

## Veränderung des genetischen Codes

Lars Merkel und Nediljko Budisa

Max-Planck-Institut für Biochemie, BioFuture Forschungsgruppe „Molekulare Biotechnologie“, Martinsried

### Der genetische Informationsfluss und der genetische Code

► „Leben“ ist ein Prozess, der durch zwei Gesetzmäßigkeiten beschrieben wird: (i) die chemischen und physikalischen Naturgesetze und (ii) das genetische Programm, das alle biologischen Aktivitäten und Phänomene bestimmt. Nucleinsäuren kodieren Informationen, die in Form von Proteinen Gestalt annehmen. Dieser einseitige Informationsfluss zwischen Nucleinsäuren und Proteinen wird als das „zentrale Dogma der molekularen Biologie“ bezeichnet: „Der Austausch von Informationen ist sowohl von Nucleinsäure zu Nucleinsäure als auch von Nucleinsäure zu Protein möglich, doch von Protein zu Protein bzw. von Protein zu Nucleinsäure unmöglich.“

Die präzise Informationsweitergabe von Nucleinsäure zu Protein ist nur möglich, wenn die im genetischen Code festgelegte Korrelation einer bestimmten Nucleotidsequenz mit einer Aminosäuresequenz im Protein streng befolgt wird (Abb. 1). Die DNA dient als Template für die RNA, die sofort nach ihrer Transkription in verschiedene RNA-Formen (tRNA, rRNA oder mRNA) umgewandelt wird. Nach dem Transport der mRNA ins Zytoplasma wird sie durch Ribosomen in ein Protein translatiert. Zuletzt wird das Protein durch posttranslationale Modifikationen so verändert, dass es seine vollständige Funktionsfähigkeit erhält – z. B. als strukturelle Komponente der Zelle oder als effiziente biochemische Maschine (Abb. 2).

### Investition der Zelle in die ribosomale Proteinsynthese

Die Gene, die für Replikation, Transkription und Translation zuständig sind, sind nahezu in allen Lebewesen identisch. Alle Typen der RNA zusammen machen 21 % der gesamten Biomasse einer durchschnittlichen Zelle aus. Das ist nicht weiter erstaunlich, da seit langem bekannt ist, dass rund 35–45 % des Genoms für die Proteinsynthesemaschinerie zuständig sind. Beim Menschen

werden ca. 5 % der täglich aufgenommenen Energiemenge für die Translation (d. h. für die Proteinsynthese) benötigt. Bei den schnell wachsenden prokaryotischen Zellen, wie *E. coli*, sind es sogar 30–50 % der produzierten Energie.

### Aminoacyl-tRNA-Synthetasen (AARS) und tRNA-Interpretation des genetischen Codes

Die Genauigkeit der Proteintranslation hängt von zwei unterschiedlichen Erkennungsmechanismen auf molekularer Ebene ab: Enzyme (Aminoacyl-tRNA-Synthetasen, AARS) müssen die tRNA mit spezifischen Aminosäuren beladen (Interpretation) und die Anticodons der tRNAs müssen zu den Codons der mRNA passen. AARS sind kom-

plexe Enzyme, die ATP als Energiequelle sowie  $Mg^{2+}$ -Ionen als Cofaktoren nutzen.

Es gibt zwanzig AARS (je eine pro Aminosäure), die für eine genaue Übersetzung der Nucleinsäure in das Protein sorgen. Die AARS muss dabei nicht nur zwischen den einzelnen kanonischen Aminosäuren unterscheiden, sondern auch zwischen kanonischen und ebenfalls in der Zelle vorliegenden, natürlich vorkommenden nichtkanonischen Aminosäuren (wie Homocystein, Ornithin oder Citrulin). Die Präzision des gesamten Prozesses (die Fehlerrate liegt bei normalen *E. coli*-Zellen in der Größenordnung von  $2 \times 10^{-3}$  bis  $2 \times 10^{-4}$ ) wird auch durch mehrere Lesekorrekturenmechanismen gewährleistet. Nach der erfolgreichen Aminosäureerkennung und der Beladung der tRNA (Aminoacylierung) wird die Aminoacyl-tRNA zum Ribosom transportiert, wo die Translation stattfindet.

### Wie man den genetischen Code erweitern kann – Umprogrammierung der Proteintranslation

Bis vor ungefähr 15 Jahren war die Erweiterung der Anzahl der Aminosäuren, die zur Proteinbiosynthese verwendet werden können, bestenfalls eine Idee. Diese Aufgabe ist

keinesfalls trivial, da ein Weg gefunden werden muss, die Proteintranslationsmaschinerie derart umzuprogrammieren, dass die Kodierungsmöglichkeiten des natürlichen genetischen Codes verändert werden. Neue Techniken

in der Grundlagenforschung und biotechnologische Anwendungen, die eine größere Zahl von einbaufähigen Aminosäuren in Proteine verfügbar machten, waren ein starker Ansporn, die

Forschungen diverser Arbeitsgruppen auf dem Gebiet der genetisch kodierten Modifikation von Proteinen durch den Einbau von nichtkanonischen Aminosäuren *in vivo* und *in vitro* zu intensivieren. Und so erfährt in letzter Zeit die Veränderung des genetischen Codes als eigenständiger Forschungsbereich einen regen Interessenzulauf. Heutzutage sind die Methoden zur Umprogrammierung der Proteintranslationsmaschinerie nahezu in jedem biochemischen oder biophysikalischen Labor bekannt.

Um eine derartige Proteinsynthese in lebenden Zellen zu erhalten, müssen folgende Voraussetzungen erfüllt sein: (1) Die Aufnahme nichtkanonischer Aminosäuren, (2) ihre intrazelluläre Anreicherung auf ein Maß,

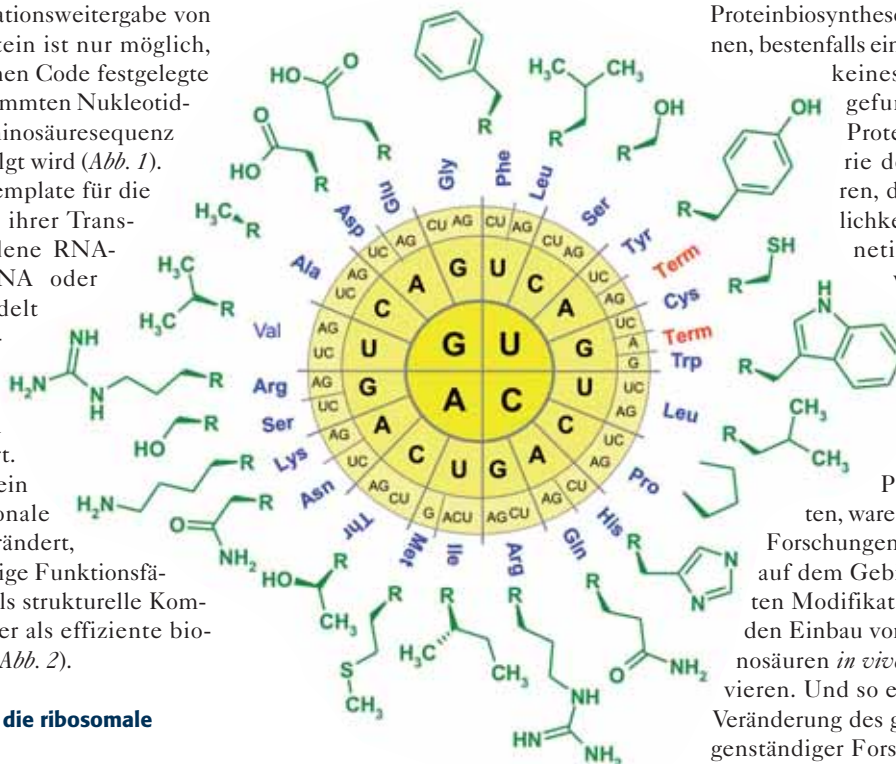


Abb. 1: Der genetische Code im mRNA-Format in einer kreisförmigen Darstellung. Die zwanzig kanonischen Aminosäuren, die hier im Dreibuchstabencode (blau) und in Form ihrer Seitenketten (grün) dargestellt sind, bilden die elementaren Bausteine für die Proteinsynthese an den Ribosomen. Es gibt 64 Dreierkombinationen (Codons) der vier Nucleotide (A, C, G, U). 61 Codons sind den 20 Aminosäuren zugeordnet, die drei verbleibenden (UAG, UAA, UGA) fungieren als Stopp-Codons.

das effizienten Umsatz durch die AARS (Aktivierung und tRNA-Acylierung) erlaubt, (3) metabolische und chemische Stabilität der aufgenommenen nichtkanonischen Aminosäure, (4) die tRNA-Beladung (Acylierung) in einer angemessenen Geschwindigkeit und (5) die Translation der nichtkanonischen Aminosäuren in Polypeptidketten durch (6) spezifische Codon-Neuzuordnung. Zellen, in denen diese Voraussetzungen erfüllt sind, bilden eine ideale Grundlage für die Variation der Aminosäureselektivität während der Proteinbiosynthese. Ihr Aminosäurerepertoire ist also erweiterbar (Abb. 2 und 3).

Es existieren zwei verschiedene Ansätze zur Veränderung des genetischen Codes. Die eine Möglichkeit, seit Mitte des letzten Jahrhunderts bekannt, basiert auf der Verwendung von auxotrophen Zellstämmen. Sie geht zurück auf die klassische Genetik mit Zellzucht, Konstruktion von Zellstämmen

und gezielter Selektion akkumulierter genetischer Veränderungen in lebensfähigen Wirtszellen, die durch die experimentellen Bedingungen gezwungen werden, Einbaufehler zu machen. Gebräuchliche Bezeichnungen für diesen experimentellen Ansatz sind Sense-Codon-Neuzuordnung oder statistischer Einbaumodus. Die aktivsten Gruppen auf diesem Gebiet sind die von D. A. Tirrell (USA), N. Budisa (D) und P. Marlier (F).

Die Umprogrammierung der genetischen Information kann durch eine neue Zuordnung von normalen kodierenden Nukleotidtripletts (Sense Codons) auf Grund von einigen „Schwachstellen“ („weak links“) im Bereich der Interpretation des genetischen Codes erzeugt werden. Das schwächste Glied in der Kette ist die Substratspezifität der AARS. Sie kann manipuliert werden, obwohl die AARS entscheidende Beiträge zur

Interpretation des genetischen Codes leisten. Außerdem sind sie im katalytischen Bereich flexibel. An ihren aktiven Zentren können weitere Reaktionen katalysiert werden, denn sie können nicht zwischen synthetischen und natürlichen Aminosäuren unterscheiden.

Zusätzlich kann die katalytische Effektivität der AARS und ihr Umsatz zugunsten von nichtkanonischen Aminosäuren verändert oder ihre Substratspezifität erweitert werden. Letztendlich ist das angestrebte Ziel der Erweiterung des Codes, „orthogonale“ AARS und tRNA herzustellen, die in der Lage sind, exklusiv eine 21. Aminosäure zu erkennen und diese in ein Polypeptid durch Benutzung eines entsprechenden kodierenden oder nicht-kodierenden Abschnitts (Triplet, Quadruplett und nicht-natürliche Basenpaare) einzubauen.

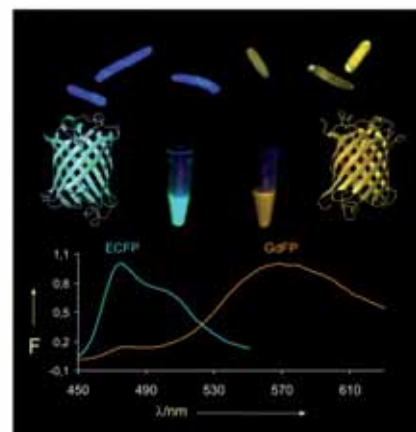
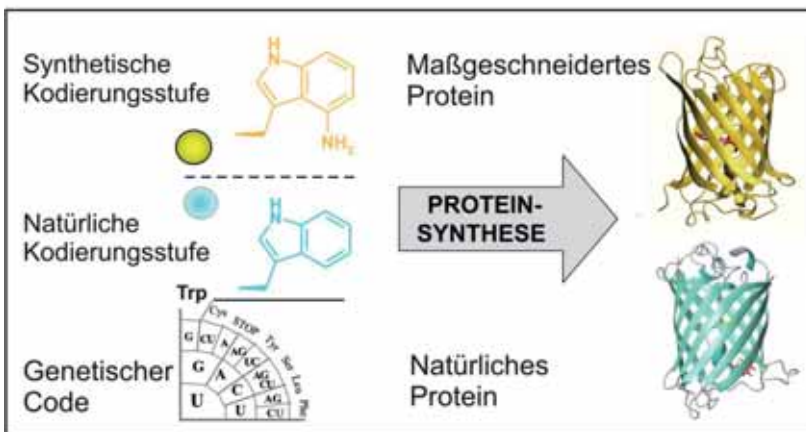
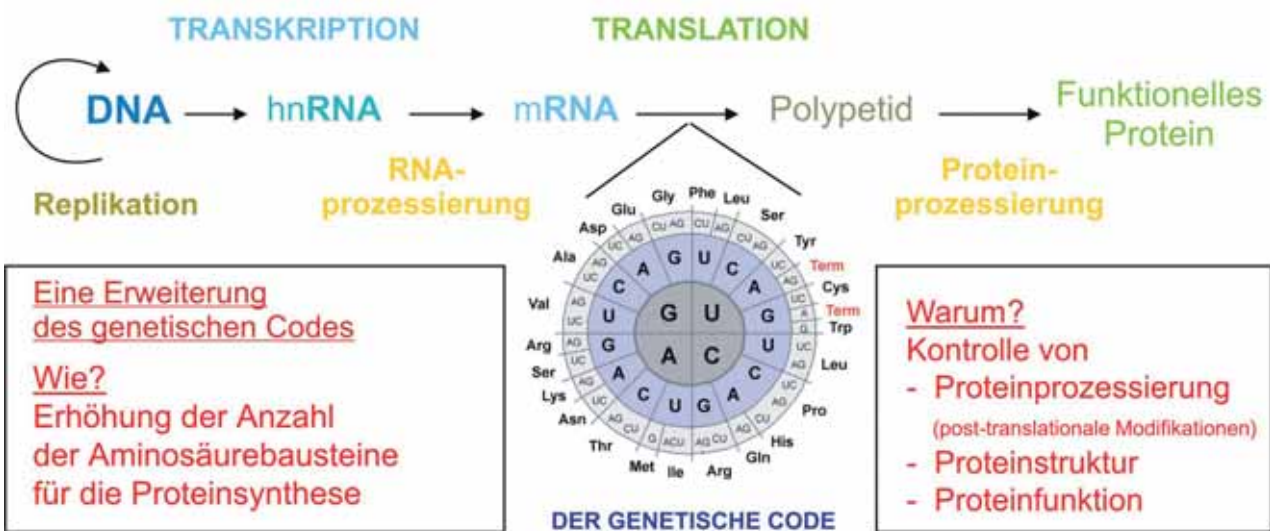
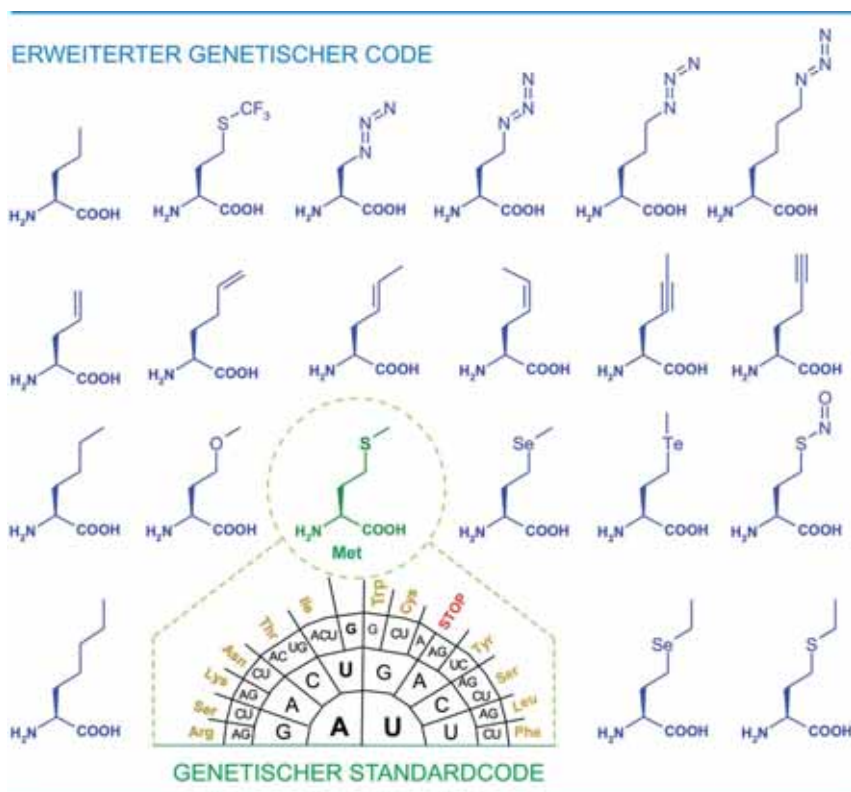


Abb. 2: Der genetische Code im Zusammenhang mit dem genetischen Informationsfluss (oberer Teil; hnRNA – engl.: homonuclear RNA), Ansätze für die Erweiterung des genetischen Codes (mittlerer Teil) und das Prinzip (unterer Teil, links) sowie ein Beispiel der Produktion maßgeschneiderter Proteine (unterer Teil rechts). Im Beispiel wird dem UGG-Codon, das als Signal für Tryptophan (Trp, (Trp, blaue Struktur)) dient, experimentell die nichtkanonische Aminosäure Aminotryptophan (gelbe Struktur) neu zugeordnet werden. So wird ein natürliches cyan-autofluoreszierendes Protein in ein synthetisches, gold-autofluoreszierendes Protein umgewandelt. Lebende Bakterienzellen, die das entsprechende Protein herstellen, sind leicht durch einfache Fluoreszenzmikroskopie zu unterscheiden. Auf diesem Wege stellt man maßgeschneiderte Proteine her, deren Eigenschaften nicht in der Natur vorkommen.



**Abb. 3:** Erweiterung der Kodierungsmöglichkeiten des AUG-Codons, das im genetischen Code der kanonischen Aminosäure Methionin (Met) zugeordnet ist. Durch das systematische Ersetzen von Methionin (grün) in der ribosomalen Proteinsynthese entstand eine Bibliothek aktiver Analoga und Surrogate von Met (blau), die alle in der Natur nicht vorkommen. Darunter sind jedoch viele interessante chemische Gruppen und Atome (z. B. Halogene, Chalcogene, Keto-, Cyano-, Azido-, Nitroso-, Silyl-Gruppen, Alkene oder Alkine). Die resultierenden Proteine sind um diese chemischen Funktionalitäten bereichert.

Der zweite methodische Ansatz ist auf die Entdeckungen der Gruppen von F. Lipmann, G. von Ehrenstein und S. Benzer aus dem Jahre 1962 zurückzuführen. Damals wurde gezeigt, dass falsch acylierte RNAs (tRNAs, die mit einer nicht ihr entsprechenden („cognate“) Aminosäure beladen wurden) dennoch an der normalen ribosomalen Peptidsynthese teilhaben können. Bei der daraus entwickelten „suppression-based“ Methode werden chemisch oder enzymatisch aminoacylierte tRNAs (meist suppressor tRNAs) verwendet, deren ursprüngliche Entsprechung der kodierenden Einheiten unterdrückt ist. Im Allgemeinen wird das durch einen positionsspezifischen Einbau erreicht. Dieser und andere Ansätze zur Produktion von unnatürlichen Proteinen und Peptiden (die z. B. auf Ribozymen, veränderten Ribosomen, neuen Basenpaaren, *in vitro* wiederhergestellter Proteinsynthese, metabolischen Veränderungen oder chemischer Manipulation der tRNAs beruhen) sind ausgiebig in der aktuellen Literatur beschrieben. Vor allem die Gruppen von P. G. Schultz, D. Dougherty, U. RajBhandary in den USA sowie S. Yokoyama und M. Sisido in Japan sind auf diesem Gebiet aktiv.

### Von neuen Technologien zu neuen Lebensformen

Die Veränderung des genetischen Codes, sei es durch Erweiterung des existierenden Aminosäurerepertoires oder durch die Einführung neuer Codierungseinheiten, ist eine von vielen Möglichkeiten des Designs/Redesigns biologischer Systeme im Grenzgebiet zwischen Physik, Chemie und Biologie. Das Engineering des genetischen Codes als neues Forschungsgebiet steht vor zwei großen Herausforderungen. Erstens der effizienten Generierung neuer, allgemeiner Klassen biologischer Monomere durch den Einsatz chemischer Synthesen, Biocomputing und zielgerichteter Veränderung des metabolischen Kreislaufs. Zweitens der Einführung dieser neuen Klassen biologischer Monomere in den genetischen Code mittels einer umprogrammierten Proteintranslation.

Die Entwicklung von Molekülen und Zellen mit neuen Eigenschaften, die verbraucherorientiert optimiert werden können, bieten enorme Möglichkeiten für die Biotechnologie des 21. Jahrhunderts. Offensichtlich ist die treibende Kraft für die Ausweitung dieses Forschungssektors die Nach-

frage nach maßgeschneiderten Biomaterialien (wie z. B. pharmakologisch wirksame Materialien) oder therapeutischen Proteinen, die *in vivo* generiert werden können. Für die akademische Grundlagenforschung sind solche maßgeschneiderten Moleküle bestens geeignet, biologische Systeme und Prozesse in ihrer ganzen Komplexität zu untersuchen, zu verstehen, und zuletzt sogar zu beeinflussen. Wir stehen am Anfang einer synthetischen Biologie, in der gerade Grundlagen für künstliche Zellen und sogar neue Lebensformen geschaffen werden.

### Danksagung

Wir danken Monika Franke für ihre Hilfe bei der Ausarbeitung dieses Manuskripts und BMBF (BioFuture, BFF132) für die finanzielle Unterstützung unserer Forschung.

### Literatur

In den letzten Jahren ist eine große Anzahl von Artikeln erschienen, die die verschiedenen Aspekte dieses neuen und dynamischen Forschungsgebietes beleuchten. Für detaillierte Informationen empfehlen wir die folgenden Übersichtsartikel und Bücher:

- [1] Anthony-Cahill, J. S. and Magliery, T. J. (2002): Expanding the Natural Repertoire of Protein Structure and Function. *Curr. Pharm. Res.* 3: 299–315.
- [2] Hoshaka, T. and Sisido, M. (2002): Incorporation of non-natural amino acids into proteins. *Curr. Opin. Chem. Biol.* 6: 809–815.
- [3] Budisa, N. (2004): Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire. *Angew. Chem. Int. Ed.* 43: 3387–3428.
- [4] Hendrickson, T. L., de Crecy-Lagard, V. and P. Schimmel, P. (2004): Incorporation of nonnatural amino acids into proteins. *Annu. Rev. Biochem.* 73: 147–176.
- [5] Wang, L. and Schultz, P.G. (2005): Expanding the genetic code. *Angew. Chem. Int. Ed.* 44: 34–66.
- [6] Link, A. J. and Tirrell, D. A. (2005): Reassignment of sense codons *in vivo*. *Methods* 36: 291–298.
- [7] Budisa, N. (2005): Engineering the Genetic Code - Expanding the Amino Acid Repertoire for the Design of Novel Proteins. Wiley-VCH, Weinheim.

### Korrespondenzadresse:

PD Dr. Nediljko Budisa  
Max-Planck-Institut für Biochemie  
BioFuture Research Group ‚Molecular  
Biotechnology‘  
Am Klopferspitz 18  
D-82152 Martinsried  
Tel.: 089-8578-2661/2344  
Fax: 089-8578-3557  
budisa@biochem.mpg.de  
www.biochem.mpg.de/budisa