



Dingermann / Winckler / Zündorf

# Gentechnik Biotechnik

Grundlagen und Wirkstoffe

4. AUFLAGE

WVG

Wissenschaftliche  
Verlagsgesellschaft  
Stuttgart





Dingermann / Winckler / Zündorf

---

# Gentechnik Biotechnik

Grundlagen und Wirkstoffe

Prof. Dr. Theodor Dingermann, Frankfurt/M.

Prof. Dr. Thomas Winckler, Jena

Dr. Ilse Zündorf, Frankfurt/M.

4., aktualisierte und erweiterte Auflage

Mit 614 Abbildungen und 73 Tabellen

## Zuschriften an

lektorat@dav-medien.de

## Anschriften der Autoren

Prof. Dr. Theodor Dinger  
Institut für Pharmazeutische Biologie  
Johann Wolfgang Goethe-Universität  
Max-von-Laue-Str. 9  
60438 Frankfurt/Main

Dr. Ilse Zündorf  
Institut für Pharmazeutische Biologie  
Johann Wolfgang Goethe-Universität  
Max-von-Laue-Str. 9  
60438 Frankfurt/Main

Prof. Dr. Thomas Winckler  
Friedrich-Schiller-Universität Jena  
Lehrstuhl für Pharmazeutische Biologie  
Institut für Pharmazie  
Simmelweisstr. 10  
07743 Jena

Um die Lesbarkeit des Werks zu verbessern, nennen wir meist nur die männliche Sprachform. Alle Formen schließen Personen jeden Geschlechts ein.

Alle Angaben in diesem Werk wurden sorgfältig geprüft. Dennoch können die Autoren und der Verlag keine Gewähr für deren Richtigkeit übernehmen. Der Verlag übernimmt trotz sorgfältiger inhaltlicher Kontrolle keine Haftung. Dies betrifft auch Inhalte genannter Internetseiten. Insbesondere hat der Verlag keinen Einfluss auf die dauerhafte Verfügbarkeit externer online-Daten und kann keinen dauerhaften Zugriff darauf gewährleisten.

Ein Markenzeichen kann markenrechtlich geschützt sein, auch wenn ein Hinweis auf etwa bestehende Schutzrechte fehlt.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <https://portal.dnb.de> abrufbar.

Jede Verwertung des Werkes außerhalb der Grenzen des Urheberrechtsgesetzes ist unzulässig und strafbar. Das gilt insbesondere für Übersetzungen, Nachdrucke, Mikroverfilmungen oder vergleichbare Verfahren sowie für die Speicherung in Datenverarbeitungsanlagen. Die Medien-

gruppe Deutscher Apotheker Verlag behält sich sämtliche Rechte an ihren Inhalten für kommerzielles Text- und Data-Mining im Sinne von § 44b UrhG ausdrücklich vor. Unsere Inhalte sind urheberrechtlich geschützt und dürfen nicht zur Entwicklung, zum Training und/oder zur Anreicherung oder Vergleichbarem von KI-Systemen, insbesondere von generativen KI-Systemen, verwendet werden. Ebenso ist die Erfassung und Speicherung der Inhalte durch digitale, fotomechanische oder durch ähnliche Verfahren zu den vorgenannten Zwecken untersagt. Für den Erwerb einer entsprechenden Nutzungslizenz kontaktieren Sie uns bitte unter [permissions@dav-medien.de](mailto:permissions@dav-medien.de).

4., aktualisierte und erweiterte Auflage 2026  
ISBN 978-3-8047-4539-1 (Print)  
ISBN 978-3-8047-4701-2 (E-Book)

© 2026 Wissenschaftliche Verlagsgesellschaft mbH  
Maybachstraße 8, 70469 Stuttgart  
[www.wissenschaftliche-verlagsgesellschaft.de](http://www.wissenschaftliche-verlagsgesellschaft.de)  
Printed in Poland

Programmplanung: Sonja Zikeli  
Lektorat: Julia Steinert  
Satz: abavo GmbH, Buchloe  
Grafiken: Dr. Ilse Zündorf  
Druck und Bindung: Drukarnia Dimograf, Bielsko-Biała  
Umschlagabbildung: [shiro\\_mofu/stock.adobe.com](https://shiro_mofu/stock.adobe.com)

## Vorwort

Seit 1999 – also seit mehr als 25 Jahren – fassen wir die Entwicklung der mit gentechnischen Methoden hergestellten Proteinarzneimittel (Biopharmazeutika) in dem Lehrbuch „Gentechnik – Biotechnik: Grundlagen und Wirkstoffe“ zusammen. Sie haben jetzt die 4. Auflage des Werks in den Händen. Sie werden vielleicht bemerkt haben, dass dieses Buch etwas umfangreicher geworden ist als die vorherige Auflage. Wir haben auch diesmal an dem Anspruch festgehalten, Ihnen alle zum Zeitpunkt der Drucklegung in Europa zugelassenen Biopharmazeutika vorzustellen. Seit der 3. Auflage hat sich der Biopharmazeutika-Markt schnell entwickelt. Waren Anfang 2018 „erst“ 242 Biopharmazeutika mit 190 Wirkstoffen zugelassen, waren es mit Stand August 2025 bereits 451 gentechnisch hergestellte Arzneimittel mit 397 Wirkstoffen. In der Zwischenzeit sind also ca. 150 Neuentwicklungen und ca. 50 Biosimilars dazugekommen, die wir in die 4. Auflage eingearbeitet haben. Wir haben auch ungefähr 30 Wirkstoffe gelöscht, vor allem, weil sie mittlerweile in Europa nicht mehr zugelassen sind.

Was ist in der 4. Auflage neu? Besonders entwickelt hat sich in den vergangenen 7 Jahren die Immuntherapie von Autoimmunerkrankungen und Tumorerkrankungen mit wichtigen Wirkstoffklassen wie den Immun-Checkpoint-Inhibitoren, CAR-T-Zellen und bispezifischen Antikörpern, die in der 4. Auflage wesentlich mehr Aufmerksamkeit erhalten haben. Natürlich dürfen auch die in der Corona-Pandemie neu entwickelten „genetischen“ Impfstoffe in Kapitel 17 nicht fehlen, und als gänzlich neue Indikation ist die Migräneprophylaxe (Kapitel 18) dazugekommen. Auch hat sich die Gentherapie in den letzten Jahren stark entwickelt und maßgeblich zur Bildung einer neuen Gruppe von Wirkstoffen beigetragen, die als „Arzneimittel für neuartige Therapien“ (*Advanced Therapy Medicinal Products*, ATMP) zusammengefasst werden. Diesen Wirkstoffen widmen wir nun ein ganzes Kapitel (Kapitel 9). Weiter ausgebaut haben wir das Kapitel 5 über Antikörper, auch weil zwischenzeitlich viele neue Antikörper-Designs entstanden sind und insbesondere die bereits erwähnten bispezifischen Antikörper eine immer größere Rolle spielen. Weiter aktualisiert haben wir auch das Kapitel 8 über Biosimilars, die als preisgünstigere Wirkstoffkopien der jeweiligen Originalia mittlerweile gut akzeptiert sind, deren Austausch gegen Originalia oder andere Biosimilars in der Praxis aber nicht ganz unproblematisch ist.

Wir haben uns entschieden, die bereits in den vorherigen Auflagen des Buches gewählte Zweiteilung in einen Methodenteil und einen Wirkstoffteil beizubehalten. Das war gar nicht einfach angesichts der vielen neu zu besprechenden Wirkstoffe, die ein erhebliches Maß an Platz in der neuen Auflage eingefordert haben.

Wir haben also, wie bereits auch in der 3. Auflage, in den Kapiteln 1–10 die eine oder andere gentechnische Methode entfernt, weil sie heute nicht mehr zeitgemäß ist. Natürlich war und ist uns bewusst, dass wir in einem Lehrbuch wie diesem nicht die ganze Breite der gentechnischen bzw. molekularbiologischen Methoden darstellen können – das ist aber auch nie unser Ziel gewesen. Dennoch halten wir es für wichtig, einige „Meilenstein-Methoden“ der Molekularbiologie weiterhin zu besprechen, damit Sie als geneigte Leserinnen und Leser sich auch weiterhin ein Bild davon machen können, wie komplex die molekularbiologische Forschung geworden ist und damit verbunden auch die Wirkstoffentwicklung, und wie man sich die Herstellung von „Wirt-Vektor-Systemen“ für die Produktion von Proteinwirkstoffen vorstellen kann und warum dafür der Einsatz sowohl von Gentechnik als auch von Biotechnik absolut unverzichtbar ist.

Der zweite Teil des Buches mit den Kapiteln 11–22 ist und bleibt ein Nachschlagewerk, in dem wir Ihnen möglichst alle derzeit zugelassenen Biopharmazeutika kurz zusammengefasst vorstellen. Wir legen dabei stets den Schwerpunkt auf die Architektur der Moleküle und ihre Wirkmechanismen. Jedem Indikationsgebiet stellen wir wieder einen kurzen Abriss der therapierten Krankheiten voran, ohne dabei den Anspruch zu erheben, die meist komplexen pathophysiologischen Zusammenhänge erschöpfend ausführen zu können. Hier sei auf die vielen hervorragenden Lehrbücher der Pharmakologie und Pathophysiologie verwiesen.

Natürlich ist es auch geboten, die zugelassenen Indikationen der einzelnen Arzneimittel aufzuführen. Da dies formale Texte sind, haben wir uns entschieden, die jeweiligen Formulierungen der Indikationsgebiete wörtlich aus den deutschen Produktinformationen der Europäischen Arzneimittel-Agentur (EMA) zu übernehmen. Die EMA erlaubt ausdrücklich die wörtliche Verwendung ihrer auf der Webseite publizierten Texte, sofern die EMA als Quelle genannt wird. Wir haben daher die Nennung der Indikationen jeweils mit dem Zusatz „Quelle: EMA; EPAR [Arzneimittelname]“ versehen. Dennoch können wir keine Gewähr auf die Vollständigkeit und Korrektheit der abgedruckten Indikationen oder eventuell angegebenen Dosierungen geben. Hier lohnt sich stets der Blick in die aktuellen Dokumente der EMA oder die Fachinformationen zu den jeweiligen Produkten.

Wir bedanken uns herzlich bei apl. Prof. Dr. Hanns-Christian Mahler für seinen Beitrag „Vom biotechnologischen Wirkstoff zum Arzneimittel“ (Kapitel 7), den wir unverändert aus der 3. Auflage weitergeführt haben.

Wir hoffen, dass wir Ihnen auch mit der erweiterten 4. Auflage des Buches „Gentechnik – Biotechnik:

Grundlagen und Wirkstoffe“ wieder eine interessante Lektüre bieten können, die Ihnen dabei hilft, nicht nur während des Pharmazie- oder Medizinstudiums, son-

dern auch im täglichen Berufsalltag diese komplexen Wirkstoffe besser zu verstehen.

Frankfurt am Main und Jena im Herbst 2025

Theodor Dingermann  
Thomas Winckler  
Ilse Zündorf

## Inhaltsverzeichnis

Vorwort	V	<b>4</b>	<b>Wirt-Vektor-Systeme zur Proteinherstellung</b>	<b>73</b>
Abkürzungsverzeichnis	XV			
<b>1 Die bio(techno)logische Revolution</b>	<b>1</b>	<b>4.1</b>	<b>Rekombinante Expressionseinheiten</b>	<b>73</b>
1.1 Von der klassischen Genetik zur Gentechnik	1	4.1.1	Informationen über Gene und Proteine: Wissenschaftliche Datenbanken	73
1.2 Gentechnik und Biotechnik	6	4.1.2	Isolierung genetischer Information	74
<b>2 DNA-Rekombinationstechnologie</b>	<b>9</b>	4.1.3	Auswahl geeigneter Produktionszellen	76
2.1 Die DNA	9	<b>4.2</b>	<b>Proteinherstellung in Bakterien</b>	<b>81</b>
2.1.1 Aufbau der DNA	9	4.2.1	Steuerung der Transkription von Transgenen	81
2.1.2 Denaturierung von DNA	11	4.2.2	Optimierung der Translation von Transgenen	85
2.1.3 Größe und Topologie der DNA	12	4.2.3	Sekretion und <i>N</i> -Glycosylierung von Proteinen durch Bakterien	85
2.1.4 DNA-Analytik	14	4.2.4	Fusionssysteme zur effizienten Produktreinigung	86
2.1.5 Chemische DNA-Synthese	17	4.2.5	Plasmidstabilisierung in Produktionsstämmen	87
2.2 Enzymatische DNA-Synthese	18	<b>4.3</b>	<b>Proteinherstellung in Hefen</b>	<b>90</b>
2.2.1 DNA-Polymerasen	18	4.3.1	Vektorsysteme für die Proteinherstellung in Hefen	91
2.2.2 Polymerase-Kettenreaktion (PCR)	19	4.3.2	Transformation von Hefezellen	92
2.2.3 Variationen der PCR	24	4.3.3	Heterologe Expression von Proteinen in Hefen	94
2.2.4 DNA-Sequenzierung	27	<b>4.4</b>	<b>Proteinherstellung in Insekten- Zellkulturen</b>	<b>97</b>
2.3 Prinzipien der Gentechnik	34	4.4.1	Baculovirus als Gentransfersystem	97
2.3.1 Gezieltes Schneiden von DNA: Restriktionsendonukleasen	35	4.4.2	Transfektion von Insekten- Zellkulturzellen	99
2.3.2 Rekombinieren von DNA-Fragmenten: DNA-Ligasen	37	<b>4.5</b>	<b>Transgene Pflanzen</b>	<b>99</b>
2.3.3 Amplifizieren rekombinierter DNA: Transformation von Bakterien	38	4.5.1	Transformationsmethoden	100
<b>3 Grundsätze der Proteinexpression</b>	<b>44</b>	4.5.2	Selektionsmarker zur Etablierung transgener Pflanzen	104
3.1 Umsetzung genetischer Information in Proteine	44	4.5.3	Regeneration transgener Pflanzen	105
3.2 Genexpression in Prokaryonten	46	4.5.4	Verhinderung der Ausbreitung transgener Pflanzen	105
3.2.1 Transkription	46	4.5.5	Manipulation von Sekundärstoffmustern in Medizinalpflanzen	107
3.2.2 Translation	48	4.5.6	Produktion von Naturstoffen in Pflanzen-Zellkulturen	107
3.3 Genexpression in Eukaryonten	54	4.5.7	Proteinexpression in transgenen Pflanzen-Zellkulturen	109
3.4 Posttranslationale Modifikationen an Proteinen	57	4.5.8	Proteinexpression in Pflanzen	109
3.4.1 Proteolytische Prozessierung von Proteinen	58			
3.4.2 Glycosylierung von Proteinen	60			
3.4.3 Glycoengineering	66			

<b>4.6</b>	<b>Proteinherstellung in Säugetier-Zellkulturen</b> .....	<b>111</b>	<b>6.2</b>	<b>Fermentation</b> .....	<b>164</b>
4.6.1	Transiente versus stabile Expression .....	112	6.2.1	Zellkulturmedien .....	164
4.6.2	Methoden zum Einführen von DNA in Säugetierzellen .....	112	6.2.2	Wachstumskinetik und Produktbildung ...	166
4.6.3	Selektionsmarker zur Identifizierung transfizierter Säugetierzellen .....	115	6.2.3	Fermentationsverfahren .....	168
4.6.4	Regulation heterologer Genexpression in Säugetierzellen .....	115	<b>6.3</b>	<b>Produkt-Aufarbeitung</b> .....	<b>169</b>
4.6.5	Produktionszelllinien .....	117	6.3.1	Zellernte .....	169
<b>4.7</b>	<b>Genetisch veränderte Säugetiere</b> .....	<b>121</b>	6.3.2	Zellaufschluss .....	171
4.7.1	Transgene Tiere .....	122	6.3.3	Produkt-Reinigung .....	172
4.7.2	Knockout-Mäuse .....	123	<b>7</b>	<b>Vom biotechnologischen Wirkstoff zum Arzneimittel</b> .....	<b>183</b>
4.7.3	Wirkstoffproduktion in transgenen Säugetieren .....	127	<b>7.1</b>	<b>Anforderungen der Arzneibücher an sterile Arzneimittel mit biotechnologischen Wirkstoffen</b> .....	<b>183</b>
<b>5</b>	<b>Monoklonale Antikörper</b> .....	<b>130</b>	<b>7.2</b>	<b>Instabilitätsreaktionen und Formulierung biotechnologischer Arzneistoffe</b> .....	<b>184</b>
<b>5.1</b>	<b>Struktur von Antikörpern</b> .....	<b>130</b>	7.2.1	Instabilitätsreaktionen .....	184
<b>5.2</b>	<b>Hybridom-Technologie</b> .....	<b>133</b>	7.2.2	Formulierung und Darreichungsform .....	186
5.2.1	Immunisierung von Mäusen .....	133	7.2.3	Produktstabilität .....	187
5.2.2	Fusion von Milzzellen mit Myelomzellen .....	134	<b>7.3</b>	<b>Besonderheiten biotechnologischer Wirkstoffe bei Herstellung, Transport, Lagerung und Anwendung</b> .....	<b>188</b>
5.2.3	Selektion der fusionierten Hybridomzellen .....	134	7.3.1	Herstellung steriler Fertigprodukte von Biopharmazeutika .....	188
5.2.4	Etablierung der antikörperproduzierenden Hybridomzellen .....	136	7.3.2	Transport .....	193
<b>5.3</b>	<b>Humanisierte Antikörper</b> .....	<b>138</b>	7.3.3	Lagerung .....	194
<b>5.4</b>	<b>Vollständig humane Antikörper</b> .....	<b>140</b>	7.3.4	Anwendung .....	194
5.4.1	Phage-Display-Technologie .....	141	<b>7.4</b>	<b>Analytische Untersuchungen am Fertigprodukt</b> .....	<b>196</b>
5.4.2	HuCal <sup>®</sup> -System .....	143	7.4.1	Analytische Methoden .....	196
5.4.3	Transgene Mäuse zur Herstellung humaner Antikörper .....	144	7.4.2	Laufzeit, Lagerbedingungen, Lagerhinweise .....	197
5.4.4	Gewinnung humaner monoklonaler Antikörper von Patienten .....	147	<b>7.5</b>	<b>Die Anforderungen des Europäischen Arzneibuchs an gentechnisch hergestellte Produkte</b> .....	<b>198</b>
<b>5.5</b>	<b>Antikörperfragmente</b> .....	<b>148</b>	7.5.1	Definition .....	198
<b>5.6</b>	<b>Nanoantikörper</b> .....	<b>149</b>	7.5.2	Herstellung .....	198
<b>5.7</b>	<b>Bispezifische Antikörper</b> .....	<b>151</b>	7.5.3	Wirt-Vektor-System .....	198
<b>5.8</b>	<b>Namensgebung bei Antikörper-Wirkstoffen</b> .....	<b>153</b>	7.5.4	Zellbanksysteme: Validierung und Kontrolle .....	200
<b>5.9</b>	<b>Fusionsproteine mit Fc-Teilen von Antikörpern</b> .....	<b>156</b>	7.5.5	Validierung des Herstellungsprozesses .....	201
<b>5.10</b>	<b>Chimäre Antigenrezeptoren</b> .....	<b>158</b>	7.5.6	Identitätsprüfung .....	202
<b>6</b>	<b>Biotechnische Produktion von Proteinwirkstoffen</b> .....	<b>161</b>	7.5.7	Reinheitsprüfung .....	202
<b>6.1</b>	<b>Bioreaktoren</b> .....	<b>161</b>	7.5.8	Gehalt und Wirksamkeit .....	203
			7.5.9	Biologische Wirksamkeit .....	203
			7.5.10	Struktur .....	203
			7.5.11	Weitere Methoden .....	203

<b>7.6</b>	<b>Die Europäische Zulassungsbehörde (EMA) .....</b>	<b>204</b>	<b>10</b>	<b>Genfunktionsanalyse, Gendiagnostik und stratifizierte Pharmakotherapie .....</b>	<b>254</b>
7.6.1	Aufgabe .....	204	<b>10.1</b>	<b>Genomprojekte .....</b>	<b>254</b>
7.6.2	Europäischer Öffentlicher Beurteilungsbericht .....	204	10.1.1	Durchführung von Genomprojekten .....	255
7.6.3	Der Orphan-Drug-Status .....	204	10.1.2	Erkenntnisse aus Genomprojekten .....	257
<b>8</b>	<b>Biosimilars .....</b>	<b>206</b>	10.1.3	Das Genom des Menschen .....	259
<b>8.1</b>	<b>Einführung .....</b>	<b>206</b>	10.1.4	Das individuelle menschliche Genom .....	261
<b>8.2</b>	<b>Definition .....</b>	<b>206</b>	<b>10.2</b>	<b>Genfunktionsanalyse .....</b>	<b>263</b>
<b>8.3</b>	<b>Zulassung eines Biosimilars .....</b>	<b>207</b>	10.2.1	Genetische Komplementation .....	263
<b>8.4</b>	<b>Die EMA als zentrales Kompetenzzentrum für Biopharmazeutika .....</b>	<b>208</b>	10.2.2	Genexpressionsanalyse .....	264
<b>8.5</b>	<b>Der Nachweis von Qualität, Wirksamkeit und Sicherheit .....</b>	<b>208</b>	10.2.3	Proteomik .....	266
<b>8.6</b>	<b>Eindeutige Identifizierung des Arzneimittels .....</b>	<b>210</b>	10.2.4	Reportergene .....	269
<b>8.7</b>	<b>Extrapolation der Indikationen .....</b>	<b>210</b>	10.2.5	Modellorganismen .....	274
<b>8.8</b>	<b>Austausch von Biosimilars (Switching) ...</b>	<b>210</b>	10.2.6	Mausmodelle .....	276
<b>9</b>	<b>Arzneimittel für neuartige Therapien (ATMP) .....</b>	<b>214</b>	<b>10.3</b>	<b>Posttranskriptionelle Genregulation .....</b>	<b>277</b>
<b>9.1</b>	<b>Gentherapie .....</b>	<b>214</b>	10.3.1	Regulation der Genexpression durch Antisense-Oligonukleotide .....	277
9.1.1	Konzept .....	214	10.3.2	Regulation der Genexpression durch siRNA .....	278
9.1.2	Nichtvirale Gentransfer-Vektoren .....	218	10.3.3	RNA-basierte Wirkstoffe .....	281
9.1.3	Retrovirale Gentransfer-Vektoren .....	219	<b>10.4</b>	<b>Diagnostik von Genmutationen .....</b>	<b>284</b>
9.1.4	Lentivirale Gentransfer-Vektoren .....	224	10.4.1	Molekulargenetische Diagnostik .....	284
9.1.5	Adenovirale Gentransfer-Vektoren .....	226	10.4.2	SNP-Bestimmung über Restriktionsfragmentlängenpolymorphismen .....	284
9.1.6	Adenoassoziierte virale Gentransfer-Vektoren .....	228	10.4.3	Allelspezifische PCR .....	285
9.1.7	Onkolytische Viren .....	230	10.4.4	SNP-Analyse durch DNA-Sequenzierung ..	286
9.1.8	Namensgebung bei Gentherapeutika .....	232	10.4.5	Massenspektrometrische SNP-Analytik .....	286
<b>9.2</b>	<b>Genreparatur .....</b>	<b>232</b>	<b>10.5</b>	<b>Genomweite Gendiagnostik .....</b>	<b>288</b>
9.2.1	Zelluläre Reparatur von DNA-Doppelstrangbrüchen .....	233	10.5.1	Zytogenetische Diagnostik .....	288
9.2.2	Zinkfinger-Nukleasen und TALENs .....	234	10.5.2	DNA-Microarrays .....	291
9.2.3	CRISPR/Cas-Systeme .....	235	10.5.3	Genomweite Assoziationsstudien .....	292
9.2.4	Therapeutische Anwendungen der Genom-Editierung mit CRISPR/Cas9 .....	243	10.5.4	Vererbungsregeln .....	294
9.2.5	Das Gene-drive-Konzept .....	245	<b>10.6</b>	<b>Stratifizierte Pharmakotherapie .....</b>	<b>296</b>
<b>9.3</b>	<b>Stammzelltherapie .....</b>	<b>246</b>	10.6.1	Arzneistoffmetabolisierende Enzyme .....	297
9.3.1	Adulte humane Stammzellen .....	247	10.6.2	Phänotypisierung .....	298
9.3.2	Humane embryonale Stammzellen .....	248	10.6.3	Genotypisierung .....	299
9.3.3	Somatischer Kerntransfer zum Klonen adulter Zellen .....	249	10.6.4	Beispiele für pharmakogenetische Interaktionen .....	300
9.3.4	Induzierte pluripotente Stammzellen .....	252	<b>11</b>	<b>Blutbildungs- und Blutgerinnungsstörungen .....</b>	<b>307</b>
			<b>11.1</b>	<b>Therapie der renalen Anämie .....</b>	<b>307</b>
			11.1.1	Erythropoetin .....	307
			11.1.2	Struktur des menschlichen Erythropoetins .....	309
			11.1.3	Erythropoetin-Rezeptor .....	310
			11.1.4	Erythropoetin-Arzneimittel .....	311

<b>11.2</b>	<b>Therapie der Polycythaemia vera .....</b>	<b>315</b>	<b>13</b>	<b>Augenkrankheiten .....</b>	<b>383</b>
<b>11.3</b>	<b>Therapie von Hämoglobinopathien .....</b>	<b>317</b>	<b>13.1</b>	<b>Therapie der Altersabhängigen Makuladegeneration .....</b>	<b>383</b>
11.3.1	Formen von Hämoglobinopathien .....	317	13.1.1	Krankheitsbild .....	383
11.3.2	Therapie von Hämoglobinopathien .....	318	13.1.2	Vaskulärer endothelialer Wachstumsfaktor (VEGF) .....	384
11.3.3	Gentherapeutische Ansätze bei Hämoglobinopathien .....	321	13.1.3	Angiopoietin-2 (ANG2) .....	384
<b>11.4</b>	<b>Therapie und Prophylaxe von Blutgerinnungsstörungen .....</b>	<b>323</b>	13.1.4	Angiogenese-Hemmer .....	385
11.4.1	Blutgerinnungskaskade .....	323	<b>13.2</b>	<b>Therapie der vitreomakulären Traktion .....</b>	<b>390</b>
11.4.2	von-Willebrand-Faktor .....	326	13.2.1	Krankheitsbild .....	390
11.4.3	Arzneimittel zur Substitution des von-Willebrand-Faktors .....	328	13.2.2	Therapieoptionen .....	390
11.4.4	Blutgerinnungsfaktor VIII .....	328	<b>13.3</b>	<b>Therapie der neurotrophen Keratitis .....</b>	<b>392</b>
11.4.5	Arzneimittel zur Substitution von FVIII .....	332	13.3.1	Nervenwachstumsfaktor .....	392
11.4.6	Blutgerinnungsfaktor IX .....	342	13.3.2	Therapeutische Verwendung von NGF .....	393
11.4.7	Arzneimittel zur Substitution von FIX .....	344	<b>13.4</b>	<b>Therapie von erblichen Netzhautdystrophien .....</b>	<b>394</b>
11.4.8	Blutgerinnungsfaktor VII .....	348	<b>14</b>	<b>Autoimmunerkrankungen .....</b>	<b>397</b>
11.4.9	Arzneimittel zur Substitution von FVII .....	349	<b>14.1</b>	<b>Therapie von rheumatischen Erkrankungen und Kollagenosen .....</b>	<b>397</b>
11.4.10	Blutgerinnungsfaktor XIII .....	351	14.1.1	Rheumatoide Arthritis .....	397
11.4.11	Arzneimittel zur Substitution von FXIII .....	352	14.1.2	Juvenile idiopathische Arthritis .....	399
<b>11.5</b>	<b>Therapie und Prophylaxe von Thrombosen .....</b>	<b>352</b>	14.1.3	Ankylosierende Spondylitis .....	399
11.5.1	Regulation der Thrombozytenaggregation .....	354	14.1.4	Psoriasis-Arthritis .....	400
11.5.2	GPIIb/IIIa-Rezeptor .....	355	14.1.5	Systemischer Lupus erythematodes .....	400
11.5.3	Thrombozytenaggregationshemmer .....	356	14.1.6	Tumornekrosefaktor .....	401
11.5.4	Hemmung der plasmatischen Blutgerinnung .....	357	14.1.7	Neutralisierung von TNF .....	403
11.5.5	Regulation der Fibrinolyse .....	358	14.1.8	Interleukin-1 $\beta$ .....	413
11.5.6	Gewebe-Plasminogenaktivator .....	360	14.1.9	Neutralisierung von IL-1 $\beta$ .....	414
11.5.7	Rekombinante Fibrinolytika .....	362	14.1.10	Interleukin-6 .....	415
<b>11.6</b>	<b>Antidote gegen direkte orale Antikoagulanzen .....</b>	<b>365</b>	14.1.11	Neutralisierung von IL-6 .....	415
11.6.1	Neutralisation direkter Thrombin-Inhibitoren .....	365	14.1.12	Neutralisierung von IL-17A .....	418
11.6.2	Neutralisation direkter FXa-Inhibitoren .....	366	14.1.13	Reduktion der T-Zell-Aktivierung .....	418
<b>12</b>	<b>Allergische Entzündungsreaktionen .....</b>	<b>369</b>	14.1.14	Depletion von B-Zellen .....	421
<b>12.1</b>	<b>Therapie des allergischen Asthmas .....</b>	<b>369</b>	14.1.15	B-Zell-aktivierender Faktor (BLys) .....	422
12.1.1	Asthma bronchiale .....	369	14.1.16	Inhibition von Typ-I-Interferonen .....	423
12.1.2	Senkung des IgE-Spiegels .....	371	<b>14.2</b>	<b>Therapie der Plaque-Psoriasis .....</b>	<b>424</b>
12.1.3	Inhibition von IL-4/13-Wirkungen .....	374	14.2.1	Neutralisierung von TNF .....	425
12.1.4	Senkung des IL-5-Spiegels .....	375	14.2.2	Interleukine IL-12 und IL-23 .....	425
12.1.5	Senkung des TSLP-Spiegels .....	379	14.2.3	Neutralisierung von IL-12 und IL-23 .....	426
<b>12.2</b>	<b>Therapie der atopischen Dermatitis .....</b>	<b>380</b>	14.2.4	Interleukin-17 .....	429
12.2.1	Atopische Dermatitis .....	380	14.2.5	Neutralisierung von IL-17 .....	430
			<b>14.3</b>	<b>Therapie der generalisierten pustulösen Psoriasis .....</b>	<b>434</b>
			14.3.1	Interleukin-36 .....	434
			14.3.2	Generalisierte pustulöse Psoriasis .....	435
			14.3.3	Neutralisation von IL-36-Wirkungen .....	435

<b>14.4</b>	<b>Therapie entzündlicher Darmerkrankungen</b> .....	<b>436</b>	15.2.13	Lysosomale $\alpha$ -Mannosidase .....	489
14.4.1	Neutralisierung von TNF .....	437	15.2.14	Saure Lipase A .....	489
14.4.2	Neutralisierung von IL-12 und IL-23 .....	437	15.2.15	Tripeptidyl-Peptidase 1 .....	491
14.4.3	Blockade von $\alpha_4\beta_7$ -Integrin .....	438	<b>15.3</b>	<b>Therapie der Phenylketonurie</b> .....	<b>493</b>
<b>14.5</b>	<b>Therapie der multiplen Sklerose</b> .....	<b>439</b>	15.3.1	Krankheitsbild .....	493
14.5.1	Klinik .....	440	15.3.2	Phenylalanin-Ammoniak-Lyase .....	493
14.5.2	Immunopathogenese .....	443	<b>15.4</b>	<b>Therapie der Hyperargininämie</b> .....	<b>495</b>
14.5.3	Therapeutische Konzepte .....	444	15.4.1	Krankheitsbild .....	495
14.5.4	Interferon beta .....	444	<b>15.5</b>	<b>Therapie des Mangels an Aromatischer L-Aminosäure-Decarboxylase</b> .....	<b>496</b>
14.5.5	Glatirameracetat .....	448	15.5.1	Krankheitsbild .....	496
14.5.6	Elimination von B- und T-Zellen .....	450	15.5.2	Gentherapie der AADCDC .....	496
14.5.7	Blockade von $\alpha_4$ -Integrinen .....	451	<b>15.6</b>	<b>Therapie der Cryopyrin-assoziierten periodischen Syndrome (CAPS)</b> .....	<b>497</b>
14.5.8	Depletion von B-Zellen .....	454	15.6.1	Familiäres kälteinduziertes auto- inflammatorisches Syndrom (FCAS) .....	497
<b>14.6</b>	<b>Therapie von Neuromyelitis-optica- Spektrum-Erkrankungen</b> .....	<b>456</b>	15.6.2	Muckle-Wells-Syndrom (MWS) .....	497
14.6.1	Krankheitsbild .....	456	15.6.3	Neonatal-onset multisystem inflammatory disease (NOMID) .....	498
14.6.2	Arzneimittel .....	457	15.6.4	Inhibition von IL-1-Wirkungen .....	498
<b>14.7</b>	<b>Therapie der Immunthrombozytopenie</b> ..	<b>458</b>	<b>15.7</b>	<b>Therapie der paroxysmalen nächtlichen Hämoglobinurie</b> .....	<b>500</b>
14.7.1	Pathophysiologie .....	458	15.7.1	Krankheitsbild .....	500
14.7.2	Thrombopoetin .....	459	15.7.2	Komplementsystem .....	501
<b>14.8</b>	<b>Therapie der Myasthenia gravis</b> .....	<b>461</b>	15.7.3	Blockade des Komplementsystems .....	503
14.8.1	Krankheitsbild .....	461	<b>15.8</b>	<b>Therapie des hereditären Angioödems</b> ..	<b>505</b>
14.8.2	Arzneimittel .....	462	15.8.1	Krankheitsbild .....	505
<b>14.9</b>	<b>Therapie der primären Kälteagglutinin-Krankheit</b> .....	<b>463</b>	15.8.2	C1-Inhibitor .....	505
14.9.1	Krankheitsbild .....	463	15.8.3	Kallikrein .....	507
14.9.2	Arzneimittel .....	464	<b>15.9</b>	<b>Therapie der Hypophosphatasie</b> .....	<b>508</b>
<b>15</b>	<b>Erbkrankheiten</b> .....	<b>465</b>	15.9.1	Krankheitsbild .....	508
<b>15.1</b>	<b>Therapie der Mukoviszidose</b> .....	<b>465</b>	15.9.2	Substitution der TNSALP .....	508
15.1.1	Krankheitsbild .....	465	<b>15.10</b>	<b>Therapie der X-chromosomalen Hypophosphatämie</b> .....	<b>509</b>
15.1.2	Chloridkanal CFTR .....	466	15.10.1	Krankheitsbild .....	509
15.1.3	Desoxyribonuklease I .....	467	15.10.2	Fibroblasten-Wachstumsfaktor 23 .....	509
<b>15.2</b>	<b>Therapie lysosomaler Speicherkrankheiten</b> .....	<b>468</b>	<b>15.11</b>	<b>Therapie der schweren kombinierten Immundefizienz</b> .....	<b>510</b>
15.2.1	Das lysosomale Kompartiment .....	469	15.11.1	Genetische Ursachen der ADA-SCID .....	510
15.2.2	Mannose-6-phosphat-Rezeptoren .....	469	15.11.2	Therapieoptionen .....	510
15.2.3	$\beta$ -Glucocerebrosidase .....	471	<b>15.12</b>	<b>Therapie der spinalen Muskelatrophie</b> ..	<b>511</b>
15.2.4	$\alpha$ -Galactosidase A .....	475	15.12.1	Krankheitsbild .....	511
15.2.5	Saure Sphingomyelinase .....	478	15.12.2	Molekularbiologische Grundlagen der SMA .....	511
15.2.6	Arylsulfatase A .....	478	15.12.3	Korrektur der SMN-Expression .....	513
15.2.7	$\alpha$ -L-Iduronidase .....	480	15.12.4	Substitution des SMN1-Proteins .....	514
15.2.8	Iduronat-2-Sulfatase .....	482			
15.2.9	N-Acetylgalactosamin-6-Sulfatase .....	483			
15.2.10	N-Acetylgalactosamin-4-Sulfatase .....	484			
15.2.11	$\beta$ -Glucuronidase .....	485			
15.2.12	Saure $\alpha$ -Glucosidase .....	486			

<b>15.13</b>	<b>Therapie der chronischen Granulomatose</b> .....	<b>515</b>	<b>17.5</b>	<b>Impfung gegen Hepatitis B</b> .....	<b>554</b>
15.13.1	Genetische Ursachen .....	515	17.5.1	Impfantigen aus Hepatitis-B-Viren .....	554
15.13.2	Interferon gamma .....	516	17.5.2	Impfstoffe gegen Hepatitis B .....	555
<b>15.14</b>	<b>Therapie der malignen Osteopetrose</b> .....	<b>517</b>	<b>17.6</b>	<b>Impfung gegen Gebärmutterhalskrebs</b> .....	<b>563</b>
15.14.1	Genetische Ursachen .....	517	17.6.1	Humane Papillomviren .....	563
<b>15.15</b>	<b>Therapie der hereditären Transthyretin-vermittelten Amyloidose</b> .....	<b>518</b>	17.6.2	Impfstoffe gegen humane Papillomviren .....	563
15.15.1	Krankheitsbild .....	518	<b>17.7</b>	<b>Impfung gegen Influenza</b> .....	<b>567</b>
15.15.2	Reduktion der TTR-Expression .....	518	17.7.1	Influenza-Viren .....	567
<b>15.16</b>	<b>Therapie der akuten hepatischen Porphyrie</b> .....	<b>521</b>	17.7.2	Gentechnisch hergestellte Impfstoffe gegen saisonale Influenza-Viren .....	568
15.16.1	Krankheitsbild .....	521	17.7.3	Gentechnisch hergestellte Impfstoffe gegen pandemische Influenza-Viren .....	571
15.16.2	Senkung der ALAS1-Expression .....	522	<b>17.8</b>	<b>Impfung gegen Herpes Zoster</b> .....	<b>572</b>
<b>15.17</b>	<b>Therapie der erworbenen thrombotisch-thrombozytopenischen Purpura</b> .....	<b>523</b>	17.8.1	Varizella-Zoster-Virus .....	572
15.17.1	Genetische Ursachen .....	523	17.8.2	Windpocken und Herpes Zoster .....	572
<b>15.18</b>	<b>Therapie der primären Hyperoxalurie Typ I</b> .....	<b>524</b>	17.8.3	Impfstoffe gegen das Varizella-Zoster-Virus .....	572
15.18.1	Krankheitsbild .....	524	<b>17.9</b>	<b>Impfung gegen Denguefieber</b> .....	<b>574</b>
15.18.2	Reduktion der Glyoxylat-Bildung .....	525	17.9.1	Denguevirus .....	574
<b>16</b>	<b>Fertilisationsstörungen</b> .....	<b>527</b>	17.9.2	Denguefieber .....	575
<b>16.1</b>	<b>Regulation des ovariellen Zyklus</b> .....	<b>527</b>	17.9.3	Impfstoff gegen Denguefieber .....	576
16.1.1	Gonadotropin-Releasing-Hormon .....	529	<b>17.10</b>	<b>Prävention von Infektionen mit dem Respiratorischen Synzytialvirus</b> .....	<b>577</b>
16.1.2	Follikelstimulierendes Hormon .....	529	17.10.1	Respiratorisches Synzytialvirus .....	577
16.1.3	Luteinisierendes Hormon .....	530	17.10.2	Passive Immunisierung gegen RSV .....	579
16.1.4	Choriongonadotropin .....	531	17.10.3	Aktive Immunisierung gegen RSV .....	580
<b>16.2</b>	<b>Assistierte Reproduktion</b> .....	<b>531</b>	<b>17.11</b>	<b>Impfung gegen Ebolafieber</b> .....	<b>581</b>
16.2.1	Rekombinante Wirkstoffe für die hormonelle Hyperstimulation .....	533	17.11.1	Das Ebolavirus .....	581
<b>17</b>	<b>Infektionskrankheiten</b> .....	<b>540</b>	17.11.2	Impfstoffe gegen Ebolafieber .....	581
<b>17.1</b>	<b>Impfstofftypen</b> .....	<b>540</b>	<b>17.12</b>	<b>Impfung gegen COVID-19</b> .....	<b>585</b>
<b>17.2</b>	<b>Adjuvanzien in Impfstoffen</b> .....	<b>542</b>	17.12.1	Humanpathogene Coronaviren .....	585
<b>17.3</b>	<b>Interferone</b> .....	<b>543</b>	17.12.2	SARS-CoV-2 .....	586
17.3.1	Einteilung der Interferone .....	543	17.12.3	mRNA-basierte Impfstoffe .....	587
17.3.2	Interferon-Rezeptoren .....	543	17.12.4	Vektorbasierte Impfstoffe .....	590
17.3.3	Wirkungen von Interferonen .....	546	17.12.5	Impfstoffe gegen SARS-CoV-2 .....	591
<b>17.4</b>	<b>Therapie von Virushepatitiden</b> .....	<b>548</b>	17.12.6	Monoklonale Antikörper gegen SARS-CoV-2 .....	596
17.4.1	Hepatitis-B-Virus .....	548	<b>17.13</b>	<b>Impfung gegen Pneumokokken</b> .....	<b>598</b>
17.4.2	Hepatitis-C-Virus .....	550	17.13.1	<i>Streptococcus pneumoniae</i> .....	598
17.4.3	Alpha-Interferon als Wirkstoff .....	551	17.13.2	Impfstoffe gegen Pneumokokken .....	599
17.4.4	Alpha-Interferon-Arzneimittel .....	552	<b>17.14</b>	<b>Impfung gegen Meningokokken</b> .....	<b>601</b>
			17.14.1	<i>Neisseria meningitidis</i> .....	601
			17.14.2	Impfstoffe gegen Meningokokken .....	602

<b>17.15</b>	<b>Impfung gegen Cholera</b> .....	<b>603</b>	<b>20.2</b>	<b>Therapie des Kurzdarmsyndroms</b> .....	<b>652</b>
17.15.1	Vibrio cholerae .....	603	20.2.1	Kurzdarmsyndrom .....	652
17.15.2	Cholera-Toxin .....	603	20.2.2	Glucagon-like peptide-2 .....	652
17.15.3	Impfstoffe gegen Cholera-Bakterien .....	604	<b>20.3</b>	<b>Therapie der Hypercholesterinämie</b> .....	<b>653</b>
<b>17.16</b>	<b>Prävention von Clostridioides-difficile-Infektionen</b> .....	<b>605</b>	20.3.1	Krankheitsbild .....	653
17.16.1	<i>Clostridioides difficile</i> .....	605	20.3.2	PCSK9 .....	655
17.16.2	Clostridioides-difficile-Antitoxin .....	606	20.3.3	Wirkstoffe zur Senkung der PCSK9-Aktivität .....	655
<b>17.17</b>	<b>Therapie von Milzbrand</b> .....	<b>607</b>	<b>20.4</b>	<b>Therapie des Familiären Chylomikronämie-Syndroms</b> .....	<b>658</b>
17.17.1	Milzbrand (Anthrax) .....	607	20.4.1	Apolipoprotein C-III .....	658
17.17.2	Milzbrand-Antitoxin .....	608	<b>20.5</b>	<b>Therapie der Lipodystrophie</b> .....	<b>659</b>
<b>18</b>	<b>Migräne</b> .....	<b>609</b>	20.5.1	Rekombinantes Leptin .....	660
<b>18.1</b>	<b>Krankheitsbild</b> .....	<b>609</b>	<b>20.6</b>	<b>Therapie der Osteoporose</b> .....	<b>661</b>
<b>18.2</b>	<b>Calcitonin gene-related peptide (CGRP) ...</b>	<b>610</b>	20.6.1	Regulation der Mineralisierung des Knochens .....	661
18.2.1	CGRP-Rezeptor .....	611	20.6.2	Osteoporose .....	662
18.2.2	Arzneimittel für die Migräneprophylaxe .....	611	20.6.3	Parathormon .....	663
<b>19</b>	<b>Organtransplantation</b> .....	<b>614</b>	20.6.4	RANK-Ligand .....	665
<b>19.1</b>	<b>Ursache von Gewebe-Unverträglichkeit</b> ....	<b>614</b>	20.6.5	Sclerostin .....	666
<b>19.2</b>	<b>T-Zellen und Abstoßungsreaktion</b> .....	<b>615</b>	<b>20.7</b>	<b>Therapie von Knochenheilungsstörungen</b> .....	<b>667</b>
<b>19.3</b>	<b>Reduktion der Aktivität von T-Zellen</b> .....	<b>619</b>	20.7.1	Bone morphogenetic proteins (BMPs) .....	667
19.3.1	Depletion aktivierter T-Zellen .....	619	20.7.2	Rezeptoren und Aktivierung .....	668
19.3.2	Reduktion der T-Zell-Aktivierung .....	620	<b>20.8</b>	<b>Therapie des Hypoparathyreoidismus</b> ....	<b>670</b>
<b>19.4</b>	<b>Reduktion von donorspezifischen Antikörpern</b> .....	<b>621</b>	20.8.1	Parathormon .....	671
<b>20</b>	<b>Stoffwechselstörungen</b> .....	<b>624</b>	<b>21</b>	<b>Tumorerkrankungen</b> .....	<b>672</b>
<b>20.1</b>	<b>Therapie des Diabetes mellitus</b> .....	<b>624</b>	<b>21.1</b>	<b>Tumorentstehung</b> .....	<b>672</b>
20.1.1	Insulin .....	626	21.1.1	Onkogene .....	673
20.1.2	Insulinrezeptor .....	626	21.1.2	Tumorsuppressor-Gene .....	678
20.1.3	Regulation der Insulinfreisetzung .....	627	21.1.3	MicroRNA .....	678
20.1.4	Konsequenzen eines Insulinmangels .....	628	<b>21.2</b>	<b>Zytokine als Tumorthapeutika</b> .....	<b>679</b>
20.1.5	Tierisches Insulin als Grundlage von Insulin-Arzneimitteln .....	629	21.2.1	Interleukin-2 .....	680
20.1.6	Biochemische Humanisierung von Schweineinsulin .....	629	21.2.2	Tumornekrosefaktor .....	681
20.1.7	Gentechnische Herstellung von Humaninsulin .....	631	21.2.3	Interferon alfa-2a .....	683
20.1.8	Produktion von Humaninsulin in Bakterien .....	631	<b>21.3</b>	<b>Anti-Zytokin-Strategien in der Tumorthherapie</b> .....	<b>684</b>
20.1.9	Produktion von Humaninsulin in Hefen ...	632	21.3.1	Castleman-Krankheit .....	684
20.1.10	Gentechnische Herstellung von Humaninsulin-Analoga .....	635	21.3.2	Neutralisierung von IL-6 .....	684
20.1.11	Schnellwirksame Humaninsulin-Analoga ..	637	<b>21.4</b>	<b>Zelloberflächenmarker als Angriffspunkte für Tumortheraeutika</b> .....	<b>685</b>
20.1.12	Langwirksame Humaninsulin-Analoga ....	640	21.4.1	Zielstruktur HER1 .....	686
20.1.13	Glucagon .....	644	21.4.2	Zielstruktur HER2 .....	690
20.1.14	Inkretine .....	645	21.4.3	Zielstruktur Trop-2 .....	695
			21.4.4	Zielstruktur RANK-Ligand .....	696

<b>21.5</b>	<b>Zellen markierende Wirkstoffe .....</b>	<b>697</b>	<b>21.11</b>	<b>Diagnostikum nach einer Thyreoidektomie .....</b>	<b>760</b>
21.5.1	Zielstruktur SLAMF7 .....	697	21.11.1	Schilddrüsenkarzinom .....	760
21.5.2	Zielstruktur Gangliosid GD2 .....	698	21.11.2	Thyreotropin (TSH) .....	761
21.5.3	Zielstruktur CD19 .....	699	<b>22</b>	<b>Therapie von Wachstumsstörungen .....</b>	<b>763</b>
21.5.4	Zielstruktur CD20 .....	704	<b>22.1</b>	<b>Achondroplasia .....</b>	<b>763</b>
21.5.5	Zielstruktur CD22 .....	711	22.1.1	Genetische Ursachen .....	763
21.5.6	Zielstruktur CD30 .....	712	22.1.2	Therapie .....	764
21.5.7	Zielstruktur CD33 .....	713	<b>22.2</b>	<b>Auf Somatotropin ansprechende Formen des Minderwuchses .....</b>	<b>765</b>
21.5.8	Zielstruktur CD38 .....	714	22.2.1	Hypophysärer Minderwuchs .....	765
21.5.9	Zielstruktur CD52 .....	716	22.2.2	Ullrich-Turner-Syndrom und SHOX-Mangel .....	765
21.5.10	Zielstruktur CD79b .....	716	22.2.3	Prader-Willi-Syndrom .....	767
21.5.11	Zielstruktur CD123 .....	717	22.2.4	Chronische Niereninsuffizienz bei Kindern .....	768
21.5.12	Zielstruktur CCR4 .....	719	22.2.5	Vorgeburtliche Wachstumsverzögerung (SGA) .....	769
21.5.13	Zielstruktur BCMA .....	719	<b>22.3</b>	<b>Humanes Wachstumshormon .....</b>	<b>770</b>
21.5.14	Zielstruktur GPRC5D .....	722	22.3.1	Regulation der Somatotropin- Freisetzung .....	770
21.5.15	Zielstruktur CTLA-4 .....	723	22.3.2	Funktion von Somatotropin .....	771
21.5.16	Zielstruktur PD-1/PD-L1 .....	725	22.3.3	Somatotropin-Wirkstoffe .....	772
21.5.17	Zielstruktur LAG-3 .....	734	<b>22.4</b>	<b>Akromegalie .....</b>	<b>779</b>
21.5.18	Zielstrukturen c-MET und EGFR .....	734	22.4.1	Krankheitsbild .....	779
21.5.19	Zielstruktur Nectin-4 .....	735	22.4.2	Somatotropin-Antagonisten .....	780
<b>21.6</b>	<b>Inhibitoren der Nährstoffzufuhr .....</b>	<b>736</b>	<b>22.5</b>	<b>Insulinähnlicher Wachstumsfaktor IGF-1 .....</b>	<b>781</b>
21.6.1	Prinzip der Angiogenese .....	736	22.5.1	Primärer IGF-1-Mangel .....	783
21.6.2	Zielstruktur VEGF/VEGF-Rezeptor .....	737	<b>Sachregister .....</b>	<b>785</b>	
21.6.3	VEGF-Blocker .....	738	<b>Die Autoren .....</b>	<b>809</b>	
21.6.4	Asparaginase .....	741			
<b>21.7</b>	<b>Bispezifische T-Zell-Engager .....</b>	<b>743</b>			
21.7.1	Bispezifische Antikörperfragmente .....	745			
21.7.2	Bispezifische Vollantikörper .....	746			
21.7.3	T-Zell-Rezeptor-Antikörper-Fusionen (ImmTACs) .....	749			
<b>21.8</b>	<b>Virale Onkolyse .....</b>	<b>751</b>			
<b>21.9</b>	<b>Adjuvante Therapeutika .....</b>	<b>752</b>			
21.9.1	Erythropoetin .....	752			
21.9.2	Granulozyten-Wachstumsfaktor .....	752			
<b>21.10</b>	<b>Wirkstoffe zur Herabsetzung der Toxizität von Chemotherapien .....</b>	<b>757</b>			
21.10.1	Urat-Oxidase .....	757			
21.10.2	Carboxypeptidase G <sub>2</sub> .....	759			

## Abkürzungsverzeichnis

### A

μM	mikromolar
A <sub>260</sub> /A <sub>280</sub>	Absorption bei 260 nm bzw. 280 nm
AAV	Adenoassoziertes Virus
AcMNPV	<i>Autographa californica Multiple Nucleopolyhedrovirus</i>
ACR	<i>American College of Rheumatology</i>
ADA	Adenosin-Deaminase
ADCC	<i>antibody-dependent cell-mediated cytotoxicity</i> ; antikörperabhängige, zellvermittelte Zytotoxizität
ADCP	<i>antibody-dependent cell-mediated phagocytosis</i> ; antikörperabhängige zellvermittelte Phagozytose
ADA	<i>anti-drug antibodies</i>
ALL	akute lymphatische Leukämie
AMD	altersabhängige feuchte Makuladegeneration
AMP	Adenosinmonophosphat
amp <sup>R</sup>	Ampicillin-Resistenz-Gen (β-Lactamase)
ATMP	<i>Advanced Therapy Medicinal Products</i> , Arzneimittel für neuartige Therapien
AS04	Adjuvans aus hydratisierten Aluminiumhydroxid und MPL
ASCT	autologe Stammzelltransplantation

### B

BAC	<i>bacterial artificial chromosome</i> , bakterielles artifizielles Chromosom
BAFF	<i>B cell activating factor of the TNF family</i>
BDD-rFVIII	<i>B-domain deleted recombinant FVIII</i>
BlyS	<i>B lymphocyte stimulator</i> (auch BAFF genannt)
BFU-E	<i>erythroid-burst forming units</i>
BHK	<i>baby hamster kidney cells</i>
BiTE	<i>Bi-specific T cell Engager</i>
BMP	<i>bone morphogenetic protein</i> ; Knochenwachstumsfaktor
bp	<i>base pairs</i> , Basenpaare in einem DNA-Molekül
Bq	Becquerel
BSA	bovines Serumalbumin
BSC	<i>best supportive care</i> ; bestmögliche Begleitbehandlung

### C

C1INH	C1-Esterase-Inhibitor
CAPS	Cryopyrin-assoziertes periodisches Syndrom
CARs	<i>chimeric antigen receptor</i> ; chimärer Antigenrezeptor
CAT	<i>Committee for Advanced Therapies</i> ; Ausschuss für neuartige Therapien
CD	<i>cluster of differentiation</i>
CDC	<i>complement-dependent cytotoxicity</i> ; komplementabhängige Zytotoxizität
CDK	cyclinabhängige Proteinkinase

cDNA	complementary DNA, auch <i>copy DNA</i> , DNA-Kopie einer mRNA
CDR	<i>complementarity determining region</i> , hypervariable Region eines Antikörpers
CF	Cystische Fibrose oder Mucoviszidose
CFTR	<i>cystic fibrosis transmembrane conductance regulator</i>
CFU-E	<i>erythroid colony forming unit</i>
CGD	chronische Granulomatose
CHMP	<i>Committee for Medicinal Products for Human Use</i> , Ausschuss für Humanarzneimittel der EMA
CHOP	Chemotherapie mit Cyclophosphamid, Doxorubicin, Vincristin, Prednisolon
CHO	<i>chinese hamster ovary</i>
CIN	<i>cervical intraepithelial neoplasia</i> , intraepithelialen Neoplasie
CIS	<i>clinically isolated syndrome</i> ; klinisch isoliertes Syndrom
CLL	chronische lymphatische Leukämie
CML	chronische myeloische Leukämie
CMP	Cytidinmonophosphat
CMV	Cytomegalievirus
CTLA-4	zytotoxisches T-Lymphozyten-Antigen-4
CXCL	<i>C-X-C motif chemokine receptor ligand</i>
CXCR	<i>C-X-C motif chemokine receptor</i>
CYP	Cytochrom-P450-Enzym

### D

Da	Dalton
dAMP	Desoxyadenosin-5'-monophosphat
DAP I	Dipeptidylaminopeptidase I
dCMP	Desoxycytidin-5'-monophosphat
ddNTPs	2',3'-Didesoxynukleosidtriphosphate, Mischung der artifiziiellen Nukleotide ddATP, ddCTP, ddGTP und ddTTP
dGMP	Desoxyguanosin-5'-monophosphat
DHFR	Dihydrofolat-Reduktase
DMARD	<i>disease-modifying antirheumatic drug</i> ; Basistherapeutikum bei Rheumatoider Arthritis
DNA	Desoxyribonukleinsäure
DNase	Desoxyribonuklease
dNTPs	2'-Desoxynukleosidtriphosphate, Mischung der DNA-Bausteine dATP, dCTP, dGTP und dTTP
DOI	<i>Digital Object Identifier</i>
dsRNA	doppelsträngige RNA
DU	Polio-D-Antigen-Einheit
dTMP	Desoxythymidin-5'-monophosphat
dUMP	Desoxyuridin-5'-monophosphat

### E

<i>E. coli</i>	<i>Escherichia coli</i>
ECOG	<i>Eastern Cooperative Oncology Group</i>
EDSS	<i>expanded disability status scale</i>

EGF(R)	<i>epidermal growth factor (receptor)</i>	GM-CSF	Granulozyten-Monozyten-koloniestimulierender Faktor
eIF	eukaryontischer Initiationsfaktor der Translation	GMP	Good Manufacturing Practices
ELISA	<i>enzyme-linked immunosorbent assay</i>	GHRH	<i>growth hormone releasing hormone</i>
EMA	<i>European Medicines Agency</i> , zentrale europäische Arzneimittelzulassungsstelle	GnT	N-Acetylglucosamin-Transferase
ENV	<i>envelope</i> (Gen in Retroviren)	gp	Glykoprotein
EPAR	<i>European Public Assessment Report</i> , Europäischer Öffentlicher Beurteilungsbericht	GPI-Anker	Glycosylphosphatidylinositol-Anker
EpCAM	<i>epithelial cell adhesion molecule</i>	GRAS	<i>generally recognized as safe</i>
EPO	Erythropoetin	GST	Glutathion-S-Transferase
ER	Endoplasmatisches Retikulum	GvHD	<i>graft-versus-host disease</i>
ERK	<i>extracellular signal-regulated kinase</i>	GVO	gentechnisch veränderter Organismus
ES	<i>embryonic stem cells</i> ; embryonale Stammzellen	<b>H</b>	
ESE	<i>exonic splicing enhancer</i>	HAE	hereditäres Angioödem
ESchG	Embryonenschutzgesetz	HAMA	humaner Anti-Maus-Antikörper
ESS	<i>exonic splicing silencer</i>	HAT	Hypoxanthin, Aminopterin, Thymidin
<b>F</b>		HBV	Hepatitis-B-Virus
Fab	<i>fragment antigen-binding</i> (von Antikörpern)	hCG	humanes Choriongonadotropin
Fc	<i>fragment crystalizable</i> (Effektordomäne von Antikörpern)	HEK	<i>human embryonic kidney cells</i>
FcγR	Rezeptor für Fc-Teil von IgG-Antikörpern	HER	<i>human epidermal growth factor receptor</i>
FCS	<i>fetal calf serum</i> ; fetales Kälberserum	HERV	endogenes humanes Retrovirus
FDA	<i>Food and Drug Administration</i> ; US-amerikanische Arzneimittelzulassungsstelle	hGH	<i>human growth hormone</i> ; humanes Wachstumshormon
FGE	<i>formylglycine-generating enzyme</i>	HGP	Humangenomprojekt
FISH	<i>fluorescent in situ hybridization</i>	HGPRT	Hypoxanthin-Guanin-Phosphoribosyltransferase
FOLFIRI	Chemotherapie mit Folinsäure + 5-Fluorouracil + Irinotecan	HHV	humanes Herpesvirus
FOLFOX	Chemotherapie mit Folinsäure + 5-Fluorouracil + Oxaliplatin	Hib	<i>Haemophilus influenzae</i> Typ b
FRET	Fluoreszenz-Resonanz-Energietransfer	His <sub>6</sub>	Hexahistidin-Peptid
FSH	follikelstimulierendes Hormon	HIT	Heparininduzierte Thrombozytopenie
Fuc	Fucose	HIV	humanes Immundefizienz-Virus
FVC	forcierte Vitalkapazität	HLA	humanes Leukozyten-Antigen
<b>G</b>		hMG	humanes Menopausen-Gonadotropin
GA	<i>gene activation</i> (Genaktivierung in humanen Zellen)	HPV	humanpathogenes Papillomavirus
GAG	<i>group-specific antigen</i> (Retroviren)	HR	<i>hazard ratio</i>
Gal	Galactose	HSV	Herpes-simplex-Virus
GalNAc	N-Acetylgalactosamin	<b>I</b>	
GALNS	N-Acetylgalactosamin-6-Sulfatase	ICAM-1	<i>intercellular adhesion molecule-1</i>
GBq	Giga-Becquerel	ICH	<i>International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use</i>
gc	<i>genome copies</i> ; Genomkopien (Dosierung virale Genterapeutika)	IE	Internationale Einheit
GC	Gemcitabin	IFN	Interferon
G-CSF	Granulozyten-koloniestimulierender Faktor	IgA	Immunglobulin A
GF	<i>growth factor</i> ; Wachstumsfaktor	IgD	Immunglobulin D
GFP	Grün fluoreszierendes Protein	IgE	Immunglobulin E
Glc	Glucose	IGF	<i>insulin-like growth factor</i> ; insulinähnlicher Wachstumsfaktor
GlcNAc	N-Acetylglucosamin	IGFBP	<i>insulin-like growth factor-binding protein</i> ; IGF-Bindeprotein
GLP	<i>glucagon-like-peptide</i>	IGFR	<i>insulin-like growth factor receptor</i>
		IgG	Immunglobulin G
		IgM	Immunglobulin M

IHRES	interne ribosomale Eintrittsstelle	NBCD	<i>non-biological complex drug</i> ; nichtbiologisches komplexes Arzneimittel
IL	Interleukin		
IL-1Ra	Interleukin-1-Rezeptorantagonist		
INN	<i>International Nonproprietary Name</i>	<b>N</b>	
iPS	<i>induced pluripotent stem cells</i>	NeuAc	<i>N</i> -Acetylneuraminsäure (Sialinsäure)
IPTG	Isopropyl- $\beta$ -D-thiogalactosid	NeuGC	<i>N</i> -Glycolylneuraminsäure
IRS	Insulinrezeptor-Substrat	<i>neo</i> <sup>R</sup>	Resistenz gegen Neomycin bzw. G418 (Aminoglycosid-Phosphotransferase)
ISE	<i>Intronic splicing enhancer</i>	NF $\kappa$ B	<i>nuclear factor <math>\kappa</math> B</i>
ISS	<i>Intronic splicing silencer</i>	NGF	<i>nerve growth factor (receptor)</i>
ITP	idiopathische thrombozytopenische Purpura	NHL	Non-Hodgkin-Lymphom
IVF	In-vitro-Fertilisation	NIH	<i>National Institutes of Health</i>
		NK	Natürliche Killerzellen
<b>J</b>		nM	nanomolar
JAK	Janus-Kinase	NSCLC	<i>non-small cell lung cancer</i> ; nichtkleinzelliges Lungenkarzinom
JNK	Jun-N-terminale Kinase	NYHA I-IV	New York Heart Association I–IV
<b>K</b>		<b>O</b>	
kbp	<i>kilo base pairs</i> , 1000 Basenpaare (bp)	OPG	Osteoprotegerin
KGF	<i>keratinocyte growth factor (receptor)</i>	ORF	offener Leserahmen
kg KG	Kilogramm Körpergewicht	<i>ori</i>	<i>origin of replication</i> , Replikationsursprung
KI	Konfidenz-Intervall		
KRAS	<i>Kirsten rat sarcoma</i>		
		<b>P</b>	
<b>L</b>		p53	Tumorsuppressorprotein 53 kDa
LAK	Lymphokin-aktivierte Killerzellen	PACE	<i>paired basic amino acid cleaving enzyme</i>
LH	luteinisierendes Hormon	PBMC	<i>peripheral blood mononuclear cells</i>
LTR	<i>long terminal repeat</i> (in Retroviren)	PBS	Primer-Bindestelle
		PCR	Polymerase-Kettenreaktion
<b>M</b>		PD-1(L)	<i>programmed death-1 (ligand)</i>
MAE-hGH	rekombinantes Wachstumshormon mit den zusätzlichen N-terminalen Aminosäuren Methionin, Alanin und Glutaminsäure	PDGF	<i>platelet-derived growth factor (receptor)</i> ; Plättchen-Wachstumsfaktor(-Rezeptor)
Man	Mannose	pdFIX	<i>plasma-derived factor IX</i>
MAP-Kinase	Mitogen-aktivierte Proteinkinase	PEG	Polyethylenglycol
Mbp	Megabasenpaare, 10 <sup>6</sup> bp	PFU	plaquebildende Einheiten
MBq	Mega-Becquerel	PI3K	Phosphatidylinositol-3-Kinase
mCi	Milli-Curie	POL	<i>polymerase</i> (Gen in Retroviren)
MCS	<i>multiple cloning site</i>	PPMS	primär progrediente MS
MHC	<i>major histocompatibility complex</i>	PRAC	<i>Pharmacovigilance Risk Assessment Committee</i> ; Ausschuss für Risikobewertung im Bereich der Pharmakovigilanz
M-hGH	rekombinantes Wachstumshormon mit einem zusätzlichen N-terminalem Methionin		
miRNA	<i>microRNA</i>	pRB	Retinoblastom-Protein
MLV	Maus-Leukämie-Virus	PTH	Parathormon
mM	millimolar	PVX	Potatovirus X
MMAE	Monomethylauristatin E	PWS	Prader-Willi-Syndrom
MMP	Matrixmetalloproteinase		
MoMLV	Moloney-Maus-Leukämievirus	<b>R</b>	
MPL	3- <i>O</i> -Desacyl-4'-monophosphoryl-Lipid A	RANK(L)	<i>receptor activator of nuclear factor kappa B (ligand)</i>
MPR	Mannose-6-phosphat-Rezeptor	PIGF	<i>placental growth factor</i> ; plazentarer Wachstumsfaktor
MPS	Mukopolysaccharidose		
MRT	Magnetresonanztomographie	RNAi	<i>RNA interference</i>
MS	multiple Sklerose	RNase	Ribonuklease
MTX	Methotrexat	RRE	<i>rev response element</i>
mU	Milli-Units	RMS	<i>relapsing multiple sclerosis</i>
NAT	Nukleinsäure-Amplifikationstechniken		

RRMS	<i>relapsing remitting multiple sclerosis</i> ; schubförmig remittierende MS	$T_m$	halbmaximale Denaturierungstemperatur von DNA
rsCT	rekombinant hergestelltes Lachs-Calcitonin	TMV	Tabakmosaikvirus
RSV	Respiratorisches Synzytial-Virus	Tn	Zelllinien aus <i>Trichoplusia ni</i>
RT	Reverse Transkriptase	TNF	Tumornekrosefaktor
RT-PCR	Reverse Transkription mit anschließender Amplifikation einer DNA durch PCR	TNSALP	<i>tissue non-specific alkaline phosphatase</i> ; gewebeunspezifische Alkalische Phosphatase
<b>S</b>		tPA	Gewebe-Plasminogenaktivator
scFv	<i>single-chain variable fragment</i> (eines Antikörpers)	TPMT	Thiopurin-S-Methyltransferase
SCID	<i>severe combined immunodeficiency</i> ; schwere kombinierte Immundefizienz	TPO	Thrombopoetin
SCNT	<i>somatic cell nuclear transfer</i> , somatischer Kerntransfer	TRH	<i>thyrotropin-releasing hormone</i>
SD	<i>standard deviation</i> , Standardabweichung	TSH	schilddrüsenstimulierendes Hormon
SDF-1	<i>stroma cell-derived factor 1</i> , auch CXCL-12	TSLP	<i>thymic stromal lymphopoietin</i>
SDS	<i>standard deviation score</i>	<b>U</b>	
Sf	Zelllinien aus <i>Spodoptera frugiperda</i>	UDP	Uridindiphosphat
SGA	<i>small for gestational age</i> ; vorgeburtliche Wachstumsverzögerung	U	unit, Einheit
SHOX	<i>shortsStature homeobox</i>	<b>V</b>	
siRNA	<i>small interfering RNA</i>	VCAM	<i>vascular cell adhesion molecule</i>
SLAMF7	<i>signaling lymphocyte activation molecule family member 7</i>	VEGF(R)	<i>vascular endothelial growth factor (receptor)</i> ; vaskulärer endothelialer Wachstumsfaktor (-Rezeptor)
SLE	systemischer Lupus erythematoses	VLA-4	<i>very late antigen-4</i> , ein Protein aus der Integrin-Familie
SNP	(sprich: „snip“): <i>single nucleotide polymorphism</i>	vWF	von-Willebrand-Faktor
SPF	<i>specific pathogen-free</i> ; spezifiziert pathogenfrei (Status von Tierherden für die Wirkstoffherstellung)	<b>W</b>	
SPMS	sekundär progrediente MS	WCB	<i>working cell bank</i> , Arbeitszellbank
SRIF	<i>somatotropin release-inhibiting factor</i> ; Somatostatin	WHO	<i>World Health Organisation</i> , Weltgesundheitsorganisation
SRP	<i>signal recognition particle</i>	<b>X</b>	
STAT	<i>signal transducers and activators of transcription</i>	X-Gal	5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranosid
STIKO	Ständige Impfkommision des Robert Koch-Instituts	Xyl	Xylose
StZG	Stammzellgesetz	<b>Y</b>	
<b>T</b>		YAC	<i>yeast artificial chromosome</i> , Hefe-artifizielles Chromosom
TEP	<i>tissue-engineered product</i> ; biotechnologisch bearbeitetes Gewebeprodukt	YE <sub>p</sub>	<i>Yeast Episomal plasmid</i>
TGF	<i>transforming growth factor</i>	YI <sub>p</sub>	<i>Yeast Integrating plasmid</i>
TIL	tumorinfiltrierende Lymphozyten	<b>Z</b>	
TK	Thymidinkinase	ZNS	Zentralnervensystem

## 1

## Die bio(techno)logische Revolution

1.1	Von der klassischen Genetik zur Gentechnik . . . . .	1
1.2	Gentechnik und Biotechnik . . . . .	6

Wir befinden uns inmitten einer „biologischen Revolution“, oder besser gesagt, einer „molekulargenetischen Revolution“. Seit gerade einmal etwas mehr als 70 Jahren wissen wir um die prinzipielle Struktur der Erbinformation, der DNA. Bereits kurze Zeit später haben wir verstanden, dass der genetische Code in allen auf der Erde lebenden Organismen gleich ist und in allen Organismen auf ähnliche Weise in Proteine übersetzt wird. In den frühen 1970er-Jahren wurden die ersten DNA-Moleküle verschiedener Organismen über Artgrenzen hinweg zu neuen funktionalen Einheiten zusammengefügt. Diese neu entwickelte DNA-Rekombinationstechnologie, die auch unter der vielleicht geläufigeren Bezeichnung Gentechnik bekannt ist, ist heute die Grundlage für die Erforschung komplexer physiologischer und pathophysiologischer Zusammenhänge auf molekularer Ebene, der Entwicklung neuer biomedizinischer Theorien und Therapieansätze, der Suche nach neuen Zielen für Wirkstoffe und der großtechnischen Produktion von Proteinwirkstoffen. Biopharmazeutika, die in gentechnisch veränderten Organismen mit biotechnologischen Verfahren hergestellte Proteine enthalten, nehmen ein stetig wachsendes Segment im Arzneimittelschatz ein, sodass sich Arzneimittelfachleute neben ihrem Fachwissen über die Gewinnung und Anwendung chemisch-synthetischer oder natürlicher niedermolekularer Wirkstoffe zunehmend auch mit den Konzepten zur Produktion von und dem Umgang mit derartigen Makromolekülen befassen müssen. Immerhin haben wir es heute mit 451 zugelassenen Biopharmazeutika und 397 Proteinwirkstoffen zu tun (Stand August 2025), die ca. 35% des gesamten deutschen Pharmamarktes ausmachen – Tendenz steigend.

Die Methoden und Konzepte zur gentechnischen Herstellung von Proteinwirkstoffen nehmen in diesem Buch einen breiten Raum ein. Allerdings hat die Gentechnik eine weitaus größere Bedeutung in der biomedizinischen und pharmazeutischen Anwendung als „nur“ ihre Nutzung zur Herstellung von Proteinwirkstoffen. In diesem einführenden Kapitel wollen wir daher die

„biologische Revolution“ zusammenfassen, die vor gut 150 Jahren mit Darwins Evolutionstheorie begann, deren Fundament die Etablierung zunächst der „klassischen“ und danach der „molekularen“ Genetik ist, die zunächst in der Erfindung der Gentechnik gipfelte und die möglicherweise erst mit dem gezielten Editieren des menschlichen Genoms zu Enhancement-Zwecken endet. Nicht alles heute oder morgen Machbare der neuen molekularmedizinischen Entwicklungen (z.B. reproduktionsmedizinische Maßnahmen) hat etwas mit Gentechnik zu tun – allerdings wäre die Erforschung neuer molekularmedizinischer Therapiekonzepte ohne die Gentechnik mit Sicherheit nicht möglich gewesen. Heute wird es kein Medikament geben – übrigens auch keines mit einem chemisch-synthetischen Wirkstoff – in dessen Entwicklung von der biologischen Grundlagenforschung bis zur klinischen Entwicklung nicht zu irgendeinem Zeitpunkt gentechnische Methoden zum Einsatz kommen. Nachfolgend wollen wir kurz die Entwicklung der Genetik und der Gentechnik rekapitulieren, bevor wir den wichtigen Unterschied zwischen „Gentechnik“ und „Biotechnik“ thematisieren, der für das Verständnis der Entwicklung und Herstellung von Proteinwirkstoffen von Bedeutung ist.

### 1.1 Von der klassischen Genetik zur Gentechnik

Molekularbiologie ist von seinen Ursprüngen her ein Zweig der Biologie, der darauf abzielt, durch das Studium (makro)molekularer Strukturen Erkenntnisse über zelluläre Vorgänge zu erlangen. Die Geschichte der Molekularbiologie, der modernen biomedizinischen Forschung und der Entwicklung von Biopharmazeutika ist eng verbunden mit der Entdeckungsgeschichte der DNA als dem Träger der Erbinformation. Alles begann im Jahr 1865, als Gregor Mendel (1822 – 1884) seine Entdeckung veröffentlichte, dass Merkmale von Organismen als vererbare Einheiten in

klaren mathematischen Zusammenhängen über Generationen weitergegeben werden, auch wenn sich die Anwesenheit dieser Merkmale nicht immer am Erscheinungsbild der Organismen ablesen lässt. Mendels „Versuche über Pflanzenhybriden“, die im Jahr 1866 veröffentlicht wurden, zeigten, dass Merkmale in den Nachkommen entweder hervortreten und **dominant** sind oder zurücktreten und **rezessiv** sind. Mendel war seiner Zeit weit voraus – und seine Entdeckungen fanden zunächst kaum Beachtung. Sie sind jedoch heute ebenso fundamental wie damals, gelten sie doch beispielsweise auch für die Vererbung von Krankheiten des Menschen und bilden die Grundlage für die Identifizierung von „Krankheitsgenen“ durch Stammbaum-Analysen in betroffenen Familien.

Ende des 19. Jahrhunderts wurde allmählich klar, dass die Zellkerne die Träger der Erbinformation sind. Im Zellkern fand Walter Flemming (1843 – 1901) eine anfärbbare Substanz, die er **Chromatin** nannte. Weitere Forschungen ergaben, dass sich das Chromatin während der Zellteilung verdichtet. Heinrich Wilhelm Waldeyer (1836 – 1921) sprach erstmals von **Chromosomen** und beschrieb 1888, dass sich Chromosomen während der Zellteilung aufspalten und die einzelnen Chromosomen-Fäden auf die entstehenden Tochterzellen verteilt werden. In diesem Zeitraum wurde auch klar, dass jedes Chromosom in einer Körperzelle zweimal vorhanden ist – außer in den Keimzellen, die jeweils nur eine Kopie enthalten.

Obwohl bereits andere Wissenschaftler um die Wende zum 20. Jahrhundert herum die Arbeiten von Mendel wiederentdeckt hatten, war es vor allem William Bateson (1861 – 1926), der im Jahr 1900 Mendels Publikation gelesen und die Mendel'schen Regeln in Wissenschaftskreisen bekannt gemacht hatte. Bateson hat auch den Begriff **Genetik** eingeführt (1906). Es wurde klar, dass die Chromosomen die Träger der Mendel'schen Merkmale sein müssen, und es war unter anderen Theodor Boveri (1862 – 1915), der aus dem Wissen der damaligen Zeit ableitete, dass die vererbten Merkmale im Mendel'schen Sinne auf den Chromosomen lokalisiert sein müssten. Walter S. Sutton (1877 – 1916) erkannte 1902, dass während der Bildung der Keimzellen und der Reduktion des doppelten Chromosomensatzes (heute nennen wir diesen Vorgang Meiose) die Chromosomenpaare in zufälliger Verteilung in die Keimzellen gelangen und dass dies die Grundlage für die von Mendel beschriebenen statistischen Verteilungsmuster der beobachteten Merkmale sein könnte. Diese Hypothese ist als **Chromosomentheorie der Vererbung** bzw. **Boveri-Sutton-Theorie** in die Geschichtsbücher eingegangen.

Der Botaniker Wilhelm Johannsen (1857 – 1927) benutzte in seinem Buch „Elemente der exakten Erbliehkeitslehre“ im Jahre 1909 erstmals den Begriff **Gen**. Das war damals etwas Abstraktes, ein Symbol, das dafür

stand, dass die Merkmale von Organismen in den Keimzellen festgelegt sind und als unabhängige Einheiten an nachfolgende Generationen vererbt werden. Von Johannsen stammt auch der abgeleitete Begriff **Genotyp**, also die Gesamtheit der Gene eines Organismus. Am **Phänotyp**, dem Erscheinungsbild eines Organismus, ist der Genotyp nicht unbedingt zu erkennen, was Mendel bereits als dominante und rezessive Merkmale beschrieben hatte.

Zur endgültigen Akzeptanz der Mendel'schen Regeln trug insbesondere Thomas Hunt Morgan (1866 – 1945) bei. Morgan arbeitete mit der Fruchtfliege *Drosophila melanogaster*, einem überaus wertvollen Modellsystem der biologischen Forschung. Morgan identifizierte viele *Drosophila*-Mutanten mit bestimmten Merkmalen und zeigte, dass diese teilweise gekoppelt vererbt werden. Da Morgan genauso viele Kopplungsgruppen fand wie die Fruchtfliege Chromosomen hat, lag der Schluss nahe, dass Gene auf den Chromosomen liegen und daher in der Regel gemeinsam vererbt werden, sofern sie sich auf demselben Chromosom befinden. Morgan bemerkte aber auch Ausnahmen von dieser Regel, also Kopplungsbrüche, die offenbar dadurch zustande kamen, dass homologe Chromosomen während der Meiose Stücke austauschen. Schließlich war es Alfred H. Sturtevant (1891 – 1970), ein Mitarbeiter von Morgan, der 1913 erkannte, dass die Wahrscheinlichkeit, dass gekoppelt vererbte Merkmale während der Keimzellbildung voneinander getrennt werden, mit der Entfernung der Gene zueinander auf den jeweiligen Chromosomen zusammenhängt. Morgan fasste die Ergebnisse seiner Forschungen im Jahr 1928 in dem berühmten Buch *The Theory of the Gene* zusammen. Noch heute werden Entfernungen zwischen Genen auf Chromosomen als **Centimorgan** angegeben.

Auf dem Weg zur Beantwortung der Frage, was ein Gen eigentlich ist, hatte zunächst die Zytogenetikerin Barbara McClintock (1902 – 1992) im Jahr 1931 erstmals am Beispiel der Maispflanze gezeigt, dass die Entkopplung genetischer Marker mit einem Austausch von genetischem Material zwischen homologen Chromosomen einhergeht. Etwa um die gleiche Zeit (1927) berichtete Herman J. Muller (1890 – 1967), ein Schüler von Morgan, über seine Experimente zur mutationsauslösenden Wirkung von Röntgenstrahlung. Die Erbinformation konnte also von außen beeinflusst (verändert) werden. Max Delbrück (1906 – 1981) und Salvador E. Luria (1912 – 1991) zeigten 1943 in bahnbrechenden Experimenten mit Bakteriophagen, dass Mutationen zufällig auftreten.

Oswald T. Avery (1877 – 1955) hat 1944 gefunden, dass Desoxyribonukleinsäure (DNA) Träger genetischer Information ist. Avery prägte den Begriff **Transformation** und meinte damit die Übertragung neuer Eigenschaften auf einen Organismus durch DNA. Alfred Kühn (1885 – 1968) sowie später George W. Beadle

(1903 – 1989) und Edward Tatum (1909 – 1975) konnten um 1950 durch biochemische Experimente zeigen, dass Gene die Information zur Bildung von Enzymen (Proteinen) tragen: Die Hypothese „Ein Gen = ein Enzym“ war geboren. Im Jahr 1952 zeigte Alfred Hershey (1908 – 1997), dass bei der Infektion von Bakterien mit einem Bakteriophagen nur die virale DNA, nicht aber seine Proteine, in die infizierte Zelle übertragen werden. Somit war klar, dass in der DNA die genetische Information zur Bildung neuer Phagenpartikel enthalten sein musste.

James D. Watson (1928 – 2025) und Francis Crick (1916 – 2004) trafen sich erstmals im Jahr 1951 und waren fortan von der Idee getrieben, die Struktur der DNA als dem offensichtlichen Träger der genetischen Information zu entschlüsseln. Die experimentellen Daten dazu kamen eigentlich von der DNA-Strukturforscherin Rosalind E. Franklin (1920 – 1958) und wurden Watson und Crick von Maurice H. F. Wilkins (1916 – 2004) ohne Wissen Franklins zur Verfügung gestellt. Die berühmte, von Franklin 1952 erstellte Röntgenbeugungsaufnahme der DNA mit der Laborbezeichnung „Photograph 51“ war letztlich die Grundlage für die Lösung der DNA-Struktur durch Watson und Crick, die in einer unscheinbaren einseitigen Publikation in *Nature* im Jahr 1953 erschien: Die DNA besteht aus zwei DNA-Strängen, die sich helixartig umeinander winden und antiparallel verlaufen; die Phosphatreste befinden sich außen und die DNA-Basen zeigen nach innen.

Für die endgültige Lösung der DNA-Struktur war noch eine weitere Information unerlässlich, und zwar die Lage der DNA-Basen beider Einzelstränge zueinander. Erwin Chargaff (1905 – 2002) hatte mit eleganten biochemischen Methoden etwa um das Jahr 1950 gezeigt, dass die Verhältnisse der DNA-Basen Thymin (T) und Adenin (A) sowie Cytosin (C) und Guanin (G) in einer DNA immer gleich sind. Daraus schlossen Watson und Crick treffend, dass sich T und A bzw. C und G in der Doppelhelix gegenüberliegen und über spezifische Wasserstoffbrückenbindungen verbunden sind. Sie schlussfolgerten auch, dass die spezifische Basenpaarung einen Replikationsmechanismus der DNA beinhaltet, denn durch die spezifische Basenpaarung sind die DNA-Stränge komplementär. Wenn die Basensequenz des einen DNA-Strangs bekannt ist, kann man aufgrund der **Komplementaritätsregeln** auf die Basenfolge des gegenüber liegenden Strangs schließen. Heute sind die Chargaff'schen Regeln bzw. die Komplementaritätsregeln Grundlage der DNA-Rekombinationstechnologie und damit der biologischen Revolution, die nach der Entdeckung der DNA-Doppelhelix erst so richtig in Gang kam.

Kurz nach der Entdeckung der DNA-Struktur führten Matthew Meselson (geb. 1930) und Franklin W. Stahl (1929 – 2025) im Jahr 1958 ein heute als klassisch anerkanntes Experiment durch, mit dem sie zeigten,

dass die DNA durch semikonservative Replikation vermehrt wird. Damit ist gemeint, dass die beiden DNA-Stränge geöffnet werden und ein Strang jeweils als Matrize zur Neubildung des Gegenstrangs verwendet wird. Etwa um die gleiche Zeit (1956) isolierte Arthur Kornberg (1918 – 2007) erstmals eine DNA-abhängige DNA-Polymerase, die diese Art der enzymatischen DNA-Synthese durchführen kann. DNA-Polymerasen gehören heute zu den unverzichtbaren Werkzeugen in der Gentechnologie.

Da nun klar war, dass die genetische Information in der DNA gespeichert ist und bei der Zellteilung von Zelle zu Zelle weitergegeben werden kann, stand die Frage im Raum, wie die in der Basenfolge der DNA „codierte“ Information in die Aminosäureabfolge der Proteine übersetzt wird. Wieder war es Francis Crick, der den entscheidenden Hinweis gab. Basierend auf Experimenten mit Bakteriophagen schlug Crick in einer 1961 erschienenen *Nature*-Publikation vor, dass die genetische Information für Aminosäuren durch jeweils drei DNA-Basen codiert wird und dass diese Information „degeneriert“ sei, also dass mehrere Basen-Triplets für eine bestimmte Aminosäure stehen könnten. Welches Triplet für welche Aminosäure codiert, fanden zuerst Marshall Nirenberg (1927 – 2010) und Heinrich Matthaei (1929 – 2025) heraus, die im Jahr 1961 ein kurzes Stück RNA herstellten, das ausschließlich Uracil-Basen enthielt. Als sie mit dieser RNA ein In-vitro-Translations-Experiment durchführten, erhielten sie ein kurzes Protein, das ausschließlich aus Phenylalanin bestand: Das Triplet UUU codiert für die Aminosäure Phenylalanin. Der gesamte genetische Code wurde in den folgenden Jahren entziffert, und es wurde festgeschrieben, was heute eine weitere Grundlage der Anwendung der DNA-Rekombinationstechnologie bildet: Der genetische Code ist auf der Erde universell gültig.

Das zunehmende Interesse an der Erforschung von Genfunktionen und ihrer Beteiligung am Auftreten von Krankheiten wurde zunächst gedämpft durch einen Mangel an Werkzeugen, mit denen DNA außerhalb von Zellen bearbeitet und Gene gezielt isoliert werden konnten. Man brauchte Methoden, um die langen chromosomalen DNA-Abschnitte gezielt in kürzere Fragmente zu zerlegen, solche Fragmente zu trennen und die gesuchten Gene zu charakterisieren. Das war die Geburtsstunde der Gentechnik. Zunächst wurden von Werner Arber (geb. 1929), Daniel Nathans (1928 – 1999) und Hamilton O. Smith (1931 – 2025) die sogenannten Restriktionsendonukleasen als fundamentale Werkzeuge der Gentechnologie entdeckt und zum gezielten Fragmentieren von DNA verwendet. Paul Berg (1926 – 2023) erzeugte 1972 erstmals rekombinante DNA, indem er Stücke des SV40-Virus mithilfe DNA-modifizierender Enzyme in bakterielle Plasmide einbaute. Kurz darauf (1973) zeigten Herbert W. Boyer (geb. 1936) und Stanley N. Cohen (geb.

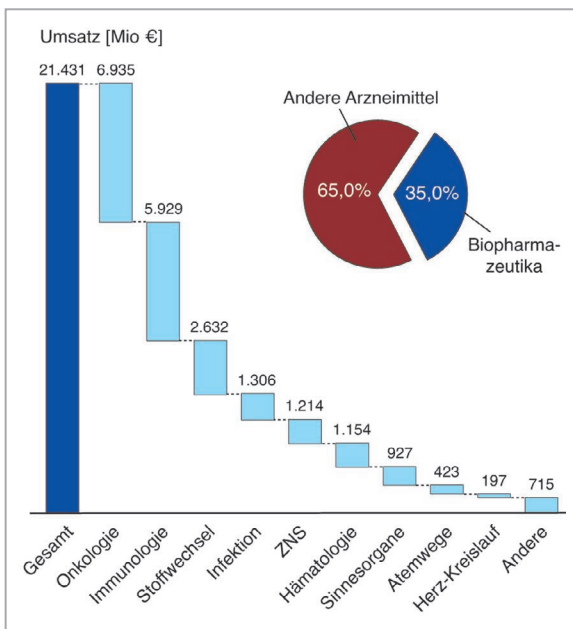
1935), dass man in vitro rekombinierte Plasmide in Bakterienzellen einbringen und dort vermehren kann. Etwas später (1974) zeigten sie, dass man auch DNA einer fremden Spezies in solche Plasmide einbauen und von Bakterien replizieren und auch transkribieren lassen kann.

Bereits kurz nach ihrer Entwicklung wurde das kommerzielle Potenzial der DNA-Rekombinationstechnologie erkannt. Im Jahr 1976 gründete Herbert Boyer die Firma Genentech (*Genetic engineering technology*), die bereits ab 1978 mit Humaninsulin den weltweit ersten rekombinanten Proteinwirkstoff herstellte (Zulassung 1982). Ebenfalls im Jahr 1978 wurde von Walter Gilbert (geb. 1932), einem Pionier der DNA-Sequenzierung, die Firma Biogen gegründet, die mit Erfolg Proteine wie Hepatitis-B-Antigen, Alpha-Interferon und Beta-Interferon produzierte. Die 1980 gegründete Firma Amgen stellte unter anderem Erythropoetin her. Am 14. Oktober 1980 ging die damals 4 Jahre alte kalifornische Firma Genentech an die New Yorker Börse – also noch 2 Jahre vor der Zulassung des ersten von der Firma hergestellten Produkts. Damit hatte man sozusagen die kommerzielle Nagelprobe der Gentechnik bestanden. Unter Investoren initiierten diese Erfolge unglaubliche Fantasien: Binnen 20 Minuten stieg der Preis der Genentech-Aktie von 35 auf 89 Dollar. Bei Börsenschluss lag der Wert der Aktie bei 71,25 Dollar. Die Möglichkeiten der Biotechnik, einer

sehr alten, aber keineswegs populären Wissenschaft, waren durch die Gentechnik deutlich erweitert worden, und das erregte damals unglaubliche Aufmerksamkeit.

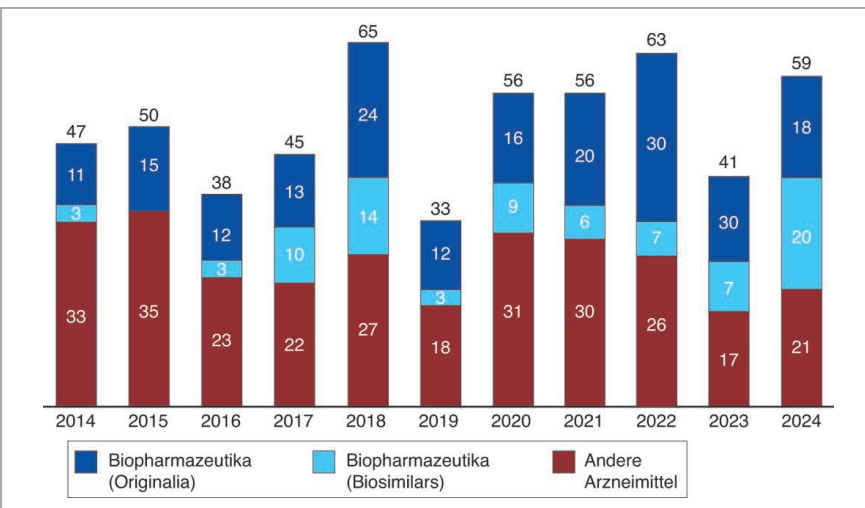
Ein bedeutendes Ziel aller biotechnologischen Forschung ist die Entwicklung kommerzieller Produkte. Aus diesem Grund wird im Gegensatz zu vielen anderen wissenschaftlichen Disziplinen auch die gentechnische bzw. biotechnische Forschung zu einem großen Teil durch eine ökonomische Motivation getragen. Die Erwartungen eines signifikanten finanziellen Gewinns war auch der Grund für das große Interesse und die Begeisterung während der Anfangszeit der industriellen Gentechnik. Bereits im Jahre 1985, also kaum 10 Jahre nach der Entwicklung der DNA-Rekombinationstechnologie, existierten über 400 Biotech-Firmen in den USA. Im Jahr 2022 waren es nach dem *Biotechnology Report 2023* von Ernst & Young insgesamt 949 Firmen. Die weltweite Biotechnologie-Industrie konnte im Jahr 2022 ca. 215 Milliarden Dollar Umsatzerlöse erzielen und beschäftigte über 280 000 Mitarbeiter.

Dieser QR-Code führt zur Information des Verbands Forschender Arzneimittelhersteller zur Medizinischen Biotechnologie in Deutschland.

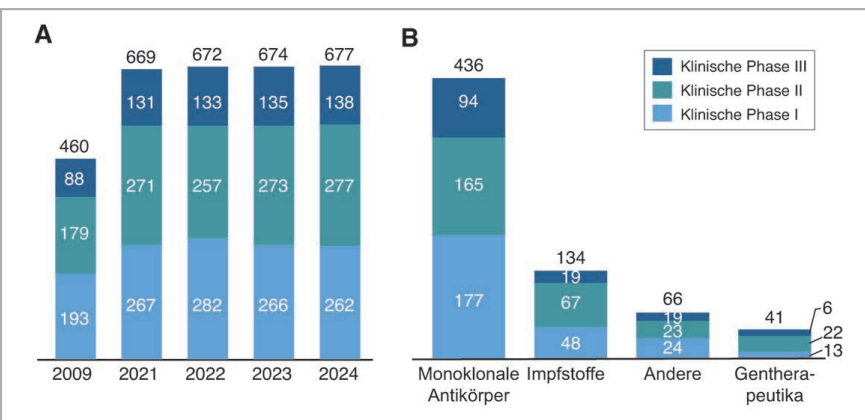


• **Abb. 1.1** Umsatz mit Biopharmazeutika in Deutschland nach Indikationsgebieten. Kreisdiagramm: Biopharmazeutika hatten 2024 einen Netto-Gesamtumsatz im Apotheken- und Krankenhaus-Markt (unter Berücksichtigung des Herstellerabschlages) von 21,4 Mrd. Euro und einen Anteil am gesamten Arzneimittelmarkt von 35,0 %. Balkendiagramm: Umsätze aufgeschlüsselt nach Indikationsgruppen. Quelle: Medizinische Biotechnologie in Deutschland 2025.

Auch in Deutschland hat der Umsatz mit biotechnologisch erzeugten Proteinwirkstoffen stark an Bedeutung gewonnen. In dem von der *Boston Consulting Group* gemeinsam mit dem vfa (Verband forschender Arzneimittelhersteller) herausgegebenen Biotech-Report „Medizinische Biotechnologie in Deutschland 2025“ wird erwähnt, dass der Anteil an Biopharmazeutika am Gesamtpharmamarkt (Apotheken und Kliniken) im Jahr 2024 bei 35,0 % lag, das entspricht einem Netto-Umsatz von ca. 21,4 Milliarden Euro (• Abb. 1.1) und einer Steigerung von 14,1 % im Vergleich zum Vorjahr. Den größten Anteil am Umsatz haben Therapeutika im Bereich der Onkologika und Immuntherapeutika sowie Arzneimittel bei Stoffwechselerkrankungen (• Abb. 1.1). Im Jahr 2024 gab es in Deutschland 144 im biopharmazeutischen Segment aktive Unternehmen mit zusammen etwa 52 000 Beschäftigten. Im August 2025 waren 451 Biopharmazeutika mit 397 Wirkstoffen auf dem europäischen Markt zugelassen (Quelle: vfa). Biopharmazeutika stellen einen erheblichen Anteil der jährlichen Arzneimittel-Neuzulassungen in Europa dar (64 % im Jahr 2024; • Abb. 1.2). Und die Pipeline der pharmazeutischen Industrie ist gut gefüllt: 2024 befanden sich 677 Biopharmazeutika in klinischen Prüfungen (Phasen I-III), davon 138 in Phase III (• Abb. 1.3). Mehr als die Hälfte der Neuentwicklungen fallen in den Bereich der monoklonalen Antikörper (• Abb. 1.3), dem Wachstumsmarkt in den vergangenen Jahren und auch in absehbarer Zukunft.



• **Abb. 1.2** Anteil von Biopharmazeutika an jährlichen Neuzulassungen in Deutschland. Aufgegliedert in „Andere Arzneimittel“ (keine Proteinwirkstoffe) und Biopharmazeutika (Originalia und Biosimilars). Quelle: Medizinische Biotechnologie in Deutschland 2025.



• **Abb. 1.3** Biopharmazeutika-Pipeline: (A) Gesamtzahl der Wirkstoffe in der klinischen Entwicklung zwischen 2021 und 2022 im Vergleich zu 2009 (B) Aufgliederung der Arzneimittelstoffe in Entwicklung im Jahr 2024 nach Wirkstoffarten. Quelle: Medizinische Biotechnologie in Deutschland 2025.

## Zusammenfassung

- Die bio(techno)logische Revolution begann vor etwa 150 Jahren mit Charles Darwins Theorie über die Entstehung der Arten. Die Vererbungslehre (Genetik) gründet auf klassischen Arbeiten von Gregor Mendel aus dem Jahr 1865.
- In den folgenden etwa 100 Jahren wurden die Grundlagen für neue naturwissenschaftliche Disziplinen geschaffen, die wir heute als Molekularbiologie und Molekulare Genetik kennen.
- Die Erfindung der DNA-Rekombinationstechnologie (Gentechnik) durch Herbert Boyer und Stanley Cohen Anfang der 1970er-Jahre hat unter anderem den Weg für die systematische Erforschung von Genfunktionen und ganzen Genomen gebnet.
- Gentechnische Methoden sind heute unerlässlich für die biomedizinische Forschung über pathophysiologische Mechanismen und die Identifizierung klinisch relevanter Zielstrukturen für Wirkstoffe. Sie bilden ebenso die Grundlage für die Herstellung von Proteinwirkstoffen (Biopharmazeutika).
- Eine ganze Industrie hat sich seit Ende der 1970er-Jahre etabliert, die heute weltweit ca. 280 000 Mitarbeitende beschäftigt und die mit biotechnologischen Verfahren potente Arzneimittel gegen Krankheiten liefert, die bis vor Kurzem noch als wenig oder gar nicht therapierbar galten.

## 1.2 Gentechnik und Biotechnik

Häufig werden die Begriffe **Molekularbiologie** und **Gentechnologie** synonym verwendet. Das ist jedoch nicht korrekt. Der Begriff **Molekulare Biologie** wurde zuerst von dem amerikanischen Wissenschaftler Warren Weaver (1894 – 1978) benutzt – und zwar bereits im Jahr 1938. Damals wollte die Rockefeller-Stiftung grundlegende biologische Fragestellungen im Bereich der biologischen Forschung und insbesondere die Frage nach der Natur des Gens finanziell unterstützen und dazu die zwei zu der Zeit vorherrschenden Denkweisen und experimentellen Ansätze zusammenbringen: die klassische Genetik und die Strukturbiologie. Daraus hat sich eine ganz neue Disziplin der biologischen Forschung geformt, und zwar die **Molekulare Genetik**, die sich mit den molekularen Grundlagen der Vererbung von Genen befasst, wie etwa der Struktur und Funktion von Nukleinsäuren oder den Replikations- und Genexpressionsmechanismen. Der Begriff **Molekulare Biologie** hat sich erst im Lauf der Zeit entwickelt und überlappt heute mit der Molekularen Genetik; Molekularbiologie ist heute eine Disziplin innerhalb der Biologie, die sich mit der Struktur und dem Stoffwechsel von Nukleinsäuren und deren Interaktionen mit zellulären Komponenten auf molekularer Ebene beschäftigt. Die Überlappungen mit anderen biologischen Disziplinen wie der Zellbiologie, Genetik oder Biochemie sind dabei fließend. Wenn sich heute fast alle Forscher aus verschiedenen biologischen Disziplinen übergreifend als **Molekularbiologen** bezeichnen, dann meinen sie damit vermutlich, dass sie Vorgänge innerhalb einer Zelle oder auch Interaktionen zwischen Zellen auf molekularer Ebene untersuchen. Dafür ist die Anwendung gentechnischer Methoden zwar unerlässlich, allerdings ist Molekularbiologie nicht mit Gentechnik gleichzusetzen. Das erkennt man auch daran, dass Molekularbiologie eine Wissenschaft ist, während Gentechnik eine Methodensammlung darstellt, die es uns erlaubt, die DNA eines Organismus zu isolieren, gezielt zu analysieren und gegebenenfalls auch zu verändern und neu zu kombinieren.

Auch die Begriffe **Gentechnik** und **Biotechnik** bedeuten nicht das Gleiche. Der Begriff Biotechnik wurde im Jahr 1917 durch den ungarischen Ingenieur Karl Ereky (1878 – 1952) geprägt. Er beschrieb damit einen von ihm entwickelten Prozess, mit dem er die Schweinemast verbesserte, indem er aus billigen Zuckerrüben wertvolles Schweinefutter herstellte. Unter Biotechnik verstand Ereky solche Prozesse, in denen ein Rohmaterial mithilfe lebender Organismen zu einem Produkt veredelt werden konnte. Fast 45 Jahre lang blieb der Begriff Biotechnik recht vage definiert. Das änderte sich im Jahr 1961, als der schwedische Mikrobiologe Carl-Göran Hedén (1920 – 2009) für ein wis-

senschaftliches Journal, das sich der angewandten Mikrobiologie und industriellen Fermentation widmete, den Titel *Biotechnology and Bioengineering* vorschlug. Seit dieser Zeit verstehen wir die Biotechnologie (Biotechnik) als industriellen Prozess, bei dem biologische Organismen, biologische Systeme und biochemische Reaktionen zur Herstellung von Produkten eine entscheidende Rolle spielen. Biotechnik gründet sich also unter anderem auf den Erfahrungen aus Mikrobiologie, Biochemie und Verfahrenstechnik.

Ein typischer biotechnischer Prozess mit Zellen als Produzenten eines Produkts besteht in Anlehnung an Karl Ereky's ursprüngliche Beschreibung aus drei Phasen:

- Im **Upstream-Processing** genannten Abschnitt der biotechnischen Produktion wird die verwendete Produktionskultur zunächst für das Wachstum und die Produktion vorbereitet.
- Anschließend folgt die **Fermentation**, in der die Produktionskultur in großem Maßstab vermehrt wird und in der ein Produkt gebildet oder durch Biotransformation biochemisch verändert wird.
- Es schließt sich das **Downstream-Processing** an, also die Reinigung des Produkts entweder aus den Produktionszellen oder dem Kulturmedium.

In der biotechnischen Forschung wird versucht, jeden einzelnen dieser Schritte zu optimieren. In den 1960er- und 1970er-Jahren konzentrierte sich die Forschung auf das Upstream-Processing, auf Bioreaktor-Konzepte und das Downstream-Processing. Diese Studien führten zu enormen Fortschritten, vor allem im Anlagenbau und in der Überwachung und Steuerung von Fermentationsprozessen. Als weitaus schwieriger erwies sich die Suche nach neuen Organismen für die Optimierung von Biotransformationsprozessen. Zwar bieten natürlich vorkommende Stämme ein unglaubliches Spektrum biosynthetischer Möglichkeiten mit bemerkenswerter Spezifität und Selektivität. Allerdings sind diese biochemischen Reaktionen für den biotechnischen Einsatz keineswegs optimiert.

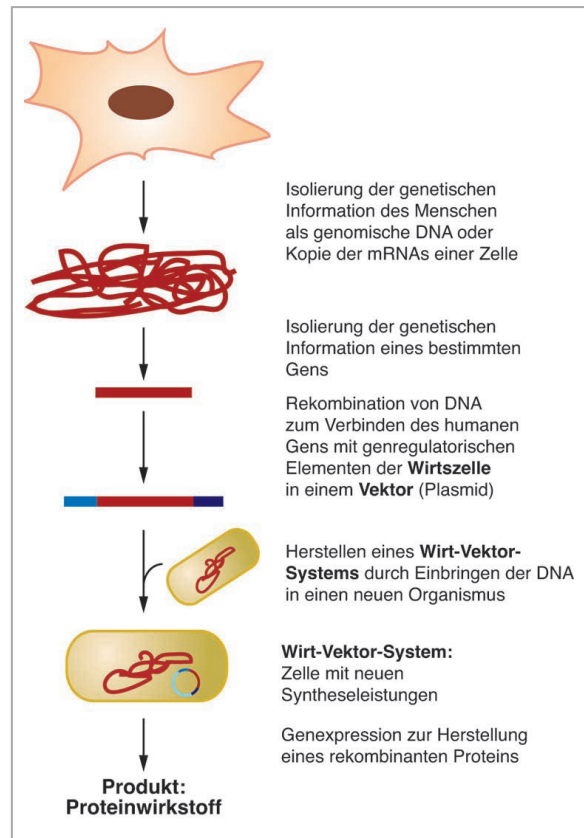
Dies wurde teilweise mit Methoden der klassischen Genetik korrigiert. Durch die Einführung von Mutationen in die Genome von Organismen gelang es, die Produktivität bestimmter Prozesse beachtlich zu steigern. Allerdings wird in solchen klassischen Selektionsverfahren die DNA nicht aus dem ausgewählten Organismus isoliert, um sie gezielt zu verändern (dies würde gentechnische Methoden voraussetzen). Mutationen werden vielmehr ungerichtet durch physikalische oder chemische Mutagene erzeugt und die gewünschte Eigenschaft danach durch Selektion von Mutanten isoliert. Das sind langwierige Prozesse, die noch dazu darauf limitiert sind, dass man lediglich die vorhandenen Syntheseleistungen der ausgewählten Organismen opti-

mieren, aber keine neuen Syntheseleistungen hinzufügen kann.

Die Produkte biotechnischer Verfahren sind aus dem täglichen Leben nicht mehr wegzudenken. Sie reichen von Bier, Wein und anderen Lebens- und Genussmitteln über Metabolite wie Vitamine und Aminosäuren zu Produkten des Sekundärstoffwechsels wie Antibiotika oder Alkaloiden. Durch Biotransformation können häufig leicht zugängliche Ausgangsstoffe in hochwertige Therapeutika umgewandelt oder komplexe sekundäre Pflanzenstoffe wie Taxane hergestellt werden, die aus natürlichen Quellen nicht in ausreichenden Mengen zur Verfügung stünden. Als biologische Quellen für biotechnische Verfahren wird heute eine breite Palette von Organismen in Form sogenannter Hochleistungsstämme etabliert – Bakterien, Pilze, Algen sowie pflanzliche und tierische Zellen sind in solchen Prozessen weit verbreitet.

Was verstehen wir unter Gentechnik? Gentechnik ist eine Sammlung von Methoden, mit denen man DNA aus einem Organismus isolieren, *in vitro* rekombinieren und dann in einen anderen Organismus einschleusen kann, beispielsweise um dort ein bestimmtes Produkt herzustellen (Abb. 1.4). Insofern können die Bereiche der Biotechnik und Gentechnik durchaus überlappen. Das Ergebnis einer mit gentechnischen Methoden optimierten biotechnischen Produktion kann dem entsprechen, was man auch durch den Einsatz von klassischen Selektionsverfahren erreichen kann. Beispielsweise kann ein Kontrollmechanismus sowohl durch ungerichtete Mutagenese als auch durch gezielte Modifikation an einer isolierten DNA außer Kraft gesetzt werden, um so eine Produktionssteigerung eines intrazellulären Metaboliten zu bewirken. Darüber hinaus können durch gentechnische Methoden aber auch die natürlicherweise vorhandenen Syntheseleistungen von Organismen erweitert werden. So kann mit gentechnischen Methoden erreicht werden, dass etablierte Produktionsstämme durch das Einführen fremder genetischer Information gezielt in einer Weise verändert werden, dass sie praktisch jede beliebige fremde Syntheseleistung durchführen können (Abb. 1.4). Woher die genetischen Korrelate für diese Syntheseleistungen stammen, ist letztlich nicht mehr von Bedeutung, da mit gentechnischen Methoden der Austausch genetischer Informationen nicht nur über Artgrenzen hinweg möglich ist, sondern selbst zwischen sehr unterschiedlichen Organismen beispielsweise aus dem Tier- und Pflanzenreich oder zwischen Pro- und Eukaryonten.

Die Vereinigung von Gentechnik und Biotechnik zu einer **Molekularen Biotechnologie** ist besonders markant in der Produktion von Proteinwirkstoffen (Biopharmazeutika; Abb. 1.4). Die Nutzung dieser Wirkstoffe, die sich in der Herstellung so fundamen-



• **Abb. 1.4** Kombination von Gentechnik und Biotechnik zur Herstellung von Biopharmazeutika. Zunächst muss die genetische Information für das als Wirkstoff entwickelte Protein isoliert werden, in der Regel aus dem menschlichen Genom. Die genetische Information dafür steht entweder in Form der genomischen DNA im Zellkern oder als Pool aller mRNAs im Zytoplasma einer Zelle zur Verfügung. Die genetische Information für das gesuchte Gen wird aus diesem Pool isoliert und in einem Vektor (Plasmid) mit Kontrollelementen für die Expression in einer geeigneten Wirtszelle rekombiniert. Diese rekombinierte DNA wird dann in eine geeignete Produktionszelle eingeschleust. Es entsteht ein Wirt-Vektor-System, das in einem geeigneten biotechnischen Verfahren zur Produktion des rekombinanten Proteins genutzt werden kann.

tal von chemisch-synthetischen „kleinen Molekülen“ unterscheiden, wäre ohne Anwendung gentechnischer Methoden nicht denkbar, denn fast immer geht es um die Produktion menschlicher Proteine in nichtmenschlichen Produktionszellen, seien es Bakterien, eukaryontische Mikroorganismen oder kultivierte tierische oder pflanzliche Zellen, und damit um die Hinzufügung einer artfremden Syntheseleistung in den Produzenten. In den folgenden Kapiteln werden wir die Konzepte der Entwicklung und Herstellung von Biopharmazeutika mit gentechnischen Methoden ausführlich darstellen – vom Gen zum Wirkstoff.

### Zusammenfassung

- Die Begriffe Molekularbiologie, Biotechnik und Gentechnik sollten nicht verwechselt oder synonym verwendet werden.
- Der Begriff Molekularbiologie beschreibt in der engen Auslegung eine Wissenschaft, die Nukleinsäuren von Zellen und die Mechanismen der DNA-Replikation und Genexpression untersucht. Eine Erweiterung des Begriffs schließt im Prinzip alle Konzepte und Methoden mit ein, die uns heute die Analyse zellulärer Vorgänge – die ja letztlich immer genetisch gesteuert sind – auf molekularer Ebene ermöglichen.
- Biotechnik (Biotechnologie) befasst sich mit der Etablierung industrieller Prozesse zur Produktion von Stoffen mit biologischen Organismen oder biochemischen Reaktionen zur Herstellung von Produkten. Biotechnik kommt im Prinzip ohne Gentechnik aus.
- Allerdings kann die Kombination von Gentechnik und Biotechnik die Selektion von Hochleistungs-Produktionsstämmen erleichtern und die Produktion von Stoffen ermöglichen, die eigentlich nicht der natürlichen Syntheseleistung eines Organismus entsprechen. Letzteres ist insbesondere im Bereich der Produktion therapeutischer Proteine von Belang.
- Die Produktion von Proteinwirkstoffen in gentechnisch veränderten Organismen ist nur durch eine Symbiose von Biotechnik und Gentechnik möglich, die heute mit Begriffen wie Pharmazeutische Biotechnologie oder Molekulare Biotechnologie umschrieben wird.

## 2 DNA-Rekombinationstechnologie

2.1	Die DNA	9
2.2	Enzymatische DNA-Synthese	18
2.3	Prinzipien der Gentechnik	34

### 2.1 Die DNA

In allen lebenden Organismen bildet die Desoxyribonukleinsäure (*desoxyribonucleic acid*, DNA) den molekularen Träger der Erbinformation. Der Begriff DNA beschreibt nicht etwa ein bestimmtes Molekül, sondern lediglich eine Molekülart. Es ist klar, dass die genetische Information in jeder Spezies verschieden sein muss. Doch ist es nicht die Molekülart DNA an sich, die in jeder Spezies variiert, sondern die in der DNA codierte Information.

DNA-Moleküle sind das Zielobjekt der Gentechnik. Dabei verstehen wir die Gentechnik als eine Technologie, die es uns erlaubt, DNA aus einem Organismus zu isolieren, die in der DNA codierte Information zu verstehen, DNA-Moleküle gegebenenfalls miteinander zu rekombinieren und diese Information in einem anderen Organismus zu realisieren. Insbesondere streben wir zwei Ziele an, wenn wir das Instrumentarium der Gentechnik einsetzen:

- Die DNA wird molekular charakterisiert. Das bedeutet, dass wir die Abfolge der DNA-Basen ermitteln, also den Informationsinhalt einer DNA verstehen wollen und untersuchen, wann und wo im komplexen Leben eines Organismus diese Information abgerufen wird.
- Teilbereiche von DNA-Molekülen werden neu kombiniert. Ziel ist es entweder, den Informationsgehalt der DNA zu variieren oder ein bestimmtes Gen mit neuen Kontrollelementen auszustatten, um zum Beispiel die in der DNA gespeicherte Information in einem anderen Organismus abrufbar zu machen.

#### 2.1.1 Aufbau der DNA

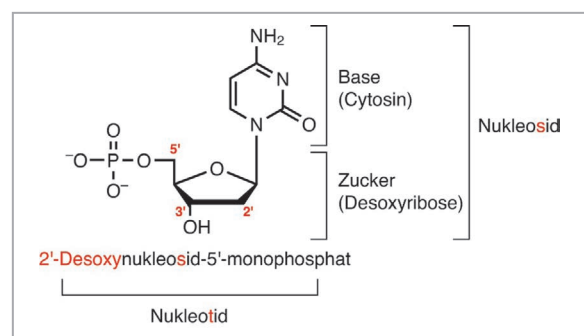
Bei der DNA handelt es sich um sehr große, strukturell sehr einheitlich gebaute Moleküle. Die monomeren Bausteine der DNA sind die Desoxynukleotide (Abb. 2.1). Jeder dieser Bausteine besteht aus drei Komponenten:

- einer Purin- oder Pyrimidinbase,
- einer Desoxyribose, die *N*-glycosidisch mit der Purin- bzw. Pyrimidinbase verknüpft ist,
- einem anorganischen Phosphatrest, der esterartig mit der 5'-OH-Gruppe der Desoxyribose verknüpft ist.

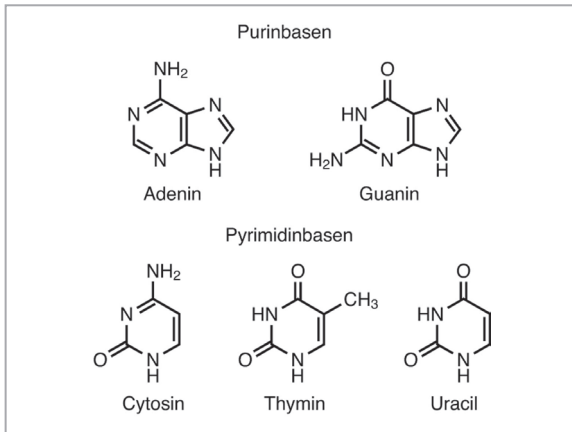
Durch die *N*-glycosidische Verknüpfung einer DNA-Base mit einer Desoxyribose entsteht ein Nukleosid. Wird an die 5'-OH-Gruppe der Desoxyribose eines Nukleosids mindestens eine Phosphatgruppe gebunden, spricht man von einem Nukleotid. In der DNA kommen jeweils zwei Purinbasen (Adenin und Guanin) und zwei Pyrimidinbasen (Thymin und Cytosin) vor (Abb. 2.2). Die Bezeichnungen der in einem DNA-Molekül vorkommenden Nukleotide lauten:

- Desoxycytidin-5'-monophosphat (dCMP),
- Desoxyguanosin-5'-monophosphat (dGMP),
- Desoxythymidin-5'-monophosphat (dTMP),
- Desoxyadenosin-5'-monophosphat (dAMP).

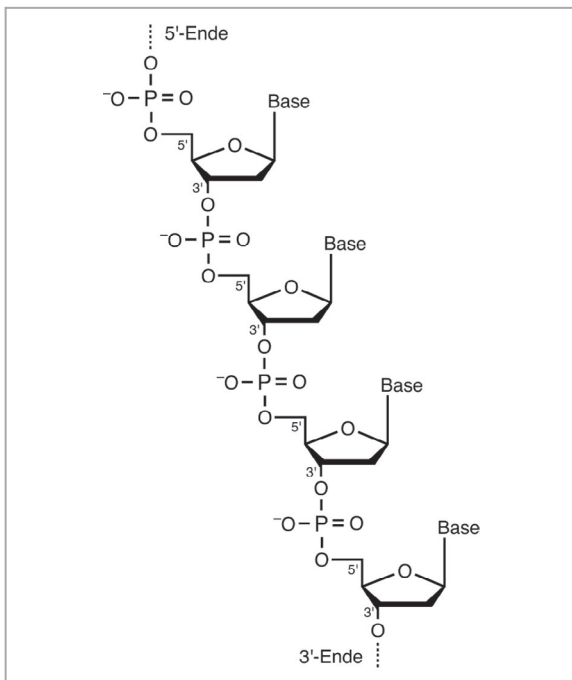
Im normalen Sprachgebrauch werden die DNA-Nukleotide mit den Buchstaben C, G, T und A abgekürzt.



● **Abb. 2.1** DNA-Bausteine. Als Beispiel für einen DNA-Baustein ist das Desoxynukleotid dCMP dargestellt. Die Base ist Cytosin, das Nukleosid ist die Base mit dem Zuckeranteil (Desoxyribose), also 2'-Desoxycytidin. Das Nukleotid umfasst die Base, den Zuckeranteil und die Phosphatgruppe, also 2'-Desoxycytidin-5'-monophosphat.



• **Abb. 2.2** Nucleinsäurebasen. Dargestellt sind die fünf Nucleinsäurebasen Adenin, Guanin, Cytosin, Thymin und Uracil. Thymin und Uracil sind funktionell äquivalent. Allerdings kommt Thymin nur in DNA und Uracil nur in RNA vor.



• **Abb. 2.3** Verknüpfung von Nucleotiden. Die Verknüpfung der Nucleotide erfolgt in der RNA wie in der DNA als Phosphorsäurediester über die 5'- und 3'-OH-Gruppen der Ribosen bzw. Desoxyribosen. Damit erhalten Nucleinsäuren eine eindeutige Polarität, da sie durch diesen Verknüpfungsmechanismus ein 5'-Ende und ein 3'-Ende besitzen.

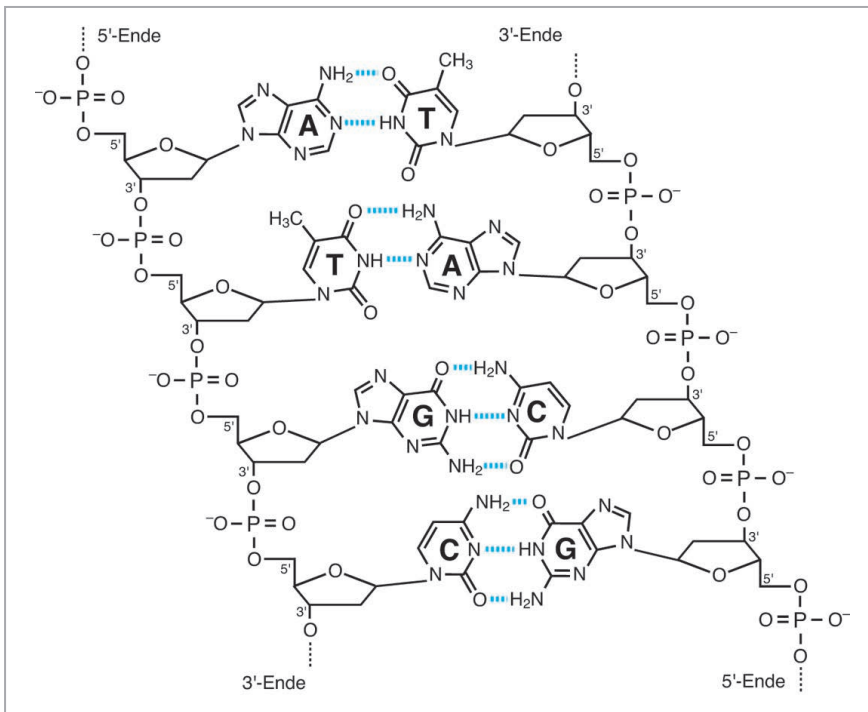
Ribonucleinsäure (RNA) unterscheidet sich von DNA durch das Vorhandensein von Ribose anstelle der Desoxyribose. Außerdem ist in RNA Thymin durch die dritte Pyrimidinbase, Uracil, ersetzt. Bis auf die Tatsache, dass die DNA-Polymerasen 2'-Desoxythymidintriphosphat und die RNA-Polymerasen Uridintriphosphat als Substrat erkennen, sind Thymin und Uracil biologisch äquivalent. Dies gilt insbesondere für die spezifische Basenpaarung mit Adenin.

Innerhalb der DNA sind die einzelnen Nucleotide durch Phosphodiesterbindungen zwischen dem 5'-C-Atom der Desoxyribose des **einen** Bausteins und dem 3'-C-Atom der Desoxyribose des **anderen** Bausteins verknüpft (• Abb. 2.3). Das bedeutet, dass jede DNA-Kette eine eindeutige Richtung hat. Sie hat ein 5'-Ende und ein 3'-Ende, wobei das 5'-Ende der DNA als „DNA-Anfang“ und das 3'-Ende als „DNA-Ende“ definiert wird. Dies ist auch biologisch sinnvoll, da alle Nucleinsäuren – DNA ebenso wie RNA – vom 5'-Ende zum 3'-Ende synthetisiert werden. Daraus leitet sich eine allgemein gültige Konvention ab: Ein einzelner DNA-Strang oder auch eine RNA-Sequenz wird von links nach rechts immer vom 5'-Ende zum 3'-Ende geschrieben. Bei doppelsträngigen DNA-Fragmenten gilt diese Regel immer für den „oberen“ Strang. Wird von dieser Konvention abgewichen, ist es wichtig, das 5'-Ende und das 3'-Ende entsprechend zu kennzeichnen.

Zwei einander komplementäre DNA-Moleküle bilden eine DNA-Doppelhelix, wobei die Zucker-Phosphatketten in entgegengesetzter, antiparalleler Richtung verlaufen. Mit **komplementär** ist gemeint, dass jedem dCMP ein dGMP und jedem dTMP ein dAMP gegenübersteht. Für die gegenseitige Bindung bilden die komplementären Basen schwache Wechselwirkungen in Form von Wasserstoffbrücken aus. Ein G/C-Basenpaar bildet drei und ein A/T-Basenpaar zwei Wasserstoffbrückenbindungen (• Abb. 2.4).

Wegen der Komplementarität von A/T und G/C ist innerhalb eines DNA-Moleküls die Anzahl der G-Nucleotide identisch mit der Anzahl der C-Nucleotide. Ebenso ist die Anzahl der A-Nucleotide gleich der Anzahl der T-Nucleotide. Andererseits ist aber in der Regel die Anzahl der A/T-Basenpaare nicht gleich der Anzahl der G/C-Basenpaare. Vielmehr unterscheidet sich der sogenannte G/C-Gehalt, ausgedrückt in Prozent, in Organismen verschiedener Arten. Generell können wir festhalten, dass der G/C-Gehalt des genetischen Materials eines Organismus nichts über dessen Komplexität aussagt. Denn letztlich ist es die Abfolge der DNA-Basen in den einzelnen, für Informationen codierenden Einheiten (den Genen), die die Individualität eines Organismus bestimmen.

Durch Verdrillung der beiden komplementären DNA-Stränge wird die typische DNA-Doppelhelix gebildet (• Abb. 2.5). Sie erinnert an eine sich nach rechts windende Wendeltreppe. Die Basenpaare bilden in diesem Bild die Stufen der Treppe. Das Rückgrat aus Phosphat- und Zuckermolekülen bildet gewissermaßen das Gelände. Jede helikale Windung der DNA-Doppelhelix besteht aus ca. 10,5 Basenpaaren, sodass jedes Basenpaar gegenüber seinem benachbarten Basenpaar um einen Winkel von ca.  $36^\circ$  verdreht ist. Die Basenpaare bilden einen Winkel von  $90^\circ$  zur Helixlängsachse. Durch diese typische Verdrillung entstehen eine kleine



• **Abb. 2.4** Komplementarität von Nucleinsäuren: Nucleotide sind paarweise zueinander komplementär. Die komplementären Nucleotidpaare A/T (in DNA) bzw. A/U (in RNA) werden durch zwei Wasserstoffbrücken zwischen den Basen stabilisiert. Dagegen bildet das komplementäre Nucleotidpaar G/C drei Wasserstoffbrücken zwischen den Basen aus. Komplementäre Nucleinsäurestränge sind immer antiparallel zueinander angeordnet, d. h. am Ende eines Nucleinsäuredoppelstrangs befindet sich das 5'-Ende des einen und das 3'-Ende des komplementären anderen Nucleinsäurestrangs.

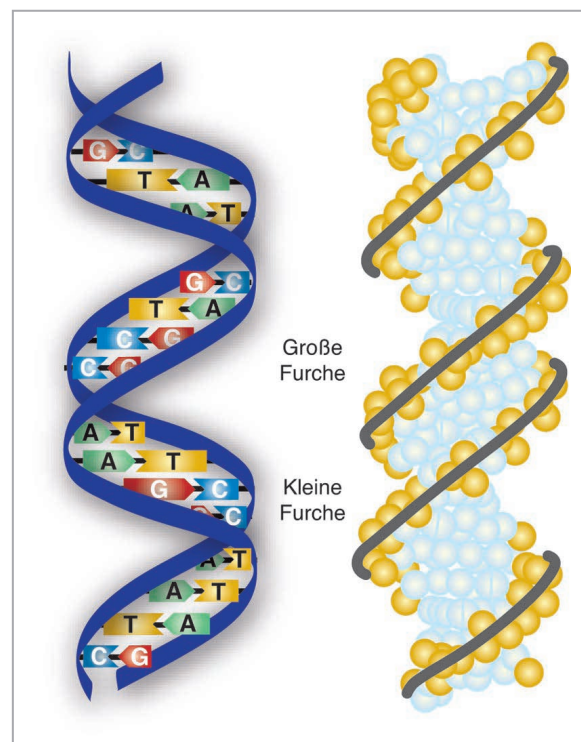
und eine große Furche (• Abb. 2.5). Die beiden Doppelstränge werden ausschließlich über die bereits angesprochenen Wasserstoffbrückenbindungen zwischen G/C- und A/T-Basenpaaren zusammengehalten. Wegen der Vielzahl dieser Wasserstoffbrückenbindungen entlang einer DNA-Doppelhelix ist der Energiegehalt so groß, dass eine DNA-Doppelhelix häufig als ein einziges Molekül angesehen wird, obwohl es sich formal um zwei, nichtkovalent miteinander verbundene Moleküle handelt.

### 2.1.2 Denaturierung von DNA

Die in der DNA-Doppelhelix angeordneten Einzelstränge lassen sich wie beim Öffnen eines Reißverschlusses voneinander trennen. Dieser Prozess wird bei der Replikation und Transkription durch spezielle Enzyme, die DNA-Helikasen, vermittelt. Doch auch durch Erhitzen und durch Zugabe bestimmter Chemikalien ist eine Aufspaltung von DNA-Doppelsträngen experimentell möglich. Diesen Vorgang nennt man in der Gentechnik **Denaturieren** von DNA.

Die Stabilität der DNA hängt vom prozentualen Anteil an G/C-Nucleotidpaaren ab, da diese drei Wasserstoffbrücken ausbilden, während A/T-Paare nur über zwei Wasserstoffbrücken miteinander wechselwirken (• Abb. 2.4). Die Temperatur, bei der 50 % einer doppelsträngigen DNA denaturiert vorliegen, ist die Schmelztemperatur  $T_m$ . Je größer der G/C-Gehalt einer DNA ist, desto höher liegt ihr  $T_m$ -Wert.

Die Gesamtheit der im Zellkern eukaryontischer Zellen vorhandenen DNA ist das **Genom**. Man kann



• **Abb. 2.5** DNA-Doppelhelix: Schematische Darstellung (links) und raumfüllende Darstellung aller Atome (rechts) einer DNA-Doppelhelix. Es wechseln sich immer eine große Furche und eine kleine Furche ab. Eine helikale Windung enthält ca. 10,5 Basenpaare und jedes Basenpaar ist um ca.  $36^\circ$  zueinander verdreht. Die Basenpaare sind im  $90^\circ$ -Winkel zur Helixachse angeordnet. Die Doppelhelix wird durch die parallele Anordnung der flachen Basenpaare signifikant über sogenannte Stacking-Kräfte zwischen den Basen stabilisiert.

den G/C-Gehalt ganzer Genome experimentell bestimmen (● Abb. 2.6). Hierfür wird das als Hyperchromizität bezeichnete Phänomen ausgenutzt, dass einzelsträngige DNA eine höhere Absorption bei 260 nm zeigt als doppelsträngig vorliegende DNA. Setzt man die genomische DNA eines Organismus steigenden Temperaturen aus, wird die DNA zunehmend denaturiert und man misst eine ansteigende Absorption. Aus der sich ergebenden Kurve lässt sich dann der  $T_m$ -Wert ablesen. Vergleicht man diesen gemessenen  $T_m$ -Wert mit dem anderer Organismen, deren G/C-Gehalte vorher schon einmal bestimmt wurden, kann man den G/C-Gehalt des untersuchten Genoms an einer Standardgeraden ablesen (● Abb. 2.6). Heute wird der G/C-Gehalt vieler Organismen im Rahmen von Genomprojekten – also der Aufklärung der DNA-Sequenzen von Gesamt-Genomen – quasi nebenbei bestimmt.

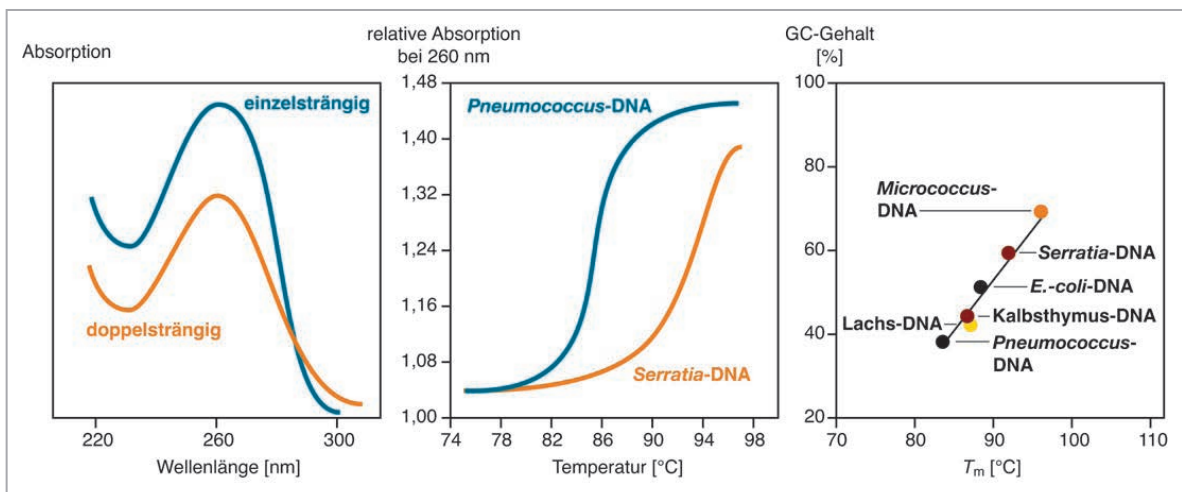
Ebenso wie DNA-Doppelstränge denaturiert werden können, lassen sich komplementäre Einzelstränge auch wieder zu einer Doppelhelix zusammenfügen. Dieser Vorgang wird **Renaturierung** genannt. Die Kinetik, mit der komplementäre DNA-Stränge renaturieren, erlaubt Rückschlüsse auf den G/C-Gehalt und auf die Komplexität der DNA-Moleküle. Erstaunlicherweise erfolgt die Renaturierung auch in den komplexesten Gemischen verschiedener DNA-Spezies hochspezifisch und effizient. Dabei bildet sich immer dann – und nur dann – ein DNA-Doppelstrang aus, wenn zwei komplementäre Partner in einem DNA-Gemisch vorhanden sind. Dieses Prinzip wird uns in vielen Bereichen der Gentechnik

und der Gendiagnostik wieder begegnen, wie zum Beispiel bei der Initiation enzymatischer DNA-Synthese durch sogenannte Primer oder dem spezifischen Nachweis von DNA-Sequenzen in der Gendiagnostik.

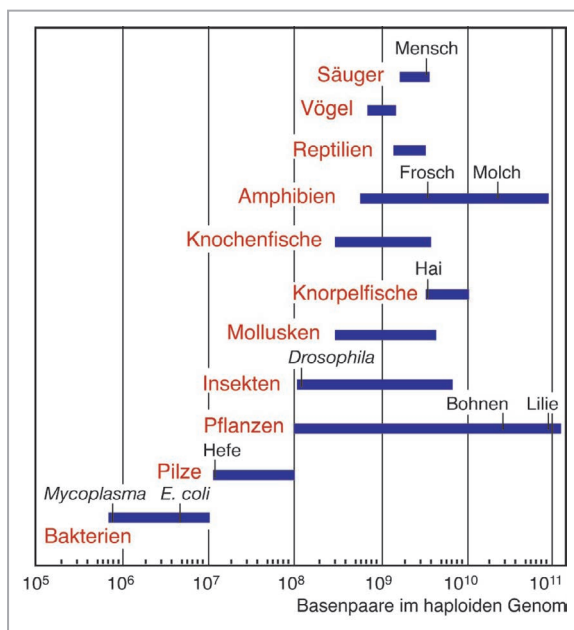
### 2.1.3 Größe und Topologie der DNA

Die Größe der DNA-Moleküle stellt beim gentechnischen Arbeiten mit isolierter DNA eine besondere Herausforderung dar (● Abb. 2.7). Ein Genom wie das des Darmbakteriums *Escherichia coli* besteht aus einem einzigen Molekül in der Größenordnung von ca. 4,6 Millionen Basenpaaren. Dieses Molekül ergibt gespreizt eine Kantlänge von ca. 1,5 mm. Die abgekürzte Schreibweise für die Länge einer DNA in Basenpaaren ist „bp“ (*base pairs*). Das Genom von *E. coli* ist demnach ca. 4 600 000 bp oder 4600 kbp oder 4,6 Mbp groß. Das menschliche Genom ist ungefähr 600-mal größer als das von *E. coli*. Es besteht aus 23 einzelnen Molekülen (Chromosomen), die zusammen etwa 3 270 000 000 bp (3270 Mbp) enthalten und gespreizt ca. 99 cm lang sind. Da die Zellen diploider Organismen jeweils zwei Chromosomensätze enthalten, befindet sich in jeder einzelnen menschlichen Zelle die genetische Information auf DNA-Molekülen, die aneinandergereiht ca. 2 m lang sind.

DNA-Moleküle gibt es in verschiedenen Formen. Sie können fadenförmig linear sein, wie beispielsweise die DNA einiger Viren, die Bakterien befallen (z. B. die Bakteriophagen T7 und T4). Die DNA der Bakteriophagen M13 und f1, die als Vektorsysteme eine Rolle in der Gentechnik spielen, ist ein ringförmiger DNA-Einzelstrang,



● **Abb. 2.6** Bestimmung des G/C-Gehalts eines Genoms: Links: Absorptionsspektren einzelsträngiger und doppelsträngiger DNA. Man erkennt den sogenannten Hyperchromizität-Effekt, also die Zunahme der Absorption bei 260 nm, beim Übergang von doppelsträngiger zu einzelsträngiger DNA. Mitte: Den Hyperchromizität-Effekt kann man ausnutzen, um den G/C-Gehalt eines Genoms experimentell zu bestimmen. Dazu setzt man eine genomische DNA ansteigenden Temperaturen aus und misst die Absorption bei 260 nm. Ändert sich die Absorption bei steigender Temperatur nicht mehr, ist die DNA maximal denaturiert. Aus der Kurve kann man die halbmaximale Denaturierungstemperatur  $T_m$  ableiten. Rechts: Der G/C-Gehalt eines Genoms steht in linearer Beziehung zu  $T_m$ . Trägt man den Messwert  $T_m$  gegen die bereits zuvor bestimmten G/C-Gehalte anderer Organismen auf, kann man den G/C-Gehalt eines untersuchten Organismus auf der Geraden ablesen.

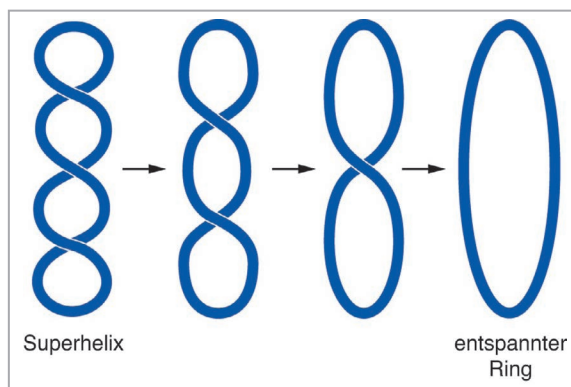


• Abb. 2.7 Vergleich der Genomgrößen verschiedener Organismen

der nach der Infektion von den Enzymen der Wirtszelle zu einem DNA-Doppelstrang komplettiert wird. Dagegen liegt die DNA des Bakteriophagen Lambda in einem linearen, doppelsträngigen DNA-Molekül vor, das erst nach der Infektion ringförmig geschlossen wird. Bakterielle Genome sind in der Regel ringförmig geschlossene, doppelsträngige DNA-Moleküle, während sich die genomische DNA im Zellkern der Eukaryonten auf mehrere lineare Moleküle, die Chromosomen, verteilt. Allerdings kommen in Eukaryonten zusätzlich auch noch ringförmige DNA-Moleküle in Form der DNA der Mitochondrien und der Chloroplasten vor.

Ringförmige DNA-Moleküle wie beispielsweise die Genome von Bakterien, Mitochondrien oder Chloroplasten haben eine wichtige topologische Besonderheit: Die beiden DNA-Stränge der Doppelhelix können nicht vollständig voneinander getrennt werden. Auch lineare DNA-Moleküle sind ab einer gewissen Länge nicht mehr vollständig in zwei Einzelstränge trennbar. Ringförmig geschlossene DNA liegt zudem meist verdrillt vor (•Abb. 2.8). Diese Verdrillungen überlagern die eigentliche DNA-Doppelhelix und sind als Superhelices sichtbar. Die Zahl der Verdrillungen (*supercoils*) ist dynamisch und kann durch zelluläre Enzyme beeinflusst werden. DNA-Moleküle, die bei identischer DNA-Sequenz in verschiedenen Konformationen vorliegen, nennt man DNA-Topoisomere. Dementsprechend nennt man Enzyme, die die Konformation einer DNA verändern können, DNA-Topoisomerasen.

Supercoils sind nicht nur in zirkulären DNA-Molekülen vorhanden, sondern auch innerhalb von Teilbereichen sehr großer, linearer Moleküle, beispielsweise



• Abb. 2.8 DNA-Topoisomere: Ringförmig geschlossene, doppelsträngige DNA-Moleküle liegen häufig verdrillt als Superhelix vor. Die Verdrillungen verhindern, dass die beiden Stränge einer DNA-Helix über eine längere Distanz hinweg zu Einzelsträngen geöffnet werden können, wie es beispielsweise für die Replikation oder Transkription notwendig ist. Um Verdrillungen aus der DNA zu entfernen, öffnen Typ-1-Topoisomerasen einen der beiden DNA-Stränge, drehen diesen Strang einmal um die helikale Achse und verknüpfen die freien DNA-Enden wieder kovalent. Dadurch werden sukzessive Windungen aus der DNA entfernt. Typ-2-Topoisomerasen spalten beide DNA-Stränge gleichzeitig und fädeln den intakten DNA-Doppelstrang durch den entstandenen DNA-Doppelstrangbruch, sodass ebenfalls Verdrillungen entfernt werden.

der Chromosomen von Säugetierzellen. Die Verdrillungen entstehen im Prinzip durch eine Drehung der DNA-Doppelhelix um ihre Längsachse. Dies führt zu Drehungsspannungen, die sich in der Bildung einer Superhelix äußern. Anders ausgedrückt entstehen Supercoils, wenn man das eine Ende der DNA um ihre Längsachse dreht, während das andere Ende festgehalten wird. Eine Konsequenz der Supercoils ist, dass ein DNA-Doppelstrang nicht über längere Strecken in Einzelstränge getrennt werden kann, da die Drehungsspannung innerhalb der Superhelix dem Prozess entgegenwirkt.

In der Zelle kann die Überdrillung einer DNA nicht dadurch verringert werden, dass die DNA-Doppelhelix um ihre Längsachse rotiert. Dazu ist das DNA-Molekül in der Regel schlicht zu lang. Damit es zur Einführung oder Entfernung einer Überdrillung und zur Ausbildung superhelikaler Formen kommen kann, muss an bestimmten Stellen der DNA mindestens einer der beiden DNA-Stränge hydrolytisch geöffnet werden. Dieser geöffnete Strang kann sich nun lokal frei um den intakten Strang drehen und damit den Verdrillungsgrad ändern. Dieser Vorgang wird durch DNA-Topoisomerasen katalysiert, die somit eine große Rolle bei der Transkription, Rekombination und Replikation spielen und für jede Zelle überlebenswichtig sind. Aus diesem Grund sind sie auch wichtige therapeutische Ziele für Antibiotika, wie beispielsweise die Gyrase-Inhibitoren.

Inaktiviert man die bakterielle Topoisomerase durch einen Gyrase-Hemmer, werden die DNA-Replikation und andere wichtige Stoffwechselprozesse pathogener Prokaryonten rasch blockiert.

Es gibt im Wesentlichen zwei Typen von DNA-Topoisomerasen:

- **Typ-1-DNA-Topoisomerasen** binden an die DNA und öffnen **einen** der beiden DNA-Stränge der Doppelhelix. Es lassen sich zwei Subtypen von Typ-1-DNA-Topoisomerasen unterscheiden:
  - Die bakteriellen **Typ-1A-Topoisomerasen** spalten einen DNA-Strang, indem sie eine Phosphodiesterbindung öffnen und eine Phosphatbrücke zwischen dem 5'-Ende der DNA und einem Tyrosinrest im aktiven Zentrum des Enzyms bilden. Durch diese kovalente Bindung an das Enzym wird die Energie der Phosphodiesterbindung konserviert und der geöffnete DNA-Strang wird am Enzym fixiert. Das Enzym veranlasst eine vollständige Drehung der DNA an der Stelle der Strangöffnung um seine eigene zentrale Achse und entfernt somit eine superhelikale Windung. Durch Transesterifizierung wird dann unter Freisetzung des Enzyms die 5'-Phosphatgruppe wieder auf die freie 3'-OH-Gruppe des gespaltenen DNA-Strangs übertragen und dadurch der zuvor geöffnete DNA-Strang wieder geschlossen. Dieser Vorgang benötigt keine chemische Energie in Form von ATP, da bei der Übertragung des DNA-Strangs auf den Tyrosinrest des Enzyms keine Energie verloren geht.
  - **Typ-1B-Topoisomerasen** kommen in Eukaryonten vor. Sie spalten ebenfalls einen der beiden Stränge einer DNA, allerdings übertragen sie dabei eine 3'-Phosphatgruppe auf einen Tyrosinrest im aktiven Zentrum des Enzyms. Nach der Rotation um die Längsachse der DNA wird die 3'-Phosphatgruppe wieder auf die verbliebene 5'-OH-Gruppe des gespaltenen DNA-Strangs zurückübertragen und der DNA-Strang dadurch geschlossen.
- Eine **Typ-2-DNA-Topoisomerase** bindet an die DNA und öffnet **beide** DNA-Stränge gleichzeitig. Die bakterielle Typ-2-DNA-Topoisomerase, die auch als Gyrase bekannt ist, besteht aus einem Tetramer aus jeweils zwei GyrA- und GyrB-Untereinheiten. Die Gyrase wirkt über folgende Teilschritte (Abb. 2.9): Zunächst bindet das Enzym an die DNA, wobei sich eine Schleife der DNA von ca. 140 Basenpaaren um das Enzym schlingt. Dann spaltet das Enzym den DNA-Doppelstrang und die GyrA-Untereinheit des Enzyms wird kovalent an das 5'-Ende des gespaltenen DNA-Strangs gebunden. Als nächster Schritt wird ein benachbart gelegenes

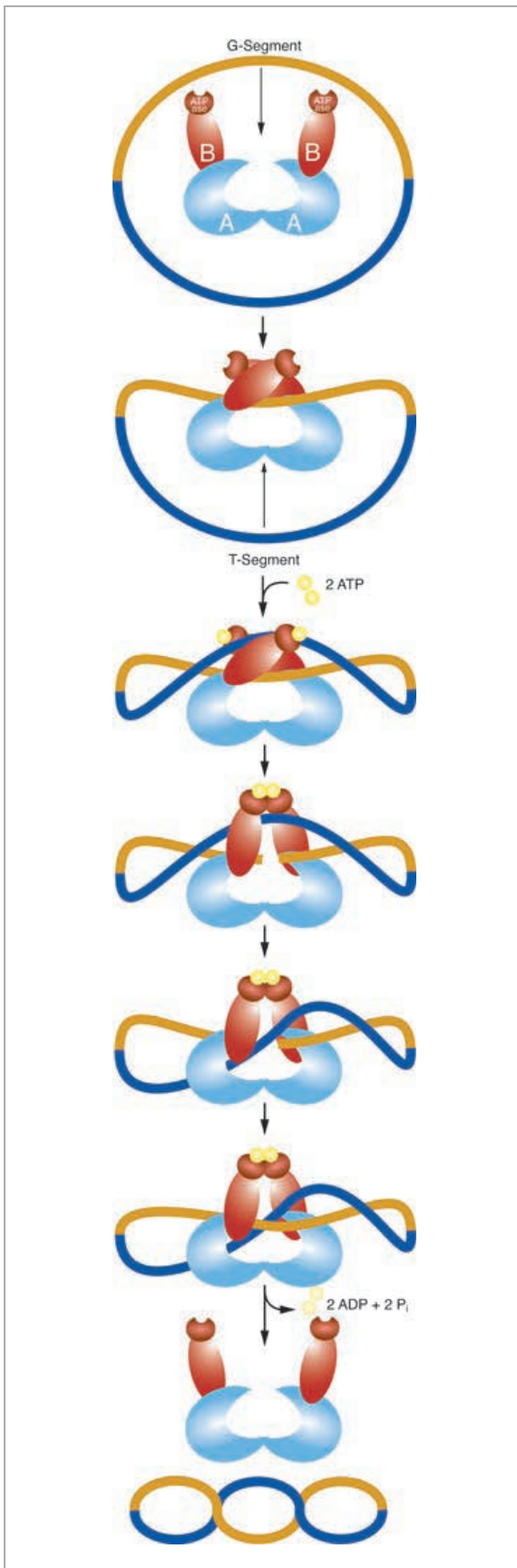
DNA-Segment durch die Öffnung gezogen. Schließlich wird das kovalent gebundene Protein abgelöst und die gespaltenen Doppelstränge werden wieder verknüpft. Bei dieser Reaktionsfolge wird unter Verbrauch eines Moleküls Adenosintriphosphat (ATP) eine superhelikale Windung in die DNA eingeführt, wie in Abb. 2.9 dargestellt. In Abwesenheit von ATP kann das Enzym auch die umgekehrte Reaktion katalysieren und somit zur Relaxation superhelikaler DNA führen.

#### 2.1.4 DNA-Analytik

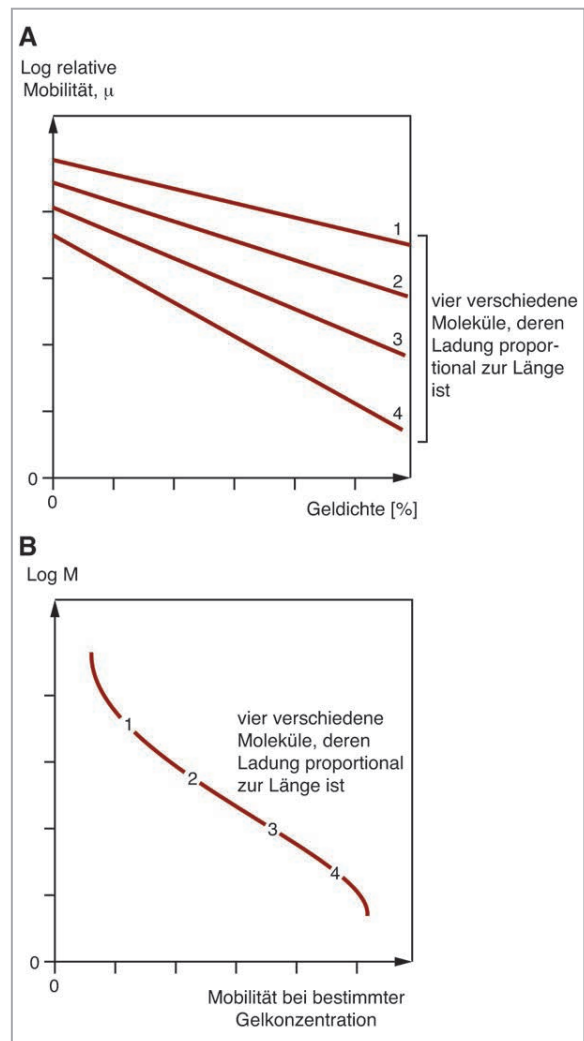
Beim gentechnischen Arbeiten wird ständig mit DNA umgegangen. Dabei werden verschiedene Standardmethoden der DNA-Analytik angewendet, die wir hier kurz besprechen wollen. Als geladene Makromoleküle können DNA oder RNA durch **Gelelektrophorese** getrennt werden. In der Regel werden dazu Matrices aus gelierter Agarose oder aus einem Polyacrylamid-Netzwerk verwendet. Der Logarithmus der relativen Mobilität des Makromoleküls ist linear abhängig von der Gelkonzentration und kann durch einen sogenannten Ferguson-Plot dargestellt werden (Abb. 2.10). Dabei trennt man gleiche Proben in Gelen unterschiedlicher Porengröße auf. Trägt man dann die Logarithmen der relativen Laufstrecken gegen die eingesetzten Gelkonzentrationen auf, erhält man eine Gerade. Die Steigung der Geraden im Ferguson-Plot ist damit ein Maß für die Molekülgröße.

Polyelektrolyte wie Nukleinsäuren tragen an jeder Monomer-Einheit eine Ladungseinheit. Daraus folgt, dass das Verhältnis der Molekülmasse der Nukleinsäure zur Gesamtladung stets konstant ist. Für die relative Mobilität in Gelmatrices ist somit der Siebeffekt – und damit die Moleküllänge – bestimmend (Abb. 2.10). Über einen relativ weiten Größenbereich gibt es eine lineare, experimentell bestimmbare Abhängigkeit zwischen dem Logarithmus ( $\log_{10}$ ) der Länge (in Basenpaaren) von DNA-Molekülen und ihrer relativen Wanderungsdistanz (in cm).

In der Praxis verwendet man 0,5- bis ca. 2 %ige Agarosegele. In 0,5 %igen Gelen lassen sich DNA-Fragmente im Bereich von 1000 bis 15000 bp gut auftrennen. Gele mit einer höheren Agarosekonzentration sind geeignet, um DNA-Fragmente zwischen 100 und 2000 bp zu trennen. Trennt man ringförmig geschlossene Plasmide, so wirkt sich die Konformation der DNA-Moleküle stärker auf die relative Mobilität aus als bei der Trennung linearer Nukleinsäurefragmente. Ringförmige, superhelikale DNA wandert nämlich schneller als ringförmig offene (relaxierte) oder lineare DNA. Das kommt daher, dass sich die dichter gepackte superhelikale DNA schneller durch die Poren des Agarosegels bewegen kann (Abb. 2.11).



• **Abb. 2.9** Reaktionsmechanismus der Gyrase: Gyrase ist eine Typ-2-Topoisomerase, die nacheinander zwei doppelsträngige DNA-Bereiche (G- und T-Segment) bindet. Im Verlauf weiterer Konformationsänderungen des Enzyms, für die zwei Moleküle ATP benötigt werden, wird das G-Segment durch die GyrA-Untereinheit des Enzyms gespalten. Das T-Segment wird dann durch die Lücke im G-Segment gefädelt, sodass – wie in diesem Beispiel gezeigt – eine Verdrillung eingeführt wird. Das Enzym kann auf gleiche Weise auch Verdrillungen aus der DNA entfernen.



• **Abb. 2.10** Ferguson-Plot für verschiedene Polyelektrolyte: (A) Polyelektrolyte wie Nukleinsäuren tragen an jeder Monomer-Einheit eine Ladungseinheit. Demnach tragen verschieden große Nukleinsäuren eine Ladung, die zu der jeweiligen Moleküllänge proportional ist. Für die relative Mobilität in Gelmatrices ist somit der Siebeffekt – und damit die Moleküllänge – bestimmend. (B) Bei einer konstanten Gelkonzentration ist in einem bestimmten Bereich der Logarithmus der Molekülmasse (Log M) proportional zur Mobilität der Moleküle in der Gelmatrix.

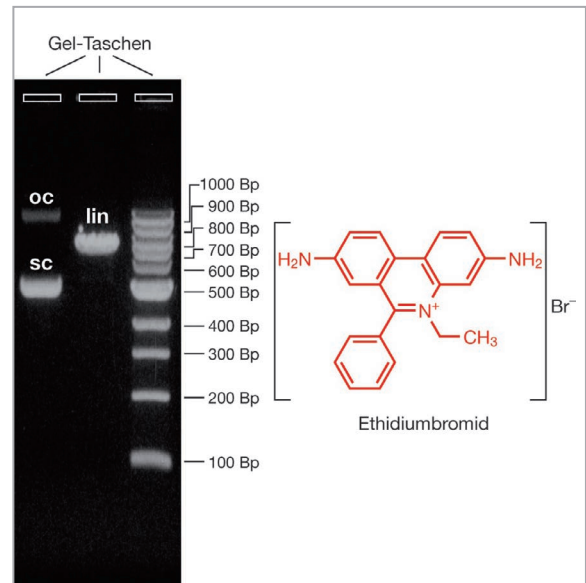
Nachweisen lassen sich die DNA-Fragmente in Gelmatrices beispielsweise durch Anfärbung mit Ethidiumbromid (● Abb. 2.11). Einzelne Ethidiumbromid-Moleküle interkalieren zwischen die DNA-Basen, wobei die aromatischen Ringe des Farbstoffs mit den heteroaromatischen Ringen der Nukleinsäurebasen wechselwirken. Der Farbstoff zeigt nach Anregung mit UV-Licht (254 – 366 nm) eine Fluoreszenz bei 590 nm. Da die Fluoreszenz des Farbstoffs nach der Bindung an DNA stark zunimmt, ist die Orange-Fluoreszenz nur an den Stellen eines Agarosegels deutlich sichtbar, an denen sich DNA befindet, selbst wenn der Farbstoff zuvor gleichmäßig in das Agarosegel eingegossen wurde.

Neben der Größenbestimmung ist ein weiterer wichtiger Parameter im Umgang mit DNA deren Konzentration in einer DNA-Lösung. Nukleinsäuren absorbieren im UV-Bereich mit einem Maximum bei 260 nm (angegeben als  $A_{260}$ ). Demzufolge erfolgt die Bestimmung der DNA-Konzentration in der Regel durch UV-Spektroskopie. Aufgrund des Phänomens der Hyperchromizität (Einzelstrang-DNA absorbiert stärker als Doppelstrang-DNA) ergeben sich aus der UV-spektroskopischen Messung verschiedener Nukleinsäure-Typen unterschiedliche Konzentrationen. Danach entspricht 1  $A_{260}$ :

- 50 µg/ml Doppelstrang-DNA,
- 40 µg/ml Einzelstrang-DNA,
- oder 33 µg/ml (Einzelstrang)-RNA.

Bezüglich der Reinheit einer gegebenen DNA sind Verunreinigungen mit anderen DNAs oder mit Proteinen von Interesse. Verunreinigungen einer DNA mit anderen DNAs können in Grenzen durch Agarosegel-Elektrophorese nachgewiesen werden. Die Nachweisbarkeitsgrenze für eine einzelne Bande in Ethidiumbromid-Gelen liegt bei ca. 10 ng Doppelstrang-DNA. Verunreinigungen mit Proteinen können über den Quotienten  $A_{260}/A_{280}$  abgeschätzt werden. Proteine absorbieren über ihre aromatischen Aminosäuren bei einem Maximum von 280 nm. Eine „saubere“ DNA sollte einen  $A_{260}/A_{280}$ -Quotienten nahe 1,8 aufweisen.

Die Nachweisbarkeitsgrenze der UV-spektroskopischen Vermessung von DNA liegt bei ca. 250 ng/ml. In vielen Fällen hat man es mit wesentlich weniger DNA zu tun, sodass empfindlichere Methoden zur Bestimmung einer DNA-Konzentration erforderlich sind. Eine solche Methode basiert auf der DNA-Hybridisierung. DNA-Doppelstränge können durch Zufuhr von Energie (Temperaturerhöhung) in Einzelstränge überführt werden (► Kap. 2.1.2). Dieser Vorgang wird Denaturierung genannt. Wird die Temperatur wieder abgesenkt, können die Einzelstränge renaturieren. Wenn nun während der Renaturierung ein kurzes DNA-Molekül im Überschuss zugesetzt wird, das zu einem bestimmten Teil der denaturierten DNA komplementär ist, kann



● **Abb. 2.11** Gelelektrophoretische Trennung von DNA-Topoisomeren: Die DNA-Fragmente werden in „Taschen“ am oberen Gelrand aufgetragen. Dann wird an das Gel eine Gleichspannung angelegt, wobei die Anode unten und die Kathode oben angeschlossen wird. Als Polyanionen wandern die DNA-Fragmente nach unten auf die Anode zu, wobei kleinere Fragmente schneller wandern als größere. Zum Nachweis der DNA enthält das Agarosegel gleichmäßig verteilt Ethidiumbromid. Während der Elektrophorese „sammelt“ die DNA Ethidiumbromid-Moleküle auf, da diese bei Kontakt mit DNA zwischen die Basenpaare interkalieren. Durch die Interkalation wird die Fluoreszenz des Ethidiumbromids stark erhöht, sodass im UV-Licht nur dort eine intensiv orange Fluoreszenz sichtbar ist, wo sich im Gel DNA befindet. In der rechten Spur ist ein DNA-Größenstandard aufgetragen worden. In der linken Spur lief eine Präparation von Plasmid-DNA. Die ringförmigen Plasmide liegen entweder als entspannte Ringe (*open circle*, oc) oder als superspiralisierte Form (*supercoiled*, sc) vor. Da die Supercoiled-DNA wesentlich kompakter ist als entspannte DNA-Ringe, läuft sie im Agarosegel schneller. In der mittleren Spur haben wir die Plasmid-DNA vorher mit einer Restriktionsendonuklease geschnitten, sodass ein linearer DNA-Strang (lin) resultiert. Die Mobilität dieser DNA liegt zwischen der Konformation *open circle* und *supercoiled*.

das kurze Oligonukleotid mit einem der längeren DNA-Einzelstränge bei der Bildung von Doppelsträngen konkurrieren. Bindet auf diese Weise gezielt ein kurzes DNA-Molekül an einen längeren DNA-Einzelstrang, spricht man bei diesem Vorgang von Hybridisieren. Die Hybridisierung von Nukleinsäuren ist eine fundamentale Methode der Gentechnologie. Sie wird beispielsweise zum qualitativen Nachweis von DNA und für die Konzentrationsbestimmung von DNA verwendet und auch, um die enzymatische In-vitro-DNA-Synthese mit sogenannten Primern zu starten.

Wie können wir die Hybridisierung einsetzen, um die Konzentration einer DNA-Lösung zu bestimmen? Stellen wir uns vor, dass wir von der DNA-Lösung, deren Konzentration bestimmt werden soll, eine Serienverdünnung herstellen und jeweils ein Aliquot davon auf eine Nitrocellulosemembran tropfen (*dot blot*). Außerdem tropfen wir auf den Filter eine Konzentrationsreihe einer DNA, deren Konzentration uns bekannt ist (Standard-Reihe). Stellen wir uns weiter vor, dass wir ein kurzes DNA-Oligonukleotid zur Verfügung haben, das an einem Ende mit einem Fluoreszenzfarbstoff gekoppelt ist (Abb. 2.12). Wir inkubieren die Nitrocellulosemembran mit einer Lösung des markierten Oligonukleotids. Mehrere Waschschriffe in Puffern mit definiertem Salzgehalt und bei definierten Temperaturen ermöglichen die Unterscheidung zwischen spezifisch und unspezifisch gebundener DNA-Sonde: Eine Sonde, die nicht „passt“ (in Abb. 2.12 rechts), wird bei geringeren Temperaturen „abgewaschen“ als die Sonde, die über die gesamte Länge spezifische Basenpaarungen eingehen kann. Nach den Waschschriffen wird die Fluoreszenz des spezifisch an der Proben-DNA gebundenen Oligonukleotids gemessen und mit der Standardreihe in Beziehung gesetzt, um die Konzentration der Ausgangslösung zu errechnen.

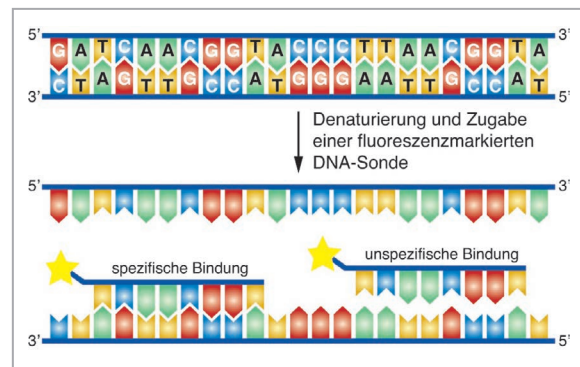


Abb. 2.12 Prinzip der Hybridisierung: Die vier DNA-Basen sind durch die unterschiedlichen Formen und Farben der Symbole dargestellt. Durch Erhitzen wird der DNA-Doppelstrang in die beiden Einzelstränge denaturiert. Wir geben ein kurzes Oligonukleotid hinzu, das für einen kleinen Bereich eines der beiden DNA-Doppelstränge exakt komplementär ist und am 5'-Ende mit einem Fluoreszenzfarbstoff gekoppelt ist. Die markierte Sonde kann über Basenpaarungen an den denaturierten DNA-Strang binden. Durch Waschen mit geeigneten Puffern und bei steigenden Temperaturen entfernt man die unspezifisch gebundenen Sonden (die Fehlpaarungen aufweisen) früher als die spezifisch gebundenen Proben (weil jede Base paaren kann). Der große Überschuss der zugegebenen Sonde sorgt in der Konkurrenz mit der Reassoziierung der beiden DNA-Stränge für eine zuverlässige Bindung der Sonde.

### Zusammenfassung

- Das Zielmolekül aller gentechnischen Methoden ist die DNA. Der Begriff DNA beschreibt allgemein eine Molekülart, während die Einzigartigkeit eines DNA-Moleküls durch die Sequenz der vier DNA-Basen bestimmt wird.
- DNA besteht aus drei Grundbausteinen: einer Purin- oder Pyrimidinbase, einer N-glycosidisch daran gebundenen Desoxyribose und einem Phosphatrest, der esterartig mit dem Zucker verbunden ist. Der Zucker und das Phosphat bilden das Rückgrat der DNA.
- Die DNA-Basen können mit einem anderen DNA-Strang komplementärer Basensequenz einen DNA-Doppelstrang ausbilden, wobei die beiden DNA-Stränge durch die Bildung von Wasserstoffbrückenbindungen zwischen komplementären Basen stabil miteinander verbunden werden. Die Einzelstränge in einer DNA-Doppelhelix verlaufen antiparallel und haben definierte Enden, die als 5'-Ende bzw. als 3'-Ende bezeichnet werden.
- Wegen der Komplementarität enthalten DNA-Moleküle immer ebenso viele A- wie T-Nukleotide und ebenso viele G- wie C-Nukleotide. Allerdings kann die Anzahl der A/T- und der G/C-Paare je nach Organismus deutlich variieren.
- Da die beiden komplementären DNA-Stränge nur durch schwache Wechselwirkungen zusammengehalten werden, lassen sie sich durch Energiezufuhr trennen („denaturieren“).
- Die Reorganisation zweier komplementärer Einzelstränge zu einem DNA-Doppelstrang wird als „renaturieren“ bezeichnet. Denaturierung und Renaturierung bilden die Basis für die wichtige gentechnische Technik des Hybridisierens.

#### 2.1.5 Chemische DNA-Synthese

Für viele gentechnische Arbeitsgänge, wie beispielsweise die enzymatische In-vitro-Synthese von DNA, sind kurze, einzelsträngige Oligonukleotide, sog. Primer, nötig. Derartige Moleküle werden chemisch-syn-

thetisch hergestellt. Heutzutage haben sich Firmen darauf spezialisiert, für wenig Geld derartige Primer innerhalb von wenigen Stunden nach der Bestellung der gewünschten Sequenz zu liefern. Die Chemie, die dahintersteckt, ist ausgeklügelt und recht aufwendig.

Das Phosphoramidit-Verfahren zur chemischen DNA-Synthese sei hier dargestellt:

Ausgangspunkt für eine chemische Synthese eines DNA-Oligonukleotids ist ein ausreichend geschütztes Nukleosid, das mit seiner 3'-OH-Gruppe über einen Spacer und eine Dicarbonsäure kovalent an eine feste Matrix gebunden ist. Die Matrix besteht in der Regel aus Silica-Gel oder aus besonderen Glas-Kügelchen (CPG, *controlled pore glass material*). Die Synthese erfolgt in Zyklen, wobei jeder Zyklus aus einer Reihe von Syntheseschritten besteht, die erforderlich sind, um das matrixgebundene Nukleotid bzw. Oligonukleotid um einen Nukleotidbaustein zu verlängern (◉ Abb. 2.13). Anders als die biologische DNA-Synthese wird bei der chemischen DNA-Synthese die Nukleotidkette vom 3'- zum 5'-Ende hin verlängert.

Der Vorteil der matrixgebundenen Oligonukleotidsynthese besteht darin, dass nach jeder chemischen Reaktion die überschüssigen Reagenzien durch intensives Waschen der Säule von dem immobilisierten Produkt getrennt werden können. Folgende Einzelschritte werden im Lauf eines Synthesesyklus bei der heute überwiegend verwendeten Desoxynukleotid-Phosphoramidit-Synthese durchlaufen (◉ Abb. 2.13):

- Im ersten Schritt wird die Dimethoxytrityl-Schutzgruppe (DMT) am 5'-Ende der mit der Matrix verbundenen Oligonukleotidkette durch Säurebehandlung entfernt.
- Das freigewordene 5'-Ende des Oligonukleotids wird an das 3'-Phosphoramidit-Derivat eines Desoxyribonukleosids gekoppelt, das die nächste Position in der Kette einnehmen soll. Das Kopplungsreagenz in dieser Reaktion ist Tetrazol.
- Alle 5'-Enden, die nicht reagiert haben, werden durch Acetylierung (*capping*) inaktiviert, um sie von den folgenden Kopplungsreaktionen auszuschließen. Dadurch wird die Verlängerung „falscher“ Oligonukleotide unterbunden.
- Der bei der Kopplungsreaktion entstandene Phosphit-Triester wird zum Phosphat-Triester oxidiert. Dabei entsteht die um ein Nukleotid verlängerte Kette.
- Dieser Reaktionszyklus kann heute in modernen Syntheseautomaten routinemäßig bis zu 100-mal wiederholt werden.
- Zum Abschluss einer Oligonukleotidsynthese wird das Oligonukleotid von der Festphase abgelöst und die diversen Schutzgruppen werden entfernt. Danach sollte das Oligonukleotid durch eine HPLC-Trennung gereinigt werden, um zu kurze Oligonukleotide abzutrennen und damit die nachfolgenden enzymatischen Reaktionen nicht durch verunreinigende Chemikalien in ihrer Effizienz herabgesetzt werden.

Trotz der vielen Einzelschritte sind die Ausbeuten erstaunlich gut, da die Effizienzen der Einzelreaktionen für chemische Reaktionen so extrem hoch sind. Trotzdem können chemisch hergestellte Oligonukleotide nicht beliebig lang sein, da die kumulativen Ausbeuten der einzelnen Synthesesyklen für das Endprodukt maßgebend sind. Dies wird an folgenden Beispielrechnungen deutlich:

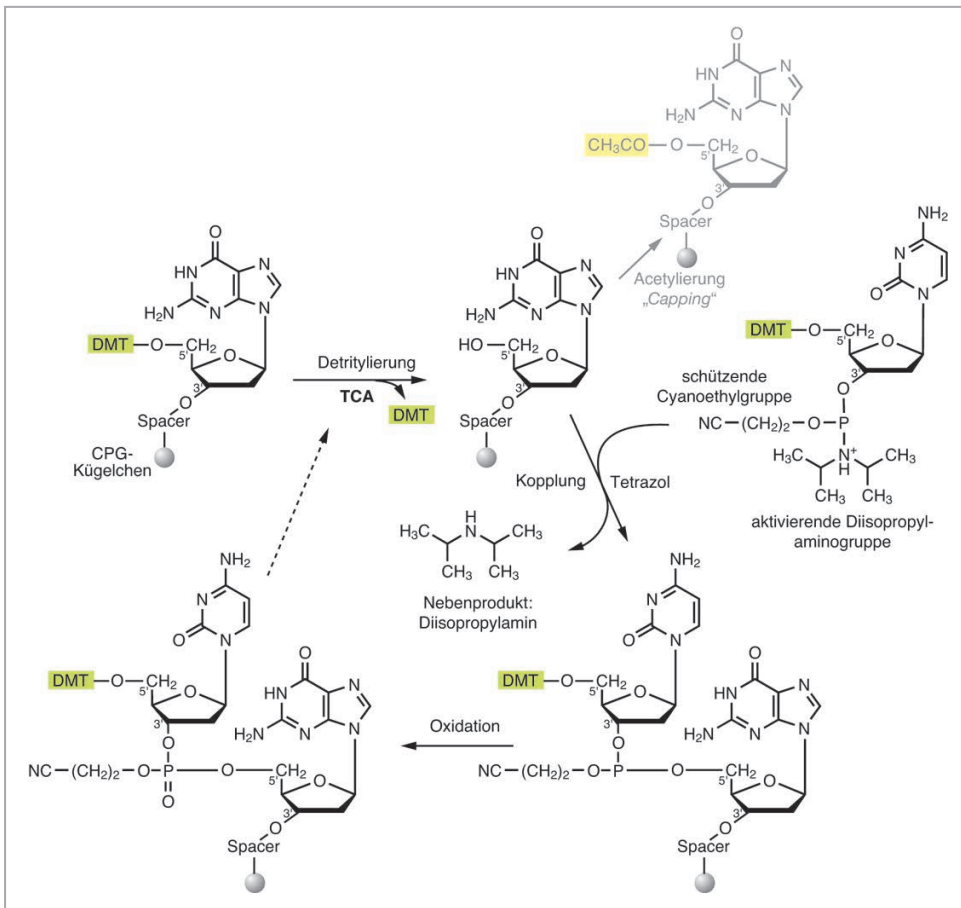
- Wird ein 20 Nukleotide langes Oligonukleotid hergestellt, haben selbst bei 99 % Kopplungseffizienz der Einzelzyklen nur 82 % ( $0,99^{20} \times 100$ ) die korrekte Länge von 20 Nukleotiden. Wäre die Kopplungseffizienz jedoch in jedem Schritt nur 90 %, hätten nur noch 12 % der Produkte die korrekte Länge.
- Die Ausbeute für ein 60-mer-Oligonukleotid liegt selbst bei 99 % Kopplungseffizienz nur bei 55 % ( $0,99^{60} \times 100$ ), bei einem 100-mer sind es nur noch 36 %.

Da die Kopplungseffizienz selbst in gut eingestellten Syntheseautomaten nicht in jedem Schritt nahe 100 % liegen wird, ist die Gesamtausbeute bei Oligonukleotiden mit über 100 Nukleotiden Länge sehr begrenzt. Es sind also in jedem Fall aufwendige Reinigungsverfahren notwendig, um korrekt lange Oligonukleotide von weniger langen Produkten zu trennen. In der Praxis ist die chemische DNA-Synthese heute auf ca. 200 Basen lange Oligonukleotide beschränkt.

## 2.2 Enzymatische DNA-Synthese

### 2.2.1 DNA-Polymerasen

Die DNA synthetisierenden Enzyme nennt man DNA-Polymerasen. Eine ihrer Funktionen ist die semikonservative Replikation der DNA vor der Zellteilung. Für die Details wollen wir an dieser Stelle auf Lehrbücher der molekularen Genetik verweisen. Eine wichtige Eigenschaft der DNA-Polymerasen ist, dass die Enzyme einen einzelsträngigen DNA-Matrizenstrang entsprechend den Komplementaritätsregeln kopieren und dadurch zum DNA-Doppelstrang komplettieren. Daher nennt man diese Enzyme korrekt DNA-abhängige DNA-Polymerasen. Die Enzyme benötigen ein kurzes Stück doppelsträngiger DNA mit einem freien 3'-Hydroxylende als Startpunkt für die DNA-Synthese. Dieses kurze Stück eines doppelsträngigen DNA-Moleküls wird dadurch generiert, dass ein einzelsträngiges Oligonukleotid (Primer) an die 5'-Seite des zu synthetisierenden DNA-Abschnitts hybridisiert wird (◉ Abb. 2.14). Die DNA-Synthese durch eine DNA-Polymerase erfolgt dann durch Verlängerung des Primers an dessen 3'-Ende, d. h. die enzymatische DNA-Synthese verläuft immer in 5' → 3'-Richtung, bezogen auf den neu synthetisierten DNA-Strang.



• **Abb. 2.13**  
Chemische DNA-Synthese nach dem Phosphoramidit-Verfahren

Die DNA-Polymerase I aus *E. coli* besitzt neben einer 5' → 3'-DNA-Polymeraseaktivität auch zwei Exonuklease-Aktivitäten. Diese bauen die DNA entweder in 5' → 3'-Richtung oder in 3' → 5'-Richtung von den Enden her ab. Die 3' → 5'-Exonuklease-Aktivität wird auch Proofreading-Aktivität genannt, da durch diese enzymatische Aktivität vom Enzym quasi überprüft werden kann, ob der erfolgte Polymerisationsschritt korrekt war. Das sogenannte Klenow-Fragment der DNA-Polymerase I aus *E. coli* ist eine modifizierte Variante der DNA-Polymerase I. Es handelt sich um das größere von zwei enzymatisch aktiven Peptidfragmenten, das bei der partiellen Proteolyse der DNA-Polymerase I durch eine bestimmte Protease entsteht. Das Klenow-Fragment besitzt nur noch die 5' → 3'-DNA-Polymeraseaktivität und die 3' → 5'-Exonukleaseaktivität. Die 5' → 3'-Exonukleaseaktivität ist hingegen verlorengegangen.

Welche Bestandteile müssen sich in einem Reaktionsansatz befinden, damit eine In-vitro-DNA-Synthese effizient abläuft? Obwohl jede DNA-Polymerase ihre individuellen Anforderungen an ein Puffersystem stellt, gibt es doch einige allgemeine Regeln:

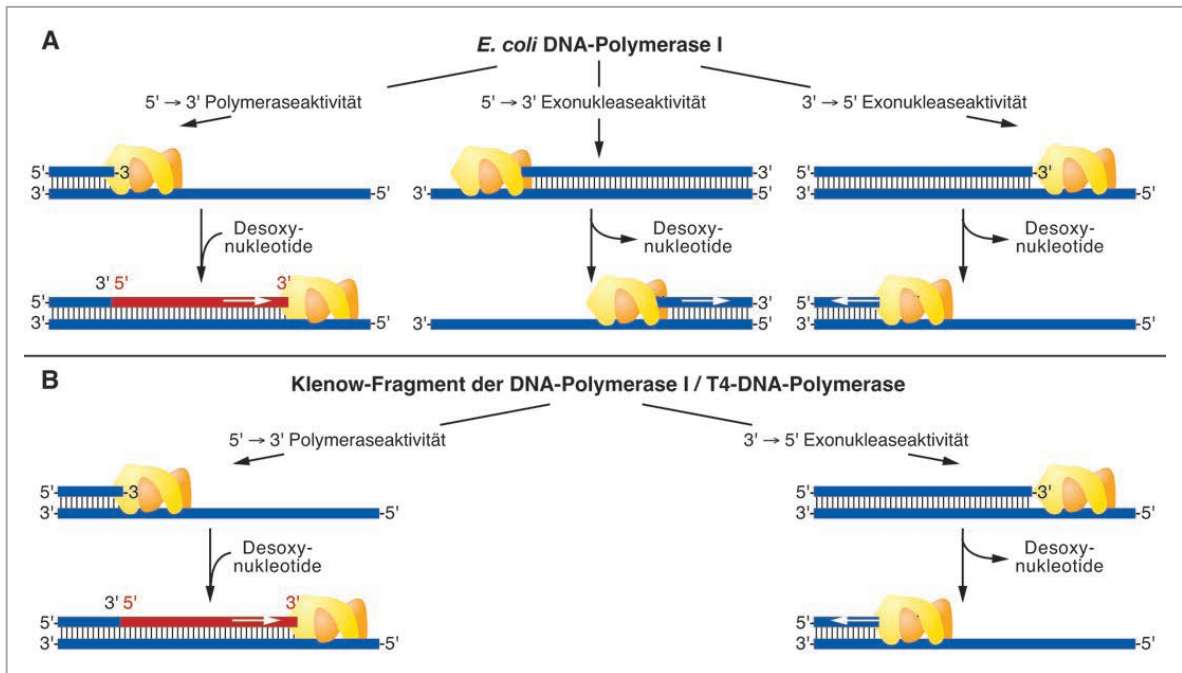
- Zum einen wird natürlich die DNA-Polymerase selbst gebraucht. Heute stammen kommerziell erhältliche

DNA-Polymerasen übrigens aus gentechnischer Produktion.

- Dann muss eine einzelsträngige DNA vorhanden sein, die als Matrize (im Englischen *template*) zum Kopieren durch die DNA-Polymerase dienen kann.
- Ein kurzes Oligonukleotid ist als Primer der DNA-Polymerisation nötig. Diese Primer stammen heute aus der kommerziell von spezialisierten Firmen durchgeführten chemischen DNA-Synthese.
- Außerdem müssen alle vier DNA-Bausteine im Reaktionsansatz vorliegen. Eine Mischung aller vier DNA-Nucleotide wird im molekularbiologischen Sprachgebrauch als Desoxynucleosidtriphosphate oder kurz als dNTPs bezeichnet.
- Alle enzymatischen Reaktionen benötigen ein geeignetes Puffersystem, das in diesem Fall auch Mg<sup>2+</sup>-Ionen enthalten muss, da die eigentlichen Substrate der DNA-Polymerase die Mg<sup>2+</sup>-Komplexe der dNTPs sind.

### 2.2.2 Polymerase-Kettenreaktion (PCR)

Eine der wichtigsten Methoden in der modernen Molekularbiologie ist die Polymerase-Kettenreaktion (PCR). Daher wollen wir diese Methode ausführlich besprechen. Die PCR wurde Mitte der 1980er-Jahre von dem Ameri-



• **Abb. 2.14** DNA-Polymerasen aus *E. coli*: (A) Das Holoenzym DNA-Polymerase I aus *E. coli* besitzt drei unterschiedliche Aktivitäten. Die Polymerase-Aktivität synthetisiert matrizenabhängig DNA vom 5'- zum 3'-Ende. Dementsprechend wird die Matrize vom 3'-Ende zum 5'-Ende hin abgelesen (oben links). Die 5' → 3'-Exonukleaseaktivität hydrolysiert DNA vom 5'-Ende in Richtung des 3'-Endes, Nucleotid für Nucleotid (oben Mitte). Demgegenüber greift die 3' → 5'-Exonukleaseaktivität die DNA an den 3'-Enden an und hydrolysiert die entsprechenden DNA-Stränge in Richtung des 5'-Endes (oben rechts). Die 5'-Enden selbst bleiben intakt. (B) Durch partielle Proteolyse wird eine Enzymvariante generiert, die als das Klenow-Fragment der DNA-Polymerase I bezeichnet wird. Diesem Fragment fehlt, ähnlich wie der DNA-Polymerase aus dem Bakteriophagen T4, die 5' → 3'-Exonukleaseaktivität des Holoenzym. Hingegen besitzen die Klenow-Polymerase und die T4-DNA-Polymerase sowohl die 5' → 3'-Polymeraseaktivität als auch die 3' → 5'-Exonukleaseaktivität (unten links und rechts).

kaner Kary Mullis (1944 – 2019) erfunden, der dafür 1993 zusammen mit Michael Smith mit dem Nobelpreis für Chemie geehrt wurde. Die Abkürzung PCR leitet sich von der englischen Bezeichnung *polymerase chain reaction* ab. Das wirklich Bemerkenswerte an der PCR besteht darin, dass Mullis ein Prinzip aufzeigte, mit dem DNA in vitro nicht einfach nur synthetisiert, sondern zugleich auch vervielfältigt (amplifiziert) werden kann.

Wie wir bereits gehört haben, waren DNA-Polymerasen lange bekannt und die enzymatische In-vitro-Synthese von DNA gehörte damals bereits zum Standard-Repertoire der Molekularbiologen. Das Neue war die exponentielle Amplifikation der DNA auf der Basis eines verblüffend einfachen Prinzips. Ausgehend von ganz wenigen Einzelmolekülen ließ sich DNA nun in einen Größenordnungsbereich vervielfältigen, der ausreichte, um die DNA nachzuweisen, zu charakterisieren oder für DNA-Rekombinationen einzusetzen.

Die PCR ist heute aus den gentechnischen Forschungslaboratorien nicht mehr wegzudenken – sie hat die molekularbiologische Forschung revolutioniert. Sie wird nicht nur eingesetzt, um DNA-Fragmente für DNA-Rekombinationen zu gewinnen, sondern ist dar-

über hinaus im Lauf der Zeit eine der wichtigsten molekularbiologischen Nachweismethoden geworden, die in praktisch jedem Zweig der Biowissenschaften eingesetzt werden kann. Unter anderem hat diese Technik auch die humangenetische Diagnostik und die forensische Analytik bereichert.

Die Basis der DNA-Amplifikation durch die PCR ist die enzymatische DNA-Synthese. Es werden daher für den Reaktionsansatz alle Komponenten gebraucht, die bei jeder DNA-Synthese nötig sind:

- Wie bei jeder durch DNA-Polymerasen katalysierten Reaktion ist auch bei der PCR eine Matrize nötig, also ein Abschnitt einer DNA, der kopiert werden soll. Diese als Matrize (engl. *template*) bezeichnete DNA muss in diesem Fall nicht einzelsträngig vorliegen, da während der PCR zyklische Temperaturerhöhungen eine Denaturierung der Matrizen-DNA in Einzelstränge bewirken.
- Bekanntermaßen braucht eine DNA-Polymerase einen Startpunkt für die DNA-Synthese. Nachdem die in die Reaktion eingesetzte doppelsträngige DNA denaturiert wurde, muss also ein Primer an die ein-

zelsträngige Matrize hybridisiert werden, um einen kurzen doppelsträngigen Bereich als Startpunkt für die DNA-Polymerisation zu bilden.

- Die vier Desoxynukleosidtriphosphate dATP, dCTP, dGTP und dTTP müssen als Substrate für die DNA-Polymerase vorhanden sein. Im molekularbiologischen Sprachgebrauch wird eine Mischung der vier Desoxynukleosidtriphosphaten als dNTPs („N“ für eine beliebige Base) bezeichnet.
- Geeignete Puffer- und Salzbedingungen ermöglichen die Arbeit des Enzyms. So dienen  $Mg^{2+}$ -Ionen als Cofaktoren für die DNA-Polymerase, deren eigentliche Nukleotid-Substrate die  $Mg^{2+}$ /dNTP-Komplexe sind.
- Eine geeignete DNA-abhängige DNA-Polymerase katalysiert dann die DNA-Synthese.

Die Genialität der von Mullis gemachten Erfindung bestand darin, dass bei der In-vitro-DNA-Synthese gleichzeitig zwei Primer eingesetzt werden und die DNA-Polymerisation mehrmals hintereinander durchgeführt wird. Eigentlich griff Mullis eine Idee des indisch-amerikanischen Biochemikers und Nobelpreisträgers Har Gobind Khorana (1922 – 2011) auf, der das Prinzip einer wiederholten DNA-Polymerisation bereits 1971 publiziert hatte. Allerdings variierte Mullis diese Idee dadurch, dass er zwei Startpunkte für die DNA-Synthese definierte. Diese beiden Startpunkte liegen in einem gewissen Abstand voneinander entfernt und begrenzen dadurch den DNA-Bereich, der amplifiziert werden soll. Außerdem werden die Primer so gewählt, dass die beiden von ihnen ausgehenden DNA-Synthesereaktionen aufeinander zulaufen.

### Prinzip der PCR

Eine schematische Übersicht über das Prinzip der PCR ist in [Abb. 2.15](#) gezeigt. Wir können nur DNA-Bereiche zwischen bekannten Nukleotidsequenzen amplifizieren. Für diese bekannten Nukleotidsequenzen werden Primer synthetisiert, die sich optimal an die bekannten Sequenzen anlagern können. Die Sequenz des amplifizierten DNA-Bereichs zwischen den beiden Primern muss hingegen nicht bekannt sein.

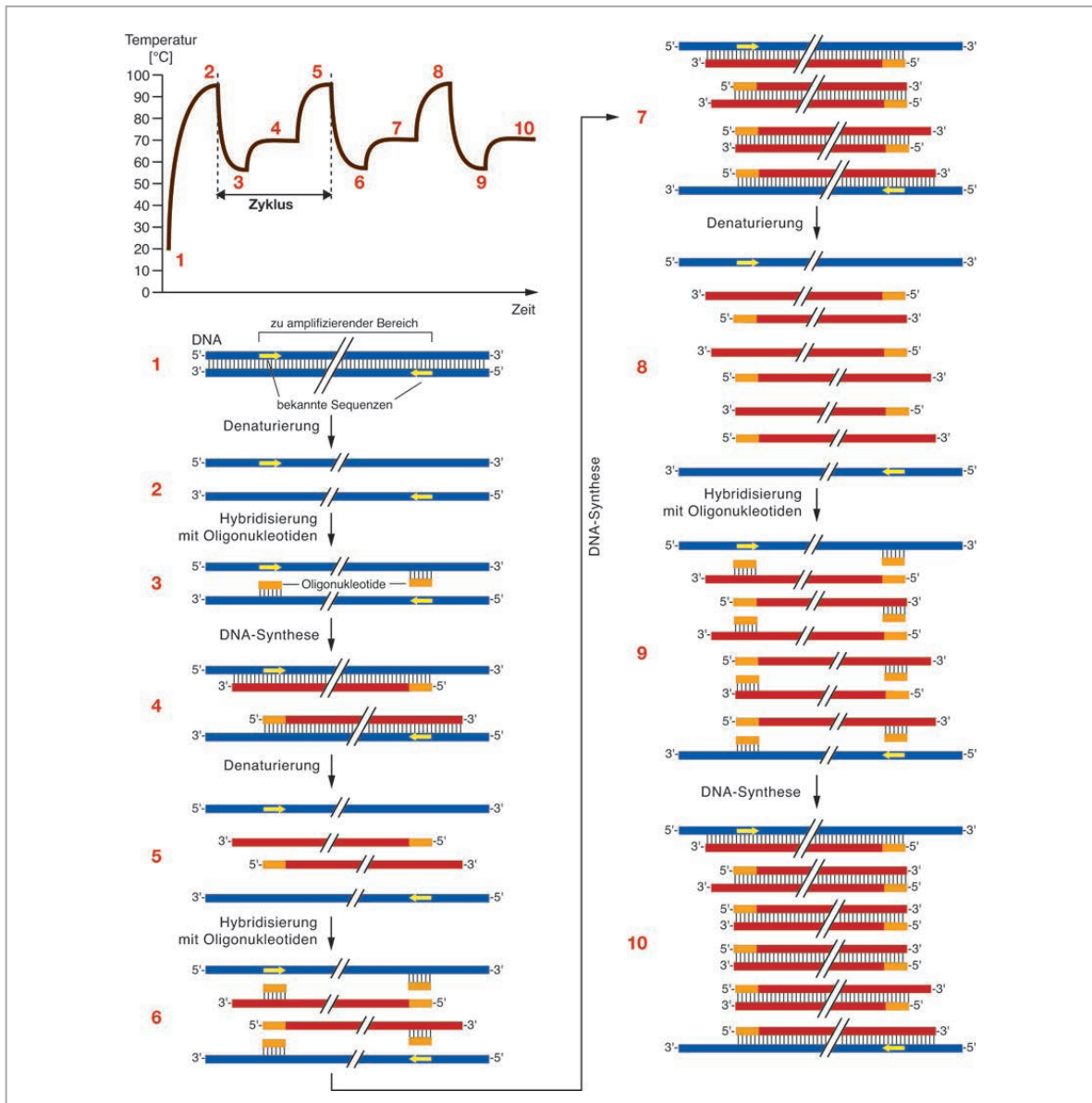
Wir benötigen immer zwei PCR-Primer. Der eine Primer bindet an den „oberen“ Strang der Ziel-DNA, wohingegen der andere Primer ininigem Abstand vom ersten Primer an den „unteren“ Strang bindet. Dadurch haben wir zwei Startpunkte für DNA-Polymerase-Reaktionen definiert, wobei die Richtungen dieser beiden Polymerisationsreaktionen aufeinander zu und schließlich überlappend verlaufen.

Die Primer und die zu amplifizierende DNA mischen wir in einer Pufferlösung zusammen mit den vier Desoxynukleosidtriphosphaten,  $Mg^{2+}$  und einer thermostabilen DNA-Polymerase. Solche thermostabilen DNA-Polymerasen hat man ursprünglich aus verschiedenen

thermophilen Bakterien isoliert, beispielsweise die Taq-Polymerase aus dem thermophilen Bakterium *Thermus aquaticus*. Heute werden diese Polymerasen gentechnisch erzeugt und sind kommerziell erhältlich. Sobald dem Reaktionsansatz alle erforderlichen Komponenten zugesetzt wurden, stellt man das Gefäß in einen sogenannten Thermocycler. Mithilfe eines solchen Gerätes wird ein komplexes Inkubationsprogramm kontrolliert und wiederholt durchlaufen. Typisch für dieses Inkubationsprogramm sind periodisch wechselnde Inkubationstemperaturen ([Abb. 2.15](#)):

- Zunächst wird die Temperatur für ca. 5 Minuten auf  $95\text{ °C}$  eingestellt. Unter den vorgegebenen Pufferbedingungen werden während dieser Zeit alle doppelsträngigen DNA-Bereiche zu Einzelsträngen denaturiert.
- Danach wird die Temperatur auf einen Bereich von  $50 - 60\text{ °C}$  gesenkt, wodurch die Primer mit den komplementären Bereichen auf der Ziel-DNA doppelsträngige Hybride ausbilden können. Die beiden Primer liegen im großen Überschuss zur DNA-Matrize vor. Deshalb bilden sich bevorzugt Doppelstrangbereiche zwischen den Primern und dem jeweiligen DNA-Einzelstrang der Ziel-DNA. Die Renaturierung der beiden denaturierten Einzelstränge der Ziel-DNA ist hingegen unter diesen Bedingungen nicht favorisiert. Die Hybridisierung der Primer (*annealing*) ist in der Regel bereits nach 15 – 30 Sekunden abgeschlossen.
- Im nächsten Schritt wird die Inkubationstemperatur auf ca.  $72\text{ °C}$  angehoben. Bei dieser Temperatur arbeiten die thermostabilen DNA-Polymerasen optimal. Je nach Länge des zu amplifizierenden DNA-Bereichs wird die Temperatur zwischen 30 Sekunden und mehreren Minuten bei  $72\text{ °C}$  gehalten.
- Nun wird durch Anheben der Temperatur für 15 – 30 Sekunden auf  $> 90\text{ °C}$  wieder die gesamte im Reaktionsansatz vorhandene doppelsträngige DNA denaturiert.
- Der Wechsel zwischen den drei verschiedenen Temperaturen ( $50 - 60\text{ °C} \rightarrow 72\text{ °C} \rightarrow 95\text{ °C}$  bzw.  $95\text{ °C} \rightarrow 50 - 60\text{ °C} \rightarrow 72\text{ °C}$ ) wird als **Zyklus** bezeichnet. Je nach gewünschtem Grad der Amplifikation der Ausgangs-DNA werden in der Praxis meist zwischen 25 und 35 solcher Zyklen nacheinander durchlaufen.

Die Produkte des ersten Amplifikationszyklus sind nur an einem Ende durch den eingesetzten Primer klar begrenzt. Sie besitzen jedoch keine genau definierte Länge; sie werden deshalb als „lange Produkte“ bezeichnet ([Abb. 2.15](#)). Ab der zweiten Amplifikationsrunde fungieren neben der Originalmatrize auch die „langen Produkte“ als Matrizen. Deshalb werden zusätzlich zu zwei weiteren „langen Produkten“ auch DNA-Fragmente definierter Länge gebildet. Diese neu entstehenden



• **Abb. 2.15** Prinzip der Polymerase-Kettenreaktion (PCR): In der ersten Abbildung ist der Temperaturverlauf der PCR dargestellt. Wir bezeichnen die Menge an Matrizen-Molekülen mit  $N_0$  und die Zahl der Zyklen mit  $z$ . (1) Die dunklen Balken zeigen einen Ausschnitt aus einer DNA, die als Matrize dient. Der Bereich, der amplifiziert werden soll, ist mit gelben Pfeilen markiert. Dieser Bereich wird von zwei Primern (orangefarbene Balken) eingerahmt, die jeweils entweder am oberen oder am unteren Strang der Matrizen-DNA binden. (2) Denaturieren der DNA durch Erhitzen auf  $> 95^\circ\text{C}$ . (3) Abkühlen auf  $50 - 60^\circ\text{C}$ , um das Anlagern der Primer (*annealing*) zu ermöglichen. (4) Einstellen der Temperatur auf  $72^\circ\text{C}$ , um die DNA-Synthese durch die Taq-Polymerase zu starten. Einen Umlauf von Denaturieren/Annealing/DNA-Synthese ( $95^\circ\text{C} \rightarrow 60^\circ\text{C} \rightarrow 72^\circ\text{C}$ ) bezeichnet man als PCR-Zyklus. Beachte, dass die DNA-Synthese weiterläuft als bis zur gegenüberliegenden Primer-Bindestelle. Solche „langen Produkte“ vermehren sich linear ( $N_0 \times z$ ). (5) Erhitzen auf  $95^\circ\text{C}$ . (6) Erneutes Annealing der Primer. (7) Erneute DNA-Synthese. Beachte, dass die ersten DNA-Moleküle entstehen, die genau so lang sind wie das gewünschte Endprodukt. Diese Moleküle amplifizieren sich jetzt exponentiell ( $N_0 \times 2^z$ ). (8) Denaturieren. (9) Annealing. (10) DNA-Synthese. Das Endprodukt ist genau so lang wie der Abstand zwischen den PCR-Primern (inklusive der Primer selbst).

den Transkripte sind an der einen Seite durch den eingesetzten Primer und an der anderen Seite durch das Ende der Matrize begrenzt. Somit liegen am Ende der PCR-basierten DNA-Amplifikation drei verschiedene Molekulararten vor (• Abb. 2.16):

- die Stränge der Originalmatrize,
- einseitig durch jeweils einen Primer begrenzte „lange Produkte“ und
- beidseitig durch die Primer begrenzte Transkripte, die eigentlichen PCR-Amplimere.

Alle diese Moleküle dienen auch in den jeweils folgenden Amplifikationszyklen als Matrizen für die DNA-Polymerisation.

### Effizienz und Limitation der PCR

Nachdem zuvor das Prinzip der PCR diskutiert wurde, ist natürlich die Frage interessant, wie stark eine DNA mit dieser Methode amplifiziert werden kann. Ausgehend von  $N_0$  Molekülen DNA, die für eine PCR-Amplifikation eingesetzt wurden, sieht die theoretische Bilanz wie folgt aus: Die Zahl der Ausgangsmatrizen ist konstant geblieben, während die Zahl der „langen Produkte“ (Abb. 2.16) auf  $N_0 \times z$  gestiegen ist ( $z = \text{Anzahl der Zyklen}$ ). Die Zahl der PCR-Amplimere definierter Länge nach  $z$  Zyklen ( $N_z$ ) errechnet sich als  $N_z = N_0 \times 2^z$ . Ein Beispiel soll den Amplifikationsgrad verdeutlichen: Nehmen wir an, dass wir 1000 Moleküle einer DNA als Ausgangsmaterial (Matrize,  $N_0$ ) haben. Nehmen wir weiterhin an, dass wir ein Stück aus dieser DNA mit 30 PCR-Zyklen amplifizieren. Wir bekommen nach der oben genannten Formel  $1000 \times 2^{30}$ , also ca.  $10^{12}$  PCR-Amplimere. Das entspricht einem Amplifikationsgrad von  $10^9$ . Interessant für die Praxis ist, dass wir 1000 Moleküle auch mit modernsten Methoden der DNA-Analytik kaum nachweisen können, während nach der Amplifikation durch die PCR aus diesen 1000 Molekülen ca.  $2 \times 10^{-12}$  mol oder  $1,3 \times 10^{-6}$  g (1,3 µg) DNA geworden sind (bei einer angenommenen Amplimerlänge von 1000 bp). Diese DNA-Menge kann über konventionelle Agarosegel-Elektrophorese leicht nachgewiesen oder in ihrer DNA-Sequenz bestimmt werden. Ebenso bekommen wir mit dieser Methode genug DNA „in die Hand“, um sie mit gentechnologischen Methoden bearbeiten zu können.

Verschiedenste Parameter können die Effizienz der PCR beeinflussen. Einer der limitierenden Faktoren ist die Länge der DNA, die durch die Primer eingeschlossen wird. Generell nimmt die Effizienz der PCR mit steigender Länge des zu amplifizierenden DNA-Bereichs kontinuierlich ab. Dies kann teilweise durch die Inkubationsdauer während des Polymerisationsschrittes ausgeglichen werden. Optimal lassen sich Bereiche in der PCR vervielfältigen, die kleiner als 2000 bp sind. Allerdings können bei entsprechender Auswahl der DNA-Polymerasen und der Inkubationsbedingungen auch bis zu 40000 bp amplifiziert werden.

In der Regel werden die theoretischen Ausbeuten an PCR-Produkt nicht erreicht, weil nach ca. 20 Zyklen die Effizienz der Reaktion kontinuierlich abnimmt. Das kann folgende Gründe haben:

- Substrate und Primer werden langsam, aber stetig verbraucht.
- Das Enzym lässt durch die starken Temperaturschwankungen in seiner Aktivität langsam nach.

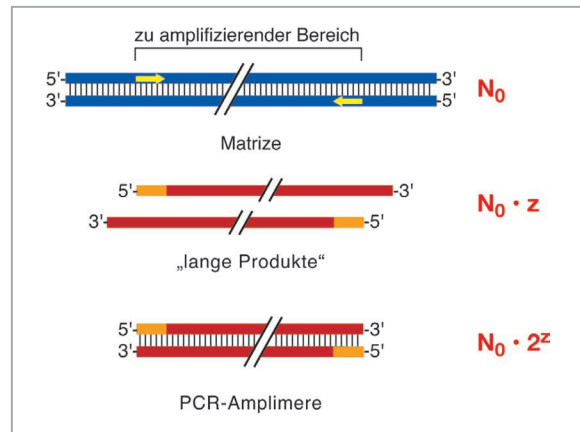


Abb. 2.16 Gegenüberstellung der in einer PCR gebildeten Produkte. Die blauen Balken zeigen einen Ausschnitt aus einer DNA, die als Matrize dient. Der zu amplifizierende Bereich ist mit gelben Pfeilen begrenzt. Die Primer sind orange dargestellt. Die Zahl der eingesetzten Matrizen-Moleküle ist ( $N_0$ ) und bleibt während der PCR konstant. Beachte, dass die eingesetzten Primer Teil des PCR-Amplimers sind. Die sogenannten „langen Produkte“ bilden sich durch Abschreiben der Matrizen-Moleküle in jedem Zyklus ( $z$ ). Da sich die Matrizen-Menge nicht ändert, vermehren sich die langen Produkte linear ( $N_0 \times z$ ). Die langen Produkte wie auch die eigentlichen PCR-Amplimere vermehren sich in jedem Zyklus und dienen wieder selbst als Matrize. Daher vermehren sich die PCR-Amplimere exponentiell ( $N_0 \times 2^z$ ).

- Es lagern sich vermehrt komplementäre Stränge zu Doppelsträngen aneinander und verhindern damit, dass sich neue Initiationsstellen für die DNA-Polymerisation bilden.

Eine der großen Stärken der Polymerase-Kettenreaktion – nämlich ihre enorme Sensitivität – kann auch zu einem großen Problem werden. Das muss besonders dann bedacht werden, wenn die PCR im Rahmen diagnostischer Verfahren eingesetzt wird. Unter Sensitivität versteht man die Möglichkeit der PCR, aus sehr geringen Matrizen-Mengen analysierbare Mengen an Amplimern zu generieren. Man kann sich leicht vorstellen, dass hierbei Kontaminationen der Analysenprobe mit fremdem biologischem Material unter allen Umständen vermieden werden müssen. Ansonsten besteht die große Gefahr der Isolierung einer falschen Amplimer-DNA und daraus folgend eines falsch positiven Ergebnisses in der genetischen Diagnostik. Dies ist gerade für solche Krankheiten eine inakzeptable Komplikation, bei denen heute bevorzugt die PCR-Analytik als Diagnostikmethode eingesetzt wird, wie beispielsweise bei Krebserkrankungen oder viralen Infektionen. Ähnliches gilt natürlich auch für die Erstellung sogenannter genetischer Fingerabdrücke, die im Rahmen forensischer Untersuchungen zur Identifizierung bestimmter Personen eingesetzt werden. Die hohe Sensitivität der

Methode bedingt, dass eine Kontamination mit DNA-Spuren, die zum Beispiel vom forensischen Personal eingetragen werden, eine Analyse sofort unbrauchbar machen würde.

### 2.2.3 Variationen der PCR

Mittlerweile gibt es so viele Anwendungsvarianten der PCR-Technik, dass eine vollständige Aufzählung und Erklärung dieser Varianten den Rahmen dieses Buches sprengen würde. Deshalb sollen nur einige Beispiele erwähnt werden, die bei der Isolierung von Genen zur Gewinnung rekombinanter Proteinwirkstoffe oder in der molekularen Diagnostik besondere Bedeutung haben.

#### Reverse-Transkription-PCR

Die Reverse Transkriptase der Retroviren (meist als RT abgekürzt) ist insofern eine ungewöhnliche DNA-Polymerase, als sie vor allem RNA als Matrize erkennt und diese in DNA kopiert. Die RT ist also eine RNA-abhängige DNA-Polymerase. Im Labor ist die RT das maßgebliche Enzym zur Herstellung von mRNA-Kopien, den cDNAs. Der Begriff cDNA steht für *complementary DNA* und deutet an, dass das erste Produkt der reversen Transkription ein DNA-Einzelstrang ist, der zur kopierten mRNA komplementär ist. Im Sprachgebrauch der Molekularbiologen wird der Begriff cDNA oft auch als *copy DNA* übersetzt, was als Synonym für „kopierte mRNA“ gesehen werden kann.

Obwohl das Prinzip der PCR darauf beruht, DNA *in vitro* zu amplifizieren, schließt dies nicht aus, dass wir auch RNA aus Ausgangssubstanz für die Amplifikation nutzen können. Wir können RNA allerdings nicht direkt als Matrize in der PCR einsetzen, da sie aufgrund ihrer chemischen Instabilität beim Erhitzen auf 95 °C in den ersten PCR-Zyklen zerstört werden würde. Wir müssen also vor der PCR die RNA durch reverse Transkription in DNA umschreiben, und diese DNA-Kopie ist dann das Ausgangsmaterial für die Amplifikation per PCR. Dieses Verfahren einer gekoppelten reversen Transkription mit der PCR nennt man Reverse-Transkription-PCR, kurz RT-PCR. In der molekularen biologischen Forschung hat diese Methode große Bedeutung:

- Mit relativ geringem Aufwand lässt sich für ein beliebiges Gen die Expressionsrate, also die Gleichgewichtsmenge von mRNA in einer Zelle, messen.
- Nur die mRNA enthält die genetische Information für ein Protein ohne Unterbrechung des Triplet-Leserahmens durch Introns. Das ist von enormer praktischer Bedeutung, wenn man therapeutische Proteine aus dem menschlichen Genom beispielsweise in Bakterien exprimieren möchte, da Bakterien keine Introns kennen und eine durch Introns unterbrochene genetische Information nicht in ein Protein umsetzen können.

Konkret gehen wir folgendermaßen vor, um mRNA in cDNA zu kopieren (● Abb. 2.17): Zunächst lagern wir an das 3'-Ende einer isolierten mRNA einen DNA-Primer an. Eine Möglichkeit, die Reverse Transkription zu starten, ist die Nutzung eines spezifischen Primers, der nur zu einer bestimmten Gensequenz passt. Der Vorteil dieser Methode ist, dass aus der Auswahl von bis zu 50 000 verschiedenen mRNAs einer Zelle theoretisch nur eine einzige cDNA-Spezies hergestellt wird, nämlich die des untersuchten Gens. Eine andere, wesentlich häufiger verwendete Methode ist, die gleichzeitige Reverse Transkription **aller** in der RNA-Präparation enthaltenen mRNAs durch die Nutzung unspezifisch bindender Primer zu starten. Es gibt zwei verschiedene Arten solcher unspezifischen Primer:

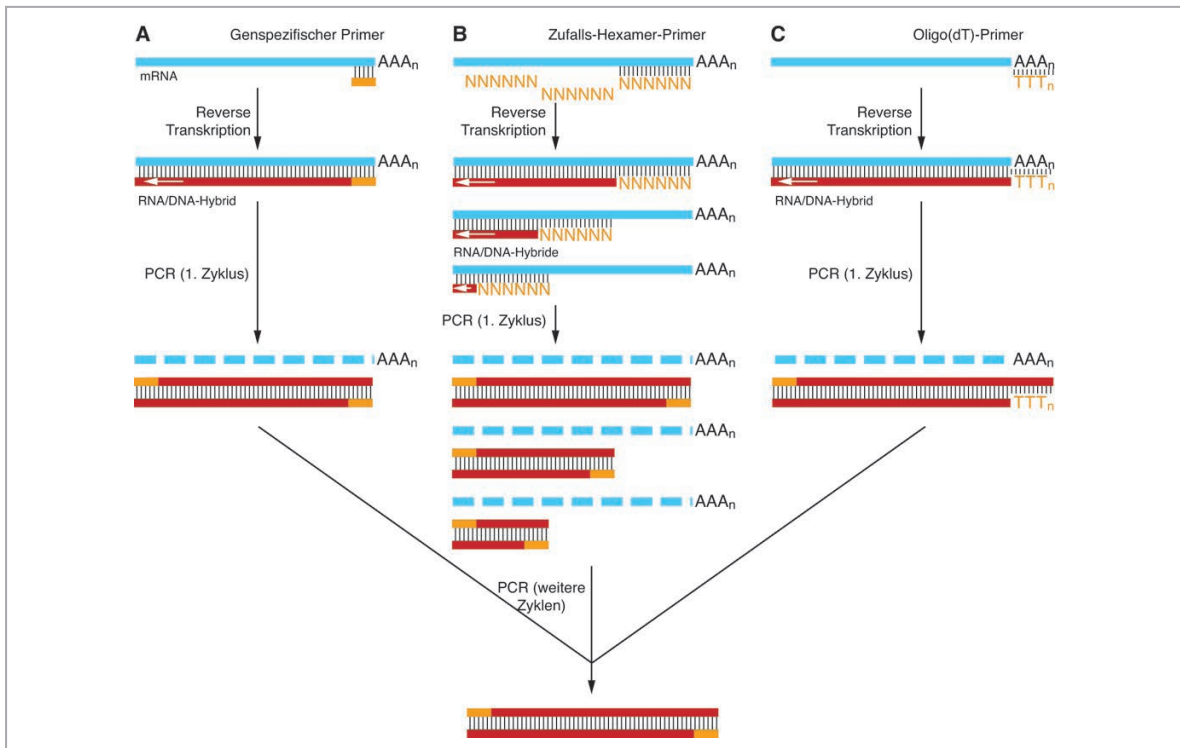
- Zum einen kann man ausnutzen, dass alle eukaryontischen mRNAs an ihren 3'-Enden polyadenyliert werden. Wir können also einen Oligo(dT)-Primer einsetzen, der an die Oligo(A)-Enden aller mRNAs hybridisiert. Damit wird die Reverse Transkription **jeder** zellulären mRNA von deren 3'-Ende aus gestartet und man erhält eine komplexe Mischung von cDNAs.
- Alternativ können sogenannte Zufalls-Hexamere eingesetzt werden. Das sind Oligonukleotide von sechs Nukleotiden Länge, die in jeder Position jedes der vier möglichen DNA-Basen enthalten können. Sie sind also theoretisch zu jeder denkbaren Gensequenz komplementär. Die Reverse Transkription wird dann an irgendeiner Position einer mRNA gestartet und man erhält ebenfalls eine komplexe cDNA-Mischung. Mit dieser Methode können auch RNAs revers transkribiert werden, die keine Oligo(A)-Enden enthalten.

Nach der Hybridisierung von Oligo(dT)- oder Hexamer-Primern an die mRNA inkubieren wir den Ansatz mit einer Reversen Transkriptase. Am Ende dieser Reaktion liegen alle mRNAs einer Zelle als RNA-DNA-Hybride vor. Soll aus dieser Mischung die cDNA eines speziellen Gens isoliert werden, führen wir anschließend eine PCR mit zwei genspezifischen Primern durch. Dabei wird im ersten Zyklus der PCR durch Erhitzen auf 95 °C der RNA-Anteil der vorhandenen RNA-DNA-Hybride zersetzt. Die einzelsträngige cDNA dient nun als Matrize für die PCR.

#### Quantitative PCR

Wir haben bereits in einem der vorausgehenden Abschnitte einige Grenzen der PCR diskutiert. Im Bereich der Anwendung der PCR in der Diagnostik sind folgende Limitationen der konventionellen PCR von praktischer Bedeutung:

- PCRs sind relativ zeitaufwendig, da einerseits die Reaktion selbst einige Zeit beansprucht, andererseits



• **Abb. 2.17** Gekoppelte Reverse-Transkription-PCR (RT-PCR): In einer RT-PCR wird eine RNA (hellblaue Linien) durch das Enzym Reverse Transkriptase (RT) in eine DNA kopiert. Es entsteht eine einzelsträngige cDNA (rote Linien), die zunächst einmal als RNA-DNA-Hybrid vorliegt. In den hier gezeigten Beispielen geht die RT-PCR von polyadenylierter RNA, also mRNA, aus. Primer sind in Orange dargestellt.

(A) Starten der RT-PCR mit einem genspezifischen Primer: In diesem Beispiel wird aus einem Pool isolierter mRNAs entsprechend der Spezifität des Primers nur eine ganz bestimmte RNA amplifiziert. Nach der Erststrang-cDNA-Synthese wird die RNA in dem RNA-DNA-Hybrid durch die RNase-H-Aktivität der RT abgebaut. Mit einem zweiten Primer, der am 3'-Ende der Erststrang-cDNA bindet, wird durch PCR ein DNA-Doppelstrang erzeugt.

(B) Starten der RT-PCR mit Zufalls-Hexameren. Diese Primer bestehen aus einer Mischung von Hexamer-Oligonukleotiden, die in jeder Position jedes der vier möglichen DNA-Basen enthalten können. Zufalls-Hexamere sind demnach theoretisch zu jeder denkbaren Gensequenz komplementär. Allerdings kann dadurch die Reverse Transkription an irgendeiner Position des zu analysierenden Gens gestartet werden, ohne dass man dies beeinflussen kann. Man bekommt also eine Schar von Erststrang-cDNAs, die sich in ihren Längen unterscheiden. Allerdings lässt sich aus dieser Schar die cDNA mit der korrekten Länge amplifizieren, indem die entsprechenden genspezifischen Primer eingesetzt werden.

(C) Starten der RT-PCR mit einem Oligo(dT)-Primer. Hier nutzt man aus, dass alle mRNAs polyadenyliert sind. Der mRNA-Pool wird mit einem Oligo(dT)-Primer hybridisiert und so die Reverse Transkription aller mRNAs gestartet. Danach folgt eine PCR mit genspezifischen Primern.

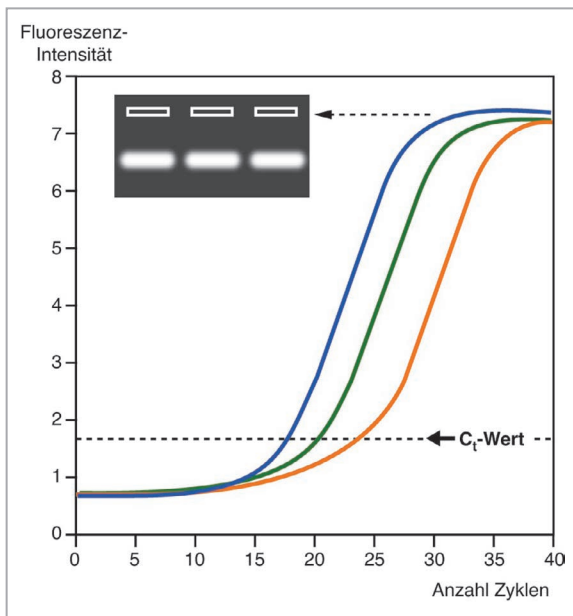
jedoch die Analyse der Amplimere durch eine nachgeschaltete Agarosegel-Elektrophorese erfolgen muss.

- Obwohl die PCR an sich für einen großen Probenumsatz beispielsweise in der klinischen Diagnostik automatisierbar wäre, muss die anschließende Auswertung durch Agarosegel-Elektrophorese händisch durchgeführt werden. Es wäre wünschenswert, eine Auswertungsmethode zu haben, die wie die PCR selbst automatisierbar wäre.
- Die konventionelle PCR ist keine quantitative Methode. Sie ist normalerweise so angelegt, dass die Reaktion möglichst vollständig abläuft. Daraus folgt, dass bei starker Amplifikation einer Matrizen-DNA die Ressourcen in der Reaktion schneller limitierend werden, als wenn nur eine geringe Amplifikation

stattfindet. Mit großer Wahrscheinlichkeit wird die Reaktion am Ende der PCR in der Sättigung gewesen sein. Daraus folgt, dass eine quantitative Abschätzung der zu Beginn der PCR vorliegenden Matrizenmenge durch Vergleich mit der am Ende einer PCR vorliegenden Amplimer-Mengen in verschiedenen Proben kaum möglich ist, denn man wird nicht beobachten können, in welchem Zyklus der PCR die Reaktion in einer Probe bereits in die Sättigung gegangen war, während sie in parallelen Proben noch exponentiell abließ.

Hieraus ergibt sich die Überlegung, ob es nicht möglich wäre, die Akkumulation der Produkte in einer PCR in Echtzeit zu verfolgen, sodass die Menge gebildeter

Amplimere während der exponentiellen Phase der PCR festgestellt werden kann. Diese Abwandlung der konventionellen PCR wird als quantitative PCR oder auch Realtime-PCR bezeichnet (● Abb. 2.18). Die Grundlage für die quantitative PCR (qPCR) ist, dass dem PCR-Ansatz ein für doppelsträngige DNA spezifischer, interkalierender Farbstoff zugesetzt wird, dessen Fluoreszenz nach Bindung an die gebildeten Amplimere zunimmt. Das gemessene Fluoreszenzsignal in jedem Zyklus ist dann proportional zur gebildeten Menge an Amplimer und kann nach jedem PCR-Zyklus abgefragt werden. Der Farbstoff der Wahl ist hier CYBR-Green®.



● **Abb. 2.18** Quantitative PCR: Diese Methode erlaubt es, die Bildung von PCR-Amplimeren in Echtzeit zu verfolgen. Dazu wird dem PCR-Ansatz ein Fluoreszenzfarbstoff zugegeben, der Doppelstrang-DNA spezifisch anfärbt (CYBR-Green®). In der Abbildung ist der theoretische Reaktionsverlauf paralleler PCRs gezeigt, in die in der Reihenfolge braune → grüne → blaue Kurve steigende Menge Matrizen-DNA eingesetzt wurden. Am Ende der Reaktionen wird in allen Reaktionen eine ähnliche Fluoreszenz erreicht, d. h. die Reaktionen waren in der Sättigung. Würde man die gebildete DNA-Menge am Ende der Reaktionen auf einem Agarosegel analysieren, würde man also etwa gleich starke Banden (Insert) erhalten und fälschlicherweise schlussfolgern, dass alle PCRs gleiche Matrizen-Mengen enthalten haben und daher die Reaktionen vergleichbar abgelaufen sind. Tatsächlich sind die Reaktionen aber unterschiedlich schnell verlaufen, was auf unterschiedliche Mengen an vorhandenem Ausgangsmaterial (Matrize) zurückzuführen ist. Den Unterschied an eingesetzter Matrize kann man bestimmen, indem man die Anzahl von PCR-Zyklen vergleicht, die bis zum Erreichen eines zuvor festgelegten Schwellenwerts ( $C_t$ ) benötigt werden, wobei der Unterschied in Zyklen ( $z$ ) im Wesentlichen einer Differenz der Matrizen-Menge von  $2^z$  entspricht.

Wie Ethidiumbromid interkaliert CYBR-Green® in doppelsträngige DNA und bildet einen sehr stabilen Komplex, wobei nach Anregung mit Licht einer Wellenlänge von 497 nm ein Fluoreszenzlicht von 521 nm emittiert wird. Die Bestimmung von DNA-Mengen mit CYBR-Green® ist wesentlich empfindlicher als die DNA-Färbung mit Ethidiumbromid.

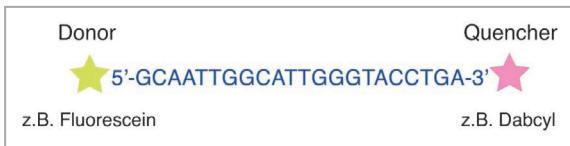
Führen wir eine PCR in Gegenwart von CYBR-Green® durch, erhalten wir einen in ● Abb. 2.18 dargestellten Verlauf der Reaktion in Abhängigkeit von der eingesetzten Matrizen-Menge. Je mehr Matrize in der PCR eingesetzt wird, desto früher erreicht die Fluoreszenz-Intensität einen Schwellenwert, den wir früh in der exponentiellen Phase der Reaktion festlegen (● Abb. 2.18). Wir können nun die Matrizen-Menge in unbekanntenen Proben abschätzen, indem wir bestimmen, bei welchem Zyklus der Schwellenwert erreicht wird. Der Unterschied an Zyklen bis zum Erreichen des Schwellenwerts steht über die bereits diskutierte Formel

$$N_z = N_0 \times 2^z$$

mit  $N_z$ : Zahl der PCR-Amplimere nach  $z$  Zyklen,  $N_0$ : Ausgangsmenge Matrize und  $z$ : Zahl der PCR-Zyklen in Beziehung zu der in der Probe befindlichen Matrizen-Menge, die sich als  $N_0 = N_z$  geteilt durch  $2^z$  bzw.  $N_0 = 2^{-z} \times N_z$  berechnet. In der Praxis können wir  $N_0$  am besten dadurch bestimmen, dass wir für eine Vergleichs-DNA eine Standardkurve wie in ● Abb. 2.18 herstellen und die Zykluszahl am Schwellenwert in den unbekanntenen Proben mit dieser Standardkurve vergleichen.

Die Verwendung von CYBR-Green® hat jedoch auch einen Nachteil: Wir detektieren **jede** in der PCR synthetisierte doppelsträngige DNA, also auch nicht gewünschte Amplimere, die durch unspezifisches Hybridisieren der Primer an falsche Matrizen entstehen können. Um dies zu vermeiden, wurde eine Modifikation der quantitativen PCR eingeführt, die als TaqMan®-PCR oder 5'-Nuklease-test bekannt geworden ist. TaqMan®-Primer tragen fluoreszierende Farbstoffe, sodass die Messung des Reaktionsverlaufs nicht auf dem Zusatz von CYBR-Green® basiert. Ein TaqMan®-Primer wird zusätzlich zu zwei konventionellen Primern eingesetzt, die für die Bildung des Amplimers sorgen. Der TaqMan®-Primer bindet spezifisch an eine DNA-Sequenz zwischen diesen beiden konventionellen Primern (● Abb. 2.19).

Das Besondere an TaqMan®-Primern ist die Nutzung des Prinzips des Förster-Resonanzenergietransfers (FRET). Die Fluoreszenz eines Chromophors am 5'-Ende des TaqMan®-Primers (Donor) wird von einem am 3'-Ende des Primers verankerten Quencher unterdrückt. Die Energie-Übertragung vom Donor zum Quencher erfolgt durch FRET. Durch Lichteinstrahlung in einen Chromophor (hier der Donor) wird ein Elektron in einen elektronisch angeregten Zustand versetzt. Die Energie des angeregten Elektrons kann auf einen



• **Abb. 2.19** Aufbau eines TaqMan®-Primers: Ein TaqMan®-Primer trägt an beiden Enden chromophore Gruppen. An das 5'-Ende des Primers ist ein Fluoreszenz-Donor gekoppelt, der bei entsprechender Anregung die eingestrahlte Energie durch Förster-Resonanzenergietransfer (FRET) an einen an das 3'-Ende des Primers gekoppelten Chromophor weitergibt. Wenn letzterer nicht selbst fluoresziert, sondern die Fluoreszenz des Donors praktisch auslöscht, spricht man von einem Quencher. Eine Fluoreszenz beobachten wir immer dann, wenn der Donor vom Quencher getrennt wird (siehe 5'-Nukleasetest).

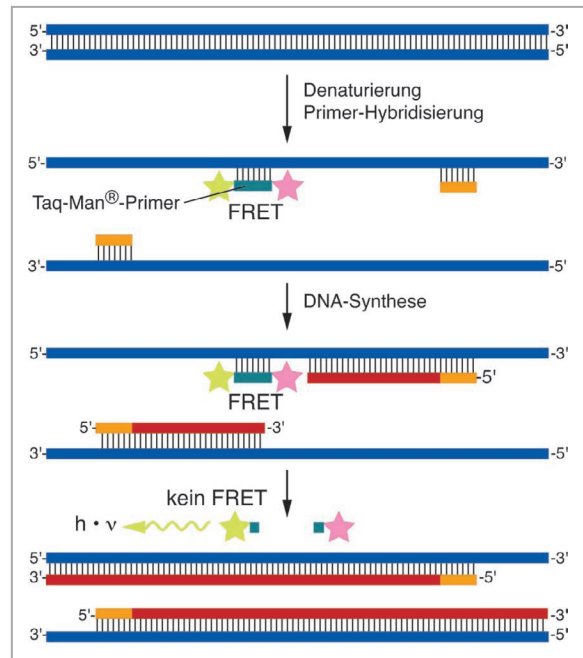
Akzeptor-Chromophor übertragen werden, der weniger als 10 nm entfernt ist. Der Energietransfer ist strahlungslos, die Geschwindigkeit und Intensität von FRET hängen unter anderem von der Größe der Überlappung des Emissionsspektrums des Donors und des Absorptionsspektrums des Akzeptors ab. In unserem Fall nutzen wir FRET, um die Fluoreszenz des Donors durch Absorption im Akzeptor, der selber nicht fluoresziert, zu unterdrücken. Wenn wir also während der PCR die Zunahme eines Fluoreszenzsignals als Maß für die Menge entstehender Amplimere messen wollen, muss der FRET innerhalb des TaqMan®-Primers irgendwie aufgehoben werden.

Was geschieht nun während der TaqMan®-PCR? Nach der Denaturierung der Matrizen-DNA binden die zwei konventionellen PCR-Primer und der TaqMan®-Primer an die Matrizen-DNA (• Abb. 2.20). Zu diesem Zeitpunkt ist FRET zwischen Donor- und Quencher-Chromophor möglich, wir bekommen kein Fluoreszenz-Signal. Trifft nun die DNA-Polymerase während der folgenden DNA-Synthese auf den TaqMan®-Primer, wird dieser durch die 5' → 3'-Exonuklease-Aktivität der DNA-Polymerase abgebaut. Die dadurch bedingte Trennung der Chromophore hebt FRET auf und der Donor-Chromophor fluoresziert messbar (daher der Name 5'-Nukleasetest). In der TaqMan®-PCR ist also das Fluoreszenz-Signal proportional zur aktuell gebildeten Menge eines Amplimers. Wir messen aufgrund der Spezifität des TaqMan®-Primers nur die Zunahme des gewünschten Amplimers und nicht die Menge unspezifischer Nebenprodukte.

#### 2.2.4 DNA-Sequenzierung

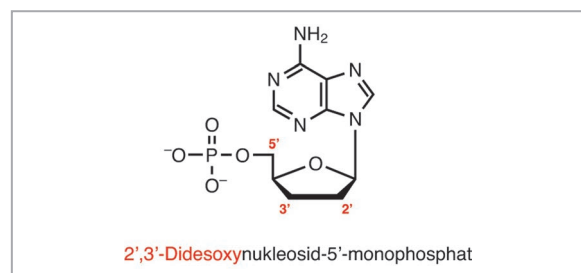
##### DNA-Sequenzierung nach Sanger

Die vom britischen Biochemiker Frederick Sanger (1918 – 2013) Mitte der 1970er-Jahre eingeführte Sequenzierung von DNA erfolgt mit einer Modifikation der normalen enzymatischen DNA-Synthese. Sangers Konzept besteht darin, DNA in Gegenwart geringer Konzentrationen



• **Abb. 2.20** Prinzip der TaqMan®-PCR (5'-Nukleasetest): Der 5'-Nukleasetest ist eine PCR mit drei Primern. Zwei Primer (orange) sind konventionelle PCR-Primer, der dritte ist ein TaqMan®-Primer. Der TaqMan®-Primer bindet zwischen den beiden normalen PCR-Primern. Die Matrizen-DNA (blaue Linien) wird denaturiert und alle drei Primer hybridisieren an die DNA. Erreicht die DNA-Polymerase den TaqMan®-Primer, baut sie ihn über ihre 5' → 3'-Exonuklease-Aktivität ab. Dadurch werden Fluoreszenz-Donor und -Quencher voneinander getrennt und es entsteht ein Fluoreszenz-Signal.

sogenannter Ketten-Terminatoren zu synthetisieren. Ketten-Terminatoren sind 2',3'-Dideoxynukleosidtriphosphate (ddNTPs), die sich von den normalen 2'-Desoxynukleosidtriphosphaten (dNTPs) darin unterscheiden, dass ihnen neben der 2'-OH-Gruppe auch die 3'-OH-Gruppe der Ribose fehlt (• Abb. 2.21). Wird ein ddNTP von der DNA-Polymerase als Substrat verwendet, kann die DNA-Kette nach dieser Position wegen des Fehlens einer 3'-OH-Gruppe nicht weiter verlängert werden.



• **Abb. 2.21** Dideoxynukleotide für die DNA-Sequenzierung: Dideoxynukleotide unterscheiden sich von gewöhnlichen Desoxynukleotiden durch das Fehlen der 3'-OH-Gruppe.

Für die Sanger-Sequenzierung vorbereitete DNAs liegen normalerweise in Plasmiden kloniert vor. Ein Reaktionsansatz zur DNA-Sequenzierung enthält alle Komponenten, die auch für die normale DNA-Synthese benötigt werden. Der zu sequenzierende Bereich einer DNA kann als doppelsträngige Plasmid-DNA eingesetzt werden, da diese DNA im Sinne eines *cycle sequencing* in Einzelstränge überführt wird. Cycle-Sequencing bedeutet, dass DNA-Sequenzierungen ähnlich wie eine PCR durchgeführt werden, jedoch enthält die Sequenzierreaktion nur einen Primer. Im ersten Schritt des jeweiligen Cycle-Sequencing-Zyklus wird also die doppelsträngige DNA denaturiert, im zweiten Schritt der Primer an die DNA hybridisiert und im dritten Schritt wird die DNA-Synthese gestartet. Wie in einer normalen PCR wird dieser Vorgang 25- bis 35-mal wiederholt.

Die Sequenzierreaktion enthält dNTPs und ddNTPs in einem definierten Verhältnis. Als Folge davon kann die DNA-Polymerase statistisch gesehen an jeder Position der Matrizen-DNA mit einer gewissen Wahrscheinlichkeit ein ddNTP statt eines dNTPs einbauen, sodass eine weitere Verlängerung der DNA nicht mehr möglich ist. Als Resultat der Reaktion entsteht eine Schar von DNA-Fragmenten, die alle denselben Anfang haben (den Primer) und mit einem ddNTP enden. Da der Kettenabbruch an jeder Position der synthetisierten DNA erfolgen kann, unterscheiden sich die erhaltenen DNA-Fragmente in der Länge um jeweils eine Base (Abb. 2.22).

Wie lässt sich jetzt herausfinden, mit welchem Didesoxynukleotid der Kettenabbruch erfolgt ist? Der Trick besteht darin, dass jedes der vier in der Reaktion eingesetzten ddNTPs mit einem anderen Fluoreszenzfarbstoff markiert wird. Die Trennung einer Schar von fluoreszenzmarkierten DNA-Molekülen erfolgt durch Kapillarelektrophorese. Die Trennschärfe ist dabei so groß, dass einzelne DNA-Fragmente getrennt werden können, die sich in der Länge um eine einzige Base unterscheiden. Am Ende der Trennstrecke regt ein Laser die vier Fluoreszenzfarbstoffe an (Abb. 2.23). Die jeweilige Emissionswellenlänge korrespondiert mit der Didesoxybase, die das entsprechende DNA-Fragment am Ende durch Kettenabbruch markiert hat. Ein Computer erstellt aus den Daten ein Elektropherogramm (Abb. 2.23) sowie ein Textdokument über die Reihenfolge der Basen, also der DNA-Sequenz.

Die DNA-Sequenzierung nach Sanger ist die Standard-Methode, wenn es beispielsweise darum geht, mit gentechnologischen Methoden bearbeitete DNA-Fragmente als eine Art „Identitätsprüfung“ zu charakterisieren. Mit einer einzigen Sequenzierreaktion erreicht man mit der Sanger-Methode in der Praxis „Leseweiten“ von bis zu 1000 Nukleotiden.

Die Sanger-Sequenzierung spielt aber auch eine große Rolle für die Identifizierung neuer DNA-Sequen-

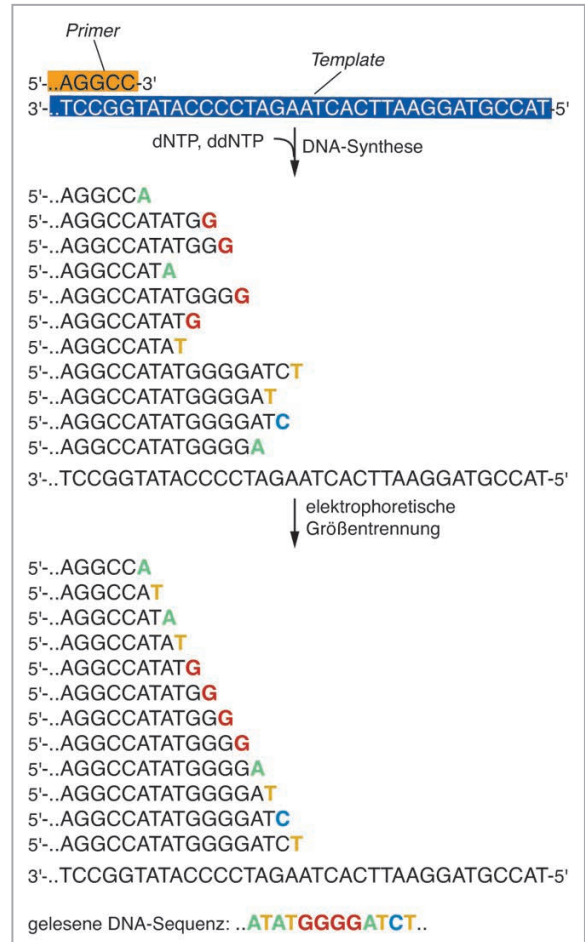
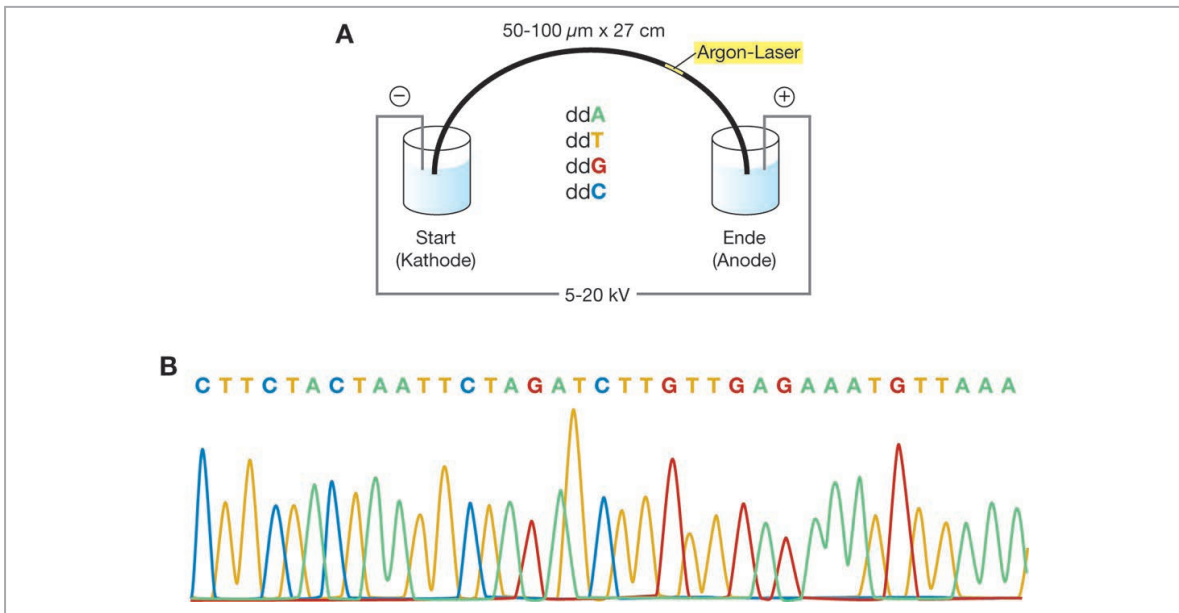


Abb. 2.22 Prinzip der DNA-Sequenzierung nach Sanger: Eine Sequenzierreaktion enthält neben einem Primer und einer Matrizen-DNA eine definierte Mischung aus normalen Desoxynukleotiden (dNTPs) und 2',3'-Didesoxynukleotiden (ddNTPs). Jedes der vier ddNTPs ist mit einem anderen Fluoreszenzfarbstoff gekoppelt. Die DNA-Synthese wird mit einem aus der PCR bekannten zyklischen Protokoll aus Denaturierung, Hybridisierung des Primers und DNA-Synthese durchgeführt. Während der vom Primer ausgehenden DNA-Synthese wird die DNA-Polymerase zu irgendeinem Zeitpunkt statt eines normalen dNTP ein ddNTP in die wachsende Nucleinsäurekette einbauen. An dieser Stelle bricht die DNA-Synthese ab. Da der Zeitpunkt des Kettenabbruchs Zufall ist, bekommt man eine Schar von DNA-Fragmenten, die statistisch alle Längen der Matrize abdecken und sich jeweils um genau eine Base unterscheiden. Diese Schar von DNA-Fragmenten kann man mit geeigneten Methoden auftrennen. An jedem Ende eines solchen Fragments befindet sich ein fluoreszenzmarkiertes ddNTP. Man kann also durch Auslesen der Fluoreszenzsignale vom Primer ausgehend eine DNA-Sequenz ableiten.

zen in der Genfunktionsanalyse. Die weitgehende Automatisierung dieser Technologie hat auch die Entschlüsselung ganzer Genome von Organismen ermöglicht, auch wenn zwischenzeitlich wesentlich effizientere Sequenzierungen der nächsten Generation (*next*



• **Abb. 2.23** Methoden zur Trennung von DNA-Fragmenten bei der DNA-Sequenzierung: (A) In modernen Sequenzierautomaten wird die Kapillarelektrophorese zur Trennung der fluoreszenzmarkierten Sequenzierprodukte eingesetzt. Am Ende der Trennstrecke werden die Fluoreszenzfarbstoffe der ddNTPs angeregt und ausgelesen. (B) Die Reihenfolge der eluierten DNA-Fragmente wird vom Computer ausgewertet und als Elektropherogramm dargestellt. Die unterschiedliche Farbgebung der einzelnen ddNTPs erleichtert die Interpretation der Peaks.

*generation sequencing*, NGS) für die Entschlüsselung kompletter Genome entwickelt wurden.

### Hochdurchsatz-DNA-Sequenzierung

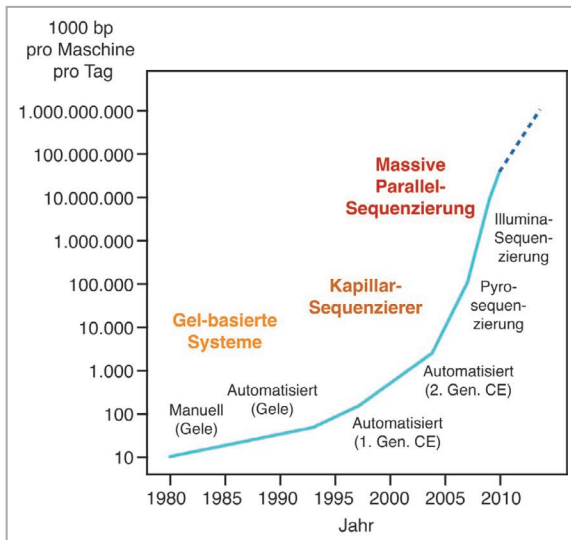
Seit der Publikation der Terminator-Technologie zur DNA-Sequenzierung durch Frederick Sanger wurde die Methode schrittweise immer weiterentwickelt (•Abb. 2.24). Diese Entwicklung wurde maßgeblich angetrieben durch die Mitte der 1990er-Jahre getroffene Vereinbarung, in einem großen internationalen Konsortium die komplette Sequenz des menschlichen Genoms zu erarbeiten. Die Auftrennung fluoreszenzmarkierter DNA-Fragmente durch Kapillarelektrophorese wurde zunächst dadurch effizienter, dass man in Sequenzierautomaten nicht eine Kapillare einsetzte und jede Sequenzierreaktion nacheinander analysierte, sondern Automaten mit beispielsweise 96 parallel zu betreibenden Kapillaren entwickelte. Dadurch konnte eine ungleich größere Datenmenge generiert werden. Die Sequenzierautomaten wurden dann auf immer höheren Probendurchsatz optimiert.

Ein Vergleich verdeutlicht, welche Fortschritte in den letzten ca. 30 Jahren auf diesem Gebiet gemacht wurden: Das etwa 3 Milliarden Basenpaare umfassende menschliche Genom wurde praktisch komplett nach der Sanger-Methode sequenziert. In dem mit öffentlichen Geldern geförderten Humanen Genomprojekt wurden ungefähr 23 Milliarden Basen sequenziert. Das hat ca. 15 Jahre gedauert und Kosten in der Größenordnung von Milliarden Dollar verursacht. Heute würde ein Sequenzierautomat, der nach dem Sanger-Prinzip

arbeitet, für ca. 1 Million Basenpaare Sequenzinformation am Tag produzieren – ein einzelner solcher Automat hätte also immer noch 23 000 Tage oder 63 Jahre gebraucht, um die Rohinformation für das menschliche Genom zu erstellen. Bedenken muss man natürlich, dass viele Sequenzierautomaten weltweit mit dem Projekt beschäftigt gewesen wären, sodass beispielsweise 63 Automaten diese große Herausforderung im Prinzip in einem Jahr hätten erledigen können. Wäre das humane Genom jedoch mit einem der modernen Hochdurchsatz-Sequenzierautomaten mit Methoden der Sequenzierung der nächsten Generation (*next generation sequencing*, NGS) erstellt worden, wären die 23 Milliarden Basenpaare in der Theorie auf einem einzigen Gerät innerhalb weniger Tage analysierbar gewesen.

An diesem Beispiel wird klar, dass man die neuen NGS-Methoden zu Recht als Hochdurchsatz-Sequenzierung bezeichnen kann. Diese Methoden eröffnen heute völlig neue Ansätze in der Genomforschung und dadurch auch in der Arzneistoffforschung. Obwohl es mehrere NGS-Methoden gibt, wollen wir an dieser Stelle nur die derzeit wichtigsten kurz vorstellen: die **Illumina**-Sequenzierung, die **PacBio**-Sequenzierung und die **Nanoporen**-Sequenzierung.

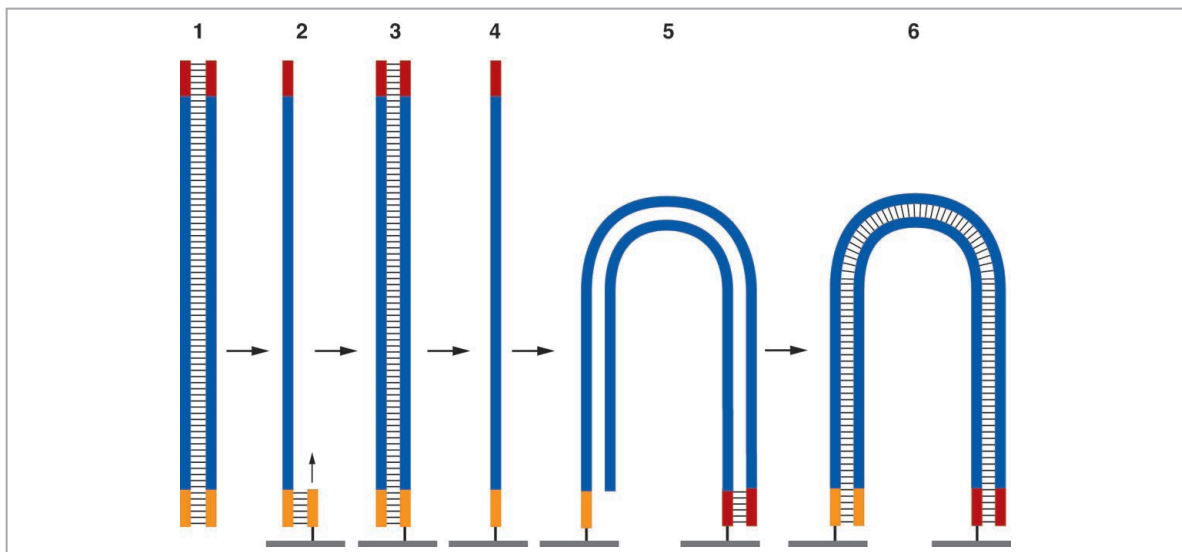
Ende der 1990er-Jahre experimentierten Shankar Balasubramanian und David Klenerman mit einem neuen Sequenzierverfahren, das auf dem Einbau von fluoreszenzmarkierten Terminator-Nukleotiden basiert. Das kommt uns bekannt vor, denn so funktioniert dem Prinzip nach auch die Sanger-Sequenzierung. Das neue



• **Abb. 2.24** Entwicklung der durch verschiedene Sequenziermethoden generierten Datenmengen: Von den Zeiten Anfang der 1980er-Jahre, als Polyacrylamid-Gele für die Trennung der DNA-Fragmente aus Sequenzierreaktionen noch individuell von Hand gefertigt wurden, bis zur Entwicklung von automatisierten Hochdurchsatz-Sequenzierungen, wurde der Output an Sequenzinformationen durch sehr viele Wissenschaftler um mehr als das Millionenfache gesteigert. CE: Kapillarelektrophorese. Quelle: Stratton et al.: Nature 458 (2009), 719 – 724.

Sequenzierverfahren wurde Anfang der 2000er-Jahre von der Firma Solexa zur Marktreife weiterentwickelt, sodass 2006 der erste Sequenzierautomat zur Verfügung stand. Das Sequenzierverfahren war demzufolge zunächst als Solexa-Methode bekannt, bevor die Firma 2007 von Illumina übernommen wurde. Seither wird meist von **Illumina-Sequenzierung** gesprochen, wenn diese Sequenzieretechnologie angewendet wird. Ein solcher Sequenzierautomat produziert heute in einem einzigen Lauf, der etwa sechs Tage dauert, bis zu 8 Milliarden Einzelsequenzen von beispielsweise 125 bp Länge, d. h. bis zu 1000 Milliarden Basenpaare Sequenzinformation pro Woche.

Bei der Illumina-Methode wird eine PCR-Vervielfältigung der zu sequenzierenden DNA auf einem soliden Träger durchgeführt, was als Bridge-PCR bezeichnet wird (• Abb. 2.25). Zunächst werden an die zu sequenzierenden DNA-Moleküle Adapter ligiert. Die DNAs werden dann denaturiert und mit beiden Enden kovalent an die Oberfläche der Fließzelle gebunden. Nach einer PCR-Runde werden die Amplifikate denaturiert und am Ort ihrer Entstehung über ihre Enden wiederum an die Oberfläche der Fließzelle fixiert. Somit entstehen in der Fließzelle dichte Cluster von doppelsträngigen DNA-Molekülen, die anschließend separat von anderen Clustern parallel sequenziert werden. Die



• **Abb. 2.25** Prinzip der Bridge-PCR bei der Illumina-Sequenzierung: (1) An die zu sequenzierende doppelsträngige DNA (blau) werden zunächst zwei unterschiedliche Adapter-Primer ligiert (orange und rot). (2) Die DNA wird dann in eine Reaktionskammer (*flow cell*) gegeben, in der die beiden Adapter-Primer als Einzelstränge kovalent an die Oberfläche gebunden sind. Die zu sequenzierende DNA wird denaturiert und über Hybridisierung mit dem fixierten Adapter-Primer (orange) in der Reaktionskammer gebunden. (3) Der DNA-Strang wird nun von einer DNA-Polymerase kopiert. (4) Danach wird der entstandene DNA-Doppelstrang denaturiert, sodass der entstehende Einzelstrang über den ersten Adapter-Primer (orange) an die Oberfläche gebunden ist und am anderen Ende den zweiten Adapter-Primer trägt (rot). (5) Über diesen kann die DNA jetzt auch mit dem zweiten Ende an die Oberfläche binden (Bridge-Bildung), indem die roten Adapter-Primer hybridisieren. (6) Nun wird die DNA vom roten Primer ausgehend kopiert. Wird diese DNA wieder denaturiert, hängen die Einzelstränge mit einem Ende jeweils an der Oberfläche und können nach erneuter Bridge-Bildung wieder kopiert und die DNA im Sinne einer konventionellen PCR exponentiell vermehrt werden. Es entstehen dichte Cluster von Amplimeren, die dann durch Terminator-Technologie sequenziert werden.

eigentliche DNA-Sequenzierung basiert auf der *Sequencing-by-Synthesis*-Methode mit fluoreszenzmarkierten Nukleotiden. Dabei wird mit modifizierten Basen gearbeitet, deren 3'-Hydroxylgruppen durch eine Schutzgruppe blockiert sind, sodass bei der DNA-Synthese jeweils nur ein einziges neues Nukleotid eingebaut werden kann. Die vier Terminator-Nukleotide werden mit unterschiedlichen Fluoreszenzfarbstoffen markiert, wobei der Farbstoff entweder direkt an die 3'-Hydroxylgruppe gehängt wird und gleichzeitig als Schutzgruppe funktioniert oder der Farbstoff mit der Base des Nukleotids verbunden wird und die 3'-Hydroxylgruppe mit einer separaten Schutzgruppe versehen wird. Man gibt in einer Sequenzreaktion alle vier Terminator-Nukleotide zu und erlaubt den Einbau der jeweils komplementären Base durch eine DNA-Polymerase. Die eingebaute Base wird anhand der Fluoreszenz des Terminator-Nukleotids bestimmt und dann werden die Schutzgruppe am 3'-OH-Ende und der Fluoreszenzfarbstoff abgespalten und ausgewaschen, bevor der nächste Zyklus beginnt.

Aufgrund der im Vergleich mit der Sanger-Sequenzierung großen produzierbaren Datenmengen und der immer wirtschaftlicher werdenden Illumina-Sequenzierung eröffnen sich viele Möglichkeiten der Anwendung dieser Technologie in der Genomforschung. Hier nur einige Anwendungsbeispiele:

- Komplexe Lebensgemeinschaften in interessanten ökologischen Systemen, z. B. in Biofilmen, können allein durch Sequenzierung der darin enthaltenen Genom-DNA charakterisiert werden, d. h. ohne die in den Proben vorhandenen Lebewesen durch Kultivierung anreichern zu müssen. Dieser neue Zweig der Genomik wird auch als Metagenomik bezeichnet.
- Genome von Organismen lassen sich komplett aufklären, deren wirtschaftliche Sequenzierung allein aufgrund ihrer Genomgröße bisher nicht denkbar war. Dies gilt insbesondere auch für tierische Genome, deren Analyse unser Wissen über Funktionen menschlicher Gene bereichern wird (Modellorganismen), was wiederum zum Verständnis pathophysiologischer Prozesse und damit zur Arzneimittelentwicklung beitragen kann.
- Für die Suche nach krankheitsrelevanten DNA-Sequenz-Polymorphismen des Menschen können Genome von Individuen ganz oder teilweise resequenziert werden.
- Die Expression aller zu einem bestimmten Zeitpunkt in einer Zelle exprimierten mRNAs kann in einem einzigen Experiment bestimmt werden. Diese als mRNA-Sequenzierung oder RNA-seq bekannte Methode hat große Bedeutung, beispielsweise für die Charakterisierung von Tumoren, da man heute weiß, dass jeder Tumor ein individuelles (vom gesunden

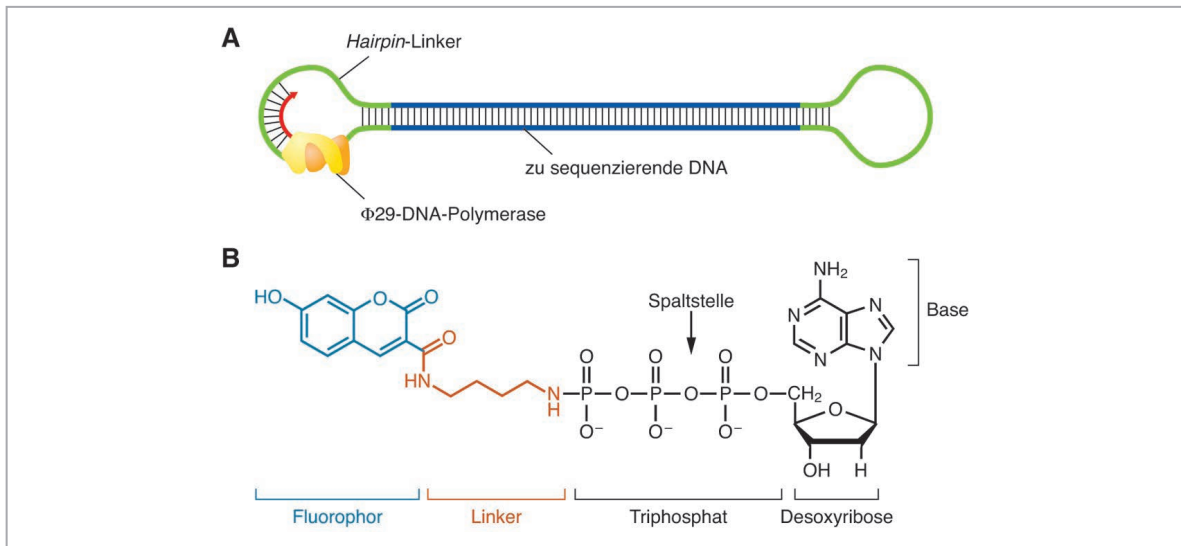
Gewebe verschiedenen) Genexpressionsmuster aufweist. Das tumorspezifische Genexpressionsmuster könnte dann als Biomarker für die Entscheidung über ein geeignetes Therapieregime herangezogen werden. Übrigens wird bei der RNA-seq nicht die mRNA selbst sequenziert. Vielmehr werden alle mRNAs einer Zelle durch die Reverse Transkriptase in doppelsträngige cDNA umgewandelt, die dann sequenziert wird.

- Auch zur Validierung der Herstellungsverfahren von Biopharmazeutika wird die Hochdurchsatz-DNA-Sequenzierung wertvolle Beiträge leisten, z. B. durch die vollständige Resequenzierung der Genome von Produktionszelllinien vor und nach einem Produktionsprozess.

Hochdurchsatz-Sequenzierung wird ständig weiterentwickelt, meist mit dem Ziel, noch mehr Sequenzinformation in kürzerer Zeit und mit wirtschaftlicheren Methoden zu generieren. Die Illumina-Sequenzierung wird mittlerweile als „2. Generation“ der Sequenziermethoden angesehen (die Sanger-Sequenzierung ist die 1. Generation), um sie von weiteren Neuentwicklungen wie beispielsweise der PacBio-Sequenzierung abzugrenzen, die als Drittgenerationen-Sequenziermethode bezeichnet wird.

Bei der **PacBio-Sequenzierung** steht nicht primär die Menge an produzierter Sequenzinformation im Vordergrund, sondern es wird darauf optimiert, möglichst lange, fortlaufende DNA-Sequenzen zu erzeugen. Das grundsätzliche Problem mit der Illumina-Sequenzierung ist nämlich, dass sie zwar große Menge an Einzelsequenzen liefert, diese aber mit ca. 100 – 300 bp recht kurz sind. Wenn man beispielsweise in Genomprojekten versucht, unbekannte genomische Sequenzen anhand von überlappenden Einzelsequenzen virtuell zusammensetzen (zu „assemblieren“), funktioniert das gut, solange die erzeugten Illumina-Sequenzen einzigartig sind und nicht mehrmals an verschiedenen Stellen des untersuchten Genoms vorkommen. Sind die Illumina-Sequenzen aber so kurz, dass sie nur repetitive und keine einzigartige Sequenzinformation enthalten, kann man sie keiner bestimmten Position im Genom zuordnen. PacBio-Sequenzen schaffen hier einen Ausweg, da sie mehrere 10 000 bp lange, ununterbrochene Sequenzen erzeugen und damit auch repetitive Abschnitte eines Genoms überspannen können.

Die PacBio-Sequenzierung wurde von der Firma Pacific BioSciences entwickelt, wodurch sich der Name dieser Drittgenerationen-Sequenziermethode erklärt. Bei dieser Technik werden DNA-Einzelmoleküle sequenziert und die Sequenzen werden in Echtzeit (*real time*) aufgezeichnet, weshalb die Methode auch als *single-molecule real-time sequencing* oder SMRT-Sequenzierung bezeichnet wird. Zur Vorbereitung der Sequen-



• **Abb. 2.26** PacBio-Sequenzierung: (A) SMRTbell-Matrizen bestehen aus einem zu sequenzierenden DNA-Doppelstrang (blau), an den Hairpin-Linker (grün) ligiert werden. Diffundieren diese Moleküle in eine ZMW, werden sie von der kovalent an die Oberfläche der ZMW fixierte  $\Phi 29$ -Polymerase gebunden. Die Polymerase kann im einzelsträngigen Bereich der Hairpins mit der DNA-Polymerisation beginnen. Da die SMRTbell-Matrizen über die Hairpins geschlossen sind, kann die Polymerase mehrere Umläufe der DNA-Synthese absolvieren, sodass konkatenierte Sequenzen der Matrizen-DNA entstehen. Die Syntheserichtung ist durch den roten Pfeil angedeutet. (B) Beispiel für ein in der PacBio-Sequenzierung genutztes, über das Triphosphat-Ende fluoreszenzmarkiertes Desoxynukleotid. Beim Einbau eines solchen Nukleotids während der DNA-Synthese wird der Pyrophosphatrest mitsamt dem Fluoreszenzfarbstoff abgespalten.

zierung werden an einzelne DNA-Moleküle Hairpin-Adapter ligiert (• Abb. 2.26). Die dadurch erhaltenen SMRTbell-Matrizen werden in eine sogenannte SMRT-Zelle gegeben, die aus bis zu 150 000 einzelnen Vertiefungen besteht und als *zero-mode waveguides* (ZMW) bezeichnet werden. Am Boden jedes ZMW ist ein **einziges** Molekül einer DNA-Polymerase gebunden. Bevorzugt wird hier die DNA-Polymerase des Bakteriophagen  $\Phi 29$ , die aus einer einzelnen Proteinkette besteht und eine hohe Prozessivität besitzt, also sehr lange DNA-Polymere herstellen kann. Der Sequenzierreaktion werden die vier erforderlichen Desoxynukleosidtriphosphate (dNTPs) zugesetzt, wobei jedes Nukleotid über das Triphosphat-Ende mit einem anderen Fluoreszenzfarbstoff verbunden ist (Beispiel in • Abb. 2.26). Die  $\Phi 29$ -Polymerase bindet im einzelsträngigen Hairpin-Bereich der SMRTbell-Matrizen und beginnt die DNA-Synthese. Die ZMW sind so gestaltet, dass das Fluoreszenzlicht der dNTPs nur in dem extrem kleinen Volumen um die DNA-Polymerase herum erfasst wird. Während die DNA-Polymerase die SMRTbell-Matrizen kopiert, wird das jeweils erforderliche dNTP kurzzeitig an das Enzym gebunden und damit das Fluoreszenzsignal des jeweiligen dNTPs gemessen. Sobald das Enzym das gebundene Nukleotid in den bereits polymerisierten DNA-Strang einbaut, wird das Pyrophosphat mitsamt dem Fluoreszenzfarbstoff abgespalten und diffundiert aus dem Detektionsvolumen, sodass das Fluoreszenzsignal erlischt. Aus dem zyklischen Binden und Hydrolyse-

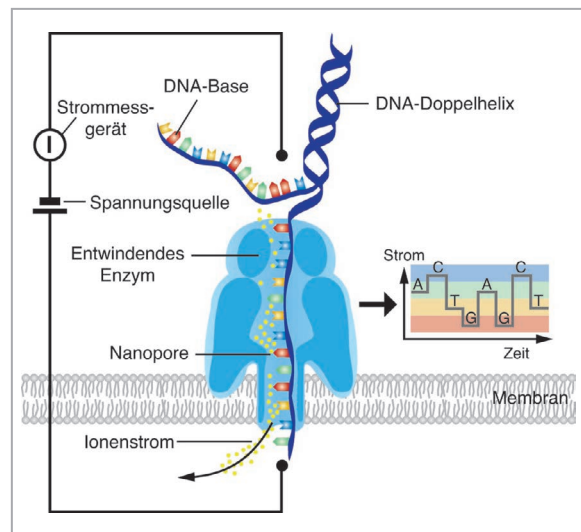
ren eines dNTPs entsprechend der Komplementarität des kopierten Matrizen-Strangs lässt sich somit die Sequenzinformation ableiten. Da die SMRTbell-Matrizen nicht jedes ZMW besetzen bzw. auch mehrere DNA-Moleküle in dasselbe ZMW diffundieren können, sind in einem Lauf pro SMRT-Zelle meist nur zwischen 35 000 und 75 000 ZMWs mit Einzelmolekülen besetzt und liefern auswertbare Sequenzinformationen. Dennoch werden mit der Methode innerhalb weniger Stunden bis zu 1000 Milliarden Basen an Sequenzinformation gewonnen, wobei jede Einzelsequenz mehrere 10 000 bp lang sein kann.

Ein Nachteil der PacBio-Sequenzierung ist derzeit noch die hohe Fehlerrate der Sequenzierung, die bei 10 – 15 % liegt. Das ist viel zu hoch, um valide Information zu liefern, wenn man bedenkt, dass beispielsweise maximal 0,01 % Fehlerrate in Genomprojekten als akzeptabel angesehen wird. Die hohe Fehlerrate der PacBio-Sequenzierung wird dadurch relativiert, dass die  $\Phi 29$ -Polymerase je nach Länge der kopierten SMRTbell-Matrizen in der Lage ist, mehrere „Runden zu drehen“, indem sie am Ende des DNA-Moleküls den Hairpin kopiert und auf dem Gegenstrang weiter synthetisiert. Die resultierenden konkatenierten DNA-Sequenzen werden im Computer übereinandergelegt und ein „Konsensus“ erstellt, der dann der am häufigsten ermittelten und deshalb fehlerlosen Sequenz entspricht.

Die **Nanoporen-Sequenzierung** (*nanopore sequencing*) unterscheidet sich von anderen Sequenziermetho-

den darin, dass die DNA direkt, d. h. ohne die Beteiligung von PCR-basierten Amplifikationsschritten, sequenziert werden kann. Ein zweiter Vorteil der Nanoporen-Sequenzierung ist, dass sie ähnlich wie die PacBio-Sequenzierung zum Teil sehr lange Einzelsequenzen von 10 – 100 kbp generieren kann. Grundlage der Nanoporen-Sequenzierung sind biologische Poren, die in eine für Ionen undurchlässige Membran eingebettet sind. Als biologische Poren können Proteine wie das *Staphylococcus aureus*  $\alpha$ -Hämolysin oder das *Mycobacterium smegmatis* porin A (MspA) verwendet werden. Über die Nanoporen wird eine elektrische Spannung angelegt, sodass sich Ionen durch die Poren bewegen. Die zu sequenzierende DNA wird zunächst durch eine DNA-Helikase entwunden. Die Größe der Nanoporen ist so gewählt, dass ein DNA-Einzelstrang gerade so hindurchpasst. Da jede der vier DNA-Basen unterschiedlich groß ist, wird der Ionenstrom durch die Poren unterschiedlich stark verändert, wenn eine der vier Basen die Pore passiert. Aus den Änderungen der Ionenströme kann direkt eine DNA-Sequenz abgeleitet werden (● Abb. 2.27).

Die größten Vorteile der Nanoporen-Sequenzierung sind die Möglichkeit der direkten DNA-Sequenzierung ohne die Notwendigkeit einer vorherigen Amplifikation der DNA durch PCR und die vergleichsweise langen, kontinuierlichen Leseweiten. Die Nanoporen-Sequenzierung ist deshalb immer dann gut einzusetzen, wenn beispielsweise Abschnitte genomischer DNA aus komplexen und repetitiven Sequenzen bestehen. Solche Sequenzabschnitte sind aus kurzen Einzel-Reads, wie sie beispielsweise die Illumina-Sequenzierung liefert, nur unvollständig zusammensetzbar (assemblierbar)



● Abb. 2.27 Prinzip der Nanoporen-Sequenzierung

und fehlen daher oft in den Genomsequenzen von Organismen. Eine Resequenzierung von Genomen mit der Nanoporen-Technologie kann also helfen, Lücken in Genomsequenzen zu schließen.

Ein Nachteil der Nanoporen-Sequenzierung ist die immer noch recht hohe Fehlerquote, insbesondere weil während der Sequenzierung Basen „übersehen“ werden können, was zu vermeintlichen Deletionen führt, oder auch mehrere Nukleotide gleichzeitig ein Spannungssignal erzeugen und daher eine Nukleotidposition nicht eindeutig bestimmbar ist. Dieser technische Nachteil wird jedoch durch die hohe erreichbare Datenmenge ausgeglichen, sodass die DNA mit einer hohen Cover-

### Zusammenfassung

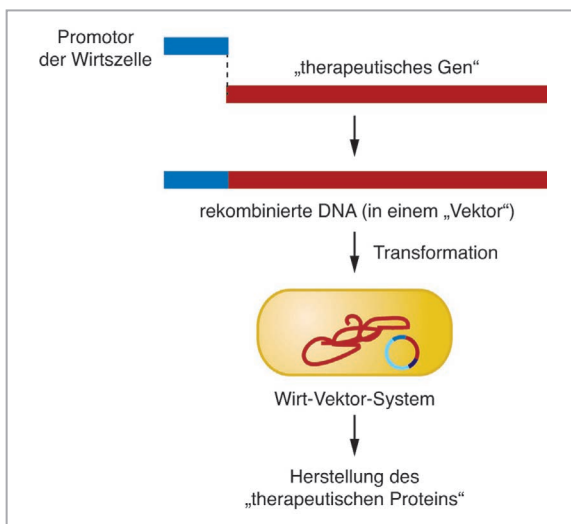
- Für die Synthese größerer DNA-Abschnitte *in vitro* sind DNA-Polymerasen das Mittel der Wahl. Diese Enzyme können einzelsträngige DNA-Abschnitte kopieren, indem sie den fehlenden DNA-Strang entsprechend der Komplementaritätsregeln mit Nukleotiden auffüllen.
- DNA-Polymerasen benötigen dabei einen Startpunkt aus doppelsträngiger DNA. Dieser Startpunkt kann künstlich festgelegt werden, indem gezielt kurze Oligonukleotide (Primer) entworfen werden, die an die zu kopierenden Einzelstrang-DNAs hybridisieren.
- Eine der wichtigsten Methoden der molekularbiologischen Forschung und der Gendiagnostik ist die Polymerase-Kettenreaktion (PCR).
- Die PCR kann eine DNA-Sequenz in einem beliebig komplexen genetischen Hintergrund sehr empfindlich nachweisen und den untersuchten DNA-Abschnitt zu Mengen vervielfältigen, die eine Analyse der DNA-Sequenz und damit eine Verifizierung der Identität der DNA ermöglicht.
- Die Analyse der Basensequenz einer DNA beruht auf einer Modifikation der enzymatischen DNA-Synthese, bei der eine definierte Mischung von normalen Desoxynukleotiden und artifiziiellen Didesoxynukleotiden während des Kopierens eines DNA-Abschnitts vorliegen. Dabei werden DNA-Fragmente erzeugt, die sich in ihrer Länge um jeweils eine Base unterscheiden.
- Sensitive Nachweismethoden erlauben heute eine automatisierte Hochdurchsatz-Sequenzierung, die die Grundlage für die Durchführung komplexer Genomprojekte bildet. Die Sequenzieretechniken werden ständig weiterentwickelt mit dem Ziel, möglichst viel DNA-Information mit geringerem Aufwand und Kosten zu produzieren.

age sequenziert werden kann. Das bedeutet, dass die gleichen Nukleotidpositionen sehr oft wiederholt gemessen werden und aus den Ergebnissen mit bioinformatischen Methoden die korrekte Konsensussequenz errechnet wird.

Die mit der Nanoporen-Sequenzierung erzeugten Änderungen der Ionenströme können in Echtzeit ausgelesen und an einen Computer gestreamt werden, der die entsprechende Sequenz ausgibt. Daraus ergibt sich der Vorteil, dass das Ergebnis schnell vorliegt. Die Geräte sind außerdem mittlerweile stark miniaturisiert worden, sodass eine Sequenzierung von DNA sogar außerhalb spezialisierter Labore „im Feld“ durchgeführt werden kann. So wurden Mini-Nanoporen-Sequenzierer beispielsweise bereits dazu genutzt, Ausbrüche von Ebola direkt vor Ort „in Echtzeit“ zu verfolgen.

### 2.3 Prinzipien der Gentechnik

DNA-Rekombinationstechnologie (Gentechnologie) ist für die Herstellung von Proteinwirkstoffen von besonderer Bedeutung. Die Herstellung von Proteinwirkstoffen erfordert die Expression von Genen, die in den meisten Fällen aus dem menschlichen Genom stammen, in einem artfremden, biotechnologisch zugänglichen Zellsystem. Die Genexpression läuft in allen Lebewesen gleich ab, d. h. die in einem Gen codierte Information wird in allen



• **Abb. 2.28** Wirt-Vektor-System: Eine Einheit zur Herstellung von Biopharmazeutika besteht aus der genetischen Information (therapeutisches Gen), die als Protein exprimiert und als Wirkstoff etabliert werden soll (therapeutisches Protein). Zur Herstellung der Expressionseinheit muss das therapeutische Gen mit einem in der Wirtszelle aktiven Promotor durch DNA-Rekombination verbunden werden. Diese funktionelle Expressionseinheit wird dann durch Transformation in eine geeignete Wirtszelle eingebracht, die das therapeutische Protein produziert.

Organismen prinzipiell gleich umgesetzt (►Kap.3). Allerdings braucht jedes Gen ein Steuerelement (Promotor), damit die Transkription gestartet werden kann. Eben diese Promotoren sind aber spezifisch für Organismen und innerhalb mehrzelliger Organismen sogar zelltypspezifisch. Eine Expression eines menschlichen Gens in einem nichtmenschlichen Zellsystem kann also nur gelingen, wenn wir Methoden haben, um die jeweiligen zellspezifischen Promotoren mit der gewünschten Gensequenz zu verbinden – also DNA-Moleküle zu rekombinieren. Nicht zuletzt deshalb ist im Europäischen Arzneibuch in der Monographie „DNA-rekombinationstechnisch hergestellte Produkte“ von „Wirt-Vektor-Systemen“ die Rede, wobei der „Vektor“ jeweils die rekombinante Transkriptionseinheit beschreibt (•Abb.2.28).

Obwohl die generelle Struktur der DNA seit den klassischen Arbeiten von James Watson und Francis Crick seit Mitte der 1950er-Jahre bekannt war und es auch klar wurde, dass die genetische Information in der Sequenz der Basen innerhalb einer DNA festgelegt ist, war es lange undenkbar, die riesigen DNA-Moleküle detailliert zu charakterisieren, die Basenfolgen zu ermitteln oder gar unterschiedliche DNA-Moleküle miteinander zu verbinden. Es fehlte schlicht die erforderliche Technologie. Möglich wurde dies erst, als Anfang der 1970er-Jahre zwei Voraussetzungen geschaffen waren:

- Enzyme wurden entdeckt und charakterisiert, die DNA als Substrat erkennen und modifizieren.
- Methoden wurden entwickelt, mit denen es möglich war, verschiedene DNA-Moleküle zu rekombinieren.

Um die Werkzeuge der Gentechnologen kennenzulernen, wollen wir uns zunächst mit den Enzymen befassen, die DNA als Substrat erkennen. Danach werden wir sehen, dass das Bakterium *Escherichia coli* mit Abstand der wichtigste Organismus innerhalb der molekularbiologischen Forschung mit gentechnischen Methoden ist. Fast jede gentechnische Modifikation wird zunächst in *E. coli* realisiert, auch wenn geplant ist, die modifizierte genetische Information später in einem anderen Organismus zu exprimieren.

Warum müssen wir neu kombinierte DNA in *E. coli* amplifizieren? Rekombinationen von DNA-Fragmenten müssen auf der Ebene einzelner Moleküle durchgeführt werden, was ein äußerst ineffizienter Vorgang ist. Man steht daher regelmäßig vor dem Problem, dass man untersuchen muss, ob ein gewünschtes Rekombinationsereignis überhaupt stattgefunden hat und man benötigt ein Werkzeug, um diese rekombinierte DNA von negativen Rekombinationsereignissen zu trennen. Es wäre durchaus eine Herausforderung, auf der Ebene einzelner Moleküle zu analysieren, ob das Zusammenfügen eines DNA-Moleküls mit einem anderen DNA-Molekül wirklich stattgefunden hat. Geschweige denn,

dass man diese rekombinierte DNA in ausreichender Menge zur weiteren Verwendung in die Hand bekäme. Aus diesem Dilemma gibt es aber einen Ausweg: die Durchführung der DNA-Rekombination in Plasmiden und eine Amplifikation dieser Plasmide in *E. coli*-Zellen. Wird ein Plasmid in eine Bakterienzelle eingebracht, wird nicht nur das Plasmid in der einzelnen Bakterienzelle amplifiziert, sondern auch die Bakterien selbst. Diese vermehren sich als Kolonie von einer einzelnen Zelle zu mehreren Millionen Zellen. Da es sich innerhalb einer Bakterienkolonie um genetisch identische Zellen, also Klone, handelt, wird der gesamte Vorgang der Amplifikation von DNA-Molekülen in *E. coli* im Rahmen der DNA-Rekombination (des gentechnischen Arbeitens) im Labor auch klonieren genannt.

An folgender Beispielrechnung wird deutlich, dass eine einfache Kultivierung von Bakterienzellen aus **einem Molekül** rekombinierter DNA, die in ein geeignetes Vektorsystem integriert und in eine Bakterienzelle eingeschleust wurde, quasi über Nacht zu einer **DNA-Menge** amplifiziert werden kann, die wir mit den gängigen DNA-modifizierenden Enzymen und DNA-analytischen Verfahren leicht erfassen können:

- Wir gehen davon aus, dass bei jeder Transformation eines Bakteriums lediglich ein **einziges** neu rekombiniertes Plasmid-Molekül in die Zelle gelangt.
- Die Plasmide werden innerhalb **einer** Bakterienzelle stark amplifiziert (100 – 700 Plasmidkopien pro Zelle).
- Wird die transformierte Bakterienzelle in Flüssigkultur vermehrt, lässt sich die Dichte der Bakterienkultur durch die Absorption (A) bei 600 nm messen: 1 ml einer Bakterienkultur der Dichte  $A_{600} = 1$  entspricht ca.  $8 \times 10^8$  Bakterien.
- Enthält jede Bakterienzelle einer solchen Kultur nur ca. 100 Kopien des Plasmids, entspricht das einem Amplifikationsfaktor von  $8 \times 10^{10}$ ; das entspricht  $1,3 \times 10^{-13}$  mol Plasmid einer Größe von 3000 bp oder 0,3 µg Plasmid. Diese **Menge** an Plasmid-DNA entspricht einer gut sichtbaren „Bande“ im Agarosegel, wenn wir die DNA mit Ethidiumbromid anfär-

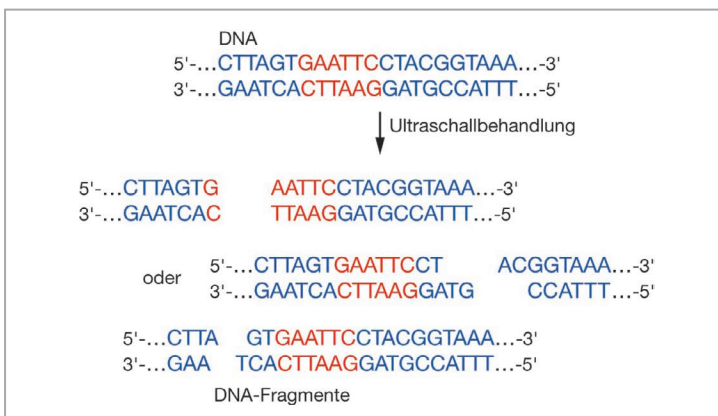
ben, d. h. wir können durch herkömmliche DNA-Analytik untersuchen, ob das isolierte Plasmid rekombinant ist oder nicht.

Gentechnisches Arbeiten bedeutet die Isolierung, Modifizierung und Verknüpfung bestimmter DNA-Moleküle zu neuen Funktionseinheiten. Damit die neu-kombinierte DNA schließlich in *E. coli* stabil erhalten bleibt, sind geeignete Vektorsysteme wie z. B. Plasmide nötig sowie geeignete Methoden, mit deren Hilfe sich die neu kombinierte DNA effizient in die *E. coli*-Zellen einschleusen lässt. Bevor wir uns diesen Methoden widmen, wollen wir uns zunächst anschauen, wie wir DNA-Moleküle in einer bestimmten, vorhersagbaren Weise miteinander verbinden können. Der Schlüssel hierfür sind die Restriktionsendonukleasen.

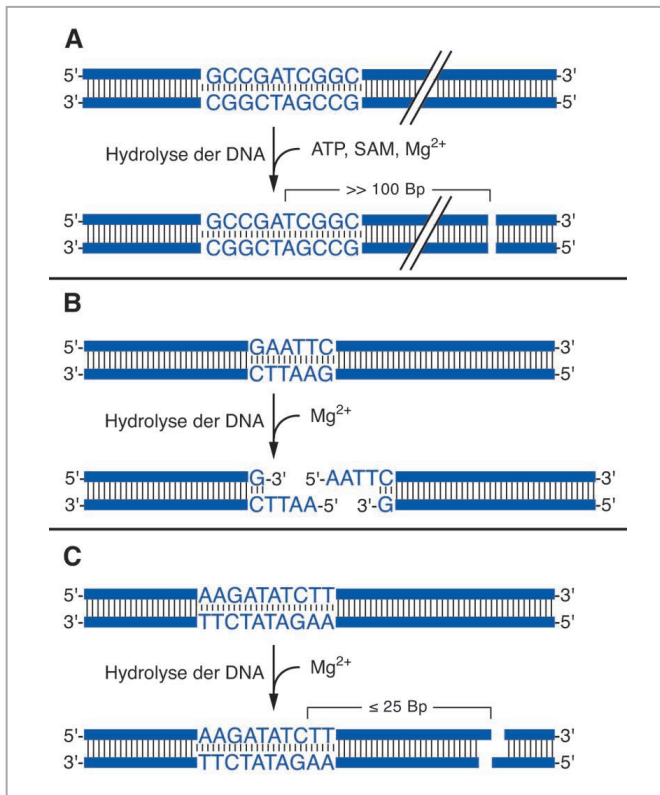
### 2.3.1 Gezieltes Schneiden von DNA: Restriktionsendonukleasen

Wie können wir DNA-Fragmente gezielt neu kombinieren? DNA kann beispielsweise durch mechanisches Scheren durch Behandlung mit Ultraschall zerbrochen werden (● Abb. 2.29). An welchen Stellen die DNA bricht, ist jedoch zufällig, sodass wir mit dieser Methode Gene nicht gezielt isolieren und klonieren können. Für die DNA-Rekombination brauchen wir eine Methode, mit der DNA spezifisch, d. h. vorhersagbar, fragmentiert werden kann. Dies erreichen wir durch Hydrolyse doppelsträngiger DNA mit Restriktionsendonukleasen (im Labor-Jargon auch kurz als Restriktionsenzyme bezeichnet). Der Einsatz dieser Enzyme hat gegenüber dem unspezifischen Scheren der DNA den Vorteil, dass Restriktionsendonukleasen ausschließlich an spezifischen DNA-Sequenzen schneiden und dabei charakteristische Enden erzeugt werden, die durch die Spezifität der verwendeten Restriktionsenzyme bestimmt werden und daher vorhersagbar sind.

Wir unterscheiden vier große Gruppen von Restriktionsendonukleasen. Allen diesen Enzymen ist gemeinsam, dass sie ausnahmslos doppelsträngige DNA als Substrate verwenden und dass sie bestimmte DNA-



- **Abb. 2.29** Unspezifisches Fragmentieren von DNA: Längere DNA-Abschnitte können unspezifisch gebrochen werden, z. B. indem man die DNA mit Ultraschall behandelt. Wäre unser Ziel, die rot eingefärbte DNA-Sequenz gezielt zu brechen, wäre dies mit dieser Methode nicht möglich, da es Zufall ist, an welchem Ort ein Doppelstrangbruch jeweils auftritt. Auch die durch die Brüche resultierenden DNA-Enden sind verschieden. Dies erschwert das gezielte Rekombinieren verschiedener DNA-Abschnitte.



• **Abb. 2.30** Die drei Hauptgruppen der Restriktionsendonukleasen: (A) Typ-I-Restriktionsendonukleasen benötigen als Co-Substrate ATP, SAM (S-Adenosylmethionin) und Mg<sup>2+</sup>. Die Hydrolyse erfolgt weit entfernt von der Erkennungssequenz. Diese Restriktionsendonukleasen spielen in der Gentechnologie praktisch keine Rolle. (B) Typ-II-Restriktionsendonukleasen hydrolysieren die DNA innerhalb der jeweils spezifischen Erkennungssequenz. Spalten sie diese Sequenz nicht versetzt, entstehen DNA-Fragmente mit „glatten Enden“. Spalten die Enzyme die Erkennungssequenz versetzt, entstehen DNA-Fragmente, bei denen entweder das 5'-Ende über das 3'-Ende oder aber das 3'-Ende über das 5'-Ende hinausragt. In der Gentechnologie gehören die Typ-II-Restriktionsendonukleasen zu den wichtigsten Enzymen. (C) Typ-III-Restriktionsendonukleasen hydrolysieren die DNA – ähnlich wie die Typ-I-Restriktionsendonukleasen – außerhalb der Erkennungssequenz. Allerdings liegen die Spaltstellen deutlich näher an der Erkennungssequenz, sodass diese Restriktionsendonukleasen teilweise auch in der Gentechnologie verwendet werden.

Sequenzmotive erkennen und binden, bevor sie die DNA schneiden (• Abb. 2.30).

- **Typ-I-Restriktionsendonukleasen** sind komplexe Enzyme, die sowohl Nuklease- als auch Methyltransferase-Aktivität besitzen. Diese Enzyme benötigen als Cofaktoren ATP, S-Adenosylmethionin und Magnesium-Ionen. Die eigentliche Hydrolyse der DNA geschieht nicht an der Erkennungsstelle des Enzyms auf der DNA, sondern in einer Entfernung von einigen hundert Basenpaaren. Aus diesem Grund werden Typ-I-Restriktionsendonukleasen praktisch nie für gezielte Klonierungsstrategien verwendet.
  - **Typ-II-Restriktionsendonukleasen** benötigen meist nur Magnesium als Cofaktor und spalten die DNA innerhalb der Erkennungssequenz. Die große Mehrzahl der in der Molekularbiologie verwendeten Enzyme sind Typ-II-Restriktionsendonukleasen. Das hängt damit zusammen, dass beliebige DNA-Moleküle, egal aus welcher Spezies sie stammen, nach Hydrolyse mit demselben Restriktionsenzym die gleichen DNA-Enden aufweisen. Wir können also DNA-Abschnitte, die wir mit derselben Restriktionsendonuklease geschnitten haben, über ihre kompatiblen Enden miteinander verbinden (rekombinieren). Die Typ-II-Restriktionsendonukleasen sind nicht mit Methyltransferasen assoziiert. Allerdings können die Erkennungsstellen für Typ-II-Restriktionsendonukleasen von separaten Methyltransferasen modifiziert werden. Dies kann dazu führen,
- dass die Restriktionsendonuklease ihre Zielsequenz nicht mehr hydrolysieren kann. Umgekehrt gibt es aber auch Restriktionsendonukleasen, die eine Spaltstelle nur dann erkennen, wenn in dieser Sequenz eine Base methyliert ist. Die unterschiedlichen Eigenschaften und Anforderungen der Typ-II-Restriktionsendonukleasen hat zu ihrer weiteren Aufteilung in elf Untergruppen (A, B, C, E, F, G, H, M, P, S und T) geführt. Die für die Klonierung von DNA relevanten Restriktionsendonukleasen sind überwiegend in der Untergruppe Typ IIP zu finden, in die Enzyme eingeteilt werden, die symmetrische Erkennungssequenzen auch symmetrisch schneiden.
- **Die Typ-III-Restriktionsendonukleasen** sind komplexe Enzyme, die eine DNA ATP-abhängig in einem spezifischen Abstand – in der Regel bis zu 25 bp – von der Erkennungssequenz entfernt hydrolysieren. Aus diesem Grund sind die überlappenden Sequenzen, die von Typ-III-Restriktionsendonukleasen erzeugt werden, nicht miteinander kompatibel. Diese Restriktionsenzyme werden daher eher selten für Klonierungen verwendet.
  - **Typ-IV-Restriktionsendonukleasen** schneiden nur modifizierte, also methylierte, hydroxymethylierte oder glucosylhydroxyethylierte DNA. Bisher ist nur wenig über diese Enzyme bekannt, ihre Erkennungssequenzen scheinen recht unspezifisch zu sein. Deshalb sind diese Enzyme für den molekularbiologischen Laboralltag irrelevant.

Für das Rekombinieren von DNA haben die Typ-II-Restriktionsendonukleasen die größte Bedeutung. An dieser Stelle sollten wir uns fragen, worin denn eigentlich eine Erkennungsstelle für eine Restriktionsendonuklease besteht. In der Regel ist die Erkennungssequenz eine palindromartige Sequenz. Palindrome sind bekanntlich Wörter oder Sätze, die von vorne und von hinten gelesen den gleichen Sinn ergeben. Beispiele sind „Otto“ oder „Anna“. So ist es auch mit den Erkennungssequenzen der Restriktionsendonukleasen (Abb. 2.31). Allerdings gilt die Spiegelbildlichkeit nicht nur für die Sequenz auf einem einzelnen DNA-Strang, sondern für die Sequenz des DNA-Doppelstrangs. Eine typische Erkennungssequenz wäre z. B. die Basenabfolge 5'-GAATTC-3' auf dem „oberen“ DNA-Strang. Da die beiden komplementären DNA-Stränge antiparallel verlaufen, lautet die Sequenz des „unteren“, komplementären Strangs ebenfalls 5'-GAATTC-3' (Abb. 2.31). Restriktionsendonukleasen funktionieren in der Regel als Homodimere, wobei ein Monomer jeweils die Zielsequenz auf einem DNA-Strang erkennt und schneidet. Schneiden beispielsweise die Dimer-Untereinheiten des Enzyms EcoRI gleichzeitig die Sequenz 5'-GAATTC-3' zwischen dem G und dem A, entstehen wie in Abb. 2.31 gezeigt kurze einzelsträngige „Überhänge“ an der Schnittstelle.

Typ-II-Restriktionsenzyme hydrolysieren doppelsträngige DNA auf bestimmte Art und Weise, sodass drei unterschiedliche Arten von Enden erzeugt werden:

- Die meisten Enzyme schneiden die Zielsequenz versetzt, aber symmetrisch zur Mitte. Dabei entstehen als Resultat DNA-Fragmente, bei denen ein Ende als einzelsträngige DNA über das andere Ende hinausragt. In Abb. 2.32 ist dies nochmals am Beispiel der Restriktionsendonuklease EcoRI gezeigt, die die DNA-Sequenz 5'-GAATTC-3' schneidet. Durch die versetzte, aber symmetrische Spaltung des DNA-Doppelstrangs entstehen DNA-Enden, bei denen das 5'-Ende übersteht. Die überstehenden Enden von zwei DNA-Fragmenten, die mit demselben Restriktionsenzym geschnitten wurden, können sich durch Hybridisierung der komplementären Basen aneinanderlagern. Man bezeichnet diese Art Endstück als „klebrige“ DNA-Enden (*sticky ends*).
- Umgekehrt kann eine Restriktionsendonuklease den DNA-Doppelstrang symmetrisch in der Weise schneiden, dass das 3'-Ende über das 5'-Ende hinausragt. Als Beispiel ist in Abb. 2.32 das Enzym KpnI genannt.
- Als dritte Möglichkeit können Enzyme die Erkennungssequenz nicht versetzt hydrolysieren, d. h. die DNA-Stränge sind an den beiden Enden gleich lang. Hierzu zählt das in Abb. 2.32 aufgeführte Enzym PvuII. Solche DNA-Enden werden als „stumpf“ (*blunt ends*) bezeichnet.



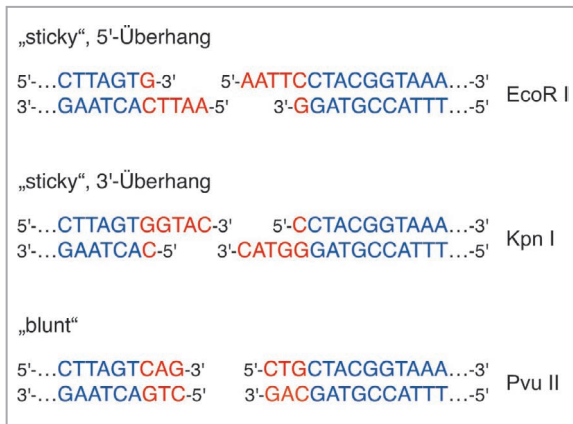
Abb. 2.31 Spezifisches Fragmentieren von DNA mit Typ-II-Restriktionsendonukleasen: Typ-II-Restriktionsendonukleasen erkennen eine spezifische DNA-Sequenz, hier rot markiert. Wir bemerken, dass diese Sequenz ein Palindrom bildet, d. h. sie lautet auf beiden DNA-Strängen – in antiparalleler Richtung gelesen – jeweils gleich. Die Restriktionsendonukleasen schneiden die DNA an der Erkennungsstelle in einer bestimmten Weise, sodass definierte Enden entstehen. Dies erleichtert das Zusammenfügen beliebiger DNA-Ende zu rekombinierten DNAs, sofern wir diese DNAs mit der gleichen Restriktionsendonuklease „geschnitten“ haben.

Restriktionsendonukleasen sind wichtige Werkzeuge in der Gentechnologie, die für viele Aufgaben eingesetzt werden können. Einige dieser Aufgaben werden wir noch kennenlernen:

- Schneiden wir DNA-Abschnitte unterschiedlicher Herkunft mit der gleichen Restriktionsendonuklease, können wir über die kompatiblen Enden z. B. ein menschliches Gen mit den Kontrollregionen eines nicht humanen Wirtsorganismus verbinden, um ein Wirt-Vektor-System für die Herstellung eines Proteinwirkstoffs für die Therapie einer Krankheit zu etablieren.
- Mithilfe verschiedener Restriktionsendonukleasen können wir das gesamte Genom eines Organismus in einzelne DNA-Fragmente zerlegen, um z. B. krankheitsrelevante Gene zu kartieren und zu isolieren oder Genome in ihrer Ganzheit zu sequenzieren.
- In der Gendiagnostik können wir das Vorhandensein oder das Fehlen von Erkennungsstellen für Restriktionsendonukleasen dazu nutzen, um Mutationen in Genen zu diagnostizieren. Man spricht von Restriktionsfragmentlängen-Polymorphismen.

### 2.3.2 Rekombinieren von DNA-Fragmenten: DNA-Ligasen

Die weitaus verbreitetste Methode, um freie DNA-Enden miteinander zu verbinden, die zuvor durch Restriktionsendonukleasen geschnitten wurden, ist die Nutzung einer DNA-Ligase. Die häufig benutzte DNA-Ligase des Bakteriophagen T4 schließt mithilfe der Cofaktoren  $Mg^{2+}$  und ATP eine Phosphodiesterbindung zwischen einer 3'-Hydroxylgruppe und einer benachbarten 5'-Phosphatgruppe in einem DNA-Strang; der Vorgang wird Ligation genannt. Dabei muss



• **Abb. 2.32** Produkte von Typ-II-Restriktionsendonukleasen: Restriktionsendonukleasen schneiden DNA meistens versetzt, sodass entweder 5'- oder 3'-überhängende einzelsträngige Enden entstehen. Ein Zusammenfügen solcher Enden wird durch Hybridisierung der vier überstehenden Basen erleichtert, weshalb man solche Enden als „klebrig“ oder *sticky* bezeichnet. Einige Restriktionsendonukleasen schneiden die DNA gerade durch, sodass „stumpfe“ Enden (*blunt*) entstehen. Da das Zusammenfügen stumpfer Enden nicht von der Hybridisierung überstehender Basen erleichtert wird, können wir beliebige stumpfe Enden zusammenfügen, um DNAs miteinander zu rekombinieren.

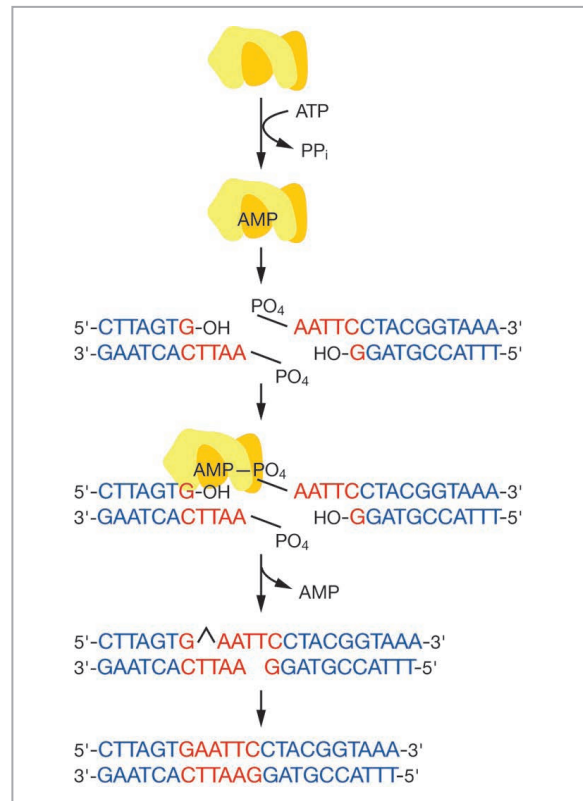
die DNA-Ligase zunächst aktiviert werden, indem ein Molekül AMP nach Hydrolyse von ATP auf einen Lysinrest im aktiven Zentrum des Enzyms übertragen wird. In diesem aktivierten Zustand kann das Enzym eine DNA nach freien 5'-Phosphatenden absuchen. Wird eine solche Stelle identifiziert, wird der AMP-Rest auf das 5'-Phosphatende der DNA in Form einer 5' → 5'-Pyrophosphatstruktur übertragen. Ist ein freies 3'-OH-Ende in der DNA verfügbar, erfolgt anschließend eine Transesterifizierung unter Freisetzung von AMP und Verknüpfung der DNA-Enden (• Abb. 2.33).

### 2.3.3 Amplifizieren rekombinierter DNA: Transformation von Bakterien

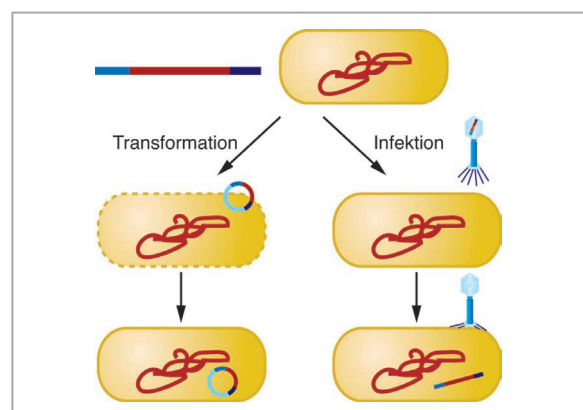
#### Transformationsmethoden

Einer der wichtigsten Schritte bei der Etablierung eines rekombinanten Organismus ist das Einbringen der zuvor *in vitro* rekombinierten DNA in eine Bakterienzelle. Dieser Prozess wird als Transformation bezeichnet, wenn die DNA auf chemischem oder physikalischem Weg in die Zelle eingebracht wird. Im Gegensatz dazu spricht man von Infektion, wenn die DNA mithilfe eines Bakteriophagen, also auf biologischem Weg, in eine *E. coli*-Zelle eingeschleust wird (• Abb. 2.34).

Transformieren lassen sich sogenannte kompetente *E. coli*-Zellen (• Abb. 2.35). Deren Zellwände und Membranen wurden durch Behandlung beispielsweise mit CaCl<sub>2</sub> temporär für Fremd-DNA durchlässig gemacht. Für die eigentliche Transformation werden die kompetente



• **Abb. 2.33** Verknüpfung von DNA-Enden durch T4-DNA-Ligase: Die T4-DNA-Ligase schließt mithilfe der Cofaktoren Mg<sup>2+</sup> und ATP eine Phosphodiesterbindung zwischen einer 3'-Hydroxylgruppe und einer benachbarten 5'-Phosphatgruppe in einem DNA-Strang. Das Enzym wird aktiviert, indem nach Hydrolyse von ATP ein Molekül AMP an einen Lysinrest im aktiven Zentrum der Ligase bindet. Der AMP-Rest wird auf 5'-Phosphatenden von DNA mit Einzelstrangbrüchen übertragen, wobei eine 5' → 5'-Pyrophosphatstruktur ausgebildet wird. Danach erfolgt eine Transesterifizierung mit einem benachbarten freien 3'-OH-Ende unter Abspaltung von AMP und Verknüpfung der DNA-Enden.

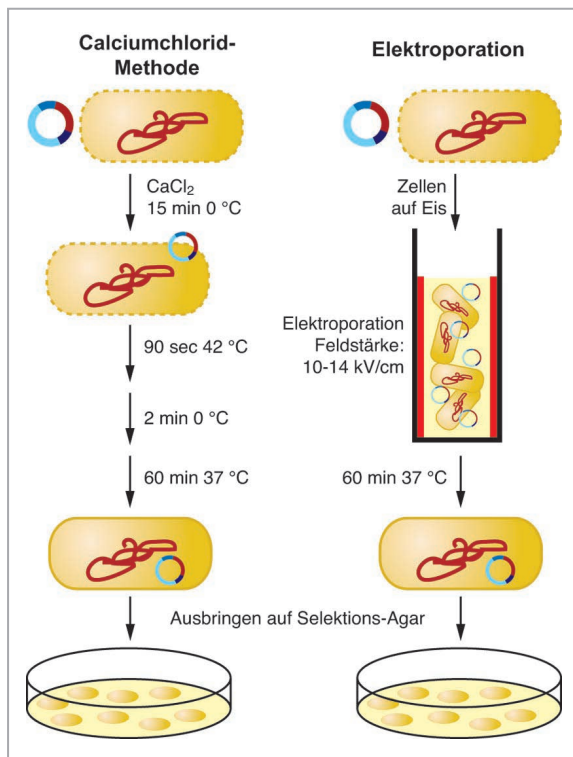


• **Abb. 2.34** Transformation versus Infektion: Wird DNA passiv über eine vorgeschädigte Zelloberfläche in eine bakterielle Zelle aufgenommen, bezeichnet man das als Transformation. Die aktive Aufnahme von Fremd-DNA mithilfe von Phagen ist dagegen eine Infektion.

tenten Zellen zusammen mit der zu transformierenden DNA in einer  $\text{CaCl}_2$ -Lösung zunächst bei  $0\text{ }^\circ\text{C}$  inkubiert. Durch die  $\text{Ca}^{2+}$ -Ionen in der Zellsuspension wird die DNA als Polyanion mikrokristallin auf der Oberfläche der kompetenten Zellen präzipitiert. Danach folgt eine Inkubation für einige Sekunden bei  $42\text{ }^\circ\text{C}$  und schließlich werden die Zellen für einige Minuten bei  $25\text{ }^\circ\text{C}$  belassen. Durch diese Temperaturwechsel wird die DNA über Mechanismen, die nicht genau bekannt sind, passiv in die Zelle aufgenommen. Zur Regeneration der Bakterienzellen wird der Transformationsan-

satz mit Medium verdünnt, 60 Minuten bei  $37\text{ }^\circ\text{C}$  inkubiert und schließlich auf eine Agarplatte transferiert (● Abb. 2.35). Die Transformationseffizienzen liegen bei diesem Verfahren bei  $10^5 - 10^7$  transformierten Zellen pro  $\mu\text{g}$  DNA. Dies hängt von der Güte der kompetenten Zellen und vom Transformationsprotokoll ab.

Deutlich höhere Transformationsraten erhält man durch ein Verfahren, das als Elektroporation bezeichnet wird (● Abb. 2.35). Hierzu werden die Zellen, die zuvor salzfrei gewaschen wurden, zusammen mit der zu transformierenden DNA in einem Feld hoher elektrischer Konduktivität inkubiert. Bei der Entladung (z. B.  $2500\text{ V}$ ,  $200\text{ }\Omega$ ,  $25\text{ }\mu\text{F}$ ) gelangt die DNA dann in die Zellen. Derartige Systeme erreichen Transformationseffizienzen von bis zu  $10^{10}$  Transformanten pro  $\mu\text{g}$  DNA.



● **Abb. 2.35** Die bakterielle Transformation: Um Bakterien zu transformieren, müssen ihre Zellwände durch eine Behandlung mit  $\text{CaCl}_2$  vorgeschädigt sein. Wir bezeichnen derartige Bakterien als „kompetent“. Links: Calciumchlorid-Methode. Die Zellen werden in der  $\text{CaCl}_2$ -Lösung zunächst 30 Minuten bei  $0\text{ }^\circ\text{C}$  inkubiert. Dann erhöhen wir die Inkubationstemperatur schnell auf  $42\text{ }^\circ\text{C}$  („Hitzeschock“) und inkubieren anschließend bei  $25\text{ }^\circ\text{C}$ . Zur Regeneration der Zellhülle verdünnen wir danach die  $\text{CaCl}_2$ -Lösung mit Medium und inkubieren bei der optimalen Wachstumstemperatur für *E. coli* von  $37\text{ }^\circ\text{C}$ . Schließlich werden die Zellen auf eine Schale mit einem geeigneten Selektionsagar ausplattiert. Rechts: Elektroporation; dafür verwenden wir Zellen, die vor der Transformation salzfrei gewaschen werden. Diese „elektrokompetenten“ Zellen werden zusammen mit der DNA in eine Küvette gegeben, die an zwei gegenüberliegenden Seiten mit dünnen Plattenelektroden ausgekleidet sind. Nach Anlegen hoher Feldstärken wird die DNA bei der Entladung in die Bakterienzellen aufgenommen. Zur Regeneration werden die Zellen für 60 Minuten bei  $37\text{ }^\circ\text{C}$  inkubiert und danach auf Selektionsagar ausplattiert.

### Plasmide

Plasmide sind natürlich vorkommende, zirkuläre DNA-Moleküle, die in Bakterienzellen unabhängig vom bakteriellen Chromosom repliziert werden. In den vergangenen Jahren wurde für die gentechnische Anwendung eine Vielzahl von Plasmiden entwickelt, die sich prinzipiell von natürlichen Plasmid-Vorbildern wie den R-Plasmiden ableiten. R-Plasmide tragen eines oder mehrere Resistenz-Gene für Antibiotika („R“ steht für Resistenz). Außerdem codieren R-Plasmide für die sogenannten *tra*-Gene (*tra* = transfer), die erforderlich sind, damit Bakterien die R-Plasmide durch einen als Konjugation bezeichneten Prozess untereinander austauschen können. Bakterien können durch Aufnahme eines einzigen R-Plasmids Multiresistenzen gegen Antibiotika erwerben.

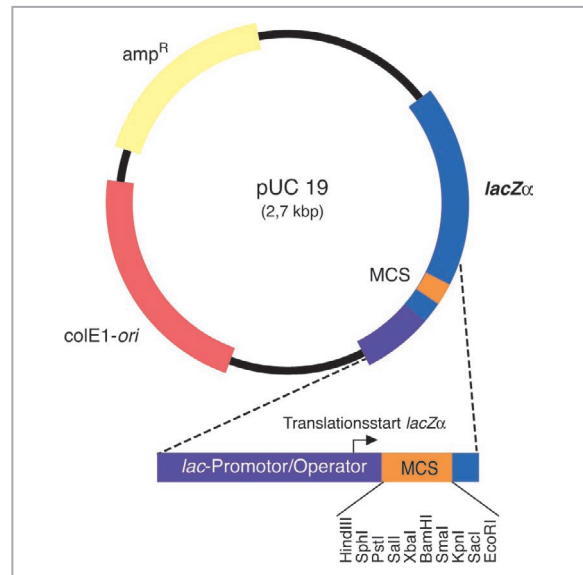
Für die gentechnische Nutzung hat man sich von den R-Plasmiden abgeschaut, was man tun muss, damit Plasmide in der Bakterienzelle episomal, also unabhängig vom Bakterienchromosom, repliziert werden können. Außerdem hat man von R-Plasmiden gelernt, wie man die Expression der auf Plasmiden codierten Antibiotika-Resistenzgene nutzen kann, um das Einbringen artifizierender Plasmide in Bakterienzellen durch Transformation zu kontrollieren. Die natürlichen R-Plasmide selbst sind allerdings aus verschiedenen Gründen für die gentechnische Anwendung ungeeignet:

- R-Plasmide kommen nur in 1 – 2 Kopien pro Bakterienzelle vor.
- R-Plasmide sind mit ca. 100000 bp zu groß und zu instabil für eine effiziente Neukombination mit fremden DNA-Fragmenten durch gentechnische Methoden.
- R-Plasmide können durch Konjugation zwischen Bakterien einer Kultur ausgetauscht werden, solange sie die *tra*-Gene tragen; dieser Vorgang ist der gewünschten klonalen Vereinzelung von gentechnisch herbeigeführten Rekombinationsereignissen natürlich abträglich und muss durch das Entfernen der *tra*-Gene unterbunden werden.

Allen in der Gentechnik eingesetzten Plasmiden ist daher gemeinsam, dass sie von R-Plasmiden nur noch die Sequenzen enthalten, die für das gentechnische Arbeiten absolut notwendig sind. Der Vorteil liegt dann darin, dass diese Vektoren wesentlich kleiner sind (3000 – 10 000 bp) als R-Plasmide, was sie deutlich besser handhabbar macht:

- Alle Vektoren besitzen eine Sequenz, die als Replikationsursprung (*origin of replication*, kurz *ori*) bezeichnet wird (● Abb. 2.36). Diese Sequenz wird von der DNA-Polymerase aus *E. coli* als Initiationspunkt für die Replikation erkannt und stellt sicher, dass das Plasmid unabhängig von der genomischen DNA des Bakteriums repliziert werden kann und zusätzlich zur genomischen DNA in der Zelle persistiert.
- Ferner enthalten alle Vektoren einen oder mehrere selektierbare Marker in Form von Resistenzgenen für Antibiotika. Der verbreitetste Marker ist das Gen für die  $\beta$ -Lactamase (● Abb. 2.36). Dieses Enzym verleiht Bakterienzellen die Fähigkeit, in Gegenwart des Antibiotikums Ampicillin zu wachsen. Somit sind Zellen, die ein Plasmid mit einem  $\beta$ -Lactamase-Gen aufnehmen, gegen Ampicillin resistent. Auch andere Antibiotikum-Resistenzen – z. B. gegen Tetracyclin, Chloramphenicol, Kanamycin oder Neomycin – können als Marker auf Plasmiden verwendet werden.
- Die dritte wichtige Komponente aller Klonierungsvektoren ist ein speziell ausgewiesener Bereich, der die zu klonierenden fremden DNA-Fragmente aufnehmen soll. Dieser Bereich, die sogenannte multiple Klonierungsstelle (*multiple cloning site*, MCS), enthält auf kleinem Raum eine Vielzahl von Erkennungssequenzen für unterschiedliche Restriktionsendonukleasen, die jeweils nur einmal in dem Plasmid schneiden (● Abb. 2.36).

Wir können uns nun vorstellen, dass wir ein beliebiges DNA-Fragment in ein Plasmid einbauen können, indem wir das DNA-Fragment und den Vektor jeweils mit einem Restriktionsenzym schneiden, dessen Erkennungssequenz Teil der MCS des Plasmids ist. Wie können wir in einem DNA-Rekombinationsexperiment unterscheiden, ob ein Plasmid eine uns interessierende DNA aufgenommen hat? Um dies zu untersuchen, wird häufig das  $\beta$ -Galactosidase-Gen (*lacZ*) aus *E. coli* als Indikator-Gen auf dem Klonierplasmid verwendet (● Abb. 2.37). Die  $\beta$ -Galactosidase ist ein 116 kDa-Protein, das in seiner aktiven Form ein Homotetramer bildet. Jedes Monomer kann funktionell in zwei Teile unterteilt werden: die N-terminale  $\alpha$ -Domäne und die C-terminale  $\omega$ -Domäne. Interessanterweise kann man  $\alpha$  und  $\omega$  gentechnisch voneinander trennen und als zwei getrennte Proteine exprimieren, wobei beide Fragmente für sich jeweils inaktiv sind. Werden aber  $\alpha$  und  $\omega$  in derselben Zelle exprimiert, können sie spontan



● **Abb. 2.36** Ein typischer Klonierungsvektor: Das Plasmid pUC19 enthält den codierenden Bereich für das  $\alpha$ -Komplement der  $\beta$ -Galactosidase als Reportergen (*lacZ $\alpha$* , blau). Die Transkription dieses Gens wird durch den regulierbaren *lac*-Promotor/Operator kontrolliert (● Abb. 4.9 in ► Kap. 4.2). Im vorderen Bereich des  $\beta$ -Galactosidase-Gens befindet sich eine multiple Klonierungsstelle (MCS, orange), die sich dadurch auszeichnet, dass sie eine Vielzahl von Erkennungssequenzen für Restriktionsendonukleasen enthält, die das Plasmid jeweils nur einmal schneiden. Wird in die MCS ein DNA-Fragment eingefügt, wird das  $\beta$ -Galactosidase-Gen zerstört. Dies ist die Basis für das spezifische Reportersystem („Blau-weiß-Selektion“), das in ● Abb. 2.37 erklärt ist. Das Plasmid enthält weiterhin ein Resistenzgen für ein Antibiotikum ( $\beta$ -Lactamase bzw. *amp<sup>R</sup>*) und einen bakteriellen Replikationsursprung für Plasmide (*colE1-ori*).

assoziiieren und so eine funktionelle  $\beta$ -Galactosidase rekonstituieren. Deshalb bezeichnet man  $\alpha$  auch als das  $\alpha$ -Komplement oder den  $\alpha$ -Donor, während  $\omega$  häufig als  $\alpha$ -Akzeptor bezeichnet wird.

Fast alle modernen Klonierungsvektoren enthalten die codierende Region des  $\alpha$ -Komplements der  $\beta$ -Galactosidase. Bemerkenswerterweise liegt die MCS eines Klonierungsvektors in der Regel innerhalb des  $\alpha$ -Komplements. Dabei wird die MCS so in das Gen für das  $\beta$ -Galactosidase inseriert, dass sie den Leserahmen für die  $\beta$ -Galactosidase-mRNA nicht zerstört. Deshalb bleibt das  $\alpha$ -Komplement trotz der inserierten MCS aktiv und kann mit dem  $\omega$ -Fragment eine aktive  $\beta$ -Galactosidase bilden. Bauen wir Fremd-DNA in die MCS ein, wird der Leserahmen für das  $\alpha$ -Komplement durch die neu eingebrachte DNA unterbrochen und die enzymatische Aktivität der  $\beta$ -Galactosidase geht verloren. Diese Inaktivierung lässt sich durch einen Farbttest nachweisen. Die transformierten Zellen werden hierzu in Kulturschalen kultiviert, deren Agar mit dem Indikator X-Gal (5-Bromo-4-chloro-

3-indolyl- $\beta$ -D-galactopyranosid) imprägniert wurde. Wird die  $\beta$ -galactosidische Bindung dieses Indikators durch die  $\beta$ -Galactosidase hydrolysiert, entsteht nach Luftoxidation der blaue Farbstoff 5,5'-Dibromo-4,4'-dichloroindigo und die Bakterienkolonien färben sich blau. Wurde allerdings durch Insertion einer Fremd-DNA in die MCS das Gen für die  $\beta$ -Galactosidase inaktiviert, bleiben die Kolonien weiß. So lassen sich diese Zellen als Träger eines rekombinanten Plasmids als weiße Kolonien identifizieren (Abb. 2.37).

Klonierungsvektoren liegen in einer Anzahl von mehr als 50 Kopien, in der Regel jedoch zu mehreren Hundert Kopien pro Zelle vor. Solche Plasmide werden als High-Copy-Vektoren bezeichnet. Hohe Kopienzah-

len sind meist vorteilhaft und erwünscht, da man die Plasmid-DNA aus relativ wenigen Bakterienzellen in guter Ausbeute isolieren kann. Manchmal ist es jedoch sinnvoll oder sogar erforderlich, mit speziellen Plasmiden zu arbeiten, die in geringerer Kopienzahl in *E. coli* persistieren. Dies ist beispielsweise dann der Fall, wenn die klonierte DNA besonders groß ist, sodass die vielen Plasmidkopien normaler Vektoren die Zelle zu stark mit DNA überfrachten würden. Bedeutsam ist der Einsatz von Low-Copy-Vektoren (Kopienzahlen < 20) aber auch, wenn das Produkt der klonierten DNA für die *E. coli*-Zelle unverträglich ist. Dies führt dazu, dass die Zellen schlecht oder gar nicht wachsen. Würden hier High-Copy-Vektoren verwendet, wäre zu viel geni-

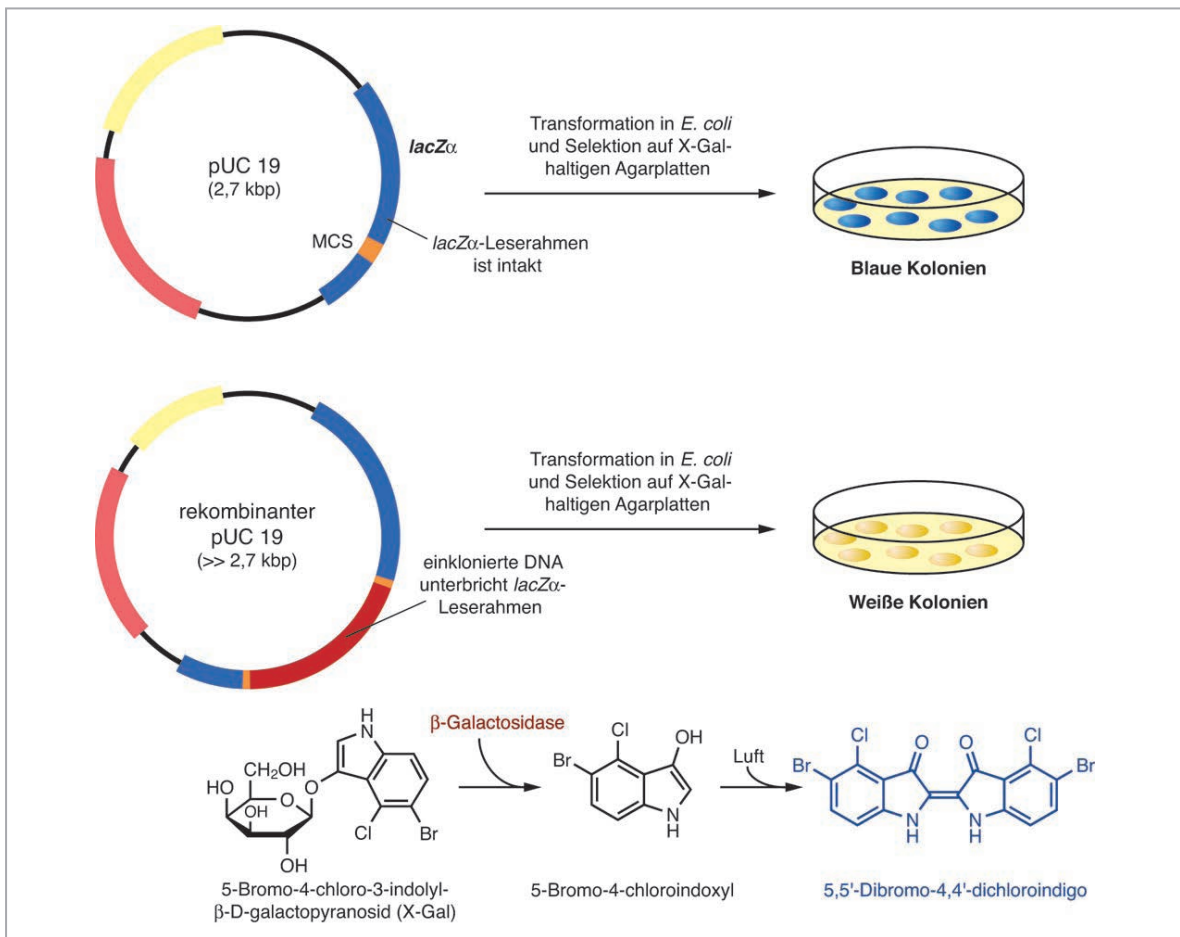


Abb. 2.37 Blau-weiß-Selektion mit dem  $\beta$ -Galactosidase-Reportergen: Zellen, die eine chromosomale Deletion im sogenannten  $\alpha$ -Komplement der  $\beta$ -Galactosidase (*lacZ $\alpha$* ) aufweisen, können kein aktives Enzym bilden. Das Gen für das fehlende  $\alpha$ -Komplement kann durch Plasmide wie pUC19 (Abb. 2.36) in Zellen eingeführt werden (blauer Bereich auf dem pUC19-Plasmid). Zellen mit derartigen Plasmiden können also wieder  $\beta$ -Galactosidase synthetisieren, und wir können die Aktivität des Enzyms durch Zugabe des Substrats 5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranosid (X-Gal) nachweisen. X-Gal wird von der  $\beta$ -Galactosidase gespalten und dann zu einem blauen Indigofarbstoff oxidiert. Die Kulturen färben sich also blau. In Plasmiden wie pUC19 ist eine multiple Klonierungsstelle (MCS) in das Gen für das  $\alpha$ -Komplement eingebaut. Die MCS unterbricht den Leserahmen für das  $\alpha$ -Komplement, sodass mit pUC19 transformierte Bakterienzellen in Gegenwart von X-Gal blaue Kolonien bilden. Die Klonierung eines DNA-Fragments in die MCS zerstört allerdings in der Regel den Leserahmen für das  $\alpha$ -Komplement, sodass Bakterienzellen, die ein derartiges rekombinantes Plasmid aufgenommen haben, keine funktionsfähige  $\beta$ -Galactosidase synthetisieren können. Die Kolonien bleiben daher auch in Gegenwart des Indikators X-Gal farblos.

ches Material in der Zelle, das für ein „unverträgliches Protein“ codieren würde.

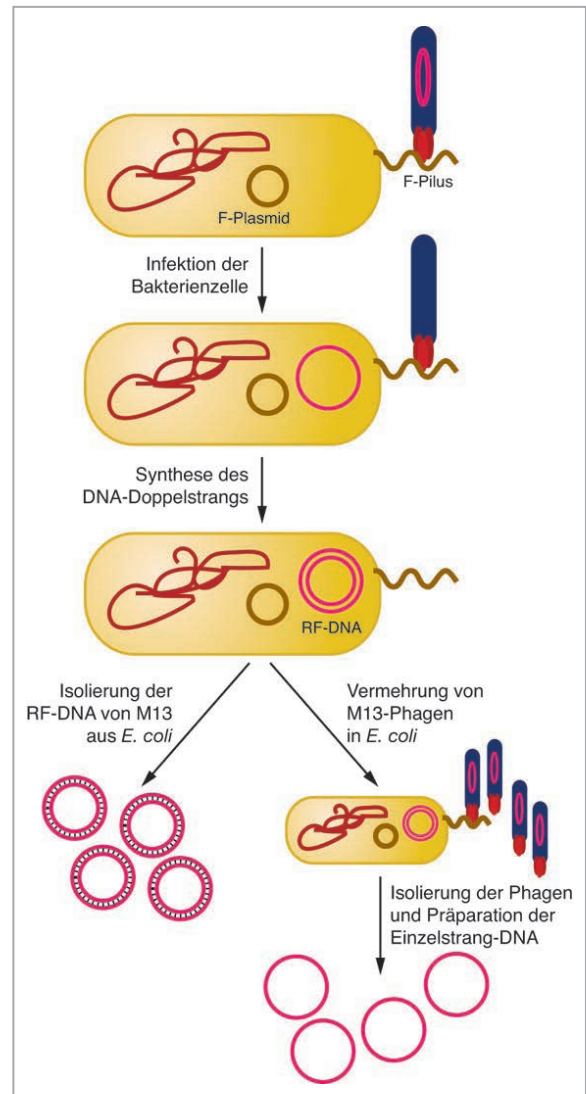
### M13-Phagen und Phagemid-Vektoren

M13 gehört zu der Gruppe der filamentösen Phagen. Diese enthalten ein einzelsträngiges, ringförmiges DNA-Molekül als Genom. Der Umgang mit M13-Phagen war zu Beginn der Gentechnologie essenziell, da die von Frederick Sanger etablierte Methode der DNA-Sequenzierung zunächst ausschließlich mit Einzelstrang-DNA als Matrize durchgeführt werden konnte. Dies hat sich mit Einführung der Cycle-Sequencing-Variante der Sanger-Sequenzierung geändert, sodass der Umgang mit M13 stark zurückging. Diese Phagen sollen hier dennoch kurz erwähnt werden, weil sie in der Entwicklung von vollständig humanen therapeutischen Antikörpern die Grundlage der sogenannten Phage-Display-Methode bilden (►Kap. 5).

Die Infektion der Bakterienzelle durch M13 erfolgt über den sog. F-Pilus (◉Abb. 2.38). Über diese Struktur können Bakterien untereinander genetisches Material austauschen. Nur „männliche“ Bakterien tragen einen F-Pilus, da nur sie ein sog. Fertilitäts-Plasmid (F-Plasmid) enthalten. Nach dem Eindringen in die Wirtszelle wird die einzelsträngige DNA des Phagen durch bakterielle Enzyme in einen DNA-Doppelstrang umgewandelt. Diese doppelsträngige DNA bezeichnet man als die replikative Form oder auch als RF-DNA (◉Abb. 2.38). Nur in der replikativen Form kann M13-DNA mit gentechnischen Methoden bearbeitet werden, da nur doppelsträngige DNA von Restriktionsenzymen geschnitten werden kann und damit für DNA-Rekombinationen zugänglich ist.

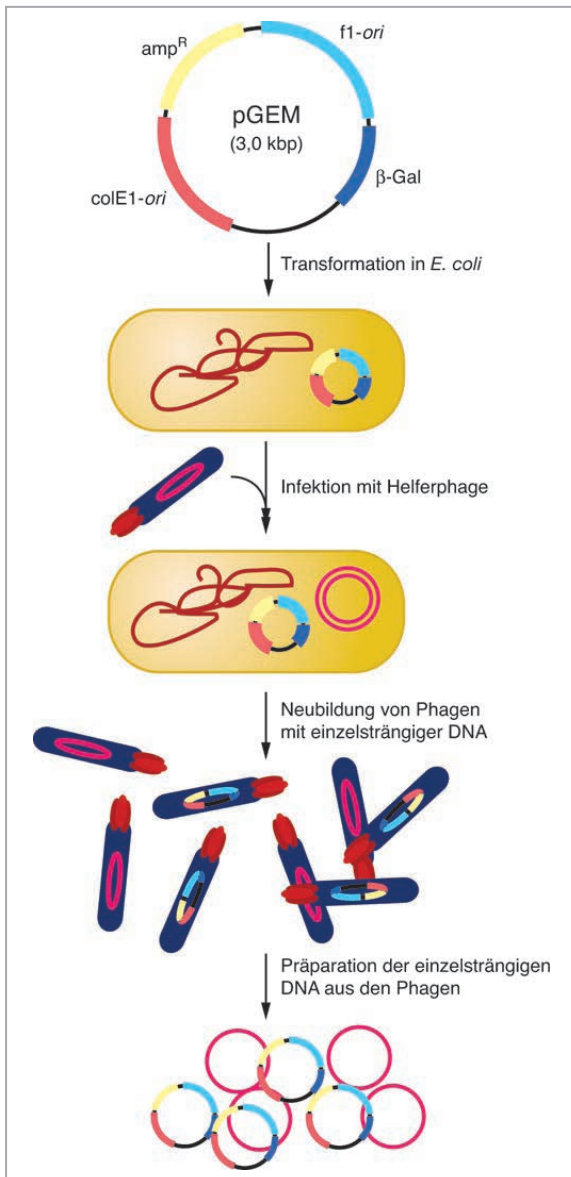
Ähnlich wie andere Vektoren, die für gentechnische Zwecke eingesetzt werden, enthalten M13-Vektoren die genetische Information für das  $\alpha$ -Komplement der  $\beta$ -Galactosidase und eine multiple Klonierungsstelle (MCS). So können auch in diesem Vektorsystem rekombinante Klone von nichtrekombinanten Klonen über die Aktivität der  $\beta$ -Galactosidase (Blau-weiß-Selektion; vgl. ◉Abb. 2.37) unterschieden werden. Infizierte Zellen erkennt man daran, dass diese langsamer wachsen als nicht infizierte Zellen. Auf einem konfluenten *E. coli*-Rasen erscheinen daher punktförmige, lichtere Bereiche infizierter Zellen als sog. Plaques, weil diese Bereiche eine geringere Zelldichte aufweisen. Diese Plaques sind trüb, da M13-Phagen die Bakterienzellen nicht lysieren. Sie verlassen vielmehr die infizierten Zellen über den Pilus, also auf dem gleichen Weg, wie sie in die Zelle eingedrungen sind.

**Phagemid** ist ein Kunstwort aus **Phage** und **Plasmid**. Phagemid-Plasmide können entweder als normales Plasmid oder als in Phagenpartikel verpackte Einzelstrang-DNA existieren. Solche Plasmide sind der RF-Form filamentöser Phagen ähnlich, tragen jedoch keine



◉ **Abb. 2.38** M13-Phagen: M13 gehört zu den filamentösen Phagen. Dies sind *E. coli*-spezifische Phagen, die jedoch nur solche *E. coli*-Zellen infizieren, die ein Fertilitätsplasmid (F-Plasmid) besitzen und daher F-Pili ausbilden. Sie binden an diese F-Pili und entlassen über die Pili ihre DNA in das Zellinnere. Die DNA der filamentösen Phagen ist einzelsträngig und wird in der *E. coli*-Zelle zu einer doppelsträngigen, replikativen Form (RF-DNA) ergänzt. Nur diese Variante ist für die Klonierung brauchbar, da Restriktionsendonukleasen nur doppelsträngige DNA hydrolysieren. Parallel zur Replikation des Phagen-genoms werden auch die Strukturproteine der Phagen synthetisiert. Filamentöse Phagen lysieren die infizierten Zellen nicht. Vielmehr verlassen sie die Zellen wieder über den Pilus. Die Phagen – und damit die einzelsträngigen DNA-Formen – können aus dem Kulturüberstand präpariert werden. Die klonierbare, doppelsträngige Phagen-DNA wird dagegen konventionell wie andere Plasmid-DNA aus den infizierten Bakterienzellen isoliert.

phagencodierten Gene und sind deshalb nicht infektiös. Phagemid-Plasmide tragen zwei Replikationsursprünge (*ori*). Der erste ist ein typischer Plasmid-*ori*



● **Abb. 2.39** Phagemid-Vektoren: Plasmide, die neben dem bakteriellen Replikationsursprung für Plasmide (ColE1-ori) auch einen Replikationsursprung eines filamentösen Bakteriophagen (hellblau) tragen, wie beispielsweise die Plasmide der pGEM-Reihe, werden als Phagemide bezeichnet. Von Phagemid-Vektoren kann man auf elegante Art einzelsträngige DNA isolieren. Hierzu muss die Zelle, die ein derartiges Plasmid trägt, mit einem sog. Helferphagen infiziert werden. Dieser Phage liefert die Proteine, die zur Herstellung einzelsträngiger DNA und der phagenspezifischen Strukturproteine erforderlich sind. Man erhält dann eine Mischung von Phagen, die das Genom des Helferphagen oder des Phagemids in einzelsträngiger Form tragen. Wenn Helferphagen wie M13K07 benutzt werden, die weitgehend replikationsdefekt sind, isoliert man praktisch reine Phagemid-Einzelstrang-DNA.

und sorgt dafür, dass die Bakterienzelle die DNA wie ein normales Plasmid in Doppelstrang-Form repliziert.

Dieses Plasmid kann mit den üblichen DNA-modifizierenden Enzymen bearbeitet werden. Zusätzlich trägt ein Phagemid-Plasmid aber noch einen Replikationsursprung für den M13-ähnlichen Phagen f1 (f1-ori; ● Abb. 2.39). Infiziert man eine Bakterienzelle, die einen Phagemid-Vektor enthält, mit einem f1-Helferphagen, lässt sich die Plasmid-DNA auch als DNA-Einzelstrang isolieren, da ein DNA-Strang des Plasmids durch den Helferphagen verpackt wird.

### Zusammenfassung

- Die DNA-Rekombinationstechnologie erlaubt die Verknüpfung zweier DNA-Fragmente zu neuen funktionellen Einheiten. Obwohl die eigentliche Verknüpfung von DNA-Fragmenten von Enzymen katalysiert wird, ist diese auf der Ebene einzelner Moleküle stattfindende Reaktion sehr ineffizient und benötigt ein Repertoire an Methoden, mit denen unterschieden werden kann, ob ein Plasmid rekombinant ist oder nicht.
- Um DNA-Rekombinationsereignisse analysieren zu können, muss die DNA zunächst von einem Einzelmolekül zu einer DNA-Menge amplifiziert werden, die mit herkömmlichen Methoden der DNA-Analytik erfassbar ist.
- Dabei bilden bakterielle Plasmide die Basis der DNA-Rekombinationstechnologie. Will man überprüfen, ob die Kombination von zwei DNA-Molekülen erfolgreich war, werden die Ligationsreaktionen zwischen Plasmiden und DNA-Fragmenten in Bakterien transformiert und die transformierten Bakterien klonal vereinzelt.
- Daher wird das Rekombinieren von DNA in der Laborsprache auch als Klonieren bezeichnet. Die am häufigsten verwendeten Vektoren sind bakterielle Plasmide, obwohl auch von Bakteriophagen abgeleitete Vektorsysteme existieren. Ein Plasmid ist dadurch gekennzeichnet, dass es einen eigenen Replikationsursprung besitzt und unabhängig vom bakteriellen Chromosom repliziert wird. In der Gentechnik verwendet man auf den Plasmiden natürlich vorkommende Resistenzgene gegen Antibiotika, um zu gewährleisten, dass die Plasmide in den Bakterienzellen verbleiben und dort repliziert werden.
- Die Rekombination von DNA-Fragmenten wird durch eine DNA-Ligation erreicht. DNA-Ligasen schließen Phosphodiesterbindungen zwischen zwei DNA-Molekülen, die danach zu einem DNA-Strang verbunden sind.

## 3 Grundsätze der Proteinexpression

3.1	Umsetzung genetischer Information in Proteine . . . . .	44
3.2	Genexpression in Prokaryonten . . . . .	46
3.3	Genexpression in Eukaryonten . . . . .	54
3.4	Posttranslationale Modifikationen an Proteinen . . . . .	57

### 3.1 Umsetzung genetischer Information in Proteine

Proteine sind Makromoleküle, die aus einer linearen Abfolge von Aminosäuren bestehen. Die Abfolge der Aminosäuren in einem Protein, also die Aminosäuresequenz, wird als die Primärstruktur bezeichnet. Letztlich ist es jedoch die räumliche Struktur eines Proteins, die Sekundär- und Tertiärstruktur, die die Funktionalität eines Proteins ausmacht. Die Natur benutzt 20 verschiedene Aminosäuren zur Herstellung von Proteinen. Die physikochemischen Eigenschaften der Seitenketten der Aminosäuren in einem Protein bestimmen, in welche dreidimensionale Struktur (Konformation) sich ein Protein faltet. Dadurch wiederum wird bestimmt, welche Oberflächen Proteine nutzen können, um z. B. mit anderen Proteinen zu interagieren oder biochemische Reaktionen zu katalysieren.

Die Information, welche Aminosäure in einem Protein an welcher Stelle der Primärsequenz steht, ist in den Genen festgelegt. In der klassischen Genetik wurde der Begriff „Gen“ ursprünglich als die kleinste vererbare Einheit definiert. In der molekularen Genetik ist es eher gebräuchlich, ein Gen als eine Informationseinheit aufzufassen, die für ein Protein codiert. Allerdings ist diese Definition unvollständig, denn manche Gene enthalten Informationen für RNAs, die nicht in Proteine übersetzt werden, sondern als RNAs Funktionen in der Zelle ausüben. Auch neigen wir in Zeiten der molekularbiologischen Erforschung von Zellfunktionen dazu, die Kontrolleinheiten (Promotoren, Terminatoren), die für die Umsetzung der genetischen Information erforderlich sind, zu den Genen hinzuzurechnen. Heute kann der Begriff „Gen“ auch als eine „Transkriptionseinheit“ verstanden werden.

Der Fluss der genetischen Information verläuft in allen Organismen von der DNA (den Genen, die die Information speichern) über RNAs (die die Information zum Ort der Translation transportieren) zu den Proteinen, die die Information ausführen. Da Nucleinsäuren und Proteine biochemisch sehr unterschiedliche

Moleküle sind, muss es einen biochemischen Prozess geben, der die „Nucleinsäure-Sprache“ in die „Aminosäure-Sprache“ übersetzt. Diesen Vorgang nennt man **Genexpression**, und man unterscheidet die **Transkription** – also das Abschreiben eines Gens in eine RNA – von der nachfolgenden **Translation** – also das Übersetzen des Nucleinsäure-Codes in eine Abfolge von Aminosäuren. Nur in wenigen Fällen kann die Richtung des Informationsflusses von RNA zu DNA quasi umgedreht sein. Dieser Vorgang, der ein besonderes Merkmal der Retroviren ist, ist die sog. **reverse Transkription**.

Zu den Hauptzielen der Gentechnik im Bereich pharmazeutisch/medizinischer Anwendungen gehört die Expression genetischer Information aus dem Menschen in Mikroorganismen bzw. pflanzlichen oder tierischen Zellen, mit anderen Worten: die Herstellung von Biopharmazeutika. Generell sprechen wir von **heterologer Genexpression**, wenn ein beliebiger Organismus eine für ihn fremde genetische Information (ein Transgen) exprimiert.

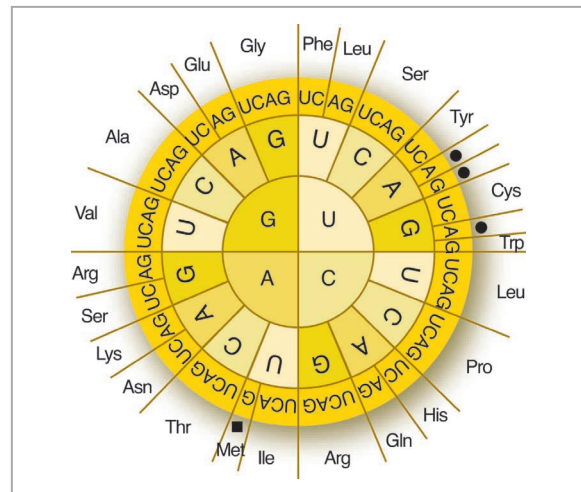
Um erfolgreich menschliche Proteinwirkstoffe durch heterologe Expression herstellen zu können, müssen wir uns die zum Teil erheblichen mechanistischen Unterschiede in der Genexpression in den jeweiligen für die heterologe Expression infrage kommenden Wirtszellen vor Augen führen. Deshalb wollen wir in den folgenden Abschnitten dieses Kapitels kurz auf die Unterschiede in der Genexpression zwischen Prokaryonten (Bakterien) und Eukaryonten eingehen. Diese Unterschiede sind letztlich der Grund, warum eine heterologe Expression ohne DNA-Rekombinationstechnologie nicht realisierbar ist, denn die speziesspezifischen Kontrollelemente der Genregulation in den Produzenten müssen in der Regel mit den zu exprimierenden Genen verbunden werden.

Eine der wichtigsten Erkenntnisse aus der molekularbiologischen Forschung der letzten Jahrzehnte ist diese: Der genetische Code gilt in der belebten Natur auf der Erde universell. Daraus folgt, dass sich die Bedeutung, also die in der DNA niedergelegte Information für ein Protein, nicht verändert – egal, in welchem

Organismus sich diese Information befindet. Nur deshalb können wir überhaupt daran denken, ein menschliches Protein als Wirkstoff z. B. in Bakterien herzustellen. Ein menschliches Gen wird in allen fremden (nicht menschlichen) Zellen in ein Protein mit identischer Primärsequenz translatiert – allerdings nur dann, wenn wir dafür sorgen, dass der jeweilige Organismus das heterologe Gen als Transkriptionseinheit verstehen und umsetzen kann.

Die in Proteinen verwendeten Aminosäuren werden in den Genen als Basentriplets codiert (● Abb. 3.1). Im molekulargenetischen Sprachgebrauch heißen diese Tripletts auch Codons. Die Natur verwendet deshalb Basentriplets, weil dadurch mithilfe der vier DNA-Basen alle 20 Aminosäuren codiert werden können. Wäre der genetische Code durch Basendupletts repräsentiert, könnte die Zelle nur 16 ( $4^2$ ) der 20 benötigten Aminosäuren codieren. Mit Tripletts sind es 64 ( $4^3$ ) Codierungsmöglichkeiten. Wie wir in ● Abb. 3.1 erkennen können, bleiben die 44 überschüssigen Codierungsmöglichkeiten nicht ungenutzt: Die Natur verwendet diese Codons, um die meisten der 20 benötigten Aminosäuren durch mehrere Tripletts zu definieren. So wird z. B. die Aminosäure Prolin durch vier Codons repräsentiert, Serin sogar durch sechs Codons (● Abb. 3.1). Das hat den Vorteil einer Pufferung von Mutationen: Unsere DNA ist ständigen Veränderungen (Mutationen) aufgrund der chemischen Instabilität der Basen und der Exposition der Zellen mit mutagenen Einflüssen (Chemikalien, Strahlung) ausgesetzt. Die Änderung eines Codons in einem Gen durch Mutagenese führt nicht zwangsläufig zu einer Änderung des codierten Proteins, wenn eine Aminosäure durch mehrere, einander ähnliche Tripletts codiert ist.

Im genetischen Code ist auch festgelegt, wo ein translatierter Bereich eines Gens anfängt und wo er endet. Wir werden den Prozess der Translation in ► Kap. 3.2.2 besprechen. Wir wollen aber bereits jetzt zur Kenntnis nehmen, dass jede translatierte Region eines Gens mit



● **Abb. 3.1** Der universelle genetische Code: Dargestellt sind die 64 Codons und die jeweils codierte Aminosäure im Drei-Buchstaben-Code in Form einer „Code-Sonne“. Die Codons sind von innen (5') und außen (3') zu lesen; sie geben die Basensequenz der mRNA-Codons wieder, die für die außerhalb des Kreises stehende Aminosäure codieren. Schwarze Punkte bezeichnen die drei Translations-Stoppocodons. Das Codon AUG (schwarzes Quadrat) steht als Translationsstart sowie innerhalb von Gensequenzen für die Aminosäure Methionin.

dem Codon AUG beginnt. Da AUG für die Aminosäure Methionin steht, bedeutet das, dass jedes neu translatierte Protein mit der Aminosäure Methionin beginnt. Es gibt außerdem drei Codons, die bestimmen, wann die Translation eines Proteins beendet ist: die sog. Stoppocodons UAG, UGA und UAA (● Abb. 3.1).

Ein Gen – wenn wir es als Transkriptionseinheit verstehen – besitzt Kontrollelemente, die bestimmen, in welchen Zellen eines Organismus ein Gen wann und in welcher Stärke transkribiert werden soll. Diese Kontrollelemente sind die Promotoren. Sie befinden sich in der Regel direkt vor dem jeweils von ihnen kontrollierten Gen und legen fest, welcher DNA-Strang von der

### Zusammenfassung

- Proteine sind Makromoleküle, die aus einer linearen Abfolge von Aminosäuren bestehen. Die Abfolge der Aminosäuren in einem Protein ist in den Gensequenzen festgelegt.
- Die Natur benutzt 20 Aminosäuren, um daraus Proteine herzustellen. Der genetische Code besteht aus Dreierkombinationen von DNA-Basen, die bestimmen, welche Aminosäure an welcher Stelle in einem Protein eingebaut wird.
- Der genetische Code gilt auf der Erde universell, d. h. die genetische Information für ein bestimmtes Protein ist immer gleich, egal, in welchem Organismus sie unabhängig von ihrer eigentlichen Herkunft realisiert wird.
- Daraus folgt, dass für einen Organismus eigentlich fremde genetische Informationen zur Wirkung gebracht werden können, vorausgesetzt, dass die für verschiedene Organismen sehr spezifischen Kontrollelemente für die Transkription mit der fremden genetischen Information kombiniert werden.

RNA-Polymerase abgelesen werden soll. Am Ende eines Gens befindet sich ein Transkriptionsterminator, der dafür sorgt, dass die Transkription am Ende des proteincodierenden Teils aufhört.

Der Strang der DNA, der als Matrize für den Transkriptionsvorgang dient, ist der codogene Strang. Als Produkt der Transkription entsteht eine Boten-RNA (*messenger RNA* oder kurz *mRNA*), die komplementär zum codogenen Strang ist. In der Praxis ist es gebräuchlich, als Basensequenz eines Gens die Sequenz der mRNA zu schreiben, also den „oberen“ (eigentlich nichtcodogenen) Strang der DNA in 5' → 3'-Richtung. Dies ist insofern logisch, weil die Basensequenz der mRNA letztlich bestimmt, welche Aminosäuren während der Translation an welcher Stelle in ein wachsendes Protein eingebaut werden.

## 3.2 Genexpression in Prokaryonten

### 3.2.1 Transkription

Das Kopieren der in der DNA gespeicherten genetischen Information in eine Transportform ist notwendige Voraussetzung für die spätere Umsetzung der Information in die Synthese von Proteinen. Der Vorgang der Transkription bezeichnet das „Abschreiben“ der DNA-Information in eine RNA-Form. Transkription ist also die Synthese von RNA, das Produkt der Transkription ist komplementär zum abgeschriebenem codogenen DNA-Strang. Die Enzyme, die die Tran-

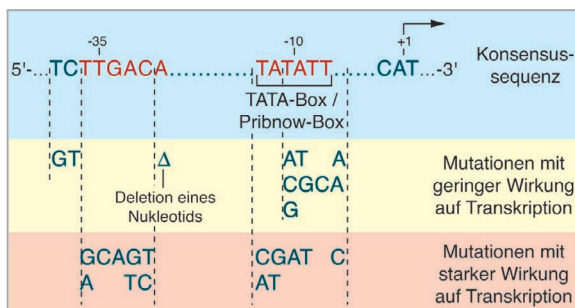
skription durchführen, sind DNA-abhängige RNA-Polymerasen. In Bakterien ist die für die Transkription verantwortliche RNA-Polymerase ein aus fünf Untereinheiten aufgebauter Multiproteinkomplex. Das Minimalenzym besteht aus vier Untereinheiten: zwei  $\alpha$ -, einer  $\beta$ - und einer  $\beta'$ - Untereinheit. Die  $\beta$ - Untereinheiten sind vor allem für die Bindung an die Matrizen-DNA zuständig, während die  $\alpha$ -Untereinheiten den Gesamtkomplex stabilisieren. Um jedoch bakterielle Promotoren zu erkennen und effizient zu transkribieren, benötigt das Enzym die fünfte Untereinheit Sigma ( $\sigma$ ). Bakterienzellen haben mehrere  $\sigma$ -Untereinheiten, die für jeweils verschiedene Sets von Genen benutzt werden.

Wie erkennt die RNA-Polymerase überhaupt einen Promotor und wie kann sie „entscheiden“, welcher der beiden DNA-Stränge transkribiert werden soll? Um diese Frage zu beantworten, müssen wir uns einen typischen bakteriellen Promotor anschauen (● Abb. 3.2).

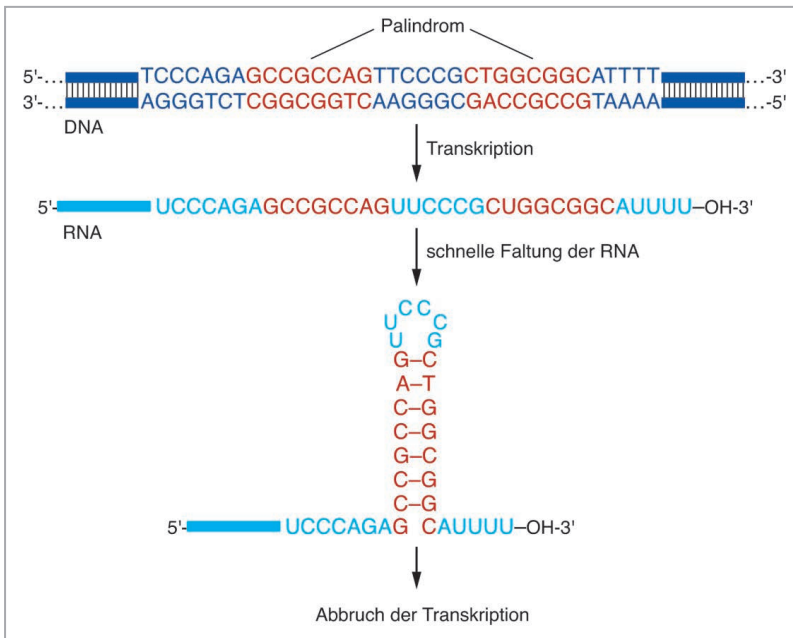
Das erste Nukleotid eines Gens, das in RNA überschrieben wird, befindet sich unterhalb („stromabwärts“ oder *downstream*) des Promotors. Es wird als Initiationsnukleotid +1 bezeichnet. Das oberhalb („stromaufwärts“ oder *upstream*) von diesem Initiationsnukleotid gelegene Nukleotid wird als Nukleotid -1 bezeichnet. Wichtig sind zwei DNA-Segmente, die gewöhnlich etwa 35 Basen beziehungsweise etwa 10 Basen vor der eigentlichen Startstelle der Transkription liegen. Diese beiden DNA-Segmente werden -35-Region bzw. als -10-Region genannt. Schreiben wir alle bekannten -35- und -10-Regionen untereinander, fällt auf, dass diese Regionen eine deutliche Sequenzähnlichkeit aufweisen. Aus dieser Sequenzähnlichkeit lässt sich eine Konsensussequenz formulieren. Die Konsensussequenz für die -35-Region lautet 5'-TTGACA-3', die Konsensussequenz für die -10-Region lautet 5'-TATATT-3'. Oft wird die -10-Region wegen ihrer A/T-reichen Sequenz auch TATA-Box oder nach ihrem Entdecker Pribnow-Box genannt.

Bindet die RNA-Polymerase an einen Promotor, entwindet sie die DNA in einem Bereich von 12 Basenpaaren um den Transkriptionsstartpunkt herum und beginnt mit der Synthese von RNA. Sobald die Transkription von dieser Initiationsphase in die Elongationsphase übergeht, fällt die  $\sigma$ -Untereinheit von der RNA-Polymerase ab und das Minimalenzym setzt die RNA-Synthese fort.

Veränderungen der Konsensussequenzen innerhalb der -35- und der -10-Region beeinträchtigen die Promotor-Stärke. Dabei wirken sich unterschiedliche Basensubstitutionen unterschiedlich stark auf die Funktion des Promotors aus (● Abb. 3.2). Letztlich hängt die Stärke eines Promotors, die als Menge an synthetisierter RNA pro Zeiteinheit definiert werden kann, davon ab, wie fest die RNA-Polymerase an den Promotor binden kann und wie schnell sie den Transkriptions-



● **Abb. 3.2** Konsenselemente für Promotoren in *E. coli*: Ein *E. coli*-Promotor befindet sich immer vor dem transkribierten Sequenzbereich. Das erste transkribierte Nukleotid wird mit +1 bezeichnet. Das Nukleotid vor der +1-Position ist Nukleotid -1. Eine signifikante Übereinstimmung (Konsensus) in allen Promotoren aus *E. coli* finden wir um die Positionen -10 und -35 (rot-markierte Nucleotide, Konsensussequenz). Den Sequenzblock um die Position -10 bezeichnet man als TATA-Box oder Pribnow-Box. Verändert man die Nucleotide innerhalb der Promotorregion, beobachtet man entweder geringe (Mutationen im gelb-unterlegten Bereich) oder aber starke Auswirkungen auf die Transkription (rosa unterlegter Bereich). Wir erkennen daran, dass die Konsensussequenz zwar Variationen zulässt, dass aber die Variationsmöglichkeiten sehr stark eingeschränkt sind.



• **Abb. 3.3** Die Transkriptionstermination wird in *E. coli* über ein RNA-Signal reguliert. Bakterielle Gene besitzen an ihrem Ende eine Sequenz, die nach Transkription – also auf der Ebene der RNA – eine Stamm-/Loop-Struktur ausbilden kann, da die zu der Loop-Region links und rechts benachbarten Nucleotide palindromartig aufgebaut sind. Unmittelbar auf diese Stamm-/Loop-Struktur folgen eine Reihe von Uridin-Nucleotiden. Wird eine solche Struktur von einer bakteriellen RNA-Polymerase synthetisiert, erkennt das Enzym dies als Ende des Gens und fällt von der DNA ab. Wir erkennen, dass das Terminationssignal nicht wie ein Promotor von einer Konsensussequenz, sondern von einer Konsensusstruktur gebildet wird.

3

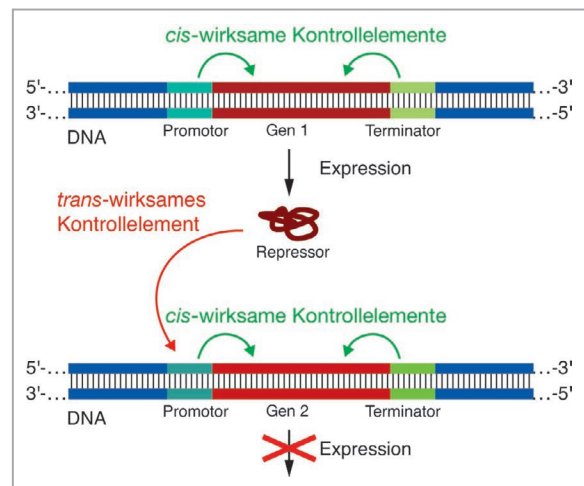
start wieder freigibt, damit neue RNA-Polymerase-Moleküle binden und die nächste Transkriptionsrunde einleiten können. Typischerweise sind immer mehrere RNA-Polymerasekomplexe gleichzeitig an demselben Promotor aktiv.

Die Transkription verläuft immer von Nucleotid +1 bis zu einer Stelle, die von der RNA-Polymerase als Transkriptionsterminator erkannt wird. An dieser Stelle fällt die RNA-Polymerase von der Matrizen-DNA ab und steht dann für eine neue Transkriptions-Runde zur Verfügung. Als Transkriptionsterminator wird in *E. coli* eine Sequenz erkannt, die aus einer Palindromsequenz aus G/C-Basen, gefolgt von mehreren U-Nucleotiden, besteht (• Abb. 3.3). Wir erkennen an den U-Nucleotiden, dass dieses Signal eigentlich auf der synthetisierten RNA liegt und nicht etwa auf der DNA. Der doppelsträngige RNA-Bereich bewirkt an der RNA-Polymerase eine Konformationsänderung, die zum Zerfallen des Komplexes aus der RNA-Polymerase, der Matrizen-DNA und der gebildeten RNA führt.

Ein Promotor ist ein typisches Beispiel für ein *cis*-wirksames Kontrollelement (• Abb. 3.4). Das bedeutet, dass dieses Element eine ganz bestimmte Position relativ zu dem Bereich einnehmen muss, den es kontrolliert. Promotoren müssen sich nämlich immer auf dem gleichen DNA-Molekül befinden wie das Gen, dessen Transkription sie kontrollieren.

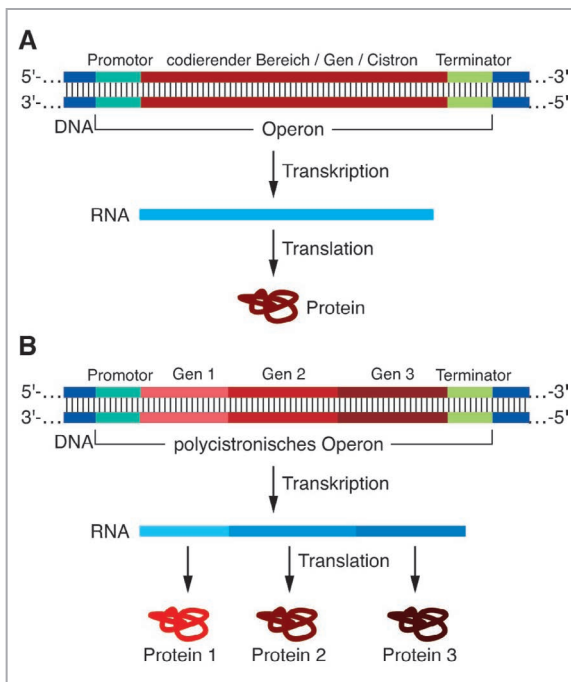
Es gibt aber auch Elemente, die die Expression eines Gens als in *trans* wirksame Kontrollelemente – sozusagen von außen her – beeinflussen (• Abb. 3.4). Ein solches Element ist beispielsweise ein Repressor, der auf einem ganz anderen DNA-Molekül codiert sein kann. Ein typischer Repressor entfaltet seine Wirkung als Pro-

tein, das an die Transkriptionseinheit bindet und mit ihr so wechselwirkt, dass das entsprechende Gen nicht mehr exprimiert werden kann.



• **Abb. 3.4** Kontrollelemente für die Transkription: In *cis* wirksame Kontrollelemente befinden sich immer auf dem gleichen Molekül und in relativ enger Nachbarschaft zu dem Element, das sie kontrollieren. Promotoren und Transkriptionsterminatoren sind typische Kontrollelemente, die in *cis* auf ein Gen wirken. Dagegen sind in *trans* wirksame Kontrollelemente eigenständige Moleküle, die durch Bindung an eine Zielsequenz eine kontrollierende Funktion ausüben. Beispielsweise ist ein Repressor-Protein ein typisches in *trans* aktives Kontrollelement. Dieses Repressorprotein wird irgendwo im Genom – bei Eukaryonten eventuell auf einem anderen Chromosom – codiert. Der Repressor bindet dann an eine Zielsequenz in der Nähe des Promotors eines bestimmten Gens und verhindert so die Expression dieses Gens.

Den Bereich, der einen Promotor, die codierende Region und einen Transkriptionsterminator beinhaltet, bezeichnet man als Operon (● Abb. 3.5). Dieser Bereich stellt gewissermaßen eine genetische Arbeitseinheit dar, die sicherstellt, dass informative DNA in RNA übersetzt werden kann. In Prokaryonten werden häufig mehrere proteincodierende Bereiche von nur einem Promotor und einem Transkriptionsterminator kontrolliert. Ein derartiges Operon wird polycistronisches Operon und die entsprechend große mRNA polycistronische mRNA genannt. Der Hintergrund für diese Bezeichnung ist, dass die funktionelle Informationseinheit des Gens auch mit dem in der klassischen Genetik gebräuchlichen Begriff Cistron bezeichnet werden kann. Zu beachten ist, dass in Bakterien von einer polycistronischen mRNA mehrere unterschiedliche Proteine translatiert werden. In Eukaryonten kommen polycistronische Gene selten vor – in diesen Zellen codiert in aller Regel eine mRNA für ein Protein.



● **Abb. 3.5** Aufbau eines Operons: Ein Operon ist der transkribierte Bereich, der von einem Promotor und einem Transkriptionsterminator kontrolliert wird.

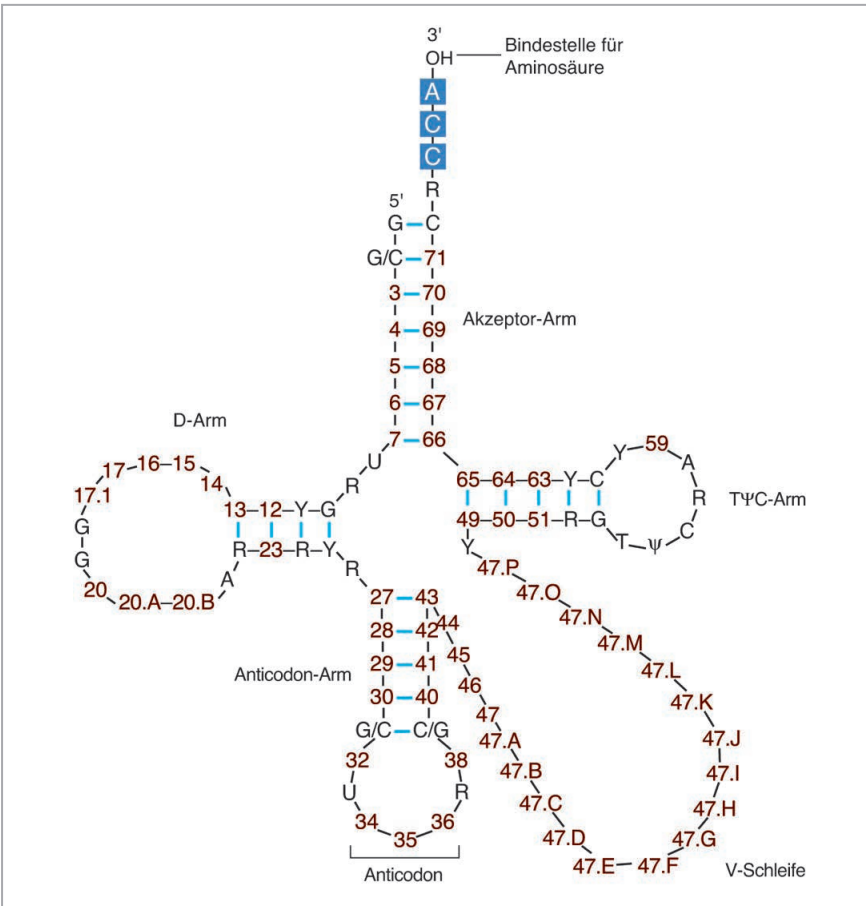
(A) Befindet sich in diesem Bereich ein einzelnes Gen (= Cistron), so handelt es sich um ein monocistronisches Operon. (B) Befinden sich zwischen den Transkriptionskontrollelementen mehrere Gene (verschiedene Rottöne), bezeichnen wir dies als ein polycistronisches Operon. Die durch Transkription von diesem Operon gebildete polycistronische mRNA besitzt für jedes Gen einen eigenen Translationsstart. Daher wird von dieser mRNA nicht etwa ein großes Vorläuferprotein gebildet, das dann in Einzelproteine zerschnitten wird, sondern es werden von der polycistronischen mRNA simultan alle codierten Proteine einzeln gebildet.

### 3.2.2 Translation

Bei der Transkription eines Gens entsteht eine mRNA, die dem Proteinsynthese-Apparat während des Vorgangs der Translation als Information für die zu bildenden Proteine zur Verfügung steht. Während der Proteinsynthese werden die Reihenfolge und die Identität der jeweils zu verknüpfenden Aminosäuren durch Adaptermoleküle festgelegt, die als Transfer-RNAs, oder kurz tRNAs, bezeichnet werden. tRNAs haben zwei für diese Aufgabe nötige Eigenschaften (● Abb. 3.6): Jede tRNA besitzt eine exponierte Sequenz aus drei Nukleotiden, die komplementär zu einem Codon der mRNA ist. Diese auf tRNAs vorkommende Sequenz wird als Anticodon bezeichnet. Die Basenpaarung eines Anticodons der tRNA mit einem Codon der mRNA am Proteinsynthese-Apparat bestimmt also, welche Aminosäure in eine wachsende Polypeptidkette eingebaut wird.

Durch genaues Studium der Strukturen verschiedener tRNAs wurden gemeinsame Eigenschaften entdeckt. So sind alle tRNAs aus 74 – 94 Basen aufgebaut. Typischerweise sind tRNAs keine reinen einzelsträngigen Moleküle. Vielmehr werden innerhalb eines tRNA-Moleküls bestimmte Basenpaarungen ausgebildet, die dem tRNA-Molekül in der zweidimensionalen Darstellung die Form eines Kleeblattes geben (● Abb. 3.6). tRNAs haben noch andere gemeinsame Eigenschaften:

- Das 5'- und das 3'-Ende bilden über eine Länge von sieben Basen einen doppelsträngigen Bereich aus, der Akzeptor-Arm genannt wird. Die 3'-Enden von tRNAs bestehen charakteristischerweise aus der Basensequenz CCA und ragen über das 5'-Ende der tRNA hinaus. Auf das Adenosin des CCA-Endes der tRNA wird eine Aminosäure übertragen, indem die Carboxylgruppe der Aminosäure mit der freien 3'-Hydroxylgruppe des Adenosins verestert wird.
- Das Anticodon liegt im Zentrum einer sieben Basen langen Schleife auf der dem Akzeptor-Arm entgegengesetzten Seite der Kleeblattstruktur, die als Anticodon-Arm bezeichnet wird.
- Etwa 10 % der Nukleotide in tRNAs sind posttranskriptionell modifiziert und entsprechen dadurch nicht exakt den vier in RNAs normalerweise vorkommenden Standardnukleotiden. Es sind mehr als 30 solche Modifikationen in tRNAs gefunden worden, aber die Funktionen dieser Modifikationen sind weitgehend unklar.
- Der in der Kleeblattstruktur links gelegene Teil der tRNA wird als D-Arm bezeichnet, weil er häufig die Base Dihydrouracil enthält.
- Der rechts gelegene Arm enthält das ungewöhnliche Nukleosid Pseudouridin (Ψ) in der Basenfolge Thyminid-Pseudouridin-Cytidin (TΨC) und wird deshalb als TΨC-Arm bezeichnet.



● **Abb. 3.6** Allgemeine Struktur einer tRNA: tRNAs können unterschiedlich lang sein, lassen sich aber in einer allgemeinen Kleeblattstruktur darstellen. Die Nummerierung der Basen beginnt am 5'-Ende der tRNAs. Der D-Arm kann unterschiedlich lang sein, sodass das auf Nucleotid 20 folgende Nucleotid mit 20.A und ggf. 20.B beziffert wird. Das gilt auch für die variable (V-)Schleife, in der die nach Nucleotid 47 folgenden Nucleotide als 47.A, 47.B, etc. gezählt werden. Neben dem Vorkommen der normalen Nucleotide A, C, G und U unterscheiden sich tRNAs in zahlreichen modifizierten Basen, die an dieser Stelle nicht im Detail diskutiert werden können. Die Funktion dieser Modifikationen ist weitgehend unbekannt (R = Purin, Y = Pyrimidin).

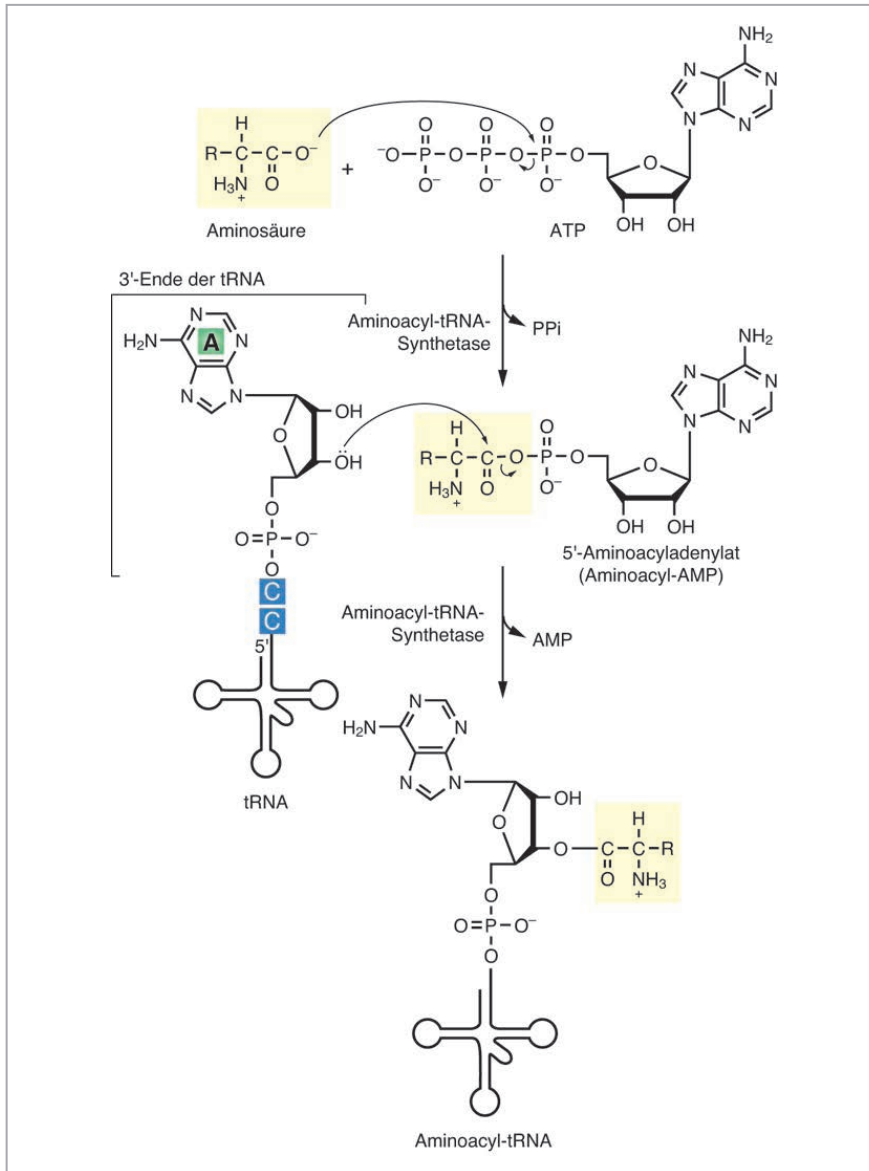
- Zwischen dem Anticodon-Arm und dem TΨC-Arm liegt eine weitere Schleife, deren Länge variabel ist und deshalb als V-Schleife bezeichnet wird.

Die Beladung der tRNAs mit Aminosäuren wird durch spezifische Enzyme katalysiert, die als Aminoacyl-tRNA-Synthetasen bezeichnet werden. In jeder Zelle gibt es mindestens 20 verschiedene Aminoacyl-tRNA-Synthetasen – eine für jede in Proteinen vorkommende Aminosäure. Die Beladung einer tRNA mit einer Aminosäure verläuft in zwei Schritten (● Abb. 3.7): Zunächst wird die Carboxylgruppe der Aminosäure biochemisch aktiviert, indem sie auf den AMP-Teil des Co-Substrats ATP übertragen wird. Danach wird das entstandene Aminoacyladenylat unter Abspaltung von AMP auf die 3'-Hydroxylgruppe der tRNA übertragen. Obwohl alle Aminoacyl-tRNA-Synthetasen die gleiche Reaktion katalysieren, unterscheiden sie sich stark in ihren Größen und Strukturen. Wichtig ist jedoch, dass jede Aminoacyl-tRNA-Synthetase sehr spezifisch die jeweils zusammengehörenden tRNAs und Aminosäuren bindet. Auf diese Weise sorgt also eine Aminoacyl-tRNA-Synthetase dafür, dass das jeweils richtige Anticodon einer tRNA mit der dazu gehörenden Aminosäure verbunden wird.

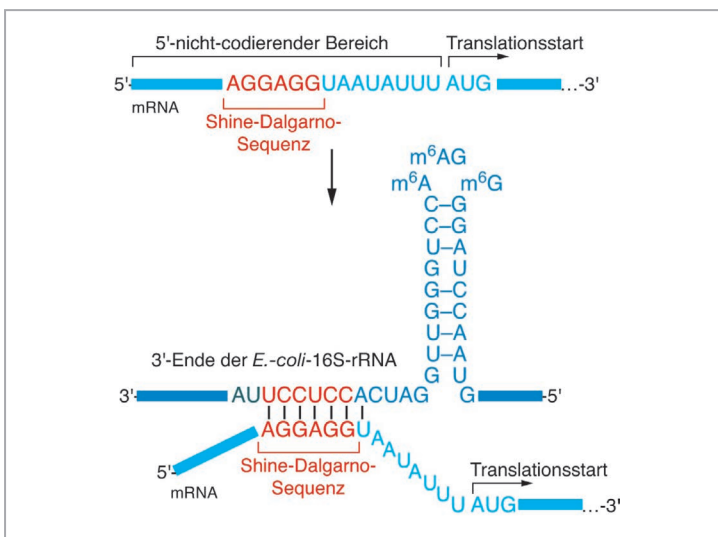
Wie wir gesehen haben, besteht die Adapterfunktion der tRNAs in der Heranführung der jeweils während der Proteinsynthese benötigten Aminosäure an den Proteinsynthese-Apparat. Danach ist die Proteinsynthese, also die Translation des Nucleinsäure-Codes in eine Abfolge von Aminosäuren im resultierenden Protein, durch zwei Schritte geprägt:

- die spezifische Basenpaarung von Codons auf der mRNA und dazu komplementären Anticodons auf den tRNAs und
- die Ausbildung von Peptidbindungen zwischen den Aminosäuren, die durch die Bindung von Aminoacyl-tRNAs in direkte Nachbarschaft zueinander gebracht wurden.

Den Proteinsynthese-Apparat einer Zelle bilden die Ribosomen, die ihrerseits aus ribosomalen RNAs (rRNAs) und Proteinen aufgebaut sind. Im funktionellen Zustand bestehen sie aus einer großen und einer kleinen Untereinheit, deren Größe – wie auch die Größe der kompletten Ribosomen – mit dem sog. Svedberg-Koeffizienten (S) angegeben wird. Der Svedberg-Koeffizient ist eine experimentell ermittelte Größe und beschreibt die Sedimentationsgeschwindigkeit sphärischer Partikel unter geometrisch vorgegebenen Bedin-



● **Abb. 3.7** Beladung von tRNAs mit Aminosäuren: Aminosäuren werden von hochspezifischen Aminoacyl-tRNA-Synthetasen auf tRNAs übertragen. Dazu wird die Carboxylfunktion der Aminosäure durch Übertragung eines AMP-Restes aus ATP unter Abspaltung von Pyrophosphat ( $\text{PP}_i$ ) biochemisch aktiviert. Danach erfolgt die Übertragung der Aminosäure auf das freie 3'-OH-Ende des Adennukleotids der tRNA. tRNAs enden am 3'-Ende immer mit der Basenfolge CCA-OH.

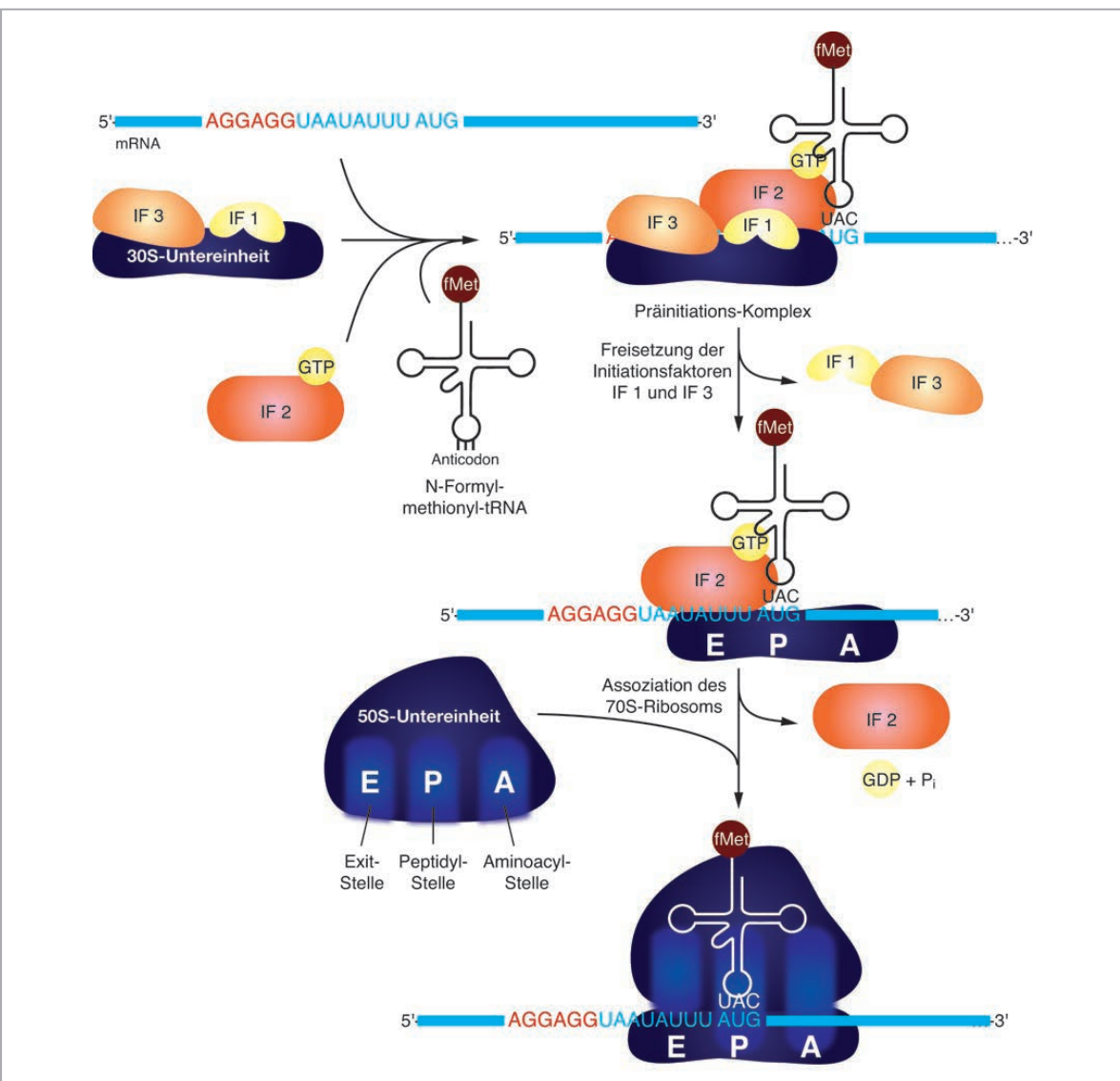


● **Abb. 3.8** Shine-Dalgarno-Sequenz: Die Shine-Dalgarno-Sequenz ist ein in *cis* wirkendes Translations-Kontrollelement, das sich nicht nur durch eine relativ konservierte Konsensussequenz auszeichnet, sondern das auch sehr exakt vor der Translations-Initiationsstelle positioniert sein muss. Die hohen Sequenz- und Positionsanforderungen ergeben sich aus der Komplementarität der Shine-Dalgarno-Sequenz mit dem 3'-Ende der ribosomalen 16S-rRNA, über die die mRNA für die Translations-Initiation korrekt am Ribosom positioniert wird.

gungen in einem Zentrifugalfeld. Wir können den Svedberg-Koeffizienten als Maß für die Größe eines Partikels interpretieren und verstehen aufgrund des experimentellen Charakters dieses Koeffizienten, warum die Größe des Gesamtribosoms von Bakterien mit 70 S angegeben wird und nicht der mathematischen Summe der großen (50 S) und der kleinen (30 S) Untereinheit entspricht. Jede Ribosomen-Untereinheit ist komplex aufgebaut. Bei Bakterien besteht die kleine Untereinheit aus der 16-S-rRNA und 21 verschiedenen Proteinen. Die große Untereinheit enthält die 5-S- und 23-S-rRNAs sowie 36 verschiedene Proteine.

Eine mRNA ist typischerweise länger als die in ihrer Codierung des zugehörigen Proteins benötigte Codon-

folge. Wir sprechen in diesem Zusammenhang von einer 5'- bzw. 3'-nichttranslatierten Region einer mRNA. Woher „weiß“ aber dann ein Ribosom, an welcher Stelle der mRNA die Translation beginnen soll? In Bakterien ist für eine effiziente Initiation der Translation ein Kontrollelement notwendig, das 1974 von den australischen Wissenschaftlern John Shine und Lynn Dalgarno entdeckt wurde. Die nach ihnen benannte Shine-Dalgarno-Sequenz kann man als Bindungsstelle für das Ribosom an die mRNA auffassen. Wichtig ist nicht nur, dass die Sequenz möglichst exakt mit der Konsensussequenz 5'-AGGAGG-3' übereinstimmt, sondern dass sich dieses Element auch in einer optimalen Position zum Translationsstart befinden muss



• **Abb. 3.9** Die Translationsinitiation: An der Ausbildung eines Präinitiationskomplexes für eine Translation beteiligen sich zunächst (i) die mRNA mit ihrer Shine-Dalgarno-Sequenz (rot), (ii) die fMet-tRNA, die über dem ersten AUG-Codon der mRNA positioniert wird, und (iii) die ribosomale 30 S-Untereinheit, an der die drei Initiationfaktoren IF1, IF2 und IF3 sowie ein Molekül GTP binden. Bevor sich die große ribosomale 50 S-Untereinheit anlagern kann, müssen die Initiationfaktoren aus dem Komplex dissoziieren.

(Abb. 3.8). Das hat den Grund, dass die Shine-Dalgarno-Sequenz komplementär zu einem bestimmten Sequenzbereich in der 16-S-rRNA ist. Über diese Komplementarität kann also die kleine ribosomale Untereinheit mit der mRNA hybridisieren und das Ribosom wird gewissermaßen auf die mRNA aufgefädelt.

Von seltenen Ausnahmen abgesehen bildet das erste Methionin-Codon (AUG) nach der Shine-Dalgarno-Sequenz das Startcodon für die Translation (Abb. 3.8). AUG kann als universelles Startcodon angesehen wer-

den, obwohl in Bakterien auch einige Proteine von alternativen Startcodons ausgehend (GUG oder UUG) gebildet werden. Da das Codon AUG für die Aminosäure Methionin codiert, können wir uns als Faustregel merken, dass alle Proteine mit einem Methionin beginnen. In Bakterien gibt es einen Unterschied zwischen dem als Start-Aminosäure verwendeten Methionin und anderen Methioninen, die intern vorkommen: Bakterien haben zwei verschiedene methioninspezifische tRNAs, die als  $tRNA_f^{Met}$  und  $tRNA_m^{Met}$  bezeichnet wer-

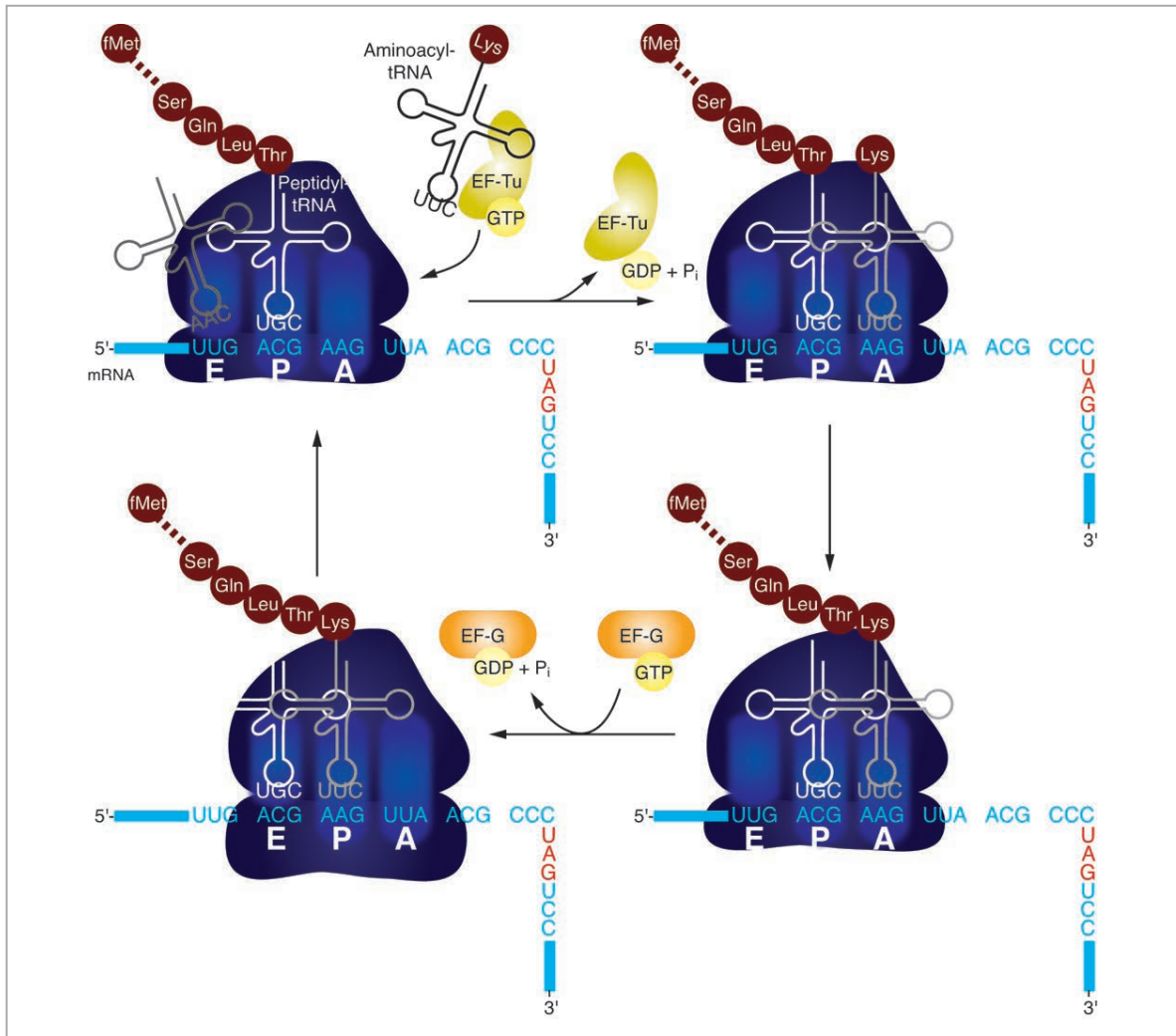


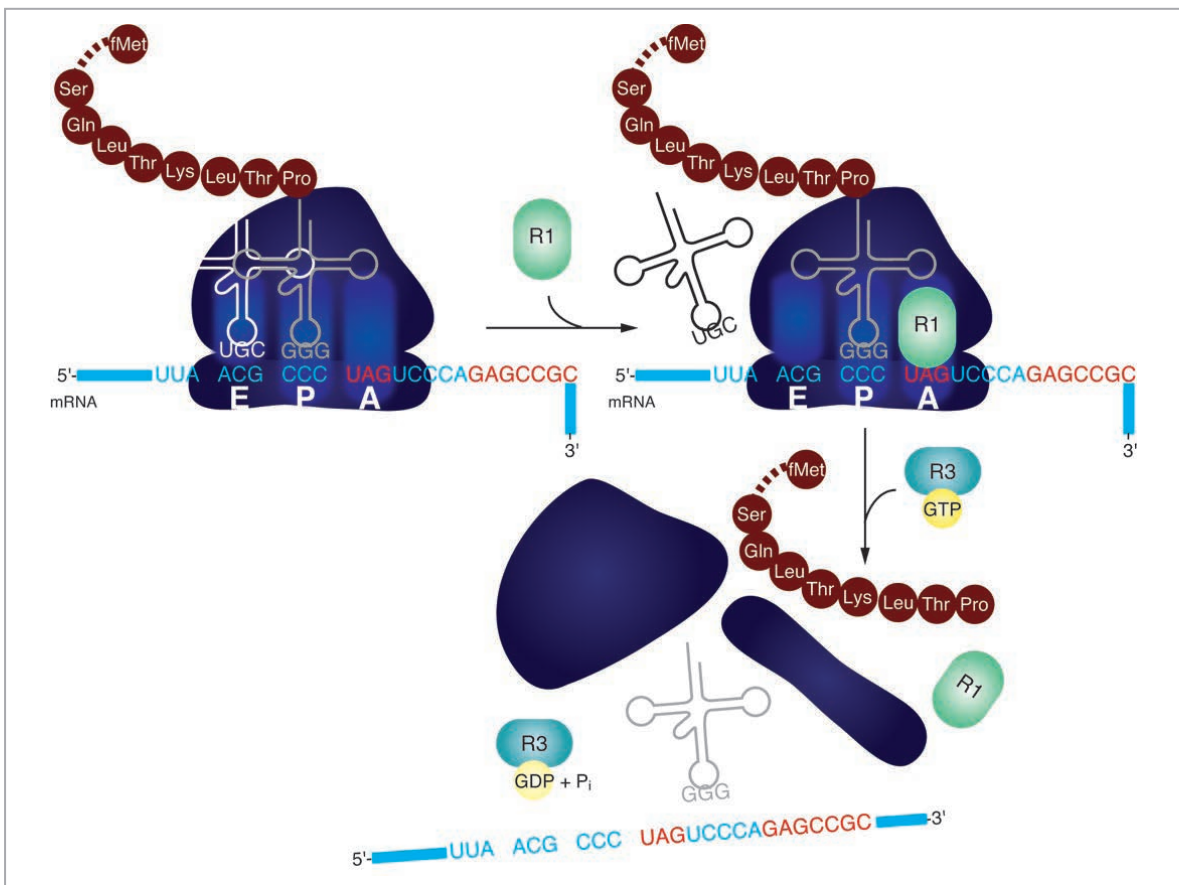
Abb. 3.10 Translationselongation: Das Ribosom bietet Platz für zwei tRNAs. Die linke Hälfte des Ribosoms bezeichnen wir als P-Stelle, da an dieser Position nur sog. Peptidyl-tRNAs binden können, d. h. tRNAs, die mit einer Peptidkette beladen sind. Auch die fMet-tRNA kann an die P-Stelle binden, da ihre  $\alpha$ -Aminogruppe nicht frei vorliegt, sondern formyliert ist. Die zweite tRNA-Bindestelle am Ribosom wird als A-Stelle bezeichnet, da hier die Aminoacyl-tRNAs binden. Hierzu bildet sich zunächst ein ternärer Komplex aus der mit einer Aminosäure beladenen Aminoacyl-tRNA, GTP und dem Elongationsfaktor EF-Tu. Bei der Positionierung der Aminoacyl-tRNA in die A-Stelle wird GTP zu GDP hydrolysiert und EF-Tu sowie GDP dissoziieren als binärer Komplex ab. Es folgt die Verlängerung der Peptidkette. Hierzu wird die Peptidkette von der Peptidyl-tRNA in der P-Stelle getrennt und komplett an die freie  $\alpha$ -Aminogruppe der Aminosäure in der A-Stelle geheftet. Die tRNA in der P-Stelle erfüllt nun nicht mehr die Anforderungen, die an eine tRNA in dieser Position gestellt werden (siehe oben) und dissoziiert deshalb über die Exit-(E)-Stelle ab. Mithilfe eines weiteren Elongationsfaktors (EF-G) und unter GTP-Hydrolyse wird nun die mit der Peptidkette beladene tRNA aus der A-Stelle in die P-Stelle transloziert. Damit ist die A-Stelle für eine neue Aminoacyl-tRNA frei und der Zyklus kann erneut beginnen.

den. Beide tRNAs werden durch dieselbe Methionyl-tRNA-Synthetase mit Methionin beladen, jedoch wird die Aminogruppe des Methionins in der tRNA<sup>fMet</sup> nachträglich mit einem Formylrest modifiziert. Diese auch als fMet-tRNA bezeichnete tRNA dient als Startermolekül für die Proteinsynthese. Der Formylrest bildet dabei quasi eine Schutzgruppe für die Aminogruppe der ersten Aminosäure Methionin, sodass die Richtung der anschließenden Peptidsynthese durch die freie Carboxylgruppe des Methionins vorgegeben ist.

Am Startcodon der mRNA formiert sich der Translations-Präinitiationskomplex (Abb. 3.9). Dazu wird der Initiationsfaktor 2 (IF2) durch Bindung von GTP aktiviert und bindet gemeinsam mit der fMet-tRNA an die 30-S-Untereinheit des Ribosoms, die zuvor bereits mit IF1 und IF3 besetzt wurde. IF1 hat unter anderem die Aufgabe, eine vorzeitige Anlagerung der großen ribosomalen Untereinheit zu unterdrücken. IF3 fördert die spezifische Anlagerung der fMet-tRNA an der ribosomalen P-Stelle und bleibt so lange gebunden, bis die fMet-tRNA durch Basenpaarung mit dem AUG-Startcodon stabil genug am Ribosom fixiert wurde. Anschlie-

ßend verlassen IF1 und IF3 den Präinitiationskomplex (Abb. 3.9). In der P-Stelle befindet sich später jeweils die an eine tRNA gebundene, wachsende Polypeptidkette, also zu Beginn der Translation nur die erste Aminosäure des gebildeten Proteins (fMet). Die korrekte Positionierung der fMet-tRNA in der P-Stelle wird von IF2 unterstützt und benötigt dazu die Energie aus der Hydrolyse von GTP. Sobald die fMet-tRNA korrekt positioniert wurde, verlässt IF2 (GDP) den Komplex und gleichzeitig wird die 50-S-Untereinheit des Ribosoms angelagert (Abb. 3.9). Der fertig gestellte 70-S-Initiationskomplex kann nun in die Translations-Elongations-Phase eintreten.

Bei der Elongation, also der Bildung einer Polypeptidkette aus Aminosäuren, bindet die jeweils zum nächsten Codon passende, beladene tRNA (allgemein auch als Aminoacyl-tRNA bezeichnet) mithilfe des Elongationsfaktors EF-Tu unter GTP-Hydrolyse an die ribosomale A-Stelle (Aminoacyl-tRNA-Bindestelle; Abb. 3.10). Durch Transacylierung wird die Aminosäure Methionin von der fMet-tRNA abgespalten und mit der  $\alpha$ -Aminogruppe der benachbarten Aminosäure



• **Abb. 3.11** Translationstermination: Folgt während der Translationselongation (Abb. 3.10) als nächstes Codon auf der mRNA ein Translations-Stoppcodon, wird der Terminationsprozess in Gang gesetzt. Mithilfe von Terminationsfaktoren (RF = *release factors*) und GTP dissoziieren alle Komponenten des Translationskomplexes und die Peptidkette wird von der letzten Peptidyl-tRNA abgespalten. Das Produkt der Translation ist ein freies Polypeptid (Protein).

verbunden. Dieser Vorgang wird als Peptidyltransferase-Reaktion bezeichnet. Nun bewegt sich das Ribosom um ein Codon auf der mRNA weiter. Für diese als Translokation bezeichnete Reaktion muss das GTP im EF-G (GTP) in GDP und Phosphat gespalten werden. Die deacylierte fMet-tRNA gelangt von der P-Stelle in die E-Stelle (*Exit*-Stelle) und verlässt das Ribosom, während sich das tRNA-gebundene Dipeptid in der P-Stelle befindet. Die A-Stelle wird dadurch für die Bindung der nächsten Aminoacyl-tRNA frei.

Dieser Prozess wiederholt sich, bis auf der mRNA eines der drei Basentriplets UGA, UAG oder UAA an

die ribosomale A-Stelle rückt. Für diese drei Codons existieren in gewöhnlichen Zellen keine spezifischen tRNAs. Daher fungieren diese drei Basentriplets als Translationsterminations-Signale, d. h. an diesen Codons bricht die Translation ab und die Ribosomen dissoziieren von der mRNA. Katalysiert wird die Termination durch die Terminationsfaktoren RF1 und RF2, die spezifisch die Codons UAA und UAG bzw. UAA und UGA erkennen. Unter dem Einfluss eines dritten Faktors, RF3 (GTP), wird unter Spaltung von GTP der gesamte Translationskomplex in seine einzelnen Bestandteile zerlegt (● Abb. 3.11).

### Zusammenfassung

- Der Fluss der genetischen Information geht von dem Speicher der Information (DNA) über eine RNA-Transportform (mRNA) zu den Proteinen.
- Zur Umsetzung der genetischen Information wird die DNA zunächst in eine Transportform umgeschrieben (transkribiert), die aus RNA besteht. Zuständig hierfür sind DNA-abhängige RNA-Polymerasen.
- Die Transkription wird an besonderen DNA-Elementen (Promotoren) reguliert, die als Bindestellen für Transkriptionsfaktoren oder die RNA-Polymerase selbst dienen und vor den Genen positioniert sind. Eine Transkriptionseinheit besteht aus einem Promotor, dem exprimierten Gen und einem Transkriptionsterminator.
- Die Translation ist die Übersetzung des Nukleinsäure-Codes in eine Sequenz aus Aminosäuren, den Proteinen. Bei der Translation nehmen tRNAs die Funktion eines Adapters zwischen der „Nukleinsäuresprache“ und der „Proteinsprache“ ein.
- tRNAs werden über hochspezifische Enzyme an ihren 3'-Enden mit einer bestimmten Aminosäure beladen. Die beladenen tRNAs binden über eine Sequenz von drei Basen, dem Anticodon, an eine komplementäre Sequenz an der mRNA (dem Codon).
- Auf diese Weise bringen sie die Aminosäuren zum Ort der Proteinsynthese, den Ribosomen, und die Codon-Anticodon-Basenpaarung definiert, welche Aminosäure in eine wachsende Polypeptidkette eingebaut wird.

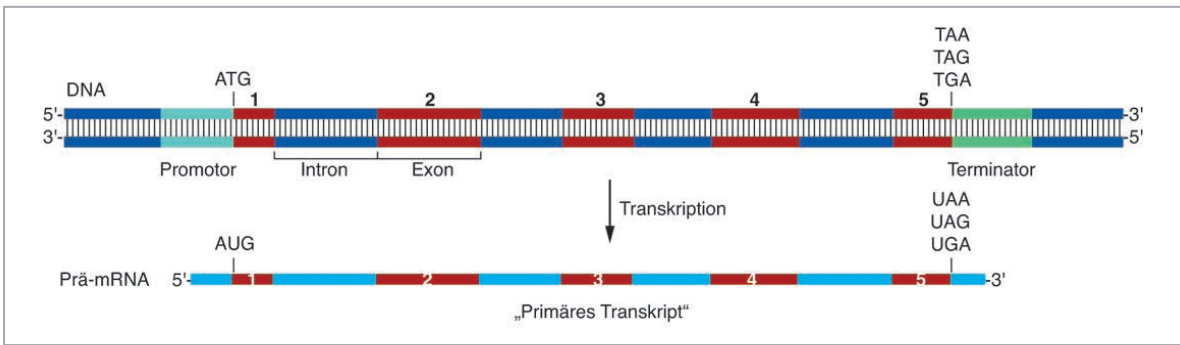
### 3.3 Genexpression in Eukaryonten

Wenn wir als Definition für den Begriff Gen die Informationseinheit zur Bildung eines Proteins nehmen, dann trifft diese Aussage für Bakterien im Wesentlichen zu, ist jedoch für Eukaryonten nicht vollständig. In Bakterien ist der proteincodierende Teil in einer einzigen, ununterbrochenen Folge von Codons festgelegt, die mit dem Startcodon beginnt und mit einem Stoppcodon endet. In Eukaryonten ist die Folge der Codons zum Teil durch nichtcodierende Bereiche von DNA unterbrochen (● Abb. 3.12). Die proteincodierenden Bereiche sind die **Exons** (*expressed regions*), während die eingelagerten, nichtcodierenden Bereiche als **Introns** (*intragenic regions*) bezeichnet werden. Eukaryontische Gene sind also in der Regel aus Exons und Introns zusammengesetzte Gene und werden daher häufig auch als Mosaikgene bezeichnet.

Das primäre RNA-Transkript (Prä-mRNA), das von einem Mosaikgen hergestellt wurde, enthält sowohl die Exons als auch die Introns. Da nur die Exons die Informationen des in dem jeweiligen Gen festgelegten Prote-

ins beinhalten, müssen die Introns während des Prozesses der Genexpression entfernt und die Exons basengenau zu einer Einheit (einen offenen Leserahmen) zusammengefügt werden. Demnach ist die nach der Transkription eines Mosaikgens gebildete RNA nicht sofort für die Translation bereit. Das Entfernen der Introns ist Teil einer mehrstufigen Prozessierung des primären Transkripts. Diese Prozessierung läuft in allen eukaryontischen Zellen prinzipiell gleich ab und hängt nicht davon ab, ob ein Gen Introns enthält oder nicht.

Zu Beginn der Prozessierung eukaryontischer Prä-mRNAs wird ein 7-Methylguanin-Nukleotid ( $m^7G$ ) als Kappe an das 5'-Ende der RNA gehängt (● Abb. 3.13). Die  $m^7G$ -Kappe wird während der laufenden Transkription an das 5'-Ende der noch wachsenden mRNA geheftet. Im Zellkern bindet ein Proteinkomplex (*cap-binding complex*) an die  $m^7G$ -Kappe und organisiert das nachfolgende Spleißen, die 3'-Prozessierung sowie den Export der Prä-mRNA ins Zytoplasma durch Rekrutierung entsprechender anderer Proteine wie beispielsweise des Spleißosoms. Die Transkription stoppt in der Nachbarschaft der Konsensussequenz 5'-AAUAAA-3',

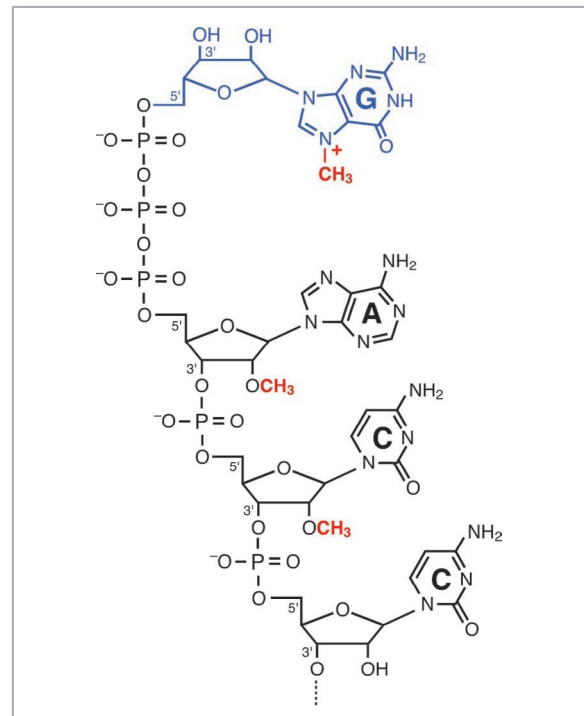


• **Abb. 3.12** Genstruktur bei Eukaryoten: Eukaryontische Transkriptionseinheiten bestehen aus einem Promotor, dem Gen und einem Transkriptionsterminator. Gene sind typischerweise in proteincodierende Exons (rot) und nichtcodierende Introns (blau) unterteilt. Nach der Transkription sind die Introns (hellblau) in der Prä-mRNA weiterhin vorhanden, d. h. diese RNA muss zunächst prozessiert werden, bevor die Translation stattfinden kann.

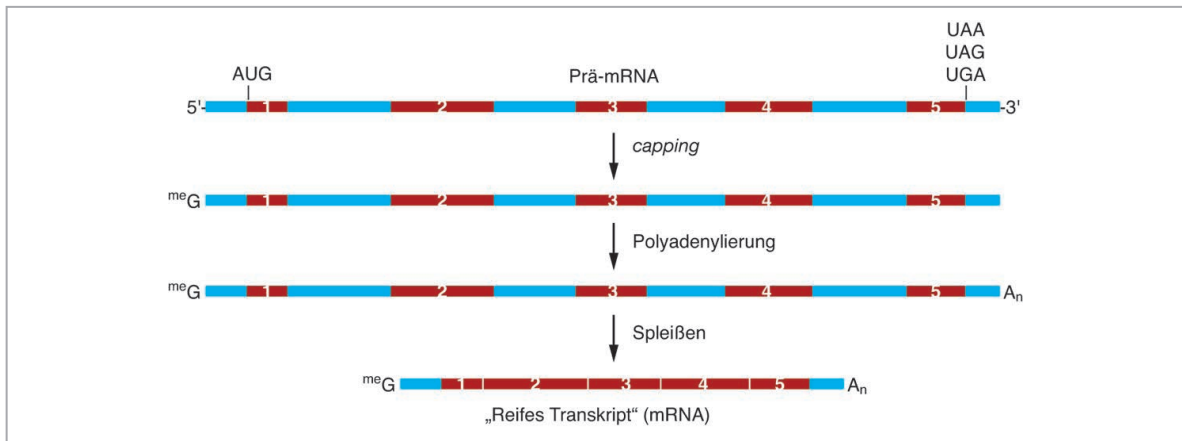
die als Erkennungsstelle für die sog. Polyadenylierung der mRNA dient. Polyadenylierung bedeutet, dass eukaryontische mRNAs an ihren 3'-Enden matrixunabhängig mit einem Poly(A)-Schwanz aus 150 – 200 Adenosin-Nukleotiden verlängert werden (• Abb. 3.14). Das Poly(A)-Ende ist für den Export der im Zellkern hergestellten und prozessierten mRNA sowie für ihre Stabilität von Bedeutung. Schließlich werden in einem letzten Prozessierungsschritt, dem Spleißen (*splicing*), die Introns aus der primären mRNA entfernt (• Abb. 3.14). Dabei werden die Introns entfernt und die Exons basengenau miteinander verbunden, sodass ein durchgehender Leserahmen entsteht, der die gesamte Aminosäuresequenz eines Proteins umfasst. Typischerweise können aus eukaryontischen Mosaikgenen durch alternatives Spleißen aus einer Prä-mRNA mehrere verschiedene mRNAs und somit verschiedene Proteine hergestellt werden (• Abb. 3.15). Im Zytoplasma rekrutiert der Cap-binding Complex den Translations-Initiationsfaktor eIF4G und die Helikase eIF4A zum 5'-Ende der mRNA. eIF4G rekrutiert das Poly(A)-bindende Protein (*poly(A)-binding protein*, PABP1), sodass ein Proteinkomplex mit einer pseudozirkulär gebundenen mRNA entsteht. Das stellt sicher, dass nur vollständig prozessierte mRNA translatiert wird. Zusammen mit anderen Faktoren bindet die mRNA schließlich an das Ribosom zur Initiation der Translation.

Promotoren in Eukaryoten unterscheiden sich grundlegend von Promotoren in Prokaryoten. Das liegt in erster Linie daran, dass eukaryontische RNA-Polymerasen nicht wie prokaryontische RNA-Polymerasen selbst den Promotor erkennen. Die Erkennung wird vielmehr durch einen sog. Präinitiationskomplex vermittelt. Das ist ein als TFIID bezeichneter Multiproteinkomplex, der aus verschiedenen Transkriptionsfaktoren besteht und der an dem Promotor sequenziell aufgebaut wird. Oft ist die Erkennungssequenz, an der der Präinitiationskomplex positioniert wird – ähnlich wie das -10-Element prokaryontischer Promotoren –, eine TATA-Box.

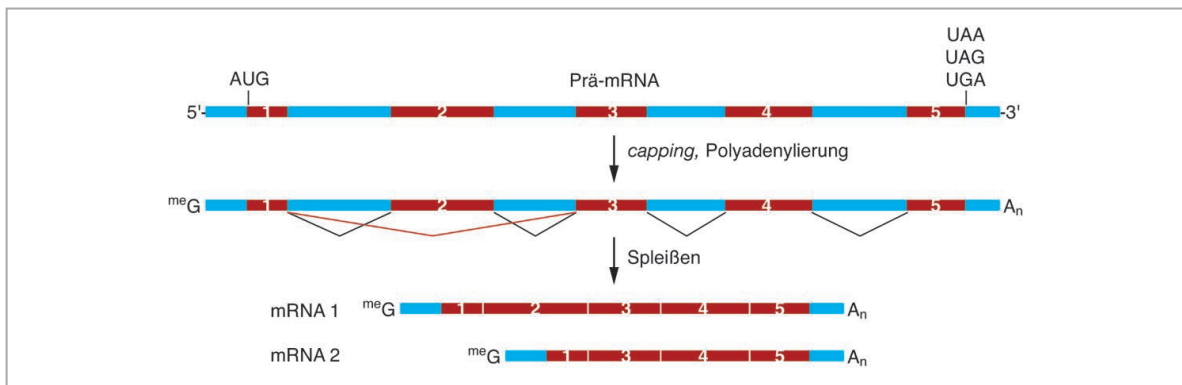
Der zentrale Transkriptionsfaktor, der TATA-Boxen sequenzspezifisch bindet, ist TBP (*TATA-binding protein*). TATA-Boxen liegen im Bereich von 26 – 34 bp oberhalb der Initiationsstelle für die Transkription. Sie können auch als *Core*-Promotoren bezeichnet werden. Die Bindung von TFIID an die DNA wird durch weitere Transkriptionsfaktoren stabilisiert, und schließlich wird durch Rekrutierung der RNA-Polymerase II ein Initia-



• **Abb. 3.13** Allgemeine Struktur der 7-Methylguanin-Kappe eukaryontischer mRNAs: Ein 7-Methylguanin ist über eine 5'-5'-Triphosphatbrücke mit dem ersten transkribierten Nucleotid verbunden. Die  $m^7GpppN$ -Struktur wird auch als Cap 0 bezeichnet. In Säugetierzellen ist die Ribose des ersten Nucleotids der mRNA am 2'-OH in der Regel methyliert (Cap 1). Es können auch die Ribosen der ersten beiden Nucleotide methyliert sein (Cap 2).



• **Abb. 3.14** Prozessierung eukaryontischer Transkripte: Die Prozessierung eukaryontischer mRNAs geschieht in drei Schritten: Capping, Polyadenylierung und Spleißen. Das Anbringen der 7-Methylguanin-Kappe ( $meG$ ) ist der erste Schritt, er passiert noch während der Transkription, wenn die Prä-mRNA gerade einmal ca. 30 Basen lang ist. Die Polyadenylierung, also die Anheftung von poly(A) an das 3'-Ende der RNA ( $A_n$ ), kann erst erfolgen, nachdem die Prä-mRNA komplett gebildet wurde. Das Spleißen kann bereits während der Transkription beginnen und mit der 3'-Prozessierung parallel verlaufen. Die Exons (1 – 5) werden dabei basengenau zusammengesetzt, sodass ein durchgehender Leserahmen für die Translation eines Proteins resultiert.

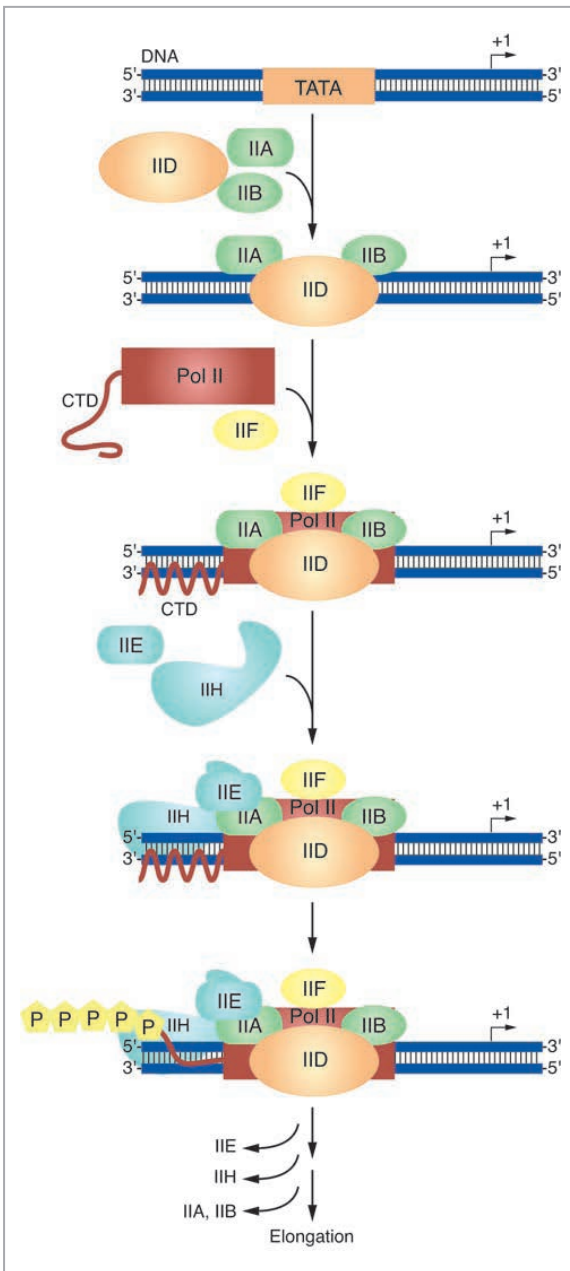


• **Abb. 3.15** Alternatives Spleißen: Sind in einem Gen mehrere Introns vorhanden, bietet sich die Möglichkeit, verschiedene Exons miteinander zu verbinden. Daraus resultieren verschiedene mRNAs bzw. Proteinprodukte. Im gezeigten Beispiel kann alternativ zur bereits bekannten mRNA 1 das Exon 1 mit Exon 3 verbunden (rote Linie) werden, sodass Exon 2 nicht in der reifen mRNA 2 vorhanden ist. Alternatives Spleißen geschieht nicht zufällig: Solche Prozesse sind in der Zelle streng reguliert.

tionskomplex gebildet (• Abb. 3.16). Die RNA-Polymerase II beginnt dann die Transkription an einer pyrimidinreichen Sequenz, die als Initiator bezeichnet wird. Zwei weitere typische Promotor-Elemente, die CCAAT-Box und die GGGCGG-Box (auch GC-Box genannt), findet man häufig oberhalb der TATA-Box. Ihre Positionen relativ zur TATA-Box sind allerdings variabel. Eukaryontische Promotoren sind bei weitem nicht so klar aufgebaut wie prokaryontische Promotoren. Beispielsweise sind eine ganze Reihe eukaryontischer Promotoren bekannt, die keine TATA-Box besitzen.

Oft wird die Aktivität eukaryontischer Promotoren durch Sequenzmotive auf der DNA beeinflusst, die zum Teil sehr weit vom Transkriptionsstart entfernt liegen. Noch überraschender war es, als man feststellte, dass

derartige Elemente auch in Introns, also mitten in einem Gen, oder gar hinter einem Gen lokalisiert sein können. Diese Elemente werden als Enhancer-Elemente bezeichnet, wenn sie die Stärke des Promotors positiv beeinflussen. Dagegen haben Silencer-Elemente einen negativen Einfluss auf die Aktivität des Promotors. Innerhalb des Enhancers lassen sich wiederum Sequenz-Motive definieren, die zum Teil als eigenständige Enhancer fungieren können, zum Teil im Enhancer selbst redundant vorhanden sind oder die auch in anderen Enhancern auftreten. An sie binden Faktoren, die mit dem Präinitiationskomplex am Kern-Promotor wechselwirken und auf diese Weise die Transkription verstärken. Dazu muss jedoch die DNA zwischen dem Kern-Promotor und dem Enhancer eine Schleife bil-



● **Abb. 3.16** Aufbau des RNA-Polymerase-II-Initiationskomplexes: TFIID bindet an die TATA-Box im Promotorbereich eines Gens. Der TFIID/DNA-Komplex wird durch TFIIB und TFIIB stabilisiert. Nur in Verbindung mit TFIIF kann die RNA-Polymerase II an diesen Komplex binden. Zwei weitere Faktoren, TFIIE und TFIIH, komplettieren den Initiationskomplex. TFIIH hat wichtige regulatorische Funktionen an der RNA-Polymerase und enthält eine DNA-Helikase, die die DNA am Transkriptionsstart entwindet und damit die Initiation der Transkription ermöglicht.

den, um die weit voneinander entfernten DNA-Bereiche in eine Nachbarschaft zueinander zu bringen. Die Interaktion zwischen Kern-Promotor und Enhancer erfolgt dann über Wechselwirkungen der Proteine, die an die Regulator-Bereiche gebunden haben.

### Zusammenfassung

- Transkription läuft bei Eukaryonten im Prinzip gleich ab wie bei den Prokaryonten. Allerdings ist die Transkription in Eukaryonten komplexer reguliert und die RNA-Transkripte werden in der Regel umfangreich prozessiert. Diese Prozessierung umfasst das Anbringen eines 7-Methyl-Guanins am 5'-Ende sowie die Polyadenylierung des 3'-Endes der primären mRNA. Die Gene von Eukaryonten bestehen meist aus einem Mosaik von proteincodierenden Bereichen (Exons) und nicht-codierenden Bereichen (Introns). Dadurch ist bedingt, dass im letzten Schritt der Prozessierung eukaryontischer Prä-mRNAs die Introns basengenau entfernt werden müssen (Spleißen).

### 3.4 Posttranslationale Modifikationen an Proteinen

3

In Eukaryonten können Proteine auf unterschiedlichste Weise während oder nach der Translation modifiziert werden. Dazu zählen beispielsweise Acylierungen, Sulfatierungen, die  $\beta$ -Hydroxylierung von Aspartatresten, die  $\gamma$ -Carboxylierung von Glutamatresten oder die Amidierung der freien Carboxylenden von Proteinen. Posttranslationale Modifikationen können Proteine auf vielfältige Weise beeinflussen, etwa deren Faltung in die korrekte Tertiärstruktur, die Stabilität eines Proteins innerhalb oder außerhalb der Zelle oder die biologische Funktionalität – beispielsweise die katalytische Aktivität oder Wechselwirkung mit einem Rezeptor.

Manche Proteine müssen posttranslational acyliert werden, um biologisch aktiv zu sein. Beispielsweise werden auf manche Proteine Fettsäuren übertragen, d. h. die Proteine werden acyliert. Die C16-Fettsäure Palmitinsäure wird in der Regel über einen Thioester an einen Cysteinrest gebunden (Palmitoylierung). Die C14-Fettsäure Myristinsäure wird dagegen exklusiv über eine Amid-Bindung an N-terminale Glycinreste gekoppelt (Myristoylierung). Durch Acylierung erwerben die Proteine meist die Fähigkeit, mit Membranen zu interagieren. Oft ist diese Fähigkeit essenziell für die Funktionsfähigkeit des entsprechenden Proteins. Hieraus ergeben sich auch attraktive Ansätze für eine Pharmakotherapie. Durch Hemmung der Acylierung könnten sich fehlregulierte Varianten derartiger Proteine in ihrer Funktion inhibieren lassen. Interessante Beispiele für einen solchen Therapieansatz sind die Proteine p60<sup>src</sup> und p21<sup>ras</sup>. In beiden Fällen handelt es sich um Onkogen-Produkte, d. h. um Proteine, die an der Entstehung von Tumoren beteiligt sind. Sie fungieren normalerweise als Teilstationen innerhalb komplexer Signalkaskaden, an deren

Enden Signale zur Zellteilung in den Kern gelangen. Sind die p60<sup>src</sup>- bzw. p21<sup>ras</sup>-codierenden Gene mutiert, lassen sich die Proteine nicht mehr inaktivieren und senden ständig Proliferationssignale aus. Beide Proteine sind in ihrer aktiven Form obligat acyliert. Es könnten daher Inhibitoren entwickelt werden, die die Acylierung dieser Proteine verhindern, um damit die gestörte Funktion der Proteine zu neutralisieren. Damit wäre der unphysiologische Teilungsreiz der betroffenen Zelle abgestellt. Die Zelle würde wieder einer geregelten Wachstumskontrolle unterliegen und sich nicht permanent teilen. Dies ist eines von vielen attraktiven neuen Konzepten, Tumore kausal zu behandeln.

Die Carboxylierung von Glutaminsäure-Seitenketten ( $\gamma$ -Carboxylierung) durch eine Vitamin-K-abhängige mikrosomale  $\gamma$ -Carboxylase ermöglicht Proteinen eine Bindung von Calcium und kann sich auf die biologische Aktivität der entsprechenden Proteine auswirken. Innerhalb der Blutgerinnungskaskade beispielsweise gibt es verschiedene Proteine, wie die Faktoren II, IX, X und das Protein C, die umfassend  $\gamma$ -carboxyliert werden. Generell verlieren diese Proteine ihre biologische Aktivität, wenn diese typische 4-Carboxyglutamat-Modifikation ausbleibt. Die mikrosomale Carboxylase steht nicht in allen für die biotechnologische Produktion von Proteinwirkstoffen etablierten Zellsystemen zur Verfügung, sodass darauf geachtet werden muss, dass die eingesetzten Produktionszellen ausreichend mit den entsprechenden Enzymen ausgestattet sind.

Besondere Beachtung bei der biotechnologischen Produktion rekombinanter Proteinwirkstoffe verdienen Proteine, die als inaktive Vorläuferproteine von der Zelle hergestellt werden und nach der Sekretion in das Plasma proteolytisch aktiviert werden müssen. Die verschiedenen Blutgerinnungsfaktoren sind auch hierfür gute Beispiele. Diese Faktoren werden physiologisch erst nach der Stimulation der Blutgerinnungskaskade durch Proteolyse aktiviert. Sollen solche Proteine als Wirkstoffe die Blutgerinnung regulieren, muss man entscheiden, ob die Faktoren in der inaktiven (aber in das Plasma sekretierten) Vorstufe produziert werden sollen (z. B. Faktor VIII) oder ob womöglich ein bereits aktivierter Faktor hergestellt werden muss, wie beispielsweise Faktor VIIa, der gegeben wird, um eine Resistenz des Patienten gegen Faktor VIII zu umgehen.

Eine Herausforderung ist die Herstellung therapeutischer Glykoproteine. Fehlende oder falsche (d. h. nicht für menschliche Zelle typische) Glycosylierungen können (müssen aber nicht) die Funktion des Proteins beeinflussen oder können den Patienten durch immunologische Reaktionen gefährden. Glycosylierungsmuster auf Proteinen sind prinzipiell in jedem Organismus unterschiedlich, und es ist äußerst schwierig, die für humane Zellen typischen Glycosylierungsmuster von Proteinen in biotechnologisch erzeugten Produkten nachzuahmen.

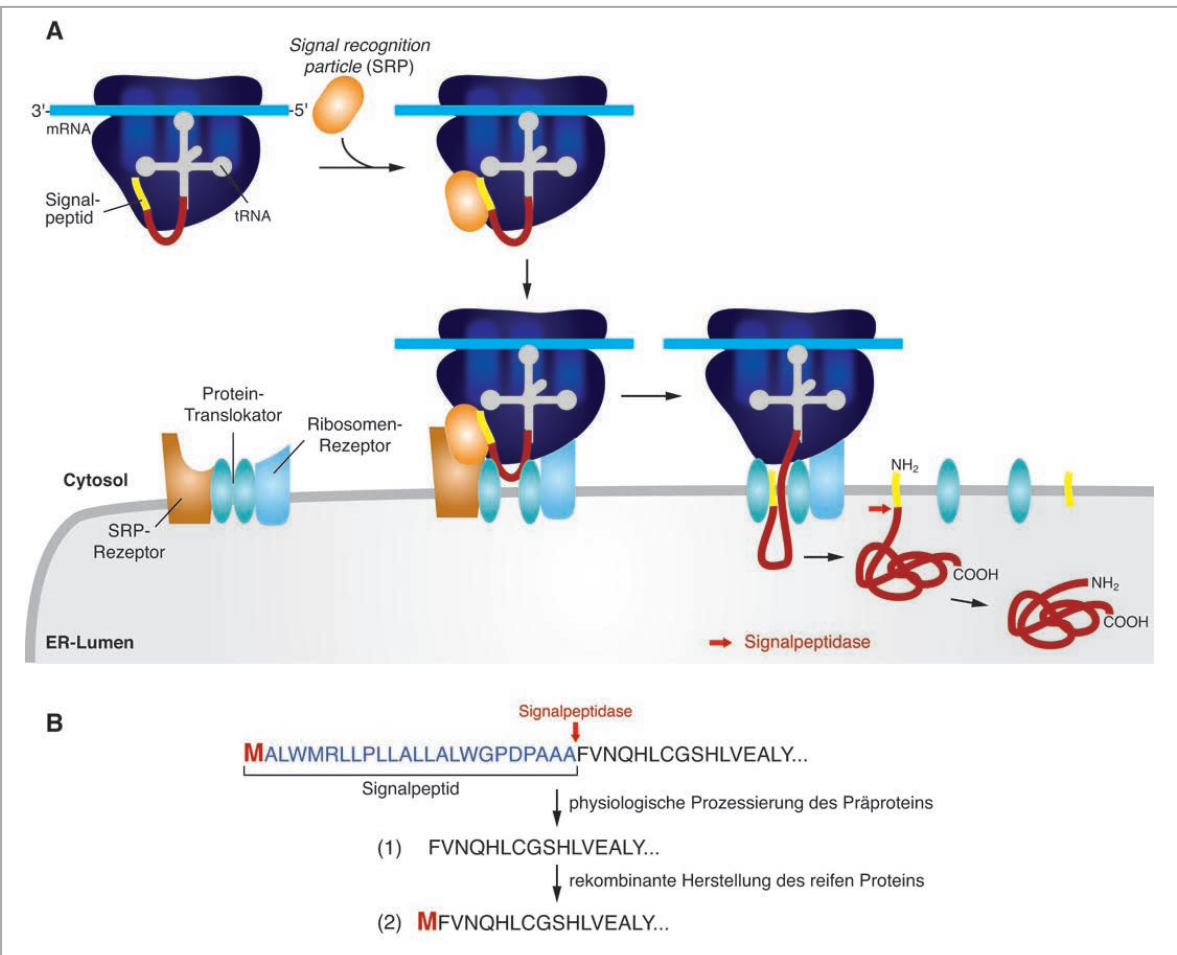
Die biotechnologisch herausfordernden Probleme der posttranslationalen proteolytischen Prozessierung und Glycosylierung von Proteinen sollen nachfolgend etwas ausführlicher betrachtet werden.

### 3.4.1 Proteolytische Prozessierung von Proteinen

Die meisten humanen Proteine, die bisher als Wirkstoffe für die Therapie entwickelt wurden, erfüllen ihre Funktionen extrazellulär, z. B. als Hormone, Zytokine oder Antikörper. Dieser Umstand bedingt einige Überlegungen hinsichtlich der biotechnologischen Produktion solcher Proteine, die wir uns kurz vor Augen führen müssen. Vor allem geht es darum, zu verstehen, dass extrazelluläre Proteine auf dem sekretorischen Weg von der Zelle ausgeschleust werden und dabei proteolytisch modifiziert werden.

Proteinsynthese findet an den im Zytoplasma lokalisierten Ribosomen statt. Somit ist leicht vorstellbar, dass neu gebildete Proteine auch im Zytoplasma entstehen. Wie verhält es sich aber mit Proteinen, die in eines der intrazellulären, von Membranen umschlossenen Kompartimente der Zelle, wie z. B. in Lysosomen oder in das endoplasmatische Retikulum, oder sogar in den Extrazellulärraum transportiert werden müssen, um ihre Funktionen ausüben? Solche Proteine werden nicht etwa nach ihrer Synthese aus dem Zytoplasma in das jeweilige Kompartiment gebracht, sondern sie werden direkt in ein Membransystem hinein synthetisiert, und zwar in das Lumen des endoplasmatischen Retikulums (ER). Von dort aus werden die Proteine dann über einen vesikulären Transport auf ihre jeweiligen Bestimmungsorte verteilt.

Die Synthese und der Transport von Proteinen in das ER sind gekoppelt. Proteine, die in ein Membrankompartiment gehören oder sekretiert werden sollen, tragen ein Signalpeptid, das von den im Zytoplasma vorhandenen Ribosomen zuerst translatiert wird (Abb. 3.17). Das entstandene Signalpeptid wird dann von einem spezialisierten Proteinkomplex (*signal recognition particle*, SRP) gebunden. Dieses hält die Proteinsynthese zunächst an, sodass der gesamte Komplex aus SRP und Ribosom über einen speziellen SRP-Rezeptor an die Außenseite der ER-Membran binden kann. ER-Membranen, die mit Ribosomen besetzt sind, nennt man aufgrund des Aussehens im Elektronenmikroskop auch das raue ER. Über den SRP-Rezeptor kommt das Ribosom in Kontakt mit einem Kanalprotein, dem Protein-Translokator. Die wachsende Polypeptidkette wird nun durch den Protein-Translokator hindurch in das ER-Lumen synthetisiert, wobei das Signalpeptid an den Translokator gebunden bleibt (Abb. 3.17). Ist das gesamte Protein durch die Pore gefädelt, schneidet die im ER-Lumen vorhandene Signalpeptidase das Signalpeptid ab und das Protein befindet sich frei im ER-Lumen, wo es sich in seine funktionale Tertiärstruktur faltet.



• **Abb. 3.17** Proteinsynthese am rauen endoplasmatischen Retikulum (ER): **(A)** Nach Beginn der Translation wird das frisch synthetisierte Signalpeptid vom SRP (*signal recognition particle*) gebunden. Die Translation wird kurzzeitig unterbrochen. Der SRP-Ribosomen-Komplex bindet an SRP-Rezeptoren an der Außenseite des ER und bringt dadurch das Ribosom in Kontakt mit einem Ribosomen-Rezeptor und einem Protein-Translokator. Das Protein wird nun durch die Pore des Protein-Translokators hindurch in das Lumen des ER synthetisiert. Das Signalpeptid bleibt dabei im Translokator hängen. Nachdem das Protein komplett translatiert wurde, zerfällt der Translokator-Komplex und das Signalpeptid wird durch eine Signalpeptidase (roter Pfeil) vom Rest des Proteins abgeschnitten. Das frei im ER-Lumen befindliche Protein kann sich nun in seine dreidimensionale Struktur falten. **(B)** Gezeigt ist der Aminoterminus des Insulins. Das Signalpeptid ist blau hervorgehoben. Sekretierte Proteine wie Insulin beginnen nicht mit der Aminosäure Methionin, da sie während des Exports in das ER von einer Signalpeptidase gespalten werden. Für die Herstellung eines rekombinanten Proteins bedeutet das, dass man entweder das Präprotein in einer Zelle herstellen muss, die den sekretorischen Weg beherrscht und das Präprotein prozessieren kann (1), oder man exprimiert das reife Protein, muss dann allerdings durch Design der cDNA dafür sorgen, dass ein artifizielles Translations-Startcodon eingeführt wird, das ein zusätzliches Methionin in das Protein einbringt (2).

Proteine, die ein Signalpeptid tragen, das während der Translation des Proteins in das ER entfernt wird, werden als Präproteine bezeichnet (• Abb. 3.17). Wir haben am Beginn dieses Kapitels gelernt, dass jedes Protein mit einem Methionin beginnt, weil die Ribosomen den Beginn eines proteincodierenden Leserahmens an dem Translations-Startcodon AUG erkennen. Präproteine sind insofern eine Ausnahme, dass sie zwar konventionell translatiert werden und mit Methionin beginnen, dass aber die Signalsequenz beim Export in das ER abgeschnitten wird und damit kein Methionin mehr am Anfang des reifen, aktiven Proteins vorhanden

ist. Das stellt für die biotechnologische Produktion solcher Proteine eine Herausforderung dar, da den Ribosomen in egal welchen Expressionszellen grundsätzlich ein Gen mit Translations-Startcodon präsentiert werden muss, damit ein rekombinantes Protein gebildet wird. Zur Lösung dieses Dilemmas gibt es zwei Ansätze:

- Die komplette genetische Information des Präproteins wird in eine Wirtszelle eingebracht. Beherrscht diese Wirtszelle den sekretorischen Weg und kann Proteine in den extrazellulären Raum transportieren, wird das Präprotein während der Translation in das

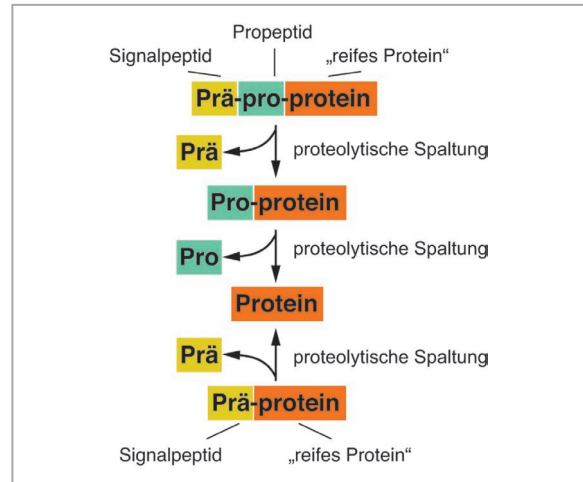
ER prozessiert und die Signalsequenz entfernt, sodass das reife Protein (ohne initiales Methionin) im Kulturüberstand der Zelle erscheint (● Abb. 3.17). Zu beachten ist, dass es sich bei dieser Art des Exports von Proteinen über das ER und den Golgi-Apparat offensichtlich um eine Eigenschaft von Eukaryonten handelt und Bakterien für diese Art der Expression von Präproteinen nicht geeignet sind. Allerdings haben auch eukaryontische Mikroorganismen wie z. B. Hefen Probleme mit der Sekretion humaner Präproteine, weil sich die Signalpeptide von Hefen und Säugetierzellen hinreichend unterscheiden. Um von Hefen effizient sekretierte Proteine zu erhalten, kann man jedoch Kunstgene herstellen, die ein Signalpeptid aus einem Hefeprotein mit der Sequenz des reifen humanen Proteins kombinieren.

- Die genetische Information des reifen Proteins wird isoliert und ein AUG-Translations-Startcodon künstlich hinzugefügt (● Abb. 3.17). Das ist die Methode der Wahl zur bakteriellen Expression von Proteinen, die eigentlich von humanen Zellen sekretiert werden. Ein Nachteil dieser Strategie ist, dass derart hergestellte und als Wirkstoffe eingesetzte Proteine ein zusätzliches Methionin am Anfang des Proteins tragen. Da dieses Methionin im physiologischen Vorbild, also dem sezernierten Protein, fehlt, kann die Anwendung solcher Wirkstoffe im Patienten zu immunologischen Überempfindlichkeitsreaktionen führen.

Beide genannten Ansätze werden bei der Konstruktion von Wirt-Vektor-Systemen für die industrielle Proteinproduktion bedacht und angewendet, wie wir später bei der Diskussion der Herstellung rekombinanter Expressionseinheiten für die Proteinproduktion und der Vorstellung einzelner Wirkstoffe sehen werden.

Viele Proteine müssen zusätzlich zur Abspaltung des Signalpeptids weiter proteolytisch prozessiert werden, um die aktive Form des Proteins zu generieren. In solchen Fällen einer zweistufigen proteolytischen Prozessierung nennt man das Vorläuferprotein Präproprotein. Nach Abspaltung des Signalpeptids entsteht zunächst ein (inaktives) Proprotein, das anschließend durch Entfernen eines Propeptids zum reifen (aktiven) Protein prozessiert wird (● Abb. 3.18).

Bei einigen Proteinen, die physiologisch durch posttranslationale proteolytische Spaltung aktiviert werden, muss bei der biotechnologischen Herstellung entschieden werden, ob das aktive, reife Protein oder das inaktive Proprotein hergestellt werden soll, wie beispielsweise bei der Produktion von Proteinen der Blutgerinnungskaskade. Gegebenenfalls muss man eine Protease, die den Wirkstoff durch Spaltung konvertieren kann, ebenfalls rekombinant herstellen. Ein Beispiel hierfür ist die unvollständige Prozessierung des Blutgerinnungsfaktors IX während der Produktion des Wirkstoffs Nonacog alfa,



● **Abb. 3.18** Proteolytische Prozessierung von inaktiven Vorläuferproteinen: Proteine, die vom Zytosol (dem Ort der Proteinbiosynthese) in bestimmte membranumhüllte Kompartimente der Zelle oder in den extrazellulären Raum exportiert werden sollen, tragen am Aminoterminus Signalsequenzen. Diese Signalsequenzen werden bereits bei der Synthese des Proteins in das Lumen des rauen endoplasmatischen Retikulums (ER) durch sogenannte Signalpeptidasen entfernt (● Abb. 3.17). Proteine, die solche Signalsequenzen tragen, nennt man Präproteine, nach Abspaltung des Signalpeptids entsteht das reife (aktive) Protein. Muss das Protein nach der Abspaltung des Signalpeptids noch zusätzlich proteolytisch gespalten werden, nennt man das Vorläuferprotein Präproprotein. Nach Abspaltung des Signalpeptids entsteht ein Proprotein, das anschließend zum reifen Protein prozessiert wird.

die durch gleichzeitige Expression der Serin-Protease PACEsol verbessert werden kann. PACE (*paired basic amino acid cleaving enzyme*) ist eine Subtilisin-ähnliche, im Golgi-Apparat lokalisierte Proprotein-konvertase, die auch unter dem Namen Furin bekannt ist. Das Enzym sorgt dafür, dass Propeptide während des Transfers von Proproteinen durch ER und Golgi zur Zelloberfläche abgespalten werden. PACEsol ist eine lösliche (d. h. nicht membrangebundene) Variante dieses Enzyms, die gemeinsam mit dem eigentlichen Proprotein-Wirkstoff exprimiert werden kann und aufgrund der Kollokalisierung mit dem sekretierten rekombinanten Propeptid-Wirkstoff in ER und Golgi für eine effiziente Abspaltung des Propeptids sorgt, sodass das sekretierte, aktive Protein aus dem Kulturüberstand gereinigt werden kann.

### 3.4.2 Glycosylierung von Proteinen

Wir unterscheiden zwei Typen von posttranslationalen Glycosylierungen an Proteinen: die *N*-Glycosylierung und die *O*-Glycosylierung. Bei der *N*-Glycosylierung werden Kohlenhydratketten mit dem Säureamid in der Seitenkette der Aminosäure Asparagin *N*-glycosidisch verbunden. Bei der *O*-Glycosylierung werden in der Regel die Hydroxylgruppen der Seitenketten von Serin oder

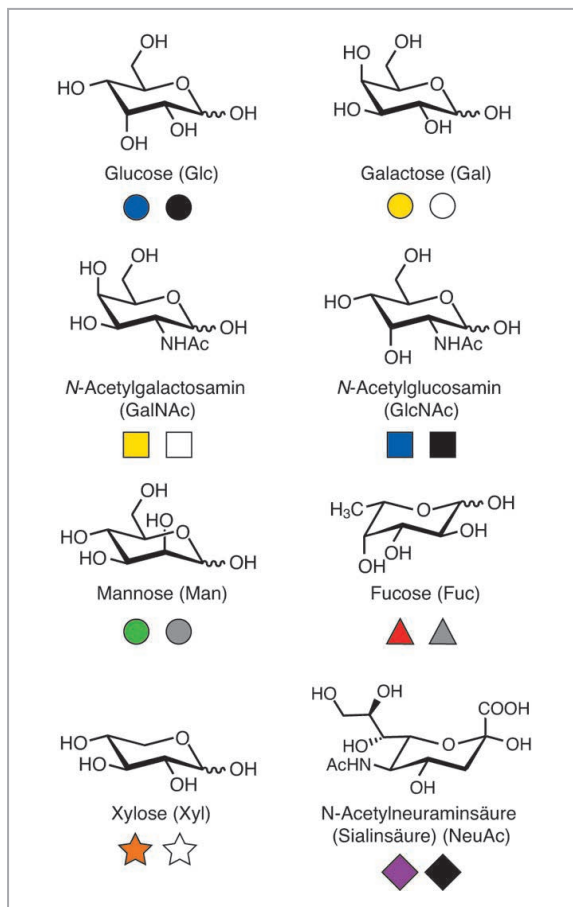
Threonin mit Oligosacchariden *O*-glycosidisch verbunden. Der Symbol-Gebrauch für die Darstellung der in den Kohlenhydrat-Ketten von Glykoproteinen häufig vorkommenden Monosaccharide ist in **Abb. 3.19** gezeigt.

Das vollständige Dokument des Consortium for Functional Glycomics ist über den nebenstehenden QR-Code abrufbar.



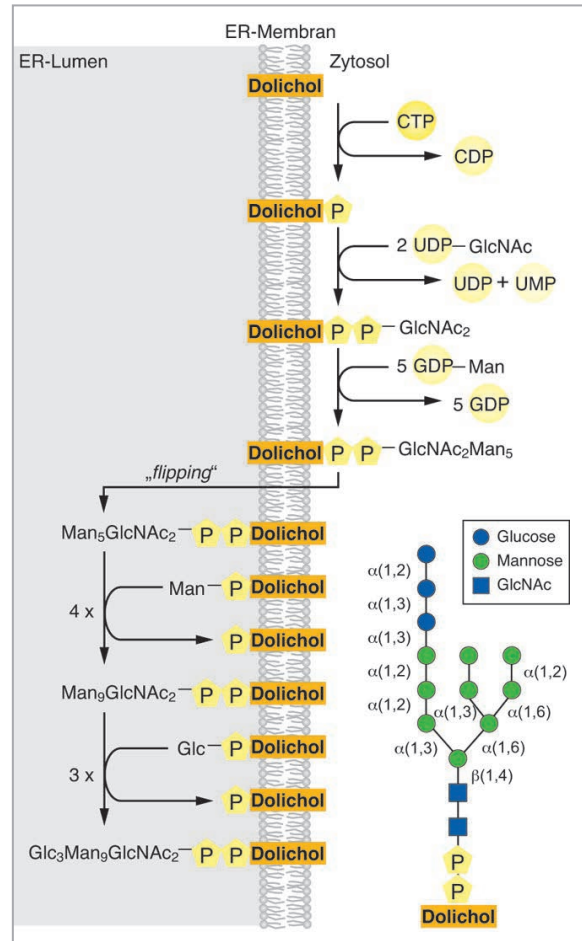
### *N*-Glycosylierung von Proteinen

Proteine, die in Vesikeln zu anderen intrazellulären Membransystemen (z. B. Lysosomen), zur Zellmembran oder in den Extrazellulärraum transportiert werden sollen, werden direkt in das Lumen des ER hinein synthetisiert (**Abb. 3.17**). Die Proteine passieren dann durch vesikulären Transport vom ER aus die verschiedenen Regionen des Golgi-Apparats, bevor sie von dort weiter verteilt werden. Bereits während der Synthese in das ER-Lumen können Proteine mit komplexen Kohlenhydrat-Ketten modifiziert werden.



**Abb. 3.19** Symbolik für Monosaccharide in *O*- und *N*-Glycanen an Proteinen (nach *Symbol Nomenclature for Glycans*). Die Symbole sind jeweils für die farbige und schwarz-weiße Darstellung angegeben.

Für die *N*-Glycosylierung von Proteinen wird zunächst ein an Dolicholpyrophosphat gebundenes Vorläufer-Oligosaccharid synthetisiert, das in allen eukaryontischen Zellen gleich aufgebaut ist (**Abb. 3.20**). Dolichol ist ein aus 22 Isopren-Einheiten aufgebauter Polyterpen-Kohlenwasserstoff, der die ER-Membran mehrmals durchspannt und für eine feste Verankerung des Oligosaccharid-Vorläufers im ER-Lumen sorgt. Das Vorläufer-Oligo-



**Abb. 3.20** Biosynthese und Struktur des universellen *N*-Glycan-Vorläufers: Der universelle *N*-Glycan-Vorläufer Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> wird Schritt für Schritt aufgebaut. Der erste Zucker (GlcNAc) wird über eine Pyrophosphatgruppe an Dolichol gehängt und damit auf der zytosomatischen Seite des ER verankert. Die Bindungsenergie der Pyrophosphatgruppe wird später benutzt, um das Glycan auf einen Asparaginrest eines Proteins zu übertragen. Zunächst wird auf der zytosomatischen Seite des ER ein Man<sub>5</sub>GlcNAc<sub>2</sub>-Glycan aufgebaut. Die Energie für die Knüpfung der Bindungen kommt jeweils aus der Hydrolyse der glycosidischen Bindung der nukleotidaktivierten Zucker. Das an Dolicholpyrophosphat gebundene Man<sub>5</sub>GlcNAc<sub>2</sub>-Glycan wird dann durch ein spezielles Transportprotein auf die Lumenseite des ER transferiert (*flipping*). Hier wird das Glycan weiter bis zum Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> aufgebaut. Dabei dienen nicht nukleotidaktivierte Zucker, sondern an Dolicholphosphat aktivierte Zucker als Substrate.

saccharid besteht aus den Zuckern Glucose (Glc), Mannose (Man) und *N*-Acetylglucosamin (GlcNAc) und hat die Zusammensetzung  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  (● Abb. 3.20).

Der erste Schritt der *N*-Glycosylierung von Proteinen ist die Übertragung des Oligosaccharid-Vorläufers von Dolichol auf einen Asparaginrest des im ER-Lumen synthetisierten Proteins (● Abb. 3.21), wobei eine  $\beta$ -ständige *N*-glycosidische Bindung mit dem amidischen Stickstoff der Asparagin-Seitenkette und der C1-Hydroxylgruppe des ersten *N*-Acetylglucosamins des Oligosaccharid-Vorläufers gebildet wird. Das modifizierte Asparagin des Proteins ist Teil einer kryptischen Erkennungssequenz Asn-X-Ser/Thr, wobei X eine beliebige Aminosäure außer Prolin sein kann. Diese kurze Sequenz wird von ER-ständigen Oligosaccharyltransferasen erkannt, muss allerdings für das modifizierende Enzym frei zugänglich sein, damit die Glycosylierung tatsächlich durchgeführt werden kann.

Während des Transports *N*-glycosylierter Proteine vom ER durch die Membranstapel des Golgi-Apparates wird der Oligosaccharid-Vorläufer vielfach modifiziert. Dabei wird der  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ -Vorläufer zuerst teilweise wieder abgebaut, was als Trimmen (*trimming*) bezeichnet wird. Im ersten Schritt des Trimmens spaltet das im ER lokalisierte Enzym  $\alpha$ -Glucosidase I die  $\alpha(1,2)$ -verknüpfte terminale Glucose vom  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ -Vorläufer ab (● Abb. 3.21, Schritt 1). Die nachfolgenden terminalen  $\alpha(1,3)$ -verknüpften Glucose-Einheiten werden ebenfalls im ER abgespalten, und zwar von dem Enzym  $\alpha$ -Glucosidase II (● Abb. 3.21, Schritt 2). Diese initialen Glucose-Abspaltungen sind nicht nur als Eingangsreaktionen für die nachfolgenden Glycan-Modifikationen von Belang, die insbesondere zu den komplexen *N*-Glycanen der Säugetierzellen führen, sondern sie sind auch wichtige Signale für die Interaktion der neu synthetisierten Proteine mit Chaperonen wie Calnexin und Calreticulin, die den Proteinen beim Falten in die korrekte Tertiärstruktur assistieren.

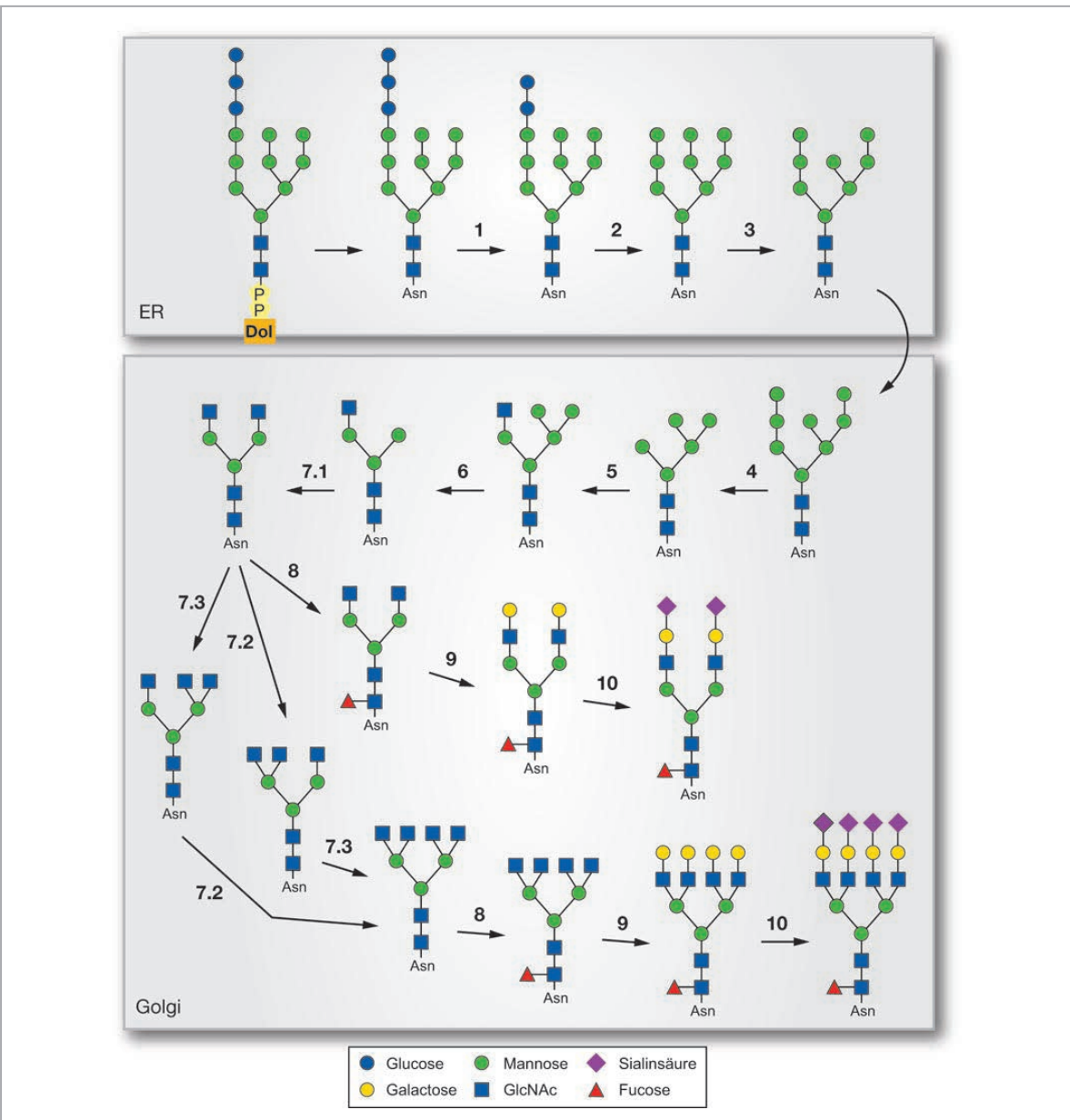
Ein weiteres Trimmen durch ER- und Golgi-ständige  $\alpha(1,2)$ -Mannosidasen produziert verschiedene Isoformen gering mannosylierter Glycane, insbesondere  $\text{Man}_5\text{GlcNAc}_2$  (● Abb. 3.21, Schritte 3 und 4). Im medialen Golgi werden solche Oligomannose-Glycane durch Hinzufügen neuer Zucker-Einheiten weiter verändert, wobei diese nachfolgenden Modifikationen in jedem Organismus unterschiedlich durchgeführt werden.

Auf dem Weg zur Bildung komplex aufgebauter *N*-Glycane in Säugetierzellen folgt als nächster Schritt die Übertragung eines Moleküls GlcNAc durch die GlcNAc-Transferase I (● Abb. 3.21, Schritt 5). Danach werden im verzweigten Teil des  $\text{GlcNAcMan}_5\text{GlcNAc}_2$  zwei Mannose-Einheiten durch  $\alpha$ -Mannosidase II entfernt. Im weiteren Verlauf wird das Glycan durch verschiedene Isoformen der GlcNAc-Transferase mit weiteren GlcNAc-Einheiten erweitert (● Abb. 3.21, Schritt 7). In diesem

Stadium kann das entstandene Glycan fakultativ am ersten GlcNAc des Glycan-Stamms durch eine  $\alpha(1,6)$ -Fucosyltransferase mit einer Einheit Fucose modifiziert werden (● Abb. 3.21, Schritt 8). Abschließend folgen zwei Modifikationen, die insbesondere für *N*-Glycane von Säugetierzellen typisch sind und häufig in Glykoproteinen aus anderen Spezies fehlen: Auf die terminalen GlcNAc-Einheiten überträgt die  $\beta(1,4)$ -Galactosyltransferase zunächst jeweils eine Einheit Galactose (● Abb. 3.21, Schritt 9), die dann als Akzeptor für terminale Sialinsäurereste dient, die wiederum von Sialyltransferasen meist in  $\alpha(2,3)$ - oder  $\alpha(2,6)$ -Verknüpfung angelagert werden.

Wie aus ● Abb. 3.22 ersichtlich ist, werden *N*-Glycosylierungen in verschiedenen Eukaryonten höchst unterschiedlich ausgeführt. In Säugetierzellen inklusive des Menschen werden extrazellulär aktive Glykoproteine in den meisten Fällen mit komplexen *N*-Glycanen modifiziert, die sich insbesondere durch die Präsenz einer terminalen Schicht aus Sialinsäure auszeichnen. Seltener werden in Säugetierzellen *N*-Glycane vom Highmannose-Typ gebildet, die das vorherrschende Glycosylierungsmuster in eukaryontischen Mikroorganismen wie Pilzen bilden. Die *N*-Glycane in Pflanzenzellen bestehen meist aus nur wenigen Mannose-Einheiten, die auch als Paucimannose bezeichnet werden. Sie können weiter mit GlcNAc modifiziert werden, enthalten aber in der Regel kaum Galactose und keine terminale Sialinsäure. Pflanzliche *N*-Glycane tragen häufig  $\alpha(1,3)$ -verknüpfte Fucose am Stamm-GlcNAc sowie  $\beta(1,2)$ -verbundene Xylosereste, was im Menschen nicht vorkommt und deshalb als antigen erkannt wird. Insektenzellen bilden ebenfalls hauptsächlich Paucimannose mit einem geringen Grad an GlcNAc-Modifikationen. Je nach Spezies fehlt Insekten zumeist das Enzymsystem zum Transfer von Galactose- und Sialinsäureresten, oder diese Enzymsysteme sind in etablierten Zellkulturen zu schwach exprimiert, um bei einer biotechnologischen Herstellung von Glykoprotein-Wirkstoffen ausreichend zu einer *N*-Glycosylierung vom komplexen Typ beizutragen. *N*-Glycane aus Insektenzellen enthalten häufig anstelle der im Menschen typischen  $\alpha(1,6)$ -Verknüpfung der Fucose am Stamm-GlcNAc eine  $\alpha(1,3)$ -Verknüpfung, was im Menschen als antigen erkannt wird.

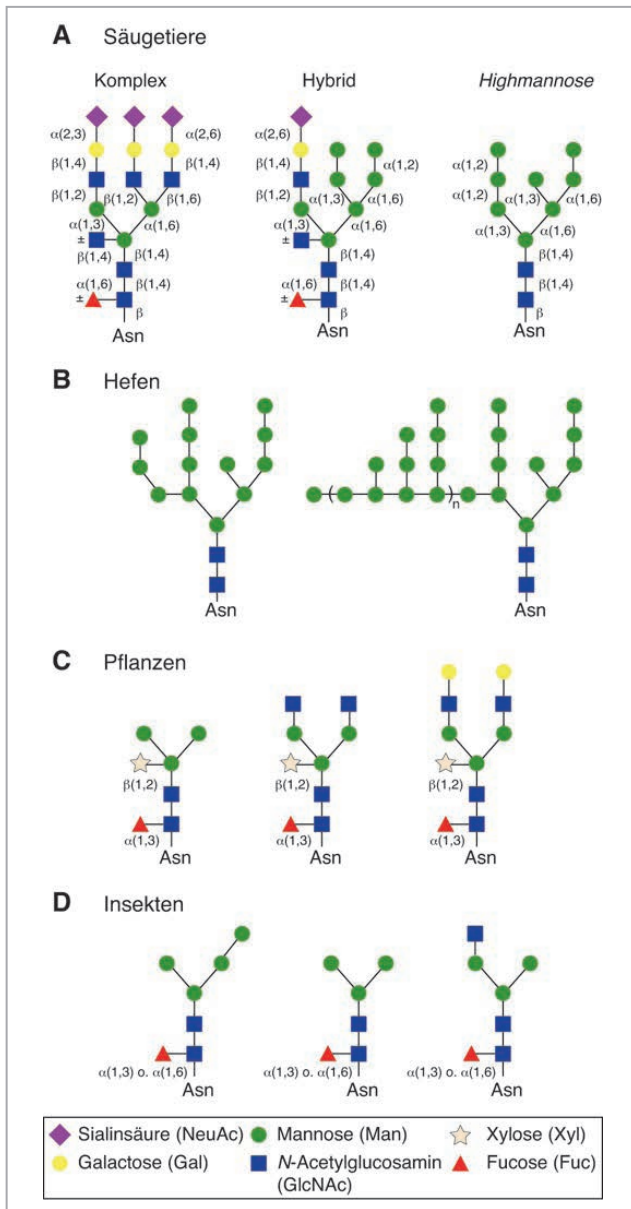
Wenn wir nachfolgend die verschiedenen, zur Verfügung stehenden Expressionssysteme für die Herstellung therapeutischer Glykoproteine diskutieren, wird es immer wieder Thema sein, welche *N*-Glycosylierungsmuster in biotechnologisch hergestellten Glykoprotein-Wirkstoffen durch die Auswahl bestimmter Produktionszellen resultieren (● Abb. 3.22) und wie diese Muster gegebenenfalls an die bevorzugte *N*-Glycosylierung vom Komplex-Typ in humanen Zellen angenähert werden könnten (► Kap. 3.4.3). Dabei muss bedacht werden, dass falsch ausgeführte, d. h. von der humanen Glycosylierungsform abweichende, Modifikationen zu Proteinen



● **Abb. 3.21** *N*-Glycosylierung auf dem sekretorischen Weg: Gezeigt ist beispielhaft der *N*-Glycosylierungsweg in CHO-Zellen. Die resultierenden *N*-Glycane sind wesentlich heterogener als hier dargestellt. Der Prozess der *N*-Glycosylierung beginnt im endoplasmatischen Retikulum (ER). Zunächst wird ein universeller  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ -Glycan-Vorläufer (●Abb. 3.20) von Dolichol (Dol) auf einen Asparaginrest eines neu synthetisierten Proteins übertragen. Im Lauf des Transports der Proteine vom ER zum Golgi-Apparat werden die Zuckerketten komplex modifiziert, bis das Protein schließlich von der Zelle sezerniert wird. Einige beteiligte Enzyme sind: (1)  $\alpha$ -Glucosidase I, (2)  $\alpha$ -Glucosidase II, (3) ER- $\alpha$ -Mannosidase, (4) Golgi- $\alpha$ -Mannosidase, (5) GlcNAc-Transferase I (GnT-I), (6) Golgi-Mannosidase II, (7.1) GlcNAc-Transferase II (GnT-II), (7.2) GnT-IV, (7.3) GnT-V, (8)  $\alpha(1,6)$ -Fucosyltransferase, (9)  $\beta(1,4)$ -Galactosyltransferase und (10)  $\beta$ -Galactosid- $\alpha(2,3)$ -Sialyltransferase. In Säugetierzellen können die terminalen Sialinsäurereste  $\alpha(2,3)$ - oder  $\alpha(2,6)$ -verbunden werden, allerdings exprimieren CHO-Zellen nur eine  $\alpha(2,3)$ -Sialyltransferase.

mit beeinträchtigter Funktion oder Plasmastabilität führen können (■Tab. 3.1). Zudem kann es beim Einsatz von Proteinwirkstoffen mit „falscher“ Glycosylierung im menschlichen Organismus zu potenten Unverträglichkeitsreaktionen kommen. Das liegt beispielsweise daran, dass Verknüpfungen von Zuckern, die im menschlichen

Körper nicht vorkommen, als fremd erkannt und vom Immunsystem attackiert werden können. Beispiele dafür sind  $\beta(1,4)$ -Fucose- und  $\beta(1,3)$ -Galactose-Verbindungen mit *N*-Acetylglucosamin. Außerdem wird geschätzt, dass ca. 1 % der zirkulierenden IgG-Antikörper im humanen Plasma gegen  $\text{Gal-}\alpha(1,3)\text{-Gal-}\beta(1,4)\text{-GlcNAc-Glycane}$



• **Abb. 3.22** Strukturen von *N*-Glycanen in verschiedenen Organismen: (A) Darstellung der drei prinzipiellen *N*-Glycan-Typen „Komplex“, „Hybrid“ und „Highmannose“. Der in Säugetierzellen typische „komplexe“ *N*-Glycosylierungs-Typ umfasst insbesondere terminale Galactose- und Sialinsäurereste. *N*-Glycane aus Säugetierzellen können am Stamm am ersten GlcNAc fucosyliert sein, wobei die Verknüpfung stets Fucose- $\alpha(1,6)$ -GlcNAc ist. Die terminalen Sialinsäurereste sind in der Regel  $\alpha(2,3)$ - oder  $\alpha(2,6)$ -verbunden. (B) *N*-Glycosylierungen von Hefe-Proteinen sind grundsätzlich vom Highmannose-Typ (Beispiel links) und enthalten keine säugetiertypischen terminalen Galactose- und Sialinsäurereste. Viele Hefe-Proteine sind zudem hypermannosyliert (Beispiel rechts). (C) Pflanzen-Proteine tragen meist gering mannosylierte *N*-Glycane (= Paucimannose), die zum Teil GlcNAc und Galactose, aber keine Sialinsäure enthalten. Pflanzliche *N*-Glycane tragen häufig  $\alpha(1,3)$ -verknüpfte Fucose am Stamm-GlcNAc sowie  $\beta(1,2)$ -verbundene Xylosereste. Beides wirkt im Menschen antigen. (D) Insektenzellen produzieren in der Regel gering hypermannosylierte *N*-Glycane (= Paucimannose). Je nach Spezies fehlt Insekten zumeist das Enzymsystem zum Transfer von Galactose- und Sialinsäureresten, oder diese Enzymsysteme sind in etablierten Zellkulturen zu schwach exprimiert, um bei einer biotechnologischen Herstellung von Glykoprotein-Wirkstoffen zu einer signifikanten *N*-Glycosylierung vom komplexen Typ beizutragen. *N*-Glycane aus Insektenzellen enthalten häufig  $\alpha(1,3)$ -verknüpfte Fucose am Stamm-GlcNAc, was im Menschen als antigen erkannt wird.

gerichtet ist, vermutlich weil solche Epitope auf verschiedenen Mikroorganismen vorkommen. Dies kann zu fatalen IgE-vermittelten allergischen Reaktionen bereits bei der ersten Anwendung eines therapeutischen Proteins führen, also vermeintlich ohne vorherige Sensibilisierungsphase durch Erstkontakt mit dem Proteinwirkstoff. Von besonderer Bedeutung für die Arzneimittelsicherheit sind auch die  $\alpha(1,3)$ -Fucosylierung des Stamm-GlcNAc, die in Pflanzen- und Insektenzellen häufig vorkommt, sowie die pflanzentypische  $\beta(1,2)$ -Verbindung von Xylose und Mannose (• Abb. 3.22). Beide bilden im Menschen potente IgE-Epitope und können damit Allergien auslösen.

Die kryptische Erkennungssequenz Asn-X-Ser/Thr (X = alle Aminosäuren außer Prolin) für potenzielle *N*-Glycosylierungsstellen in einem Protein ist keine

Garantie dafür, dass ein rekombinant hergestelltes Protein genau wie das physiologische Vorbild auch wirklich an dieser Stelle *N*-glycosyliert wird. Unter physiologischen Bedingungen wird ein Protein im Milieu des endoplasmatischen Retikulums mit der Hilfe organismusspezifischer Chaperone mit hoher Wahrscheinlichkeit in die korrekte Tertiärstruktur gefaltet. Diese korrekte Faltung ist dann die Grundlage für die korrekte Erkennung von potenziellen Glycosylierungsstellen durch die entsprechenden Glycosyltransferasen. Wird jedoch ein für den Organismus fremdes Protein exprimiert, könnte die Faltung in dem für dieses Protein nicht natürlichen Milieu dazu führen, dass bestimmte Glycosylierungsstellen, die eigentlich glycosyliert werden sollten, nicht erkannt werden. Auf der anderen Seite könnten Konsensusmotive, die bei korrekter Faltung des Pro-

■ **Tab. 3.1** Potenzielle Effekte von Glycosylierungen im Kontext therapeutischer Proteine

Rolle bzw. Effekt	Beispiel
Förderung der Proteinfaltung	Die Glycan-Ketten der Gonadotropine FSH und LH spielen eine Rolle bei der korrekten Faltung der Proteine im endoplasmatischen Retikulum und der anschließenden Sekretion der Proteine.
Förderung der Sekretion von Proteinen	Das Entfernen von zwei oder mehr <i>N</i> -Glycanen von Erythropoetin bewirkt eine stark reduzierte Sekretion des Proteins aus den Produktionszellen.
Förderung der Liganbindung bzw. des Targetings	Die terminalen Mannose-Einheiten der <i>N</i> -Glycane des Wirkstoffs Imiglucerase bewirken die Aufnahme in Makrophagen über die Bindung an Mannose-Rezeptoren.
Beteiligung an der biologischen Aktivität von Proteinen	Das Entfernen des an Asparagin-52 gebundenen <i>N</i> -Glycans in Gonadotropinen verstärkt die Rezeptorbindung, aber unterbindet die Aktivierung einer Signaltransduktion nach Rezeptorbindung. Die <i>N</i> -Glycane von IgG-Antikörpern spielen eine Rolle für die Effektorfunktion der Antikörper, wie beispielsweise die Fähigkeit zur Auslösung antikörperabhängiger, zellvermittelter Zytotoxizität (ADCC).
Stabilisierung von Proteinen	$\alpha$ -Galactosidase ist an Position 184 <i>N</i> -glycosyliert. Das Entfernen des <i>N</i> -Glycans führt zur Aggregation und Präzipitation des Proteins.
Regulation der Plasmahalbwertszeit von Proteinen	Eine Verminderung des Sialinsäuregehalts führt bei den meisten glycosylierten Proteinen (z. B. Erythropoetin) zu einer verstärkten Elimination durch die Bindung an Asialoglykoprotein-Rezeptoren auf Hepatozyten. Mannose und <i>N</i> -Acetylglucosamin exponierende Glykoproteine werden aus der Zirkulation durch Zellen entfernt, die Mannose-Rezeptoren tragen (z. B. Makrophagen).

teins eigentlich nicht zugänglich für Glycosylierungen wären, erkannt und glycosyliert werden. Dieses Phänomen nennt man Makroheterogenität, es ist ein grundsätzliches Problem bei der heterologen Expression glycosylierter Proteine. Aber auch selbst innerhalb einer Zuckerkette kann die Zusammensetzung der Zucker variieren (Mikroheterogenität), sodass sowohl biotechnologisch erzeugte Glykoproteine als auch ihre physiologischen Vorbilder eigentlich immer als eine Mischung

sehr ähnlicher, aber im Detail der *N*-Glycane leicht verschiedener Moleküle gebildet werden. Zudem wird die *N*-Glycosylierung eines heterolog exprimierten Proteins nicht nur vom verwendeten Wirtssystem und dem physiologischen Status der Wirtszellen, sondern auch von den Fermentationsbedingungen erheblich mitbestimmt.

Generell können wir festhalten, dass rekombinante humane Glykoproteine funktionell meist nur in Säugertierzellen hergestellt werden können, da alternative Expressionssysteme wie Insektenzellen oder Hefezellen die entsprechenden enzymatischen Aktivitäten nur unzureichend bilden. Wie wir im Abschnitt über Glycoengineering sehen werden, gibt es mittlerweile interessante Ansätze, um eine humantypische *N*-Glycosylierung therapeutischer Proteine auch in nicht humanen Expressionssystemen zu verwirklichen.

#### INN-Bezeichnungen für glycosylierte Proteinwirkstoffe

Rekombinante Glykoproteine können in der Proteinsequenz identisch sein, sich aber in ihren Glycosylierungen etwas unterscheiden, wenn sie in verschiedenen Zellsystemen hergestellt werden. Da die abweichenden Glycosylierungsmuster Einfluss auf die biologische und damit die therapeutische Qualität haben können, ist es sinnvoll, diese Wirkstoffe voneinander unterscheiden zu können. Daher bekommen Glykoprotein-Wirkstoffe entsprechend den Regularien der *International Nonproprietary Names (INN) for biological and biotechnological substances* (WHO/EMPI RHT/TSN/2019.1) in der Regel einen voll ausgeschriebenen griechischen Buchstaben als Namenszusatz. Die griechischen Buchstaben werden in der Reihenfolge des griechischen Alphabets verwendet, beginnend mit „alfa“, „beta“, etc. (z. B. Epoetin alfa, Epoetin beta). Diese Bezeichnungen (Schreibweise alfa!) sind nicht zu verwechseln mit der Namensgebung für physiologische Proteine wie Interferon alpha und Interferon beta oder Tumornekrosefaktor alpha und Tumornekrosefaktor beta, die unterschiedliche Aminosäuresequenzen und gegebenenfalls unterschiedliche physiologische Funktionen haben.

#### O-Glycosylierung von Proteinen

Die O-Glycosylierung von Proteinen erfolgt relativ spät im Golgi-Apparat. Im Gegensatz zur *N*-Glycosylierung erfolgt die O-Glycosylierung nicht durch Übertragung eines vorgefertigten Oligosaccharids auf Proteine, sondern die Zucker werden einzeln durch spezifische Glycosyltransferasen auf die Hydroxylgruppen der Aminosäure-Seitenketten von Serin oder Threonin bzw. die bereits vorhandenen Kohlenhydrate übertragen. O-Glycosylierungen sind schwieriger vorherzusagen als *N*-Gly-