

Alte Codons – neue Aminosäuren**

Ulrich Hahn, Gottfried J. Palm und Winfried Hinrichs*

Stichwörter:

Genexpression · Ligasen · mRNA · Proteinmodifikationen · tRNA

Professor Wolfram Saenger
zum 65. Geburtstag gewidmet

Einleitung

Heutzutage lassen sich mit gentechnischen Methoden nichtproteinogene Aminosäuren in Proteine einbauen. Dies ermöglicht, die strukturelle, chemische und funktionelle Vielfalt der Proteine zu erweitern. Die resultierenden Alloproteine^[1] können ungewöhnliche biophysikalische und biochemische Eigenschaften haben und zur Proteinanalyse in Biochemie und Zellbiologie genutzt werden. Weiterhin bestehen Möglichkeiten für weitreichende biotechnologische und medizinische Anwendungen.

In Proteinen sind zahlreiche Nichtstandard-Aminosäurereste bekannt,^[2] allerdings werden diese durch enzymatische Umwandlung nach der ribosomalen Polypeptidsynthese oder durch nichtribosomale Peptidsynthese eingebaut. Direkte chemische Modifikationen sind unselektiv und selten sequenzspezifisch, was oft zu Störungen in der dreidimensionalen Struktur bis zur Denaturierung des behandelten Proteins

führt. Festphasen-Peptidsynthesen oder erfolgreiche native chemische Umwandlungen in großen Mengen und mit zufriedenstellenden Ausbeuten sind sehr aufwändig.^[3]

Eine Strategie, den sequenzspezifischen Einbau einer gewünschten nichtnatürlichen Aminosäure in vivo zu erzwingen, erfordert zusätzliche Werkzeuge für die Proteinbiosynthese. Ein neues Paar aus tRNA und zugehöriger Aminoacyl-tRNA-Ligase (aaRL), früher Aminoacyl-tRNA-Synthetase genannt, müsste zu einer effizienten Translation führen. Die neue tRNA sollte ausschließlich mit der neuen Aminosäure verestert werden und müsste eindeutig ein Codon erkennen, das nicht von einer anderen Standard-Aminoacyl-tRNA genutzt wird. Bisher war dies immer ein Stoppcodon oder ein künstliches Vier-Basen-Codon der mRNA.

Was bisher möglich war

Der erste Fall bezieht sich auf die Ausnahme Selenocystein, das zusätzlich zu den 20 Standardamino-säuren (daher häufig als 21. Aminosäure bezeichnet) in die wachsende Polypeptidkette eingebaut wird, indem ein Stoppcodon uminterpretiert wird.^[4] Dieses Phänomen wurde in natürlich vorkommenden Enzymen entdeckt, die Selenocystein im aktiven Zentrum enthalten. Kürzlich wurde ein vergleichbarer Fall für Pyrrolysin als 22. natürlich vorkommende Aminosäure publiziert.^[5] Die undefinierten Stoppcodons sind im Leserahmen der anderen Codons der mRNA, jedoch nicht am Ende der codierenden Region und beenden die Verlängerung der Polypeptidkette nicht, wenn die zugehörige Aminoacyl-tRNA verfügbar ist.

Ein Ansatz zum Ausbau der zellulären Translationsmaschinerie beruht auf der Einführung eines Paares aus Stoppcodon und zusätzlicher tRNA.^[6] Eine spezifische tRNA reicht für den sequenzspezifischen Einbau einer synthetischen Aminosäure in die wachsende Polypeptidkette nicht aus. Zusätzlich muss eine Aminoacyl-tRNA-Ligase entwickelt werden, die ihre entsprechende tRNA mit der gewünschten Aminosäure belädt. Die neue Einheit aus tRNA und Aminoacyl-tRNA-Ligase muss orthogonal zu den 20 üblichen Aminosäuren arbeiten, d. h., nur die neue Aminosäure dient dem konstruierten Ligase/tRNA-Paar als Substrat, und die tRNA reagiert ausschließlich auf ein eindeutiges Codon.^[7] Auf dieser Grundlage konnten L-3-(2-Naphthyl)alanin sowie funktionalisierte Derivate von Phenylalanin und Tyrosin selektiv in Proteine eingebaut werden.^[6,8]

Diese Methode funktioniert ausgezeichnet beim Einbau nichtkanonischer Aminosäuren an nur einer Positionen der Sequenz, das Überlesen mehrerer Stoppcodons ist jedoch nicht effektiv. Unter Verwendung auxotropher Stämme von *Escherichia coli* mit reduzierter Spezifität der modifizierten Aminoacyl-tRNA-Ligasen wird diese Problematik umgangen. Diese künstlichen Ligasen aminoacylieren die entsprechende tRNA mit einer synthetischen Aminosäure, die ihr natürliches Pendant bei der ribosomalen Polypeptidverlängerung ersetzt. In diesem Fall nutzen die neue Aminosäure und die zu ersetzende natürliche Aminosäure dasselbe Codon.^[9]

[*] Dr. G. J. Palm, Prof. Dr. W. Hinrichs
Institut für Chemie und Biochemie
Mathematisch-Naturwissenschaftliche
Fakultät
Universität Greifswald
Soldmannstraße 16
17489 Greifswald (Deutschland)
Fax: (+49) 3834-86-4373
E-mail: hinrichs@uni-greifswald.de

Prof. Dr. U. Hahn
Abteilung Biochemie und
Molekularbiologie
Institut für Biochemie und
Lebensmittelchemie
Fachbereich Chemie
Universität Hamburg
Martin-Luther-King-Platz 6
20146 Hamburg (Deutschland)

[**] Wir danken dem Fonds der Chemischen
Industrie für die finanzielle Unterstützung.

Anwendungen

Hendrickson et al. bauten für die kristallographische Proteinstrukturanalyse das Selenhomologe von Methionin in Proteine ein.^[10] Dieser Ansatz ist heute eine Standardmethode zur direkten Bestimmung der dreidimensionalen Proteinstruktur mit MAD (multiwavelength anomalous diffraction). Selenomethionin kann leicht in beliebige Proteine eingebaut werden, indem das gewünschte Protein in einem Methioninauxotrophen, in selenomethioninhaltigem Medium wachsenden *E. coli*-Stamm überproduziert wird.

In weiteren Studien wurden Fluor-tryptophan in Annexin V, 3-Fluor- und 3-Chlortyrosin in Azurin und Annexin V sowie 4-Aminotryptophan in GFP (grünes fluoreszierendes Protein) eingebaut. Die Modifikationen erfolgten mit dem Ziel, Struktur, Stabilität, Aktivität oder spektroskopische Eigenschaften dieser Alloproteine zu studieren.^[11] Ein jüngeres Beispiel ist der Austausch aller 13 Methioninreste in der Hämdomäne von Cytochrom P450 BM-3 (463 Aminosäurereste) gegen das analoge Norleucin. Daraus resultierte eine verdoppelte Peroxygenaseaktivität und eine signifikant reduzierte Thermostabilität des Cytochroms.^[12]

Der genetische Code ist degeneriert

Eine weitere Möglichkeit, den genetischen Code zu bearbeiten, beruht auf dem Umstand, dass von 64 Nucleotidtripletts (Codons) nur drei Stoppcodons sind, während die verbleibenden 61 Kombinationen (Sense-Codons) nur für 20 proteinogene Aminosäuren codieren. Diese Tatsache wird gewöhnlich als Degeneration des genetischen Codes bezeichnet.^[13] Methionin und Tryptophan werden durch je ein Codon, die restlichen 18 Aminosäuren durch 2 bis 6 Tripletts definiert. Interessanterweise sind in den meisten Fällen die beiden ersten Nucleotide redundanter Codons identisch und definieren bereits eindeutig die einzubauende Aminosäure. Diese Situation könnte als ein (2+1)-Buchstaben-Code angesehen werden. Alternativ könnte einem ein Zwei-Buchstaben-Code als Vorläufer im Laufe der

Evolution in den Sinn kommen. Wie dem auch sei, die Degeneration des genetischen Codes ist die Basis dafür, bei der ribosomalen Verlängerung der Polypeptidkette nichtnatürliche Aminosäuren in definierte Positionen einzubauen. Im Folgenden wollen wir auf die kürzlich von Tirrell et al. beschriebene Möglichkeit aufmerksam machen, degenerierte Sense-Codons neu zuzuweisen.^[14]

Der neue Ansatz

In *E. coli* codieren die Tripletts UUC und UUU der mRNA für Phenylalanin (Phe). Zur Übertragung von Phe auf die wachsende Polypeptidkette hat dieser Organismus nur eine tRNA mit dem Anticodon GAA, das gemäß der Basenpaarung nach Watson und Crick sehr gut mit dem Codon UUC korrespondiert. Etwas schwächer verläuft die Paarung mit UUU, wobei das „Wobble“-Basenpaar G-U entsteht.^[13] Wie üblich werden die Polynucleotidsequenzen vom 5'- zum 3'-Terminus geschrieben, außerdem verlaufen Codon und Anticodon antiparallel. Die Wechselwirkung zwischen dem Codon UUC und dem Anticodon GAA ist etwa 10 kJ mol^{-1} stärker als die zwischen UUU und GAA.^[15] Trotzdem wird in *E. coli* in beiden Fällen Phe in die wachsende Polypeptidkette eingebaut.

Zur Unterscheidung der beiden Codons wurde ein exogenes Phe transferierendes, aber modifiziertes Paar tRNA/Aminoacyl-tRNA-Ligase aus Hefe in *E. coli* eingeführt. Diese tRNA enthält das Anticodon AAA (ytRNA_{AAA}^{Phe}), was in *E. coli* zur optimalen Erkennung des Zielcodons UUU über Watson-Crick-Basenpaarungen führt. Sie kann durch ihre zugehörige, aber gentechnisch modifizierte Aminoacyl-tRNA-Ligase mit einer Nichtstandard-Aminosäure beladen werden. In der hier diskutierten Arbeit wurde L-3-(2-Naphthyl)alanin (Nal) als Beispiel gewählt.^[14] Durch Selektion von Ligasen aus Hefe mit veränderten Aminosäureresten im aktiven Zentrum resultierte ein Enzym, das aufgrund geringerer Substratspezifität Nal aufnimmt.^[16] In vitro zeigt die mutierte Hefe-Ligase gegenüber der künstlichen Aminosäure eine achtfach höhere Affinität als ge-

genüber Phe.^[14] Hier muss darauf hingewiesen werden, dass der Aminoacylierungsgrad der Hefe-tRNA^{Phe} durch endogene *E. coli*-Enzyme bedeutungslos ist.^[17] Außerdem erkennt die Variante der Aminoacyl-tRNA-Ligase aus Hefe keine der tRNAs aus *E. coli* als Substrat. Diese Eigenschaften gewährleisten die Orthogonalität des hefe-stämmigen Paares tRNA^{Phe}/Phenylalanyl-tRNA-Ligase zum *E. coli*-System (Abbildung 1).

In einem weiterführenden Ansatz wurde ein Phe-auxotropher *E. coli*-Stamm etabliert, der die Gene für die exogene ytRNA_{AAA}^{Phe}, die mutierte Aminoacyl-tRNA-Ligase-Variante aus Hefe (mu-yPheRL) sowie die Dihydrofolat-Reduktase der Maus (mDHFR) trägt. Dieser Satz an Genen ermöglichte in Gegenwart von Nal im Kulturmedium die Expression des Testproteins mDHFR, das die neue Aminosäure in definierten Positionen eingebaut hatte.

Tryptische Abbauprodukte der mDHFR wurden durch MALDI-TOF-Massenspektrometrie und Flüssigchromatographie-Tandem-Massenspektrometrie (LC/MS/MS) analysiert, um den kompletten Austausch der fünf UUU-codierten Phe-Reste gegen Nal und den Einbau der vier UUC-codierten Phe-Reste zu überprüfen. Die Translation der beiden Codons verlief unterschiedlich, und Phe sowie Nal wurden spezifisch in das Polypeptid eingebaut. Folglich wird mit dem genetischen Code dieses *E. coli*-Konstruktes die ribosomale Polypeptidsynthese um eine nichtnatürliche, aber spezifische Aminosäure erweitert, die ihr eigenes Codon hat.

Orthogonalitätsbedingungen für tRNA/Aminoacyl-tRNA-Ligase-Paare

Die strukturelle Basis für die Spezifität orthogonaler Bedingungen korrespondierender tRNA/Aminoacyl-tRNA-Ligase-Paare wurde durch Kristallstrukturanalyse des Tyrosyl-tRNA/Tyrosyl-tRNA-Ligase-Komplexes des Archaeons *Methanococcus jannaschii* erhellt.^[18] Die Erkenntnisse erlauben verbesserte Anwendungen durch ortsgerichtete Mutagenese der Ligase und durch Modifikation der zugehörigen tRNA, die ein Stoppcodon unterdrückt.

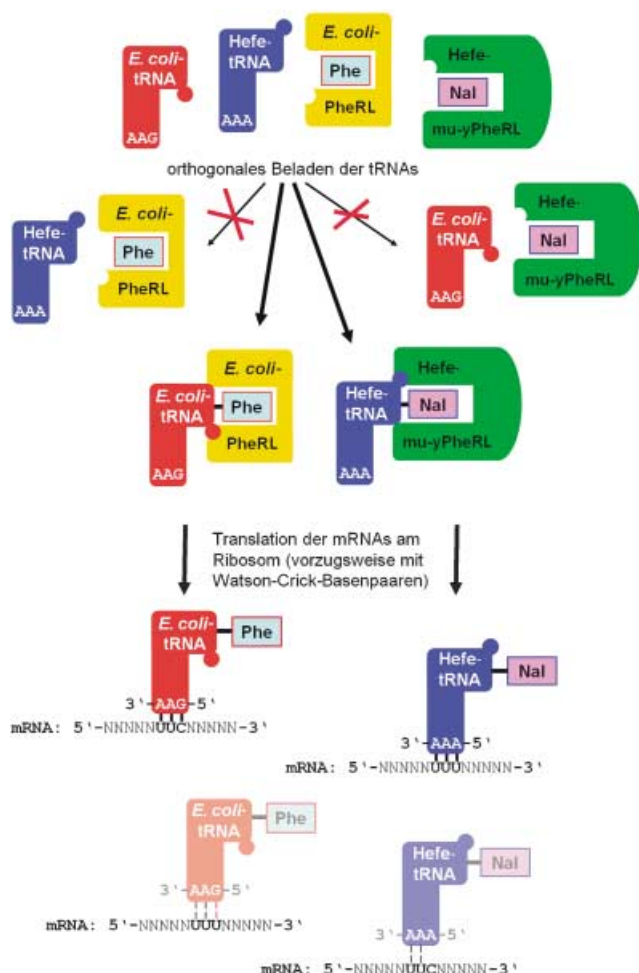


Abbildung 1. Nutzung orthogonaler tRNA/Aminoacyl-tRNA-Ligase-Paare zum Einbau von Phe und Nal bei der ribosomalen Proteinbiosynthese. Aufgrund der Orthogonalität der Paare *E. coli*-tRNA^{Phe}/PheRL und Hefe-tRNA^{Nal}/mutierte PheRL wird die tRNA^{Phe} nur mit Phe und die tRNA^{Nal} nur mit Nal aminoacyliert. Das Genom von *E. coli* umfasst 21 Aminoacyl-tRNA-Ligasen und 84 tRNAs. Es ist eine eindrucksvolle Fähigkeit, über 1600 falsche Aminoacylierungen zu vermeiden und zugleich genau die korrekten 84 zu ermöglichen. Die beladenen tRNAs werden vom Ribosom mit unterschiedlicher Präferenz für die kanonischen Phe-Codons UUC und UUU verwendet: tRNA^{Phe} interagiert besser mit UUC und tRNA^{Nal} nur mit UUU auf der mRNA. Das beruht auf den optimierten Erkennungsmustern der Wasserstoffbrücken (Watson-Crick-Paarung) zwischen Codon und Anticodon. Die unterdrückten Wechselwirkungen von tRNA und Codon sind am Ende der Abbildung mit blassen Farben markiert. Auf diese Weise wird 5'-NNN**UUCUUU**NNN-3' auf der mRNA spezifisch zu NH₂-Xaa**PheNal**Xaa-COOH translatiert.

Die Gruppen um Cusack und Yokoyama legten dar, dass Orthogonalität auf der unterschiedlichen Erkennung der tRNA-Struktur durch die zugehörige Tyrosyl-tRNA-Ligase bei bakteriellen und archaealen/eukaryotischen Systemen beruht. Sequenzen und Tertiärstrukturen bakterieller Tyrosyl-tRNA-Ligasen weisen hohe Homologien auf, unterscheiden sich jedoch von ihren funktionellen Verwandten in Archaea und Eukaryoten.

Zusammenfassung

Die neue Methode, bei der redundante Codons umdefiniert werden, um bei der Translation mit unterschiedlichen Spezifitäten erkannt zu werden, wurde unter Verwendung orthogonaler Paare aus tRNA und zugehöriger Aminoacyl-tRNA-Ligase möglich. Wie kürzlich gezeigt wurde, ist es nicht ausreichend, die genetische Information als eine Sequenzvariation der vier Nucleotidbasen anzusehen, vielmehr bestimmt auch die Biochemie des Lesesystems die biologische Funktion.^[19]

Leider ist es noch nicht möglich, die strukturellen und funktionellen Veränderungen der Alloproteine, wie Faltung, Stabilität und Liganden-Erkennung, vorauszusagen. Somit haben Strukturaufklärungen durch NMR-Spektroskopie und Röntgenstrukturanalyse sowie spektroskopische und computergestützte Moleküldynamikstudien künftig weitere Ziele.

Zusätzliche Aminosäuren können als Sonden bei diesen neuen Herausforderungen der strukturellen Biochemie, Biologie und Biotechnologie dienen. Dies war die Motivation für die Gruppe um Tirell, eine neue Methode dem Repertoire der Proteinchemie hinzuzufügen.

Jede neue Nichtstandard-Aminosäure erfordert noch immer die Selektion einer angepassten Aminoacyl-tRNA-Ligase. Hier gehen die Bemühungen der Biochemiker weiter: Es existieren Aminoacyl-tRNA-Ligase-Mutanten ohne Adenylierungsaktivität, die aber noch immer Aminosäure und ATP binden.^[20] Ebenso wurde die Änderung der Spezifität der Purinbindungstasche einer Phosphoribosyltransferase demonstriert.^[21] Können wir eine modifizierte Aminoacyl-tRNA-Ligase konstruieren, die ein aktiviertes Derivat einer beliebigen Aminosäure aufnimmt und damit eine tRNA belädt? Wir haben eine aufregende Proteinchemie zu erwarten.

- [1] H. Koide, S. Yokoyama, G. Kawai, J. M. Ha, T. Oka, S. Kawai, T. Miyake, T. Fuwa, T. Miyazawa, *Proc. Natl. Acad. Sci. USA* **1988**, 85, 6237–6241.
- [2] G. C. Barrett in *Amino Acid Derivatives: a Practical Approach* (Hrsg.: G. C. Barrett), Oxford University Press, New York, **1999**, S. 15–30.
- [3] P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, 69, 923–960.
- [4] a) F. Zinoni, J. Heider, A. Böck, *Proc. Natl. Acad. Sci. USA* **1990**, 87, 4660–4664; b) A. Böck, K. Forchhammer, J. Heider, W. Leinfelder, G. Sawers, B. Veprek, F. Zinoni, *Mol. Microbiol.* **1991**, 5, 515–520; c) M. Rother, R. Wilting, S. Commans, A. Böck, *J. Mol. Biol.* **2000**, 299, 351–358.
- [5] a) G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, 296, 1459–1462;

- b) B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, 296, 1462–1466.
- [6] a) L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, 292, 498–500; b) Z. Zhang, L. Wang, A. Brock, P. G. Schultz, *Angew. Chem.* **2002**, 114, 2964–2966; *Angew. Chem. Int. Ed.* **2002**, 41, 2840–2842; c) L. Wang, A. Brock, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, 124, 1836–1837; d) L. Wang, Z. Zhang, A. Brock, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2003**, 100, 56–61.
- [7] R. Furter, *Protein Sci.* **1998**, 7, 419–426.
- [8] S. Ohno, T. Yokogawa, I. Fujii, H. Asahara, H. Inokuchi, K. Nishikawa, *J. Biochem.* **1998**, 124, 1065–1068.
- [9] V. Döring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. de Crécy-Lagard, P. Schimmel, P. Marlière, *Science* **2001**, 292, 501–504.
- [10] W. A. Hendrickson, J. R. Horton, D. M. LeMaster, *EMBO J.* **1990**, 9, 1665–1672.
- [11] a) C. Minks, R. Huber, L. Moroder, N. Budisa, *Biochemistry* **1999**, 38, 10649–10659; b) C. Minks, S. Alefelder, L. Moroder, R. Huber, N. Budisa, *Tetrahedron* **2000**, 56, 9431–9442; c) J. H. Bae, M. Rubini, G. Jung, G. Wiegand, M. H. Seifert, M. K. Azim, J.-S. Kim, A. Zumbusch, T. A. Holak, L. Moroder, R. Huber, N. Budisa, *J. Mol. Biol.* **2003**, 328, 1071–1081.
- [12] P. C. Cirino, Y. Tang, K. Takahashi, D. A. Tirrell, F. H. Arnold, *Biotechnol. Bioeng.* **2003**, 83, 729–734.
- [13] F. H. C. Crick, *J. Mol. Biol.* **1966**, 19, 548–555.
- [14] I. Kwon, K. Kirshenbaum, D. A. Tirrell, *J. Am. Chem. Soc.* **2003**, 125, 7512–7513.
- [15] M. Meroueh, C. S. Chow, *Nucleic Acids Res.* **1999**, 27, 1118–1125.
- [16] P. Wang, N. Vaidehi, D. A. Tirrell, W. A. Goddard III, *J. Am. Chem. Soc.* **2002**, 124, 14442–14449.
- [17] E. C. T. Peterson, O. C. Uhlenbeck, *Biochemistry* **1992**, 31, 10380–10389.
- [18] T. Kobayashi, O. Nureki, R. Ishitani, A. Yaremchuk, M. Tukalo, S. Cusack, K. Sakamoto, S. Yokoyama, *Nat. Struct. Biol.* **2003**, 10, 425–432.
- [19] C. Fenske, G. Palm, W. Hinrichs, *Angew. Chem.* **2003**, 115, 626–630; *Angew. Chem. Int. Ed.* **2003**, 42, 606–610.
- [20] R. J. Leatherbarrow, A. R. Fersht, G. Winter, *Proc. Natl. Acad. Sci. USA* **1985**, 82, 7840–7844.
- [21] N. R. Munagala, C. C. Wang, *Biochemistry* **1998**, 37, 16612–16619.