Synthesis of Proteins Containing Modified Arginine Residues†

Ambar K. Choudhury, Serguei Y. Golovine, Larisa M. Dedkova, and Sidney M. Hecht*

Departments of Chemistry and Biology, University of Virginia, Charlottesville, Virginia 22904-4319

Received September 30, 2006; Revised Manuscript Received December 2, 2006

ABSTRACT: Unnatural amino acid mutagenesis provides the wherewithal to study protein function in great detail. To extend the repertoire of functionalized amino acids available for study by this technique, seven structural analogues of arginine were prepared and used to activate a suppressor $tRNA_{CUA}$. These included N^{γ}-methylarginine, N^{γ}-nitroarginine, citrulline, homoarginine, and three conformationally constrained analogues based on proline. These misacylated tRNAs were shown to be capable of introducing the arginine analogues into dihydrofolate reductase (position 22) and *Photinus pyralis* luciferase (positions 218 and 437). Most of the modified luciferases containing arginine analogues at position 218 emitted light with less efficiency and at longer wavelength than the wild type. This is consistent with the postulated role of this residue as essential for maintaining the polarity and rigidity of the luciferin-binding site. Interestingly, the luciferase containing N^{γ}-methylarginine at position 218 emitted light at the same wavelength as the wild type and was at least as efficient. Alteration of the arginine residue at position 437 had no effect on the wavelength of emitted light but afforded analogues, all of which emitted light less efficiently than the wild type. This is altogether consistent with the putative role of Arg437, which is an invariant residue within the superfamily of enzymes that includes *P. pyralis* luciferase. This amino acid is part of the linker between the two structural domains of luciferase that is believed to be essential for efficient enzyme function but not part of the substrate-binding site.

Site-directed mutagenesis of proteins permits the replacement of a specific amino acid in a protein by any of the other proteinogenic amino acids and constitutes a powerful technique for determining the role of individual amino acids in protein function (*1*−*4*). In recent years, the ability to purposefully misacylate tRNAs with unnatural amino acids (*5*−*7*) and to insert those unnatural amino acids into specific, predetermined positions in proteins (*8*, *9*) has dramatically extended the repertoire of amino acid side chains that can be introduced into a given position for study. Potentially, the use of "unnatural amino acid mutagenesis" can permit the analysis of the function of individual amino acids in a protein at a high level of resolution. In fact, numerous studies have now appeared describing the synthesis and study of proteins containing unnatural amino acids (*10*−*27*).

Logically, the analysis of amino acid contributions to protein function should be focused disproportionately on amino acids having side chains with functional groups that participate in the mechanism of action of many enzymes via substrate binding and catalysis, such as histidine, serine, threonine, cysteine, tyrosine, aspartic acid, glutamic acid, lysine, and arginine. Surprisingly, however, there has been little systematic effort to prepare sets of analogues potentially useful for defining the roles of such amino acids in many proteins. Notable exceptions are tyrosine, numerous analogues of which have been prepared and used to study the function of DNA topoisomerases I and the nicotine acetylcholine receptor (*19*, *28*, *29*), and aspartic acid, several analogues of which have been used to study the active-site aspartic acid residue in dehydrofolate reductase (DHFR)[1] (*30*).

One amino acid of special interest in this context is arginine. It is an essential element of a (serine- and arginine-rich) protein domain important in splicing (*31*), functions critically in nucleic-acid-binding proteins (*32*, *33*), and is present in the motif arginine−glycine−aspartate (RGD) that is associated with adhesive and chemotactic properties of proteins, such as osteoporin (*34*, *35*), as well as the binding of fibrinogen and platelet aggregation inhibitors to the platelet GPIIb•IIIa complex (*36*).

Presently, we describe the preparation and incorporation into protein of seven analogues of arginine that should find broad utility in the analysis of arginine function in proteins. Analogues of particular interest include *N*-methylarginine (**2**), which facilitates nuclear export when present in hnRNP proteins (*37*), is part of a modified repetitive arginine−glycine−glycine (RGG) motif in the nuclear localization signal of serine−arginine (SR) proteins that affects nuclear import (*38*), and is a constituent of histones formed by post-translational modification (*39*, *40*). Citrulline (**4**) is a protein constituent formed by enzymatic deimination of arginine (*41*)

† This study was supported by National Institutes of Health Research Grant CA77359, awarded by the National Cancer Institute.

* To whom correspondence should be addressed: Department of Chemistry, University of Virginia, Charlottesville, VA 22904-4319. Telephone: 434-924-3906. Fax: 434-924-7856. E-mail: sidhecht@virginia.edu.

[1] Abbreviations: DHFR, dihydrofolate reductase; TEMED, *N*,*N*,*N*′,*N*′-tetramethylethylenediamine; ESI, electrospray ionization; CI, chemical ionization; DMF, *N*,*N*-dimethylformamide; IPTG, isopropyl-β-D-thiogalactopyranose; DTT, dithiothreitol; BSA, bovine serum albumin; MES, 4-morpholinemethanesulfonate; TFA, trifluoroacetic acid; FAB, fast atom bombardment; DEAE, diethylaminoethyl; EDTA, ethylenediamine tetraacetic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid.

10.1021/bi062042r CCC: $37.00
Published on Web 03/09/2007

and is associated with neurodegeneration (*42*−*44*). Although not present naturally in proteins, homoarginine (**5**) has been introduced into a protein by chemical modification of lysine residues and shown, e.g., to stabilize a cold-adapted α-amylase from *Pseudoalteramonas haloplanhtis* (*45*). Also of interest for protein modification are arginine derivative **3**, containing an electron-withdrawing guanidine-linked nitro substituent, and conformationally constrained arginine analogues **6**−**8**.

The arginine analogues were incorporated into position 22 of *Escherichia coli* dihydrofolate reductase, which normally contains a tryptophan residue; this substitution was carried out to assess the efficiency of nonsense codon suppression in a sequence context found previously to provide results broadly representative of many substitutions studied. Also altered were positions Arg218 and Arg437 of *Photnus pyralis* luciferase. Arg218 has been suggested to be critical for the maintenance of the rigidity and polarity of the emitter-binding site in the enzyme. While no role has been suggested for Arg437, its presence near the putative active site of the enzyme and as a hydrophilic amino acid in an otherwise hydrophobic region suggested that it was likely to be important to luciferase function. The properties of the modified proteins were characterized to illustrate the way in which the arginine analogues can facilitate an analysis of protein function.

EXPERIMENTAL PROCEDURES

General Methods and Materials. Amino acids *S*-arginine, *S*-citrulline, *S*-homoarginine, (4*R*)-hydroxy-*S*-proline, and *S*-proline were purchased from Sigma Chemicals (St. Louis, MO). N^{γ}-Methylarginine and N^{γ}-nitroarginine were purchased from Novabiochem (San Diego, CA). The succinimidyl ester of pentenoic acid was prepared according to a reported procedure (*46*). Anhydrous acetonitrile and anhydrous *N*,*N*-dimethylformamide (DMF) were purchased from Aldrich Chemicals. Triethylamine was distilled over P_2O_5. T4 RNA ligase, T4 polynucleotide kinase, and restriction enzymes *Nco*I, *Fok*I, *Bam*HI, *Pst*I, and *Drn*I were purchased from New England Biolabs (Ipswich, MA). [^{35}S]Methionine (1000 Ci/mmol) was obtained from GE Healthcare (Piscataway, NJ). Tris, acrylamide, bis-acrylamide, urea, ammonium persulfate, *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED), ethylenediamine tetraacetic acid (EDTA), amino acids, diethylaminoethyl (DEAE) Sepharose, imidazole, NADPH, dihydrofolate, nucleoside triphosphates (NTPs), cyclic adenosine 5′-monophosphate (cAMP), phospho(enol)pyruvate, pyruvate kinase, guanosine 5′-monophosphate (GMP), folinic acid, and other components for the cell-free translation systems were purchased from Sigma Chemicals. T7 RNA polymerase, isopropyl-β-D-thiogalactopyranose (IPTG), *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (Hepes), and sodium dodecyl sulfate (SDS) were purchased from USB (Cleveland, OH). Kits for T7 transcription of RNA were purchased from Epicentre Technologies (Madison, WI). The pTrc99A vector was from Pharmacia Biotech (Piscataway, NJ). Plasmid pYRNA8 coding for yeast suppressor $tRNA^{Phe}$ was constructed by Dr. S. Mamaev. Strain Bl-21 (DE-3) was purchased from Promega (Madison, WI).

Thin-layer chromatography (TLC) was performed on 60 F_{254} silica gel plates from E. Merck (Bodman Industries, Aston, PA) and visualized using iodine or charring with 10% phosphomolybdic acid in ethanol. Silica gel 60 (230−400 mesh) was used for flash-column chromatography. Melting points were measured on a capillary melting-point apparatus and are uncorrected. All reactions were carried out under an argon atmosphere using oven-dried glassware.

^{1}H nuclear magnetic resonance (NMR) and ^{13}C NMR spectra were recorded on a General Electric QE 300 NMR instrument, and high-resolution mass spectra were recorded at the Michigan State University−NIH Mass Spectrometry Facility in East Lansing, MI.

Plasmid DNAs were isolated using a Qiagen plasmid midi kit. SDS−polyacrylamide gel electrophoresis (PAGE) was carried out using the standard Laemmli procedure (*47*). Gels were visualized and quantified by the use of a Molecular Dynamics 400E Phosphorimager with ImageQuant version 3.2. DNA sequence analysis was performed using a Sequenase version 2.0 DNA sequencing kit from USB (Cleveland, OH).

*Synthesis of Aminoacylated pdCpA Derivatives: N-(4-Pentenoyl)-S-homoarginine pdCpA Ester (**pentenoyl-5-pdCpA**).* *S*-Homoarginine (500 mg, 2.2 mmol) was dissolved in 15 mL of 1:1 water/dioxane and treated with 277 mg (3.3 mmol) of $NaHCO_3$. To this stirred solution was added 408 mg (2.06 mmol) of the succinimidyl ester of pentenoic acid (*46*). The reaction mixture was stirred at 25 °C for 18 h and then acidified to pH 4 by the addition of 1 N $NaHSO_4$. The aqueous phase was extracted with 10 mL of ethyl acetate, and the organic layer was discarded. The aqueous phase was concentrated under diminished pressure. The residue was suspended in 15 mL of methanol, and the insoluble solid was filtered. The solid was washed 2 times with 10 mL portions of methanol. The combined methanol extract was concentrated under diminished pressure. The residue was purified by flash chromatography on a silica gel column (20 × 3 cm). Elution with 1:1, 2:1, 3:1, and then 4:1 methanol/ethyl acetate afforded *N*-(4-pentenoyl)-*S*-homoarginine as a colorless foam: yield, 350 mg (59%); silica gel TLC R_f, 0.32 (4:1 methanol/ethyl acetate); ^{1}H NMR (CD_3OD) δ, 1.30−1.39 (m, 2H), 1.49−1.81 (m, 4H), 2.25−2.30 (m, 4H), 3.07−3.12 (m, 2H), 4.19−4.23 (m, 1H), 4.90−5.04 (m, 2H), and 5.74−5.83 (m, 1H); mass spectrum [chemical ionization (CI), methane], *m*/*z* 271.2 $(M + H)^+$; mass spectrum [fast atom bombardment (FAB)], *m*/*z* 271.1762 ($C_{12}H_{23}N_4O_3$ requires 271.1770).

N-(4-Pentenoyl)-*S*-homoarginine (300 mg, 1.11 mmol) was dissolved in 5 mL of DMF and heated at 45−50 °C. To this stirred solution was added 755 μL (4.4 mmol) of *N*,*N*-diisopropylethylamine followed by 140 μL (2.22 mmol) of $ClCH_2CN$. The reaction mixture was stirred at 45−50 °C for 36 h and then concentrated under diminished pressure. The residue was purified by flash chromatography on a silica gel column (25 × 3 cm). Elution with 2:1 and then 1:1 ethyl acetate/methanol afforded *N*-(4-pentenoyl)-*S*-homoarginine cyanomethyl ester as a syrup: yield, 175 mg (51%); silica gel TLC R_f, 0.58 (1:1 ethyl acetate/methanol); ^{1}H NMR (CD_3OD) δ, 1.32−1.48 (m, 2H), 1.50−1.62 (m, 2H), 1.69−1.85 (m, 2H), 2.29−2.30 (m, 4H), 3.11−3.16 (m, 2H), 4.36−4.41 (m, 1H), 4.86 (d, 2H, $J = 1.5$ Hz), 4.91−5.04 (m, 2H), and 5.74−5.83 (m, 1H); mass spectrum (chemical ionization, methane), *m*/*z* 310.7 $(M + H)^+$; mass spectrum (FAB), *m*/*z* 310.1890 ($C_{14}H_{24}N_5O_3$ requires 310.1879).

A solution of the tris(tetrabutylammonium) salt of pdCpA (*48*) (4.0 mg, 2.94 μmol) dissolved in 50 μL of anhydrous DMF was added to a dry conical vial containing 20 mg (64.6 μmol) of *N*-(4-pentenoyl)-*S*-homoarginine cyanomethyl ester in 100 μL of anhydrous DMF. To this stirred solution was added 15 μL of triethylamine. The reaction mixture was stirred at 25 °C and monitored by high-performance liquid chromatography (HPLC). Aliquots (8 μL) were removed and diluted with 42 μL of 1:1 CH_3CN/50 mM NH_4OAc at pH 4.5; 20 μL of each diluted aliquot was analyzed on a 3 μm C_{18} reversed-phase HPLC column (250 × 10 mm). The column was washed with 1 → 63% CH_3CN/50 mM NH_4-OAc at pH 4.5 over a period of 45 min at a flow rate of 3.5 mL/min (detection at 260 nm; the desired product had a retention time at 15.2 min). After 48 h, the remaining reaction mixture was diluted with 400 μL of 1:1 CH_3CN/50 mM NH_4-OAc at pH 4.5 and purified by C_{18} reversed-phase HPLC on a semipreparative column (250 × 10 mm) using the same gradient as described above. The N-protected dinucleotide derivative of **5** was recovered from the appropriate fractions as a colorless solid following lyophilization: yield, 0.7 mg (27%); mass spectrum [electrospray ionization (ESI)], *m*/*z* 889.2 $(M + H)^+$; mass spectrum (FAB), *m*/*z* 889.2743 ($C_{31}H_{47}N_{12}O_{15}P_2$ requires 889.2759).

*1-[N-(4-Pentenoyl)]-(4S)-guanidinyl-S-proline pdCpA Ester (**pentenoyl-6-pdCpA**).* To a solution of 300 mg (1.3 mmol) of 1-[*tert*-butyloxycarbonyl]-(4*S*)-amino-*S*-proline (*49*) in 10 mL of methanol was added 162 mg (1.3 mmol) of aminoiminomethanesulfonic acid (*50*). The reaction mixture was stirred at 25 °C, and a clear solution was formed after 3 h. The reaction mixture was then stirred at 25 °C for another 15 h. The solution was concentrated under diminished pressure, and the residue was applied to a silica gel column (20 × 3 cm). Elution with 2:1 and then 1:1 ethyl acetate/methanol and finally with methanol afforded the intermediate *N*-Boc-(4*S*)-guanidinyl-*S*-proline as a colorless powder; silica gel TLC R_f, 0.52 (methanol). The structure of the intermediate was confirmed from its mass spectrum [(CI, methane), *m*/*z* 273.1 $(M + H)^+$ (theoretical, 273.1)]. *N*-Boc-(4*S*)-guanidinyl-*S*-proline was treated with 4 mL of 3:1 dicholoromethane/trifluoroacetic acid (TFA), and the reaction mixture was stirred at 25 °C for 4 h. The solution was concentrated under diminished pressure. The residue was co-evaporated 3 times with 5 mL portions of toluene to afford (4*S*)-guanidinyl-*S*-proline as a syrup. The residue was dissolved in 10 mL of 1:1 water/dioxane and treated with 293 mg (3.48 mmol) of $NaHCO_3$. To this stirred solution was added 412 mg (2.1 mmol) of succinimidyl pentenoate (*46*). The reaction mixture was stirred at 25 °C for 4 h and then acidified to pH 4 by the addition of 1 N $NaHSO_4$. The aqueous phase was extracted with 10 mL of ethyl acetate, and the organic phase was discarded. The aqueous phase was concentrated under diminished pressure. The residue was suspended in 20 mL of methanol, and the insoluble solid was filtered. The solid was washed 2 times with 10 mL portions of methanol. The combined methanol extract was concentrated under diminished pressure. The residue was purified by flash chromatography on a silica gel column (18 × 3 cm). Elution with 1:1 ethyl acetate/methanol and then with methanol afforded 1-[*N*-(4-pentenoyl)]-(4*S*)-guanidinyl-*S*-proline as a colorless foam: overall yield, 250 mg (75% for three steps); silica gel TLC R_f, 0.5 (methanol); ^1H NMR (CD_3OD) δ, 2.25−2.60 (m, 6H), 3.51−3.58 (m, 1H), 3.72−3.85 (m, 1H), 4.06−4.28 (m, 2H), 4.85−5.06 (m, 2H), and 5.74−5.85 (m, 1H); mass spectrum (ESI), *m*/*z* 255.3 $(M + H)^+$; mass spectrum (FAB), *m*/*z* 255.1461 ($C_{11}H_{19}N_4O_3$ requires 255.1343).

1-[*N*-(4-Pentenoyl)]-(4*S*)-guanidino-*S*-proline (130 mg, 0.51 mmol) was dissolved in 3 mL of anhydrous DMF and stirred at 25 °C. To this stirred solution was added 350 μL (2.0 mmol) of *N*,*N*-di-isopropylethylamine followed by 75 μL (1.0 mmol) of ICH_2CN. The reaction mixture was stirred at 25 °C for 36 h and then concentrated under diminished pressure. 1-[*N*-(4-Pentenoyl)]-(4*S*)-guanidinyl-*S*-proline cyanomethyl ester was isolated as a crude product: yield, 125 mg; mass spectrum (ESI), *m*/*z* 294.3 $(M + H)^+$ (theoretical, 294.15).

A solution of the tris(tetrabutylammonium) salt of pdCpA (*48*) (2.5 mg, 1.84 μmol) dissolved in 50 μL of anhydrous DMF was added to a dry conical vial containing 15 mg (48.5 μmol) of the cyanomethyl ester in 50 μL of anhydrous DMF. To this stirred solution was added 10 μL of triethylamine. The reaction mixture was stirred at 25 °C and monitored by HPLC. Aliquots (4 μL) were removed and diluted with 23 μL of 1:1 CH_3CN/50 mM NH_4OAc at pH 4.5 and 23 μL of water; 30 μL of each diluted aliquot was analyzed on a 3 μm C_{18} reversed-phase HPLC column (250 × 10 mm). The column was washed with 1 → 63% CH_3CN/50 mM NH_4-OAc at pH 4.5 over a period of 45 min at a flow rate of 3.5 mL/min (detection at 260 nm; the desired 2′- and 3′-*O*-acyl products eluted at 12.3 and 13.1 min). After 18 h, the remaining reaction mixture was diluted with 200 μL of 1:1 CH_3CN/50 mM NH_4OAc at pH 4.5 and 200 μL of water and purified by C_{18} reversed-phase HPLC on a semipreparative column (250 × 10 mm) using the same gradient as described above. 1-[*N*-(4-Pentenoyl)]-4*S*-guanidinyl-*S*-proline pdCpA ester (**pentenoyl-6-pdCpA**) was recovered from the appropriate fractions as a colorless solid following lyophilization: yield, 0.8 mg (27%); mass spectrum (ESI), *m*/*z* 873.3 $(M + H)^+$ (theoretical, 873.2).

In Vitro Transcription of 5′-Monophosphorylated Truncated Yeast Suppressor tRNA $^{Phe}_{CUA}(-C_{OH})$ (51). Synthesis of a truncated suppressor $tRNA_{CUA}$ transcript ($tRNA^{Phe}_{CUA}-C_{OH}$) programmed by the *Fok*I restriction fragment from plasmid pYRNA8 was carried out using AmpliScribe T7 transcription kit with additional 20 mM GMP and 5% formamide at 42 °C for 12−24 h. This condition was critical for the preparation of tRNA transcripts of the correct size. The addition of 5% formamide into the transcription mixture did not influence the total yield of tRNA (5−6 mg/mL after purification) but improved the quality of the tRNA.

After transcription, the tRNA was purified by chromatography by loading 1 mL of transcription mixture on a 1 mL DEAE Sepharose column that had been equilibrated with 0.1 M NaOAc at pH 5.3 followed by NaCl-gradient elution (usually, the tRNA transcript eluted with 0.5−0.7 M NaCl in 0.1 M NaOAc at pH 5.3). Fractions containing the tRNA transcript were collected and precipitated with 2 volumes of isopropanol. For an additional purification step, the tRNA transcript was sometimes purified by gel filtration on a Sephacryl S-200 column in buffer containing 0.1 M NaCl, 1 mM EDTA, and 10 mM Tris-HCl at pH 8.0 and then precipitated with 3 volumes of ethanol in the presence 0.3

M sodium acetate at pH 5.3. Finally, the tRNA was washed with 70% ethanol, dissolved in water, and kept at −20 °C.

Synthesis of Aminoacyl-tRNAs. A mixture of 0.5 A_{260} unit of *N*-(4-pentenoyl)-aminoacyl-pdCpA and 50 μg of a tRNA$_{CUA}$ transcript lacking the 3′-terminal dinucleotide pCpA in 50 μL of 50 mM Na-Hepes at pH 7.5 containing 15 mM $MgCl_2$, 0.5 mM ATP, 20−25% dimethylsulfoxide (DMSO) (v/v), and 100 units of T4 RNA ligase was incubated at 37 °C for 90 min. The reaction mixture was then quenched with 0.1 volume of 3 M NaOAc at pH 5.0 followed by 4 volumes of cold ethanol. The resulting precipitate of *N*-(4-pentenoyl)-aminoacyl-tRNA was recovered after centrifugation. The precipitate was washed with ethanol, dried, and then dissolved in H_2O. The ligation efficiency was determined by gel electrophoresis in 8% PAGE/7 M urea/0.1 M NaOAc at pH 5.0 (*52*).

Free aminoacyl tRNA was obtained by deprotection of *N*-(4-pentenoyl)-aminoacyl-tRNA at 25 °C for 15 min at a tRNA concentration of 1 μg/μL in 10 mM I_2 [from a stock solution containing 100 mM I_2 dissolved in 1:1 tetrahydrofuran (THF)/H_2O]. After deprotection, the solution was centrifuged and the supernatant was adjusted to 0.3 M NaOAc at pH 5.0. Traces of iodine were removed by precipitating the resulting aminoacylated tRNA with 4 volumes of ethanol. The pellet of aminoacylated tRNA was then washed with 70% ethanol, dried under diminished pressure, and dissolved in 1 mM KOAc at pH 5.0 for *in vitro* suppression experiments or in formamide and 50 mM NaOAc at pH 5.0 containing 0.025% xylene cyanol for electrophoretic analysis.

Plasmid Construction. The gene for *Photinus pyralis* luciferase was excised from commercial plasmid pGL-3 (Promega) and subcloned into the pTrc99A vector. The resulting plasmid pTrcLucwt was analyzed by restriction analysis and luciferase assay of the protein product. A TAG codon was introduced at positions 218 and 437 using a QuickChange *in vitro* mutagenesis kit (Stratagene). The modifications were confirmed by DNA sequencing. Plasmids pETDHwt and pETDH$_{22}$ were constructed previously (*53*).

S-30 Preparation and in Vitro Translation. The preparation of the S-30 extract from *E. coli* was carried out as described by Short et al. (*54*) with minor modifications. Cells were grown until the optical density at 450 nm was 1.7−2.0. Cells were collected by centrifugation, washed, and pressed at 8400 psi in a French press. The S-30 extract was preincubated at 30 °C for 150 min followed by 37 °C for 15 min. Finally, the S-30 extract was dialyzed against 10 mM Tris-HCl at pH 7.4 containing 60 mM KOAc, 14 mM $MgCl_2$, and 1 mM dithiothreitol (DTT), then frozen in liquid nitrogen, and stored at −80 °C.

Translation reactions were carried out in incubation mixtures up to 1 mL in volume, containing 40% S-30 extract, 40−180 μg/mL template DNA, 300 μg/mL aminoacylated suppressor tRNA, 35 mM Tris-acetate at pH 7.4, 190 mM potassium glutamate, 30 mM ammonium acetate, 2 mM DTT, 0.2 mg/mL unfractionated *E. coli* tRNA, 2 mM IPTG, 3.5% PEG 6000, 20 μg/mL folinic acid, 20 mM ATP and GTP, 5 mM CTP and UTP, 10 mM cAMP, 100 μM amino acids, and 10 μCi/mL [^{35}S]methionine. Reactions were incubated for 45 min at 37 °C and quenched by chilling on ice. The results of the reactions were estimated by SDS gel analysis of aliquots followed by the quantitation of radioactivity in the gel using a phosphoimager.

Purification of Synthesized Proteins on Ni−Nitrilotriacetic Acid (NTA) Agarose. Aliquots (150−200 μL) of the translation mixture were diluted 2-fold with 50 mM potassium phosphate buffer at pH 7.8 containing 0.3 M NaCl and 100 μg/mL bovine serum albumin (BSA). Ni−NTA agarose columns (50 μL) were equilibrated with 500 μL of 50 mM potassium phosphate buffer at pH 7.8 containing 0.3 M NaCl. Diluted samples were applied to the columns, which were washed with 500 μL of 50 mM potassium phosphate buffer at pH 7.8 containing 0.3 M NaCl and then with 150 μL of the same buffer containing 50 mM imidazole. The proteins were eluted with 150 μL of 150 mM imidazole in 50 mM potassium phosphate buffer at pH 7.8 containing 100 μg/mL BSA and 0.3 M NaCl. In the case of luciferase, the protein was precipitated using 40% ammonium sulfate (16 h at 4 °C) and dissolved in 50 μL of 0.1 M potassium phosphate at pH 7.8 containing 1 mM $MgSO_4$, 1 mM EDTA, and 15% glycerol. The concentrations of the purified proteins were estimated using phosphorimager analysis after PAGE.

Determination of the Enzymatic Activity of Modified DHFRs. The enzymatic activity of DHFR was determined in deoxygenated 50 mM 4-morpholinemethanesulfonate (MES)/25 mM Tris/25 mM ethanolamine buffer at pH 7.0 containing 100 mM NaCl, 10 mM β-mercaptoethanol, and 0.1 mM EDTA at 37 °C and 340 nm in the presence of 100 mM DHF and 100 mM NADPH in a 1.0 cm path-length cuvette using a Perkin-Elmer Lambda Array 3840 spectrophotometer. A total of 1 unit of DHFR activity was defined as the amount of enzyme required to reduce 1 μmol of DHF/min based on a molar extinction coefficient of 11.8×10^{-3} M^{-1} cm^{-1} for NADPH plus DHF at 37 °C and pH 7.0.

Luciferase Assay. The luciferase assay was carried out using a commercial luciferase assay system (Promega). After translation, 10 μL of the translation mixture or purified samples was added to 100 μL of the luciferase assay reagent and light emission was measured immediately using a Hitachi F 2000 fluorescence spectrophotometer. The pH dependence of light emission spectra was studied in 0.1 M MES buffer, containing 5 mM $MgCl_2$, 0.5 mM luciferin, 1.5 mM ATP, and 1 mM EDTA. The pH was varied by adding solid 3-(*N*-morpholino)propanesulfonic acid (MOPS) (acidic pH) or Tris (alkaline pH).

RESULTS

Synthesis of Arginine pdCpA Esters. Arginine (**1**) and arginine analogues **2**−**5** (Figure 1) were all converted to the respective argininyl-pdCpA derivatives in a similar fashion; this is illustrated for homoarginine (**5**) in Scheme 1. The key intermediate was the N^{α}-protected cyanomethyl ester, which was obtained in two steps. Initially, homoarginine was treated with 4-pentenoic acid succinimide ester (*46*) in the presence of $NaHCO_3$ to obtain *N*-pentenoylhomoarginine in 59% yield. Subsequent treatment of the product with chloroacetonitrile in anhydrous DMF in the presence of di-isopropylethylamine at 45−50 °C afforded the desired cyanomethyl ester as a syrup in 51% yield. Coupling of the cyanomethyl ester with tris(tetrabutylammonium) salt of pdCpA in anhydrous DMF afforded the homoarginine pdCpA ester (**pentenoyl-5-pdCpA**) in 27% yield (Scheme 1). The syntheses of the

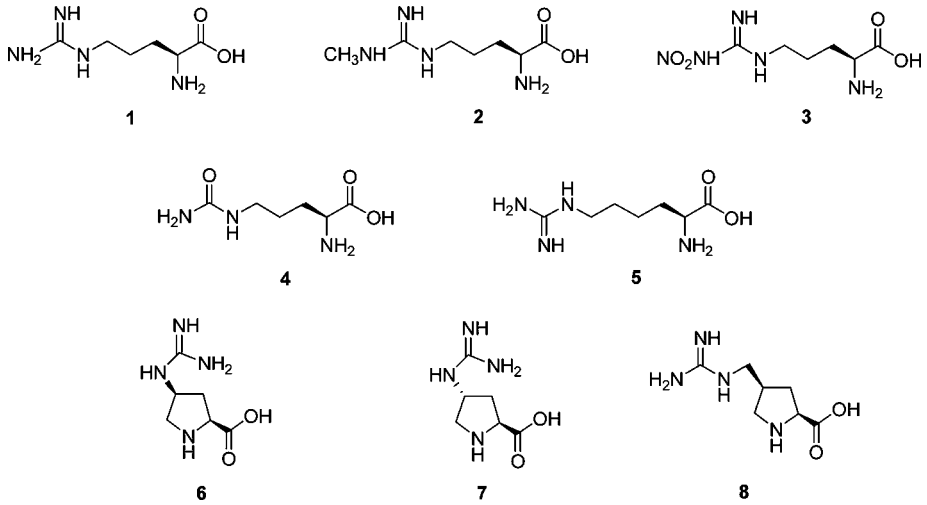
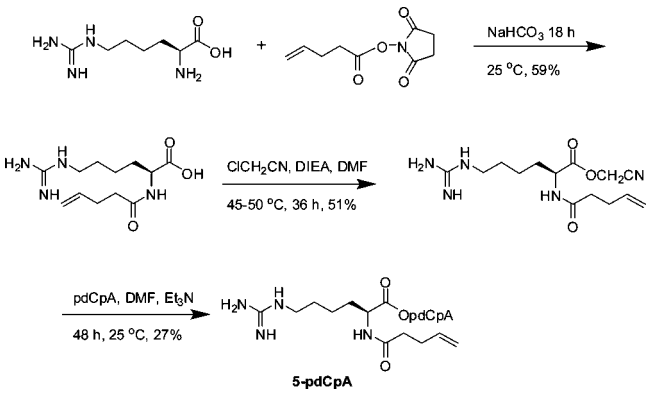


FIGURE 1: Arginine analogues attached to $tRNA^{Phe}_{CUA}$ and incorporated into proteins.

Scheme 1



pdCpA esters of *N*-pentenoylarginine and *N*-pentenoylarginine analogues **2**−**4** are described in Schemes 1 and 2 of the Supporting Information.

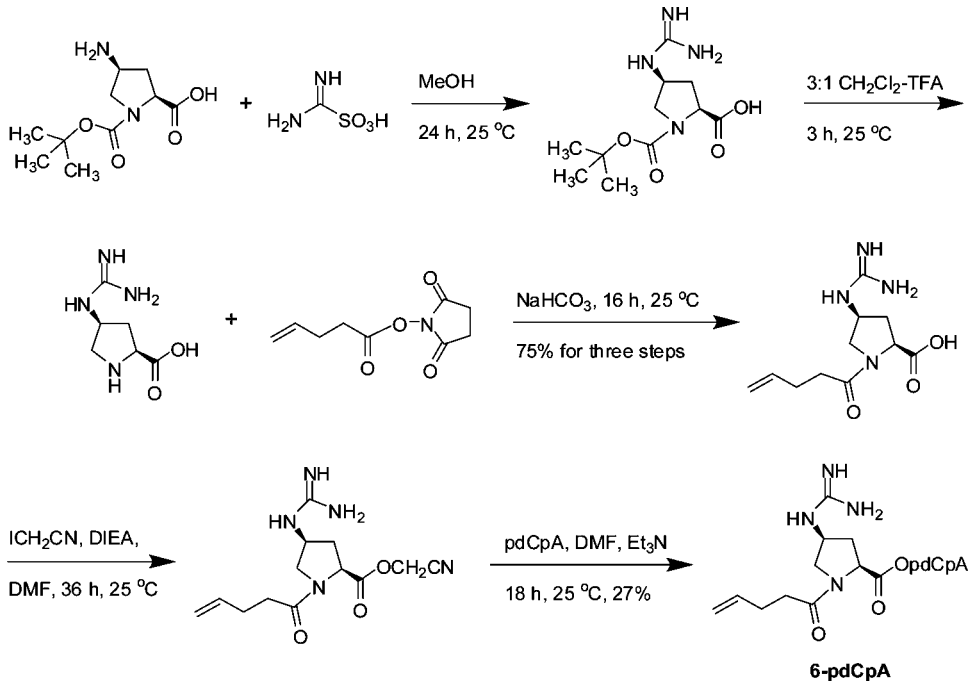
For the syntheses of conformationally constrained arginine analogues **6**−**8**, the syntheses began with 4-substituted derivatives of proline, as illustrated for arginine derivative **6** in Scheme 2. The known (*49*) (4*S*)-amino-*S*-proline was synthesized starting from (4*R*)-hydroxy-*S*-proline in 35% overall yield for five steps. Treatment of Boc-protected (4*S*)-amino-*S*-proline with aminoiminomethanesulfonic acid (*50*) afforded the desired Boc-protected (4*S*)-guanidinylproline derivative. The deprotection of the Boc group was accomplished with 3:1 CH_2Cl_2/TFA at 25 °C for 3 h. The resulting (4*S*)-guanidinylproline was treated with 4-pentenoic acid succinimide ester (*46*) in the presence of $NaHCO_3$ to obtain N^{α}-pentenoyl-(4*S*)-guanidinyproline in 75% overall yield for three steps. The cyanomethylation (*55*) reaction was carried out by treating the free acid with iodoacetonitrile in DMF at 25 °C using di-isopropylethylamine (DIEA) as a catalyst (Scheme 2). The desired cyanomethyl ester was obtained as a syrup. The cyanomethyl ester was coupled with the tris(tetrabutylammonium) salt of pdCpA in DMF in the presence of triethylamine to afford (4*S*)-guanidinylproline pdCpA ester (**pentenoyl-6-pdCpA**) in 27% yield (Scheme 2). The pdCpA esters of *N*-pentenoyl-protected (4*R*)-guanidinylproline (**pentenoyl-7-pdCpA**) and (4*S*)-guanidinylmethylproline (**pentenoyl-8-pdCpA**) and proline (**pentenoyl-9-pdCpA**) were prepared analogously, as outlined in Schemes 3−5 of the Supporting Information, respectively.

Synthesis of Argininyl-$tRNA_{CUA}$s. Misacylated suppressor $tRNA_{CUA}$ transcripts were prepared as described previously (*56*) by *in vitro* runoff transcription of a *Fok*I-linearized DNA plasmid. The transcript intentionally lacked the 3′-terminal C and A moieties common to all mature tRNAs. As illustrated in Scheme 3 for homoarginine, the N^{α}-pentenoyl-protected arginine pdCpA esters were combined with the abbreviated $tRNA_{CUA}-C_{OH}$ transcript in the presence of T4 RNA ligase and ATP, as described in refs *5*−*7*.

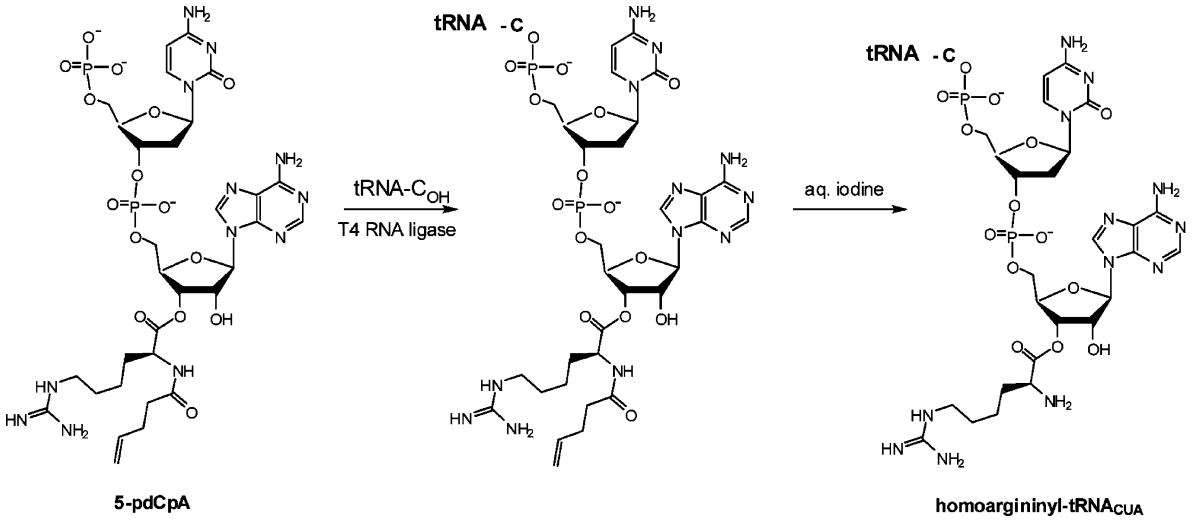
Following enzymatic ligation to afford the full-length N^{α}-pentenoyl-protected argininyl-$tRNA_{CUA}$s, the extent of the ligation reaction was determined by analyzing the tRNA samples by PAGE under acidic conditions (*52*). As shown in Figure 2, under these conditions, the full-length ($tRNA-CCA_{OH}$) and abbreviated ($tRNA_{CUA}-C_{OH}$) tRNA transcripts were resolved on the gel based on their differences in length (and hence phosphodiester content, cf. lanes 1 and 2). Transfer RNAs bearing *N*-pentenoylaminoacyl moieties (e.g., leucine or citrulline; lanes 3 and 6, respectively) migrated a smaller distance. Interestingly, the presence of a positively charged group on the aminoacyl moiety (e.g., for arginine and homoarginine; lanes 4 and 5, respectively) reduced the mobility further. The free argininyl-$tRNA_{CUA}$s were obtained by the addition of a small amount of iodine to the aqueous solution of the protected argininyl-tRNAs (*57*, *58*).

Modified DHFRs Containing Arginine Analogues at Position 22. Position 22 in DHFR normally contains tryptophan; it is a conserved residue that is part of a flexible loop that helps to define the folate-binding site (*59*−*63*). Changes in this amino acid can be tolerated but do lead to changes in enzyme efficiency (*53*, *59*−*63*). Accordingly, the argininyl-$tRNA_{CUA}$s as well as other misacylated tRNAs were used to synthesize DHFRs altered at this position, using a construct that had been shown to function efficiently in *in vitro* DHFR synthesis previously (*53*). Most of the analogues tested were incorporated reasonably efficiently, but the tRNAs activated with conformationally constrained analogues **6**−**8** functioned poorly as suppressor tRNAs (Table 1). Interestingly, all of these analogues are proline derivatives containing guanidine substituents in the 4 position of the proline ring and may thus be regarded as "hybrids" of proline and arginine. While none of these analogues could be incorporated efficiently, prolyl- and argininyl-$tRNA_{CUA}$s themselves effected rea-

Scheme 2



Scheme 3



sonably efficient suppression of the same UAG codon (Table 1). The reason for the limited incorporation of analogues **6**−**8** is unclear but is consistent with the observed trend of more selective incorporation of amino acids having side chains that include polar or charged residues.

Those DHFR analogues accessible in reasonable yield were used to assess the effects of the modifications at position 22 on DHFR function. As shown in Table 2, the replacement of tryptophan with other aromatic amino acids, such as phenylalanine or 2-naphthylglycine, had little effect on the ability of the enzyme to convert dihydrofolate to tetrahydrofolate. The introduction of an amino group on the para

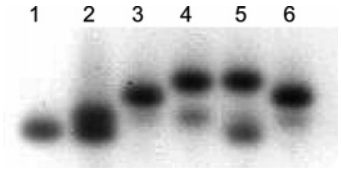


FIGURE 2: Analysis of the pentenoyl-protected aminoacylated yeast $tRNA^{Phe}_{CUA}$ transcript on a denaturing acid gel. Lane 1, truncated $tRNA^{Phe}_{CUA}$; lane 2, mixture of full-length and truncated $tRNA^{Phe}_{CUA}$s; lane 3, leucyl-$tRNA^{Phe}_{CUA}$; lane 4, argininyl-$tRNA^{Phe}_{CUA}$; lane 5, homoargininyl-$tRNA^{Phe}_{CUA}$; and lane 6, citrullinyl-$tRNA^{Phe}_{CUA}$.

Table 1: Incorporation of Arginine Analogues into Position 22 of *E. coli* DHFR and Position 218 of *P. pyralis* Luciferase

amino acid	suppression efficiency (%)	
	DHFR	firefly luciferase
S-arginine (**1**)	27 (21)[a]	29
N^{γ}-methylarginine (**2**)	7.8	
N^{γ}-nitroarginine (**3**)	5.8	
citrulline (**4**)	6.1	
homoarginine (**5**)	18	
(4*S*)-guanidinyl-*S*-proline (**6**)	2.6	3.7
(4*R*)-guanidinyl-*S*-proline (**7**)	1.9	2.5
(4*S*)-guanidinylmethyl-*S*-proline (**8**)	0.5	1.8
S-proline (**9**)	21	22

[a] Duplicate experiment.

Table 2: Comparison of Relative Activities of Modified DHFRs Having Arginine Analogues in Position 22

amino acid	enzyme activity (units/μL)
tryptophan (wild-type enzyme)	0.070
phenylalanine	0.062
2-naphthylglycine	0.065
p-aminophenylalanine	0.027
arginine (**1**)	<0.001
N^{γ}-methylarginine (**2**)	0.012
N^{γ}-nitroarginine (**3**)	<0.001
citrulline (**4**)	0.007
homoarginine (**5**)	<0.001

position of phenylalanine resulted in a 2−3-fold reduction in NADPH consumption, suggesting that a positively charged group in this position would not be well-tolerated. In fact, the introduction of arginine, homoarginine, or N^{γ}-nitroarginine essentially eliminated enzyme activity. While also relatively inefficient, modified DHFRs containing citrulline and N^{γ}-methylarginine at position 22 did exhibit some ability to function catalytically.

Firefly Luciferases Modified at Positions 218 and 437. To permit a more narrowly focused study of the influence of different modifications of the arginine guanidine group on enzyme function and also gain experience in introducing arginine analogues into other proteins, two new constructs of *P. pyralis* luciferase were prepared. These involved positions in luciferase that contain arginine in the wild-type enzyme. Plasmids pTrcLuc218 and pTrcLuc437, containing the luciferase gene with TAG codons at positions corresponding to Arg218 and Arg437, respectively, were employed for *in vitro* transcription and translation. Translation from the wild-type gene and also from the modified genes in the presence of unacylated $tRNA^{Phe}$ was carried out at the same time as controls. Although the conformationally constrained arginine analogues (**6**−**8**) were also not incorporated efficiently into this luciferase (Table 1), it was found that the remaining four arginine analogues (**2**−**5**) were incorporated with moderate to good efficiency, as shown in Figure 3. The suppression yields, relative to luciferase synthesis from wild-type mRNA, varied from 21 to 36% for the modification of position 218 and 17−50% for the modification of position 437. After purification by Ni−NTA chromatography, the modified luciferases were assayed for function.

Samples of wild-type and modified luciferases were characterized with respect to emission spectrum parameters (Em and λ_{max}) at pH 7.3 (Figure 4 and Table 3). As is clear from Figure 4, the introduction of modified arginines had

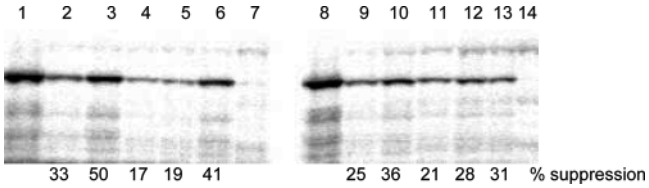


FIGURE 3: Electrophoretic analysis of samples prepared during translation of *P. pyralis* luciferase from two modified mRNAs [UAG codon in positions corresponding to Arg437 (left) and Arg218 (right)] in the presence of argininyl-tRNA (lanes 2 and 9), homoargininyl-tRNA (lanes 3 and 10), N^{γ}-methylargininyl-tRNA (lanes 4 and 11), N^{γ}-nitroargininyl-tRNA (lanes 5 and 12), citrullinyl-tRNA (lanes 6 and 13), and uncharged tRNA (lanes 7 and 14). Lanes 1 and 8, translation of luciferase from wild-type mRNA.

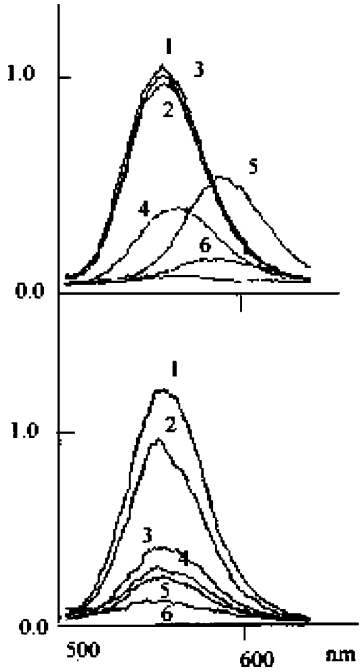


FIGURE 4: Light emission spectra of samples of *P. pyralis* luciferase, synthesized from wild-type mRNA (*1*) and two modified mRNAs [UAG codon in positions corresponding to Arg218 (upper panel) and Arg437 (lower panel)] in the presence of argininyl-tRNA (*2*), N^{γ}-methylargininyl tRNA (*3*), homoargininyl-tRNA (*4*), N^{γ}-nitroargininyl-tRNA (*5*), and citrullinyl-tRNA (*6*). The spectra in this figure are not normalized for the amount of protein measured.

Table 3: Comparison of the Emission Efficiency (Em) and Maximum Wavelength (λ_{max}) of Wild-Type (Arg437 and Arg218) and Modified Luciferases Containing Arginine Analogues in Positions 437 and 218[a]

	position 437		position 218	
amino acid	λ_{max} (nm)	Em (%)	λ_{max} (nm)	Em (%)
arginine (**1**)	559 ± 0.6[b]	100	559 ± 0.4	100
N^{γ}-methylarginine (**2**)	559 ± 2	41 ± 10	560 ± 0.8	136.2 ± 13.5
N^{γ}-nitroarginine (**3**)	557 ± 2	1.3 ± 0.2	594 ± 2.5	18.6 ± 3.7
citrulline (**4**)	558 ± 3	1.7 ± 0.5	591 ± 2	2.5 ± 0.9
homoarginine (**5**)	558 ± 3	4.5 ± 1.8	568 ± 0.7	37.6 ± 11.5
lysine	566 ± 2.9	19.2 ± 5.6	569 ± 2.2	28.7 ± 8.4

[a] Determined at pH 7.3. [b] Standard deviation based on five experiments.

rather different effects at the two positions. At position 218 of this luciferase, substitution of N^{γ}-methylarginine for arginine had little effect on the efficiency or wavelength of emitted light. In other experiments, the luciferase containing N^{γ}-methylarginine actually emitted significantly more efficiently than the wild-type enzyme, as reflected in Table 3. In comparison, arginine analogues containing **3**−**5** at position 218 all caused the luciferase to emit less efficiently and at a longer wavelength. The most extreme changes were caused by the introduction of citrulline and N^{γ}-nitroarginine, which resulted in wavelength shifts of more than 30 nm and a significant reduction in the efficiency of emitted light. In comparison, substitution of arginine at position 437 by any of the analogues produced a substantial diminution in the amount of light produced but no significant shift in the wavelength of the emitted light (Table 3).

Also measured were the thermal stabilities of the modified luciferases, as illustrated for the wild-type enzyme and modified luciferase containing N^{γ}-methylarginine at position 218 (Figure 5). As shown, no difference was observed in

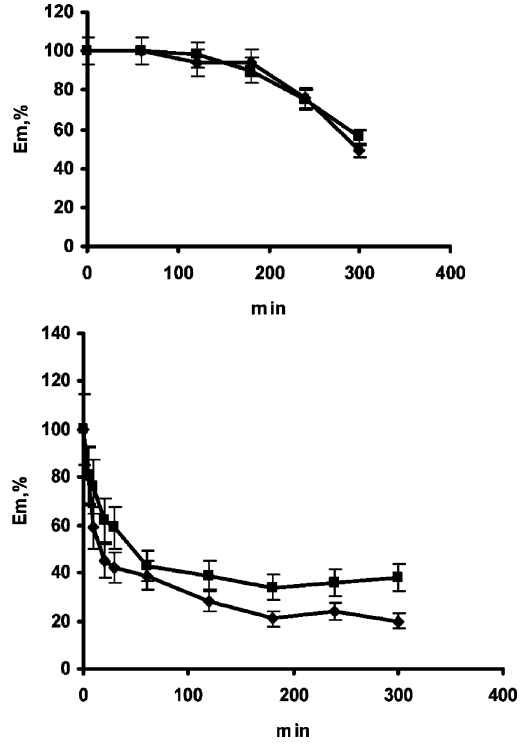


FIGURE 5: Comparison of the thermostabilities of wild-type (◆, arginine in position 218) and modified (■, N^{γ}-methylarginine in position 218) luciferases at 37 °C (upper panel) and 42 °C (lower panel).

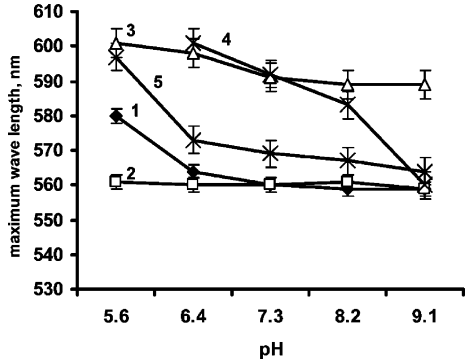


FIGURE 6: Effect of pH on the emission spectra of modified luciferases containing arginine (**1**) (wild type) and its unnatural analogues N^{γ}-methylarginine (**2**), N^{γ}-nitroarginine (**3**), citrulline (**4**), and homoarginine (**5**) in position 218.

the stability of these two enzymes at 37 or 42 °C. Luciferases can exhibit pH-dependent changes in the wavelength of emitted light, and the four modified luciferases containing arginine analogues at position 218 were characterized at different pH values from 5.6 to 9.1, in comparison with the wild type. As shown in Figure 6, with the exception of the species containing N^{γ}-methylarginine, all of the enzymes emitted at a shorter wavelength with increasing pH but the patterns of change differed for all of the enzymes. The wild-type luciferase exhibited a λ_{max} at 580 nm at the lowest pH (5.6), but the wavelength maximum shifted to 564 nm at pH 6.4. Further increases in pH produced only minimal additional changes in the emission wavelength. Of the modified luciferases, the enzyme containing homoarginine exhibited similar behavior (albeit over a wider range of wavelengths), emitting at 597 nm at pH 5.6 and 573 nm at pH 6.4. A further change of 9 nm was observed in λ_{max} as the pH was raised to 9.1. The greatest pH-dependent change in emission wavelength was observed for the luciferase containing citrulline, which emitted at 601 nm at pH 6.4 and shifted continuously at increasing pH values, emitting at 560 nm at pH 9.1. The remaining two luciferases, containing N^{γ}-methylarginine and N^{γ}-nitroarginine, exhibited relatively small changes in λ_{max} values as the pH was varied.

DISCUSSION

The goals of the present study were focused on the preparation of a set of arginine analogues and their incorporation into representative proteins. A demonstration that small changes in the amino acid structure could produce significant changes in the properties of the synthesized proteins was also of interest but was expected on the basis of numerous previous studies with other unnatural amino acids (*10*−*27*). Seven analogues of arginine were prepared, characterized fully, and used to activate a suppressor $tRNA_{CUA}$ transcript. These were all shown to be incorporated into the two model proteins chosen for study, although three of the arginine analogues were incorporated only in low yield in the proteins chosen. Because the incorporation of individual amino acid analogues is often context-dependent, these analogues may well be incorporated into other proteins with sufficient efficiency to permit their study in those cases.

The arginine analogues chosen for study included those having a modified guanidine moiety (**2**−**4**) that results in the alteration of the pK_a and nucleophilicity of this functional group. Also studied were analogues having conformationally constrained side chains (**6**−**8**) and one analogue (**5**) in which the side chain is longer than that in arginine itself. While the potential utility of each of these modified arginines will clearly differ according to the specific arginine being studied, the analogues were chosen to be broadly representative of species capable of defining the need for the positioning and orientation of the guanidine moiety of the arginine side chain in the wild-type proteins under study, as well as the nature of this functional group itself.

The modified proteins studied the most carefully in the present case were those derived from firefly luciferase. This protein has been of considerable interest as a reporter in biological studies (*64*−*67*) for a number of reasons. These include high sensitivity of detection, low background luminescence in host cells, the ability to manipulate the emission properties of the enzyme, and the absence of requirement for post-translational processing. Although related luciferases are found in many species of fireflies, the luciferase from the North American firefly, *P. pyralis* (EC 1.13.12.7), is now of special interest because the X-ray crystallographic structure has been reported (*68*). The analysis of the way in which this enzyme functions has been the subject of a number of recent studies (*68*−*74*).

P. pyralis luciferase contains 550 amino acids and has a molecular weight of 62 kDa. Its molecular architecture is quite unique and includes large N-terminal (1−436) and small C-terminal (441−550) domains. These domains are connected by polypeptide segment 436−440, which contains highly conserved amino acid residues and is disordered in the crystal structure (*68*). It has been suggested that the two domains come together to form the substrate-binding pocket (*68*). On the basis of the crystal structure and mutagenesis

analysis of luciferase function, the molecular modeling of the active site of luciferase has been carried out (*69*−*71*). By analogy with acyltransferases, which also catalyze the formation of substrate-bound adenylate, the luciferase active site was situated in the hydrophobic cleft between the two structural domains. Indeed, the most invariant residues were found on the surface of these domains opposite each other and in the polypeptide that connects the two domains (*69*−*71*). It is believed that the active site of luciferase can be formed by residues 198−206, 218, 245−247, 413−457, and 524−528 (*68*−*71*). Among these are seven absolutely conserved residues (Gly200, Lys206, Glu344, Asp422, Arg437, Gly446, and Glu455), which were invariant in 38 enzymes in this superfamily. *P. pyralis* luciferase contains 11 arginine residues, but only 3 of them are located near the proposed active site (Arg218, Arg337, and Arg437) (*68*, *69*, *74*).

At present, there are three models of the active site of luciferase, and arginines play an important role in all of them (*69*−*71*). Sundalova and Ugarova have proposed that side chains of Arg218 and Arg337 interact with luciferin and shield the substrate from the solvent (*71*). Branchini and coworkers have shown that mutagenesis of these arginine residues leads to a red shift in the emission spectrum of the enzyme (*74*). Further, alteration of Arg218 but not Arg337 resulted in a 15−20-fold increase in the K_m values for luciferin. On the basis of data from their mutagenesis and fluorescence experiments, Branchini et al. suggested that Arg218 is required to maintain the rigidity and polarity of the emitter-binding site essential for the normal yellow−green emission of this luciferase (*74*). Accordingly, Arg218 (Figure 7) was chosen for replacement with unnatural analogues of arginine.

In contrast, there are no data concerning the functional role of the invariant Arg437, which is also near the putative active site, but the orientation of the hydrophilic arginine side chain in an otherwise hydrophobic region of the protein seems unlikely to be accidental. Accordingly, this position was also chosen for replacement with arginine analogues. In a preliminary experiment to assess whether modification at this position was likely to be enlightening, replacement of Arg437 by Lys resulted in a luciferase that emitted at longer wavelength (569 nm versus 559 nm for the wild type) and with only 12% of the efficiency of the wild type (data not shown).

In the context of these earlier studies, the present results offer some additional insight. The introduction of a methyl group on the guanidine moiety of Arg218 afforded an enzyme with unaltered (or sometimes enhanced) emission properties under standard assay conditions (pH 7.3) but that also resisted alteration in the emission wavelength as the pH was varied. This argues for the ability of the methylated Arg218 to stabilize the emitter-binding site more efficiently than arginine itself. In contrast, the enzymes containing homoarginine, citrulline, and N^{γ}-nitroarginine at position 218 all emitted light less efficiently and at a longer wavelength. Thus, the alteration of the nature (citrulline and N^{γ}-nitroarginine) and orientation (homoarginine) of the guanidine functional group in these modified enzymes was not well-tolerated by the enzyme. These findings support the suggestion by Branchini et al. (*74*) that Arg218 is a key structural element within the emitter-binding site.

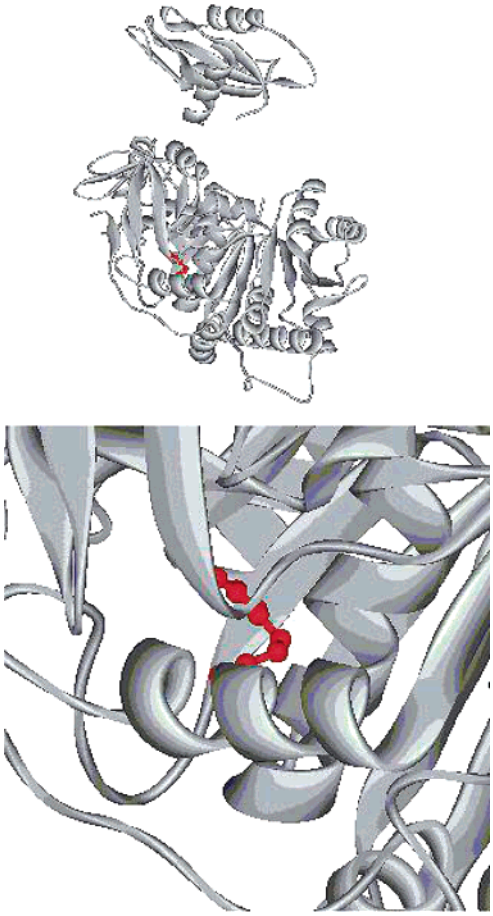


FIGURE 7: X-ray crystallographic structure of *P. pyralis* luciferase (*68*), illustrating the position of Arg218 (in red). Arg437 is part of the polypeptide that links the two folded domains and does not appear in the crystal structure because this region was disordered.

Alteration of Arg437 produced modified luciferases, all of which were much less efficient than the wild type in terms of total light emission. However, unlike the changes to position 218, none of the luciferases modified at position 437 resulted in an altered wavelength of the emitted light. This is entirely consistent with the role of Arg437 as a invariant residue essential for efficiency of enzyme function but not a residue thought to participate directly in the binding of luciferin to the enzyme.

In the aggregate, the present results affirm the utility of unnatural amino acid mutagenesis for the detailed dissection of enzyme function. Also described are experiments that provide ready access to a set of arginine analogues that can be used to explore the role of this important amino acid in protein function.

SUPPORTING INFORMATION AVAILABLE

Synthetic methods for all pentenoylaminoacyl pdCpA derivatives that do not appear in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Smith, M. (1985) *In vitro* mutagenesis, *Annu. Rev. Genet. 19*, 423−462.
2. Kunkel, T. A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection, *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.

3. Antikainen, N. M., and Martin, S. F. (2005) Altering protein specificity: Techniques and applications, *Bioorg. Med. Chem. 13*, 2701−2716.
4. Yuan, L., Kurek, I., English, J., and Keenan, R. (2005) Laboratory-directed protein evolution, *Microbiol. Mol. Biol. Rev. 69*, 373−392.
5. Hecht, S. M., Alford, B. L., Kuroda, Y., and Kitano, S. (1978) "Chemical aminoacylation" of tRNA's, *J. Biol. Chem. 253*, 4517−4520.
6. Heckler, T. G., Zama, Y., Naka, T., and Hecht, S. M. (1983) Dipeptide formation with misacylated $tRNA^{Phe}$s, *J. Biol. Chem. 258*, 4492−4495.
7. Heckler, T. G., Chang, L. H., Zama, Y., Naka, T., and Hecht, S. M. (1984) Preparation of 2′(3′)-*O*-acyl-pCpA derivatives as substrates for T4 RNA ligase-mediated "chemical aminoacylation", *Tetrahedron* 1984, *40*, 87−94.
8. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins, *Science 244*, 182−188.
9. Hecht, S. M. (2007) In *Protein Engineering* (RajBhandary, U. L., and Koehrer, C., Eds.) in press, Springer, New York.
10. Ellman, J. A., Mendel, D., and Schultz, P. G. (1992) Site-specific incorporation of novel backbone structures into proteins, *Science 255*, 197−200.
11. Ellman, J. A., Volkman, B. F., Mendel, D., Schultz, P. G., and Wemmer, D. E. (1992) Site-specific isotopic labeling of proteins for NMR studies, *J. Am. Chem. Soc. 114*, 7959−7961.
12. Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., and Schultz, P. G. (1992) Probing protein stability with unnatural amino acids, *Science 256*, 1798−1802.
13. Chung, H.-H., Benson, D. R., and Schultz, P. G. (1993) Probing the structure and mechanism of ras protein with an expanded genetic code, *Science 259*, 806−809.
14. Judice, J. K., Gamble, T. R., Murphy, E. C., deVos, A. M., and Schultz, P. G. (1993) Probing the mechanism of staphylococcal nuclease with unnatural amino acids: Kinetic and structural studies, *Science 261*, 1578−1581.
15. Mendel, D., Ellman, J. A., and Schultz, P. G. (1993) Protein biosynthesis with conformationally restricted amino acids, *J. Am. Chem. Soc. 115*, 4359−4360.
16. Cornish, V. W., Benson, D. R., Altenbach, C. A., Hideg, K., Hubbell, W. L., and Schultz, P. G. (1994) Site-specific incorporation of biophysical probes into proteins, *Proc. Natl. Acad. Sci. U.S.A. 91*, 2910−2914.
17. Thorson, J. S., Chapman, E., Murphy, E. C., Schultz, P. G., and Judice, J. K. (1995) Linear free energy analysis of hydrogen bonding in proteins, *J. Am. Chem. Soc. 117*, 1157−1158.
18. Cornish, V. W., Mendel, D., and Schultz, P. G. (1995) Probing protein structure and function with an expanded genetic code, *Angew. Chem., Int. Ed. Engl. 34*, 621−633.
19. Nowak, M. W., Kearney, P. C., Sampson, J. R., Saks, M. E., Labarca, C. G., Silverman, S. K., Zhong, W., Thorson, J., Abelson, J. N., Davidson, N., Schultz, P. G., Dougherty, D. A., and Lester, H. A. (1995) Nicotinic receptor binding site probed with unnatural amino acid incorporation in intact cells, *Science 268*, 439−442.
20. Mamaev, S. V., Laikhter, A. L., Arslan, T., and Hecht, S. M. (1996) Firefly luciferase: Alteration of the color of emitted light resulting from substitutions at position 286, *J. Am. Chem. Soc. 118*, 7243−7244.
21. Saks, M. E., Sampson, J. R., Nowak, M. W., Kearney, P. C., Du, F., Abelson, J. N., Lester, H. A., and Dougherty, D. A. (1996) An engineered $tRNA^{Gln}$ for *in vivo* incorporation of unnatural amino acids into proteins by nonsense suppression, *J. Biol. Chem. 271*, 23169−23175.
22. Steward, L. E., Collins, C. S., Gilmore, M. A., Carlson, J. E., Ross, J. B. A., and Chamberlin, A. R. (1997) *In vitro* site-specific incorporation of fluorescent probes into β-galactosidase, *J. Am. Chem. Soc. 119*, 6−11.
23. Arslan, T., Mamaev, S. V., Mamaeva, N. V., and Hecht, S. M. (1997) Structurally modified firefly luciferase. Effects of amino acid substitution at position 286, *J. Am. Chem. Soc. 119*, 10877−10887.
24. England, P. M., Lester, H. A., Davidson, N., and Dougherty, D. A. (1997) Site-specific, photochemical proteolysis applied to ion channels *in vivo*, *Proc. Natl. Acad. Sci. U.S.A. 94*, 11025−11030.
25. Hohsaka, T., Ashizuka, Y., Taira, H., Murakami, H., and Sisido, M. (2001) Incorporation of nonnatural amino acids into proteins by using various four-base codons in an *Escherichia coli in vitro* translation system, *Biochemistry 40*, 11060−11064.
26. Lu, T., Ting, A. Y., Mainland, J., Jan, I. Y., Schultz, P. G., and Yang, J. (2001) Probing ion permeation and gating in a K^+ channel with backbone mutations in the selectivity filter, *Nature Neurosci. 4*, 239−246.
27. Wang, B., Brown, K. C., Lodder, M., Craik, C. S., and Hecht, S. M. (2002) Chemically mediated site-specific proteolysis. Alteration of protein−protein interaction, *Biochemistry 41*, 2805−2813.
28. Gao, R., Zhang, Y., Choudhury, A. K., Dedkova, L. M., and Hecht, S. M. (2005) Analogues of vaccinia virus DNA topoisomerase I modified at the active site tyrosine, *J. Am. Chem. Soc. 127*, 3321−3331.
29. Gao, R., Zhang, Y., Dedkova, L., Choudhury, A. K., Rahier, N. J., and Hecht, S. M. (2006) Effects of modification of the active site tyrosine of human DNA topoisomerase I, *Biochemistry 45*, 8402−8410.
30. Karginov, V. A., Mamaev, S. V., An, A., Van Cleve, M. D., Hecht, S. M., Komatsoulis, G. A., and Abelson, J. N. (1997) Probing the role of an active site aspartic acid in dihydrofolate reductase, *J. Am. Chem. Soc. 119*, 8166−8176.
31. Graveley, B. R., and Maniatis, T. (1998) Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA, *Mol. Cell 1*, 765−771.
32. Zapp, M. L., Hope, T. J., Parslow, T. G., and Green, M. R. (1991) Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus rev protein: A dual function for an arginine-rich binding motif, *Proc. Natl. Acad. Sci. U.S.A. 88*, 7734−7738.
33. Truant, R., and Cullen, B. R. (1999) The arginine-rich domains present in human immunodeficiency virus type 1 tat and rev function as direct importin β-dependent nuclear localization signals, *Mol. Cell. Biol. 19*, 1210−1217.
34. Ruoslahti, E., and Pierschbacher, M. D. (1987) New perspectives in cell-adhesion−RGD and integrins, *Science 238*, 491−497.
35. Xuan, J. W., Hota, C., Shigeyama, Y., D'Errico, J. A., Somerman, M. J., and Chambers, A. F. (1995) Site-directed mutagenesis of the arginine−glycine−aspartic acid sequence in osteopontin destroys cell-adhesion and migration functions, *J. Cell. Biochem. 57*, 680−690.
36. Savage, B., Marzec, U. M., Chao, B. H., Harker, L. A., Maraganore, J. M., and Ruggeri, Z. M. (1990) Binding of the snake venom-derived proteins applaggin and echistatin to the arginine−glycine−aspartic acid recognition site(s) on platelet glycoprotein IIb·IIIa complex inhibits receptor function, *J. Biol. Chem. 265*, 11766−11772.
37. Shen, E. C., Henry, M. F., Weiss, V. H., Valentini, S. R., Silver, P. A., and Lee, M. S. (1998) Arginine methylation facilitates the nuclear export of hnRNP proteins, *Genes Devel. 12*, 679−691.
38. Yun, C. Y., and Fu, X.-D. (2000) Conserved Sr protein kinase functions in nuclear import and its action, is counteracted by arginine methylation in *Saccharomyces cerevisiae*, *J. Cell Biol. 150*, 707−717.
39. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications, *Nature 403*, 41−45.
40. Lee, Y. H., Coonrod, S. A., Kraus, W. L., Jelinek, M. A., and Stallcup, M. R. (2005) Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination, *Proc. Natl. Acad. Sci. U.S.A. 102*, 3611−3616.
41. Rothnagel, J. A., and Rogers, G. E. (1984) Citrulline in proteins from the enzymatic deimination of arginine residues, *Methods Enzymol. 107*, 624−631.
42. Asaga, H., and Ishigami, A. (2001) Protein deimination in the rat brain after kainate administration: Citrulline-containing proteins as a novel marker of neurodegeneration, *Neurosci. Lett. 299*, 5−8.
43. Asaga, H., Akiyama, K., Ohsawa, T., and Ishigami, A. (2002) Increased and type II-specific expression of peptidylarginine deiminase in activated microglia but not hyperplastic astrocytes following kainic acid-evoked neurodegeneration in the rat brain, *Neurosci. Lett. 326*, 129−132.
44. Ishigami, A., Ohsawa, T., Hiratsuka, M., Taguchi, H., Kobayashi, S., Saito, Y., Murayama, S., Asaga, H., Toda, T., Kimura, N., and Maruyama, N. (2005) Abnormal accumulation of citrullinated proteins catalyzed by peptidylarginine deiminase in hippocampal extracts from patients with Alzheimer's disease, *J. Neurosci. Res. 80*, 120−128.
45. Siddiqui, K. S., Poljak, A., Guilhaus, M., De Francisci, D., Curmi, P. M. G., Feller, G., D'Amico, S., Gerday, C., Uversky, V. N., and Cavicchioli, R. (2006) Role of lysine versus arginine in enzyme cold-adaptation: Modifying lysine to homo-arginine

stabilizes the cold-adapted α-amylase from *Pseudoalteramonas haloplanktis*, *Proteins 64*, 486−501.
46. Lodder, M. (2000) Ph.D. Thesis, University of Virginia, Charlottesville, Virginia.
47. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
48. Robertson, S. A., Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) The use of 5′-phospho-2-deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of transfer RNA, *Nucleic Acids Res. 17*, 9649−9660.
49. Webb, T. R., and Eigenbrot, C. (1991) Conformationally restricted arginine analogs, *J. Org. Chem. 56*, 3009−3016.
50. Kim, K., Lin, Y.-T., and Mosher, H. S. (1988) Monosubstituted guanidines from primary amines and aminoiminomethanesulfonic acid, *Tetrahedron Lett. 29*, 3183−3186.
51. Nazarenko, I. A., Harrington, K. M., and Uhlenbeck, O. C. (1994) Many of the conserved nucleotides of $tRNA^{Phe}$ are not essential for ternary complex formation and peptide elongation, *EMBO J. 13*, 2464−2471.
52. Varshney, U., Lee, C. P., and RajBhandary, U. L. (1991) Direct analysis of aminoacylation levels of tRNAs *in vivo* application to studying recognition of *Escherichia coli* initiator tRNA mutants by glutaminyl-tRNA synthetase, *J. Biol. Chem. 266*, 24712−24718.
53. Dedkova, L. M., Fahmi, N. E., Golovine, S. Y., and Hecht, S. M. (2003) Enhanced D-amino acid incorporation into protein by modified ribosomes, *J. Am. Chem. Soc. 125*, 6616−6617.
54. Short, G. F., III, Golovine, S. Y., and Hecht, S. M. (1999) Effects of release factor 1 on *in vitro* protein translation and the elaboration of proteins containing unnatural amino acids, *Biochemistry 38*, 8808−8819.
55. Bodanszky, M., and Bodanszky, A. (1994) In *The Practice of Peptide Synthesis*, 2nd ed., p 96, Springer-Verlag, New York.
56. Noren, C. J., Anthony-Cahill, S. J., Suich, D. J., Noren, K. A., Griffith, M. C., and Schultz, P. G. (1990) *In vitro* suppression of an amber mutation by a chemically aminoacylated transfer RNA prepared by runoff transcription, *Nucleic Acids Res. 18*, 83−88.
57. Lodder, M., Golovine, S., and Hecht, S. M. (1997) Chemical deprotection strategy for the elaboration of misacylated transfer RNA's, *J. Org. Chem. 62*, 778−779.
58. Lodder, M., Golovine, S., Laikhter, A. L., Karginov, V. A., and Hecht, S. M. (1998) Misacylated transfer RNAs having a chemically removable protecting group, *J. Org. Chem. 63*, 794−803.
59. Groff, J. P., London, R. E., Cocco, L., and Blakley, R. L. (1981) Nuclear magnetic resonance study of dihydrofolate reductase labeled with [γ-^{13}C]tryptophan, *Biochemistry 20*, 6169−6178.
60. Hoeltzli, S. D., and Frieden, C. (1994) ^{19}F NMR spectroscopy of [6-^{19}F]tryptophan-labeled *Escherichia coli* dihydrofolate reductase: Equilibrium folding and ligand binding studies, *Biochemistry 33*, 5502−5509.
61. Basran, J., Casarotto, M. G., Basran, A., and Roberts, G. C. K. (1997) Effects of single-residue substitutions on negative cooperativity in ligand binding to dihydrofolate reductase, *Protein Eng. 10*, 815−826.
62. Ohmae, E., Sasaki, Y., and Gekko, K. (2001) Effects of five tryptophan mutations on structure, stability and, function of *Escherichia coli* dihydrofolate reductase, *J. Biochem. 130*, 439−447.
63. Xu, Y., Feller, G., Gerday, C., and Glansdorff, N. (2003) Moritella cold-active dihydrofolate reductase: Are there natural limits to optimization of catalytic efficiency at low temperature? *J. Bacteriol. 185*, 5519−5526.
64. Gould, S. J., and Subramani, S. (1988) Firefly luciferase as a tool in molecular and cell biology, *Anal. Biochem. 175*, 5−13.
65. Kricka, L. J. (1995) Chemiluminescence and bioluminescence, *Anal. Chem. 67*, 499R−502R.
66. Price, R. L., Squirrell, D. J., and Murphy, M. J. (1998) Luciferase and recombinant luciferase labels, *J. Clin. Ligand Assay 21*, 349−357.
67. Kricka, L. J. (2000) Application of bioluminescence and chemiluminescence in biomedical sciences, *Methods Enzymol. 305*, 333−345.
68. Conti, E., Franks, N. P., and Brick, P. (1996) Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes, *Structure 4*, 287−298.
69. Franks, N. P., Jenkins, A., Conti, E., Lieb, W. R., and Brick, P. (1998) Structural basis for the inhibition of firefly luciferase by a general anesthetic, *Biophys. J. 75*, 2205−2211.
70. Branchini, B. R., Magyar, R. A., Murtiashaw, M. H., Anderson, S. M., and Zimmer, M. (1998) Site-directed mutagenesis of histidine 245 in firefly luciferase: A proposed model of the active site, *Biochemistry 37*, 15311−15319.
71. Sandalova, T. P., and Ugarova, N. N. (1999) Model of the active site of firefly luciferase, *Biochemistry* (Moscow) *64*, 962−967.
72. Branchini, B. R., Magyar, R. A., Murtiashaw, M. H., Anderson, S. M., Helgerson, L. C., and Zimmer, M. (1999) Site-directed mutagenesis of firefly luciferase active site amino acids: A proposed model for bioluminescence color, *Biochemistry 38*, 13223−13230.
73. Branchini, B. R., Murtiashaw, M. H., Magyar, R. A., and Anderson, S. M. (2000) The role of lysine 529, a conserved residue of the acyl-adenylate-forming enzyme superfamily, in firefly luciferase, *Biochemistry 39*, 5433−5440.
74. Branchini, B. R., Magyar, R. A., Murtiashaw, M. H., and Portier, N. G. (2001) The role of active site residue arginine 218 in firefly luciferase bioluminescence, *Biochemistry 40*, 2410−2418.

BI062042R