Commentary

Making sense out of nonsense

Margaret E. Saks*

Department of Biology, University of Oregon, Eugene, OR 97403

There is considerable interest among protein scientists in methods that permit unnatural amino acids to be incorporated at specific sites in proteins. Such methods would facilitate studies of protein structure and function and would allow for the synthesis of proteins having novel properties. Current methods use nonsense-suppressing tRNAs (1) to incorporate an unnatural amino acid at the desired site in a protein and chemical (rather than enzymatic) means to attach the unnatural amino acid to the nonsense-suppressing tRNA (2). However, protein yields are typically modest because the suppressor tRNA participates in only one round of translation. Protein expression could be increased by the development of *in vivo* systems, but this poses significant challenges. It requires a synthetase that activates an unnatural amino acid and only attaches it to a designated tRNA. In addition, it requires that the designated tRNA is not aminoacylated by any other synthetase in the cell and that it efficiently translates a nonsense codon. A "21st cognate pair" is essentially created if all of these criteria are met (Fig. 1). The paper by Kowal *et al.* (3) in this issue of PNAS makes an important contribution by describing a novel approach for creating tRNA-synthetase pairs that efficiently incorporate natural amino acids at specific sites in proteins *in vivo*.

To incorporate a desired amino acid at a specific codon, the translation machinery must be manipulated without disrupting its overall accuracy and efficiency. Each amino acid is normally activated by a single aminoacyl-tRNA synthetase, which subsequently attaches the amino acid to only the corresponding tRNA. The charged tRNA is then bound by an elongation factor that transports it to the ribosome. At the ribosome, complementary base-pairing between tRNA anticodons and mRNA codons dictates the order of addition of amino acids to the growing polypeptide chain. Ordinarily, there are no tRNAs that read the three nonsense (stop) codons. However, tRNAs competent to insert an amino acid in response to nonsense codons can be engineered by altering the tRNA anticodon (4–6). Coexpression of a suppressor and a gene having an introduced nonsense codon
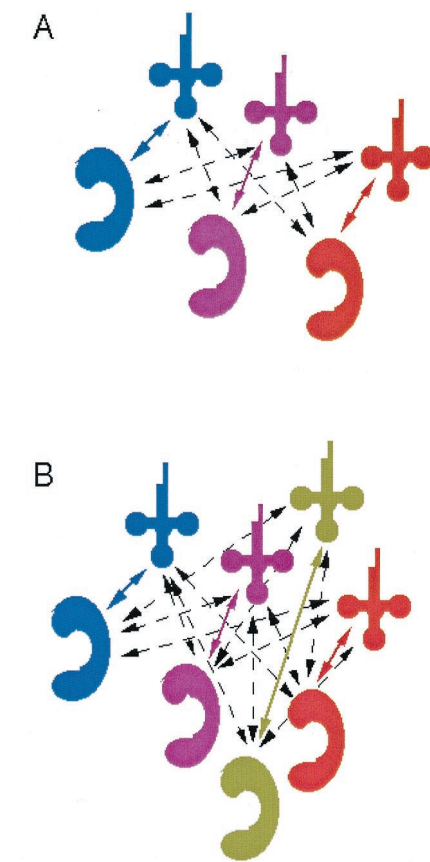

Fig. 1. Network of potential interactions among tRNAs (cloverleaf) and synthetases ("C-shape"). (*A*) Hypothetical wild-type case having three cognate pairs. In the typical cell, about 1,200 potential interactions are possible. (*B*) The increase in network complexity that arises when a new cognate pair is introduced. Colored arrows indicate cognate interactions, and dashed arrows indicate interactions that would result in amino acid misincorporation.

gives a full-length protein. Thus, this method permits a spectrum of desired amino acids to be inserted at a single site in a protein.

Engineering an endogenous tRNA with the requisite characteristics is problematic. Abundant structural and functional data (7) make it relatively easy to identify nucleotide changes in a tRNA that discourage aminoacylation by the cognate synthetase. The difficulty lies in identifying changes that do not promote recognition by another synthetase. All of the

synthetases and the 20 types of tRNA are coexpressed in the same cellular compartment, resulting in a complex network of potential interactions among these macromolecules. The subset of correct interactions that gives rise to accurate amino acid incorporation thus depends on a proper balance among the concentrations of tRNAs and synthetases (8, 9) as well as on positive elements that promote productive interactions with the cognate synthetase and negative elements that discourage interactions with the 19 noncognate synthetases (10, 11).

The sequence space available for creating a tRNA that is not recognized by the endogenous synthetases but that is recognized by a "21st" synthetase is limited (11). tRNAs are comprised of only about 74–90 nt; and typically, nucleotides in only a few regions of the tRNA are specifically contacted by synthetases (7, 12–15). Consequently, changes that discourage aminoacylation by the cognate synthetase often concomitantly encourage aminoacylation by a noncognate synthetase(s). The problem is particularly acute for nonsense suppressors (6). By definition, these tRNA variants have an altered anticodon trimer. Because most synthetases recognize at least one of the three anticodon nucleotides (15), a change in the anticodon trimer can abolish recognition by the cognate synthetase but can concomitantly promote recognition by another synthetase in the cell (6, 15, 16).

Fortunately, there are loopholes in the rules governing tRNA recognition. In particular, the identity of nucleotides that comprise the complex network of positive and negative interactions among the tRNAs and synthetases differs from organism to organism. Thus, when a tRNA is expressed in a heterologous system, it often accepts an amino acid but does so with a reduced efficiency (7, 17–19). The introduction of mutations into these debilitated tRNAs can render them unrecognizable by any endogenous synthetase in the cell, giving them one of the characteristics required of tRNA for the site-specific incorporation of amino acids.

Engineering a 21st synthetase that activates an unnatural amino acid has proven to be extremely difficult. The problem can be circumvented by chemical (2, 20), rather than enzymatic, aminoacylation of a nonsense suppressor and the subsequent introduction of the acylated tRNA into an *in vitro* translation system (2). This approach has been used extensively and successfully with a broad range of amino acid analogues (21). However, protein yields are limited by inherent properties of *in vitro* translation systems and by the inability of the engineered tRNA to participate in multiple rounds of translation. An approach was developed to overcome some of these limitations. It involved chemical acylation of a nonsense suppressor followed by its microinjection into *Xenopus* oocytes (22). Using this method, proteins such as membrane-bound receptors can be expressed at levels suitable for their characterization in intact eukaryotic cells. Nevertheless, protein expression is limited because the tRNA participates in only one round of translation.

Protein yields can be increased by using *in vivo* expression systems and by ensuring that the nonsense suppressor efficiently reads the corresponding mRNA codon. Recently, two cognate pairs were developed for use in intact *Escherichia coli* cells (23, 24). One pair is specific for glutamine and is comprised of a yeast tRNA and synthetase. The other pair is specific for tyrosine and is comprised of an archaeon tRNA and synthetase. Although these pairs can be used *in vivo*, protein yields are modest due to the low translation efficiency of the heterologous nonsense suppressors. The efficiency with which tRNA reads mRNA codons depends on proper covariations between the last (3′) nucleotide of the anticodon and nucleotides within the anticodon loop and stem (25). Thus, nonsense-suppressing tRNAs are usually designed according to these rules. Nevertheless, their efficiency is typically lower than that of wild-type tRNAs. Two alternative strategies help to circumvent this problem. In one strategy, the interaction between release factor and stop codons is impaired, making stop codons accessible to the suppressor tRNA (26, 27). Another strategy uses the unusual codon assignments of *Tetrahymena* in which the codon UAG encodes glutamine rather than stop (28, 29). An engineered variant of the tRNA[Gln] that reads this codon has a high translation efficiency, is not recognized by the endogenous syn-

thetases of *Xenopus* oocytes, and is thus a useful vehicle for the delivery of unnatural amino acids (30).

By taking a novel approach in their choice of tRNAs and synthetases for membership in a new cognate pair, RajBhandary's group has created two *in vivo* systems that substantially alleviate the protein expression problem (3). The approach is based on their previous studies comparing the sequences and properties of elongator and initiator tRNA[Met] (31, 32) and uses tRNA from one organism and a synthetase from another to create new cognate pairs. The *E. coli* system uses an engineered *E. coli* nonsense-suppressing initiator tRNA[Met] that participates in protein chain elongation rather than initiation, that is not aminoacylated by any of the *E. coli* synthetases, but that is aminoacylated by yeast tyrosyl-tRNA synthetase (TyrRS). Because yeast TyrRS also recognizes *E. coli* tRNA[Pro] it was subjected to mutagenesis and selection to eliminate recognition of this noncognate tRNA. The eukaryotic system uses an engineered human nonsense suppressing initiator tRNA[Met] that participates in protein chain elongation. It is not aminoacylated by any mammalian synthetases but is aminoacylated by *E. coli* glutaminyl-tRNA synthetase (GlnRS). Because *E. coli* GlnRS does not recognize any of the mammalian tRNAs, it is an ideal member of a new cognate pair. By deriving the nonsense suppressor from an endogenous tRNA and pairing it with a synthetase from a different organism, two cognate pairs of tRNAs and synthetases were created, one for use in *E. coli* and the other for use in yeast. The partners in each pair "know" each other but do not cross-react with other tRNAs or synthetases in the cell. The engineering of endogenous tRNAs is a crucial aspect of these new reagents because in heterologous systems, incompatibilities in the details of tRNA-ribosome interactions, and in the recognition of tRNA by the processing and modification enzymes, can limit the rate of protein synthesis. By avoiding these problems, the new cognate pairs give improved protein expression.

> By taking a novel approach in their choice of tRNAs and synthetases for membership in a new cognate pair, RajBhandary's group has created two *in vivo* systems that substantially alleviate the protein expression problem.

Now that new cognate pairs are available, attention will turn to identifying unnatural amino acids suitable for use *in vivo*. To be incorporated into protein, an unnatural amino acid must be transported into the cell and not degraded. Strategies are currently being developed to identify unnatural amino acids with these properties (33). In addition, the availability of entire genome sequences for a broad range of organisms may help in identifying "nonstandard" pathways having the desired uptake characteristics. To be a useful tool, the 21st synthetase must activate the amino acid analogue of choice and must exclude amino acids with similar characteristics. Otherwise, the expressed protein will contain a mixture of amino acids at the site of interest. In some synthetases, certain residues participate in amino acid binding and activation as well as in tRNA binding and amino acid transfer (34, 35). Thus, changes that alter amino acid recognition or activation might either disrupt the transfer of the amino acid to tRNA or redirect tRNA specificity, causing the novel amino acid to be incorporated at an unintended site(s) in the protein. Given the complexity of the problem, rational design is unlikely to yield a solution. Accordingly, several labs are combining limited mutagenesis of synthetase genes combined with *in vivo* or *in vitro* selection schemes to identify synthetases with the desired characteristics (3, 33). These screens for unnatural amino acid incorporation will undoubtedly provide new insights into the mechanisms by which synthetases recognize amino acids and tRNAs and the interactions that aminoacyl-tRNA makes with other macromolecules of the translation machinery. Once unnatural amino acids can be routinely inserted into protein, it will be possible to create proteins with novel characteristics that will be useful for addressing a broad range of basic and applied problems.

1. Steege, D. A. & Söll, D. G. (1979) in *Biological Regulation and Development*, ed. Goldberger, R. F. (Plenum, NY), pp. 433–485.
2. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. (1989) *Science* **244,** 182–188.
3. Kowal, A. K., Köhrer, C. & RajBhandary, U. L.

(2001) *Proc. Natl. Acad. Sci. USA* **98,** 2268–2273. (First Published January 23, 2001; 10.1073/pnas.031488298)
4. Normanly, J., Masson, J.-M., Kleina, L. G., Abelson, J. & Miller, J. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 6548–6552.

5. Kleina, L. G., Masson, J. M., Normanly, J., Abelson, J. & Miller, J. H. (1990) *J. Mol. Biol.* **213,** 705–717.
6. Normanly, J., Kleina, L. G., Masson, J.-M., Abelson, J. & Miller, J. H. (1990) *J. Mol. Biol.* **213,** 719–726.

7. Giegé, R., Sissler, M. & Florentz, C. (1998) *Nucleic Acids Res.* **26,** 5017–5035.

8. Yarus, M. (1972) *Nat. New Biol.* **239,** 106–108.

9. Swanson, R., Hoben, P., Sumner-Smith, M., Uemura, H., Watson, L. & Söll, D. (1988) *Science* **242,** 1548–1551.

10. Normanly, J. & Abelson, J. (1989) *Annu. Rev. Biochem.* **58,** 1029–1049.

11. McClain, W. H. & Nicholas, H. B. (1987) *J. Mol. Biol.* **194,** 635–642.

12. McClain, W. H. (1993) *J. Mol. Biol.* **234,** 257–280.

13. Saks, M. E., Sampson, J. R. & Abelson, J. N. (1994) *Science* **263,** 191–197.

14. Martinis, S. A. & Schimmel, P. (1995) in *tRNA: Structure, Biosynthesis, and Function*, eds. Söll, D. & RajBhandary, U. L. (Am. Soc. Microbiol., Washington, DC), pp. 349–370.

15. Pallanck, L., Pak, M. & Schulman, L. H. (1995) in *tRNA: Structure, Biosynthesis, and Function*, eds. Söll, D. & RajBhandary, U. (Am. Soc. Microbiol., Washington, DC), pp. 371–394.

16. Saks, M. E., Sampson, J. R. & Abelson, J. (1998) *Science* **279,** 1665–1670.

17. Shiba, K., Motegi, H. & Schimmel, P. (1997) *Trends Biochem. Sci.* **22,** 453–457.

18. Nair, S., Ribas de Pouplana, L., Houman, F., Avruch, A., Shen, X. & Schimmel, P. (1997) *J. Mol. Biol.* **269,** 1–9.

19. Wakasugi, K., Quinn, C. L., Tao, N. & Schimmel, P. (1998) *EMBO J.* **17,** 297–305.

20. Heckler, T. G., Chang, L. H., Zama, Y., Naka, T., Chorghade, M. S. & Hecht, S. M. (1984) *Biochemistry* **23,** 1468–1473.

21. Mendel, D., Cornish, V. W. & Schultz, P. G. (1995) *Annu. Rev. Biophys. Biomol. Struct.* **24,** 435–462.

22. Nowak, M. W., Kearney, P. C., Sampson, J. R., Saks, M. E., Labarca, C. G., Silverman, S. K., Zhong, W., Thorson, J., Abelson, J. N., Davidson, N., *et al* (1995) *Science* **268,** 439–442.

23. Liu, D. R., Magliery, T. J., Pastrnak, M. & Schultz, P. G. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 10092–10097.

24. Wang, L., Magliery, T. J., Liu, D. R. & Schultz, P. G. (2000) *J. Am. Chem. Soc.* **122,** 5010–5011.

25. Yarus, M. (1982) *Science* **18,** 646–652.

26. Steward, L. E., Collins, C. S., Gilmore, M. A., Carlson, J. E., Ross, J. B. A. & Chamberlin, A. R. (1997) *J. Am. Chem. Soc.* **119,** 6–11.

27. Short, G. F., Golovine, S. Y. & Hecht, S. M. (1999) *Biochemistry* **38,** 8808–8819.

28. Hanyu, N., Kuchino, Y. & Nishimura, S. (1986) *EMBO J.* **5,** 1307–1311.

29. Schüll, C. & Beier, H. (1994) *Nucleic Acids Res.* **22,** 1974–1980.

30. Saks, M. E., Sampson, J. R., Nowak, M. W., Kearney, P. C., Du, F., Abelson, J. N., Lester, H. A. & Dougherty, D. A. (1996) *J. Biol. Chem.* **271,** 23169–23175.

31. Lee, C. P. & RajBhandary, U. L. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 11378–11382.

32. Drabkin, H. J., Estrella, M. & RajBhandary, U. L. (1998) *Mol. Cell. Biol.* **18,** 1459–1466.

33. Liu, D. R. & Schultz, P. G. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 4780–4785.

34. Arnez, J. G. & Moras, D. (1997) *Trends Biochem. Sci.* **22,** 211–216.

35. Cusack, S. (1997) *Curr. Opin. Struct. Biol.* **7,** 881–889.

COMMENTARY