

## Erweiterung des genetischen Codes

## Pyrrollysyl-tRNA-Synthetase: Methanogenese und Gencode-Erweiterung

NIKOLAJ GEORG KOCH<sup>1,2</sup>, NEDILJKO BUDISA<sup>1,3</sup><sup>1</sup> INSTITUT FÜR CHEMIE, TU BERLIN<sup>2</sup> INSTITUT FÜR BIOTECHNOLOGIE-BIOANALYTIK, TU BERLIN<sup>3</sup> DEPARTMENT OF CHEMISTRY, UNIVERSITY OF MANITOBA, WINNIPEG, KANADA

**Pyrrollysyl-tRNA synthetase (PylRS) is an enzyme of some methanogenic Archaea for the natural incorporation of pyrrolysine into proteins. The discovery of PylRS as a natural tool for genetic code expansion paved the way for site-specific incorporation of non-canonical amino acids (ncAAs) into proteins, with versatile side chains useful in biotechnology. Almost 20 years after the discovery, we describe the journey which led to three distinct classes of PylRSs with unique nAA recognitions.**

DOI: 10.1007/s12268-021-1653-x

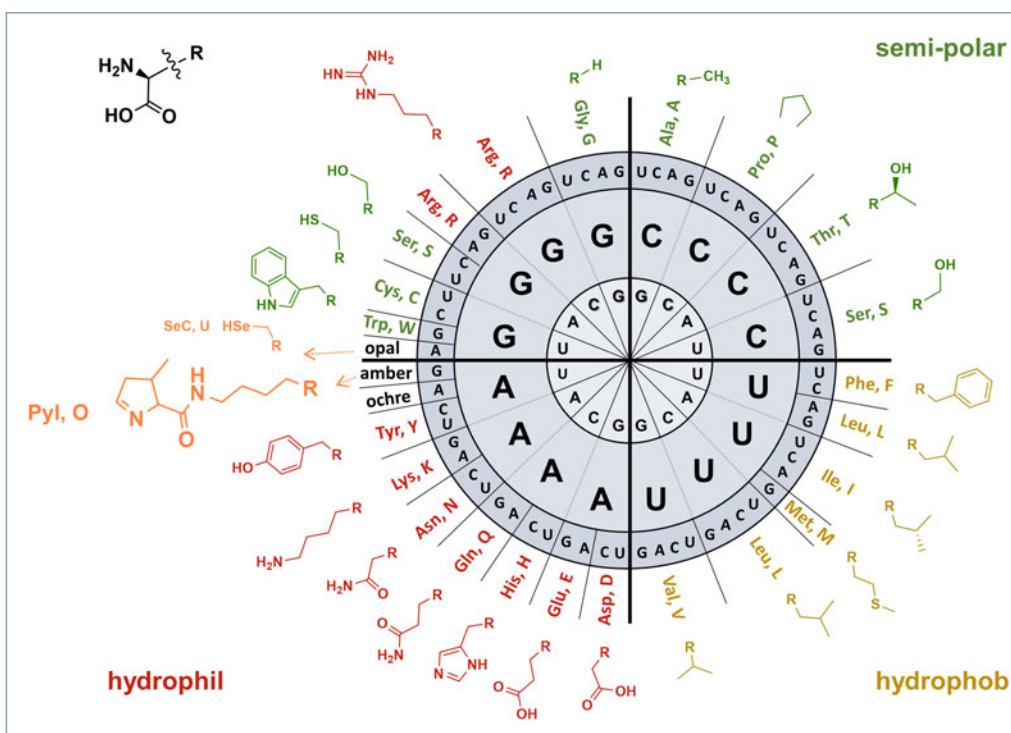
© Die Autoren 2021

Die Grundlage allen Lebens sind Proteine. Sie katalysieren alle essenziellen Reaktionen in Organismen, ohne die Leben auf der Erde unmöglich wäre. Fast alle Lebewesen verwenden dafür dieselben 20  $\alpha$ -Aminosäuren

(kanonische) als Bausteine. Der Bauplan für Proteine ist durch die DNA- und deren korrespondierende RNA-Sequenz codiert (Abb. 1). Spezifisch besteht die Information für eine Aminosäure aus einem Basentriplet,

welches ein Codon bildet. Es gibt 64 solcher Codons. 61 Codons spezifizieren, welche Aminosäure an welcher Stelle einer RNA-Sequenz eingebaut wird, und drei Codons terminieren die Proteinsequenz (Stopp-Codons). Da alle Lebewesen die gleichen 20 Aminosäuren für die Synthese aller Proteine verwenden, ist ein gezieltes Manipulieren einer spezifischen Aminosäure in einer Proteinsequenz, z. B. durch chemische Reaktionen, fast unmöglich.

Ein großes Ziel war es deshalb, andere Aminosäuren durch die Erweiterung des genetischen Codes ortsspezifisch einzuführen, um Proteine zur Funktionsaufklärung gezielt manipulieren oder ihre Funktionen erweitern zu können [1]. Eine Methode, um das zu bewerkstelligen, ist das gezielte Umprogrammieren eines der drei Stopp-Codons. Hierzu werden in einen Organismus tRNA und die dazugehörige Aminoacyl-tRNA-Synthetase eingeführt, welche nicht mit der endogenen Translationsmaschinerie



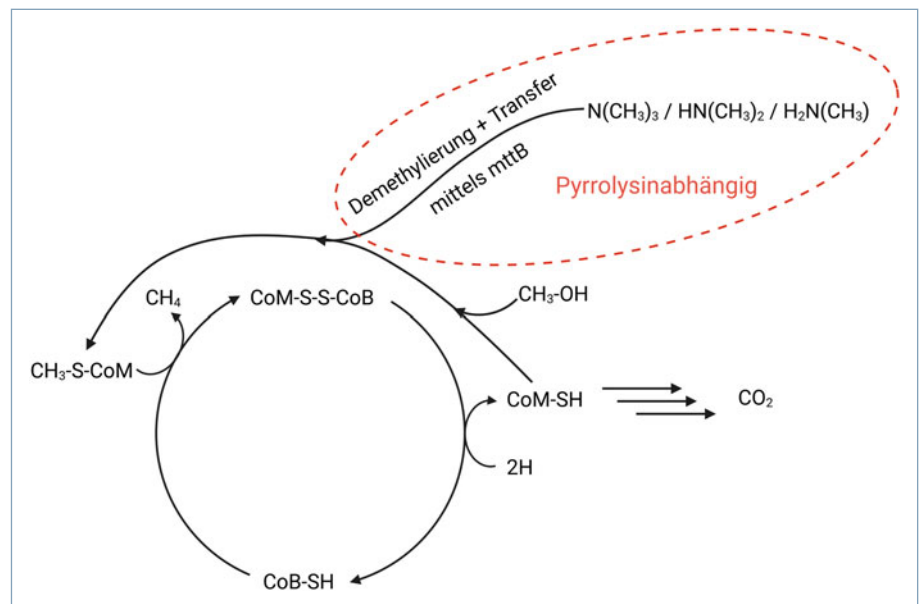
◀ **Abb. 1:** Schematische Darstellung des genetischen Codes in Form der Code-Sonne, sortiert nach der zweiten Base. Die 5'- zu 3'-Richtung des Codons ist von innen nach außen. Die Formeln der Aminosäuresseitenketten sind mit entsprechenden drei- und einbuchstabigen Abkürzungen versehen. Die Aminosäuren Selenocystein (Sec, U) und Pyrrolysine (Pyl, O) (orange hervorgehoben) sind spezielle kanonische Aminosäuren, die in der Literatur oft als die 21. (Sec) und 22. (Pyl) Aminosäuren des genetischen Codes bezeichnet werden. Sie werden entweder in einem speziellen Sequenz-Kontext (Sec, UGA, Opal-Codon) oder durch natürliche orthogonale Paare (Pyl, UAG, Amber-Codon) in Proteine eingebaut.

interagiert (also orthogonal ist) und außerdem so designt wurden, dass sie ausschließlich eine nicht kanonische Aminosäure (nkAS) in Korrespondenz zu einem Stopp-Codon in die Proteinsequenz einführen. Erstmals gelang das in *Escherichia coli* mit einem orthogonalen Phenylalanin-System aus der Hefe [2]. Später wurden auch verbesserte orthogonale Translationssysteme (OTS) aus Archaeen entwickelt [3]. Zu den von Menschen entwickelten Systemen kam durch eine serendipitäre Entdeckung ein natürlich vorkommendes hinzu. Bei der Aufklärung der Methanogenese eines speziellen Methanbildners, *Methanosarcina barkeri*, wurde in dem Enzym Monomethylamin-Methyltransferase ein Amber-Stopp-Codon in der DNA-Sequenz gefunden. Durch die Röntgenstrukturaufklärung dieses Enzyms wurde herausgefunden, dass das Amber-Stopp-Codon für eine weitere kanonische Aminosäure codiert (die 22., nach Selenocystein) [4]. Es handelte sich um ein Lysin-Derivat mit einem 4-Methyl-Pyrroline-5-Carboxylat-Ring verknüpft am N<sup>ε</sup> des Lysins, welches auf den Namen Pyrrolysin (Pyl) getauft wurde (Abb. 1).

### Verflechtung der Evolutionsgeschichte von Pyrrolysin mit „Alien“-Leben

Höchstwahrscheinlich sind wir Zeugen eines evolutionären Prozesses der Umwidmung des Amber-Stopp-Codons der *Methanosarcina*-Arten, welche in Lebensräumen gedeihen können, die für andere Organismen unwirtlich sind. Diese Umgebungen (z. B. Rinderpansen oder die tiefe Biosphäre des Meers) werden zwar auch von anaeroben Bakterien bewohnt, aber von Archaeen dominiert. Archaeen gehören einer Lebensdomäne an, die sich phylogenetisch sowohl von Eukaryoten als auch von Bakterien deutlich unterscheidet [5].

Es wird auch spekuliert, dass extraterrestrisch bewohnbare Zonen und Planeten mit Methanvorkommen in der Atmosphäre außerirdische Lebensformen beherbergen könnten, die methanogenen Archaea-Zellen ähnlich oder mit ihnen verwandt sind [6]. Da die Methanogenese ein einzigartiger Stoffwechsel ist, führt sie in diesem Zusammenhang zur Umcodierung von Organismen, indem Pyrrolysin als spezielle kanonische Aminosäure in ihren genetischen Code aufgenommen wird. Dies ist nicht überraschend, denn jüngste phylogenetische Analysen weisen eindeutig darauf hin, dass



▲ **Abb. 2:** Schematischer Mechanismus der methyloptrophen Methanogenese aus Methanol am Beispiel von Trimethylamin (TMA). Zu beachten ist, dass die Seitenkette der speziellen kanonischen Aminosäuren Pyl für den Demethylierungsprozess unerlässlich ist. Diese evolutionäre Neuerung erforderte eine Erweiterung des genetischen Codes in *Methanosarcinae* und ermöglichte so einen ungewöhnlichen Stoffwechsel in extremen Lebensräumen. CoM-SH: Coenzym-M-Thiol; CoB-SH: Coenzym-B-Thiol; CoM-S-S-CoB: Coenzym-M-Coenzym-B-Heterodisulfid; mttB: TMA-Methyltransferase-Gen. Mit BioRender.com erstellt.

PylRS ein evolutionäres Überbleibsel einer frühen archaealen Innovation des genetischen Codes ist [7].

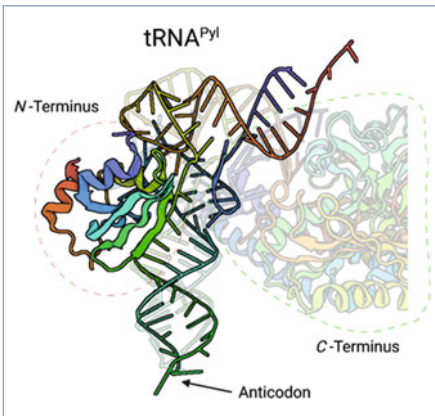
Kürzlich haben wir das Evolutionsmodell „Alanin-Welt“ vorgeschlagen, das erklärt, wie der alte genetische Code (in der RNA-Welt) mit nur vier Aminosäurebausteinen begann: Alanin (Ala), Glycin, Prolin (Pro) und Arginin/Lysin-Äquivalente (wie Ornithin oder Pyl). Die für den Stoffwechsel und die Translation ausgewählten Aminosäuren sind hauptsächlich Derivate von Ala, was durch die erdspezifische geochemische Umgebung in der Evolution des Lebens diktiert wurde. Es kann darüber spekuliert werden, ob auf der Erde oder in einem anderen Teil des Universums, bei Wiederholung der Geschichte, ein auf Pyrrolysin basierendes Leben entstehen könnte, das eine grundlegend andere Chemie (Stoffwechsel) und/oder sogar eine andere Energetik des Lebens aufweist [8].

Nichtsdestotrotz haben wir jetzt eine gute Möglichkeit, diese Optionen aus der evolutionären Vergangenheit wieder aufzugreifen und experimentell parallele biologische Welten zu schaffen, wie z. B. eine potenzielle Pyrrolysin-Welt. Es sei auch darauf hingewiesen, dass Pyrrolysin als ein zusammengesetztes Molekül betrachtet werden kann, das aus methyliertem/oxidiertem Prolin-

Analogon und Lysin besteht (siehe Pyl, **Abb. 1**). Kürzlich wurde eine Reihe von Pyl-Derivaten mit D-Prolin und Analoga erzeugt und in Proteine eingebaut, die ihnen einzigartige katalytische Eigenschaften verleihen (Dieter Söll, persönliche Kommunikation). Sowohl Pro als auch Pyl sind nicht nur die ältesten Aminosäuren im Aminosäurerepertoire des genetischen Codes, sondern auch wichtig für die Produktion von homochiralen Zuckern und die Vermittlung von chemischen Transformationen, die in den frühen Stadien der Evolution lebenswichtig waren (wie Transfer-Hydrierungen oder Transaminierungen) [9].

### Entdeckung des natürlichen Pyrrolysin-tRNA-Synthetase-Systems

Pyrrolysin spielt eine essenzielle Rolle in der Methanogenese aus Methylaminen. Generell gibt es drei Arten der Methanogenese: (i) Die am weitesten verbreitete Variante verwendet H<sub>2</sub> und CO<sub>2</sub>, um Methan zu synthetisieren. (ii) Die Bildung aus Acetat. (iii) Die am wenigsten verbreitete Variante, die methyloptrophe Methanogenese aus organischen Substraten, die eine Methylgruppe beinhalten. Sie wird hauptsächlich von Organismen aus der Ordnung der *Methanosarcinales* durchgeführt (**Abb. 2**, [10]).



▲ **Abb. 3:** Erkennungsmechanismus der PyIRS von der tRNA<sup>Pyl</sup>. Überlagerung der Kristallstruktur der N-terminalen Domäne von *Methanosarcina mazei* mit Bindung an die tRNA (PDB ID 5UD5) und der C-terminalen Domäne von *Desulfotribacterium hafniense* und dessen tRNA (PDB ID 2ZNI). N- und C-Terminus der PyIRS erkennen die tRNA im Bereich des Anticodonarms. Dieser besitzt im Gegensatz zu kanonischen tRNAs nur einen winzigen variablen Arm, was die Orthogonalität zu allen anderen kanonischen tRNA/aaRS-Paaren erklärt. Auch ist zu erkennen, dass das Anticodon nicht am Erkennungsmechanismus beteiligt ist. Mit BioRender.com erstellt.

Methanosarcinales sind die einzigen Organismen, bei denen bekannt ist, dass sie auf allen drei Wegen Methan bilden können. Untersuchungen zeigten, dass zum Einbau von Pyrrolysin in die Proteinsequenz nur die spezifische tRNA (tRNA<sup>Pyl</sup>) und die dazugehörige Pyrrolysyl-tRNA-Synthetase (PylRS) notwendig ist (**Abb. 3**). Es handelt sich also um ein natürlich entstandenes Äquivalent zu den oben beschriebenen synthetisch gemachten OTS [11].

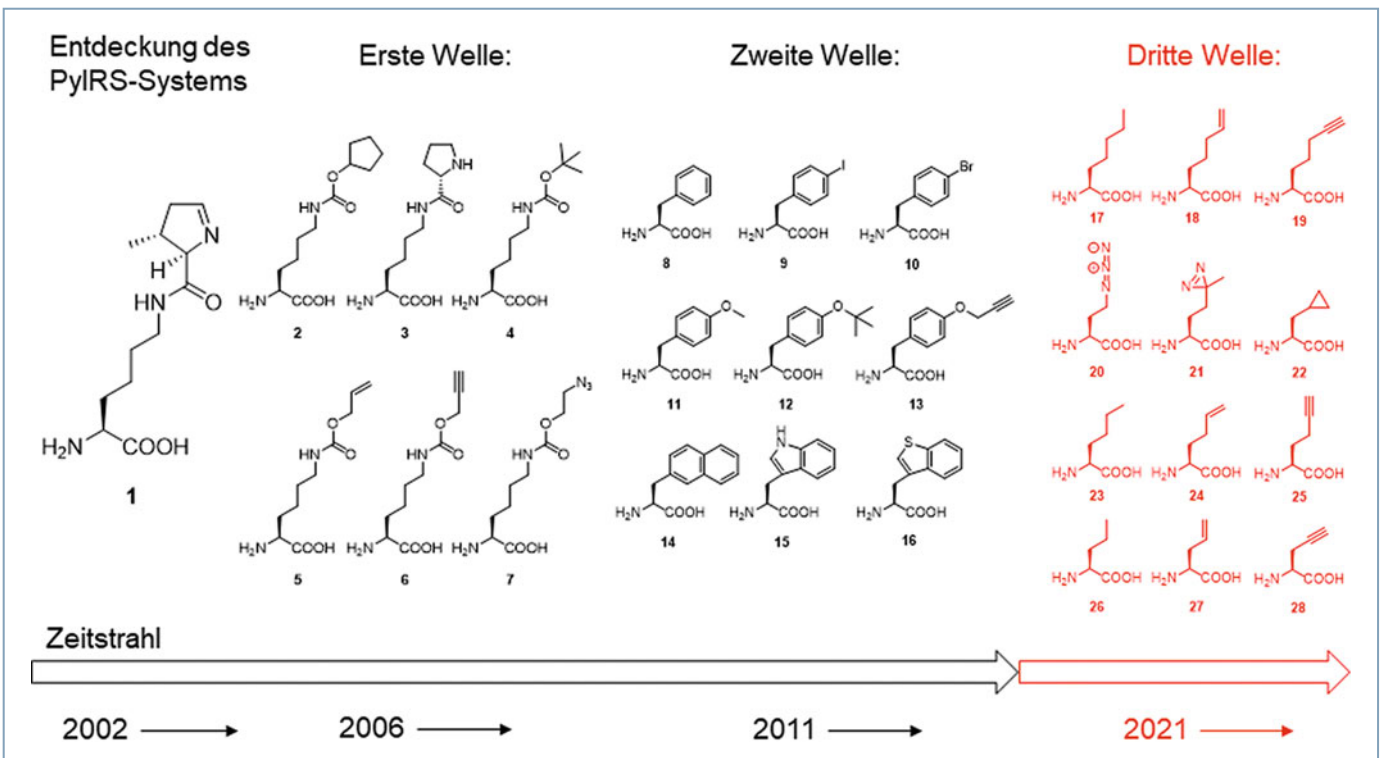
### Die erste Welle: Erste Nutzung des PyIRS-Systems für die Erweiterung des genetischen Codes

Die PylRS ist ein Enzym, welches einen vergleichsweise unspezifischen Mechanismus zur Substraterkennung besitzt. Die Erkennung findet über eine tiefe Enzymtasche und bedingt unspezifischen hydrophoben Wechselwirkungen mit dem endogenen Substrat statt. Diese Eigenschaft wurde Anfang der 2000er-Jahre ausgenutzt, um mit dem nativen oder leicht mutierten PylRS-System eine Reihe von biochemisch/strukturbioologisch nützlichen Pyrrolysin-Derivaten in rekombinante Proteine einzubauen [12]. Obgleich so einige nicht kanonische Aminosäuren einge-

baut wurden, die bioorthogonale funktionelle Gruppen aufwiesen (hauptsächlich Alkin-, Alken- und Azid-Gruppen), waren die Anwendungsgebiete doch deutlich eingeschränkt. Da es sich um Pyrrolysin-Analoga handelte, waren alle Substrate lang und voluminös. Ein Hauptproblem ist die hohe Flexibilität der langen Seitenkette. Sie verursacht z. B. Probleme in spektroskopischen Anwendungen bzw. limitiert den zu erhaltenden Informationsgehalt deutlich.

### Die zweite Welle: Engineerte PyIRS-Enzyme für den Einbau von Phenylalanin und dessen Derivate

Kavran und Kollegen fiel die hohe strukturelle Ähnlichkeit zwischen der PylRS und der bakteriellen Synthetase für den Einbau von Phenylalanin (PheRS) auf [7]. Das war eine bemerkenswerte Leistung, da die Sequenzhomologie von PylRS und PheRS sehr gering ist. Daraufhin führten die Arbeiten von Liu und Kollegen dazu, dass das Spektrum der Aminosäuren, welche mit dem PylRS-System eingebaut werden konnten, deutlich erweitert wurde [12]. Es konnten nun Phenylalanin und dessen Derivate eingebaut werden. Das löste zum Teil das Problem der hohen



▲ **Abb. 4:** Zeitliche Abfolge der möglichen einzubauenden Substrate. Rot sind die in unserer Arbeitsgruppe neu eingebauten Substrate. Dies ist nur ein kleiner Bruchteil aller möglichen nicht kanonischen Aminosäurerepertoires, die mithilfe von PyIRS-basierten orthogonalen Paaren in rekombinante Proteine eingebaut werden können. Die außerordentliche Vielseitigkeit des natürlichen PylRS-Enzyms in Verbindung mit ausgereiften Methoden für die gerichtete Evolution von Enzymen (die auch die Verwendung von KI einschließen) hat eine solch dramatische Erweiterung des Anwendungsbereichs der ribosomalen Proteinbiosynthese ermöglicht.

Flexibilität; auch konnte die Palette an funktionellen Gruppen noch erweitert werden (Cyano-Gruppen als spektroskopische Sonde). Diese Substrate waren allerdings immer noch sehr voluminös.

### Die dritte Welle: PylRS-Enzyme für den Einbau von nicht aromatischen kleinen Substraten

Die erste kleinere und nicht aromatische Aminosäure wurde von unserer Arbeitsgruppe *in vivo* eingebaut. Es handelte sich um ein Cystein-Derivat mit einer angehängten Allyl-Gruppe (S-Allyl-L-cystein, Sac) [13]. Leider war die Effizienz, mit der diese Aminosäure eingebaut werden konnte, relativ schlecht. Weiterführende Arbeiten zur Verbesserung der Effizienz konnten zeigen, dass die Mutationen, die notwendig sind, damit die PylRS das Cystein-Derivat erkennt, zu einer deutlichen Destabilisierung des Enzyms führen. Somit lag das Enzym in einer unlöslichen und unfunktionalen Konformation im Proteinproduktionsorganismus (*E. coli*) vor. Wir konnten dieses Problem beheben, indem wir ein deutlich löslicheres Protein genetisch an das PylRS-Enzym fusionierten [14]. Mithilfe dieses Fusionsproteins wurden die Positionen aufgeklärt, die wichtig zur Erkennung von Sac sind. Außerdem bildete dieses Enzym dann die Grundlage für weitere Varianten, mit deren Hilfe erstmals kleine Aminosäuren (**Abb. 4**) mit einer Reihe von biochemisch nützlichen funktionellen Gruppen in Proteinsequenzen eingebaut werden konnten [15]. Unter anderem wird der Einbau dieser Substrate in verschiedene Proteine/Enzyme dabei helfen, ihre genauen Funktionsmechanismen spektroskopisch aufzuklären zu können.

Der Einbau dieser neuen Funktionen könnte dann auch mit der *in vivo*-nkAS-Synthese kombiniert werden. So könnten Zellen mithilfe des PylRS-OTS dann in den einfachsten Nährmedien Proteine/Enzyme mit nicht natürlichen Funktionen/Eigenschaften herstellen. Zurzeit werden den Produktionskulturen von außen hauptsächlich chemisch synthetisierte Aminosäureanaloga zugeführt,

was einen erhöhten Aufwand bedeutet. Daher ist die Kopplung der billigen und effizienten metabolischen Produktion von gewünschten nkAS mit OTS eine wichtige Richtung in der zukünftigen Forschung, die sowohl für den akademischen Bereich als auch für die biotechnologische Industrie wichtig ist [16]. ■

### Literatur

- [1] Chin JW (2017) Expanding and reprogramming the genetic code. *Nature* 550: 53–60
- [2] Furter R (1998) Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in *Escherichia coli*. *Protein Sci* 7: 419–426
- [3] Wang L, Brock A, Herberich B, Schultz PG (2001) Expanding the genetic code of *Escherichia coli*. *Science* 292: 498–500
- [4] Hao B, Gong W, Ferguson TK et al. (2002) A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* 296: 1462–1466
- [5] Rothman DH, Fournier GP, French KL et al. (2014) Methanogenic burst in the end-permian carbon cycle. *Proc Natl Acad Sci* 111: 5462–5467
- [6] Taubner R-S, Pappenreiter P, Zwicker J et al. (2018) Biological methane production under putative Enceladus-like conditions. *Nat Commun* 9: 748
- [7] Kavran JM, Gundllapalli S, O'Donoghue P et al. (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc Natl Acad Sci U S A* 104: 11268–11273
- [8] Kubyshkin V, Budisa N (2019) Anticipating alien cells with alternative genetic codes: away from the alanine world! *Curr Opin Biotechnol* 60: 242–249
- [9] Kubyshkin V, Budisa N (2019) The alanine world model for the development of the amino acid repertoire in protein biosynthesis. *Int J Mol Sci* 20: 5507
- [10] Enzmann F, Mayer F, Rother M, Holtmann D (2018) Methanogens: biochemical background and biotechnological applications. *AMB Express* 8: 1
- [11] Namy O, Zhou Y, Gundllapalli S et al. (2007) Adding pyrrolysine to the *Escherichia coli* genetic code. *FEBS Lett* 581: 5282–5288

- [12] Wan W, Tharp JM, Liu WR (2014) Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim Biophys Acta – Proteins Proteomics* 1844: 1059–1070
- [13] Exner MP, Kuenzl T, To TMT et al. (2017) Design of S-allylcysteine in situ production and incorporation based on a novel pyrrolysyl-tRNA synthetase variant. *ChemBioChem* 18: 85–90
- [14] Koch NG, Baumann T, Budisa N (2021) Efficient unnatural protein production by pyrrolysyl-tRNA synthetase with genetically fused solubility tags. Eingereicht.
- [15] Koch NG, Goettig P, Rappsilber J, Budisa N (2021) Engineering pyrrolysyl-tRNA synthetase for the incorporation of non-canonical amino acids with smaller side chains. Eingereicht.
- [16] Völler J-S, Budisa N (2017) Coupling genetic code expansion and metabolic engineering for synthetic cells. *Curr Opin Biotechnol* 48:1–7

**Funding note:** Open Access funding enabled and organized by Projekt DEAL.  
**Open Access:** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

### Korrespondenzadresse:

Prof. Dr. Nediljko Budisa  
 Institut für Chemie  
 Fakultät II  
 Technische Universität Berlin  
 Müller-Breslau-Straße 10  
 D-10623 Berlin  
 nediljko.budisa@tu-berlin.de  
 nediljko.budisa@umanitoba.ca

### AUTOREN



#### Nikolaj Georg Koch

2010–2018 Studium Economics (Bachelor) und Chemie (Bachelor und Master) an der TU Berlin (TUB). 2017 Teamleiter des TU BIOMOD-Teams. Seit 2018 Doktorand im AK Biokatalyse (TUB) unter der Leitung von Prof. Dr. N. Budisa. Seine Forschung konzentriert sich auf die Erweiterung und Evolution des genetischen Codes und die Entwicklung von Technologien, die auf diesem Ansatz basieren.



#### Nediljko (Ned) Budisa

Studium Biologie & Chemie, molekulare Biologie und molekulare Biophysik. 1997 Promotion. 1997–2000 Postdoc am Max-Planck-Institut für Biochemie, Martinsried. 2001–2005 Habilitation in Biochemie, TU München. Seit Mai 2010 W3-Professor (Biokatalyse) an der TU Berlin. Seit 2018 Professor für Chemie und Inhaber des Tier 1 Canada Research Chair (CRC) für chemische synthetische Biologie an der University of Manitoba, Winnipeg, Kanada.