J. Am. Chem. Soc. **1999**, *121*, 12194−12195

Incorporation of Two Different Nonnatural Amino Acids Independently into a Single Protein through Extension of the Genetic Code

Takahiro Hohsaka, Yuki Ashizuka, Hiroshi Sasaki, Hiroshi Murakami, and Masahiko Sisido*

Department of Bioscience and Biotechnology Okayama University, 3-1-1 Tsushimanaka Okayama 700-8530, Japan

Received June 28, 1999

A method for introducing nonnatural amino acids into proteins has been developed[1,2] and employed to probe structure and function of proteins.[3−7] However, only a single nonnatural amino acid could have been introduced into a single protein, simply because no codon other than UAG amber stop codon has been available for the incorporation of nonnatural amino acid. An attempt to use UGA opal codon resulted in highly efficient readthrough.[8] The use of nonnatural codons containing an *iso*C-*iso*G base pair will be one of the solutions,[9] but this strategy requires chemically synthesized mRNA and tRNA.

As an alternative strategy to extend the genetic code, we have reported the incorporation of nonnatural amino acids through in vitro frameshift suppression. A four-base codon AGGU was translated into a nonnatural amino acid by chemically amino-acylated frameshift suppressor tRNA containing complementary four-base anticodon ACCU.[10] Subsequent studies have revealed that another four-base codon CGGG works more efficiently.[11,12] In both cases, aminoacylated frameshift tRNAs could not introduce nonnatural amino acids into proteins unless the corresponding four-base codons were present in the mRNA. This frameshift strategy overcomes the limitation of the 64 genetic codes and provides more than two extended codons that may be assigned to different nonnatural amino acids.

Here we show a successful incorporation of two different nonnatural amino acids into two different sites of a single protein by combining two four-base codons. Incorporation of two different nonnatural amino acids will allow the structure−function relationships to be probed more precisely and more complex processes on a protein framework such as site-to-site electron transfer[11] and energy transfer to be studied.

* To whom correspondence should be addressed.

(1) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182−188.
(2) Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R.; Diala, E. S. *J. Am. Chem. Soc.* **1989**, *111*, 8013−8014.
(3) Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 621−633.
(4) Thorson, J. S.; Cornish, V. W.; Barrett, J. E.; Cload, S. T.; Yano, T.; Schultz, P. G. *Methods Mol. Biol.* **1998**, *77*, 43−73.
(5) (a) Killian, J. A.; Van Cleve, M. D.; Shayo, Y. F.; Hecht, S. M. *J. Am. Chem. Soc.* **1998**, *120*, 3032−3042. (b) Arslan, T.; Mamaev, S. V.; Mamaeva, N. V.; Hecht, S. M. *J. Am. Chem. Soc.* **1997**, *119*, 10877−10887. (c) Karginov, V. A.; Mamaev, S. V.; An, H.; Van Cleve, M. D.; Hecht, S. M.; Komatsoulis, G. A.; Abelson, J. N. *J. Am. Chem. Soc.* **1997**, *119*, 8166−8176.
(6) (a) Kanamori, T.; Nishikawa, S.; Nakai, M.; Shin, I.; Schultz, P. G.; Endo, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3634−3639. (b) Kanamori, T.; Nishikawa, S.; Shin, I.; Schultz, P. G.; Endo, T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 485−490.
(7) Nowak, M. W.; Gallivan, J. P.; Silverman, S. K.; Labarca, C. G.; Dougherty, D. A.; Lester, H. A. *Methods Enzymol.* **1998**, *293*, 504−529.
(8) Cload, S. T.; Liu, D. R.; Froland, W. A.; Schultz, P. G. *Chem. Biol.* **1996**, *3*, 1033−1038.
(9) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537−539.
(10) Hohsaka, T.; Ashizuka, Y.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc.* **1996**, *118*, 9778−9779.
(11) Murakami, H.; Hohsaka, T.; Ashizuka, Y.; Sisido, M. *J. Am. Chem. Soc.* **1998**, *120*, 7520−7529.
(12) Hohsaka, T.; Kajihara, D.; Ashizuka, Y.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc.* **1999**, *121*, 34−40.
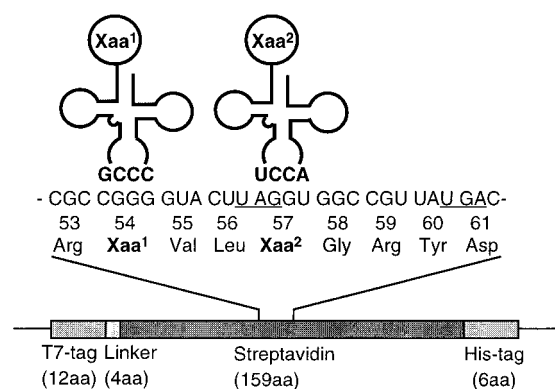

Figure 1. Nucleotide and amino acid sequence of the mutated streptavidin. Underlines indicate stop codons which appear when one of the four-base codons is decoded as a triplet by the endogenous arginyl tRNA.

As shown in Figure 1, we designed a mRNA for evaluation of double incorporation of nonnatural amino acids. Two four-base codons CGGG and AGGU were introduced into Tyr54 and Thr57 sites of streptavidin, respectively. If CGGG is decoded by aminoacyl tRNA$_{CCCG}$ and AGGU is decoded by aminoacyl tRNA$_{ACCU}$, a full-length streptavidin containing two nonnatural amino acids (**Xaa^1** and **Xaa^2**) will be produced. If the CGGG codon was undesirably read as a CGG triplet by the endogenous arginyl tRNA$_{CCG}$, the reading frame remained unshifted, resulting in an encounter of the UAG stop codon that is underlined. Similarly, if the AGGU codon was read as an AGG triplet by the arginyl tRNA$_{CCU}$, the protein synthesis will stop at the UGA stop codon underlined. As a result, the full-length streptavidin successfully produced must contain two nonnatural amino acids at directed sites.

As nonnatural amino acids, ϵ-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-L-lysine [Lys(NBD), **1**] and L-2-naphthylalanine (2-napAla, **2**) were chosen. Lys(NBD) was used for fluorescence detection
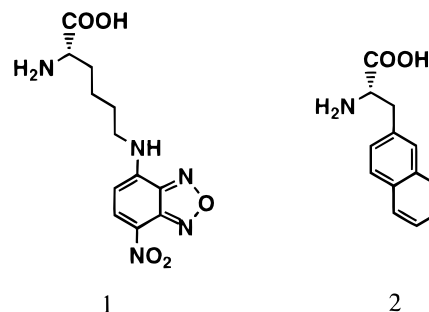
of a hexapeptide obtained by tryptic digestion at Arg53 and Arg59 of the full-length streptavidin. 2-Naphthylalanine was employed because it had been incorporated in excellent yield by using CGGG codon.[12] The tRNAs containing the four-base anticodons and charged with respective nonnatural amino acids were prepared as previously described.[12] The mutated streptavidin gene containing the CGGG and AGGT codons at Tyr54 and Thr57 sites, respectively, was generated by the PCR method and the mRNA was prepared by T7 RNA polymerase. Then, the tRNAs and mRNA were added into the *E. coli* in vitro protein synthesizing system.[12] The translation products were analyzed by Western blotting using anti-T7tag antibody and alkaliphosphatase-labeled anti-mouse IgG.[12]

As shown in Figure 2, in the absence of the two tRNAs or in the absence of either aminoacyl tRNA$_{CCCG}$ or aminoacyl tRNA$_{ACCU}$, no full-length streptavidin was synthesized, but
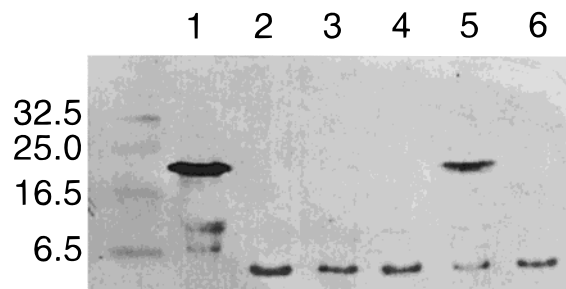
Figure 2. Western blot analysis of the translation of wild-type streptavidin mRNA (lane 1) and mutated mRNA containing CGGG and AGGU at Tyr54 and Thr57 sites, respectively, with no tRNA (lane 2), with Lys-(NBD)-tRNA$_{CCCG}$ (lane 3), with 2-napAla-tRNA$_{ACCU}$ (lane 4); with both aminoacylated tRNAs (lane 5), and with both non-aminoacylated tRNAs (lane 6).
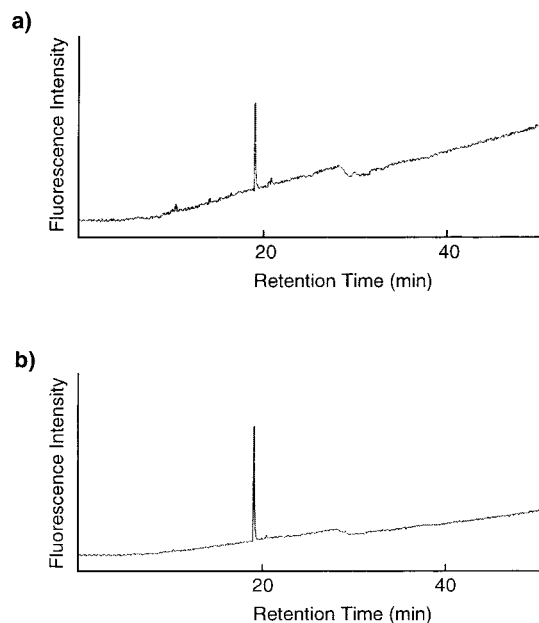

Figure 3. HPLC analyses of (a) trypsin-digested fragment of the full-length streptavidin synthesized in the presence of Lys(NBD)-tRNA$_{CCCG}$ and 2-napAla-tRNA$_{ACCU}$ followed by purification on a Ni-NTA column and (b) chemically synthesized Lys(NBD)-Val-Leu-2-napAla-Gly-Arg (1 pmol). $\lambda_{ex} = 470$ nm and $\lambda_{em} = 530$ nm.

truncated peptides were detected instead (lane 2−4). The latter truncated product may be assigned to a peptide terminated by the stop codon at the 56−57 or at the 60−61 sites. The two peptides could not be resolved by the electrophoresis. The results of Western blotting indicate that spontaneous frameshift occurs at neither CGGG nor AGGU site, and that the CGGG codon is not recognized by ACCU anticodon and the AGGU codon is not recognized by CCCG anticodon. Only in the presence of both aminoacyl tRNA$_{CCCG}$ and aminoacyl tRNA$_{ACCU}$ was the full-length streptavidin successfully synthesized (lane 5). This is interpreted in terms of the translation of the CGGG and AGGU codons into Lys(NBD) and 2-napAla by the aminoacyl tRNA$_{CCCG}$ and aminoacyl tRNA$_{ACCU}$, respectively. In the presence of two non-aminoacylated tRNAs, no full-length streptavidin was synthesized (lane 6), supporting the idea that these tRNA could not be aminoacylated by any aminoacyl tRNA synthetase as clarified earlier.[10,12] Densitometric analysis of Western blotting showed that the yield of the mutant streptavidin was estimated at 9% with respective to wild-type streptavidin. This value agrees with the product of the relative yields of the mutant streptavidin produced by Tyr83CGGG mutated mRNA and Lys(NBD)-tRNA$_{CCCG}$ (41%) and that produced by Tyr83AGGU mRNA and 2-napAla-tRNA$_{ACCU}$ (22%), supporting the idea that each incorporation occurs independently. It should be noted that the incorporation efficiency of nonnatural amino acids corresponding to AGGU codon was lower than that corresponding to CGGG codon.

The double incorporation of these nonnatural amino acids was directly confirmed by tryptic digestion followed by HPLC analysis.[13] The full-length streptavidin was purified on a Ni-NTA affinity column, digested by trypsin, and then applied to reverse-phase HPLC equipped with a fluorescence detector. As shown in Figure 3, the trypsin-digested fluorescent fragment (Figure 3a) was identical with a chemically synthesized Lys(NBD)-Val-Leu-2-napAla-Gly-Arg[14] (Figure 3b). This result demonstrates that Lys(NBD) and 2-napAla are specifically introduced into Tyr54 and Thr57 sites of streptavidin, respectively.

We introduced other nonnatural amino acid pairs into other two different sites of streptavidin. For example, incorporation of 2-pyrenylalanine and nitrophenylalanine into Tyr54 and Tyr83 sites, respectively, was accomplished by using a mRNA containing CGGG and AGGU codons at the Tyr54 and Tyr83 sites and tRNA$_{CCCG}$ and tRNA$_{ACCU}$ aminoacylated with 2-pyrenylalanine and nitrophenylalanine, respectively. The resulting streptavidin retained biotin binding activity. Such mutant will serve to study electron transfers between two specific sites in protein frameworks.

In conclusion, a protein containing two different nonnatural amino acids at two directed sites was synthesized for the first time by using the *E. coli* in vitro protein synthesizing system. This novel technique will expand the scope of the nonnatural amino acid mutagenesis.

Supporting Information Available: Experimental detail for Lys-(NBD)-pdCpA (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA992204P

(13) Translation mixture (20 µL) was diluted with 100 µL of 50 mM Tris−HCl, pH 8, containing 8 M urea, and placed on a Ni-NTA affinity column (QIAGEN, 20 µL of suspended gel solution). After washing seven times with 200 µL of 50 mM Tris−HCl, pH 8.0, containing 8 M urea, the full-length streptavidin was eluted with 20 µL of 30 mM phosphate buffer, pH 5.0, containing 8 M urea. The eluate (5 µL) was diluted with 35 µL of 50 mM Tris−HCl (pH 7.5) and 1 mM CaCl$_2$, then incubated with 0.1 µg of trypsin (Promega) at 37 °C for 40 min. Reverse-phase HPLC analysis was done with an ODS column (Waters), a linear gradient from 0.1% TFA to acetonitrile in 50 min, and a flow rate of 1 mL/min. The eluate was monitored by fluorescence detector ($\lambda_{ex} = 70$ nm, $\lambda_{em} = 530$ nm).

(14) Chemical synthesis of Lys(NBD)-Val-Leu-2-napAla-Gly-Arg was carried out as follows. NBD-Cl (0.4 mg) and α-Fmoc-Lys-Val-Leu-2-napAla-Gly-Arg (0.4 mg) synthesized on a peptide synthesizer (Pioneer, PerSeptive Biosystems) were added to a mixture of 50 µL of methanol and 10 µL of 4% aqueous NaHCO$_3$. After incubation at 55 °C for 1 h, Fmoc-Lys(NBD)-Val-Leu-2-napAla-Gly-Arg was isolated by reverse-phase HPLC. The removal of the Fmoc group was done by incubating in 20% piperidine/DMF at 37 °C for 2 h, then the desired product was purified on the reverse-phase HPLC. HRMS (FAB) calcd for C$_{44}$H$_{62}$N$_{13}$O$_{10}$ (MH$^+$) 932.4743, found 932.4781.