Rare Amino Acids

DOI: 10.1002/anie.201103769

Biosynthesis and Charging of Pyrrolysine, the 22nd Genetically Encoded Amino Acid

*Christian Hertweck**

amber codon · amino acids · biosynthesis · genetic code · protein engineering

Considering the vast number of diverse proteins and enzymes that support all kinds of cellular functions, it is amazing that a basic set of only 20 canonical amino acid residues is sufficient to meet all needs. This holds true for virtually all cases, yet a few enzymes require the 21st proteinogenic amino acid, the cysteine homologue selenocysteine. Only recently, there has been another remarkable addition to the family of genetically encoded amino acids: pyrrolysine (Pyl, **1**).[1] Krzycki and colleagues found that this unprecedented lysine homologue is incorporated into several methyltransferases from archaebacteria, for example monomethylamine methyltransferase (MtmB; Figure 1) from

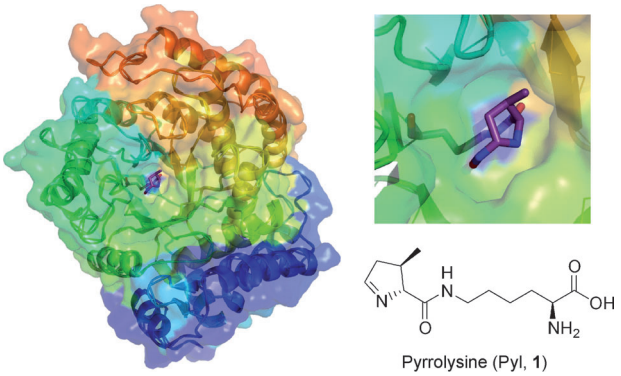
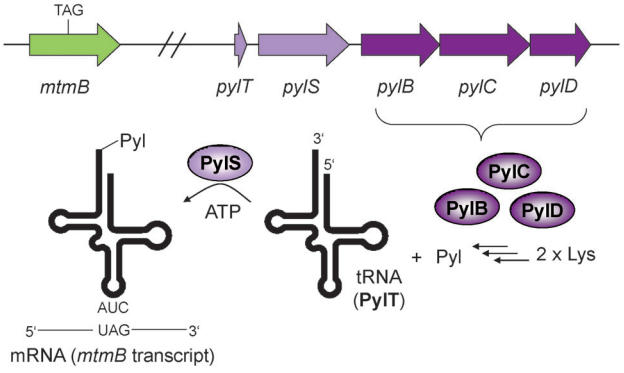


Figure 1. Molecular structures of monomethylamine methyltransferase MtmB (PDB 1V2) and pyrrolysine (Pyl, **1**). View into the substrate channel with Pyl positioned in the active site (inset: magnification).

Methanosarcina barkeri.[2] The exotic Pyl residue is placed into the active site of MtmB, where it is indispensable for the catalytic function of the corrinoid-dependent enzyme. During the past decade, genetic and biochemical investigations have provided an ample body of knowledge on the coding and incorporation of Pyl, yet little has been known about its biosynthetic origin. Two independent studies by the Krzycki[3] and Geierstanger[4] laboratories now shed more light on the Pyl biosynthetic pathway. In conjunction with previous work, these results imply an intriguing merger of two amino acids into one.

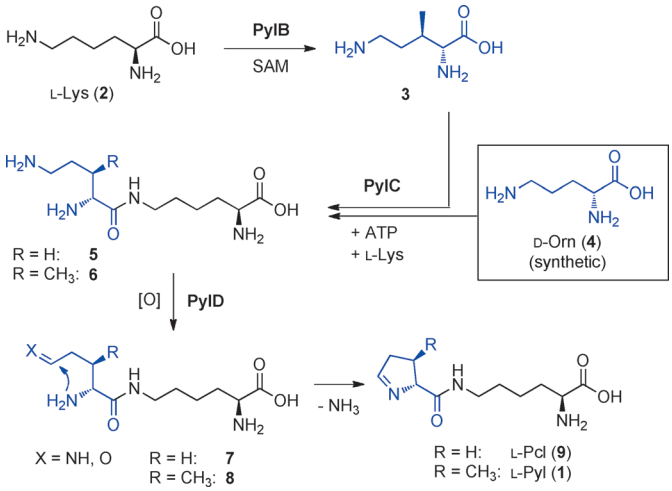
This detective work actually started in 2002 with the finding of a specific codon (TAG), which normally causes the termination of protein biosynthesis, within the reading frame of the *mtmB* gene (Scheme 1).[1] A closer examination of the *mtmB* gene product revealed the presence of Pyl, which implied that the incorporation of the rare 22nd amino acid involves suppression of the stop codon.[2] A scenario where a so-called in-frame amber codon programs the introduction of an amino acid has precedence for selenocysteine.[5] However, unlike the path known for the formation of selenocysteinyl-tRNA, charging of pyrrolysine involves a specialized aminoacyl-tRNA synthetase, PylS, which loads Pyl onto a designated tRNA, PylT.[6] Interestingly, in *Methanosarcina* spp. the genes coding for PylS and PylT are located in a small gene cluster (*pylTSBCD*) near the gene coding for MtmB (Scheme 1).

Through heterologous expression of *mtmB* together with *pylTSBCD* in *E. coli* it was shown that this "genetic code expansion cassette" is required and sufficient to confer onto the heterologous host the ability to produce functional MtmB harboring Pyl;[7] this indicates that the gene products of *pylBCD* are enzymes involved in Pyl biosynthesis. Another interesting finding was that the addition of D-ornithine (D-Orn) to the heterologous expression host seemed to increase



Scheme 1. Schematic representation of the *mtmB* gene featuring the TAG amber codon, and the genetic code expansion cassette for Pyl biosynthesis (*pylBCD*) and incorporation (*pylTS*) into MtmB.

[*] Prof. Dr. C. Hertweck
Dept. Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology, HKI
Beutenbergstr. 11a, 07745 Jena (Germany)
and
Chair for Natural Product Chemistry, Friedrich Schiller University Jena (Germany)
E-mail: christian.hertweck@hki-jena.de
Homepage: http://www.hki-jena.de



Scheme 2. Model for the biosynthesis of pyrrolysine (Pyl) from two lysine (Lys) units.

MtmB titers, and it was thus initially suggested that D-Orn was a precursor of Pyl.[8] However, a closer inspection disclosed that in lieu of Pyl, a desmethyl variant, pyrroline-carboxylysine (Pcl, **9**, Scheme 2), was produced by PylBCD from D-Orn[3] and incorporated into the enzyme.[3]

Stable isotope-labeling experiments by Krzycki and co-workers have now disclosed that D-Orn is not a biosynthetic intermediate, but instead Lys is the sole precursor of Pyl.[3] MtmB was heterologously produced by *E. coli* supplemented with [$^{13}C_6{}^{15}N_2$]lysine, and mass spectrometric analysis of protein fragments (tryptic digests) showed a mass shift of 15 Da, which pointed out that Pyl is derived from two Lys residues. Yet, one amino group is lost along the biosynthetic pathway. Since one ^{15}N label of administered [ε-^{15}N]Lys was missing in Pyl, while [α-^{15}N]Lys was fully incorporated, it became apparent that the ε-amino group is lost, likely during heterocyclization.[3] Furthermore, through MS analyses of *pylBCD*-expressing *E. coli* cultures that were grown with and without D-Orn, Geierstanger and colleagues could detect either Pcl or Pyl, revealing that these amino acids are fully synthesized before the aminoacyl-tRNA adduct is formed.[4]

Results from further in vivo and in vitro experiments that were independently performed in the Krzycki and Geierstanger laboratories support the following biosynthetic model: PylB seems to catalyze the first step in Pyl biosynthesis as its absence can be compensated by addition of D-Orn to give Pcl (**9**).[4] Moreover, addition of synthetic 3-methyl-D-Orn (**3**) can complement a strain lacking PylB.[4] Sequence comparisons suggest that PylB is related to radical SAM enzymes, some of which function as mutases. It is thus readily conceivable that PylB represents an aminomutase that could mediate the mechanistically challenging transformation of L-Lys into **3** with inversion of the configuration at the α carbon. PylC, which shows sequence homology to D-amino acid ligases, could ligate the ε-amino group of Lys to either D-Orn (**4**) or **3** to yield the dipeptides **5** or **6**. An in vitro assay with PylC showed ATP binding and turnover, but no specific activity could be observed.[4] However, mass spectrometry data indicates that PylC can generate dipeptide **5**, at least in vivo when D-Orn is added to an *E. coli* strain expressing *pylC*.[3] At the dipeptide level the putative dehydrogenase PylD could oxidize the ε-amino group to the corresponding imine, setting the stage for a subsequent, possibly spontaneous condensation–heterocyclization. Indeed, an in vitro study revealed that purified PylD transforms dipeptide **5** into Pcl (**9**) in the presence of NAD^+ and ATP.[4] Finally, Pyl (or, alternatively, Pcl) is activated by the Pyl-specific aminoacyl-tRNA synthetase PylS with consumption of ATP, and loaded onto the Pyl-specific tRNA that is encoded by *pylT* (Scheme 1).

In sum, three enzymes are apparently sufficient to transform two Lys units into one Pyl residue. From an evolutionary point of view, it is remarkable that the TAG amber codon for Pyl differs in only one position from the AAG codon for Lys. According to the co-evolution theory, amino acids emerging from the same precursors would have similar codon assignments—which is the case for Lys and Pyl. Biosynthesis and charging of Pyl are not only mechanistically intriguing, but have also practical applications since the Pyl translational machinery can be exploited to expand the genetic code.[9] To date, cotranslational insertion of synthetic pyrrolysine analogues by PylS and PylT has been successfully applied to enable, for example, protein click chemistry[10] and site-specific protein ubiquitination, respectively.[11] A deeper insight into the Pyl pathway and the factors governing substrate specificities could be employed to develop new ways for engineering Pyl analogues and functionalized proteins in vivo.

Received: June 2, 2011
Published online: July 27, 2011

[1] G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459–1462.
[2] B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462–1466.
[3] M. A. Gaston, L. H. Zhang, K. B. Green-Church, J. A. Krzycki, *Nature* **2011**, *471*, 647–650.
[4] S. E. Celitti, W. Ou, H. P. Chiu, J. Grünewald, D. H. Jones, X. Hao, Q. Fan, L. L. Quinn, K. Ng, A. T. Anfora, S. A. Lesley, T. Uno, A. Brock, B. H. Geierstanger, *Nat. Chem. Biol.* **2011**, *7*, 528–530.
[5] A. Ambrogelly, S. Palioura, D. Söll, *Nat. Chem. Biol.* **2007**, *3*, 29–35.
[6] S. K. Blight, R. C. Larue, A. Mahapatra, D. G. Longstaff, E. Chang, G. Zhao, P. T. Kang, K. B. Green-Church, M. K. Chan, J. A. Krzycki, *Nature* **2004**, *431*, 333–335.
[7] D. G. Longstaff, R. C. Larue, J. E. Faust, A. Mahapatra, L. H. Zhang, K. B. Green-Church, J. A. Krzycki, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1021–1026.
[8] O. Namy, Y. Zhou, S. Gundllapalli, C. R. Polycarpo, A. Denise, J. P. Rousset, D. Söll, A. Ambrogelly, *FEBS Lett.* **2007**, *581*, 5282–5288.
[9] T. Fekner, M. K. Chan, *Curr. Opin. Chem. Biol.* **2011**, *15*, 387–391.
[10] T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem.* **2009**, *121*, 1661–1663; *Angew. Chem. Int. Ed.* **2009**, *48*, 1633–1635.
[11] X. Li, T. Fekner, J. J. Ottesen, M. K. Chan, *Angew. Chem.* **2009**, *121*, 9348–9351; *Angew. Chem. Int. Ed.* **2009**, *48*, 9184–9187.

