Elucidation of Solvent Exposure, Side-Chain Reactivity, and Steric Demands of the Trifluoromethionine Residue in a Recombinant Protein†

Henry S. Duewel,‡ Elisabeth Daub,‡ Valerie Robinson,§ and John F. Honek*,‡

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 July 2, 2001

ABSTRACT: When incorporated into proteins, fluorinated amino acids have been utilized as ^{19}F NMR probes of protein structure and protein−ligand interactions, and as subtle structural replacements for their parent amino acids which is not possible using the standard 20-amino acid repertoire. Recent investigations have shown the ability of various fluorinated methionines, such as difluoromethionine (DFM) and trifluoromethionine (TFM), to be bioincorporated into recombinant proteins and to be extremely useful as ^{19}F NMR biophysical probes. Interestingly, in the case of the bacteriophage lambda lysozyme (LaL) which contains only three Met residues (at positions 1, 14, and 107), four ^{19}F NMR resonances are observed when TFM is incorporated into LaL. To elucidate the underlying structural reasons for this anomalous observation and to more fully explore the effect of TFM on protein structure, site-directed mutagenesis was used to assign each ^{19}F NMR resonance. Incorporation of TFM into the M14L mutant resulted in the collapse of the two ^{19}F resonances associated with TFM at position 107 into a single resonance, suggesting that when position 14 in wild-type protein contains TFM, a subtle but different environment exists for the methionine at position 107. In addition, ^{19}F and [^{1}H−^{13}C]-HMQC NMR experiments have been utilized with paramagnetic line broadening and K_2PtCl_4 reactivity experiments to obtain information about the probable spatial position of each Met in the protein. These results are compared with the recently determined crystal structure of LaL and allow for a more detailed structural explanation for the effect of fluorination on protein structure.

Bioincorporation of fluorinated amino acids into proteins and their spectroscopic study by ^{19}F NMR[1] is a powerful method for investigation of protein structure and dynamics, especially for membrane proteins and those proteins of high symmetry or large molecular size (*1*−*9*) and for study of redox-dependent conformational alterations in metalloproteins (*10*). This approach has recently been extended to the study of [5-F]Trp-labeled *Escherichia coli* F1-ATPase (M_r = 380 kDa) (*11*).

The chemical shifts of ^{19}F nuclei are extremely sensitive to their environment. In some cases, the ^{19}F NMR of incorporated [5-F]Trp can detect subtle conformational changes that are not evident from simple Trp fluorescence changes (*12*). ^{19}F NMR has been successfully employed in studying protein folding and unfolding events (*13*−*16*). The lack of endogenous fluorine compounds in cells allows for in vivo application of ^{19}F NMR without encumbrance from background ^{19}F signals (*17*−*20*). The ability to incorporate fluorinated amino acids into proteins in a site-selective manner in vitro (*21*) or in vivo (*22*) extends this technique to new problems in biochemistry.

Few studies, however, have explored the application of nonaromatic fluorine-containing amino acids as ^{19}F NMR probes for protein structural studies. We recently reported the bioincorporation of difluoromethionine (DFM) and trifluoromethionine (TFM) into a bacteriophage lysozyme (LaL) and found that these analogues do not alter enzyme activity and are useful ^{19}F NMR probes (*23*, *24*). The levels of incorporation of DFM into protein are very high (>95%) (*24*). Incorporation levels for the larger TFM analogue have ranged from 30 to 70% possibly due to the lower affinity of the *E. coli* Met-tRNA synthetase for TFM (M. D. Vaughan and J. F. Honek, unpublished results). The levels of incorporation of other fluorinated amino acids into proteins have been reported to range from 5 to 95%, depending on the protein being investigated and the amino acid analogue being used (*3*, *4*). The incorporation of TFM into LaL results in production of an ensemble of lysozymes which randomly incorporate TFM or endogenous methionine to varying extents at the three methionine positions (Met1, Met14, and Met107) (*23*). Interestingly, although the three expected distinct ^{13}C resonances were detected by HMQC experiments

† This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) (J.F.H.). H.S.D. was supported by a postgraduate scholarship from NSERC.

* To whom correspondence should be addressed: Department of Chemistry, University of Waterloo, 200 University Ave., Waterloo, Ontario, Canada N2L 3G1. Telephone: (519) 888-4567, ext. 5817. Fax: (519) 746-0435. E-mail: jhonek@uwaterloo.ca.

‡ University of Waterloo.

§ University of Guelph.

[1] Abbreviations: DFM, difluoromethionine; EDTA, ethylenediaminetetraacetic acid; ESMS, electrospray mass spectrometry; Gd·EDTA, gadolinium(III) ethylenediaminetetraacetic acid complex; HMQC, heteronuclear multiple-quantum correlation; IPTG, isopropyl 1-thio-β-D-galactopyranoside; LaL, bacteriophage lambda lysozyme; LB, Luria-Bertani broth; MQW, Milli-Q water; M_r, average mass; NMR, nuclear magnetic resonance; TFM, trifluoromethionine; wt, wild type.

10.1021/bi011381b CCC: $20.00
Published on Web 10/09/2001

in the protein in which [*methyl*-^{13}C]methionine had been incorporated (*25*), four ^{19}F NMR signals were detected for the TFM-labeled lysozymes (*23*). Of these four ^{19}F signals, two resonances (resonances A and B) each integrated to one TFM while the remaining two resonances (C and D) integrated *together* to one TFM. The relative integrated areas of resonances C and D were found to vary depending upon the extent of TFM incorporation into LaL (which depended upon the relative concentration of TFM compared to that of Met added to the growth medium during induction). This suggested that one of the TFM residues in the protein was exhibiting two different chemical environments depending on the presence of either TFM or Met at one of the other two positions in the protein.

To elucidate the underlying factor(s) responsible for the resulting ^{19}F spectra produced by this new ^{19}F NMR probe at the molecular level, which is quintessential to its further application in biochemistry, the complete ^{19}F NMR resonance assignments of TFM-labeled LaL were made. Experiments were also undertaken to probe the environment of each methionine position utilizing paramagnetic line broadening and platination reactivity studies. This report, in combination with the recently reported crystal structure of a modified form of LaL (*26*), and our recently published structure of wild-type LaL (wt LaL) with an oligosaccharide inhibitor bound (*27*), clearly presents structural information about the probable cause of the double resonance. These findings suggest further application of fluorinated methionines as subtle biophysical probes and will be useful in attempts to understand the effects of protein structure on ^{19}F resonances, a topic of much current interest (*28*−*31*).

MATERIALS AND METHODS

Construction of Met14Leu and Met107Leu Mutations in LaL. Mutations were constructed in pLR102 (*23*), using the QuikChange site-directed mutagenesis method of Stratagene Cloning Systems (La Jolla, CA) and using *Pwo* DNA Polymerase (Roche Diagnostics), *Dpn*I (New England Biolabs, Beverly, MA), and *E. coli* DH5α cells for transformation. The primers for the mutagenesis were as follows: 5′-GTAAGGCGTTCCTCGAT***CTT***CTGGCGTGGTCGGAGGG-3′ and 5′-CCCTCCGACCACGCCAG***AAG***ATCGAGGAACGCCTTAC-3′ for Met14Leu (M14L) and 5′-GGAGCGTGGCGCTTTACCT***CTG***ATTGATCGTGGTGATATC-3′ and 5′-GATATCACCACGATCAAT***CAG***AGGTAAAGCGCCACGCTCC-3′ for Met107Leu (M107L) (the mismatching bases are in bold italics). Plasmid DNA from the candidate DH5α colonies was prepared and transformed into *E. coli* B834 (λDE3) (Novagen, Madison, WI) for screening. The candidates were grown at 37 °C in 20 mL of Luria-Bertani (LB) broth with ampicillin (50 μg/mL) to an absorbance of 0.65 at 600 nm, and protein production was induced with IPTG (0.75 mM) for 3 h. Cells were collected and disrupted by sonication, and the mutant proteins were purified as described previously (*23*). The molecular masses of purified mutant proteins (with or without incorporated TFM) were determined by electrospray mass spectrometry as previously described (*23*) at the Biomedical Mass Spectrometry Centre, University of Waterloo. The plasmids encoding the M107L and M14L mutations were named pLR103 and pLR104, respectively. The LaL sequence of the mutant plasmids was determined by DNA sequencing to confirm the mutations. Oligonucleotide primer synthesis and DNA sequencing were services provided by the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University (Hamilton, ON).

Preparation and Isolation of Wild-Type and Mutant LaL Containing Met Analogues. Wild-type LaL was prepared as previously described (*25*). To prepare the [*methyl*-^{13}C]-methionine-labeled wild-type and mutant LaL proteins, the *E. coli* strains B834(λDE3)/pLR102 (wt), B834(λDE3)/pLR103 (M107L), and B834(λDE3)/pLR104 (M14L) were grown separately at 37 °C in 1 L of $M9_{Glu,amp}$ (M9 minimal medium supplemented with 0.4% glucose, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 50 μg/mL ampicillin) with 0.3 mM L-[*methyl*-^{13}C]Met (Aldrich, 99% ^{13}C) to an absorbance of 0.65 at 600 nm, then induced with IPTG (0.75 mM), and grown for 4.5 h. The production of the TFM-labeled wild-type and mutant proteins followed the procedure previously used to produce low-level incorporation of TFM into LaL (*23*). For the M107L or M14L mutants, the B834(λDE3)/pLR103 or B834(λDE3)/pLR104 cells were grown first in $M9_{Glu,amp}$ supplemented with 0.1 mM L-Met to an absorbance of 0.65 at 600 nm, collected by centrifugation, washed with $M9_{glu,amp}$, and collected again by centrifugation. The cells were then suspended in one-half of the original culture volume of $M9_{Glu,amp}$ now containing 1 mM L-TFM, 20 μM L-Met, and 0.75 mM IPTG and grown for 9.5 h. Following the indicated induction times, cells were collected by centrifugation, resuspended in 50 mM potassium phosphate (pH 7), and disrupted by sonication. All proteins were purified by sequential chromatography over S-Sepharose Fast Flow, Mono-S, and Phenyl Superose (Amersham Pharmacia-Biotech Inc.), then dialyzed against 5 mM potassium phosphate (pH 7), and lyophilized as reported previously (*23*).

^{19}F NMR and Gd·EDTA Experiments. ^{19}F NMR spectra were obtained at 376.3 MHz on a Varian Unity spectrometer fitted with a 5 mm dual-broadband probe head with the proton coil tuned to fluorine as previously described (*23*). Standard undecoupled spectra parameters were as follows: 27 740 Hz spectral width, 32K data points, 10 μs pulse width, 0.577 s acquisition time, 0.5 s relaxation delay, and 0.5 Hz line broadening. All spectra were recorded at 22 ± 0.5 °C. Samples were measured using a coaxial insert containing CD_3OD (as the lock solvent) and $CFCl_3$ as an internal frequency standard (referenced to 0.00 ppm). The insert was placed into a standard 5 mm NMR tube containing the sample. Treatment of TFM-labeled LaL with the Gd·EDTA complex was achieved following the method reported by Luck and Falke (*32*). Increasing amounts of a Gd·EDTA stock solution (prepared in D_2O containing 80 mM $GdCl_3$ and 500 mM EDTA brought to pH 7.1 with added solid NaOH dissolved in D_2O) were added to the NMR sample to produce the final concentrations as indicated in Figure 3. Protein samples were prepared with concentration ranges of 1.0−6.0 mg/mL in 600 μL of D_2O containing 50 mM potassium phosphate (pH 7).

[^{1}H−^{13}C]-HMQC NMR Experiments. [^{1}H−^{13}C]-HMQC NMR experiments for the M14L and M107L mutants were undertaken using conditions described previously for wt LaL containing L-[*methyl*-^{13}C]Met (*25*).

Reaction of LaL with K_2PtCl_4. For electrospray mass spectrometry (ESMS) studies, a solution of unlabeled wt LaL

(0.3 mM) in Milli-Q Water (Milli-Q RG Ultrapure water system; Waters Associates; MQW) was treated with various amounts of a freshly prepared aqueous K_2PtCl_4 solution (3 mM in MQW) to produce final concentrations of K_2PtCl_4 ranging from 0.30 to 1.05 mM. Preparations of K_2PtCl_4 solutions and subsequent incubations were undertaken at 25 °C under low-light conditions. At various time intervals, aliquots of the reactions were removed and diluted 100-fold with MQW, centrifuged (14000*g* for 10 min), and subjected immediately to ESMS analysis as described below. For HMQC NMR studies, various solutions of wt LaL containing [*methyl*-^{13}C]Met and K_2PtCl_4 were prepared as described above but in D_2O [50 mM potassium phosphate (pH 7)]. A series of HMQC spectra was recorded at various times for each sample under conditions described previously (*25*). The [^{1}H−^{13}C]-HMQC spectra were obtained on a Bruker AMX500 spectrometer and were externally calibrated with 2,2,3,3-tetradeutero-3-(trimethylsilyl)propionic acid (TSP).

Mass Spectrometry. ESMS was performed on a Micromass VG Quatro II triple-quadrupole instrument equipped with an electrospray ionization source. Instrument control and data analysis were implemented with the Windows-based MASSLYNX software. Samples (5−10 μL) were introduced into the source by injection via a Rheodyne valve into a continuously flowing 1:1 water/acetonitrile solvent (containing 0.1% formic acid) at a flow rate of 20 μL/min. Spectra were the averaged result of between 10 and 20 scans, and the multiply charged distribution profiles were subjected to the MaxEnt algorithm to produce true molecular mass spectra. Calibration of the mass spectrometer was achieved using the charge distribution profile of wt LaL (M_r = 17 825.217 Da).

Miscellaneous Methods. Protein concentrations were determined spectrophotometrically using the previously determined extinction coefficient of 31 712 M^{-1} cm^{-1} (*23*). The enzymatic activities of LaL and its mutants were measured by a turbidimetric assay using chloroform-treated *E. coli* cells (*25*).

RESULTS

Isolation and Characterization of M14L- and M107L-LaL. By using polymerase chain reaction site-directed mutagenesis and pLR102 as the template, two plasmids were constructed, pLR103 and pLR104, which contain the gene for LaL in which either the codon for Met107 or Met14, respectively, was replaced with a codon for Leu. The plasmids were transformed into *E. coli* B834(λDE3), a strain that is auxotrophic for methionine. Prior to incorporation of TFM or [*methyl*-^{13}C]Met into M107L- and M14L-LaL, the mutants were first purified from cultures of B834(λDE3)/pLR103 and B834(λDE3)/pLR104 in a methionine-enriched medium (LB) to obtain the unlabeled purified proteins for purposes of characterization. Both M14L- and M107L-LaL could be purified to homogeneity using the same procedure used to purify wt LaL (*23*). Although SDS−PAGE analysis of the supernatants (post cell disruption) indicated equal and high expression levels of the mutants compared to that of wt LaL expressed in B834(λDE3)/pLR102 grown in enriched medium, the yield of the purified protein obtained for M14L-LaL was less than that obtained for M107L-LaL and wt LaL, both of which yielded 30−40 mg of purified protein from a 1 L of culture. The calculated molecular masses of wt LaL and of the two Leu mutants are 17 825 and 17 807 Da, respectively. The expected molecular mass of 17 807 Da was observed for the two mutants from ESMS analysis, confirming the mutations at the protein level (data not shown). In addition, the LaL gene was sequenced in pLR103 and pLR104 to confirm the mutations and to ensure that no other mutations in the gene resulted from the PCR amplification. Enzymatic activities of the mutants were assessed turbidimetrically using chloroform-treated *E. coli* cells as the substrate (*25*). Using the unlabeled mutants and wt LaL as a standard, the respective activities of M14L-LaL and M107L-LaL are essentially equal to that of wt LaL.

Isolation and Characterization of TFM- and [methyl-^{13}C]-Met-Labeled M14L- and M107L-LaL. To incorporate TFM into the various mutants of LaL, the procedure to prepare LaL labeled at a low level with TFM was followed as reported previously (*23*). As such, *E. coli* B834(λDE3) cells harboring pLR103 or pLR104 were initially grown in minimal medium supplemented with 0.1 mM L-Met to establish cell growth. The cells were collected, washed, then introduced into minimal medium supplemented with 1.0 mM L-TFM and 20 μM L-Met, and subsequently induced with IPTG to express the TFM-labeled mutants. Incorporation of [*methyl*-^{13}C]Met into the mutant lysozymes was achieved by growing and inducing B834(λDE3)/pLR103 and B834(λDE3)/pLR104 cells in minimal medium in the presence of 0.3 mM L-[*methyl*-^{13}C]Met. The respective TFM-labeled or [*methyl*-^{13}C]Met-labeled mutant proteins was purified to homogeneity by the previously reported purification method (*23*).

In the case of TFM-labeled LaL, low-level incorporation conditions typically resulted in the isolation of 10−15 mg of purified wt TFM-labeled LaL per liter of culture. Although these yields were obtained for TFM-labeled M107L-LaL as well, the TFM-labeled M14L-LaL was quite unstable during purification (notably during the initial S-Sepharose chromatography), and only small quantities of this protein could be obtained. Characterization of the purified mutant TFM-labeled enzymes by ESMS indicated masses of 17 807, 17 861, and 17 915 Da which correspond to protein species in which either none, one, or both of the remaining Met positions in the respective mutants had incorporated TFM (data not shown).

Conditions used to produce [*methyl*-^{13}C]Met-labeled M14L- and M107L-LaL resulted in yields similar to those from enriched medium. The expected mass of 17 809 Da was observed for both of the [*methyl*-^{13}C]Met-labeled mutants from ESMS analysis. Fewer difficulties were encountered in purifying Met (or [*methyl*-^{13}C]Met)-labeled M14L-LaL than with TFM-labeled M14L-LaL.

^{19}F Resonance Assignments for TFM-Labeled LaL. The spectrum of LaL with a low level of incorporation of wt TFM (Figure 1) exhibits four ^{19}F NMR resonances, despite the presence of only three methionine positions in LaL (*23*). Previously, we had chosen to designate the four resonances, without knowledge of their sequence-specific assignments, as resonances A−D and had suggested from our findings that one of the three TFM residues was responsible for giving rise to both resonances C and D (*23*). The ^{19}F NMR spectrum of TFM-labeled M107L-LaL (Figure 1) indicates the pres-

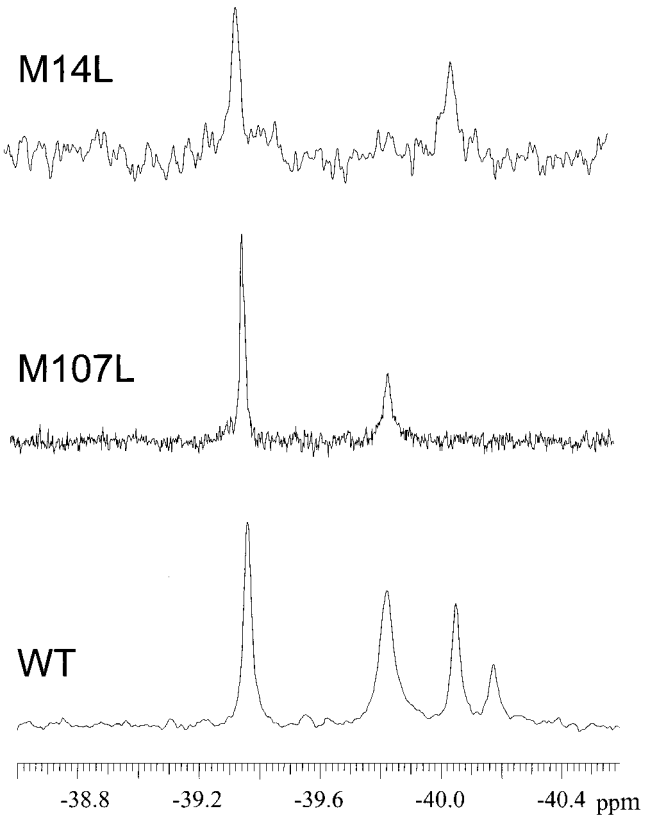


FIGURE 1: ^{19}F NMR spectrum of the TFM resonances in TFM-labeled LaL. Shown are the 376.3 MHz ^{19}F NMR spectra of M14L, M107L, and wt LaL, each labeled with TFM. Each sample was prepared in D_2O containing 50 mM potassium phosphate (pH 7.0), and the spectra are referenced to $CFCl_3$ (0.00 ppm).

ence of resonances A and B. These resonances integrated equally to each other. The corresponding TFM-labeled M14L mutant was prepared, and although its isolation resulted in low yields of purified protein, the spectrum for TFM-labeled M14L-LaL (Figure 1) was obtained and contains only resonances A and C, both with similar intensities. Thus, the assignments of the ^{19}F resonances for wt TFM-labeled LaL can be made. Resonance A at −39.36 ppm corresponds to the TFM incorporated at position Met1 (by default); resonance B at −39.88 ppm corresponds to TFM at position Met14, and resonances C and D correspond to TFM at −40.04 and −40.16 ppm, respectively, which we had previously suggested (*23*) to come from a single residue. These are indeed being produced by TFM at the Met107 position.

^{1}H and ^{13}C Resonance Assignments for [methyl-^{13}C]Met-Labeled LaL. A similar approach was utilized to obtain the sequence-specific assignment of the [*methyl*-^{13}C]Met resonances in [*methyl*-^{13}C]Met-labeled LaL, as these were to be important in defining exposed methionine residues in the platination studies. The [^{1}H−^{13}C]-HMQC spectrum of wt LaL into which [*methyl*-^{13}C]Met was incorporated (Figure 2) exhibits the three expected resonances for the three Met residues in the wt enzyme. To obtain the sequence-specific assignment of the three Met methyl group ^{13}C resonances, the M14L- and M107L-LaL proteins containing [*methyl*-^{13}C]-Met were prepared and studied. The spectra for the M14L and M107L mutant proteins that were expressed in the presence of [*methyl*-^{13}C]Met are shown in Figure 2. From these spectra, the assignment of the ^{13}C resonances from the three Met residues in LaL was made. ^{13}C resonances located in the HMQC spectrum at δ 16.80, 17.20, and 18.26 result from methionines at positions 14, 1, and 107, respectively.

Studies on Solvent Exposure of Methionine Residues. With the assignments of the ^{19}F and ^{13}C resonances established, attempts were made to probe further into the specific properties and reactivities of the methionines in the protein using the respective methionine-labeled proteins.

Effect of the Gd·EDTA Complex on ^{19}F NMR Resonances of TFM-Labeled LaL. Surface exposure of fluorinated methionine side chains was probed by paramagnetic ion NMR line broadening experiments. Previous investigations have demonstrated the application of the aqueous paramagnetic Gd(III)−EDTA complex (Gd·EDTA) in probing the conformation-dependent solvent exposure of [5-F]Trp residues in the *E. coli* D-galactose receptor (*32*). We have applied this technique to TFM-labeled LaL and have found it to be extremely useful in elucidating the surface exposure of individual TFM residues as shown in Figure 3. The spectrum demonstrates the sensitivity of the TFM signal at position 1 which broadens rapidly and produces, after addition of 3.81 mM Gd·EDTA, a line width of >250 Hz, as compared to a line width of approximately 10 Hz in the absence of the reagent. Similarly, the two resonances arising from TFM at position 107 were broadened to the same extent, and to an only slightly lesser extent than the broadening noted for TFM

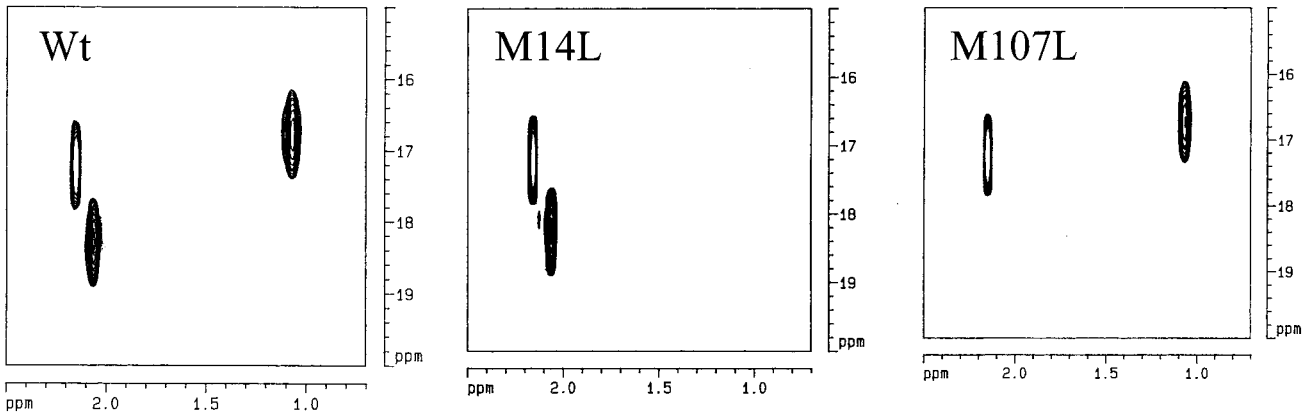


FIGURE 2: [^{1}H−^{13}C]-HMQC spectra of [*methyl*-^{13}C]Met-labeled LaL. Shown are the [^{1}H−^{13}C]-HMQC spectra of wt LaL, M14L-LaL, and M107L-LaL labeled with 100% [*methyl*-^{13}C]Met. Each sample was prepared in D_2O containing 50 mM potassium phosphate (pH 7.0) and 2,2,3,3-tetradeuterio-3-(trimethylsilyl)propionate as an internal standard (0.00 ppm).

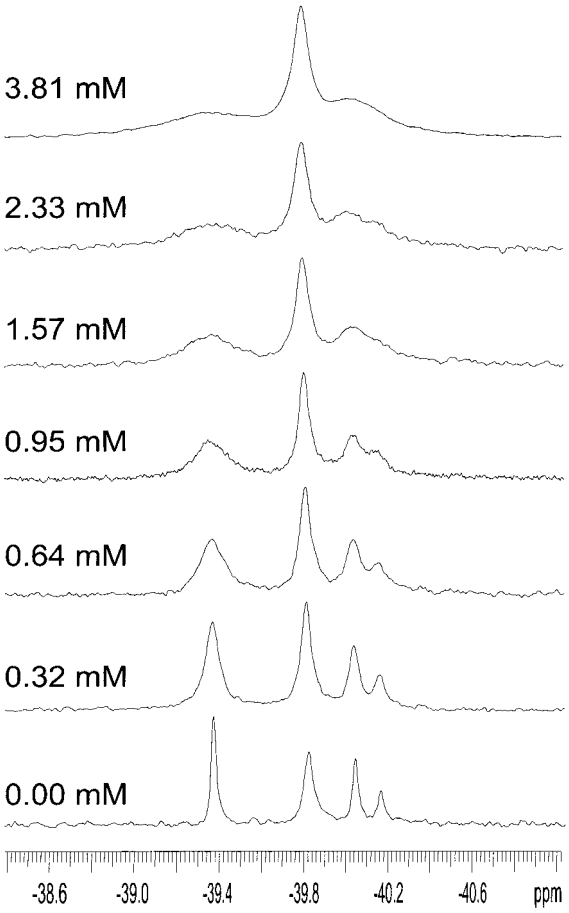


FIGURE 3: Effect of the presence of the Gd·EDTA probe on the ^{19}F NMR resonances of wt TFM-labeled LaL. Shown are a series of spectra for wt TFM-labeled LaL in the presence of increasing concentrations of the paramagnetic probe as indicated. Sample conditions were like those described in the legend of Figure 1.

at position 1. In contrast, the highest concentration of the Gd·EDTA complex only broadened the signal from TFM at position 14 by 10 Hz. It can be concluded that the N-terminus (Met1) and Met107 are solvent-exposed residues accessible to the Gd·EDTA complex whereas residue Met14 appears to be buried in the protein structure and is incapable of close interaction with the paramagnetic reagent.

Reaction of TFM- and [methyl-^{13}C]Met-Labeled LaL with K_2PtCl_4. It has been previously shown that certain platinum reagents, such as K_2PtCl_4, have a high propensity for reaction with surface methionine residues and have been useful in forming heavy atom derivatives for protein crystallography (*33*, *34*). To obtain further information about the environment surrounding the three methionine residues in LaL, the reaction of this enzyme with K_2PtCl_4 was investigated.

TFM residues in TFM-labeled LaL exhibited a lack of reaction with K_2PtCl_4 as monitored by ^{19}F NMR. Due to the sluggish nature of the reaction of TFM with K_2PtCl_4, NMR studies with K_2PtCl_4 were performed on the [*methyl*-^{13}C]-labeled lysozyme.

Incubation of [*methyl*-^{13}C]Met labeled LaL with K_2PtCl_4 (1−3.5 equiv) resulted in a rapid reaction with only the N-terminal methionine (Met1) as observed by NMR (Figure 4). The two ^{13}C resonances associated with the methyl group in Met14 and Met107 were not affected under any of the conditions that were tested. The spectrum acquired following

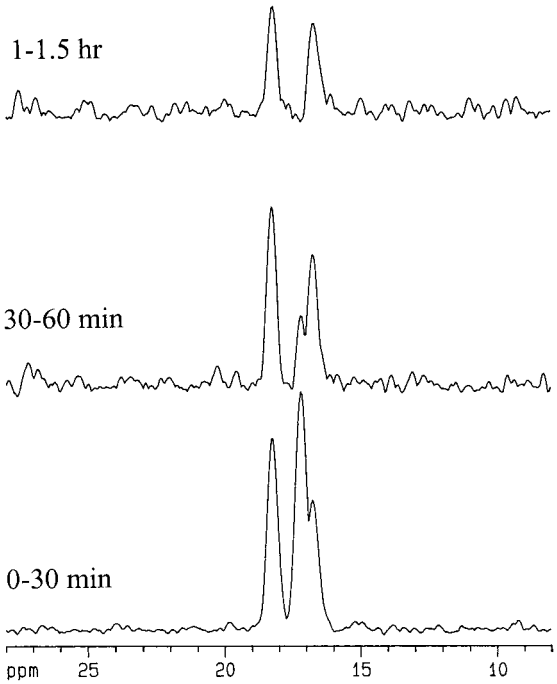


FIGURE 4: Effect of K_2PtCl_4 on the ^{13}C resonances of wt [^{13}C-Met]LaL. Shown are the spectra of [^{13}C-Met]LaL in the presence of 3 equiv of K_2PtCl_4 at various times as indicated.

a 30 min reaction time indicates that the Met1 resonance has been reduced in intensity by approximately two-thirds, and the resonance is no longer observed in the spectrum obtained after 1.5 h. Hence, the rapid reaction of K_2PtCl_4 is confined to Met1. Protein precipitation was observed upon exposure to higher K_2PtCl_4 concentrations, possibly due to the known tendency of K_2PtCl_4 to cross-link proteins at high concentrations (*35*, *36*).

ESMS was utilized to further explore the platination reaction and to analyze the time course of the reaction between wt LaL and K_2PtCl_4 (1−3.5 equiv) (Figure 5). The formation of an initial noncovalent complex between $[PtCl_4]^{2-}$ and LaL is consistent with a species with an observed mass of 18 162 Da. This is followed by a series of chloride ion losses over time to produce an apparent molecular species corresponding to the molecular mass of LaL plus a single Pt(II) atom. No evidence of two LaL molecules cross-linked by Pt(II) was found, but if formed, this species may have precipitated and been removed by centrifugation of the sample prior to ESMS analysis.

DISCUSSION

Almost all of the reported ^{19}F NMR studies on proteins containing fluorinated amino acids have utilized aromatic analogues; however, reports have recently appeared on the application of monofluoroleucine (*37*), difluoromethionine (*24*, *38*), and trifluoromethionine (TFM) analogues (*23*) to this area. TFM studies have focused upon its incorporation into a model protein, the lysozyme from bacteriophage lambda (LaL) (*23*). Four ^{19}F NMR resonances, A and B (each integrating to one TFM) and C and D (integrating together to one TFM), were observed for the LaL enzymes expressed in the presence of TFM, although there are only three Met residues in the amino acid sequence. With regard to resonances C and D, low levels of incorporation (∼31% TFM

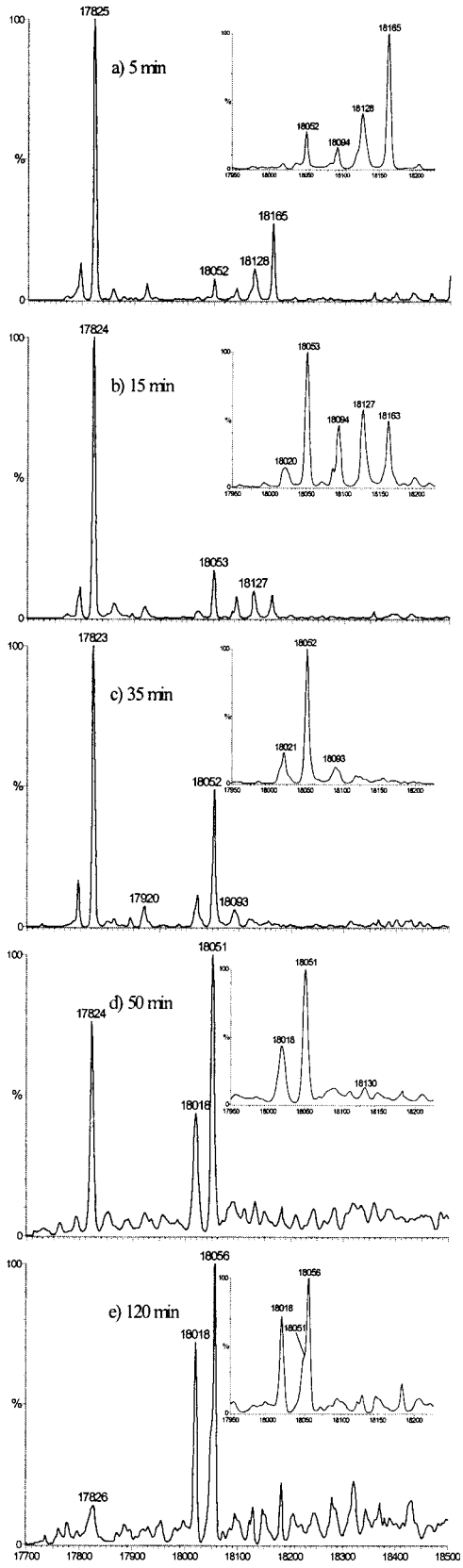


FIGURE 5: Reconstructed electrospray mass spectra of wt LaL treated with 3 equiv of K_2PtCl_4. The enzyme (0.3 mM) was treated with 0.9 mM K_2PtCl_4, and at the indicated times, an aliquot was diluted 100-fold and the spectrum immediately acquired.

incorporation) resulted in a predominance of resonance C; however, with high levels of TFM incorporation (~70% TFM incorporation), resonance D predominated for that position. Variable-temperature and phase-sensitive 2D exchange experiments with TFM-labeled LaL exhibited no evidence of chemical exchange for resonances C and D (*23*). These and other experiments indicated the possible existence of distinct, noninterconverting conformers of TFM-labeled LaL whose relative populations are dictated by levels of TFM incorporation. The overall ^{19}F NMR spectrum is the composite of the ensemble of variously TFM-substituted protein molecules in solution (e.g., mono-, bis-, and tris-incorporated TFM); a maximum of eight different proteins are produced, including the nonfluorinated enzyme. To understand, in molecular terms, the exact reasons for the multiple ^{19}F NMR resonances and, in a more general sense, to evaluate the impact that the introduction of TFM could have on macromolecular structure, the assignment of ^{19}F NMR resonances to specific residues in LaL was essential.

Leucine was chosen as a conservative replacement for methionine. The NMR resonance assignments were accomplished by preparation of the TFM-labeled mutants M14L and M107L. Resonances C and D were assigned to the position 107 TFM residue from analysis of the ^{19}F NMR spectrum of the fluorinated M107L mutein, which resulted in loss of resonances C and D (resonances A and B *each* integrate to one TFM residue in this mutein). Incorporation of TFM into the M14L mutein resulted in the loss of resonance B (position 14); also, only resonance C (integrating to one TFM molecule) and not resonance D from position 107 was observed. Since resonance C dominates under conditions with low-level TFM incorporation and is the sole resonance produced from TFM at position 107 in the M14L mutein, the extra resonance that is observed (resonance D) in the fluorinated wild-type enzyme is due to a subtle effect on position 107 by incorporation of a slightly larger TFM residue at position 14. Reducing the levels of TFM incorporation implies a greater population of protein molecules that contain predominantly Met over TFM at the three Met positions. The Met → Leu mutants artificially mimic the low-level incorporation conditions, and a smaller population of multiply TFM-labeled protein species results. From these observations, it can now be concluded that only when TFM occupies position 14 does a TFM residue at position 107 *in the same protein molecule* produce resonance D. Correspondingly, the presence of a Met residue at position 14 will result in a TFM at position 107 exhibiting only resonance C. Thus, it appears that two subtly different environments exist around position 107, the exact nature being dependent upon the particular amino acid (Met or TFM) at position 14. This phenomenon can be explained if one considers a tight packing environment around Met14. The replacement of a methyl group with a trifluoromethyl group in a tightly packed region could be sufficient to induce a conformational change that is transmitted through the protein to position 107. This hypothesis is supported by analysis of the recently determined crystal structures of LaL (see below) (*26*, *27*).

Reaction of TFM-Labeled LaL with the Gd·EDTA Complex. To further elucidate the connection between positions 14 and 107 from a structural perspective and how TFM could produce a perturbation in the protein, it was important to determine the spatial location of the three methionines as the crystal structure of the protein was unavailable at the time these studies were undertaken. Although little sequence homology exists between LaL and other lysozymes whose

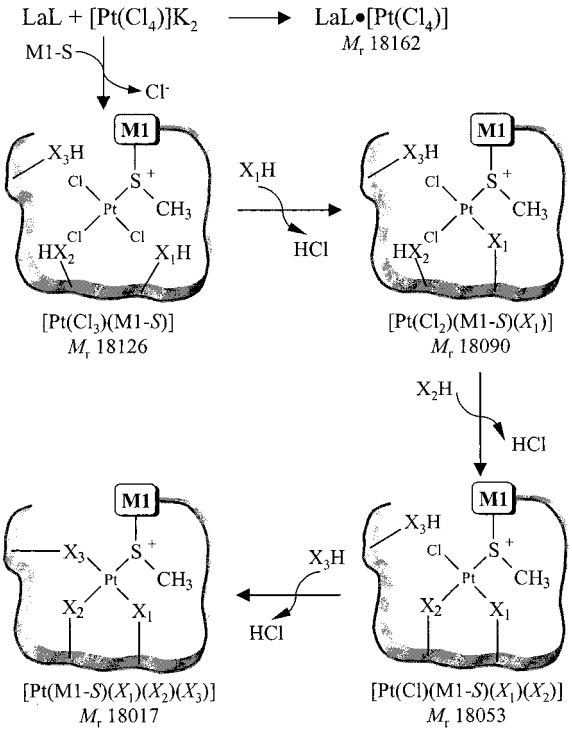
X-ray structures have already been determined (*39*), visual inspection of the primary sequence along with secondary sequence predictions of LaL suggested that Met1 could possibly be exposed to solvent, Met14 could be buried in the protein possibly in an α-helix, and Met107, which was found in a relatively hydrophilic sequence proximal to a proline residue, might well be in a solvent-exposed area of the protein. The application of paramagnetic line broadening reagents, such as Gd(III) ion, which act in a distance-dependent manner ($1/r^6$), has been shown previously to effectively probe the solvent exposure of fluorinated aromatic residues in proteins (*32*, *40*). The ^{19}F NMR resonances corresponding to Met1 (resonance A) and Met107 (resonances C and D) in wt TFM-labeled LaL exhibit concentration-dependent line broadening with increasing concentrations of the Gd·EDTA complex. There is no apparent line broadening effect on resonance B, which corresponds to Met14, suggesting a buried location for this residue. The Gd·EDTA complex has a diameter of approximately 11 Å (*32*). From this information, it can be concluded that positions corresponding to Met1 and Met107 are solvent-exposed whereas the Met14 position is buried in the protein and that the paramagnetic gadolinium reagent is incapable of close interaction with this residue.

Platination Reactions. Selective reaction of methionine residues with K_2PtCl_4 and *cis*-$[PtCl_2(NH_3)_2]$ can yield information about the solvent exposure of the thiomethyl group (*33*, *34*, *41*−*43*). Reaction of K_2PtCl_4 with TFM-labeled LaL, however, did not produce any TFM adduct (<5%) even under forcing conditions, as monitored by ^{19}F NMR. Reaction of trifluoromethionine itself with K_2PtCl_4 is much slower than that with normal methionine (M. D. Vaughan and J. F. Honek, unpublished results). This result is in agreement with the reduced reactivity exhibited by trifluoromethionine with other electrophilic reagents such as hydrogen peroxide and CNBr (*23*, *44*).

Reaction of K_2PtCl_4 was successfully performed on the [*methyl*-^{13}C]Met-labeled LaL. M14L and M107L mutations were used to assign the ^{13}C NMR resonances of [*methyl*-^{13}C]Met-labeled LaL (Figure 2). Since no steric changes are introduced into the protein upon incorporation of [*methyl*-^{13}C]Met, only three resonances were observed. Interaction of K_2PtCl_4 under a variety of concentrations (1−3.5 equiv) indicated rapid reaction of the N-terminal methionine with disappearance of the ^{13}C resonance corresponding to Met1 over time (Figure 4). The two ^{13}C resonances produced by Met14 and Met107 were not affected under these conditions, a result that initially appeared to be in disagreement with the suggested solvent exposure of the Met107 position as determined by the Gd·EDTA experiments on TFM-labeled LaL.

The platination reaction was further investigated by ESMS (Figure 5). The short time needed to obtain the ESMS data permitted the rapid acquisition of many spectra over a narrow time period and, therefore, favored the possible identification of intermediates. Transient intermediates may not be observed from HMQC NMR experiments due to the extended time requirements needed to obtain an individual spectrum. A tentative reaction sequence of LaL with K_2PtCl_4 is shown in Scheme 1 based on the time-dependent effects on mass peaks determined by ESMS (Figure 5). Peaks with masses corresponding very closely to those predicted for the suggested intermediates are seen to be present at various times and in varying intensities in the respective mass spectra. Although we cannot address with certainty the nature of the reaction products, we propose the following. The first step most likely involves reaction between the sulfur of Met1 and K_2PtCl_4 in which the sulfur becomes a ligand to Pt(II) after displacement of one of the four chlorine ligands. There is ample precedent for this type of interaction with methionine as well as methionyl residues in proteins (*33*, *34*, *41*, *45*−*48*). Following the formation of this initial monoadduct of platinum {i.e., [Pt(Cl$_3$)(M1-*S*)]; M_r = 18 126 Da}, a total of three additional functional groups (one of which is probably the α-NH_2 group of Met1) in the vicinity of the newly formed Pt(II)−protein sulfonium adduct also appear to become ligands to platinum, each successively displacing the remaining three chlorine atoms and forming the di-, tri-, and tetraprotein adducts (M_r = 18 090, 18 053, and 18 017 Da, respectively) with Pt(II) (Scheme 1). Recently, *S*,*N*-chelates of platinum (sulfur atom from an exposed methionine and a nitrogen from the protein) have been detected for reaction of cisplatin with serum albumin and transferrin (*42*, *43*). Subsequent formation of the 18 053 and 18 017 Da species suggests that other amino acid side chains of LaL displace the remaining two chlorine atoms. As indicated in Figure 5, there is a significant reduction in the relative intensity of the mass peak corresponding to unreacted protein (M_r = 17 825 Da) after 50 min, and after 120 min, essentially no unreacted protein is detected. This is similar to the time dependency regarding the disappearance of the Met1 resonance revealed from the NMR experiments noted above. After 24 h, no evidence of the involvement of the two additional methionines was seen by ESMS, confirming the NMR data that had dismissed the reaction of Met14 and Met107 with K_2PtCl_4. Therefore, sole and complete reaction

Scheme 1: Suggested Structures of the Intermediates Detected by ESMS (Figure 5) in the Time-Dependent Reaction of wt LaL with K_2PtCl_4



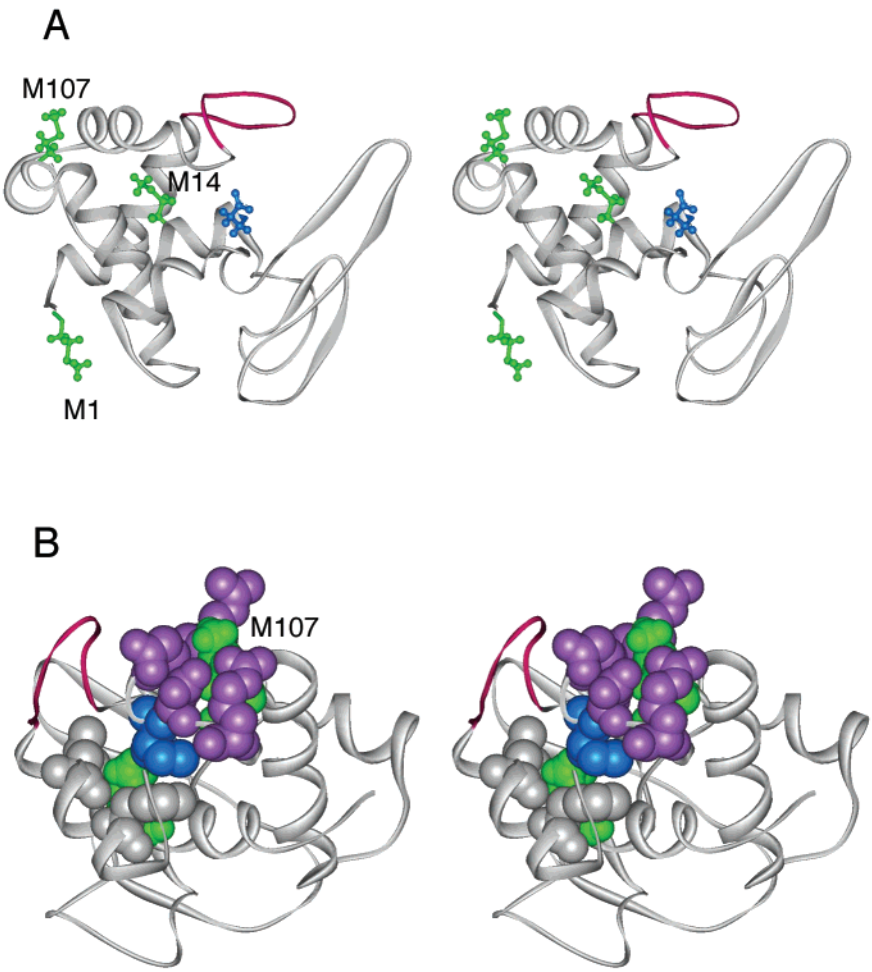


FIGURE 6: Stereoviews of the X-ray structure of wt LaL, showing the positions of the three Met residues. The structure is one of the monomers found in the unit cell of the chitohexaose-bound enzyme (*27*). The oligosaccharide has been removed for clarity. (A) The Met residues (green) and the catalytic general acid, Glu19 (blue), are shown. The mobile loop (shown closed) is red. (B) Reorientation of the view in panel A showing M14 and M107 in green, Leu142 and Phe146 in gray, Ile113 in blue, and Arg110, Asp112, Gln115, and Arg119 in purple. Hydrogen atoms have not been added to heavy atoms in this figure except for the thiomethyl groups of Met14 and Met107. This figure was generated using WebLab ViewerPro, version 3.7 (Molecular Simulations Ltd.).

under these conditions occurs between Met1 of the protein and K_2PtCl_4.

On the basis of the paramagnetic line broadening experiments and the platination studies, we can conclude that the N-terminal methionine is completely exposed to solvent. However, although TFM at position 107 is greatly affected by the paramagnetic line broadening agent, the Gd·EDTA complex, which indicates that this residue is close to the surface, reaction of the platination reagent with [*methyl*-^{13}C]-Met-labeled LaL indicates that the reactive lone pairs on the sulfur atom are not accessible for reaction. Position 14, on the other hand, does not react with the platinum reagent, nor are the ^{19}F signals substantially broadened by the Gd·EDTA complex. This suggests that this methionine residue is buried in the protein.

Crystal Structure of LaL. With the recent determination of the crystal structures of 7-azaTrp-substituted LaL (*26*) and wt LaL with bound chitohexaose (*27*), it is now possible to look at the structural environment of the three methionines in the protein. Because the structure of 7-azaTrp-substituted LaL was determined with three monomers in the unit cell, and that of wt LaL was determined with two monomers in the unit cell, it is possible to compare the structures of five different monomers of LaL. It is evident from the crystal structure (Figure 6) that Met1 and Met107 are solvent-exposed and Met14 is buried in the interior of the protein, which explains the effects of the Gd·EDTA reagent upon the ^{19}F NMR resonances. Similarly, the positions of Met1 and Met14 explain the results with K_2PtCl_4, in that Met1 reacts completely and Met14 does not react at all due to its position in the interior. Although the Met107 position was affected by the Gd·EDTA reagent, its lack of reactivity with the platinum reagent was surprising, and the structure provides the possible reason for this paradox. The sulfur atom of Met107 in all five structures of the various forms of LaL is turned inward toward the protein, and the surrounding residues (Pro106, Arg110, Asp112, Gln115, Arg119, and Gly103) provide possible steric protection from reagents such as K_2PtCl_4. Preliminary molecular dynamics simulations indicate that this protected conformation of Met107 is a predominant conformation, and little time is spent in alternate conformations which are less sterically encumbered (data not shown).

Possible Explanation for the Effect of TFM at Position 14 on Position 107. This work indicates that the incorporation of TFM into position 14 affects the ^{19}F resonance of TFM at position 107, possibly through a steric/electrostatic effect that is transmitted in some manner to position 107. In the

crystal structure of LaL (Figure 6), the specific side chains of amino acids Ile113, Leu142, and Phe146 are in close contact with the methyl group of Met14. As the size of the Met14 methyl group is increased by fluorination, the protein structure would need to accommodate this additional increase in size [a CF_3^- group has been suggested to be similar in size to an isopropyl moiety (*49*)]. Ile113 is also on the helix that contains Arg110, Asp112, Gln115, and Arg119, all of which surround residue Met107. It may be that the alteration of the ^{19}F resonance of TFM107 occurs due to the steric interaction between TFM14 and Ile113 (among other residues), and this interaction is transmitted to the helix residues proximal to TFM107. Although in general it has been found that fluorine is a nonperturbing atom in proteins containing fluorinated amino acids, it has been reported that the presence of even the single fluorine atom of [2-F]Phe (but not [3-F]-Phe or [4-F]Phe) in avian egg white lysozymes (*40*) and of [3-F]Tyr in rat glutathione transferase M1-1 can slightly perturb protein structure (*50*). In the latter case, X-ray crystallography has shown that the local conformational changes that occur for some of the [3-F]Tyr residues can even be propagated to non-neighboring residues (*51*).

Possible Explanation for ^{19}F Shifts upon Oligosaccharide Binding. It was previously observed that the ^{19}F NMR resonance from now identified TFM14 was affected in a major fashion upon oligosaccharide binding to TFM-labeled LaL. On the basis of comparison of the LaL structure with (*27*) and without (*26*) bound chitohexaose, it can be concluded that the helix in which Met14 is located is the same helix that contains the active site Glu19, believed to be the general acid that is essential for catalytic activity. If this residue and/or helix experiences a change in conformation, however slight, to accommodate the oligosaccharide, the movement may affect Met14, and hence could explain the upfield chemical shift change and the line broadening of the TFM14 ^{19}F NMR resonance (*23*). In addition, a loop containing residues 127−139 closes over the active site when oligosaccharide is bound. This alteration may also affect the TFM14 resonance. The slight effects on the ^{19}F NMR resonance of TFM107 resulting from oligosaccharide binding may be the indirect result of the changes in the Met14 environment. In addition, local effects produced by the loop movement directly alter the TFM107 environment.

Our results clarify some of the atomic level effects that are induced by this new ^{19}F NMR probe and the subtle structural perturbations that can occur upon introduction of a CF_3 group into a protein. Our results accurately reflect the topology of the methionine residues as observed in the recently determined crystal structures of this enzyme and further establish the applicability of this probe to problems in structural biochemistry.

ACKNOWLEDGMENT

We gratefully acknowledge Dr. Sandra Mooibroek for kind assistance with several aspects of the NMR experiments, Mr. Mark Vaughan for helpful discussions, and Dr. Gilles Lajoie for access to the ESMS facilities.

REFERENCES

1. Sykes, B. D., and Hull, W. E. (1978) *Methods Enzymol. 49*, 270−295.
2. Sykes, B. D., and Weiner, J. H. (1980) *Magn. Reson. Biol. 1*, 171−196.
3. Gerig, J. T. (1994) *Prog. NMR Spectrosc. 26*, 293−370.
4. Danielson, M. A., and Falke, J. J. (1996) *Annu. Rev. Biophys. Biomol. Struct. 25*, 163−195.
5. Sun, Z.-Y., Pratt, E. A., and Ho, C. (1996) *^{19}F-Labeled Amino Acids as Structural and Dynamic Probes in Membrane-Associated Proteins*, Vol. 639, American Chemical Society, Washington, DC.
6. Crabb, J. W., Carlson, A., Chen, Y., Goldflam, S., Intres, R., West, K. A., Hulmes, J. D., Kapron, J. T., Luck, L. A., Horwitz, J., and Bok, D. (1998) *Protein Sci. 7*, 746−757.
7. Luck, L. A., Barse, J. L., Luck, A. M., and Peck, C. H. (2000) *Biochem. Biophys. Res. Commun. 270*, 988−991.
8. Peteranderl, R., Rabenstein, M., Shin, Y. K., Liu, C. W., Wemmer, D. E., King, D. S., and Nelson, H. C. (1999) *Biochemistry 38*, 3559−3569.
9. Truong, H. T., Pratt, E. A., and Ho, C. (1991) *Biochemistry 30*, 3893−3898.
10. Li, D., Soriano, A., and Cowan, J. A. (1996) *Inorg. Chem. 35*, 1980−1987.
11. Lee, H. W., Sohn, J. H., Yeh, B. I., Choi, J. W., Jung, S., and Kim, H. W. (2000) *J. Biochem. 127*, 1053−1056.
12. Zemsky, J., Rusinova, E., Nemerson, Y., Luck, L. A., and Ross, J. B. (1999) *Proteins 37*, 709−716.
13. Sun, Z. Y., Pratt, E. A., Simplaceanu, V., and Ho, C. (1996) *Biochemistry 35*, 16502−16509.
14. Bai, P., Luo, L., and Peng, Z. (2000) *Biochemistry 39*, 372−380.
15. Ropson, I. J., and Frieden, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7222−7226.
16. Hoeltzli, S. D., and Frieden, C. (1998) *Biochemistry 37*, 387−398.
17. Williams, S. P., Fulton, A. M., and Brindle, K. M. (1993) *Biochemistry 32*, 4895−4902.
18. Williams, S. P., Haggie, P. M., and Brindle, K. M. (1997) *Biophys. J. 72*, 490−498.
19. Haggie, P. M., and Brindle, K. M. (1999) *J. Biol. Chem. 274*, 3941−3945.
20. Mehta, V. D., Kulkarni, P. V., Mason, R. P., Constantinescu, A., and Antich, P. P. (1994) *Bioconjugate Chem. 5*, 257−261.
21. Thorson, J. S., Shin, I., Chapman, E., Stenberg, G., Mannervik, B., and Schultz, P. G. (1998) *J. Am. Chem. Soc. 120*, 451−452.
22. Furter, R. (1998) *Protein Sci. 7*, 419−426.
23. Duewel, H., Daub, E., Robinson, V., and Honek, J. F. (1997) *Biochemistry 36*, 3404−3416.
24. Vaughan, M. D., Cleve, P., Robinson, V., Duewel, H. S., and Honek, J. F. (1999) *J. Am. Chem. Soc. 121*, 8475−8478.
25. Duewel, H. S., Daub, E., and Honek, J. F. (1995) *Biochim. Biophys. Acta 1247*, 149−158.
26. Evrard, C., Fastrez, J., and Declercq, J. P. (1998) *J. Mol. Biol. 276*, 151−164.
27. Leung, A. K.-W., Duewel, H. S., Honek, J. F., and Berghuis, A. M. (2001) *Biochemistry 40*, 5665−5673.
28. de Dios, A. C., Pearson, J. G., and Oldfield, E. (1993) *Science 260*, 1491−1496.
29. Pearson, J. G., Montez, B., Le, H., Oldfield, E., Chien, E. Y., and Sligar, S. G. (1997) *Biochemistry 36*, 3590−3599.
30. Lau, E. Y., and Gerig, J. T. (2000) *J. Am. Chem. Soc. 122*, 4408−4417.
31. Arnold, W. D., Mao, J., Sun, H., and Oldfield, E. (2000) *J. Am. Chem. Soc. 122*, 12164−12168.
32. Luck, L. A., and Falke, J. J. (1991) *Biochemistry 30*, 6484−6490.
33. Dickerson, R. E., Eisenberg, D., Varnum, J., and Kopka, M. L. (1969) *J. Mol. Biol. 45*, 77−84.
34. Howe-Grant, M. E., and Lippard, S. J. (1980) *Met. Ions Biol. Syst. 11*, 63−125.
35. Zhicheng, X., Ziyong, S., Zhihao, Y., Wenxia, T., and Dexu, Z. (1992) *J. Inorg. Biochem. 48*, 197−202.
36. Trynda, L., and Kuduk-Jaworska, J. (1994) *J. Inorg. Biochem. 53*, 249−260.

37. Feeney, 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., and Oldfield, E. (1996) *J. Am. Chem. Soc. 118*, 8700−8706.
38. McIntyre, D. J., Yuan, T., and Vogel, H. J. (1996) *Prog. Biophys. Mol. Biol. 65*, P-A1-27.
39. Jespers, L., Sonveaux, E., and Fastrez, J. (1992) *J. Mol. Biol. 228*, 529−538.
40. Lian, C., Le, H., Montez, B., Patterson, J., Harrell, S., Laws, D., Matsumura, I., Pearson, J., and Oldfield, E. (1994) *Biochemistry 33*, 5238−5245.
41. Melius, P., and Friedman, M. E. (1977) *Inorg. Perspect. Biol. Med. 1*, 1−18.
42. Ivanov, A. I., Christodoulou, J., Parkinson, J. A., Barnham, K. J., Tucker, A., Woodrow, J., and Sadler, P. J. (1998) *J. Biol. Chem. 273*, 14721−14730.
43. Cox, M. C., Barnham, K. J., Frenkiel, T. A., Hoeschele, J. D., Mason, A. B., He, Q. Y., Woodworth, R. C., and Sadler, P. J. (1999) *J. Biol. Inorg. Chem. 4*, 621−631.
44. Duewel, H. S., and Honek, J. F. (1998) *J. Protein Chem. 17*, 337−350.
45. Boswell, A. P., Moore, G. R., and Williams, R. J. (1982) *Biochem. J. 201*, 523−526.
46. Gummin, D. D., Ratilla, E. M. A., and Kostic, N. M. (1986) *Inorg. Chem. 25*, 2429−2433.
47. Wilson, C., Scudder, M. L., Hambley, T. W., and Freeman, H. C. (1992) *Acta Crystallogr. C48*, 1012−1015.
48. Norman, R. E., Ranford, J. D., and Sadler, P. J. (1992) *Inorg. Chem. 31*, 877−888.
49. O'Hagan, D., and Rzepa, H. S. (1997) *J. Chem. Soc., Chem. Commun.*, 645−652.
50. Parsons, J. F., Xiao, G., Gilliland, G. L., and Armstrong, R. N. (1998) *Biochemistry 37*, 6286−6294.
51. Xiao, G., Parsons, J. F., Tesh, K., Armstrong, R. N., and Gilliland, G. L. (1998) *J. Mol. Biol. 281*, 323−339.

BI011381B