proteinspotlight
> ONE MONTH, ONE PROTEIN <

Issue 25, August 2002

www.proteinspotlight.org

Life's jokers

Vivienne Baillie Gerritsen

There is more to the genetic code than meets the eye. We are acquainted with the dogma: 'One codon, one amino acid'. Life, however, has found a way of wriggling out of this straight jacket by using its stop codons as jokers. One example is given by our own mitochondrial DNA where the translational machinery recognises the UGA stop codon as tryptophan, and not as the classical stop codon as in the nucleus. Tryptophan though is an old-timer. As are the great majority of the now classical twenty amino acids which have been on the scene for over a century. Threonine was the last of the classics to be discovered in 1936. 1986, though, was another scoop year, when it was discovered that the UGA codon could produce a new amino acid altogether: selenocysteine. Selenocysteine, the 21st amino acid, is found in archaea, eubacteria and animals. Selenocysteine may just sound like a modified cysteine, but it is not because it has its own tRNA which is like granting it a passport. Similarly, in May 2002, the existence of a 22nd amino acid was reported: pyrrolysine.

Pyrrolysine was discovered in the sequences of three enzymes which participate in the production of methane in the archaeon *Methanosarcina barkeri*: mono-, di- and trimethylamine methyltransferases. Like selenocysteine, it is also encoded by a stop codon – the UAG or amber codon – and has its own tRNA. The scientists who reported the existence of a 22nd amino acid had known for months that the amber codon in *M.barkeri* was used to produce a lysine residue. There remained doubts however as to whether this residue was modified or not. The structure of the monomethylamine methyltransferase was then determined in the hope that the lysine residue would reveal a little intimacy. And it did. It turned out to be a derivative of a lysine residue to which the group (4R,5R)-4-substituted-pyrroline-5-carboxylate was added. The modified lysine was subsequently called pyrrolysine. What happens is that a lysine residue which has already settled on its tRNA is modified just before squeezing into the ribosome to be added to the growing protein.

The quest for amino acids started in the beginning of the 19th century. The first two were discovered around 1820 by the French chemist Henri Braconnot (1780-1855) who was, at the time, studying the effects of sulfuric acid on animal substances. In so doing, gelatin produced what he called '*sucre de gélatine*', literally 'gelatin sugar', which was later named glycocoll or glycine. Sulfuric acid converted wool and muscle fibre to a white substance which he called leucine, from the Greek '*leukos*' meaning white. And it took a further century to discover the remaining eighteen amino acids.

1st letter	Second Letter: U	Second Letter: C	Second Letter: A	Second Letter: G	3rd letter
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA Stop UAG Stop	UGU Cys UGC Cys UGA Stop UGG Trp	U C A G
C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G

The genetic code as we knew it

How did they know that these amino acids were the naturally occurring constituents of proteins? They did not. In fact, in the days of Braconnot, the very designation 'amino acid', or indeed 'protein', was still unknown. The Swedish chemist Jöns Jakob Berzelius (1779-1848) suggested the term 'amino acid' for nitrogen-containing organic acids in 1820. In 1902, a revolutionary breakthrough was made independently – and during the same meeting! –

by two German scientists: the organic chemist Emil Fischer (1852-1919) and the physiologist Franz Hofmeister (1850- 1922). The breakthrough was the notion of the peptide bond – a notion that has made protein chemistry what it is today. Scientists could finally make some sense out of the chemical formulae they had so laboriously calculated, and proteins became the strings of amino acids as we understand them today. In the 1940s, the Canadian chemist Hubert Bradford Vickery (1893-1978) sorted all the known amino acids into four groups, one of which turned out to be the list of amino acids produced by way of protein hydrolysis.

So, in 2002, the list has become longer still. Why a 22nd amino acid? Are twenty not sufficient? *M.barkeri* methylamine transferases are hexamers of barrel-shaped monomers. Each barrel has a cavity in which is lodged the novel amino acid pyrrolysine, which suggests a role in catalysis. *M.barkeri* breaks down methylamines, the ultimate product of which is methane. The methyltransferases assist in the first transfer of the methylamine methyl group to cofactors, and the pyrrolysine residue may promote this transfer by literally feeding the methyl group to the cofactor.

So life has its teasers and its jokers. And scientists too are doing their best to trick the laws of life. What if they could construct original amino acids – along with their tRNAs – with novel structural, chemical and physical properties, and pop them into the sequence of an enzyme to enhance, or perform, a specific activity? Enzymes are widely used in the food, textile and paper industry already. Expanding the genetic code to fit our fancies is tantalising. A first attempt has already been carried out in *Escherichia coli*, and what is more successfully; its ribosome accepted a novel amino acid seated in its own tRNA.

It has taken the best part of two centuries to root out the first 22 amino acids which make up the genetic code. Life clearly has many more surprises in store; there is little doubt that genetically encoded amino acids – still unknown today – will surface in the future.

Cross-references to Swiss-Prot

Monomethylamine methyltransferase, *Methanosarcina barkeri* : O30642

References

1. Hao B., Gong W., Ferguson T.K., James C.M., Krzycki J.A., Chan M.K.
   A new UAG-encoded residue in the structure of a methanogen methyltransferase
   Science 296:1462-1466(2002)
   PMID: 12029132

2. Atkins J.F., Gesteland R.
   The 22nd amino acid
   Science 296:1409-1410(2002)
   PMID: 12029118

3. Wang L, Brock A., Herberich B., Schultz P.G.
   Expanding the genetic code of *Escherichia coli*
   Science 292:498-500(2001)
   PMID: 11313494

4. Tanford C., Reynolds J.
   Nature's Robots, A history of proteins, pp. 11-20, pp. 29-41
   Oxford University Press (2002)

Protein Spotlight (ISSN 1424-4721), http://www.proteinspotlight.org, is published by the Swiss-Prot group at the Swiss Institute of Bioinformatics (SIB).