DOI: 10.1002/cbic.200900302

Hacking the Genetic Code of Mammalian Cells

Dirk Schwarzer*[a]

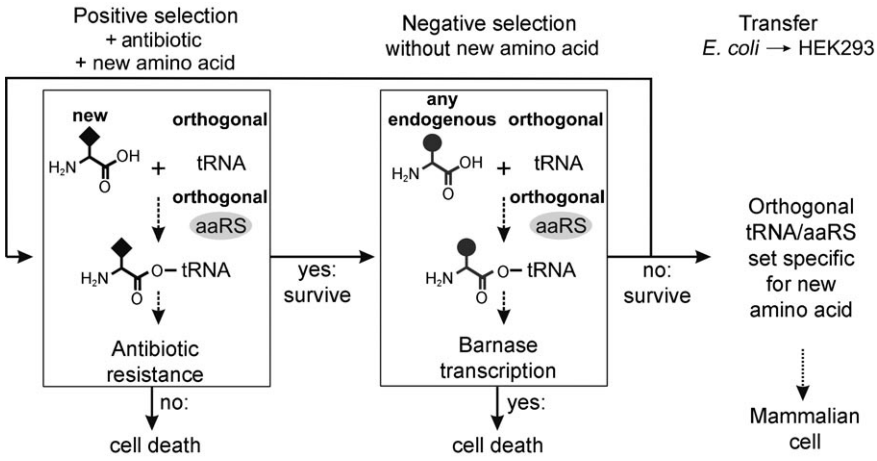
The proteins of all living organisms consist of the same canonical twenty amino acids encoded by the universal genetic code. Although these standard amino acids are sufficient to support the biochemical processes of life, additional amino acids with new chemical properties serve as excellent tools to study biological systems or even enhance certain properties of whole organisms. In this regard a general method has been developed that allows the addition of new amino acids to the genetic repertoire of living organisms.[1]

The first genetically enhanced organism was an *E. coli* strain, which was able to incorporate the unnatural amino acid *o*-methyl-L-tyrosine with high fidelity into its proteins.[2] This was achieved by installing an additional aminoacyl-tRNA synthetase (aaRS) and tRNA set that was specific for the new amino acid. The requirements for such an orthogonal tRNA/aaRS pair are stringent because specific incorporation of the new amino acid is only possible when neither the new tRNA is aminoacylated by an endogenous aaRS nor does the new aaRS charge any endogenous tRNAs. Furthermore, the orthogonal tRNA needs a dedicated codon for specific incorporation of the new amino acid. This can be achieved by suppression of the amber codon (UAG), which usually serves as one of three stop codons. However, so-called amber suppressor tRNAs exist, which efficiently incorporate amino acids in response to the amber codon, and orthogonal tRNAs can be engineered to suppress the codon by incorporating the new amino acid at this site.

[a] *Dr. D. Schwarzer*
Section for Chemical Biology
Leibniz-Institut für Molekulare Pharmakologie
Robert-Rössle-Strasse 10
13125 Berlin (Germany)
Fax: (+49) 30-94793-159
E-mail: schwarzer@fmp-berlin.de

Orthogonal aaRS and tRNA pairs used in *E. coli* are often derived from Archaea because they differ significantly from their eubacterial counterparts. The orthogonal tRNA/aaRS pairs still need to be evolved specifically for the new amino acids. This process begins with a randomized library of aaRS with active-site mutation affecting their substrate recognition. This library is passed through a series of positive and negative selections to identify mutants that can incorporate the desired amino acid with high fidelity (Scheme 1). The initial positive selection is based on amber suppression in an antibiotic-resistance gene. The cells are then cultivated on media supplemented with both the antibiotic and the new amino acid, and survive when the orthogonal tRNA/aaRS set incorporates the new amino acid in response to the amber codon and the antibiotic-resistance conferring protein is expressed. In the following negative selection, orthogonal tRNA/aaRS pairs that suppress the amber codon with an endogenous amino acid are removed from the library. In this case the cells contain a toxic barnase gene disrupted by amber codons. When these cells are cultivated on media without the new amino acid, incorporation of any endogenous amino acids results in cell death due to expression of the full-length barnase protein. After repetitive rounds of positive and negative selection tRNA/aaRS pairs can be obtained that are highly specific for the new amino acid, which is now encoded by the amber codon. This strategy has become a general tool for expanding the genetic code of *E. coli* and is also used to optimize the incorporation efficiency of new amino acids and the orthogonality of tRNAs.[1]

Based on this technology a manifold of unnatural amino acids have been added to the genetic code of *E. coli*. A prominent example is a benzophenone-derived amino acid, which serves as a photocrosslinker for studying protein–protein interactions.[3] This methodology has been demonstrated for a mutant glutathione *S*-transferase (GST) with the photocrosslinker amino acid incorporated at the dimer interface. Irradiation of cells expressing the modified GST resulted in efficient cross-linking of the dimers in the cytoplasm of *E. coli*.[4] This amino acid was accommodated by an evolved *Methanocaldococcus jannaschii* tyrosyl–tRNA/TyrRS set, which has become the scaffold for encoding many novel amino



Scheme 1. General selection strategy for evolving orthogonal tRNA/aaRS pairs.

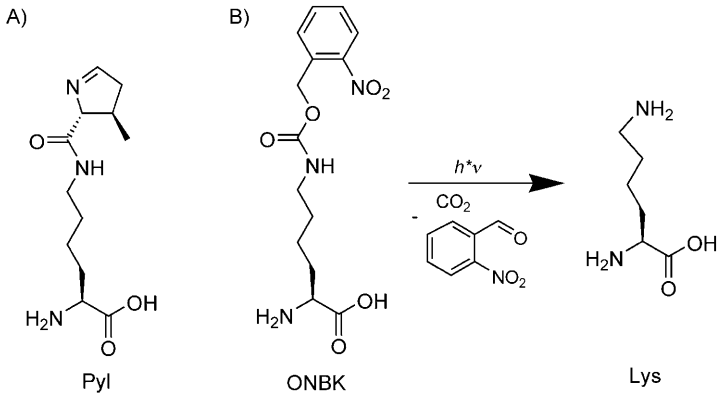
acids in *E. coli*. Other variants with useful applications include photocaged amino acids, fluorescent amino acids and residues carrying bioorthogonal functional groups for bioconjugation reactions.[1]

Efforts have also focused on enlarging the genetic repertoire of eukaryotic organisms. The genetic tools for eukaryotes, and for mammalian cells in particular, are more limited than those for *E. coli*. Ideally, it should be possible to simply transfer an orthogonal tRNA/aaRS set evolved in *E. coli* into mammalian cells without further alterations. However, this is not feasible in the case of the *M. jannaschii* tyrosyl–tRNA/TyrRS pair because it is not orthogonal in most eukaryotes. Alternative strategies use tRNA/aaRS pairs, which are orthogonal in eukaryotes, from other bacteria. To this end a tyrosyl–tRNA/TyrRS pair from *E. coli* and *Bacillus stearothermophilus* could be engineered by rational design to recognize and incorporate 3-iodo-tyrosine into mammalian proteins.[5] Furthermore, a scheme for evolving bacterial tRNA/aaRS sets in yeast has been developed and used for subsequent transfer into mammalian cells.[6]

Despite these recent advances, the genetic background of *E. coli* still provides advantages for evolving proteins, and a system in which a tRNA/aaRS pair can be evolved in *E. coli* and utilized in mammalian cells without further modifications is highly desirable. Schultz, Geierstanger et al. have now linked the *E. coli* selection system to amber suppression in mammalian cells in a "shuttle system".[7] The basic requirement is a tRNA/aaRS set that is orthogonal in both *E. coli* and mammalian cells. A pair fulfilling this prerequisite is the well studied pyrrolysyl–tRNA/PylRS, which incorporates pyrrolysine (Pyl) by amber suppression in proteins of several Archaea (Scheme 2 A).[8] Previously, the *Methanosarcina mazei* pyrrolysyl–tRNA/PylRS set was shown to maintain orthogonality in mammalian cells and could be used to include Pyl derivatives into proteins of this host.[9] In *E. coli* orthogonality of an analogous pyrrolysyl–tRNA/PylRS set from *M. barkeri* was reported and used to incorporate Pyl analogues into proteins.[10] Addition of N^{ε}-acetyllysine through amber suppression after altering the specificity of the *M. bakeri* PylRS has also been demonstrated.[11] Furthermore, the specificity of the *M. mazei* PylRS could be altered by rational design and used to incorporate N^{ε}-protected lysine derivatives into proteins expressed in *E. coli*.[12]

In the initial step of establishing the shuttle system the authors reconfirmed the orthogonality of the *M. mazei* pyrrolysyl–tRNA/PylRS pair in *E. coli* and mammalian cells, and continued to evolve a mutant PylRS specific for *o*-nitrobenzyloxycarbonyl-N^{ε}-L-lysine (ONBK; Scheme 2 B). This photocaged lysine derivative is a useful tool for controlling the onset of lysine modifications in vitro and in vivo. Based on the selection scheme described above for *E. coli* and the solved crystal structure, a focused library of mutant PylRS proteins with randomized substrate-recognition residues was created and subjected to several selection cycles in either the presence or absence of ONBK.[7,12,13] After three positive and two negative rounds of selection, two ONBK specific mutant pyrrolysyl–tRNA/PylRS pairs were obtained. The ONBK incorporation in *E. coli* proceeded efficiently, and the more potent pair—referred to as NBK-1—was shuttled into human HEK293 cells. ONBK-dependent amber suppression was monitored by the expression of enhanced green fluorescent protein (EGFP). Expression of full-length EGFP in HEK293 cells was only observed in the presence of ONBK, and ESI-MS analysis confirmed the incorporation of ONBK into EGFP. Finally, the EGFP protein was irradiated to remove the caging group of ONBK and the uncaging yielded a lysine residue at the site of ONBK incorporation.

The establishment of this "shuttle system" represents an important step toward controlling the chemical composition of mammalian proteins. With respect to the enormous scientific task of uncovering the complex networks of biochemical and signaling processes in mammals the chemical tools provided by new amino acids will be very useful. At this point it remains to be seen if the *M. mazei* PylRS can give rise to the same functional diversity as was realized by evolving the *M. jannaschii* TyrRS. Encouragingly, protein evolution and rational design have already demonstrated the flexibility of the PylRS scaffold; this gives rise to the expectation that further amino acids can be added to the genetic repertoire of mammalian cells in the near future.[7,11,12]



Scheme 2. Structures recognized by natural and evolved PylRSs: A) pyrrolysine (Pyl), B) *o*-nitrobenzyloxycarbonyl-N^{ε}-L-lysine (ONBK), which can be uncaged to lysine.

Acknowledgements

I would like to thank Michael Gerrits, Christian Hackenberger and Christian Freund for critically reading the manuscript. I also gratefully acknowledge funding by the Emmy Noether Program of the Deutsche Forschungsgemeinschaft (DFG).

Keywords: amber suppression · amino acids · cage compounds · genetic code · protein evolution

[1] a) L. Wang, J. Xie, P. G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225–249; b) J.

Xie, P. G. Schultz, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775–782; c) N. Budisa, *Angew. Chem.* **2004**, *116*, 6586–6624; *Angew. Chem. Int. Ed.* **2004**, *43*, 6426–6463.
[2] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498–500.
[3] J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11020–11024.
[4] J. W. Chin, P. G. Schultz, *ChemBioChem* **2002**, *3*, 1135–1137.
[5] a) D. Kiga, K. Sakamoto, K. Kodama, T. Kigawa, T. Matsuda, T. Yabuki, M. Shirouzu, Y. Harada, H. Nakayama, K. Takio, Y. Hasegawa, Y. Endo, I. Hirao, S. Yokoyama, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9715–9720; b) K. Sakamoto, A. Hayashi, A. Sakamoto, D. Kiga, H. Nakayama, A. Soma, T. Kobayashi, M. Kitabatake, K. Takio, K. Saito, M. Shirouzu, I. Hirao, S. Yokoyama, *Nucleic Acids Res.* **2002**, *30*, 4692–4699.
[6] W. Liu, A. Brock, S. Chen, S. Chen, P. G. Schultz, *Nat. Methods* **2007**, *4*, 239–244.
[7] P. R. Chen, D. Groff, J. Guo, W. Ou, S. Cellitti, B. H. Geierstanger, P. G. Schultz, *Angew. Chem.* **2009**, *121*, 4112–4115; *Angew. Chem. Int. Ed.* **2009**, *48*, 4052–4055.
[8] J. A. Krzycki, *Curr. Opin. Microbiol.* **2005**, *8*, 706–712.
[9] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818–822.
[10] C. R. Polycarpo, S. Herring, A. Bérubé, J. L. Wooda, D. Söll, A. Ambrogelly, *FEBS Lett.* **2006**, *580*, 6695–6700.
[11] H. Neumann, S. Y. Peak, S. Y. Chew, J. W. Chin, *Nat. Chem. Biol.* **2008**, *4*, 232–234.
[12] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187–1197.
[13] T. Yanagisawa, R. Ishii, R. Fukunaga, O. Nurekia, S. Yokoyama, *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* **2006**, *62*, 1031–1033.

Received: May 13, 2009
Published online on June 16, 2009

