Biosynthesis of Proteins Incorporating a Versatile Set of Phenylalanine Analogues

Kent Kirshenbaum, Isaac S. Carrico, and David A. Tirrell*[a]

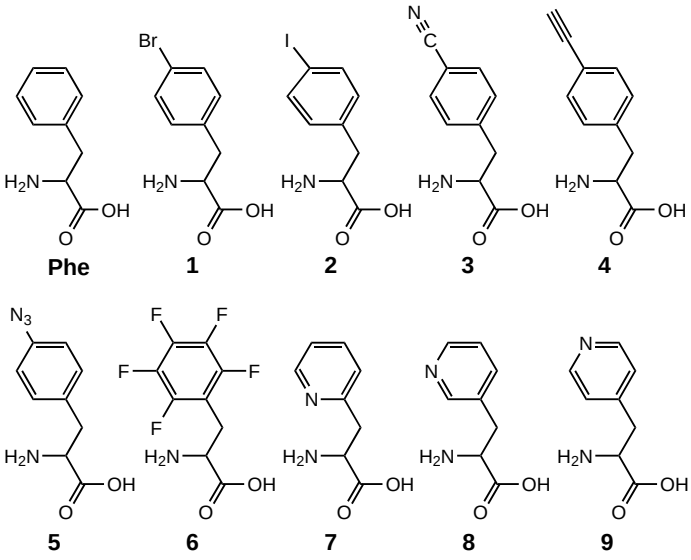
KEYWORDS:

amino acids · biosynthesis · mutagenesis · photoaffinity labeling · protein modifications

Macromolecular chemistry faces a dichotomy. Chemists can prepare polymers with a wide variety of functional groups, but cannot attain the sequence-specificity and monodispersity of proteins and nucleic acids. Conversely, the chemical diversity of proteins is severely constrained by the small number of amino acids specified by the genetic code. Can we find ways to combine the diversity of synthetic polymer chemistry with the precision of protein biosynthesis?

One approach is to enhance the capability of the protein biosynthetic apparatus to utilize monomers other than the twenty canonical amino acids.[1, 2] Particular attention has been focused on the aminoacyl-tRNA synthetases (aaRS), which conjugate amino acids to their cognate tRNAs. The specificity of tRNA charging is pivotal for ensuring the fidelity of translation of genetic information into protein sequence.[3] Techniques have been developed for engineering aaRS to catalyze acylation of tRNA by amino acid analogues, facilitating incorporation of novel side chains into recombinant proteins in vivo.[4, 5] Herein we describe the elaboration of the use of a mutant form of the *Esherichia coli* phenylalanyl-tRNA synthetase (Ala 294 → Gly; termed PheRS*), which has an enlarged substrate binding pocket[4] and which has been shown to effect incorporation of *p*-bromophenylalanine (**1**) into a recombinant protein expressed in a bacterial host.[4c, 6] We now find that *p*-iodo-, *p*-cyano-, *p*-ethynyl-, and *p*-azido-phenylalanine (**2** – **5**) and 2-, 3-, and 4-pyridylalanine (**7** – **9**) can also be substituted for Phe in bacterial hosts outfitted with PheRS*.

The *E. coli* strain AF-IQ[pQE-FS] is a Phe auxotroph that harbors a plasmid encoding PheRS* and the test protein murine dihydrofolate reductase (DHFR).[6] Figure 1 shows sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of cell lysates from 10 mL cultures of AF-IQ[pQE-FS] following induction of DHFR expression in minimal media supplemented with Phe or with one of the analogues **1** – **9**. Expression of DHFR is evident in all cultures except that supplemented with pentafluorophenylalanine (**6**) and the negative control lacking supplementation. Cultures of the control strain[6] AF-IQ[pQE-15], which lacks the gene for PheRS*, showed efficient target protein expression only in media supplemented with Phe or with one of the isosteric analogues **7** – **9** (data not shown).



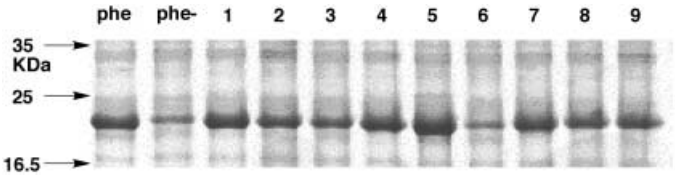


Figure 1. *SDS-PAGE of cell lysates of AF-IQ[pQE-FS] 4 hr after induction with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Efficient expression of target protein DHFR (24 kDa) is observed for cultures supplemented with Phe or with one of the analogues **1** – **5** or **7** – **9**. The lane labeled Phe – is the control lane which lacked supplementation.*

Cell lysates were subjected to nickel-affinity chromatography for purification of DHFR through binding to an N-terminal hexahistidine tag (Qiagen). Amino acid analyses demonstrated that the extent of analogue substitution for Phe in DHFR coexpressed with PheRS* varied between 45 and 90% (Table 1).

Table 1. *Extent of substitution of Phe by analogues **2** – **9** in DHFR coexpressed with a mutant Phe-tRNA synthetase (PheRS*) or expressed in a control strain (wild-type PheRS), as determined by amino acid analysis.*

Phe analogue	% Substitution	
	PheRS*	wild-type PheRS
2	45	n.d.[a]
3	48	n.d.
4	62	n.d.
5	67	n.d.
6	n.d.	n.d.
7	77	81
8	90	90
9	89	84

[a] n.d.: not detected.

[a] *Prof. D. A. Tirrell, Dr. K. Kirshenbaum, I. S. Carrico*
Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, CA 91125 (USA)
Fax: (+1) 626-793-8472
E-mail: tirrell@caltech.edu

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

In agreement with the SDS-PAGE analysis, analogue **6** was not detected. Only Phe and analogues **7** – **9** were detected in samples of DHFR expressed in the control strain lacking PheRS*.

Incorporation of Phe analogues was confirmed by tryptic digestion of purified DHFR followed by analysis of the resultant peptide fragments by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry. For DHFR prepared in Phe-supplemented media, two peptides with masses between 1550 and 1820 Daltons were observed and assigned to residues 34 – 47 and 93 – 106, respectively (Figure 2 a). Each of these fragments includes one of the nine Phe residues of DHFR. The corresponding mass spectra of tryptic peptides incorporating analogues **2** – **5** and **7** – **9** showed additional signals consistent with the increased masses of the analogues relative to Phe (Figure 2 b, Table 2). No new peaks were observed in the spectrum of DHFR expressed in media supplemented with **6**, as anticipated.

Table 2. *MALDI-TOF data for tryptic peptides derived from DHFR expressed in media supplemented with Phe or analogues **2** – **9**. Values shown are for major peaks between 1550 and 1820 Da.*

Amino acid	*m/z* Peptide 1	*m/z* Peptide 2	Δ*m/z* Observed (calculated)
Phe	1591.93	1681.88	
2	1592.69, 1718.58	1682.65, 1808.53	125.89, 125.88 (125.90)
3	1592.84, 1617.82	1682.80, 1707.80	24.98, 25.00 (25.00)
4	1592.75, 1616.74	1682.72, 1706.72	23.99, 23.99 (24.00)
5	1591.93, 1606.99	1681.91, 1696.91	15.06, 15.00[a] (41.00)
6	1592.67	1682.63	n.d.[b] (89.95)
7	1592.87, 1593.86	1682.84, 1683.84	0.99, 1.00 (1.00)
8	1592.84, 1593.84	1682.82, 1683.81	1.00, 0.99 (1.00)
9	1592.79, 1593.78	1682.76, 1683.76	0.99, 1.00 (1.00)

[a] Masses observed for **5** are consistent with photodecomposition to the aryl nitrene upon laser irradiation of the MALDI sample. [b] n.d.: not detected.

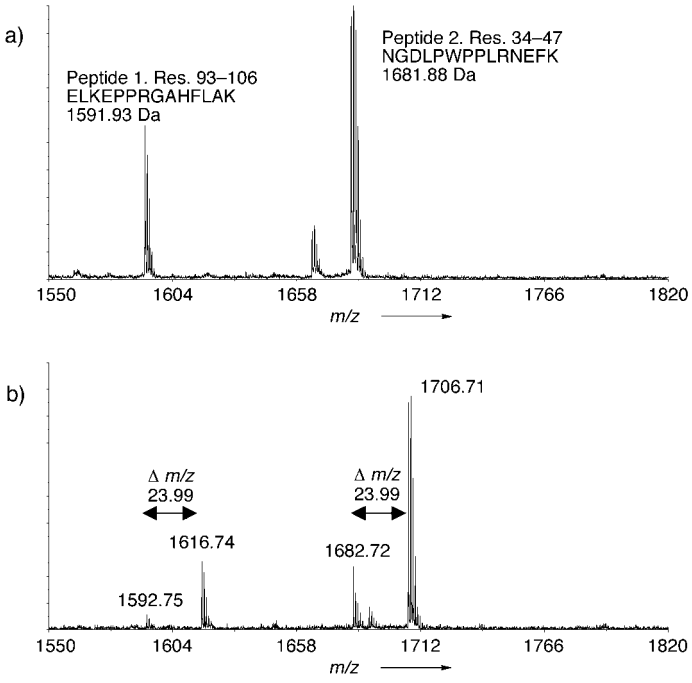


Figure 2. *MALDI-TOF mass spectra of tryptic peptides derived from DHFR expressed in media supplemented with a) Phe or b) analogue **4**. Two prominent mass peaks in (a) correspond to peptides 34 – 47 and 93 – 106, each containing one Phe residue. Two new mass peaks are observed in (b) with a Δm/z of 23.99, consistent with the increased mass of **4** relative to Phe.*

Large scale expressions were similarly performed in 0.1 L cultures of AF-IQ[pQE-FS] in media supplemented with Phe or with one of the translationally active analogues **2** – **5** or **7** – **9**. The resultant purified proteins were termed DHFR-Phe, DHFR-**2**, etc. Yields were in the range of 6 – 18 mg L^{-1}, as determined by a dye-binding assay (BioRad) with DHFR-Phe used as a calibration standard. The UV absorption spectra of DHFR solutions prepared under denaturing conditions were obtained (Figure 3). Samples containing analogues with extended conjugation showed enhanced absorption in the region between 240 and 280 nm. New absorption maxima were observed for solutions of DHFR-**4**, -**5**, -**7**, -**8**, and -**9**. The positions and intensities of the maxima were consistent with the UV spectra of the free amino acid analogues; this indicates that the novel functional groups were not modified by the bacterial host or by photodegradation.[7, 8]

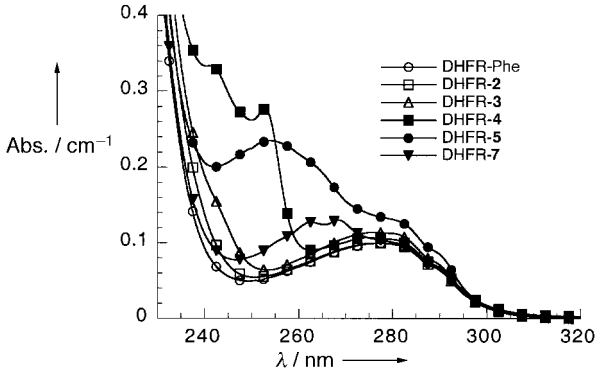


Figure 3. *UV spectra of purified DHFR expressed in media supplemented with Phe or with one of the analogues **2** – **5** or **7**. Spectra were obtained in buffer (pH 4.5) containing 8 M urea, 100 mM NaH_2PO_4, 10 mM tris(hydroxymethyl)aminomethane (Tris), at 25 °C with 4.2 μM protein. Spectra for DHFR-**8** and -**9** were similar to -**7**, with some variation of peak positions.*

An additional characteristic feature of DHFR-**3** was revealed in the Fourier transform IR (FT-IR) spectrum of a dried film of the protein. Vibrational excitation of the nitrile group can give rise to absorption in a region of the IR spectrum typically regarded as "silent" with respect to protein chromophores. The FT-IR spectrum of DHFR-**3** displayed a distinct peak at 2228 cm^{-1}, consistent with the presence of the aryl nitrile; no such peak was evident in this region of the FT-IR spectrum of DHFR-Phe.

The above results demonstrate the biosynthesis of proteins incorporating chemical functionality not typically present in biological macromolecules. Introduction of such functional groups should enable a variety of new techniques in structural biology, proteomics, biomaterials science, and bioconjugate chemistry.

Analogues **1** – **5** and **7** – **9** display distinct photophysical properties in the X-ray, UV, and IR regions that may facilitate techniques such as phasing of crystallographic diffraction data,

  1439-4227/02/03/02-03 $ 17.50+.50/0 ChemBioChem **2002**, No. 02-03

rapid screening of protein ligands, and biophysical studies by vibrational spectroscopy. In particular, the aryl azide **5** provides an intrinsic capacity for intramolecular photoactivated cross-linking and intermolecular photoaffinity labeling.[9] Proteins bearing ethynyl- and halo-aryl groups are subject to palladium-mediated coupling reactions that are orthogonal to existing methods for protein modification.[10]

The extent of structural and functional perturbation caused by analogue incorporation is currently under investigation in a variety of protein systems. In those cases where such perturbation is problematic, the strategy reported here will permit partial replacement of phenylalanine by the analogue of choice,[6] followed by affinity selection of properly folded species. In addition, this study should lead to new methods for site-specific incorporation of nonnatural amino acids in vivo.

This work was supported by the National Institutes of Health through a grant (no.: RO1-GM62523-01), a postdoctoral fellowship to K.K. (no.: F32-GM29474), and a training grant stipend to I.S.C. (no.: T32-GM08501). We thank Nandita Sharma, Ian Suydam, Steven Boxer, Melanie Bennett, Mona Shahgholi, Hanna Rapaport, and Yi Tang for their advice and assistance.

Experimental Section

Materials: **1**, **2**, and **5** were obtained from Chem-Impex. **3**, **6, 7**, **8** and **9** were obtained from PepTech. **4** was synthesized as described by Kayser et al.[11] The pQE-FS expression plasmid is derived from pQE-15 (Qiagen) and encodes, in addition to DHFR, a mutant form of the α-subunit of *E. coli* PheRS (Ala 294 $\rightarrow$Gly) under control of a *lac* promoter.[6]

Protein expression: Cultures of AF-IQ[pQE-FS] and AF-IQ[pQE-15] were grown in M9 minimal medium supplemented with glucose (0.2 wt %), thiamine (5 mg L^{-1}), $MgSO_4$ (1 mM), $CaCl_2$ (0.1 mM), 20 amino acids (20 mg L^{-1} Phe, 40 mg L^{-1} other amino acids), and antibiotics (ampicillin and chloramphenicol). At an optical density of 0.8 – 1.0 at 600 nm (OD_{600}), the cultures were sedimented by centrifugation for 10 min (3000 *g*) at 4 °C and the cell pellets were washed twice with NaCl (0.9 wt %). The cells were resuspended in M9 minimal medium as above, but without chloramphenicol or Phe. Aliquots were transferred to culture flasks into which one of the amino acid supplements was added: L-Phe, L-**1**, L-**2**, L-**3**, L-**6**, L-**7**, L-**8**, or L-**9** (0.25 g L^{-1}); D,L-**4** or D,L-**5** (0.5 g L^{-1}); or no additional supplementation.

After a 10 min incubation, IPTG (1mM) was added to induce protein expression. The OD_{600} of the cultures was determined 4 h after induction, and the cells were harvested by centrifugation. The cells were lysed in buffer (pH 8) containing urea (8 M), NaH_2PO_4 (100 mM), and Tris (10 mM) and subjected to a freeze/thaw cycle. Protein expression was evaluated by Tricine SDS-PAGE with Coomassie blue staining. Loading of the gel was normalized for cell densities as determined by OD_{600}. The target proteins were purified by nickel-affinity chromatography on nickel – nitrilotriacetate (Ni-NTA) resin following the manufacturer's protocols (Qiagen). The target protein was eluted in buffer (pH 4.5) containing urea (8 M), NaH_2PO_4 (100 mM), and Tris (10 mM).

Amino acid analyses: Purified DHFR solutions were subjected to exchange of buffer against water by ultrafiltration (Millipore Ultrafree, molecular weight cut-off of 5 000). Samples were supplied to the Molecular Structure Facility at the University of California, Davis, for analyses on a Beckman 6300 instrument with Li cation exchange based columns and buffers (Pickering). Standard chromatograms of all Phe analogues were obtained before and after application of an HCl (6 N) hydrolysis solution. Quantitation was by reference to the standard chromatograms of the hydrolysis products.

Tryptic digestion and MALDI-TOF mass spectrometry: An aliquot (12.5 μL) of protein in elution buffer containing urea (8 M), NaH_2PO_4 (100 mM), and Tris (10 mM) at pH 4.5 was added to NH_4OAc solution (112.5 μL, 50 mM). Modified trypsin (Promega; 2 μL, 0.2 g L^{-1}) was added, and the solution was allowed to stand at room temperature overnight. Trifluoroacetic acid (0.1 M) was used to quench the reaction. Chromatography on $ZipTip_{C18}$ columns (Millipore) provided purified peptide samples (2 μL), which were added to a α-cyano-β-hydroxycinnamic acid MALDI matrix (10 μL, 10 g L^{-1} in 1:1 H_2O/CH_3CN). The samples were analyzed on an Applied Biosystems Voyager DE Pro instrument.

[1] For examples utilizing chemical misacylation of tRNA, see: a) S. M. Hecht, *Acc. Chem. Res.* **1992**, *25*, 545 – 552; b) J. D. Bain, C. G. Glabe, T. A. Dix, A. R. Chamberlin, E. S. Diala, *J. Am. Chem. Soc.* **1989**, *111*, 8013 – 8014; c) D. Mendel, V. W. Cornish, P. G. Schultz, *Angew. Chem.* **1995**, *107*, 677 – 690; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 621 – 633; d) M. W. Nowak, J. P. Gallivan, S. K. Silverman, C. G. Labarca, D. A. Dougherty, H. A. Lester, *Methods Enzymol. Part B* **1998**, *293*, 504 – 529.

[2] For examples utilizing the inherent permissiveness of wild type aaRS in vivo, see a) D. B. Cowie, G. N. Cohen, *Biochim. Biophys. Acta* **1957**, *26*, 252 – 261; b) M. H. Richmond, *Bacteriol. Rev.* **1962**, *26*, 398 – 420; c) G. Hortin, I. Boime, *Methods Enzymol.* **1983**, *96*, 777 – 784; d) M. J. Wilson, D. L. Hatfield, *Biochim. Biophys. Acta* **1984**, *781*, 205 – 215; e) N. Budisa, C. Minks, F. J. Medrano, J. Lutz, R. Huber, L. Moroder, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 455 – 459; f) K. L. Kiick, J. C. M. van Hest, D. A. Tirrell, *J. Am. Chem. Soc.* **2000**, *122*, 1282 – 1288; g) Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado, D. A. Tirrell, *Angew. Chem.* **2001**, *113*, 1542 – 1544; *Angew. Chem. Int. Ed.* **2001**, *40*, 1494 – 1496.

[3] M. Ibba, D. Söll, *Science* **1999**, *286*, 1893 – 1897.

[4] a) P. Kast, H. Hennecke, *J. Mol. Biol.* **1991**, *222*, 99 – 124; b) M. Ibba, P. Kast, H. Hennecke, *Biochemistry* **1994**, *33*, 7107 – 7112; c) M. Ibba, H. Hennecke, *FEBS Lett.* **1995**, *364*, 272 – 275; d) C. Behrens, J. J. Nielsen, X.-J. Fan, X. Doisy, K.-H. Kim, M. Praetorius-Ibba, P. E. Nielsen, M. Ibba, *Tetrahedron* **2000**, *56*, 9443 – 9449.

[5] a) R. Furter, *Protein Sci.* **1998** *7*, 419 – 426; b) L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498 – 500; c) V. Doring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. de Crecy-Lagard, P. Schimmel, P. Marliere, *Science* **2001**, *292*, 501 – 504.

[6] N. Sharma, R. Furter, P. Kast, D. A. Tirrell, *FEBS Lett.* **2000**, *467*, 37 – 40.

[7] Behrens and co-workers[4d] have recently reported that **5** inhibits the synthesis of Phe-$tRNA^{Phe}$ by PheRS*, and suggested that **5** should be accommodated in the active site of the mutant synthetase.

[8] Cultures supplemented with the photoreactive analogue **5** were maintained in the dark. Protein purification was performed under diminished light conditions.

[9] S. A. Fleming, *Tetrahedron* **1995**, *51*, 12 479 – 12 520.

[10] J. P. Genet, M. J. Savignac, *J. Organomet. Chem.* **1999**, *576*, 305 – 317.

[11] B. Kayser, J. Altman, W. Beck, *Tetrahedron* **1997**, *53*, 2475 – 2484.

Received: October 10, 2001 [Z 306]