Published in final edited form as:

J Biomol NMR. 2012 March ; 52(3): 191–195. doi:10.1007/s10858-012-9606-9.

An efficient protocol for incorporation of an unnatural amino acid in perdeuterated recombinant proteins using glucose-based media

Vincenzo Venditti, Nicolas L. Fawzi, and G. Marius Clore*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

Abstract

The *in vivo* incorporation of unnatural amino acids into proteins is a well-established technique requiring an orthogonal tRNA/aminoacyl-tRNA synthase pair specific for the unnatural amino acid that is incorporated at a position encoded by a TAG amber codon. Although this technology provides unique opportunities to engineer protein structures, poor protein yields are usually obtained in deuterated media, hampering its application in the protein NMR field. Here, we describe a novel protocol for incorporating unnatural amino acids into fully deuterated proteins using glucose-based media (which are relevant to the production, for example, of amino acid-specific methyl-labeled proteins used in the study of large molecular weight systems). The method consists of pre-induction of the pEVOL plasmid encoding the tRNA/aminoacyl-tRNA synthetase pair in a rich, H₂O-based medium prior to exchanging the culture into a D₂O-based medium. Our protocol results in high level of isotopic incorporation (~95%) and retains the high expression level of the target protein observed in Luria-Bertani medium.

Keywords

Site-specific labeling; Isotopic labeling; Unnatural amino acids; spin-labeling

In vivo incorporation of unnatural amino acids into proteins (Wang et al., 2001; Liu et al., 2007) has significantly expanded frontiers in protein engineering. To date, over 50 unnatural amino acids have been genetically encoded (Liu and Schultz, 2006; Xie and Schultz, 2006; Brustad et al., 2008; Brustad et al., 2008; Cellitti et al., 2008), permitting the manipulation of physicochemical, biological, and pharmacological properties of proteins with exquisite control over structure.

The advantages of creating proteins with unnatural amino acids for NMR studies have been recently discussed (Jones et al., 2010). For example, the introduction of unnatural amino acids bearing post-translational modifications (Liu and Schultz, 2006; Xie et al., 2007; Guo et al., 2008) can be particularly important for structural and dynamic investigations of signaling pathways. In addition, site-specific incorporation of NMR active (¹H, ¹³C, ¹⁵N and ¹⁹F) unnatural amino acids into silent backgrounds dramatically reduces spectral complexity, thereby facilitating the study of large proteins by *in-vitro* and *in-vivo* NMR techniques (Cellitti et al., 2008; Jones et al., 2010; Lampe et al., 2010; Li et al., 2010; Watt et al., 2011). Particularly attractive are methods that employ unnatural amino acids to design binding sites for paramagnetic metal ions (Lee and Schultz, 2008; Lee et al., 2009) or nitroxide spin-labels (Fleissner et al., 2009). Indeed, incorporation of unnatural amino acids

*Corresponding authors: mariusc@mail.nih.gov.

having bio-orthogonal chemical functionality in the context of the native protein can facilitate production of site-specific, paramagnetically-labeled samples, critical for the characterization of transient, sparsely-populated states (Clore and Iwahara, 2009) and very valuable for resonance assignment (Pintacuda et al., 2004; Venditti et al., 2011) in large proteins and protein complexes.

Incorporation of unnatural amino acids is accomplished by co-expressing a bio-orthogonal tRNA/aminoacyl-tRNA synthetase pair specific for the relevant unnatural amino acid together with the protein of interest whose DNA sequence has been modified by substitution of a TAG amber codon at the site of desired incorporation (Wang et al., 2001). However, lower protein yields are generally obtained by this method owing to translation termination. Thus, to reduce truncation problems, an optimized *E. coli* incorporation system consisting of a plasmid that overexpresses the aminoacyl-tRNA synthetase at the time of induction has been developed (Cellitti et al., 2008), and the effects of the degree, timing, and temperature of induction have been extensively tested (Young et al., 2010). These studies demonstrated that the simultaneous overexpression of the aminoacyl-tRNA synthetase and the target protein at 30°C dramatically reduces truncation problems, enabling production of sufficient quantities of protein for NMR studies using minimal or rich media (Cellitti et al., 2008; Young et al., 2010). More recently, the expression of the protein FAS-TE incorporating the unnatural amino acid *p*-methoxy-phenylalanine (*p*-OMePhe) in deuterated M9 medium supplemented with 0.5% d-glycerol and 0.1 mM FeCl₃ has been also reported (Jones et al., 2010). However, the yield of the protein was modest (~25% of the level obtained in rich medium) and, to the best of our knowledge, no other successful incorporation of unnatural amino acids into deuterated proteins have been reported to date.

Here, we investigate the expression levels for a mutant (C575A) of the C-terminal domain (A261-C575) of the *E. coli* enzyme I (EIC) of the bacterial phosphotransfer system (PTS), incorporating the unnatural amino acid *p*-acetyl-L-phenylalanine (*p*AcF) at position M469 in protonated and deuterated media (Table 1). *p*AcF is particularly relevant to the NMR community since it reacts quantitatively with a new hydroxylamine functionalized nitroxide spin-label, known as HO-4120 (Fleissner et al., 2009), enabling production of site-directed, paramagnetically-labeled proteins even in the presence of surface exposed Cys residues which interfere with standard site-directed spin-labeling techniques. We focused our attention on glucose-based deuterated media since these are routinely employed for the production of amino acid-specific, methyl labeled proteins (Tugarinov et al., 2006) that are extremely useful in NMR studies of high molecular weight systems (Mittermaier and Kay, 2006).

We initially followed the optimized protocol for incorporation of *p*AcF into proteins using a protonated medium (Young et al., 2010). Briefly, BL21 star (DE3) *E. coli* cells were co-transformed with pET11a/EIC (containing the desired TAG mutant) and the pEVOL plasmid that contains the orthogonal tRNA and aminoacyl-tRNA synthetase specific for *p*AcF (Young et al., 2010). Transformed bacteria were then plated onto an LB-agar plate containing ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml) for selection. A single-colony was used to inoculate 50 ml of Luria-Bertani (LB) medium containing the aforementioned antibiotics and grown overnight at 37°C. The following day the starter culture was diluted to A₆₀₀ ~0.05 in 50 ml of each medium listed in Table 1 and grown at 37°C to A₆₀₀ ~0.6. *p*AcF was then added to each culture to a final concentration of 2 mM. After 1 hour, the cultures were moved to 30°C and the pET and pEVOL plasmids were induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and 0.02% L-arabinose, respectively. After 12 hours, 1 ml of each culture was harvested by centrifugation, resuspended in 300 µL of B-PER (Pierce) and the soluble fraction was checked by SDS-page.

Thick overexpression bands are clearly observed for both the full length and truncated proteins in LB medium (Fig. 1A). In contrast, only a small amount of truncated EIC with no detectable band for the full-length protein was obtained from cultures in protonated and deuterated minimal media (Figs. 1B–D). This finding cannot be ascribed to an intrinsically low expression of EIC in minimal media since ~ 15 mg of U-[²H-¹⁵N] wild type EIC was obtained from 50 ml cultures in M9^{D2O} (table 1). An alternative explanation may be the reduced expression of the tRNA/aminoacyl-tRNA synthetase pair which is under control of the *ara* promoter and could be largely inhibited at the high glucose concentration used in the tested minimal media. To test this hypothesis the starter culture was diluted to A₆₀₀ ~ 0.05 in 50 ml of LB medium and grown at 37°C to an A₆₀₀ ~ 0.4. The pEVOL plasmid was then induced with 0.02% L-arabinose for 2 hours, while still in protonated, LB medium. Bacteria were collected by gentle centrifugation (170 g for 25 min) and resuspended in 50 ml of two different recipes of deuterated media, M9^{D2O} and M9^{ISO} (Table 1), to a final A₆₀₀ of ~0.8. Subsequently, 0.5 g/l *pAcF* (final concentration 2 mM) and 0.02% L-arabinose were added to each culture. After 1 hour of incubation at 30°C the expression of EIC was induced with 1 mM IPTG and the two cultures were shaken at 30°C for 12 hours. In stark contrast to the near zero yield afforded by the standard protocol, a large amount of full-length protein was obtained using the latter protocol, with yields comparable to the one obtained in LB medium (Figs. 1E,F and Table 2). These data clearly indicate that pre-inducing the pEVOL plasmid in LB medium is extremely beneficial for protein expression.

The degree of isotopic incorporation in the proposed protocol was tested by liquid chromatography/mass spectrometry on purified EIC M469*pAcF*/C575A. The results are summarized in Table 3, demonstrating that high levels of isotopic incorporation (~95%) are obtained in both the deuterated media tested.

As a final quality test on the produced protein, a comparison of the ¹H-¹⁵N TROSY spectra of the wild type and mutant proteins is reported in Fig. 2. The two spectra show excellent overlap (Fig. 2A), differing only in cross-peaks assigned to residues in the vicinity of the introduced mutation (Fig. 2B), indicating that the introduction of the unnatural amino acid *pAcF* at position 469 does not alter the overall folding of EIC.

In conclusion, we have presented a new protocol for incorporating the unnatural amino acid *pAcF* into deuterated EIC using the recently developed pEVOL vector (Young et al., 2010). We have shown that pre-inducing the expression of the tRNA/aminoacyl-tRNA synthetase pair in LB medium prior to switching the culture to a deuterated, glucose-based medium results in protein yields comparable to that obtained in rich medium, with no significant effects on the isotopic incorporation level. Similar results were also obtained for an EIC construct incorporating *pAcF* at position A494, as well as for the full-length EI (subunit molecular weight of 64 kDa) incorporating *pAcF* at position M469 (Fig. 3), suggesting that our procedure is of general applicability. We therefore fully expect the overall protocol presented here to be applicable to other unnatural amino acids and protein constructs. It is worth noting, however, that the reduced expression levels obtaining with standard protocols is probably related to the high glucose content of the tested minimal media. Thus, pre-induction of the pEVOL plasmid in LB media may not be needed when using other carbon sources.

Acknowledgments

We thank Drs. Mark Fleissner and Wayne Hubbell for generously providing *pAcF* and the Scripps Research Institute for the pEVOL plasmid. This work was supported by funds from the Intramural Program of the NIH, NIDDK, and the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.).

References

- Brustad E, Bushey ML, Brock A, Chittuluru J, Schultz PG. A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins. *Bioorg Med Chem Lett*. 2008; 18:6004–6006. [PubMed: 18845434]
- Brustad E, Bushey ML, Lee JW, Groff D, Liu W, et al. A genetically encoded boronate-containing amino acid. *Angew Chem Int Ed Engl*. 2008; 47:8220–8223. [PubMed: 18816552]
- Cellitti SE, Jones DH, Lagpacan L, Hao X, Zhang Q, et al. In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. *J Am Chem Soc*. 2008; 130:9268–9281. [PubMed: 18576636]
- Clore GM, Iwahara J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chemical reviews*. 2009; 109:4108–4139. [PubMed: 19522502]
- Fleissner MR, Brustad EM, Kalai T, Altenbach C, Cascio D, et al. Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proc Natl Acad Sci U S A*. 2009; 106:21637–21642. [PubMed: 19995976]
- Guo J, Wang J, Lee JS, Schultz PG. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angew Chem Int Ed Engl*. 2008; 47:6399–6401. [PubMed: 18624319]
- Jones DH, Cellitti SE, Hao X, Zhang Q, Jahnz M, et al. Site-specific labeling of proteins with NMR-active unnatural amino acids. *J Biomol NMR*. 2010; 46:89–100. [PubMed: 19669620]
- Lampe JN, Brandman R, Sivaramakrishnan S, de Montellano PR. Two-dimensional NMR and all-atom molecular dynamics of cytochrome P450 CYP119 reveal hidden conformational substates. *J Biol Chem*. 2010; 285:9594–9603. [PubMed: 20097757]
- Lee HS, Schultz PG. Biosynthesis of a site-specific DNA cleaving protein. *J Am Chem Soc*. 2008; 130:13194–13195. [PubMed: 18788806]
- Lee HS, Spraggon G, Schultz PG, Wang F. Genetic incorporation of a metal-ion chelating amino acid into proteins as a biophysical probe. *J Am Chem Soc*. 2009; 131:2481–2483. [PubMed: 19193005]
- Li C, Wang GF, Wang Y, Creager-Allen R, Lutz EA, et al. Protein (19)F NMR in *Escherichia coli*. *J Am Chem Soc*. 2010; 132:321–327. [PubMed: 20050707]
- Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nature Biotechnol*. 2006; 24:1436–1440. [PubMed: 17072302]
- Liu W, Brock A, Chen S, Chen S, Schultz PG. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nature methods*. 2007; 4:239–244. [PubMed: 17322890]
- Mittermaier A, Kay LE. New tools provide new insights in NMR studies of protein dynamics. *Science* (New York, NY). 2006; 312:224–228.
- Patel HV, Vyas KA, Mattoo RL, Southworth M, Perler FB, et al. Properties of the C-terminal domain of Enzyme I of the *Escherichia coli* Phosphotransferase system. *J Biol Chem*. 2006; 281:17579–17587. [PubMed: 16547354]
- Pintacuda G, Keniry MA, Huber T, Park AY, Dixon NE, et al. Fast structure-based assignment of 15N HSQC spectra of selectively 15N-labeled paramagnetic proteins. *J Am Chem Soc*. 2004; 126:2963–2970. [PubMed: 14995214]
- Teplyakov A, Lim K, Zhu PP, Kapadia G, Chen CC, et al. Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. *Proc Natl Acad Sci U S A*. 2006; 103:16218–16223. [PubMed: 17053069]
- Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Protoc*. 2006; 1:749–754. [PubMed: 17406304]
- Venditti V, Fawzi NL, Clore GM. Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear overhauser enhancement spectroscopy. *J Biomol NMR*. 2011; 51:319–328. [PubMed: 21935714]
- Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of *Escherichia coli*. *Science* (New York, NY). 2001; 292:498–500.
- Watt ED, Rivalta I, Whittier SK, Batista VS, Loria JP. Reengineering rate-limiting, millisecond enzyme motions by introduction of an unnatural amino acid. *Biophys J*. 2011; 101:411–420. [PubMed: 21767494]

- Xie J, Schultz PG. A chemical toolkit for proteins--an expanded genetic code. *Nature Rev Mol Cell Biol.* 2006; 7:775–782. [PubMed: 16926858]
- Xie J, Supekova L, Schultz PG. A genetically encoded metabolically stable analogue of phosphotyrosine in *Escherichia coli*. *ACS Chem Biol.* 2007; 2:474–478. [PubMed: 17622177]
- Young TS, Ahmad I, Yin JA, Schultz PG. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J Mol Biol.* 2010; 395:361–374. [PubMed: 19852970]

Fig. 1.

Expression of EIC M469pAcF/C575A in different media. Lanes A, B, C and D correspond to expression in LB, M9^{H2O}, M9^{D2O} and M9^{ISO} media (see Table 1), respectively. Lanes E and F correspond to expression performed by pre-inducing the pEVOL plasmid in LB medium prior to moving the culture to the deuterated M9^{D2O} and M9^{ISO} media, respectively. The left-hand most lane displays standard molecular weight markers.

Fig 2.
 (A) 600MHz ^1H - ^{15}N TROSY spectra of wild-type (blue) and M469pAcF/C575A (red) EIC. The NMR spectra were acquired at 37°C. Sample conditions were as follow: 400 μM EIC, 20 mM Tris, pH 7.4, 100 mM NaCl, 4 mM MgCl_2 , 1 mM EDTA. Under these conditions EIC is present as a 70 kDa dimer (Patel et al., 2006). Resonance assignment for cross-peaks showing large chemical shift differences in the two EIC constructs is reported on the wild-type spectrum. (B) X-ray structure of the wild-type EIC dimer (PDB code 2HWG) (Teplyakov et al., 2006). The monomeric subunits are colored in cyan and orange. Residues whose resonances show large $^1\text{H}_\text{N}/^{15}\text{N}$ chemical shift differences (>0.1 ppm) as a consequence of the M469pAcF and C575A mutations are indicated as blue spheres. M469 is shown as yellow sticks.

Fig 3.

Expression of full-length (residues 1-575) EI M469 *pAcF/C575A* (left panel) and EIC A494 *pAcF/C575A* (right panel) in different media. Lanes A, B and C correspond to expression in LB, M9^{H2O} and M9^{ISO} media (see Table 1), respectively. Lane D corresponds to expression performed by pre-inducing the pEVOL plasmid in LB medium prior to moving the culture to the deuterated M9^{ISO} medium. The left-hand most lane in each panel displays standard molecular weight markers.

Table 1Composition of the various culture media used in the present work.^a

Medium	Composition (per 1 L)
LB	10.0 g bacto-tryptone, 5.0 g yeast extract, 5.0 g NaCl, 200 μ l 5 M NaOH
M9 ^{H2O}	6.5 g Na ₂ HPO ₄ , 3.0 g KH ₂ PO ₄ , 0.5 g NaCl, 1.0 g ¹⁵ NH ₄ Cl, 2.0 g D-glucose, 120 mg MgSO ₄ , 11 mg CaCl ₂ , 10 ml 100X MEM vitamins mix (VWR)
M9 ^{D2O}	6.5 g Na ₂ HPO ₄ , 3.0 g KH ₂ PO ₄ , 0.5 g NaCl, 1.0 g ¹⁵ NH ₄ Cl, 2.0 g D-glucose-1,2,3,4,5,6-d ₇ , 120 mg MgSO ₄ , 11 mg CaCl ₂ , 10 ml 100X MEM vitamins mix (VWR)
M9 ^{ISO}	6.5 g Na ₂ HPO ₄ , 3.0 g KH ₂ PO ₄ , 0.5 g NaCl, 1.0 g ¹⁵ NH ₄ Cl, 2.0 g D-glucose-1,2,3,4,5,6-d ₇ , 120 mg MgSO ₄ , 11 mg CaCl ₂ , 10 ml 100X MEM vitamins mix (VWR), 1.0 g ¹⁵ N- ² H IsoGro (Isotec)

^aThe LB and M9^{H2O} media were prepared in water, while M9^{D2O} and M9^{ISO} media were prepared in 99.9% D₂O. The MEM vitamin mix used in the deuterated media was lyophilized and resuspended in 99.99% D₂O before usage.

Table 2

Final optical density (A_{600}) and yield of EIC M469pAcF/C575A measured for the 50 ml test cultures. LB-M9^{D2O} and LB-M9^{ISO} indicate cultures pre-induced in LB medium with L-arabinose and then moved to deuterated media (see text).

Medium	Final A_{600}	Yield (mg) ^a
LB	~2.4	~9.1
M9 ^{H2O}	~2.0	-
M9 ^{D2O}	~1.5	-
M9 ^{ISO}	~1.6	-
LB-M9 ^{D2O}	~2.0	~8.4
LB-M9 ^{ISO}	~2.2	~10.2

^a Measured for purified EIC M469pAcF/C575A (residues 261–575) using an extinction coefficient $\epsilon_{280} = 19.94 \text{ mM}^{-1} \text{ cm}^{-1}$. In all cases conversion to mg units was carried out by using the theoretical mass for the EIC domain at natural abundance (35,206 Da).

LC-MS analysis of the natural abundance and U-[²H, ¹⁵N]-labeled EIC M469pAcF/C575A samples. Data were acquired under denaturing conditions with a liquid chromatography (Zorbax 300SB-C3 column) mass spectrometry electrospray ionization time-of-flight (LC-MS ESI-TOF) instrument (HP 1100 HPLC/MSD). Protein samples (6 μM final concentration) were diluted in water and acidified with 10% acetic acid. Mass spectra were recorded in the positive ion mode (mass range 500–2000) and results were analyzed using the HP ChemStation software. In the last row (WT-M9D2O) LC-MS data for U-[²H, ¹⁵N]-labeled wild-type (WT) EIC expressed in the M9D2O medium are reported.

Table 3

Medium	Theoretical mass (Da) ^a	Observed mass (Da)	Theoretical number of isotopic substitutions ^b	Observed number of isotopic substitutions ^c	Isotopic incorporation (%)
LB	35,206	35,205	-	-	-
LB-M9D2O	37,551	36,428	2,345	2,223	94.8
LB-M9ISO	37,551	37,441	2,345	2,236	95.4
WT-M9D2O	37,547	37,530	2,368 ^d	2,351 ^d	99.3

^aTheoretical masses were calculated assuming 100% isotopic incorporation and complete back exchange to ¹H at labile hydrogen positions.

^bThis is the expected number of ¹⁵N and ²H atoms incorporated into EIC and is calculated by taking the difference between the theoretical mass in LB and deuterated media.

^cThis is the observed number of ¹⁵N and ²H atoms incorporated into EIC and is calculated by taking the difference between the observed mass in LB and deuterated media.

^dCalculated using the theoretical mass in LB for the natural abundance wild-type EIC (35,179 Da).