

Oligonucleotid-Synthese, bis vor einigen Jahren noch eher ein exotisches Betätigungsgebiet weniger Experten, ist mittlerweile integraler Bestandteil des Arsenals molekularbiologischer Techniken. Gerade im letzten Jahrzehnt hat sich auf dem Gebiet der DNA-Synthese eine unglaublich rasante Entwicklung abgespielt, welche zur Automatisierung der Synthesen führte und schließlich in der Herstellung von Genen mit einer Länge von über 1000 Basenpaaren gipfelte. Maßgeschneiderte synthetische Gene ermöglichen heute auch die gentechnische Herstellung veränderter und sogar völlig neuartiger Proteine („de-novo-Protein-Design“). Im Zusammenspiel mit den modernen Methoden der Isolierung, Sequenzierung und Expression von Genen trug die Gensynthese in den letzten Jahren wesentlich zum enormen Fortschritt in der Gentechnologie bei.

1. Einleitung

Seit der grundlegenden Erkenntnis von *Watson* und *Crick* im Jahre 1953^[1] ist die Struktur und chemische Beschaffenheit der DNA in ihren Grundzügen bekannt. Ausgehend vom Aufbau der Nucleotidketten aus Zucker, Phosphat und Heterocyclen (Abb. 1) läßt sich die Primärstruktur der Doppelhelix aus den beiden antiparallel laufenden Einzelsträngen ableiten. Aus der jeweiligen Komplementarität der Heterocyclen und der damit möglichen Wasserstoffbrückenbindungen resultiert ein geordneter Verbund. Um auch komplexe Moleküle dieser Art in einer verständlichen Kurzform schreiben zu können, ist folgende Nomenklatur eingeführt worden: Die in der DNA vorkommenden Heterocyclen, auch Nucleobasen genannt, werden mit den Abkürzungen A (Adenin), C (Cytosin), G (Guanin) und T (Thymin) bezeichnet, der zugrundeliegende Zucker, die 2'-Desoxyribose, mit d, und er hat die Numerierung 1' bis 5'. Das Formelbild von Abbildung 1 zeigt einen Ausschnitt aus dem Dodecamer, dessen Röntgenstrukturanalyse^[2] die Grundlage für unser derzeitiges Verständnis der B-DNA, eine der räumlichen Anordnungen von DNA, bildet. Bei der chemischen DNA-Synthese werden die beiden komplementären Stränge separat aufgebaut, d. h. es kann nur einzelsträngige DNA (Oligonucleotide) synthetisiert werden, die dann hybridisiert (Bildung von Wasserstoffbrückenbindungen) und durch enzymatische Verknüpfung zu größeren DNA-Einheiten verbunden wird, um Gene oder deren Regulationseinheiten herzustellen. Wenngleich es mehrere Möglichkeiten der Definition eines Gens gibt, so soll an dieser Stelle als Arbeitsdefinition^[3] gelten, Gene sind die DNA-Sequenzen, die für die Produktion von Polypeptiden verantwortlich sind. Mit Hilfe von DNA-Rekombinationstechniken ist es derzeit möglich, beliebige, sei es synthetische oder natürliche DNA, zu verknüpfen und zu vermehren und im zellulären Verband in Polypeptide umzusetzen.

Ausgehend von den chemischen Grundlagen der Oligonucleotidsynthese und deren Automatisierung beschreiben wir

die eigentliche chemoenzymatische Gensynthese^[4–6] und ihre Anwendungen, die das Tor zu völlig neuen Peptiden und Proteinen öffnet. Die Grundlagen für den Einsatz der hier geschilderten Methoden wurden in dieser Zeitschrift schon eingehend beschrieben^[7, 8].

2. Synthese und Anwendung von Oligonucleotiden

2.1. Strategien zur Synthese von Oligonucleotiden

Oligonucleotide sind Polymere, die durch Polykondensation von Nucleosidphosphaten entstehen. Hierbei ist die Syntheserichtung nach zwei Seiten – 3'-5'- und 5'-3'-Phosphatverknüpfung – möglich. Aufgrund der höheren Reaktivität der primären Hydroxygruppe hat sich überwiegend die Syntheserichtung 3'-5' bewährt, während die Natur, basierend auf den 5'-Triphosphaten, die umgekehrte Syntheserichtung benutzt.

Die Synthese eines langen Oligomers oder gar Polymers kann entweder durch schrittweise Addition einzelner Monomere oder durch die Verknüpfung oligomerer Bausteine erfolgen. Da DNA aus nur vier verschiedenen Nucleotiden besteht, können 16 dimere^[9–11] oder 64 trimere^[12, 13] Oligonucleotidblöcke (4ⁿ) hergestellt und für die Synthese jeder beliebigen DNA-Sequenzen als Bausteine verwendet werden. Beide Konzepte sind verwirklicht worden. Abzuschätzen ist der Syntheseaufwand für die Oligonucleotidblöcke auf der einen Seite und die Ausbeuten der Monomeradditionen auf der anderen Seite. Die Graphik (Abb. 2), welche die Kupplungsausbeute mit der Länge der DNA korreliert, zeigt überzeugend, daß bei einer Ausbeute von 99 % pro Stufe – die erstaunlicherweise für alle vier Nucleotide erreicht werden kann – ausreichend lange Stücke für die Gensynthese in akzeptablen Ausbeuten hergestellt werden können. In der Praxis der Festphasen-Synthese hat sich die schrittweise Addition monomerer Bausteine durchgesetzt.

2.2. Synthesemethoden für Oligonucleotide

Die Methoden zur Synthese definierter Oligonucleotide unterscheiden sich im wesentlichen durch die Art, wie die Phosphorsäureesterbindung erhalten wird. Je nach Art

[*] Prof. Dr. J. W. Engels
Institut für Organische Chemie der Universität
D-6000 Frankfurt am Main 50
Dr. E. Uhlmann
Hoechst AG
D-6230 Frankfurt am Main 80

5' C G A A T
5' _____

5' C G A A T
3' G C T T A
5' _____
3' _____

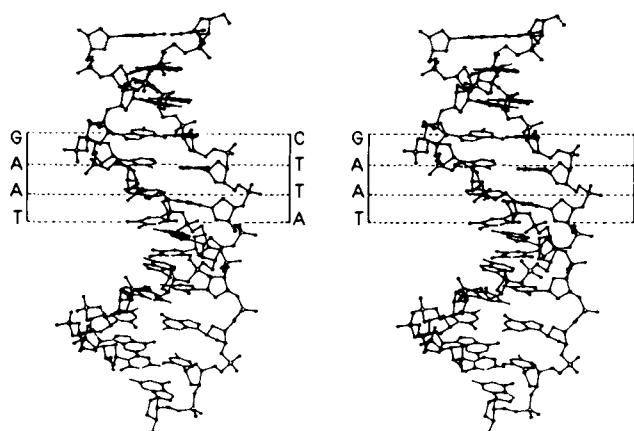
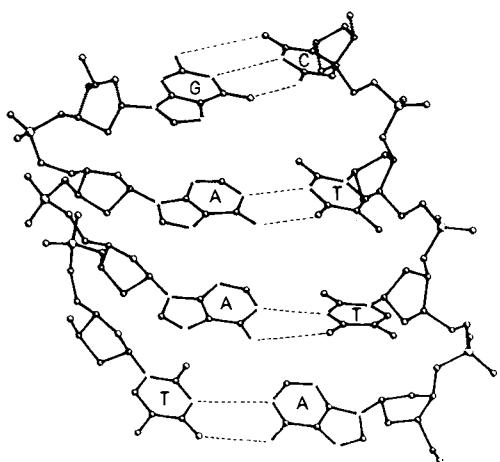
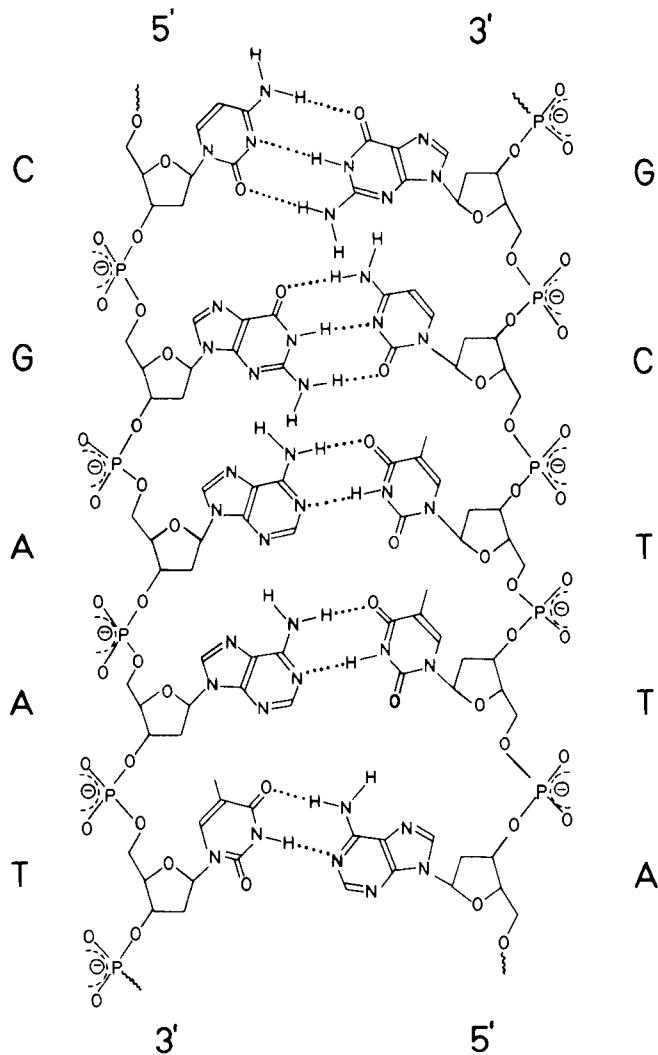
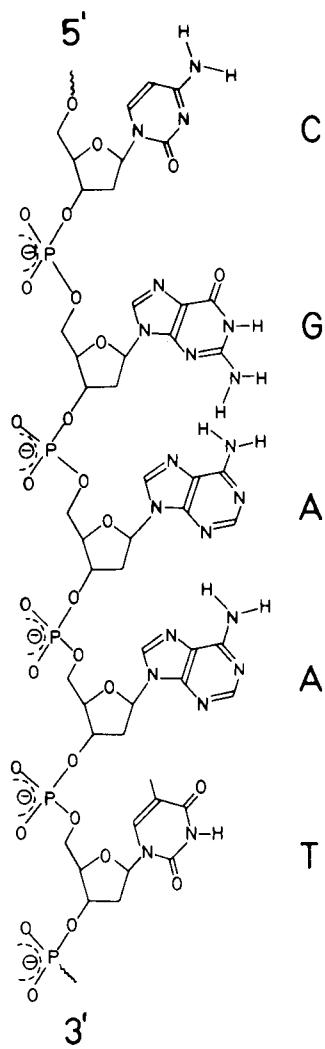


Abb. 1. Einzelsträngige und doppelsträngige B-DNA (Dodecamer) CGCGAATTTCGCG. Oben links: 5'CGAAT-Einzelstrang; oben rechts: idealisierte Watson-Crick-Basenpaarung; unten links: Ausschnitt GAAT unter Beibehalt der Basenpaarung; unten rechts: Stereobild des Dodecamers.

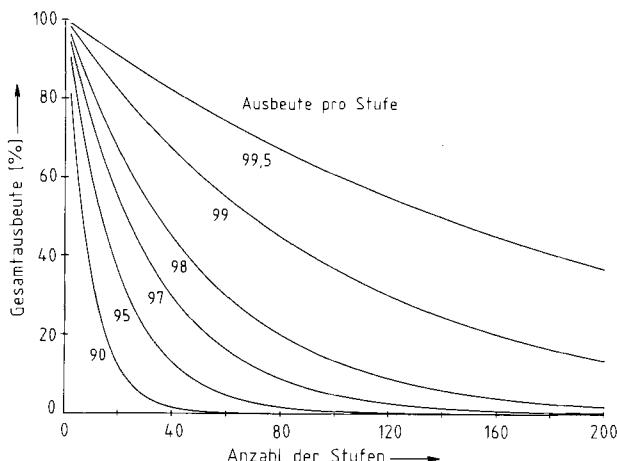


Abb. 2. Korrelation der Gesamtausbeute in Abhängigkeit von der Ausbeute pro Schritt mit der Anzahl der Reaktionsstufen, d.h. Länge der DNA.

der Phosphorkomponente spricht man von der Triester- (Abb. 3a), Phosphit- (Abb. 3b, $X = \text{Cl}$), Phosphoramid- (Abb. 3b, $X = \text{NiPr}_2$) und H-Phosphonatmethode (Abb. 3c). Als Ausgangsverbindung der Synthese fungiert das entsprechend geschützte Nucleosidderivat. Obgleich eine enorme Fülle von Schutzgruppen^[14] bekannt ist und auch für diese Chemie getestet wurde^[15-18], so haben sich doch

aufgrund der Automatisierung (und Kommerzialisierung) der DNA-Synthese einige wenige Schutzgruppen durchgesetzt. Es sind dies im wesentlichen die von *Khorana* et al.^[19] eingeführten Acylgruppen für die exocyclischen Aminofunktionen an Adenin, Guanin und Cytosin sowie die Dimethoxytritylgruppe (DMT) an der 5'-Hydroxygruppe der Desoxyribose. Sie lassen sich heute leicht und in guten Ausbeuten über die „transient protection“-Methode von *Ti* et al.^[20] einbauen. Durch intermediären Schutz der Zucker-OH-Gruppen als Trimethylsilylether lassen sich gezielt die Aminogruppen durch Acylierungsreagentien schützen. Zunehmend werden auch labilere Acyl-^[21] und Amidenschutzgruppen^[22, 23] eingesetzt, was besonders bei der Synthese von modifizierter DNA mit labilerer Funktionalität von Vorteil ist. Eine recht vollständige Aufstellung aller in der Oligonucleotidsynthese bis 1986 verwendeten Schutzgruppen gab *Sonveaux*^[18]. Speziell bei den Synthesemethoden mit sehr reaktiven Phosphorderivaten kann eine unerwünschte Derivatisierung der Amidfunktionen im Guanin und Thymin eintreten^[24-26]. Auch hier sind mehrere Schutzgruppen eingesetzt worden, die das Amid in den Iminoether überführen^[27-30]. Im Falle der Phosphoramiditmethode haben wir zeigen können, daß eine optimierte Syntheseführung der exzessiven Schutzgruppenverwendung vorzuziehen ist^[31].

Die geschützten Nucleoside werden dann mit den geeigneten Reagentien phosphoryliert. Diese können entweder mono- oder bifunktionell sein. Für die Synthese am festen Träger nach der Merrifield-Strategie hat sich jedoch bewährt, stabile monomere Nucleotidbausteine zu synthetisieren, die als lagerfähige Komponenten bei Bedarf in die Synthese eingebracht werden können. Als Schutzgruppen am Phosphor ist eine Vielzahl verschiedener Funktionen eingesetzt worden^[18], aber auch hier hat die Automatisierung eine Vereinfachung mit sich gebracht. Ein vom chemisch-mechanistischen Verständnis erwähnenswerter Gesichtspunkt ist die Art, in der die Schutzgruppe vom Phosphor abgespalten wird. Reaktionen am Phosphat-Phosphor sind seit den Untersuchungen von *Westheimer* et al.^[32] dafür bekannt, daß sie über Additions-Eliminierungs-Mechanismen unter möglicher Beteiligung von Umorientierungen, wie Pseudorotationen ablaufen. Um diese Problematik – eventuell uneinheitliche Hydrolysen – zu umgehen, werden Schutzgruppen bevorzugt, bei denen der Angriff des Abspaltungsreagens direkt an der Schutzgruppe ohne Beeinträchtigung des Phosphatzentrums stattfindet^[33]. Hierbei wurden sowohl Alkylreste wie Methyl und Benzyl als auch Reste wie β -Cyanethyl, Arylsulfonylethyl, Arylethyl und Halogenethyl, die durch β -Eliminierung entfernt werden können, benutzt. Durch diese Strategie ist eine Isomerisierung der Phosphatbindung weitestgehend ausgeschlossen und Strangbrüche werden vermieden.

2.2.1. Die Triestermethode

Seit ihrer Einführung durch *Todd* et al.^[34] basiert diese Synthesemethode (Abb. 3a) auf der geeigneten Aktivierung eines geschützten Nucleosidphosphorsäurediesters mit Säurederivaten vom Typ der Arylsulfonsäureazolide. Die so aktivierte Phosphorsäurediester reagieren dann mit den 5'-OH-ungeschützten Nucleosiden zu den entsprechenden Dinucleosidphosphorsäuretriestern. Eine Vielzahl solcher

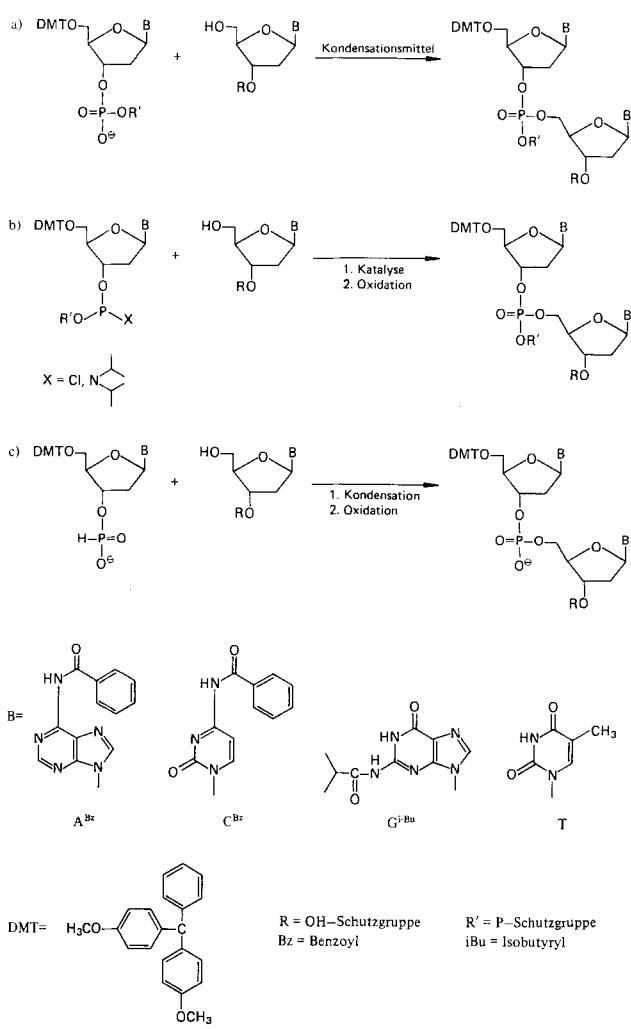


Abb. 3. Drei Methoden zur Oligonukleotidsynthese: a) Triester-, b) Amidit- und c) H-Phosphonatmethode.

Reagentien, wie Phosphorsäure- und Sulfonsäurederivate^[35–38], wurden eingesetzt. Bei den Azoliden haben sich Tetrazol und 3-Nitrotriazol besonders bewährt. Die geringere Reaktivität der P^V-Verbindungen kann so genügend ge steigert werden. Als entscheidender Vorteil dieser Methode können sicher die gute Stabilität der Ausgangsdiester sowie deren potentielle Rückgewinnung (z. B. als Barium-Salze)^[39] angesehen werden. Ein Nachteil sind die Nebenreaktionen, die durch die Kondensationsmittel hervorgerufen werden, z. B. Sulfonierung der 5'-OH-Position.

2.2.2. Phosphoramiditmethode

Hierbei wird ein geeignet geschütztes Nucleosidphosphoramidit (Abb. 3b) in Gegenwart von Tetrazol mit einem 5'-OH-freien Nucleosid zu einem Phosphitriester umgesetzt. Nachfolgende Oxidation führt zum analogen Triester. Die höhere Reaktivität der P^{III}-Verbindungen hat diese schon früh als attraktive Edukte erscheinen lassen^[40]. Für eine Routine-Oligonucleotidsynthese stand ihre schlechte Handhabbarkeit und Zugänglichkeit im Weg, insbesondere die geringe Lagerfähigkeit der monomeren Bausteine. Mit der Einführung der durch Säure aktivierbaren Phosphoramidite durch Caruthers et al.^[41, 42] gelang hier der Durchbruch. Während zunächst die Wahl auf die Dimethylaminogruppe als Amiditkomponente und die Methoxyverbindung^[43] als Schutzgruppe fiel, wird heute beim weiteren Ausfeilen der optimalen Reaktionsbedingungen die Diisopropylaminoverbindung^[44] als der akzeptierte Kompromiß in Verbindung mit der β -Cyanethylschutzgruppe^[45] für den Phosphatschutz angesehen. Wichtigster Vorteil dieser Methode ist die nahezu nebenreaktionsfreie Aktivierung durch das milde Reagens Tetrazol.

2.2.3. H-Phosphonatmethode

Todd et al.^[46] hatten schon früh das Synthesepotential der H-Phosphonate erkannt und genutzt. Hierbei wird ein geschütztes Nucleosid-H-phosphonat (Abb. 3c) in Gegenwart eines Kondensationsmittels, meist eines Säurechlorids^[47, 48], mit einem 5'-OH-Nucleosid zu einem Dinucleosid-H-phosphonsäurediester umgesetzt. Anschließende Oxidation führt zum analogen Diester. Als Vorteile sind hier zu erwähnen, daß keine intermediäre Schutzgruppe am Phosphor vonnöten ist und daß die Oxidation der H-Phosphonate erst am Ende der Synthese stattfindet, nicht wie bei der Amiditmethode nach jeder Addition. Andererseits ist hier, wie auch bei der Triestermethode, durch das Kondensationsmittel (z. B. Pivaloylchlorid) mit Nebenreaktionen (Acylierungen) zu rechnen. Es zeichnet sich ab, daß die Ausbeuten der H-Phosphonatmethode, trotz weiterer Optimierung, wie etwa durch den Einsatz von Adamantoylchlorid^[49] als Kondensationsmittel, mit denen der Phosphoramiditmethode nicht ganz konkurrieren können.

2.3. Oligonucleotidsynthesen an festen Trägern

Da es sich bei der Oligonucleotidsynthese um eine repetitive Technik handelt, hat sich die Synthese am festen Träger gegenüber der Synthese in Lösung oder der Polykonden-

sation^[50] durchgesetzt. Seit der Einführung der Festphasensynthese durch Merrifield^[51] und Letsinger^[52, 53] ist folgendes grundlegende Konzept beibehalten worden: 1) Die polymere Matrix soll im verwendeten Lösungsmittel unlöslich und möglichst inert sein. 2) Funktionelle Gruppen am polymeren Träger sollen die kovalente Anknüpfung des zu synthetisierenden Produktes ermöglichen. 3) Die Struktur des Polymers soll die chemische Reaktion nicht zu sehr beeinträchtigen, z. B. durch Behinderung der Diffusion der reagierenden Moleküle.

Eine Vielzahl von Materialien wurde seitdem getestet, Polystyrol^[54], Polyacrylamid^[55], Cellulose^[56], Polytetrafluorethylen^[57] und Kieselgel^[9, 58, 59]. Letzteres und speziell definierte Glasperlen, sogenanntes controlled-pore-glass-(CPG)-Material^[44, 60, 61], finden heute die weiteste Verwendung. Die Verbindung zwischen dem festen Träger und dem Startnucleosid wird durch einen Spacer (am einfachsten ein Alkylamin) erreicht (Abb. 4). Da die Synthese in der 3'-5'-Richtung abläuft, wird das Startnucleosid über seine 3'-Hydroxyfunktion und eine Dicarbonsäure (z. B. Bernsteinsäure)^[59] mit dem Spacer verknüpft. Derzeit werden Träger mit den vier Nucleosiden A, G, C, T oder universelle Anker^[62] eingesetzt.

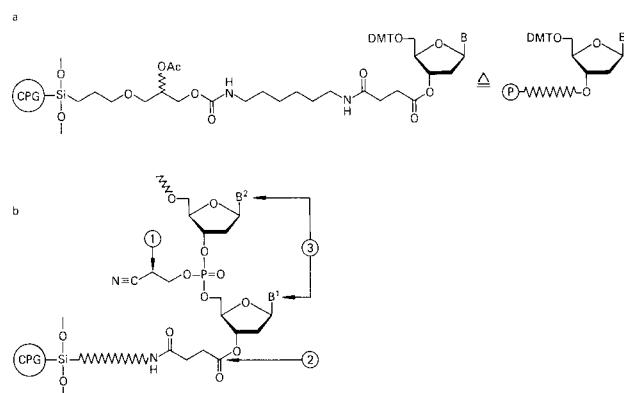


Abb. 4. a) Fixierung des Startnucleosids am festen Träger; b) Abspaltung 1. der Schutzgruppe am Phosphor, 2. des Trägers und 3. der Basenschutzgruppen.

Von den geforderten Voraussetzungen zur Festphasensynthese sind Punkt 1) und 2) gut einzuhalten. Die dritte Forderung, daß der feste Träger möglichst keinen Einfluß auf das Wachstum der Kette haben sollte, bereitet jedoch Probleme. Bei den derzeit erfolgreich verwendeten CPG-Materialien ist ein sehr guter Kompromiß gefunden worden. Interessant ist die Beobachtung, daß beim Einsatz von 500 Å-CPG-Material ein Abfall der individuellen Kupplungsausbeute auftritt, sobald eine Länge von 80 Basen überschritten wird. Ein 1000 Å-Material indes läßt Synthesen mit einer Kettenlänge von bis zu 200 Basen zu^[63]. Das Trägermaterial scheint eine entscheidende Limitierung der Methode zu sein.

Zu erwähnen an dieser Stelle ist auch die Taktik, Synthesen simultan durchzuführen. Für die Festphasensynthese sind bereits Anfang 1980 Automaten eingeführt worden. Damit können ein oder mehrere (drei oder vier) Oligonucleotide gleichzeitig synthetisiert werden. Da prinzipiell nur vier verschiedene Reaktanten (A, G, C, T) bei der Synthese natürlicher DNA infrage kommen, hat sich als Alternative dazu die Filterblättchen-Methode^[64–66] (20 bis 100 Oligonucleotide)

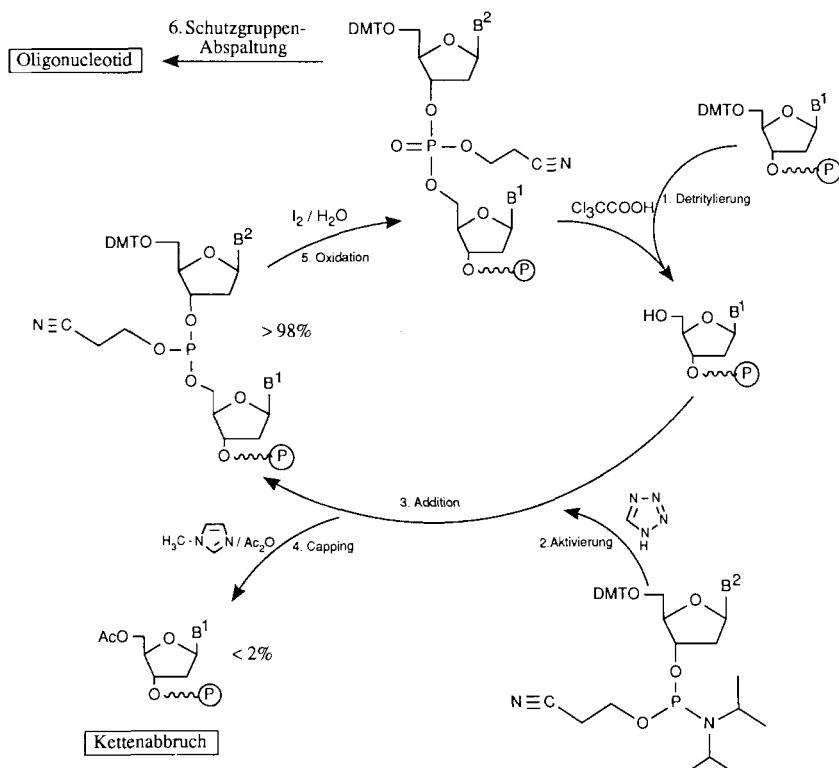


Abb. 5. Synthesecyclus der Phosphoramiditmethode.

bewährt. Das Handicap dieser Variante sind etwas geringere Ausbeuten und Schwierigkeiten bei der Automatisierung.

Die Automatisierung war der entscheidende Schritt, um die Oligonucleotidsynthese breit anwenden zu können. So wurden ausgetüftelte Syntheseprotokolle zu dem in Abbildung 5 schematisch gezeigten Reaktionszyklus erstellt. Die wesentlichen Schritte eines solchen Cyclus sind: 1) Abspaltung der Schutzgruppe von der 5'-OH-Position durch Säure; 2) Aktivierung der Nucleotidkomponente mit Tetrazol; 3) Addition des Nucleotids; 4) „Capping“ nicht abreagierter 5'-OH-Gruppen; 5) Oxidation (im Falle von P^{III} -Verbindungen).

Zwischen diesen entscheidenden chemischen Schritten finden mehrere Spülungen statt, die zum Auswaschen überschüssiger Reagentien und zum Äquilibrieren der jeweiligen Reaktionen dienen. Während die Schritte 1)–3) allgemeine Routine aller Methoden sind, ist die Notwendigkeit von Schritt 4), das Capping nicht abreagierter 5'-OH-Gruppen, umstritten. Obwohl die einzelnen Additionen nahezu quantitativ sind, hat sich dieser Schritt zur Vermeidung von Fehlsequenzen sehr bewährt. Nach vollendeter Synthese werden die am festen Träger synthetisierten Oligonucleotide direkt im Automaten von ihren Schutzgruppen befreit (Schritt 6 in Abb. 5). Demzufolge wurden auch hier die Abspaltungsbedingungen vereinfacht. Im Falle der Phosphoramiditmethode – die wir bevorzugt verwenden – wird mit konzentrierter wässriger Ammoniaklösung abgespalten. Das so erhaltene Nucleotidgemisch wird dann durch préparative Gelelektrophorese oder Hochdruckflüssigkeitschromatographie (HPLC)^[67] gereinigt.

2.4. Anwendung synthetischer Oligonucleotide

Oligodesoxynukleotide können aufgrund des von *Watson* und *Crick* formulierten Komplementaritätsprinzips sowohl

mit DNA als auch mit RNA hybridisieren (Duplex-Bildung). Der Duplex mit der exakten Basenpaarkomplementarität hat den höchsten Schmelzpunkt. Somit ist bei gradueller Hybridisierung eine äußerst spezifische Analytik durchzuführen. Dieses Prinzip scheint sich besonders erfolgversprechend in der Diagnose von Gendefekten, der Genom-Analytik^[68], zu bewähren. Diese Aspekte werden hier nicht weiter behandelt, vielmehr wird der Einsatz der synthetisierten Oligonucleotide als Bausteine für die Gensynthese im Detail besprochen.

3. Konstruktion synthetischer Gene

3.1. Historische Entwicklung der Gensynthese

Bevor die Methoden zum Design und zur Synthese von Genen im Detail diskutiert werden, soll anhand einiger instruktiver Beispiele die enorme Entwicklung der DNA-Synthese in den letzten zwei Jahrzehnten veranschaulicht werden (Tabelle 1). Bereits in den sechziger und frühen siebziger Jahren vollbrachten *Khorana* et al. entscheidende Pionierarbeiten, indem sie die Gene für die Alanin-spezifische Transfer-RNA (tRNA) aus Hefe^[69] und für die Tyrosin-Suppressor-tRNA aus *E. coli*^[70–82] synthetisierten. Hierbei wurden grundlegende Methoden und Strategien entwickelt, die lange Zeit als beispielhaft galten und bis heute zur Synthese zahlreicher Gene herangezogen wurden.

Trotz der aufsehenerregenden Arbeiten und außergewöhnlichen Erfolge *Khoranas* war die Zeit noch nicht reif für eine konsequente Anwendung synthetischer Gene, da viele der wichtigen molekularbiologischen Hilfsmittel und Techniken (z. B. Restriktionsenzyme, DNA-Sequenzierung) erst später gefunden oder entwickelt wurden. Der eigentliche

Tabelle 1. Historische Entwicklungen in der Gensynthese.

Gen für	Größe des Gens [bp]	Anzahl der Oligonucleotide	Synthesemethode [a]	Länge der synthetisierten Oligonucleotide	Länge des Überlapps	Publikationsjahr	Lit.
Alanin-spez. tRNA Hefe	77	15	Diester	5–20	4–7	1970	[69]
Angiotensin II	33	7	Diester	6–11	5–6	1975	[83]
Tyrosin-Suppressor-tRNA aus <i>E. coli</i>	126	26	Diester	6–13	4–7	1976	[70–82]
Somatostatin	56	8	Triester	11–16	5–8	1977	[84]
Insulin							
B-Kette	104	18	Triester-blöcke	10–12	4–6	1978	[85]
A-Kette	77	12	Triester/Fp	10–15	5–8		
Interferon- α epidermaler	514	67	Triester/Fp	14–21	7–10	1981	[87]
Wachstumsfaktor	170	10	Amidit/Fp	11–59	7–24	1983	[97]
Interferon- γ	453	66	Chloridit/Fp	9–14	5–8	1984	[116]
Interferon- γ	452	34	Amidit/Fp	18–33	6 [b]	1984	[115]
Wachstumshormon	584	78	Triester/Fp	7–26	4–10	1984	[89]
Interleukin-2-Rezeptor	753	38	Amidit/Fp/A	27–51	6–12 [b]	1986	[90]
Rinderrhodopsin	1057	72	Amidit/Fp/A	15–40	4–5 [b]	1986	[168]
Rinderprochymosin [d]	1108	28		59–102	4 [b]		
	1126	26	Amidit/Fp/A	72–102	7–8 [b]	1987	[92]
HMG-CoA-Reduktase	1400	19	Amidit/Fp/A	86–147	[c]	1988	[93]
t-Plasminogen-Aktivator	1610	101	Amidit/Fp/A	17–37	4–20	1988	[91]

[a] Fp = Festphase; A = DNA-Syntheseautomat. [b] Vorgeformte Duplex-DNA. [c] Nicht durch Ligieren der Oligonucleotide. [d] Zwei Synthesen ausgeführt.

Durchbruch gelang erst etwa eine Dekade später mit der Synthese und Expression von Strukturgenen, d.h. von Genen, die wertvolle Peptide oder Proteine codieren. Unter Expression eines Gens versteht man dessen Transkription in eine Boten-RNA (mRNA) und die anschließende Translation in das jeweilige Polypeptid.

Obwohl Köster et al.^[83] bereits im Jahre 1975 die erste Totalsynthese eines Strukturgens gelang – es codierte das Octapeptid Angiotensin II –, ließ der Durchbruch synthetischer Gene noch einige Jahre auf sich warten, da dieses Gen seinerzeit nicht exprimiert werden konnte. Wenn auch der Weg zur Expression eines synthetischen Gens vorgezeichnet war, so war es zweifellos einer der Höhepunkte der Gentechnik überhaupt, als im Jahre 1977 das erste synthetische Gen von Itakura et al.^[84] in Bakterien zur Expression gebracht werden konnte. Das so gewonnene Tetradecapeptid Somatostatin war zudem das erste Peptidhormon, das nach DNA-Rekombination in *E. coli* nachgewiesen werden konnte. Im folgenden Jahr berichtet dieselbe Arbeitsgruppe in einer weiteren spektakulären Arbeit über die erfolgreiche Synthese von Genen für die Insulin-A- und Insulin-B-Ketten^[85], die später in *E. coli* exprimiert wurden^[86].

Ein weiterer Meilenstein in der Historie der Gensynthese war die Totalsynthese des Gens für Leukocyten-Interferon (IFN- α) im Jahre 1981^[87]. Beeindruckend war einerseits die Länge des Gens von 514 Basenpaaren (bp), das 166 Aminosäuren codiert; zum anderen war es das erste Mal, daß ein synthetisches Gen nicht aus der Peptidsequenz, sondern vielmehr aus der publizierten Nucleotidsequenz (cDNA)^[88] abgeleitet wurde. Zuvor waren die Gene noch manuell und in Lösung hergestellt worden, eine Verfahrensweise, die für ein Gen solcher Größe nicht praktikabel gewesen wäre. Erst durch die Anwendung der Festphasentechnik ist die Synthese größerer Gene, etwa der Länge des Interferon-Gens, Realität geworden. Mit der Einführung der Technik des sequentiellen Klonierens, d.h. der Aufteilung eines Gens in Subfragmente, die getrennt kloniert und erst nach Verifizierung der Teilsequenzen zum Gesamtgen zusammengesetzt wer-

den, gelang die Konstruktion immer längerer Gene. 1984 erfolgte beispielsweise die Herstellung eines Gens für das humane Wachstumshormon (192 Aminosäuren), das aus 78 Oligonucleotiden über zwei Subfragmente zusammengesetzt wurde^[89]. Die nach 1984 publizierten Gensynthesen wurden größtenteils mit Hilfe von DNA-Syntheseautomaten durchgeführt.

Das erste Beispiel eines Gens für ein Rezeptorprotein wurde 1986 vorgestellt^[90]. Es besteht aus 753 bp und konnte aus nur 38 synthetischen Oligonucleotiden über drei Subklone hergestellt werden. Das längste synthetische Gen, das wir heute kennen, umfaßt 1610 bp und codiert für den humanen t-Plasminogen-Aktivator (h-tPA)^[91]. Dieses Gen erforderte die Synthese von 101 Oligonucleotiden mit einer durchschnittlichen Länge von 32 Nucleotideinheiten. Bei einer Cycluszeit von 12.5 min pro Addition eines Nucleotids errechnet sich somit eine Gesamtsynthesezeit von etwa 30 Tagen für das tPA-Gen. Die Hauptarbeit war das Zusammenfügen der 101 Oligonucleotide zum Gesamtgen^[91].

Indirekt spiegelt Tabelle 1 den rasanten Fortschritt in der Oligonucleotidsynthese der letzten Jahre wider, welcher in der wachsenden Kettenlänge der verwendeten Oligonucleotide zum Ausdruck kommt. Wenn wir einmal vom tPA-Gen absehen, so geht der Trend eindeutig hin zur Verwendung längerer Oligonucleotide^[92, 93], welche das Zusammensetzen des Gens vereinfachen und zudem die Entwicklung neuer Strategien beim Genaufbau zulassen. Es läßt sich unschwer vorhersagen, daß sich in Zukunft noch längere synthetische Gene mit einem überschaubaren und zu rechtfertigenden Aufwand herstellen lassen.

Bis heute wurden über 150 synthetische Gene hergestellt, die unter anderem für Hormone^[83–85, 89, 94–113], Interferone und Lymphokine^[87, 114–126] sowie andere immunologisch wichtige Proteine^[90, 127–130], Wachstumsfaktoren^[131–140], Enzyme^[92, 93, 141–152], Neuropeptide^[153–157], calciumbindende Proteine^[158, 159], Gastrointestinal-Peptide^[160], Proteinase-Inhibitoren^[161–167] und andere Produkte codieren^[69, 82, 91, 168–188].

3.2. Wege zum Aufbau doppelsträngiger DNA

3.2.1. Grundlegende enzymatische Reaktionen: DNA-Ligase- und DNA-Polymerase-Reaktion

Die genetische Information ist in Form der Nucleotidtriplets im codierenden Strang der DNA und indirekt im nichtcodierenden Strang niedergelegt. Die Doppelhelix-Struktur der DNA (siehe Abb. 1) dient zum einen dazu, die DNA zu stabilisieren und zum anderen ist sie eine Voraussetzung für die Replikation. Die natürliche Fähigkeit von einzelsträngigen Oligo- oder Polynucleotiden, sich aufgrund der Basenpaarung in definierter Art und Weise aneinanderzulagern, lässt sich zum Aufbau doppelsträngiger DNA aus geeigneten chemisch synthetisierten Oligonucleotiden ausnutzen.

Prinzipiell unterscheiden wir zwei Typen enzymatischer Reaktionen, die zum Aufbau des DNA-Doppelstranges herangezogen werden können: die DNA-Ligase- und die DNA-Polymerase-Reaktion. Das Enzym T4-DNA-Ligase katalysiert die Bildung von Phosphodiesterbindungen zwischen benachbarten 5'-Phosphat- und 3'-Hydroxy-Enden in einem DNA-Duplex in Gegenwart von ATP (Abb. 6). Ein-

glatte Ende ligiert werden, wird in der Gensynthese normalerweise nicht verwendet, da hier die Selektivität der Reaktion aufgrund der fehlenden Basenpaarung niedrig ist und diese sogenannten „blunt end“-Ligationen zudem viel weniger effizient sind als „sticky end“-Ligationen. Falls die zu verknüpfenden Oligonucleotide nicht bereits chemisch als 5'-Phosphate synthetisiert wurden^[189, 190], ist eine Phosphorylierung der 5'-Hydroxy-Oligonucleotide mit Polynucleotid-Kinase und ATP erforderlich. In diesem enzymatischen Schritt können die Oligonucleotide bei Verwendung von $[\gamma-32\text{P}]$ ATP radioaktiv markiert werden. Dies kann bei der Ligation subnanomolarer Mengen von Vorteil sein. Das Enzym DNA-Polymerase I (Klenow-Fragment), dem die 5'-3'-Exonucleaseaktivität des Gesamtenzyms fehlt, kann in Gegenwart der Desoxynucleosid-5'-triphosphate zu einem einzelsträngigen Polynucleotid (Templat) beginnend am 3'-Hydroxy-Ende eines anhybridisierten kurzen Oligonucleotids (Primer) eine komplementäre Kopie synthetisieren (Abb. 6 b). Prinzipiell kann man alle die DNA-Polymerasen aus anderen Organismen für die „fill in“-Reaktion verwenden, die man auch bei der DNA-Sequenzierung nach Sanger^[202] einsetzt.

3.2.2. Strategien zur Gensynthese

Retrosynthetisch kann ein zu synthetisierendes Gen auf den in Abbildung 7 gezeigten Wegen A)–E) in einzelsträngige Oligonucleotide zerlegt werden, deren Anzahl von der Länge dieses Oligonucleotide abhängt.

Weg A): Ligation sich überlappender Oligonucleotide nach Khorana

Das klassische Verfahren der Ligation nach Khorana^[69, 191] ist Weg A). Es basiert auf dem Ligationstyp A von

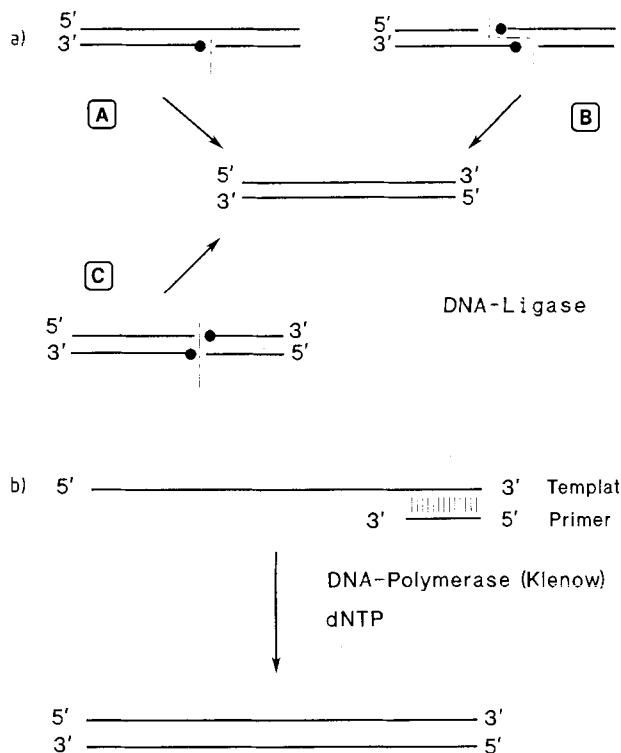


Abb. 6. Grundlegende enzymatische Reaktionen zum Aufbau doppelsträngiger DNA aus Oligonucleotiden: a) DNA-Ligase-Reaktion; b) DNA-Polymerase-Reaktion.

zelsträngige Oligonucleotide sind dagegen keine Substrate der DNA-Ligasen. Im Reaktionstyp A werden die zu verknüpfenden Oligonucleotide durch einen intakten Bereich im komplementären Oligonucleotid zusammengehalten. Dagegen werden im Reaktionstyp B die überlappenden Enden in beiden Strängen verknüpft. Ähnlich wie bei den DNA-Fragmenten, die durch Restriktionsenzym-Spaltung entstehen, können die Enden 5'-überhängend oder aber 3'-überhängend ligiert werden. Die Reaktionsvariante C, bei der

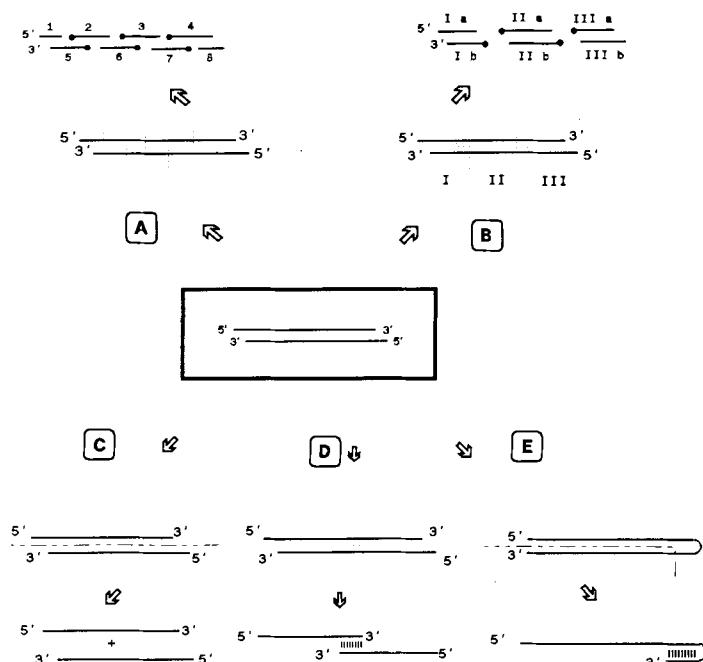


Abb. 7. Strategien zur Synthese eines Gens: A) Ligationsschema nach Khorana; B) Sequentielle Ligation präformierter DNA-Duplices; C) Direkte Hybridisierung und Klonierung langer Oligonucleotide; D) „Fill in“-Methode überlappender Oligonucleotide; E) „Fill in“-Methode „autoprimerender“ Oligonucleotide.

Abbildung 6a, wobei sich die zu schließenden Lücken alternierend im oberen und unteren Strang befinden. Diese Strategie erfordert die Synthese eines Satzes sich überlappender Oligonucleotide (hier beispielsweise acht Stücke) und deren Ligation zum Gesamtgen oder Genfragment, wobei nur die internen Oligonucleotide (2–7) in Form ihrer 5'-Phosphate vorliegen müssen. Durch Erhitzen einer Lösung aller Oligonucleotide und langsames Abkühlen erreicht man aufgrund der Basenkomplementarität das korrekte Aneinanderlagern („Annealing“) dieser Oligonucleotide unter Bildung des noch nicht kovalent verknüpften Hybridmoleküls. In Gegenwart von ATP knüpft dann die DNA-Ligase die gewünschten Phosphodiesterbindungen. Auf diese Weise konnten beispielsweise 20 Oligonucleotide in einer einzigen Ligase-Reaktion zu einem 280 bp-Strukturen verknüpft werden^[131]. In unseren Händen erwies sich jedoch eine Anzahl von sechs bis zehn Oligonucleotiden als optimal hinsichtlich der Ausbeute und Größe des Gens. Allgemein gilt, daß die Ausbeute an komplett ligierten Produkten mit steigender Anzahl der in einer Reaktion zu verknüpfenden Oligonucleotide abfällt^[176]. Die Verwendung langer Oligonucleotide mit etwa 40 bis 60 Basen führte zu einer Erhöhung der Effizienz der enzymatischen Ligationen^[192].

Beispielhaft soll an dieser Stelle die Synthese eines Gens für die Tyrosin-Suppressor-tRNA aus *E. coli* erläutert werden, über die Khorana et al. im Jahre 1976 in zwölf aufeinanderfolgenden Publikationen^[70–81] berichtet haben. Das 126 bp umfassende Gen wurde zunächst in vier Fragmente, diese wiederum in insgesamt 26 Oligonucleotide unterteilt. Die Aufteilung erfolgte so, daß sich mindestens vier bis sechs Basenpaare zwischen den einzelnen Oligonucleotiden überlappen konnten. Ein weiterer Gesichtspunkt war die Ökonomie der Synthese, beispielsweise die mehrfache Verwendbarkeit von Syntheseblöcken. Die Totalsynthese des Gens erforderte also die chemische Synthese von 26 Oligonucleotiden nach der Phosphodiestermethode, was zu jener Zeit der arbeitsintensivste Teil war. Diese Oligonucleotide wurden dann in Gruppen zu vier Fragmenten ligiert, welche nachfolgend zum Gesamtgen verknüpft werden konnten. Dabei erwies sich die Verknüpfung der Fragmente als schneller und effizienter im Vergleich zur vorangegangenen Ligation der einzelsträngigen Oligonucleotide. Dieser Tatsache wurde in späteren Entwicklungen (vgl. Strategie B) Rechnung getragen. Auch heute noch ist das Ligationsverfahren nach Khorana das gängigste Konzept, insbesondere für Oligonucleotide zwischen 15 und 50 Einheiten. Da die Filterblättchen-Methode^[64–66] hauptsächlich in diesem Längenbereich verwendet wird, und mit ihr leicht eine große Anzahl von Oligonucleotiden simultan synthetisiert werden kann, ist die Ligationsstrategie A) hier die Methode der Wahl. Größere Gene wird man heutzutage in mehrere Subklone (Unterfragmente) aufteilen^[115] und nicht, wie im Falle des Leukozyten-Interferon-Gens^[87], in einer Ligationskaskade ohne intermediäre Klonierung aufbauen. Nach dieser Strategie wurde unter anderem das 1610 bp enthaltende Gen für tPA^[91] aus 101 Oligonucleotiden über drei Subklone synthetisiert.

Weg B): Sequentielle Ligation präformierter DNA-Duplices

Ähnlich wie beim Schneiden von DNA mit Restriktionsenzymen läßt sich ein Gen auch retrosynthetisch in kleinere DNA-Fragmente aufteilen. Dieser Ansatz setzt jedoch eine

gewisse Kettenlänge (etwa 30 bis 80 Nucleotide) der synthetischen Oligonucleotide voraus, damit stabile DNA-Duplices gebildet werden. Strategie B) hat gegenüber A) den Vorteil, daß die Oligonucleotide paarweise zu kleinen DNA-Doppelsträngen hybridisiert werden, welche dann sequentiell oder in einem Schritt in guter Ausbeute zum Gesamtgen ligiert werden können. Damit werden unerwünschte Hybridisierungen (beispielsweise Ia mit IIb in Abb. 7) ausgeschlossen. Die DNA-Doppelstränge lassen sich zudem leicht über Agarose-Gele (2–4 %) reinigen. Die überlappenden Enden können nach Belieben 5'- oder 3'-überhängend geplant werden. Meistens werden Überlappungen von vier bis acht Basenpaaren gewählt, jedoch sind auch Überlappungen bis zu zwölf Basenpaaren^[90] erfolgreich verwendet worden. Obwohl längere Überlappungen zu stabileren Hybriden führen, handelt man sich mit langen Überhängen potentiell wieder Sekundärstrukturprobleme (vgl. Abschnitt 3.4.1) und falsche intermolekulare Wechselwirkungen ein, so daß hier ein Kompromiß zwischen den gegenläufigen Effekten zu suchen ist.

Weg C): Direkte Hybridisierung und Klonierung langer Oligonucleotide

Die Aufteilung eines Gens in den codierenden und nicht-codierenden Strang ist der denkbar einfachste Weg zur Totalsynthese eines Gens, der in der Vergangenheit jedoch nicht praktikabel war, da die chemische Synthese keine Oligonucleotide ausreichender Länge liefern konnte. Heute ist jedoch die Synthese von Oligonucleotiden mit einer Länge bis zu etwa 200 Einheiten möglich. Dies bedeutet, daß durch die Synthese von nur zwei Oligonucleotiden und deren Hybridisierung Gene oder Genfragmente zugänglich sind, die bis zu 65 Aminosäuren codieren. Im letzten Jahr wurde erstmals über die Synthese und Klonierung eines Calcitonin-Gens nach diesem Verfahren berichtet^[109]. Der Vorteil dieser Strategie besteht darin, daß die manuell zu verrichtende Arbeit beträchtlich reduziert ist, da keine enzymatische Ligation notwendig ist, die chemische Synthese vom Synthesearmat ausgeführt wird und die Reinigung auf zwei Oligonucleotide beschränkt ist. Unter Anwendung der sequentiellen Klonierung (Subklonierung) von Genfragmenten stellt diese Strategie eine attraktive Methode zur Synthese auch größerer Gene dar. Im Prinzip ist diese Methodik eine konsequente Weiterentwicklung der Präformation von DNA-Duplices (Strategie B). In Abhängigkeit von der Kettenlänge der verwendeten Oligonucleotide und von der Festlegung der retrosynthetischen Schnittpositionen gibt es eigentlich einen fließenden Übergang von der Strategie A) über B) nach C). So betrachtet, ist die Subklonierung nichts anderes als das Verlagern der sequentiellen Ligation präformierter DNA-Duplices in ein Plasmid, jedoch mit dem außerordentlichen Vorteil, daß die Sequenz der synthetischen Fragmente intermediär verifiziert werden kann und zudem eine unerschöpfliche Ressource für das jeweilige Genfragment vorliegt. Eine Kombination der Strategien B) und C) finden wir in der Synthese des etwa 1100 bp großen Gens für Rinder-Prochymosin wieder^[92].

Weg D): „Fill in“-Methode überlappender Oligonucleotide

Das sogenannte „fill in“-Verfahren, bei dem man sich der DNA-Polymerase bedient, wurde bereits 1968 von Khor-

na^[191] weitblickend vorgeschlagen, ist jedoch erst 1982 von Rossi et al.^[193] realisiert worden. Es beruht auf der Synthese von zwei längeren Oligonucleotiden, die an ihrem 3'-Ende eine perfekt basenpaarende Sequenz gemeinsam haben. Nach der korrekten Hybridisierung dieser Oligonucleotide kann die in Abbildung 6b skizzierte DNA-Polymerase(Klenow)-Reaktion zur Anwendung kommen, wobei jeweils das 3'-Ende des einen Stranges als Primer zur templatabhängigen Synthese einer Kopie des anderen Stranges dient. Auf diese Art und Weise kann man etwa 40% der chemischen Synthesearbeit sparen. Trotzdem konnte sich diese Strategie bisher nicht wirklich durchsetzen; nur wenige Beispiele sind bekannt^[161, 162, 178, 194–196]. Die Größe der Genfragmente betrug dabei etwa 60 bis 100 bp. Bei unseren eigenen Arbeiten bei der Synthese eines Genfragmentes für Interferon- γ ^[3] stießen wir auf folgende Probleme: 1) Der Einbau des letzten Nucleotids am 5'-Ende bereitet manchmal Schwierigkeiten. Deshalb ist es ratsam, stets zwei bis drei Nucleotide mehr zu synthetisieren, die dann durch ein geeignetes Restriktionsenzym wieder abgespalten werden können. Vorteilhaft verwendet man dabei solche Restriktionsenzyme, die beim Schneiden überhängende Enden erzeugen, da diese im Vergleich zu glatten Enden besser ligierbar sind. 2) Die Qualität der Produkte war selten mit der aus Ligase-Reaktionen erhaltenen DNA vergleichbar. Wir sind deshalb der Auffassung, daß die Bildung unerwünschter Sekundärstrukturen das Hauptproblem dieser Synthesestrategie ist, besonders dann, wenn sehr lange Oligonucleotide aus ökonomischen Gründen zur Anwendung kommen sollen.

Weg E): „Fill in“-Methode mit Hilfe von Autoprimer-Oligonucleotiden

Eine unmittelbare Folge dieser Nachteile war die Einführung einer neuen Methode, welche ebenfalls von der an sich ökonomischen DNA-Polymerase-Reaktion Gebrauch macht^[197, 198]. Diese Strategie basiert auf der Verwendung langer Oligonucleotide, an deren 3'-Ende eine kurze Sequenz zur Bildung einer Haarnadelstruktur geeignet ist; die außergewöhnlich stabile Haarnadelstruktur^[198–201] dient dann als Primer in der DNA-Polymerase-katalysierten Bildung des komplementären Stranges. Diesen Typus von Oligonucleotid charakterisieren wir folglich als „Autoprimer“. Nach der „fill in“-Reaktion mit den Desoxynucleotidtriphosphaten, erzeugt man durch Spaltung mit geeigneten Restriktionsenzymen die gut ligierbaren überhängenden Enden, wobei gleichzeitig die Haarnadelstruktur aufgehoben wird. Die Verwendung von Autoprimer-Oligonucleotiden (Strategie E) bietet gegenüber der von überlappenden Oligonucleotiden (Strategie D) zahlreiche Vorteile: 1) Es ist keine spezielle Hybridisierung notwendig, da die Autoprimer-Oligonucleotide bereits aus der Synthese und Reinigung in Form ihrer entropisch günstigen Haarnadelstruktur anfallen. 2) Die Strangseparations-Aktivität der DNA-Polymerase^[249] wird ausgeschaltet, da der Primer kovalent an das Templat gebunden ist. 3) Man hat die Option, den codierenden oder aber den nicht codierenden DNA-Strang chemisch zu synthetisieren und den komplementären Strang mit Hilfe der problemlosen enzymatischen Polymerase-Reaktion aufzubauen. Auf diese Weise kann man der chemischen Synthese langer purinreicher Sequenzen aus dem Wege gehen, welche des öfteren Probleme bereitet.

Aus lediglich einem 140mer Oligonucleotid konnte so ein 117 bp großes Gen für Calcitonin-Gly(33) hergestellt werden^[197]. Mit Hilfe der Subklonierungstechnik können jedoch Gene jeder Größe dargestellt werden. Beispielsweise gelang es, aus nur drei Autoprimer-Oligonucleotiden durch enzymatisches Prozessieren ein 310 bp großes Genfragment herzustellen^[198]. Ähnlich wie bei der DNA-Sequenzierung nach Sanger^[202], der ebenfalls die DNA-Polymerase-Reaktion zugrunde liegt, kann man dabei ohne Schwierigkeiten vier bis acht Reaktionen gleichzeitig ausführen.

3.2.3. Auswahlkriterien: Gengröße, Oligonucleotidlänge, Mutationshäufigkeit

Um die Synthese eines Gens zu planen, sind zunächst folgende Betrachtungen anzustellen. Wie lang ist das zu synthetisierende Gen? Kann es in einem Block synthetisiert werden oder ist es in Genfragmente aufzuteilen? Welche chemische Methode steht zur Synthese der Oligonucleotide zur Verfügung und welche Oligonucleotidlängen können erzielt werden?

Prinzipiell wird man versuchen, ein Gen mit möglichst langen Oligonucleotiden in einem Klonierungsschritt herzustellen, da hierbei eine minimale Anzahl von Oligonucleotiden zu reinigen ist, welche über eine möglichst niedrige Anzahl von Ligationspunkten zu verknüpfen sind. Gegenläufig dazu ist jedoch die Mutationshäufigkeit, die mit wachsender Kettenlänge der Oligonucleotide zunimmt und mit steigender Länge eines Gens zum limitierenden Faktor wird.

Kürzlich berichtete Mutationsfrequenzen liegen zwischen 0.05% und 0.32% pro Nucleotid^[92, 173, 198, 203, 204], bei einem Mittelwert von 0.15% ist bei 1000 Nucleotiden mit etwa ein bis zwei Fehlern zu rechnen. Die Chance, bei der direkten Klonierung eines 500 bp-Gens einen richtigen Klon zu finden, ist somit bestenfalls 50%. In der Praxis muß man also mindestens zwei Klone sequenzieren, um einen richtigen Klon zu identifizieren. Im allgemeinen findet man alle Arten von Mutationen: Transitionen, Transversionen, Insertionen und Deletionen. Allerdings scheint die Mutation eines GC-Basenpaares in ein AT-Basenpaar mit der größten Häufigkeit aufzutreten^[139, 204]. Eine mögliche Erklärung hierfür ist die chemische Modifikation der Säureamidfunktion (O⁶) des Desoxyguanosins während der Oligonucleotidsynthese. Bei der Ammoniak-Behandlung könnte an den entsprechenden Positionen der DNA die Base 2,6-Diaminopurin gebildet werden, welche sowohl mit Cytosin als auch mit Thymin eine Basenpaarung eingehen kann. Das bislang längste subkli- nierte Genfragment mit 710 bp codiert für die Serinprotease-Domäne von tPA^[91]. Nach dessen Klonierung wurden drei Klone sequenziert, wobei ein Klon richtig war, während die anderen zwei Klone insgesamt drei Mutationen aufwiesen. Zum gegenwärtigen Zeitpunkt ist es empfehlenswert, größere Gene in Fragmente von etwa 250 bis 500 Basenpaaren zu unterteilen, damit die bei der Genkonstruktion eingesparte Arbeit nicht anschließend durch die DNA-Sequenzierung überkompenziert wird. In Tabelle 2 sind die Mutationshäufigkeiten, bezogen auf Oligonucleotidpaare, aus kürzlich publizierten Arbeiten einander gegenübergestellt. Während bei kurzen Oligonucleotiden (33 bp im Durchschnitt)^[173] nur jeder zwölften Klon einen Fehler aufweist, ist dies bei längeren Oligonucleotiden (85 bis 132 bp)^[92, 203] in etwa jedem vier-

Tabelle 2. Mutationshäufigkeit bei der Klonierung synthetischer Oligonucleotide.

Zahl der Oligonucleotidpaare	Gensynthese-Strategie (vgl. Abb. 7)	Mutationshäufigkeit pro Oligonucleotidpaar x/y [b]	Lit. [%]
33	A	1/12	[173]
105	E	2/18	[198]
85	B, C [a]	34/139	[92]
132	C	10/36	[203]
80	C	8/20	[204]
100	B	6/9	[139]

[a] Kombination verschiedener Strategien, inklusive „Shotgun“-Verfahren.
[b] x = Anzahl falscher Klone, y = Gesamtzahl sequenziert Klonen.

ten Klon der Fall. Eine auffallend niedrige Fehlerhäufigkeit ergibt sich bei der Verwendung von Autoprimer-Oligonucleotiden in einer „fill in“-Reaktion. Dies ist verständlich, da nur ein Strang der DNA chemisch synthetisiert wird, während der Gegenstrang in einer nahezu fehlerlosen enzymatischen Reaktion gebildet wird. Insgesamt spiegelt die Mutationshäufigkeit jedoch in starkem Maße die Qualität der verwendeten Oligonucleotide wieder. Daher dürfte eine optimale retrosynthetische Aufteilung eines Gens erst nach Sammlung einiger Erfahrung möglich sein. Neben den schon genannten werden in der Praxis weitere Auswahlkriterien Bedeutung erlangen, die später diskutiert werden. Beispielsweise kann die Aufteilung eines Gens in diverse kleine Subklone trotz des zunächst größeren Aufwands beim Genaufbau lohnend sein, wenn zu einem späteren Zeitpunkt viele Varianten des Gens hergestellt werden sollen.

3.3. Schnellverfahren zur Gensynthese

In diesem Kapitel sollen einige spezielle Variationen der vorher vorgestellten prinzipiellen Verfahren erläutert werden, welche die rasche Synthese von Genen ermöglichen. Diese Ausführungsformen der Gensynthese, die nicht mehr dem Ideal der ursprünglichen Khorana-Arbeiten unter Charakterisierung von Zwischenprodukten entsprechen, wurden erst durch die enormen Fortschritte in der Nucleotid-Synthese und DNA-Sequenzierung, sowie durch die Ausnutzung effizienter Selektions-Verfahren im Klonierungsschritt möglich.

3.3.1. „Shotgun“-Verfahren

Diese Verfahrensweise beruht darauf, daß das Zusammensetzen des Gens aus einer Anzahl von sich überlappenden Oligonucleotiden direkt und ohne intermediäre Kontrolle oder Reinigung vorgenommen wird, indem diese unmittelbar in einen geeigneten M13-Vektor ligiert und in *E. coli* kloniert werden (Abb. 8 a). *Grundström* et al. [205] berichteten über die Synthese einer Anzahl von SV40-DNA-Segmenten, die aus einem Satz von 15 Oligonucleotiden zusammengesetzt wurden. Da das synthetische DNA-Fragment an den beiden Enden die überhängenden Sequenzen für den Einbau in einen entsprechenden Vektor besitzt, werden nur vollständig ligierte Produkte, die die korrekten Enden aufweisen, letztlich in den Vektor integriert. Alle nicht vollständig ligierten Produkte führen nicht zur Recircularisierung des Vektors und bleiben damit stumm. Es wurde gefunden, daß etwa die Hälfte aller analysierten Klone die gewünschte Sequenz ent-

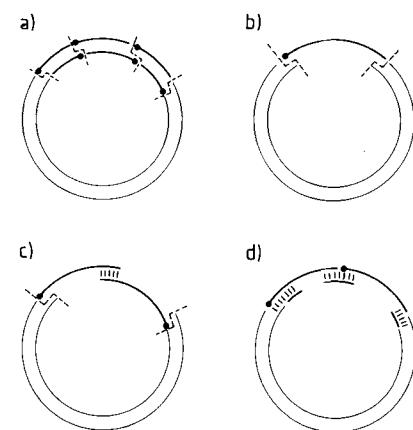


Abb. 8. Schnellverfahren zur Synthese von Genen: a) „Shotgun“-Verfahren; b) Einzelstrang-Klonierung eines Oligonucleotids; c) Einzelstrang-Klonierung 3'-überlappender Oligonucleotide; d) Einzelstrang-Klonierung mittels Hilfsoligonucleotid.

hielt [205]. Ausgehend von fünf Sätzen, einem aus elf und drei aus sieben sowie einem aus zwei Oligonucleotiden konnten *Brodin* et al. [158] ein Gen für ein Ca-bindendes Protein vom Rind herstellen. In der Zwischenzeit wurden zahlreiche weitere Gensynthesen nach diesem Schnellverfahren durchgeführt.

Die „in-vitro“-Klonierung [206] ist eine Weiterentwicklung des „Shotgun“-Verfahrens, bei der zusätzlich vom enzymatischen Apparat der Zelle Gebrauch gemacht wird, um die synthetischen Oligonucleotide zu phosphorylieren und zu ligieren. Da das Verfahren die Merkmale von in-vitro- und in-vivo-Reaktionen in sich vereinigt, erscheint der Terminus „in vitro“ gerechtfertigt. Im Grunde verfährt man wie beim Shotgun-Verfahren, nur daß man die Oligonucleotide direkt mit einem geeignet geöffneten Vektor hybridisiert und ohne in-vitro-Phosphorylierung und -Ligation in kompetente Zellen transformiert.

3.3.2. Einzelstrang-Klonierung unter Reparatur der DNA-Lücken

Im Prinzip ist die genetische Information durch einen DNA-Strang eines Gens definiert, während der zweite DNA-Strang zur Aufrechterhaltung der Doppelhelix-Struktur, zur zusätzlichen Sicherheit und zur Replikation dient. Es läßt sich deshalb nahezu die Hälfte der chemischen Synthesearbeit einsparen, wenn man einzelsträngige DNA in hierfür geeignete Vektoren ligiert. Anschließend werden Bakterien mit solch einem lückenhaften Vektor transformiert, wobei die Reparatur-Maschinerie der Wirtszellen den Vektor vervollständigt („in vivo gap repair“). Dieses Verfahren wurde bei der Synthese eines Genfragments realisiert [207], das für die Aminosäuren 198–231 von Bakteriorhodopsin codiert. Dabei wird ein Oligonucleotid pflasterartig in einen Vektor ligiert, der auf der einen Seite eine 3'-überhängende und auf der anderen Seite eine 5'-überhängende einzelsträngige DNA-Sequenz hat, die jeweils zum 5'-Ende und 3'-Ende des Oligonucleotids komplementär ist (Abb. 8 b). Bislang konnte diese Methode nur auf relativ kurze Oligonucleotide angewendet werden, wobei eine niedrige Effizienz beobachtet wurde [208]. Durch eine in-vitro-„fill in“-Reaktion oder durch eine trickreiche sequentielle Ligationsprozedur ließen

sich mit einem großen Überschuß an Oligonucleotid deutlich erhöhte Ausbeuten an gewünschten Klonen erzielen^[209]. Restriktionsenzympaare zum Öffnen des Vektors, die stets einen 5'- und 3'-Überhang erzeugen, sind nicht mehr limitierend, wenn man zwei einzelsträngige Oligonucleotide verwendet, die über ihre komplementären 3'-Enden hybridisiert sind (Abb. 8c). In diesem Fall kann man auch einen Vektor verwenden, der auf beiden Seiten 5'-überhängende Enden zur Ligation besitzt^[207]. Zur Erhöhung der Ligationsausbeute können auch Hilfsoligonucleotide eingesetzt werden, die den doppelsträngigen Bereich an den Ligationspunkten verlängern. So wurde kürzlich über die Klonierung eines 290 bp-Gens aus zwei langen Oligonucleotiden (144mer und 139mer) berichtet^[185], die mit Hilfe von drei kleineren Oligonucleotiden (11- bis 16mer) miteinander verknüpft und in einen Vektor ligiert wurden (Abb. 8d). Über die Effizienz dieser Variante wurde mitgeteilt, daß von 48 analysierten Klonen sieben Klonen das gewünschte Gen enthielten. Von drei vollständig sequenzierten Klonen zeigte lediglich einer die korrekte Sequenz.

Obwohl sich die Konzepte der Abbildungen 8c und 8d durch Verwendung einer noch größeren Anzahl von Oligonucleotiden theoretisch beliebig ausbauen ließen, sind in der Praxis Grenzen gesetzt, da mit steigender Anzahl von Oligonucleotiden die Ausbeuten fallen. Hier könnte unter Umständen die Einzelstrang-Ligation von Oligonucleotiden, die an ihrem 3'-Ende eine Ribonucleotid-Einheit aufweisen, mit Hilfe von T4-RNA-Ligase von Vorteil sein^[210-212].

Die M13-Insertionstechnik ist eine spezielle Ausführungsform der in-vitro-Mutagenese^[213]. Das Prinzip der oligonucleotidgesteuerten ortsspezifischen M13-Mutagenese (vgl. Abschnitt 4.2) beruht darauf, daß man zu einem einzelsträngigen M13-Templat ein Oligonucleotid entwirft, das zu diesem bis auf eine Basenmißpaarung komplementär ist, welche die gewünschte Mutation erzeugt (siehe Abb. 11). Nach dem Anhybridisieren des Oligonucleotids an das Templat wird mit Hilfe der DNA-Polymerase-Reaktion der zweite Strang gebildet. Anschließend wird mit DNA-Ligase der Vektor zirkularisiert. Um nun eine längere DNA-Sequenz in eine definierte Stelle im Vektor einschieben zu können (Gensynthese)^[214], muß sich diese anstelle der Mißpaarung befinden. Nach der Replikation des partiell basenungepaarten Vektors werden die Klonen mit dem gewünschten Einschub von den simultan entstehenden Wildtyp-Klonen mittels geeigneter Verfahren selektiert. Eine interessante Weiterentwicklung der Einzelstrang-Klonierung zur Gensynthese basiert auf der oligonucleotidgesteuerten Reparatur eines DNA-Doppelstrangbruches^[215].

3.3.3. Festphasen-Gensynthese

Im Abschnitt 3.2 wurde konstatiert, daß in einer Ligase-Reaktion die Ausbeute an vollständig ligierten Produkten mit steigender Anzahl der zu verknüpfenden Oligonucleotide abnimmt. Daher scheint eine sequentielle Verknüpfung von kleinen Genfragmenten vorteilhaft zu sein, wobei ein Überschuß des zu ligerenden Teilstücks die Ausbeute stark zu erhöhen vermag. Ähnlich wie bei der Synthese eines Oligonucleotids am festen Träger können auch ganze Gene durch repetitive Ligation von Genfragmenten aufgebaut werden. Die Vorteile der Festphasen-Gensynthese wurden erstmals von Hostomský et al.^[164] am Beispiel eines Colo-

strum-Trypsininhitor-Gens des Rindes beschrieben. Im ersten Schritt wurde ein Starter-Oligonucleotid über das 3'-Ende kovalent an Sephadryl S-500 via Bromcyan-Aktivierung verankert. Bei der anschließenden schrittweisen Addition wurden DNA-Duplices, die durch Hybridisierung aus drei bis vier Oligonucleotiden präpariert wurden, über relativ lange überhängende Enden von 15 Nucleotiden an komplementäre Oligonucleotide hybridisiert. Dabei wird also die wachsende Kette lediglich durch die Wasserstoffbrückenbindungen der Basen zusammengehalten. Im letzten Elongationsschritt wurde schließlich das Gesamtgen mit Hilfe von DNA-Ligase kovalent verknüpft. Der linearisierte Klonierungs-Vektor wird vorteilhaft noch am Träger addiert, bevor das Gesamtkonstrukt von diesem abgespalten und zirkularisiert wird. Diese Verfahrensweise hat den Vorteil, daß nur die vollständig gewachsenen DNA-Ketten mit den entsprechenden Enden in den Vektor integriert werden können. Nachteilig erscheinen die aufwendige Derivatisierung des festen Trägers mit dem Anker-Oligonucleotid und die ungeschützten langen Überhänge der Oligonucleotide, die potentiell zu Selbststrukturen und damit zur ineffizienten Ligation oder gar Fehlligation führen können.

Probleme dieser Art können jedoch leicht umgangen werden, indem man partiell doppelsträngige Anker-Oligonucleotide über entsprechend modifizierte ungepaarte Nucleotide an den Träger fixiert^[216]. Weiter kann man für die Kettenverlängerung vorteilhaft die im vorangehenden Kapitel beschriebenen Autoprimer-Oligonucleotide verwenden, wenn man nach der „fill in“-Reaktion durch Restriktionsenzym-Verdau nur auf einer Seite die überhängenden Enden erzeugt und auf der anderen Seite die Loop-Struktur beibehält, so daß diese Fragmente unter Verwendung eines Überschusses sehr effizient an die wachsende Kette ligiert werden können^[198]. Die Haarnadel-Struktur, welche eine Mehrfachaddition verhindert, kann durch Schneiden mit einem geeigneten Restriktionsenzym entfernt werden, um eine weitere Kettenverlängerung zu ermöglichen.

3.4. Konstruktionskonzepte für synthetische Gene

Die Güte der Konzeption eines synthetischen Gens ist maßgeblich verantwortlich für das Gelingen der Gensynthese sowie für eine erfolgreiche Expression des Gens. Daher ist eine systematische Planung unter Einbeziehung von Computer-Programmen, beispielsweise nach dem in Abbildung 9 gezeigten Fließdiagramm, zu empfehlen. Je größer ein zu synthetisierendes Gen ist, desto aufwendiger und ausgefeilter wird die Syntheseplanung sein. Die Konzepte zur Expression werden in Abschnitt 3.4.2 diskutiert. Es ist selbstverständlich, daß diese Betrachtungen mit in das endgültige Design eines Gens eingehen müssen.

3.4.1. Design eines synthetischen Gens

Genetischer Code und Wirtspezifität der DNA

Die Totalsynthese eines Gens bietet sich besonders dann an, wenn die Aminosäuresequenz eines Proteins bestimmt wurde, die DNA-Sequenz des entsprechenden Gens jedoch noch nicht bekannt ist. Kennt man den genetischen Code (Tabelle 3), so kann die Aminosäuresequenz in eine DNA-

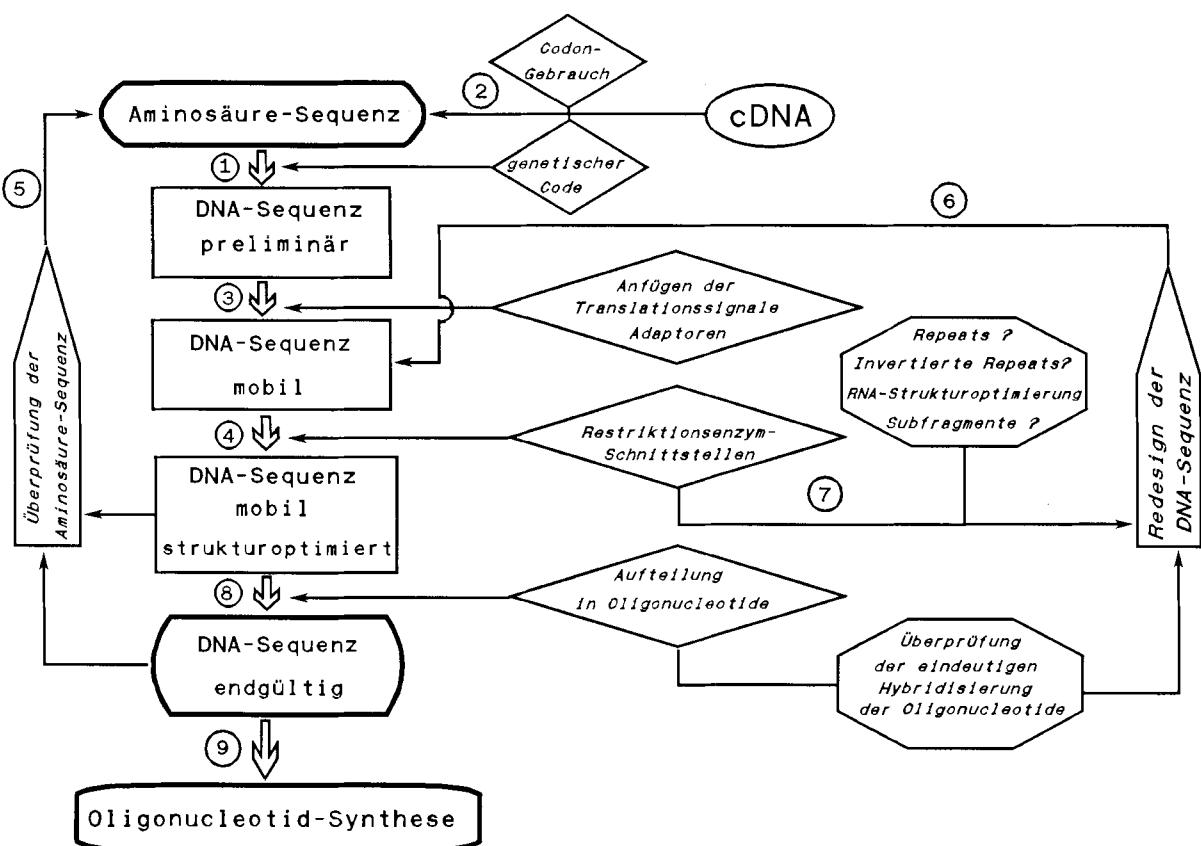


Abb. 9. Planung und Design eines synthetischen Gens.

Tabelle 3. Codon-Gebrauch in stark und schwach exprimierten Genen in *E. coli* (nach Grosjean und Fiers^[217]). Codon-Paare, die nur von einer tRNA gelesen werden, sind cingerahmt. Amber, ochre und opal entsprechen Stop-Codons.

U		C		A		G	
stark schwach		stark schwach		stark schwach		stark schwach	
U	Phe	39 151 113 102	Ser	93 36 87 49	Tyr	34 96 98 65	U C A G
	Lys	12 71 16 64		6 37 12 62	ochre amber	25 66	
C	Leu	26 73 33 69 → 3 22 345 294	Pro	21 29 2 46	His Gln	19 95 75 59 38 90 169 166	U C A G
		262 118		26 45 162 101		223 99 101 133 → 3 27 → 1 42	
A	Ile	67 156 262 118 → 2 27	Thr	103 46 137 119	Asn Lys	13 101 159 98 259 163 106 44	U C A G
	Met	140 130		15 32 28 76		10 56 49 61 → 3 28 → 1 17	
G	Val	192 108 41 66 119 48 83 123	Ala	173 87 48 178 119 107 129 149	Asp Glu	116 183 204 106 333 210 106 98	U C A G
						226 124 174 140 → 4 42 → 14 66	

Sequenz übersetzt werden (Stufe 1, Abb. 9). Der genetische Code ist jedoch degeneriert, das heißt, für die meisten Aminosäuren stehen mehrere Nucleotidtriplets zur Auswahl. Weil in natürlicher DNA die synonymen Codons nicht in der gleichen Häufigkeit benutzt werden und zudem der nicht statistische Codon-Gebrauch mit der Stärke der Genexpression zusammenhängt, scheint es einen Satz von „optimalen“ Codons für die synthetischen Gene zu geben. Je nach Art des Wirtsorganismus wird man deshalb den in hochexprimierten Genen benutzten Codon-Gebrauch von beispielsweise

E. coli^[217–219] oder anderen Wirten^[220, 221] berücksichtigen. Allerdings sollte der Codon-Gebrauch zumindest im Falle von *E. coli* nicht überbewertet werden, da bislang keine schlüssigen experimentellen Daten zu diesem Problem vorliegen, obwohl der Codon-Gebrauch für die meisten synthetischen Gene angewandt wird.

Die chemische Gensynthese ist jedoch auch dann sinnvoll, wenn die Aminosäuresequenz durch Rückübersetzung aus der cDNA-Sequenz abgeleitet wurde. Von Vorteil ist dabei neben der Anpassung optimaler Codons (Stufe 2) die Schnelligkeit, mit der maßgeschneiderte Gene zugänglich sind. Da Proteine in der Natur oft als Vorläufer synthetisiert werden, fehlen der cDNA oft die entsprechenden Translationssignale. Die Anfügung eines Start-Codons (ATG) für Methionin, eines oder zweier Translations-Stop-Codons sowie zusätzlicher Adaptor-Sequenzen für die Integration des synthetischen Gens in ein Plasmid an eine codierende Sequenz lässt sich in synthetischen Genen leicht realisieren und führt zu einer „mobilen“ DNA-Sequenz (Stufe 3). Die Umkonstruktion isolierter cDNA erfordert oft einen beträchtlichen Aufwand, um zu einer mobilen DNA-Sequenz zu gelangen.

Restriktionsenzym-Schnittstellen

Die Einführbarkeit singulärer Restriktionsenzym-Schnittstellen in ein synthetisches Gen (Stufe 4) ist einer der entscheidenden Vorteile synthetischer Gene gegenüber den natürlichen isolierten Genen. Solche Restriktionsenzym-Schnittstellen, die im allgemeinen gleichmäßig über die Gensequenz verteilt werden, sind nicht nur zur Subklonierung bei der Konstruktion eines Gens von Vorteil, sondern sie

Tabelle 4. Erkennungssequenzen für einige ausgewählte Restriktionsenzyme.

Erkennungssequenz [a]	Restriktionsenzym
5'-GAATTC- 3'-CTTAAAG- ↑	<i>Eco R</i> I
5'-CTGCAG 3'-GACGTC- ↑	<i>Pst</i> I
5'-GGCC- 3'-CCGG- ↑	<i>Hae</i> III
5'-GAAGA(N),N'- 3'-CTTCT(N),N'N'- ↑	<i>Mbo</i> II
5'-GAAN N N' N' TTC- 3'-CTTN'N'N'AAG- ↑	<i>Xmn</i> I
5'-GTPyPuAC- 3'-CAPuPyTG- ↑	<i>Hind</i> II

[a] Schnittposition mit Pfeil markiert. N = G, A, T oder C, N' = dazu komplementäre Base; Py = C oder T, Pu = G oder A

erlauben zudem eine einfache Veränderung (Mutagenese) der ursprünglichen DNA-Sequenz (vgl. Abschnitt 4). Dabei werden Schnittstellen bevorzugt, bei deren Spaltung kohäsive Enden gebildet werden (Tabelle 4). Zur Klonierung von Genen bzw. Subklonierung von Genfragmenten bedient man sich vorteilhaft solcher Schnittstellen, die auch in kommerziell erhältlichen Plasmiden vorhanden sind, beispielsweise in den pUC-Plasmiden^[222] oder den M13mp-Plasmiden^[223], welche Polylinker (multiple Klonierungsstellen) enthalten.

Wenn die ausgesuchten Restriktionsenzym-Schnittstellen jedoch in den käuflichen Plasmiden nicht zur Verfügung stehen, können temporäre Schnittstellen verwendet werden. Eine Möglichkeit der Verwendung temporärer Schnittstellen besteht darin, daß man das Gen oder Genfragment um einige Nucleotide verlängert, um eine klonierbare (temporäre) Erkennungssequenz zu erhalten. Nach der Integration des Gens in ein entsprechendes Plasmid und Vervielfältigung kann die gewünschte Gensequenz durch Schneiden mit den ursprünglich geplanten Restriktionsenzymen isoliert werden. Eine andere Möglichkeit zur Erzeugung temporärer Schnittstellen besteht darin, daß man in die DNA-Sequenz eine kurze Nucleotidfolge einbaut, welche eine klonierbare Restriktionsenzym-Erkennungssequenz ergibt. Die zusätzlichen Nucleotide können zu einem späteren Zeitpunkt durch sukzessive Spaltung mit dem jeweiligen Restriktionsenzym und der einzelstrangspezifischen Nuclease S1 entfernt werden^[195].

Die Planung singulärer Restriktionsenzym-Schnittstellen ist relativ komplex und gelingt nur mit Computer-Programmen^[224, 225]. Zur Zeit sind etwa 600 Restriktionsenzyme bekannt^[226–228].

Sekundärstruktur-Betrachtungen

Nach der Optimierung einer DNA-Sequenz hinsichtlich Codon-Gebrauch und Restriktionsenzym-Schnittstellen muß die Sekundärstruktur der korrespondierenden mRNA überprüft werden (Stufe 7). Direkte Wiederholungen von Nucleotidsequenzen („Repeats“) können beim Aufbau des Gens stören, während invertierte Wiederholungen zusätzlich zu unerwünschten Sekundärstrukturen der mRNA führen können, die eine effiziente Expression des Gens unmöglich

machen. Haarnadelstrukturen, welche die Initiierung der Translation inhibieren, können die Expression drastisch verringern^[229, 230]. Bei der Aufteilung eines Gens in Oligonucleotide (Stufe 8) muß jedes der Oligonucleotide auf störende intra- und intermolekulare Wechselwirkungen überprüft werden. Zur Berechnung der erwünschten und unerwünschten Sekundärstrukturen und Wechselwirkungen stehen zahlreiche Programme zur Verfügung^[224, 231–233]. Die Erfordernis mehr oder weniger umfangreicher Berechnungen hängt natürlich stark von der zum Genaufbau verwendeten Strategie (Abb. 7) ab. Bei der Verwendung einer relativ großen Anzahl kurzer Oligonucleotide in einer Hybridisierungsstufe sind naturgemäß viel mehr unerwünschte Wechselwirkungen zu erwarten als bei der Hybridisierung und nachfolgenden Klonierung langer Oligonucleotide^[203].

Beim Design eines synthetischen Gens müssen alle bislang diskutierten Punkte berücksichtigt werden. Wenn man dabei schrittweise, wie in Abbildung 9 gezeigt, vorgeht, kann immer wieder eine Änderungen der DNA-Sequenz (Stufe 6: Redesign) notwendig sein, beispielsweise nach den Stufen 4, 7 und 8. Da sich hierbei unter Umständen Fehler einschleichen können, ist es sinnvoll, von Zeit zu Zeit durch Rückübersetzung der DNA-Sequenz die Aminosäuresequenz zu überprüfen (Stufe 5). Das eigentliche Experiment beginnt mit der chemischen Synthese der Oligonucleotide (Stufe 9).

3.4.2. Expressionskonzepte für synthetische Gene

Zur Expression synthetischer oder natürlicher Gene gibt es drei grundlegende Verfahren (Abb. 10): a) Direktxpression, b) Expression als Fusionsprotein, c) Transportexpression.

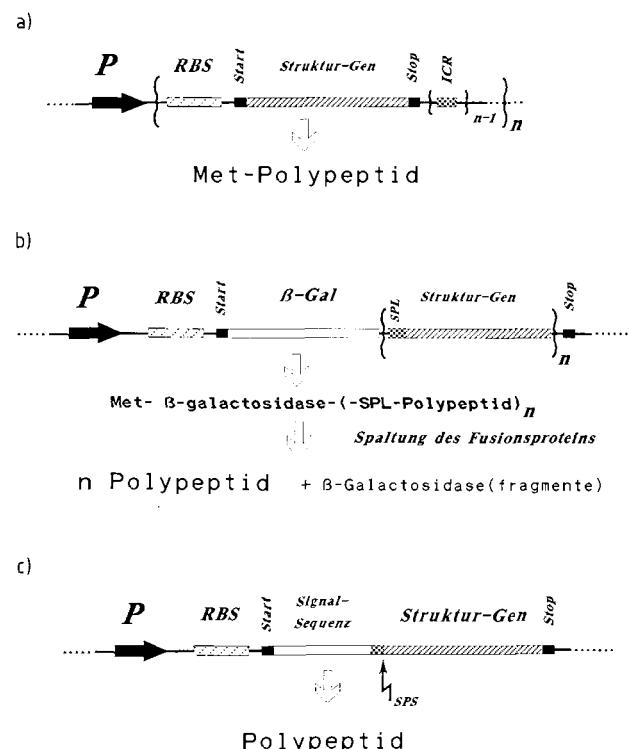


Abb. 10. Expressionskonzepte für synthetische Gene. a) Direktxpression; b) Expression als Fusionsprotein; c) Transportexpression. ICR: Intercistronische Region, P: Promotor, RBS: Ribosomenbindungsstelle, SPL: Spaltungslinker, SPS: Signalpeptidase.

Direktxpression

Die Direktxpression eines Gens unter der Kontrolle einer konstitutiven oder induzierbaren^[115] Regulationseinheit ist der einfachste Weg zur Expression in Prokaryonten, vorausgesetzt, das entstehende Produkt ist stabil und für den Wirtsorganismus untoxisch (Abb. 10 a, $n = 1$). Da dem Strukturgen ein Start-Codon für die Translation, welches für Methionin codiert, vorangestellt werden muß, wird ein Methionyl-Protein gebildet. In den meisten Fällen stört das zusätzliche Methionin nicht die Aktivität des Proteins; es kann jedoch durch Spaltung mit Bromcyan oder spezifischen Aminopeptidasen^[140] entfernt werden. Um die Expressionsausbeute zu erhöhen, kann die Direktxpression auch als polycistronisches Tandem-Gen-System ausgeführt werden. Diese Variante, bei der das gleiche Strukturgen n -mal (Abb. 10 a, $n = 2, 3 \dots$) durch intercistronische (ICR) Bereiche getrennt hintereinander geschaltet wird, nutzt den Gen-Dosis-Effekt aus. Ähnlich wie bei den polycistronischen bakteriellen Lactose- oder Tryptophan-Operons wird durch die gemeinsame Transkriptionseinheit, die aus Promotor (P) und Ribosomenbindungsstelle (RBS) besteht, lediglich eine mRNA gebildet. Die Verfahrensweise wurde erfolgreich zur Expression von Leukocyten-Interferon^[234] angewendet. Dabei war die Menge an exprimiertem Interferon direkt proportional zur Anzahl der Interferon-Gene im Plasmid ($n = 1$ bis 4).

Expression als Fusionsprotein

Die Direktxpression von Polypeptiden mit einer Länge von bis zu 50–80 Aminosäuren ist oft nicht erfolgreich, da die relativ kleinen Proteine in dem Wirtsorganismus zu schnell proteolytisch abgebaut werden. So ließ sich das Tetradecapeptid Somatostatin^[84] nur durch Stabilisierung in Form eines Fusionsproteins mit β -Galactosidase exprimieren (Abb. 10 b). Aus dem Fusionsprotein, bestehend aus etwa 1000 aminoterminalen Aminosäuren der β -Galactosidase, gefolgt von Methionin und dem Tetradecapeptidanteil des Somatostatins, konnte durch Bromcyanspaltung das Somatostatin isoliert werden. Wenn das gewünschte Protein ein internes Methionin enthält, sind auf der Genebene alternative Spaltungssequenzen einzuplanen, z. B. AspPro (säure-labil)^[235], AsnGly (Hydroxylamin)^[236], ArgArg (Trypsin, Carboxypeptidase B)^[237] oder IleGluGlyArg (Blutgerinnungsfaktor Xa)^[238]. Oft genügt auch schon ein kleinerer Fusionsanteil für eine Stabilisierung, etwa Teile von Trp E^[99], Interferon- γ ^[100], Wachstumshormon^[144] oder kurze Homooligopeptid-Schwänze^[239]. Eine Variante ist der Einbau mehrerer Gene, die jedoch nicht wie vorher beschrieben durch intercistronische Bereiche getrennt sind, sondern vielmehr über Spaltungssequenzen verbrückt sind. Im Falle von Proinsulin, das aufgrund seiner kleinen Halbwertszeit in *E. coli* ($t_{1/2} = 2$ min) direkt nicht in guten Ausbeuten exprimiert werden kann, ließen sich durch Aneinanderkuppeln von drei bis vier Genen (Abb. 10 b, $n = 3$ oder 4) über einen für ArgArgAsnSerMet codierenden Spaltungslinker (SPL) stabile Fusionsproteine herstellen^[240]. Durch Spaltung des Fusionsproteins mit Bromcyan und nachfolgende Behandlung mit Trypsin/Carboxypeptidase B konnte Proinsulin gewonnen werden. Dieses Multigenkopie-Verfahren wurde unter anderem für Substanz P^[153], den Wachstumshormon-Releasing-Faktor^[102] und Met-Encephalin^[154] beschrieben

und hat gegenüber β -Galactosidase-Fusionen den offensichtlichen Vorteil, daß die Wirtszellen weniger Ballast zu produzieren haben, welcher zudem bei der Spaltung und Aufarbeitung des Fusionsproteins störend sein kann.

Transportexpression

Ein entscheidender Nachteil der Expression als Fusionsprotein besteht darin, daß die Proteine im allgemeinen in einer biologisch inaktiven Form anfallen. Die Faltung der Proteine zu biologisch aktiven Strukturen kann mitunter schwierig sein. Nun hat man mit der Transportexpression (Abb. 10 c) eine Verfahrensweise zur Hand, mit der sich die zuvor diskutierten Probleme (zusätzliches Methionin, Proteinfaltung, Disulfidbrücken-Bildung, Protease-Labilität) weitgehend umgehen lassen. Der Vorteil basiert auf der Tatsache, daß die wachsende Polypeptidkette mit Hilfe einer Signalsequenz durch die cytoplasmatische Membran aus der reduzierenden Umgebung des Cytoplasmas ausgeschleust wird. Nach der Hypothese von Blobel und Dobberstein^[241] wird die Signalsequenz beim Durchtritt der Polypeptidkette durch die Membran von einer Signalpeptidase (SPS) abgespalten. Beispielsweise konnten der epidermale Wachstumsfaktor (EGF)^[101], Hirudin^[242] und Wachstumshormon^[243] mit Hilfe der Signalsequenz der alkalischen Phosphatase in das Periplasma von *E. coli* transportiert werden. Wir berichteten früher über ein synthetisches DNA-Fragment für das Signalpeptid der alkalischen Phosphatase, das die Sekretion von Proinsulin bewirkt^[244]. Zum Zwecke der Transportexpression ist *E. coli* jedoch nicht unbedingt der Wirtsorganismus der Wahl. Oft erhält man in anderen Wirtsorganismen wie Staphylokokken^[245], *B. subtilis*^[246] oder Hefe bessere Ausbeuten, weil das Produkt hierbei in das Medium sezerniert wird. Besondere Bedeutung kommt *Saccharomyces cerevisiae* in Verbindung mit der α -Faktor-Signalsequenz^[247, 248] zu, womit sich beispielsweise die synthetischen Gene für das atriale natriuretische Peptid^[98] und das „connective tissue activating“-Peptid^[131] exprimieren ließen. Ein nicht zu unterschätzendes Problem bei der Transportexpression ins Medium können jedoch wirtseigene, extrazelluläre Proteasen darstellen, wenn diese das exprimierte Produkt abzubauen vermögen. In Hefe kann es im Falle glycosylierbarer Proteine zudem zur Bildung von Glycoproteinen kommen, deren Glycosyl-Anteil nicht identisch ist mit dem des natürlichen Produktes.

4. Anwendungen und Vorteile synthetischer Gene

4.1. Maßgeschneiderte Gene

Gerade für kleine Proteine mit bis zu etwa 200 Aminosäuren, deren Primärstruktur durch Protein-Sequenzierung aufgeklärt wurde, bietet die chemisch-enzymatische Gensynthese eine vorteilhafte Alternative zur Isolierung des natürlichen Gens. Synthetische Gene für Proteine dieser Größe können heutzutage in wenigen Wochen bis Monaten hergestellt und kloniert werden. Im Gegensatz zu natürlichen Genen, die nach der Isolierung für eine Expression (beispielsweise in Bakterien) erst umkonstruiert werden müssen, lassen sich synthetische Gene gleich so planen, daß sie unmittelbar nach der Klonierung zur Expression gebracht werden können.

Wie im vorangegangenen Abschnitt gezeigt, kann dabei die Wirtspezifität durch Verwendung bevorzugter Codons berücksichtigt werden. Zudem lassen sich bei synthetischen Genen in vernünftigen Abständen (etwa alle 50 bis 100 bp) singuläre Restriktionsenzym-Erkennungsstellen einbauen, welche die nachträgliche Veränderung von Genen und Genprodukten stark vereinfachen. Schließlich kann die Gensynthese regulatorische DNA-Sequenzen mit einschließen, welche die Transkription und Translation des Strukturgens bewirken. Die gezielte Veränderung von Strukturgenen oder von regulatorischen DNA-Sequenzen ist eine wertvolle Methode zum Studium von Struktur-Aktivitäts- und Struktur-Funktions-Beziehungen.

4.2. Gezielte Veränderungen und Optimierung von Protein-Eigenschaften

Obwohl man heutzutage mit Hilfe der sogenannten oligonucleotidgesteuerten ortsspezifischen M13-Mutagenese (vgl. Abschnitt 3.3.2) jede DNA beliebigen Ursprungs gezielt verändern kann^[249], ist jedoch synthetische DNA prädestiniert für ein solches Vorhaben. Während man mit den verschiedenen Ausführungen der M13-Mutagenese^[213, 249-255] neben der gewünschten Mutante stets auch das Wildtyp-Gen er-

hält, führt die modulare oder Kassetten-Mutagenese bei korrekter Ausführung ausschließlich zum gezielt veränderten Gen (Abb. 11). Bei richtiger Planung eines synthetischen Gens erlauben die entsprechend plazierten Restriktions-enzym-Schnittstellen eine schnelle Veränderung des Protein-Produktes, indem man ein DNA-Segment, das von zwei singulären Restriktionsenzym-Schnittstellen flankiert wird, durch ein entsprechendes, die Mutation enthaltendes, DNA-Fragment ersetzt (modulare Mutagenese^[225]). So wurden bestimmte Gene^[96, 141, 159, 169] speziell im Hinblick auf nachfolgende Veränderungen geplant. Im Falle des Wachstumshormon-Releasing-Hormons (GHRH) haben wir das Codon für Methionin 27 durch ein Codon für Leucin ersetzt, um so die Freisetzung des GHRH-Derivats aus einem Fusionsprotein mit β -Galactosidase durch Bromcyan-Spaltung zu ermöglichen^[96]. Bei der sogenannten Kassetten-Mutagenese wird das zu verändernde DNA-Fragment durch eine Kassette von nicht eindeutig synthetisierten Oligonucleotiden ersetzt^[256, 257]. Dies führt zur Erzeugung zahlreicher Mutanten in einem Klonierungsschritt. Spezielle Ausführungsformen der Kassetten-Mutagenese erlauben die Klonierung einzelsträngiger Oligonucleotide^[208, 209, 258].

Die Veränderung von Enzymen kann diese beispielsweise resistent gegenüber Oxidation^[259] machen, ihre pH-Abhängigkeit verschieben^[260], oder gar die Umkehrung einer allo-

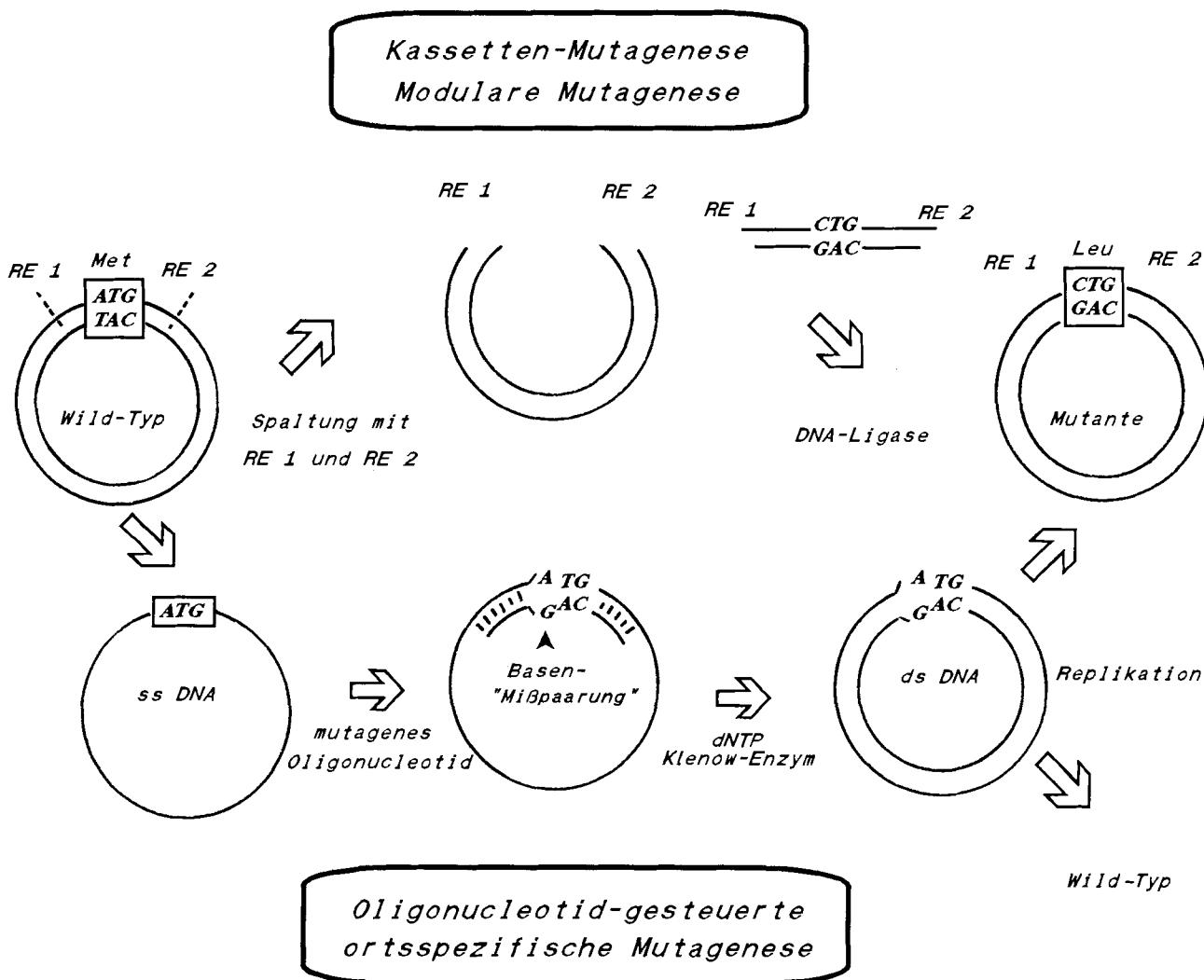


Abb. 11. Alternative Wege zur gezielten Veränderung eines Gens.

sterischen Inhibition in eine Aktivierung bewirken^[261]. Bei pharmakologisch interessanten Proteinen wie Interferon- γ ^[262] oder Interleukin-2^[263, 264] wurden Struktur-Aktivitäts-Beziehungen studiert. Im günstigsten Fall führen Mutationen zur Erhöhung der biologischen Aktivität, wie etwa bei bestimmten Mutanten von Interleukin-1^[265] oder Interleukin-2^[266].

Gene lassen sich zudem so verändern, daß eine selektive posttranskriptionale Modifikation des exprimierten Vorläufer-Proteins möglich ist. So treten in der Natur viele kleinere Proteine (< 50 Aminosäuren) in Form ihrer C-terminalen Carboxamide auf, welche durch Expression in Prokaryonten nicht direkt zugänglich sind. Wir konnten kürzlich zeigen^[96, 267], daß man durch Planung auf der DNA-Ebene einen GHRH-Vorläufer herstellen kann, welcher am C-Terminus ein zusätzliches Glycin enthält. Dieses dient als Signal für ein amidierendes Enzym^[268] zur Umwandlung des Vorläufer-Proteins in das gewünschte GHRH-Carboxamid.

Eine besondere Variante der Gensynthese, auch „Hybrid“- oder „Mißpaarungs“-Verfahren genannt, erlaubt die simultane Herstellung nahe verwandter Gene. Dabei wird ein partiell basenungepaartes Gen hergestellt, dessen beide Strände für unterschiedliche Proteine codieren^[95]. So wurde ein Gen hergestellt, in dem ein Strang für den humanen epidermalen Wachstumsfaktor (hEGF) codiert, während der andere Strang indirekt für den EGF der Maus codiert^[132].

4.3. Kontrolle der Genexpression durch synthetische DNA

Ein weiterer außerordentlicher Vorteil synthetischer Gene liegt in der Tatsache begründet, daß zusätzlich zum Strukturgen auch noch Transkriptions- oder Translations-Kontrollregionen mit in den Syntheseplan einbezogen werden können, welche die Expression des synthetischen Gens steuern. Im Baukastensystem lassen sich die einzelnen Kontrollelemente modular zu einer gesamten Regulationseinheit zusammensetzen. Wir verwendeten zum Beispiel eine synthetische Regulationseinheit^[3], bestehend aus einem idealisierten Promotor, einem Operator und einer Ribosomenbindungsstelle, zur Expression von Interferon- γ . Es soll allerdings nicht der Eindruck erweckt werden, daß eine Kombination von individuell sehr effizienten Modulen stets zu einer optimalen Genexpression führt, auch das Gegenteil kann in der Praxis auftreten. Der hohe Wert des Baukastensystems liegt vielmehr darin, daß sehr viele unterschiedliche Modulkombinationen schnell und effizient herstellbar sind, um die optimale Kombination zur Genexpression ausfindig machen zu können.

Die Transkription kann durch synthetische Promotoren, Operatoren und Terminatoren kontrolliert werden. Mit Hilfe synthetischer Ribosomenbindungsstellen und der Planung einer günstigen mRNA-Sekundärstruktur läßt sich die Translation positiv beeinflussen. Ein Promotor ist eine spezielle DNA-Sequenz, die mit der RNA-Polymerase in Wechselwirkung tritt, um die Synthese der mRNA zu ermöglichen. Die Effizienz dieser Kontrollregion, die sich im allgemeinen im nichttranslatierten 5'-Bereich befindet, ist entscheidend für die Menge der gebildeten mRNA. Es wurden schon zahlreiche Promotoren synthetisiert^[3, 196, 269 – 273].

Es gibt Promotoren, die stets angeschaltet sind (konstitutiv), und solche, die induziert werden. Die regulierbaren Pro-

motoren enthalten zusätzlich zur vorher beschriebenen RNA-Polymerase-Wechselwirkungsregion einen sogenannten Operator. Das wohl am besten untersuchte System ist das lac-Operon^[275], das negativ reguliert ist, indem der Start der Transkription durch die Bindung des lac-Repressor-Proteins an die lac-Operator-Sequenz inhibiert wird. Durch Zugabe von Isopropyl- β -D-thiogalactopyranosid (IPTG), das den Repressor in eine den Operator nicht bindende Form zwingt, wird der lac-Promotor induziert. Interessanterweise konnten idealisierte synthetische lac-Operatoren hergestellt werden^[276, 277], welche den lac-Repressor bis zu achtmal stärker binden als der natürliche Operator. Durch Kupplung eines für sich konstitutiven Promotors mit einer lac-Operator-Sequenz konnte daraus ein induzierbares System hergestellt werden^[278].

Insbesondere dann, wenn zum Transkriptionsstart starke Promotoren verwendet werden, ist für die Termination der Transkription ein effizienter Terminator von Vorteil. So konnte die Expressionsausbeute für Interferon- α um den Faktor zwei^[279] und für EGF um den Faktor fünf bis zehn^[280] erhöht werden, wenn starke Terminatoren „flußabwärts“ zum Strukturgen angefügt wurden. Synthetische Terminatoren werden sinnvollerweise gleich mit dem entsprechenden Gen mitsynthetisiert^[158].

Die Effizienz der Translation einer mRNA wird durch die Ribosomenbindungsstelle (RBS) und die mRNA-Struktur kontrolliert. Die RBS weist in Prokaryonten die Konsensus-Sequenz AAGGA, auch Shine-Dalgarno(SD)-Sequenz^[281] genannt, auf. Sie ist für den Start der Proteinbiosynthese durch ihre Bindung an das 3'-Ende der 16S-RNA verantwortlich. Die SD-Sequenz ist im allgemeinen vier bis neun Nucleotide lang und befindet sich etwa drei bis zwölf Nucleotide „flußaufwärts“ zum Translationsstart-Codon. Gene eukaryontischen Ursprungs haben keine RBS, die von prokaryontischen Ribosomen erkannt werden. Daher werden seit einiger Zeit synthetische RBS zur Expression eukaryontischer Gene in Bakterien verwendet^[3, 279, 280, 282 – 284]. Dabei scheint es wichtig zu sein, daß die RBS durch die 16S-RNA leicht zugänglich ist. Man nimmt an, daß die RBS auch noch den Teil des Strukturgen umfaßt, der für den aminoterminalen Bereich des Proteins codiert. Es erscheint daher sinnvoll, mit Hilfe entsprechender Computer-Programme^[231 – 233] störende Sekundärstrukturen aufzuspüren. Bei der Expression von Interferon- α wurden Schwankungen um den Faktor 50 beobachtet, wenn an bestimmten Positionen in der Nähe der Translations-Startregion bestimmte Nucleotide ausgetauscht wurden. Diese Resultate konnten auf der Basis von mRNA-Sekundärstrukturberechnungen erklärt werden^[230]. Aus unseren eigenen Arbeiten beim Studium der Expression von Interferon- γ haben wir jedoch gelernt, daß Aussagen über die Höhe der Expression anhand von mRNA-Strukturberechnungen allein kaum möglich sind. Die Expression muß schließlich doch experimentell optimiert werden.

5. Ausblick

Die Rekombinationstechniken und besonders die Möglichkeit der Herstellung synthetischer Gene und Oligonucleotide, haben viele biologisch aktive Proteine verfügbar gemacht. Es ist jedoch zusehends klarer geworden, daß unser

Verständnis von Struktur und Funktion komplexer Polypeptide sehr beschränkt ist. Um aber eine Antwort auf die Frage „Welche Proteinsequenz ergibt welche Eigenschaften?“ zu finden, kann die Polypeptidsynthese über synthetische Gene die chemische Peptidsynthese in idealer Weise ergänzen^[285, 286]. Durch die allgemeine Gültigkeit des genetischen Codes ist es möglich, für ein beliebiges Polypeptid oder besser, eine definierte Reihenfolge verknüpfter Aminosäuren, eine entsprechende DNA zu synthetisieren und diese nach Rekombination in einem geeigneten Vektor im jeweiligen Wirtsorganismus zur Expression zu bringen. Hierbei ist besonders die Synthese von neuen und unbekannten Polypeptiden^[143, 180], das sogenannte de-novo-Protein-Design^[287-291] hervorzuheben, das es ermöglicht unnatürliche Proteine aufzubauen und die Hypothesen über Struktur und Eigenschaften zu testen (Abb. 12). In dem abgebildeten Schema läßt sich wie in einem Kreisprozeß die Beziehung zwischen Proteinsequenz und Funktion und der zugrundeliegenden DNA-Sequenz verdeutlichen.

Protein-Design

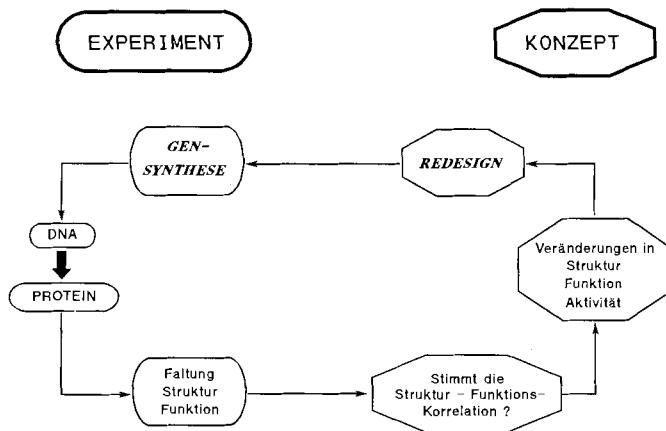


Abb. 12. De-novo-Protein-Design mit Hilfe synthetischer Gene.

Noch ist die Vorhersage der Tertiärstruktur von Proteinen ein mehr oder weniger ungelöstes Problem^[292]. Aktuell ist die Synthese bestimmter Sekundärstrukturmotive^[293], wie α -Helices und β -Faltblätter, die bereits besser vorhersagbar sind^[294] und die sich dann zu Supersekundärstrukturen oder sogenannten Motiven arrangieren. Ein solches Experiment wird derzeit am Beispiel des aus vier α -Helices bestehenden Helix-Bündels sowohl peptidchemisch als auch durch Rekombinationstechniken angegangen. Ein dafür codierendes Gen ist synthetisiert und das artifizielle Produkt läßt sich tatsächlich durch Expression in *E. coli*^[295] herstellen. Für zukünftige Forschung auf diesem Gebiet ist ermutigend festzustellen, daß hier ein experimenteller Ansatz existiert, eine der größten Herausforderungen der Molekularbiologie zu begegnen und den Zusammenhang zwischen Peptidsequenz, -struktur und -funktion zu erarbeiten.

Für die finanzielle Unterstützung dieser Arbeiten danken wir der Deutschen Forschungsgemeinschaft, dem Fonds der Chemischen Industrie, dem Bundesministerium für Forschung und Technologie sowie der Hoechst AG. Besonders danken wir E. Groeneveld für die Texterfassung und H. Zimmermann für die Reinzeichnungen.

Eingegangen am 26. September 1988 [A 720]

- [1] J. D. Watson, F. H. C. Crick, *Nature (London)* 171 (1953) 737.
- [2] R. E. Dickerson, H. R. Drew, *J. Mol. Biol.* 149 (1981) 761.
- [3] J. Engels, M. Leineweber, E. Uhlmann in J. Streith, H. Prinzbach, G. Schill (Hrsg.): *Organic Synthesis: an Interdisciplinary Challenge*. Blackwell, Oxford 1985, S. 205.
- [4] J. Engels, E. Uhlmann in A. Fiechter (Hrsg.): *Advances in Biochemical Engineering/Biotechnology*. Springer, Berlin 1988, S. 75.
- [5] R. Frank, H. Blöcker in B. Wittmann-Liebold (Hrsg.): *Advanced Methods in Protein Microsequence Analysis*. Springer, Berlin 1986, S. 387.
- [6] M. Edwards, *Int. Biotechnol. Lab.* 5 (1987) 19.
- [7] J. E. Davies, H. G. Gassen, *Angew. Chem.* 95 (1983) 26; *Angew. Chem. Int. Ed. Engl.* 22 (1983) 13.
- [8] F. Wengenmayer, *Angew. Chem.* 95 (1983) 874; *Angew. Chem. Int. Ed. Engl.* 22 (1983) 842.
- [9] F. Chow, T. Kempe, G. Palm, *Nucleic Acids Res.* 9 (1981) 2807.
- [10] G. Kumar, M. S. Poonian, *J. Org. Chem.* 49 (1984) 4905.
- [11] R. Letsinger, E. P. Groody, N. Lander, N. Tanaka, *Tetrahedron* 40 (1984) 137.
- [12] T. Hirose, R. Crea, K. Itakura, *Tetrahedron Lett.* 20 (1979) 2449.
- [13] K. Miyoshi, T. Miyake, T. Hozumi, K. Itakura, *Nucleic Acids Res.* 8 (1980) 5473.
- [14] T. W. Greene: *Protective Groups in Organic Synthesis*. Wiley, New York 1981.
- [15] H. Kössel, H. Seliger, *Fortschr. Chem. Org. Naturst.* 32 (1975) 297.
- [16] V. Amarnath, A. D. Broom, *Chem. Rev.* 77 (1977) 183.
- [17] S. A. Narang, *Tetrahedron* 39 (1983) 3.
- [18] E. Sonveaux, *Bioorg. Chem.* 14 (1986) 274.
- [19] H. G. Khorana, *Bioorg. Chem.* 1 (1978) 351.
- [20] G. S. Ti, B. L. Gaffney, R. A. Jones, *J. Am. Chem. Soc.* 104 (1982) 1316.
- [21] J. C. Schulhof, D. Molko, R. Teoule, *Nucleic Acids Res.* 15 (1986) 397.
- [22] L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* 24 (1983) 2953.
- [23] B. C. Froehler, M. D. Matteucci, *Nucleic Acids Res.* 22 (1983) 8031.
- [24] C. B. Reese, A. Ubasawa, *Tetrahedron Lett.* 21 (1980) 2265.
- [25] B. L. Gaffney, L. A. Marky, R. A. Jones, *Tetrahedron* 40 (1984) 3.
- [26] A. D. Barone, J. Y. Tang, M. H. Caruthers, *Nucleic Acids Res.* 12 (1984) 4051.
- [27] C. B. Reese, P. A. Skone, *J. Chem. Soc. Perk. Trans. 1* 1984, 1263.
- [28] F. Himmelsbach, B. S. Schulz, T. Trichtinger, R. Charubala, W. Pfeiderer, *Tetrahedron Lett.* 40 (1984) 59.
- [29] S. S. Jones, C. B. Reese, S. Sibande, A. Ubasawa, *Tetrahedron Lett.* 22 (1981) 4755.
- [30] T. Kamimura, M. Tsuchiya, K. Koura, M. Sekine, T. Hata, *Tetrahedron Lett.* 24 (1983) 2775.
- [31] M. Mag, J. W. Engels, *Nucleic Acids Res.* 16 (1988) 3525.
- [32] F. H. Westheimer, *Acc. Chem. Res.* 1 (1968) 70.
- [33] a) J. Engels, *Habilitationsschrift*, Universität Konstanz 1978; b) E. Uhlmann, W. Pfeiderer, *Nucleic Acids Res. Symp. Ser.* 4 (1978) 25; c) E. Uhlmann, W. Pfeiderer, *Helv. Chim. Acta* 64 (1981) 1688.
- [34] A. M. Michelson, A. R. Todd, *J. Chem. Soc.* 1955, 2632.
- [35] S. Stawinski, T. Hozumi, S. A. Narang, *Can. J. Chem.* 54 (1976) 670.
- [36] J. H. van Boom, P. M. J. Burgers, G. van der Marel, C. H. M. Verdegaal, G. Wille, *Nucleic Acids Res.* 4 (1977) 1047.
- [37] J. B. Chattopadhyaya, C. B. Reese, *Tetrahedron Lett.* 20 (1979) 5059.
- [38] J. Engels, U. Krahmer, L. Zsolnai, G. Huttner, *Liebigs Ann. Chem.* 1982, 745.
- [39] G. R. Gough, M. J. Brunden, P. T. Gilham, *Tetrahedron Lett.* 22 (1981) 4177.
- [40] R. L. Letsinger, *J. Am. Chem. Soc.* 97 (1975) 3278.
- [41] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* 22 (1981) 1859.
- [42] M. H. Caruthers, A. D. Barone, S. L. Beaucage, D. R. Dodds, E. F. Fisher, L. J. McBride, M. D. Matteucci, Z. Stabinsky, J.-Y. Tang, *Methods Enzymol.* 154 (1987) 287.
- [43] G. W. Daub, E. van Tamelen, *J. Am. Chem. Soc.* 99 (1977) 3526.
- [44] S. P. Adams, *J. Am. Chem. Soc.* 105 (1983) 661.
- [45] N. D. Sinha, J. Biernat, H. Köster, *Tetrahedron Lett.* 24 (1983) 5843.
- [46] R. H. Hall, A. Todd, R. F. Webb, *J. Chem. Soc.* 1957, 3291.
- [47] B. C. Froehler, P. G. Ng, M. D. Matteucci, *Nucleic Acids Res.* 14 (1986) 5399.
- [48] P. J. Garegg, I. Lindh, T. Regberg, J. Stawinski, R. Stromberg, C. Henrichson, *Tetrahedron Lett.* 27 (1986) 4051.
- [49] Applied Biosystems, *User Bull.* 44 (1987) 1.
- [50] W. S. Zielinski, L. E. Orgel, *Nature (London)* 327 (1987) 346.
- [51] R. B. Merrifield, *J. Am. Chem. Soc.* 85 (1963) 2149.
- [52] R. L. Letsinger, M. J. Kornet, *J. Am. Chem. Soc.* 85 (1963) 3045.
- [53] R. L. Letsinger, V. Mahadevan, *J. Am. Chem. Soc.* 87 (1965) 3526.
- [54] H. Ito, Y. Ike, S. Ikuta, K. Itakura, *Nucleic Acids Res.* 10 (1982) 1755.
- [55] K. Miyoshi, R. Arentzen, T. Huang, K. Itakura, *Nucleic Acids Res.* 8 (1980) 5507.
- [56] R. Crea, T. Horn, *Nucleic Acids Res.* 8 (1980) 2331.
- [57] V. K. Potapov, V. P. Veiko, O. N. Koroleva, Z. A. Shabarova, *Nucleic Acids Res.* 6 (1979) 2041.
- [58] H. Köster, *Tetrahedron Lett.* 1972, 1527.
- [59] M. D. Matteucci, M. H. Caruthers, *J. Am. Chem. Soc.* 103 (1981) 3185.

- [60] G. R. Gough, M. J. Brunden, P. T. Gilham, *Tetrahedron Lett.* 22 (1981) 4177.
- [61] G. R. Gough, M. J. Brunden, P. T. Gilham, *Tetrahedron Lett.* 24 (1983) 5321.
- [62] J. S. deBaer, J. A. Hayes, M. P. Koleck, G. R. Gough, *Nucleosides & Nucleotides* 6 (1987) 821.
- [63] J. W. Efcavitch, L. J. McBride, S. C. Eadie in K. S. Bruzik, W. J. Stec (Hrsg.): *Biophosphates and Their Analogues – Synthesis, Structure, Metabolism and Activity*, Elsevier, Amsterdam 1987, S. 65.
- [64] a) R. Frank, W. Heikens, G. Heisterberg-Moutsis, H. Blöcker, *Nucleic Acids Res.* 11 (1983) 4365; b) R. Frank, A. Meyerhans, K. Schwellnus, H. Blöcker, *Methods Enzymol.* 154 (1987) 221.
- [65] a) H. W. D. Matthes, W. M. Zeneke, T. Grundstroem, *EMBO J.* 3 (1984) 801; b) H. W. D. Matthes, A. Staub, P. Chambon, *Methods Enzymol.* 154 (1987) 250.
- [66] W. Bannwarth, P. Iaiza, *DNA* 5 (1986) 413.
- [67] a) M. J. Gait: *Oligonucleotide Synthesis, a Practical Approach*, IRL Press, Oxford 1984; b) Applied Biosystems, *User Bull.* 13 (1987) 1.
- [68] K. Itakura, J. J. Rossi, R. B. Wallace, *Annu. Rev. Biochem.* 53 (1984) 323.
- [69] K. L. Agarwal, H. Büchi, M. H. Caruthers, N. Gupta, H. G. Khorana, K. Kleppe, A. Kumar, E. Ohtsuka, U. L. Rajbhandary, J. H. van de Sande, V. Sgaramella, H. Weber, T. Yamada, *Nature (London)* 227 (1970) 27.
- [70] H. G. Khorana, K. L. Agarwal, P. Besmer, H. Büchi, M. H. Caruthers, P. Cashion, M. Fridkin, E. Jay, K. Kleppe, R. Kleppe, A. Kumar, P. Loewen, R. C. Miller, K. Minamoto, A. Panet, U. Rajbhandary, B. Ramamoorthy, T. Sekiya, T. Takeya, J. H. van de Sande, *J. Biol. Chem.* 251 (1976) 565.
- [71] J. H. van de Sande, M. H. Caruthers, A. Kumar, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 571.
- [72] K. Minamoto, M. H. Caruthers, B. Ramamoorthy, J. H. van de Sande, N. Sidorova, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 587.
- [73] K. L. Agarwal, M. H. Caruthers, M. Fridkin, A. Kumar, J. H. van de Sande, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 599.
- [74] E. Jay, P. J. Cashion, M. Fridkin, B. Ramamoorthy, K. L. Agarwal, M. H. Caruthers, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 609.
- [75] K. L. Agarwal, M. H. Caruthers, H. Büchi, J. H. van de Sande, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 624.
- [76] T. Sekiya, P. Besmer, T. Takeya, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 634.
- [77] P. C. Loewen, R. C. Miller, A. Panet, T. Sekiya, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 642.
- [78] A. Panet, R. Kleppe, K. Kleppe, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 651.
- [79] M. H. Caruthers, R. Kleppe, K. Kleppe, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 658.
- [80] R. Kleppe, T. Sekiya, P. C. Loewen, K. Kleppe, K. L. Agarwal, H. Büchi, P. Besmer, M. H. Caruthers, P. J. Cashion, M. Fridkin, E. Jay, A. Kumar, R. C. Miller, K. Minamoto, A. Panet, U. L. Rajbhandary, B. Ramamoorthy, N. Sidorova, T. Takeya, J. H. van de Sande, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 667.
- [81] B. Ramamoorthy, R. G. Lees, D. G. Kleid, H. G. Khorana, *J. Biol. Chem.* 251 (1976) 676.
- [82] E. L. Brown, B. Ramamoorthy, M. Ryan, H. G. Khorana, *Methods Enzymol.* 68 (1979) 109.
- [83] H. Köster, H. Blöcker, R. Frank, S. Geussenhainer, W. Kaiser, *Hoppe-Seylers Z. Physiol. Chem.* 356 (1975) 1585.
- [84] K. Itakura, T. Hirose, R. Crea, A. D. Riggs, H. Heyneker, F. Bolivar, H. Boyer, *Science (Washington)* 198 (1977) 1056.
- [85] R. Crea, A. Kraszewski, T. Hirose, K. Itakura, *Proc. Natl. Acad. Sci. USA* 75 (1978) 5765.
- [86] D. Goeddel, D. G. Kleid, F. Bolivar, H. Heyneker, D. Y. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, A. D. Riggs, *Proc. Natl. Acad. Sci. USA* 76 (1979) 106.
- [87] M. Edge, A. R. Greene, G. R. Heathcliffe, P. A. Meacock, W. Schuch, D. B. Scanlon, T. C. Atkinson, C. R. Newton, A. Markham, *Nature (London)* 292 (1981) 756.
- [88] N. Mantei, M. Schwarzenstein, M. Streuli, S. Panem, S. Nagata, C. Weissmann, *Gene* 10 (1980) 1.
- [89] M. Ikebara, E. Ohtsuka, T. Tokunaga, S. Iwai, K. Kitano, S. Miyamoto, T. Ohgi, Y. Sakuragawa, K. Fujiyama, T. Ikari, M. Kobayashi, T. Miyake, S. Shibahara, A. Ono, T. Ueda, T. Tanaka, H. Baba, A. Sakurai, T. Oishi, O. Chisaka, K. Matsubara, *Proc. Natl. Acad. Sci. USA* 81 (1984) 5956.
- [90] J. Engels, W. Becker, T. Beckers, D. Hüskens, W. Wetekam, *Pure Appl. Chem.* 59 (1987) 437.
- [91] L. D. Bell, J. C. Smith, R. Derbyshire, M. Finlay, I. Johnson, R. Gilbert, P. Slocum, E. Cook, H. Richards, P. Clissold, D. Meredith, C. H. Powell-Jones, K. M. Dawson, B. L. Carter, K. G. McCullagh, *Gene* 63 (1988) 155.
- [92] M. A. Wosnick, R. W. Barnett, A. M. Vicentini, H. Erfle, R. Elliott, M. Summer-Smith, N. Mantei, R. W. Davies, *Gene* 60 (1987) 115.
- [93] E. Uhlmann, F. Hein, S. Hein, L. Hornung, R. Schardt, unveröffentlicht.
- [94] R. Wetzel, D. G. Kleid, R. Crea, H. L. Heyneker, D. G. Yansura, T. Hirose, A. Kraszewski, A. D. Riggs, D. V. Goeddel, *Gene* 63 (1981) 63.
- [95] W. L. Sung, D. M. Zahab, F.-L. Yao, C. S. Tam, *Biochem. Cell Biol.* 64 (1986) 133.
- [96] J. W. Engels, J. Glauder, H. Müllner, E. Uhlmann, W. Wetekam, T. Hashimoto Gotoh, B. Scheikl-Lenz, *Nucleosides & Nucleotides* 6 (1987) 185.
- [97] M. S. Urdea, J. P. Merryweather, G. T. Mullenbach, D. Coit, U. Heberlein, P. Valenzuela, P. J. Barr, *Proc. Natl. Acad. Sci. USA* 80 (1983) 7461.
- [98] G. P. Vlasuk, G. H. Bencen, R. S. Scarborough, P.-K. Tsai, J. L. Whang, T. Maack, M. J. F. Camargo, S. W. Kirsher, J. A. Abraham, *J. Biol. Chem.* 261 (1986) 4789.
- [99] J. Smith, E. Cook, I. Fortheringham, S. Pheby, R. Derbyshire, M. A. E. Eaton, M. Doel, D. M. J. Lilley, J. F. Pardon, T. Patel, H. Lewis, L. D. Bell, *Nucleic Acids Res.* 10 (1982) 4467.
- [100] I. Ivanov, L. Gigova, E. Jay, *FEBS Lett.* 210 (1987) 56.
- [101] T. Oka, S. Sakamoto, K.-I. Miyoshi, T. Fuwa, K. Yoda, M. Yamasaki, G. Tamura, T. Miyake, *Proc. Natl. Acad. Sci. USA* 82 (1985) 7212.
- [102] T. Kempe, F. Chow, S. M. Peterson, P. Baker, W. Hays, G. Opperman, J. J. L'Italien, G. Long, P. Paulson, *Bio/Technology* 4 (1986) 565.
- [103] F. Kishimoto, H. Gomi, M. Kanaoka, T. Nakatani, A. Ito, T. Katoh, H. Agui, S. Sumida, S. Ogino, *Gene* 45 (1986) 311.
- [104] Y. Ovchinnikov, V. A. Efimov, I. N. Ivanova, S. V. Reverdatto, N. P. Skiba, O. G. Chakhmakhcheva, *Gene* 31 (1984) 65.
- [105] R. Brousseau, R. Scarpulla, W. Sung, H. M. Hsiung, S. A. Narang, R. Wu, *Gene* 17 (1982) 279.
- [106] K.-I. Miyoshi, A. Hasegawa, M. Tomiyama, T. Miyake, *Nucleic Acids Res. Symp. Ser.* 10 (1981) 197.
- [107] A. Cravador, P. Jacobs, A. van Elsen, C. Lacroix, B. Colau, P. van Alphen, A. Herzog, A. Bollen, *Biochimie* 67 (1985) 829.
- [108] H. Seliger, A. Herold, U. Kotschi, J. Lyons, G. Schmidt, F. Eisenbeiß, *Nucleosides & Nucleotides* 6 (1987) 137.
- [109] F. Hein, H. W. Jansen, E. Uhlmann, *Nucleosides & Nucleotides* 6 (1987) 489.
- [110] I. Ivanov, J. Tam, P. Wishart, E. Jay, *Gene* 59 (1987) 223.
- [111] M. Lennick, J. R. Haynes, S.-H. Shen, *Gene* 61 (1987) 103.
- [112] S. A. Rabbani, T. Yasuda, H. P. J. Bennett, W. L. Sung, D. M. Zahab, C. S. Tam, D. Goltzman, G. Hendy, *J. Biol. Chem.* 263 (1988) 1307.
- [113] S. Villa, G. De Fazio, S. Donini, G. Tarchi, U. Canosi, *Eur. J. Biochem.* 171 (1988) 137.
- [114] K. Alton, Y. Stabinski, R. Richards, B. Ferguson, L. Goldstein, B. Altruck, L. Miller, N. Stebbing in E. DeMaeyer, E. Schellekens (Hrsg.): *The Biology of the Interferon System*, Elsevier, Amsterdam 1983, S. 119.
- [115] J. Engels, M. Leineweber, E. Uhlmann, *Angew. Chem.* 96 (1984) 969; *Angew. Chem. Int. Ed. Engl.* 23 (1984) 961.
- [116] E. Jay, D. MacKnight, C. Lutze-Wallace, D. Harrison, P. Wishart, W.-Y. Liu, Y. Asundi, L. Pommeroy-Cloney, J. Rommens, L. Eglington, J. Pawlak, F. Jay, *J. Biol. Chem.* 259 (1984) 6311.
- [117] S. Tanaka, T. Oshima, K. Ohsuye, T. Ono, A. Mizono, A. Ueno, H. Nakazato, M. Tsujimoto, N. Higashi, T. Noguchi, *Nucleic Acids Res.* 11 (1983) 1707.
- [118] P. W. Gray, B. B. Aggarwal, C. V. Benton, T. S. Bringman, W. J. Henzel, J. A. Jerrett, D. W. Leung, B. Moffat, P. Ng, L. P. Svedersky, M. A. Palladino, G. E. Nedwin, *Nature (London)* 312 (1984) 721.
- [119] M. D. Edge, A. R. Greene, G. R. Heathcliffe, V. E. Moore, N. J. Faulkner, R. Camble, N. N. Petter, P. Trueman, W. Schuch, J. Hennam, T. C. Atkinson, C. R. Newton, A. F. Markham, *Nucleic Acids Res.* 11 (1983) 6419.
- [120] S. M. Zurawski, K. Pope, H. Cherwinski, G. Zurawski, *Gene* 49 (1986) 61.
- [121] A. G. Porter, L. D. Bell, J. Adair, G. H. Catlin, J. Clarke, J. A. Davies, K. Dawson, R. Derbyshire, S. M. Doel, *DNA* 5 (1986) 137.
- [122] R. Camble, N. N. Petter, P. Trueman, C. R. Newton, F. J. Carr, R. C. Hockney, V. E. Moore, A. R. Greene, D. Holland, M. D. Edge, *Biochem. Biophys. Res. Commun.* 134 (1986) 1404.
- [123] W. C. Leung, A. P. Liu, X. Wu, S. E. Hasnain, G. Ritzel, G. Jing, in H. Kirchner (Hrsg.): *Biol. Interferon Syst. Proc. TNO-ISIR Meet.*, Elsevier, Amsterdam 1985, S. 419.
- [124] Y. Nagase, N. Nakamura, J. Tohyama, S. Watanabe, H. Ogino, K. Horikoshi, A. Nii, M. Soma, M. Nobuhara, *Nucleic Acids Res. Symp. Ser.* 12 (1983) 83.
- [125] D. C. Humber, M. Allsopp, B. A. Coomber, L. G. D'Urso, C. W. Dykes, A. N. Hobden, M. Mitchell, S. A. Noble, *Nucleosides & Nucleotides* 6 (1987) 413.
- [126] D. P. Williams, K. Parker, P. Bacha, W. Bishai, M. Borowski, F. Genbauffe, T. B. Strom, J. R. Murphy, *Protein Eng.* 1 (1987) 493.
- [127] R. Wetzel, H. L. Heyneker, D. V. Goeddel, P. Jhurani, J. Shapiro, R. Crea, T. L. K. Low, J. E. McClure, G. B. Thurman, A. L. Goldstein, *Biochemistry* 19 (1980) 6096.
- [128] W. Mandelk, K. W. Mollison, T. J. Bolling, B. S. Powell, G. W. Carter, J. L. Fox, *Proc. Natl. Acad. Sci. USA* 82 (1985) 3543.
- [129] G. E. Armah, S. Nishikawa, C. Kondo, K.-I. Tomita, *Nucleic Acids Res. Symp. Ser.* 19 (1988) 165.
- [130] A. Calenda, A. Cordonnier, F. Lederer, D. K. H. Le, J. M. Pleau, *Biootechnol. Lett.* 10 (1988) 155.
- [131] G. T. Mullenbach, A. Tabrizi, R. W. Blacher, K. S. Steimer, *J. Biol. Chem.* 261 (1986) 719.
- [132] W. L. Sung, D. M. Zahab, F.-L. Yao, S. A. Narang, *Nucleic Acids Res.* 14

- (1986) 6159.
- [133] B. S. Sproat, M. J. Gait, *Nucleic Acids Res.* **13** (1985) 2959.
- [134] A. Elmblad, L. Fryklund, L.-O. Heden, E. Holmgren, S. Josephson, M. Lake, B. Löwenadler, G. Palm, A. Skottner-Lundin in *3rd Eur. Congr. Biotechnol. Vol. III*, Verlag Chemie, Weinheim 1984, S. 287.
- [135] G. Buell, M.-F. Schulz, G. Selzer, A. Chollet, N. R. Movva, D. Semon, S. Escanez, E. Kawashima, *Nucleic Acids Res.* **13** (1985) 1923.
- [136] G. Allen, C. A. Paynter, M. D. Wirth, *J. Cell. Sci. Suppl.* **3** (1985) 29.
- [137] M. A. Peters, E. P. Lau, D. L. Snitkin, J. J. van Wyk, L. E. Underwood, W. E. Russel, M. E. Svoboda, *Gene* **35** (1985) 83.
- [138] M. L. Bayne, M. C. Cascieri, B. Kelder, J. Applebaum, G. Chicchi, J. A. Shapiro, F. Pasleau, J. J. Kopchick, *Proc. Natl. Acad. Sci. USA* **84** (1987) 2638.
- [139] P. Denefle, S. Kovarik, J.-D. Guitton, T. Cartwright, J.-F. Mayaux, *Gene* **56** (1987) 61.
- [140] R. Shapiro, J. W. Harper, E. A. Fox, H.-W. Jansen, F. Hein, E. Uhlmann, *Anal. Biochem.* **175** (1988) 450.
- [141] K. N. Nambiar, J. Stackhouse, D. M. Stauffer, W. Poindexter Kennedy, J. K. Eldredge, S. A. Benner, *Science (Washington)* **223** (1984) 1299.
- [142] Y. Jigami, M. Muraki, N. Harada, H. Tanaka, *Gene* **43** (1986) 273.
- [143] L. Bülow, K. Mosbach, *FEBS Lett.* **210** (1987) 147.
- [144] M. Ichihara, E. Ohtsuka, T. Tokunaga, S. Nishikawa, S. Uesugi, T. Tanaka, Y. Aoyama, S. Kikyodani, K. Fujimoto, K. Yanase, K. Fuchimura, H. Morioka, *Proc. Natl. Acad. Sci. USA* **83** (1986) 4695.
- [145] D. P. Gearing, G. McMullen, P. Nagley, *Biochem. Int.* **10** (1985) 907.
- [146] T. Horn, J. S. Allen, M. S. Urdea, *Nucleosides & Nucleotides* **6** (1987) 335.
- [147] J. Phipps, J. J. Michniewicz, F.-L. Lao, S. A. Narang, *Biochem. Biophys. Res. Commun.* **145** (1987) 190.
- [148] S. A. Narang, F.-L. Yao, J. J. Michniewicz, G. Dubuc, J. Phipps, R. L. Samorjai, *Protein Eng.* **1** (1987) 481.
- [149] K. P. Nambiar, J. Stackhouse, S. R. Presnell, S. A. Benner, *Eur. J. Biochem.* **163** (1987) 67.
- [150] S. Nishikawa, H. Morioka, T. Kimura, Y. Ueda, T. Tanaka, S. Uesugi, T. Hakoshima, K.-I. Tomita, E. Ohtsuka, M. Ichihara, *Eur. J. Biochem.* **173** (1988) 389.
- [151] T. Tanaka, S. Kimura, Y. Ota, *Gene* **64** (1988) 257.
- [152] P. S. Vermersch, G. N. Bennett, *DNA* **7** (1988) 243.
- [153] T. Kempe, S. B. H. Kent, F. Chow, S. M. Petersom, W. I. Sundquist, J. J. L'Italien, D. Harbrecht, D. Plunkett, W. J. DeLorbe, *Gene* **39** (1985) 239.
- [154] Z. Hostomsky, J. Smrt, V. Paces, *Gene* **39** (1985) 269.
- [155] J. Tank, X. Zhou, Z. Liang, B. Zon, X. Gao, X. Yang, C. Chen, D. Wang, *Sci. Sin. Ser. B (Engl. Ed.)* **27** (1984) 466.
- [156] S. Tanaka, T. Oshima, K. Ohsue, T. Ono, S. Oikawa, I. Takano, T. Noguchi, K. Kanagawa, N. Minamino, H. Matsuo, *Nucleic Acids Res.* **10** (1982) 1741.
- [157] V. N. Dobrynin, V. G. Korobko, I. V. Severtsova, V. P. Vlasov, N. S. Bystron, M. N. Kolosov, *Biorg. Khim.* **7** (1981) 1745.
- [158] P. Brodin, T. Grundström, T. Hofmann, T. Drakenberg, E. Thulin, S. Forsen, *Biochemistry* **25** (1986) 5371.
- [159] D. M. Roberts, R. Crea, M. Malecha, G. Alvarado-Urbina, R. H. Chiarello, D. M. Watterson, *Biochemistry* **24** (1985) 5090.
- [160] N. Y. Theriault, C.-S. Tomich, W. Wierenga, *Nucleosides & Nucleotides* **5** (1986) 15.
- [161] C. Bergmann, J. Dodd, S. Köhler, E. Fink, H. G. Gassen, *Biol. Chem. Hoppe-Seyler* **367** (1986) 731.
- [162] H. Rink, M. Liersch, P. Sieber, F. Meyer, *Nucleic Acids Res.* **12** (1984) 6369.
- [163] E. Fortkamp, M. Rieger, G. Heisterberg-Moutses, S. Schweitzer, R. Sommer, *DNA* **5** (1986) 511.
- [164] a) Z. Hostomsky, J. Smrt, L. Arnold, Z. Tocik, V. Paces, *Nucleic Acids Res.* **12** (1987) 4849; b) Z. Hostomsky, J. Smrt, *Nucleic Acids Res. Symp. Ser.* **18** (1987) 241.
- [165] E. Auerswald, W. Schröder, M. Kotick, *Biol. Chem. Hoppe-Seyler* **367** (1987) 1413.
- [166] P. Flecker, *Eur. J. Biochem.* **166** (1987) 151.
- [167] M. Strauss, F. O. Bartsch, J. Stollwerk, M. Trstenjak, A. Böhning, H. G. Gassen, W. Machleidt, V. Turk, *Biol. Chem. Hoppe-Seyler* **369 Suppl.** (1988) 209.
- [168] L. Ferretti, S. S. Karnik, H. G. Khorana, M. Nassal, D. D. Oprian, *Proc. Natl. Acad. Sci. USA* **83** (1986) 599.
- [169] M. T. Doel, M. T. Eaton, E. A. Cook, H. Lewis, T. Patel, N. H. Carey, *Nucleic Acids Res.* **8** (1980) 4575.
- [170] V. P. Kumarev, M. I. Rivkin, V. S. Bogachev, L. V. Baranova, V. M. Merkulov, V. N. Rybakov, *FEBS Lett.* **114** (1980) 273.
- [171] J. B. Ohrlogge, P. D. Beremand, D. N. Kuhn, P. E. Parker, *Biochem. Soc. Trans.* **14** (1986) 579.
- [172] K. Miura, Y. Inoue, H. Nakamori, S. Iwai, E. Ohtsuka, M. Ichihara, S. Noguchi, S. Nishimura, *Jpn. J. Cancer Res.* **77** (1986) 45.
- [173] S. Heaphy, M. Singh, M. J. Gait, *Protein Eng.* **1** (1987) 425.
- [174] A. Rosenthal, R. Ulrich, H. Billwitz, R. Frank, H. Blöcker, *Nucleic Acids Res. Symp. Ser.* **18** (1987) 233.
- [175] D. D. Oprian, R. S. Molday, R. J. Kaufman, H. G. Khorana, *Proc. Natl. Acad. Sci. USA* **84** (1987) 8874.
- [176] M. Nassal, T. Mogi, S. S. Karnik, H. G. Khorana, *J. Biol. Chem.* **262** (1987) 9264.
- [177] K. D. Preuss, H. Sachse, H. S. Sharma, G. Hagendorff, K. H. Scheit, *Biol. Chem. Hoppe-Seyler* **368** (1987) 501.
- [178] J. Biernat, H. Hasselmann, B. Hofer, N. Kennedy, H. Köster, *Protein Eng.* **1** (1987) 345.
- [179] J. Biernat, H. Köster, *Protein Eng.* **1** (1987) 353.
- [180] R. Moser, S. Frey, K. Münger, T. Hehlgans, S. Klauser, H. Langen, E.-L. Winnacker, R. Mertz, B. Gutte, *Protein Eng.* **1** (1987) 339.
- [181] D. J. Ecker, T. R. Butt, J. Marsh, E. J. Sternberg, N. Margolis, B. P. Monia, S. Jonnalagadda, M. I. Khan, P. L. Weber, L. Mueller, S. T. Crooke, *J. Biol. Chem.* **262** (1987) 14213.
- [182] V. M. Reyes, J. Abelson, *Anal. Biochem.* **166** (1987) 90.
- [183] J.-M. Masson, P. Meuris, M. Grunstein, J. Abelson, J. H. Miller, *Proc. Natl. Acad. Sci. USA* **84** (1987) 6815.
- [184] D. J. Ecker, M. I. Khan, J. Marsh, T. R. Butt, S. T. Crooke, *J. Biol. Chem.* **262** (1987) 3524.
- [185] S. E. Adams, I. D. Johnson, M. Braddock, A. J. Kingsman, R. M. Edwards, *Nucleic Acids Res.* **16** (1988) 4287.
- [186] J.-H. Lee, J. L. Weickmann, R. K. Koduri, P. Ghosh-Dastidar, K. Saito, L. C. Blair, T. Date, J. S. Lai, S. M. Hollenberg, R. L. Kendall, *Biochemistry* **27** (1988) 5101.
- [187] T. P. Sakmar, H. G. Khorana, *Nucleic Acids Res.* **16** (1988) 6361.
- [188] A. G. Stewart, H. Richards, S. Roberts, J. Warwick, K. Edwards, L. Bell, J. Smith, R. Derbyshire, *Nucleic Acids Res.* **11** (1983) 6597.
- [189] E. Uhlmann, J. Engels, *Tetrahedron Lett.* **27** (1986) 1023.
- [190] E. Uhlmann, J. Engels, *Chem. Scr.* **26** (1986) 217.
- [191] H. G. Khorana, *Pure Appl. Chem.* **17** (1968) 349.
- [192] N. Y. Theriault, J. B. Carter, S. P. Pulaski, *BioTechniques* **6** (1988) 470.
- [193] J. J. Rossi, R. Kierzek, T. Huang, P. A. Walker, K. Itakura, *J. Biol. Chem.* **257** (1982) 9226.
- [194] R. C. Scarpulla, S. Narang, R. Wu, *Anal. Biochem.* **121** (1982) 356.
- [195] O. G. Chakhmakhcheva, V. A. Efimov, O. V. Mirskikh, R. V. Reverdatto, A. A. Buryakova, Y. A. Ovchinnikov, *Chem. Scr.* **26** (1986) 31.
- [196] J. J. Rossi, X. Soberon, Y. Marumoto, J. McMahon, K. Itakura, *Proc. Natl. Acad. Sci. USA* **80** (1983) 3203.
- [197] E. Uhlmann, F. Hein, *Nucleic Acids Res. Symp. Ser.* **18** (1987) 237.
- [198] E. Uhlmann, *Gene* **71** (1988) 29.
- [199] D. H. Hare, B. R. Reid, *Biochemistry* **25** (1986) 5341.
- [200] S. Ikuta, R. Chattopadhyaya, *Biochemistry* **25** (1986) 4840.
- [201] R. Chattopadhyaya, S. Ikuta, K. Grzeskowiak, R. E. Dickerson, *Nature (London)* **334** (1988) 175.
- [202] a) F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. USA* **74** (1977) 5462; b) Nobel-Aufsat: F. Sanger, *Angew. Chem.* **93** (1981) 937.
- [203] F. Hein, H. W. Jansen, E. Uhlmann, *Nucleosides & Nucleotides* **7** (1988) 497.
- [204] W. H. McClain, K. Foss, K. L. Mittelstadt, J. Schneider, *Nucleic Acids Res.* **14** (1986) 6770.
- [205] T. Grundström, W. M. Zenke, M. Wintzterith, H. W. D. Matthes, A. Staub, P. Chambo, *Nucleic Acids Res.* **13** (1985) 3305.
- [206] S. A. Narang, G. Dubuc, F. L. Yao, J. J. Michniewicz, *Biochem. Biophys. Res. Commun.* **134** (1986) 407.
- [207] O. G. Chakhmakhcheva, V. A. Efimov, Y. A. Ovchinnikov, *Nucleosides & Nucleotides* **6** (1987) 321.
- [208] A. R. Oliphant, A. L. Nussbaum, K. Struhl, *Gene* **44** (1986) 177.
- [209] K. M. Derbyshire, J. J. Salvo, N. D. F. Gringley, *Gene* **46** (1986) 145.
- [210] C. A. Brennan, R. I. Gumpert, *Nucleic Acids Res.* **13** (1985) 8665.
- [211] R. Cosstick, L. W. McLaughlin, F. Eckstein, *Nucleic Acids Res.* **12** (1984) 1791.
- [212] D. Zeh, H. Seliger, *Nucleosides & Nucleotides* **6** (1987) 485.
- [213] M. J. Zoller, M. Smith, *Nucleic Acids Res.* **10** (1982) 6487.
- [214] H. J. Fritz, *Biol. Chem. Hoppe-Seyler* **367 Suppl.** (1986) 86.
- [215] a) W. Mandecki, *Proc. Natl. Acad. Sci. USA* **83** (1986) 7177; b) W. Mandecki, T. J. Bolling, *Gene* **68** (1988) 101; c) M. A. Hayden, W. Mandecki, *DNA* **7** (1988) 571.
- [216] S. Pochet, B. Arcangoli, T. Huynh-Dinh, *Nucleic Acids Res.* **16** (1988) 1619.
- [217] H. Grosjean, W. Fiers, *Gene* **18** (1982) 199.
- [218] T. Ikemura, *J. Mol. Biol.* **158** (1982) 573.
- [219] T. Togunaga, S. Iwai, H. Gomi, K. Kodama, E. Ohtsuka, M. Ichihara, O. Chisaka, K. Matsubara, *Gene* **39** (1985) 117.
- [220] a) T. Maruyama, T. Gojobori, S.-I. Aota, T. Ikemura, *Nucleic Acids Res.* **14** (1986) Suppl. r151; b) S.-I. Aota, T. Gojobori, T. Maruyama, T. Ikemura, *ibid.* **16** (1988) Suppl. r315.
- [221] G. O'Neill, S. H. Goh, R. A. J. Warren, D. G. Kilburn, R. C. Miller Jr., *Gene* **44** (1986) 325.
- [222] J. Vieira, J. Messing, *Gene* **19** (1982) 259.
- [223] J. Messing, B. Gronenborn, B. Müller-Hill, P. H. Hofsneider, *Proc. Natl. Acad. Sci. USA* **74** (1977) 3642.
- [224] a) *Nucleic Acids Res.* **10** (1982) 1–456; b) *ibid.* **12** (1984) 1–853; c) *ibid.* **14** (1986) 1–621; d) *ibid.* **16** (1988) 1655–1877.
- [225] S. R. Presnell, S. A. Benner, *Nucleic Acids Res.* **16** (1988) 1693.
- [226] a) R. J. Roberts, *Nucleic Acids Res.* **13** (1985) Suppl. r165; b) *ibid.* **15** (1987) Suppl. r189.
- [227] C. Kessler, H. J. Höltke, *Gene* **47** (1986) 1.
- [228] R. J. Roberts, *Nucleic Acids Res.* **16** (1988) Suppl. r271.

- [229] A. C. Looman, J. Bodlaender, M. de Gruyter, A. Vogelaar, P. H. van Knippenberg, *Nucleic Acids Res.* 14 (1986) 5481.
- [230] L.-H. Tessier, P. Sondermeyer, T. Faure, D. Dreyer, A. Benavente, D. Villeval, M. Courtney, J.-P. Lecocq, *Nucleic Acids Res.* 12 (1984) 7663.
- [231] M. Zucker, P. Stiegler, *Nucleic Acids Res.* 9 (1981) 133.
- [232] R. Nussinov, G. Pieczenik, J. R. Riggs, D. J. Kleitman, *SIAM J. Appl. Math.* 35 (1978) 68.
- [233] C. Papanicolaou, M. Gouy, J. Ninio, *Nucleic Acids Res.* 12 (1984) 31.
- [234] N. Lee, J. Cozzitorto, N. Wainwright, D. Testa, *Nucleic Acids Res.* 12 (1984) 6797.
- [235] a) P. R. Szoka, A. B. Schreiber, H. Chan, J. Murthy, *DNA* 5 (1986) 11; b) M. Landon, *Methods Enzymol.* 47 (1977) 145.
- [236] J. K. Blodgett, G. M. Loudon, K. D. Collins, *J. Am. Chem. Soc.* 107 (1985) 4305.
- [237] C. Lazure, N. G. Seidah, D. Pelaprat, M. Chretien, *Can. J. Biochem. Cell. Biol.* 61 (1983) 501.
- [238] K. Nagai, H. C. Thoegersen, *Nature (London)* 309 (1984) 810.
- [239] W. L. Sung, F.-L. Yao, D. M. Zahab, S. A. Narang, *Proc. Natl. Acad. Sci. USA* 83 (1986) 561.
- [240] S.-H. Shen, *Proc. Natl. Acad. Sci. USA* 81 (1984) 4627.
- [241] B. Blobel, B. Dobberstein, *J. Cell. Biol.* 67 (1975) 835.
- [242] J. Dodt, T. Schmitz, T. Schäfer, C. Bergmann, *FEBS Lett.* 202 (1986) 373.
- [243] G. L. Grey, J. S. Baldridge, K. S. McKeown, H. L. Heyneker, C. Nan Chang, *Gene* 39 (1985) 247.
- [244] E. Uhlmann, J. Engels, W. Wetekam, *Nucleosides & Nucleotides* 4 (1985) 259.
- [245] B. Nilsson, E. Holmgren, S. Josephson, S. Gatenbeck, L. Philipson, M. Uhlen, *Nucleic Acids Res.* 13 (1985) 1151.
- [246] K. Ohmura, K. Nakamura, H. Yamazaki, T. Shiroza, K. Yamane, Y. Jigami, H. Tanaka, K. Yoda, M. Yamasaki, G. Tamura, *Nucleic Acids Res.* 13 (1984) 5307.
- [247] K. M. Zsebo, H.-S. Lu, J. C. Fieschko, L. Goldstein, J. Davis, K. Duker, S. V. Suggs, P. H. Lai, G. A. Bitter, *J. Biol. Chem.* 261 (1986) 5858.
- [248] S. D. Emr, R. Schekman, M. C. Flessel, J. Thorner, *Proc. Natl. Acad. Sci. USA* 80 (1983) 7080.
- [249] M. Smith, *Annu. Rev. Genet.* 19 (1985) 423.
- [250] W. Kramer, V. Drutsa, H. W. Jansen, B. Kramer, M. Pflugfelder, H. J. Fritz, *Nucleic Acids Res.* 12 (1984) 9441.
- [251] M. J. Zoller, M. Smith, *DNA* 3 (1984) 479.
- [252] W. Kramer, K. Schughart, H. J. Fritz, *Nucleic Acids Res.* 10 (1982) 6475.
- [253] P. Carter, H. Bedouelle, G. Winter, *Nucleic Acids Res.* 13 (1985) 4431.
- [254] T. A. Kunkel, *Proc. Natl. Acad. Sci. USA* 82 (1985) 488.
- [255] J. W. Taylor, J. Ott, F. Eckstein, *Nucleic Acids Res.* 13 (1985) 8764.
- [256] M. D. Matteucci, H. L. Heyneker, *Nucleic Acids Res.* 11 (1983) 3113.
- [257] J. A. Wells, M. Vasser, D. B. Powers, *Gene* 34 (1985) 315.
- [258] S. W. Cline, M. Yarus, P. Wier, *DNA* 5 (1986) 37.
- [259] D. A. Estell, T. P. Graycar, J. A. Wells, *J. Biol. Chem.* 260 (1985) 6518.
- [260] P. G. Thomas, A. J. Russel, A. R. Fersht, *Nature (London)* 318 (1985) 375.
- [261] K. T.-K. Lau, A. R. Fersht, *Nature (London)* 326 (1987) 811.
- [262] Y.-R. Hsu, B. Ferguson, M. Narachi, R. M. Richards, Y. Stabinsky, N. K. Alton, N. Stebbing, T. Arakawa, *J. Interferon Res.* 6 (1986) 663.
- [263] F. E. Cohen, P. A. Kosen, I. D. Kuntz, L. B. Epstein, T. L. Ciardelli, K. A. Smith, *Science (Washington)* 234 (1986) 349.
- [264] A. Wang, S.-D. Lu, D. F. Mark, *Science (Washington)* 224 (1984) 1431.
- [265] J. J. Huang, R. C. Newton, R. Horuk, J. B. Matthew, M. Covington, K. Pezzella, Y. Lin, *FEBS Lett.* 223 (1987) 294.
- [266] T. Arakawa, T. Boone, J. M. Davis, W. C. Kenney, *Biochemistry* 25 (1986) 8274.
- [267] J. W. Engels, J. Glauder, H. Müllner, D. Tripier, E. Uhlmann, W. Wetekam, *Protein Eng.* 1 (1987) 195.
- [268] A. Bradbury, M. D. F. Finnie, D. G. Smyth, *Nature (London)* 298 (1982) 686.
- [269] H. J. Fritz, *Nucleic Acids Res. Spec. Publ.* 4 (1978) s243.
- [270] V. N. Dobrynin, V. G. Korobko, I. V. Severtsova, N. S. Bystrov, S. A. Chuvilo, M. N. Kolosov, *Nucleic Acids Res. Symp. Ser.* 7 (1980) 365.
- [271] P. L. DeHaseth, R. A. Goldman, C. L. Cech, M. H. Caruthers, *Nucleic Acids Res.* 11 (1983) 773.
- [272] J. Rommens, D. MacKnight, L. Pomeroy-Cloney, E. Jay, *Nucleic Acids Res.* 11 (1983) 5921.
- [273] E. Jay, J. Rommens, L. Pomeroy-Cloney, D. MacKnight, C. Lutze-Wallace, P. Wishart, D. Harrison, W. Y. Lui, V. Asundi, M. Dawood, F. Jay, *Proc. Natl. Acad. Sci. USA* 81 (1984) 2290.
- [274] J. D. Windass, C. R. Newton, J. De Maeyer-Guignard, V. E. Moore, A. F. Markham, M. D. Edge, *Nucleic Acids Res.* 10 (1982) 6639.
- [275] F. Jacob, J. Monod, *J. Mol. Biol.* 3 (1961) 318.
- [276] J. R. Sadler, H. Sasmor, J. L. Betz, *Proc. Natl. Acad. Sci. USA* 80 (1983) 6785.
- [277] A. Simons, D. Tils, B. von Wilcken-Bergmann, B. Müller-Hill, *Proc. Natl. Acad. Sci. USA* 81 (1984) 1624.
- [278] J. Brosius, A. Holy, *Proc. Natl. Acad. Sci. USA* 81 (1984) 6929.
- [279] G. Simons, E. Remaut, B. Allet, R. Devos, W. Fiers, *Gene* 28 (1984) 55.
- [280] J.-I. Flock, I. Fotheringham, J. Light, L. Bell, R. Derbyshire, *MGG Mol. Genet.* 195 (1984) 246.
- [281] J. Shine, L. Dalgarno, *Nature (London)* 254 (1975) 34.
- [282] E. Dworkin-Rastl, P. Swetly, M. B. Dworkin, *Gene* 21 (1983) 237.
- [283] P. Stanssens, E. Remaut, W. Fiers, *Gene* 36 (1985) 211.
- [284] D. M. Marquis, J. M. Smolec, D. H. Katz, *Gene* 42 (1986) 175.
- [285] R. E. Offord, *Protein Eng.* 1 (1987) 151.
- [286] M. Mutter, *Angew. Chem. 97* (1985) 639; *Angew. Chem. Int. Ed. Engl.* 24 (1985) 639; S. Vuilleumier, M. Mutter, *ibid.* 101 (1989) 551 bzw. 28 (1989) 535.
- [287] R. Moser, S. Klauser, T. Leist, H. Langen, T. Epprecht, B. Gutte, *Angew. Chem. 97* (1985) 737; *Angew. Chem. Int. Ed. Engl.* 24 (1985) 719.
- [288] V. L. Rath, R. J. Fletterick, *Cell* 49 (1987) 583.
- [289] T. L. Blundell, B. L. Sibanda, M. J. E. Sternberg, J. Thornton, *Nature (London)* 326 (1987) 347.
- [290] W. V. Shaw, *Biochem. J.* 246 (1987) 1.
- [291] T. Blundell, D. Carney, S. Gardner, F. Hayes, B. Howlin, T. Hubbard, J. Overington, D. A. Singh, B. L. Sibanda, M. Sutcliffe, *Eur. J. Biochem.* 172 (1988) 513.
- [292] R. Jaenicke, *Prog. Biophys. molec. Biol.* 49 (1987) 117; *Angew. Chem. 96* (1984) 385; *Angew. Chem. Int. Ed. Engl.* 23 (1984) 395.
- [293] J. van Brunt, *Bio/Technology* 6 (1988) 655.
- [294] J. F. Gibrat, J. Garnier, B. Robson, *J. Mol. Biol.* 198 (1987) 425.
- [295] L. Regan, W. F. DeGraff, *Science (Washington)* 241 (1988) 976.