DOI: 10.1002/cbic.200500249

Biosynthetic Incorporation of Fluorohistidine into Proteins in *E. coli*: A New Probe of Macromolecular Structure

Jack F. Eichler,[a] John C. Cramer,[b] Kenneth L. Kirk,[b] and James G. Bann*[a]

Histidine is important for carrying out a number of protein functions. For instance, it can act as a general acid/base in enzymatic catalysis,[1] it plays a role in ligating metal cations in metalloproteins,[2] and stabilizes protein structure by metal binding, hydrogen bonding, or electrostatic interactions.[3] One noteworthy example of histidine's potential role in protein stability is observed in the pathogenesis of the anthrax toxin, where protonation of His side chains in one of the proteins leads to a large structural perturbation and subsequent toxicity.[4] Given the unique role of this amino acid in such processes, the development of methods that probe the structural and mechanistic features of His in proteins would be extremely valuable.

The biosynthetic incorporation of unnatural amino acids into proteins provides the experimentalist with a variety of methods that can probe protein structure and function.[5] In particular, fluorinated amino acids can be used to achieve a relatively isosteric change (by replacing a single hydrogen with fluorine) that results in quite different electronic properties.[6] Additionally, ^{19}F NMR can be used to monitor changes in protein conformation in response to changes in the environment that are sometimes not detectable by other techniques.[7] Although incorporating fluorine-labeled amino acids is not a new idea,[8] access to an expanding tool box of fluorinated protein building blocks has promoted a renaissance in this area of research.[5,9]

Over 30 years ago, a photochemical Schiemann reaction was developed for synthesizing 2-fluorohistidine (2-FHis) and 4-fluorohistidine (4-FHis).[10] To our knowledge this still represents the only fluorination procedure available for accessing these imidazole derivatives. The pK_a of the side chain of both 2-FHis and 4-FHis has been measured previously, and is decreased from approximately 6.0–6.5 to 1 and 3, respectively.[11] Because of this, these analogues provide a means for verifying the role of native His in pH-dependent processes. To this end, 4-FHis has been incorporated into the S peptide of ribonuclease, and into full-length ribonuclease A with chemical synthetic methods.[8d,12]

Early experiments demonstrated that tritium-labeled 2-FHis could be incorporated into bacterial protein, and these analogues (2-FHis more so than 4-FHis) had an inhibitory effect on *E. coli* growth.[13] However, in order to achieve high levels of incorporation for structural studies, it is typical to employ the use of bacterial auxotrophs. While the use of auxotrophs for the biosynthetic incorporation of novel His analogues into *E. coli* has been reported,[14] a similar protocol for the incorporation of fluorohistidine derivatives has yet to be described. Herein, we provide unequivocal evidence for the incorporation of both 4-FHis and 2-FHis into a mutant form of the chaperone PapD by using an *E. coli* strain that is auxotrophic for His.

PapD is the prototype for a wide variety of highly homologous chaperones that utilize the chaperone-usher pathway for the assembly of P-pili, and has been previously labeled with fluorophenylalanine and fluorotryptophan in protein-folding studies.[7a,b] The wild-type (WT) protein does not contain any His residues. Thus, site-specific labeling can be accomplished by the introduction of a single His residue by site-directed mutagenesis, and biosynthetic labeling can be performed according to previously described protocols.[15]

In this work, we used site-directed mutagenesis to introduce a single His residue at Arg200 in PapD. Among the chaperones that are homologous to PapD, the most similar is PmFD from *Proteus mirabilis* (47% identity), which possesses a His residue at position 200.[16] Therefore, an R200H substitution was not expected to alter the structure and stability of PapD. To confirm this, urea denaturation studies were performed on PapD (R200H) and it was found to have a similar stability as that previously reported for WT PapD; PapD (WT): $\Delta G^\circ = 8.95$ kcal

[a] *Dr. J. F. Eichler, Prof. Dr. J. G. Bann*
Department of Chemistry, Wichita State University
Wichita, KS 67226 (USA)
Fax: (+1) 316-978-7373
E-mail: jim.bann@wichita.edu

[b] *Dr. J. C. Cramer, Dr. K. L. Kirk*
Laboratory of Bioorganic Chemistry
National Institute of Diabetes & Digestive & Kidney Diseases
National Institutes of Health
Bethesda, MD 20892 (USA)

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

mol^{-1}, $m = 2.7$ kcal mol^{-1} M, midpoint = 3.31 M;[7b] PapD(R200H): $\Delta G^{\circ} = 6.58$ kcal mol^{-1}, $m = 2.12$ kcal mol^{-1} M, midpoint = 3.10 M (see Experimental Section).

After sequencing the PapD(R200H) plasmid DNA to verify the incorporation of His, it was transformed into the His-auxotrophic bacterial strain, UTH780.[17] Standard labeling and expression procedures[15] were then used to incorporate 2-FHis and 4-FHis into PapD(R200H). Figure 1 shows the results from

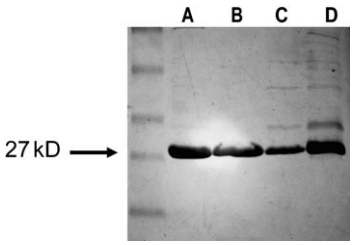

Figure 1. SDS-PAGE analysis of A) PapD(WT); B) PapD(R200H); C) PapD-(R200H-4F); D) PapD(R200H-2F). All samples were obtained after consecutive ion exchange and gel-filtration column purifications. Differences in purity of proteins are presumed to be due to variances in the periplasmic extraction process.

the SDS-PAGE analysis of PapD(WT), PapD(R200H-2F), and PapD(R200H-4F) proteins. It is evident from the gel that both labeling experiments resulted in PapD production.

In order to determine if fluorohistidine was incorporated into PapD protein, PapD(R200H), PapD(R200H-2F), and PapD(R200H-4F) were also analyzed by using ESI and MALDI-TOF mass spectrometry. The results from the MALDI-TOF analyses are illustrated in Figure 2. The mass spectra of PapD(R200H-2F) and PapD(R200H-4F) showed that both His analogues were indeed incorporated, as evidenced by the expected increase in mass compared to unlabeled PapD(R200H). A summary of all mass spectrometry analyses is presented in Table 1.

Based on the data shown in Figure 3 and Table 1, it is evident that there was only partial incorporation of 4-FHis into PapD(R200H). However, a lower amount of this reagent was used in the bacterial growth (4-FHis: 0.010 g; 2-FHis: 0.020 g) because only limited quantities were available for these experiments. Hence, at this time we cannot be certain whether the partial incorporation of 4-FHis is due to an insufficient amount of amino-acid analogue in the growth media or some other mechanism (for instance less efficient binding to the histidyl-tRNA synthetase).[18]

The incorporation of fluorohistidine into PapD was also verified by ^{19}F NMR. The ^{19}F spectrum of PapD(R200H-2F) was obtained before and after treatment with urea (5 M). Figure 3 illustrates the chemical shift that was obtained upon denaturation of fluorine-labeled PapD (0 M urea: $\delta = -37.3$; 5 M urea: $\delta = -35.9$). The observation of a single broad resonance for PapD(R200H-2F) at 0 M urea is consistent with the presence of only one His residue. Addition of urea (up to 5 M) caused the resonance to sharpen and shift downfield toward that of free 2-FHis ($\delta = -31.9$). This indicates an increased flexibility of the His side chain after protein denaturation. The presence of a shoulder at 5 M urea also suggests that some intermediate

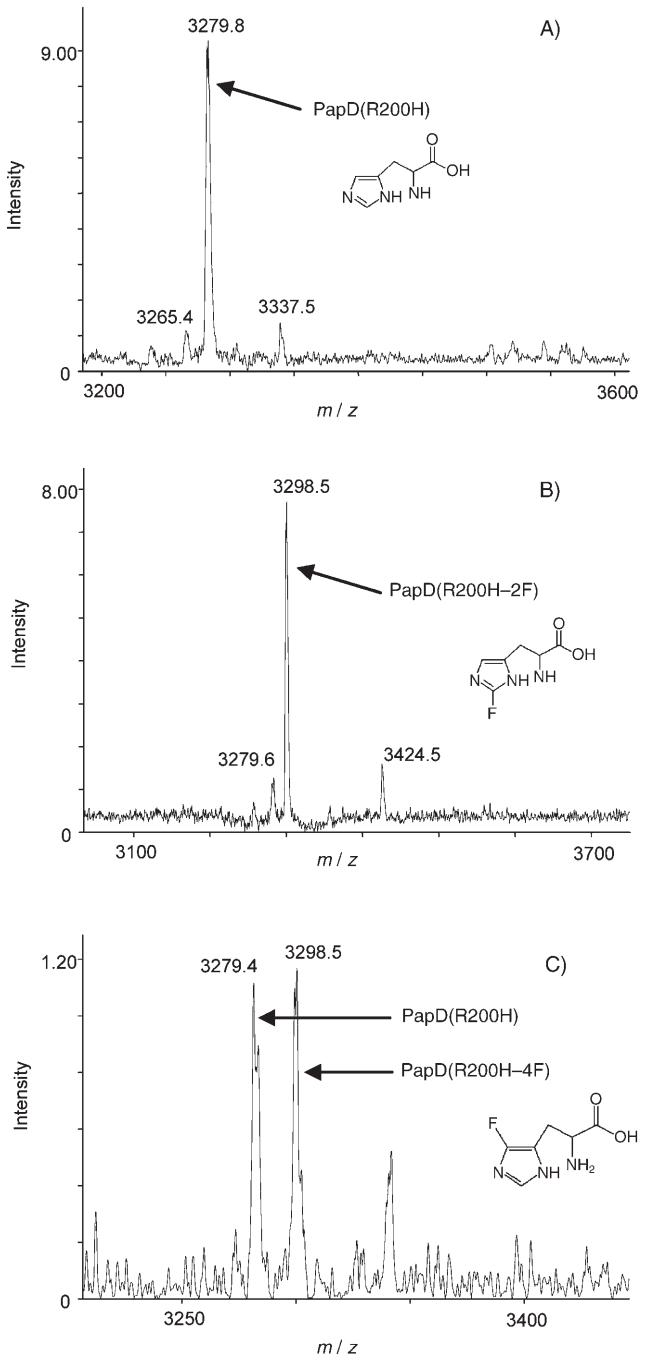

Figure 2. MALDI-TOF mass spectrometry analyses of A) PapD(R200H), B) PapD(R200H-2F), and C) PapD(R200H-4F). The MALDI-TOF spectra were performed on trypsin digests of the parent proteins. Peptides corresponding to amino acids 204–232, which contain the R200H substitution, were analyzed. When Arg is substituted with His, the theoretical monoisotopic mass of the peptide (SANYNTPYLSYINDYGG**H**PVLSFICNGSR) is 3279.5 D.

Table 1. Summary of mass spectrometry analyses.

	PapD (R200H)	PapD (R200H-2F)	PapD (R200H-4F)
ESI [*m/z*]	24543.0	24557.0	24560.0 (~55%) 24544.0 (~45%)
MALDI-TOF [*m/z*] (trypsin digest)	3279.8	3298.5	3298.5 (~55%) 3279.4 (~45%)

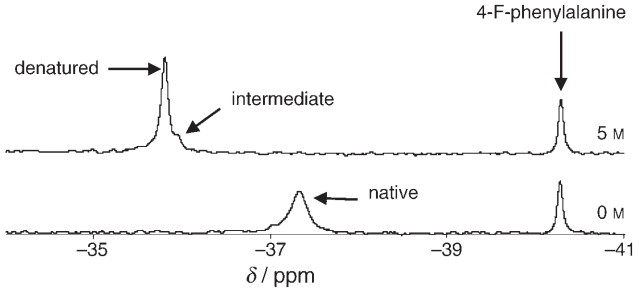

Figure 3. ^{19}F NMR spectra of PapD(R200H-2F) at 0 and 5 M urea. Spectra were obtained at 20 °C with protein (50 μM) in MOPS-HCl (30 mM, pH 7.0)/D_2O (10%, *v/v*)/4-fluorophenylalanine (30 μM); internal reference, $\delta = -40.29$.

unfolded structure might be present. All of these results are in agreement with previous studies on PapD.[7a,b]

This work represents the first unambiguous demonstration of the successful incorporation of both 2-FHis and 4-FHis into a protein by using biosynthetic methods. These results open a new avenue for the structural and functional characterization of proteins that utilize His for carrying out physiological processes. In addition, the incorporation of these His analogues should provide a means for determining the role of the imidazole ring in protein stability, for example, in the pH-dependent conformational change observed in the pore formation of the anthrax toxin. The biosynthetic incorporation of 2-FHis and 4-FHis into proteins that undergo pH-dependent conformational changes will be pursued with the aim of determining whether these fluorohistidine analogues have an impact on such processes.

Experimental Section

Materials: 2-FHis was prepared as previously described.[19] A new enantioselective procedure, based on our previously reported synthesis[8] was used to prepare 4-FHis.[20] The His auxotrophic bacterial strain, UTH780 (CGSC Strain #5954), was obtained from the Yale University *E. coli* Genetic Stock Center.[17]

Site-directed mutagenesis: The PapD(R200H) mutant was generated according to the protocol provided with the Stratagene Quick-Change Site-Directed Mutagenesis Kit, by using the following primers:
Forward primer: 5′-GACTATGGTGGTCACCCGGTACTGTCGTTTATC-3′
Reverse primer: 5′-GATAAACGACAGTACCGGGTGACCACCATAGTC-3′

The sequence of the PapD(R200H) mutant DNA was verified by the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at Washington University (St. Louis, USA).

Protein expression: The pQE80 plasmid containing the PapD (R200H) gene was transformed into UTH780. A single colony was inoculated into Lauria–Bertani (LB; 5 mL) media that was supplemented with ampicillin (0.1 mg mL^{-1}), and grown for approximately 8 h at 37 °C. This culture was diluted (1:100) in fresh media and grown at 37 °C, overnight. The culture was then further diluted (1:100) in fresh media (600 mL), grown to $OD_{600} = 0.8$ by using a platform shaker, and then IPTG was added (final concentration 1 mM). The culture was grown for an additional 45 min and then the cells were harvested. The protein was isolated from the periplasm by resuspending the bacterial pellet (5 mL per gram of bacteria) in sucrose (20%)/Tris pH 8.0 (20 mM). EDTA (100 mM) and lysozyme (10 mg mL^{-1}) were added to give final concentrations of 100 mM and 50 μg mL^{-1}, respectively. After incubation for 10 min in an ice bath, $MgCl_2$ was added (final concentration 20 mM). The solution was centrifuged and the supernatant containing the PapD (R200H) was stored at −20 °C until further purification was carried out.

Fluorohistidine labeling: The labeling of PapD(R200H) was generally performed according to the protocol reported for the site-specific labeling of PapD with fluorophenylalanine.[15] The following changes were implemented: a culture of UTH780 (50 mL) that carried the pQE80–PapD(R200H) plasmid was grown in LB media at 37 °C for approximately 10 h; the culture was centrifuged and the bacterial pellet resuspended in minimal media (600 mL for 2-FHis labeling, 300 mL for 4-FHis labeling) that was supplemented with His (0.2 mM); this culture was grown at 37 °C to $OD_{600} = 3$. The bacteria were harvested and washed with NaCl (0.9%, 250 mL). This step was repeated, the bacteria were reharvested, and then resuspended in fresh minimal media (600 mL for 2-FHis, 300 mL for 4-FHis) that contained fluorohistidine (0.20 mM). The cultures were incubated for 10 min and then induced with IPTG (final concentration 1 mM). After 45 min of growth, the cultures were harvested and the protein was isolated from the periplasm as described above. Glucose concentration was monitored during all phases of bacterial growth and was found not to fall below 0.2%.

Protein purification: Periplasm solutions were dialyzed twice for approximately 8 h in 2-morpholinoethanesulfonic acid (MES; 2 L, 20 mM, pH 6) at 4 °C. The protein samples were purified sequentially on a HiTrap SP FF (Amersham Biosciences) ion exchange column and H*i*Prep Sephacryl S-100 gel filtration column. Fractions containing PapD(R200H), PapD(R200H-2F), and PapD(R200H-4F) were analyzed by using SDS-PAGE and found to be adequately pure for analysis by mass spectrometry. Further purification was not attempted so as to ensure that sufficient concentrations of protein were available for ^{19}F NMR studies.

Urea denaturation: PapD(R200H) was diluted in 3-morpholinopropanesulfonic acid (MOPS)-HCl (pH 7.0; final concentration 1 μM). Samples were treated with the appropriate concentration of urea (0–5.5 M) and incubated at room temperature, overnight. Fluorescence-emission intensities (350 nm) were measured on a Cary Eclipse fluorescence spectrophotometer at 20 °C by using an excitation wavelength of 295 nm. Slits were set at 5 nm for excitation and 10 nm for emission. The ΔG°, cooperativity index (*m*), and midpoint were obtained from fitting the plot of emission intensity as a function of urea concentration to a two-state model.[21] Errors were $< 1\%$ and $< 0.5\%$ for ΔG° and *m*, respectively. The raw data used to calculate these values can be found in the Supporting Information.

Mass spectrometry: ESI and MALDI-TOF mass spectrometry analyses were performed at the PNACL at Washington University, St. Louis, and the Protein Purification Laboratory in the Structural Biology Center at the University of Kansas, respectively.

^{19}F NMR: NMR spectra were recorded on a Varian Unity-Plus 500 MHz spectrometer operating at 470.3 MHz, which was equipped with a Varian Cryo-Q dedicated 5 mm ^{19}F probe, as previously described.[7a,b] The data shown in Figure 3 represent 1024 transients processed with 20 Hz line broadening.

Acknowledgements

This work was supported by National Institutes of Health grant U54 AI057160 to the Midwest Regional Center of Excellence for Biodefense and Emerging Infectious Diseases Research (J.G.B), as well as the National Science Foundation Research Sites for Educators in Chemistry (NSF-RSEC) program, which supported J.F.E's postdoctoral studies. We also gratefully acknowledge Professor Carl Frieden for his generosity in providing access to the ^{19}F NMR cryoprobe at Washington University, St. Louis.

Keywords: amino acids · biosynthesis · fluorine · histidine · protein modifications

[1] Z. Fisher, J. A. Hernandez Prada, C. Tu, D. Duda, C. Yoshioka, H. An, L. Govindasamy, D. N. Silverman, R. McKenna, *Biochemistry* **2005**, *44*, 1097 – 1105.
[2] a) M. Hernick, C. A. Fierke, *Arch. Biochem. Biophys.* **2005**, *433*, 71 – 84; b) D. R. Brown, H. Kozlowski, *Dalton Trans.* **2004**, 1907 – 1917; c) S. Abuskhuna, M. McCann, J. Briody, M. Devereux, V. McKee, *Polyhedron* **2004**, *23*, 1731 – 1737.
[3] a) A. B. Patel, E. Crocker, P. J. Reeves, E. V. Getmanova, M. Eilers, H. G. Khorana, S. O. Smith, *J. Mol. Biol.* **2005**, *347*, 803 – 812; b) A. Toyama, Y. Takahashi, H. Taeuchi, *J. Am. Chem. Soc.* **2004**, *126*, 753 – 765; c) A. Toyama, Y. Takahashi, H. Taeuchi, *Biochemistry* **2004**, *43*, 4670 – 4679.
[4] a) R. O. Blaustein, T. M. Koehler, R. J. Collier, A. Finkelstein, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2209 – 2213; b) J. Miller, J. L. Elliott, R. J. Collier, *Biochemistry* **1999**, *38*, 10432 – 10441; c) E. Santelli, L. A. Bankston, S. H. Leppla, R. C. Liddington, *Nature* **2004**, *430*, 905 – 908; d) J. G. Bann, S. J. Hultgren, *Nature* **2004**, *430*, 843 – 844.
[5] a) N. Budisa, *Angew. Chem.* **2004**, *116*, 6586 – 6624; *Angew. Chem. Int. Ed.* **2004**, *43*, 6426 – 6463; b) L. Wang, P. G. Schultz, *Angew. Chem.* **2004**, *116*, 34 – 68; *Angew. Chem. Int. Ed.* **2004**, *44*, 34 – 66; c) A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603 – 609.
[6] P. P. Pal, J. H. Bae, M. K. Azim, P. Hess, R. Friedrich, R. Huber, L. Moroder, N. Budisa, *Biochemistry* **2005**, *44*, 3663 – 3672.
[7] a) J. G. Bann, J. Pinkner, S. J. Hultgren, C. Frieden, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 709 – 714; b) J. G. Bann, C. Frieden, *Biochemistry* **2004**, *43*, 13775 – 13786; c) Q. Shu, C. Frieden, *J. Mol. Biol.* **2005**, *345*, 599 – 610; d) M. H. J. Seifert, D. Ksiazek, M. K. Azim, P. Smialowski, N. Budisa, T. A. Holak, *J. Am. Chem. Soc.* **2002**, *124*, 7932 – 7942.
[8] a) K. L. Kirk, *Biochemistry of Haologenated Organic Compounds*, **1991**, Plenum, New York; b) D. C. Klein, J. L. Weller, K. L. Kirk, R. W. Hartley, *Mol. Pharmacol.* **1977**, *13*, 1105 – 1110; c) C. R. Creveling, W. L. Padgett, E. T. McNeal, L. A. Cohen, K. L. Kirk, *Life Sci.* **1992**, *51*, 1197 – 1204; d) B. M. Dunn, C. DiBello, K. L. Kirk, L. A. Cohen, I. M. Chaiken, *J. Biol. Chem.* **1974**, *249*, 6295 – 6301.
[9] A. Sutherland, C. L. Willis, *Nat. Prod. Rep.* **2000**, *17*, 621 – 631.
[10] a) K. L. Kirk, L. A. Cohen, *J. Am. Chem. Soc.* **1971**, *93*, 3060 – 3061; b) K. L. Kirk, L. A. Cohen, *J. Am. Chem. Soc.* **1973**, *95*, 4619 – 4623; c) K. L. Kirk, W. Nagai, L. A. Cohen, *J. Am. Chem. Soc.* **1973**, *95*, 8389 – 8392.
[11] H. J. C. Yeh, K. L. Kirk, L. A. Cohen, *J. Chem. Soc. Perkin Trans. 2* **1975**, 928 – 934.
[12] D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243 – 247.
[13] S. Nagshineh, L. A. Cohen, K. L. Kirk, unpublished work.
[14] Y. Ikeda, S. Kawahara, M. Taki, A. Kuno, T. Hasegawa, K. Taira, *Protein Eng.* **2003**, *16*, 699 – 706.
[15] C. Frieden, S. D. Hoeltzli, J. G. Bann, *Methods Enzymol.* **2004**, *380*, 400 – 415.
[16] D. L. Hung, S. D. Knight, R. M. Woods, J. S. Pinkner, S. J. Hultgren, *EMBO J.* **1996**, *15*, 3792 – 3805.
[17] For more detailed information about the His auxotroph, UTH780, consult the following references: a) E. P. Goldschmidt, M. S. Cater, T. S. Matney, M. A. Butler, A. Greene, *Genetics* **1970**, *66*, 219 – 229; b) L. Garrick-Silversmith, P. E. Hartman, *Genetics* **1970**, *66*, 231 – 244.
[18] Y. Tang, D. A. Tirrell, *Biochemistry* **2002**, *41*, 10635 – 10645.
[19] E. De Clerq, A. Billiau, V. G. Eddy, K. L. Kirk, L. A. Cohen, *Biochem. Biophys. Res. Commun.* **1978**, *82*, 840 – 846.
[20] In this new approach, which will be published separately, diastereoselective alkylation of a key 4-bromomethyl-5-fluoroimidazole precursor obviated enzymatic resolution, as was performed in the earlier procedure.
[21] C. N. Pace, J. M. Scholtz, *Protein Structure: A Practical Approach* (Ed.: T. E. Creighton), Oxford University Press, Oxford, **1997**, pp. 299 – 321.

Received: June 10, 2005
Published online on October 31, 2005

