Difluoromethionine as a Novel ^{19}F NMR Structural Probe for Internal Amino Acid Packing in Proteins

Mark D.Vaughan,[†] Paul Cleve,[†] Valerie Robinson,[‡] Henry S. Duewel,[†] and John F. Honek*,[†]

Contribution from the Department of Chemistry, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, and Department of Chemistry and Biochemistry, University of Guelph, Guelph, Ontario, Canada N1G 2W1

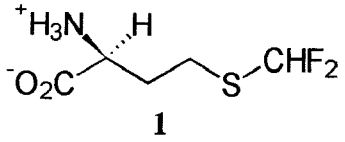
Received April 12, 1999. Revised Manuscript Received July 13, 1999

Abstract: The successful incorporation of difluoromethionine (DFM), a novel ^{19}F NMR probe of internal amino acid packing, into the three methionine positions (1, 14, and 107) of a recombinant protein, the lysozyme from bacteriophage λ (LaL), is reported. The anisochronous ^{19}F NMR signals of the diastereotopic fluorines showed a variation in the degree of chemical shift difference when present at relatively free surface positions (Met1 and Met107) versus the tightly packed protein core (Met14), with the anisochronicity greatly enhanced for DFM incorporated at this latter position. The increased magnetic nonequivalence of the two fluorines at position 14 is thought to be a consequence of the restricted environment of DFM at this position. The anisochronicity of these two fluorines is further manifested in a differential chemical shift change for these two fluorines upon binding of an oligosaccharide inhibitor to LaL, with one of the two fluorines experiencing a significant upfield shift compared to the other. This differential variation is thought to be associated with a very subtle change in the protein conformation surrounding one fluorine at position 14, which is not significantly translated to the environment of the other fluorine.

Introduction

The small steric size of fluorine has allowed its use as a replacement for hydrogen in many biologically relevant molecules.[1] In addition, its ^{19}F NMR properties render it particularly suitable for protein biophysical studies. Bioincorporation of fluoroamino acids into recombinant proteins has been useful in probing protein structure, function, and dynamics, particularly for those proteins which are either membrane bound or whose symmetry or molecular size hinders structural investigations by conventional NMR analysis.[2] Fluorinated derivatives of the aromatic amino acids such as 4-, 5-, and 6-F−Trp, 3- and 4-F−Phe, and 3-F−Tyr have been utilized predominantly in the literature. Currently, much interest is being focused on their biophysical use and structural effects on proteins containing these analogues.[3] To date, however, the application and effects of fluorinated aliphatic amino acids on protein structure and function, and their utility in protein biophysical studies, has not been as extensively investigated. Recent published reports on ^{19}F NMR investigations of proteins containing monofluoroleucine,[4] as well as our own work on trifluoromethionine (TFM),[5] have attempted to address these issues. On the basis of the successes of these previous studies, we felt that a fluoroamino acid which could serve as an ^{19}F NMR probe of rotational freedom and internal packing in a protein would be a powerful addition to the repertoire of biophysical probes which make use of the ease and simplicity of ^{19}F NMR approaches.

It is well-known that diastereotopic nuclei such as geminal hydrogen or fluorine atoms in a chiral molecule can not only exhibit different (anisochronous) NMR chemical shifts but also that this chemical shift difference can vary depending on the rotational mobility of the functional group.[6] We felt that the diastereotopic nature of the two fluorine atoms in L-difluoromethionine (L-*S*-(difluoromethyl)homocysteine; DFM, **1**)[7] could be



1

exploited in the investigation of protein structure and function. Although methyl groups in proteins have been shown to have little rotational restriction,[8] the presence of the two fluorine atoms introduces a slightly larger size and shape to the methyl group.[9] This altered size could be useful in exploring the three-dimensional space surrounding the methionyl methyl group,

* To whom correspondence should be addressed. Telephone: (519) 888-4567, ext. 5817. Fax: (519) 746-0435. E-mail: jhonek@uwaterloo.ca.

[†] University of Waterloo.

[‡] University of Guelph.

(1) (a) Welch, J. T.; Eswarakrishnan, S. *Fluorine in Bioorganic Chemistry*; John Wiley and Sons: New York, 1991. (b) O'Hagan, D.; Rzepa, H. S. *J. Chem. Soc., Chem. Comm.* **1997**, 645−652.

(2) (a) Gerig, J. T. *Prog. NMR Spectrosc.* **1994**, *26*, 293−370. (b) Danielson, M. A.; Falke, J. J. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 163−195.

(3) (a) Hoeltzli, S. D.; Frieden, C. *Biochemistry* **1998**, *37*, 387−398. (b) Thorson, J. S.; Injae, S.; Chapman, E.; Stenberg, G.; Mannervik, B.; Schultz, P. G. *J. Am. Chem. Soc.* **1998**, *120*, 451−452. (c) Andersen, O. S.; Greathouse, D. V.; Providence, L. L.; Becker, M. D.; Koeppe, R. E. *J. Am. Chem. Soc.* **1998**, *120*, 5142−5146. (d) Peteranderl, R.; Rabenstein, M.; Shin, Y.-K.; Liu, C. W.; Wemmer, D. E.; King, D. S.; Nelson, H. C. M. *Biochemistry* **1999**, *38*, 3559−3569.

(4) Feeny, J.; McCormick, 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.

(5) Duewel, H. S.; Daub, E.; Robinson, V.; Honek, J. F. *Biochemistry* **1997**, *36*, 3404−3416. Presented in part at the 78th CSC Conference, Guelph, Ontario, Canada, May 1995.

(6) Eliel, E. L.; Wilen, S. H. *Stereochemistry of Organic Compounds*; John Wiley and Sons, Inc.: New York, 1994; pp 488−507.

(7) (a) Houston, M. E.; Honek, J. F. *J. Chem. Soc., Chem. Commun.* **1989**, 761−762. (b) Tsushima, T.; Ishihara, S.; Fujita, Y. *Tetrahedron Lett.* **1990**, *31*, 3017−3018.

10.1021/ja9911418 CCC: $18.00
Published on Web 08/25/1999

which is difficult to achieve using any of the 19 other standard amino acids or currently available unnatural analogues. It was thought that rotational hindrance of a difluoromethyl group in a protein would be evident in the ^{19}F NMR spectrum, giving a characteristic splitting pattern and providing information about the density of packing in the surrounding region. This effect was indeed observed on investigation of a recombinant protein, the lysozyme from bacteriophage λ (LaL), labeled with DFM. We report herein on the dramatic differences in ^{19}F resonances of the bioincorporated DFM residues as a function of the environment within this model protein.

Experimental Section

Materials. L-DFM was synthesized as previously described.[7] Hexa-*N*-acetylchitohexaose (fine grade, purity >95%) was purchased from Seikagaku America, Inc. (Rockville, MD). The bacterial strain *Escherichia coli* B834(DE3) was obtained from Novagen (Madison, WI). All other chemicals were reagent grade and were used without further purification.

Incorporation of L-DFM into LaL. The *E. coli* methionine auxotroph B834(DE3)[5] containing the LaL-encoding plasmid pLR102 was grown at 37 °C in M9 minimal medium supplemented with 0.1 mM L-methionine, 0.4% glucose, and 40 mg/L ampicillin. At the midexponential phase of cell growth ($OD_{600} = 0.6-0.7$), the cells were harvested by centrifugation (10 000*g*), washed with methionine-free M9 medium to remove any residual L-Met, and resuspended in half the original culture volume of methionine-free M9 medium containing 2 mM L-DFM. Protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside to 0.75 mM, and the cells were collected after 8 h. The harvested cells were resuspended in 25 mL 50 mM potassium phosphate (pH 7) and disrupted by sonication. The labeled protein was purified by methods previously described[10] utilizing successive chromatography over S-Sepharose Fast Flow, Mono-S, and Phenyl Superose columns (Pharmacia-LKB Biotechnology AB, Uppsala, Sweden). Following dialysis against 5 mM potassium phosphate (pH 7), the protein was lyophilized and stored at −20 °C.

Mass Spectrometry. Electrospray (ESMS) mass spectra were obtained using a Fisons VG Quatro II triple-quadrupole mass spectrometer fitted with an electrospray ionization source with a delivery solvent of 1:1 H_2O/CH_3CN containing 0.1% TFA. Data were analyzed using Masslynx software, which employs the MaxEnt algorithm to transform the multiply charged electrospray spectrum to the true mass scale.

NMR Spectroscopy. NMR data were collected at 376.3 MHz using a Bruker Avance spectrometer fitted with a 5 mm inverse broad-band probehead with the proton coil tuned to fluorine. Standard undecoupled parameters were 18 832 Hz sweep width, 0.877 s acquisition time, and a 1.0 s relaxation delay. A 3 Hz line broadening was applied. Unless otherwise stated, all spectra were recorded at 22 ± 0.5 °C. Spectra were referenced to an external sample of $CFCl_3$ (set at 0.00 ppm). Protein samples were prepared with concentrations ranging from 2.5 to 6.0 mg/mL in 600 μL D_2O. GdEDTA$^-$ line broadening was achieved by adding 8 μL of a stock solution of GdEDTA$^-$ (80 mM $GdCl_3$, 500 mM EDTA, pH 7.1 with NaOH)[11] to the sample. Oligosaccharide binding was investigated by adding 5.2 mg hexa-*N*-acetylchitohexaose to the sample. In both cases, it was assumed that the additions did not significantly change the concentration of the protein sample. Variable temperature experiments on L-DFM itself were performed using a sample concentration of 22 mM (3.2 mg in 800 μL CD_3OD). The DFM−LaL two-dimensional ^{19}F−^{19}F COSY spectrum was accumulated in absolute value mode (512 increment points in t_2 for each of 64 t_1 values, 1000 scans per increment). A 1.0 s relaxation delay was used. Zero filling in the F1 dimension produced a 512 × 512 data matrix. A sine bell squared window function was applied in both dimensions.

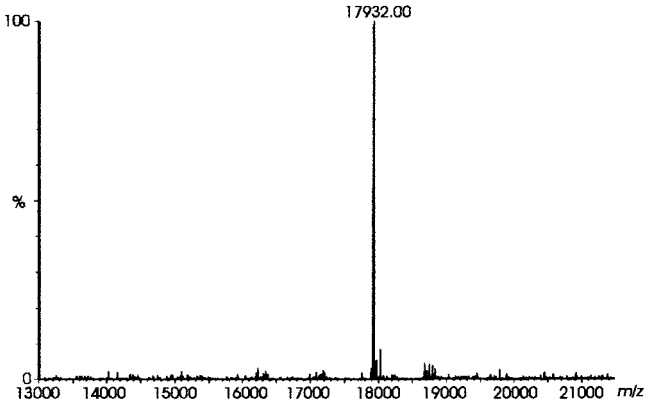


Figure 1. Electrospray mass spectrum of DFM−LaL, transformed to the true mass scale using the MaxEnt algorithm. The peak at 17 932.0 Da represents LaL in which all three Met residues have been replaced by DFM, and the absence of any other significant peaks indicates that the incorporation level is essentially 100%.

Results and Discussion

In comparison to the lower protein yields and incomplete levels of TFM incorporation obtained when LaL was expressed in the presence of TFM (1−2 mg/L of culture at 74% incorporation and 15 mg/L at 30% incorporation), utilization of DFM, containing one less fluorine, increased the yields of fluorinated protein to close to wild type, nonfluorinated levels (20−25 mg/L).[12] In addition, DFM did not exhibit the high level of cytotoxicity associated with TFM, and the methionine auxotrophic *E. coli* host could produce DFM−LaL with DFM as the sole methionine supplement during the induction phase. This afforded a much higher level of analogue incorporation. ESMS indicated the incorporation level to be essentially 100%, with all three methionine positions (Met 1, 14 and 107) being completely labeled (Figure 1). DFM−LaL retains complete catalytic activity relative to the nonfluorinated protein,[10] suggesting that there are no large scale conformational changes associated with analogue incorporation. The two fluorine atoms in DFM itself are diasterotopically related, and are therefore not chemically equivalent. Consequently, the free amino acid exhibits the expected ABX spectrum for the CHF_2 group, with a proton−fluorine coupling of $J_{H,F} = 54$ Hz and a fluorine−fluorine coupling of $J_{F,F} = 196$ Hz in D_2O and chemical shifts centered at −92.00 ppm and −92.02 ppm for the two diasterotopic fluorines (Figure 2A). Purified and homogeneous DFM−LaL exhibited a ^{19}F NMR spectrum (Figure 2B) containing two sets of apparent doublets (centered at approximately −91.2 ppm and −92.5 ppm; actually ABX patterns which are broadened due to protein tumbling motion) each integrating to one DFM residue, and two sets of quartets (centered at approximately −94.2 and −95.5 ppm; also an ABX pattern), which together integrated to the third DFM. Site-directed mutagenesis was used to assign the DFM107 resonance in an approach similar to that described by Jarema et al. for the assignment of the ^{19}F signals of 3-fluorotyrosine residues incorporated into the *lac* repressor.[13] The Met107Leu mutant enzyme[14] was overproduced in the

(8) (a) Jones, W. C.; Rothgeb, T. M.; Gurd, F. R. N. *J. Biol. Chem.* **1976**, *251*, 7452−7460. (b) Wittebort, R. J.; Rothgeb, T. M.; Szabo, A.; Gurd, F. R. N. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1059−1063. (c) Nicholson, L. K.; Kay, L. E.; Baldisseri, D. M.; Arango, J.; Young, P. E.; Bax, A.; Torchia, D. A. *Biochemistry* **1992**, *31*, 5253−5263. (d) Lee, A. L.; Flynn, P. F.; Wand, J. *J. Am. Chem. Soc.* **1999**, *121*, 2891−2902.

(9) Houston, M. E.; Harvath, L.; Honek, J. F. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 3007−3012.

(10) Duewel, H. S.; Daub, E.; Honek, J. F. *Biochim. Biophys. Acta* **1995**, *1247*, 149−158.

(11) Luck, L.; Falke, J. J. *Biochemistry* **1991**, *30*, 6484−6490.

(12) Although no detailed accounts of DFM in proteins have appeared, a successful incorporation of DFM into calmodulin has been reported: McIntyre, D. D.; Yuan, T.; Vogel, H. J. *Prog. Biophys. Mol. Biol.* **1996**, *65*, P-A1−27. Abstracts of the XIIth International Biophysics Congress, 11−16 August, 1996, Amsterdam, The Netherlands.

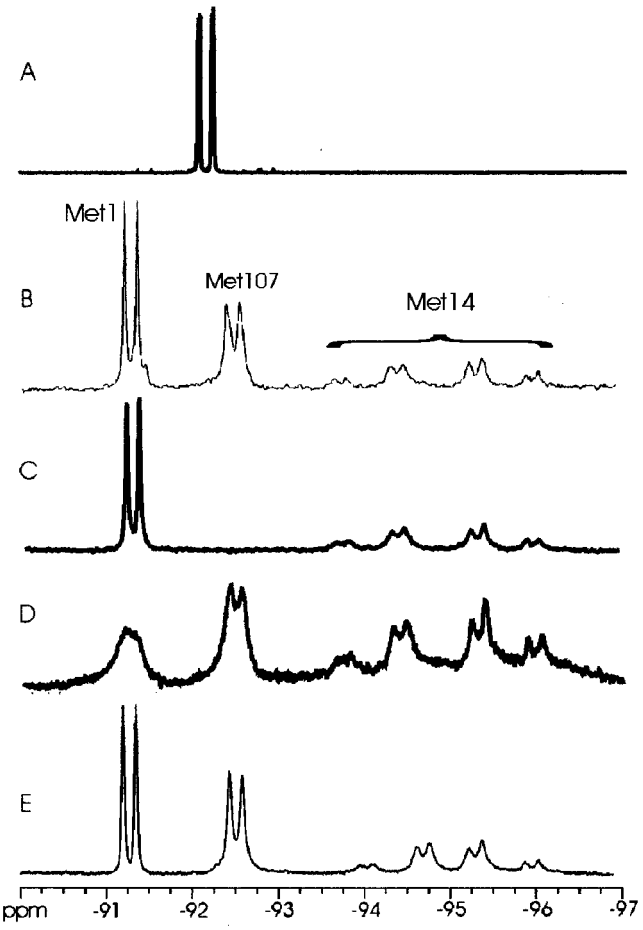


Figure 2. The 376.3 MHz ^{19}F NMR spectra of various DFM species: (A) free DFM, illustrating the ABX splitting pattern; (B) completely labeled DFM−LaL, indicating the identities of the three spin systems; (C) DFM−LaL Met107Leu mutant, which lacks the Met107 signal; (D) DFM−LaL in the presence of 1 mM $GdEDTA^-$; (E) DFM−LaL in the presence of 7 mM hexa-*N*-acetyl chitohexaose.

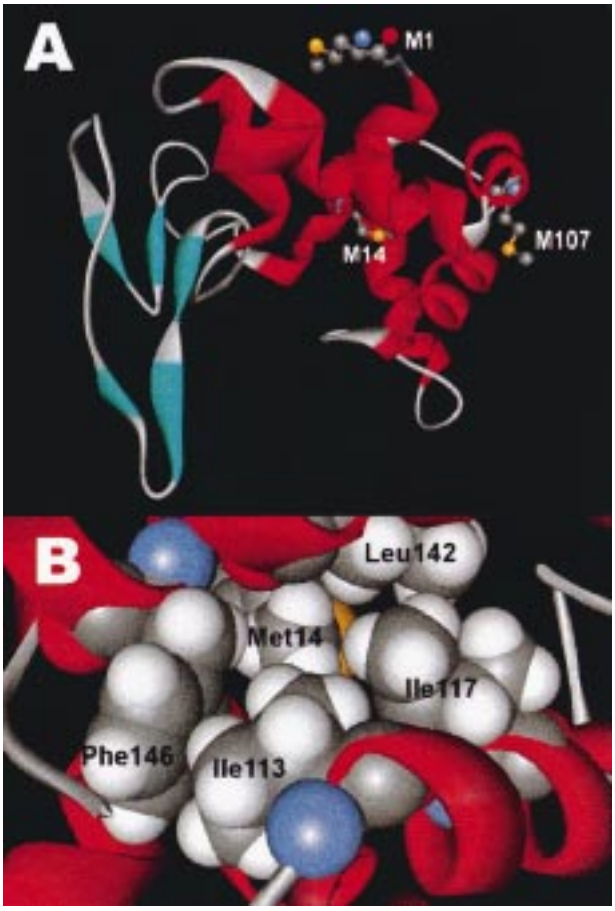


Figure 3. (A) X-ray crystal structure of LaL, showing the positions of the three Met residues. (B) Space-filling representation of the environment surrounding the Met14 methyl group, illustrating the dense packing in the vicinity of this group. Graphics generated using WebLab ViewerPro, version 3.0 (1997, Molecular Simulations Ltd.)

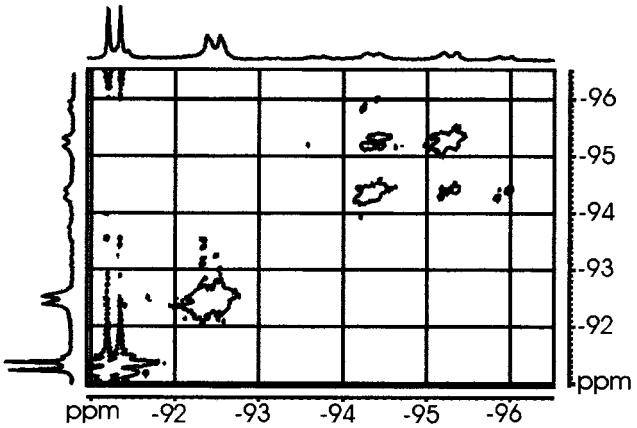


Figure 4. Two-dimensional $^{19}F-^{19}F$ COSY spectrum of DFM−LaL.

presence of DFM, and ^{19}F NMR indicated the disappearance of the resonance centered at −92.5 ppm (Figure 2C), suggesting that this signal arises from incorporation of DFM at position 107. The remaining assignments were readily made utilizing the known paramagnetic NMR line broadening effects on surface/exposed residues by Gd(III) complexed to EDTA[11] and the recently published crystal structure of the λ lysozyme.[15] The X-ray structure indicates that Met1 and Met107 are solvent exposed, whereas Met14 is found to reside in the protein core (Figure 3A). Addition of $GdEDTA^-$ resulted in line broadening of the ^{19}F NMR resonances centered at −91.2 and −92.5 ppm (Figure 2D). Given that the ^{19}F resonance at −92.5 ppm was assigned to DFM107, the signal at −91.2 ppm was assigned to DFM1. By default, the two quartets at −94.2 and −95.5 ppm were assigned to DFM14.

The hypothesis that the two upfield quartets both arise from DFM14 was confirmed by two-dimensional $^{19}F-^{19}F$ COSY experiments (Figure 4). This suggests that the chemical shift nonequivalency of these two fluorines is enhanced in the protein interior. The chemical shift difference between the geminal fluorines in the amino acid DFM does not vary extensively under a range of solvent conditions.[16] We propose that the enhancement of the chemical shift difference between the fluorines originating from DFM14 is attributable to restrictions in the rotations about the DFM side chain bonds at this position. Investigation of the ^{19}F signals of the amino acid DFM at various temperatures in deuterated methanol (Figure 5) shows an increase in the chemical shift difference for the two fluorines

(13) Jarema, M. A. C.; Lu, P.; Miller, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2707−2711.

(14) Polymerase chain reaction mutagenesis was utilized to prepare a gene encoding the Met107Leu site-directed mutant of LaL (experimental methods to be discussed in a forthcoming publication).

(15) Evrard, C.; Fastrez, J.; Declerq, J.-P. *J. Mol. Biol.* **1998**, *276*, 151−164. Note that the structure solved is that of LaL containing 7-azatryptophan at native Trp positions, although this is not thought to significantly perturb the structure.

(16) D_2O, methanol-d_4, DMSO-d_6, benzyl alcohol, and D_2O containing high concentrations of L-phenylalanine.

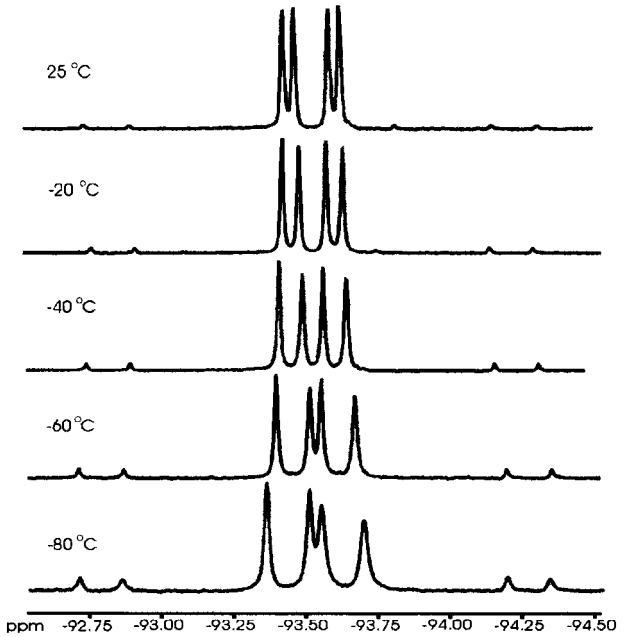


Figure 5. ^{19}F NMR spectra of free DFM in CD_3OD at various temperatures.

as the temperature is decreased and the rates of rotation about the S−CHF_2 and other bonds are correspondingly lowered. It is therefore not unlikely that an analogous phenomenon occurs within the tightly packed protein interior. Chemical shift differences between the two fluorines at either of the solvent-exposed positions (1 and 107) are not as prominent, presumably due to less restricted environments for the DFM residues at these sites.

Although we currently do not have information concerning the degree of rotational freedom for the nonfluorinated Met14 methyl group, previous studies have indicated that core aliphatic side chains can have considerable conformational freedom, including methyl group rotation.[8] However, it is possible that the chemical shift nonequivalency observed for DFM14 may reflect an intrinsic rotational hindrance for the nonfluorinated methyl group. As shown in Figure 3B, the Met14 methyl group exists in a region of dense side chain packing which may decrease its rate of rotation. Alternatively, it is quite possible that the fluorination of the methyl group contributes to the observed rotational restriction. While the van der Waals' radius of fluorine is only slightly larger than that of hydrogen (1.47 Å for F vs 1.20 Å for H), the volume of the methyl group increases significantly upon fluorination, and the size of the trifluoromethyl group itself has been estimated to be close to that of an isopropyl group.[1b] We have previously performed ab initio calculations at the RHF/6-31G* level which give van der Waals' volumes of 71.2, 80.9, and 84.4 $Å^3$ for CH_3SCH_3, CH_3SCHF_2, and CH_3SCF_3, respectively.[9] Clearly the increased steric bulk of the difluoromethyl group relative to the nonfluorinated methyl group could lead to unfavorable interactions with surrounding amino acid side chains. It has been shown that the incorporation of 3-fluorotyrosine into glutathione *S*-transferase causes slight structural perturbations,[17] despite the very small increase in volume upon fluorination of an aromatic ring (∼0.7%).[2b] One would therefore expect similar consequences upon introduction of the difluoromethyl group into the protein core. This may lead to a preferred rotational conformation for this group which minimizes the disruptive effects of its incorporation.

In addition to the steric contributions to rotational restriction discussed above, there may be hydrogen bonding interactions involved as well. The ability of fluorine to act as a hydrogen bond acceptor has been the subject of ongoing debate, although its potential in this regard has been estimated to be weak at best.[18] It has also been suggested that the hydrogen atom of the difluoromethyl group may be a weak hydrogen bond donor whose hydrogen bonding strength is approximately 1 kcal/mol.[19] Some contribution of the difluoromethyl group to hydrogen bonding interactions with proximal main chain or side chain groups, although weak, could contribute to rotational constraints for DFM14 in the protein core.[20] Currently, we have no evidence supporting the involvement of DFM14 in hydrogen bonding; however, we are working toward elucidating the crystal structure of DFM−LaL which we hope will provide some insight into this matter.

The reduced rotational freedom of DFM in a tightly packed protein environment could lead to differential changes in chemical shift for each fluorine when the protein undergoes a conformational change or ligand binding event. We have previously shown that chitooligosaccharides can serve as inhibitors of λ lysozyme[10] and that there is a ^{19}F chemical shift change upon oligosaccharide binding to TFM−LaL.[5] Interestingly, on interaction of the oligosaccharide hexa-*N*-acetylchitohexaose with DFM−LaL, the signal of only one of the two fluorine atoms at position 14 experiences a significant upfield shift, while there is little visible effect on the chemical shift of any other resonance (Figure 2E). This may indicate that any structural effect associated with oligosaccharide binding to DFM−LaL does not globally affect the DFM14 residue itself but rather influences, in a subtle fashion, the predominant environment proximal to one of the two fluorines. A significant shift of this ^{19}F signal upon a small conformational change is not unlikely given the high paramagnetic term in the ^{19}F shielding parameter, which renders ^{19}F NMR acutely sensitive to changes in the environment of the fluorine atom.[2]

Conclusions

As we have shown, DFM can be utilized to probe amino acid packing in protein interiors, but it could potentially be used in more specific applications in the study of methionine and its implications in protein structure and function.[21] In addition, DFM may find merit as a highly sensitive probe to detect ligand binding, and the extent of resonance shift may give additional information concerning the nature and degree of protein structural change occurring upon ligand binding. The ease and high levels of incorporation of DFM into recombinant proteins should ensure its widespread applicability to problems in structural biochemistry.

Acknowledgment. We thank Dr. Elisabeth Daub for helpful discussions and generation of the Met107Leu LaL mutant plasmid, Dr. Gilles Lajoie for access to ESMS facilities, Dr. Michael Houston and Mr. Roy Park for synthesis of DFM, and NSERC (Canada) for financial support.

JA9911418

(17) Xiao, G.; Parsons, J. F.; Tesh, K.; Armstrong, R. N.; Gilliland, G. L. *J. Mol. Biol.* **1998**, *281*, 323−339.

(18) (a) Dunitz, J. D.; Taylor, R. *Chem.−Eur. J.* **1997**, *3*, 89−98. (b) Barbarich, T. J.; Rithner, C. D.; Miller, S. M.; Anderson, O. P.; Strauss, S. *J. Am. Chem. Soc.* **1999**, *121*, 4280−4281.

(19) Erickson, J. A.; McLoughlin, J. I. *J. Org. Chem.* **1995**, *60*, 1626−1631.

(20) The oxygen atom from the peptide bond between Leu142 and Ile143 is in close contact with the methyl group of Met14 in the crystal structure of the protein.

(21) Gellman, S. H. *Biochemistry* **1991**, *30*, 6633−6636.