Efficient Incorporation of Nonnatural Amino Acids with Large Aromatic Groups into Streptavidin in In Vitro Protein Synthesizing Systems

Takahiro Hohsaka, Daisuke Kajihara, Yuki Ashizuka, Hiroshi Murakami, and Masahiko Sisido*

Contribution from the Department of Bioscience and Biotechnology, Faculty of Engineering, Okayama University, 3-1-1 Tsushimanaka, Okayama 700-8530, Japan

Received April 17, 1998

Abstract: Efficiencies of the incorporation of various nonnatural amino acids carrying aromatic side groups into streptavidin were examined. The aromatic amino acids were linked to a mixed dinucleotide, pdCpA, and the resulting aminoacyl pdCpAs were coupled with tRNACCCG(−CA) to afford chemically aminoacylated tRNACCCG's. Mutant streptavidin mRNA containing a CGGG 4 base codon at the Tyr83 site was prepared and added to an *Escherichia coli* in vitro translation system with the aminoacyl tRNACCCG. The expression of the full-length mutant streptavidins was confirmed by a Western blot analysis, and their biotin binding activity was examined by a dot blot analysis. The Western blot analysis indicated that the efficiencies of the incorporation were higher for aromatic groups with straight configurations than those with widely expanded or bend configurations. The incorporation efficiencies were also examined in a rabbit reticulocyte lysate. In the latter system, the efficiencies were markedly improved for nonnatural amino acids with large side groups such as pyrene and anthraquinone.

Site-specific incorporation of nonnatural amino acids into proteins through in vitro protein biosynthesizing systems is becoming an important technique for structure−function analysis and for artificial functionalization of proteins.[1] By incorporating nonnatural amino acids into proteins, we can introduce a variety of specialty functions depending on the side groups of the amino acids. Particularly, nonnatural amino acids carrying various aromatic groups may serve as fluorescent probes for the analysis of microenvironment of proteins and as electron donors and acceptors for building up pathways for electron transfers in proteins.[2] For this technique to be used widely in the field of protein engineering, however, it is essential to know what type of nonnatural amino acids will be successfully incorporated and what type will be rejected in the *Escherichia coli* and other protein biosynthesizing systems. The nonnatural amino acids that have been examined so far were restricted to those carrying rather small side groups, and the incorporation efficiencies of the nonnatural amino acids with large aromatic groups have not been studied extensively. From a few examples reported so far, however, it is suggested that the efficiency may depend sharply on the side groups. For instance, *p*-benzoylphenylalanine has been successfully incorporated into T4 Lysozyme[3] and cytochrome *b*2[4] in good yield, whereas the incorporation of the amino acid carrying oxyltetramethylpyrroline was unsuccessful.[3] In this study we have explored a relationship between the structure and incorporation efficiency of nonnatural amino acids carrying 19 different side groups in the *E. coli* and rabbit reticulocyte in vitro systems.

Since the amino acid selectivity is mainly governed by the amino acid adaptability of the ribosomal systems[5] and the adaptability depends on the structure of ribosomes, the different ribosomal systems are expected to show different amino acid selectivity. And the results will be useful for designing proteins incorporated with nonnatural amino acids carrying large side groups with specialty functions.

Another important step for the incorporation of nonnatural amino acids is the assignments of codons to the amino acids. Most of the workers have been using nonsense codons for this purpose. But the number of nonsense codons that can be assigned to nonnatural amino acids is limited to be less than two. We have proposed in the previous paper that a four-base codon−anticodon pair, AGGU-ACCU, worked effectively for this purpose in the *E. coli* in vitro system.[6] The frameshift suppressor tRNAACCU's chemically aminoacylated with non-natural amino acids carrying *p*-nitrophenyl, 1- and 2-naphthyl, 2-anthryl, and *p*-phenylazophenyl groups, respectively, success-

* Corresponding author. Fax: +81-86-251-8219. E-mail: sisido@cc.okayama-u.ac.jp.

(1) (a) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182−188. (b) Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R.; Diala, E. S. *J. Am. Chem. Soc.* **1989**, *111*, 8013−8014. For recent reports, see: (c) Mamaev, S. V.; Laikhter, A. L.; Arslan, T.; Hecht, S. M. *J. Am. Chem. Soc.* **1996**, *118*, 7243−7244. (d) 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. (e) Steward, L. E.; Collins, C. S.; Gilmore, M. A.; Carlson, J. E.; Ross, J. B. A.; Chamberlin, A. R. *J. Am. Chem. Soc.* **1997**, *119*, 6−11. (f) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314−11322. (g) England, P. M.; Lester, H. A.; Davidson, N.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11025−11030. (h) van Hest J. C.; Tirrell D. A. *FEBS Lett.* **1998**, *428*, 68−70.

(2) (a) Sisido, M. *Prog. Polym. Sci.* **1992**, *17*, 699−764. (b) Sisido, M. *Adv. Photochem.* **1997**, *22*, 197−228.

(3) Cornish, V. W.; Benson, D. R.; Altenbach, C. A.; Hideg, K.; Hubbell, W. L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2910−2914.

(4) Kanamori, T.; Nishikawa, S.; Shin, I.; Schultz, P. G.; Endo, T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 485−490.

(5) Hohsaka, T.; Sato, K.; Sisido, M.; Takai, K.; Yokoyama, S. *FEBS Lett.* **1993**, *335*, 47−50.

(6) Hohsaka, T.; Ashizuka, Y.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc.* **1996**, *118*, 9778−9779.

10.1021/ja9813109 CCC: $18.00 © 1999 American Chemical Society
Published on Web 12/19/1998

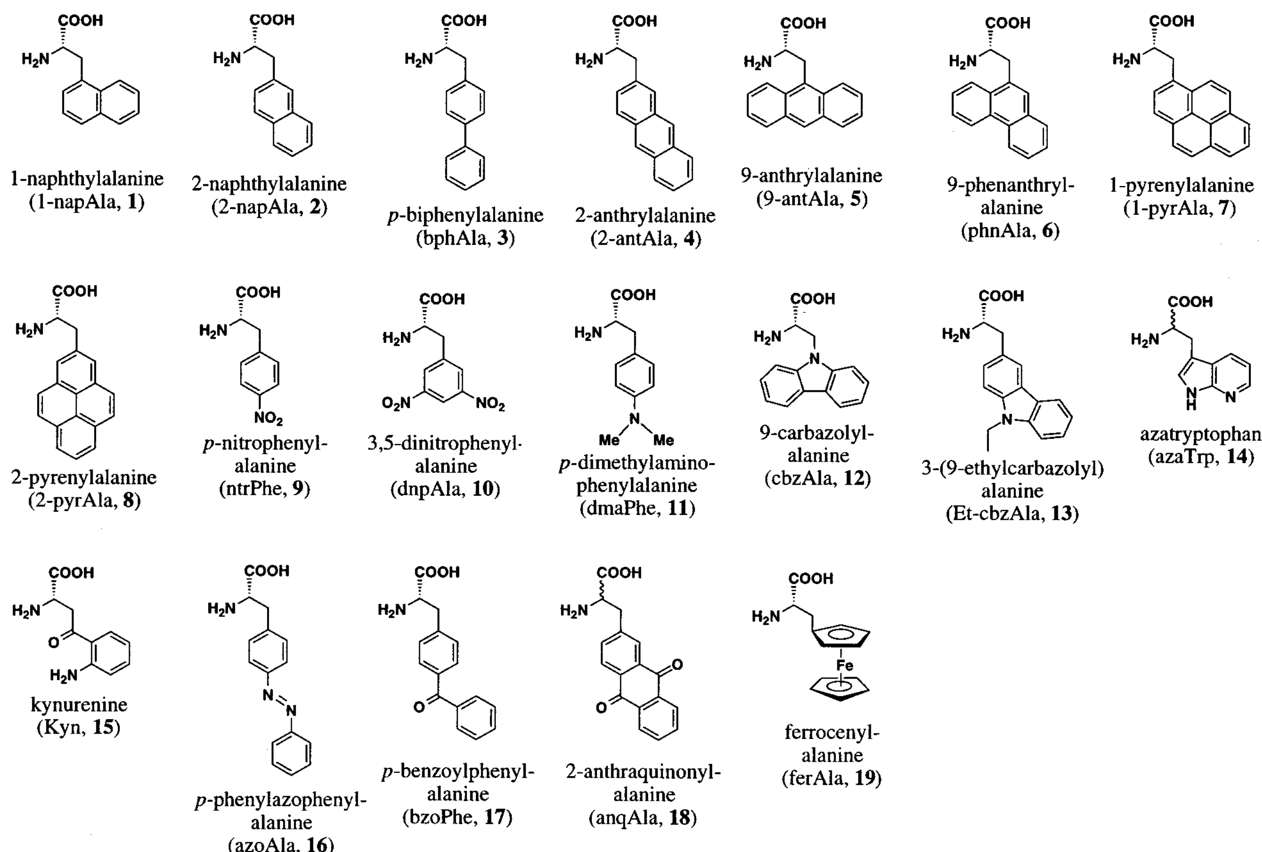

Figure 1. Structure of aromatic nonnatural amino acids examined in this study.

fully suppressed the frameshift mutation in a mRNA and produced mutant proteins containing the nonnatural amino acids. In the case of nitrophenylalanine, the suppression efficiency was about 20%.[6] The frameshift strategy using the four-base codon−anticodon pairs will be much more advantageous than the nonsense codon strategy because of the possible extension to incorporate two or more different nonnatural amino acids into single proteins. In this study, we have used another four-base codon−anticodon pair, CGGG-CCCG, because of its higher efficiency than the AGGU-ACCU pair.

Streptavidin was selected as a target protein, since its structure and function have been characterized extensively, and its activity can be easily evaluated by using biotinylated enzymes or biotinylated fluorophores. In this paper, we evaluated incorporation efficiencies of nonnatural amino acids with various aromatic side groups as listed in Figure 1. The aromatic amino acids were of the L-form except for DL-anthraquinonylalanine and DL-azatryptophan. They were incorporated at the Tyr83 site of streptavidin by using a CGGG-CCCG codon−anticodon pair. The incorporation efficiencies were evaluated from the Western blotting of the reaction mixture of the in vitro system.

Materials and Methods

L-*p*-Nitrophenylalanine, L-1-naphthylalanine, L-2-naphthylalanine, DL-azatryptophan, and L-kynurenine were purchased from Sigma (St. Louis, MO). Boc-L-*p*-benzoylphenylalanine was from Bachem (Switzerland). L-*p*-Biphenylalanine,[7a] L-9-anthrylalanine,[7b] L-9-phenanthrylalanine,[7c] L-1-pyrenylalanine,[7d] L-2-pyrenylalanine, and L-3-(9-ethylcarbazolyl)-alanine[7e] were synthesized from the corresponding arylaldehydes and acetylglycine. L-2-Anthrylalanine, L-3,5-dinitrophenylalanine, DL-2-anthraquinonylalanine,[7f] and L-ferrocenylalanine[7g] were synthesized from the corresponding arylmethyl halides and diethyl acetamidomalonate. L-9-Carbazolylalanine[7h] was synthesized from carbazole and 2,3-dibromopropanoic acid. L-*p*-Dimethylaminophenylalanine[7i] and L-*p*-phenylazophenylalanine[7j] were synthesized from L-*p*-nitrophenylalanine. Of the above 19 amino acids, L-2-anthrylalanine, L-2-pyrenylalanine, and L-3,5-dinitrophenylalanine were newly synthesized in this work. The details of the syntheses of the first two are described in the Supporting Information. DL-3,5-Dinitrophenylalanine was synthesized from 3,5-dinitrobenzyl chloride and diethyl acetamidomalonate and selectively deacetylated with acylase. The detail will be published elsewhere.

Vent DNA Polymerase, T7 RNA polymerase, and Prestained Protein Marker were purchased from New England Biolabs (Beverly, MA). T4 RNA ligase was from Takara Shuzo (Kyoto, Japan). *E. coli* S-30 extract, Rabbit Reticulocyte Lysate Systems, and ProtoBlot II AP System for the Western blot analysis were from Promega (Madison, WI). Plasmid encoding streptavidin was from R&D Systems Europe (Abingdon, U.K.). Immun-Blot PVDF Membrane was from Bio-Rad (Hercules, CA). RNase Inhibitor was from Wako Chemicals (Osaka, Japan). T7-Tag Antibody was from Novagen (Madison, WI). Biotinylated alkaliphosphatase was from Zymed Laboratories (San Francisco, CA). Nitrocellulose membrane was from Toyo Roshi (Tokyo, Japan). Other biochemicals were from Sigma.

Synthesis of Aminoacyl tRNA. The syntheses of aminoacyl pdCpAs and tRNA(-CA)s are described in the Supporting Information. The ligation reaction was carried out in a mixture that contained 1 nmol of tRNA(-CA), 20 nmol of aminoacyl pdCpA in DMSO, 1 mM ATP, 15

(7) (a) Kuragaki, M.; Sisido, M. *J. Phys. Chem.* **1996**, *100*, 16019−16025. (b) Egusa, S.; Sisido, M.; Imanishi, Y. *Bull. Chem. Soc. Jpn.* **1986**, *59*, 3175−3178. Sisido, M. *Macromolecules* **1989**, *22*, 4367−4372. (c) Sasaki, H.; Sisido, M.; Imanishi, Y. *Langmuir* **1991**, *7*, 1949−1952. (d) Egusa, S.; Sisido, M.; Imanishi, Y. *Chem. Lett.* **1983**, 1307−1310. Egusa, S.; Sisido, M.; Imanishi, Y. *Macromolecules* **1985**, *18*, 882. (e) Taku, K.; Sasaki, H.; Kimura, S.; Imanishi, Y. *Amino Acids* **1994**, *7*, 311−316. (f) Matsubara, T.; Shinohara, H.; Sisido, M. *Macromolecules* **1997**, *30*, 2651−2656. (g) Kira, M.; Sisido, M. *Chem. Lett.* **1997**, 89−90. (h) L-9-Carbazolylalanine was synthesized accroding to the private communication from Professor Norikazu Nishino of Kyushu Institute of Technology. (i) Sisido, M.; Tanaka, R.; Inai, Y.; Imanishi, Y. *J. Am. Chem. Soc.* **1989**, *111*, 6790−6796. (j) Goodman, M.; Kossoy, A. *J. Am. Chem. Soc.* **1966**, *88*, 5010−5015.

mM $MgCl_2$, 3.3 mM DTT, 20 μg/mL of BSA, and 60 units of T4 RNA ligase in a 40 μL of 55 mM Hepes−Na (pH 7.5). The mixture was incubated at 4 °C for 2 h, then diluted with 1 vol of prechilled 0.6 M potassium acetate (pH 4.5). The solution was extracted with phenol/chloroform and chloroform, then the aminoacyl tRNA was precipitated with 3 vol of ethanol. The pellet was dissolved in 4 μL of prechilled 1 mM potassium acetate (pH 4.5), and immediately the solution was added to the reaction mixture of the in vitro translation. The purity of the tRNA was confirmed by a 10% PAGE with 7 M urea.

In Vitro Protein Biosynthesis of Mutant Streptavidins and Evaluation of the Incorporation Efficiency. The preparation of mRNA was described in the Supporting Information. In vitro translation was carried out in a 10 μL of a reaction mixture containing 55 mM Hepes−KOH (pH 7.5), 210 mM potassium glutamate, 6.9 mM ammonium acetate, 9 mM magnesium acetate, 1.7 mM DTT, 1.2 mM ATP, 0.28 mM GTP, 26 mM phosphoenolpyruvate, 1 mM spermidine, 1.9% poly(ethylene glycol)-8000, 35 μg/mL folinic acid, 0.1 mM each of amino acids except arginine, 0.01 mM arginine, 16 μg of mRNA, 0.1 nmol of aminoacyl tRNA, and 2 μL of *E. coli* S30 extract. The mixture was incubated at 37 °C for 60 min.

Each of 1 μL of the reaction mixture was mixed with 19 μL of 50 mM Tris-HCl (pH 6.8), 4% SDS, 2% 2-mercaptoethanol, 12% glycerol, and 0.01% bromophenol blue. The resulting solution was incubated at 95 °C for 5 min, then 5 μL of the solution was applied to a 15% SDS-polyacrylamide gel electrophoresis. After electroblotting to a PVDF membrane, the membrane was incubated at 37 °C for 30 min with 1% BSA in TBST (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.1% Tween20), then with 1/10 000 diluted T7-Tag Antibody in TBST. After washing with TBST for 5 min three times, the membrane was incubated with 1/5000 diluted alkaliphosphatase-labeled anti-mouse IgG in TBST. The membrane was washed three times with TBST for 5 min, once with TBS (20 mM Tris-HCl (pH 7.5) and 150 mM NaCl), then soaked in NBT/BCIP solution at 37 °C for 30 min.

For quantitative evaluation of the incorporation efficiency, the reaction mixture of wild-type streptavidin was serially diluted with the translation mixture without mRNA, and the resulting 12 samples were applied to the SDS-PAGE (16 × 16 cm, 22 lanes) together with the mutant streptavidins. The band intensity of the Western blot was evaluated by using the public domain NIH Image program (developed at the U. S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/) on Macintosh. Incorporation efficiency of each nonnatural amino acid was expressed as a value obtained from the calibration curve. The eight separete Western blot analyses were made, and the efficiencies were determined by at least four assays.

In a rabbit reticulocyte system, the reaction mixture contained 7 μL of nuclease-treated rabbit reticulocyte lysate, 1.6 μg of mRNA, and 0.1 mM each of amino acids except arginine in 10 μL. The mixture was incubated at 30 °C for 60 min. A 10-μL volume of the reaction mixture was mixed with 10 μL of 100 mM Tris-HCl (pH 6.8), 8% SDS, 4% 2-mercaptoethanol, 24% glycerol, and 0.02% bromophenol blue, then 4 μL of the solution was applied to a 15% SDS-PAGE, followed by the Western blotting.

Biotin Binding Assay. The binding activity of the mutant streptavidin was evaluated by a dot blot analysis using biotinylated alkaliphosphatase. The concentrations of the mutant streptavidins in the *E. coli* in vitro translation mixture were obtained from the incorporation efficiencies determined by the Western blot, and 5 ng of each streptavidin was spotted onto a nitrocellulose membrane by using a microfiltration blotting apparatus. After washing with TBS twice for 5 min, the membrane was incubated at 37 °C with 3% gelatin/TBS for 30 min and then with 1/500 diluted biotinylated alkaliphosphatase in 1% gelatin/TBS for 30 min. After washing for 5 min with TBST three times and once with TBS, the membrane was soaked in NBT/BCIP solution at 37 °C for 30 min, then washed with water and dried.

Fluorescence polarization was measured on a BEACON 2000 system (PanVera Corp., Madison, WI) equipped with the filters of 490 nm for excitation and 520 nm for emission. One microliter portion of 1/10 diluted wild-type streptavidin in the *E. coli* in vitro translation mixture

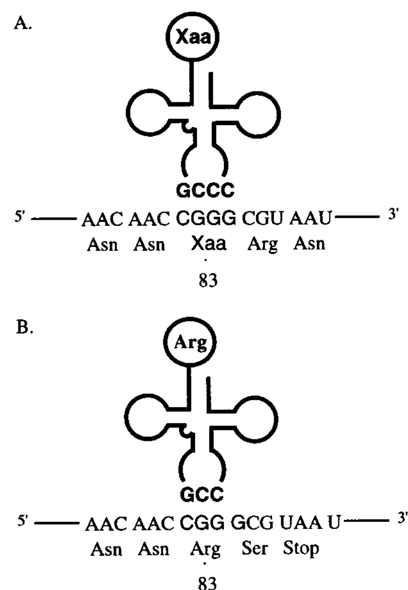

Figure 2. Nucleotide and amino acid sequence of the mutated region of streptavidin. (A) In the presence of aminoacyl tRNACCCG, the CGGG four-base codon is read by the tRNA. (B) In the absence of the tRNACCCG, the codon is read as the CGG three-base codon and the protein synthesis stops at the UAA stop codon.

was directly added into 100 μL of 0.5 nM fluorescein biotin in TBS. The solution was incubated for 5 min at 25 °C before each measurement.

Results and Discussion

Expression of Mutant Streptavidins. The synthetic streptavidin gene was inserted to T7 tag sequence under the control of T7 promoter. The N-terminal T7 tag is introduced for efficient translation in the *E. coli* system and for easy detection by anti T7 tag monoclonal antibody. At the C-terminal, histidine hexamer was attached for the purification of the protein. Since only the full-length protein that carries the terminal histidine hexamer binds to the Ni−NTA column, the truncated peptide that failed to incorporate the nonnatural amino acid can be removed by this procedure.

The mutation was introduced at the Tyr83 site. The resulting sequence is shown in Figure 2. In the presence of the frameshift suppressor tRNACCCG, CGGG will be recognized as a four-base codon and translated to the nonnatural amino acid, then the downstream sequence is translated correctly. On the other hand, in the absence of the tRNACCCG, CGGG is recognized as a three-base codon by Arg-tRNACCG and the next GCG is translated to Ala, then a UAA stop codon appears. As the result, only and every full-length protein contains the nonnatural amino acid, provided that the tRNACCCG is not aminoacylated enzymatically with other amino acids during the in vitro protein biosynthesis. One of the advantages of the frameshift strategy is that serious competition with the releasing factor that is encountered in the amber suppression strategy can be avoided.

The coding region including T7 promoter was amplified by a PCR reaction, and the mRNA was prepared by T7 RNA polymerase. The product was analyzed by a denaturing poly-

(8) (a) Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. *J. Biol. Chem.* **1978**, *253*, 4517−4520. (b) Heckler, T. G.; Zama, Y.; Naka, T.; Hecht, S. M. *J. Biol. Chem.* **1983**, *258*, 4492−4495. (c) Heckler, T. G.; Chang, L.-H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1984**, *23*, 1468−1473. (d) Heckler, T. G.; Chang, L.-H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M. *Tetrahedron* **1984**, *40*, 87−94. (e) Lodder, M.; Golovine, S.; Hecht, S. M. *J. Org. Chem.* **1997**, *62*, 778−779.

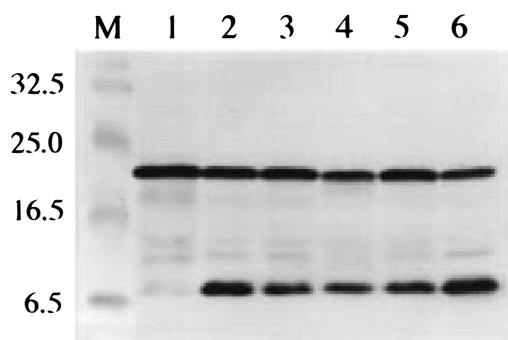

Figure 3. Western blot of the reaction mixture of the *E. coli* in vitro translation. The band at 19 kDa is a streptavidin monomer: lane 1, wild-type mRNA; lane 2, mRNA containing CGGU + nitrophenylalanyl tRNAACCG; lane 3, mRNA containing CGGG + nitrophenylalanyl tRNACCCG; lane 4, mRNA containing CGGC + nitrophenylalanyl tRNAGCCG; lane 5, mRNA containing CGGA + nitrophenylalanyl tRNAUCCG; lane 6, mRNA containing AGGU + nitrophenylalanyl tRNAACCU. Lane M contained prestained molecular weight marker.

acrylamide gel electrophoresis, which showed two bands corresponding to a mRNA terminated at the T7 terminator and that terminated at the 3′ end of the template. The latter will contain the same sequence as the former except for the extra 78 bases after the T7 terminator.

The preparation of aminoacyl tRNA was accomplished by the chemical misacylation method originally developed by Hecht and co-workers.[8] At first, the template of the tRNAACCU(-CA) under T7 promoter was synthesized from two oligonucleotides and then cloned into pUC18. The template of tRNANCCG(-CA) was generated from that of tRNAACCU(-CA) by site-directed mutagenesis. For transcription, the coding region of the plasmid was amplified by the PCR reaction using M13 primer and 3′ terminal primer that defines the 3′ terminal of the tRNA to be lacking a CA dinucleotide. The T7 transcription reaction yielded about 3 mg of the tRNA(-CA) in 1 mL of the reaction mixture after purification with ion-exchange chromatography.

Aminoacyl tRNA was obtained by the coupling of the tRNA-(-CA) with aminoacyl pdCpA catalyzed by T4 RNA ligase. Denaturing polyacrylamide gel electrophoresis showed that about a half of the tRNA(-CA) was linked with the aminoacyl pdCpA irrespective of the type of amino acid. Degradation of aminoacyl pdCpA under the condition of the ligation was followed by HPLC analysis for four typical amino acids, i.e., DL-anthraquinonylalanine, L-ferrocenylalanine, L-2-naphthylalanine, and L-*p*-nitrophenylalanine. The degradation took place, but the rate was not dependent on the type of the amino acids; about 20% after 2 h and 50% after 10 h. These results suggest that the aminoacyl tRNAs charged with different amino acids have been produced by about the same amount and they show similar lifetimes in aqueous media.

Protein synthesis was carried out in *E. coli* S30 in vitro translation system. Amino acids excluding arginine were added at the concentration of 0.1 mM. The concentration of arginine was 0.01 mM. The amount of aminoacyl tRNA was 0.1 nmol per 10 μL of a reaction mixture for all amino acids examined. The reaction mixture was incubated at 37 °C for 60 min, then the mixture was applied to SDS 15% polyacrylamide gel electrophoresis, followed by the transfer to a PVDF membrane. Streptavidins were detected by using anti T7-tag monoclonal antibody and alkaliphosphatase-labeled anti-Mouse IgG. The amount of wild-type streptavidin was estimated from the measurement of fluorescence polarization. The reaction mixture of the in vitro translation was added to a solution of biotinylated fluorescein (5×10^{-10} M). The amount of streptavidin was determined from a turning point of the fluorescence polarization. It was 10 μg per 1 mL of the reaction mixture.

Figure 3 shows the results of the Western blotting of the frameshift suppression of AGGU and CGGN four-base codons inserted at the Tyr83 site by the *p*-nitrophenyl tRNAACCU and tRNANCCG, respectively. The slower migrating band at 19 kDa corresponds to a full-length streptavidin, and the faster migrating band corresponds to a truncated peptide of 9 kDa that is formed when the three of four nucleotide is read by the internal Arg-tRNA. The frameshift suppression of the CGGN codons gave higher yields of the mutant streptavidin than the suppression of the AGGU codon. Since, the CGGG codon gave the highest yield in the CGGN codons, the following experiments were carried out by using the CGGG four-base codon.

Incorporation Efficiencies of Various Nonnatural Amino Acids in the *E. coli* In Vitro System. Figure 4 shows results of the Western blotting of the expressions of the mutant streptavidins in the presence of the aminoacyl tRNAs carrying various nonnatural amino acids. In the absence of the tRNA and in the presence of nonaminoacylated tRNA, a negligible amount of full-length streptavidin was synthesized and truncated peptide of 9 kDa was formed, instead. These results show that the in vitro translation system cannot suppress the frameshift mutation unless an aminoacylated tRNA carrying the four-base anticodon is present and that the tRNACCCG cannot be recognized by any aminoacyl tRNA synthetases.

The Western blotting indicates that the incorporation efficiency markedly depends on the structure of amino acids. On one hand, some aromatic nonnatural amino acids, including those with large side groups such as 2-anthrylalanine and

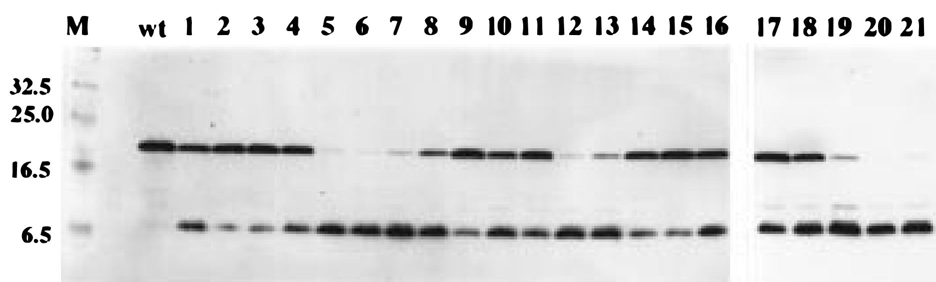

Figure 4. Western blot of the reaction mixture of the *E. coli* in vitro translation containing tRNACCCG's charged with various nonnatural amino acids: lane wt, wild-type mRNA; lane 1, 1-naphthylalanine; lane 2, 2-naphthylalanine; lane 3, *p*-biphenylalanine; lane 4, 2-anthrylalanine; lane 5, 9-anthrylalanine; lane 6, 9-phenanthrylalanine; lane 7, 1-pyrenylalanine; lane 8, 2-pyrenylalanine; lane 9, *p*-nitrophenylalanine; lane 10, 3,5-dinitrophenylalanine; lane 11, *p*-dimethylaminophenylalanine; lane 12, 9-carbazolylalanine; lane 13, 3-(9-ethylcarbazolyl)alanine; lane 14, azatryptophan; lane 15, kynurenine; lane 16, *p*-phenylazophenylalanine; lane 17, *p*-benzoylphenylalanine; lane 18, 2-anthraquinonylalanine; lane 19, ferrocenylalanine; lane 20, no tRNA; lane 21, nonacylated tRNA. Lane M contained prestained molecular weight marker.

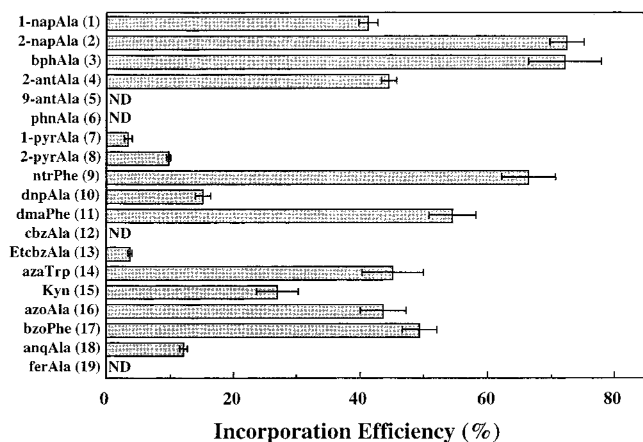

Figure 5. Incorporation efficiencies in the *E. coli* in vitro translation by tRNAcccg's charged with various aromatic nonnatural amino acids at Tyr 83 site of streptavidin. The incorporation efficiencies are determined as described in the text. Data are mean ± sem of at least four assays. ND indicates less than 2% efficiency.

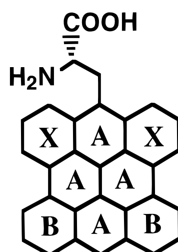

Figure 6. Hypothetical model for the adaptability of nonnatural amino acids to the *E. coli* ribosome. The nonnatural amino acids carrying benzene rings in the region A are allowed. Those carrying rings in the region B may also be allowed. Those carrying benzene rings in the region X are rejected.

2-pyrenylalanine, are efficiently incorporated. On the other hand, 9-phenanthrylalanine and ferrocenylalanine are absolutely rejected.

The suppression efficiency was quantitatively evaluated by comparing the band intensities of mutant streptavidins in the Western blotting with the intensities of the serially diluted wild-type streptavidin. Since the detection depends on the binding of the antibody to the N-terminal T7 tag, possible denaturation or partial unfolding of the mutants will not influence the results. The results are summarized in Figure 5.

The most efficient suppression was observed in the case of 2-naphthylalanine (72%). In contrast, 1-naphthylalanine was incorporated in a 30% efficiency. Similarly, 2-anthrylalanine (45%) and 2-pyrenylalanine (10%) were more effectively incorporated than 9-anthrylalanine (less than 2%) and 1-pyrenylalanine (3%), respectively. In the case of carbazole side groups, 3-(9-ethylcarbazolyl)alanine showed a little higher efficiency than 9-carbazolylalanine. These differences suggest that the protein biosynthesizing system discriminates nonnatural amino acids not by their sizes or hydrophobicities but by their shapes of the side groups. It appears that amino acids with linearly expanded aromatic groups such as 2-naphthylalanine, *p*-biphenylalanine, 2-anthrylalanine, *p*-benzoylphenylalanine, and *p*-phenylazophenylalanine are favored. On the other hand, those with rather widely expanded or bend aromatic groups such as 9-anthrylalanine, 9-phenanthrylalanine, and 9-carbazolylalanine are strongly rejected. It should be noted that the efficiencies of 2-anthraquinonylalanine and azatryptophan increase if L-amino acids are used instead of the racemic ones, since only the L-form is expected to be incorporated.

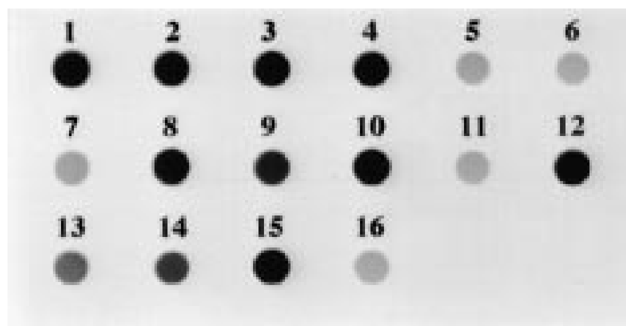

Figure 7. Dot blot analysis of wild-type and mutant streptavidins in the reaction mixture of the *E. coli* in vitro translation. A 5 ng sample of each streptavidin was applied. The intensity of each dot reflects the biotin binding activity of the streptavidin. Spot 1 contained the wild-type streptavidin, and spots 2−16 contained mutant streptavidins containing nonnatural amino acids: 2, 1-naphthylalanine; 3, 2-naphthylalanine; 4, *p*-biphenylalanine; 5, 2-anthrylalanine; 6, 1-pyrenylalanine; 7, 2-pyrenylalanine; 8, *p*-nitrophenylalanine; 9, 3,5-dinitrophenylalanine; 10, *p*-dimethylaminophenylalanine; 11, 3-(9-ethylcarbazolyl)alanine; 12, azatryptophan; 13, kynurenine; 14, *p*-phenylazophenylalanine; 15, *p*-benzoylphenylalanine; 16, 2-anthraquinonylalanine.

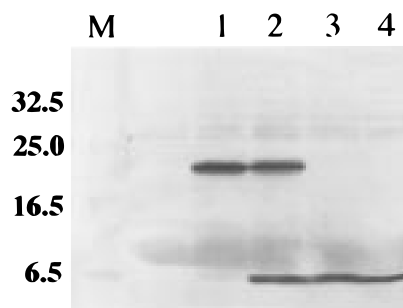

Figure 8. Western blot of the reaction mixture of the rabbit reticulocyte in vitro translation: lane 1, wild-type mRNA; lane 2, in the presence of nitrophenylalanyl tRNAcccg; lane 3, in the presence of nonacylated tRNAcccg; lane 4, in the absence of tRNA. Lane M contained prestained molecular weight marker.

Since the amounts and the lifetimes of the aminoacyl tRNAs are about the same for all amino acids, the sharp amino acid dependence of the incorporation efficiency may be governed by the adaptability of the *E. coli* ribosome to each type of nonnatural amino acid. We have reported previously the inhibitory effects of puromycin analogues carrying various nonnatural amino acids instead of *O*-methyl tyrosine.[5] The efficiency of the inhibition by the puromycin analogues and the efficiency of the incorporation in Figure 5 show a good correlation for almost all the nonnatural amino acids tested. The parallel relationship suggests that a selection of the nonnatural amino acids is occurring at the ribosomal A-site.

A close inspection of the amino acid selectivity in Figure 5 suggests a hypothesis for the allowed and excluded regions of the aromatic groups of L-arylalanine-type amino acids by the *E. coli* ribosome (Figure 6). Since *p*-biphenylalanine and 2-naphthylalanine are almost freely accepted by the ribosome, the regions indicated as A must be allowed by the ribosome. Indeed, 2-pyrenylalanine is incorporated moderately despite the large side group. One of the benzene rings of 1-naphthylalanine indicated as X is less favored by the ribosome. A similar tendency is observed in the comparison of 1- and 2-pyrenylalanine. The rejection of 9-phenanthrylalanine is also explained in terms of the excluded benzene ring X. The benzene ring indicated by B may belong to an allowed region because it is allowed both in 2-anthrylalanine and in *p*-benzoylphenylalanine.

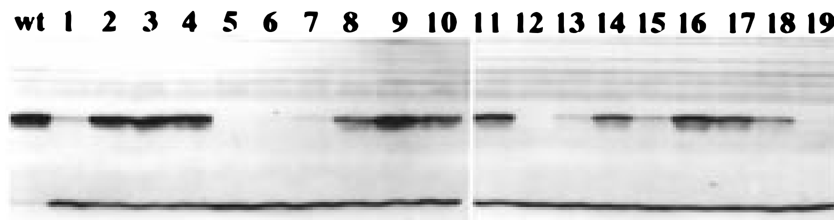

Figure 9. Western blot of the reaction mixture of the rabbit reticulocyte in vitro translation containing aminoacyl tRNACCCG's carrying various nonnatural amino acids: lane wt, wild-type mRNA; lane 1, 1-naphthylalanine; lane 2, 2-naphthylalanine; lane 3, *p*-biphenylalanine; lane 4, 2-anthrylalanine; lane 5, 9-anthrylalanine; lane 6, 9-phenanthrylalanine; lane 7, 1-pyrenylalanine; lane 8, 2-pyrenylalanine; lane 9, *p*-nitrophenylalanine; lane 10, 3,5-dinitrophenylalanine; lane 11, *p*-dimethylaminophenylalanine; lane 12, 9-carbazolylalanine; lane 13, 3-(9-ethylcarbazolyl)alanine; lane 14, azatryptophan; lane 15, kynurenine; lane 16, *p*-phenylazophenylalanine; lane 17, *p*-benzoylphenylalanine; lane 18, 2-anthraquinonylalanine; lane 19, ferrocenylalanine.

The information on the allowed and excluded rings will serve for designing new nonnatural amino acids that can be incorporated in the in vitro system.

Activities of Mutant Streptavidins Carrying Various Nonnatural Amino Acids at the Tyr83 Site. The biotin binding activities of the mutant streptavidins were evaluated by dot blot analysis using biotinylated alkaliphosphatase. The concentrations of the mutant streptavidins in the in vitro translation mixture were obtained from the efficiencies in Figure 5, and 5 ng of each streptavidin was applied to the dot blot analysis. As shown in Figure 7, the incorporation of small amino acids such as naphthylalanines and *p*-nitrophenylalanine did not affect the binding activity very much. However, the incorporation of large aromatic groups such as anthryl and pyrenyl groups reduced the binding activity. Although the Tyr83 site is far from the biotin binding site, the introduction of large amino acids would influence the structure of the binding site and reduce the binding activity.

Incorporation Efficiencies of Various Nonnatural Amino Acids in the Rabbit Reticulocyte Lysate System. Incorporation efficiencies were also examined in rabbit reticulocyte lysate to explore the toleration in the eukariote protein biosynthetic system. As shown in Figure 8, the full-length streptavidin was synthesized only in the presence of tRNACCCG aminoacylated with nitrophenylalanine, indicating that the nonnatural amino acid was incorporated into the protein under direction of the four-base codon in the rabbit reticulocyte system as well as in the *E. coli* system. The Western blotting in Figure 9 shows that various aromatic nonnatural amino acids are incorporated into streptavidin in the reticulocyte lysate system. It must be noted that 9-anthrylalanine, 9-phenanthrylalanine, 9-carbazolylalanine, and ferrocenylalanine could be incorporated neither in the rabbit system nor in the *E. coli* system.

The efficiencies quantitatively evaluated from the Western blotting were summarized in Figure 10. The comparisons between 1- and 2-naphthylalanine, 2- and 9-anthrylalanine, 1- and 2-pyrenylalanine, and 3-(9-ethylcarbazolyl)alanine and 9-carbazolylalanine suggest that nonnatural amino acids with linearly expanded aromatic groups are more favorable than those with widely expanded or bend aromatic groups as has been observed in the *E. coli* system. The parallel amino acid dependence in the two in vitro systems suggests that the ribosome of rabbit reticulocyte recognizes the nonnatural amino acids in the same manner as that of *E. coli*. Presumably, the structures of the A-sites of the two types of ribosomes are well preserved.

A detailed comparison of the incorporation efficiencies in the two systems, however, discloses a subtle but important difference. The ratios of the incorporation efficiencies in the rabbit system to those in the *E. coli* are shown in Figure 11. It is seen that the amino acids with large side groups such as pyrene, dinitrobenzene, and anthraquinone are incorporated in the rabbit system more efficiently than in the *E. coli* system. This suggests that the molecular recognition of the rabbit ribosomal A-site is less tight than that of *E. coli* and the former system is more appropriate for preparing proteins incorporated with nonnatural amino acids carrying relatively large side groups.

The efficient incorporation of the aromatic nonnatural amino acids described in this paper will open a way to investigate fluorescence analyses, photoinduced electron transfers, photoenergy transfers, and other specialty functions on the protein framework.

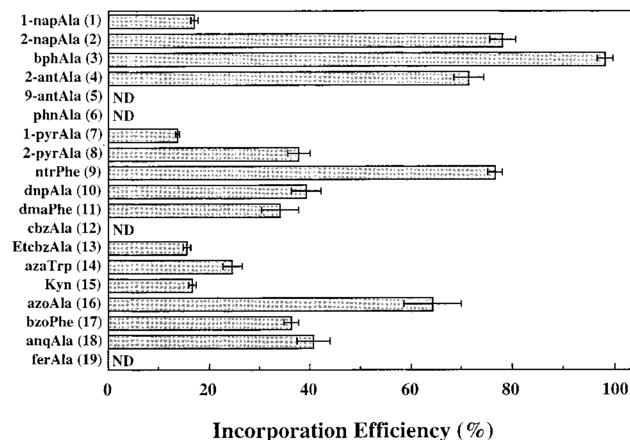

Figure 10. Incorporation efficiencies in rabbit reticulocyte in vitro translation by tRNACCCG's charged with various aromatic nonnatural amino acids at the Tyr 83 site of streptavidin. ND indicates less than 10% efficiency.

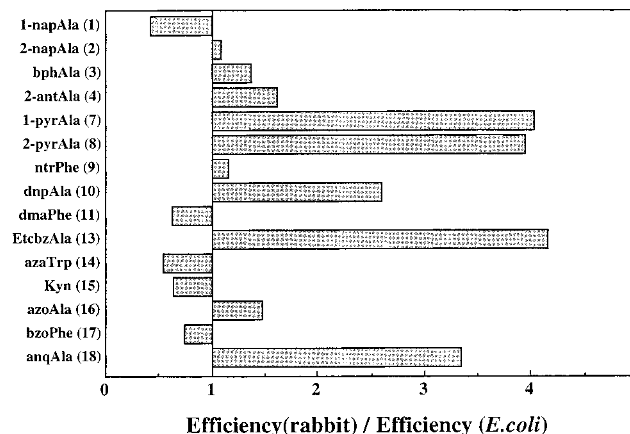

Figure 11. Ratios of the incorporation efficiencies of nonnatural amino acids in rabbit reticulocyte system to those in the *E. coli* system.

Acknowledgment. We thank Dr. Kazuyuki Takai (Chiba Institute of Technology) and Professor Shigeyuki Yokoyama (The University of Tokyo) for their helpful discussions. We also thank Dr. Bong Lee (Kawamura Institute of Chemical Research, Chiba, Japan), Dr. Akira Tsuchida (Gifu University), and Professor Masahide Yamamoto (Kyoto University) for kindly giving us 2-pyrenecarboxaldehyde. HRMS were measured at Okayama University of Science by Dr. Shinji Toyota. This work was supported by the Grants-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan (06403034).

Supporting Information Available: Experimental details of the syntheses of nonnatural amino acids, aminoacyl pdCpAs, tRNA(-CA)s, and mRNAs and HPLC profiles of *p*-nitrophenylalanyl pdCpA (13 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA9813109