Assignment and Analysis of Fluorine Nuclear Magnetic Resonance Spectra of 4-Fluorotryptophan Myoglobins and Hemoglobins†

John G. Pearson, Bernard Montez, Hongbiao Le, and Eric Oldfield*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

Ellen Y. T. Chien and Stephen G. Sligar

Biophysics Program, Beckman Institute, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, Illinois 61801

Received July 9, 1996; Revised Manuscript Received January 27, 1997⊗

ABSTRACT: We have obtained the 470 MHz ^{19}F NMR spectra of wild type [4-F]Trp-labeled myoglobins (MbCO, MbO_2, deoxyMb, metMb, and MbCN) and hemoglobins (HbCO, HbO_2, and deoxyHb), as well as those of several mutants (W7F Mb, βW15F Hb, βW37S Hb, and βY130F Hb, all as the carbonmonoxy adducts), prepared via site-directed mutagenesis. The maximum observed chemical shift range induced by folding is 6.4 ppm. Using a multipole shielding polarizability−local reaction field approach, we have computed the electrostatic field contributions to the fluorine shielding. For residues which do not have F atoms in contact with neighboring groups, we find an ∼1 ppm mean square deviation in shift from experiment, with the R2-like structure of HbCOA being in very close accord with experiment.

Folding a protein into its native conformation generates a large range of nuclear magnetic resonance (NMR)[1] chemical shift nonequivalencies, especially for the higher Z nuclei such as ^{13}C, ^{15}N, ^{17}O, and ^{19}F. While the shifts of all NMR active nuclei are encoded with structural information, ^{19}F-labeled aromatic amino acids in particular have been used since the 1970s to probe protein structure (Browne & Otvos, 1976; Chaiken et al., 1973; Falke et al., 1992; Gamcsik & Gerig, 1986; Gamcsik et al., 1986, 1987; Gerig et al., 1983; Ho et al., 1989; Hull & Sykes, 1975, 1976; Kimber et al., 1977, 1978; Luck & Falke, 1991a−c; Peersen et al., 1990; Robertson et al., 1977; Taylor et al., 1981), due to both the fluorine chemical shift's strong environmental dependence and the spectral simplicity afforded by having few resonances when compared to the "natural" nuclei, ^{1}H and ^{13}C (and^{15}N). In this article, we show that combining rigorously assigned spectra with chemical shielding calculations opens new avenues to structural analysis in general, and the potential of fluorotryptophan as an environmental probe for heme proteins in particular.

Prediction of folding-induced changes in protein chemical shifts is a potentially powerful method of testing the accuracy of proposed protein structures. However, successful prediction of protein NMR spectra requires both a reliable relationship between structure and chemical shift and the availability of accurate protein structures. Since most NMR studies of proteins are conducted in aqueous solution while most protein structures are from diffraction experiments in the solid state, there is some question as to whether NMR data and diffraction data from a particular protein refer to the *same* protein structure (Smith et al., 1994). Indeed, many proteins have several crystal structures differing in symmetry and tertiary conformation, raising concerns that there may be significant crystal−solution differences which will confound detailed analysis of solution NMR chemical shifts. One particularly interesting case is that of ligated human hemoglobin A, which has two reported crystal structures, R and R2. Since the two structures differ in and around the α_1/β_2 interface, comparison of predicted and observed spectra of fluorine-labeled hemoglobins might allow a determination of which crystal structure is a better representation of the protein in solution, which in turn should lead to a better understanding of allosteric control mechanisms in hemoglobin.

In this paper, we present ^{19}F NMR results on a variety of [^{19}F]Trp-labeled hemoglobins and myoglobins, and we analyze the observed chemical shifts in terms of the effects of weak electric fields generated by the protein on the NMR peak position. This approach has been used by us previously (Pearson et al., 1993) to analyze ^{19}F chemical shifts in the galactose binding protein, so we limit discussion of the method to a brief review.

The effect of an electric field on a molecular electronic property, *P*, such as the chemical shielding, was first suggested for small molecules in fluids (Buckingham, 1960; Stephen, 1957) as being expressible as a power series in the uniform field, **F**:

$$P_{\alpha\beta} = P_{\alpha\beta}^{(0)} + P_{\alpha\beta,\gamma}^{(1)}F_\gamma + P_{\alpha\beta,\gamma,\delta}^{(2)}F_\gamma F_\delta + ... \quad (1)$$

assuming that the field is a uniform one arising from the

† This work was supported by the National Institutes of Health (Grants HL-19481, HL-51084, GM-31756, and GM-50694). J.G.P. was supported by an American Cancer Society Postdoctoral Fellowship (Grant PF-4141). H.L. was supported by a University of Illinois Graduate Fellowship.

⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1997.

[1] Abbreviations: NMR, nuclear magnetic resonance; [4-F]Trp, 4-fluorotryptophan; WT, wild type; Mb, myoglobin; metMb, metmyoglobin; MbCN, cyanometmyoglobin; MbCO, carbonmonoxymyoglobin; [4-F]Trp Mb, myoglobin with 4-fluorotryptophan in place of tryptophan; Hb, recombinant adult human hemoglobin; HbCO, carbonmonoxy-Hb; HbO_2, oxy-Hb; [4-F]Trp Hb, Hb with 4-fluorotryptophan in place of tryptophan; MD, molecular dynamics; CHF, coupled Hartree−Fock; LRF, local reaction field; MSP, multipole shielding polarizability; vdW, van der Waals.

S0006-2960(96)01664-9 CCC: $14.00

neighboring molecules. Here, P is a molecular property (e.g. the chemical shift), and $P_{\alpha\beta}^{(0)}$ is the $\alpha\beta$ component of the tensorial observable in the absence of a field. $P_{\alpha\beta,\gamma}^{(1)}$ is the derivative of $P_{\alpha\beta}$ with respect to a uniform field along γ, the so-called dipole shielding polarizability. $P_{\alpha\beta,\gamma}^{(1)} F_\gamma$ is the first-order change of P due to the field, F_γ, etc. However, electric fields in macromolecules such as proteins are in general nonuniform, requiring additional terms:

$$P_{\alpha\beta} = P_{\alpha\beta}^{(0)} + P_{\alpha\beta,\gamma}^{(1)} F_\gamma + P_{\alpha\beta,\gamma\delta}^{(1)} F_{\gamma\delta} + P_{\alpha\beta,\gamma,\delta}^{(2)} F_\gamma F_\delta + P_{\alpha\beta,\gamma\delta,\epsilon\zeta}^{(2)} F_{\gamma\delta} F_{\epsilon\zeta} + \ldots \quad (2)$$

where $F_{\gamma\delta}$ is an element of the field gradient tensor, as suggested previously (Batchelor, 1975; Buckingham & Lawley, 1960). Knowledge of the terms in eq 2 therefore permits calculation of the electric field contributions to the chemical shift. Dykstra and colleagues (Augspurger et al., 1991, 1992; Augspurger & Dykstra, 1991) have calculated ^{19}F shielding polarizability tensors for fluorobenzene, which we have adopted as a reasonable model of the shielding polarizability tensors of [4-^{19}F]tryptophan. Warshel and colleagues have developed a molecular dynamics program, ENZYMIX (Lee & Warshel, 1992; Warshel & Creighton, 1989), for rapid calculation of molecular dynamics trajectories of electric field and field gradient tensors at sites of interest. The application of multipole shielding polarizability coefficients to the MD electric field and field gradient trajectories thus generates chemical shift trajectories, which may be averaged to simulate the electric field response of the total chemical shift (Pearson et al., 1993). The convergence properties shown in eq 2 have been investigated previously, and it has been found that at typical interatomic separations only the terms linear in the field and field gradient are of interest, with the uniform field component dominating (Augspurger et al., 1992; Pearson et al., 1993). The MSP−LRF approach neglects the possibility that van der Waals dispersion forces make a major contribution to ^{19}F shielding in proteins, a view which is supported by the ability to compute ^{19}F shifts in proteins and model systems at the coupled Hartree−Fock (CHF) level (de Dios & Oldfield, 1994; de Dios et al., 1993), which also omits van der Waals dispersion. Indeed, such methods even permit the accurate prediction of complete ^{19}F shielding tensors in small molecules, with no adjustable parameters (de Dios & Oldfield, 1994). Thus, it seems reasonable to attempt to simulate ^{19}F spectra of fluorotryptophan-labeled heme proteins by use of the MSP−LRF approach, as discussed below.

As in our previous work (Pearson et al., 1993), we evaluate solely the electrostatic contributions to shielding, σ, which are related to the experimentally observed chemical shifts, δ, as

$$\delta = x - \sigma$$

where x is related to the absolute ^{19}F shieldings. Assuming the dominance of electrostatic field effects, δ and σ will be correlated, with a slope of -1 (and an R^2 value of 1 also).

MATERIALS AND METHODS

Biochemical Aspects. The W7F Mb mutant was constructed using PCR-mediated site-directed mutagenesis, for the purpose of making a specific W14F assignment. The alternate strategy of selective ^{1}H decoupling of the 5-H resonances was not pursued since the effects of ^{19}F substitution on the ^{1}H chemical shifts is not known. The mutagenic primer 5′-AACTGCAGATAACTAACTAAAGGAGAA-CAACAACAATGGTTCTGTCTGAAGGTGA-A**TTC**CAGCTGGTTC-3′ contains the codon for phenylalanine in place of tryptophan at position 7. The PCR product was digested with *Pst*I and *Hin*dIII, gel purified, treated as a cassette, and subcloned directly into the *Pst*I and *Hin*dIII sites of pMb122. pMb122 contains the synthetic gene for wild type sperm whale Mb. It is a modified pMb413a with the correct amino acid (aspartic acid) at position 122 (Phillips et al., 1990; Springer & Sligar, 1987).

βW15F, βW37S, and βY130F hemoglobin mutants were also constructed using cassette mutagenesis (Hernan & Sligar, 1995). βW15F was cloned into the *Nco*I and *Hpa*I sites, the βW37S into *Sal*I and *Bgl*II sites, and βY130F into *Eco*RI and *Mlu*I sites of pWHS486. All work on hemoglobin is based upon the bacterially expressed synthetic human α- and β-globin genes with the first amino acid (valine) replaced by methionine, for both α- and β-chains (Hernan & Sligar, 1995).

The oligonucleotides were synthesized at the University of Illinois Biotechnology Center Genetic Facility. Restriction enzymes, ligases, and buffers were purchased from New England Biolabs (Beverly, MA) and Bethesda Research Laboratories (Grand Island, NY). Mutant sequences were confirmed using the Sequenase double-stranded DNA sequencing reagents and protocols from United States Biochemicals (Cleveland, OH).

The plasmids were transformed into *Escherichia coli* strain NCM533 (F^+ lacA::Tn5, labIq, λ^+, Kan^-). The cells were grown on a shaker at 37 °C in Moore's medium (Chiu, 1986) [1 mM $MgCl_2$, 2.2 μM $FeSO_4$, 40.1 mM KH_2PO_4, 65.4 mM Na_2HPO_4, 25.9 mM NaCl, 17.5 mM $(NH_4)_2SO_4$, and glycerol (10 g/L) as carbon source], supplemented with ampicillin (200 mg/L) and thiamine (10 mg/L). At mid-log phase, 0.2 mM isopropyl β-D-thiogalactopyranoside was added to induce protein expression, and the medium was further supplemented with 1 μM 5-aminolevulinic acid, tryptophan (6 mg/L), 4-fluorotryptophan (50 mg/L), phenylalanine (50 mg/L), tyrosine (50 mg/L), glyphosate (Aldrich, Milwaukee, WI) (2 mL/L) to block the cell's tryptophan biosynthesis (Bode et al., 1985, 1986; Kunze et al., 1989), ampicillin (200 mg/L), and glycerol (10 g/L). Growth was continued for another 9−12 h on a shaker, at 37 °C. Cells were harvested by centrifugation and stored frozen at −80 °C until ready for protein isolation.

Myoglobins were purified as described in Abadan et al. (1995) and hemoglobins as described in Hernan and Sligar, (1995). The α- and β-chains of hemoglobin were separated following the procedure by Bucci and Fronticelli (1965).

Kinetic Measurements. CO recombination rate constants to myoglobins and hemoglobins were measured using conventional flash photolysis and stopped-flow rapid mixing techniques, respectively, as described in detail by Rohlfs et al. (1990). Hemoglobin oxygen binding equilibrium concentrations and partial pressures were measured using the hemox-analyzer described in Hernan and Sligar (1995), which is similar to the automatic oxygenation apparatus described by Imai (1981). All ligand bindings to hemoglobin were measured in 50 mM Tris and 0.1 M Cl^- at pH 7.4 and 20 °C.

NMR Spectroscopy. Proton-coupled ^{19}F NMR spectra were measured at 470 MHz on a "home-built" spectrometer which consists of a Nicolet (Madison, WI) model 1280 computer, an 11.7 T, 2.0 in. bore superconducting solenoid magnet (Oxford Instruments, Osney Mead, Oxford, U.K.), and a Cryomagnet Systems (Indianapolis, IN) 5 mm ^{19}F NMR probe, basically as described previously (Lian et al., 1994). Purified proteins were concentrated to 0.5 mM in 50 mM Tris, 1 mM EDTA, pH 8 aqueous solutions. The expressed met proteins were reduced with sodium dithionite (1 mM) and then purged with CO, N_2, or O_2 to generate the respective carbonmonoxy-, deoxy-, and oxyheme proteins. NMR experiments were performed at room temperature (~24 °C) for all myoglobin samples and for wild type and βY130F hemoglobin samples, at 15 °C for βW15F HbCO, and at 4 °C for βW37S HbCO. For βW37S HbCO, we also used a Hahn spin echo experiment to reduce the probe background signal. Spectral widths were ±6.25 kHz, and 4096 complex data points were collected for each spectrum. The recycle time was 2.0 s between each acquisition. All ^{19}F chemical shifts are referenced to neat external trifluoroacetic acid at 0 ppm. The convention used is that high-frequency, low-field, paramagnetic, or deshielded values are more positive (IUPAC δ-convention for chemical shifts). Myoglobin spectra required about 90 min to acquire, while hemoglobin spectra required about 12 h, the additional acquisition time being needed to compensate for the relatively high fraction of denatured Hb and the increased line widths, and the lower effective concentrations due to α/β nonequivalence.

Chemical Shift Calculations. ^{19}F NMR spectra were simulated via the MSP−LRF approach, using a modified version of the ENZYMIX program. All computations were performed on IBM RISC/6000 computers (IBM Corp., Austin, TX). [4-F]Trp residues were created by replacing $H^{\epsilon 3}$ of each Trp residue with a fluorine atom. The fluorine vdW radius was approximated by using a typical oxygen value, and the charges on the fluorine and its bonded carbon were set to $-0.25e$ and $+0.25e$, respectively (Pearson et al., 1993). Time steps of 2 fs were used, and heavy atoms were subjected to a harmonic constraint of 0.1 kcal mol^{-1} $Å^{-2}$. Each trajectory was generated by first subjecting the system to 1000 steps at 500 K for randomization, followed by 1000 steps at 300 K for equilibration, and ending with 10 000 steps at 300 K for observation, producing a 20 ps trajectory. Since the NMR time scale is much longer than our trajectories, we calculated multiple trajectories for each site, with values reported typically being averaged over six trajectories. Starting structures were taken from the Protein Data Bank of the Brookhaven National Laboratory (Abola et al., 1987; Bernstein et al., 1977), with structure 1hho (Shaanan, 1983) taken to represent the R state of hemoglobin A, structure 2hhb (Fermi et al., 1984) taken to represent the T state of hemoglobin A, structure 1bbb (Arnone & Silva, 1992) taken to represent the R2 state of hemoglobin A, structure 1mbc (Kuriyan et al., 1986) taken to represent sperm whale carbonmonoxymyoglobin, and structure 1mbd (Phillips & Schoenborn, 1981) taken to represent sperm whale deoxymyoglobin. Additionally, we modified structure 1bbb (Arnone & Silva, 1992) by replacing tyrosine β130 with phenylalanine, generating a starting structure suitable for simulations of CO-ligated βY130F mutant Hb.

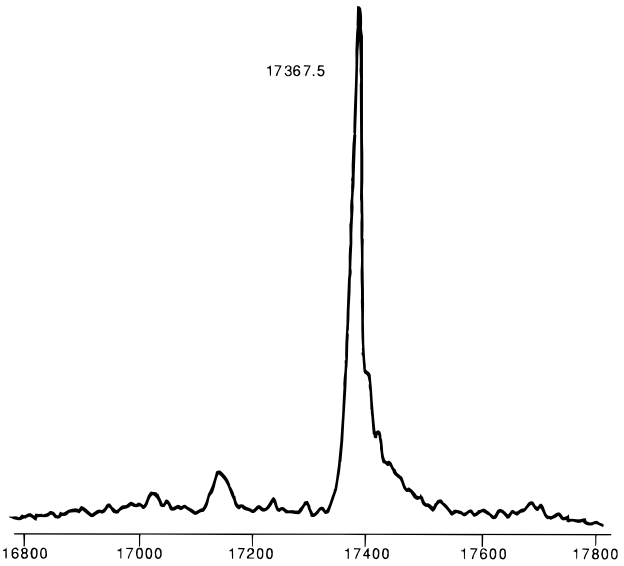

FIGURE 1: Electrospray mass spectrum of wild type [4-F]Trp myoglobin. The calculated value for unlabeled Mb is 17 334 Da and for fully labeled Mb is 17 370 Da. The result indicates the sample has greater than 90% [4-F]Trp incorporation.

Ring current and peptide group magnetic anisotropy contributions to the chemical shielding were also calculated, using the SHIFTS program of Case et al., (1994). Since such contributions were expected to be about 1 order of magnitude smaller than those expected for the electrostatic contributions, due to large distances to the heme ring, the SHIFTS program was used only on the starting X-ray structures (i.e. the SHIFTS program was not incorporated into the molecular dynamics trajectories).

RESULTS AND DISCUSSION

^{19}F NMR Spectra of Myoglobin and Hemoglobin. The construction of synthetic human hemoglobin and sperm whale myoglobin genes for bacterial expression (Hernan et al., 1992; Springer & Sligar, 1987) permits the ready expression of highly ^{19}F-labeled heme proteins, as shown in the electrospray mass spectrum (ESMS) of Figure 1, as well as permitting a ready route to spectral assignment. From the ESMS shown in Figure 1, we conclude that the WT Mb sample has >90% [4-F]Trp incorporation, and similar high levels of incorporation are observed for the other proteins investigated as well. Figure 2A shows the 470 MHz ^{19}F NMR spectrum of [4-F]Trp MbCO (pH 8.0, 24 °C, and 0.5 mM). There are two resonances observed, which originate from Trp 7 and Trp 14 of the wild type protein. The chemical shift separation is 4.75 ppm, only about $^1/_2$ of that seen in the galactose binding protein (Luck & Falke, 1991b; Pearson et al., 1993), implying a rather large challenge for ^{19}F NMR chemical shift calculations, which can only be expected to be accurate to about 1−2 ppm (Pearson et al., 1993). Figure 2B shows the ^{19}F NMR spectrum of the W7F MbCO mutant, in which the low-field peak has essentially disappeared, enabling its assignment to Trp 7, with the more shielded resonance being attributable to Trp 14.

The assignment of Mb W14 is based on the near identity of the W14 chemical shift in WT MbCO and the W7Y mutant. The difference in chemical shift is ~0.5 ppm. This small change in shift must be attributed to the W → Y mutation at position 7, two helix turns away.

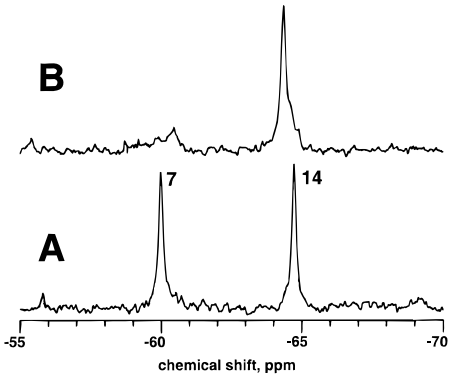

FIGURE 2: 11.7 T proton-coupled ^{19}F NMR spectra of 0.5 mM [4-F]Trp-labeled (A) wild type and (B) W7F mutant carbonmonoxymyoglobin (sperm whale) in 50 mM Tris, 1 mM EDTA, pH 8 aqueous solutions at 24 °C. Each spectrum contains 4096 complex data points, has a spectral width of ±6.25 kHz, and is referenced to neat external trifluoroacetic acid at 0 ppm. The spectra were signal averaged over 50 700 scans with a recycle delay between scans of 2.0 s.

Table 1: Spectral Assignments of ^{19}F Resonances in [4-F]Trp Myoglobins and Hemoglobins

	chemical shift, δ (ppm)[a]	
system	7	14
MbCO	−60.0	−64.4
MbO_2	−60.0	−64.2
deoxyMb	−59.8	−64.4
metMb	−59.8	−64.0
MbCN	−60.2	−64.8

	chemical shift, δ (ppm)[a]		
system	α14	β15	β37
WT HbCO	−66.2	−62.9	−61.4
βY130F HbCO	−66.2	−63.3	−61.4
HbO_2	−66.3	−63.3	−61.5
deoxyHb	−66.3	−61.5	−59.9

[a] Chemical shifts are in parts per million from the external standard of neat trifluoroacetic acid (IUPAC δ-convention).

The alternative assignment would involve a remarkable 5 ppm deshielding for W14 in the WT MbCO and a remarkably coincidental 5 ppm increase in shielding for W14, due to the W → Y mutation.

The following facts therefore form the basis for our Trp assignments.

(1) There is a peak at the right position in W7F MbCO (within 0.5 ppm of the WT shift).

(2) It is at the calculated chemical shift (see below).

(3) While the chemical shift is slightly different in the mutant (∼0.5 ppm), this is about the same difference seen across the range of systems, MbCO, MbO_2, Mb, metMb, and MbCN, due to small structural changes (Table 1).

(4) There is a major peak at about this chemical shift position in α_4 of HbCO.

(5) There have never been 5 ppm shift changes seen previously involving mutations >6 Å from ^{19}F.

(6) There are only minor (1 ppm) changes seen in [4-F]-Trp shielding, due to multiple mutations, in other proteins, e.g. lysozyme (Lian et al., 1994).

Thus, all available information supports the assignments given.

As noted above, we also obtained the ^{19}F NMR spectra of oxymyoglobin, deoxymyoglobin, metmyoglobin, and cyanometmyoglobin, and the chemical shifts obtained are reported in Table 1. As can be seen, there are only minor differences between the shifts observed in the paramagnetic and diamagnetic proteins, consistent with the large distances (∼15 Å) of the ^{19}F sites from the heme iron. For MbO_2, an interesting time dependence of the spectra was seen, as shown in Figure 3. The initial spectrum represents pure MbO_2, but over time, the sample converts to deoxyMb, and finally to metMb. Figure 3 shows the ^{19}F NMR spectra for wild type oxymyoglobin at 6 h intervals. Resonance assignments were achieved by comparing the observed spectra at 1.5 h time intervals (only every fourth spectrum is shown) to the pure deoxy- and metMb ^{19}F NMR spectra (data not shown). Spectra of MbO_2 and deoxyMb taken at short time values indicate a ∼0.2 ppm separation between the oxy and deoxy states. Since the peaks are therefore in slow exchange, these results indicate that intermolecular O_2 exchange occurs at a rate much slower than 100 s^{-1}. We now move to hemoglobin spectra and assignments.

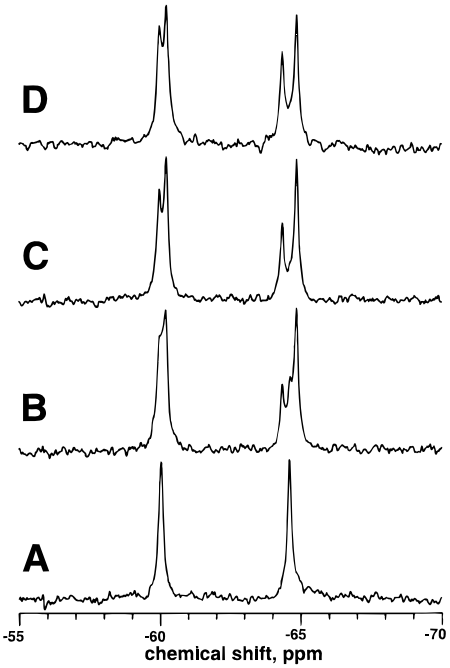

FIGURE 3: 11.7 T proton-coupled ^{19}F NMR spectra of 0.5 mM [4-F]Trp-labeled wild type oxymyoglobin in 50 mM Tris, 1 mM EDTA, pH 8 aqueous solutions at 24 °C after (A) 0 h and 1000 scans, (B) 6 h and 2500 scans, (C) 12 h and 2500 scans, and (D) 18 h and 2500 scans. New resonances in later spectra are thought to be due to deoxymyoglobin and met(aquo)myoglobin. Each spectrum contains 8192 complex data points, has a spectral width of ±6.25 kHz, and is referenced to neat external trifluoroacetic acid at 0 ppm. The recycle delay between scans was 2.0 s.

We show in Figure 4 the ^{19}F NMR spectra for wild type HbCO, the α-subunit of HbCO, and βW15F HbCO. The four resonance peaks in wild type hemoglobin spectra are due to the three chemically distinct tryptophans in hemoglobin, α14, β15, and β37, plus a fourth resonance arising from unfolded protein. The presence of a large denatured peak in Hb (but not Mb) spectra may be due to folding or assembly errors occurring during Hb tetramer assembly from rHb subunits (Hernan & Sligar, 1995). The denatured peak was readily assigned a chemical shift of −60.2 ppm by observing that all the peaks collapsed into it upon thermal denaturation (data not shown). To assign α14, we separated the α- and β-chains of labeled hemoglobin and measured the spectrum of the isolated α-chains, shown in Figure 4B.

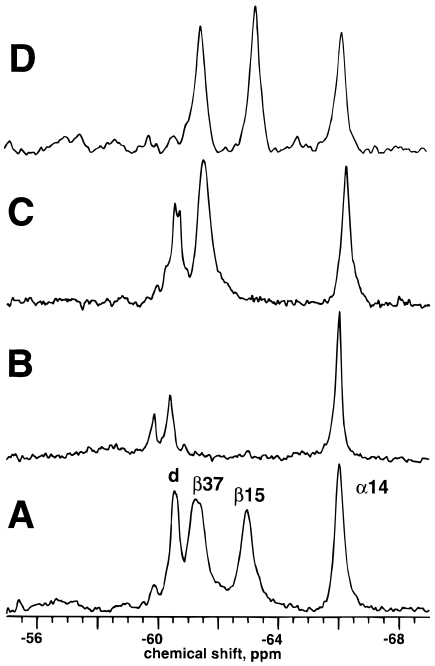

FIGURE 4: 11.7 T proton-coupled ^{19}F NMR spectra of 0.5 mM [4-F]Trp-labeled hemoglobins in 50 mM Tris, 1 mM EDTA, pH 8 aqueous solutions. (A) Wild type [4-F]Trp HbCO, 28 440 scans, 25 Hz line broadening due to exponential multiplication, and 24 °C. (B) [4-F]Trp HbCO α subunit, 8000 scans, 25 Hz line broadening due to exponential multiplication, and 24 °C. (C) [4-F]-Trp βW15F mutant HbCO, 30 000 scans, 25 Hz line broadening due to exponential multiplication, and 15 °C. (D) βY130F mutant [4-F]Trp HbCO, in 50 mM Tris, 1 M ammonium sulfate (pH 8.0), and 1 M AMS, 32 000 scans, 25 Hz line broadening due to exponential multiplication, and 24 °C. Each spectrum contains 8192 complex data points, has a spectral width of ±6.25 kHz, and is referenced to neat external trifluoroacetic acid at 0 ppm. The recycle delay between experiments was 2.0 s.

The dominant resonance at −66.2 ppm, the most upfield, is assigned to α14, with the smaller peaks being due to denatured protein. Next, we used site-directed mutagenesis to assign β15 and β37. In particular, we were able to assign β15 to the peak at −62.9 ppm and β37 to the peak at −61.4 ppm by comparing the spectrum for βW15F HbCO (Figure 4C) to the wild type HbCO spectrum, with the peak deleted between spectra being attributed to β15, and the remaining peak being attributed to β37. The spectrum of [4-F]Trp βW37S mutant Hb was also measured, and while consistent with the above assignment, the reduced stability of the βW37S mutant Hb prevented the acquisition of enough scans for generation of a high *S*/*N* ratio spectrum.

In principle, it might be possible to further substantiate these assignments via titration of labeled α-subunits with unlabeled β-subunits to show that the chemical shift is unperturbed as the second subunit is added. However, if free α-subunits are titrated with unlabeled β-subunits, the fraction of the α-subunits in the sample must remain stable during data collection, and α-subunits are not very stable. This is one reason this experiment was not attempted. Furthermore, the W14 residue is located at the side of the molecule away from the $\alpha\beta$ interface so that its chemical shielding should not be affected regardless of its interaction with the β-chain, assuming that the structure of the α-subunit is not significantly perturbed when the α- and β-subunits form Hb tetramers. If this assumption is not valid, then all the previously reported kinetic measurements on the α- and β-subunits would also be invalid.

On the basis of this assumption, we predicted that there would one major peak in the α_4 spectrum. A major peak was found, at ∼66 ppm, which corresponded almost exactly with one of the WT HbCO peaks (Figure 4A,B). In WT HbCO, βY130F HbCO, HbO_2, and Hb (see below), the chemical shift range was from 66.2 to 66.3 ppm (Table 1), essentially that seen in the α_4 tetramer, strongly supporting an assignment to α14. Moreover, this assignment is also consistent with the most upfield peak in MbCO being assigned to the homologous W14, as discussed above.

Since the crystallographic structures of monomeric α- or β-subunits are not known, further studies of the interactions between α- and β-subunits were not pursued, since (in the absence of *p*-mercuribenzoate) neither spectra nor spectral predictions can be obtained.

Similar ^{19}F NMR spectra were also obtained for wild type oxy- and deoxyhemoglobins and for βY130F mutant carbonmonoxy- and βY130F mutant deoxyhemoglobins. HbCO assignments were sufficient to assign HbO_2 spectra, for the spectra are very similar and the ligated proteins are assumed to be in the same state, either R or R2. WT HbCO assignments were also assumed to be sufficient in assigning the spectra for the βY103F mutant, since the only significant difference between the two spectra was the disappearance of the denatured peak in the CO-ligated βY130F mutant spectrum, and a 0.4 ppm change in the position of the peak assigned to Trp β15, as can be seen in Figure 4D. Additionally, since HbCO may exist in either the R (high-salt) or R2 (low-salt) states, we measured the ^{19}F NMR spectra of the CO-ligated βY130F mutant in solutions containing zero (low-salt), 1 M, 2 M, and 3 M (high-salt) ammonium sulfate. We found no significant variation between the spectra (not shown), suggesting the presence of a single major conformer. This point is discussed in more detail below.

It is important to note here that salt concentrations *per se* are not expected to influence ^{19}F chemical shifts directly, that is to say by modifying the local charge field, for a given conformation. Unlike e.g. histidine residues (whose pK_a values and hence chemical shifts are expected to shift with ionic strength), tryptophan resiudes do not titrate or have measurable pK_as. Their ^{19}F shifts are therefore dominated by through-space polarization effects due to structural changes, rather than to changes in electronic structure due to protonation. General support for this picture includes e.g. the lack of pH titration effects in lysozyme in the pH range of ∼2−9 and the observation that removal of surface charges on lysines via acetylation causes <1 ppm shift changes (Lian et al., 1994).

In the case of deoxyHb, a single upfield resonance in the deoxyHb spectrum was found to be at essentially the same frequency as α14 in HbCO and is therefore attributed to α14. A comparison of spectra of WT deoxyHb and βW15F deoxyHb was then used to differentiate β15 and β37, with the deleted peak again being attributed to β15 (spectra not shown). The spectrum of the deoxy βY130F mutant was found to be similar to that of the wild type, including the presence of a denatured peak (spectra not shown). The denatured peak disappeared when CO was re-introduced to the deoxy βY130F sample, and the original CO-ligated spectrum was obtained. A compilation of all of the chemical shift assignments is given in Table 1.

Line Width Considerations. The ^{19}F line widths we observe are much broader than typical ^{1}H line widths, especially for hemoglobins. This is because the large ^{19}F chemical shift anisotropy (CSA) and long τ_c for hemoglobin results in a large CSA contribution to the line width at high field. For example, many of the Hb lines are ≈200 Hz, much greater than in ^{1}H NMR. This is because the ^{19}F CSA is ∼200 ppm and not ∼20 ppm (as for ^{1}H), so the relaxation contribution in $(200/20)^2$ or ∼100 × larger for ^{19}F. This necessitated relatively long acquisition times for achieving acceptable *S*/*N* ratios. In hemoglobin (Figure 4A), there are small differential line broadenings observed with β37W and β15W being clearly broader than the signal attributed to α14W. These differential line widths can arise from a number of mechanisms, including dipolar and CSA broadening, conformational averaging, or a distribution of conformational substates. At present, we cannot quantify these various contributions accurately; however, the effects appear to be real, and the narrow α14W line width may be associated with less disorder at this position since, as shown below, its shift is less scattered in the chemical shift calculations.

Kinetic Measurements. Kinetic measurements were carried out in order to check for functional perturbations of the proteins upon [4-F]tryptophan incorporation or mutation. The CO association rates for labeled and unlabeled proteins were all found to be approximately 1×10^5 M^{-1} s^{-1} for Hb and 5×10^5 M^{-1} s^{-1} for Mb, indicating that mutations and fluorine incorporation did not significantly alter the function of the proteins.

A note must be made at this point regarding the βY130F mutant. The ^{19}F NMR spectra and the stopped-flow kinetics measured for this mutant thus far seem to suggest that this mutant might be closer to native human hemoglobin A (HbA) than is the recombinant Hb and are also consistent with the suggestion of Hernan et al. (1992) which stated that the recombinant human hemoglobin expressed in bacteria is either partially misassembled or misfolded. Our results however suggest that the "misfolded" fraction of Hb is markedly smaller or nonexistent in βY130F mutant Hb. Examination of the hemoglobin crystal structure (1hho) indicates that the hydroxyl of the Tyr β130 residue in the H-helix is hydrogen bonded to the carbonyl backbone of Val β11 in the A-helix. It is therefore possible that Tyr β130 disrupts the packing between the A- and the H-helix during protein folding. A detailed characterization of the Tyr β130 mutant will be reported in a separate paper.

Chemical Shift Calculations. We now consider the origins of the chemical shift nonequivalencies due to folding in myoglobin and hemoglobin. Here, one possible complication of using fluoroaromatic residues as probes of protein structure is the generation of local conformational changes due to incorporation of the fluorine atom (Pratt & Ho, 1975). Since substitution of an aromatic hydrogen atom by fluorine significantly increases the C−X bond length (by ∼0.3 Å) and the van der Waals radius at that site (by ∼0.2 Å), it is certainly conceivable that a [4-F]Trp-labeled protein may exhibit local structure around the tryptophan sites different from that found in the unlabeled protein, due to the introduction of van der Waals overlaps of up to 0.5 Å in the native structures. The van der Waals overlap of the 4 position of each tryptophan, both before and after insertion of a fluorine atom, was therefore evaluated at each site using an AMBER force field (Weiner et al., 1986) within Insight II (Biosym Technologies, San Diego, CA), with the results shown in Table 2. As can be seen, while none of the native $H^{\epsilon 3}$ sites have a significant overlap with surrounding atoms, several of the fluorinated $F^{\epsilon 3}$ sites do have a significant overlap with neighboring atoms. Examining individual structures, we find none of the [4-F]Trp fluorine atoms of sperm whale myoglobin (1mbc or 1mbd) structures have significant overlaps. For [4-F]Trp wild type hemoglobin, $F^{\epsilon 3}$ of [4-F]Trp α14 has a significant overlap with H^{α} of the same residue in 1hho, but no significant overlap in the other structures; $F^{\epsilon 3}$ of [4-F]Trp β15 has significant overlap with the O^{η} of Tyr β130 in all of the structures, as shown for example in Figure 5, and $F^{\epsilon 3}$ of [4-F]Trp β37 has significant overlap in the 1hho (R-state) and 2hhb (T-state) structures, as shown for example in Figure 6. Here, we show the backbone of the $\alpha_1\beta_2$-subunits from both R and R2 structures. The backbones are superimposable for almost all residues. The main difference is observed at the C terminus of the α-subunit, near the Tyr α140 residue. In the R structure, the Tyr is very close to βW37, while in the R2 structure, it is not. In the R structure, the Tyr α140 is so close to the βW37 in the β2 subunit that there is van der Waals contact, when the H at the fourth position is replaced by F, which we believe is one of the reasons why the calculated chemical shielding does not agree well with the observed chemical shift.

Table 2: van der Waals Overlap of Trp $H^{\epsilon 3}/F^{\epsilon 3}$ Sites in Native and [4-F]Trp Sperm Whale Myoglobin, R-State Hemoglobin, T-State Hemoglobin, and R2-State Hemoglobin

		vdW overlap[b]		
structure[a]	residue	native	[4-F]Trp	neighboring atom[c]
1mbc	Trp 7	none	none	
1mbc	Trp 14	none	none	
1mbd	Trp 7	none	none	
1mbd	Trp 14	none	none	
1hho (R)	Trp α14	none	0.07	α14 $H^{\alpha 1}$
1hho (R)	Trp β15	none	0.20	β130 $O^{\eta 1}$
1hho (R)	Trp β37	none	0.18	α140 $C^{\epsilon 2}$
2hhb (T)	Trp $\alpha_1$14	none	none	
2hhb (T)	Trp $\beta_1$15	none	0.16	$\beta_1$130 $O^{\eta 1}$
2hhb (T)	Trp $\beta_1$37	none	0.14	$b_1$37 H^{N}
2hhb (T)	Trp $\alpha_2$14	none	none	
2hhb (T)	Trp $\beta_2$15	none	0.20	$\beta_1$130 $O^{\eta 1}$
2hhb (T)	Trp $\beta_2$37	none	0.04	$\alpha_1$140 Hβ2
1bbb (R2)	Trp $\alpha_1$14	none	none	
1bbb (R2)	Trp $\beta_1$15	none	0.16	$\beta_1$130 $O^{\eta 1}$
1bbb (R2)	Trp $\beta_1$37	none	none	
1bbb (R2)	Trp $\alpha_2$14	none	none	
1bbb (R2)	Trp $\beta_2$15	none	0.14	$\beta_2$130 $O^{\eta 1}$
1bbb (R2)	Trp $\beta_2$37	none	none	
m1bbb (mR2)	Trp $\alpha_1$14	none	none	
m1bbb (mR2)	Trp $\beta_1$15	none	none	
m1bbb (mR2)	Trp $\beta_1$37	none	none	
m1bbb (mR2)	Trp $\alpha_2$14	none	none	
m1bbb (mR2)	Trp $\beta_2$15	none	none	
m1bbb (mR2)	Trp $\beta_2$37	none	none	

[a] The four-character code identifies the structure in the Protein Data Bank (PDB) of the Brookhaven National Laboratory (Abola et al., 1987, Bernstein et al., 1977). m1bbb indicates a mutant 1bbb structure, with Tyr β130 replaced with Phe. The corresponding R, T, R2, and modified R2 (mR2) structure designations are given after the PDB file number, in parentheses. [b] van der Waals overlap was calculated as $^1/_2$ of the difference between the internuclear distance and the sum of van der Waals radii found in an AMBER force field. [c] Atom with the largest overlap with the fluorine electron cloud. The fluorine in [4-F]Trp β37 of 1hho also overlaps with several atoms in the Tyr α140 ring.

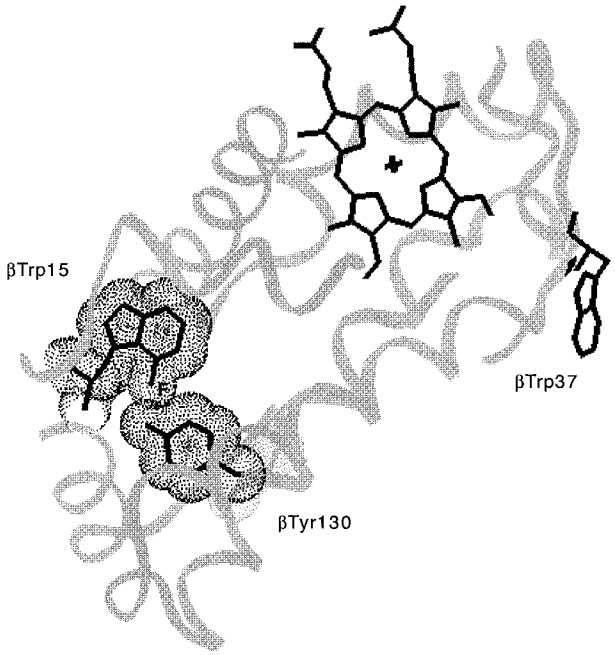

FIGURE 5: Partial representation of the R-state structure using PDB file 1hho, showing the vdW overlap between $F^{\epsilon 3}$ of $\beta 15$ and O^{η} of Tyr $\beta 130$.

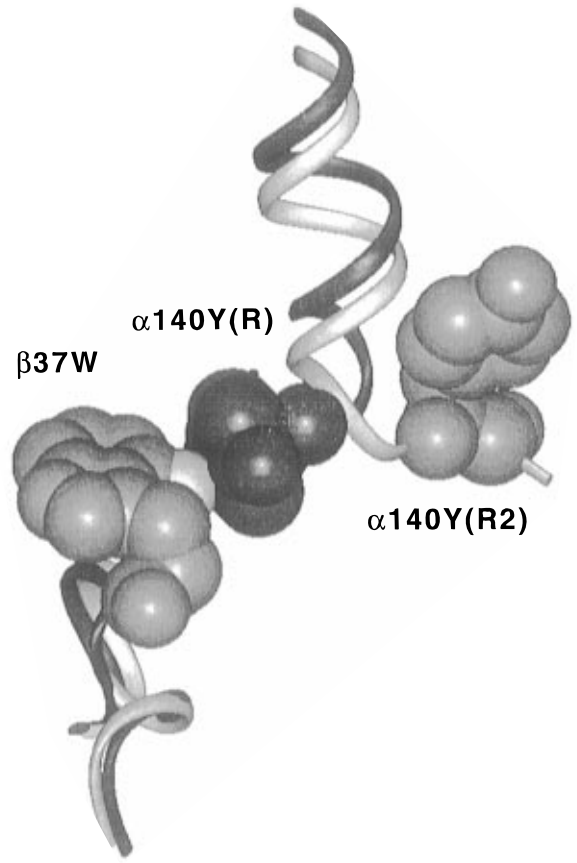

FIGURE 6: Partial representation of the R-state and R2-state structures using PDB files 1hho and 1bbb. Structures are superimposed at the β37W residue, with the backbone ribbon and residue α140Y in structure 1hho (R) shown in black and the backbone ribbon and residues β37W and α140Y of structure 1bbb (R2) shown in gray. Hydrogens are not shown. Incorporation of [4-F]Trp results in insertion of a fluorine atom directly between residues β37W and α140Y in the 1hho (R) structure, generating a steric contact.

In the 1bbb structure, the overlap between the side chain hydroxyl group of Tyr $\beta 130$ and $F^{\epsilon 3}$ of [4-F]Trp $\beta 15$ is the only one present, so the βY130F mutant of 1bbb yields a hemoglobin structure with no significant overlaps through use of a fairly conservative mutation. The structure of the resulting fluorinated mutant protein is thus closely related to the structure of the unlabeled wild type protein, with the tyrosine hydroxyl group exchanged for an aromatic (tryptophan) fluorine group occupying roughly the same space. We now discuss the shieldings computed, bearing in mind the overlaps discussed above due to [4-F]Trp incorporation.

We show in Table 3 the average values for the electrostatic field and field gradient contributions to the chemical shift for each of the ^{19}F sites in carbonmonoxymyoglobin, deoxymyoglobin, R-state hemoglobin, T-state hemoglobin, R2-state hemoglobin, and R2-state βY130F mutant hemoglobin, as calculated using fluorinated PDB structures in ENZYMIX and the MSP−LRF method described above. During these calculations, we noted that the trajectories for the $\beta_2 15$ site in βY130F 1bbb indicated the presence of an isolated solvent water in the hydrophobic pocket immediately adjacent to the fluorine atom. Close examination of the actual trajectory structures confirmed the presence of this water molecule. Both the lack of radical change in the ^{19}F spectra and the lack of crystallographic water at this position in all hemoglobin X-ray structures led us to suspect that this might be an artifact of the MD simulation. These suspicions were confirmed upon observing that extension of the trajectories caused the water molecule to be expelled from the pocket and that calculation of similar trajectories differing only in random number seed did not reproduce the presence of a water molecule at this location. Therefore, the results reported in Table 3 for the $\beta_2 15$ site in βY130F 1bbb are taken from 20 ps trajectories following the expulsion of this single water molecule. The magnetic anisotropy and ring current contributions to chemical shielding at each site were calculated using the SHIFTS program of Case (1994), with values obtained also being reported in Table 3.

We now examine the correlation between the experimentally observed chemical shifts and those evaluated from the X-ray structures using the MSP−LRF method. When examining all structures (Figure 7A), we find that there is very little correlation between theory and experiment, in sharp contrast with our previous results on galactose binding protein (GBP) (Pearson et al., 1993). The reasons for this lack of correlation appear to be that 9 of the 25 points shown in Figure 7A (●) originate from fluorine sites which have steric interactions with adjacent residues, as shown in Table 2 and Figures 5 and 6. Such sites can be expected to undergo structural changes to relax the repulsive interactions, but since the nature of the structural changes is not known, it is not possible to accurately evaluate the electric fields and field gradients for these specific sites. Removal of all points having overlaps, as shown in Table 2, results in a much improved correlation between experiment and prediction, as shown in Figure 7B. We find an R^2 value of 0.78, a slope of −0.67, and an rmsd between theory and experiment of 0.9 ppm, close to but clearly worse than that reported previously for GBP (0.93, −0.93, and 1.0 ppm) (Pearson et al., 1993), most likely due to the fact that the hemoglobin structures are less accurate than those of the smaller GBP molecule. Since the accuracy of a MSP−LRF calculation is largely dependent upon the accuracy of the starting structure used, any inaccuracies in the initial structure will translate to poor shift predictions. For proteins the size of

Table 3: Experimental ^{19}F Chemical Shifts and Calculated ^{19}F Chemical Shieldings of [4-F]Trp Sperm Whale Myoglobins, R-State Hemoglobin, T-State Hemoglobin, and R2-State Hemoglobin

protein	ligand	residue	δ (experimental)[a]	structure[b]	σ(E)[c]	$\sigma(\nabla\mathbf{E})$[c]	σ(M)[d]	$\sigma(\mathbf{E}+\mathrm{M})$[e]
Mb	CO/O_2	Trp 7	−60.0/−60.0	1mbc	5.6	−0.7	0.3	5.2
Mb	CO/O_2	Trp 14	−64.4/−64.2	1mbc	9.6	−2.0	0.8	8.4
Mb	none	Trp 7	−59.8	1mbd	4.1	0.4	0.3	4.8
Mb	none	Trp 14	−64.4	1mbd	5.4	0.3	0.8	6.5
Hb	CO/O_2	Trp α14	−66.2/−66.3	1hho (R)[f]	6.9	1.6	0.2	8.7
Hb	CO/O_2	Trp β15	−62.9/−62.9	1hho (R)[f]	7.9	−1.0	−0.1	6.7
Hb	CO/O_2	Trp β37	−61.4/−61.5	1hho (R)[f]	−1.1	11.6	1.6	12.2
Hb	none	Trp $\alpha_1$14	−66.3	2hhb (T)	6.1	1.7	0.0	7.8
Hb	none	Trp $\beta_1$15	−61.4	2hhb (T)[f]	9.3	−1.0	−0.3	8.0
Hb	none	Trp $\beta_1$37	−59.9	2hhb (T)[f]	−1.8	7.7	−0.2	5.6
Hb	none	Trp $\alpha_2$14	−66.3	2hhb (T)	5.9	2.6	0.0	7.5
Hb	none	Trp $\beta_2$15	−61.4	2hhb (T)[f]	8.4	−1.0	−0.3	7.0
Hb	none	Trp $\beta_2$37	−59.9	2hhb (T)[f]	0.9	9.2	−0.2	9.8
Hb	CO/O_2	Trp $\alpha_1$14	−66.2/−66.3	1bbb (R2)	10.6	−1.0	−0.1	9.6
Hb	CO/O_2	Trp $\beta_1$15	−62.9/−62.9	1bbb (R2)[f]	11.3	−1.6	−0.1	9.6
Hb	CO/O_2	Trp $\beta_1$37	−61.4/−61.5	1bbb (R2)	4.5	1.2	−0.1	5.6
Hb	CO/O_2	Trp $\alpha_2$14	−66.2/−66.3	1bbb (R2)	13.3	−3.2	−0.1	9.9
Hb	CO/O_2	Trp $\beta_2$15	−62.9/−62.9	1bbb (R2)[f]	13.1	−4.2	−0.2	8.8
Hb	CO/O_2	Trp $\beta_2$37	−61.4/−61.5	1bbb (R2)	3.3	0.9	−0.1	4.1
Hb	CO	Trp $\alpha_1$14	−66.2	m1bbb (mR2)	8.0	0.6	−0.1	8.6
Hb	CO	Trp $\beta_1$15	−63.3	m1bbb (mR2)	5.1	1.6	−0.1	6.6
Hb	CO	Trp $\beta_1$37	−61.4	m1bbb (mR2)	3.2	0.2	−0.1	3.3
Hb	CO	Trp $\alpha_2$14	−66.2	m1bbb (mR2)	9.7	−1.5	−0.1	8.1
Hb	CO	Trp $\beta_2$15	−63.3	m1bbb (mR2)	5.1	2.2	−0.2	7.1
Hb	CO	Trp $\beta_2$37	−61.4	m1bbb (mR2)	3.0	1.5	−0.1	4.4

[a] Experimental chemical shifts are reported in parts per million from the ^{19}F resonance observed in neat, external trifluoroacetic acid. [b] The four-character code identifies the structure in the Protein Data Bank of the Brookhaven National Laboratory (Abola et al., 1987; Bernstein et al., 1977). m1bbb indicates the 1bbb structure with Tyr β130 replaced with Phe. The corresponding R, T, R2, and modified R2 (mR2) structure designations are given after the PDB file number, in parentheses. [c] Electrostatic field and gradient components of shielding, reported in parts per million. Zero shielding represents no *electrostatic* contribution to the chemical shielding. Each value is averaged over six 20 ps trajectories for structures 1mbc, 1mbd, 1hho, 2hbb, and 1bbb and twenty-four 20 ps trajectories for structure m1bbb. [d] Magnetic component of shielding, reported in parts per million. Zero shielding represents no *magnetic* contribution to the chemical shielding. [e] Total electrostatic and magnetic anisotropy contributions to shielding. [f] This structure contains a van der Waals overlap upon incorporation of [4-F]Trp at this position.

hemoglobin and with crystallographic resolutions of ≈2 Å, the pooled standard deviations for ϕ,ψ and χ can be expected to be ∼15−20° (Morris et al., 1992), making shift predictions a challenging proposition. This, together with the fact that the dipole and quadrupole shielding polarizabilities are unlikely to be accurate to better than 20−30% (Augspurger et al., 1992), places, we believe, an error of ∼1−2 ppm on individual shift calculations, about twice that observed previously with the galactose binding protein. Given that it is also not at present possible to predict accurate absolute shieldings in a protein, it is important to view both the patterns of shielding in hemoglobin and the myoglobin shift calculations (Figure 7B) when drawing structural conclusions. Given that, our results for HbCO are summarized in Table 4, and clearly indicate that the R2-state HbCO structure yields predicted shifts in best accord with the ligated hemoglobin shifts. The generally good agreement between the R2-state crystal structure and the experimental shifts on the one hand and the generally poor agreement between the deoxy and R-state crystal structures and the experimental shifts (Tables 3 and 4) on the other therefore strongly suggest the existence, at least on a local scale, of [4-F]Trp-labeled carbonmonoxyhemoglobin in the R2 state.

As can be seen from Table 4, the ^{19}F chemical shifts of both β15 and β37 are significantly different between CO-liganded and deoxy states, as might be expected due to their different conformations. With the change of βY130 to Phe to remove steric interactions due to ^{19}F, and by using the R2 structure, the calculated chemical shieldings are in good accord with experiment, with quite clearly the R2-state CO structure being the only one to give not only the correct ordering but also the correct overall chemical shift range. In addition, this structure also yields computed shieldings *which lie on the Mb correlation* (Figure 7B). Using the R2 structure, we therefore find the following. First, there is good accord with the experimental shifts for βY130F HbCO (all three sites). Second, these points fall on the same curve as the Mb data, where the X-ray structures can be expected to be significantly more accurate.

The conclusion from Tables 3 and 4 and Figures 6 and 7 is therefore that only the R2 structure give shifts in close accord with experiment. For residues which do not have F atoms in contact with neighboring groups, we find an ∼1 ppm mean square deviation in shift from experiment, with the R2-like structure of [4-F]Trp HbCOA giving results in very close accord with experiment. Of course, the possibility does exist that the presence of the ^{19}F labels might stabilize an R2-like structure over an R structure, in solution. Indeed, the lack of any major chemical shift changes on transition from no salt to high salt would support this idea, since salt induces an observable shift in the C-terminal histidine pK_a values [see e.g. Busch et al. (1991) and Kilmartin et al. (1973) for discussions] which has been attributed to a major conformational change in the interfacial region (Silva et al., 1992). That is, while the structure of [4-F]Trp HbCO is indeed R2-like in solution, it does not necessarily follow that the native unlabeled protein in solution would be entirely R2. However, since the structural perturbations induced by F substitution can be expected to be small, on the order of 1−2 kcal, then our results would support the presence of a

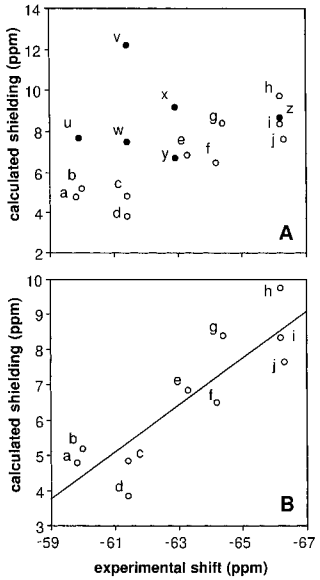

FIGURE 7: Comparison of calculated electrostatic field contributions to chemical shieldings with experimental chemical shifts for carbonmonoxymyoglobin, deoxymyoglobin, R-state hemoglobin, T-state hemoglobin, and R2-state hemoglobin, using data from (A) all 4-fluorotryptophan sites, both sterically hindered (●) and unhindered (○) (from Table 2): (a) W7 of 1mbd, (b) W7 of 1mbc, (c) βW37 of 1bbb, (d) βW37 of βY130F 1bbb, (e) βW15 of βY130F 1bbb, (f) W14 of lmbd, (g) W14 of lmbc, (h) αW14 of 1bbb, (i) αW14 of βY130F 1bbb, (j) αW14 of 2hhb, (u) βW37 of 2hhb, (v) βW37 of 1hho, (w) βW15 of 2hhb, (x) βW15 of 1bbb, (y) βW15 of 1hho, and (z) αW14 of 1hho. (B) Unhindered fluorotryptophan sites. Linear regression analysis of unhindered sites yields an R^2 of 0.78, a slope of −0.67 (where 1.0 and −1.0 would be ideal), and an rmsd of 0.9 ppm.

Table 4: Summary of Experimental Shifts and Calculated Electrostatic Contribution to Shielding for Hb and HbCO Using R, R2, and βY130F Structures for HbCO

ligand form	residue	experimental shift (ppm)[a]	calculated shielding (ppm)[b]	structure[c]
	α14	−66.3	7.8, 7.5	
deoxy	β15	−61.4	8.0, 7.0	T
	β37	−59.9	5.6, 9.8	
		(α14 > β15 > β37)	(no clear order)	
	α14	−66.2, −66.3	8.7	
CO	β15	−62.9	6.7	R
	β37	−61.4, −61.5	12.2	
		(α14 > β15 > β37)	(wrong order)	
	α14	−66.2, −66.3	9.6, 9.9	
CO	β15	−62.9	9.6, 8.8	R2
	β37	−61.4, −61.5	5.6, 4.1	
		(α14 > β15 > β37)	(α14 ≥ β15 > β37)	
	α14	−66.2	8.6, 8.1	
CO	β15	−63.3	6.6, 7.1	R2 with βY130F
	β37	−61.4	3.3, 4.4	
		(α14 > β15 > β37)	(α14 > β15 > β37)	

[a] Experimental chemical *shift* measured in parts per million from an external standard of trifluoroacetic acid. [b] Computed electrostatic contributions to *shielding*. Note that the shift (δ) is proportional to the negative shielding ($-\sigma$). [c] The structure files used for the T, R, R2, and modified R2 structures are as in the footnotes to Tables 2 and 3, as discussed in the text.

significant population of R2-like conformers for native HbCO in solution as well.

CONCLUSIONS

The results we have presented above indicate that it is possible to compute ^{19}F chemical shifts in heme proteins and that these shifts are dominated by the electrostatic field contribution to shielding, as noted previously for the galactose binding protein from *E. coli* (Pearson et al., 1993). We find overall an R^2 of 0.78, a slope of −0.67, and a rmsd versus experiment for MbCO, deoxyMb, HbCO, and deoxyHb of 0.9 ppm, but only when ^{19}F sites having steric overlaps with adjacent residues are omitted. Related steric effects were noted by us previously in a more extreme case, that of [F]Phe labeling of hen egg white lysozyme (Lian et al., 1994). There, we observed that *o*-fluorine substitution of phenylalanine residues caused a peak doubling of all three Phe sites [not seen with *m*- or *p*-Fluorine labeling (Lian et al., 1994)], confirming that un-natural steric interactions due to ^{19}F labeling can alter local structure. Thus, while ^{19}F NMR can be a useful tool for probing protein structure and electrostatics, the effects of fluorine incorporation need to be very carefully assessed.

Overall, our results support the hypothesis that [4-F]Trp HbCO exists in the R2 state, for the following reasons. First, simulated ^{19}F chemical shieldings based upon the R2 structure give best accord with experimental shifts. Second, the lack of disordered protein in βY130 FHbCO (no steric congestion of Trp β15 due to the OH → H mutation), coupled with the reversible appearance of disordered protein upon removal of the CO ligand, indicates that the ^{19}F sites in the carbonmonoxy mutant are free of significant steric overlaps. This supports the idea of an R2 structure, because only the R2 structure is free of such interactions. Third, the similarity of resonance frequencies for the ^{19}F peaks assigned to specific residues for HbCO and βY130F HbCO indicates that, as with the mutant, [4-FTrp] HbCO is also in the R2 state.

ACKNOWLEDGMENT

We thank Arthur Arnone for advice concerning crystal structures and both referees for their constructive criticisms.

REFERENCES

Abadan, Y., Chien, E. Y. T., Chu, K., Eng, C. D., Nienhaus, G. U., & Sligar, S. G. (1995) *Biophys. J. 68*, 2497−2504.

Abola, E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) Protein Data Bank, in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) p 107, Data Commission of the International Union of Crystallography, Bonn, Cambridge, and Chester.

Arnone, A., & Silva, M. M. (1992) *J. Biol. Chem. 267*, 17248−17256.

Augspurger, J. D., & Dykstra, C. E. (1991) *J. Phys. Chem. 95*, 9230−9238.

Augspurger, J. D., Dykstra, C. E., & Oldfield, E. (1991) *J. Am. Chem. Soc. 113*, 2447−2451.

Augspurger, J. D., Pearson, J., Oldfield, E., Dykstra, C. E., Park, K. D., & Schwartz, D. (1992) *J. Magn. Reson. 100*, 342−357.

Batchelor, J. G. (1975) *J. Am. Chem. Soc. 97*, 3410−3415.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, J. E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Bode, R., Kunze, G., & Birnbaum, D. (1985) *Biochem. Physiol. Pflanzen. 180*, 613−619.

Bode, R., Schauer, F., & Birnbaum, D. (1986) *Biochem. Physiol. Pflanzen. 181*, 39−46.

Browne, D. T., & Otvos, J. D. (1976) *Biochem. Biophys. Res. Commun. 68*, 907−913.

Bucci, E., & Fronticelli, C. (1965) *J. Biol. Chem. 240*, PC55.

Buckingham, A. D. (1960) *Can. J. Chem. 38*, 300−307.

Buckingham, A. D., & Lawley, K. P. (1960) *Mol. Phys. 3*, 219−222.

Busch, M. R., Mace, J. E., Ho, N. T., & Ho, C. (1991) *Biochemistry 30*, 1863−1877.
Case, D. A., Dyson, H. J., & Wright, P. E. (1994) *Methods Enzymol. 239*, 392−416.
Chaiken, I. M., Freedman, M. H., Lyerla, J. R. J., & Cohen, J. S. (1973) *J. Biol. Chem. 248*, 884−891.
Chiu, M. L. (1986) Ph.D. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.
de Dios, A. C., & Oldfield, E. (1994) *J. Am. Chem. Soc. 116*, 5307−5314.
de Dios, A. C., Pearson, J. G., & Oldfield, E. (1993) *Science 260*, 1491−1496.
Falke, J. J., Luck, L. A., & Scherrer, J. (1992) *Biophys. J. 62*, 82−86.
Fermi, G., Perutz, M. F., Shaanan, B., & Fourme, R. (1984) *J. Mol. Biol. 175*, 159−174.
Gamcsik, M. P., & Gerig, J. T. (1986) *FEBS Lett. 196*, 71−74.
Gamcsik, M. P., Gerig, J. T., & Swenson, R. B. (1986) *Biochim. Biophys. Acta 874*, 372−374.
Gamcsik, M. P., Gerig, J. T., & Gregory, D. H. (1987) *Biochim. Biophys. Acta 912*, 303−316.
Gerig, J. T., Klinkenborg, J. C., & Nieman, R. A. (1983) *Biochemistry 22*, 2076−2087.
Hernan, R. A., & Sligar, S. G. (1995) *J. Biol. Chem. 270*, 26257−26264.
Hernan, R. A., Hui, H. L., Andracki, M. E., Noble, R. W., Sligar, S. G., Walder, J. A., & Walder, R. Y. (1992) *Biochemistry 31*, 8619−8628.
Ho, C., Pratt, E. A., & Rule, G. S. (1989) *Biochim. Biophys. Acta 988*, 173−184.
Hull, W. E., & Sykes, B. D. (1975) *J. Mol. Biol. 98*, 121−153.
Hull, W. E., & Sykes, B. D. (1976) *Biochemistry 15*, 1535−1546.
Imai, K. (1981) *Methods Enzymol. 76*, 438−449.
Kilmartin, J. V., Breen, J. J., Roberts, G. C. K., & Ho, C. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 1246−1249.
Kimber, B. J., Griffiths, D. V., Birdsall, B., King, R. W., Scudder, P., Feeney, J., Roberts, G. C. K., & Burgen, A. S. V. (1977) *Biochemistry 16*, 3492−3500.
Kimber, B. J., Feeney, J., Roberts, G. C. K., Birdsall, B., Griffiths, D. V., & Burgen, A. S. V. (1978) *Nature 271*, 184−185.
Kunze, G., Bode, R., Rintala, H., & Hofemeister, J. (1989) *Curr. Genet. 15*, 91−98.
Kuriyan, J., Wilz, S., Karplus, M., & Petsko, G. A. (1986) *J. Mol. Biol. 192*, 133−154.
Lee, F. S., & Warshel, A. (1992) *J. Chem. Phys. 97*, 3100−3107.
Lian, C., Le, H., Montez, B., Patterson, J., Harrell, S., Laws, D., Matsumura, I., Pearson, J., & Oldfield, E. (1994) *Biochemistry 33*, 5238−5245.
Luck, L. A., & Falke, J. J. (1991a) *Biochemistry 30*, 6484−6490.
Luck, L. A., & Falke, J. J. (1991b) *Biochemistry 30*, 4248−4256.
Luck, L. A., & Falke, J. J. (1991c) *Biochemistry 30*, 4257−4261.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992) *Proteins 12*, 345−364.
Pearson, J. G., Oldfield, E., Lee, F. S., & Warshel, A. (1993) *J. Am. Chem. Soc. 115*, 6851−6862.
Peersen, O. B., Pratt, E. A., Truong, H.-T. N., Ho, C., & Rule, G. S. (1990) *Biochemistry 29*, 3256−3262.
Phillips, G. N., Jr., Arduini, R. M., Springer, B. A., & Sligar, S. G. (1990) *Proteins: Struct., Funct., Genet. 7*, 358−365.
Phillips, S. E. V., & Schoenborn, B. P. (1981) *Nature 292*, 81−82.
Pratt, E. A., & Ho, C. (1975) *Biochemistry 14*, 3035−3040.
Robertson, D. E., Kroon, P. A., & Ho, C. (1977) *Biochemistry 16*, 1443−1451.
Rohlfs, R. J., Mathews, A. J., Carver, T. E., Olson, J. S., Springer, B. A., Egeberg, K. D., & Sligar, S. G. (1990) *J. Biol. Chem. 265*, 3168−3176.
Shaanan, B. (1983) *J. Mol. Biol. 171*, 31−59.
Silva, M. M., Rogers, P. H., & Arnone, A. (1992) *J. Biol. Chem. 267*, 17248−17256.
Smith, L. J., Redfield, C., Smith, R. A. G., Dobson, C. M., Clore, G. M., Gronenborn, A. M., Walter, M. R., Naganbushan, T. L., & Wlodawer, A. (1994) *Nat. Struct. Biol. 1*, 301−310.
Springer, B. A., & Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8961−8965.
Stephen, M. J. (1957) *Mol. Phys. 1*, 223−232.
Taylor, H. C., Richardson, D. C., Richardson, J. S., Wlodawer, A., Komoriya, A., & Chaiken, I. M. (1981) *J. Mol. Biol. 149*, 313−317.
Warshel, A., & Creighton, S. (1989) *Computer Simulation of Biomolecular Systems* (van Gunsteren, W. F., & Weiner, P. K., Eds.) pp 120−138, ESCOM Science Publishers, Leiden.
Weiner, S. J., Kollman, P. A., Hguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230−252.

BI961664H