HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2017 December 11.

Published in final edited form as:

J Am Chem Soc. 2015 September 09; 137(35): 11206–11209. doi:10.1021/jacs.5b03135.

Ribosome-Mediated Incorporation of Dipeptides and Dipeptide Analogues into Proteins in Vitro

Rumit Maini, **Larisa M. Dedkova**, **Rakesh Paul**, **Manikandadas M. Madathil**, **Sandipan Roy Chowdhury**, **Shengxi Chen**, and **Sidney M. Hecht***

Center for BioEnergetics, Biodesign Institute, and Department of Chemistry & Biochemistry, Arizona State University, Tempe, Arizona 85287, United States

Abstract

Plasmids containing 23S rRNA randomized at positions 2057−2063 and 2502−2507 were introduced into *Escherichia coli*, affording a library of clones which produced modified ribosomes in addition to the preexisting wild-type ribosomes. These clones were screened with a derivative of puromycin, a natural product which acts as an analogue of the 3′-end of aminoacyl-tRNA and terminates protein synthesis by accepting the growing polypeptide chain, thereby killing bacterial cells. The puromycin derivative in this study contained the dipeptide *p*-methoxyphenylalanylglycine, implying the ability of the modified ribosomes in clones sensitive to this puromycin analogue to recognize dipeptides. Several clones inhibited by the puromycin derivative were used to make S-30 preparations, and some of these were shown to support the incorporation of dipeptides into proteins. The four incorporated species included two dipeptides (Gly-Phe (**2**) and Phe-Gly (**3**)), as well as a thiolated dipeptide analogue (**4**) and a fluorescent oxazole (**5**) having amine and carboxyl groups approximately the same distance apart as in a normal dipeptide. A protein containing both thiolated dipeptide **4** and a 7-methoxycoumarin fluorophore was found to undergo fluorescence quenching. Introduction of the oxazole fluorophore **5** into dihydrofolate reductase or green fluorescent protein resulted in quite strong enhancement of its fluorescence emission, and the basis for this enhancement was studied. The aggregate results demonstrate the feasibility of incorporating dipeptides as a single ribosomal event, and illustrate the lack of recognition of the central peptide bond in the dipeptide, potentially enabling the incorporation of a broad variety of structural analogues.

In recent years, many laboratories have employed ribosome-mediated protein synthesis to produce proteins containing non-proteinogenic amino acids, enabling a more detailed study of protein structure, function, and dynamics.[1] While a broad range of amino acids can now be incorporated into proteins, the incorporation of α-amino acids containing large and

***Corresponding Author**: sidhecht@asu.edu.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b03135.

Synthetic methods, compound characterization for compounds prepared, and additional characterization of the elaborated proteins (PDF)

Notes

The authors declare no competing financial interest.

negatively charged side chains is still challenging,[2] as is protein biosynthesis with non α-amino acids.[3]

Recently, we have described the facilitated incorporation of non-proteinogenic amino acids by the use of modified ribosomes in which key regions of the *Escherichia coli* 23S rRNA have been altered. Initial efforts involved the successful incorporation of D-amino acids[4] and β-amino acids.[5] The strategy employed for the incorporation of β-amino acids was of special interest, involving the use of a puromycin analogue containing a β-amino acid to enable the selection of promising ribosomal architectures at the level of bacterial cells harboring the plasmids with the modified 23S rRNAs.[5a] This study argued that it may well be possible to mediate the incorporation of other types of amino acids using ribosomes selected by the use of puromycin analogues containing the amino acid structures of interest. Presently, we employ a puromycin derivative (**1**) containing the dipeptide *p*-methoxyphenylalanylglycine, and demonstrate that the selected ribosomes successfully incorporated three dipeptides (**2**−**4**, Figure 1), including one dipeptide having a thioamide linkage (**4**), as well as a fluorescent dipeptidomimetic analogue (**5**, Scheme S1) reminiscent of the fluorophores in natural fluorescent proteins.[6]

The bacterial clones employed for screening were part of a previously described library containing modifications in specific regions of 23S rRNA known to be involved in peptide bond formation.[5a,7] These were screened for inhibition by dipeptidylpuromycin **1** (Figure 1), the preparation of which is outlined in Scheme S2. A total of 419 clones were screened, and 13 were inhibited by at least 50% in the presence of 100 μg/mL puromycin derivative **1**. Nine of these clones, containing the sequence 2057UGCGUGG2063 in their 23S rRNA to confer erythromycin resistance, had also been randomized in the region 2502−2507, and are summarized in Table 1. As shown, four of these nine clones proved to be identical, having the sequence 2502ACGAAG2507, and another two shared the sequence 2502CUACAG2507. Clone 010326R6 was chosen for further study based on its inhibition by **1** and the favorable properties of the S-30 system prepared from this clone in cell-free protein synthesis experiments.

As shown in Scheme S3, glycylphenylalanine (**2**) was protected as its *N*-pentenoyl derivative and used to esterify the dinucleotide pdCpA.[8] Bacteriophage T4 RNA ligase was then employed to ligate the dinucleotide to an in vitro RNA transcript, affording glycylphenylalanyl-$tRNA_{CUA}$. Phenylalanylglycine (**3**) was used to prepare phenylalanylglycyl-$tRNA_{CUA}$ analogously (not shown). When utilized in an S-30 system prepared from clone 010326R6, and in the presence of *E. coli* dihydrofolate reductase (DHFR) mRNA containing a UAG codon corresponding to position 10 of DHFR, the dipeptidyl$tRNA_{CUA}$s effected suppression of the UAG codon, producing DHFRs containing **2** (8% suppression) and **3** (14% suppression) at position 10 (Figure 2). Verification of the incorporation of **2** was accomplished by MALDI-MS analysis of a tryptic digest of the elaborated DHFR (Figure S2) which contained an ion at m/z 1377, corresponding to the peptide fragment MISLIAALA**GF**DR, while DHFR containing phenylalanine at position 10 had the analogous ion at m/z 1320, corresponding to MISLIAALA**F**DR. A complete analysis of fragment ions is summarized in Table S1. It may be noted that, apart from clone 010326R6, the S-30 preparations derived from an additional three of the nine unique clones

identified could also incorporate **2** and **3** into DHFR with comparable suppression efficiency (not shown). Further, the DHFR prepared from clone 010326R6 containing **2** at position 10 (Figure 2) was ~84% as active as authentic wild-type DHFR, providing a measure of the fidelity with which this S-30 system was able to incorporate α-amino acids.[9] The lower fidelity of the modified ribosomes is unsurprising, and will likely prove to be general for such ribosomes.

Thiopeptide moieties have been shown to be useful as fluorescence quenchers in peptides and proteins, functioning both by Förster resonance energy transfer (FRET)[10] and by photoinduced electron transfer (PET).[11] This property has been utilized very effectively by the Petersson laboratory for characterization of protein structure following introduction of the requisite thioamide by native chemical ligation of a synthetic peptide fragment containing the thioamide to the remainder of the ribosomally produced protein.[10b,12,13] Since the modified ribosomes described here could plausibly provide an alternative route to proteins containing thioamides at predetermined positions, we incorporated a thiolated dipeptide (**4**) into position 16 of DHFR and the fluorescent probe L-(7-methoxycoumarinyl-4-yl)ethylglycine, which is known to be sensitive to environment,[14] into position 49.[15]

The requisite thiolated dipeptide was obtained as outlined in Scheme S4. Fully protected phenylalanylglycine was converted to the respective thiodipeptide (**24**) by treatment with Lawesson's reagent.[16] Following conversion to the *N*-pentenoyl protected cyanomethyl ester (**26**), condensation with pdCpA afforded the thiolated dipeptidyl-pdCpA ester, the latter of which was converted to thiophenylalanylglycyl-tRNA$_{CCCG}$ via the agency of T4 RNA ligase. In analogy with an earlier study,[17] thiophenylalanylglycyl-tRNA$_{CCCG}$ was used to introduce **4** into position 16 of DHFR, and (7-methoxycoumarin-4-yl)-ethylglycyl-tRNA$_{CUA}$ was used to introduce the fluorophore into position 49. Another modified DHFR was prepared containing phenylalanylglycine at position 16 and the 7-methoxycoumarin fluorophore at position 49. As shown in Figure 3A, when equimolar amounts[18] of the two samples were excited at 310 nm, only the DHFR containing thiophenylalanine at position 16 underwent fluorescence quenching.[19] While the mechanism was not studied in detail, given the distances involved (Figure 3B)[15] it seems likely that this occurred by photoinduced electron transfer.

In addition to the foregoing three dipeptides, dipeptidomimetic analogue **5** was also investigated. This compound lacks the peptide bond which connects the amino acids in dipeptides **2**−**4**, but has a similar distance between the amine and carboxylate groups which participate in incorporation of the dipeptide into the protein backbone. Oxazole **5** is fluorescent, having λ_{ex} at 302 nm, and λ_{em} at 403 nm in water.[20] It was prepared and used to activate tRNA$_{CUA}$ as outlined in Scheme S1. The incorporation of this oxazole within DHFR at position 10 was verified by MALDI-MS of a tryptic digest (Figure S5).[21] Incorporation by the S-30 preparations from two different clones was achieved with 14−15% suppression efficiency (Table 2). It was also found to result in a dramatic increase in the intensity of fluorescence emission and a shift in the emission maximum to ~395 nm, reflecting its sensitivity to its environment (Figure 4). We posit that the inclusion of the oxazole fluorophore within the protein backbone may be sufficient to increase its

fluorescence signal significantly. To test this hypothesis, we incorporated the same oxazole at two positions of green fluorescent protein (GFP) (Figure 5).

In one construct, GFP was rendered nonfluorescent by mutating Tyr66 (which is essential for fluorophore generation) to glycine, and then introducing the oxazole into position 39 of GFP (GFP66Gly39oxazole5), which is outside (but close to) the β-barrel structure in which the GFP fluorophore normally resides.[6] This species emitted at 410 nm (Figure 5) with an intensity 4−5-fold greater than that of wild-type blue fluorescent protein (BFP, chosen for comparison due to its shorter emission wavelength than GFP).

Since the BFP fluorophore is known to be greatly enhanced by its inclusion within the β-barrel structure, this suggests that the inclusion of the oxazole fluorophore within the protein backbone alone is sufficient to increase its fluorescence signal significantly. When the oxazole was introduced into position 66 of GFP in lieu of Tyr66 (GFP66oxazole5), the resulting GFP derivative had fluorescence emission several-fold stronger than BFP or GFP, and had its emission shifted to shorter wavelength, undoubtedly due to its presence within the β-barrel structure. It may be noted that, unlike the typical fluorophores formed in the natural and modified fluorescent proteins, the oxazole fluorophore is completely stable, and requires no preactivation to exhibit fluorescence.

The absorption extinction coefficients and quantum yields of the GFP analogues containing **5** were determined in comparison with that of wild type GFP (Table 3). The GFP analogues had absorption extinction coefficients and quantum yields greater than those of wild-type GFP, and entirely consistent with the fluorescence emission data shown in Figure 5.

As illustrated by the foregoing examples, we anticipate that it will be possible to incorporate a variety of dipeptide analogues into proteins and other biomaterials. The lack of participation of the central amide bond of the dipeptide in the process by which the dipeptide is incorporated implies that it is essentially expendable structurally, as illustrated by analogues **4** and **5**. Thus, the incorporation of dipeptide mimetics can afford structures not accessible by the successive incorporation of two amino acids. Numerous applications, including the creation of libraries of fluorescent proteins having a range of photophysical properties[23] and the metabolic stabilization of proteins of therapeutic interest,[24] can be readily envisioned. While structure **5**, in particular, represents a fairly significant departure from any amino acid whose ribosomal incorporation into protein has been reported previously, it may ultimately prove feasible to select ribosomes capable of incorporating even more complex and potentially useful substrates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health Research Grant GM103861, awarded by the National Institute of General Medical Sciences. We thank Dr. Sriloy Dey for assistance in preparing a synthetic intermediate.

References

1. (a) Hendrickson TL, de Crecy-Lagard V, Schimmel P. Annu Rev Biochem. 2004; 73:147–176. [PubMed: 15189139] (b) Liu CC, Schultz PG. Annu Rev Biochem. 2010; 79:413–444. [PubMed: 20307192]
2. (a) Karginov VA, Mamaev SV, An H, Van Cleve MD, Hecht SM, Komatsoulis GA, Abelson JN. J Am Chem Soc. 1997; 119:8166–8176.(b) Hohsaka T, Kajihara D, Ashizuka Y, Murakami H, Sisido M. J Am Chem Soc. 1999; 121:34–40.(c) Rothman DM, Petersson EJ, Vazquez ME, Brandt GS, Dougherty DA, Imperiali B. J Am Chem Soc. 2005; 127:848–847. [PubMed: 15656618]
3. (a) Bain JD, Diala ES, Glabe CG, Wacker DA, Lyttle MH, Dix TA, Chamberlin AR. Biochemistry. 1991; 30:5411–5421. [PubMed: 2036409] (b) Killian JA, Van Cleve MD, Shayo YF, Hecht SM. J Am Chem Soc. 1998; 120:3032–3042.(c) Eisenhauer BM, Hecht SM. Biochemistry. 2002; 41:11472–11478. [PubMed: 12234190]
4. (a) Dedkova LM, Fahmi NE, Golovine SY, Hecht SM. J Am Chem Soc. 2003; 125:6616–6617. [PubMed: 12769555] (b) Dedkova LM, Fahmi NE, Golovine SY, Hecht SM. Biochemistry. 2006; 45:15541–15551. [PubMed: 17176075]
5. (a) Dedkova LM, Fahmi NE, Paul R, del Rosario M, Zhang L, Chen S, Feder G, Hecht SM. Biochemistry. 2012; 51:401–415. [PubMed: 22145951] (b) Maini R, Nguyen D, Chen S, Dedkova LM, Roy Chowdhury S, Alcala-Torano R, Hecht SM. Bioorg Med Chem. 2013; 21:1088–1096. [PubMed: 23375097] (c) Maini R, Roy Chowdhury S, Dedkova LM, Roy B, Daskalova SM, Paul R, Chen S, Hecht SM. Biochemistry. 2015; 54:3694–3706. [PubMed: 25982410]
6. Tsien RY. Annu Rev Biochem. 1998; 67:509–544. [PubMed: 9759496]
7. Modifications to the 23S rRNA also included one of eight new sequences for nucleotides 2057−2063, introduced to confer moderate resistance to erythromycin in support of the selection protocol.[5a]
8. Robertson SA, Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. Nucleic Acids Res. 1989; 17:9649–9660. [PubMed: 2602139]
9. While the S-30 preparation contained both modified and wild-type ribosomes, only the modified ribosomes can incorporate the dipeptide, and so must be the source of all DHFR polypeptide chains containing the dipeptide. Position 10 of DHFR has no known function,[2a] so incorporation of Gly-Phe at position 10 per se would not be expected to alter DHFR activity. See Table S2 for complete data.
10. (a) Goldberg JM, Batjargal S, Petersson EJ. J Am Chem Soc. 2010; 132:14718–14720. [PubMed: 20886849] (b) Wissner RF, Batjargal S, Fadzen CM, Petersson EJ. J Am Chem Soc. 2013; 135:6529–6540. [PubMed: 23594264]
11. (a) Goldberg JM, Speight LC, Fegley MW, Petersson EJ. J Am Chem Soc. 2012; 134:6088–6091. [PubMed: 22471784] (b) Goldberg JM, Batjargal S, Chen BS, Petersson EJ. J Am Chem Soc. 2013; 135:18651–18658. [PubMed: 24266520]
12. Batjargal S, Wang YJ, Goldberg JM, Wissner RF, Petersson EJ. J Am Chem Soc. 2012; 134:9172–9182. [PubMed: 22468862]
13. Muir TW. Annu Rev Biochem. 2003; 72:249–289. [PubMed: 12626339]
14. Murakami H, Hohsaka T, Ashizuka Y, Hashimoto K, Sisido M. Biomacromolecules. 2000; 1:118–125. (Figures 7−9 in this report employ three amino acids containing the 7-methoxycoumarin fluorophore, each of which was attached to the backbone of streptavidin at position 120 via one of three different linkers. Their fluorescence emission maxima differed by about 20 nm, establishing the environmental sensitivity of this fluorophore.). [PubMed: 11709833]
15. In an earlier study two pyrenylalanines incorporated into these positions were within several angstroms, and close enough to undergo excimer formation. See:Chen S, Wang L, Fahmi NE, Benkovic SJ, Hecht SM. J Am Chem Soc. 2012; 134:18883–18885. and references therein. [PubMed: 23116258]
16. Thomsen I, Clausen K, Scheibye S, Lawesson SO. Org Synth. 1984; 62:158–164.
17. Chen S, Fahmi NE, Wang L, Bhattacharya C, Benkovic SJ, Hecht SM. J Am Chem Soc. 2013; 135:12924–12927. [PubMed: 23941571]

18. Quantification of protein concentrations was carried out by Coomassie Brilliant Blue R-250 staining of bands of the proteins following separation by SDS−PAGE, in comparison with known concentrations of wild-type DHFR (Figure S3).
19. Also prepared in parallel was a modified DHFR containing thiodipeptide 4 at position 10. Analysis of a tryptic digest by MALDI-MS verified incorporation of the thiolated dipeptide, as well as the absence of any detectable exchange of O in lieu of S (Figure S4).
20. For oxazole derivative 12, the molar absorptivity in MeOH was 26 800 M^{-1} cm^{-1} and the quantum yield was 0.59.
21. To date, we have been unable to verify dipeptide incorporation by MS/MS analysis, relying instead on MS data derived from proteolytic fragments. We note that these do not reflect any increase in microheterogeneity occasioned by the use of modified ribosomes.
22. Patterson GH, Knobel SM, Sharif WD, Kain SR, Piston DW. Biophys J. 1997; 73:2782–2790. [PubMed: 9370472]
23. We have recently succeeded in incorporating a fluorescent thiazole amino acid into GFP. Roy Chowdhury et al. submitted.
24. Kamionka M. Curr Pharm Biotechnol. 2011; 12:268–274. [PubMed: 21050165]

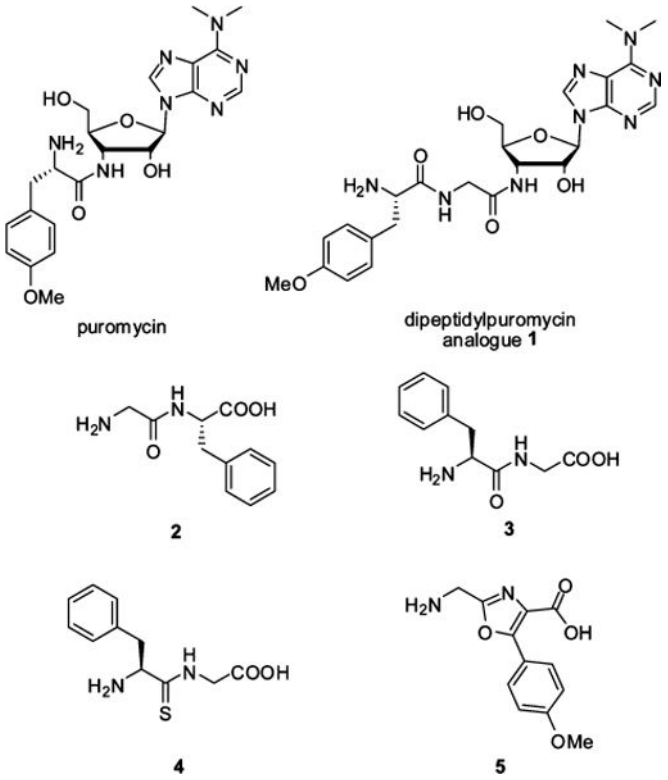

Figure 1.

Puromycin derivative **1** used for selection of modified ribosomes, and dipeptides (**2** and **3**) and dipeptide analogues (**4** and **5**) incorporated into proteins using the modified ribosomes.

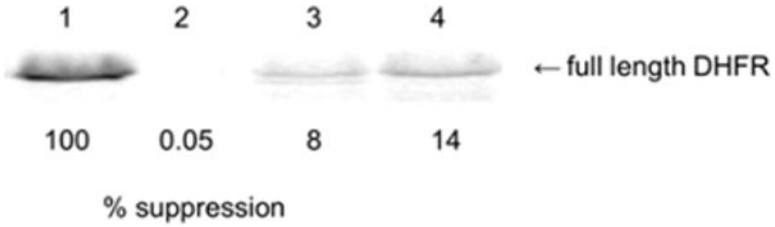

Figure 2.

Autoradiogram of a SDS−polyacrylamide gel showing the translation of DHFR from wild-type (lane 1) and modified (lanes 2−4) (UAG codon in position 10) mRNA in the presence of different suppressor $tRNA_{CUA}$s using an S-30 system prepared from clone 010326R6. Lane 2, non-acylated $tRNA_{CUA}$; lane 3, glycylphenylalanyl-$tRNA_{CUA}$; lane 4, phenylalanylglycyl-$tRNA_{CUA}$. The suppression efficiency relative to wild type is shown below each lane. Additional incorporation data is provided in Figure S1.

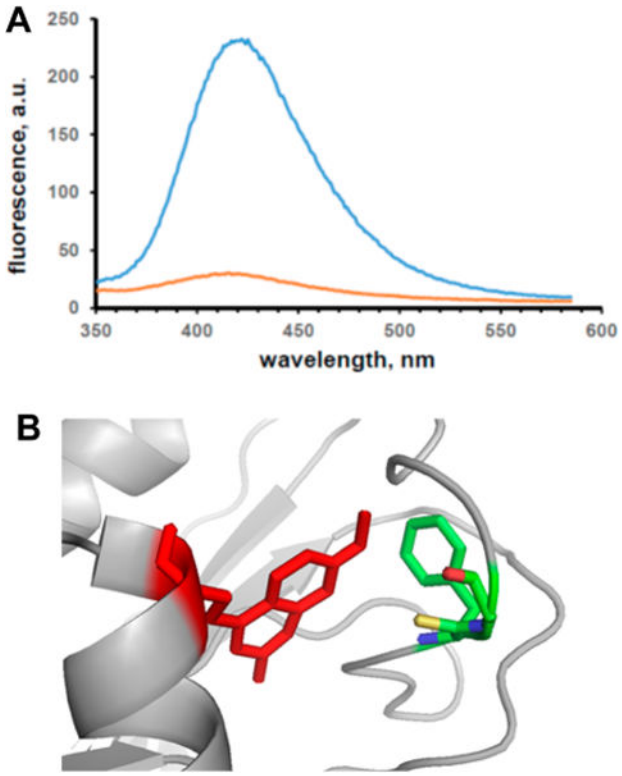

Figure 3.

(A) Fluorescence emission of two samples of DHFR (10 ng/μL), each having 7-methoxycoumarin in position 49 and Phe-Gly (**3**) (blue trace) or thioPhe-Gly (**4**) (orange trace) in position 16 after excitation at 310 nm in 25 mM Tris-HCl, pH 7.4, containing 0.5 M NaCl. (B) Model of the DHFR construct containing the 7-methoxycoumarin fluorophore in position 49 and thio-Phe-Gly (**4**) at position 16, based on DHFR structure PDB 1RA1. The distance from the oxygen of Met16 (corresponding to the S atom of ThioFG in the figure, yellow color) to the α-carbon of Ser49 is 7.7 Å.

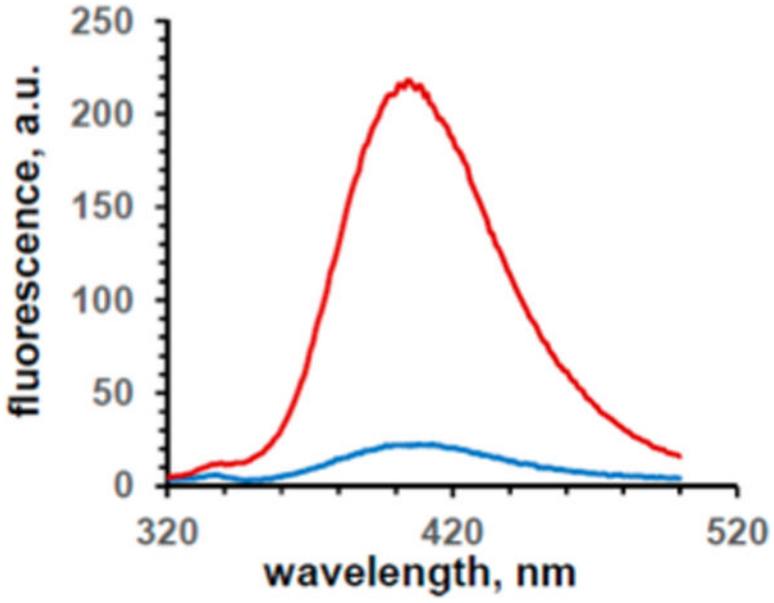

Figure 4.

Fluorescence emission spectra of (blue trace) free oxazole **5** and (red trace) oxazole **5** present at position 10 of DHFR (both were present at 100 nM concentration in 25 mM Tris-HCl, pH 7.4, containing 0.5 M NaCl; excitation at 302 nm). The spectrum of oxazole **5** was obtained on the protected oxazole (**12**) to mimic the form of the oxazole present within full length DHFR. The emission spectrum of oxazole **12** was also recorded at higher and lower concentrations (Figure S6).

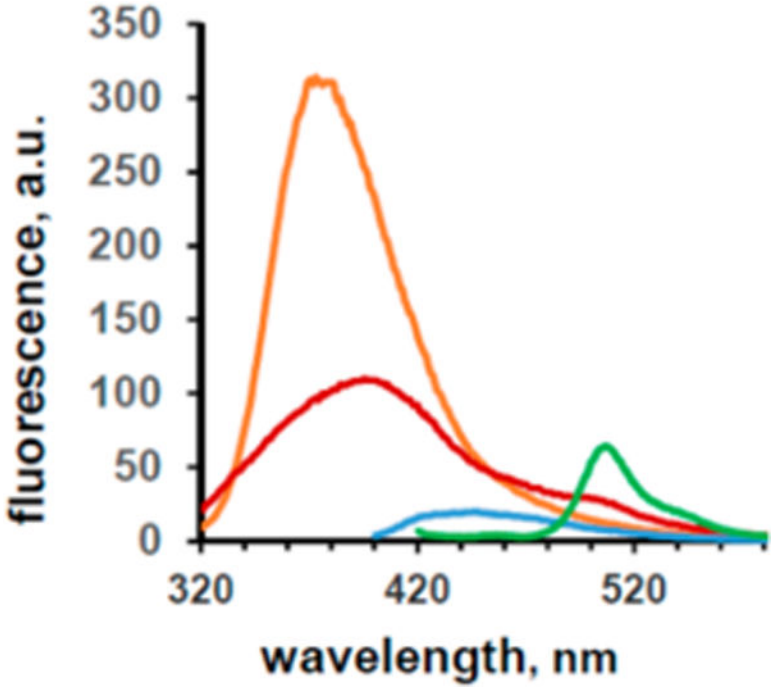

Figure 5.

Fluorescence emission spectra of (orange trace) green fluorescent protein containing oxazole **5** at position 66 (excitation at 310 nm); (red trace) green fluorescent protein containing oxazole **5** at position 39 (excitation at 310 nm); (blue trace) blue fluorescent protein (excitation at 375 nm); (green trace) green fluorescent protein (excitation at 395 nm), all present at 2.5 ng/μL concentration in 25 mM Tris-HCl, pH 7.4, containing 0.5 M NaCl.

Table 1

Sequence in Region 2502−2507 of Clones Having the Same Sequeuce (UGCGUGG) in Region 2057−2063

clones	nucleotide in position					
	2502	2503	2504	2505	2506	2507
010309R3						
010326R6						
010328R4						
010322						
010310R4	C	G	C	A	C	G
010326R5	C	U	A	U	G	U
010310R1	C	G	C	A	A	U
010328R2						
010326R1						

J Am Chem Soc. Author manuscript; available in PMC 2017 December 11.

Table 2

Incorporation of Oxazole 5 into Position 10 of E. coli DHFR by the Use of S-30 Systems Having Different Modified Ribosomes

sequence in 23S rRNA of modified ribosomes		suppression efficiency in different S-30 systems having modified ribosomes (%)	
region 1	region 2	amino acid	
(2057–2063)	(2502–2507)	–	5
UGCGUGG	ACGAAG	0.8 ± 0.2	15.3 ± 2.5
AGUGAGA	AUCCGA	2.2 ± 1.0	14.0 ± 3.0

Table 3

Estimation of Quantum Yields of Modified Green Fluorescent Proteins in Comparison with GFPwt

protein	λ_{ex}/abs, nm	abs extinction coeff, M^{-1} cm^{-1}	quantum yield, λ_{em}
GFPwt	395	25,000	0.79 (509 nm)[22]
GFP66oxazole5	310	90,300	0.91 (378 nm)
GFP66Gly39oxazole5	310	50,200	0.84 (407 nm)