenzen des Matrizenstrangs zu RNA-Kopien um. Das RNA-Produkt wird als Transkript bezeichnet. Der Prozess ähnelt der DNA-Replikation durch DNA-Polymerasen. Die Transkription beginnt an einer Promotorsequenz in der DNA und läuft so lange weiter ab, bis eine Terminationssequenz erreicht ist. Bei Eukaryoten werden spezielle Gengruppen von verschiedenen Polymerasen transkribiert.

#### RNA-Prozessierung

Das erste prä-mRNA-Transkript ist groß und enthält Introns. Diese werden anschließend in Spleißosomen aus dem Transkript entfernt. Durch unterschiedliches Spleißen können aus demselben Gen somit verschiedene Formen eines Proteins hergestellt werden. Prokaryotische mRNA muss nicht prozessiert werden. Modifikationen an den 5'- und 3'-Bereichen der mRNA erfolgen ebenfalls während der RNA-Prozessierung.

#### Translation

Die Translation findet an den Ribosomen statt und beginnt an einem AUG-Start-Codon. Die Nukleotidsequenz wird von tRNA-Molekülen, die jeweils mit der entsprechenden Aminosäure beladen sind, in Triplets gelesen. Das Anticodon des tRNA-Moleküls bindet an das komplementäre Codon in der mRNA. Die Peptidyltransferase bildet Peptidbindungen zwischen den ausgerichteten Aminosäuren, und es entstehen wachsende Polypeptide. Nach der Translation können die Proteine chemisch modifiziert werden.



### Tipp

#### Verwandte Themen:

- (Abschnitt 1.2) Gene
- (Abschnitt 1.7) Regulation der Genexpression
- (Abschnitt 5.3) Gene und Krebs



## 1.5 DNA-Mutation

### 1.5.1 Mutationen

Die DNA-Sequenz eines Gens bestimmt die Aminosäuresequenz des codierten Proteins. Es ist sehr wichtig, dass die DNA-Sequenz bewahrt wird, da Veränderungen der Aminosäuresequenz die Funktionsfähigkeit des Proteins beeinflussen können. Dies wiederum kann Auswirkungen auf die Gesundheit des Organismus haben. Veränderungen der DNA-Sequenz können die Folge der Einwirkung chemischer oder physikalischer Agenzien auf die DNA sein. Sie können auch durch seltene Fehler bei der DNA-Replikation entstehen. Derartige Veränderungen nennt man **Mutationen**. Wenn sie einmal eingeführt sind, werden diese DNA-Sequenzänderungen durch DNA-Replikation dauerhaft gemacht. Sie werden nach der Zellteilung an die Tochterzellen weitergegeben.

Zwei wichtige Begriffe, mit denen beschrieben wird, ob ein Organismus eine Mutation trägt oder nicht, sind der **Genotyp** und der **Phänotyp**. Der Genotyp dient zur Beschreibung der Mutation und des Gens, in dem sie stattfindet. Der Phänotyp beschreibt die Auswirkung der Mutation auf den Organismus. Ein Organismus, der den üblichen Phänotyp für diese Spezies aufweist, heißt **Wildtyp**. Einen Organismus, dessen üblicher Phänotyp infolge einer Mutation verändert wurde, nennt man **Mutante**. Es gibt zwei Arten von Mutationen: **Punktmutationen**, bei denen eine Base an irgendeiner Position in einem Gen verändert wird, und **Großmutationen**, bei denen längere Bereiche der DNA-Sequenz verändert werden. Der Ort der Mutation innerhalb eines Gens spielt eine wichtige Rolle. Nur Mutationen innerhalb der codierenden Region können Auswirkungen auf das Protein haben. Mutationen in nicht codierenden oder intergenen Regionen haben normalerweise keine Konsequenzen.

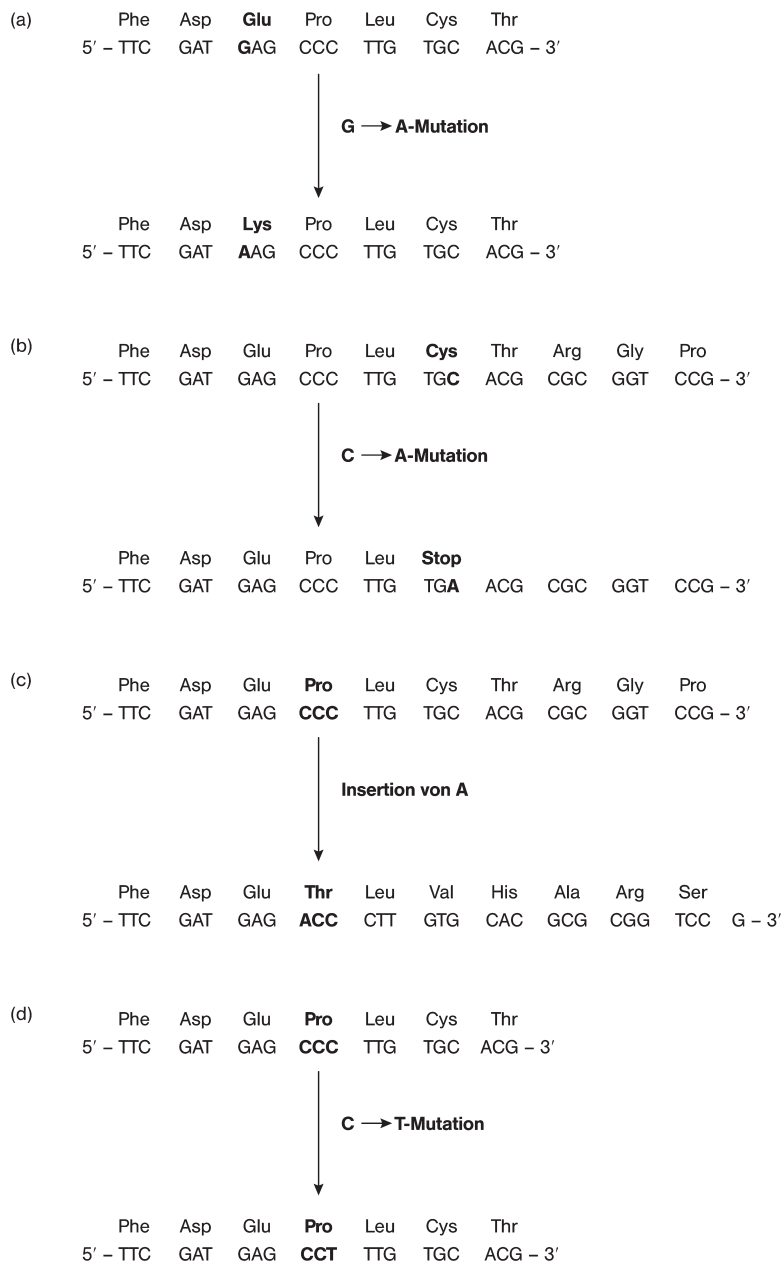
### 1.5.2 Punktmutationen

Punktmutationen fallen in mehrere Kategorien. Jede davon hat unterschiedliche Konsequenzen für das Protein, das von dem Gen codiert wird.

#### 1.5.2.1 Sinnverändernde Mutationen

Bei diesen Punktmutationen wird eine einzelne Base verändert. Dadurch verändern sich das entsprechende Codon und folglich auch die codierte Aminosäure (Abb. 1.20a). Derartige Mutationen treten üblicherweise in einer der beiden ersten Basen eines Codons auf. Durch die Redundanz (Degeneration) des genetischen

26 | 1 Molekulargenetik



**Abb. 1.20** (a) Sinnverändernde Punktmutation. (b) Sinnentstellende Punktmutation. (c) Raster-schub-Punktmutation. (d) Stille Punktmutation.

Codes führen Mutationen der dritten Base seltener zur Veränderung der Aminosäure. Die Auswirkungen sinnverändernder Mutationen auf den Organismus sind unterschiedlich. Die meisten Proteine tolerieren eine Veränderung ihrer Aminosäuresequenz. Veränderungen der Aminosäuresequenz in Bereichen des Proteins, die wichtig für die Struktur oder Funktion sind, haben eher gesundheitsbeeinträchtigende Folgen und führen zu einem Mutanten-Phänotyp.

#### 1.5.2.2 Sinnentstellende Mutationen

Diese Punktmutationen führen dazu, dass ein Codon, das für eine Aminosäure codiert, in ein Stopp-Codon umgewandelt wird (Abb. 1.20b). Die Translation der mRNA bricht infolgedessen vorzeitig ab. So entsteht ein verkürztes Protein, dem Teile der C-terminalen Region fehlen. Sinnentstellende Mutationen haben üblicherweise ernste Auswirkungen auf die Aktivität des codierten Proteins und verursachen oft einen Mutanten-Phänotyp.

#### 1.5.2.3 Rasterschubmutationen

Rasterschubmutationen entstehen, wenn zusätzliche Basen eingefügt oder wenn Basen aus der DNA-Sequenz eines Gens herausgeschnitten werden. Wenn die Anzahl der eingefügten oder entfernten Basen nicht ein Vielfaches von drei ist, verändert sich das Leseraster, und das Ribosom liest stromabwärts der Mutation eine andere Codon-Zusammenstellung. Dadurch verändert sich die Aminosäuresequenz des codierten Proteins beträchtlich (Abb. 1.20c). Rasterschubmutationen haben in der Regel schwerwiegende Auswirkungen auf das codierte Protein und führen zu Mutanten-Phänotypen.

#### 1.5.2.4 Stille Mutationen

Mutationen können an der dritten Base eines Codons auftreten. Wegen des degenerierten genetischen Codes wird die Aminosäure nicht verändert (Abb. 1.20d). Stille Mutationen haben keine Auswirkung auf das codierte Protein und führen zu keinem Mutanten-Phänotyp. Sie neigen dazu, sich in der DNA der Organismen anzuhäufen und führen zum **Polymorphismus**. Sie tragen zur Variabilität der DNA-Sequenz von Individuen einer Art bei.

Punktmutationen, bei denen ein Purin durch ein Pyrimidin ersetzt wird oder umgekehrt, bezeichnet man als **Transversionen**. Wenn durch eine Mutation eine Purinbase gegen eine andere Purinbase bzw. eine Pyrimidinbase gegen eine andere Pyrimidinbase ausgetauscht werden, spricht man von **Transitionen**.

#### 1.5.3 Großmutationen

Großmutationen verursachen erhebliche Veränderungen der DNA. Oft sind lange Sequenzabschnitte betroffen. Es kann eine Reihe von Großmutationen auftreten.

### 1.5.3.1 Deletionen

Hierbei geht ein Teil der DNA-Sequenz verloren (Abb. 1.21a). Die Größe dieses Anteils kann stark variieren. Deletionen können nur eine einzige Base oder auch ein viel größeres Fragment betreffen. In einigen Fällen kann eine komplette Gensequenz verloren gehen.

### 1.5.3.2 Insertionen

In diesem Fall wird die Mutation durch das Einfügen zusätzlicher Basen – üblicherweise von einer anderen Stelle eines Chromosoms – verursacht (Abb. 1.21b). Wie bei Deletionen können nur ein oder zwei Basen oder auch ein viel längeres DNA-Stück eingefügt werden.

Der Ausdruck „Indel“ bezeichnet kleine Deletionen und Insertionen als Gruppe.

### 1.5.3.3 Umgruppierungen

Bei dieser Mutation werden Abschnitte der DNA-Sequenz innerhalb oder außerhalb eines Gens miteinander vertauscht (Abb. 1.21c). Ein einfaches Beispiel ist die Inversionsmutation, bei der ein Teil der DNA-Sequenz ausgeschnitten und anschließend an derselben Stelle, aber in entgegengesetzter Richtung wieder eingefügt wird.

Bei Großmutationen sind stets längere Gensequenzen von der Veränderung betroffen. Aus diesem Grund haben sie gravierende Auswirkungen auf das codierte Protein und sind häufig mit einem Mutanten-Phänotyp verbunden.

## 1.5.4 Mutationen und Krankheiten

Bei Menschen und anderen höheren Organismen spielen DNA-Mutationen bei der Entwicklung von Krankheiten eine wichtige Rolle. Insbesondere Erbkrankheiten wie die Bluterkrankheit und Mukoviszidose werden durch Mutationen verursacht. Aber auch an der Entstehung von Krebs sind Mutationen beteiligt.

Erbkrankheiten werden von Mutationen verursacht, die vererbt und von den Eltern an ihre Nachkommen weitergegeben werden. Mutationen können zufällig in irgendeiner Zelle auftreten. Da Zellen jedoch kontinuierlich ersetzt werden, hat eine Mutation in einer einzelnen Zelle normalerweise keine Auswirkung auf den Organismus. Eine Mutation in einer Keimzelle (Spermium oder Eizelle) kann jedoch nach der Befruchtung weitergegeben werden, sodass sie in jeder Zelle des Nachkommens vorhanden ist. Der Nachkomme wird dann zum Träger der Mutation und kann wiederum die Mutation an die nächsten Generationen weitergeben. Bei Erbkrankheiten ist normalerweise ein einzelnes Gen von der Mutation betroffen. Die mutierten Gene unterscheiden sich weitgehend voneinander. Das trägt zur unterschiedlichen Natur von Erbkrankheiten bei.

Mutationen, die an der Entwicklung von Krebs beteiligt sind, kommen in den normalen Körperzellen (somatische Zellen) vor. Krebs entwickelt sich oft in Verbindung mit Mutationen in Genen, die die Zellteilung regulieren. Derartige

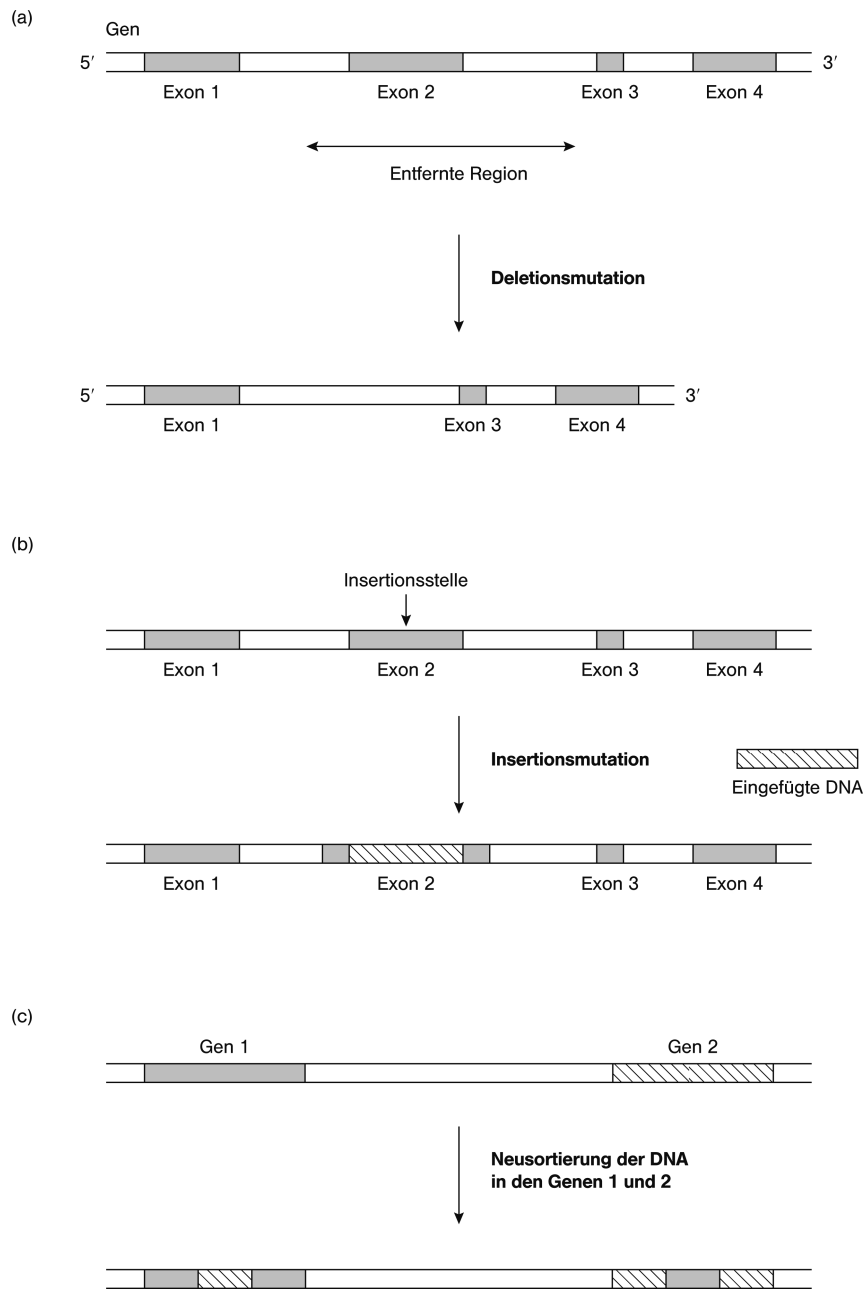


Abb. 1.21 (a) Deletionsmutation. (b) Insertionsmutation. (c) Umgruppierungsmutation.

30 | 1 Molekulargenetik

Mutationen verleihen den Zellen oft die Fähigkeit, sich auf eine unkontrollierte Art und Weise zu teilen, und führen so zur Entwicklung eines Tumors.

### 1.5.5 Mutationen auf der Organismusebene

Man kann Mutationen nicht nur hinsichtlich der Veränderung der DNA-Sequenz definieren, sondern auch über die Beschreibung des Phänotyps des mutierten Organismus. Dieser Ansatz ist sowohl auf prokaryotische als auch auf eukaryotische Zellen anwendbar. Meistens wird er jedoch im Zusammenhang mit bakteriellen Mutanten verwendet. Es wurden etliche Mutanten-Phänotypen definiert.

#### 1.5.5.1 Auxotrophe Mutanten

Diesen Mutanten fehlt ein Genprodukt, das an der Synthese eines essenziellen Metaboliten wie zum Beispiel einer Aminosäure beteiligt ist. Sie können nur isoliert werden, wenn man sie in einem Medium kultiviert, dem dieser Metabolit zugesetzt wurde. Ein Tryptophan-auxotropher *E. coli*-Stamm zum Beispiel kann Tryptophan nicht selbst synthetisieren und muss es aus dem Wachstumsmedium aufnehmen. Das Gegenteil von auxotroph ist prototroph. Ein prototropher Organismus hat keine speziellen Nahrungsansprüche.

#### 1.5.5.2 Temperatursensitive Mutanten

Diese Mutanten sind durch die Bedingungen, die zur Aufzucht der Zellen eingesetzt wurden, eingeschränkt. Sie können nur überleben, wenn sie bei einer bestimmten Temperatur kultiviert werden. Wenn die Temperatur erhöht wird, sind die Zellen nicht mehr in der Lage zu wachsen und sterben ab. Die Mutation einer temperatursensitiven Mutante betrifft üblicherweise eine Aminosäure, die für die Aufrechterhaltung einer Proteinstruktur wichtig ist. Wenn die Temperatur erhöht wird, denaturiert das mutierte Protein und verliert seine Aktivität. So entsteht der Mutanten-Phänotyp.

#### 1.5.5.3 Antibiotikaresistente Mutanten

Antibiotika töten Wildtyp-Bakterien ab. Sie haben aber keine Wirkung auf resistente Mutanten. Antibiotikaresistenz kann auf verschiedenen Wegen entstehen. Oft ist jedoch ein Gen von der Mutation betroffen, das für ein Protein codiert, auf das ein Antibiotikum abzielt.

#### 1.5.5.4 Regulatorische Mutanten

In diesem Fall verlieren die Mutanten die Fähigkeit, die Expression eines Gens oder eines Operons zu regulieren. Es gibt zum Beispiel regulatorische Mutanten von *E. coli*, die Gene exprimieren, die für den Lactose-Metabolismus benötigt werden, selbst wenn keine Lactose im Medium vorhanden ist. Man spricht hier von konstitutiven Mutanten. Sie entstehen in der Regel durch Mutation des Lac-Repressorproteins.



## Noch einmal in Kürze

### Mutationen

Unter Mutationen versteht man Veränderungen der üblichen DNA-Sequenz eines Organismus. Dies kann durch Einwirkung chemischer oder physikalischer Agenzien verursacht werden oder die Folge von DNA-Replikationsfehlern sein. Mutationen bleiben durch die Zellteilung erhalten. Die Beschaffenheit einer Mutation und ihre Auswirkung auf einen Organismus werden durch den Genotyp und den Phänotyp beschrieben. Organismen können Wildtyp-Phänotypen oder Mutanten-Phänotypen besitzen. Es gibt Punktmutationen und Großmutationen. Nur wenn Mutationen in den codierenden Regionen der Gene vorkommen, ist die Wahrscheinlichkeit groß, dass sie Auswirkungen auf die Proteinfunktion haben.

### Punktmutationen

Unter Punktmutationen versteht man Veränderungen einer einzelnen Base. Sinnverändernde Mutationen verändern eine einzige codierte Aminosäure. Durch sinnentstellende Mutationen entstehen Stopp-Codons, und es werden verkürzte Polypeptide produziert. Durch Rasterschubmutationen wird eine Base eingefügt oder entfernt. So entsteht ein verändertes Leseraster. Stille Mutationen treten an der dritten Base eines Codons auf und verändern die codierte Aminosäure nicht. Stille Mutationen häufen sich in der DNA als Polymorphismen an. Punktmutationen haben unterschiedliche Konsequenzen für die biologische Aktivität des codierten Proteins.

### Großmutationen

Bei Großmutationen werden längere DNA-Sequenzen verändert. Bei Deletionsmutationen werden Basen entfernt. Die Länge der entfernten Sequenz kann zwischen einer einzigen Base und einer gesamten Gensequenz variieren. Bei Insertionsmutationen werden zusätzliche Basen eingefügt. Kleine Insertionen und Deletionen bezeichnet man kollektiv als „Indels“. Bei Neusortierungen werden Segmente der DNA-Sequenz innerhalb oder außerhalb eines Gens miteinander vertauscht. Großmutationen führen in der Regel zu einem vollständigen Verlust der biologischen Aktivität des codierten Proteins.

### Mutationen und Krankheiten

Die Ursache für genetische Erkrankungen und für Krebs sind DNA-Mutationen. Genetische Erkrankungen werden durch vererbte Mutationen hervorgerufen. Normalerweise ist ein einzelnes Gen betroffen. Die Mutation erfolgt in einer Keimzelle, und es entsteht ein Nachkomme, der die Mutation nun in allen Zellen trägt und sie an Folgegenerationen weitergeben kann. Mutationen, die mit der Entstehung von Krebs zu tun haben, treten in somatischen Zellen auf – oft in Genen, die die Zellteilung kontrollieren.

### Mutationen auf der Organismusebene

Mutationen können über den Phänotyp, den sie hervorrufen, definiert werden. Man bezieht sich dabei üblicherweise auf bakterielle Mutanten. Auxotrophe Mutanten sind nicht in der Lage, essenzielle Metaboliten zu synthetisieren. Temperatursensitive Mutanten können nicht wachsen, wenn die Temperatur erhöht wird. Antibiotikaresistente Mutanten überleben in Anwesenheit des Antibiotikums. Regulatorische Mutanten verlieren ihre Fähigkeit, die Expression von Genen oder von Operons zu regulieren.



#### Tipp

##### Verwandte Themen:

- (Abschnitt 1.3) Der genetische Code
- (Abschnitt 5.1) Erbkrankheiten
- (Abschnitt 5.3) Gene und Krebs

## 1.6 Mutagene und DNA-Reparatur

### 1.6.1 Mutagene

Mutationen können durch Fehler bei der DNA-Replikation entstehen. Das ist normalerweise sehr selten. Bei Bakterien kommt dies etwa alle  $10^{10}$  Basen einmal vor, bei höheren Organismen wahrscheinlich häufiger. Man spricht von der **spontanen Mutationsrate**. Die Mutationsrate steigt, wenn die Zellen chemischen oder physikalischen Agenzien ausgesetzt werden, die direkt mit der DNA reagieren und die Struktur der einzelnen Nukleotide verändern. Man bezeichnet diese Agenzien als **Mutagene**. Veränderte Nukleotide führen zu einer geänderten oder falschen Basenpaarung, sodass bei der DNA-Replikation eine falsche Base gegenüber der modifizierten Base eingesetzt wird. Dadurch wird die DNA-Sequenz verändert. Nachfolgende Replikationsrunden machen die Veränderung dauerhaft. Die DNA-Sequenz ist mutiert (Abb. 1.22). Einige Mutagene wirken auf eine andere Weise. Sie verursachen schwerwiegende physikalische Deformationen in der DNA, die die DNA-Replikation oder die Transkription blockieren.

Viele verschiedene natürliche oder synthetische, organische oder anorganische Chemikalien können mit DNA reagieren, ihre Struktur verändern und so Mutationen hervorrufen. Die meisten chemischen Mutagene sind **karzinogen** und verursachen Krebs. Darüber hinaus kann eine Reihe von physikalischen Agenzien Mutationen hervorrufen. Dazu zählen ionisierende Strahlung in Form von Röntgen- und Gammastrahlung, nichtionisierende Strahlung, insbesondere ultraviolettes Licht und auch Hitze.

### 1.6.2 Chemische Mutagene

Viele unterschiedliche Chemikalien können DNA-Mutationen verursachen. Einige chemische Mutagene sind **Basenanaloga**. Ihre Struktur ähnelt der von normalen,





**Abb. 1.22** Mutationsmechanismen. (a) Mutationen entstehen aufgrund von Fehlern, die während der DNA-Replikation auftreten.

(b) Mutationen entstehen nach strukturellen Veränderungen in Nukleotiden, die eine veränderte Basenpaarung zur Folge haben.

34 | 1 Molekulargenetik

in der DNA vorkommenden Basen, und sie können von DNA-Polymerasen während der Replikation in die DNA eingebaut werden. Sie rufen Mutationen hervor, da sie zu einer veränderten Basenpaarung führen. Ein Beispiel dafür ist **5-Bromuracil (5BU)**, ein Basenanalogen, das sich von Thymin ableitet. 5BU geht normalerweise mit Adenin eine Basenpaarung ein. Es kann jedoch eine geringfügige Strukturveränderung erfahren, die man als Tautomerie bezeichnet. Dies führt dazu, dass 5BU statt mit Adenin mit Guanin Basenpaarung eingeht. Nach der DNA-Replikation ist das ursprüngliche TA-Basenpaar auf einem der beiden Tochterstränge durch ein GC ersetzt. Es hat eine Punktmutation stattgefunden (Abb. 1.23).

Eine andere Gruppe von Mutagenen sind die **interkalierenden Agenzien**. Dabei handelt es sich um flache Moleküle, die die DNA-Replikation stören, indem sie sich zwischen benachbarte Basenpaare der Doppelhelix schieben. Ein Beispiel dafür ist **Ethidiumbromid**, ein Molekül, das aus vier Ringen besteht und eine ähnliche Abmessung wie ein Purin-Pyrimidin-Basenpaar hat. Ethidiumbromid schiebt sich in die Doppelhelix (interkaliert) und führt dazu, dass sich die angrenzenden Basenpaare etwas auseinander bewegen (Abb. 1.24). Es ist noch nicht genau bekannt, wie dadurch die Replikation zum Erliegen kommt. Bekannt ist aber die Gesamtwirkung: Der Einschub von Ethidiumbromid führt zur Insertion eines einzelnen Nukleotids an der Interkalationsstelle und verursacht eine Rasterschubmutation in dem entsprechenden Gen.

Viele Mutagene wirken, indem sie Basen chemisch modifizieren. Einige führen zur Addition von Alkyl- oder Arylgruppen. Beispiele hierfür sind **Methylmethan-**

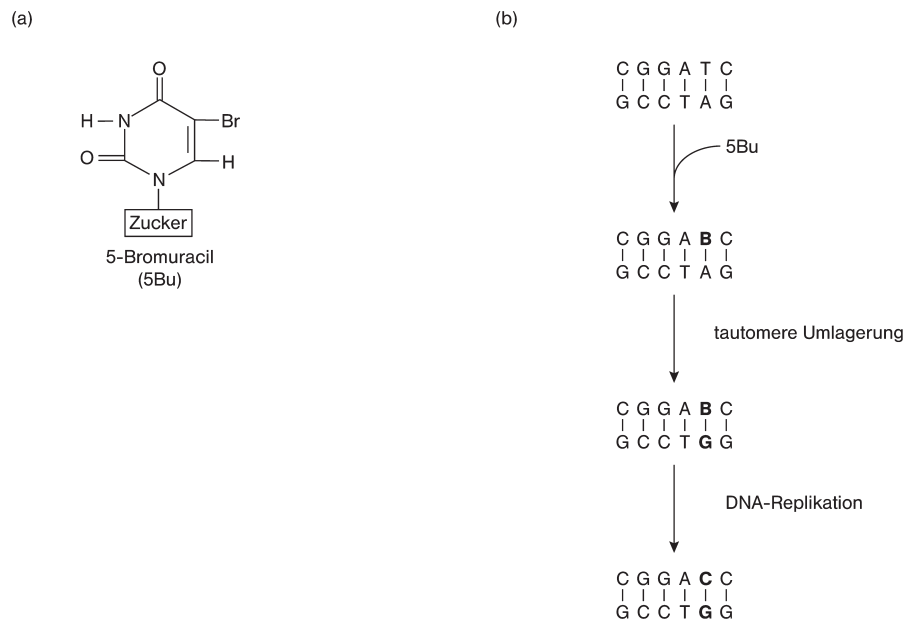


Abb. 1.23 (a) Struktur von 5-Bromuracil. (b) DNA-Mutation durch 5-Bromuracil.

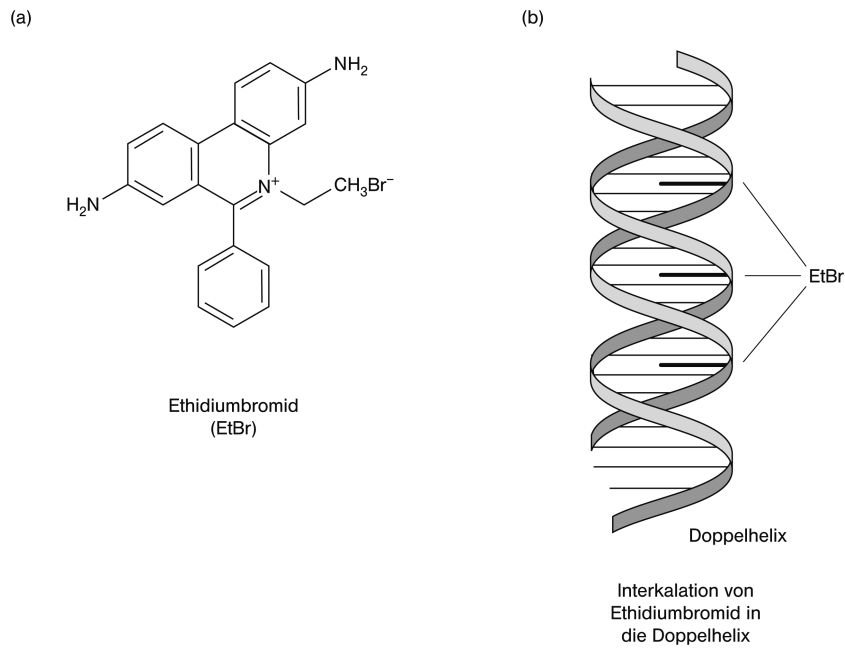


Abb. 1.24 (a) Struktur von Ethidiumbromid. (b) Mutagene Wirkung von Ethidiumbromid.

**sulfonat** und **Ethylnitrosoharnstoff**. Diese Substanzen übertragen Methylgruppen an verschiedenen Positionen der Basen (Abb. 1.25). Eine andere Art der Basenmodifizierung ist die **Desaminierung**. **Salpetrige Säure** zum Beispiel desaminiert Cytosin, und es entsteht Uracil (Abb. 1.26). Uracil geht eine Basenpaarung mit Adenin ein. Nach der Replikation ist das ursprüngliche G–C-Basenpaar durch A–T ersetzt. Die Desaminierung von Adenin zu einem Guaninanalogen namens Hypoxanthin führt dazu, dass ein A–T- durch ein G–C- Basenpaar ersetzt wird.

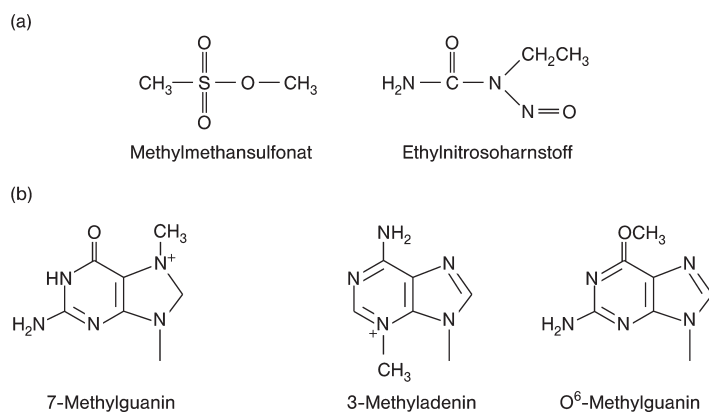


Abb. 1.25 Beispiele für (a) alkylierende Agenzien und (b) alkylierte Basen.

36 | 1 Molekulargenetik

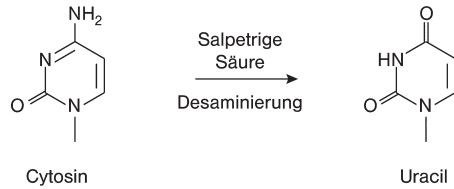


Abb. 1.26 Desaminierung von Cytosin zu Uracil mit salpetriger Säure.

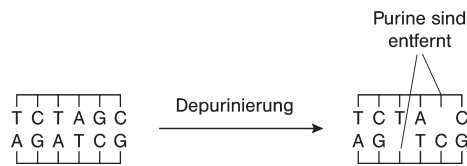


Abb. 1.27 Depurinierung von DNA.

Durch chemische Reaktion der DNA mit normalen chemischen Verbindungen, die üblicherweise in den Zellen vorkommen, können Mutationen auch spontan auftreten. Cytosin kann zum Beispiel spontan zu Uracil desaminieren. Cytosin, das manchmal als 5-Methylcytosin vorliegt, kann außerdem zu Thymin desaminieren. Eine sehr häufige, spontan auftretende Modifikation ist die **Depurinierung**. Dabei bricht die Verbindung zwischen dem Desoxyribosezucker und seiner Purinbase (Abb. 1.27). Basen können auch durch reaktive Sauerstoffspezies (ROS, *reactive oxygen species*) geschädigt werden. Dies passiert in allen aeroben Zellen. Zu den ROS gehören Superoxid, Wasserstoffperoxid und Hydroxylradikale.

### 1.6.3 Physikalische Mutagene

Energiereiche ionisierende Strahlung wie Röntgen- oder Gammastrahlen verursachen große Schäden an DNA-Molekülen. Sie führen zu Strangbrüchen und zerstören Zucker und Basen. Nicht ionisierende Strahlung in Form von ultraviolettem (UV) Licht wird von den Basen absorbiert und kann strukturelle Veränderungen auslösen. UV-Licht kann insbesondere die Entstehung von **Cyclobutylidimeren** zwischen benachbarten Pyrimidinen, vor allem zwischen Thyminmolekülen, hervor-

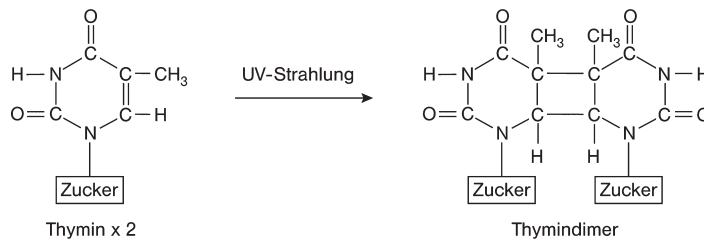


Abb. 1.28 Entstehung von Thymindimeren durch UV-Strahlung.

rufen (Abb. 1.28). Dimerisierung führt dazu, dass sich die Basen dichter stapeln. Dies wiederum kann nach der DNA-Replikation zu Deletionsmutationen führen. Hitze ist auch ein bedeutendes Umweltmutagen. Hitze kann im DNA-Polynukleotid Depurinierung hervorrufen. Es entstehen sog. AP-Stellen (*apurinic sites*), an denen nach der DNA-Replikation Punktmutationen oder Deletionsmutationen auftreten können.

#### 1.6.4 DNA-Reparatur

Es ist sehr wichtig, dass die DNA-Sequenz eines Gens trotz der Anwesenheit vieler Agenzien, die die DNA mutieren können, konserviert bleibt. Indem sie beschädigte DNA reparieren, haben die Organismen Methoden entwickelt, Mutationen zu verhindern. Die Reparaturmechanismen sind kompliziert. Es gibt im Wesentlichen drei Haupttypen.

##### 1.6.4.1 Exzisionsreparatur

Hierbei handelt es sich um ein komplexes System, das wahrscheinlich die üblichste Form der DNA-Reparatur darstellt. Es repariert viele verschiedene Schadensarten, einschließlich der Pyrimidindimere. Zuerst erkennt eines von mehreren Enzymen beschädigte Nukleotide und markiert sie für die Reparatur. Die Markierung kann ein Bruch in einem der beiden Stränge der Doppelhelix neben dem geschädigten Bereich sein. Eine geschädigte Base kann aber auch entfernt werden, sodass eine Lücke zurückbleibt. In der nächsten Phase entfernt eine Nuklease das markierte Nukleotid sowie einige benachbarte Nukleotide. Eine DNA-Polymerase (DNA-Polymerase I in *E. coli*) synthetisiert dann neue DNA, um die fehlenden Basen zu ersetzen. DNA-Ligase verbindet dann die neue DNA mit dem bereits existierenden Molekül und stellt so die ursprüngliche Struktur der DNA wieder her (Abb. 1.29).

##### 1.6.4.2 Direkte Reparatur

Dies ist eine viel seltenere Form der DNA-Reparatur, bei der strukturelle Veränderungen der Nukleotide rückgängig gemacht werden. Ein wichtiges Beispiel für direkte Reparatur ist die **Photoreaktivierung**, bei der Pyrimidindimere, die durch UV-Strahlung entstanden sind, repariert werden. Enzyme, die als DNA-Photolyasen bezeichnet werden, werden durch sichtbares Licht induziert und reparieren Pyrimidindimere: Sie spalten die Bindungen, die die Dimerisierung verursachen. DNA-Photolyasen kommen in Bakterien, mikrobiellen Eukaryoten und in Pflanzen vor.

##### 1.6.4.3 Fehlpaarungsreparatur

Dieses System erkennt fehlgepaarte Nukleotide und korrigiert so Fehler, die während der DNA-Replikation auftreten. Mehrere Enzyme markieren die Fehlpaarung oder reparieren sie direkt. Es ist wichtig, dass das System bestimmen kann, welche der fehlgepaarten Basen die richtige ist. Das System unterscheidet,

38 | 1 Molekulargenetik

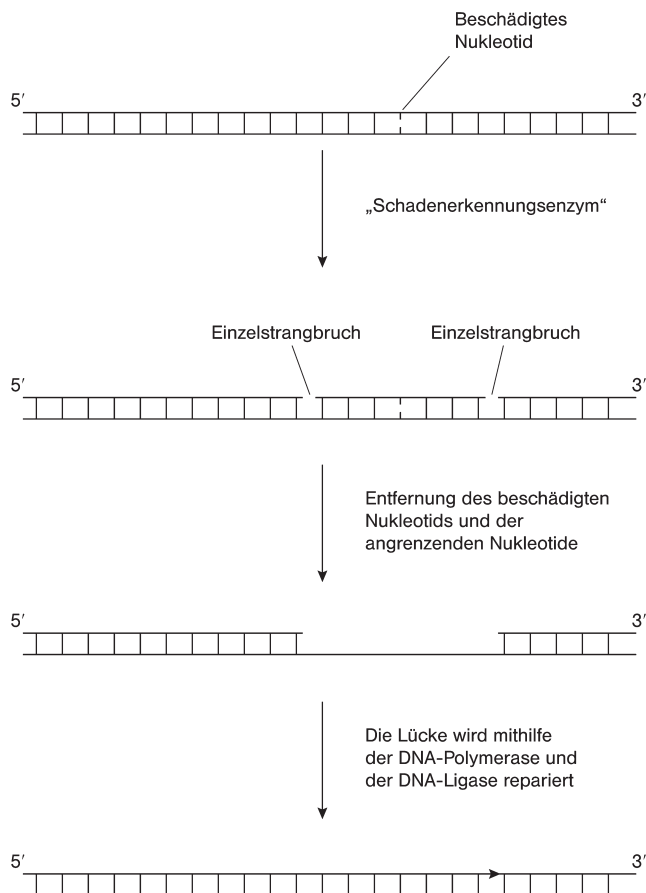


Abb. 1.29 DNA-Exzisionsreparatur.

welches der DNA-Ausgangsstrang mit der korrekten Sequenz ist und welches der Tochterstrang mit der mutierten Sequenz ist. In *E. coli* ist der Ausgangsstrang einfach zu erkennen, da seine Adeninbasen mit Methylgruppen markiert sind.

### 1.6.5 Genotoxizität

Hierbei handelt es sich um einen vor kurzem entwickelten Zweig der Toxikologie, bei dem Mutagene aus der Umwelt identifiziert werden. Die meisten krebserzeugenden Agenzien, die **Karzinogene**, sind auch Mutagene. Man hat erkannt, dass die Identifizierung von Mutagenen in vielen Bereichen sehr wichtig ist. Dazu zählen Pharmazeutika, Nahrungszusätze, Landwirtschaft, Verschmutzungsanalyse und viele industrielle Prozesse. Deshalb war die Entwicklung von Tests notwendig, um diese gefährlichen Verbindungen zu identifizieren. Frühere Testsysteme waren auf kleine Säugetiere wie Ratten oder Mäuse angewiesen. Diese Tests waren

aber zeitintensiv, sehr teuer und unter ethischen Gesichtspunkten bedenklich. Es wurden einige *in vitro*-Tests entwickelt, bei denen Bakterien oder Tierzellen aus Gewebekulturen verwendet werden. In vielen Ländern müssen neue Chemikalien von Gesetzes wegen mithilfe einer Reihe verschiedener Tests untersucht werden, bevor sie von den Behörden zugelassen werden.

Der bekannteste Test ist der Ames-Test. Dies ist ein sehr schneller Mutagentest. Er verwendet *Salmonella typhimurium*-Bakterien, die eine Mutation im Histidin-Operon (Abschnitt 1.7) besitzen und daher kein Histidin synthetisieren können. Man bezeichnet sie als Histidin-auxotroph (*his*<sup>-</sup>). Verbindungen werden daraufhin untersucht, ob sie die Umkehrung dieser Mutation induzieren können. Der Test ist sehr einfach durchzuführen. Auf einer Agarplatte, die nur eine Spur von Histidin enthält, werden *his*<sup>-</sup>-Mutanten ausplattiert. Ein Glas oder eine Filterscheibe mit der zu testenden Verbindung wird auf die Oberfläche des Agars gegeben. Die Bakterien wachsen eine kurze Zeit lang, bis das Histidin aufgebraucht ist. Die einzigen Bakterien, die danach weiter wachsen und Kolonien bilden können, sind diejenigen, bei denen eine Reversionsmutation stattgefunden hat und die daher in der Lage sind, ihr eigenes Histidin zu synthetisieren. Sie sind Histidin-prototroph (*his*<sup>+</sup>). Wenn die zu testende Verbindung nicht mutagen ist, findet man ebenfalls einige revertante Kolonien, die zufällig über die Agarplatte verteilt sind. Wenn die Verbindung mutagen ist, steigt die Anzahl der Kolonien an und sie häufen sich um den Punkt herum, an dem die Verbindung auf die Platte aufgebracht wurde. Der Test kann quantitativ aufgebaut werden, sodass man für jede getestete Verbindung eine Dosis-Wirkungs-Kurve erhält.

Um die größtmögliche Anzahl an Mutagenen zu identifizieren, werden zwei Typen der *his*<sup>-</sup>-Mutanten verwendet. Bei einer handelt es sich um eine Mutation, bei der eine einzelne Base ausgetauscht wurde, bei der anderen handelt es sich um eine Rasterschubmutation (Abschnitt 1.5). So kann man Mutagene nachweisen, die auf verschiedene Weise auf die DNA einwirken. Man kann sowohl die Permeabilität des Bakteriums für die zu untersuchende Verbindung verändern, als auch seine Reparaturfähigkeit geschädigter DNA senken. Dies erhöht auch die Nachweiswahrscheinlichkeit für mutagene Aktivität.

Einige Verbindungen, von denen man weiß, dass sie Krebs auslösen, sind dazu erst in der Lage, nachdem sie im Körper enzymatisch in Mutagene umgewandelt wurden. Diese Substanzen nennt man **Prokarzinogene**. Enzymatische Aktivität wandelt sie in **endgültige Karzinogene** um. Wenn Prokarzinogene im Ames-Test eingesetzt werden, ist das Testergebnis negativ. Der Test kann jedoch entsprechend angepasst werden. Die Leber ist eine Fundgrube für aktivierende Enzyme. Man kann also Leberextrakt, der diese Enzyme enthält, zusammen mit der zu testenden Verbindung zugeben. Dann wird das Prokarzinogen aktiviert, und man kann eine erhöhte Anzahl von Revertanten nachweisen. Der Ames-Test ist schnell. Er kann innerhalb von 48 Stunden durchgeführt werden. Er ist billig und einfach quantifizierbar. Mit seiner Hilfe wurden viele Verbindungen, wie bestimmte Haarfärbemittel, Flammschutzmittel und Lebensmittelfarbstoffe, als Mutagene identifiziert.

Eine Schwäche des Ames-Tests ist, dass es sich bei dem Zielorganismus um ein Bakterium und nicht um ein Säugetier handelt. Aus diesem Grund wurden viele Tests mit kultivierten tierischen Zelllinien entwickelt. Sie sind dem Ames-Test sehr ähnlich, verwenden jedoch andere Selektionssysteme und Testgene.

Agenzien, die Chromosomen schädigen, nennt man **Clastogene**. Auch sie werden durch Testsysteme nachgewiesen. Die Tests können an Zelllinien, an Labortieren oder sogar an Pflanzen durchgeführt werden. Bei den Tests werden Chromosomenaberrationen wie Brüche, Austausch von Basen, Ringchromosomen, dizentrische Chromosomen und Translokationen ausgewertet. Die Aberrationshäufigkeit ist oft gering. Eine Alternative zur Auszählung solcher Chromosomenaberrationen ist die Auswertung des **Schwesterchromatidaustausches** (SCE, *sister chromatid exchange*). Unter SCE versteht man den Austausch von Material zwischen den beiden Chromatiden des gleichen Chromosoms. Dieser Prozess kommt spontan und mit geringer Häufigkeit in allen Zelltypen vor. Er kann sowohl in der Mitose als auch in der Meiose auftreten. Er ist deshalb nützlich, weil die Häufigkeit, mit der SCEs als Reaktion auf die Behandlung mit einem Clastogen auftreten, viel schneller zunimmt als große Chromosomenaberrationen. SCE kann mit mehreren Verfahren nachgewiesen werden. Alle beruhen auf der semikonservativen Natur der DNA-Replikation (Abschnitt 1.1). Das Wesen dieser Technik besteht darin, die beiden Schwesterchromatiden unterschiedlich anzufärben, sodass ein Austausch leicht beobachtet werden kann. Man erreicht dies mithilfe einer komplizierten Färbemethode. Zunächst werden die Zellen über zwei Teilungszyklen hinweg in Gegenwart des Thymidinanalogs Bromdesoxyuridin (BrdU) kultiviert und anschließend gefärbt. Anstelle von Thymidin wird BrdU in die neu synthetisierte DNA eingebaut und verändert die Färbeeigenschaften des Chromatids, sodass eines dunkel und das andere hell gefärbt ist.



#### Noch einmal in Kürze

##### Mutagene

Mutationen entstehen durch seltene Fehler bei der DNA-Replikation oder infolge der Einwirkung chemischer oder physikalischer Agenzien auf die DNA. Derartige Agenzien nennt man Mutagene. Veränderungen der Nukleotidstruktur ändern die Basenpaarung. Diese wird nach der DNA-Replikation dauerhaft. Andere Mutagene verursachen eine physikalische Deformation der DNA, die die Replikation oder Transkription blockiert.

##### Chemische Mutagene

Viele verschiedene Chemikalien wirken als Mutagene. Basenanaloga ersetzen während der DNA-Replikation die normalen Basen und führen zu Mutationen, weil sie veränderte Basenpaarungsmuster besitzen. Interkalierende Substanzen schieben sich zwischen die Basen in der Doppelhelix. Sie bewirken die Insertion einer zusätzlichen Base. Während der Replikation entsteht eine Raster-schubmutation. Viele chemische Mutagene modifizieren Basen. Oft geschieht dies durch Anheften von Alkyl- oder Arylgruppen oder durch Desaminierung.



DNA unterliegt auch der spontanen Mutation durch Reaktionen mit normalen Chemikalien in der Zelle. Auch reaktive Sauerstoffspezies (ROS, *reaktive oxygen species*), die in aeroben Zellen vorkommen, können Basen schädigen.

#### **Physikalische Mutagene**

Ionisierende Strahlung in Form von Röntgenstrahlen oder Gammastrahlen schädigen DNA-Moleküle beträchtlich. Ultraviolette Strahlung wird von den Basen absorbiert und führt zur Bildung von Cyclobutyldimeren zwischen benachbarten Pyrimidinbasen. Auch Hitze ist ein bedeutendes Mutagen.

#### **DNA-Reparatur**

Die Gegenwart zahlreicher Agenzien, die die DNA mutieren, hat dazu geführt, dass Organismen umfangreiche DNA-Reparaturmechanismen entwickelt haben. Die Reparaturmechanismen sind komplex. Es gibt jedoch im Wesentlichen drei Haupttypen: Exzisionsreparatur, direkte Reparatur und Fehlpaarungsreparatur.

#### **Exzisionsreparatur**

Ein Reparaturenzym schneidet die DNA neben einem beschädigten Nukleotid, sodass Einzelstrangbrüche entstehen. Eine Nuklease entfernt die beschädigte und die benachbarten Basen. Die entstandene Lücke wird dann von einer DNA-Polymerase mit neuer DNA aufgefüllt und von einer DNA-Ligase geschlossen.

#### **Direkte Reparatur**

Bei diesem Mechanismus werden strukturelle Veränderungen der Nukleotide rückgängig gemacht. Ein wichtiges Beispiel dafür ist die Photoreaktivierung. Dabei werden Thyminidimere von DNA-Photolyasen enzymatisch repariert. Dies geschieht durch Spaltung der durch die Dimerisierung entstandenen Bindungen.

#### **Fehlpaarungsreparatur**

Dieses System korrigiert Fehler der DNA-Replikation, indem es fehlgepaarte Nukleotide erkennt und die falsch eingefügte Base ersetzt. Das System erkennt, welcher der beiden DNA-Stränge der Ausgangsstrang ist – dieser Strang ist methyliert – und bestimmt so, welche der fehlgepaarten Basen die richtige ist.

#### **Genotoxizität**

Genotoxizität bezieht sich auf den Nachweis von Agenzien, die DNA schädigen und so Mutationen auslösen. Kommerzielle Tests mit Tieren wurden inzwischen weitgehend durch Systeme mit Bakterien- oder Tierzellkulturen ersetzt.

Der bekannteste Test ist der Ames-Test. Dabei wird untersucht, ob ein potenzielles Mutagen die Fähigkeit besitzt, Mutationen im Histidin-Operon von *Salmonella typhimurium* rückgängig zu machen. Durch Mutagene entstehen erhöhte Revertanzahlen. Wenn man Rattenleberextrakt zugibt, kann der Test für den Nachweis von Prokarzinogenen modifiziert werden. Tierzelllinien

können in ähnlichen Tests eingesetzt werden. Clastogene sind Mutagene, die eine Chromosomenschädigung verursachen. Sie können nachgewiesen werden, indem man ihre Fähigkeit zur Induktion des Schwesterchromatidaustauschs (SCE, *sister chromatid exchange*) abschätzt. SCEs treten spontan auf, aber die Induktionsrate ist sehr empfindlich gegenüber Mutagenen. SCEs sind nur nachweisbar, wenn die beiden Chromatiden desselben Chromosoms unterschiedlich angefärbt sind. Dies erreicht man, indem man Zellen über zwei Teilungszyklen mit Bromdesoxyuridin markiert, bevor man sie färbt.



#### Tipp

##### Verwandte Themen:

- (Abschnitt 1.1) DNA-Struktur
- (Abschnitt 1.5) DNA-Mutation

## 1.7 Genexpression

### 1.7.1 Regulation der Genexpression

Abgesehen von wenigen Ausnahmen codieren alle Gene für Proteine, und die Genexpression ist dafür verantwortlich, welche Proteine in einer Zelle oder in einem Gewebe vorhanden sind. Einige Proteine kommen in allen Zellen vor, andere liegen nur zu bestimmten Zeiten oder in bestimmten Zelltypen vor. Daher muss das Genexpressionsmuster reguliert werden. Das gilt sowohl für einzellige Prokaryoten als auch für vielzellige Eukaryoten. Gene können „an-“ oder „abgeschaltet“ werden, je nachdem, ob die Proteine, für die sie codieren, vorhanden sind oder nicht. Expressionsmuster können auch quantitativ verändert werden.

Genexpression kann auf unterschiedlichen Niveaus reguliert werden. Dazu gehören die Proteinaktivität und -stabilität und die Translation der mRNA an Ribosomen. Der wichtigste Mechanismus betrifft jedoch die Transkription (Abschnitt 1.4). Sowohl bei Prokaryoten als auch bei Eukaryoten sind Gene beteiligt, die für Proteine codieren, die die Transkription anderer Gene regulieren. Die Mechanismen unterscheiden sich in den beiden Reichen jedoch sehr.

### 1.7.2 Regulation der Transkription bei Prokaryoten

Ein entscheidendes Merkmal, das festlegt, wie die Gentranskription in Bakterien reguliert wird, ist die Organisation der Gene in **Operons**. Das sind Transkriptionseinheiten, in denen mehrere Gene, die üblicherweise für Proteine mit verwandten Funktionen codieren, zusammen reguliert werden. Es gibt auch andere Gene, die für regulatorische Proteine codieren, die die Genexpression in Operons kontrollieren. In *E. coli* wurden viele verschiedene Operons identifiziert. Die meisten davon enthalten Gene, die für Proteine codieren, die an der Biosynthese von Aminosäuren oder am Metabolismus von Nährstoffen beteiligt sind. Operons werden als **induzierbar** oder **reprimierbar** eingestuft. Induzierbare Operons enthalten Gene,

die für Enzyme codieren, die an Stoffwechselwegen beteiligt sind. Die Expression der Gene wird durch ein Substrat des Stoffwechselwegs kontrolliert. Ein Beispiel für ein induzierbares Operon ist das **lac-Operon**, das für Enzyme codiert, die für den Lactosestoffwechsel benötigt werden. Reprimierbare Operons enthalten Gene, die für an Biosynthesewegen beteiligte Enzyme codieren. Die Genexpression wird durch das Endprodukt des Stoffwechselwegs kontrolliert. Es kann die Expression des Operons unterdrücken oder das Operon durch einen alternativen Mechanismus, der **Attenuierung** genannt wird, kontrollieren. Ein Beispiel für ein reprimierbares Operon ist das **trp-Operon**, das für Enzyme codiert, die an der Biosynthese von Tryptophan beteiligt sind.

### 1.7.3 Das lac-Operon

Dieses Operon enthält drei Gene, die für Enzyme codieren, die *E. coli* für die Verwertung des Disaccharids Lactose benötigt. Diese sind die Lactose-Permease, die Lactose in die Zelle transportiert, die  $\beta$ -Galactosidase, die Lactose in seine beiden Zuckerkomponenten (Glucose und Galactose) hydrolysiert, und die  $\beta$ -Galactosid-Transacetylase, die auch an der Hydrolyse von Lactose beteiligt ist. Diese Enzyme liegen normalerweise in *E. coli* in sehr geringen Konzentrationen vor. Bei Anwesenheit von Lactose steigt ihre Konzentration jedoch schnell an. Die drei Gene im lac-Operon heißen *lac Z*, *Y* und *A*. Sie codieren für  $\beta$ -Galactosidase, Lactose-Permease und  $\beta$ -Galactosid-Transacetylase. Die Gene sind aufeinanderfolgend und werden als einzige mRNA von einem einzigen Promotor aus transkribiert. Ein anderes regulatorisches Gen, *lac I*, das gesondert exprimiert wird, liegt stromaufwärts des Operons und codiert für das Lac-**Repressor**protein. Der Lac-Repressor reguliert die Expression der *lac Z*-, *Y*- und *A*-Gene. Wenn keine Lactose vorhanden ist, bindet der Lac-Repressor an eine DNA-Sequenz, die man als **Operator** bezeichnet. Der Operator liegt zwischen dem *lac*-Promotor und dem Beginn des *lac Z*-Gens. Wenn der Lac-Repressor an den Operator gebunden hat, blockiert er den Weg der RNA-Polymerase, die stromaufwärts an den *lac*-Promotor gebunden hat, und verhindert die Transkription der *lac*-Gene. Wenn die Zelle auf Lactose trifft, kann sie die Lactose mithilfe der wenigen Lac-Enzyme, die in der Zelle vorhanden sind, aufnehmen und metabolisieren. **Allolactose**, ein Lactose-Isomer, das als Intermediat im Verlauf des Lactosestoffwechsels gebildet wird, wirkt als Induktor. Allolactose bindet an den Lactoserepressor und verändert seine Konformation, sodass dieser nicht mehr an den Operator binden kann. Der Weg der RNA-Polymerase ist nicht mehr blockiert und das Operon wird transkribiert. Es werden viele Enzymmoleküle gebildet, die die Lactose aufnehmen und verstoffwechseln. Vorhandene Lactose induziert somit die Expression der Enzyme, die für ihre Metabolisierung nötig sind. Wenn die Lactose aufgebraucht ist, kehrt der Lac-Repressor in seine ursprüngliche Konformation zurück, bindet wieder an den *lac*-Operator und verhindert so die Transkription. Das Operon wird abgeschaltet (Abb. 1.30).

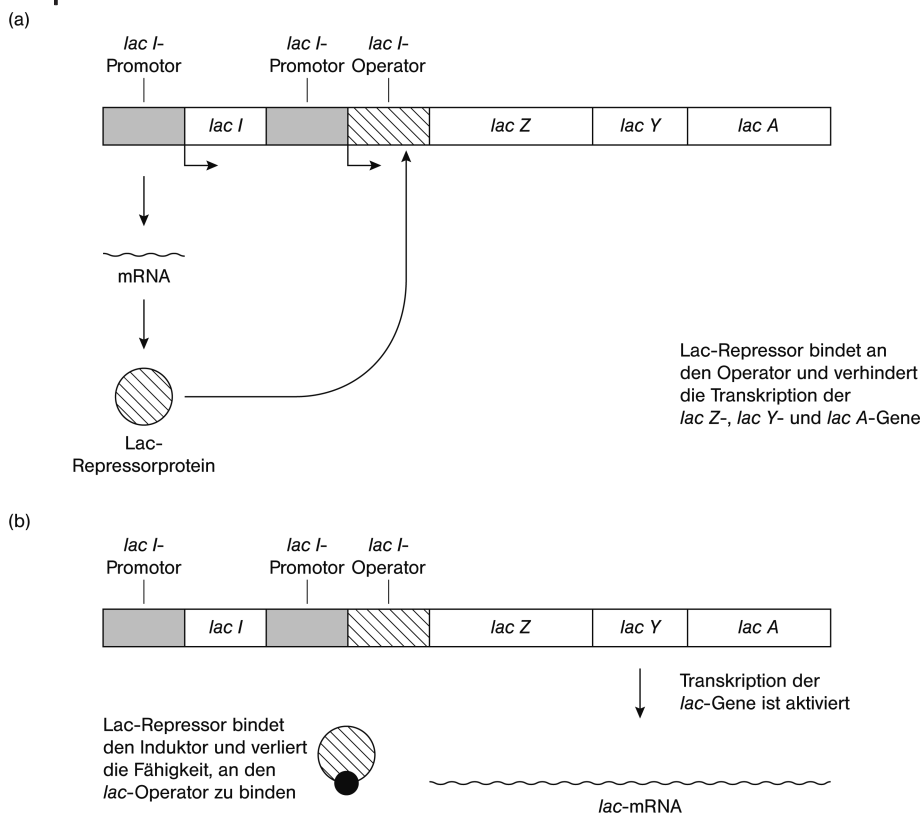


Abb. 1.30 Regulierung des *lac*-Operons (a) in Abwesenheit und (b) in Anwesenheit eines Induktors.

### 1.7.4 Katabolitrepression

Dieser Ausdruck beschreibt einen weiteren regulatorischen Mechanismus, der es dem *lac*-Operon ermöglicht, auf die Anwesenheit von Glucose zu reagieren. Glucose ist eine bevorzugte Energiequelle und eine Alternative zur Lactose. Wenn sowohl Glucose als auch Lactose vorhanden sind, verbrauchen die Zellen zuerst die Glucose und wenden keine Energie dafür auf, Lactose in ihre Zuckerkomponenten zu spalten. Die Anwesenheit von Glucose schaltet das *lac*-Operon durch einen Mechanismus ab, den man **Katabolitrepression** nennt. An der Katabolitrepression ist ein regulatorisches Protein, das **Katabolit-Aktivatorprotein** (KAP), beteiligt (Abb. 1.31). KAP bindet an eine DNA-Sequenz stromaufwärts des *lac*-Promotors und verstärkt die Bindung der RNA-Polymerase. Das bewirkt eine gesteigerte Transkription des Operons. KAP bindet aber nur in Gegenwart von zyklischem Adenosinmonophosphat (cAMP), einem ATP-Derivat. Die cAMP-Konzentration wiederum wird von Glucose beeinflusst. Das Enzym **Adenylat-Cyclase** katalysiert die Bildung von cAMP und wird durch Glucose gehemmt. Wenn

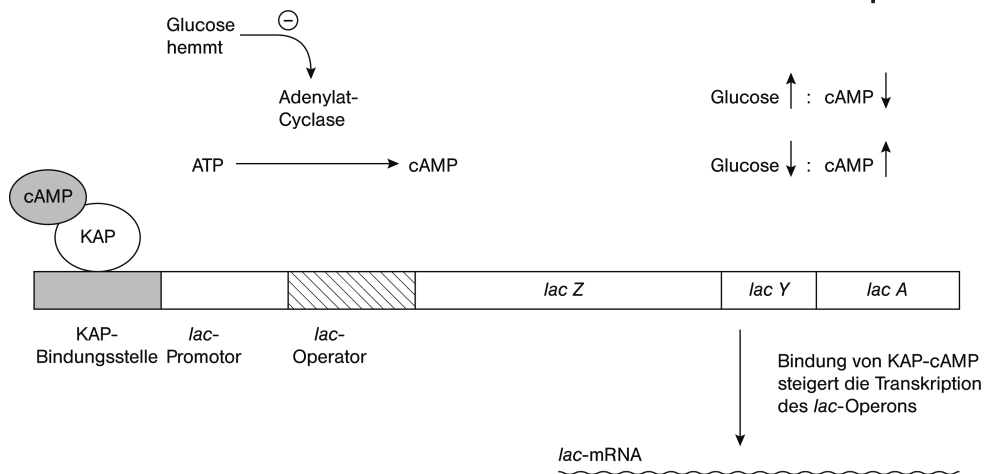


Abb. 1.31 Katabolitrepression des *lac*-Operons.

Glucose in der Zelle verfügbar ist, wird die Adenylat-Cyclase gehemmt, und der cAMP-Spiegel ist niedrig. Unter diesen Bedingungen bindet KAP nicht stromaufwärts vom Promotor, und das *lac*-Operon wird nur auf sehr niedrigem Niveau transkribiert. Im umgekehrten Fall, wenn wenig Glucose vorhanden ist, wird die Adenylat-Cyclase nicht gehemmt, die cAMP-Konzentration ist höher und KAP bindet. Dadurch steigert es die Transkription des Operons. Wenn Glucose und Lactose gleichzeitig vorliegen, wird das *lac*-Operon nur auf niedrigem Niveau transkribiert. Wenn die Glucose jedoch aufgebraucht ist, endet die Katabolitrepression, und die Transkription des *lac*-Operons steigt. So kann die verfügbare Lactose verbraucht werden.

### 1.7.5 Das *trp*-Operon

Dieses Operon enthält fünf Gene, die für Enzyme codieren, die an der Biosynthese der Aminosäure Tryptophan beteiligt sind. Die Gene werden als eine einheitliche mRNA von einem stromaufwärts gelegenen Promotor transkribiert. Die Expression des Operons wird durch die Tryptophankonzentration in der Zelle reguliert (Abb. 1.32). Ein regulatorisches Gen stromaufwärts des *trp*-Operons codiert für ein Protein, das man den **Trp-Repressor** nennt. Dieses Protein bindet an eine DNA-Sequenz, den ***trp*-Operator**. Diese Sequenz liegt direkt stromabwärts des *trp*-Promotors und überlappt teilweise mit ihm. Wenn in der Zelle Tryptophan vorhanden ist, bindet es an das Trp-Repressorprotein, sodass dieses an die *trp*-Operatorsequenz zu binden. Dadurch blockiert es die Bindung der RNA-Polymerase an den *trp*-Promotor und verhindert die Transkription des Operons. In Abwesenheit von Tryptophan ist der Trp-Repressor nicht in der Lage, den *trp*-Operator zu binden. Die Transkription läuft ab. Tryptophan, das Endprodukt der Enzyme, die vom *trp*-

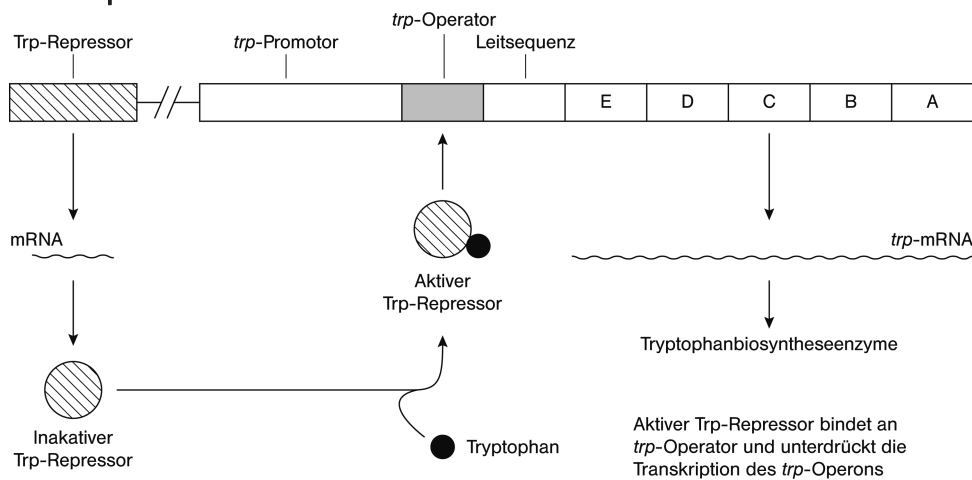
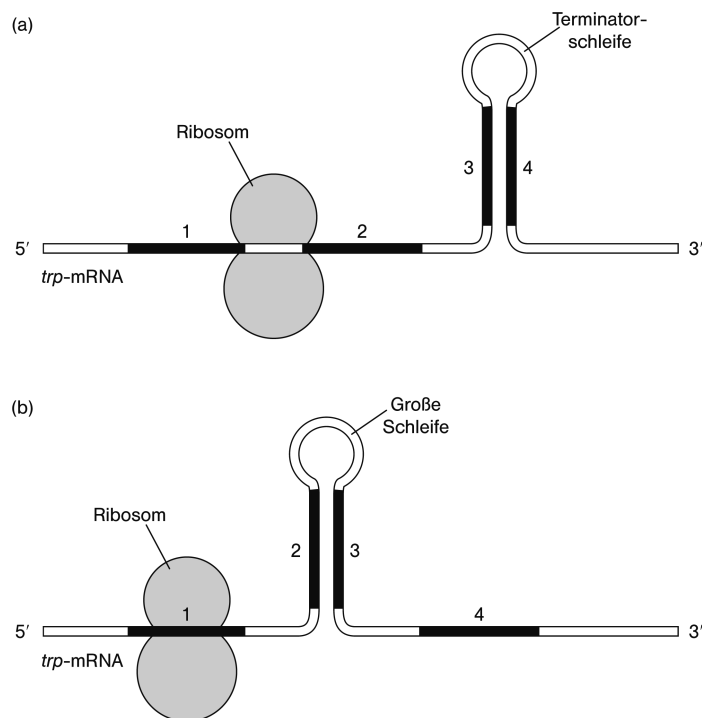


Abb. 1.32 Das *trp*-Operon.

Operon codiert werden, wirkt zusammen mit dem Trp-Repressorprotein als Co-Repressor und hemmt über Endprodukthemmung seine eigene Synthese.

### 1.7.6 Attenuierung

Das *trp*-Operon verwendet eine alternative Strategie, um die Transkription zu kontrollieren. Diese Strategie nennt man Attenuierung. Sie ermöglicht die Feinabstimmung der Expression (Abb. 1.33). Die transkribierte mRNA-Sequenz zwischen dem *trp*-Promotor und dem ersten *trp*-Gen kann zwei Haarnadelschleifenstrukturen bilden. Die beiden Sequenzen sind so positioniert, dass die beiden Haarnadelschleifen nicht gleichzeitig entstehen können. Es kann jeweils nur eine der beiden Schleifen gebildet werden. Die größere, stabilere Struktur beeinflusst die Transkription nicht und entsteht stromaufwärts von der kleineren Haarnadelschleife, die als Transkriptionsterminator wirkt. Wenn sich diese kleinere Struktur bildet, wird die Transkription gestoppt, noch bevor das erste Gen erreicht ist, und die Genexpression wird verhindert. Attenuierung basiert auf der Tatsache, dass in Bakterien Transkription und Translation miteinander verknüpft sind: Ribosomen heften sich an die mRNA noch während diese synthetisiert wird und beginnen, sie in Protein zu übersetzen. An eine mRNA, die gerade transkribiert wird, können bereits ein oder mehrere Ribosomen angelagert sein. Die Bindung von Ribosomen an die *trp*-mRNA hat Einfluss darauf, welche der beiden Haarnadelschleifenstrukturen entstehen kann, und bestimmt so, ob die Transkription beendet wird oder nicht. Direkt stromaufwärts der Haarnadelschleifenregion befindet sich ein kurzes offenes Leseraster aus 14 Codons und einem anschließenden Stopp-Codon, die vor den Strukturgenen translatiert werden. Zwei dieser 14 Codons codieren für Tryptophan. Wenn die Tryptophankonzentration entsprechend hoch ist, translatiert das Ribosom die codierende Region und folgt dabei dicht der RNA-Polymerase. Unter



**Abb. 1.33** Attenuierung im *trp*-Operon. (a) Wenn Tryptophan im Überschuss vorhanden ist, verhindert das Ribosom die Bildung der großen Schleife. Die Terminatorschleife kann entstehen. (b) Wenn wenig Tryptophan vorhanden ist, gerät das Ribosom ins Stocken. Dadurch kann sich die große Schleife bilden und die Entstehung der Terminatorschleife wird verhindert.

diesen Umständen verhindert die Anwesenheit des Ribosoms die Bildung der größeren Haarnadelschleife und ermöglicht die Bildung der Terminatorschleife. Die Transkription endet. Bei Tryptophanmangel kommt das Ribosom zum Stillstand, während es die codierende Region translatiert. Die RNA-Polymerase wandert weiter, und die erste Haarnadelschleife kann entstehen. Die Bildung der Terminatorschleife ist dann blockiert, und die Transkription des Operons kann weiter ablaufen. Die Geschwindigkeit, mit der das Ribosom die codierende Region übersetzt, ist nicht bei jedem Transkript gleich. Wenn Tryptophan im Medium in mittlerer Konzentration vorhanden ist, werden manche Transkripte beendet und andere nicht. So ist eine Feineinstellung der Transkription des Operons möglich. Insgesamt bestimmt der Trp-Repressor, ob das Operon an- oder abgeschaltet wird, und die Attenuierung bestimmt, wie effizient es transkribiert wird. Beides hängt von der Tryptophankonzentration in der Zelle ab. Attenuierung erlaubt der Zelle eine exakt auf ihre Bedürfnisse abgestimmte Tryptophansynthese.

Attenuierung ist nicht auf das *trp*-Operon beschränkt. Sie kommt mindestens in sechs weiteren Operons vor, die für Enzyme der Aminosäurebiosynthese codieren. Einige Operons wie zum Beispiel das *trp*- und das *phe*-Operon werden über

Repressoren und Attenuierung reguliert. Andere wie das *his*-, *leu*- und *thr*-Operon stützen sich nur auf Attenuierung.

### 1.7.7 Regulation durch alternative Sigmafaktoren

Die bakterielle RNA-Polymerase besteht aus fünf einzelnen Polypeptiduntereinheiten (zwei  $\alpha$ ,  $\beta$ ,  $\beta'$  und  $\omega$ ). Eine weitere Untereinheit, der **Sigma( $\sigma$ )faktor**, ist für die Einleitung der Transkription verantwortlich. Er ist an der Erkennung der bakteriellen Promotor-DNA-Sequenzen beteiligt. Bakterien, wie z. B. *E. coli*, besitzen alternative Sigmafaktoren, die verschiedene Promotoren erkennen und bewirken, dass die RNA-Polymerase verschiedene Gengruppen transkribiert. Das ist eine Methode, um die Genexpression zu regulieren, wo Umweltbedingungen bedeutende Veränderungen im Genexpressionsmuster auferlegen. Der  $\sigma^{70}$ -Faktor ist der von *E. coli* am häufigsten verwendete Faktor. Alternative  $\sigma$ -Faktoren kommen in vielen Situationen zum Tragen, wie zum Beispiel bei der Reaktion auf Hitzeschock in *E. coli* oder bei der Sporenbildung in *Bacillus subtilis*.

### 1.7.8 Regulation der Transkription bei Eukaryoten

Die Regulation der Gentranskription wird durch die Wechselwirkung von Genpromotoren und DNA-bindenden Proteinen, den **Transkriptionsfaktoren**, erreicht (Abb. 1.34). Die Transkription eines Gens durch die RNA-Polymerase wird am Promotor initiiert. Die Effizienz der Initiation kann durch Wechselwirkungen zwischen kurzen regulatorischen DNA-Sequenzen im Promotor und in Transkriptionsfaktor-Proteinen verändert werden. Die regulatorischen Sequenzen im Promotor liegen auf dem gleichen Chromosom (DNA-Doppelhelix) wie die codierende Sequenz. Man spricht von **cis-aktiven** Sequenzen.

In eukaryotischen Zellen werden alle proteincodierenden Gene von der RNA-Polymerase II transkribiert. Die Transkription wird durch die Bildung des Transkriptionsinitiationskomplexes (TIC) initiiert. Dabei binden die RNA-Polymerase II und eine Reihe assoziierter Proteine, die man als **basale Transkriptionsfaktoren** bezeichnet, an eine charakteristische Sequenz in der DNA des Promotors. Sie hat die Sequenz 5' TATA(A/T)A(A/T)3' und wird als **TATA-Box** bezeichnet. Sie kommt in den meisten, aber nicht in allen eukaryotischen Genen vor und liegt etwa 25 Bp stromaufwärts der Transkriptionsinitiationsstelle. Ihre Funktion be-

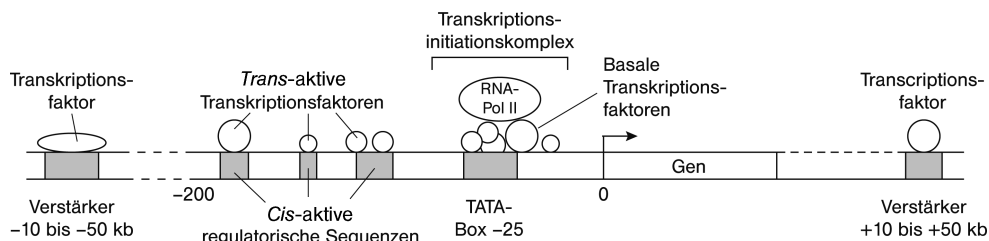


Abb. 1.34 Regulation der Genexpression in Eukaryoten.



steht darin, die RNA-Polymerase in die korrekte Position zu bringen, um die Transkription einzuleiten. Einige Gene, insbesondere die, die nur in bestimmten Geweben oder Zellen exprimiert werden, besitzen keine TATA-Box. Stattdessen haben sie eine Initiatorsequenz, die sich üblicherweise über dem Transkriptionsstartpunkt befindet. Andere Gene, Gene, die üblicherweise auf niedrigem Niveau exprimiert werden, besitzen weder eine TATA-Box noch ein Initiatorelement.

Wie effizient die Transkriptionsinitiation ist und wie viel mRNA dementsprechend gebildet wird, hängt noch von anderen Transkriptionsfaktoren ab. Diese Faktoren binden an andere DNA-Sequenzen im Promotor und können mit den Proteinen des TIC in Wechselwirkung treten und so dessen Stabilität beeinflussen. Transkriptionsfaktoren können die Transkriptionsrate steigern oder senken. Es gibt viele verschiedene Transkriptionsfaktoren. Jeder erkennt und bindet eine DNA-Sequenz im Promotor. Die erkannte Sequenz kann von Promotor zu Promotor variieren. Die Bindungsstelle wird üblicherweise als Konsensussequenz mit möglichen Abweichungen beschrieben. Transkriptionsfaktoren werden im Cytoplasma synthetisiert, wirken aber im Zellkern. Von daher werden sie oft als **transaktive** Faktoren bezeichnet.

Die Transkriptionsrate eines Gens kann auch durch Sequenzelemente, die als **Enhancer** (Verstärker) bezeichnet werden, beeinflusst werden. Ein Enhancer kann sich Tausende Basenpaare entfernt vom Startpunkt der Transkription befinden. Enhancer sind typischerweise 100-200 Bp lang und enthalten Sequenzen, die Transkriptionsfaktoren binden und die Transkription des verknüpften Gens stimulieren. Die Position des Enhancers relativ zum Gen, das er beeinflusst, kann variieren. Sie kann stromaufwärts oder stromabwärts liegen. Enhancer arbeiten unabhängig von ihrer Ausrichtung und sind gleichermaßen wirksam, egal, ob sie vorwärts oder rückwärts ausgerichtet sind. Wechselwirkung zwischen dem Enhancer und seinem Promotor erfolgt über Schleifenbildung der dazwischenliegenden DNA, um die beiden in unmittelbare Nähe zu bringen. Einige Enhancer enthalten Sequenzen, die Transkriptionsfaktoren binden, die wiederum die Transkription negativ beeinflussen. Man nennt sie **Silencer** (Dämpfer). Sie können für die Beschränkung der Expression auf spezifische Zelltypen verantwortlich sein. Enhancer und Silencer wirken im Allgemeinen, indem sie die Chromatinstruktur öffnen oder schließen (Abschnitt 1.8). Es gibt noch andere entfernt liegende Sequenzen, die **Locus-Kontrollelemente** genannt werden. Sie beeinflussen die Expression gesamter Gencluster, indem sie den Zugang der Transkriptionsproteine zur DNA kontrollieren. Beispiele dafür sind die Locus-Kontrollregionen, die die Expression der Globin-Genfamilien regulieren (Abschnitt 3.8).

### 1.7.9 Transkriptionsfaktoren

Dazu zählen viele Proteine, die die Expression von Genen regulieren. Sie unterscheiden sich von den basalen Transkriptionsfaktoren, die mit der RNA-Polymerase II in Wechselwirkung treten, um den TIC zu bilden. Transkriptionsfaktoren haben vielfältige Expressionsmuster: Einige kommen nur in speziellen Zelltypen vor, während andere in allen Zelltypen zu finden sind. Transkriptions-

faktoren wirken kombinatorisch. Sie beeinflussen sich gegenseitig. Oft müssen sie sich mit anderen Molekülen verbinden. Sie können Homodimere bilden, wenn sie sich mit anderen Molekülen desselben Transkriptionsfaktors zusammenlagern, oder Heterodimere, wenn es sich um andere Transkriptionsfaktoren handelt.

Transkriptionsfaktoren besitzen eine modulare Struktur. Sie sind aus einzelnen Proteindomänen mit speziellen Funktionen zusammengesetzt. Üblicherweise kommen drei Arten von Domänen vor:

- DNA-bindende Domänen – Beispiele dafür sind Helix-Turn-Helix, Zinkfinger und einfache Domänen;
- Dimerisierungsdomänen – Beispiele dafür sind Leucin-Zipper und HLH;
- Transaktivierungsdomänen – diese Domänen enthalten viele saure Aminosäuren.

Einige Transkriptionsfaktoren können die Transkription unterdrücken. Das kann auf viele Weisen erreicht werden. Einige reagieren direkt mit dem Transkriptionsinitiationskomplex. Andere können auf zahlreichen Wegen indirekt wirken: (i) durch Blockade der DNA-Bindungsstelle eines aktivierenden Transkriptionsfaktors; (ii) durch Bildung eines Dimers, dem eine DNA-Bindungsdomäne fehlt, oder (iii) durch Binden eines Repressorproteins an die Aktivierungsdomäne eines anderen Transkriptionsfaktors.

### 1.7.10 Regulation der Genexpression durch Hormone und Cytokine

Hormone sind Stoffe, die von Zellen gebildet werden und die Funktion anderer Zellen auf verschiedene Arten beeinflussen. Dazu gehört die Veränderung des Gentranskriptionsmusters. Hormone können kleine Moleküle sein, oft Steroide wie Östrogene und Glucocorticoide oder Polypeptide wie das Insulin. Cytokine sind Proteine, die ähnlich wirken wie Hormone. Oft haben sie Blutzellen zum Ziel. Hormone und Cytokine regulieren die Genexpression in den Zielzellen auf unterschiedliche Weise. Steroidhormone sind fettlöslich und können so durch die Zellmembran ins Cytoplasma gelangen, wo sie an einen Transkriptionsfaktor, einen **Steroidhormonrezeptor**, binden. Die Bindung bewirkt, dass der Steroidhormonrezeptor von einem inhibitorischen Protein freigesetzt wird. Er dimerisiert dann und gelangt in den Zellkern, wo er die Transkription der Zielgene aktiviert, indem er an Promotorsequenzen bindet. Polypeptidhormone und Cytokine haben eine andere Wirkweise als Steroidhormone. Sie binden an Rezeptorproteine auf der Oberfläche der Zielzelle. Genaktivierung wird durch einen Prozess ausgelöst, den man als **Signaltransduktion** bezeichnet. Dabei wird nacheinander ein Netzwerk von Proteinen durch Proteinphosphorylierung oder Proteolyse aktiviert. Schließlich stimuliert die Bindung von Transkriptionsfaktoren an Genpromotorsequenzen die Transkription der Zielgene.

### 1.7.11 Posttranskriptionelle Regulation der Genexpression durch RNA-Interferenz

Genexpression kann auch reguliert werden, indem die Translation der mRNA verhindert wird. RNA-Interferenz (RNAi) tritt auf, wenn ein RNA-Molekül Basenpaarungen mit einem mRNA-Molekül eingeht und so einen Doppelstrang bildet. Das kann die Translation blockieren und/oder den Abbau der mRNA zur Folge haben. Die interferierende RNA bezeichnet man als **Antisense-RNA**, da sie komplementär zur „Sinn“-mRNA (Sense-RNA) ist, die für Aminosäuren codiert. Dieses Verfahren wurde zur Entwicklung der ersten kommerziellen genetisch veränderten Pflanze, der FlavrSavr-Tomate, eingesetzt. Bei dieser Tomate wurde die Aktivität des Enzyms Polygalacturonase reduziert. Dieses Enzym lässt reife Tomaten weich werden. Ohne dieses Enzym kann man die Tomate reifen lassen, bevor man sie erntet (Abschnitt 6.3). Die Technik wird häufig eingesetzt, um die Genexpression experimentell zu reduzieren, damit sich Veränderungen im Phänotyp beobachten lassen.

Eukaryoten verwenden RNAi-Mechanismen zur Regulation der Genexpression. Doppelsträngige RNA bildet sich, wenn sie sich mit sich selbst paart (Foldback) oder durch Hybridisierung zwischen mRNA und Antisense-RNA. Diese doppelsträngige RNA wird durch ein Enzym in 21–28 Bp lange, kurze, interferierende RNAs (siRNA, *short interfering RNA*) geschnitten. Dieses Enzym wird „Dicer“ genannt, da es die RNA in kleine Stücke schneidet (*to dice* – in Würfel schneiden). Die siRNAs lagern sich mit Proteinen, darunter ein Mitglied der *Argonauten*-Familie, zusammen, werden einzelsträngig und bewahren den Antisense-Strang. Der RNA-Protein-Komplex wird als **RNA-induzierter Silencing-Komplex (RISC)** bezeichnet. Er paart sich mit komplementärer mRNA. Wenn die Übereinstimmung vollkommen oder nahezu vollkommen ist, wird die mRNA gespalten und dann abgebaut. Der RISC kann die Spaltung anderer mRNA fortsetzen. Wenn es Fehlpaarungen gibt, wird die mRNA nicht gespalten, aber die Translation kann gestoppt werden. Kleine RNAs mit perfekter Basenpaarung bezeichnet man als **kurze interferierende RNAs (siRNAs)**. Kleine RNAs mit einigen Fehlpaarungen nennt man **mikro-RNAs (miRNAs)**. Die Zielsequenzen in der mRNA befinden sich häufig, aber nicht ausschließlich in den nicht translatierten Regionen an einem der beiden Enden.

Mikro-RNAs werden von *mir*-Genen gebildet und kommen in Pflanzen, Invertebraten (etwa 100 Gene), Vertebraten (etwa 250 Gene) und einigen Pilzen vor. Die Gene besitzen invertierte Sequenzen, die sich zum Stiel einer Haarnadelstruktur aneinander lagern können. Diese Struktur wird von *Drosha* geschnitten. Im Anschluss daran kürzt *Dicer* auf 22 Bp lange Stränge (18 gepaarte Basen und zwei aus je 2 Nukleotiden bestehende 3'-endständig überstehende Enden). Genomsequenzierungsdaten werden eingesetzt, um siRNAs zu konstruieren, mit deren Hilfe die Genexpression experimentell gehemmt werden kann. Doppelsträngige (ds) RNA kann durch Einfügen der Sequenz zwischen Promotoren hergestellt werden, sodass beide Stränge transkribiert werden. Die dsRNA kann in Kulturzellen transfiziert oder in Zellen oder Organismen injiziert werden und RNAi

auslösen. So erreicht man den phänotypischen Effekt eines Genverlusts, ohne die Zeit und Kosten für die Herstellung inaktivierter Mutanten aufwenden zu müssen. Dieses Verfahren kann bei genetisch schwierig zu manipulierenden kultivierten humanen Zellen oder Organismen eingesetzt werden. Es ist sehr nützlich bei der Untersuchung von Genen mit unbekannter Funktion.

Einige Viren bilden dsRNA, die das inhibierende System zur Zerstörung viraler RNA veranlasst. Der gleiche Mechanismus ist auch beteiligt an der Methylierung von DNA, der Kondensation von Heterochromatin des Centromers und der Modifizierung von Chromatin, das transponierbare Elemente oder wiederholte Sequenzen enthält. Verursacht wird dies anscheinend von einer geringen Transkription von Sequenzen mit entgegengesetzter Orientierung, möglicherweise an verschiedenen Loci, bei der komplementäre RNA-Stränge entstehen, die sich paaren können. RNAi ist an der epigenetischen Regulierung beteiligt (Abschnitt 1.8).

### 1.7.12 Differenzierung und Entwicklung

Voraussetzung für die Transkription von Genen ist der Zugang der RNA-Polymerase zur DNA. Eukaryoten schalten manche Gene aus, indem sie das Chromatin so modifizieren, dass es unzugänglich wird. Dazu gehören die Methylierung von DNA, üblicherweise von Cytosin, und die Methylierung spezifischer Aminosäuren in den Histonen. Das Chromatin wird dadurch so kompakt, dass es solange nicht zugänglich ist, bis ein Chromatin-Umformungskomplex die Veränderungen rückgängig macht. Das Ausschalten ist „halbdauerhaft“. Es bleibt während der somatischen Teilung bestehen und hält in der Entwicklung über die Lebensdauer des Organismus an. In einigen Fällen bleibt es über nachfolgende Generationen bestehen (Abschnitt 1.8 und 3.8).



#### Noch einmal in Kürze

##### Regulation der Genexpression

Einige Proteine kommen in allen Zellen eines Organismus vor, während andere nur in einigen Zellen oder zu bestimmten Zeiten vorkommen. Welche Gene „an-“ oder „ab“geschaltet sind, wird durch eine Reihe regulatorischer Wege festgelegt. Diese wirken auf unterschiedlichen Ebenen. Am wichtigsten ist jedoch die Kontrolle der Transkription. Dafür sind einige Gene nötig, die die Transkription anderer Gene kontrollieren.

##### Regulation der Transkription bei Prokaryoten

Viele bakterielle Gene sind in Form von koordiniert regulierten Operons angeordnet, die für Proteine mit verwandten Funktionen codieren. Induzierbare Operons wie das *lac*-Operon codieren für Enzyme, die an Stoffwechselwegen beteiligt sind und durch das Substrat des Stoffwechselwegs induziert werden. Reprimierbare Operons wie das *trp*-Operon codieren für Enzyme, die

an Biosynthesewegen beteiligt sind. Die Genexpression wird durch das Endprodukt des Stoffwechselwegs oder durch Attenuierung reguliert.

### Das *lac*-Operon

Dieses Operon enthält drei Gene (*lac Z*, *Y* und *A*), die für Enzyme codieren, die *E. coli* für den Lactose-Stoffwechsel benötigt. Die Gene werden von einem einzigen Promotor transkribiert. Ihre Expression wird durch Lactose induziert. In Gegenwart von Lactose bindet Allolactose den Lac-Repressor und verhindert, dass er an den Operator bindet. Das Operon kann transkribiert werden. Wenn Lactose aufgebraucht ist, kann der Lac-Repressor wieder an den *lac*-Operator binden und die Transkription ist blockiert.

### Katabolitrepression

Dieser regulatorische Mechanismus ermöglicht es *E. coli*, das *lac*-Operon in Gegenwart von Glucose zu reprimieren. Das Katabolit-Aktivatorprotein (KAP) bindet cAMP und stimuliert die Transkription des *lac*-Operons, indem es stromaufwärts des *lac*-Promotors bindet. Glucose wiederum hemmt die Adenylat-Cyclase und reguliert so die cAMP-Konzentration. Wenn Glucose verfügbar ist, ist der cAMP-Spiegel niedrig, KAP bindet nicht an den *lac*-Promotor, und das Operon wird in geringem Ausmaß transkribiert. Wenn die Glucosekonzentration niedrig ist, steigt der cAMP-Spiegel, KAP bindet den *lac*-Promotor und stimuliert die Transkription des Operons. Katabolitrepression stellt sicher, dass zunächst Glucose verbraucht wird, falls sowohl Glucose als auch Lactose verfügbar sind.

### Das *trp*-Operon

Dieses Operon enthält fünf Gene, die von einem einzigen Promotor transkribiert werden. Sie codieren für Enzyme, die für die Biosynthese von Tryptophan benötigt werden. Der Trp-Repressor bindet in Gegenwart von Tryptophan an den *trp*-Operator und blockiert die Transkription des Operons. In Abwesenheit von Tryptophan bindet der Trp-Repressor nicht, und das Operon wird transkribiert.

### Attenuierung

Dieser Regulationsmechanismus ermöglicht die Feineinstellung der Expression des *trp*-Operons und anderer Operons. mRNA-Sequenzen zwischen dem *trp*-Promotor und dem ersten *trp*-Operon-Gen können entweder eine große Haarnadelschleifenstruktur bilden, die die Transkription nicht beeinflusst, oder eine kleinere Terminationsschleife. Eine kurze codierende Region stromaufwärts enthält Tryptophan-Codons. Bei ausreichender Tryptophankonzentration transkribiert die RNA-Polymerase die Region. Der RNA-Polymerase folgt direkt ein Ribosom, das verhindert, dass sich die größere Haarnadelschleife bildet. So kann die Terminationsschleife entstehen und die Transkription wird beendet. Bei Tryptophanmangel kommt das Ribosom zum Stillstand, die RNA-Polymerase wandert weiter und die große Haarnadelschleife bildet sich. Die Ent-

stehung der Terminationsschleife wird blockiert, und die Transkription des Operons wird fortgesetzt.

#### **Regulation durch alternative Sigmafaktoren**

Alternative  $\sigma$ -Faktoren verändern die Spezifität der bakteriellen RNA-Polymerase, sodass sie verschiedene Genpromotoren erkennt.

#### **Regulation der Transkription bei Eukaryoten**

Eukaryotische Zellen regulieren die Genexpression meistens, indem sie die Geschwindigkeit der Gentranskription variieren. Wechselwirkungen zwischen RNA-Polymerase II und basalen Transkriptionsfaktoren führen zur Entstehung des Transkriptionsinitiationskomplexes (TIC) an der TATA-Box. Andere Transkriptionsfaktoren verändern die Geschwindigkeit der Transkriptionsinitiation, indem sie an Promotorsequenzen binden und die Stabilität des TIC beeinflussen. Entfernte Sequenzen, die als Enhancer (Verstärker) oder Silencer (Dämpfer) bezeichnet werden, beeinflussen ebenfalls die Transkriptionsgeschwindigkeit.

#### **Transkriptionsfaktoren**

Genpromotoren besitzen viele Bindungsstellen für Transkriptionsfaktoren. Jeder dieser Faktoren kann die Transkription beeinflussen. Die Gesamtwirkung hängt von der Gesamtheit der gebundenen Transkriptionsfaktoren ab. Transkriptionsfaktoren besitzen eine Modulstruktur, zu der DNA-Bindung, Dimerisierung und Transaktivierungsdomänen mit charakteristischen Strukturmotiven gehören. Transkriptionsfaktoren arbeiten kombinatorisch. Sie bilden Hetero- und Homodimere. Transkriptionsfaktoren können die Transkription auch durch direkte oder indirekte Mechanismen unterdrücken.

#### **Regulation der Genexpression durch Hormone und Cytokine**

Hormone und Cytokine beeinflussen die Zielzellen, indem sie die Gentranskriptionsmuster verändern. Steroidhormone gelangen in die Zellen und binden an das Steroidhormon-Rezeptorprotein. Durch die Bindung wird das Rezeptorprotein von einem inhibitorischen Protein gelöst. Der Rezeptor dimerisiert und wird in den Zellkern transloziert. Dort bindet er an Promotoren der Zielgene und aktiviert die Transkription. Polypeptidhormone und Cytokine binden Rezeptorproteine auf der Oberfläche der Zielzellen. Die Genaktivierung wird durch Signaltransduktion ausgelöst. Dabei wird ein Netzwerk aus Proteinen der Reihe nach durch Proteinphosphorylierung aktiviert.

#### **Posttranskriptionelle Regulation durch RNA-Interferenz**

RNA-Interferenz (RNAi) tritt auf, wenn doppelsträngige RNA durch die Enzyme Droscha und Dicer in mikro-RNA (miRNA) oder kurze interferierende RNA (siRNA, *short interfering RNA*)-Fragmente geschnitten wird. Diese Fragmente sind üblicherweise 22 Bp lang. Sie lagern sich mit Proteinen, zu denen ein Argonautenprotein gehört, zusammen und bilden einen RNA-induzierten Still-

egungskomplex (RISC, *RNA-induced silencing complex*). Die RNA wird einzelsträngig und leitet den Komplex zu komplementärer mRNA, an die er bindet. Anschließend katalysiert er das Schneiden der mRNA. Wenn die Basenpaarung fehlerhaft ist, kann nicht geschnitten werden. Der Komplex blockiert jedoch die Translation. Genomsequenzdaten können verwendet werden, um Antisense-RNA herzustellen. Dies hat sich als sehr effektive Methode erwiesen, um experimentell die Expression von Genen zu senken, damit man Auswirkungen auf den Phänotyp untersuchen kann.

#### Differenzierung und Entwicklung

Differenzierung und Entwicklung von Zellen und Geweben geschieht bei vielzelligen Organismen über langfristige epigenetische Regulierung der Gene. Dazu gehören molekulare Modifikationen der DNA und Histone, die DNA unzugänglich für RNA-Polymerasen machen.

#### Tipp

##### Verwandte Themen:

- (Abschnitt 1.2) Gene
- (Abschnitt 1.4) Von der DNA zum Protein
- (Abschnitt 1.8) Epigenetik und Chromatinmodifikation
- (Abschnitt 2.2) Prokaryotengenome
- (Abschnitt 2.3) Eukaryotengenome
- (Abschnitt 3.8) Gene in der Entwicklung



## 1.8 Epigenetik und Chromatinmodifikation

### 1.8.1 Überblick

Im Lauf der Zeit hat sich die Bedeutung des Begriffs „Epigenetik“ geändert. Aktuell wird er verwendet, um Veränderungen der genetischen Expression zu beschreiben, die zu einem gewissen Grad erblich sind, bei denen es sich aber nicht um Veränderungen der DNA-Sequenz handelt. Sie sind dauerhafter als die normale Genregulation durch Repressoren (Abschnitt 1.7). Die DNA muss auf eine Art und Weise markiert werden, die über die DNA-Replikation und die Zellteilung hinweg bestehen bleibt. Es gibt drei Hauptmanifestationen: (i) Differenzierung – die gewebespezifische Regulation der Gene, die festlegt, zu welchem bestimmten Zelltyp eine Zelle während der gesamten Lebenszeit des Organismus gehört; (2) Prägung – die monoallele Expression bestimmter Gene, bei der das Allel von Elternteilen des einen Geschlechts unterdrückt wird; (iii) sporadische Chromosomenumordnung, die einen epigenetischen Kontroll-Locus in die Nähe eines Gens bringt, das normalerweise nicht unter dessen Kontrolle steht. Die Erkennung einer Stelle zur epigenetischen Modifikation wird von RNA-tragenden Proteinen gesteuert, die dann das Chromatin langfristig modifizieren.

### 1.8.2 Chromatinmodifikation und der Histon-Code

Epigenetische Effekte fallen bei vielzelligen Tieren und bei Pflanzen auf. Inaktives Chromatin besitzt höhere Konzentrationen von 5-Methylcytosin. Bei Vertebraten kann das Cytosin in 5' CpG 3'-Dinukleotiden enzymatisch methyliert werden, während bei Pflanzen das Hauptziel 5' CNG 3'-Triplets sind (N = ein beliebiges Nukleotid). Bei Invertebraten gibt es keine flächige Cytosinmethylierung. Sie besitzen aber gleichwertiges kondensiertes Heterochromatin.

Man dachte, die Modifikation von Histonen sei einfach und Acetylierung entspreche aktivem Chromatin. Es gibt jedoch eine komplexe Reihe von Modifikationen, die alternative Zustände signalisieren. Man nennt dies den „Histon-Code“, da spezifische Modifikationen der Histone dazu führen, dass spezifische Proteine binden und spezifische Gene regulieren. Am besten untersucht sind die N-terminalen Enden der Histone H3 und H4. H3 besitzt an den Positionen 4, 9, 14, 18, 23 und 27 die Aminosäure Lysin, die mono-, di- oder trimethyliert oder -acetyliert werden kann. Trimethylierung von Lysin 4 und Acetylierung von Lysin 9 scheint zu aktivieren, während Methylierung von Lysin 9 die Bindung von Heterochromatinprotein 1 (HP1) begünstigt und inaktiviert. Das Polycomb-Genprodukt ist in *Drosophila* an der Aufrechterhaltung der Gewebedifferenzierung beteiligt. Es bindet an trimethyliertes Lysin 27 und grenzt HP1 aus. Die Serinmoleküle an den Positionen 10 und 28 können phosphoryliert werden. Weitere Forschung wird die Bedeutung weiterer Modifikationen aufdecken.

Gene im Heterochromatin sind normalerweise inaktiv, vermutlich weil die Struktur der RNA-Polymerase den Zugang verwehrt. 5-Methylcytosin bildet 1,8-mal festere Wasserstoffbrückenbindungen mit Guanin als Cytosin und stabilisiert die Doppelhelix. Die DNA von Invertebraten wird nicht ausgiebig methyliert, falls sie überhaupt methyliert wird. Die Markierung von Chromatin bei Invertebraten funktioniert vermutlich alleine über RNA und Proteinmodifikationen.

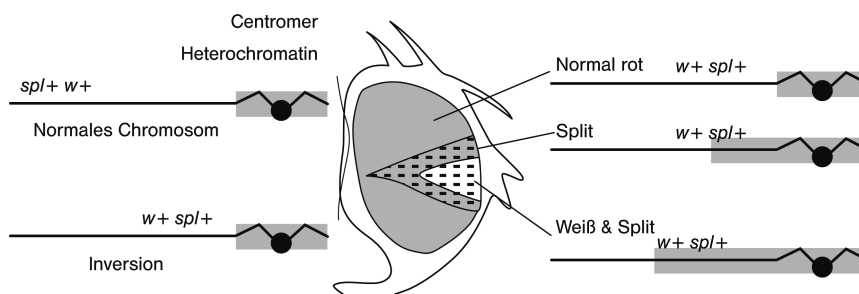
Nach der DNA-Replikation (Abschnitt 1.1) ist nur der alte Strang methyliert. Die hemimethylierten mCpG·CpG-Palindrome in der DNA sind Hauptziele für die Erhaltungsmethylasen, Enzyme, die die Cytosine des neuen Strangs gegenüber der mCGs auf dem alten Strang methylieren. In Säugetierzellen gibt es verschiedene Methyltransferasen. DNMT1 kommt am häufigsten vor und besitzt eine große Präferenz für hemimethylierte DNA. Es ist am Replikationsfokus konzentriert und ist die wichtigste Erhaltungsmethylase. DNMT3A und DNMT3B sind die wichtigsten *de novo*-Methyltransferasen. Methyl-CpG-bindende Proteine (MECPs) treten mit methylierter DNA in Wechselwirkung. MECP2 hemmt direkt den Transkriptionsfaktor IIB und kann Histon H1 von Nukleosomen verdrängen. Es reagiert mit anderen Proteinen mit Methyl-CpG-Bindungsdomänen (MBD1, MBD2, MBD3) und mit Histon-Deacetylasen, um die Chromatinstruktur umzuformen und zu kondensieren. Kondensiertes inaktives Chromatin enthält auch methylierte Histone. Den ersten Hinweis darauf, wie die Chromatinmodifikation gesteuert wird, gab es 2002, als man entdeckte, dass RNA-Interferenz beteiligt ist. Doppelsträngige RNAs (dsRNAs) werden in etwa 22 Bp lange Fragmente zerlegt. Diese reagieren mit Proteinen und leiten sie zu homologer mRNA, die dann zerstört wird.



(Abschnitt 1.7). In der Hefe *Schizosaccharomyces pombe* entstand durch geringe Transkription centromerer Wiederholungssequenzen (Abschnitt 2.4) in beiden Richtungen lange dsRNA, die nach der Prozessierung Proteine zu solchen DNA-Sequenzen leitete und heterochromatinbildende Modifikationen einleitete. RNA ist auch an der Inaktivierung des X-Chromosoms beteiligt, und Antisense-Transkription spielt bei der Prägung eine Rolle (s. unten).

### 1.8.3 Positionseffekt-Variation

Die DNA-Kondensierung (Heterochromatinbildung) breitet sich von Inaktivierungszentren durch kooperative Bindung von Heterochromatinproteinen und chromatinmodifizierenden Enzymen aus, bis sie auf ein Grenzelement oder auf einen Isolator trifft. Wenn die Gene durch chromosomale Neusortierung näher an Inaktivierungszentren rücken oder wenn Grenzelemente durch Deletion entfernt werden, kann die sich ausbreitende Inaktivierung (Heterochromatin) solche Gene erreichen und sie ausschalten. Das Muster wird klonal vererbt, sodass manche Klone das Gen exprimieren und andere nicht. So entsteht ein vielfältiges Erscheinungsbild. Die klassischen, erkennbaren Beispiele dafür sind die Split- und White-Loci auf dem X-Chromosom von *Drosophila*. Wenn sie durch Inversion nahe an das Centromer gelangen, kann sich das Heterochromatin über diese Gene ausbreiten und sie in einem keilförmigen Klon inaktivieren (Abb. 1.35). Das Gen, das näher am Centromer sitzt (*split*, *spl*), wird zuerst inaktiviert, sodass sich in einem Split-Zell-Klon immer weiße Zellen bilden. Eine Erkrankung des Menschen, fazio-scapulo-humerale Muskeldystrophie (FSHD) entsteht durch eine Deletion von Wiederholungssequenzen (Abschnitt 2.3) in der Nähe des Telomers eines kurzen Arms von Chromosom 13. Die Deletion scheint keine aktiv codierenden Sequenzen zu enthalten. Nach Verlust der Wiederholungen werden jedoch einige nahegelegene Gene hypomethyliert und überexprimiert. Das lässt auf eine verminderte Chromatinkondensation schließen.



**Abb. 1.35** Ausdehnung der Chromatinkondensation vom Centromer her inaktiviert zuerst das Split-Allel und dann das Weiß-Allel. So entsteht ein weißer Keil innerhalb des Klons mit dem Split-Phänotyp.

#### 1.8.4 Desaminierung von Methylcytosin

5-Methylcytosin desaminiert zu Thymin, das keine Wasserstoffbrückenbindungen mit Guanin bildet. Replikation oder ungeeignete Fehlpaarungsreparatur des Guanins zu Adenin ersetzt das ursprüngliche C–G-Paar durch ein T–A. Man nimmt an, dass durch diesen Vorgang 80 % der ursprünglichen CpG-Dinukleotide aus der DNA der Säugetiere entfernt wurden. Sie bleiben nur dort erhalten, wo entweder keine Methylierung vorkommt – in den Keimbahnzellen – oder dort, wo die Selektion sie bewahrt, indem sie Mutationen in essenziellen Codons und in Kontrollregionen entfernt, die an codierende Sequenzen angrenzen.

#### 1.8.5 CpG-Inseln

Es gibt CpG-Dinukleotid-Cluster, die man als **CpG-Inseln** bezeichnet. Sie liegen in der Nähe des Startpunktes vieler Säugetiergene, wo sie an der Transkriptionskontrolle beteiligt sind. Man findet sie vor allem am 5' (mRNA)-Ende des Gens in der Promotorregion, oft im ersten Exon. Sie können aber auch überall in der codierenden Sequenz vorkommen. Die Restriktionsnuklease HpaII schneidet CCGG, wenn das zentrale C nicht methyliert ist. In DNA von Zellen, in denen ein Gen aktiv ist, schneidet sie CpG-Inseln in kleine Fragmente. Experimente mit einigen Genen ergaben, dass ihre DNA in Geweben, in denen das Gen nicht transkribiert wird, nicht so häufig oder gar nicht geschnitten wird. Das lässt vermuten, dass die inaktive DNA methyliert ist. Sequenzdaten zeigen, dass etwa 40 % der Start-Codons in einer CpG-Insel liegen und dass die ersten Exons eine höhere CpG-Dichte besitzen als die folgenden Exons. Die Sequenzierung hat auch ergeben, dass manche Gene weit ausgedehntere und CG-reichere Inseln besitzen als andere; und es gibt einen fließenden Übergang zwischen Genen mit ausgedehnten Inseln und Genen ohne CpG-Inseln, statt einer klaren Unterscheidung zwischen Genen mit und ohne CpG-Inseln.

#### 1.8.6 Inaktivierung des X-Chromosoms

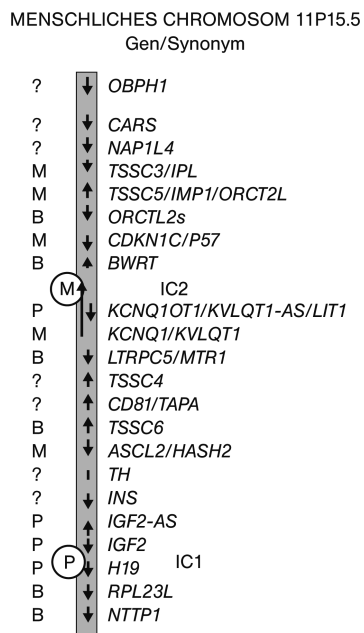
Da weibliche Säuger und *Drosophila* zwei X-Chromosomen besitzen, ist ein Mechanismus erforderlich, damit ihre mRNA-Produktion an die des einzelnen X-Chromosoms der männlichen Organismen angeglichen wird, damit ein Gleichgewicht mit den diploiden Autosomen bewahrt wird. Bei weiblichen Säugetieren bleibt nur ein X-Chromosom pro Zelle aktiv. Das zweite X-Chromosom und alle zusätzlichen Kopien bei polysomen Individuen werden inaktiviert (Abschnitt 2.4 und 3.6). An der X-Chromosomen-Inaktivierung bei Säugetieren spielt die Methylierung von CpG-Inseln eine Rolle. Bei *Drosophila* wird das einzelne X der Männchen hochreguliert, um die Transkription der mRNAs zu verdoppeln. Sowohl Säugetiere als auch *Drosophila* verwenden einen Mechanismus, der darauf basiert, dass reife RNA, die vom betroffenen X transkribiert wird, an das gleiche X bindet (in *cis*-Position). Bei Säugetieren wird der Prozess durch das X-Inaktivierungszentrum (*Xic*) kontrolliert, das etwa 1 Mb lang ist und verschiedene Elemente besitzt: Das X-inaktive spezifische Transkript (*Xist*) codiert für eine 17 kb lange nicht codierende RNA, die

vom inaktiven X transkribiert wird und es bedeckt. *TsiX* (*Xist* in entgegengesetzter Richtung) wird von dem Strang, der dem *Xist* gegenüber liegt, transkribiert und ist folglich komplementär zu *Xist*. *TsiX* beginnt in der Nähe des kurzen *DXPas34*-Gens. Eine Deletion von *DXPas34* führt immer zur Inaktivierung des X-Chromosoms. *Tsix* und *DXPas34* wirken vermutlich *Xist* entgegen. Ein anderes Element, *Xce* (X-Chromosom-kontrollierendes Element), spielt eine Rolle bei der Wahl zwischen den beiden Chromosomen. Einige Allele von *Xce* erhöhen die Wahrscheinlichkeit, dass ihr X-Chromosom aktiv bleibt. Nachdem das inaktive X mit stabiler *Xist*-RNA bedeckt ist, wird es zu spät repliziert. Man findet dann in einigen Nukleosomen die Histonvariante MakroH2A1.2, Histone werden hypoacetyliert und DNA wird hypermethyliert. Bei allen Beuteltieren und in den extraembryonalen Membranen bei Säugetieren ist das väterliche X immer inaktiviert. Das lässt vermuten, dass der Prozess als Form der Prägung begonnen hat. Bei *Drosophila melanogaster* (Fruchtfliege) gibt es zwei kleine RNAs, roX2 (1,1 kb) und roXi (3,5 kb), und fünf männchenspezifische letale (MSL) Proteine, MLE (maleless), MSL-1, MSL-2, MSL-3 und MOF (*males absent on first*). Diese bilden mit den RNAs einen Komplex und binden an Hunderten Stellen an das einzige X. MOF ist eine Histon-Acetyltransferase.

### 1.8.7 Prägung

Prägung ist eine Form der Allelaktivierung und -inaktivierung, die davon abhängt, von welchem Elternteil ein bestimmtes Chromosom stammt. Nur das Allel eines Chromosoms wird exprimiert, entweder das des ursprünglich mütterlichen oder das des ursprünglich väterlichen Chromosoms. Das inaktive Allel wird im Allgemeinen methyliert. Man hat bislang herausgefunden, dass bei Mäusen und Menschen über 30 Gene durch Prägung beeinflusst werden. Man spricht auch von **monoalleler Expression** oder von **Allelexklusion**. Sie wird in speziellen Geweben durchgehend beibehalten. In einem frühen Entwicklungsstadium wird das Allel eines Elternteils selektiv inaktiviert. Das entsprechende Identifizierungsverfahren ist jedoch nicht bekannt. DNA wird in den Gameten methyliert. Sechs bis acht Stunden nach der Befruchtung wird der Vorkern des Spermiums jedoch aktiv demethyliert. Die mütterliche DNA wird nach der Replikation nicht methyliert. Sie wird also bei jeder Mitose durch Ausdünnung demethyliert (passive Demethylierung). Damit dies geschehen kann, müssen sich das väterliche und das mütterliche Chromatin voneinander unterscheiden. Es ist aber nicht bekannt, in welchem Stadium die Allele für die Prägung erkannt werden. Es wurden keine Signale identifiziert. Die Demethylierung lässt aber vermuten, dass die Methylierung der Spermien-DNA nicht das wichtigste Signal für die Prägung ist. Die Prägung wird in Prägezentren kontrolliert. In einigen Fällen, wenn es Konkurrenz zwischen zwei benachbarten Promotoren um die Bindung von Transkriptionsfaktoren gibt, kann die Methylierung, durch die ein Gen supprimiert wird, die Expression des anderen zulassen.

Beim Menschen gibt es zwei große Prägungsregionen, eine auf Chromosom 11p15.5 und die andere auf Chromosom 15q12. Jede trägt ein Cluster geprägter Gene, aber einige Gene in jedem Cluster werden in entgegengesetzter Richtung



**Abb. 1.36** Die geprägte Region des menschlichen Chromosoms 11p15.5 zeigt die Richtung der Transkription (Pfeile) und die Prägung. Die Buchstaben auf der linken Seite bezeichnen die Expression väterlicher Allele (P), mütterlicher Allele (M) oder beider Allele (B). Die Kreise markieren die Position mütterlicher und väterlicher Inaktivierungszentren (IC2 bzw. IC1).

geprägt oder überhaupt nicht geprägt (Abb. 1.36). Ein Verlust des einzigen funktionsfähigen Allels führt zu einer Mangelerkrankung, da das andere Allel geprägt und somit inaktiv ist. Im Fall von 15q12 verursacht der Verlust des väterlichen Chromosomensegments das Prader-Willi-Syndrom. Der Verlust des mütterlichen Segments hat das Angelman-Syndrom zur Folge. Diese beiden Krankheiten haben unterschiedliche Symptome, da verschiedene Gengruppen derselben Region inaktiv sind (Abschnitt 5.1).

Es gibt eine Hypothese, dass die Konkurrenz zwischen Allelen vom Mann und von der Frau oder zwischen verschiedenen männlichen Elternteilen einen Selektionsvorteil für die Prägung schafft. Angeschaltete Gene vom männlichen Elternteil tendieren dazu, den einzelnen Nachkommen zu begünstigen (Plazentawachstum vereinnahmt zum Beispiel zusätzliche Nährstoffe von der Mutter) und können das Verhalten beeinflussen, indem sie die Konkurrenz unter den Geschwistern fördern. Angeschaltete Gene von der Mutter wirken dem eher entgegen, wodurch die Ressourcen unter den Nachkommen gleichverteilt werden.

In der Region 11p15.5 ist das Gen für den insulinartigen Wachstumsfaktor Typ 2 (*Insulin-like growth factor* Typ 2, IGF2) auf dem väterlichen Chromosom aktiv. Ein IGF-Rezeptor (IGFR), der IGF2 entfernt und ihm entgegenwirkt, ist im mütterlichen Genom aktiv. Diploide Zygoten, die zwei vom Vater stammende Genome enthalten, entwickeln massenhaft extraembryonales und plazentales Material. Entsprechende Zygoten, die zwei Genome von der Mutter enthalten, entwickeln embryonales Gewebe, aber sehr wenig Plazenta. Untersuchungen mit Mäusen bestätigen zufällige Ereignisse beim Menschen. Einige geprägte Gene werden nur

im Gehirn unterschiedlich exprimiert, wodurch das Verhalten kontrolliert wird. Sie können nach der Geburt die Konkurrenz unter den Geschwistern (väterlich) oder die Kooperation (mütterlich) unterstützen. Prägung wurde in Pflanzen nachgewiesen, wo beide Allelsätze in der Embryonalentwicklung (analog zu den Säugetieren) aktiv sind. Sie wurde aber nicht in *Drosophila*, *C. elegans* oder dem Zebrafisch entdeckt, wo der Beitrag des Weibchens zum Embryo abgeschlossen ist, bevor das Ei befruchtet wird.

### 1.8.8 Sporadische Effekte und Krebs

Anomale DNA-Methylierungsmuster sind in Krebszellen üblich. Sie verursachen anomale Expression von Genen, die die Zellteilung fördern (Onkogene), oder die Hemmung von Zellzyklus-regulierenden Genen (Krebssuppressoren). Methylierung des funktionsfähigen Allels in einer Zelle, die heterozygot für ein inaktives Allel ist, führt zum totalen Funktionsverlust. Dies kommt in Tumorzellen häufig vor und trägt zur Entwicklung von Krebs bei (Abschnitt 5.3). Dies kann nach der Neuordnung durch Positionseffekte auftreten oder weil die Chromatinmodifikation im Klon der Krebszellen ungenau wurde. Methylierung eines transponierbaren Elements, das in die Promotorregion eines Gens eingefügt wurde, kann die Expression dieses Gens verändern. Die Wahrscheinlichkeit, dass ein Gen inaktiviert wird, kann davon abhängen, ob das Chromosom von der Mutter oder vom Vater vererbt wurde. Die Viable-Yellow-Mutation von Mäusen ist zum Beispiel dominant, wenn sie aktiv ist. Wenn sie von Spermien übertragen wird, ist sie in etwa 19 % der Nachkommen inaktiv. Wenn sie jedoch nacheinander von Weibchen weitergegeben wird, steigt die Häufigkeit der Methylierung und damit die Inaktivierung mit jeder Generation an.

Die normale Entwicklung erfordert ein exaktes Methylierungssystem. Mehrere Erkrankungen beim Menschen sind mit einer fehlerhaften Methylierung verbunden (Abschnitt 5.1). Das ICF-Syndrom (Immundefizienz mit Centromerinstabilität und fazialen Dysmorphien) hängt mit dem Verlust der katalytischen Aktivität in der DNA-Methyltransferase 3B (DNMT3B) zusammen. Dieses Enzym ist an der Kontrolle der DNA-Methylierung beteiligt. Das Rett-Syndrom ist die häufigste sporadische Ursache für geistige Unterentwicklung bei Frauen (etwa 1:12.000). Es wird durch Mutationen im Methyl-CpG-Bindungsprotein 2 (MECP2)-Gen, das auf dem X-Chromosom liegt, hervorgerufen. Betroffene Frauen sind heterozygot und inaktivieren das funktionsfähige Allel in etwa der Hälfte ihrer Zellen. Einige Funktionen von MECP2 sind überlebensnotwendig. Daher überleben betroffene Männer nicht. Die häufigste Ursache geistiger Unterentwicklung bei Männern ist das Fragiles-X-Syndrom. Das FMR1 (*fragile X mental retardation 1*)-Gen enthält normalerweise 29 CGG-Wiederholungen. Träger besitzen 50 oder mehr Kopien. Bei einigen Müttern ist das an den Nachkommen übertragene Allel hypermethyliert und mit einer auf über 200 Kopien erhöhten Anzahl an Wiederholungen verbunden.

### 1.8.9 Ursprung der Methylierung

Eukaryotische Methylierung ähnelt bakteriellen Restriktionssystemen, die für die Abwehr von Bakteriophagen (Viren) eingesetzt werden. Man vermutet, dass die ursprüngliche Funktion der Methylierung darin bestand, Genomparasiten wie zum Beispiel Retroviren oder transponierbare Elemente zu inaktivieren. Etwa ein Drittel des menschlichen Genoms gehört zu diesem Typ. Es enthält etwa 90 % des Methylcytosins. Verschiedene Kopien der transponierbaren Elemente werden wahrscheinlich in verschiedenen Richtungen transkribiert, sodass sich zwei komplementäre RNA-Moleküle paaren können. So entsteht dsDNA, die den RNA-Interferenz-Mechanismus auslöst (Abschnitt 1.7). Pflanzen und Pilze (und vermutlich auch Tiere) können DNA-Elemente, die in Mehrfachkopien vorliegen, dadurch identifizieren, methylieren und inaktivieren. Pflanzen inaktivieren oft Transgene, die durch Gentechnik eingefügt wurden. Hypomethylierung von transponierbaren Elementen in Krebszellen aktiviert ihre Rekombinase-Gene und verursacht Mobilität, Chromosomenbrüche und -umordnung sowie weitere Mutationen.



#### Noch einmal in Kürze

##### Überblick

Unter epigenetischen Auswirkungen versteht man (nach derzeitigem Sprachgebrauch) langfristige Veränderungen der Genexpression durch Modifikation der Chromatinstruktur, ohne Veränderungen der DNA-Sequenz. Sie können die Differenzierung in Zelllinien in einem vielzelligen Organismus (Augengewebe oder Knochenzellen) aufrechterhalten oder die Expression von einer Generation zu nächsten betreffen. Viele epigenetische Auswirkungen werden in der Gametogenese oder in der Embryogenese gelöscht.

##### Chromatinmodifikation und der Histon-Code

Kondensation (Verdichtung/Heterochromatinbildung) eines Chromatinbereichs kann die Transkription der Gene in dieser Region verhindern, wahrscheinlich dadurch, dass der erfolgreiche Zugriff der Transkriptionsmaschinerie blockiert wird (Transkriptionsrepression). Die Kondensation geht einher mit der Modifikation von DNA und Histonen, was wiederum eine Kaskade von Proteinbindungen antreibt, die das Chromatin verdichten. DNA wird durch Methylierung von Cytosin zu 5-Methylcytosin in 5' CpG3'-Dinukleotiden bei Vertebraten oder in 5' CNG3'-Trinukleotiden bei Pflanzen modifiziert. (Bei Invertebraten wird Cytosin nicht methyliert.) Nach der DNA-Replikation erkennen Erhaltungsmethylasen hemimethylierte metCpG·CpG-Palindrome (CpG liest auf jedem Strang dasselbe) und methylieren den neuen Strang, sodass er zum alten passt. *De novo*-Methylasen methylieren nicht methylierte DNA. Histone werden hauptsächlich an N-terminalen Schwänzen modifiziert, die aus den Nukleosomen hervorragen und andere Proteine binden. Zur Kontrolle gehören Acetylierung oder Methylierung von Lysin und auch die Phosphorylie-

rung von Serin. Es gibt einen komplizierten Code spezifischer Modifikationen, die spezifische Gene regulieren. Man nennt ihm den „Histon-Code“. Die Inaktivierung wird über doppelsträngige RNA auf spezifische Sequenzen gelenkt. Der entsprechende Mechanismus ähnelt der RNA-Inhibition (Abschnitt 1.7).

#### **Positionseffekt-Variation**

Kondensation kann sich von Inaktivierungszentren her ausbreiten, bis sie von einem Grenzelement oder von einem Isolator gestoppt wird. Gene, die durch chromosomale Neusortierung näher an Inaktivierungszentren gerückt sind, können durch sich ausbreitendes Heterochromatin inaktiviert werden, wenn es ihre neue Position erreicht. Wachsende Klone betroffener Zellen ergeben vielfältige (variierte) Expressionsmuster.

#### **Desaminierung von Methylcytosin**

5-Methylcytosin kann spontan zu Thymin desaminieren. So entsteht eine T-G-Fehlpaarung. Diese kann falsch zu T-A repariert werden und dadurch eine Mutation verursachen. Man vermutet, dass auf diese Weise 80 % der ursprünglichen CpGs in der DNA von Säugetieren verloren gegangen sind. Daraus folgt, dass CpG nur dann überlebt, wenn es entweder nicht methyliert in Keimbahnzellen vorkommt oder wenn es für Codierungszwecke oder für die Kontrolle der Genexpression nötig ist, sodass Mutanten durch Selektion aussortiert werden.

#### **CpG-Inseln**

CpG-reiche Regionen kommen am Startpunkt vieler Säugetiergene vor. Sie haben dort anscheinend eine Regulationsfunktion. CpGs in diesen Bereichen sind vergleichsweise unmethyliert, während sie zu RNA transkribiert werden. In Zellen, in denen diese Gene inaktiv sind, können sie jedoch methyliert sein.

#### **Inaktivierung des X-Chromosoms**

Eine ausgewogene Genexpression benötigt einen Dosiskompensationsmechanismus, um die Transkription der beiden X-Chromosomen bei Frauen der des einzelnen X bei Männern anzugleichen. Bei Säugetieren ist nur ein X aktiv, die anderen sind inaktiviert und kondensiert. Der Zustand jedes X-Chromosoms wird über die somatische Mitose hinweg bewahrt. Bei *Drosophila melanogaster* wird das einzige X-Chromosom mit einer höheren Rate transkribiert, um es an die beiden X-Chromosomen des Weibchens anzugleichen (Abschnitt 3.6).

#### **Prägung**

Prägung kommt in Säugetieren und in Pflanzen vor. Sie ist zurückzuführen auf die embryonale Identifizierung des ursprünglichen Elternteils eines jeden Allels, gefolgt von der Expression nur eines Allels (uniparentale oder monoallele Expression). Bei Säugetieren wurden mehr als 30 geprägte Gene mit ähnlichen, aber nicht identischen Mustern bei Mäusen und Menschen identifiziert. Der Verlust des aktiven Allels führt zu einem Defizit und zur Erkrankung, obwohl

es noch ein gesundes ausgeschaltetes Allel auf dem anderen Chromosom gibt (zum Beispiel das Prader-Willi-Syndrom nach dem Verlust des väterlichen 15q12 und das Angelman-Syndrom nach dem Verlust des mütterlichen 15q12). Man vermutet, dass die evolutionär-selektive Ursache dafür die Konkurrenz zwischen Genen vom Vater und von der Mutter oder zwischen verschiedenen männlichen Elternteilen ist. Angeschaltete Gene vom männlichen Elternteil tendieren dazu, den einzelnen Nachkommen zu begünstigen (Plazentawachstum vereinnahmt zum Beispiel zusätzliche Nährstoffe von der Mutter) und können das Verhalten beeinflussen, indem sie die Konkurrenz unter den Geschwistern fördern. Angeschaltete Gene von der Mutter wirken dem eher entgegen, wodurch die Ressourcen unter den Nachkommen gleichverteilt werden.

#### **Sporadische Effekte und Krebs**

Krebs zeigt häufig eine anomale Methylierung und den Aktivitätsverlust in Tumorsuppressorgenen. Bei Genen, die ein transponierbares Element in der Promotorregion besitzen, kann die Expression durch Methylierung des Transposons modifiziert werden. Dies erfolgt oft geschlechtsspezifisch.

#### **Ursprung der Methylierung**

Die Methylierung von DNA bei Säugetieren und Pflanzen ähnelt den antiviralen Restriktionssystemen von Bakterien. Man vermutet, dass sich die DNA-Methylierung zur Inaktivierung von Retroviren und transponierbaren Elementen entwickelt hat. Das stimmt damit überein, dass sie unter der Regie von dsRNA auf hoch repetitive Sequenzen abzielt. Dieses Verfahren ist vor allem bei Pflanzen wirksam.



#### **Tipp**

##### **Verwandte Themen:**

- (Abschnitt 1.7) Regulation der Genexpression
- (Abschnitt 2.4) Chromosomen
- (Abschnitt 3.6) Geschlechtsbestimmung
- (Abschnitt 5.3) Gene und Krebs



#### **Wissen testen**

##### 1.1. Die Polarität von DNA:

- A) bezieht sich auf die Orientierung von Wasserstoffbrückenbindungen zwischen komplementären Basen.
- B) bezieht sich auf die beiden Stränge der Doppelhelix, die mit derselben Orientierung mit ihren 3'-Enden zusammenliegen.



- C) bezieht sich auf die Verknüpfung der 3'- und 5'-Atome der Desoxyribose und gibt die Syntheserichtung an.  
D) bezieht sich auf die elektrostatischen Ladungen der Basen in der Doppelhelix.
- 1.2. Wenn der Cytosinanteil in einem DNA-Molekül 0,3 beträgt:  
A) ist der Thyminanteil 0,2.  
B) ist der Guaninanteil 0,2.  
C) ist der Adeninanteil 0,6.  
D) ist der Thyminanteil 0,3.
- 1.3. Einige Schritte der DNA-Replikation lauten (1) Ligation des Okazaki-Fragments, (2) Trennung der beiden Stränge der Doppelhelix, (3) Replikation von kurzen Fragmenten in Richtung des Ursprungs auf dem 3'-komplementären Strang und (4) Initiation kontinuierlicher Replikation von Leitsträngen auf der 5'-Matrize. Wie lautet die richtige Reihenfolge dieser Schritte?  
A) 1,3,4,2.  
B) 2,4,3,1.  
C) 4,3,1,2.  
D) 4,3,2,1.
- 1.4. Welche der folgenden Aussagen ist **falsch**?  
A) Die meisten bakteriellen Gene enthalten Introns und Exons.  
B) Ein Gen trägt die kodierte Information für die Aminosäuresequenz in einem Polypeptid.  
C) Prokaryotische (bakterielle) Gene tendieren dazu, in koordiniert kontrollierten Operons zusammengefasst zu werden.  
D) Die Expression von Genen wird durch stromaufwärts gelegene Promotorregionen reguliert, die mit Transkriptionsfaktoren in Wechselwirkung treten.
- 1.5. Einfache Multigenfamilien:  
A) sind identische Kopien von Genen, die für Produkte kodieren, die von der Zelle in großer Menge benötigt werden.  
B) sind Kopien von Genen, die Proteine bilden, die sich ähneln, aber geringfügig unterschiedliche Funktionen besitzen.  
C) sind Ansammlungen von Genen, die dann mutiert sind, sodass sie nicht funktionsfähig sind.  
D) bestehen immer aus einem Tandemblock identischer Sequenzen.
- 1.6. Der genetische Code steht für:  
A) die von Watson und Crick entdeckte Doppelhelixstruktur der DNA.  
B) die drei aufeinanderfolgenden Nukleotide, die für eine spezifische Aminosäure kodieren.  
C) das spezifische Genom eines jeden Organismus.  
D) die Sequenzen des Genoms eines Organismus, die für Aminosäuren kodieren.
- 1.7. Welche der folgenden Aussagen ist NICHT richtig?  
A) Der genetische Code ist degeneriert.  
B) Der genetische Code ist mit seltenen Ausnahmen universell.

66 | 1 Molekulargenetik

- C) Der genetische Code enthält drei Stopp-Codons.
  - D) Die dritte Base in jedem Codon für eine bestimmte Aminosäure ist am wenigsten variabel.
- 1.8. In welche Richtung wird RNA während der Transkription verlängert?
- A) Von 5' nach 3'.
  - B) Von 3' nach 5'.
  - C) Von 2' nach 5'.
  - D) Von 5' nach 2'.
- 1.9. Welche der folgenden Aussagen ist NICHT richtig?
- A) Bei Eukaryoten beginnt die Polypeptidsynthese mit dem AUG-Codon.
  - B) Anticodons auf mRNA werden von Codons der tRNA erkannt.
  - C) tRNAs werden durch die Aminoacyl-tRNA-Synthetase kovalent mit den entsprechenden Aminosäuren verbunden.
  - D) Die nächste Aminosäure wird durch die Peptidyltransferase von der tRNA auf die wachsende Polypeptidkette übertragen.
- 1.10. Welche der folgenden Aussagen über Mutationen einzelner Basen ist **falsch**?
- A) Sinnverändernde Mutationen verändern eine Aminosäure im kodierten Polypeptid.
  - B) Sinnentstellende Mutationen verwandeln ein Codon in ein Stopp-Codon und verkürzen so das Polypeptid.
  - C) Eine Rasterschub-Mutation fügt Base ein oder entfernt Basen und verändert den Leserahmen der stromabwärts gelegenen Codons.
  - D) Eine stille Mutation tritt in einem Intron auf und wird nicht exprimiert.
- 1.11. Auxotrophe Mutaten:
- A) können bei hoher Temperatur nicht überleben.
  - B) sind resistent gegenüber Antibiotika.
  - C) besitzen ein Gen, das unangemessen an- oder ausgeschaltet wird.
  - D) benötigen Nährstoffergänzungen.
- 1.12. Der Austausch eines Thyminnukleotids durch ein Guanin ist ein Beispiel für:
- A) Translokation.
  - B) Transition.
  - C) Transversion.
  - D) Vorwärtsmutation.
- 1.13. Thymindimere werden hervorgerufen durch:
- A) UV-Licht.
  - B) Röntgenstrahlen.
  - C) Freie Radikale (z.B. Sauerstoff).
  - D) Desaminierung.
- 1.14. Einzel- oder Doppelstrangbrüche in der DNA werden verursacht durch:
- A) UV-Licht.
  - B) Röntgenstrahlen.
  - C) Freie Radikale (z.B. Sauerstoff).
  - D) Desaminierung.

- 1.15. Excisionsreparatur:
- A) identifiziert und entfernt einen doppelsträngigen Bereich beschädigter DNA.
  - B) identifiziert und entfernt einen einzelsträngigen Bereich beschädigter DNA.
  - C) identifiziert, entfernt und ersetzt eine falsche Base aus einem Nukleotid.
  - D) identifiziert, entfernt und korrigiert eine Transversionsmutation.
- 1.16. Ein induzierbares Gen:
- A) wird durch die Anwesenheit des Substrats angeschaltet.
  - B) wird durch die Anwesenheit des Substrats abgeschaltet.
  - C) wird durch die Anwesenheit des Produkts angeschaltet.
  - D) wird durch die Anwesenheit des Produkts abgeschaltet.
- 1.17. Welche der folgenden Aussagen beschreibt die Regulation der *lac*- und *trp*-Operons am besten?
- A) Das *lac*-Operon wird durch Lactose reprimiert und durch Glucose induziert, während das *trp*-Operon durch Tryptophan reprimiert wird.
  - B) Das *lac*-Operon wird durch Lactose induziert und durch Glucose reprimiert, während das *trp*-Operon durch Tryptophan induziert wird.
  - C) Das *lac*-Operon wird durch Lactose induziert und durch Glucose reprimiert, während das *trp*-Operon durch Tryptophan reprimiert wird.
  - D) Das *lac*-Operon wird durch Lactose reprimiert und durch Glucose induziert, während das *trp*-Operon durch Tryptophan induziert wird.
- 1.18. Bei Eukaryoten ist eine TATA-Box:
- A) die Stelle, an der Transkriptionsrepressoren binden.
  - B) die Stelle, an der Transkriptions-Enhancer binden.
  - C) die cis-aktive Stelle, an der Transkriptionsfaktoren und RNA-Polymerase binden.
  - D) ein trans-aktiver Transkriptionsaktivator.
- 1.19. Attenuierung des *trp*-Operons:
- A) verhindert die Transkription in Anwesenheit von Tryptophan.
  - B) verlangsamt die Transkriptionsrate und nimmt deren Feinabstimmung vor in Anwesenheit von Tryptophan.
  - C) verhindert die Transkription in ruhenden Zellen.
  - D) erhöht die Transkriptionsrate wenn Aminosäurekonzentrationen hoch sind und die Proteinsynthese schnell fortschreitet.
- 1.20. Welche der folgenden Aussagen beschreibt den Mechanismus der epigenetischen Kontrolle am besten?
- A) Spezifische DNA-Sequenzveränderungen in speziellen Zelltypen.
  - B) Chemische Modifikationen von DNA, Histonen und anderen Chromatinkomponenten.
  - C) Methylierung von Cytosinen in DNA.
  - D) Methylierung von Histonen.

68 | 1 Molekulargenetik

- 1.21. Prägung bezieht sich auf langfristige Veränderungen der Chromatinstruktur:
- A) in spezifischen Zelltypen wie z.B. Knochen.
  - B) in einem X-Chromosom von weiblichen Säugetieren.
  - C) die sich bei Spermien- und Eizellchromosomen unterscheiden.
  - D) die vom umweltbedingten Hintergrund, wie z.B. einer Hungerperiode in der Kindheit, verursacht werden.
- 1.22. Welche der folgenden Aussagen beschreibt die Dosiskompensation von Geschlechtschromosomen bei Säugetieren und *Drosophila* korrekt?
- A) Säugetiere regulieren die Transkription des einzelnen X bei Männchen hoch, *Drosophila* inaktiviert ein X bei Weibchen unter Verwendung nicht-translatierter RNA.
  - B) *Drosophila* reguliert die Transkription des einzelnen X bei Männchen hoch, Säugetiere inaktivieren ein X bei Weibchen ohne Verwendung nicht-translatierter RNA.
  - C) Säugetiere regulieren die Transkription des einzelnen X bei Männchen hoch, *Drosophila* inaktiviert ein X bei Weibchen ohne Verwendung nicht-translatierter RNA.
  - D) *Drosophila* reguliert die Transkription des einzelnen X bei Männchen hoch, Säugetiere inaktivieren ein X bei Weibchen unter Verwendung nicht-translatierter RNA.