Preparation of site-specifically labeled fluorinated proteins for ^{19}F-NMR structural characterization

Jared T Hammill, Shigeki Miyake-Stoner, Jennifer L Hazen, Jennifer C Jackson & Ryan A Mehl

Chemistry Department, Franklin and Marshall College, PO Box 3003, Lancaster, Pennsylvania 17604, USA. Correspondence should be addressed to R.A.M. (rmehl@fandm.edu).

Published online 18 October 2007; doi:10.1038/nprot.2007.379

A straightforward protocol for the site-specific incorporation of a ^{19}F label into any protein *in vivo* is described. This is done using a plasmid containing an orthogonal aminoacyl-tRNA synthetase/tRNA$_{CUA}$ that incorporates L-4-trifluoromethylphenylalanine in response to the amber codon UAG. This method improves on other *in vivo* methods because the ^{19}F label is incorporated into only one location on the protein of interest and that protein can easily be produced in large quantities at low cost. The protocol for producing ^{19}F-labeled protein is similar to expressing protein in *Escherichia coli* and takes 4 d to obtain pure protein starting from the appropriate vectors.

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

INTRODUCTION

Site-specific introduction of fluorine labels into proteins is a powerful tool for monitoring protein conformational changes and interactions[1–7]. The most common methods for incorporation of fluorinated amino acids involve semisynthetic incorporation or use of natural translational machinery[1,8–10]. Semisynthetic incorporation enables high fidelity at specific sites but becomes impractical when medium to large proteins are needed for study. Use of natural translational machinery to force fluorinated mimics into their natural codons can produce large ^{19}F-labeled proteins *in vivo*. Unfortunately, incorporation with natural machinery or slightly altered natural machinery rarely approaches 95% efficiency at a single location and results in background incorporation of ^{19}F-labeled amino acid in place of other natural amino acids[7,11]. Varied incorporation level at more than one site also results in multiple different protein molecules all with potentially different structural perturbations[12]. The clean, site-specific labeling of large quantities of protein *in vivo* with fluorine nuclei would allow the detailed dynamic characterization of protein structure *in vivo* and *in vitro*. It would also allow the use of fluorinated proteins, of any size, in areas of protein structural characterization, protein processing and protein stability studies[13–15].

We have developed an alternative method for the incorporation of a fluorine label into proteins that allows uniform labeling of a single site in a protein *in vivo* with a ^{19}F-labeled amino acid[16]. The method relies on an orthogonal aminoacyl-tRNA synthetase/tRNA$_{CUA}$ that

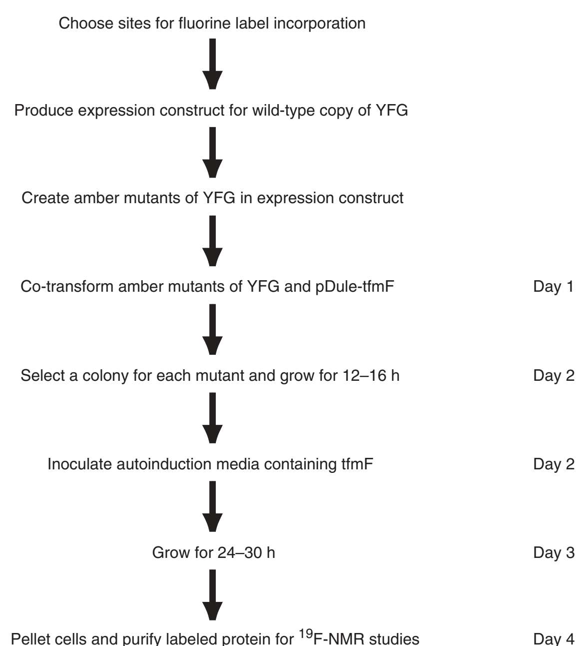

Figure 2 | Flowchart of the steps to express a protein site specifically incorporating tfmF in *E. coli*. A timeline is included to the right of the flowchart.

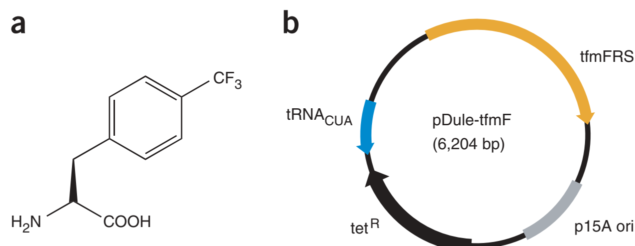

Figure 1 | tfmF and pDule plasmid. (**a**) The chemical structure of tfmF. (**b**) pDule-tfmF plasmid. The tfmF-specific aminoacyl-tRNA synthetase is expressed between an lpp promoter and an rrnB terminator. The tRNA$_{CUA}$ is expressed between an lpp promoter and an rrnC terminator. The plasmid contains a p15A origin and a tetracycline resistance marker.

incorporates L-4-trifluoromethylphenylalanine (tfmF; see **Fig. 1a**) in response to the amber (stop) codon UAG in *E. coli*[15,17]. The system in this protocol uses two plasmids: pDule-tfmF (**Fig. 1b**), a p15A tetracycline-resistant plasmid to express the orthogonal tRNA and synthetase, and a second plasmid containing an amber mutant of the gene of interest. To produce the ^{19}F-labeled protein, cultures of *E. coli* carrying both plasmids are grown in the presence of the unnatural amino acid (**Fig. 2**). As this method incorporates tfmF with high fidelity in response to a UAG codon *in vivo*, complementary studies with labeled protein can easily be performed *in vitro* (with crude cell extract or purified protein components). The protocol outlines the use of the pDule-tfmF plasmid in conjunction with the pBAD plasmid carrying the amber mutant gene of interest.

This protocol shows tfmF incorporation into three different proteins of different sizes (**Fig. 3**). These ^{19}F-labeled proteins enable

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

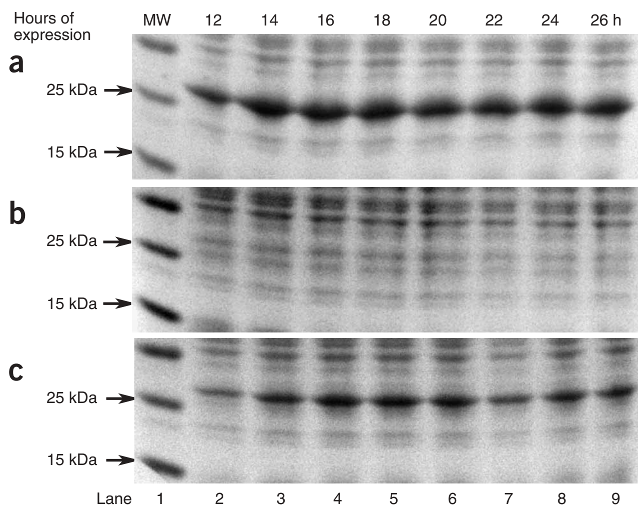

Figure 3 | Analysis of crude cell pellet expressions for NTR and NTR-tfmF from **Table 2**. The molecular weight of NTR is 27 kDa. (**a**) Positive control for protein production (native protein) separated by SDS-PAGE and Coomassie-stained. (**b**) Negative control for protein production (tfmF withheld) separated by SDS-PAGE and Coomassie-stained. (**c**) ^{19}F-labeled protein production (tfmF added) separated by SDS-PAGE and Coomassie-stained. Lane 1 contains broad-range prestained markers (Bio-Rad). Lanes 2–9 show protein expression results from the time points collected during protein expression. The number of hours since inoculation is shown at the top of each lane.

the monitoring of conformational changes and dynamics: in the hydrophobic interface of *E. coli* histidinol dehydrogenase (HDH), a 50 kDa homodimer; in the active site of *E. coli* nitroreductase (NTR), a 27 kDa homodimer; and in the mobile loop of *Salmonella typhimurium* nitrogen-regulating protein N-terminal receiver domain, a 14 kDa monomer (NtrC^{r}).

When this pair of plasmids (pBAD and pDule-tfmF) is used to produce protein in autoinduction media[18], exceptional yields of pure ^{19}F-labeled protein result (**Fig. 4**). This pair of plasmids can be used to produce labeled protein also in standard rich media formulations, but the yield is often reduced at least fivefold compared to what is observed in the autoinduction media described. The arabinose promoter system on the pBAD plasmid also allows control over levels of protein production if more native-like levels of protein are needed for *in vivo* ^{19}F-NMR studies.

MATERIALS

REAGENTS

- Arabinose, natural amino acids, antibiotics, salts for media and buffers (see below and REAGENT SETUP; Sigma-Aldrich)
- Arabinose autoinduction media[18] prepared as described in text (Step 5) with 100 μg ml^{-1} ampicillin and 25 μg ml^{-1} tetracycline
- BD Talon Metal Affinity Resin (BD Biosciences)
- Coomassie protein assay reagent (protein concentration kit, Pierce)
- *E. coli* strain DH10B prepared as electrocompetent cells (Invitrogen)
- Elution buffer: 50 mM sodium phosphate, 300 mM sodium chloride and 150 mM imidazole (pH 8)
- Lysis buffer: 50 mM sodium phosphate, 300 mM sodium chloride (pH 8), 0.25 mg/ml lysozyme (pH 8)
- tfmF (Peptech) solubilized in sterile water with 8 M NaOH solution
- PAGEr Gold Precast Polyacrylamide Gels, 4–12% Tris-glycine, 9 cm × 10 cm, 1 mm thick
- PBS: 10 mM sodium phosphate and 140 mM sodium chloride (pH 7.2)
- Plasmid vectors: pBAD/His is available from Invitrogen. Plasmids pDule-tfmF, pBAD-HDH, pBAD-HDH-225TAG, pBAD-NTR and pBAD-NTR-124TAG are available from Ryan A. Mehl (ryan.mehl@fandm.edu)
- Poly-Prep Chromatography Columns (0.8 cm × 4 cm; Bio-Rad)
- Selective media: LB agar and LB media containing 100 μg ml^{-1} of ampicillin and 25 μg ml^{-1} of tetracycline
- Sodium hydroxide (Sigma-Aldrich)
- SOC rescue media
- QuikChange Site-Directed Mutagenesis Kit (Stratagene)
- Wash buffer: 50 mM sodium phosphate and 300 mM sodium chloride (pH 8)

EQUIPMENT

- Bruker stem coaxial insert tube (2 mm outer diameter)
- Sonic dismembrator (Fisher: model 100)
- Incubators and shakers for cell growth
- Electroporator
- Centrifuge
- Rocker
- PAGE-gel rig apparatus and power supply

REAGENT SETUP

50× M salt solution Combine the following salts in a total volume of 500 ml water with gentle warming. Once dissolved, autoclave the solution to sterilize.

50x M content	Per 500 ml
Na_2HPO_4 (1.25 M)	88.73 g
KH_2PO_4 (1.25 M)	85.1 g
NH_4Cl (2.5 M)	66.9 g
Na_2SO_4 (0.25 M)	17.8 g

Note: This solution can recrystallize at room temperature (21 °C). Microwave solution to redissolve before use. This 50× solution can be stored at room temperature for years and is stable even after repeated microwaving.

25× 18-amino-acid solution Combine 5 g of the following 18 amino acids in a total volume of 1 liter Millipore water. Sterile-filter this stock solution after the amino acids have dissolved. The 18-amino-acid solution should be stored at 4 °C. This solution can darken in color but no precipitate should form, and, provided sterility is maintained, it can be stored for years at 4 °C. This 25× 18-amino-acid solution will provide each amino acid at a final concentration of 400 μg ml^{-1} in the growth media (leave out tyrosine and cysteine).

Amino acid	MW g/mole
Glutamic acid sodium salt	169.1
Aspartic acid	133.1
Lysine-HCl	182.6
Arginine-HCl	210.7
Histidine-HCl-H_2O	209.6
Alanine	89.1
Proline	115.1
Glycine	75.1
Threonine	119.1
Serine	105.1
Glutamine	146.1
Asparagine-H_2O	150.1
Valine	117.1
Leucine	131.2
Isoleucine	131.2
Phenylalanine	165.2
Tryptophan	204.2
Methionine	149.2

5,000× trace metal stock solution To prepare this solution, it is convenient to make individual sterile metal stock solutions and then combine and dilute to the correct 5,000× concentration. Add each of the metal salts (**Table 1**) to 30 ml of Millipore water and autoclave, with the exception of $FeCl_3$, which must be dissolved in 0.1 M HCl and then filtered (through a 0.2 μm filter) to remove insoluble residues. Combine the 30 ml stock salt solutions together according to the volumes reported in **Table 1** and dilute to 50 ml with sterile water. This stock solution might show minor precipitation over time but is stable at 15–25 °C for years.

PROCEDURE

Preparation of mutant construct and transformation

● TIMING 1–2 weeks

1| Choose a protein of interest (your favorite protein (YFP)) and choose amino acids for replacement with the fluorine probe tfmF. Clone the gene (your favorite gene (YFG)) into pBAD, a standard arabinose-inducible overexpression vector. Note that the control plasmids pBAD-HDH, pBAD-HDH-225TAG, pBAD-NTR, pBAD-NTR-124TAG, pBAD-NtrC^r and pBAD-NtrC^r-99 used to obtain the data discussed in ANTICIPATED RESULTS are derived from pBAD/Myc-His A (Invitrogen).

▲ CRITICAL STEP The overexpression vector for your protein should be a pBAD vector. We have tested pTrc and pET vectors with lactose induction[18] as well as IPTG induction, and their performance is considerably worse in autoinduction and/or rich media. If you choose a vector other than pBAD, a tetracycline resistance marker should not be used, as its presence will not allow the selection of this vector to be distinguished from that of pDule-tfmF (see INTRODUCTION). The alternate vector to pBAD should also contain an origin of replication compatible with the pDule's p15A origin of replication so that both plasmids can be maintained in the cell.

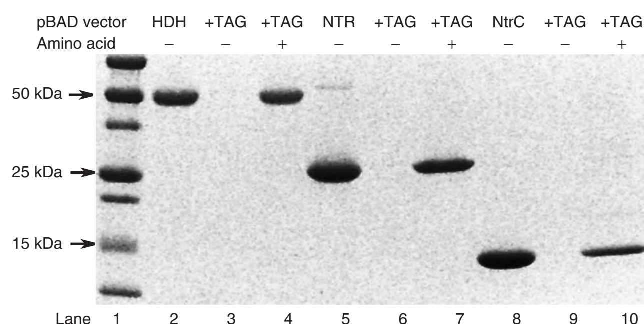

Figure 4 | Analysis of native and tfmF-protein produced for HDH, NTR and NtrC^r. The molecular weight of proteins HDH, NTR and NtrC^r is 50, 27 and 14 kDa, respectively. Positive and negative controls for protein production separated by SDS-PAGE and Coomassie stained are shown. Lane 1 contains broad-range prestained markers (Bio-Rad). Lanes 2–10 show protein expression results from the plasmids indicated and match the controls in **Table 2**.

2| Mutate the codon corresponding to the selected amino acid in YFG to the amber codon (UAG) using a standard oligonucleotide-directed mutagenesis method to create the plasmid pBAD-YFG-TAG. We use the QuikChange method from Stratagene, but other methods are acceptable.

▲ CRITICAL STEP Ideal sites for mutagenesis are 'permissive', that is, mutagenesis to tfmF will not alter protein function or expression. tfmF is an aromatic hydrophobic amino acid. It is likely that tyrosine, phenylalanine or tryptophan residues will be the most permissive to replacement with tfmF. Because it is often not clear a priori whether an amino acid will be permissive to substitution, it is worth considering constructing several mutants in parallel. Before proceeding with the incorporation of the ^{19}F label, if the site chosen for tfmF insertion does not contain a tyrosine, we recommend using site-directed mutagenesis to incorporate a tyrosine at that location. We have experienced that tyrosine best mimics the structural aspects of tfmF and most problems with protein stability and structural integrity will be observed following tyrosine incorporation.

3| Prepare two different protein production strains of *E. coli* (DH10B). One strain should contain YFG in the pBAD-His plasmid (pBAD-YFG, ampicillin resistant) and the pDule-tfmF plasmid (tetracycline resistant) and the second should contain the amber mutant pBAD-YFG-TAG plasmid (ampicillin resistant) and the pDule-tfmF plasmid (tetracycline resistant). For delivering two plasmids, we find it most convenient to perform a co-transformation into DH10B electrocompetent cells followed by selection with both antibiotics on rich medium agar plates. To prepare each expression strain, mix 0.2 µg of plasmid DNA with 50 µl of electrocompetent *E. coli* DH10B cells and transform via electroporation (2 mm cuvette at 2.5 kV). Rescue the cells in 1 ml of SOC media at 37 °C for 1 h with shaking at 250 r.p.m.

▲ CRITICAL STEP *E. coli* cell types DH10B or Top10 should be used. These strains are capable of transporting L-arabinose but not metabolizing it. The pBAD vector uses L-arabinose as an inducer of protein expression. This is important for autoinduced expression of protein, as the level of L-arabinose will increase inside the cell and not decrease over time.

TABLE 1 | Preparation of individual trace metal stock solutions and 5,000× stock solution.

Amount for 30 ml stock solution	Amount of 30 ml stock solution for 50 ml (5,000x) solution	1x media concentration
$CaCl_2 \cdot 2H_2O$ (8.82 g)	500 µl	4 µM
$MnCl_2 \cdot 4H_2O$ (5.93 g)	500 µl	2 µM
$ZnSO_4 \cdot 7H_2O$ (8.62 g)	500 µl	2 µM
$CoCl_2 \cdot 6H_2O$ (1.32 g)	500 µl	0.4 µM
$CuCl_2$ (807 mg)	500 µl	0.4 µM
$NiCl_2$ (777 mg)	500 µl	0.4 µM
$Na_2MoO_4 \cdot 2H_2O$ (1.45 g)	500 µl	0.4 µM
Na_2SeO_3 (1.03 g)	500 µl	0.4 µM
H_3BO_3 (371 mg)	500 µl	0.4 µM
$FeCl_3$ (486 mg)	25 ml	10 µM
	Dilute to 50 ml with sterile water	

4| Select cells carrying both the pDule plasmid and the pBAD plasmid by plating the transformed cells from Step 3 onto selective LB agar medium containing 25 mg liter^{-1} tetracycline and 100 mg liter^{-1} ampicillin. If the expression vector has a marker other than ampicillin, substitute ampicillin with the appropriate antibiotic at the relevant concentration.
? TROUBLESHOOTING

5| For each transformed strain, select a fresh colony grown on a rich agar plate containing the appropriate antibiotics and grow at 37 °C to saturation in 6 ml of LB media containing 25 mg liter^{-1} of tetracycline and 100 mg liter^{-1} of ampicillin. It should take the cell culture about 12 h to reach saturation (optical density (OD) (at 600 nm) greater than 1).

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

Expression of mutant protein containing tfmF ● TIMING 24–30 h

6| For each expression trial, prepare 500 ml of arabinose autoinduction media in a 2-liter baffled flask (one flask with 500 ml for each of the three trials described in **Table 1**) as follows:

Glycerol (10% (wt/vol))	25 ml	(autoclave)
Glucose (40% (wt/vol))	0.6 ml	(autoclave)
Arabinose (20% (wt/vol))	1.25 ml	(sterile-filter)
50× M salts	10 ml	
$MgSO_4$ (1 M)	1 ml	(autoclave)
Aspartate (5%, pH 7.5)	25 ml	(autoclave)
Leucine (4 mg ml^{-1}, pH 7.5)	10 ml	(autoclave)
18-amino-acid mix (25×)	20 ml	
Trace metal stock solution	100 µl	
Water	To 500 ml	(autoclave)

7| Prewarm each flask (500 ml of media in a 2-liter flask) at 37 °C for 1 h with shaking at 250 r.p.m. Add the antibiotics ampicillin (100 µg ml^{-1}) and tetracycline (25 µg ml^{-1}). Add 2.5 ml of each of the saturated cell cultures (from Step 5) to the appropriate flask.

8| After cell cultures have grown for 30 min, add tfmF to flask 3 (see **Table 2**). For 500 ml of media, dissolve 117 mg of tfmF in 4 ml of sterile water by adding 80 µl of 8 M NaOH. Add the amino-acid solution to the growing culture in flask 3 to give a final concentration of 1 mM.
▲ CRITICAL STEP Prepare the amino-acid stock solution fresh. Basic conditions can epimerize the unnatural amino acid. The amount of time after dissolving the amino acid in sterile water with base should be kept to a minimum. Excessive use of base can damage the amino acid and can retard cell growth.

9| Incubate the cultures at 37 °C with 250 r.p.m. shaking, and protein production will commence without the need for external inducers. Continue incubation at 37 °C for an additional 25–30 h with shaking at 250 r.p.m. Cell culture density reaches an OD of 4–6 within 12–15 h. Generally, a constant cell density for 6–10 h indicates that protein expression is finished. When growing multiple cultures in parallel, it is common for them to double at slightly different rates. Monitor cell density by diluting 0.2 ml of media with 0.8 ml of water and recording OD reading at 600 nm every 2 h after 12 h of expression. To estimate the optimal time length of protein expression, retain 0.2 ml of cells when taking OD readings. Centrifuge the 0.2 ml culture time points at 8,000*g* for 5 min at room temperature and remove the supernatant. These cell pellets should be stored at −20 °C until their use in Step 11. Note that the OD may not be constant but decreases slightly after 12–15 h of growth.

10| Harvest each culture by centrifugation at 4 °C for 20–30 min at 5,000*g*. Because of the high yield of cells per volume of media, it may be convenient at this point to centrifuge fractions of the total volume. To run ^{19}F-NMR spectra of labeled protein, 25 ml of media is a convenient volume for future purification.
■ PAUSE POINT Cell pellets may be stored for months at −80 °C for use in Step 12.

Crude gel of protein expressions ● TIMING 4 h

11| To assess the level of expression of the protein before embarking on the purification step, analyze the crude cells via SDS-polyacrylamide gel electrophoresis (SDS-PAGE). For this purpose, resuspend the 0.2 ml cell pellets of expression time points in 100 µl of water and mix with 100 µl of 2× SDS loading buffer. Heat the samples at 100 °C for 10 min (mixing intermittently) and then remove cellular debris by centrifugation at room temperature for 5 min at 16,000*g*. Load 20 µl samples onto a 4–20% Tris-glycine gel and carry out electrophoresis at 120 V. Stain the gel with the Coomassie stain. Note that example results using the control constructs are shown in **Figure 3**.

TABLE 2 | Parallel overexpression series for each protein site studied.

Expression trial	Flask 1 Native protein (positive control)	Flask 2 No tfmF control (negative control)	Flask 3 tfmF-containing protein
tfmF amino acid	None	None	1 mM
YFP expression vector	pBAD-YFG	pBAD-YFG-TAG	pBAD-YFG-TAG
pDule vector	pDule-tfmF	pDule-tfmF	pDule-tfmF
Antibiotics	Ampicillin (100 µg ml^{-1}) Tetracycline (25 µg ml^{-1})	Ampicillin (100 µg ml^{-1}) Tetracycline (25 µg ml^{-1})	Ampicillin (100 µg ml^{-1}) Tetracycline (25 µg ml^{-1})

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

▲ CRITICAL STEP The yield of proteins containing tfmF produced with the pDule plasmid may vary. Expression of protein containing tfmF in NTR, HDH and NtrC^r as described gave yields of 20–78 mg liter^{-1} pure tfmF-protein from the media described. The protein yield normally reflects the production levels of native protein in this medium.
? TROUBLESHOOTING

Purification of protein expressions ● TIMING 4 h
12| Purify protein from each of the three cultures, using the BD Talon Affinity Resin Native Purification protocol VI B. Detailed protocols for affinity purification of 6-His-tagged protein are provided with the BD Talon Affinity resin. Resuspend a 25 ml pellet from each culture (Step 10) in 15 ml of lysis buffer and allow it to stand on ice for 20 min. Sonicate the samples at 4 °C until the viscosity is visibly reduced, and then centrifuge samples at 4 °C for 30 min at 20,000*g*.
▲ CRITICAL STEP Protein expression can be monitored by many purification/detection methods. The methods using Coomassie staining or silver staining function with UAA-protein as with any natural protein. The incorporation of tfmF into proteins, as in any substitution with natural amino acids, may alter the stability of the protein, thereby affecting the ease of protein purification and analysis.

13| Prepare the Talon resin by suspending 750 µl of resin in 7.5 ml of wash buffer (ten times the bed volume). Centrifuge the resin at 700*g* for 5 min at room temperature, remove and discard the supernatant and then repeat the wash process twice. Divide this washed resin into three equal samples of resin, one for each lysate. Note that the following steps deal with the purification of one such sample.

14| Add the cleared lysate (Step 12) to 250 µl of washed Talon resin and gently mix at 15–25 °C for 20 min. Centrifuge the resin at 700*g* for 5 min at room temperature to pellet the resin, and then carefully remove and discard the supernatant.

15| Resuspend the resin in 15 ml of wash buffer and transfer the sample to a Bio-Rad Poly-Prep Chromatography Column. Allow the column to drain, and wash by gravity flow with 15 ml of wash buffer (repeat the wash process twice). Elute the protein in 2.5 ml of elution buffer.
? TROUBLESHOOTING

Pure protein SDS-PAGE gel ● TIMING 4 h
16| Analyze an aliquot of each sample by SDS-PAGE. The results from cell lysates of strains carrying our control plasmids are shown in **Figure 4**. The purified lysate carrying native HDH yields 51 mg liter^{-1} of pure protein, HDH225-tfmF yields 44 mg liter^{-1} of pure protein, native NTR yields 134 mg liter^{-1} of pure protein, NTR124-tfmF yields 78 mg liter^{-1} of pure protein, whereas native NtrC^r yields 53 mg liter^{-1} of pure protein, NtrC^{r}99-tfmF yields 25 mg liter^{-1} of pure protein (all protein samples are $>$95% pure).
? TROUBLESHOOTING

17| Using an Amersham PD10 column, exchange the buffer of the eluted purified protein (2.5 ml) with a buffer appropriate for the NMR study of your protein of interest. Purified proteins have a wide range of optimal buffer conditions and those optimal conditions will vary with changes in protein concentration and your choice of NMR experiment. Measure the protein concentration using the Coomassie protein assay kit and dilute or concentrate the protein samples for NMR (in the case of HDH225-tfmF and NTR124-tfmF, sample concentrations were 1–3 mg ml^{-1} in a 20 mM phosphate buffer (pH 7.5)).

18| To prevent the contamination of ^{19}F-labeled protein with deuterated solvents, use a stem coaxial insert tube that contains lock and reference solvents. The use of a stem coaxial insert tube prevents the possibility of a deuterium isotope effect on the ^{19}F chemical shift. It also allows the protein sample to be used in future experiments where deuterated solvents are not desired. In a glass vial, add 2 µl of 4-fluorotoluene to 1 ml of toluene-d_8. Add 100 µl of this reference and lock solution to a Bruker stem coaxial insert tube (2 mm outer diameter). This insert can be placed into a standard 5 mm NMR tube that contains 500 µl of protein solution. The NMR can be locked on toluene-d_8, and a reference peak will appear at −120.771 ppm (as referenced from trifluorotoluene at −65.000 ppm).

● TIMING
Steps 1–5, preparation of mutant construct and media: 1–2 weeks
Transformation of constructs and preparation of cultures for protein expression: 2 d
Steps 6–10, expression of mutant protein containing tfmF: 24–30 h
Step 11, crude gel of protein expressions: 4 h
Steps 12–15, purification of protein expressions: 4 h
Steps 16–18, pure protein SDS-PAGE gel: 4 h

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 3**.

TABLE 3 | Troubleshooting table.

Step	Problem	Possible reason	Solution
4	Colonies not obtained on co-transformation	Co-transformation with both plasmids and direct selection of colonies that are resistant to both ampicillin and tetracycline will have a lower transformation efficiency than when transformed with single plasmids	Perform sequential transformation into the cells. First make competent cells containing pDule plasmids and then transform the pBAD-YFG-TAG plasmid into the cells containing the pDule plasmid
11	No expressed protein observed by SDS-PAGE crude cell analysis	Protein may be at too low a concentration to be easily visualized by Coomassie stain on cell lysate	Purify the protein to concentrate it and allow the protein to be visualized by silver staining of SDS-PAGE gel
			Denatured protein purification and gel analysis can be used if protein insolubility is suspected
15	High-mobility (low molecular weight) bands are seen on SDS-PAGE	Some protein will be made that terminates at the incorporated UAG stop codon, resulting in truncated protein. Truncated protein can be purified if the location of the amber codon allows production of a soluble protein fragment and the affinity tag is on the N terminus of the protein	Use of a C-terminal affinity tag will only allow purification of full-length protein. As a C-terminal affinity tag is at the end of the protein being expressed, it can only be added if the unnatural amino acid was incorporated and full-length protein is formed
16	Little or no protein is produced	The native or mutant protein is toxic or poorly expressed	Use rich media expression conditions (LB media) for the parallel overexpressions in **Table 2** and induce protein production at an OD of 0.6–1 by the addition of arabinose (final concentration 0.2%). Express for the length of time needed to produce the mutant protein
	The mutant tfmF-containing proteins are poorly expressed	There is a structural problem with incorporation of tfmF	Change the desired site to tyrosine using site-directed mutagenesis. Express this protein as described herein to check for structural problems that adversely affect expression or purification
		Incorporation of an amino acid into a stop codon may decrease protein yield relative to wild-type protein. The nucleotides proximal to the amber codon can have an effect on unnatural amino-acid incorporation efficiency[11]	If a C or T directly follows the amber codon in your gene, then change this nucleotide to A or G and repeat the expression experiment with the new mutant
			Use another more permissive site in the protein for mutagenesis

ANTICIPATED RESULTS

The pDule-tfmF plasmid allows the *in vivo*, site-specific incorporation of the ^{19}F-labeled amino acid tfmF. The amino acid is incorporated with high efficiency, fidelity and specificity into proteins of almost any length, resulting in a single protein product with a ^{19}F label in only one position. We have successfully produced proteins between 14 and 93 kDa, but there is no indication that proteins outside this range would not be feasible. When the pDule-tfmF plasmid is used in conjunction with DH10B cells and the common high-copy protein production plasmid pBAD vector, purified protein yields of 10–80 mg of protein per liter of media have been obtained. Whereas standard LB media work for preparing protein with the pBAD/pDule-tfmF plasmid pair, the high protein yields described here are due largely to the high cell densities provided by the autoinduction media. This medium is also beneficial in that it removes the experimental burden of being present for adding the inducing agent. As the autoinduction medium is a defined rich medium, individual components can easily be altered to allow isotopic labeling of protein[18]. This medium and general protocol are also compatible with incorporation of other unnatural amino acids into proteins provided the pDule plasmid for that particular amino acid has been developed[17,19]. tfmF incorporation results in a single, sharp ^{19}F-NMR signal around −65 ppm for the denatured protein or for a protein whose label is exposed to solvent[16]. Folded proteins with a tfmF amino acid may display a variety of ^{19}F-NMR signal types depending on the number of different electrostatic environments to which the trifluoromethyl group is exposed and the dynamics of the labeled protein.

Source

This protocol was provided directly by the authors listed on the title page. For further details on the composition of media, standard buffers and standard procedures, see Sambrook, J. & Russell, D.W. (eds.) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2001).

ACKNOWLEDGMENTS We thank Cinthia Kinsland for her helpful discussions on molecular biology and media. This work was supported by F&M Hackman and Eyler funds and NSF-MCB-0448297, Research Corporation (CC6364) and ACS-PRF (42214-GB4).

Published online at http://www.natureprotocols.com
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Danielson, M.A. & Falke, J.J. Use of ^{19}F NMR to probe protein structure and conformational changes. *Annu. Rev. Biophys. Biomol. Struct.* **25**, 163–195 (1996).
2. Ulrich, A.S. Solid state ^{19}F NMR methods for studying biomembranes. *Prog. Nucl. Magn. Reson. Spectrosc.* **46**, 1–21 (2003).
3. Gerig, J.T. Fluorine NMR of proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **26**, 293–370 (1994).
4. Bai, P., Luo, L. & Peng, Z. Side chain accessibility and dynamics in the molten globule state of alpha-lactalbumin: a (19)F-NMR study. *Biochemistry* **39**, 372–380 (2000).
5. Bann, J.G. & Frieden, C. Folding and domain–domain interactions of the chaperone PapD measured by ^{19}F NMR. *Biochemistry* **43**, 13775–13786 (2004).
6. Hoeltzli, S.D. & Frieden, C. Refolding of [6-^{19}F]tryptophan-labeled *Escherichia coli* dihydrofolate reductase in the presence of ligand: a stopped-flow NMR spectroscopy study. *Biochemistry* **37**, 387–398 (1998).
7. Furter, R. Expansion of the genetic code: site-directed *p*-fluoro-phenylalanine incorporation in *Escherichia coli*. *Protein Sci.* **7**, 419–426 (1998).
8. Vaughan, M.D., Cleve, P., Robinson, V., Duewel, H.S. & Honek, J.F. Difluoromethionine as a novel ^{19}F NMR structural probe for internal amino acid packing in proteins. *J. Am. Chem. Soc.* **121**, 8475–8478 (1999).
9. Duewel, H., Daub, E., Robinson, V. & Honek, J.F. Incorporation of trifluoromethionine into a phage lysozyme: implications and a new marker for use in protein ^{19}F NMR. *Biochemistry* **36**, 3404–3416 (1997).
10. Feeney, J.M. *et al.* ^{19}F nuclear magnetic resonance chemical shifts of fluorine containing aliphatic amino acids in proteins: studies on *Lactobacillus casei* dihydrofolate reductase containing (2*S*,4*S*)-5-fluoroleucine. *J. Am. Chem. Soc.* **118**, 8700–8706 (1996).
11. Frieden, C., Hoeltzli, S.D. & Bann, J.G. The preparation of ^{19}F-labeled proteins for NMR studies. *Methods Enzymol.* **380**, 400–415 (2004).
12. Luck, L.A. & Falke, J.J. ^{19}F NMR studies of the D-galactose chemosensory receptor. 2. Ca(II) binding yields a local structural change. *Biochemistry* **30**, 4257–4261 (1991).
13. Lee, K.H., Lee, H.Y., Slutsky, M.M., Anderson, J.T. & Marsh, E.N. Fluorous effect in proteins: *de novo* design and characterization of a four-alpha-helix bundle protein containing hexafluoroleucine. *Biochemistry* **43**, 16277–16284 (2004).
14. Higuchi, M. *et al.* ^{19}F and ^{1}H MRI detection of amyloid beta plaques *in vivo*. *Nat. Neurosci.* **8**, 527–533 (2005).
15. Wang, L., Brock, A., Herberich, B. & Schultz, P.G. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498–500 (2001).
16. Jackson, J.C., Hammill, J.T. & Mehl, R.A. Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **129**, 1160–1166 (2007).
17. Xie, J. & Schultz, P.G. An expanding genetic code. *Methods* **36**, 227–238 (2005).
18. Studier, F.W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
19. Wang, L. & Schultz, P.G. Expanding the genetic code. *Angew. Chem. Int. Ed. Engl.* **44**, 34–66 (2004).

© 2007 Nature Publishing Group http://www.nature.com/natureprotocols
npg