Published in final edited form as:

Nature. 2011 March 31; 471(7340): 583–584. doi:10.1038/471583a.

How two amino acids become one

Stephen W. Ragsdale

Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, Michigan 48109-0606, USA

Stephen W. Ragsdale: sragdsal@umich.edu

Abstract

Twenty amino acids form the basis of all proteins, but another two genetically encoded amino acids have also been discovered. The biosynthesis of one of these, pyrrolysine, has now been elucidated.

The first genetically encoded amino acid was identified more than two centuries ago, but new ones are still being found. The most recently reported one — the twenty-second — is pyrrolysine (Pyl), which was found^{1,2} in 2002 at the active sites of methyltransferase enzymes obtained from a methane-producing archaeon. Like the 20 common amino acids that are incorporated into cellular proteins, Pyl is synthesized in the cytoplasm and incorporated at a specific position in a growing polypeptide chain during translation³. However, it was the only genetically encoded amino acid for which a biosynthetic pathway had not been established. On page 647 of this issue, Krzycki and colleagues⁴ report that the essential amino acid lysine is the sole precursor of Pyl, and they define the enzymatic steps for the conversion of two L-lysine molecules into one molecule of L-Pyl.

Translation occurs on ribosomes, and involves decoding a series of nucleotide triplets (codons) on a messenger RNA strand into a corresponding series of amino acids. Prior to translation, an aminoacyl-tRNA synthetase enzyme catalyses the attachment of each amino acid to a transfer RNA, forming an aminoacyl-tRNA molecule. During translation, the ribosome transfers the growing protein chain carried on the preceding tRNA (the peptidyl-tRNA) to the next incoming amino acyl-tRNA. Translation continues step-wise in this way until the ribosome reaches a stop codon, which triggers specific factors to release the polypeptide chain into the cell.

In previous studies^{1,2}, Krzycki and coworkers found that, surprisingly, a specific stop codon (UAG) in the sequence of methyltransferase enzymes from methanogenic archaea encodes Pyl. They went on to show⁵ that the biological machinery associated with Pyl's synthesis and incorporation into proteins is encoded by the *pylTSBCD* cluster of genes, which can be thought of as a genetic-code expansion cassette — a gene cluster that, when transferred into an organism, enables that organism's ribosomes to recognize UAG and translate it into Pyl. The ribosome therefore 'reads through' this stop codon and catalyses a reaction between the Pyl-tRNA and the preceding peptidyl-tRNA to insert Pyl into a protein. The protein PylS was found to be the Pyl-tRNA synthetase⁶, whereas PylT was identified⁷ as the Pyl-tRNA. The remaining *pylBCD* genes in the cassette were therefore expected to encode the enzymatic pathway for Pyl biosynthesis.

It has been proposed that Pyl derives from lysine and some other cellular compound — possibly one of the amino acids D-ornithine⁸, D-glutamate⁹, D-isoleucine⁸ or D-proline^{5,8}. But Krzycki and colleagues⁴ have now demonstrated that the protein products of the *pylBCD* genes catalyse the synthesis of Pyl from two lysines using the pathway shown in Figure 1.

The authors began by genetically engineering a common laboratory strain of the bacterium *Escherichia coli* to include the *pylTSBCD* expansion cassette from the archaeon *Methanosarcina acetivorans* and the methyltransferase gene *mtmB1* from another archaeon, *Methanosarcina barkeri*. The mRNA sequence of *mtmB1* contains the UAG stop codon that specifies Pyl. Krzycki *et al.* then provided the engineered *E. coli* with lysine in which all six carbon atoms and both nitrogen atoms were isotopically labelled, and later purified the methyltransferase produced by the organism.

To decipher the biosynthetic pathway for Pyl, the authors used mass spectrometry to accurately measure the masses of peptide fragments (produced *in situ* in the mass spectrometer) of the purified methyltransferase. By comparison with a similar analysis of methyltransferase purified from engineered *E. coli* grown in unlabelled lysine, they identified a single labelled Pyl-containing peptide fragment. Further mass spectrometry experiments unambiguously revealed that all 12 carbon atoms in the Pyl residue and all three of its nitrogen atoms were isotopically labelled. Because lysine contains six carbons and two nitrogens, the results conclusively demonstrated that two molecules of lysine combine to produce Pyl, and that one of the lysines eliminates a nitrogen atom during the PylBCD-catalysed biosynthetic pathway (Fig. 1). In other words, no precursor other than lysine is used in the biosynthesis of Pyl.

These results are surprising in light of a report⁸ that D-ornithine stimulates UAG read-through in an *E. coli* strain similarly engineered to contain the *pylTSBCD* expansion cassette — a finding that suggests that D-ornithine is a precursor of Pyl. To investigate the apparent disparity, Krzycki and co-workers⁴ performed mass spectrometric analysis of the methyltransferase obtained from engineered *E. coli* cultures grown in a medium supplemented with both unlabelled D-ornithine and labelled lysine. They discovered that some of this methyltransferase contained labelled Pyl, as before. However, part of the protein contained desmethylpyrrolysine, an amino acid in which the methyl group of Pyl has been replaced by a hydrogen atom. Desmethyl pyrrolysine can be made from one lysine and one D-ornithine, suggesting that D-ornithine was charged onto Pyl-tRNA by the Pyl-tRNA synthetase and thereby misincorporated into the methyltransferase. This implies that the Pyl biosynthetic cassette could be used to incorporate useful modified amino-acid residues into proteins — something that is of interest to many research laboratories.

One limitation of Krzycki and colleagues' study⁴ is that the *pylBCD*-encoded proteins were not purified and used to demonstrate their proposed activities. But on the basis of the similarity of the amino-acid sequences of PylB, PylC and PylD to other proteins whose functions are known, the biosynthetic pathway proposed by the authors is reasonable and chemically feasible. The door is now open for enzymologists to study the Pyl bio-synthetic pathway in detail. A prime target for investigation is the PylB-catalysed lysine mutase reaction, in which an aminoethyl group ($\text{CH}_2\text{CH}_2\text{NH}_2$) shifts from one part of the molecule to another (Fig. 1a). This is particularly interesting because the amino-acid sequence of PylB suggests that it is a member of the radical S-adenosylmethionine protein family¹⁰, which is not currently known to catalyse this reaction. More broadly, these findings will help us to better understand the relationship between the evolution of the genetic code and of amino-acid biosynthetic pathways.

References

1. Hao B, et al. Science. 2002; 296:1462–1466. [PubMed: 12029132]
2. Srinivasan G, James CM, Krzycki JA. Science. 2002; 296:1459–1462. [PubMed: 12029131]
3. Rother M, Krzycki JA. Archaea. 2010; 2010:453642.10.1155/2010/453642 [PubMed: 20847933]

4. Gaston MA, Zhang L, Green-Church KB, Krzycki JA. *Nature*. 2011; 471:647–650. [PubMed: 21455182]
5. Longstaff DG, et al. *Proc Natl Acad Sci USA*. 2007; 104:1021–1026. [PubMed: 17204561]
6. Polycarpo C, et al. *Proc Natl Acad Sci USA*. 2004; 101:12450–12454. [PubMed: 15314242]
7. Blight SK, et al. *Nature*. 2004; 431:333–335. [PubMed: 15329732]
8. Namy O, et al. *FEBS Lett*. 2007; 581:5282–5288. [PubMed: 17967457]
9. Krzycki JA. *Curr Opin Chem Biol*. 2004; 8:484–491. [PubMed: 15450490]
10. Frey PA, Hegeman AD, Ruzicka FJ. *Crit Rev Biochem Mol Biol*. 2008; 43:63–88. [PubMed: 18307109]

Figure 1. Proposed biosynthesis of pyrrolysine

Krzycki and colleagues⁴ report that ϵ -pyrrolysine forms from two ϵ -lysine molecules in archaea, and propose the following biosynthetic pathway. **a**, In the presence of a cofactor (*S*-adenosylmethionine, SAM), the protein PylB catalyses the conversion of ϵ -lysine to 3-methyl-D-ornithine, a molecular-rearrangement reaction. **b**, PylC then catalyses the ATP-dependent combination of 3-methyl-D-ornithine with another ϵ -lysine to make 3-methyl-D-ornithyl- ϵ -lysine (ATP is an energy-carrying cofactor). **c**, **d**, Finally, PylD catalyses an oxidative deamination reaction (in which an NH₂ group is eliminated as ammonia, NH₃), which is followed by cyclization and dehydration steps to yield ϵ -pyrrolysine. It is not currently clear whether PylD catalyses the transformation shown in **d**, or whether this is a spontaneous process. Fragments of the molecules are colour-coded to make the reactions easier to follow.