DOI: 10.1002/cbic.200700578

N-Terminal Specific Fluorescence Labeling of Proteins through Incorporation of Fluorescent Hydroxy Acid and Subsequent Ester Cleavage

Takayoshi Watanabe, Yoichi Miyata, Ryoji Abe, Norihito Muranaka, and Takahiro Hohsaka*[a]

We have developed a novel method to attach a fluorescent label at the N terminus of proteins through a four-base codon-mediated incorporation of a fluorescent hydroxy acid and subsequent cleavage of the ester bond in a cell-free translation system. We found that a fluorescent-labeled p*-amino-*L*-phenyllactic acid was successfully incorporated downstream of N-terminal tag peptides in response to a CGGG codon, and the tag peptides could be removed through ester cleavage to leave the fluorescent hydroxy acid at the N terminus of the proteins. Immunoprecipitation analysis revealed that ester cleavage occurred spontaneously during the translation reaction. The efficiency of the ester cleavage and the yield of the labeled proteins were dependent on the peptide tag sequence. We demonstrate that the insertion of an asparagine residue between the N-terminal T7 tag and the fluorescent hydroxy acid achieved both quantitative ester cleavage and efficient expression of the labeled proteins. The present method is a potential tool for N-terminal specific labeling of proteins with various compounds.*

Introduction

Fluorescence labeling of proteins is a powerful and versatile technique for the structural and functional analysis of proteins; however, position-specific and quantitative labeling cannot be easily achieved by commonly used chemical modification. Excess or insufficient fluorescence labeling results in decreased protein activity and yield. The use of green fluorescent protein (GFP) derivatives is available for the N- or C-terminal specific labeling of proteins, but the large molecular weight of GFP derivatives might interfere with protein function. Recently, semisynthetic approaches such as native chemical ligation and protein trans-splicing have been used for position-specific labeling of proteins.[1–3]

In contrast to these techniques, incorporation of fluorescent unnatural amino acids[4–9] into proteins by using unnatural mutagenesis[10–13] can be used for position-specific and quantitative fluorescence labeling of proteins. We have reported that well-designed fluorescent unnatural amino acids that can be excited by visible wavelengths are efficiently incorporated into proteins in an *E. coli* cell-free translation system.[14] This position-specific and quantitative fluorescence-labeling method was used for structural analysis by fluorescence resonance energy transfer.[14]

Incorporation of fluorescent labels into the internal positions of proteins, however, might affect the structure and function of proteins to some extent. For fluorescence labeling of proteins without steric hindrance, fluorescent labels should be attached at the N or C terminus of proteins. At the N terminus, tag peptides are often attached to enhance the initiation of translation.[15] In such cases, fluorescent amino acids need to be incorporated downstream of the tag peptide. To eliminate the influence of the additional tag peptide on protein structure and function, it is necessary to remove the tag peptide to leave the fluorescent label at the N terminus.

In this study, we propose the use of a fluorescent hydroxy acid instead of fluorescent amino acids for the N-terminal specific fluorescence labeling of proteins. Hydroxy acids were incorporated into proteins in cell-free translation systems as analogues of amino acids.[16–23] Because an ester bond is much more labile during hydrolysis than an amide bond, the ester bond between a hydroxy acid and an adjacent amino acid can be easily cleaved. Under the weak basic conditions of a translation reaction, the ester hydrolysis is expected to occur spontaneously after the incorporation of a fluorescent hydroxy acid into a polypeptide chain. Therefore, the incorporation of a fluorescent hydroxy acid and subsequent ester cleavage will be available for N-terminal specific fluorescence labeling of proteins that do not have N-terminal tag peptides (Scheme 1 A).

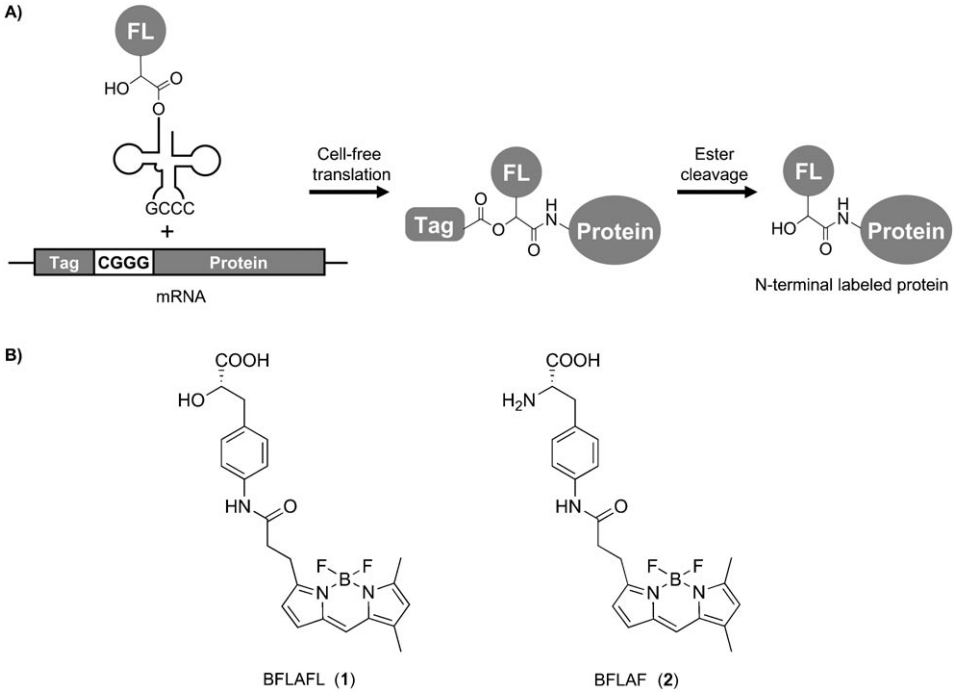
Results and Discussion

Design and synthesis of fluorescent hydroxy acid

As a fluorescent hydroxy acid, we designed *p*-(BODIPYFL-amino)-L-phenyllactic acid (BFLAFL, **1**; Scheme 1 B). Based on our previous results, which indicated that *p*-substituted phenylalanine derivatives are good substrates for the translation machinery, *p*-(BODIPYFL-amino)-L-phenylalanine (BFLAF, **2**) was synthesized and was found to be efficiently incorporated into

[a] *T. Watanabe, Y. Miyata, R. Abe, Dr. N. Muranaka, Prof. T. Hohsaka*
School of Materials Science (Japan)
Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923-1292 (Japan)
Fax: (+81) 761-51-1149
E-mail: hohsaka@jaist.ac.jp

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.



Scheme 1. A) Schematic illustration of N-terminal specific fluorescence labeling through incorporation of a fluorescent hydroxy acid and subsequent ester cleavage. B) Structures of fluorescent hydroxy acid and amino acid used in this study.

proteins.[14] BFLAFL, which has a hydroxyl group in place of the amino group of BFLAF, is also expected to be acceptable as a substrate for the translational machinery.

p-Aminophenyllactic acid was synthesized by substitution of the amino group of *p*-nitrophenylalanine with a hydroxy group, followed by reduction of the nitro group into an amino group. The carboxyl group of aminophenyllactic acid was activated by forming cyanomethyl ester, and was attached to a dinucleotide, pdCpA. Aminophenyllactyl-pdCpA was then treated with BODIPYFL-succinimide ester at pH 5 to obtain pdCpA, which was acylated with BFLAFL (BFLAFL-pdCpA). The product was purified by HPLC and identified by electrospray ionization mass spectrometry (ESI-MS). According to the chemical aminoacylation method,[24] BFLAFL-pdCpA was ligated with a T7-transcribed yeast phenylalanine tRNA that had a four-base anticodon (CCCG) and lacked the 3′ terminal dinucleotide, by using T4 RNA ligase. For comparison, a tRNA that was aminoacylated with BFLAF was prepared as described previously.[14]

Incorporation of fluorescent hydroxy acid into the N terminus of a protein fused with various N-terminal tag peptides

A streptavidin gene that was fused with the His tag at the C terminus was chosen for evaluating the incorporation of the fluorescent hydroxy acid. Five types of tag peptides, T7,[25] RGSHis,[26] VSV-G,[27] Avi,[28] and EE[29] (Table 1), were fused to the N terminus of the streptavidin gene in which a sequence CGGG AGT AAC was inserted between the tag sequence and the GAC codon that encodes Asp1 of streptavidin. For comparison, a streptavidin gene that contained the CGGG AGT AAC sequence immediately after the initiation codon was prepared. When the four-base codon CGGG is decoded by BFLAFL–tRNA, which has the corresponding four-base anticodon, BFLAFL will be incorporated in response to the CGGG codon, and as a consequence, a full-length fluorescent streptavidin will be obtained. In contrast, the three-base decoding of CGG by a naturally occurring Arg-tRNA will terminate the protein synthesis at the downstream stop codon TAA of the −1 reading frame.

Each of the tagged streptavidin mRNA was added to an *E. coli* cell-free translation system along with BFLAFL–tRNA or BFLAF–tRNA. The translation product was analyzed by SDS-PAGE, followed by fluorescence imaging of the gel. The

Table 1. Sequence of tags attached at the N terminus of streptavidin gene.

Tag	Sequence
T7	ATG GCT AGC ATG ACT GGT GGA CAG CAA ATG GGT
RGSHis	ATG AGA GGA TCG CAT CAT CAT CAT CAT CAT
VSV-G	ATG GCA TAC ACT GAT ATC GAA ATG AAC CGC CTG GGT AAG
Avi	ATG GGT CTG AAC GAC ATC TTC GAG GCT CAG AAA ATC GAA TGG-CAC GAA
EE	ATG GAA TAC ATG CCA ATG GAA

fluorescent SDS-PAGE analysis allowed us to quantitatively evaluate the incorporation of BFLAFL and the ester cleavage. As shown in Figure 1 A, the addition of BFLAF–tRNA gave fluorescent bands of about 20 kDa, which correspond to the full-length tagged protein. On the other hand, the addition of BFLAFL–tRNA gave slightly smaller protein bands than BFLAF–tRNA. The translation products for T7-tagged streptavidin were further analyzed by Western blotting. Western blotting by using an anti-T7 tag antibody showed that the addition of BFLAFL–tRNA did not produce a band, whereas BFLAF–tRNA gave a distinct band with the same mobility as T7-tagged wild-type streptavidin (Figure 1 C). In Western blots that were probed by using an antibody against the C-terminal His tag, the translation product in the presence of BFLAFL–tRNA gave a band with the same mobility as wild-type streptavidin that did not have the N-terminal tag (Figure 1 D). The fluorescence image and the Western blot images indicate that BFLAFL is incorporated into streptavidin, and the ester bond between the T7 tag and BFLAFL is cleaved. The BFLAFL-containing streptavi-

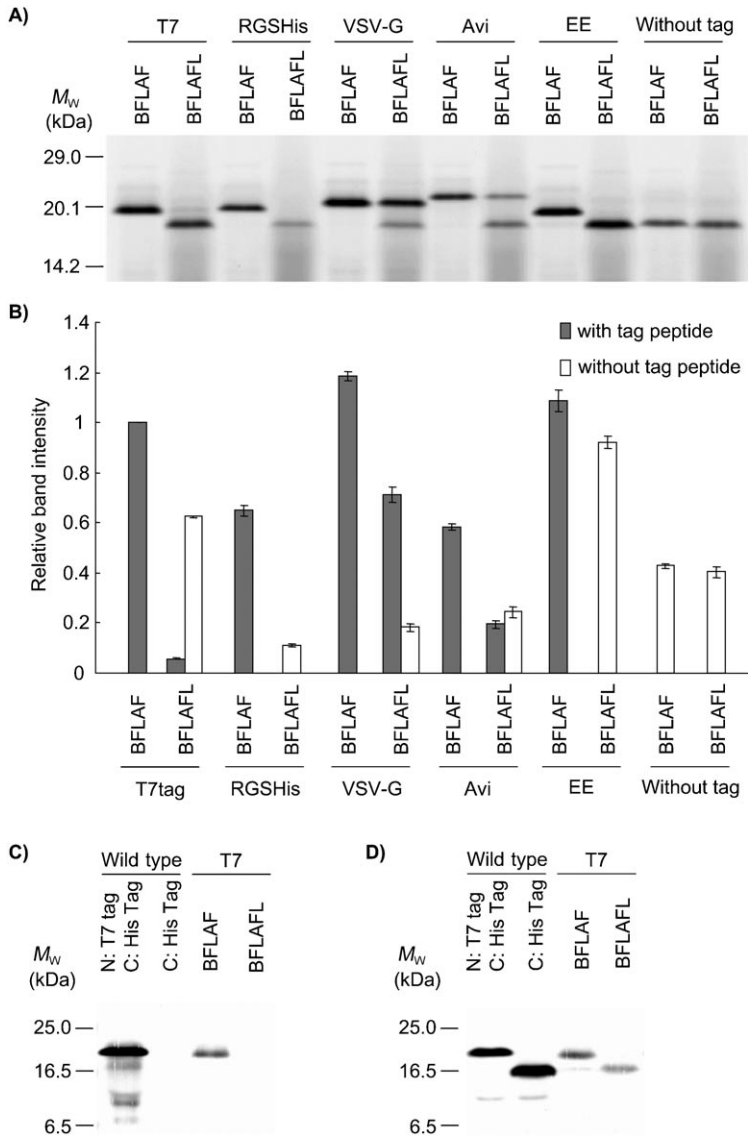


Figure 1. Incorporation of BFLAF and BFLAFL into streptavidin fused with various N-terminal tag peptides. A) Fluorescence image ($\lambda_{ex}=488$ nm and $\lambda_{em}=520$ nm) of SDS-PAGE for cell-free translation products that were expressed in the presence of either BFLAF–tRNA or BFLAFL–tRNA. B) Relative intensity of the fluorescent bands. Data are represented as values relative to BFLAF-containing T7-tagged streptavidin (mean ± standard deviation, $n=3$). Western blot analysis of T7-tagged streptavidin by using C) an anti-T7 tag antibody, and D) an anti-His tag antibody.

din was then purified by using Ni-NTA beads and analyzed by MALDI-TOF MS (Figure S1 in the Supporting Information). The mass value that corresponded to the BFLAFL-containing streptavidin without the T7 tag was identified (calcd, 17931.7; found: 17929.8); this demonstrates that the tag peptide was removed to leave BFLAFL at the N terminus of streptavidin.

Quantification of the fluorescent bands on the SDS-PAGE gel indicates that the ratio of tagged to nontagged protein and the yield of the fluorescently labeled protein are different and depend on the type of tag peptides used (Figure 1B). T7, RGSHis, and EE tags showed highly efficient ester cleavage, and in addition, T7 and EE tags gave a higher yield of BFLAFL-containing proteins than the nontagged protein obtained from the mRNA without the tag sequence. This demonstrates the effectiveness of the N-terminal tags for the efficient production of proteins that are labeled at the N terminus. The incorporation efficiencies of BFLAFL for the EE-tagged and nontagged proteins were comparable to those of BFLAF; this indicates that the fluorescent hydroxy acid could be potentially accepted as a substrate for the translation machinery with a similar efficiency as the corresponding fluorescent amino acid. The total yield of the BFLAFL-containing streptavidin can be estimated to be 5–10 $\mu g\,mL^{-1}$ based on previous results, which have reported the yield of the T7-tagged wild-type streptavidin to be 25 $\mu g\,mL^{-1}$[30] and the incorporation efficiency of BFLAF to be 46%.[14]

Cleavage of ester bond during cell-free translation

Ester cleavage is expected to occur during incubation in the cell-free translation system, but it could occur during the sample preparation for SDS-PAGE. To examine whether the ester cleavage occurs during the translation reaction, the BFLAFL-containing T7-tagged protein was immunoprecipitated by an anti-T7 tag antibody immediately after the translation reaction, followed by analysis by SDS-PAGE. In this experiment, monomeric fatty acid binding protein (FABP) was used instead of tetrameric streptavidin, because the tetramerization of streptavidin made it difficult to separate T7-tagged monomer from nontagged monomer. The fluorescence image of SDS-PAGE indicates that the BFLAF-containing protein was not observed to any significant extent in the supernatant, but was present in the eluate fraction (Figure 2). On the other hand, the BFLAFL-containing protein remained in the supernatant and was not present in the eluate from the immunoresin; this indicates that the BFLAFL-containing protein did not have the T7 tag peptide. These results show that the removal of the T7 tag peptide is completed during the translation reaction.

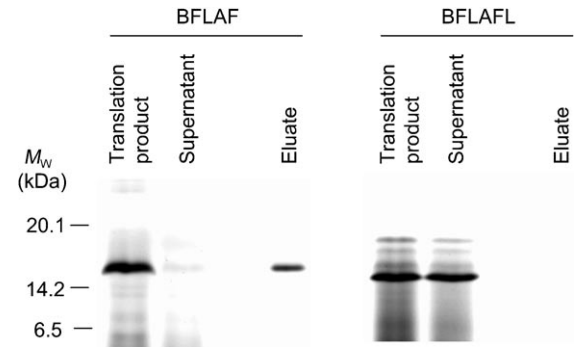


Figure 2. Immunoprecipitation analysis of T7-tagged FABP that contained either BFLAF or BFLAFL. The supernatant and eluate from the immunoresin were analyzed by fluorescence imaging of SDS-PAGE.

In the cases of VSV-G and Avi tags, however, a significant amount of tag-containing protein was observed in the SDS-PAGE (Figure 1 A). The VSV-G tagged protein was completely hydrolyzed when the cell-free translation solution was mixed with 1 vol of 1 M $NaHCO_3$ and further incubated at 37 °C for 1 h (Figure S2); this indicates that the VSV-G tag is not effectively removed during the translation reaction even though it is fused to a protein via an alkali-labile ester bond. These results indicate that the type of tag peptide might determine whether the ester bond is cleaved during the translation reaction.

Effect of upstream amino acid on ester cleavage and translation efficiency

There is a possibility that the cleavage of the ester bond between the tag peptide and the fluorescent hydroxy acid depends on the type of amino acid upstream of the fluorescent hydroxy acid. To investigate the effect of the upstream amino acid on the ester cleavage, each of 20 codons, which correspond to the 20 amino acids, was inserted between the T7 tag and the CGGG codon in the streptavidin gene. Fluorescence images of SDS-PAGE for the translation products showed that the ratio of tagged to nontagged proteins was significantly different depending on the type of upstream amino acids (Figure 3). In particular, asparagine and proline allowed quantitative ester cleavage, but hydrophobic residues, such as isoleucine and methionine, strongly prevented ester cleavage.

The ester bond between asparagine and BFLAFL can be cleaved via a succinimide intermediate as was observed in a protein-splicing reaction in which an amide nitrogen of asparagine at the C terminus of spliced inteins nucleophilically attacked the carbonyl carbon atom of the peptide bond and formed a succinimide.[31] The nucleophilic substitution could easily occur when the peptide bond was replaced by an ester bond. It has been reported that an asparagine residue in a methyl ester of tetrapeptide YAHN was converted to the succinimide.[32] The ester bonds between BFLAFL and amino acids other than asparagine could be hydrolyzed depending on the structural properties of the amino acid residues. The proline ring and small residues, such as glycine and serine, might not interfere with the attack of hydroxyl ions on the ester bond. On the other hand, large hydrophobic residues probably protect the ester bond from the attack by hydroxyl ions.

In previous studies, it was found that ester bond cleavage often requires alkaline treatments and its efficiency depends on the adjacent amino acid sequences (Tables S1 and S2).[16–23] However, it should be noted that in these studies, the hydroxy acids were incorporated into the internal regions and not the N termini of proteins, which comprise a limited variety of upstream amino acids. In one report, where the upstream amino acid was asparagine, quantitative ester cleavage was observed in SDS-PAGE without alkaline treatment.[17] To evaluate the cleavage of BFLAFL-derived ester bonds in the internal regions, BFLAFL was incorporated into the internal regions of maltose-binding protein (MBP).[33] As shown in Figure S3, ester cleavage was largely inhibited at Y70, Y167, and Y176 positions of MBP, despite the upstream amino acids being glycine and lysine, whose ester cleavage efficiencies at the downstream of T7 tag were 78% and 56%, respectively. This result is consistent with previous studies that suggested that ester bonds at the internal regions of proteins are stable. The stability of these ester bonds can be explained by considering the fact that they are protected by the surrounding protein structure in the internal regions, but they are less protected downstream of the N-terminal tag peptide. However, substituting the upstream amino acids with asparagine at these three internal positions of MBP resulted in nearly quantitative ester cleavage (Figure S3), as reported previously, where the ester bond was easily cleaved when the upstream amino acid was asparagine.[17] The efficient ester cleavage might be because of the same mechanism as that described for T7 tag with asparagine insertion.

The upstream amino acid also affects the yield of the fluorescent-labeled proteins. The insertion of asparagine increased the yield of the labeled protein by 3.1-fold compared with the labeled protein without insertion. On the other hand, the insertion of histidine resulted in a very low yield, which is consistent with the low yield of RGSHis-tagged protein (Figure 1 B). In other upstream amino acids, total yields of the tagged and nontagged proteins were different depending on the type of upstream amino acids. As discussed below, these results raise the possibility that a slight change in the sequence of

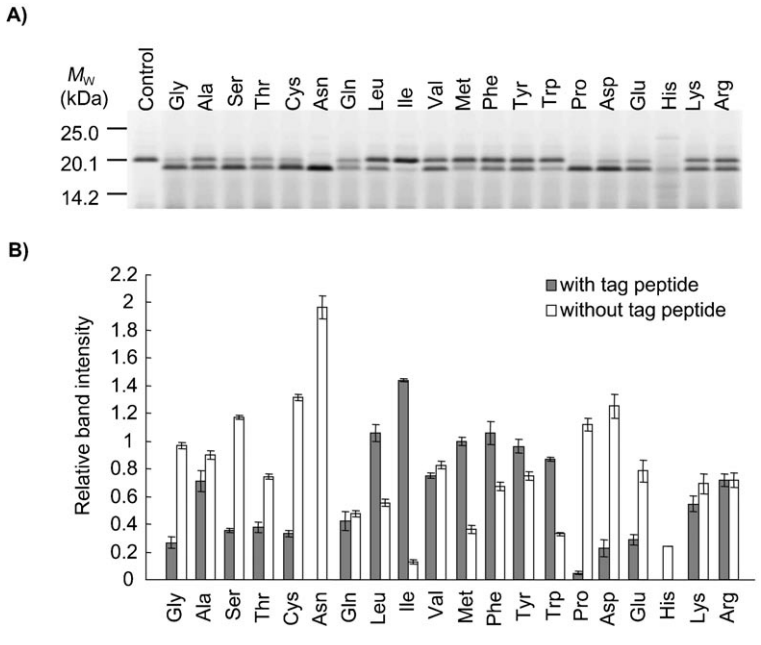


Figure 3. Incorporation of BFLAFL into T7-tagged streptavidin with insertion of 20 types of amino acids between the T7 tag and CGGG codon. A) Fluorescence image (λ_{ex} = 488 nm and λ_{em} = 520 nm) of SDS-PAGE for cell-free translation products obtained in the presence of BFLAFL–tRNA. B) Relative intensity of the fluorescent bands. Data are represented as values relative to BFLAF-containing T7-tagged streptavidin without insertion (described as "Control"; mean ± standard deviation, n = 3).

tag peptides might influence the yield of the labeled proteins.

To confirm the effect of the upstream asparagine residue on the ester cleavage in various tag peptides, an AAC codon was inserted upstream of the CGGG codon, or introduced in place of the C-terminal codons of the tags. As shown in Figure 4,

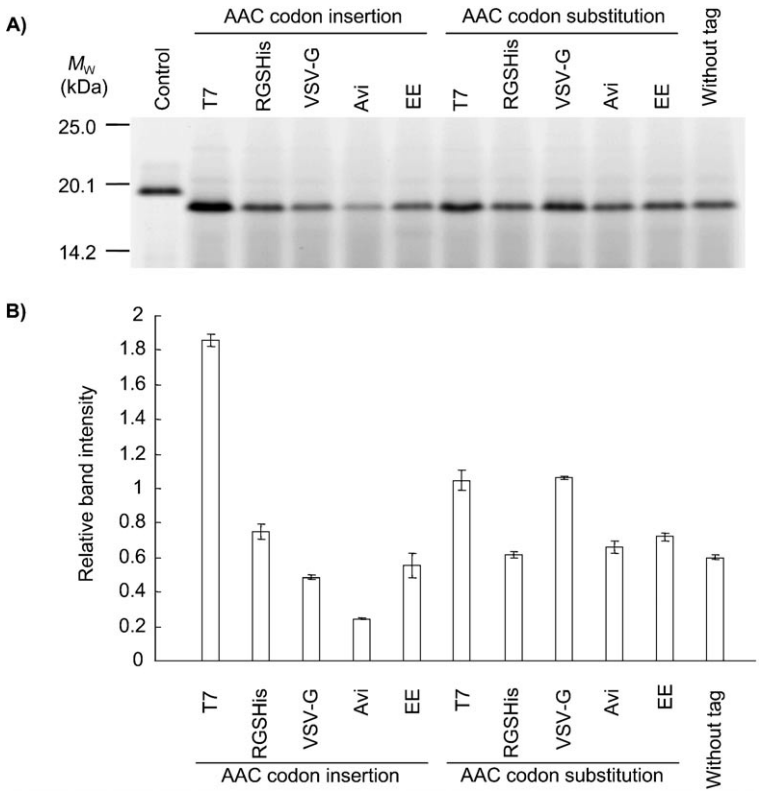


Figure 4. Incorporation of BFLAFL into N-terminal tagged streptavidin in which asparagine was either inserted downstream of the C-terminal codons of the tags, or introduced in place of the C-terminal codons. A) Fluorescence image (λ_{ex} = 488 nm and λ_{em} = 520 nm) of SDS-PAGE for cell-free translation products expressed in the presence of BFLAFL–tRNA. B) Relative intensity of the fluorescent bands. Data are represented as values relative to BFLAF-containing T7-tagged streptavidin without asparagine insertion (described as "Control"; mean ± standard deviation, n = 3).

ester cleavage occurred during the translation reaction in all cases. This demonstrates that the upstream asparagine residue is effective for quantitative ester cleavage regardless of the type of tag peptides. The yield of the labeled proteins was, however, different depending on the tag peptide sequence. The T7 tag with the asparagine insertion showed the highest yield (3.1-fold compared with the nontagged protein), although substitution with asparagine resulted in moderate yield. On the other hand, the VSV-G, Avi, and EE tags with the asparagine insertion showed lower yields than that for the nontagged protein, but the substitution into asparagine increased the yields to higher levels. The difference between the insertion and substitution suggests that the yield of the labeled proteins is significantly affected by only a single codon change (Figure 3).

The yield of the labeled proteins will reflect both the decoding efficiency of the CGGG codon and the translation efficiency of mRNAs. In our previous work, the structure of unnatural amino acids was identified as the major factor that determines the decoding efficiency of the CGGG codon.[4] The incorporation position also influenced the decoding efficiency of the CGGG codon when BFLAF was incorporated,[14] while rather small residues such as *p*-aminophenylalanine were equally incorporated at various positions.[33] In the present study, a significant difference between the insertion and substitution was also observed when BFLAF was incorporated (Figure S4), but was not observed when tyrosine was incorporated (Figure S5). These facts indicate that the decoding of the CGGG codon might be highly dependent on the neighboring sequence for amino or hydroxy acids that carry large side chains.

Asparagine insertion and substitution might influence the translation efficiency of mRNAs. N-terminal tag sequences have often been used to enhance the translation efficiency of mRNAs.[15] In addition, silent mutations of the N-terminal nucleotide sequence were sometimes effective in increasing translation efficiency in a cell-free translation system.[34] In the present study, even a single codon insertion or substitution for the N-terminal tag sequence might significantly increase or decrease the translation efficiency of mRNAs.

Utility of T7 tag with asparagine insertion for other proteins

The utility of the T7 tag with asparagine insertion was examined for several proteins other than streptavidin. The T7 tag with the AAC codon insertion was fused with the N terminus of calmodulin, FABP, and λ-Cro repressor genes. For comparison, the CGGG AGT AAC sequence was added to these genes immediately after the initiation codon. A fluorescence image of the SDS-PAGE for the translation products showed that the T7 tag with asparagine insertion gave labeled proteins that had the same mobility as the corresponding nontagged proteins (Figure 5). This indicates that the ester bond was quantitatively cleaved in these proteins, as with streptavidin. Moreover, the yields of the labeled proteins were higher than those of the corresponding nontagged proteins. These results suggest that the T7 tag with asparagine insertion is universally applicable for efficient expression of N-terminal labeled proteins.

Conclusions

A novel method was developed for the N-terminal specific fluorescence labeling of proteins through the incorporation of a fluorescent hydroxy acid and spontaneous cleavage of the resulting ester bond. For N-terminal specific labeling, initiation codon-mediated incorporation of fluorescent amino acids[35–37] and fluorescent carboxylic acids[38] have been reported, but they cannot be very effective for genes in which the N-termi-

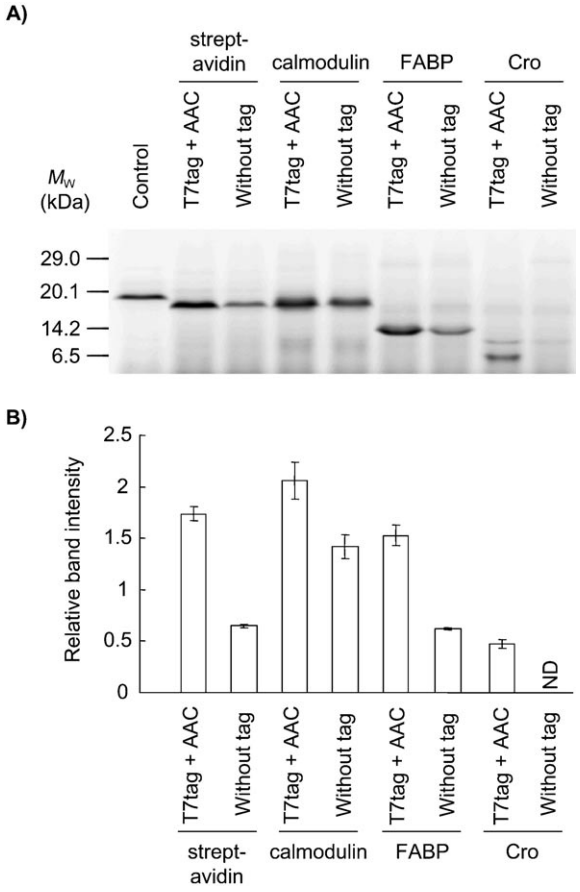


Figure 5. Incorporation of BFLAFL into several proteins that contained the T7 tag that had asparagine inserted between the T7 tag and CGGG codon. A) Fluorescence image ($\lambda_{ex}=488$ nm and $\lambda_{em}=520$ nm) of SDS-PAGE for cell-free translation products expressed in the presence of BFLAFL–tRNA. B) Relative intensity of the fluorescent bands. Data are represented as values relative to BFLAF-containing T7-tagged streptavidin without asparagine insertion (described as "Control"; mean ± standard deviation, $n=3$); ND: not detected.

nal nucleotide or amino acid sequence is inappropriate for protein expression. The present method has a unique advantage in that the protein expression can be enhanced by peptide tags that are spontaneously removed from the expressed proteins; this leaves the fluorescent label at the N terminus. For practical applications, the T7 tag that contains the insertional asparagine upstream of the hydroxy acid was found to be suitable for the efficient synthesis of fluorescently labeled proteins without tag peptide. Further optimization of tag peptides to achieve both efficient expression and quantitative ester cleavage will be essential to improve this method. Moreover, hydroxy acids that are labeled with other labeling compounds, such as biotin, will make this method a general and useful tool for the N-terminal specific labeling of proteins.

Experimental Section

Materials: *p*-Nitro-L-phenylalanine was purchased from Tokyo Chemical Industry (Tokyo, Japan); BODIPYFL-succinimide ester was purchased from Molecular Probes; T4 RNA ligase was purchased from Takara BIO (Osaka, Japan); RTS *E. coli* linear template generation set was purchased from Roche Diagnostics; *E. coli* S30 extract for linear templates, alkaline phosphatase-labeled anti-mouse IgG, and MagneHis Ni-Particles was purchased from Promega; anti-T7 tag antibody, anti-His tag antibody, and GrabIt T7-Tag Kit were purchased from Novagen; XTerra C18 and XBridge C18 columns were purchased from Waters; Poros R2/10 column was purchased from Applied Biosystems; and ZipTipC18 was purchased from Millipore.

Synthesis of *p*-nitro-L-phenyllactic acid: An aqueous solution of sodium nitrite (2.2 M; 10 mL) was gradually added to a solution of *p*-nitro-L-phenylalanine (1.49 g, 7.07 mmol) in aqueous H_2SO_4 (0.55 M; 22 mL) and acetone (22 mL), over 30 min with cooling on ice. The mixture was stirred on ice for 1.5 h, and then at room temperature for 16 h. After evaporation of the acetone, the mixture was extracted with EtOAc. The organic layer was washed once with water, once with saturated aqueous NaCl, and then dried over Na_2SO_4. After evaporation of the solvent, isopropyl acetate and hexane were added to precipitate the product, which was collected by filtration and dried under vacuum to afford *p*-nitro-L-phenyllactic acid (809.8 mg, 3.84 mmol; 54% yield): ^{1}H NMR (300 MHz, [D$_6$]DMSO): $\delta=2.93$ (dd, $J=8.4$, 13.8 Hz, 1H), 3.12 (dd, $J=4.2$, 13.8 Hz, 1H), 4.23 (tt, $J=4.2$, 8.4 Hz, 1H), 5.47 (br s, 1H), 7.53 (d, $J=$ 8.7 Hz, 2H), 8.15 (d, $J=8.7$ Hz, 2H), 12.64 (br s, 1H).

Synthesis of *p*-amino-L-phenyllactic acid: Anhydrous EtOH (10 mL) was added to a mixture of *p*-nitro-L-phenyllactic acid (507 mg, 2.40 mmol) and 10% Pd/C (275 mg) under H_2 gas. The resulting mixture was stirred at room temperature for 1.5 h. After removal of Pd/C by filtration followed by evaporation of the solvent, MeOH was added to precipitate the product, which was collected by filtration and dried under vacuum to afford *p*-amino-L-phenyllactic acid (173.8 mg, 0.96 mmol; 40% yield): ^{1}H NMR (300 MHz, [D$_6$]DMSO): $\delta=2.57$ (dd, $J=8.1$, 13.8 Hz, 1H), 2.76 (dd, $J=4.5$, 13.8 Hz, 1H), 3.98 (tt, $J=4.5$, 8.1 Hz, 1H), 6.45 (d, $J=8.1$ Hz, 2H), 6.86 (d, $J=8.1$ Hz, 2H).

Synthesis of *p*-(Boc-amino)-L-phenyllactic acid cyanomethyl ester: A solution of $(Boc)_2O$ (3.7 M) in dioxane (67 μL, 0.25 mmol) was added to a mixture of *p*-amino-L-phenyllactic acid (21.7 mg, 0.12 mmol) in aqueous $NaHCO_3$ (0.1 M; 2.7 mL) and dioxane (2.7 mL). The mixture was stirred at room temperature for 3 h. After removal of dioxane by evaporation, the resulting mixture was extracted with EtOAc. The organic layer was washed once with aqueous $KHSO_4$ (5%), once with saturated aqueous NaCl, and then dried over anhydrous $MgSO_4$. The solvent was removed by evaporation to afford *p*-(Boc-amino)-L-phenyllactic acid (23.2 mg) as a crude product. To a mixture of *p*-(Boc-amino)-L-phenyllactic acid (23.2 mg, 0.082 mmol) in triethylamine (690 μL, 5 mmol) and acetonitrile (1.5 mL), chloroacetonitrile (520 μL, 8.3 mmol) was gradually added. After being stirred at room temperature for 4 h, triethylamine (69 μL, 0.5 mmol) and chloroacetonitrile (52 μL, 0.83 mmol) were added. The resulting mixture was stirred at room temperature for a further 9 h. The mixture was acidified with aqueous $KHSO_4$ (5%) and extracted with EtOAc. The organic layer was washed once with aqueous $KHSO_4$ (5%), and then the solvent was removed by evaporation. The desired product was purified by a preparative reversed-phase HPLC (XBridge C18, 5 μm, 10×50 mm, flow rate: 3.0 mL min^{-1} with a linear gradient of 0–100% MeOH in 0.38% formic acid over 15 min) to give *p*-(Boc-amino)-L-phenyllactic acid cyanomethyl ester (10.0 mg, 31.1 μmol; 26% yield, 2 steps): ^{1}H NMR (300 MHz, $CDCl_3$): $\delta=1.51$ (s, 9H), 2.54 (br s, 1H), 2.99 (dd, $J=6.3$, 14.1 Hz, 1H), 3.11 (dd, $J=4.8$, 14.1 Hz, 1H), 4.53 (br s, 1H), 4.77 (d, $J=3.0$ Hz, 2H), 6.47 (br s, 1H), 7.13 (d, $J=8.4$ Hz, 2H), 7.32 (d, $J=8.7$ Hz, 2H).

Synthesis of *p*-amino-L-phenyllactyl-pdCpA: Acylation of pdCpA with *p*-(Boc-amino)-L-phenyllactic acid was carried out by the addition of *p*-(Boc-amino)-L-phenyllactic acid cyanomethyl ester (5.90 mg; 18.4 μmol) to a 88 mM DMF solution of pdCpA tetra-*n*-butyl-ammonium salt (105.5 μL; 9.30 μmol). The resulting solution was kept at 37 °C for 6.5 h. Et_2O (1 mL) was added to the solution, and the precipitate was collected by centrifuging. The resulting precipitate was washed twice with Et_2O (1 mL), and dried under vacuum. The desired product was purified with an analytical scale reversed-phase HPLC (XTerra C18, 2.5 μm, 4.6×20 mm, flow rate 1.5 mL min^{-1} with a linear gradient of 0–100% MeOH in 0.38% formic acid, over 10 min). The purified product was dissolved in trifluoroacetic acid (200 μL) and placed on ice for 10 min to remove the Boc group. After evaporation of trifluoroacetic acid by vacuum centrifuge, the pellet was washed twice with Et_2O (1 mL), and dried under vacuum to afford *p*-amino-L-phenyllactyl-pdCpA (1.53 μmol; 16% yield). The product was identified by ESI-TOF MS. *p*-Amino-L-phenyllactyl-pdCpA; calcd: 798.1655; found: 798.1626 $[M-H]^-$.

Synthesis of *p*-(BODIPYFL-amino)-L-phenyllactyl-pdCpA (BFLAFL-pdCpA): A DMSO solution of BODIPYFL-succinimide ester (50 mM; 36 μL, 1.8 μmol) was added to a mixture of a DMSO solution of *p*-amino-L-phenyllactyl-pdCpA (180 μL, 0.4 μmol) and aqueous pyridine–HCl (1 M; pH 5.0, 216 μL). After incubation at 37 °C for 1.5 h, a DMSO solution of BODIPYFL-succinimide ester (50 mM; 18 μL, 0.9 μmol) was added, and the resulting solution was further incubated for 1.5 h. The mixture was washed three times with Et_2O (1 mL) to remove unreacted BODIPYFL. The desired product was purified with an analytical scale reversed-phase HPLC (XTerra C18, 2.5 μm, 4.6×20 mm, flow rate: 1.5 mL min^{-1} with a linear gradient of 0–100% MeOH in 0.38% formic acid, over 10 min) to afford *p*-(BODIPYFL-amino)-L-phenyllactyl-pdCpA (149 nmol, 37% yield). The product was identified by ESI-TOF MS. *p*-(BODIPYFL-amino)-L-phenyllactyl-pdCpA; calcd: 1072.2739; found: 1072.2775 $[M-H]^-$.

Preparation of acyl-tRNA: Acylated tRNA with *p*-(BODIPYFL-amino)-L-phenyllactic acid was synthesized by the chemical aminoacylation method. A yeast phenylalanine tRNA that contained a CCCG anticodon and lacked the 3′-terminal dinucleotide was synthesized as described previously.[4] Ligation of the truncated tRNA and *p*-(BODIPYFL-amino)-L-phenyllactyl-pdCpA was carried out in a reaction mixture (10 μL) that contained Hepes–Na (pH 7.5, 5.5 mM), ATP (1 mM), $MgCl_2$ (15 mM), DTT (3.3 mM), BSA (2 μg mL^{-1}), tRNA (0.25 nmol), *p*-(BODIPYFL-amino)-L-phenyllactyl-pdCpA (2.2 nmol) in DMSO (1 μL), and T4 RNA ligase (30 units). The mixture was incubated at 4 °C for 2 h. After the incubation, KOAc (pH 4.5) was added to a final concentration of 0.3 M. The acyl-tRNA was isolated by extraction with phenol/$CHCl_3$ and $CHCl_3$, and by EtOH precipitation. The precipitate was dissolved in prechilled KOAc (1 mM, pH 4.5) just before addition to the cell-free translation system. The *p*-(BODIPYFL-amino)-L-phenylanalyl-tRNA was also prepared as described previously.[14] The acyl-tRNAs were analyzed by HPLC (Poros R2/10, 4.6×100 mm), flow rate: 1.0 mL min^{-1} with a linear gradient of 0–100% of acetonitrile in aqueous triethylammonium acetate (0.1 M; pH 6.0) over 15 min.[35] The acylation yield was determined to be 80% for BFLAFL and 61% for BFLAF.

Preparation of mRNAs that have various tag peptides: Streptavidin genes fused with His-tag at the C terminus and one of the five types of tag peptide (T7, RGSHis, VSV-G, Avi, and EE; Table 1) at the N terminus, in which the sequence CGGG AGTAAC was inserted between the tag sequence and GAC codon encoding Asp1 of streptavidin, were prepared with RTS *E. coli* linear template generation set by using a streptavidin gene pGSH[4] as a template. For comparison, a streptavidin gene that contained the CGGGAGTAAC sequence just after the initiation codon was also prepared. As 5′ gene-specific primers, 5′-CTTTAAGAAGGAGATATACCATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGAGTAACGACCCGTCCAAGGACTCCAAAGC-3′ (T7), 5′-CTTTAAGAAGGAGATATACCATGAGAGGATCGCATCATCATCATCATCATCGGGAGTAACGACCCGTCCAAGGACTCCAAAGC-3′ (RGSHis), 5′-CTTTAAGAAGGAGATATACCATGGCATACACTGATATCGAAATGAACCGCCTGGGTAAGCGGGAGTAACGACCCGTCCAAGGACTCCAAAGC-3′ (VSV-G), 5′-CTTTAAGAAGGAGATATACCATGGGTCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAACGGGAGTAACGACCCGTCCAAGGACTCCAA-3′ (Avi), 5′-CTTTAAGAAGGAGATATACCATGGAATACATGCCAATGGAACGGGAGTAACGACCCGTCCAAGGACTCCAAAGC-3′ (EE), and 5′-GAAGGAGATATACCATGCGGGAGTAACGACCCGTCC-3′ (without tag) were used. As a 3′ gene-specific primer, 5′-TGATGATGAGAACCCCCCCCTTAGTGGTGGTGGTGGTGGTGTTGCTGAACAGCGTCTAGAG-3′ was used. The mRNAs were prepared by using T7 RNA polymerase as described previously.[4] Insertion and substitution mutants were prepared by using 5′ primers that contained the corresponding mutations. Genes that encoded FABP, calmodulin, and λ-Cro repressor were prepared in a similar manner.

Cell-free translation: The acyl-tRNA and the CGGG-containing tagged mRNA were added to an *E. coli* cell-free translation system. Translation was carried out in a reaction mixture (10 μL) that contained HEPES-KOH (55 mM, pH 7.5), potassium glutamate (210 mM), ammonium acetate (6.9 mM), magnesium acetate (12 mM), DTT (1.7 mM), ATP (1.2 mM), GTP (0.28 mM), phosphoenolpyruvate (26 mM), spermidine (1 mM), polyethyleneglycol-8000 (1.9%), folinic acid (35 μg mL^{-1}), 20 amino acids (0.1 mM of each), mRNA (8 μg), acyl-tRNA (0.2 nmol), and *E. coli* S-30 extract (2 μL). The mixture was incubated at 37 °C for 60 min.

SDS-PAGE and Western blotting: The reaction mixture (1 μL) was mixed with 2×SDS-PAGE sample buffer (10 μL) and water (9 μL), and then the sample (5 μL) was subjected to SDS-PAGE (15%). The fluorescence image of the SDS-PAGE gel was measured with a fluorescence scanner (FMBIO-III, Hitachi Software Engineering) with excitation at 488 nm and emission at 520 nm. The same gel was analyzed by Western blot analysis by using an anti-T7 tag antibody or anti-His tag antibody and alkaline phosphatase-labeled anti-mouse IgG.

MALDI-TOF MS analysis: A cell-free translation reaction mixture for T7-tagged streptavidin that was obtained in the presence of BFLAFL–tRNA (50 μL) was diluted with phosphate buffer (50 μL; 50 mM Na_2HPO_4–NaH_2PO_4, pH 7.0, 150 mM NaCl), and loaded onto MagneHis Ni-Particles (20 μL). The beads were washed five times with wash buffer (200 μL; 50 mM Na_2HPO_4–NaH_2PO_4, pH 7.0, 300 mM NaCl, 5 mM imidazole), and the His-tagged protein was eluted with elution buffer (25 μL; 50 mM Na_2HPO_4-NaH_2PO_4, pH 7.0, 300 mM NaCl, 500 mM imidazole). The eluate was diluted with phosphate buffer (25 μL) and desalted, and concentrated by using ZipTipC18. The proteins were eluted with a matrix solution that contained 2,5-dihydroxybenzoic acid (10 mg mL^{-1}) in acetonitrile/0.1% TFA (2:3). The mass measurement was performed by MALDI TOF-MS (Voyager DE-Pro, Applied Biosystems) in the positive mode by using apomyoglobin as an external calibrant (16952.6 for MH^+).

Immunoprecipitation analysis: The T7-tagged FABP that contained either BFLAF or BFLAFL was immunoprecipitated by using a GrabIt T7-Tag Kit. A cell-free translation reaction mixture (10 μL) was loaded onto GrabIt T7-tag antibody agarose (10 μL), and mixed at room temperature for 1 h. The supernatant was collected as the unbound protein fraction. The agarose was washed with

wash buffer (200 μL), and heated at 80 °C for 5 min in 1×SDS sample buffer (50 μL) to elute bound protein. The supernatant and eluate were applied to SDS-PAGE and analyzed by fluorescence imaging.

Acknowledgements

We thank Dr. Kenzo Fujimoto for measurements by MALDI-TOF MS. This work was partially supported by a Grant-in-Aid of Exploratory Research (17651122) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Keywords: cell-free translation · fluorescent probes · four-base codon · hydroxy acid · protein engineering

[1] B. Schuler, L. K. Pannell, *Bioconjugate Chem.* **2002**, *13*, 1039–1043.
[2] C. Ludwig, M. Pfeiff, U. Linne, H. D. Mootz, *Angew. Chem.* **2006**, *118*, 5343–5347; *Angew. Chem. Int. Ed.* **2006**, *45*, 5218–5221.
[3] V. Muralidharan, T. W. Muir, *Nat. Methods* **2006**, *3*, 429–438.
[4] T. Hohsaka, D. Kajihara, Y. Ashizuka, H. Murakami, M. Sisido, *J. Am. Chem. Soc.* **1999**, *121*, 34–40.
[5] T. Hohsaka, N. Muranaka, C. Komiyama, K. Matsui, S. Takaura, R. Abe, H. Murakami, M. Sisido, *FEBS Lett.* **2004**, *560*, 173–177.
[6] V. W. Cornish, D. R. Benson, C. A. Altenbach, K. Hideg, W. L. Hubbell, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2910–2914.
[7] G. Turcatti, K. Nemeth, M. D. Edgerton, U. Meseth, F. Talabot, M. Peitsch, J. Knowles, H. Vogel, A. Chollet, *J. Biol. Chem.* **1996**, *271*, 19991–19998.
[8] B. E. Cohen, T. B. McAnaney, E. S. Park, Y. N. Jan, S. G. Boxer, L. Y. Jan, *Science* **2002**, *296*, 1700–1703.
[9] R. D. Anderson, III, J. Zhou, S. M. Hecht, *J. Am. Chem. Soc.* **2002**, *124*, 9674–9675.
[10] A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603–609.
[11] N. Budisa, *Angew. Chem.* **2004**, *116*, 6586–6624; *Angew. Chem. Int. Ed.* **2004**, *43*, 6426–6463.
[12] T. L. Hendrickson, V. de Crécy-Lagard, P. Schimmel, *Annu. Rev. Biochem.* **2004**, *73*, 147–176.
[13] L. Wang, P. G. Schultz, *Angew. Chem.* **2005**, *117*, 34–68; *Angew. Chem. Int. Ed.* **2005**, *44*, 34–66.
[14] D. Kajihara, R. Abe, I. Iijima, C. Komiyama, M. Sisido, T. Hohsaka, *Nat. Methods* **2006**, *3*, 923–929.
[15] D. S. Waugh, *Trends Biotechnol.* **2005**, *23*, 316–320.
[16] J. D. Bain, E. S. Diala, C. G. Glabe, D. A. Wacker, M. H. Lyttle, T. A. Dix, A. R. Chamberlin, *Biochemistry* **1991**, *30*, 5411–5421.
[17] J. A. Ellman, D. Mendel, P. G. Schultz, *Science* **1992**, *255*, 197–200.
[18] H. H. Chung, D. R. Benson, P. G. Schultz, *Science* **1993**, *259*, 806–809.
[19] J. T. Koh, V. W. Comish, P. G. Schultz, *Biochemistry* **1997**, *36*, 11314–11322.
[20] E. Chapman, J. S. Thorson, P. G. Schultz, *J. Am. Chem. Soc.* **1997**, *119*, 7151–7152.
[21] I. Shin, A. Y. Ting, P. G. Schultz, *J. Am. Chem. Soc.* **1997**, *119*, 12667–12668.
[22] P. M. England, H. A. Lester, D. A. Dougherty, *Biochemistry* **1999**, *38*, 14409–14415.
[23] M. C. T. Hartman, K. Josephson, C. W. Lin, J. W. Szostak, *PLoS ONE* **2007**, *2*, e972.
[24] S. M. Hecht, B. L. Alford, Y. Kuroda, S. Kitano, *J. Biol. Chem.* **1978**, *253*, 4517–4520.
[25] F. W. Studier, A. H. Rosenberg, J. J. Dunn, J. W. Dubendorff, *Methods Enzymol.* **1990**, *185*, 60–89.
[26] J. Crowe, H. Dobeli, R. Gentz, E. Hochuli, D. Stuber, K. Henco, *Methods Mol. Biol.* **1994**, *31*, 371–387.
[27] T. E. Kreis, *EMBO J.* **1986**, *5*, 931–941.
[28] D. Beckett, E. Kovaleva, P. J. Schaltz, *Protein Sci.* **1999**, *8*, 921–929.
[29] T. Grussenmeyer, K. H. Scheidtmann, M. Hutchinson, W. Eckhart, G. Walter, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 7952–7954.
[30] T. Hohsaka, Y. Ashizuka, H. Murakami, M. Sisido, *J. Am. Chem. Soc.* **1996**, *118*, 9778–9779.
[31] M. Q. Xu, T. C. Evans Jr. , *Curr. Opin. Biotechnol.* **2005**, *16*, 440–446.
[32] Y. Shao, M. Q. Xu, H. Paulus, *Biochemistry* **1995**, *34*, 10844–10850.
[33] T. Watanabe, N. Muranaka, I. Iijima, T. Hohsaka, *Biochem. Biophys. Res. Commun.* **2007**, *361*, 794–799.
[34] M. Hoffmann, C. Nemetz, K. Madin, B. Buchberger, *Biotechnol. Annu. Rev.* **2004**, *10*, 1–30.
[35] S. Gite, S. Mamaev, J. Olejnik, K. Rothschild, *Anal. Biochem.* **2000**, *279*, 218–225.
[36] W. Kudlicki, O. W. Odom, G. Kramer, B. Hardesty, *J. Mol. Biol.* **1994**, *244*, 319–331.
[37] S. Mamaev, J. Olejnik, E. K. Olejnik, K. J. Rothschild, *Anal. Biochem.* **2004**, *326*, 25–32.
[38] N. Muranaka, M. Miura, H. Taira, T. Hohsaka, *ChemBioChem* **2007**, *8*, 1650–1653.

Received: September 28, 2007
Published online on April 16, 2008

