RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Genetic incorporation of D-amino acids into green fluorescent protein based on polysubstrate specificity

Hairong Ma[1,2], Nan Liu[1,2] Shaobo Shi[1], Shuzhen Wang[1] and Yijun Chen[1,*]

[1]State Key Laboratory of Natural Medicines and Laboratory of Chemical Biology, China Pharmaceutical University

24 Tongjia Street, Nanjing, Jiangsu Province, 210009, People's Republic of China

[2] These authors contributed equally to this work.

*Corresponding Author: Yijun Chen.

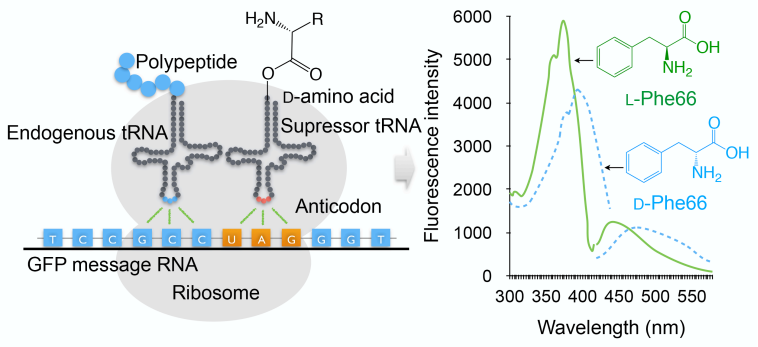
Fax: 86-25-83271031.

Tel: 86-25-83271045.

E-mail: yjchen@cpu.edu.cn

Table of Contents

A number of D-amino acids were genetically incorporated into green fluorescent protein, and the GFPuv mutant containing D-phenylalanine in the fluorophore at residue 66 was characterized.



Abstract

D-amino acids are widely distributed in living organisms and suggested to play important roles in protein folding and function. Genetic incorporation of D-amino acids through protein synthesis machinery has been an exploratory task in protein engineering with great enticement and difficulty. In the present work, a number of D-amino acids were genetically incorporated into green fluorescent protein using a polysubstrate specific tRNA synthetase with cognate tRNA in *Escherichia coli*, and the GFPuv mutant containing D-phenylalanine in the fluorophore at residue 66 was characterized. Stereochemical switch of phenylalanine at position 66 resulted in red shifts in emission and excitation maxima and significantly improved thermal stability of the protein. Molecular modeling further revealed that opposite configurations of Phe66 in GFPuv produced two respective isomeric fluorophores to exhibit distinctive spectral properties and thermal stability. The present study expands the backbone stereochemistry of protein molecules by *in vivo* ribosomal translation to facilitate protein engineering.

Introduction

The perpetuation of homochirality in macromolecular world is likely to be essential for cell survival. However, extensive distribution of functional D-amino acids (D-AAs) in various living organisms has been documented, such as the resistance against proteolytic enzymes, osmoregulation, hormone synthesis and neurotransmission[1-4]. It has been a challenging task for ribosomal incorporation of D-AAs into proteins, both *in vitro* and *in vivo*, owing to the multiple checkpoints that favorably allow L-amino acids (L-AAs) to be incorporated during the translation process. These checkpoints and major limiting factors include aminoacyl-tRNA synthetase, elongation factor Tu, the ribosome and D-aminoacyl-tRNA deacylase[5-8].

Despite these restrictions, the formation of D-aminoacyl-tRNAs by corresponding aminoacyl-tRNA synthetases has

been succeeded *in vitro*[9-11], indicating that the specificity of aminoacyl-tRNA synthetases and the ribosome toward L-AAs for protein synthesis are not absolute. D-AAs can bind both in the peptidyltransferase acceptor and donor sites of the ribosome and compete with respective L-isomers[12]. Synthesis of dipeptides *in vitro* has shown that some D-AAs, such as D-Phe and D-Tyr, could take part in peptide bond formation on the ribosome with lower efficiency than their L-isomers[5]. Furthermore, not all D-AAs suffer from D-Tyr-tRNA^{Tyr} deacylase, an editing enzyme that removes D-Tyr and other D-AAs from charged tRNAs in *E. coli* and yeast. Among 19 D-AAs, only D-Tyr-tRNAs, D-Trp-tRNAs, D-Asp-tRNAs, D-Ser-tRNAs, D-Glu-tRNAs are deacylated by D-Tyr-tRNA^{Tyr} deacylase in *E. coli*[13,14]. Thus, D-AAs might be more favored by tRNAs in cells than we have ever considered.

Thus far, extensive efforts have been made to explore specific mechanism for the synthesis of proteins containing D-AAs. To permit the incorporation of D-AAs into proteins in *in vitro* protein-synthesis systems, Dedkova *et al.* modified the ribosome by altering the peptidyltransferase center and helix 89 of 23S rRNA, and then D-Phe and D-Met were incorporated into target proteins to replace five L-amino acid residues[15,16]. Similarly, 19 D-AAs were charged onto initiator tRNA$^{fMet}_{CAU}$ to test the initiation in a reconstituted *E. coli* translation system lacking methionine[17]. Additionally, preacetylation of D-AA-tRNA$^{fMet}_{CAU}$ and modification of the ribosome dramatically increased expression level *in vitro*[15,17]. However, all these attempts have only been achieved in cell-free protein synthesis rather than in cells. To date, successful incorporation of D-AAs in cells is only limited to the incorporation of D-Lys into a reductase using lysyl-tRNA synthetase in our previous study[18].

Previous studies in the field indicated that several tRNA synthetases can aminoacylate multiple unnatural amino acids with various functional groups to exhibit polysubstrate specificity. These polysubstrate specific tRNA synthetases, such as pyrrolysyl-tRNA synthetase from *Methanosarcinamazei* and phenylalanyl-tRNA synthetase from *Thermus thermophiles,* typically possess a large pocket for amino acid binding, recognizing natural amino acid derivative as well as their

analogues[19,20]. Wild type pyrrolysyl-tRNA synthetase was able to recognize substrates with a variety of α-substitutedside chain including a D-amino acid (D-Boc-lysine) for aminocylation[21], suggesting that the polysubstrate specificity of this tRNA synthetase could provide an opportunity to aminoacylate D-AAs for the incorporation of D-AAs in cells.

In particular, *p*-cyanophenylalanine specific aminoacyl-tRNA synthetase (pCNFRS) derived from *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase together with its cognate amber nonsense suppressor tRNA was reported to exhibit high incorporation efficiency and substrate diversity compared to other aminoacyl-tRNA synthetases[22,23]. Such a polysubstrate specificity allowed to incorporate dozens of unnatural L-AAs, and maintained the ability to discriminate 20 canonical amino acids. On the basis of these unique characteristics, pCNFRS was chosen for screening to explore the possibility of incorporating D-AAs in the present work. Consequently, eight D-AAs were successfully incorporated into green fluorescence protein (GFP) in *E. coli*. Furthermore, aromatic D-Phe was used to substitute L-Tyr in the fluorophore of GFP to investigate the functional changes, including fluorescent spectra and protein stability.

Results and discussion

Substrate permissivity of pCNFRS/tRNA for D-AAs. Taking the advantages of orthogonal pCNFRS/tRNA pair on its polysubstrate specificity, we screened all 19 D-AAs for their possible incorporation into GFPuv, a widely used variant of green fluorescent protein[24]. First, the plasmid pCNFRSII-tRNA containing two copies of pCNFRS gene and a copy of suppressor tRNA gene was constructed similar to our previous work[25]. Then, the construct together with pET-GFPuv-S2TAG in which the codon encoding Serine residue at position 2 in the original sequence of GFPuv is changed to a stop codon TAG was transformed to *E. coli*BL21 (DE3) cells. The plasmid pET-GFPuv-S2TAG carries an IPTG-inducible His-tag fused GFPuv with an amber mutation to replace Ser18 residue (residue 2 in the original sequence of GFPuv) that is located at N-terminus without alteration of overall structure[26]. The transformed BL21(DE3) cells were used to examine the recognition of pCNFRS/tRNA toward 19 D-AAs through the expression of GFPuv in M9 media. During the

course of culturing, we found that the *E. coli* cells did not grow at all in the presence of 1 mM D-Glu. This might be due to the interference to glutamate synthesis by inhibiting glutamate dehydrogenase[27]. For bacterial grown in the media containing other D-AAs, the wet weight from 200 mL broth had no significant difference (ranging 2.5-3.0g), suggesting that 1 mM of the rest 18 D-AAs did not have negative impact on cell growth. Subsequently, these 18 D-AAs were investigated for the possibility of being incorporated into GFPuv. After induction, cell lysates from 200 mL cultures for each D-AA were heated at 65°C for 10 min, and then purified by Ni-NTA affinity chromatography. The yields of purified D-AA-containing proteins ranged from 0 to 2.25mg·L^{-1} cultures. After diluting to a final concentration of 0-40 μg·mL^{-1}, the relative fluorescence intensity to represent respective GFPuv expression level is shown in Figure 1a. To rule out the interference by background fluorescence, the same BL21 (DE3) cells grown in M9 media without addition of D-AAs were used as a control. Mutants containing D-Asn, D-Ile, D-Met, D-Arg, D-Phe, D-Ala, D-Pro and D-Val displayed remarkable increase of fluorescence intensity whereas others showed low intensity with certain degree of variation. In addition, the expression of GFPuv containing these 8 D-AAs at residue 18 was clearly seen on SDS-PAGE after purification (Fig. 1b). It is worth to note that, in Miyake-Stoner's work, when pCNFRS/tRNA were used for expression of sfGFP containing UAAs at position 150, very low background fluorescence was detected in the absence of UAA[28]. However, this phenomenon was not observed in the present study (Fig. 1b, lane 10), which could be due to relatively lower yield of D-AA-containing proteins compared to the protein concentrations of 100-1,500 mg·mL^{-1} used in their work. Combination of the strong and consistent fluorescence intensity and SDS-PAGE analysis unambiguously verified that these 8 D-AAs could be well recognized by the orthogonal pair of pCNFRS/tRNA to produce GFPuv mutants. To further confirm the incorporation of D-AAs in GFPuv, the mutant proteins were digested by trypsin, and resulting peptide fragments were analyzed by MALDI-TOF-MS (Supplementary Table 1; Supplementary Fig. 1a-h,). In the case of mutant $S18^{D}R$ containing a D-Arg at residue18, the proteolytic cleavage site remained at Lys19, revealing that trypsin can not recognize the site of D-Arg. The peptide

fragments containing respective D-AA for other 7 D-AAs were observed from MS analyses and are listed in Supplementary Table 1.

Incorporation of D-phenylalanine into the fluorophore of GFPuv. Tyr66 in GFP participates in the formation of the fluorophore, and changes of this residue, particularly the electron donating capability, could alter spectral properties of the protein[26]. In addition, previous studies by conventional mutagenesis indicated that an aromatic residue is necessary at this position for the generation of fluorescence[29]. To exploit the functional changes by replacing this residue with a D-amino acid, we first substituted Tyr66 in GFPuv to L-Phe by conventional mutagenesis, resulting in a mutant GFPuv-Y66F to serve as a control. Meanwhile, D-Phe was used to replace Tyr66 by introducing a TAG codon to encode residue 66 (pET-GFPuv-Y66TAG) and supplementing D-Phe in the cultures during the induction. Although D-Phe (up to 45 mM) had no apparent effects on *E. coli* cell growth[30], we assessed the effects of higher concentrations of D-Phe on the growth of recombinant *E. coli* cells because both inhibition of cell viability and incorporation of D-Phe into GFPuv should be in a dose-dependent manner[31,32]. After transformation of pCNFRSII-tRNA along with pET-GFPuv Y66TAG into BL21 (DE3) cells, the engineered *E. coli* cells grew in M9 media under different concentrations of D-phe, and were induced by 1 mM IPTG, 0.2% (w/v) arabinose at OD_{600} of 0.8 at 30°C. As anticipated, D-Phe did not inhibit the growth of the engineered *E. coli* cells in both adaptive and logarithmic growth phases (Supplementary Fig. 2). This might attribute to the high orthogonality between pCNFRS and tRNA, which can efficiently utilize D-Phe, and ultimately reduce the accumulation of D-Phe in cells to decrease its toxicity. In stable growth phase, higher concentrations of D-Phe had moderate inhibitory effects on cell growth, which could be resulted from the insertion of D-Phe at the stop codon (UAG) during protein synthesis (Supplementary Fig. 2). Next, mutant GFPuv containing D-Phe at residue 66 (GFPuv-Y66^{D}F) was induced and expressed in the engineered *E. coli* cells in M9 media supplemented with 2 mM D-Phe at 30°C. The incorporation of D-Phe at residue 66 was confirmed by tryptic digestion and MALDI-TOF-MS (Supplementary Table 1;Supplementary Fig. 1i).

Characterization of spectral properties and thermal stability of GFPuv-Y66^{D}F. Similar to previous report on the replacement of Tyr66 by L-Phe in wild-type GFP, replacement of Tyr66 by L-Phe in GFPuv shifted both excitation and emission maxima to shorter wavelengths and the fluorescence intensity was distinctly lower than native GFPuv (Fig. 2a) [33]. This indicated that the switch from L-Tyr to L-Phe did not change structural integrity and fluorophore formation. In contrary, according to previous notion that incorporation of D-AAs into proteins would result in misfolding and drastic changes in structure and function of proteins[15,34,35], one would expect to see a fluorescent quench when L-Phe was switched to D-Phe. To our surprise, the mutant GFPuv-Y66^{D}F still exhibited similar strong fluorescence as GFPuv-Y66F, suggesting that opposite configuration at residue 66 had no effects on protein folding and the formation of fluorophore. Nevertheless, the spectral property of the mutant GFPuv-Y66^{D}F did significantly differ from those of GFPuv-Y66F (Fig. 2a). GFPuv-Y66^{D}F displayed an obvious red shift in emission maximum (λ_{em}^{max}=452nm) and excitation maximum (λ_{ex}^{max}=387nm) compared to GFPuv-Y66F (λ_{em}^{max}=430nm, λ_{ex}^{max}=370nm).

The backbone change from the incorporation of D-Phe may change thermal stability of the protein. Therefore, we compared thermal stability of GFPuv, GFPuv-Y66F and GFPuv-Y66^{D}F. The proteins were heated at 60°C, 70°C and 80°C for 80 min and their respective fluorescence intensity was measured. Although three proteins had approximately similar decrease in fluorescence intensity at 60°C as shown in Figure 2b, GFPuv-Y66^{D}F seemed more stable than GFPuv-Y66F and GFPuv at 70°C (Fig. 2c), and further increase of thermal stability was observed at 80°C (Fig. 2d). Therefore, the change of phenylalanine configuration at the fluorophore in GFPuv actually improved thermal stability..

Mechanistic insight of GFPuv-Y66F and GFPuv-Y66^{D}F by molecular modeling. In order to rationalize the changes in spectral properties and thermal stability, we modeled GFPuv-Y66F and GFPuv-Y66^{D}F to compare their architectural differences based on the crystal structure of wild-type GFP[36] as described in the Experimental section. Ramachandran plot statistics of mutant models were evaluated using the RamPageserve. More than 99% of the dihedral

angles of all residues in each mutant were located either in the most favored or in additionally allowed regions (Supplementary Fig. 3). The VERIFY-3D scores ensured good compatibility of the atomic model (3D) with respective amino acid sequences (Supplementary Fig. 4).

From molecular simulation, the structure models showed that the formation of the imidazolinone intermediate and dehydration were completed spontaneous (Fig. 3a), in which the amide nitrogen of Gly67 attacked the carbonyl carbon of Ser65 to promote internal cyclization. However, the hydroxymethyl group of Ser65 in GFPuv-Y66^{D}F was rotated by a large angle relative to that in GFPuv-Y66F, owing to the formation of an amide bond between the amino group of ^{D}Phe66 and the carboxyl group of Ser65 (Fig. 3b-c). On the other hand, no marked distinction to D-Phe itself was observed and the different configuration of Phe66 caused deflection of the main plane of imidazolinone, leading Gly67 far away from Asn121 and closer to Val68 in GFPuv-Y66^{D}F (Supplementary Fig. 5). Additionally, the carbonyl group in imidazolinone interacted with both Arg96 and Gln69 in GFP-Y66^{D}F to form hydrogen bonds (Supplementary Fig. 5). Since the amino acids surrounding the fluorophore play pivotal roles during the maturation process by stabilizing functional groups involved in the catalysis[37], the positional shift of amino acids in the fluorophore changed the form of interactions between the fluorophore and surrounding residues as shown in Supplementary Figure 5.

Regardless of the altered interactions between the fluorophore and surrounding amino acids, the opposite configuration should be the major cause for the changes of spectral properties. Because of different orientations of Phe at residue 66, C4 in the imidazolinone intermediate possessed an opposite configuration (Fig. 4). The α-β bond of Phe66 in two enantiomers were dehydrated and oxidized by molecular oxygen to produce two respective isomeric fluorophores of 4-benzylidene-5-imidazolinone with H_2O_2 as a by-product. As a result, *R, S*-intermediates produced Z- and *E*-isomers correspondingly. Consequently, the emission maxima of *E*-isomer was observed to be red shifted by 22 nm under our experimental condition (in polar solvent) compared to *Z*-isomer. This wavelength dependent *trans* to *cis* photoisomerization

has previously been reported by V. Raj Gopal et al[38]. The fluorescence band observed at longer wavelength region could be assigned to the state of isomer/charge transfer [39]. In this particular case, the change might be due to preferential light absorption and excitation of the *E*-isomer[38]. Taken together, the introduction of D-Phe into the fluorophore could display altered spectral characteristics from the conformational change in the fluorophore.

Despite the existence of D-amino acids in proteins in various living organisms has been known for a long period of time, the biological significance and roles remain exclusively unclear[40,41]. To probe the biological functions of D-AAs in a diverse array of proteins, it is a perquisite to develop appropriate methodologies on the incorporation of D-AAs into proteins. Given that D-AAs can be aminoacylated by a number of natural or evolved aminoacyl-tRNA synthetases, protein synthesis has been mainly utilized to attempt the incorporation of D-AA into proteins *in vitro*, and great success has been accomplished[9,11,42]. Meanwhile, modification of the ribosome has been an alternative approach to incorporate D-AA into proteins, and several successful examples have also been reported[16,17]. However, these methods are only limited to *in vitro* protein synthesis and relatively short protein sequences. Certainly, there is an ultimate need for an *in vivo* system that is able to incorporate D-AAs into various proteins in order to obtain large amount of proteins regardless of the inserting sites and the length of proteins. Because of some major limiting factors, such as the specificity of aminoacyl-tRNA synthetase[6], the competition of elongation factor Tu[5], the recognition by the ribosome[7] and the presence of D-aminoacyl-tRNA deacylase[13], *in vivo* incorporation of D-AAs into protein has been a difficult, if not impossible, task in the field.

Since 2001, incorporation of unnatural amino acids into proteins by utilizing a stop codon to encode a noncanonical amino acid at a specific site with an orthogonal pair of evolved aminoacyl-tRNA synthetase/tRNA has paved a way for genetic incorporation of various unnatural amino acids into proteins[43,44]. This expansion of genetic code has shown great promise on molecular imaging for proteins in live cells and the development of therapeutic proteins[45,46]. With the development of efficient methods on the evolution of aminoacyl-tRNA synthetase specificity[42,47,48], several orthogonal pairs

have been used to acylate suppressor tRNAs or A/P-site substrates to test *in vitro* ability of protein synthesis machinery on accepting backbone analogues with L-configuration, including α-hydroxy acids[49,50], N-methyl amino acids[51,52], α, α-disubstituted amino acids[49], β-amino acids[49,53,54], and D-amino acids[10,49,55,56]. Under the circumstances of cellular process, due to stereochemical difference, D-amino acids, another group of unnatural amino acids, are even more difficult to be inserted into proteins than their L-counterparts in cells given above-stated limiting factors. However, based on the polysubstrate specific nature of the evolved tRNA syntheteases, we speculated that genetic incorporation of D-AAs could be feasible by taking the advantages of polysubstrate specificity of the orthogonal pairs, which was demonstrated in our previous work to introduce D-Lysine into diketoreductase in *E. coli* cells using *Pyrococcus horikoshii* lysyl-tRNA synthetase[18]. Compared to previous study, the present work further expanded the scope of incorporating D-AAs into GFPuv using pCNFRS/tRNA pair. As a result, 8 types of D-AAs were inserted to GFPuv and experimentally verified by MS analyses, and L-Tyr at residue 66 was further replaced to D-Phe by the same strategy to show that the spectral properties and thermal stability were significantly changed after the substitution in the fluorophore-forming cassette, which provides the possibility of using protein synthesis machinery for template encoded synthesis of novel backbone proteins with defined length and composition. However, in the presence of suppressor tRNAs, due possibly to the stereochemical proofreading during the translation process, D-amino acids generally gave much lower yields than those for unnatural amino acids reported by Miyake-Stoner *et al*. using the same pCNFRS *in vitro*[28]. This is consistent with the fact that tRNA synthetase recognizes D-amino acids at very low efficiency compared to L-amino acids[5,21]. Thus, such lowered efficiency requires further improvement to balance all potential factors, such as the interaction of unnatural tRNAs, optimization of aminoacyltRNA synthetase, reduction or knock-out of release factors and the interaction between EF-Tu and ribosome[57-61]. It is conceivable to expect that D-AAs could be more efficiently incorporated into various proteins by an improved protein translation system in cells.

Experimental

Materials. Primers used in site-directed mutagenesis were synthesized by Invitrogen Bio. Inc., (Shanghai, China). Mutations were confirmed by DNA sequencing with an ABI Genetic Analyzer 3730 (Invitrogen Bio. Inc., Shanghai, China). *Escherichia coli* strains DH5α and BL21 (DE3) were obtained from Tiangen Biotech Co., Ltd. (Beijing, China). The Axyprep™ Plasmid Miniprep kit was from Axygene Biotech Ltd. (USA). The pGFPuv plasmid was a gift from Professor Rongyue Cao at China Pharmaceutical University, China. The pEVOL plasmid containing *ONBYRS* gene was a gift from Professor Peter. G. Schultz at Scripps Research Institute (La Jolla, CA, USA). All D-AAs were purchased from Qiude Biochemical Engineering Co., Ltd. (Shanghai, China). Isopropyl-β-D-thiogalactoside (IPTG) was purchased from Sigma Chemical Co. (St. Louis, USA). D-arabinose was from Nanjing SunShine Biotechnology Co., Ltd. (Nanjing, China).

Plasmid construction. The *GFPuv* gene was cloned from plasmid pGFPuv. The *GFPuv*, *GFPuv-Y66F*, *GFPuv-S2TAG*, *GFPuv-Y66TAG* genes were obtained by QuikChange (Stratagene) mutagenesis following manufacture's procedures. Primers used in this work were listed in Supplementary Table 2. The genes to encode GFPuv, GFPuv-Y66F, GFPuv-S2TAG, GFPuv-Y66TAG were then ligated to plasmid pETDuet-1 using cloning sites of *EcoR* I and *Hind* III, and the resulting plasmids are pET-GFPuv, pET-GFPuv-Y66F, pET-GFPuv-S18TAG, pET-GFPuv-Y66TAG. A 6 × His-tag in pETDuet-1 was fused to the N-terminus of the mutated *GFPuv* gene in frame. The entire *GFPuv* gene is under the control of *lac* operon.

Protein expression. BL21 (DE3) cells containing pET-GFPuv and pET-GFPuv-Y66F were grown in 5 mL LB media supplemented with 100 $\mu g \cdot mL^{-1}$ ampicillin at 37°C overnight. The 5 mL culture was transferred to 200 mL LB media in the presence of ampicillin and grown at 37°C to an OD_{600} of 0.8. Protein expression was then induced by addition of 1 mM IPTG at 30°C for 12h. Cells were harvested by centrifugation at 5,000×g for 10 min, and then resuspended in 50 mMTris-HCl buffer (pH8.0). The cell suspensions were disrupted through a High Pressure Disruptor (Constant Systems

Ltd., UK). Cell-free extracts were obtained by centrifugation at 15,000×g for 30 min at 4°C. Proteins were purified by Ni-NTA affinity chromatography under native conditions. After equilibrating with 50 mMTris-HCl buffer (pH8.0), the bound proteins were washed with the same buffer containing 20 mM imidozole, and then eluted with the same buffer containing 100 mM imidozole. Purified proteins were then exchanged into 50 mM Tris-HCl buffer (pH 8.0) with a Hi-Trap desalting column (GE Healthcare Life Sciences, USA) and stored at 4°C.

To express GFPuv containing D-AAs, *E. coli* BL21 (DE3) cells co-transformed with pCNFRSII-tRNA and pET-GFPuv-S2TAG or pET-GFPuv-Y66TAG were grown in M9 media containing 1% (v/v) glycerol in the presence of 34 $\mu g \cdot mL^{-1}$ chloramphenicol, 100 $\mu g \cdot mL^{-1}$ampicillin and 1 mM respective D-amino acid at 37°C. When cells reached an OD_{600} of 0.8, 1 mM IPTG and 0.2% (w/v) arabinose were added to induce protein expression at 30°C. After 30 h, cells were pelleted and disrupted, and the cell-free extracts in 50 mM Tris-HCl buffer (pH8.0) were purified by Ni-NTA affinity chromatography as described above. Next, the partially purified sample was applied onto a MonoQ-Sepharose (5/50 GL, GE Healthcare Life Sciences, USA) equilibrated with 50mM Tris-HCl buffer (pH8.0). The bound proteins were eluted with a gradient of 0-1.0 M NaCl in 30 mL of 50mM Tris-HCl buffer (pH 8.0), and the fractions containing the desired mutants were collected and concentrated. Purified proteins were exchanged into assay buffer (50 mM Tris-HCl buffer, pH 8.0) with a Hi-Trap desalting column (GE Healthcare Life Sciences, USA).

Protein concentrations were determined by Bradford assay (BCA kit, Generay Biotech Co., Ltd, Shanghai, China). Protein purity was examined by SDS-PAGE.

Fluorescence spectroscopy and thermal stability assay. Fluorescence spectra of GFP and mutants were determined using Infinite M200PRO Multifunctional Microplate Reader (Tecan, Switzerland) in 50 mM Tris-HCl buffer at pH 8.0. Since the standard curve exhibited a good linear relationship between fluorescence intensity and GFPuv concentration from 1-40 $\mu g \cdot mL^{-1}$(R^2=0.9925). GFP mutants containing D-AA at residue 18 from 200 mL broth were diluted in 50 mM

Tris-HCl buffer (pH 8.0) to the same volume with the final concentrations of 1-40 μg·mL^{-1}. 200 μL of each diluted protein was added into the microplate, and fluorescence emission intensities were detected at 512 nm with an excitation wavelength of 402 nm. Emission spectra were recorded with excitation wavelengths of each fluorescent protein. To examine thermal stability, GFPuv and mutants was diluted in 50 mM Tris-HCl buffer (pH 8.0) with a final concentration of 40 μg·mL^{-1}, and incubated in a water bath at 60°C, 70°C and 80°C. The fluorescence intensity was recorded at various time intervals over 80 min. Fluorescence emission intensity was recorded with excitation wavelengths of each fluorescent protein and expressed as relative intensity after subtracting the background fluorescence from the control. All measurements were repeated for three times, and data are averages with standard deviations.

Molecular modeling. Structure models of GFPuv-Y66F and GFPuv-Y66DF were generated by homology modeling using wild-type GFP (PDB code 1GFL)[36] as a template. Wild-type GFP and GFPuv-Y66F are different on only four amino acid residues at positions 66, 99, 153 and 163. Replacement of D-Phe at residue 66 was generated by Molecular Builder in the Molecular Operating Environment[62,63]. All models were subjected to Amber 99 energy minimization until the RMS of the conjugated gradient was 0.05 kcal·mol^{-1}·Å. Reduced units were used with a time step of 0.0005 ps, and simulation was performed until the potential energy U of atomic system and kinetic energy K of the atoms stabilized. Default values were applied for other parameters. Obtained models were evaluated by RamPageserve[64] and VERIFY-3D[65] for geometry.

Conclusion

In summary, we have demonstrated the site-specific incorporation of eight D-AAs into GFPuv using a polysubstrate specific orthogonal pair of amber suppressor tRNA/pCNFRS. D-Phe was further introduced into the fluorophore of GFPuv to probe spectral properties and thermal stability of the protein previously impossible to obtain. The present study has exemplified a feasible way to investigate structure-function relationship with regard to backbone change for various proteins and expanded the backbone diversity of protein molecules through *in vivo* ribosomal translation.

Acknowledgements

We thank Professor Peter G. Schultz at Scripps Research Institute, USA and Professor R. Cao at China Pharmaceutical University, China, for providing the plasmids of pEVOL-ONBYRS and pGFPuv. This work was supported by the grants of "111 Project" by the Ministry of Education of China and the State Administration of Foreign Experts Affairs of China (No.111-2-07), National Science Foundation of China (No. 81473126), the PAPD of Jiangsu Province, China and the Innovation Fund Project for Graduate Students of Jiangsu Province, China (CXZZ13_0333).

References

1. E. Okuma, K. Watanabe and H. Abe, *Fish Sci.,* 1998, **64**, 606-611.
2. M. Friedman, *J. Agric. Food Chem.*, 1999, **47**, 3457-3479.
3. Y. Nagata, H. Homma, J. A. Lee and K. Imai, *FEBS Lett.*, 1999, **444**, 160-164.
4. N. Ota, T. Shi and J. V. Sweedler, *Amino acids*, 2012, **43**, 1873-1886.
5. T. Yamane, D. Miller and J. Hopfield, *Biochemistry*, 1981, **20**, 7059-7064.
6. A. C. Dock-Bregeon, B. Rees, A. Torres-Larios, G. Bey, J. Caillet and D. Moras, *Mol. Cell*, 2004, **16**, 375-386.
7. I. Agmon, M. Amit, T. Auerbach, A. Bashan, D. Baram, H. Bartels, R. Berisio, I. Greenberg, J. Harms and H. A. Hansen, *FEBS Lett.*, 2004, **567**, 20-26.
8. O. Soutourina, J. Soutourina, S. Blanquet and P. Plateau, *J. Biol. Chem.*, 2004, **279**, 42560-42565.
9. P. M. England, Y. Zhang, D. A. Dougherty and H. A. Lester, *Cell*, 1999, **96**, 89-98.
10. L. M. Dedkova, N. E. Fahmi, S. Y. Golovine and S. M. Hecht, *J. Am. Chem. Soc.*, 2003, **125**, 6616-6617.
11. T. Takayama, T. Ogawa, M. Hidaka, Y. Shimizu, T. Ueda and H. Masaki, *Biosci. Biotechnol. Biochem.*, 2005, **69**, 1040.
12. K. Quiggle, G. Kumar, T. W. Ott, E. K. Ryu and S. Chladek, *Biochemistry*, 1981, **20**, 3480-3485.

13. J. Soutourina, P. Plateau, F. Delort, A. Peirotes and S. Blanquet, *J. Biol. Chem.*, 1999, **274**, 19109-19114.

14. J. Soutourina, P. Plateau and S. Blanquet, *J. Biol. Chem.*, 2000, **275**, 32535-32542.

15. L. M. Dedkova, N. E. Fahmi, S. Y. Golovine and S. M. Hecht, *Biochemistry*, 2006, **45**, 15541-15551.

16. L. M. Dedkova, N. E. Fahmi, S. Y. Golovine and S. M. Hecht, *J. Am. Chem. Soc.*, 2003, **125**, 6616-6617.

17. Y. Goto, H. Murakami and H. Suga, *RNA*, 2008, **14**, 1390-1398.

18. Z. Liu, X. Yang, D. Yi, S. Wang and Y. Chen, *Amino Acids*, 2012, **43**, 2553-2559.

19. Y. S. Wang, X. Fang, A. L. Wallace, B. Wu and W. R. Liu, *J. Am. Chem. Soc.*, 2012, **134**, 2950-2953.

20. L. Stokes, S. J. Miyake-Stoner, J. C. Peeler, D. P. Nguyen, R. P. Hammer and R. A. Mehl, *Mol. BioSyst.*, 2009, **5**, 1032-1038.

21. T. Kobayashi, T. Yanagisawa, K. Sakamoto and S. Yokoyama, *J. Mol. Biol.*, 2009, **385**, 1352-1360.

22. D. D. Young, T. S. Young, M. Jahnz, I. Ahmad, G. Spraggon and P. G. Schultz, *Biochemistry*, 2011, **50**, 1894-1900.

23. D. D. Young, S. Jockush, N. J. Turro and P. G. Schultz, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 7502-7504.

24. A. Crameri, E. Whitehorn, E. Tate, W. Stemmer, A. Crameri, P. Kitts and P. Kitts, *Nat. Biotechnol.*, 1996, **14**, 315-319.

25. H. Ma, X. Yang, Z. Lu, N. Liu and Y. Chen, *PLoS One*, 2014, **9**, e103792.

26. F. Wang, W. Niu, J. Guo and P. G. Schultz, *Angew. Chem. Int. Ed.*, 2012, **51**, 10132-10135.

27. H. Arkin and N. Grossowicz, *J. Gen. Microbiol.*, 1970, **61**, 255-261.

28. S. J. Miyake-Stoner, C. A. Refakis, J. T. Hammill, H. Lusic, J. L. Hazen, A. Deiters and R. A. Mehl, *Biochemistry*, 2010, **49**, 1667-1677.

29. R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509-544.

30. M. Caparros, A. Pisabarro and M. De Pedro, *J. Bacteriol.*, 1992, **174**, 5549-5559.

31. G. Zheng, W. Liu, Y. Gong, H. Yang, B. Yin, J. Zhu, Y. Xie, X. Peng, B. Qiang and J. Yuan, *Biochem. J*, 2009, **417**, 85-94.

32. A. R. Parrish, X. She, Z. Xiang, I. Coin, Z. Shen, S. P. Briggs, A. Dillin and L. Wang, *ACS Chem. Biol.*, 2012, **7**, 1292-1302.

33. B. Cubitt, R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross and R. Y. Tsien, *Trends Biochem. Sci.*, 1995, **20**, 448-455.

34. Y. Demizu, H. Yamashita, N. Yamazaki, Y. Sato, M. Doi, M. Tanaka and M. Kurihara, *J. Org. Chem.*, 2013, **78**, 12106-12113.

35. S. Y. Hong, J. E. Oh and K.-H. Lee, *Biochem. Pharmacol.*, 1999, **58**, 1775-1780.

36. F. Yang, L. G. Moss and G. N. Phillips Jr, *Nat. Biotechnol.*, 1996, **14**, 1246-1251.

37. V. Sample, R. H. Newman and J. Zhang, *Chem. Soc. Rev.*, 2009, **38**, 2852-2864.

38. V. R. Gopal, A. M. Reddy and V. J. Rao, *J. Org. Chem.*, 1995, **60**, 7966-7973.

39. A. B. Naik, L. Naik, V. J. Rao and H. Pal, *Open J. Phy. Chem.*, 2012, **2**, 212-220.

40. L. Bai, S. Sheeley and J. V. Sweedler, *Bioanal. Rev.*, 2009, **1**, 7-24.

41. N. Fujii, Y. Kaji, N. Fujii, T. Nakamura, R. Motoie, Y. Mori and T. Kinouchi, *Chem. Biodivers.*, 2010, **7**, 1389-1397.

42. S. W. Santoro, L. Wang, B. Herberich, D. S. King and P. G. Schultz, *Nat. Biotechnol.*, 2002, **20**, 1044-1048.

43. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413-444.

44. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498-500.

45. J. Y. Kang, D. Kawaguchi, I. Coin, Z. Xiang, D. D. O'Leary, P. A. Slesinger and L. Wang, *Neuron*, 2013, **80**, 358-370.

46. K. C. Schultz, L. Supekova, Y. Ryu, J. Xie, R. Perera and P. G. Schultz, *J. Am. Chem. Soc.*, 2006, **128**, 13984-13985.

47. T. G. Heckler, L. H. Chang, Y. Zama, T. Naka, M. S. Chorghade and S. M. Hecht, *Biochemistry*, 1984, **23**, 1468-1473.

48. J. R. Roesser, C. Xu, R. C. Payne, C. K. Surratt and S. M. Hecht, *Biochemistry*, 1989, **28**, 5185-5195.

49. J. A. Ellman, D. Mendel and P. G. Schultz, *Science*, 1992, **255**, 197-200.

50. P. M. England, H. A. Lester and D. A. Dougherty, *Biochemistry*, 1999, **38**, 14409-14415.

51. C. Merryman and R. Green, *Chem. Biol.*, 2004, **11**, 575-582.

52. A. Frankel, S. W. Millward and R. W. Roberts, *Chem. Biol.*, 2003, **10**, 1043-1050.

53. B. M. Eisenhauer and S. M. Hecht, *Biochemistry*, 2002, **41**, 11472-11478.

54. M. S. Gilreath, H. Roy, T. J. Bullwinkle, A. Katz, W. W. Navarre and M. Ibba, *FEBS Lett.*, 2011, **585**, 3284-3288.

55. J. D. Bain, E. S. Diala, C. G. Glabe, D. A. Wacker, M. H. Lyttle, T. A. Dix and A. R. Chamberlin, *Biochemistry*, 1991, **30**, 5411-5421.

56. S. R. Starck, X. Qi, B. N. Olsen and R. W. Roberts, *J. Am. Chem. Soc.*, 2003, **125**, 8090-8091.

57. K. Wang, H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Biotechnol.*, 2007, **25**, 770-777.

58. Y. Ryu and P. G. Schultz, *Nat. Methods*, 2006, **3**, 263-265.

59. W. An and J. W. Chin, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 8477-8482.

60. Y. Doi, T. Ohtsuki, Y. Shimizu, T. Ueda and M. Sisido, *J. Am. Chem. Soc.*, 2007, **129**, 14458-14462.

61. B. Johnson, J. Xu, Z. Shen, J. K. Takimoto, M. D. Schultz, R. J. Schmitz, Z. Xiang, J. R. Ecker, S. P. Briggs and L. Wang, *Nat. Chem. Biol.*, 2011, **7**, 779-786.

62. P. Anand, A. Schug and W. Wenzel, *FEBS Lett.*, 2013, **587**, 3231-3235.

63. M. Banerjee and T. Chen, *Biochem. Pharmacol.*, 2013, **86**, 824-835.

64. S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson and D. C. Richardson, *Proteins*, 2003, **50**, 437-450.

65. Eisenberg, R. Lüthy and J. U. Bowie, *Methods Enzymol.*, 1997, **277**, 396.

Figure legends

Figure 1. D-amino acid incorporation into GFPuv at residue 18. (**a**) Fluorescence intensity of GFPuv mutants containing D-AA. Data represent the fluorescence intensity from a mutant after subtracting background emission. The background emission ranging from 4595 to 5514 was from the same treatment of cells in the absence of D-AAs. Eighteen D-AAs are shown as one-letter abbreviation in the axis label. Mutants displaying higher fluorescence intensity are show as bars filled with green lattice. Data are averages of 3 independent experiments with standard deviations. (**b**) SDS-PAGE analysis of GFPuv mutants containing D-AA. Eighteen D-AAs are shown in one-letter abbreviation. Lane Cont. represents the control with the same treatment of cells in the absence of D-AA, and Lane M is protein molecular weight marker.

Figure 2. Comparison of GFPuv, GFPuv-Y66F and GFPuv-Y66^{D}F on spectral change and thermal stability. (**a**) Fluorescent excitation and emission spectra of GFPuv and mutants. (**b**) Thermal stability of the mutants at 60°C. (**c**) Thermal stability of the mutants at 70°C. (**d**) Thermal stability of the mutants at 80°C. Data represent the fluorescence intensity normalized by the mutants without heating. Data are averages of 3 independent experiments with standard deviations.

Figure 3. Structural comparison of GFPuv-Y66F and GFPuv-Y66^{D}F. (**a**) Overlay of the fluorophore of GFPuv-Y66F (Cyan) and GFPuv-Y66^{D}F (Magentas). (**b**) Dihedral angle between C'_{65}-$C_{\alpha 65}$-$C_{\beta 65}$-O_{65} (black dashed line, 60.4°) and N_{66}-C'_{65}-$C_{\alpha 65}$-$C_{\beta 65}$（-0.8°）in the fluorophore of GFPuv-Y66F. (**c**) Dihedral angle between C'_{65}-$C_{\alpha 65}$-$C_{\beta 65}$-O_{65} (-53.4°) and N_{66}-C'_{65}-$C_{\alpha 65}$-$C_{\beta 6}$5（black dashed line, 24.9°）in the fluorophore of GFPuv-Y66^{D}F.

Figure 4. Comparison of proposed mechanism on fluorophore formation between GFPuv-Y66F and GFPuv-Y66^{D}F. Stereochemical properties of L/D-phe at position 66 were highlight in green (primary translation products), blue (products from S65-F66-G67) and red (products from S65-^{D}F66-G67), respectively.

Figures

Figure 1

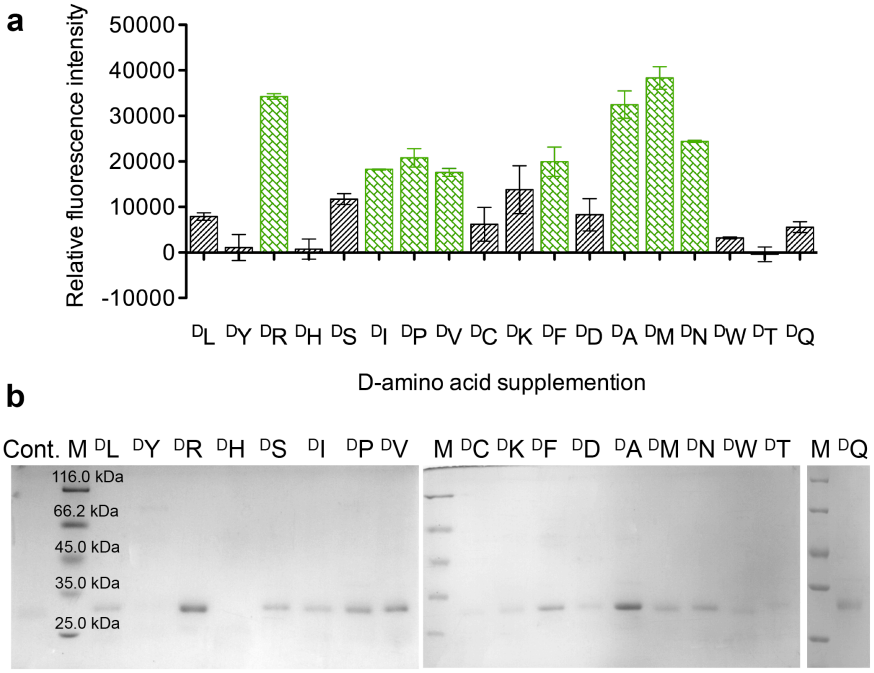


Figure 2

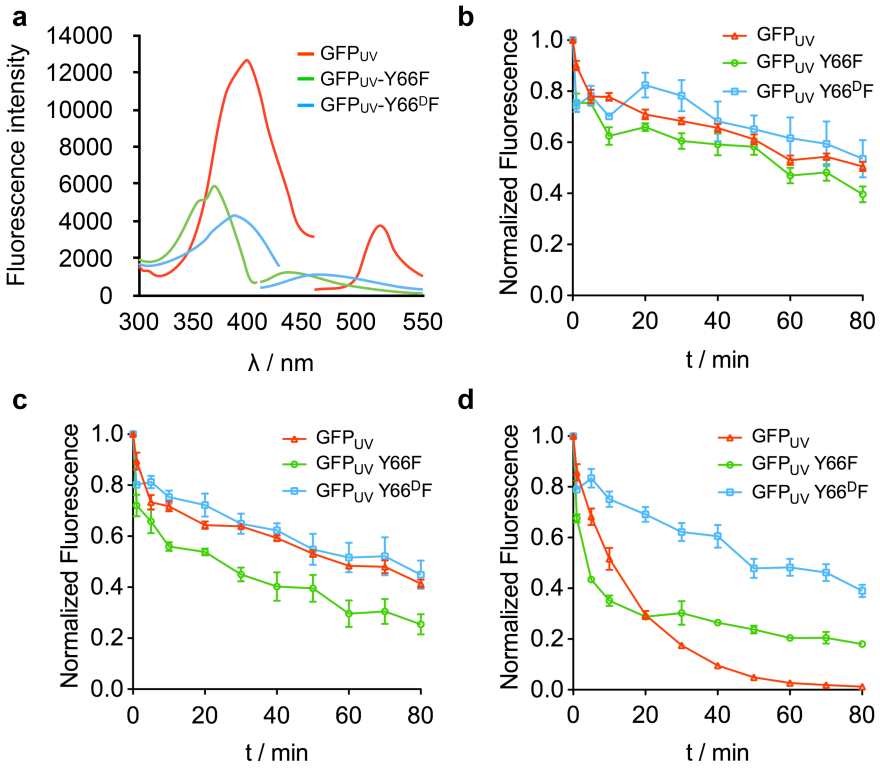


Figure 3

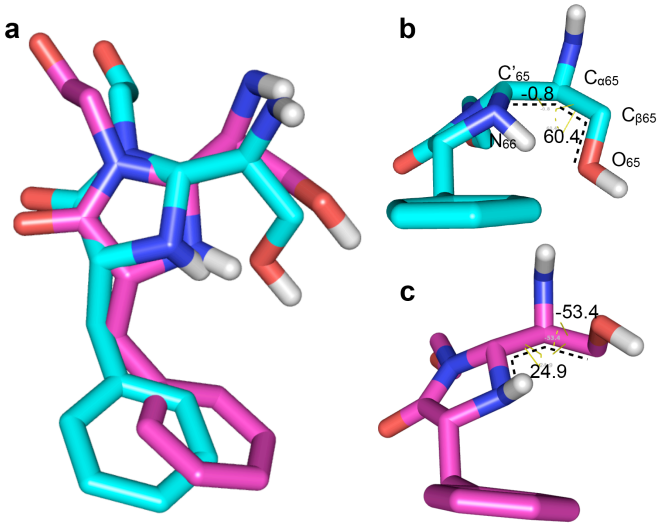


Figure 4