Position-specific incorporation of dansylated non-natural amino acids into streptavidin by using a four-base codon

Takahiro Hohsaka[a,b,*], Norihito Muranaka[a,c], Chie Komiyama[c], Kinue Matsui[c], Satomi Takaura[c], Ryoji Abe[c], Hiroshi Murakami[c], Masahiko Sisido[c,*]

[a] *School of Materials Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Tatsunokuchi, Ishikawa 923-1292, Japan*
[b] *PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*
[c] *Department of Bioscience and Biotechnology, Okayama University, 3-1-1 Tsushimanaka, Okayama 700-8530, Japan*

Received 17 November 2003; revised 15 January 2004; accepted 15 January 2004

First published online 4 February 2004

Edited by Gianni Cesareni

Abstract **Novel non-natural amino acids carrying a dansyl fluorescent group were designed, synthesized, and incorporated into various positions of streptavidin by using a CGGG four-base codon in an *Escherichia coli* in vitro translation system. 2,6-Dansyl-aminophenylalanine (2,6-dnsAF) was found to be incorporated into the protein more efficiently than 1,5-dansyl-lysine, 2,6-dansyl-lysine, and 1,5-dansyl-aminophenylalanine. Fluorescence measurements indicate that the position-specific incorporation of the 2,6-dnsAF is a useful technique to probe protein structures. These results also indicate that well-designed non-natural amino acids carrying relatively large side chains can be accepted as substrates of the translation system.**

Key words: Non-natural amino acid; Dansyl; Four base codon; In vitro protein synthesis; Streptavidin

1. Introduction

Incorporation of non-natural amino acids into proteins is a powerful and versatile technique for protein structural and functional analyses. A number of researchers have synthesized proteins containing non-natural amino acids at desired positions by using an amber suppression technique [1–10]. We have developed a four-base codon technique to introduce non-natural amino acids into proteins with high efficiency [11–13]. Incorporation of various non-natural residues has been examined, and we have found that non-natural amino acids with vertically extended side groups are good substrates for translation systems [12].

One of the most important applications of this technique is to probe protein structures. So far, external fluorescent groups have been incorporated into proteins by chemical modification, although the chemical technique suffers the disadvantages that the modification is not sufficiently specific and quantitative. Quantitative modification may not be accomplished even in the case of a highly selective modification of cysteine residue with maleimide derivatives. We have reported the position-specific and quantitative incorporation of relatively small fluorescent non-natural amino acids into proteins, and showed that the method is useful to probe protein structures [14,15]. Here, we report the incorporation of a non-natural amino acid carrying a relatively large dansyl group, and show the usefulness of the dansylated non-natural amino acid to probe protein structures.

A dansyl group is a common fluorescent probe for probing protein structures. It shows strong and environmentally sensitive fluorescence compared with anthraniloyl and anthryl groups successfully incorporated into proteins in our previous studies [14,15]. Incorporation of ε-1,5-dansyl-lysine (1,5-dnsLys) into proteins through the amber suppression has been examined [16,17], but unfortunately 1,5-dnsLys is a poor substrate of an *Escherichia coli* protein synthesis system. The suppression yield was below 5% when 1,5-dnsLys was incorporated into T4 lysozyme [16] and β-galactosidase [17]. The protein synthesis system may reject this amino acid because of its widely expanded dansyl group. We have found that aromatic non-natural amino acids with vertically extended side groups are good substrates of protein synthesis systems [12]. In this study, alternative dansylated amino acids were designed, synthesized, and incorporated into proteins. Fluorescence analysis of proteins containing the dansylated amino acid at specific positions was also investigated.

2. Materials and methods

2.1. Materials

α-Boc-4-aminophenylalanine was purchased from Bachem, and α-Boc-lysine was from Sigma, 1,5-dansyl chloride was from Tokyo Kasei, and 2,6-dansyl chloride was from Molecular Probes. Anti-T7 tag antibody, biotinylated alkaline phosphatase, and TALON metal affinity columns were obtained from Novagen, Calbiochem, and Clontech, respectively.

2.2. Chemical synthesis of dansylated aminoacyl-pdCpA

Dansyl-lysine derivatives were synthesized as follows. To 10.4 mg of α-Boc-lysine in 120 μl of water and 120 μl of dimethylformamide (DMF), 2,6-dansyl chloride (10.7 mg) in 120 μl of DMF and 5 μl of triethylamine were added in an ice bath. After 4 h stirring on ice, the reaction mixture was diluted with ethyl acetate and washed with 5% $KHSO_4$ and then saturated NaCl. The organic phase was dried over $MgSO_4$, and solvent was evaporated. The resulting product was dried under vacuum to yield 12.0 mg of yellow material. The cyanomethylation of the dansyl-lysine was done by adding chloroacetonitrile (60 μl) and triethylamine (8 μl) in 400 μl of acetonitrile. After 24 h stirring at room temperature, the reaction mixture was diluted with ethyl acetate and washed with 4% $NaHCO_3$, 5% $KHSO_4$, and then saturated NaCl. The organic phase was dried over $MgSO_4$, and solvent was evaporated. The resulting product was dried under vacuum to yield 13.7 mg (65%) of yellow material. Aminoacylation of pdCpA

*Corresponding author.
E-mail addresses: hohsaka@jaist.ac.jp (T. Hohsaka), sisido@cc.okayama-u.ac.jp (M. Sisido).

0014-5793/04/$30.00 © 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(04)00099-7

was carried out as previously described. α-Boc-ε-(2,6-dansyl)-lysine cyanomethyl ester (1.1 μmol) was added to pdCpA (0.22 μmol) in 5 μl of DMF, and the resulting solution was incubated at 37°C for 18 h. The reaction was monitored by reverse-phase high performance liquid chromatography (Waters μBondasphere C18, flow rate 0.6 ml/min, linear gradient of 0.1 M $AcONH_4$ pH 4.5 to MeOH over 50 min). The Boc-protected aminoacyl-pdCpA was collected as a precipitate by adding ether, washed with ether and then dried. To remove the Boc group, the product was dissolved in 100 μl of trifluoroacetic acid (TFA) and the resulting solution was incubated for 10 min on ice. TFA was flushed off with N_2 gas, and the resulting pellet was washed with ether. 1,5-Dansyl-lysyl-pdCpA was obtained in a similar manner. The products were identified by mass spectroscopy (Mariner ESI-TOF mass spectrometer). 2,6-Dansyl-lysyl-pdCpA: calcd for $C_{37}H_{49}O_{16}N_{11}SP_2$ ($M-H^-$) 996.2477; found 996.2394. 1,5-Dansyl-lysyl-pdCpA: calcd for $C_{37}H_{49}O_{16}N_{11}SP_2$ ($M-H^-$) 996.2477; found 996.2521.

Dansylated aminophenylalanine derivatives were synthesized as follows. To 5.0 mg of α-Boc-4-aminophenylalanine in 250 μl of 2% aqueous $NaHCO_3$ and 125 μl of 1,5-dioxane, 2,6-dansyl chloride (10.7 mg) in 125 μl of acetonitrile and 25 μl of DMF were added in an ice bath. After 2 h stirring at room temperature, the reaction mixture was diluted with water and washed with ethyl acetate to remove unreacted dansyl chloride. The aqueous layer was acidified with 5% $KHSO_4$ and the product was extracted with ethyl acetate. After washing with saturated aqueous NaCl, the organic phase was dried over $MgSO_4$, and solvent was evaporated. The recrystallization was done with ethyl acetate and *n*-hexane to yield 7.8 mg of α-Boc-ε-(2,6-dansyl)-aminophenylalanine. Cyanomethylation and aminoacylation reactions were done as described above. 2,6-Dansyl-aminophenylalanyl-pdCpA: calcd for $C_{40}H_{47}O_{16}N_{11}SP_2$ ($M-H^-$) 1030.2319; found 1030.2307. 1,5-Dansyl-aminophenylalanyl-pdCpA: calcd for $C_{40}H_{47}O_{16}N_{11}SP_2$ ($M-H^-$) 1030.2319; found 1030.2278.

2.3. In vitro translation

Dansylated aminoacyl-pdCpA was coupled with a tRNA(-CA) containing the CCCG anticodon by using T4 RNA ligase as described previously [12]. In vitro translation was carried out by adding the aminoacyl-$tRNA_{CCCG}$ and a streptavidin mRNA containing a CGGG codon at specific positions into an *E. coli* S-30 system (Promega, for linear templates). The streptavidin gene was fused with a T7 tag and a histidine tag at the N- and C-termini, respectively. The product was analyzed by Western blotting using anti-T7 tag antibody. Incorporation efficiency of each dansylated amino acid was quantified by comparing the band intensity of the full-length streptavidin obtained in the presence of the dansylated aminoacyl-tRNA to those of serially diluted wild-type streptavidin. The biotin binding activity was evaluated by a dot blot analysis using a biotinylated alkaline phosphatase as described previously [12].

2.4. Fluorescence measurement

The reaction mixture (50 μl) of the in vitro translation was applied to a TALON metal affinity column (15 μl) equilibrated with a phosphate buffer (50 mM sodium phosphate, 300 mM NaCl, pH 7.0). The column was washed with a phosphate buffer containing 5 mM imidazole (total 7 ml), and then the full-length streptavidin was eluted with a phosphate buffer containing 0.5 M imidazole and 0.1% PEG8000 (30 μl). The eluate was diluted 10 times with a phosphate buffer containing 0.02% PEG8000, and the fluorescence spectrum was recorded on a Spex Jobin-Yvon Fluoromax-2. The slit width was 1.5 nm for excitation and 5 nm for emission.

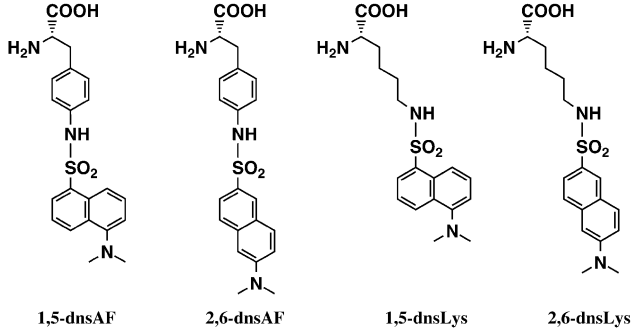

Fig. 1. Structures of dansylated amino acids.

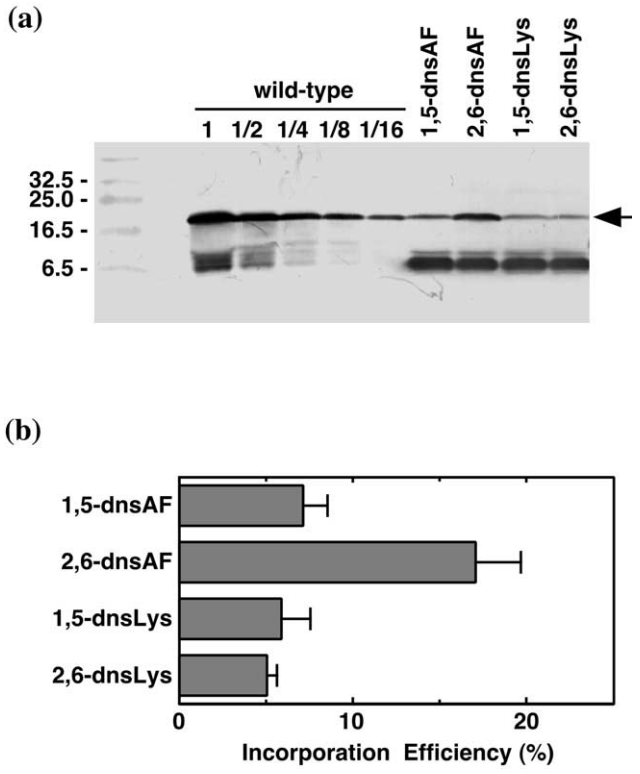

Fig. 2. Incorporation of dansylated amino acids into streptavidin by using a CGGG four-base codon. The in vitro translation products were analyzed by Western blotting using anti-T7 tag antibody (a). The incorporation efficiency was evaluated by comparing band intensities of the full-length streptavidin with serially diluted wild-type streptavidin (b). The data are represented as means ± S.D. of at least three assays.

3. Results and discussion

3.1. Incorporation of dansylated amino acids into a protein in an E. coli in vitro translation system

Fig. 1 shows structures of newly designed dansylated amino acids. We have found that non-natural amino acids with linearly extended aromatic side chains such as *p*-biphenyl and *p*-phenylazophenyl groups are good substrates of an *E. coli* in vitro translation system [12]. According to this information, *p*-aminophenylalanine was employed as the building block. Dansylated aminophenyl residues are expected to be accepted as substrates by an *E. coli* in vitro translation system because of their linearly extended side chain structures. In addition to a conventional 1,5-dansyl group, a 2,6-dansyl group was considered, since the structure will be more favored by the translation system. For comparison, lysine derivatives of the two dansyl groups were also examined.

Incorporation of these dansylated amino acids into proteins was tested by using a streptavidin expression system developed previously [11,12]. A streptavidin gene containing a CGGG codon at position Tyr83 was employed in this system. When the CGGG codon was decoded by the $tRNA_{CCCG}$ aminoacylated with non-natural amino acids, a full-length streptavidin is synthesized. On the other hand, when the amino-

acyl-tRNA$_{CCCG}$ was not accepted by the translation system, a CGG sequence is decoded as arginine and the reading frame shifts to a +1 direction, which causes an irregular termination of peptide elongation at a stop codon downstream. Thus, the incorporation of the non-natural amino acid into position Tyr83 gives full-length streptavidin containing the amino acid and an undesired three-base decoding gives a truncated protein.

Fig. 2a shows the result of Western blotting of the in vitro translation. The full-length streptavidin observed around 20 kDa was obtained in the presence of dansylated aminoacyl-tRNAs, whereas no full-length streptavidin was obtained in the absence of the tRNA. This result indicates that the dansylated amino acids are incorporated into streptavidin. The incorporation efficiency of the dansylated amino acids, however, strongly depends on the structure of their side chains. The relative yield of the full-length streptavidin obtained in the presence of the aminoacyl-tRNA was determined by comparing the band intensity of the full-length product with those of serially diluted wild-type streptavidin as described previously [12]. The results are summarized in Fig. 2b. The most efficient substrate was 2,6-dansyl-aminophenylalanine (2,6-dnsAF, 17%). On the other hand, 1,5-dansyl-aminophenylalanine (1,5-dnsAF) was a poor substrate (7%). Lysine modified with a 1,5-dansyl group (1,5-dnsLys) was also unfavorable (5%) as reported for the amber suppression [16,17]. Lysine modified with a 2,6-dansyl group (2,6-dnsLys) was not incorporated (5%) as efficiently as 2,6-dnsAF.

The result that 2,6-dnsAF is incorporated into streptavidin with relatively high yield indicates that 2,6-dnsAF is a preferable substrate of the in vitro translation system as expected. The ineffectiveness of 1,5-dnsAF suggests that a small difference in side chain structure causes a significant influence on amino acid recognition of the translation system. Both types of dansylated lysine were not accepted as substrates for the translation system, indicating that the benzene ring in 2,6-dnsAF is another important factor in the amino acid recognition. It had been reported that puromycin analogs containing phenylalanine and tyrosine instead of *O*-methyltyrosine showed higher activity than a lysine analog [18]. The ribosomal A-site may prefer amino acids carrying aromatic rings to those carrying alkyl chains. The result that 2,6-dnsLys was

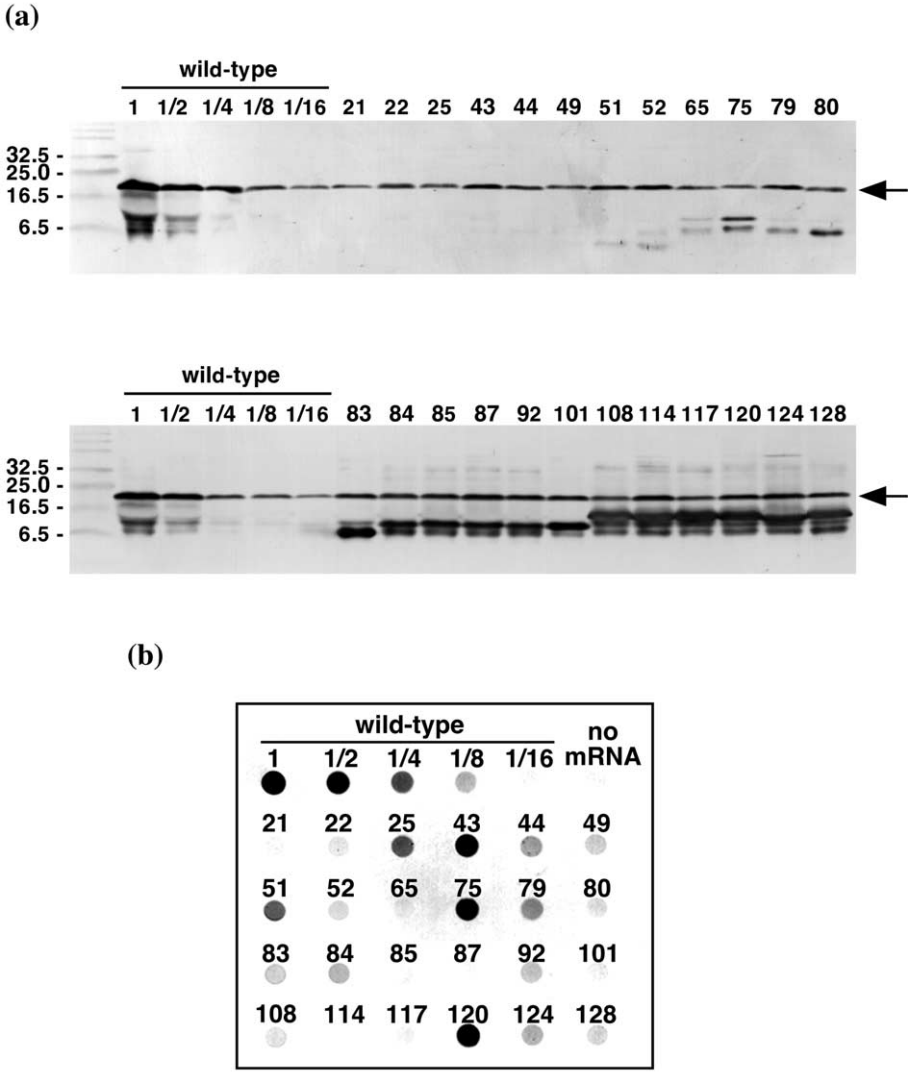

Fig. 3. Incorporation of 2,6-dnsAF into various positions of streptavidin. The expression of full-length streptavidin containing 2,6-dnsAF was examined by Western blot analysis (a), and the biotin binding activity was evaluated by a dot blot analysis using a biotinylated alkaline phosphatase (b).

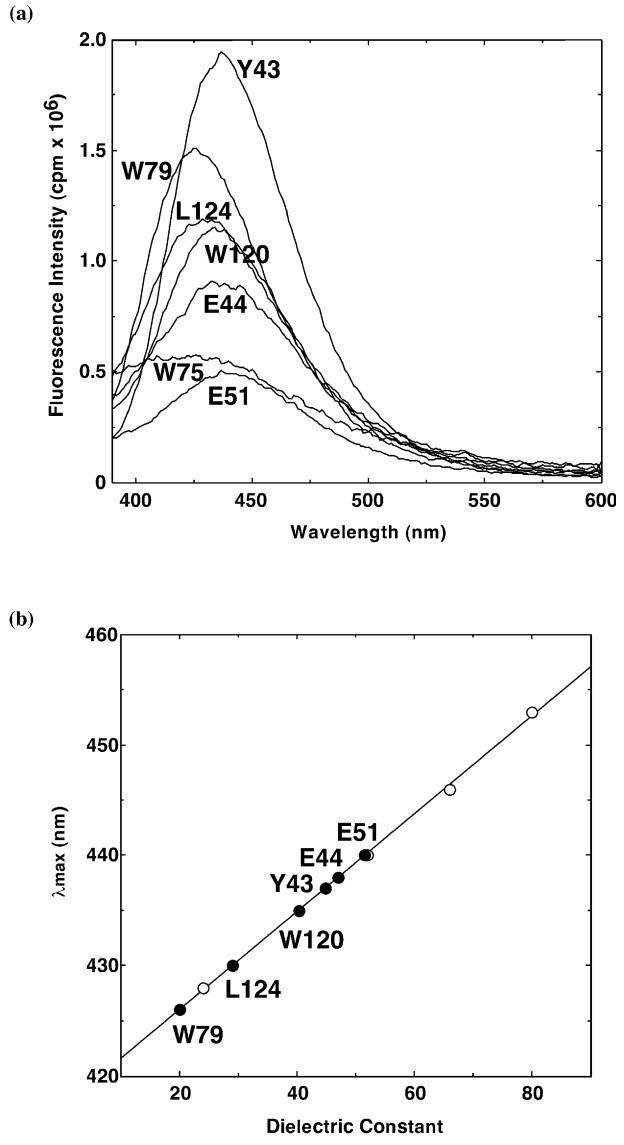

Fig. 4. a: Fluorescence spectra of streptavidin containing 2,6-dnsAF at seven different positions in phosphate buffer pH 7. $\lambda_{ex} = 330$ nm. b: Dependence of λ_{max} of dansyl fluorescence of free 2,6-dnsAF on dielectric constant in 100, 50, 25, and 0% ethanol in phosphate buffer (○). The λ_{max} of dansyl fluorescence of streptavidin mutants was fitted to the dependence (●).

poorly incorporated even though it contained the linearly extended 2,6-dansyl group can be explained in terms of this idea.

Then, 2,6-dnsAF was incorporated into various positions of streptavidin. A series of streptavidin mRNAs containing the CGGG codon at 24 different positions were translated in the in vitro translation system in the presence of 2,6-dnsAF-tRNA. The expression of the dansylated streptavidin was confirmed by Western blotting using anti-T7 tag antibody. As shown in Fig. 3a, a full-length streptavidin was produced at all positions, indicating that streptavidin mutants containing 2,6-dnsAF at respective positions were successfully synthesized. The densitometric analysis of the band indicated that the relative yields of the full-length streptavidins containing 2,6-dnsAF were 10–30%, with respect to the wild-type. The biotin binding activities of these products were evaluated by dot blotting using a biotinylated alkaline phosphatase [12]. The results shown in Fig. 3b indicate that the streptavidin mutants containing 2,6-dnsAF at positions L25, Y43, E44, E51, W75, W79, Y83, R84, Y92, W120, L124, and D128 retained strong or relatively strong biotin binding activity. These positions are located at surface regions of the β-barrel structure or at an interface region of a biotin binding pocket. Other mutants showed no or very low activity, suggesting that the incorporation of 2,6-dnsAF disturbs the binding of biotin or the folding of the streptavidin structure.

3.2. Fluorescence analysis of streptavidin containing 2,6-dnsAF

The 2,6-dnsAF mutants of streptavidin were purified and their fluorescence was measured to evaluate the feasibility of 2,6-dnsAF as a fluorescence probe. To this end, Y43, E44, E51, W75, W79, W120, and L124 mutants were chosen. Since the full-length streptavidin has a histidine tag at the C-terminus, it can be easily purified by using a metal affinity column.

Fluorescence spectra of the seven dansylated streptavidin mutants are shown in Fig. 4. As shown in Fig. 4a, the dansyl fluorescence was observed in all mutants, but λ_{max} of the fluorescence differed. It should be noted that the fluorescence intensity was not quantitative because the concentration of the proteins varied depending on the purification yield. As shown in Fig. 4b, the fluorescence wavelength of 2,6-dnsAF is sensitive to its environment. The λ_{max} value shifted to a longer wavelength, from 428 nm in ethanol (dielectric constant = 24) to 453 nm in phosphate buffer (dielectric constant = 80). Thus the blue-shifted fluorescence observed in the cases of positions W75, W79 and L124 suggests that these positions are rather hydrophobic. The X-ray crystallographic structure of apo-streptavidin (PDB: 1SWB, Fig. 5) [19] supports this result. On the other hand, E44, E51, and W120 mutants showed red-shifted fluorescence, suggesting that these residues are exposed to an aqueous solvent. These residues are located at the surface in the 3D structure, supporting the red-shifted fluorescence (Fig. 6). In the case of position Y43, a red-shifted fluorescence is observed, whereas this residue is located in an inner hydrophobic region. The hydroxyl group of Y43, how-

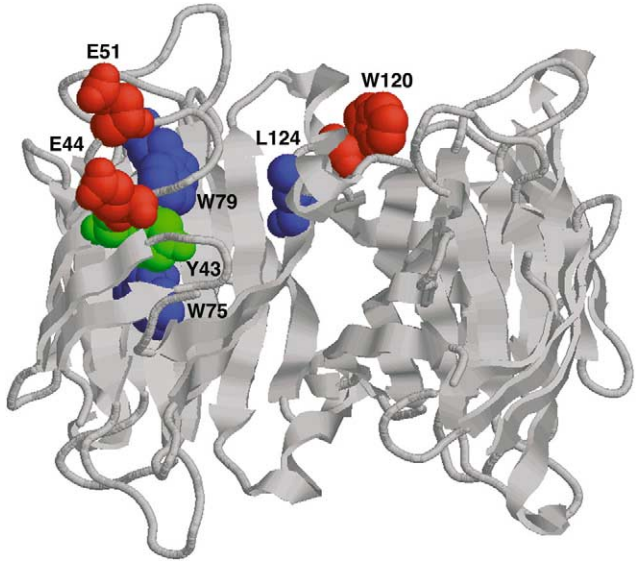

Fig. 5. Three-dimensional structure of apo-streptavidin (PBD: 1SWB). Residues where 2,6-dnsAF was incorporated in this study are represented as a spacefill model.

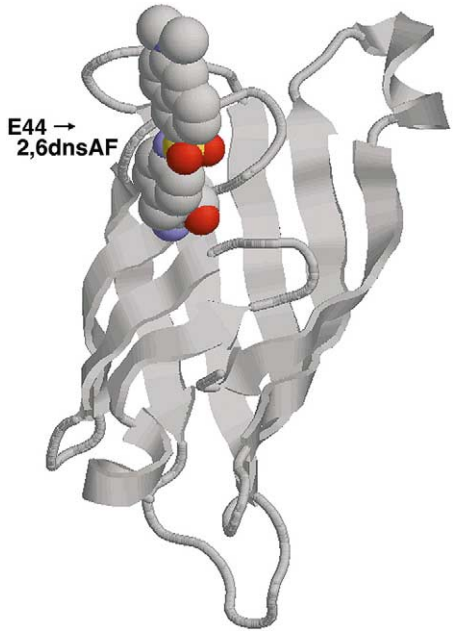

Fig. 6. Predicted structure of apo-streptavidin containing 2,6-dnsAF at position E41. The orientation of 2,6-dnsAF is optimized by varying the side chain angles of the 2,6-dnsAF and those of the neighboring amino acids located within 8 Å from 2,6-dnsAF [20].

ever, is located at a cavity region in apo-streptavidin. The dansyl group may be exposed to solvent through this cavity.

Fluorescence resonance energy transfer from tryptophan to the dansyl group was observed in all cases. Because streptavidin has six tryptophan residues in a monomer unit, multiple energy transfer may occur and no simple interpretation was given. Monitoring the fluorescence change induced by biotin binding was also examined, but quenching of dansyl fluorescence by the addition of biotin was observed and further information was not obtained.

4. Conclusions

Dansylated non-natural amino acids were designed, synthesized and incorporated into streptavidin using a CGGG four-base codon. Among four different amino acids, we found that 2,6-dnsAF was the most efficient substrate of the in vitro translation system. This is the first example of efficient incorporation of a relatively large dansylated amino acid into proteins. 2,6-dnsAF was then incorporated into various positions of streptavidin, and fluorescence measurements were carried out for the mutants that retained biotin binding activity. The λ_{max} of the dansyl fluorescence reflected the microenvironment of the probe positions. This work shows that the position-specific incorporation of 2,6-dnsAF is a useful technique to probe protein environments.

Acknowledgements: This work was supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, Sports and Culture, Japan (No. 11102003), and by Industrial Technology Research Grant Program in 2000 from the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

References

[1] Ellman, J., Mendel, D., Anthony-Cahill, S.J., Noren, C.J. and Schultz, P.G. (1991) Methods Enzymol. 202, 301.
[2] Cornish, V.W., Mendel, D. and Schultz, P.G. (1995) Angew. Chem. Int. Ed. Engl. 34, 621.
[3] Gillmore, M.A., Steward, L.E. and Chamberlin, A.R. (1999) Top. Curr. Chem. 202, 77.
[4] Arslan, T., Mamaev, S.V., Mamaeva, N.V. and Hecht, S.M. (1997) J. Am. Chem. Soc. 119, 10877.
[5] Wang, B., Brown, K.C., Lodder, M., Craik, C.S. and Hecht, S.M. (2002) Biochemistry 41, 2805.
[6] Beene, D.L., Dougherty, D.A. and Lester, H.A. (2003) Curr. Opin. Neurobiol. 13, 264.
[7] Kimata, Y., Shimada, H., Hirose, T. and Ishimura, Y. (1995) Biochem. Biophys. Res. Commun. 208, 96.
[8] Kanamori, T., Nishikawa, S., Nakai, M., Shin, I., Schultz, P.G. and Endo, T. (1999) Proc. Natl. Acad. Sci. USA 96, 3634.
[9] Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C. and Schultz, P.G. (1989) Science 244, 182.
[10] Bain, J.D., Glabe, C.G., Dix, T.A., Chamberlin, A.R. and Diala, E.S. (1989) J. Am. Chem. Soc. 111, 8013.
[11] Hohsaka, T., Ashizuka, Y., Murakami, H. and Sisido, M. (1996) J. Am. Chem. Soc. 118, 9778.
[12] Hohsaka, T., Kajihara, D., Ashizuka, Y., Murakami, H. and Sisido, M. (1999) J. Am. Chem. Soc. 121, 34.
[13] Hohsaka, T., Ashizuka, Y., Taira, H., Murakami, H. and Sisido, M. (2001) Biochemistry 40, 11060.
[14] Murakami, H., Hohsaka, T., Ashizuka, Y., Hashimoto, K. and Sisido, M. (2000) Biomacromolecules 1, 118.
[15] Taki, M., Hohsaka, T., Murakami, H., Taira, K. and Sisido, M. (2001) FEBS Lett. 507, 35.
[16] Cornish, V.W., Benson, D.R., Altenbach, C.A., Hideg, K., Hubbell, W.L. and Schultz, P.G. (1994) Proc. Natl. Acad. Sci. USA 91, 2910.
[17] Steward, L.E., Collins, C.S., Gilmore, M.A., Carlson, J.E., Ross, J.B.A. and Chamberlin, A.R. (1997) J. Am. Chem. Soc. 119, 6.
[18] Vanin, E.F., Greenwell, P. and Symons, R.H. (1974) FEBS Lett. 40, 124.
[19] Freitag, S., Le Trong, I., Klumb, L., Stayton, P.S. and Stenkamp, R.E. (1997) Protein Sci. 6, 1157.
[20] Murakami, H., Hohsaka, T., Ashizuka, Y. and Sisido, M. (1998) J. Am. Chem. Soc. 120, 7520.