

Das genetische Gedächtnis hat seinen Sitz in spezifischen Nucleinsäuremolekülen. Die Information ist in Form einer linearen Sequenz aus vier Arten von Basen chiffriert, der Sequenzen aus 20 Arten von Aminosäuren im Protein entsprechen. Die Übersetzung von der Nucleinsäure zum Protein vollzieht sich in sequenzieller Weise nach einem systematischen Code mit relativ einfachen Gesetzen. Jede Nucleinsäureinheit definiert die Molekülspezies, die ausgewählt werden soll, ihre Position relativ zum vorher ausgewählten Molekül und den Zeitunterschied zwischen diesem und dem vorhergehenden Vorgang. Die Nucleinsäure dient daher sowohl als Matrize für andere Moleküle als auch als biologische Uhr.

Die Information wird in Form eines eindimensionalen Stranges chiffriert und dechiffriert. Das Produkt der Übersetzung, das Polypeptid, faltet sich dann in einer spezifischen Weise, die durch die Aminosäuresequenz vorherbestimmt ist, zu einem komplexen dreidimensionalen Protein.

Konzept eines Gen-Protein-Codes

Die Fortschritte auf dem Gebiet der biochemischen Genetik gehen auf die Anstrengungen von Forschern auf praktisch jedem Gebiet der Wissenschaften zurück. Zu den Meilensteinen gehören die Identifizierung der DNA als genetisches Material durch *Avery*, *MacLeod* und *McCarty*^[1], das Ein-Gen-ein-Protein-Konzept von *Beadle* und *Tatum*^[2] sowie die Pionierexperimente von *Brachet*^[3] sowie *Caspersson*^[4] über den Zusammenhang zwischen RNA und Protein-Synthese. Abgesehen davon wurde das Rätsel der Protein-Synthese Schritt für Schritt gelöst, und es wurden in-vitro-Systeme für die Protein- und Nucleinsäure-Synthese ausgearbeitet.

Das Konzept eines einfachen Codes, der die Basensequenz von Nucleinsäuren und die Aminosäuresequenz von Proteinen in Beziehung setzt, geht auf drei unabhängige Quellen Anfang der fünfziger Jahre zurück. Die Intuition war spektakulär, wenn man sie im

Zusammenhang mit den damals verfügbaren Informationen betrachtet. *Caldwell* und *Hinshelwood*^[5] vermuteten, daß die RNA aus fünf Arten von Einheiten zusammengesetzt ist, den vier Basen und dem Ribosephosphat, und daß zwei benachbarte Einheiten in der RNA einer Aminosäure im Protein entsprechen. *Dounce*^[6] schlug vor, daß drei benachbarte Basen in der RNA einer Aminosäure im Protein entsprechen sollten. Außerdem wurden die Konzepte für die Polarität der Translation und der Aktivierung der Aminosäuren in bemerkenswerten Einzelheiten formuliert. *Dounces* Überzeugung, daß Matrizen für die Synthese von Protein nötig seien, entstand während seines Rigosums, als er von *James Sumner* zu einer Äußerung über das Problem aufgefordert wurde, wie Proteine andere Proteine synthetisieren.

Zur gleichen Zeit vermutete *George Gamow*^[7], daß ein DNA-Doppelstrang Bindungsstellen für Aminosäuren besitze, wobei jede durch ein Basenpaar und angrenzende nichtkomplementäre Basen gegenüberliegender DNA-Stränge bestimmt werde. Ihm kam diese Idee beim Lesen des Aufsatzes von *Watson* und *Crick* über die Basenpaarung in der DNA^[8]. Weitere Überlegungen über die Natur des Codes wurden von vielen Forschern gegen Ende der fünfziger Jahre angestellt (vgl. die neuere Übersicht von *Woese*^[9]).

Obwohl das Konzept der RNA als Matrize für das Protein allgemein anerkannt war, fehlten doch direkte biochemische Beweise. *Hersheys*^[10] Ergebnis, daß eine RNA-Fraktion in mit T2-Bakteriophagen infizierten *E. coli* schnell synthetisiert und dann abgebaut wird, und der Hinweis von *Volkin* und *Astrachan*^[11], daß diese RNA-Fraktion in der Zusammensetzung eher der Phagen-DNA als der Coli-DNA entspreche, waren erregend, weil sie vermuten ließen, daß die instabile RNA-Fraktion die Matrize für die Synthese des Phagen-Proteins sein könne.

Ich stürzte mich einige Zeit nach meiner Promotion in das Problem der Protein-Synthese. In meinem fortgeschrittenen Biochemie-Studium leitete mich *James Hogg* an; direkt nach der Promotion arbeitete ich mit *DeWitt Stetten* und mit *William Jakoby* an den National Institutes of Health. Dann kam ich in *Gordon Tompkins* Abteilung und begann mich mit den Schrit-

[*] Prof. Dr. M. Nirenberg
Laboratory of Biochemical Genetics, National Heart Institute,
National Institutes of Health, Bethesda, Maryland 20014
(USA)

[**] Copyright © The Nobel Foundation 1969. — Wir danken der Nobel-Stiftung, Stockholm, für die Genehmigung zum Druck dieser Übersetzung.

[1] O. T. Avery, C. M. MacLeod u. M. McCarty, J. exp. Medicine 79, 137 (1944).

[2] G. W. Beadle u. E. L. Tatum, Proc. nat. Acad. Sci. USA 27, 499 (1941).

[3] J. Brachet, Arch. Biol (Liège) 53, 207 (1942).

[4] T. Caspersson, Naturwissenschaften 29, 33 (1941).

[5] P. C. Caldwell u. C. Hinshelwood, J. chem. Soc. (London) 1950, 3156.

[6] A. L. Dounce, Enzymologia 15, 251 (1952).

[7] G. Gamow, Nature (London) 173, 318 (1954).

[8] J. D. Watson u. F. H. C. Crick, Nature (London) 171, 737 (1953).

[9] C. R. Woese: Genetic Code. Harper und Row Publ., New York 1967, Kap. 2.

[10] A. D. Hershey, J. Dixon u. M. Chase, J. gen. Physiol. 36, 777 (1953).

[11] E. Volkin u. L. Astrachan, Virology 2, 149 (1956).

ten zu beschäftigen, die DNA, RNA und Protein verbinden. Die Ausbildung in Enzymologie und die anregende Umgebung beeinflussten den zukünftigen Verlauf meiner Arbeit sehr stark.

Umfassende Studien über den Verlauf der Protein-Synthese hatten viele Ergebnisse gebracht, und es schien damals so, als ob in den kommenden zehn Jahren die Synthese eines Enzyms in Zellextrakten möglich sein sollte. Da ein solches System viele Möglichkeiten zum Studium des Informationsflusses von der Nucleinsäure zum Protein bieten würde, entschied ich mich, über die zellfreie Synthese von Penicillinase zu arbeiten. Pollock und Mitarbeiter [12, 13], die viele Informationen über die Regulation der Penicillinase-Synthese in vivo erhalten hatten, konnten zeigen, daß das Molekulargewicht des Enzyms relativ niedrig ist und daß es kein Cystein enthält. Es schien, als könne man die Synthese der Proteine, die Cystein benötigen, selektiv hemmen, und gleichzeitig die Penicillinase-Synthese in vitro durch Zusatz von Nucleinsäure-Matrizen zum Zellextrakt stimulieren.

In den nächsten beiden Jahren studierte ich die Eigenschaften des Systems, besonders den Effekt der Reaktionsbedingungen, der Nucleinsäure und anderer Faktoren auf die Geschwindigkeit der zellfreien Protein-Synthese. In dieser Zeit veröffentlichten Lamborg und Zamecnik [14], Tissieres, Schlessinger und Gros [15] sowie andere äußerst interessante Ergebnisse über die Protein-Synthese in *E. coli*-Extrakten. Tissieres, Schlessinger und Gros [15], Kameyama und Novelli [16] sowie Nisman und Fukuhara [17] berichteten, daß DNase den in-vitro-Einbau von Aminosäuren in Protein hemmt. Auch ich hatte dieses Phänomen beobachtet und war daran sehr interessiert, weil die Resultate klar darauf hindeuteten, daß die zellfreie Protein-Synthese letztlich von DNA-Matrizen abhängig war.

Bei diesen Arbeiten wurde Heinrich Matthaei mein Mitarbeiter. Wir konnten bald zeigen, daß aus Ribosomen präparierte RNA den Aminosäure-Einbau in Protein stimulierte [18]. Allerdings wurden Aminosäuren auch ohne zugesetzte RNA schnell in Protein eingebaut [19], so daß es schwierig war, den RNA-abhängigen Anteil der Protein-Synthese zu bestimmen. Dieses Problem wurde durch Inkubation der *E. coli*-Extrakte mit den für die Protein-Synthese nötigen Komponenten unter DNase-Zusatz gelöst, um die

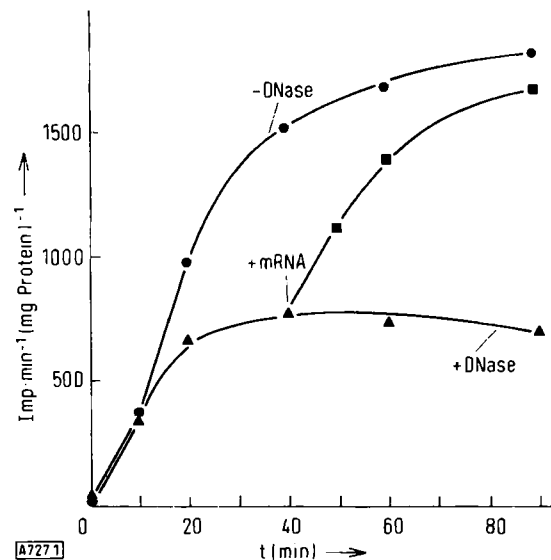


Abb. 1. Einfluß von DNase und mRNA auf den Einbau von ^{14}C -Valin in Protein in *E. coli*-Extrakt.

● keine Zugabe; ▲ mit $10\ \mu\text{g}$ DNase pro ml Reaktionsgemisch; ■ mit $10\ \mu\text{g}$ DNase und $0.5\ \text{mg}$ einer mRNA-Fraktion pro ml Reaktionsgemisch.

Menge der endogenen RNA-Matrizen zu verringern (Abb. 1). Nach kurzer Inkubationszeit hört die Protein-Synthese auf; eine weitere Protein-Synthese findet nur nach Zusatz von Matrizen-RNA statt. Transfer-RNA ersetzt die Matrizen-RNA nicht.

Es wurde eine Analysen-Methode ausgearbeitet, die den Zeitbedarf für jedes Experiment auf ungefähr ein Viertel verringerte. Sie beruhte auf der Filtration der ^{14}C -Protein-Niederschläge. Wir präparierten RNA aus vielen Quellen, um ihre Spezifität und Aktivität als Matrize für die Protein-Synthese zu bestimmen. RNA aus Hefe, Ribosomen und Tabakmosaikvirus stimulierte den Einbau jeder getesteten Aminosäure-Spezies stark. Im Gegensatz dazu stimulierte Poly-U ziemlich spezifisch den Einbau von Phenylalanin in Protein, wobei Polyphenylalanin entstand. Einzelsträngige Poly-U war eine aktive Matrize für den Phenylalanin-Einbau, während Doppel- und Tripel-Stränge von Poly-U/Poly-A-Helices nicht als Matrize für die Protein-Synthese wirkten [18].

Diese Resultate zeigten, daß RNA eine Matrize für Protein ist, daß U-Reste in Poly-U dem Phenylalanin im Protein entsprechen und daß die Übersetzung der mRNA sowohl durch die Primär- als auch durch die Sekundärstruktur der RNA beeinflusst wird.

1961 war die Rolle der tRNA noch umstritten. Die meisten Forscher nahmen zwar an, daß die tRNA an der Protein-Synthese beteiligt ist, aber es fehlte noch der direkte Beweis, daß die tRNA für diesen Prozeß nötig ist. Lipmann und Nathans gaben uns freundlicherweise ein gereinigtes Transfer-Enzym-Präparat, und damit fanden wir, daß Phe-tRNA ein obligatorisches Intermediat der Polyphenylalanin-Synthese ist, und daß Transfer-Enzyme und GTP ebenfalls für die Synthese dieses Polypeptids nötig sind [20].

[20] M. W. Nirenberg, J. H. Matthaei u. O. W. Jones, Proc. nat. Acad. Sci. USA 48, 104 (1962).

[12] M. R. Pollock, Proc. Roy. Soc. (London), Ser. B. 148, 340 (1958).

[13] M. R. Pollock in I. C. Gunsalus u. R. Y. Stanier: The Bacteria. Academic Press, New York 1962, Bd. 4, S. 121 ff.

[14] M. R. Lamborg u. P. C. Zamecnik, Biochim. biophysica Acta 42, 206 (1960).

[15] A. Tissieres, D. Schlessinger u. F. Gros, Proc. nat. Acad. Sci. USA 46, 1450 (1960).

[16] T. Kameyama u. G. D. Novelli, Biochem. biophysic. Res. Commun. 2, 393 (1960).

[17] B. Nisman u. H. Fukuhara, C. R. hebd. Séances Acad. Sci. 249, 2240 (1959).

[18] M. W. Nirenberg u. J. H. Matthaei, Proc. nat. Acad. Sci. USA 47, 1588 (1961).

[19] J. H. Matthaei u. M. W. Nirenberg, Proc. nat. Acad. Sci. USA 47, 1580 (1961).

Basenzusammensetzung der Codons

Der genetische Code wurde in zwei experimentellen Phasen in annähernd sechs Jahren entziffert. Während der ersten Phase wurden die Basenzusammensetzung der Codons und die allgemeine Natur des Codes durch Steuerung der Protein-Synthese mit statistisch geordneten RNA-Matrizen, die verschiedene Basenkombinationen enthielten, erforscht. Solche Polymeren wurden mit Hilfe der Polynucleotid-Phosphorylase synthetisiert, die *Grunberg-Manago*, *Ortiz* und *Ochoa*^[21] entdeckt hatten.

Die von *Ochoa* und Mitarbeitern^[23] und von uns^[24] erhaltenen Daten sind in Tabelle 1 zusammengestellt. Nur Polynucleotide, die eine für den Einbau einer bestimmten Aminosäure in das Protein nötige Mindestzahl von Basen-Spezies enthalten, sind aufgeführt. Poly-U, Poly-C und Poly-A stimulieren den Einbau

Tabelle 1. Mindestanzahl von Basen-Spezies für mRNA-Codons. Es ist die Spezifität statistisch geordneter Polynucleotid-Matrizen bei der Stimulation des Einbaus von Aminosäuren in Proteine in Gegenwart von *E. coli*-Extrakten angegeben. Viele Aminosäuren, die auf Polymere aus zwei oder mehr Basen-Spezies ansprechen, sind weggelassen.

Polynucleotide	Aminosäuren			
U	Phe			
C	Pro			
A	Lys			
G	—			
UC	Leu	Ser		
UA	Leu	Tyr	Ile	Asn
UG	Leu	Val	Cys	Trp
CA	His	Thr	Gln	Asn
CG	Arg	Ala		
AG	Arg	Glu		
UAG	Asp	Met		
CAG	Asp	Ser		

von Phenylalanin, Prolin bzw. Lysin. Für Poly-G konnte keine Matrizen-Aktivität gefunden werden. In späteren Arbeiten wiesen *Maxine Singer*, *Bill Jones* und ich nach, daß G-reiche Poly-(U, G)-Präparate wegen ihres hohen Grades an Sekundärstruktur in Lösung nicht als Matrizen für die Protein-Synthese wirken^[22].

Poly-(U, C), Poly-(C, G) und Poly-(A, G) sind Matrizen für zwei zusätzliche Aminosäuren pro Polynucleotid, während Poly-(U, A), Poly-(U, G) und Poly-(C, A) als Matrizen für vier zusätzliche Aminosäuren pro Polynucleotid dienen. Jedes Polynucleotid, das aus drei Basen-Spezies zusammengesetzt ist, fungiert als Matrize für zehn oder mehr Aminosäuren.

Statistisch geordnete Polynucleotide, die aus 1, 2, 3 oder 4 Arten von Basen zusammengesetzt sind, enthalten 1, 8, 27 bzw. 64 Arten von Triplets. Die relative

Häufigkeit jedes Triplets kann leicht berechnet werden, wenn die Basenzusammensetzung des statistisch geordneten Polynucleotids bekannt ist. Man kann sowohl die *Arten* der Basen, die einer Aminosäure entsprechen, als auch deren *Zahl* ableiten, da die Menge jeder einzelnen Aminosäure, die ins Protein eingebaut wird, vom Zusatz des Polynucleotid-Präparates abhängt, und weil dessen Basenzusammensetzung experimentell bestimmt werden kann. Auf diese Weise wurden die Basenzusammensetzungen von ungefähr 50 Codons den Aminosäuren zugeordnet^[23,24]. Die Ergebnisse zeigten, daß einer Aminosäure mehrere Codons entsprechen können und demnach der Code hochgradig degeneriert ist. In den meisten Fällen unterscheiden sich synonyme Codons nur durch eine Base. Man nahm daher an, daß die nichtvariablen Basen die gleichen Positionen innerhalb synonyme Wörter einnehmen. Durch genetische Studien wiesen *Crick*, *Barnett*, *Brenner* und *Watts-Tobin*^[25] nach, daß der Code ein Triplett-Code ist; die biochemischen Studien stützten dieses Ergebnis. Die Analyse des Hüll-Proteins von Tabakmosaikvirus-Mutanten bewies, daß in der mRNA Triplets in nicht überlappende Weise übersetzt werden, weil der Ersatz einer Base durch eine andere in der mRNA gewöhnlich nur zum Austausch einer einzigen Aminosäure im Protein führte^[26].

Basensequenz der Codons

Obwohl die Basenzusammensetzung der Codons bestimmt war, so war doch die *Reihenfolge* der Basen innerhalb der Codons nicht bekannt. Wir prüften viele potentielle Methoden für die Bestimmung der Basen-Sequenz in den Codons. Einen Schlüssel für die Lösung des Problems bot das wichtige Ergebnis von *Arlinghaus*, *Favelukes* und *Schweef*^[27] sowie von *Kaji* und *Kaji*^[28], daß Phe-tRNA in Abhängigkeit von Poly-U an Ribosomen gebunden wird, bevor Peptidbindungen gebildet werden. Möglicherweise würden Trinucleotide oder Hexanucleotide mit bekannter Basen-Sequenz ebenfalls die Bindung von Aminoacyl-tRNA an Ribosomen bewirken. Um diese Möglichkeit zu probieren, arbeiteten *Philip Leder* und ich eine Schnellmethode zur Trennung ribosomengebundener AA-tRNA von ungebundener AA-tRNA aus, die auf der selektiven Retention des ribosomalen Zwischenproduktes auf Cellulosenitrat-Filtern beruhte, und fanden dann, daß Trinucleotide als spezifische Matrize für die Bindung von AA-tRNA an Ribosomen wirkten^[29]. Wie Tabelle 2 zeigt, stimuliert das Trinucleotid AAA die Bindung von Lys-tRNA an Ribosomen und ist ebenso aktiv als Matrize für Lys-tRNA wie das

[21] M. Grunberg-Manago, P. J. Ortiz u. S. Ochoa, *Biochim. biophysica Acta* 20, 259 (1956).

[22] M. Singer, O. W. Jones u. M. W. Nirenberg, *Proc. nat. Acad. Sci. USA* 49, 392 (1963).

[23] J. F. Speyer, P. Lengyel, C. Basilio, A. J. Wahba, R. S. Gardner u. S. Ochoa, *Cold Spring Harbor Sympos. quantit. Biol.* 28, 559 (1963).

[24] M. W. Nirenberg, O. W. Jones, P. Leder, B. F. C. Clark, W. S. Sly u. S. Pestka, *Cold Spring Harbor Sympos. quantit. Biol.* 28, 549 (1963).

[25] F. H. C. Crick, L. Barnett, S. Brenner u. R. J. Watts-Tobin, *Nature (London)* 192, 1227 (1961).

[26] H. G. Wittmann u. B. Wittmann-Leibold, *Cold Spring Harbor Sympos. quantit. Biol.* 28, 589 (1963).

[27] R. Arlinghaus, G. Favelukes u. R. Schweef, *Biochem. biophys. Res. Commun.* 11, 92 (1963).

[28] A. Kaji u. H. Kaji, *Biochem. biophys. Res. Commun.* 13, 186 (1963).

[29] M. W. Nirenberg u. P. Leder, *Science* 145, 1399 (1964).

Tabelle 2. Einfluß von Oligo-A-Präparationen auf die Bindung von Lys-tRNA aus *E. coli* an Ribosomen. Die Bestimmungsmethode für die AA-tRNA-Bindung an Ribosomen ist an anderer Stelle beschrieben [29]. 50 µl Reaktionsgemisch enthielten 0.4 nmol des angegebenen Oligonucleotids sowie 7.0 pmol ¹⁴C-Lys-tRNA (0.150 E²⁶⁰-Einheiten), 1.1 E²⁶⁰-Einheiten Ribosomen aus *E. coli*, 0.05 mol/l Tris-acetat, pH = 7.2, 0.03 mol/l Magnesiumacetat, 0.05 mol/l Kaliumchlorid. In Abwesenheit von Oligo-A werden 0.49 pmol ¹⁴C-Lys-tRNA an Ribosomen gebunden; dieser Betrag ist von den angegebenen Werten abgezogen.

Zusatz	Bindung von ¹⁴ C-Lys-tRNA (pmol nach Oligo-A-Zugabe)
ApA	0.01
ApApA	1.92
ApApApA	1.92
ApApApApA	1.92
ApApApApApA	2.71

Tetra- oder Pentanucleotid. Das Dublett AA hat keinen Einfluß auf die Lys-tRNA-Bindung. Demnach entsprechen drei aufeinanderfolgende Basen in der mRNA einer Aminosäure im Protein.

Diese Experimente ermöglichten relativ einfach die Bestimmung der Basensequenz der Codons. Die Fraktionierung der Abbauprodukte von Poly-(U, G) lieferte die drei Trinucleotide GUU, UGU und UUG, die als Codons für Valin, Cystein bzw. Leucin erkannt wurden [75].

Die Trinucleotid-Synthese erwies sich als unser experimentelles Hauptproblem. Damals waren erst 20–30 der 64 Trinucleotide in der Literatur beschrieben. Philip Leder begann mögliche enzymatische Methoden für die Synthese von Trinucleotiden auszuarbeiten, und wir baten Leon Heppel und Maxine Singer um Rat. Während unserer gesamten Arbeiten am Code halfen uns Heppel und Singer bei Nucleinsäure-Problemen. Jeder Besuch in ihrem Laboratorium wurde für mich so etwas wie eine Wallfahrt nach Delphi. Der Hauptunterschied war, daß die Hinweise von einem der beiden Orakel stets klar und genau waren.

Marianne Grunberg-Manago besuchte damals die National Institutes of Health für einige Tage, und sowohl sie als auch Maxine Singer beteiligten sich an Philip Leders Studien der Oligonucleotid-Synthese, die durch die primer-abhängige Polynucleotid-Phosphorylase katalysiert wurde (Abb. 2). Schließlich fanden

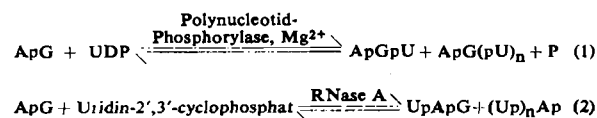


Abb. 2. Trinucleotid-Synthese, katalysiert durch Polynucleotid-Phosphorylase [Gl. (1)] und Pancreas-RNase A [Gl. (2)].

Leder, Singer und Brimacombe [30] sowie Thach und Doty [31] Bedingungen für die Oligonucleotid-Synthese.

Heppel schlug eine andere von ihm, Whitfield und Markham [32] gefundene Synthese vor, die auf der

[30] P. Leder, M. F. Singer u. R. Brimacombe, *Biochemistry* 4, 1561 (1965).

[31] R. E. Thach u. P. Doty, *Science* 147, 1310 (1965).

[32] L. A. Heppel, P. R. Whitfield u. R. Markham, *Biochem. J.* 60, 8 (1955).

Fähigkeit der Pancreas-Ribonuclease A zur Synthese von Oligonucleotiden aus Pyrimidin-2',3'-cyclophosphaten und Mono- oder Oligonucleotiden als Acceptoren beruhte. Merton Bernfield untersuchte verschiedene Aspekte dieser Reaktion und synthetisierte viele Trinucleotide mit diesem Enzym [33–35].

In einer bemerkenswerten, sich über viele Jahre erstreckenden Serie von Experimenten arbeiteten Khorana und Mitarbeiter chemische Methoden für die Oligo- und Polynucleotid-Synthese aus [38]. Sie waren in der Lage, die 64 Trinucleotide auf chemischem Weg zu synthetisieren, während in unserem Laboratorium enzymatische Methoden benutzt wurden.

Die Basensequenz der Codons wurde sowohl durch Stimulation der Bindung von AA-tRNA an Ribosomen durch Trinucleotide bekannter Sequenz [36, 37] als auch durch Stimulation der in-vitro-Protein-Synthese durch Polyribonucleotide mit sich wiederholenden Di-, Tri- oder Tetrameren bekannter Sequenz bewiesen, wie sie Khorana im folgenden Aufsatz beschreibt.

Der genetische Code ist in Abbildung 3 gezeigt. Die meisten Tripletts entsprechen Aminosäuren. Wenn mehrere Codons für die gleiche Aminosäure vorliegen, unterscheiden sie sich in der Regel nur durch die Base in der dritten Position des Tripletts. Synonyme Codons stehen also in systematischer Beziehung zuein-

UUU Δ ○ Phe	UCU Δ ○	UAU Δ ○ Tyr	UGU Δ ○ Cys
UUC Δ ○	UCC Δ ○ Ser	UAC Δ ○	UGC Δ
UUA ○ Leu	UCA Δ	UAA Δ Term	UGA Δ Term
UUG Δ ○	UCG Δ ○	UAG Δ	UGG Δ ○ Trp
CUU Δ ○	CCU Δ ○	CAU Δ ○ His	CGU Δ ○
CUC Δ ○	CCC Δ ○	CAC Δ ○	CGC Δ ○
CUA Δ Leu	CCA Δ ○ Pro	CAA Δ ○ Gln	CGA Δ ○ Arg
CUG Δ	CCG Δ ○	CAG Δ	CGG Δ
AUU Δ ○	ACU Δ ○	AAU Δ ○ Asn	AGU Δ
AUC Δ ○ Ile	ACC Δ ○	AAC Δ ○	AGC Δ ○ Ser
AUA Δ ○	ACA Δ ○ Thr	AAA Δ ○ Lys	AGA Δ ○ Arg
AUG Δ ○ Met	ACG Δ	AAG Δ	AGG Δ
GUU Δ ○	GCU Δ ○	GAU Δ ○ Asp	GGU Δ ○
GUC Δ	CCC Δ ○ Ala	GAC Δ ○	GGC Δ ○ Gly
GUA Δ ○ Val	GCA Δ ○	GAA Δ ○ Glu	GGA Δ ○
GUG Δ	GCG Δ	GAG Δ	GGG Δ

Abb. 3. Der genetische Code.

Δ Basensequenz von mRNA-Codons, ermittelt durch Stimulation der Bindung von *E. coli*-AA-tRNA an *E. coli*-Ribosomen mit Trinucleotid-Matrizen; ○ Basensequenz von mRNA-Codons, bestimmt durch Stimulation des Einbaus von Aminosäuren in Protein mit statistisch geordneten Polynucleotid-Matrizen in *E. coli*-Extrakten. Term: Kettenabbruch-Codons (Kettenabbruch- und Start-Codons sind in Tabelle 3 angegeben).

[33] M. Bernfield, *J. biol. Chemistry* 240, 4753 (1965).

[34] M. Bernfield, *J. biol. Chemistry* 241, 2014 (1966).

[35] M. Bernfield u. F. M. Rottman, *J. biol. Chemistry* 242, 4134 (1967).

[36] D. Söll, J. Cherayil, D. S. Jones, R. D. Faulkner, H. Hampel, R. M. Bock u. H. G. Khorana, *Cold Spring Harbor Sympos. quantitat. Biol.* 31, 51 (1966).

[37] M. W. Nirenberg, T. Caskey, R. Marshall, R. Brimacombe, D. Kellogg, B. Doctor, D. Hatfield, J. Levin, F. Rottman, S. Pestka, M. Wilcox u. F. Anderson, *Cold Spring Harbor Sympos. quantitat. Biol.* 31, 11 (1966); B. P. Doctor, J. E. Loebel u. D. A. Kellogg, *ibid.* 31, 543 (1966); D. Hatfield, *ibid.* 31, 619 (1966); S. Pestka u. M. W. Nirenberg, *ibid.* 31, 641 (1966).

[38] H. G. Khorana, H. Buchi, H. Ghosh, N. Gupta, T. M. Jacob, H. Kössel, R. Morgan, S. A. Narang, E. Ohtsuka u. R. O. Wells, *Cold Spring Harbor Sympos. quantitat. Biol.* 31, 39 (1966).

ander. Man findet fünf Arten der Codon-Degeneration, wobei jede Art durch die Basen in der dritten Position der synonymen Triplets bestimmt ist. Die dritte Base jedes degenerierten Triplets ist in Schema 1 gezeigt; die Striche entsprechen der ersten und zweiten Base.

(1) --G	(4) --U
(2) --U	--C
--C	--A
(3) --A	(5) --U
--G	--C
	--A
	--G

Schema 1

Die letzte Art der Codon-Degeneration ist als Summe von zwei Arten aufzufassen (Diskussion s. u.).

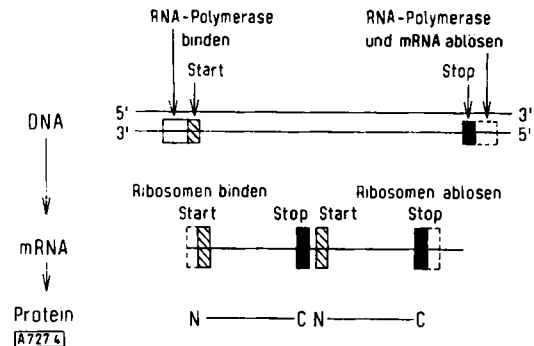


Abb. 4. Schematische Darstellung der Signalisierung von Transkription und Translation. Ribosomen-Untereinheiten werden an die mRNA nahe dem 5'-Ende gebunden und nahe dem 3'-Ende abgelöst. Spekulationen sind durch gestrichelte Linien angedeutet. N und C bedeuten die N- bzw. C-terminale Aminosäure des Proteins.

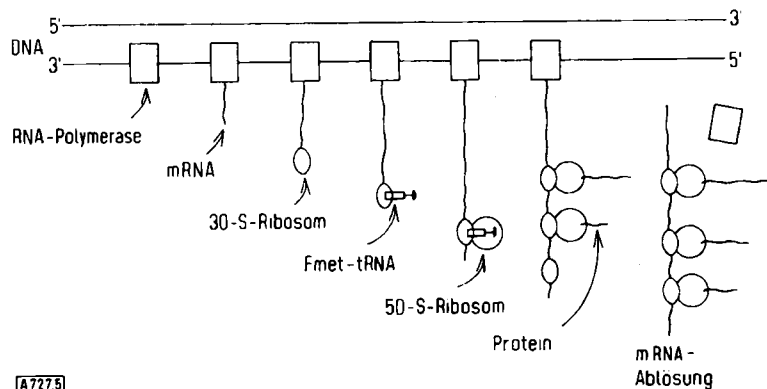


Abb. 5. Schematische Darstellung der ersten Schritte bei der Protein-Synthese.

Die Ergebnisse mit Trinucleotiden bestätigen 43 von den 50 Codon-Basenzusammensetzungen, die vorher mit den statistisch geordneten Polynucleotiden und dem zellfreien protein-synthetisierenden System gewonnen worden waren.

Ein bis sechs Codons können einer Aminosäure entsprechen, je nach dem, um welche Aminosäure es sich handelt. Eine Konsequenz der systematischen Degeneration ist die, daß der Austausch einer Base gegen eine andere in der DNA oft nicht zum Austausch einer Aminosäure gegen eine andere im Protein führt. Viele Mutationen bleiben deshalb „still“. Der Code scheint so angelegt zu sein, daß die Auswirkungen eines Basenaustauschs in der DNA oder einer irrtümlichen Übersetzung von mRNA-Basen möglichst vermindert werden. Der Austausch einer Aminosäure im Protein als Folge des Austauschs einer Base in der Nucleinsäure kann aus Abbildung 3 entnommen werden, wenn man von der fraglichen Aminosäure aus in waagerechter oder senkrechter (nicht aber diagonal) Richtung ausgeht.

Signale

Die Signalisierung der Transkription und Translation ist in den Abbildungen 4 und 5 schematisch wiedergegeben. Die RNA-Polymerase wird an eine oder mehrere spezifische Bindungsstellen der DNA gebunden und legt dabei den zu transkribierenden DNA-

Strang, die Richtung der Transkription und die erste zu transkribierende Base fest. Über den Start der RNA-Synthese bleiben noch viele Fragen zu beantworten.

Die Richtung der mRNA-Synthese ist der Ableserichtung des DNA-Stranges entgegengesetzt. Die erste in die wachsende mRNA eingebaute Base bildet das 5'-Ende, die zuletzt eingebaute das 3'-Ende. In ähnlicher Weise wird die RNA-Matrize während der Protein-Synthese abgelesen, wobei am oder nahe beim 5'-Ende der RNA begonnen und schrittweise um jeweils zugleich drei Basen in Richtung auf das 3'-Ende der RNA weitergegangen wird. Die mRNA wird demnach in der gleichen Richtung synthetisiert und abgelesen. Die erste Aminosäure entspricht dem N-Ende der Peptidkette; die C-terminale Aminosäure wird als letzte eingebaut.

Start

Die Protein-Synthese in *E. coli* wird durch eine besondere tRNA-Spezies, die von Marcker und Sanger^[39] entdeckte N-Formyl-tRNA_f, eingeleitet. Ein ribosomales 30-S-Partikel lagert sich an die wachsende mRNA-Kette nahe bei ihrem 5'-Ende an, noch bevor sich die mRNA von der DNA-Matrize ablöst. Mindestens drei nichtdialysierbare Faktoren und GTP sind für den Start der Protein-Synthese notwendig.

[39] K. Marcker u. F. Sanger, J. molecular Biol. 8, 835 (1964).

Die Reaktionen sind noch nicht vollständig geklärt; immerhin lassen die verfügbaren Ergebnisse vermuten, daß sich ein Faktor (F3) an der Bindung der ribosomalen 30-S-Untereinheit an die wachsende mRNA-Kette beteiligt und daß andere Faktoren (F1 und F2) und GTP für die Bindung der *N*-Formyl-met-tRNA an den 30-S-Ribosomen/mRNA-Komplex in Abhängigkeit von einem Start-Codon (AUG oder GUG^[62,63]) nötig sind. Die ribosomale 50-S-Untereinheit wird dann an den ribosomalen 30-S-Komplex gebunden, bevor das nächste Codon von AA-tRNA abgelesen wird. *N*-Formyl-met-tRNA wählt also das erste zu übersetzende Codon aus und bestimmt die Phase für die Übersetzung der folgenden Codons.

Eine andere tRNA-Spezies aus *E. coli*, Met-tRNA_m, bindet keine Formyl-Reste, spricht ausschließlich auf das Codon AUG an und entspricht der Aminosäure Methionin in inneren Positionen des Proteins.

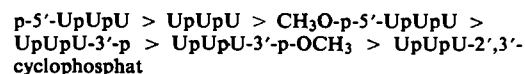
Der Typ der Degeneration, den man mit *N*-Formyl-met-tRNA beobachtet, weicht von der anderer AA-tRNA-Spezies ab, weil im Start-Codon die Basen an der ersten und nicht an der dritten Stelle variieren.

Jedes Triplet kann in drei strukturell verschiedenen Formen vorkommen: am 5'-Ende, am 3'-Ende oder in der Kette. Substituenten, die an die Hydroxygruppen der Ribose in den Codons gebunden sind, können die Matrizeneigenschaften der Codons grundlegend beeinflussen. Der Zusammenhang zwischen Codon-Struktur und Matrizenaktivität ist von meinem Mitarbeiter *Fritz Rottman*^[83] erforscht worden (Abb. 6).

Oligonucleotid	rel. Matrizenaktivität
p-5'-UpUpU	510
UpUpU	100
MeO-pUpUpU	74
UpUpU-3'-p	48
UpUpU-p-OMe	18
UpUpU-2',3'-cyclophosphat	17
(2'-5')-UpUpU	0
Oligodesoxy-T	0
Oligonucleotid	rel. Matrizenaktivität
p-5'-ApApA	181
ApApA	100
ApApA-3'-p	57
ApApA-2'-p	15
(2'-5')-ApApA	0
Oligodesoxy-A	0

Abb. 6. Relative Matrizenaktivität substituierter Oligonucleotide, näherungsweise erhalten durch Vergleich der in Gegenwart limitierender Oligonucleotidmengen an Ribosomen gebundenen AA-tRNA-Menge mit UpUpU für ¹⁴C-Phe-tRNA und ApApA für ¹⁴C-Lys-tRNA (jeweils gleich 100 % gesetzt); Daten von *Rottman* und *Nirenberg* [83].

Die Matrizenaktivität von Oligo-U-Präparaten bei limitierenden Oligonucleotid-Konzentrationen nimmt in der Reihe



ab. Trimere mit (2'-5')-Phosphodiesterbindungen, (2'-5')-UpUpU und (2'-5')-ApApA, eignen sich nicht als Matrizen für Phe- bzw. Lys-tRNA. Die rela-

tiven Matrizenaktivitäten von Oligo-A-Präparaten sind:



Ikehara und *Ohtsuka*^[70] zeigten, daß N⁶-DiMeApApA die Bindung von Lys-tRNA an Ribosomen nicht stimuliert, während das Tubercidin(7-Desazaadenosin)-analogon, TupApA, als Matrice für Lys-tRNA dienen kann.

Die RNA-Polymerase katalysiert die Synthese von mRNA mit 5'-terminalem Triphosphat. Es sind auch viele Enzyme beschrieben worden, die den Transfer von Molekülen zu oder von Hydroxygruppen von Nucleinsäuren katalysieren. Es ist deshalb möglich, daß bestimmte Modifikationen der Ribose- oder Desoxyribose-Hydroxygruppe von Nucleinsäuren eine Möglichkeit zur Regulation der Geschwindigkeit von Transkription oder Translation bieten.

Da die mRNA- und die Protein-Synthese vorübergehend über die Bildung eines DNA/mRNA-Ribosomen-Intermediates gekoppelt sind, ist es denkbar, daß die Synthese bestimmter mRNA-Spezies selektiv durch Vorgänge auf der Ebene des Ribosoms geregelt wird [76-79].

Kettenabbruch

Den ersten Beweis für „nonsense“-Codons erbrachten *Benzer* und *Champe* 1962^[40]. Sie erhielten eine Mutante des Bakteriophagen T4 mit einer Deletion, die sich über einen Teil des A-Gens und einen Teil des angrenzenden B-Gens der rII-Region erstreckte. Vermutlich sind die verbleibenden Segmente der Gene A und B miteinander verbunden und bilden auf diese Weise ein Gen. Und trotzdem wurde ein funktionelles Produkt des B-Gens gefunden. Allerdings hatte eine zweite im A-Gen lokalisierte Mutation den Verlust des funktionellen B-Gen-Produktes zur Folge. Diese Resultate lassen vermuten, daß ein „sense“-Codon durch eine Mutation in ein „nonsense“-Codon umgewandelt wurde, das nicht abgelesen werden kann. Also können auch auf das Gen folgende Regionen nicht abgelesen werden. *Sarabhai*, *Stretton*, *Brenner* und *Bolle*^[41] wiesen dann darauf hin, daß Mutationen an verschiedenen Stellen innerhalb des Gens für das Kopf-Protein des Bakteriophagen T4 einen Einfluß auf die Kettenlänge des entsprechenden Polypeptids haben. Diese entscheidenden Ergebnisse zeigten, daß „nonsense“-Codons dem Kettenabbruch bei der Protein-Synthese entsprechen. Zusätzliche Ergebnisse von *Brenner*^[42,43] sowie von *Garen*^[44,45] und ihren Mitarbeitern

[40] *S. Benzer* u. *S. P. Champe*, Proc. nat. Acad. Sci. USA 48, 1114 (1962).

[41] *A. S. Sarabhai*, *A. O. W. Stretton*, *S. Brenner* u. *A. Bolle*, Nature (London) 201, 13 (1964).

[42] *S. Brenner*, *A. O. W. Stretton* u. *S. Kaplan*, Nature (London) 206, 994 (1965).

[43] *S. Brenner*, *L. Barnett*, *E. R. Katz* u. *F. H. C. Crick*, Nature (London) 213, 449 (1967).

[44] *M. G. Weigert* u. *A. Garen*, Nature (London) 206, 992 (1965).

[45] *M. G. Weigert*, *E. Lanka* u. *A. Garen*, J. molecular Biol. 23, 391 (1967).

zeigten, daß drei Codons, UAA, UAG und UGA, die Kettenabbruch-Codons bei der Protein-Synthese sind (vgl. auch die neuere Übersicht von Garen^[46]).

Der Mechanismus des Kettenabbruchs wurde durch Stimulation der zellfreien Protein-Synthese mit statistisch geordneten Polynucleotiden^[47-49], Oligonucleotiden^[50] und Polynucleotiden bekannter Sequenz^[51, 52] sowie Viren-RNA^[53, 54] erforscht. *Capecchi* hat gezeigt, daß die Ablösung der Peptide von den Ribosomen sowohl von einem „release“-Faktor als auch von einem Kettenabbruch-Codon abhängt^[53]. Die Codons UAA, UAG und UGA stimulieren nicht die Bindung von AA-tRNA an Ribosomen (obwohl Bakterienmutanten mit AA-tRNA-Spezies gefunden worden sind, die Kettenabbruch-Codons entsprechen).

Kürzlich haben meine Mitarbeiter *Caskey*, *Tompkins* und *Scolnick*^[55, 56] gefunden, daß der Kettenabbruch mit Hilfe von Trinucleotiden untersucht werden kann. Die Inkubation eines Kettenabbruch-Trinucleotids und des „release“-Faktors mit dem Komplex *N*-Formyl-met-tRNA/AUG/Ribosom führt zur Ablösung von freiem *N*-Formyl-methionin aus dem ribosomalen Intermediat. Der „release“-Faktor aus *E. coli* wurde in zwei Komponenten zerlegt, die jeweils einer Gruppe von Codons entsprechen: R1 wirkt gemeinsam mit UAA oder UAG, R2 mit UAA oder UGA. Damit steht fest, daß Kettenabbruch-Codons von spezifischen Molekülen erkannt werden. Die einfachste Hypothese ist die, daß R1 und R2 mit den Kettenabbruch-Codons auf den Ribosomen in Wechselwirkung treten. Allerdings ist der Schritt der Codon-Erkennung und der Mechanismus des Kettenabbruchs bisher nicht völlig klar.

Tabelle 3. Kettenstart- und Kettenabbruch-Codons der Protein-Synthese in *E. coli*. Die „release“-Faktoren 1 und 2 sind für den Kettenabbruch mit den angegebenen Codons nötig, doch es ist nicht bekannt, ob sie direkt mit den Kettenabbruch-Codons reagieren.

Start (<i>N</i> -Formyl-met-tRNA)	AUG oder GUG
Kettenabbruch („release“-Faktor 1)	UAA oder UAG
Kettenabbruch („release“-Faktor 2)	UAA oder UGA

Wie Tabelle 3 zeigt, ähnelt die Art der mit R1 (UAA und UAG) gefundenen Codon-Degeneration der bei einigen Spezies von AA-tRNA; d.h. A und G in dritter Position des Codons sind äquivalent. Dagegen findet man mit R2 (UAA und UGA) eine andere Art

der Codon-Degeneration, weil A und G hier an zweiter und nicht an dritter Stelle der Triplets gegeneinander ausgetauscht sind.

Multiple Codon-Erkennung

1962 hatten Versuche mit statistisch geordneten RNA-Matrizen gezeigt, daß der Code weithin degeneriert ist und daß synonyme Codons sich oft nur um eine Base unterscheiden. Es wurde angenommen, daß die nicht ausgetauschten Basen innerhalb synonymen Triplets die gleiche Position annehmen. Eine systematische Art der Degeneration war anzunehmen, weil oft U mit C und A mit G äquivalent waren. Es wurde versucht, die Gesetze der Degeneration aus den Daten über die Basenzusammensetzung der Codons und den Austausch von Aminosäuren im Protein abzuleiten^[57, 58]. Es wurden zwei Spezies von Leu-tRNA gefunden, die verschiedenen mRNA-Codons entsprechen^[59]. Allerdings war weitere Arbeit nötig, um festzustellen, ob eine tRNA-Spezies nur auf ein Codon oder auf zwei oder mehr Codons anspricht.

Als die Reihenfolge der Basen innerhalb der Codons feststand, wurde es sehr deutlich, daß synonyme Codons in systematischer Beziehung zueinander stehen. Wie oben besprochen, unterscheiden sich synonyme Triplets nur in der Besetzung der dritten Position. Da nur wenige Arten der Degeneration bei den Codons für die 20 Aminosäuren gefunden worden sind, ist es wahrscheinlich, daß entsprechend wenige Arten der Codon-Erkennung beteiligt sind^[60].

Der Beweis, daß ein AA-tRNA-Molekül auf zwei Codons ansprechen kann, wurde dadurch geliefert, daß die meisten Phe-tRNA-Moleküle sowohl auf UUU als auch auf UUC^[61] ansprechen. Weitere Beweise erhielt man durch Bestimmung der Spezifität gereinigter tRNA-Fractionen für Trinucleotid-Codons. Die Resultate zeigen, daß eine reine tRNA-Spezies entweder ein, zwei oder drei Codons erkennt^[63-67, 37]. In Tabelle 4 sind unsere Arbeiten mit

[57] C. Woese, *Nature* (London) 194, 1114 (1962).

[58] R. Eck, *Science* 140, 477 (1963).

[59] B. Weisblum, S. Benzer u. R. W. Holley, *Proc. nat. Acad. Sci. USA* 48, 1449 (1962).

[60] M. W. Nirenberg, P. Leder, M. Bernfield, R. Brimacombe, J. Trupin, F. Rottman u. C. O'Neal, *Proc. nat. Acad. Sci. USA* 53, 1161 (1965); J. Trupin, F. Rottman, R. Brimacombe, P. Leder, M. Bernfield u. M. W. Nirenberg, *ibid.* 53, 807 (1965); R. Brimacombe, J. Trupin, M. W. Nirenberg, P. Leder, M. Bernfield u. T. Jaouni, *ibid.* 54, 954 (1965).

[61] M. R. Bernfield u. M. W. Nirenberg, *Science* 147, 479 (1965).

[62] B. F. C. Clark u. K. A. Marcker, *J. molecular Biol.* 17, 394 (1966).

[63] D. A. Kellogg, B. P. Doctor, J. E. Loebel u. M. W. Nirenberg, *Proc. nat. Acad. Sci. USA* 55, 912 (1966).

[64] D. Söll, D. S. Jones, E. Ohtsuka, R. D. Faulkner, R. Lohrmann, H. Hayatsu, H. G. Khorana, J. D. Cherayil, A. Hampel u. R. M. Bock, *J. molecular Biol.* 19, 556 (1966).

[65] D. Söll, J. D. Cherayil u. R. M. Bock, *J. molecular Biol.* 29, 97 (1967); D. Söll, E. Ohtsuka, D. S. Jones, R. Lohrmann, H. Hayatsu, S. Nishimura u. H. G. Khorana, *Proc. nat. Acad. Sci. USA* 54, 1378 (1965).

[66] D. Söll u. U. L. RajBhandary, *J. molecular Biol.* 29, 113 (1967).

[67] C. T. Caskey, A. Beaudet u. M. W. Nirenberg, *J. molecular Biol.* 37, 99 (1968).

[46] A. Garen, *Science* 160, 149 (1968).

[47] M. Takanami u. Y. Yan, *Proc. nat. Acad. Sci. USA* 54, 1450 (1965).

[48] M. S. Bretscher, H. M. Goodman, J. R. Menninger u. J. D. Smith, *J. molecular Biol.* 14, 634 (1963).

[49] M. C. Ganoza u. T. Nakamoto, *Proc. nat. Acad. Sci. USA* 55, 162 (1966).

[50] J. A. Last, W. M. Stanley jr., M. Salas, M. B. Hille, A. J. Wahba u. S. Ochoa, *Proc. nat. Acad. Sci. USA* 57, 1062 (1967).

[51] A. R. Morgan, R. D. Wells u. H. G. Khorana, *Proc. nat. Acad. Sci. USA* 56, 1899 (1966).

[52] H. Kössel, *Biochim. biophysica Acta* 157, 91 (1968).

[53] M. R. Capecchi, *Proc. nat. Acad. Sci. USA* 58, 1144 (1967).

[54] M. S. Bretscher, *J. molecular Biol.* 34, 131 (1968).

[55] C. T. Caskey, R. Tompkins, E. Scolnick, T. Caryk u. M. W. Nirenberg, *Science* 162, 135 (1968).

[56] E. Scolnick, R. Tompkins, T. Caskey u. M. W. Nirenberg, *Proc. nat. Acad. Sci. USA* 61, 768 (1968).

Tabelle 4. Codons, die von AA-tRNA-Spezies aus *E. coli* erkannt werden. Aminoacyl-tRNA-Präparationen aus *E. coli* wurden durch Säulenchromatographie mit umgekehrten Phasen fraktioniert und ihr Verhalten gegenüber Trinucleotid-Matrizen [67] bestimmt. Zusätzliche Ergebnisse erhielten Khorana und Mitarbeiter [64–66]. Ein Strich bedeutet Leu-tRNA-Fractionen, die auf kein Trinucleotid-Codon ansprechen. Die in Klammern stehende Zahl gibt die Zahl der gefundenen AA-tRNA-Maxima wieder.

Aminosäure	AA-tRNA-Spezies				
	1	2	3	4	5
Leu	CUU CUC	CUA CUG	mRNA-Codons CUG	UUG	—
Ser	UCU UCC	UCA UCG(2)	UCG	AGU AGC	
Arg	CGU CGC CGA	AGA AGG	CGG		
Ala	GCU GCC	GCA GCG			
Val	GUU GUC	GUA GUG			
Trp	UGG				
Met	AUG				
Ile	AUU AUC(2)				
Phe	UUU UUC(2)				
Tyr	UAU UAC(2)				
Cys	UGU UGC(3)				
His	CAU CAC				
Lys	AAA AAG(3)				
Glu	GAA GAG				

gereinigten tRNA-Fractionen aus *E. coli* zusammengefaßt. Es wurden vier, möglicherweise fünf Arten synonymen Codongruppen gefunden (s. Schema 2). Hier ist die dritte Base jedes synonymen Triplets genannt, die Striche stellen die ersten und zweiten Basen jeden Triplets dar.

- (1) —G (4) —U
 (2) —U —C
 —C —A
 (3) —A (5) —A
 —G —G
 —U

Schema 2

Die fünfte Degenerationsgruppe wurde mit Ser-tRNA aus *E. coli* (möglicherweise auch mit Val-tRNA) gefunden, bis jetzt aber noch nicht mit AA-tRNA aus anderen Organismen.

Die Zahl der Wörter oder Wortgruppen im Code entspricht eher der Zahl der tRNA-Anticodons als der Zahl der Aminosäuren. Da multiple tRNA-Spezies für die gleiche Aminosäure oft verschiedenen Gruppen von Codons entsprechen, besteht der tRNA-Code aus mehr Wortgruppen als der Aminosäure-Code.

Für die gleiche Aminosäure wurden mehrere AA-tRNA-Fractionen gefunden, die sich in den chromatographischen Eigenschaften unterschieden, aber ähnlich auf die Codons reagierten. Solche AA-tRNA-Fractionen mögen Produkte des gleichen Gens sein,

die auf verschiedene Weise durch Enzyme in vivo oder vielleicht auch in vitro während der Reinigungsprozedur verändert worden sind. Die Alternative wäre, daß diese AA-tRNA-Fractionen Produkte verschiedener Gene sind.

Crick hat vorgeschlagen, daß die Codon-Degeneration eine Folge schnell wechselnder Basenpaarungen zwischen einer Base in einem tRNA-Anticodon und Basen in der dritten Position synonymen mRNA-Anticodons ist [68]. Vermutlich bilden die erste und zweite Base der mRNA-Codons antiparallele Watson-Crick-Basenpaare mit den entsprechenden Basen im tRNA-Anticodon. „Veränderliche“ Basenpaare, wie sie Crick vorschlägt, sind in Tabelle 5 angegeben.

Tabelle 5. Schnell wechselnde Basenpaarung zwischen einer Base in einem tRNA-Anticodon (links) und der (den) Base(n) in der dritten Position synonymen mRNA-Codons. Die Paarung besteht nach Crick [68] in der Bildung antiparalleler, schnell wechselnder Wasserstoffbrücken.

tRNA-Anticodon (1. Base)	mRNA-Codons (3. Base)
U	A G
C	G
G	C U
I	U C A

U im tRNA-Anticodon paart entweder mit A oder G, das die dritte Position im synonymen mRNA-Codon einnimmt; C paart mit G, G paart mit C oder U, und I paart mit U, C oder A.

Die Aufklärung der Basensequenz der Hefe-Ala-tRNA durch Holley et al. [69] bot die Möglichkeit, eine Beziehung zwischen der Basensequenz des tRNA-Anticodons und dem mRNA-Codon zu finden. Holley überließ uns freundlicherweise ein tRNA^{Ala}-Präparat bekannter Sequenz und hoher Reinheit, und Philip Leder und ich, und zugleich Khorana und Söll et al. [65] konnten nachweisen, daß die Ala-tRNA auf die Codons GCU, GCC und GCA ansprach. Die Ergebnisse stützten Holleys Voraussage, daß die Sequenz IGC als tRNA^{Ala}-Anticodon dient. Inosin im Anticodon paart demnach abwechselnd mit U, C oder A in der dritten Position des mRNA-Codons. Es ist auch die Basensequenz anderer tRNA-Spezies aufgeklärt worden, und in jedem Fall ist die Codon-Anticodon-Beziehung in Einklang mit der schnell wechselnden Basenpaarung.

Universalität

Die Ergebnisse vieler Arbeiten lassen vermuten, daß verschiedene Lebensformen sich im wesentlichen der gleichen genetischen Sprache bedienen. Allerdings

[68] F. H. C. Crick, J. molecular Biol. 19, 548 (1966).

[69] R. W. Holley, J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee, S. H. Merrill, J. R. Penswick u. A. Zamir, Science 147, 1462 (1965).

[70] M. Ikehara u. E. Ohtsuka, Biochem. biophysic. Res. Commun. 21, 257 (1965).

kann die Zuverlässigkeit der Codon-Übersetzung sich ziemlich drastisch als Folge von Veränderungen an den für die Protein-Synthese nötigen Komponenten ändern. So zeigen Zellen manchmal eine unterschiedliche Spezifität der Codon-Übersetzung.

Richard Marshall, Thomas Caskey und ich untersuchten das Verhalten von AA-tRNA aus Bakterien, Amphibien und Säugetieren (*E. coli*, *Xenopus laevis* bzw. Meerschweinchenleber) gegenüber Trinucleotid-Codons. Fast identische Übersetzungen von Nucleotidsequenzen in Aminosäuresequenzen wurden mit AA-tRNA aus Bakterien, Amphibien und Säugetieren gefunden^[71]. Allerdings sprechen AA-tRNA-Präparationen aus *E. coli* auf manche Codons nicht merklich an, die wirksame Matrizen für AA-tRNA aus Metazoen sind.

Unser Interesse an der spezies-spezifischen Codon-Erkennung wurde durch die Möglichkeit angeregt, daß solche Phänomene als Regulatoren der Zelldifferenzierung dienen könnten. AA-tRNA-Präparate wurden deshalb säulenchromatographisch fraktioniert und das Verhalten der tRNA-Fractionen gegenüber Trinucleotid-Codons bestimmt^[67]. Eine Zusammenfassung unserer Ergebnisse gibt Abbildung 7. Zusätzliche Informationen kamen von Khorana und Söll et al.^[64-66]. Es sind viele „universelle“ AA-tRNA-Spezies gefunden worden. Allerdings wurden auch sieben tRNA-Spezies aus Säugetieren gefunden, die in *E. coli*-tRNA-Präparationen nicht vorkamen, und

umgekehrt fünf Spezies von *E. coli*-tRNA, die nicht in Präparationen von tRNA aus Säugetieren nachzuweisen waren.

Große Differenzen traten in der Konzentration der bestimmten Codons entsprechenden tRNA auf. Einige Organismen enthalten anscheinend keine AA-tRNA für bestimmte Codons. Zum Beispiel spricht Ile-tRNA aus Säugetieren gut auf AUU, AUC und AUA an, während Ile-tRNA aus *E. coli* nur auf AUU und AUC anspricht. Auch wurde eine Arg-tRNA-Spezies in Säugetieren, die auf AGG ansprach, gefunden, aber keine auf AGA ansprechende Arg-tRNA.

Obwohl tatsächlich einige Variationen bei der Codon-Übersetzung vorkommen, so führt doch die bemerkenswerte Ähnlichkeit der Basensequenzen in den Codons, die durch AA-tRNA aus Bakterien, Amphibien und Säugetieren erkannt werden, zur Vermutung, daß die meisten, vielleicht alle Lebensformen auf diesem Planeten im wesentlichen die gleiche Sprache benutzen, und daß diese Sprache nach universellen Gesetzen übersetzt wird.

Es ist über Funde von Mikroorganismen in Fossilien berichtet worden, die auf ein Alter von $3.1 \cdot 10^9$ Jahren geschätzt werden^[72]. Die ersten Vertebraten erschienen ungefähr vor $0.5 \cdot 10^9$ Jahren, die Amphibien und Säugetiere vor 350 bzw. 180 Millionen Jahren. So entstand der genetische Code möglicherweise vor mehr als $0.6 \cdot 10^9$ Jahren. Hinegardner und Engelberg^[73] sowie Sonneborn^[74] vermuten, daß der Code seit der

UUU ● ▲ 2 Phe	UCU ● ▲ Ser	UAU ● ▲ 2 Tyr	UGU ● 3 Cys
UUC ● ▲	UCC ● ▲	UAC ● ▲	UGC ● Cys
UUA — Leu	UCA 2 Ser	UAA Term	UGA Term
UUG — Leu	UCG ● ● ▲	UAG Term	UGG ● Trp
CUU ● Leu	CCU ● Pro	CAU ● His	CGU ● ▲ Arg
CUC ● Leu	CCC ● Pro	CAC ● His	CGC ● ▲ Arg
CUA ● Leu	CCA ● Pro	CAA Gln	CGA ● ▲ 2 Arg
CUG ● ● Leu	CCG ● Pro	CAG Gln	CGG ● 2 Arg
AUU ● 2 Ile	ACU ● Thr	AAU Asn	AGU ● ▲ Ser
AUC ● 2 Ile	ACC ● Thr	AAC Asn	AGC ● ▲ Ser
AUA — Ile	ACA Thr	AAA 3 Lys	AGA ● — Arg
AUG ● Met	ACG Thr	AAG ● 3 Lys	AGG ● ▲ Arg
GUU ● Val	GCU ● ▲ Ala	GAU Asp	GGU ● Gly
GUC ● Val	GCC ● ▲ Ala	GAC Asp	GGC ● Gly
GUA ● Val	GCA ● ▲ Ala	GAA Glu	GGA ● 2 Gly
GUG ● Val	GCG ● Ala	GAG Glu	GGG ● 2 Gly

Abb. 7. Zusammenfassung der in unserem Laboratorium mit gereinigten Aminoacyl-tRNA-Fractionen erhaltenen Ergebnisse. Weitere Ergebnisse wurden von Khorana und Söll et al. [64-66] veröffentlicht. Synonyme Codongruppen wurden durch Stimulation der Bindung gereinigter Fractionen von AA-tRNA aus *E. coli*, Hefe oder Meerschweinchenleber an *E. coli*-Ribosomen mit Trinucleotid-Codons bestimmt. Die miteinander verbundenen Symbole bedeuten synonyme Codons, die durch eine einzige gereinigte AA-tRNA-Fraktion aus ● *E. coli*, ▲ Hefe oder ■ Meerschweinchenleber erkannt worden sind. Die Zahlen zwischen den Symbolen geben die Zahl der AA-tRNA-Maxima an (AA-tRNA-Fractionen mit gleicher Spezifität für Codons).

[71] R. E. Marshall, C. T. Caskey u. M. W. Nirenberg, Science 155, 820 (1967).

Entwicklung derart komplexer Organismen wie Bakterien „eingefroren“ ist, weil größere Veränderungen des Codes die Aminosäuresequenz der meisten von der Zelle synthetisierten Proteine beeinflussen und somit wahrscheinlich letal sein würden.

[72] E. Barghoorn u. J. Schopf, Science 152, 758 (1966).

[73] R. Hinegardner u. J. Engelberg, Science 144, 1031 (1964).

[74] T. M. Sonneborn in V. Bryson u. H. J. Vogel: Evolving Genes and Proteins. Academic Press, New York 1965, S. 377.

Wenn man die Anzahl der für die Synthese eines einzigen Proteins nötigen Molekülspezies bedenkt, und wenn man berücksichtigt, daß die zelluläre Maschinerie für die Synthese komplex, heterogen und nicht zuverlässig ist, dann scheint die präzise Protein-Synthese ein außerordentlich schwieriges Problem zu sein. Um ein Molekül eines Proteins mit 400 Aminosäureresten zu synthetisieren, müssen 400 AA-tRNA Moleküle in der richtigen Reihenfolge ausgewählt werden. Für die Synthese des entsprechenden mRNA-Moleküls müssen mindestens 1206 Ribonucleosidtriphosphate nacheinander ausgesucht werden.

Man muß unterscheiden zwischen schrittweisen Operationen, d.h. aufeinanderfolgenden Schritten, und parallelen Operationen, d.h. simultanen Schritten. Gewöhnlich verschlechtert sich die Gesamtpräzision eines vielstufigen Prozesses mit steigender Anzahl aufeinanderfolgender Schritte sehr schnell. Zwei oder mehr aufeinanderfolgende Schritte sind für die Synthese jedes AA-tRNA-Moleküls nötig, weil eine AA-tRNA-Ligase zuerst die Synthese eines Aminoacyl-adenylats und dann den Transfer des Aminosäurerestes auf die richtige tRNA-Spezies unter Bildung von AA-tRNA katalysiert. Viele AA-tRNA-Moleküle können gleichzeitig synthetisiert werden. Obwohl hunderte von aufeinanderfolgenden Selektionsschritten für die Synthese eines Proteinmoleküls nötig sind, so ist die Protein-Synthese innerhalb der Zelle doch so organisiert, daß jede Aminosäure gewöhnlich *unabhängig* von anderen Aminosäuren ausgewählt wird. Auf diese Weise beeinflußt ein Translations-Irrtum gewöhnlich nicht die Genauigkeit der Translation anderer Codons, und die Fehler akkumulieren sich gewöhnlich nicht. Wenn allerdings ein Übersetzungsfehler die Phase der Ablesung ändert oder einen vorzeitigen Kettenabbruch bewirkt, werden die folgenden Schritte offenbar beeinflusst.

Baldwin und Berg^[80] haben gezeigt, daß die Ile-tRNA-Ligase aus *E. coli* die Synthese einer AA-tRNA nur dann katalysiert, wenn sowohl die Aminosäure als auch die tRNA richtig ausgewählt sind. Ist einmal ein falsches Aminoacyl-adenylat synthetisiert worden, so korrigiert das Enzym den Irrtum, indem es die Hydrolyse des Aminoacyl-adenylats katalysiert. 1960 vermuteten Yanofsky und St. Lawrence^[81], daß bestimmte Mutationen zur Produktion strukturell veränderter tRNA- oder Aminoacyl-tRNA-Synthetase mit veränderter Spezifität in bezug auf den Amino-

säure-Einbau in Protein führen können. Heute sind wir gut über Suppressor-Mutationen informiert, die die Komponenten der Protein-Synthese betreffen^[82]. Außerdem sind Faktoren, die die Präzision der Protein-Synthese beeinflussen, ausgiebig mit synthetischen Polynucleotid-Matrizen und protein-synthetisierenden in-vitro-Systemen untersucht worden, ferner durch Bestimmung der Bindung von AA-tRNA an Ribosomen in Abhängigkeit von Tri- oder Polynucleotid-Matrizen. Die Ergebnisse zeigen, daß die Genauigkeit der Codon-Ablesung durch die Inkubationstemperatur, den pH-Wert, die Konzentration der tRNA-Spezies, die Magnesiumionen-Konzentration, durch aliphatische Amine wie Putrescin, Spermidin, Spermin sowie durch Streptomycin, verwandte Antibiotica und durch andere Verbindungen beeinträchtigt wird.

Die meisten Codons werden vermutlich relativ fehlerfrei (0.1 bis 0.01 % Fehler oder weniger) übersetzt; die Fehler können allerdings bei bestimmten Codons bis 50 % betragen. Das heißt, die Genauigkeit der Translation kann von einem Codon zum anderen um mindestens den Faktor 5000 schwanken. Die meisten Irrtümer bei der Codon-Translation haben keinen beliebigen Aminosäure-Ersatz im Protein zur Folge, weil zwei von drei Basen im Codon gewöhnlich richtig abgelesen werden (dies bedeutet: Wenn die Genauigkeit der Ablesung abnimmt, dann mag ein Codon wie UUU zu 80 % als Phenylalanin, zu 15 % als Isoleucin und zu 5 % als Leucin abgelesen werden). Ein Codon wird dann durch relativ wenige AA-tRNA-Spezies übersetzt.

Über die biologische Bedeutung eines flexiblen, leicht modifizierbaren Apparates zur Übersetzung der Codons kann man nur spekulieren. Eine außerordentlich interessante Möglichkeit ist die, daß der Apparat der Codon-Erkennung sich in einer gesetzmäßigen, voraussagbaren Weise zu bestimmten Zeiten während des Zellwachstums und der Differenzierung ändert, und daß solche Veränderungen selektiv die Geschwindigkeit der Synthese bestimmter Protein-Spezies regulieren.

Geschwindigkeit der Translation

Ein *E. coli*-Chromosom ist aus $3 \cdot 10^6$ Basenpaaren zusammengesetzt; sie enthalten genügend Information, um die Sequenz von $1 \cdot 10^6$ Aminosäuren im Protein zu determinieren (das entspricht annähernd 2500 bis 3000 Protein-Spezies oder weniger, da manche Gene doppelt anwesend sein können).

Annähernd 20 bis 80 mRNA-Triplets werden pro Ribosom und pro Sekunde bei 37 °C übersetzt. Eine Zelle mag 1000 bis 15000 Ribosomen pro Chromosom enthalten, je nach ihrer Wachstumsgeschwindigkeit; Proteine werden demnach an vielen Stelle gleichzeitig synthetisiert. Parallele Vorgänge erhöhen die Leistungsfähigkeit der Zelle bei der Protein-Synthese stark.

[75] P. Leder u. M. W. Nirenberg, Proc. nat. Acad. Sci. USA 52, 420 (1964); 52, 1521 (1964).

[76] H. Bremer u. M. W. Konrad, Proc. nat. Acad. Sci. USA 51, 801 (1964).

[77] G. S. Stent, Science 144, 816 (1964).

[78] R. Byrne, J. G. Levin, H. A. Bladen u. M. W. Nirenberg, Proc. nat. Acad. Sci. USA 52, 140 (1964).

[79] H. A. Bladen, R. Byrne, J. G. Levin u. M. W. Nirenberg, J. molecular Biol. 11, 78 (1965).

[80] A. N. Baldwin u. P. Berg, J. biol. Chemistry 241, 839 (1966).

[81] C. Yanofsky u. P. St. Lawrence, Annu. Rev. Microbiology 14, 311 (1960).

[82] L. Gorini u. J. R. Beckwith, Annu. Rev. Microbiology 20, 401 (1966).

[83] F. Rottman u. M. W. Nirenberg, J. molecular Biol. 21, 555 (1966).

Abschließende Bemerkungen

Der genetische Code ist heute im wesentlichen entziffert. Ich hatte das Glück, im Laufe unserer Arbeiten viele begeisterte Mitarbeiter zu haben. Den Jahren der Mühe und den wichtigen Beiträgen von Mitarbeitern und zahllosen Kollegen überall auf der Welt Rechnung zu tragen, ist in der verfügbaren Zeit unmöglich. Man braucht nur auf die zusammenfassenden Berichte

auf den Cold Spring Harbor Symposien über quantitative Biologie 1963 und 1966 zu verweisen, um die Weite dieses Gebietes und das Ausmaß der heute verfügbaren Informationen zu sehen. Zusätzliche Angaben können den neu erschienenen Büchern von *Woese*^[9] und *Jukes*^[84] entnommen werden.

Eingegangen am 21. Juli 1969 [A 727]
Übersetzt von Dr. Th. Höpner, Heidelberg

[84] T. Jukes: *Molecules and Evolution*. Columbia University Press, New York 1967.

Nucleinsäure-Synthese als Werkzeug für das Studium des genetischen Codes (Nobel-Vortrag)^[**]

Von H. G. Khorana^[*]

I. Einführung

Der neuerliche Fortschritt im Verständnis des genetischen Codes ist der Erfolg der Anstrengungen zahlreicher Forscher, die eine Vielfalt wissenschaftlicher Disziplinen vertreten. Es scheint mir deswegen angebracht, einen kurzen Überblick über die wichtigsten Schritte bei der Entwicklung des genetischen Codes zu geben, bevor ich meinen Beitrag diskutiere, der stets zum großen Teil auf den Anstrengungen einer ganzen Gruppe basierte. Ich möchte auch in Ihr Gedächtnis zurückrufen, daß *Crick* einen Überblick über das Problem des genetischen Codes bis 1962 in seinem Nobelvortrag^[1] gegeben hat.

Es ist immer schwierig, vielleicht unmöglich, den Ausgangspunkt eines wissenschaftlichen Gebietes zu determinieren oder klar zu definieren; die Idee, daß Gene Proteine herstellen, war aber auf jeden Fall ein wichtiger Schritt, und dieses Konzept wurde besonders durch die Ein-Gen-ein-Protein-Hypothese von *Beadle* und *Tatum*^[2] in den Brennpunkt des Interesses gerückt. Damit war die biochemische Genetik geboren. Der nächste Schritt war getan, als feststand, daß Gene Nucleinsäuren sind. Die Transformationsversuche von *Avery* und Mitarbeitern^[3] und danach die Bakteriophagen-Experimente von *Hershey* und *Chase*^[4] sicherten dies für die DNA, die Arbeit mit der Tabakmosaik virus-RNA einige Jahre später auch für die RNA^[5,6].

[*] Prof. Dr. H. G. Khorana

The University of Wisconsin, Madison, Wisc. 53706 (USA)

[**] Copyright © The Nobel Foundation 1969. — Wir danken der Nobel-Stiftung, Stockholm, für die Genehmigung zum Druck dieser Übersetzung.

[1] F. H. C. Crick in: *Les Prix Nobel En 1962* (Norstedt, Stockholm 1963), S. 179; *Angew. Chem.* 75, 425 (1963).

[2] G. W. Beadle u. E. L. Tatum, *Proc. nat. Acad. Sci. USA* 27, 499 (1941).

[3] O. T. Avery, C. M. Macleod u. M. McCarty, *J. exp. Medicine* 79, 137 (1944).

[4] A. D. Hershey u. M. Chase, *J. gen. Physiol.* 36, 39 (1952).

[5] A. Gierer u. G. Schramm, *Nature (London)* 177, 702 (1956).

[6] H. Fraenkel-Conrat, B. Singer u. R. C. Williams, *Biochim. biophysica Acta* 25, 87 (1957).

Anfang der fünfziger Jahre stand damit fest, daß Gene Nucleinsäuren sind, und daß Nucleinsäuren die Protein-Synthese steuern, wobei die direkte Beteiligung der RNA an diesem Prozeß in den frühen Arbeiten von *Caspersson*^[7] sowie *Brachet*^[8] vermutet wurde. Auf dieser Stufe war es wichtig, mehr über die Chemie der Nucleinsäuren zu wissen, und tatsächlich waren die nun immer schneller aufeinanderfolgenden Entdeckungen weitgehend eine Folge der Arbeit an den Nucleinsäuren auf biochemischer und chemischer Ebene.

Die Strukturchemie der Nucleinsäuren, die sich während etwa 70 Jahren in vielen Ländern entwickelte, ging Schritt für Schritt von der Chemie der Bestandteile aus: von Purinen, Pyrimidinen und Zuckern zu den Nucleosiden und schließlich den Nucleotiden. Ein Höhepunkt war 1952 mit der Aufklärung der Internucleotidbindung in Nucleinsäuren durch *Brown* und *Todd* und ihre Mitarbeiter^[9] erreicht. (Es war ein großes Glück für mich, Mitarbeiter von Professor, jetzt Lord *Todd* vor dem Beginn meiner Arbeit auf dem Nucleotidgebiet gewesen zu sein.) Kurz danach wurde die Watson-Crick-Struktur^[10] für DNA vorgeschlagen, die die Aufmerksamkeit besonders auf die biologische Bedeutung ihrer physikalischen Struktur lenkte.

Etwa um diese Zeit entstand die Hypothese, daß eine lineare Sequenz von Nucleotiden in der DNA die lineare Sequenz von Aminosäuren und Proteinen bestimmt. Einige Jahre später kam die Enzymologie der DNA durch die Arbeiten von *Kornberg* und Mitarbeitern^[11] in Gang: Ihre Entdeckung und Charakterisierung des Enzyms DNA-Polymerase war ein Triumph der modernen Enzymologie, und die Methoden, die

[7] T. Caspersson, *Naturwissenschaften* 28, 33 (1941).

[8] J. Brachet, *Arch. Biol. (Liège)* 53, 207 (1942).

[9] D. M. Brown u. A. R. Todd in E. Chargaff u. J. N. Davidson: *Nucleic Acids*. Academic Press, New York 1955, Bd. 1, S. 409.

[10] J. D. Watson u. F. H. C. Crick, *Nature (London)* 171, 737 (1953).

[11] I. R. Lehman, M. J. Bessman, E. S. Simms u. A. Kornberg, *J. biol. Chemistry* 233, 163 (1958).