J|A|C|S
ARTICLES
Published on Web 01/17/2007

Site-Specific Incorporation of a ^{19}F-Amino Acid into Proteins as an NMR Probe for Characterizing Protein Structure and Reactivity

Jennifer C. Jackson, Jared T. Hammill, and Ryan A. Mehl*

Contribution from the Department of Chemistry, Franklin and Marshall College, P.O. Box 3003, Lancaster, Pennsylvania 17604

Received June 30, 2006; E-mail: rmehl@fandm.edu

Abstract: ^{19}F NMR is a powerful tool for monitoring protein conformational changes and interactions; however, the inability to site-specifically introduce fluorine labels into proteins of biological interest severely limits its applicability. Using methods for genetically directing incorporation of unnatural amino acids, we have inserted trifluoromethyl-L-phenylalanine (tfm-Phe) into proteins in vivo at TAG nonsense codons with high translational efficiency and fidelity. The binding of substrates, inhibitors, and cofactors, as well as reactions in enzymes, were studied by selective introduction of tfm-Phe and subsequent monitoring of the ^{19}F NMR chemical shifts. Subtle protein conformational changes were detected near the active site and at long distances (25 Å). ^{19}F signal sensitivity and resolution was also sufficient to differentiate protein environments in vivo. Since there has been interest in using ^{19}F-labeled proteins in solid-state membrane protein studies, folding studies, and in vivo studies, this general method for genetically incorporating a ^{19}F-label into proteins of any size in *Escherichia coli* should have broad application beyond that of monitoring protein conformational changes.

Introduction

The ability to monitor inhibitor binding events in protein structure/activity relationship studies as well as conformational changes and dynamics of proteins that are too large for conventional protein NMR would advance biochemical research.[1−3] The high sensitivity of ^{19}F to surrounding environments, 100% natural abundance, and high sensitivity to NMR detection (83% that of ^{1}H) has made ^{19}F NMR spectroscopy useful for investigating protein structure and dynamics.[1,4−7] The simplicity of observing hypersensitive ^{19}F chemical shifts by NMR makes it an exquisite tool for monitoring protein movements resulting from small-molecule binding, covalent modification, or protein interactions.[8−11] The ability to uniformly label any single site in a protein in vivo with a ^{19}F-amino acid will enable the study of large proteins with chemical clarity. Site-specific in vivo incorporation of fluorine into proteins at high yield will advance protein structural work,[3] solid-state studies,[12] and inhibitor design.[13,14]

Current methods to incorporate ^{19}F-labeled amino acids into proteins suffer from certain limitations. For example, semisynthetic incorporation enables high fidelity at specific sites but becomes impractical when medium to large proteins are needed. The use of natural translational machinery to force fluorinated mimics of Tyr, Trp, Phe, Met, and Leu into their natural codons can produce large, ^{19}F-labeled proteins. However, this method of incorporation alters all locations of one amino acid simultaneously, resulting in structural perturbation and overlapping of ^{19}F signals in large proteins.[1,5,15,16] Incorporation of fluorinated mimics rarely approaches 95% at one site when relying on natural machinery. When labeling does exceed 90%, special growing conditions and cell lines are needed.[17] With natural machinery, ^{19}F-amino acids are incorporated at different levels throughout the protein due to variation in codon usage.[18] High levels of ^{19}F-amino acid incorporation allows lower protein concentrations to be used and avoids a heterogeneous population of labeled protein.[17] Unfortunately, when replacing all of one amino acid in a protein, the heterogeneity increases with protein size, resulting in greater diversity of protein stabilities.

(1) Danielson, M. A.; Falke, J. J. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 163−195.
(2) Ulrich, A. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **2003**, *46*, 1−21.
(3) Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293−370.
(4) Bai, P.; Luo, L.; Peng, Z. *Biochemistry* **2000**, *39*, 372−380.
(5) Vaughan, M. D.; Cleve, P.; Robinson, V.; Duewel, H. S.; Honek, J. F. *J. Am. Chem. Soc.* **1999**, *121*, 8475−8478.
(6) Bann, J. G.; Frieden, C. *Biochemistry* **2004**, *43*, 13775−13786.
(7) Hoeltzli, S. D.; Frieden, C. *Biochemistry* **1998**, *37*, 387−398.
(8) Bourret, R. B.; Drake, S. K.; Chervitz, S. A.; Simon, M. I.; Falke, J. J. *J. Biol. Chem.* **1993**, *268*, 13089−13096.
(9) Luck, L. A.; Falke, J. J. *Biochemistry* **1991**, *30*, 4248−4256.
(10) Hinds, M. G.; King, R. W.; Feeney, J. *Biochem. J.* **1992**, *287* (Pt. 2), 627−632.
(11) Luck, L. A.; Mason, A. B.; Savage, K. J.; MacGillivray, R. T. A.; Woodworth, R. C. *Magn. Reson. Chem.* **1997**, *35*, 477−481.
(12) Goetz, J. M.; Schaefer, J. *J. Magn. Reson.* **1997**, *127*, 147−154.
(13) Pellecchia, M.; Sem, D. S.; Wuthrich, K. *Nat. Rev. Drug Discovery* **2002**, *1*, 211−219.
(14) Leone, M.; Rodriguez-Mias, R. A.; Pellecchia, M. *ChemBioChem* **2003**, *4*, 649−650.
(15) Feeney, J.; McCormich, J. E.; Bauer, C. J.; Birdsall, B.; Moody, C. M.; Starkmann, B. A.; Young, D. W.; Francis, P.; Havlin, R. H.; Arnold, W. D.; Oldfield, E. *J. Am. Chem. Soc.* **1996**, *118*, 8700−8706.
(16) Duewel, H.; Daub, E.; Robinson, V.; Honek, J. F. *Biochemistry* **1997**, *36*, 3404−3416.
(17) Frieden, C.; Hoeltzli, S. D.; Bann, J. G. *Methods Enzymol.* **2004**, *380*, 400−415.
(18) Luck, L. A.; Falke, J. J. *Biochemistry* **1991**, *30*, 4257−4261.

10.1021/ja064661t CCC: $37.00

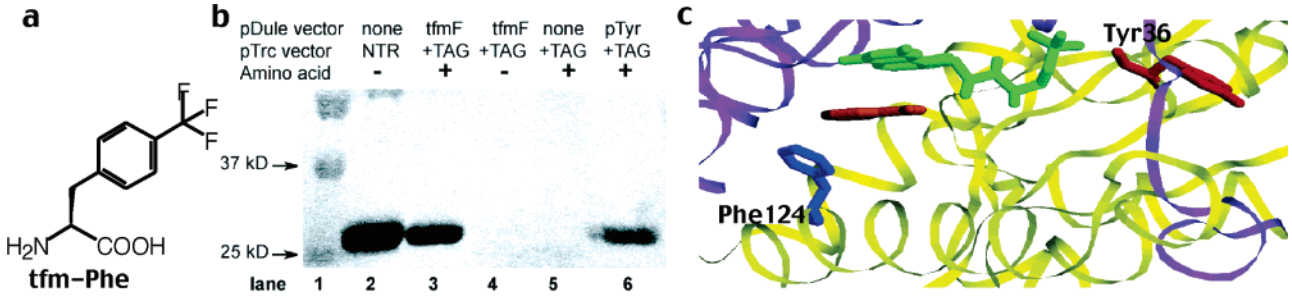

Figure 1. Selection of a modified *M. jannaschii* TyrRS that specifically incorporates tfm-Phe into proteins in response to a TAG stop codon in *E. coli*. **(a)** Structure of *p*-trifluoromethylphenylalanine (tfm-Phe). **(b)** SDS−PAGE analysis of tfm-Phe incorporation into NTR-124 as compared to native NTR and NTR-124 suppression with original MjTyrRS/tRNA$^{Tyr}_{CUA}$. Silver-stained gel of purified NTR samples by immobilized metal affinity chromatography. Lane 1 is a molecular weight marker, while protein production conditions are shown at the top of lanes 2−6. Lane 4 shows that, when tfm-Phe was withheld, no NTR was detectable by silver stain after SDS−PAGE. **(c)** Cross section of NTR dimer with subunits in yellow and purple (PBD code 1ICR). The active site Phe124 is blue, bound FMN is green, nicotinic acid is orange, and the interface Tyr36 is red.

In order to place a ^{19}F label at a single site in a protein genetically, Furter used the first nonessential tRNA/synthetase pair to incorporate a *p*-fluorophenylalanine and clearly points out the advantage of site-directed incorporation to many areas of protein structure study.[19] While this research was a great step in site-directed unnatural amino acid incorporation—producing up to 75% incorporation at a single site—the advance did not result in the site-specific incorporation that is needed for homogeneous protein structural characterization. This site-directed incorporation method also causes 3−7% of all phenylalanine residues to be replaced by *p*-fluorophenylalanine. Improvements on Furter's site-directed incorporation yields have been made by changing protein production conditions, but these also result in higher levels of background contamination.[17]

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.[20−22] Since it has been demonstrated by Schultz that an unnatural amino acid (UAA) can be site-specifically encoded with a unique, orthogonal tRNA/aminoacyl-tRNA synthetase pair, one could use this technology to cleanly label a single site with a fluorinated amino acid.[23−25]

In this article, we present the development of a simple tool for site-specific incorporation of trifluoromethyl-L-phenylalanine (tfm-Phe) into proteins and show that it is feasible to monitor protein structural changes by ^{19}F NMR with this fluorinated amino acid. An orthogonal aminoacyl-tRNA synthetase/tRNA pair capable of incorporating tfm-Phe into protein was selected. This synthetase/tRNA pair functions with high translational efficiency and fidelity when incorporating tfm-Phe into TAG nonsense codons in *Escherichia coli*. Further optimization of this genetic machinery allowed for production of protein at high yield for structural studies by ^{19}F NMR. Tfm-Phe was incorporated at multiple interface sites in two obligate homodimers, nitroreductase (NTR) and histidinol dehydrogenase (HDH). Both NTR and HDH have interface active sites and cannot function without dimeric quaternary structure, therefore allowing the use of activity as a measure of structural integrity. Since both NTR and HDH have active sites at their dimeric interfaces, structural changes are expected at the interface due to interactions at the active sites. We monitored the chemical shifts of the ^{19}F NMR signal that resulted from substrate and inhibitor binding to the active site of interface-labeled proteins. The ability to characterize interface conformational changes in these systems shows that tfm-Phe can be used to monitor structural changes in proteins by ^{19}F NMR. The signal quality of the three NMR-equivalent atoms on tfm-Phe was also sufficient to resolve different positional labeling within tfm-Phe NTR by in vivo ^{19}F NMR.

Results and Discussion

Selection of a tfm-Phe Synthetase. Previous site-specific incorporation of ^{19}F-amino acids failed since the UAAs chosen were permissive to more than one natural synthetase in the cell. For in vivo site-specific incorporation to succeed in *E. coli*, new ^{19}F-UAAs need be sufficiently structurally different from the natural amino acids so they are not recognized by natural *E. coli* synthetases. This structural differentiation from the natural amino acids is key to making the unnatural amino acid site-specific and not just site-directed. Tfm-Phe was chosen because it is not permissive, is nontoxic, gives a strong single ^{19}F signal resulting from a single rigid CF_3 group (Figure 1a), and should functionally replace at least hydrophobic amino acids.[2]

The in vivo, high-fidelity incorporation of tfm-Phe into proteins is accomplished by starting with a previously generated orthogonal tRNA-synthetase pair.[23] The modified *Methanococcus jannaschii* suppressor tRNA/aminoacyl-tRNA synthetase pair (MjTyrRS/tRNA$^{Tyr}_{CUA}$) has been used to site-specifically incorporate UAAs at sites specified by a TAG nonsense codon. The orthogonality of the MjTyrRS/tRNA$^{Tyr}_{CUA}$ with all the *E. coli* synthetase/tRNA pairs ensures that there is no background incorporation of tfm-Phe into any other amino acid sites and that incorporation into the desired TAG site approaches natural fidelity levels.[25] In order to generate a MjTyrRS that is selective for tfm-Phe, we started with a 6×10^7 library of MjTyrRS mutants in which six residues were randomized.[26] The mutant library was passed through three rounds of alternating

(19) Furter, R. *Protein Sci.* **1998**, *7*, 419−426.
(20) Lee, K. H.; Lee, H. Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. *Biochemistry* **2004**, *43*, 16277−16284.
(21) Jackel, C.; Seufert, W.; Thust, S.; Koksch, B. *Chembiochem* **2004**, *5*, 717−720.
(22) Higuchi, M.; Iwata, N.; Matsuba, Y.; Sato, K.; Sasamoto, K.; Saido, T. C. *Nat. Neurosci.* **2005**, *8*, 527−533.
(23) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498−500.
(24) Xie, J.; Schultz, P. G. *Methods* **2005**, *36*, 227−238.
(25) Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. *Nat. Biotechnol.* **2004**, *22*, 1297−1301.

positive and negative selection as previously described.[24,27] By challenging the cells to grow in the presence and in the absence of tfm-Phe in the positive and negative selections, respectively, only mutants expressing synthetases that efficiently incorporate tfm-Phe in response to the TAG codon survive.

The two top-performing synthetases were subcloned into the pDule vector (pDule-tfm-Phe1 and pDule-tfm-Phe2) for further analysis of protein containing tfm-Phe.[28] The medium-copy pDule plasmid is designed to contain both the $tRNA^{Tyr}_{CUA}$ and new synthetase, allowing a standard high-copy protein expression plasmid for use in production of tfm-Phe-containing protein. The resulting plasmids, pDule-tfm-Phe1 and pDule-tfm-Phe2, with their respective sequence changes [(Y32L, L65A, F108S, H109H, D158A, L162M) and (Y32Q, L65Q, F108Q, H109E, D158S, L162A)], performed equally well at incorporating tfm-Phe into proteins. Arbitrarily, pDule-tfm-Phe1 was used for all further NMR and MS protein characterization. Full-length NTR was produced when cells containing both plasmids, pDule-tfm-Phe1 (tfm-PheRS/$tRNA^{Tyr}_{CUA}$) and pTrc-NTR-124TAG (nitroreductase with TAG codon at Phe124), were induced to produce protein in modified minimal media supplemented with tfm-Phe (Figure 1b).[28] The purified yield of NTR-tfmPhe-124 ranged from 12 to 24 mg/L, and NTR-tfmPhe-36 yielded from 10 to 16 mg/L, while positive controls native NTR and NTR-Tyr-124 from the original MjTyrRS yielded 17−30 and 7−12 mg/L, respectively. The high yield of labeled protein resulting from expression of pDule-tfmPhe with a standard protein expression plasmid is a great advantage when performing protein structure studies.

Native NTR and NTR-tfmPhe-124 were analyzed by electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-Q-Tof Ultima) to further demonstrate that only a single tfm-Phe is inserted in response to the TAG codon. The native NTR and NTR-tfmPhe-124 showed the calculated mass shift of 68 Da, having masses 27 674.1 ± 0.2 and 27 742.9 ± 0.4 Da, respectively (Supporting Information, Supplemental Figure 2). To verify that the discrete incorporation of tfm-Phe for Phe takes place only at site 124, isolated proteins were digested by trypsin, and the fragments were analyzed by Micromass Q-Tof. The only peptide fragment containing tfm-Phe was the expected F(tfm-Phe)ADMHR fragment or (tfm-Phe)ADMHR fragment with 96.2% confidence and >80% coverage (Supporting Information, Supplemental Figure 3). No peaks were observed in the mass spectra corresponding to natural amino acid incorporation at site 124. These results confirm the high fidelity and efficiency of tfm-Phe incorporation into proteins.

Characterization of tfm-Phe at Protein Interfaces. To explore the utility of the ^{19}F signal from tfm-Phe in probing a range of protein conformational changes by NMR, tfm-Phe was incorporated into multiple interface locations within NTR and HDH. Since NTR and HDH are obligate homodimers with active sites at the interface, substrates or inhibitors can be added in defined orders to show that ^{19}F signal changes result from changes in protein interface structure and not from nonspecific binding to tfm-Phe. Incorporated at two sites in NTR, tfm-Phe would monitor the extent of interface change through a reaction cycle, one at the dynamic active-site residue F124 and the other at the stationary Y36 residue, which has its para position exposed to solvent (Figure 1c).[29,30] To determine if tfm-Phe can monitor ordered binding events and resulting subtle conformational changes at long distances, two sites 20−25 Å from the binding sites in HDH and >10 Å from any reported structural change were incorporated at the HDH hydrophobic interface (L225 and T377, Supporting Information, Supplemental Figure 1).[31]

Short-Distance Active-Site Probes. NTR is a homodimeric, 48 kDa nonspecific nitroreductase that uses a bound flavin mononucleotide (FMN) in two symmetrically placed interface active sites.[29,30] The structure, mechanism, and optimization of NTR have been studied extensively because it is currently being used in clinical trials to activate nitroaryl prodrugs for gene-directed cancer therapy.[32,33] NTR uses NADH or NADPH in a ping−pong, Bi−Bi mechanism to reduce a broad range of substrates.[29,30] NTR conformational changes are observed upon binding nicotinic acid (a competitive inhibitor of NADH). These conformational changes show interface site Phe-124 moving as a result of interacting with nicotinic acid, while a nearby interface site Tyr-36 is unaffected by inhibitor binding. Currently there are no reported structures of *E. coli* NTR in the reduced form—the form binding the electron-receiving substrate. Since the active-site cofactor, FMN, is bound at the interface, it could significantly perturb the active-site amino acids and protein interface. The ^{19}F probe, tfm-Phe, incorporated at Phe-124 and Tyr-36 should report on structural changes due to ligand binding and redox state via changes in ^{19}F NMR chemical shift.

To confirm that incorporation of tfm-Phe had minimal structural effect on NTR, native NTR activity and K_I values of the competitive nicotinic acid were compared to those for NTR-tfmPhe-124 and NTR-tfmPhe-36 by monitoring the consumption of NADH when reducing menadione. Native NTR activity and K_I are $7.7 \times 10^5\ M^{-1}\ s^{-1}$ and 20 μM, respectively. NTR-tfmPhe-124 was 6.5% more active, with a K_I of 21 μM, and NTR-tfmPhe-36 was 27% less active, with a K_I of 78 μM. Considering the close proximity of tfm-Phe to the interface active site, the alteration of these sites has surprisingly little effect on NTR activity.

All spectra were collected on a standard broadband probe with the proton coil tuned to the ^{19}F frequency for 1−18 h. To lock the ^{19}F signal, the common practice of adding D_2O to the buffer was avoided since there was a noticeable deuterium isotope effect on ^{19}F chemical shift. To eliminate unwanted chemical interactions and overlapping reference signals, all NMR samples contained a capillary tube containing lock and reference solvents. A mixture of 4-fluorotoluene in toluene-d_8 (a 0.2% solution) was used in an internal capillary reference tube because it provided a stable reference and lock signal and was free of thermally induced chemical shifts.

As expected, completely solvent-exposed tfm-Phe peptides, like those of denatured protein, show a narrow chemical shift

(26) Bose, M.; Groff, D.; Xie, J.; Brustad, E.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 388−389.
(27) Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 935−939.
(28) Farrell, I. S.; Toroney, R.; Hazen, J. L.; Mehl, R. A.; Chin, J. W. *Nat. Methods* **2005**, *2*, 377−384.
(29) Lovering, A. L.; Hyde, E. I.; Searle, P. F.; White, S. A. *J. Mol. Biol.* **2001**, *309*, 203−213.
(30) Haynes, C. A.; Koder, R. L.; Miller, A. F.; Rodgers, D. W. *J. Biol. Chem.* **2002**, *277*, 11513−11520.
(31) Barbosa, J. A.; Sivaraman, J.; Li, Y.; Larocque, R.; Matte, A.; Schrag, J. D.; Cygler, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1859−1864.
(32) Hu, L.; Yu, C.; Jiang, Y.; Han, J.; Li, Z.; Browne, P.; Race, P. R.; Knox, R. J.; Searle, P. F.; Hyde, E. I. *J. Med. Chem.* **2003**, *46*, 4818−4821.
(33) Grove, J. I.; Lovering, A. L.; Guise, C.; Race, P. R.; Wrighton, C. J.; White, S. A.; Hyde, E. I.; Searle, P. F. *Cancer Res.* **2003**, *63*, 5532−5537.

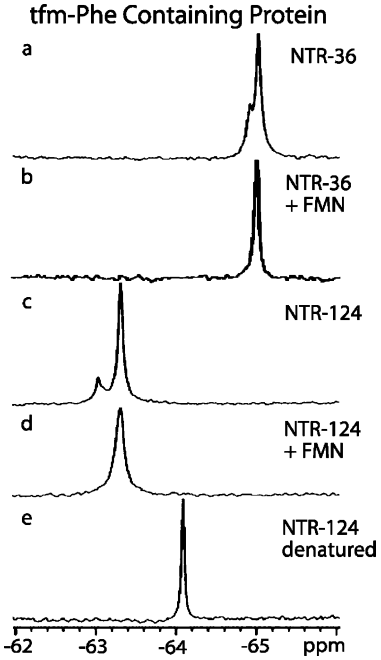

Figure 2. ^{19}F NMR spectra of tfm-Phe-containing protein. Spectra are referenced to an internal sealed capillary standard of 4-fluorotoluene in toluene-d_8 at −120.771 ppm.

range (−64.1 to −64.4 ppm), while tfm-Phe at internalized sites shifts appreciably (Figure 2). Since the CF_3 group on tfm-Phe is free to rotate on NMR time scales, resulting in a sharp single peak, spectra containing anything but a single peak have been a result of multiple protein conformations. The left shoulder on NTR-tfmPhe-36 and NTR-tfmPhe-124 is due to pure protein devoid of bound FMN, which collapses to a single peak upon prolonged incubation with FMN (Figure 2a−d).

To test the chemical shift sensitivity of tfm-Phe at NTR-124 and NTR-36 due to conformational changes at the protein interface, a competitive inhibitor was added and the active-site FMN was reduced and oxidized. Reduced NTR samples were prepared by the addition of NADH. Subsequent oxidation by addition of menadione in an argon environment was done to prevent noticeable NTR−FMN oxidation by oxygen. Fully reduced NTR protein can be easily visualized, as the samples change from bright yellow to colorless due to a reduction of FMN. While all spectral sets shown are from the same enzyme preparation, protein samples were run at least twice from different enzyme preparations to confirm the reproducibility of the ^{19}F resonances.

As predicted by crystal structures, NTR-tfmPhe-36 is not affected by the addition of inhibitor at concentrations greater than 10 times the K_I or by reduction (Figure 3, column 1). In contrast to NTR-tfmPhe-36, the signal from NTR-tfmPhe-124 is very sensitive to conformational changes in the active site. The reduction of FMN in NTR-tfmPhe-124 causes an upfield shift from −63.32 to −63.60 ppm (Figure 3, column 2). Since this conformational change involves a tightly bound cofactor, one is able to monitor separate peaks as protein is oxidized by stepwise addition of substrate. The NTR-tfmPhe-124 signal is also affected by nicotinic acid binding, as predicted by crystal structures (Figure 3, column 3).

In contrast to NTR reduction, a single exchange-averaged resonance is observed due to the rapid interconversion rate of bound and unbound nicotinic acid NTR forms relative to the NMR frequency difference for these forms (Figure 3, column 3). Because nicotinic acid is a competitive inhibitor for the oxidized form of NTR, the displacement of this average signal from NTR-124 should represent the amount of enzyme−inhibitor complex in solution. The normalized ^{19}F signal shift matches well with %ES complex calculated from $K_I = E^*I/EI$ (Supporting Information, Supplemental Figure 4).

In Vivo Probe. While a few ^{19}F imaging studies have labeled a protein in vitro and then reintroduced it into a living organism for monitoring, no study has labeled and monitored a ^{19}F-protein in vivo.[34,35] The method of ^{19}F-labeling presented here is truly site-specific for one position in one protein, therefore changes to that protein can be monitored in vivo. Labeled NTR was monitored in vivo by placing washed cell paste into a standard 5 mm NMR tube. The signal quality of tfm-Phe is sufficient to resolve NTR-tfmPhe-36 and NTR-tfmPhe-124 from each other in vivo (Figure 3, column 4). Changes in the chemical shifts of their ^{19}F NMR signal were not detected upon the addition of nicotinic acid or NADH to the cell paste. The inability to monitor chemical shifts in vivo, as seen with in vitro NTR-tfmPhe-124 studies, could be due to the poor cell permeability of nicotinic acid and NADH. Attempts to reduce the large peak width of NTR-tfmPhe-36 by varying the level of protein in the cell and cell density in the NMR tube were unsuccessful. The severe line broadening of the NTR-tfmPhe-36 signal could be due to the exposure of the trifluoromethyl group to the solvent when incorporated at this location in the interface. This method of labeling proteins provides adequate ^{19}F signal to distinguish between different labeled sites in vivo. Expressing labeled ^{19}F-protein in this manner and monitoring washed cells should provide access to protein conformational changes for proteins that cannot be purified or that need a natural setting for processing.

Long-Distance Active-Site Probe. Site-specific incorporation of tfm-Phe into HDH also served to monitor the binding of substrates at significant distance from its active site (Figure 4). HDH is a homodimeric, 98 kDa metalloenzyme with one metal and substrate binding site per monomer. HDH oxidizes L-histidinol to L-histidine following a Bi−Uni−Uni−Bi reaction mechanism in two consecutive oxidation reactions.[31] Placing tfm-Phe at two locations (20−25 Å from the active site) within the extensive hydrophobic interface between the homodimers allowed for monitoring of ordered binding events in the HDH reaction cycle.[31] Both sites make contact with the opposite subunit and are at least 10 Å from any section of the protein that is reported to move upon binding substrates. Site T377 is completely removed from solvent, whereas L225 is on the edge of the hydrophobic interface and is partially exposed to solvent (Supplemental Figure 1). For a reaction in HDH to take place, the metal (in this study Mn^{2+}) must bind first, then histidinol, and finally NAD^+.[36] We have shown with both labeled sites that no change in ^{19}F signal (chemical shift or peak shape) takes place upon the addition of metal ion or histidinol independently; only when both cofactor and substrate are combined with the protein do ^{19}F signals change. Both labeled HDH samples

(34) Shimizu, M.; Kobayashi, T.; Morimoto, H.; Matsuura, N.; Shimano, T.; Nomura, N.; Itoh, S.; Yamazaki, M.; Iriguchi, N.; Yamamoto, T.; Yamai, S.; Furuta, T.; Maki, T.; Mori, T. *Magn. Reson. Med.* **1987**, *5*, 290−295.
(35) Mehta, V. D.; Kulkarni, P. V.; Mason, R. P.; Constantinescu, A.; Antich, P. P. *Bioconjugate Chem.* **1994**, *5*, 257−261.
(36) Kanaori, K.; Uodome, N.; Nagai, A.; Ohta, D.; Ogawa, A.; Iwasaki, G.; Nosaka, A. Y. *Biochemistry* **1996**, *35*, 5949−5954.

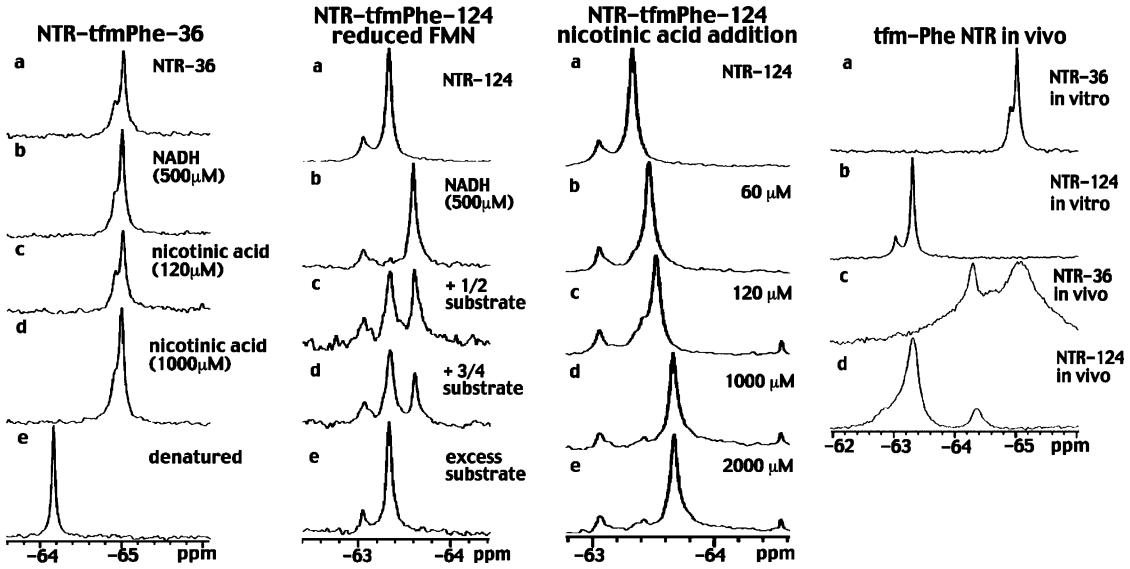

Figure 3. In vitro ^{19}F NMR spectra of NTR-tfmPhe-36 and NTR-tfmPhe-124 with the addition of inhibitor and substrates. Column 1 shows changes in the ^{19}F NMR spectrum of NTR-tfmPhe-36 upon reacting with NADH and binding nicotinic acid to the active site and changes upon denaturation (shoulder on peak at −65.00 ppm results from NTR-tfmPhe-36 without bound FMN). Column 2 **(**a,b**)** shows the effects of reducing the active site FMN in NTR-tfmPhe-124 by addition of excess NADH. Stepwise reoxidation of NTR-tfmPhe-124 is shown by addition of menadione **(**c−e**)**. Column 3 shows the stepwise addition of nicotinic acid to NTR-tfmPhe-124. As expected, the fraction of NTR-tfmPhe-124 without bound FMN is unaffected by active-site binding (peak at −63.04 ppm) or denatured protein (peak at −64.56 ppm). Column 4 shows comparison of in vivo vs in vitro ^{19}F NMR spectra for NTR-tfmPhe protein. The peak at −64.3 ppm in both in vivo spectra **(**c,d**)** likely represent residual free tfmPhe in the cell or denatured protein. The^{19}F signal from buffered free tfm-Phe is at −64.65 ppm, while addition of free amino acid to in vivo samples increased the −64.3 ppm signal.

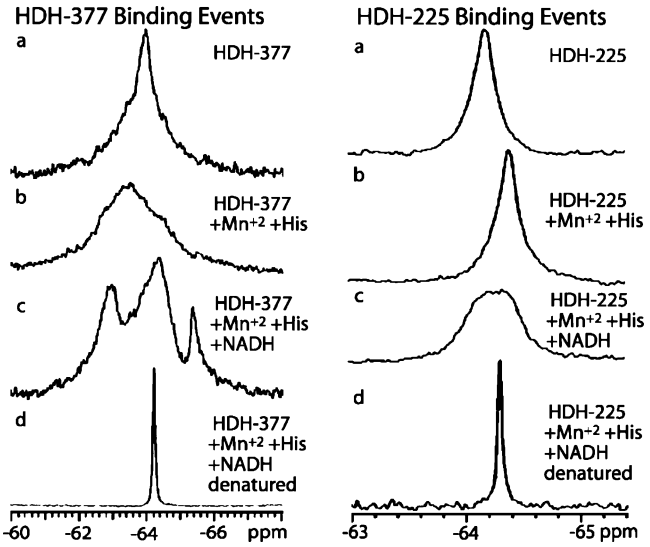

Figure 4. In vitro ^{19}F NMR spectra of HDH-tfmPhe-377 and HDH-tfmPhe-225 with the sequential addition of substrates. Column 1a, apo HDH-tfmPhe-377 protein, peak at −64.00 ppm; column 1b, HDH-tfmPhe-377 with 1.0 mM Mn^{2+} and 2.0 mM histidinol, peak at −63.60 ppm; column 1c, HDH-tfmPhe-377 with 1.0 mM Mn^{2+}, 2.0 mM histidinol, and 10 mM NADH, peaks at −62.80, −64.40, and −65.80 ppm; column 2a, apo HDH-tfmPhe-225, peak at −64.20 ppm; column 2b, HDH-tfmPhe-225 with 1.0 mM Mn^{2+} and 2.0 mM histidinol, peak at −64.40 ppm; column 2c, HDH-tfmPhe-225 with 1.0 mM Mn^{2+}, 2.0 mM histidinol, and 10 mM NADH, peak at −64.30 ppm; columns 1d and 2d, the resulting spectra from adding guanidinium-HCl to the respective sample from **(**c**)**.

exhibit a ^{19}F peak change, reflecting the predicted conformational change of binding Mn^{2+} and histidinol (Figure 4, column 1, and Figure 2a,b). The sharper peak of HDH-225 shows a 0.2 ppm shift upfield, while the broader peak of HDH-377 shows a 0.4 ppm shift downfield and broadening indicative of an increase in conformational states or a fast equilibrium between bound and unbound states. When sufficient second substrate, NAD^+, was added to the NMR sample (Figure 4b) to convert all of the histidinol substrate to product histidine, the signal for both HDH sites reverted to the unbound form, shown in Figure 4a. If the electron-acceptor product, NADH, is added to each mixture in Figure 4b, the binding of NADH is expected instead of turnover of histidinol. The addition of NADH to the HDH-377-tfm-Phe sample caused it to divide into what appears to be three distinct conformational states, none of which match the previous forms in chemical shift (Figure 4, column 1c). The addition of NADH to the HDH-255-tfm-Phe sample shows an averaged signal apo form and Mn^{2+}/histidinol-bound form (Figure 4, column 2c). Denaturation of the samples from Figure 4c, with the addition of buffered guanadinium-HCl, produces the expected sharp singlet at −64.22 ppm in both samples (Figure 4d). Dialysis of the sample (Figure 4c) returns the peak to the form shown in Figure 4a. While the HDH tfm-Phe signals are broader than the NTR tfm-Phe signals, likely due to the lack of rigidity in the hydrophobic interface, the shift changes reflect the proper binding order of substrates, even 20−25 Å from the active site. There is also the possibility that Mn^{2+} could be acting as a spin-relaxing agent, broadening the lines shown in Figure 4.[37]

Conclusion

These studies describe the first site-specific in vivo incorporation of a fluorinated amino acid. The tfm-Phe was incorporated into *E. coli* protein with high yield and was used to monitor small-molecule binding and protein conformational changes. Tfm-Phe protein shows strong sharp ^{19}F NMR signals for in vivo and in vitro studies with the standard advantages of environmentally sensitive chemical shifts. The high-fidelity labeling resulting from this method also allows access to in vivo

(37) Troughton, J. S.; Greenfield, M. T.; Greenwood, J. M.; Dumas, S.; Wiethoff, A. J.; Wang, J.; Spiller, M.; McMurry, T. J.; Caravan, P. *Inorg. Chem.* **2004**, *43*, 6313−6323.

experiments and quantification of protein conformational states. Because tfm-Phe is incorporated genetically, any protein that is expressed in *E. coli*, irrespective of size or sequence, can be studied using this technique. The ease with which labeled protein can be produced should facilitate the study of protein conformations, interactions, and processing. While this is the first fluorinated amino acid incorporated with high translational efficiency and fidelity, this method of incorporating unnatural amino acids has proven to be general. This general method has facilitated incorporation of >30 diverse UAA structures and will provide access to incorporation of other fluorinated UAAs.[23,24,38] Ongoing efforts in our laboratory focus on tfm-Phe protein in 2D NMR studies, membrane proteins studies, and application to ^{19}F MRI.

Experimental Section

General Methods. Chemical reagents were purchased from Sigma-Aldrich and Fisher Scientific and used without further purification. L-4-Trifluoromethylphenylalanine was purchased from Peptech. Oligonucleotides, dH10B cells, and pTrcHisA were purchased from Invitrogen.

Selection of tfm-Phe-Specific tRNA-Aminoacyl Synthetase. The general selection method of Xie and Schultz was used.[24,26] The library of tRNA-aminoacyl synthetases was encoded on a kanamycin (Kn)-resistant plasmid, pBK-JM-Lib, under control of the constitutive *E. coli* GlnRS promoter and terminator. The library was transformed by electroporation into DH10B *E. coli* containing the tetracycline (Tet)-resistant positive selection plasmid, pREP2/YC-JYCUA. The positive selection plasmid encodes a mutant $tRNA^{Tyr}_{CUA}$, a TAG stop codon-disrupted chloramphenicol acetyltransferase, and an amber-disrupted T7 RNA polymerase that drives the production of green fluorescent protein. The first positive selection was performed in the presence of 1 mM tfm-Phe, 70 μg/mL chloramphenicol (Cm), 50 μg/mL Kn, and 25 μg/mL Tet in a glycerol based minimal media supplemented with leucine and metals (modified GMML)[28] at 37 °C with 250 rpm shaking until the OD was >1.0. Plasmid DNA was extracted, and pBK-JM-Lib was separated from pREP2/YC-JYCUA by agarose gel electrophoresis. The purified pBK-JM-Lib was then transformed into DH10B *E. coli* containing the ampicillin (Amp)-resistant negative selection plasmid, pJ17B3, encoding a mutant $tRNA^{Tyr}_{CUA}$ and an amber-disrupted barnase gene under control of an arabinose promoter and rrnC terminator. The rescued negative selection cells were plated on Luria−Bertani (LB) agar plates containing 100 μg/mL Amp, 50 μg/mL Kn, and 0.2% arabinose. Following 18 h of growth at 37 °C, the plates were scraped into LB medium, plasmid DNA was extracted, and pBK-JM-Lib was purified by gel electrophoresis. The remaining pBK-JM-Lib was transformed into positive selection cells and selected as described in modified GMML medium but on agar plates instead of in liquid medium. After two more rounds of positive and negative selection, remaining pBK-JM-Lib members were transformed into positive selection cells and grown on modified GMML medium plates in the presence of 1 mM tfm-Phe, 100 μg/mL Cm, 50 μg/mL Kn, 25 μg/mL Tet, and 0.002% arabinose. Single individual colonies (56) were selected from the surviving library and screened in this medium in the presence and in the absence of tfm-Phe and varying concentrations of Cm from 0 to 120 μg/mL. The efficiency and fidelity of the synthetases (7) that survived in >120 μg/mL Cm in the presence of tfm-Phe, and <10 μg/mL Cm in the absence of tfm-Phe, were carried on to tfm-Phe-protein production trials. The individual tfm-PheRS plasmids were transformed into DH10B cells containing pBAD/JYAMB-4TAG (Tet resistance), which encodes the mutant $tRNA^{Tyr}_{CUA}$ and the amber-disrupted sperm whale myoglobin gene.[27] The C-terminal His-6 tag on the protein facilitated monitoring of protein yield as well as protein purity through silver-stained SDS−PAGE analysis. The two synthetases that produced the highest yield of purified tfm-Phe-myoglobin when tfm-Phe was added to the medium were subcloned into the pDule vector under control of the lpp promoter and the rrnB terminator.[28] These synthetases showed no detectable background myoglobin production when tfm-Phe was withheld from the medium. The plasmid pDule-tfm-Phe1 was used for all further tfm-Phe protein production and is available upon request.

Expression and Purification of tfm-Phe-Containing Nitroreductase and Histidinol Dehydrogenase. The standard overexpression plasmid pTrcHisA, containing the *E. coli* nitroreductase gene, *nfnB*, or histidinol dehydrogenase gene, *hisD*, had TAG mutations incorporated to replace F124 and T36 or L225 and T377, respectively. The TAG-containing overexpression plasmids created were pTrc-NTR-124TAG, pTrc-NTR-36TAG, pTrc-HDH-225TAG, and pTrc-HDH-377TAG. *E. coli* DH10B cells containing pDule-tfm-Phe1 and the appropriate pTrc-TAG plasmid were grown in modified minimal medium in growth volumes of 0.1−1.0 L (medium was supplemented with 1 mM tfm-Phe where applicable and flavin at 10 μM for NTR protein expressions). Expressions were induced to produce protein with 1.0 mM IPTG at an OD of 0.8, as previously described.[28] Cells were harvested 18 h after induction by centrifugation and stored at −80 °C. Cells (from 0.5 L of medium) for in vivo NMR measurements were thawed on ice, resuspended in 50 mL of PBS buffer (10 mM NaH_2PO_4, 140 mM NaCl, pH 7.8) at 0 °C, and repelleted by centrifugation at 5000*g*. The cells for in vivo NMR spectra, shown for NTR-tfmPhe-36 and NTR-tfmPhe-124, were washed three and five times, respectively. Protein for in vitro studies was purified to >95% by Co ion affinity chromatography using BD Talon resin. Pure protein was desalted into 10 mM Tris-HCl buffer, pH 7.8, using PD10 columns. Tfm-Phe-protein yield varies largely due to induction OD and expression scale.[28] Pure protein solutions were concentrated by centrifugation to 3.0 mg/mL for NMR studies. The purified yield of HDH-225-tfm-Phe and HDH-377-tfm-Phe varied between 10 and 20 mg/L. No full-length HDH protein was detectable by silver stain after Co ion affinity chromatography and SDS−PAGE when tfm-Phe was withheld from medium during expression of protein from pTrc-HDH-225TAG and pTrc-HDH-377TAG in the presence of pDule-tfm-Phe1 (as shown with the NTR expressions described in text and Figure 1).

NMR Measurements. NMR data were collected using a Varian Unity INOVA 500 MHz spectrometer fitted with a 5 mm broadband probe. The proton coil was tuned to fluorine frequency (470.114 MHz). Standard decoupling parameters were used with a 100 000 Hz spectral width, 9 μs pulse length (approximately 45°), 2.00 s acquisition time, and 1.00 s relaxation delay. A 10 Hz line broadening was applied, and all spectra were recorded at 25 ± 0.2 °C. All spectra were locked and referenced to a solvent mixture (0.2% solution of 4-fluorotoluene in toluene-d_8) in a capillary tube within every NMR sample. The 4-fluorotoluene ^{19}F signal in this internal reference capillary tube was set at −120.771 ppm, referenced from trifluorotoluene at −65.000 ppm. The common practice of adding D_2O to the buffer was avoided due to a noticeable deuterium isotope effect on ^{19}F chemical shift. In vivo samples contained 0.4 mL of washed cell paste and 0.1 mL of PBS buffer. Additions to NMR samples came from buffered stocks, and the total volume of solution added was always less than 2% of the total sample volume. Reduced NTR samples were prepared by the addition of NADH and subsequent oxidation by addition of menadione in an argon environment to prevent noticeable NTR−FMN oxidation by oxygen. All purified protein samples were at 0.1 mM, with the exception of the denatured samples. Denatured samples contained 0.05 mM protein and 4.2 M guanidinium-HCl at pH 7.8 and were heated to 70 °C for 5 min, with the exception of HDH-377-tfmPhe (Figure 4c), which required 6.3 M guanidinium-HCl to ensure complete denaturation. Completely bound FMN−protein samples (Figure 2) were generated by incubating pure protein with 100 μM FMN for 1−2 weeks at room temperature.

(38) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225−249.

NTR Kinetics. Assays were performed spectrophotometrically in the presence of 10 mM Tris-HCl buffer (pH 7.8), 150 μM NADH, and varying amounts of menadione (10−700 μM). DMSO was used to prepare a 100 mM stock solution of menadione. For menadione, the reaction was initiated with 10 μL of NTR to give a final concentration of 2 nM. The progress of the reaction was monitored at 340 nm by observing the oxidation of NADH. This was converted to a rate of reduction of menadione using the molar absorbance of NADH (ϵ = 6200 M^{-1} cm^{-1}), assuming 2 mol of NADH consumed per mole of reduced menadione and normalized for enzyme concentration. Michaelis constants were calculated from initial velocity measurements (the observed rate of the reaction for the first minute) at different substrate concentrations using Wilman4 Kinetic software. For nicotinic acid inhibition studies, nicotinic acid concentrations of 7.5 and 30 μM were used. Menadione was maintained at 400 μM, and NADH was varied from 5 to 500 μM.

Acknowledgment. We thank Beth Buckwalter for her technical assistance and Claude Yoder, Dave Wemmer, and Peter Fields for their helpful discussions. This work was supported by F&M Hackman and Eyler funds, NSF MCB-0448297, Research Corporation (CC6364), and ACS-PRF (42214-GB4).

Supporting Information Available: Description of the mass spectrometry of proteins and %ES complex calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

JA064661T