ACS chemical biology

Point of VIEW

Site-Specific Incorporation of Chemical Probes into Proteins for NMR

A Quamrul Hassan*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

ABSTRACT The ability to incorporate chemical probes into peptides is of great importance because it can render novel functionality to proteins and greatly expand our capacity to investigate complex biological systems. A methodology developed by the Schultz laboratory provides a unique strategy to incorporate chemical probes as unnatural amino acids into proteins by "expanding the genetic code" of the host cell. A recent application of this methodology that allows the site-specific incorporation of three NMR-active probes into proteins demonstrates the potential for researchers to explore avenues that are not easily achievable with existing methods.

*Corresponding author, aqhassan@mit.edu.

Published online September 19, 2008
10.1021/cb800215d CCC: $40.75

Although site-directed mutagenesis is a powerful method for studying the structure and function of proteins, it allows incorporation of only the 20 naturally occurring amino acids, a limitation researchers have sought to overcome. The Schultz laboratory developed a general methodology to incorporate unnatural amino acids (UAAs) into proteins site-specifically by exploiting the natural biosynthetic machinery of protein synthesis (*1*). In a recent paper, Cellitti *et al.* (*2*) report an exciting application of this methodology by site-specifically incorporating three NMR-active probes into fatty acid synthase (FAS). This allowed for the study of the enzyme's dynamics, structure, and ligand-binding properties by NMR. This article represents steady and exciting progress in the past few years toward the robust application of this methodology.

The strategy to incorporate UAAs is often based on introducing a nonsense codon, TAG, into a desired position of a plasmid, which allows incorporation with the help of a suppressor tRNA (tRNA) charged with UAAs. To ensure that the aminoacylated tRNA does not serve as a substrate for aminoacyl-tRNA synthetases (AARS) present in the *in vitro* transcription/translation system from *Escherichia coli* used to test the methodology, the authors used $tRNA^{Phe}_{CUA}$ from yeast as a template, which was chemically acylated with desired UAAs. This strategy allowed Noren *et al.* (*3*) to successfully incorporate six structurally and electronically dissimilar phenylalanine analogs site-specifically into β-lactamase. This study presented a groundbreaking demonstration of how to hijack the natural biosynthetic machinery to incorporate UAAs site-specifically into proteins. Although the methodology has subsequently been used to introduce nearly 80 UAAs (*4*), several challenges, including the requirement for the semisynthetic preparation of charged tRNA and difficulty with folding/expression of proteins *in vitro*, have limited its robust application.

Wang *et al.* (*5*) achieved the next monumental advancement in methodology by incorporating *O*-methyl-Tyr site-specifically into dihydrofolate reductase (DHFR) *in vivo* using *E. coli* as a host. Unlike the *in vitro* strategy, the charging of tRNA with UAA in this case was carried out by heterologous expression of an AARS inside the cell (Figure 1). The required orthogonal $tRNA^{Tyr}_{CUA}$/tyrosyl-tRNA synthetase (TyrRS) pair was derived from the archaeal organism *Methanococcus jannaschii* (*Mj*). The orthogonality of the pair was further optimized through random or structure-based mutagenesis to ensure that it did not cross-react with any of the endogenous tRNA/AARS pairs, amino acids, or sense codons present in *E. coli*. The amber codon was once again used to define the site of incorporation. Using the evolved $tRNA^{Tyr}_{CUA}$/*Mj* TyrRS pair, the authors found the fidelity of *O*-methyl-Tyr incorporation into DHFR to be >99%. The great potential of the methodology has subsequently been demonstrated by introducing >30 UAAs site-specifically into various proteins *in vivo* (*1, 6, 7*). The

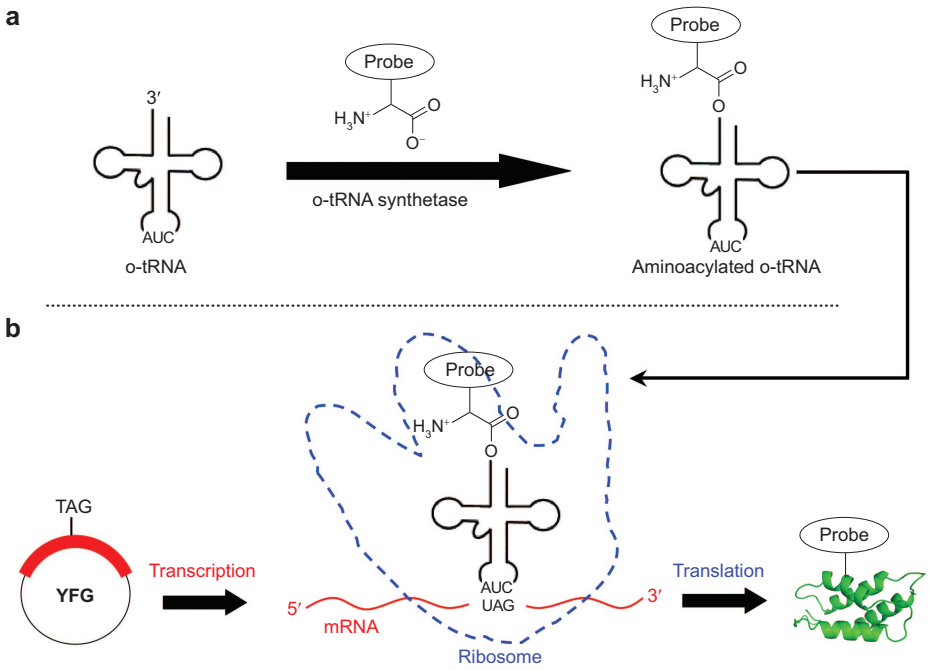

Figure 1. The strategy to site-specifically incorporate UAAs into protein using natural biosynthetic machinery. a) Orthogonal tRNA/tRNA synthetase pair at work: aminoacylation of orthogonal tRNA (o-tRNA) by orthogonal tRNA synthetase. b) Suppression of nonsense codon by aminoacylated o-tRNA.

Schultz laboratory then developed a general and powerful strategy based on directed evolution to generate an orthogonal tRNA/ARS pair for a UAA of interest (*1, 8*). Significant efforts have also been devoted to improving the yield of full-length product using multiple copies of tRNA and/or ARS in the expression system (*2, 9, 10*).

Researchers have sought to use the UAA incorporation methodology to tackle challenging issues. NMR spectroscopy is a useful method for studying the dynamics, structure, and function of protein in solution (*11*). Incorporation of amino acids isotopically labeled with ^{13}C, ^{15}N, or ^{19}F is often necessary in order to simplify complex NMR spectra of proteins, especially for large proteins. Because many methods to incorporate NMR active labels into proteins result in heterogeneous protein samples, incorporation of NMR-active labels by the UAA methodology is an attractive alternative. Earlier, Deiters *et al.* (*7*) used a previously developed tRNA/ARS pair (*5*) to introduce the NMR-active probe ^{15}N-*p*-methoxyphenylalanine (Figure 2, **1**) into myoglobin with modest yield (1 mg L^{-1}) but high fidelity (>99%). Mehl *et al.* (*12*) focused on incorporating an ^{19}F-labeled probe with high fidelity into nitroreductase and histidinol dehydrogenase. The study documented the development of an ARS for a trifluoromethyl-L-phenylalanine (Figure 2, **2**) probe and its successful site-specific incorporation into the proteins.

Now, Cellitti *et al.* have taken a further step and incorporated three NMR probes with three different isotopes (^{15}N, ^{13}C, and ^{19}F) site-specifically into FAS (Figure 2). This study demonstrates (i) the design and synthesis of the probes, (ii) the evolution of AARS for the probes, and (iii) the optimization of conditions to improve incorporation efficiency. Three NMR-active probes, $^{13}C/^{15}N$-*p*-methoxyphenylalanine (Figure 2, **3**), 2-amino-3-(4-(trifluoromethoxy)phenyl) propanoic acid (Figure 2, **4**), and ^{15}N-*o*-nitrobenzyl-tyrosine (Figure 2, **5**), were utilized in this study. Using a new expression system, the authors introduced all three probes site-specifically into each of 11 positions of FAS, with yields varying from 5 to 144 mg L^{-1}. These FAS variants, each labeled with a single probe, were used to study conformational changes in FAS upon ligand binding using $^1H-^{15}N$ HSQC, $^1H-^{13}C$ HSQC, and ^{19}F spectroscopy. The beauty of this paper is that it facilitates the study of large proteins by NMR

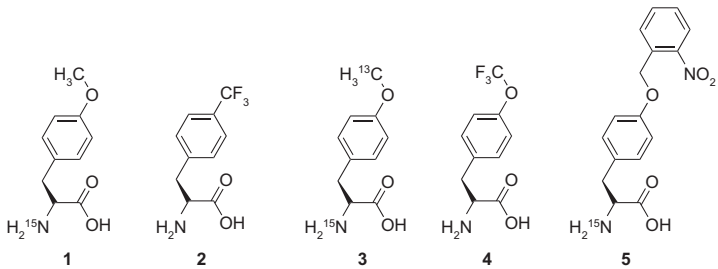

Figure 2. NMR-active probes site-specifically introduced into various proteins using UAA methodology.

methods, something otherwise not achievable. It also presents a significant advancement toward the robust application of this methodology in site-specific incorporation of chemical probes into proteins.

Acknowledgment: The author acknowledges helpful discussion with Ryan Mehl.

REFERENCES

1. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys. Biomol. Struct. 35*, 225–249.
2. Cellitti, S. E., Jones, D. H., Lagpacan, L., Hao, X., Zhang, Q., Hu, H., Brittain, S. M., Brinker, A., Caldwell, J., Bursulaya, B., Spraggon, G., Brock, A., Ryu, Y., Uno, T., Schultz, P. G., and Geierstanger, B. H. (2008) *In vivo* incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy, *J. Am. Chem. Soc. 130*, 9268–9281.
3. Noren, C. J., Anthonycahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins, *Science 244*, 182–188.
4. Cornish, V. W., Mendel, D., and Schultz, P. G. (1995) Probing protein structure and function with an expanded genetic code, *Angew. Chem., Int. Ed. 34*, 621–633.
5. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *E. coli*, *Science 292*, 498–500.
6. Seyedsayamdost, M. R., Xie, J., Chan, C. T. Y., Schultz, P. G., and Stubbe, J. (2007) Site-specific insertion of 3-aminotyrosine into subunit α2 of *E. coli* ribonucleotide reductase: direct evidence for involvement of Y-730 and Y-731 in radical propagation, *J. Am. Chem. Soc. 129*, 15060–15071.
7. Deiters, A., Geierstanger, B. H., and Schultz, P. G. (2005) Site-specific *in vivo* labeling of proteins for NMR studies, *ChemBioChem 6*, 55–58.
8. Cropp, T. A., and Schultz, P. G. (2004) An expanding genetic code, *Trends Genet. 20*, 625–630.
9. Ryu, Y. H., and Schultz, P. G. (2006) Efficient incorporation of unnatural amino acids into proteins in *E. coli*, *Nat. Methods 3*, 263–265.
10. Farrelli, I. S., Toroney, R., Hazen, J. L., Mehl, R. A., and Chin, J. W. (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone, *Nat. Methods 2*, 377–384.
11. Tugarinov, V., Hwang, P. M., and Kay, L. E. (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins, *Annu. Rev. Biochem. 73*, 107–146.
12. Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a ^{19}F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity, *J. Am. Chem. Soc. 129*, 1160–1166.