Determination of the Zero-Field Splitting Constant for Proton NMR Chemical Shift Analysis in Metaquomyoglobin. The Dipolar Shift as a Structural Probe

Yung-Hsiang Kao and Juliette T. J. Lecomte*

Contribution from the Department of Chemistry and Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802

Received May 26, 1993®

Abstract: Two-dimensional ^{1}H NMR experiments were carried out on sperm whale metaquomyoglobin (paramagnetic complex, $S = {}^5/_2$) and carbonmonoxymyoglobin (diamagnetic complex) in order to determine the zero-field splitting constant (D) of the Fe(III) atom in metaquomyoglobin. The paramagnetic shift contains contact and dipolar contributions. The former is negligible for the protons included in the study while the latter is directly proportional to the geometric factor and D, and can be estimated by taking the difference of the chemical shifts in the two myoglobin complexes. The value so obtained is accurate if the carbonmonoxymyoglobin shift is an adequate diamagnetic reference for metaquomyoglobin. In order to determine D, 37 backbone resonances were selected for which this assumption was likely to be closely satisfied. The selection was achieved by identifying protons with a constant ring current shift as calculated from the solid-state structures of metaquomyoglobin and carbonmonoxymyoglobin. With this restricted set of dipolar shifts and the geometric factors calculated from the solid state coordinates, D was found to be 9.66 ± 0.08 cm^{-1}. This constant was then used to evaluate the dipolar contribution to the chemical shift of the other 104 assigned protons. Comparison of the predicted and the observed values provided a description of the structural alterations occurring upon change in complexation and dissolution. In nearly 80% of the cases, good agreement was observed, which indicated that the X-ray structure is the same as the solution structure and that the dipolar shift can be used to confirm spectral assignments. However, discrepancies were noted in the A–G–H and G–H interfaces and in turns. For several regions, the structural perturbations could be rationalized with features specific to the metaquomyoglobin solid-state structure, such as proximity to an ion binding site and unallowed steric interactions. The dipolar shift was demonstrated to be a sensitive probe for the elucidation of the conformational properties of metaquomyoglobin in solution and for a comparison to the solid-state structure.

Introduction

Metaquomyoglobin (metMbH_2O)[1] is a nonfunctional form of the oxygen-storage protein with the iron atom in the +3 oxidation state and water as the sixth ligand. MetMbH_2O from most mammalian sources is stable over a wide range of salt concentrations, pH, and temperatures. These thermodynamic properties make it an interesting protein for the experimental characterization of electrostatic interactions. The protein extracted from sperm whale skeletal muscles contains 12 histidines out of 153 residues. The unusually high histidine content and the interest in monitoring the ionization behavior of individual imidazole groups have prompted numerous NMR studies of myoglobins.[2-6] To analyze completely the results of such studies and to assess the computational methods applied to the evaluation of acidity constants, it is necessary to describe the structure that myoglobin adopts in solution, in particular the environment of ion binding sites and ionizing groups.[2,4,7-9] The solid-state structure of sperm whale myoglobin has been solved in several complexation forms, including the metaquo form.[10] Solid-state and solution structures are expected to be nearly identical; however, for our purpose—that of providing an adequate basis for electrostatic calculations—minor structural perturbations such as side-chain reorientation might have nontrivial consequences. Thus, the effects of disruption of the crystal lattice and release of ions bound from the mother liquor must be inspected carefully. The present report summarizes our efforts to find a simple and sensitive NMR probe of the conformational properties of Mb under the conditions of the pH titrations.

MetMbH_2O is a paramagnetic high-spin complex ($S = {}^5/_2$) with an octahedral ligand field at the Fe(III) ion. The paramagnetism is manifested in two NMR properties: chemical shift and relaxation. The effect of the five unpaired electrons on the spin–lattice relaxation parameter depends on R^{-6} and has predominance over other mechanisms within a small range of distances.[11,12] Paramagnetic relaxation is therefore most useful in the direct environment of the iron atom where it can be exploited to arrive at and confirm assignments on the basis of distance evaluation.[13-15] In contrast, the paramagnetic chemical shift is felt relatively far away from the iron atom and reflects both

* To whom correspondence should be addressed at the Department of Chemistry, The Pennsylvania State University, University Park, PA 16802.

® Abstract published in *Advance ACS Abstracts*, October 1, 1993.

(1) Abbreviations: apoMb, apomyoglobin; des-Fe Mb, protoporphyrin IX–globin complex; 2D, two-dimensional; DIPSI, decoupling in the presence of scalar interactions; DQF-COSY, two-dimensional double-quantum-filtered correlated spectroscopy; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate; GF, geometric factor, holoMb, holomyoglobin; Mb, myoglobin; MbCO, carbonmonoxymyoglobin; metMb, ferrimyoglobin; metaquoMb, metMbH_2O, metMb^2H_2O, metaquomyoglobin; metMbCN, metcyanomyoglobin; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser spectroscopy; PDB, Protein Data Bank; RCS, ring current shift; rms, root-mean-square; TOCSY, total correlation spectroscopy; TPPI, time proportional phase incrementation; 2Q, two-quantum spectroscopy.

(2) Botelho, L. H. Ph.D. Thesis, Indiana University, 1975.
(3) Wilbur, D. J.; Allerhand, A. *J. Biol. Chem.* **1977**, *252*, 4968–4975.
(4) Botelho, L. H.; Gurd, F. R. N. *Biochemistry* **1978**, *17*, 5188–5196.
(5) Bradbury, J. H.; Carver, J. A. *Biochemistry* **1984**, *23*, 4905–4913.
(6) Cocco, M. J.; Kao, Y.-H.; Phillips, A. T.; Lecomte, J. T. J. *Biochemistry* **1992**, *31*, 6481–6491.
(7) Botelho, L. H.; Friend, S. H.; Matthew, J. B.; Lehman, L. D.; Hanania, G. I. H.; Gurd, F. R. N. *Biochemistry* **1978**, *17*, 5197–5205.
(8) Garcia-Moreno E., B.; Chen, L. X.; March, K. L.; Gurd, S. R.; Gurd, F. R. N. *J. Biol. Chem.* **1985**, *260*, 14070–14082.
(9) Friend, S. H.; Gurd, F. R. N. *Biochemistry* **1979**, *18*, 4620–4630.
(10) Takano, T. *J. Mol. Biol.* **1977**, *110*, 537–568.
(11) Solomon, I. *Phys. Rev.* **1955**, *99*, 559–565.
(12) Swift, T. J. In *NMR of Paramagnetic Molecules. Principles and Applications*; La Mar, G. N., Horrocks, W. DeW., Jr., Holm, R. H., Eds.; Academic: New York, 1973; pp 53–83.

0002-7863/93/1515-9754$04.00/0

distance and angular properties.[14,15] It is the parameter that we chose to use here.

In paramagnetic compounds, the chemical shift arises from two contributions

$$\delta = \delta_{dia} + \delta_{para} \quad (1)$$

δ_{dia} is the diamagnetic term, in which electrostatic and ring current effects are included; δ_{para} is the paramagnetic (or hyperfine) term, which is due to unpaired electron effects. This latter contribution can be partitioned into two components: contact and dipolar

$$\delta_{para} = \delta_{con} + \delta_{dip} \quad (2)$$

The contact shift is observable for nuclei onto which the unpaired electrons are delocalized, in the case of myoglobin mainly those of the heme group and of the ligated histidine (His-93, proximal). The contact contribution is steeply attenuated for more remote residues and need not be considered here. The dipolar contribution to the hyperfine shift results from the anisotropic magnetic moment caused by zero-field splitting at the metal ion. This term is given by

$$\delta_{dip} = -\frac{28g^2\beta^2(3\cos^2\theta - 1)}{9k^2T^2}\frac{D}{R^3} \quad (3)$$

where g is the electron g-factor, β the Bohr magneton, k the Boltzmann constant, T the absolute temperature, and D the zero-field splitting (ZFS) constant.[16,17] A positive δ_{dip} corresponds to a downfield shift. Equation 3 is valid in the high temperature limit when the paramagnetic metal ion has a ground state with A symmetry, which gives rise to an isotropic g tensor.[16,17] The relative coordinate information is included in the geometric factor $(3\cos^2\theta - 1)/R^3$, where R is the length of the vector connecting the point of interest to the iron atom and θ is the angle between that vector and the z-axis of the ligand field at the Fe(III) ion. The geometric factor in eq 3 demonstrates that δ_{dip} is a sensitive property of the structure if D is sufficiently large.

We show that D can be determined with experimental δ_{dip} values obtained from proton NMR data collected for the metaquo form and for a diamagnetic reference form, chosen here as the ferrous carbonmonoxy form (MbCO). This D value, in turn, can aid in the assignment of shifted resonances and, more importantly, help in identifying the regions of structure which differ in the X-ray structure and in solution. The small dipolar shift in this paramagnetic molecule with an A-type orbital ground state appears to be a promising tool for the refinement of local structure in myoglobin.

Experimental Section

Protein Samples. Sperm whale Mb (Sigma) was used without further purification. Samples for low ionic strength measurements were prepared by extensive dialysis against distilled water. To exchange labile protons against deuterons, the protein was dissolved into ^2H_2O for 24–36 h and then lyophilized; this was repeated two to three times. MetaquoMb samples for 2D NMR experiments in ^2H_2O were also prepared by reconstitution from apoMb, which was prepared and exchanged as described previously.[6] Final protein concentrations ranged between 3 and 6 mM. The pH (or pH*, uncorrected meter reading in ^2H_2O) was adjusted to 5.7 with 0.1 M ^2HCl and NaO^2H. The MbCO sample for the measurement of diamagnetic reference chemical shift was prepared using the method of Mabbutt and Wright[18] with minor changes: a Sephadex G-25 column was used to remove dithionite, and no phosphate buffer was used. During the preparation of MbCO, the protein sample was kept under a carbon monoxide atmosphere as much as possible. The protein eluted from the gel filtration column was collected directly into a centricon-10 microconcentrator and concentrated to ca. 4.5 mM. The pH was adjusted to 5.7 with 0.01 M ^2HCl.

NMR Experiments. One- and two-dimensional proton NMR spectra were acquired as previously described on a Bruker AM-500 spectrometer.[6] Standard homonuclear 2D NMR experiments were recorded in the phase-sensitive mode with the TPPI method.[19,20] These included DQF-COSY,[21] TOCSY,[22] 2Q,[23,24] and NOESY[25] with or without Hahn echo.[26] Spin locking in TOCSY experiments was achieved with a DIPSI-2 sequence;[27] the 90° spin-locking pulse was ca. 43 μs, and the mixing times were between 68 and 73 ms. NOESY mixing times were 50 and 110 ms. Sixty-four or ninety-six transients were collected for each of the 480 or 512 t_1 values. Both TOCSY and NOESY experiments were sine-modulated in the t_1 domain.[28] Two-dimensional data were processed as previously described.[6] MetMbH_2O data sets were also collected at 293, 303, 308, and 313 K to resolve spectral overlap and observe the effect of temperature on the chemical shift. Chemical shifts were referenced to DSS through H_2O at 4.81 ppm (293 K), 4.76 ppm (298 K), 4.65 ppm (308 K), and 4.60 ppm (313 K). Although the digital resolution is usually less than 4 Hz/point, the accuracy of the chemical shift is estimated to be ±0.01 ppm owing to experimental factors such as pH and temperature.

Dipolar Shift Calculations. A computer program was written to evaluate the contribution of the paramagnetic dipolar shift to the total chemical shift according to eq 3. The calculations were carried out by using the X-ray structure of sperm whale metaquoMb (PDB file 4mbn).[29] The hydrogen coordinates were generated with the molecular modeling program QUANTA (Molecular Simulations, Inc.).

A. Geometric Factor. The first step in the calculation is to determine the best plane through the heme group. This is achieved with a least-squares method[30] on the 16 atoms of the heme macrocycle. The geometric factor, $(3\cos^2\theta - 1)/R^3$, for each proton in the molecule is then calculated by using the normal to the fitted plane for the determination of θ and the iron atom coordinates for the determination of R.

B. Treatment of the Phenylalanine and Tyrosine Aromatic Rings and the Methyl Groups. Phe and Tyr rings as well as all methyl groups were taken as freely rotating at room temperature. Previous NMR studies on several complexes of Mb[31,32] demonstrate that even for the aromatic rings in contact with the heme, unhindered rotation occurs. In the dipolar shift and ring current shift calculations, methyl groups are treated as a collection of rotamers distributed according to a potential function of the form

$$V(\phi) = \left(\frac{V_0}{2}\right)\sin\left(3\phi - \frac{\pi}{2}\right) \quad (4)$$

where ϕ is the angle of rotation from the lowest energy position (staggered conformation) and V_0 is the barrier height for the rotation, taken to be 12.6 kJ/mol. The methyl proton shift is evaluated by taking the Boltzmann-weighted average of the shifts at 12 proton positions corresponding to $\phi = n\pi/6$, $n = 1, 2, 3, \ldots, 12$. A fast flipping motion between two equally weighted positions is assumed for the Tyr and Phe rings so that the shifts of δ and ϵ ring protons are the average of the shifts of two equivalent ring protons.

C. Determination of the Zero-Field Splitting Constant. Spectral assignments were obtained for metMbH_2O and for MbCO as described

(13) Unger, S. W.; Lecomte, J. T. J.; La Mar, G. N. *J. Magn. Reson.* **1985**, *64*, 521–526.
(14) Rajarathnam, K.; La Mar, G. N.; Chiu, M. L.; Sligar, S. G.; Singh, J. P.; Smith, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 7886–7892.
(15) Satterlee, J. D. *Annu. Rep. NMR Spectrosc.* **1986**, *17*, 79–179.
(16) Kotani, M. *Suppl. Progr. Theor. Phys.* **1961**, *17*, 4–13.
(17) Kurland, R. J.; McGarvey, B. R. *J. Magn. Reson.* **1970**, *2*, 286–301.
(18) Mabbutt, B. C.; Wright, P. E. *Biochim. Biophys. Acta* **1985**, *832*, 175-185.
(19) Drobny, G.; Pines, A.; Sinton, S.; Weitekamp, D. P.; Wemmer, D. *Symp. Faraday Soc.* **1979**, *13*, 49–55.
(20) Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967–974.
(21) Rance, M.; Sørensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479–485.
(22) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521–528.
(23) Braunschweiler, L.; Bodenhausen, G.; Ernst, R. R. *Mol. Phys.* **1983**, *48*, 535–560.
(24) Rance, M.; Wright, P. E. *J. Magn. Reson.* **1986**, *66*, 372–378.
(25) Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
(26) Davis, D. G. *J. Magn. Reson.* **1989**, *81*, 603–607.
(27) Shaka, A. J.; Lee, C. A. P. *J. Magn. Reson.* **1988**, *77*, 274–293.
(28) Otting, G.; Widmer, H.; Wagner, G.; Wüthrich, K. *J. Magn. Reson.* **1986**, *66*, 187–193.
(29) Takano, T. In *Methods and Applications in Crystallographic Computing*; Hall, S. R., Ashida, T., Eds.; Oxford University Press: Oxford, U.K., 1984; pp 262-272.
(30) Schomaker, V.; Waser, J.; Marsh, R. E.; Bergman, G. *Acta Crystallogr.* **1959**, *12*, 600–604.
(31) Dalvit, C.; Wright, P. E. *J. Mol. Biol.* **1987**, *194*, 313–327.
(32) Emerson, S. D.; Lecomte, J. T. J.; La Mar, G. N. *J. Am. Chem. Soc.* **1988**, *110*, 4176–4182.

below and elsewhere.[6,31] By assuming that the diamagnetic contribution to the total shift in metMbH$_2$O is equal to the MbCO chemical shift, the observed dipolar shifts can be evaluated from eqs 1 and 2 combined with $\delta_{con} = 0$ to yield eq 5:

$$\delta_{dip} = \delta(\text{metMbH}_2\text{O}) - \delta(\text{MbCO}) \quad (5)$$

Equation 3 is rearranged into

$$\delta_{dip,i} = DC(GF_i) \quad (6)$$

where $\delta_{dip,i}$ is the observed dipolar shift obtained from eq 5 for proton i, GF_i is the geometric factor of proton i, and C is a constant equal to $28g^2\beta^2/9k^2T^2$, in which $g = 2$, $\beta = 9.27 \times 10^{-21}$ erg G^{-1}, and $k^2 = (1.38 \times 10^{-16})0.70$ erg cm^{-1} K^{-2}. The observed dipolar contribution ($\delta_{dip,i}$) is then plotted versus GF_i, and D (in cm^{-1}) is calculated from the slope of the line, DC. For this purpose, a limited set of $\delta_{dip,i}$ values was chosen as described in the results. According to eq 6, the line should go through the origin; however, the best linear fit returns a small y-intercept represented below as b.

To avoid possible errors from the difference between the heme normal and the true z-axis of the ligand field, the D value is then refined in the following way. The limited set of $\delta_{dip,i}$ values and the associated i coordinates are used in a nonlinear least-squares fitting program which treats the z-direction ($x_{1,n}$, $x_{2,n}$, $x_{3,n}$) as an adjustable vector while the normalization condition $x_{1,n}^2 + x_{2,n}^2 + x_{3,n}^2 = 1$ is retained. The nonlinearly constrained minimization subroutine DNCONF from the International Mathematical and Statistical Library is applied to minimize the function

$$F(b,D,x_{1,n},x_{2,n},x_{3,n}) = \sum_i (\delta_{dip,i} - \delta_{dip,i}^{calc} - b)^2 \quad (7)$$

where b corresponds to the y-intercept of the $\delta_{dip,i}$ versus GF_i line. Convergence with this method is achieved readily and reproducibly. The fit provides an optimized z-direction and D value used in the rest of the calculations.

D. Uncertainty of Calculated δ_{dip}. The geometrical dependence contained in eq 3 results in different degrees of sensitivity of δ_{dip} to displacement about the atomic position. The sensitivity of individual protons can be estimated through an error analysis by taking the derivative of eq 3 with respect to x_j. The expression for the uncertainty is given by the following equation:

$$\Delta\delta_{dip} = \sum_{j=1}^{3} \left(\left| \frac{\partial \delta_{dip}}{\partial x_j} \right| \Delta x_j \right) \quad (8)$$

We assume that the Cartesian coordinate uncertainty (Δx_j) is 0.2 Å.[33] In fact, this value is anisotropic at each point and varies over the structure. However, an arbitrary and uniform Δx_j allows us to illustrate the relative sensitivity of the assigned protons to a small variation in position relative to the iron atom. The vertical error bars in Figures 3 and 4 represent the upper and lower limits of δ_{dip} calculated within a cube of dimension $2\Delta x_j$ centered on the x_j coordinates given in the PDB file.

Ring Current Shift Calculations. Ring current shift (RCS) calculations were performed with the refined neutron diffraction structure of sperm whale MbCO (PDB file 2mb5)[34] and the X-ray structure of the metaquo form.[29] To utilize the neutron diffraction data consistently with the X-ray data, protons and deuterons were eliminated from the coordinate file 2mb5 and regenerated in the same fashion as that for the metMbH$_2$O structure. The ring current shifts induced by aromatic side chains and the heme group were evaluated according to the Johnson–Bovey[35] and Haigh–Mallion models[36] with a FORTRAN program employing the five- and eight-loop heme ring models[37,38] and the ring current intensities recently listed by Ösapay and Case.[39] We report Haigh–Mallion data; Johnson–Bovey data are comparable.

Results

Structural Differences between MbCO and MetMbH$_2$O in the Solid State. The observable signals in our NMR experiments arise from protons that are not relaxed efficiently by the unpaired electrons. These protons generally experience relatively small dipolar shifts. To obtain the pure dipolar shift contributions needed for the determination of D, it is necessary to select protons for which the diamagnetic term is reliably known. MbCO provides an appropriate set of δ_{dia} values if it can be ascertained that the local structure around the protons of interest is unchanged upon altering oxidation state and ligation. One criterion that can be applied to determine which protons are acceptable for evaluating D is the difference in the RCS induced by neighboring aromatic protein residues and the heme group as evaluated with the metaquo and the carbonmonoxy solid-state structures. Figure 1 presents a plot of the calculated RCS difference, ΔRCS = RCS(metMbH$_2$O) – RCS(MbCO), versus residue number for the backbone amide protons (A), the C$^\alpha$ protons (B), and the side-chain protons (C). A comparison of the three panels illustrates that the two solid-state structures differ mostly in the side-chain positions. We assume that this pattern of ΔRCS values is preserved in solution. Although ring current calculations have limited accuracy, a ring current shift difference larger than an arbitrary 0.01 ppm for a given proton in metMbH$_2$O and MbCO is taken as an indication that the proton is not suitable for the determination of D.

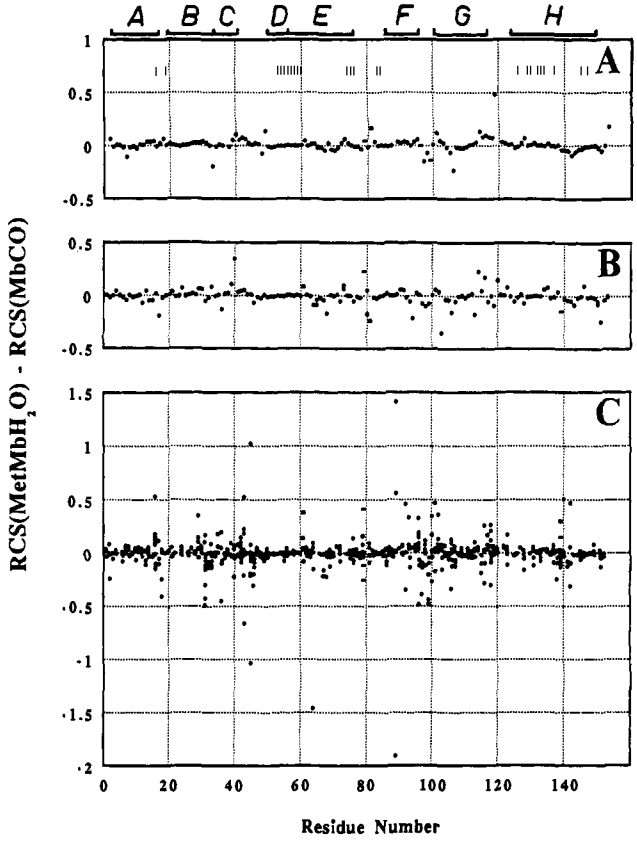

Figure 1. Ring current shift differences calculated from the solid-state structures of metaquoMb and MbCO versus residue number for (A) backbone amide protons, (B) C$^\alpha$ protons, and (C) side-chain protons. The ring current shift was obtained with the Haigh–Mallion model[36] as described in the text. On the average the ring current shift difference is larger for side-chain protons than for backbone protons. Helices extend as follows: A, 3–18; B, 20–35; C, 36–42; D, 51–57; E, 58–77; F, 86–95; G, 100–118; H, 124–149. The vertical bars in panel A indicate the location of basic set protons (see text).

Spectral Assignments. The RCS calculations point to the backbone protons as most appropriate for the determination of D, and we first concentrate on these. Several assignments have already been reported for metMbH$_2$O[6] and constitute starting points for further analysis of the spectra. Thus, the side chain of Val-17 was assigned on the basis of its NOE connectivities to

(33) Luzzati, V. *Acta Crystallogr.* **1952**, *5*, 802–810.
(34) Cheng, X.; Schoenborn, B. P. *J. Mol. Biol.* **1991**, *220*, 381–399.
(35) Johnson, C. E.; Bovey, F. A. *J. Chem. Phys.* **1958**, *29*, 1012–1014.
(36) Haigh, C. W.; Mallion, R. B. *Prog. NMR Spectrosc.* **1980**, *13*, 303–344.
(37) Perkins, S. J. *J. Magn. Reson.* **1980**, *38*, 297–312.
(38) Cross, K. J.; Wright, P. E. *J. Magn. Reson.* **1985**, *64*, 220–231.
(39) Ösapay, K.; Case, D. A. *J. Am. Chem. Soc.* **1991**, *113*, 9436–9444.

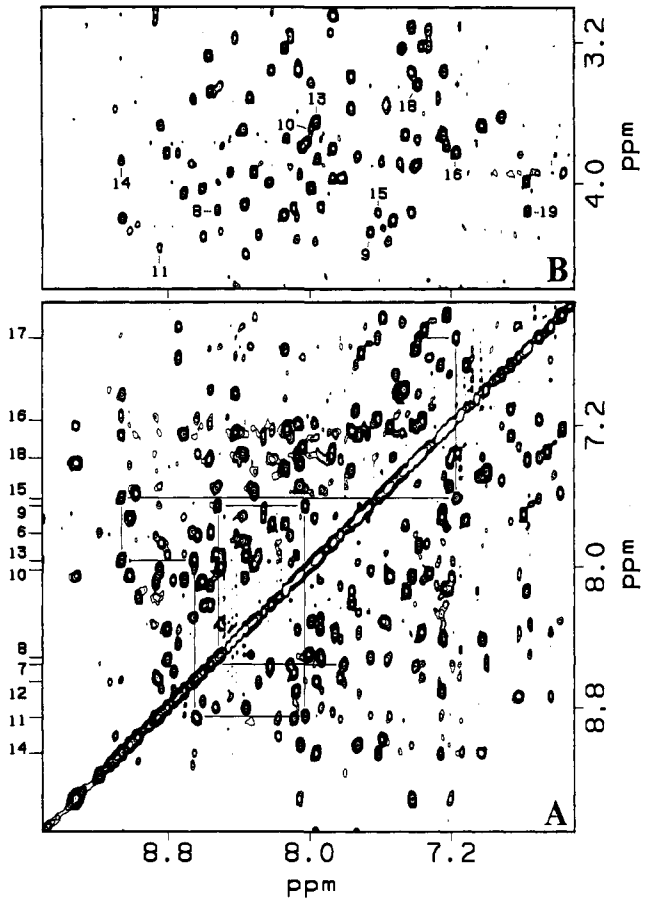

Figure 2. (A) NH–NH region of a phase-sensitive NOESY spectrum of sperm whale metMbH_2O recorded at 500 MHz, 313 K, and pH 5.7 in 90% H_2O/10% ^2H_2O with a mixing time of 110 ms. Chemical shifts are referenced to DSS through H_2O at 4.60 ppm. A-Helix amide protons from residue 6 to residue 18 are marked on the left side, and the NH–NH connectivities are indicated with the lines connecting cross peaks. (B) NH–C$^\alpha$H region of a TOCSY spectrum collected under the same conditions with a mixing time of 73 ms. The NH–C$^\alpha$H cross peaks arising from several residues in the A-helix are labeled. Assignments are contained in Table I and the supplementary material.

His-24 and Trp-14. The NH proton is readily found in TOCSY experiments and provides an entry into the A-helix. Figure 2A presents the NH–NH section of a metMbH_2O NOESY spectrum collected at pH 5.7 and 313 K. From the 17 NH–18 NH cross peak, the NH_i-to-NH_{i+1} connectivities for almost the entire A-helix (Gly-6 to Glu-18) are traced. Figure 2B illustrates the NH–C$^\alpha$H region of a TOCSY spectrum collected under the same conditions. The connectivities pertaining to the A-helix are annotated. Confirmation of the assignments is obtained by pursuing *J*-connectivities further in the side chain or by relying on NOEs, as in the case of aromatic side chains. The continuation of NH–NH connectivities from the C-terminus of the A-helix leads to the assignments of the NH of residues 21 (Val) and 22 (Ala). There is only one Val-Ala sequence in the protein, and the cross-peak pattern found in the TOCSY data supports the interpretation. These assignments were also confirmed by the following NOE connectivities: Val-21 ↔ Val-66; Val-21 ↔ Leu-69; Val-21 ↔ Val-17; Ala-22 ↔ Val-66. Starting from residues 21 and 22, the B-helix backbone connectivities can be traced up to Ile-30. The NH and C$^\alpha$H resonances were assigned for other helices with the same strategy. The assigned backbone protons include those of residues 53–57 (in the D-helix), 58–64 and 73–77 (in the E-helix), 78–85 (in the E–F turn), 101–102 and 113–119 (in the G-helix), and 127–147 (in the H-helix). Assignments and chemical shifts at 298 K are presented in Table I and in the supplementary material.

In addition to the backbone and side-chain protons described above, several residues that have large paramagnetic shifts were also assigned. For example, broad signals occur at –0.84 and 1.45 ppm that are *J*-correlated (DQF-COSY data not shown) and indicate a shifted alanine spin system. The temperature dependence of these signals is strong, and it, as well as the line width, reveals the strong influence of the paramagnetic center. In the NOESY spectrum this alanine is in dipolar contact with Tyr-151, which was assigned previously by comparing sperm whale Mb and horse Mb spectra.[6] NOEs are also observed from this alanine to a residue identified as another tyrosine spin system with ring proton resonances at 4.89 and 5.72 ppm (detected at 313 K, while only one set of ring protons is observed at 298 K). These two residues were assigned as Ala-94 and Tyr-146, since the X-ray studies show that Ala-94 is the only alanine close to Tyr-151 and a second Tyr, Tyr-146. A strong NOE is observed between the C$^\delta$H of Tyr-146 and a resonance at –0.04 ppm arising from a residue identified as an isoleucine in TOCSY and COSY spectra. The assignment of this isoleucine as Ile-101 is consistent with the X-ray structure of metMbH_2O[29] and NMR studies on MbCO.[31] The NOEs observed from the side chain of Ile-101 to backbone protons of Ile-101, Lys-102, and Ala-143 support this assignment. A relatively strong NOE cross peak was also found between C$^\delta$H of Ile-101 and a resonance at –1.21 ppm (at 298 K, shifting to –0.1 ppm at 313 K). The large temperature dependence of the chemical shift and the line width indicate proximity to the paramagnetic center. The NH and C$^\alpha$H resonances of this residue were found in a TOCSY experiment, and their backbone connectivities suggest that this residue is Ile-142. Only part of the Ile-142 spin system (NH, C$^\alpha$H, C$^\beta$H, and C$^{\gamma 2}H_3$) is observed in TOCSY data under our conditions probably because the remaining protons are too close (≤10 Å) to the iron. Other side chains were also assigned, and chemical shifts are available in the supplementary material.

The assignments in MbCO were obtained with the same procedure. In this case, the initial information is found in the early work of Dalvit and Wright[31] and in that of Hughson et al.[40] The assignment procedure was repeated for the residues already reported and extended further to obtain the needed chemical shifts at room temperature. These and other chemical shifts are listed in Table I and in the supplementary material.

Evaluation of *D* and Prediction of the Paramagnetic Shift. In order to determine the ZFS constant *D*, we first select a set of protons whose local environment is not altered by the change of ligand. This set is referred to as the basic set. Of the 141 assigned resonances in both complexes of Mb (which exclude the pH-sensitive histidine ring signals), 53 have a calculated ΔRCS value less than 0.01 ppm. Most of those are C$^\alpha$ and backbone amide protons (Figure 1). Several side-chain protons also display a small ΔRCS; this is probably due to fortuitous cancellation of various contributions, since in those residues the entire side chain does not display a uniformly small ΔRCS. In view of this behavior, all side-chain resonances were excluded from the basic set, which at this point contains 39 backbone protons. A linear least-squares fit of the observed $\delta_{dip,i}$ versus GF_i for the set leaves two of the 39 points more than 2 rms deviations from the best line. They arise from the NH resonances of residues 59 and 138. We therefore exclude these two points as well. The final basic set retains 37 protons located between 13 and 29 Å from the iron and displaying δ_{dip} values between –0.6 and +0.6 ppm (Table I). The 37-point set is well fitted with a straight line having a small nonzero *y*-intercept (b = –0.01 ppm) likely due to a systematic chemical shift referencing error and readily compensated for provided that the same data sets are used throughout the analysis. The slope gives a *D* value of 9.69 cm^{-1}.

The procedure described above relies on a precise knowledge of the effective ligand field *z*-axis to calculate θ (eq 3). In order

(40) Hughson, F. M.; Wright, P. E.; Baldwin, R. L. *Science* **1990**, *249*, 1544-1548.

Table I. Paramagnetic Shift Parameters for metMbH$_2$O at 298 K and pH 5.7

residue	proton[a]	R^b	θ^b	GF[c]	ΔRCS[d]	δ(metMbH$_2$O)[e]	δ(MbCO)[f]	δ_{dip}(obs)[g]	δ_{dip}(cal) + b^h	$\Delta\delta_{dip}$(cal)[i]	deviation[j]
						Basic Set Points					
Lys-16	H	21.6	60.4	−0.03	−0.01	7.17	7.14	0.03	0.02	0.01	0.01
Ala-19	H	20.6	50.3	0.03	0.01	6.78	6.83	−0.05	−0.04	0.02	0.01
Ala-19	HA	21.9	47.2	0.04	0.00	4.16	4.20	−0.04	−0.05	0.01	0.01
Ala-53	H	22.3	26.9	0.13	0.00	8.22	8.37	−0.15	−0.16	0.02	0.01
Ala-53	HA	22.2	19.6	0.15	0.00	3.72	3.92	−0.20	−0.19	0.02	0.01
Glu-54	H	21.5	26.7	0.14	0.00	7.51	7.69	−0.18	−0.18	0.02	0.00
Glu-54	HA	21.2	23.8	0.16	0.00	3.69	3.90	−0.21	−0.20	0.03	0.01
Met-55	H	18.9	25.0	0.22	0.00	7.78	8.05	−0.27	−0.27	0.04	0.00
Lys-56	H	18.8	17.7	0.26	0.00	8.19	8.53	−0.34	−0.32	0.04	0.02
Lys-56	HA	18.0	9.8	0.33	0.01	3.17	3.59	−0.42	−0.41	0.05	0.01
Ala-57	H	19.7	15.1	0.24	0.00	6.76	7.07	−0.31	−0.29	0.03	0.02
Ala-57	HA	20.9	11.1	0.21	0.00	3.94	4.20	−0.26	−0.26	0.02	0.00
Ser-58	H	17.9	15.8	0.31	0.00	6.38	6.78	−0.40	−0.38	0.05	0.02
Ser-58	HA	18.1	20.9	0.27	0.00	4.10	4.43	−0.34	−0.34	0.04	0.00
Glu-59	HA	17.2	18.1	0.34	0.01	3.61	4.03	−0.42	−0.42	0.04	0.00
Asp-60	H	15.9	26.4	0.35	0.01	7.96	8.42	−0.46	−0.43	0.06	0.03
Asp-60	HA	14.0	31.0	0.44	0.00	3.93	4.47	−0.54	−0.54	0.10	0.00
Ala-74	HA	15.1	89.5	−0.29	0.01	4.35	4.03	0.32	0.35	0.03	0.02
Ile-75	H	12.6	94.2	−0.50	0.01	7.79	7.23	0.56	0.59	0.09	0.03
Ile-75	HA	14.1	104.6	−0.29	0.00	3.85	3.51	0.34	0.34	0.07	0.01
Leu-76	H	13.9	92.1	−0.37	−0.01	9.10	8.63	0.47	0.44	0.06	0.03
Glu-83	H	18.0	126.5	0.01	0.00	8.82	8.82	0.00	−0.02	0.02	0.02
Ala-84	H	17.2	126.7	0.01	0.01	8.52	8.54	−0.02	−0.03	0.03	0.01
Ala-84	HA	16.4	132.6	0.09	0.00	3.92	4.01	−0.09	−0.11	0.05	0.02
Asp-126	H	28.5	76.1	−0.04	−0.01	8.65	8.61	0.04	0.04	0.00	0.00
Asp-126	HA	27.5	81.5	−0.05	−0.01	4.30	4.26	0.04	0.05	0.01	0.01
Gly-129	H	24.5	80.6	−0.06	0.00	8.15	8.09	0.06	0.07	0.01	0.01
Ala-130	H	23.7	84.4	−0.07	0.01	7.85	7.77	0.08	0.08	0.01	0.00
Ala-130	HA	22.8	90.6	−0.08	0.00	3.97	3.87	0.10	0.09	0.01	0.01
Asn-132	H	19.9	85.2	−0.12	0.00	8.93	8.79	0.14	0.14	0.02	0.00
Lys-133	H	20.5	92.7	−0.12	0.01	8.13	8.00	0.13	0.13	0.01	0.00
Lys-133	HA	20.1	100.5	−0.11	0.00	4.16	4.02	0.14	0.13	0.02	0.01
Ala-134	H	18.9	94.3	−0.15	0.00	8.67	8.50	0.17	0.17	0.02	0.00
Leu-137	H	17.0	107.9	−0.15	−0.01	8.10	7.93	0.17	0.17	0.04	0.00
Leu-137	HA	17.5	117.2	−0.07	−0.01	4.22	4.13	0.09	0.08	0.03	0.01
Lys-145	HA	16.3	161.9	0.40	0.00	3.61	4.08	−0.48	−0.49	0.06	0.02
Lys-147	H	16.4	156.3	0.34	−0.01	7.23	7.63	−0.40	−0.42	0.05	0.02
						Deviating Points[k]					
Leu-2	HD2	21.2	96.6	−0.10	0.08	0.45	0.36	0.09	0.11	0.01	0.02
Val-13	HA	22.4	63.5	−0.04	0.04	3.67	3.61	0.06	0.04	0.01	0.03
Val-13	HB	22.3	68.5	−0.05	0.02	2.50	2.42	0.08	0.06	0.01	0.03
Val-13	HG2	22.4	62.8	−0.03	−0.11	1.19	1.14	0.05	0.03	0.01	0.02
Trp-14	HE1	14.8	68.4	−0.18	−0.03	10.47	10.31	0.16	0.21	0.06	0.05
Trp-14	HE3	16.3	68.6	−0.14	0.01	7.28	7.09	0.19	0.16	0.04	0.03
Ala-15	H	21.0	67.3	−0.06	0.04	7.61	7.59	0.02	0.06	0.01	0.04
Val-17	HB	17.2	57.6	−0.03	0.12	0.90	0.84	0.06	0.02	0.03	0.04
Val-17	HG1	15.5	51.5	0.04	−0.05	−0.54	−0.51	−0.03	−0.06	0.04	0.03
Glu-18	H	19.1	56.1	−0.01	0.07	7.38	7.40	−0.02	0.00	0.02	0.02
Leu-40	HG	13.4	56.3	−0.03	−0.03	1.14	1.20	−0.06	0.03	0.07	0.09
Met-55	HE	16.2	39.1	0.19	−0.10	1.43	1.72	−0.29	−0.24	0.06	0.05
Ala-57	HB	21.1	16.3	0.19	0.00	1.15	1.37	−0.22	−0.24	0.03	0.02
Glu-59	H	18.1	21.5	0.27	0.00	8.90	9.18	−0.28	−0.33	0.04	0.05
Ala-74	H	12.6	84.3	−0.48	0.03	8.01	7.52	0.49	0.58	0.09	0.09
Ala-74	HB	13.2	92.7	−0.43	0.00	2.04	1.58	0.46	0.51	0.06	0.06
Gly-80	H	18.4	107.8	−0.12	−0.01	9.37	9.29	0.08	0.13	0.03	0.05
Ala-94	HB	11.6	169.7	1.23	0.05	−0.86	0.33	−1.19	−1.51	0.25	0.32
Ile-101	H	12.2	130.4	0.14	0.12	9.21	9.15	0.06	−0.18	0.14	0.24
Ile-101	HA	10.7	126.9	0.07	0.06	3.62	3.55	0.07	−0.09	0.18	0.16
Ile-101	HB	13.6	128.6	0.07	0.04	1.89	1.85	0.04	−0.09	0.08	0.13
Ile-101	HB2	13.5	124.6	−0.02	0.01	0.84	0.77	0.07	0.01	0.07	0.07
Lys-102	H	13.0	118.0	−0.15	0.05	9.13	8.78	0.35	0.18	0.10	0.17
Lys-102	HA	13.8	109.0	−0.26	0.04	4.25	4.14	0.11	0.31	0.08	0.19
Leu-115	H	17.5	54.9	0.00	0.08	8.39	8.42	−0.03	−0.01	0.02	0.02
Met-131	HE	18.0	73.2	−0.13	−0.05	2.42	2.22	0.20	0.15	0.03	0.05
Phe-138	H	14.8	109.8	−0.20	0.01	7.95	7.78	0.17	0.24	0.06	0.07
Arg-139	H	13.3	113.5	−0.22	−0.04	8.59	8.42	0.16	0.26	0.09	0.10
Lys-140	H	15.0	121.8	−0.05	−0.04	8.83	8.71	0.12	0.05	0.05	0.07
Asp-141	HB1	15.2	136.2	0.16	−0.12	2.53	2.67	−0.14	−0.21	0.07	0.07
Asp-141	HB2	14.7	130.6	0.09	−0.08	2.32	2.55	−0.23	−0.11	0.07	0.12
Ile-142	HG2	9.8	146.4	1.15	−0.09	−1.21	−0.12	−1.09	−1.40	0.43	0.31
Ala-143	H	13.7	139.0	0.28	−0.06	7.87	8.13	−0.26	−0.35	0.10	0.09
Ala-144	H	15.7	144.8	0.26	−0.04	7.35	7.62	−0.27	−0.33	0.07	0.06

[a] Proton names as in the PDB file. [b] R (in Å) is the length of the vector connecting the proton to the Fe(III) atom, and θ (in deg) is the angle between that vector and the normal to the heme plane; both R and θ are calculated from the X-ray structure of sperm whale Mb (4mbn). In the cases of methyl and aromatic ring protons, the average R and θ are reported. [c] Geometric factor (in 10^{21} cm^{-3}) is defined as $(3\cos^2\theta - 1)/R^3$. [d] Ring current shift differences calculated from the solid-state structures of metMbH$_2$O and MbCO with the Haigh–Mallion model;[36] ΔRCS (ppm) = RCS(metMbH$_2$O) − RCS(MbCO). [e] Observed chemical shift (ppm) in metMbH$_2$O. [f] Observed chemical shift (ppm) in MbCO. [g] Observed dipolar shift, δ_{dip}(obs) = δ(metMbH$_2$O) − δ(MbCO). [h] Calculated dipolar shift (ppm) obtained from eq 3 with metMbH$_2$O X-ray structure 4mbn; 0.01 ppm was subtracted from the calculated value to account for the small negative nonzero intercept from the linear fit. [i] Uncertainty in δ_{dip}(cal) calculated with eq 7. [j] Absolute deviation between the calculated and observed dipolar shifts. [k] The predicted and observed dipolar shifts are considered to be inconsistent according to the criterion described in the Results. The remaining 70 data points (not part of the basic set and not "deviating") are listed in the supplementary material.

to test the adequacy of the normal to the heme plane as z-axis, the 37-point basic set information ($\delta_{dip,i}$ and $x_{j,i}$) was subjected to a 5-parameter fit that determines simultaneously the optimal z-direction, D value, and b-intercept. The fit yielded a D of 9.66 cm^{-1}, in excellent agreement with other ZFS constants reported for metMbH_2O[41-43] and not significantly different from the linear fit value. The standard deviation of D was estimated to be ±0.08 cm^{-1} by using the new z-axis in the linear fit of δ_{dip} versus geometric factor for the basic set. The scatter of the points and the rms deviation are small, and D so obtained is insensitive to the removal of any single point from the data set. The new z-direction is tilted by less than 3° from the normal to the heme plane defined by the macrocycle or by the four pyrrole nitrogens. The optimized parameters were employed in further calculations to determine the dipolar contribution at any point in space. The refinement allows us to conclude that the normal to the heme plane corresponds well to the z-axis and, alternatively, that the D value can be obtained without using the heme plane coordinates if the basic set is large enough. In our method the calculations performed with the heme normal helped in selecting the best set of basic points.

In an ideal case, there should be perfect correlation between calculated and observed shifts. The correlation plot for the basic set is shown in Figure 3A. As anticipated from the fit of observed δ_{dip} versus geometric factor, all points cluster near the line of slope unity. Figure 3B contains all the data points that were not considered adequate for the evaluation of D. The scatter is larger for this latter set than for the basic set. Calculation of D based on the 141 points yields D_{all} = 8.7 cm^{-1}. D_{all} and D, although close, are not within one standard deviation of one another.

The uncertainty in the calculated dipolar shift depends on its sensitivity to the atomic position. The crystal structure coordinates for metMbH_2O (4mbn) carry an uncertainty related to the resolution (2.0 Å, with an R factor of 0.172) and the thermal factor (average B value of 13.1 $Å^2$). As a representative value, we use Δx_j = 0.2 Å.[33] Calculation of the resulting uncertainty in δ_{dip} is achieved by assuming that an isotropic displacement of amplitude 0.4 Å is possible at each proton location. Thus, the value represents the maximum range of δ_{dip} corresponding to this Δx_j. The range is plotted as a vertical error bar in Figures 3 and 4. In general, when a proton is close enough to the iron atom to experience a large dipolar shift ($|\delta_{dip}| > 0.5$ ppm), the sensitivity to the atomic position is high. While the basic set is composed mostly of points with a low sensitivity to displacement (Figure 3A), several points in the extended set display a strong dependence on the atomic coordinates (Figure 3B). The error in the observed shift arises directly from the two measurements (MbCO shift and metMbH_2O shift) and is estimated at ±0.02 ppm. The uncertainty boxes therefore have the same 0.04-ppm basis and varying height. A point is considered within error of the expected value if the correlation line bisects the box in two areas, both larger than 15% of the total box area.

In Figure 4 we illustrate the center portion of Figure 3B (small δ_{dip}). Figure 4A contains the data points for which the calculation is successful, exclusive of the basic set; Figure 4B presents the remainder. Overall, there are 34 points for which calculated and observed values do not agree well within the sensitivity limits (vertical error bars) and chemical shift error (horizontal error bars). These deviations are not the consequence of a misoriented z-axis, since a nonlinear least-squares fit carried out with the 34 points alone does not converge to a reasonable z-axis reducing the deviations. This implies that an accurate prediction of δ_{dip} for these protons requires their position relative to the heme to be displaced from that given by the X-ray data. The residues represented by one or more protons in Figure 4B include Leu-2, Val-13, Trp-14, Ala-15, Val-17, Glu-18, Leu-40, Met-55, Ala-57, Glu-59, Ala-74, Gly-80, Ala-94, Ile-101, Lys-102, Leu-115, Met-131, Phe-138, Arg-139, Lys-140, Asp-141, Ile-142, Ala-143, and Ala-144.

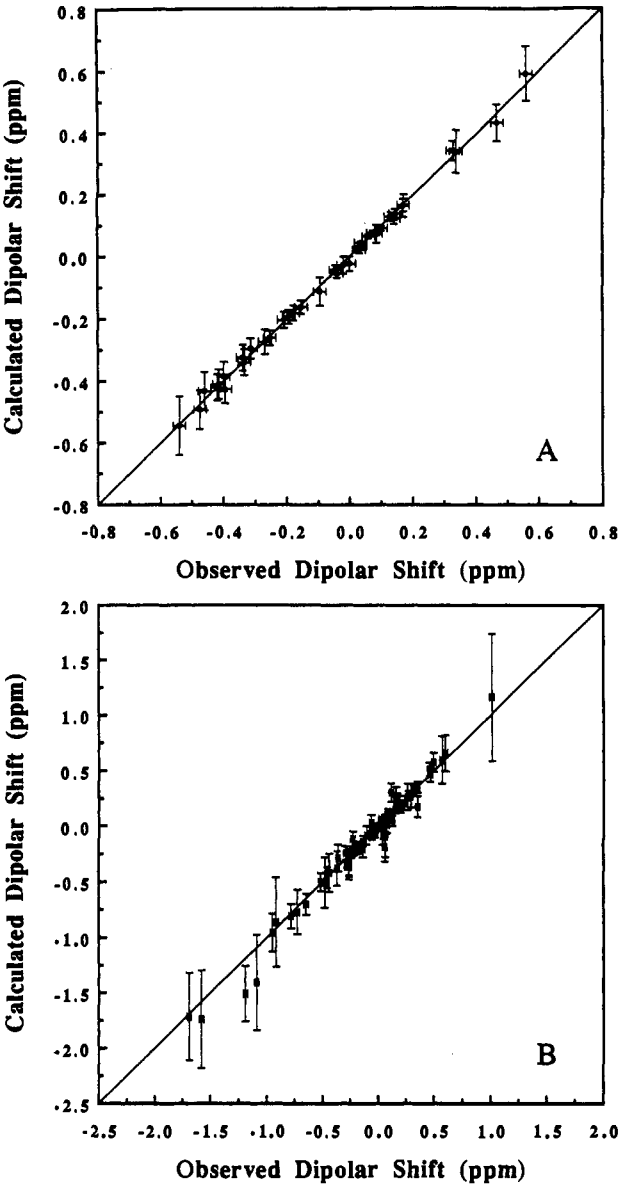

Figure 3. Calculated versus observed dipolar shifts in sperm whale metMbH_2O (A) for the protons in the basic set and (B) for 104 protons (or groups of protons) that are not included in the basic set. The straight lines are perfect correlation lines. The vertical error bars mark the sensitivity of the dipolar shift to the variance of the atomic position; they are calculated with the assumption that the uncertainty of the coordinates given in the PDB file 4mbn is ±0.2 Å. The horizontal error bars (±0.02 ppm) represent the estimated uncertainty in the dipolar shift measurement. A strong sensitivity to the atomic position is observed for the points with a large δ_{dip} (i.e., $|\delta_{dip}| > 0.5$ ppm); these arise from protons in the proximity of the iron atom. The large deviations are covered by the error bars, and the corresponding protons are not considered significantly perturbed by complexation or dissolution. The methyl group of Ala-94 (-1.51 ppm; -1.19 ppm) is an exception discussed in the text.

Discussion

In NMR spectroscopy of proteins, the chemical shift of proton resonances is a common indicator of ordered structure.[44] A recent

(41) Brackett, G. C.; Richards, P. L.; Caughey, W. S. *J. Chem. Phys.* **1971**, *54*, 4383–4401.
(42) Uenoyama, H.; Ilzuka, T.; Morimoto, H.; Kotani, M. *Biochim. Biophys. Acta* **1968**, *160*, 159–166.
(43) Scholes, C. P.; Isaacson, R. A.; Feher, G. *Biochim. Biophys. Acta* **1971**, *244*, 206–210.
(44) Bundi, A.; Wüthrich, K. *Biopolymers* **1979**, *18*, 285–297.

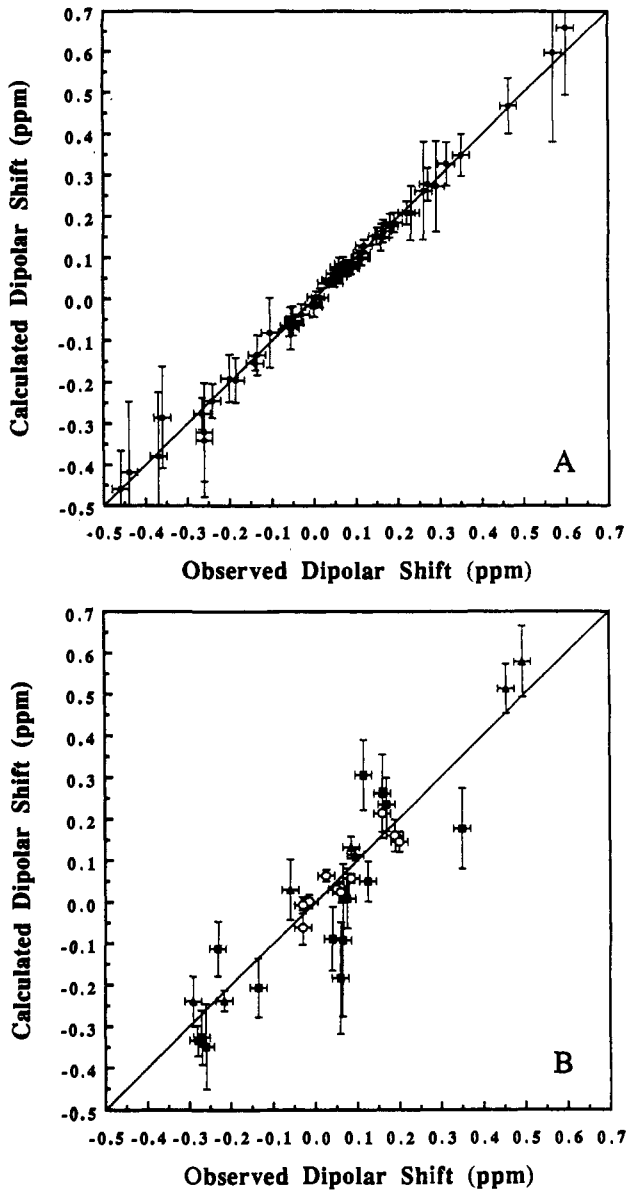

Figure 4. Expanded region (–0.5 to 0.5 ppm) of Figure 3B: (A) Points for which the error bars cross the perfect correlation line and structural agreement is most likely. (B) Points for which the predicted and observed values do not agree within the error box (15% area criterion, see Results). The deviating points are classified into three groups: Ile-101, Lys-102, Phe-138, Arg-139, Lys-140, Asp-141, Ile-142, Ala-143, and Ala-144 (filled squares), residues which are located at the G–H interface; Val-13, Trp-14, Ala-15, Val-17, Glu-18, Met-131, and Leu-115 (open circles), residues which are located at the A–G interface; and Leu-2, Leu-40, Met-55, Glu-59, and Gly-80 (filled triangles), residues which are isolated from other assigned residues. Chemical shifts are listed in Table I.

analysis of the available NMR data bases has established empirical trends relating the chemical shift and the type of secondary structure in which a residue is involved.[45,46] Calculations of the conformation-dependent components of the shift have included anisotropy of the peptide group and backbone electrostatic effects.[39] Two main contributions can be precisely evaluated from a knowledge of the structure: ring current effects[35] and paramagnetic effects.[15,47–49] Both have practical interest in the comparison of solid-state and solution structures.

While RCS calculations are successful for aromatic side chains,[35,36,50] the heme contribution is more complex to model. Cross and Wright[38] have reported a method relying on the five- and eight-loop Johnson–Bovey and Haigh–Mallion models which yields acceptable results. The calibration constants have been recently readjusted by Ösapay and Case,[39] and these are the constants we used. Paramagnetic shift calculations have been performed on several proteins: cytochrome *c*,[51] metMbCN, a low-spin complex of myoglobin,[52,53] and more recently a cobalt-substituted zinc finger,[49] to name a few. These studies indicate that when a reliable estimate of the diamagnetic contribution is available, the paramagnetic shift calculation is a useful tool for investigating the structure of a protein and its response to external conditions. $MetMbH_2O$ makes a good model system for NMR chemical shift and structural studies, since its crystal structure is available to an appropriate resolution. Furthermore, unlike that in metMbCN,[52,53] the dipolar shift in the high-spin $metMbH_2O$ complex has a simple mathematical form (eq 3) due to the nearly totally symmetrical ground state of Fe(III). The paramagnetic shift calculation for $metMbH_2O$ therefore requires no knowledge of susceptibilities.

In our study, the paramagnetic shift is used as the major structural probe while the RCS calculations assist in identifying the regions of the structure where the diamagnetic reference is likely to be incorrect. There are two reasonable underlying assumptions in this approach. (1) The ring current generated by aromatic rings is likely to dominate the observed difference in diamagnetic shift. Other chemical shift perturbations, such as those due to reorientation within a nonaromatic environment, are expected to be of lesser magnitude. (2) The ring current shift *difference* is a sensitive marker of conformational alteration even though the calculated ring current shift in individual structures does not always account accurately for the chemical shifts observed in solution because of the inherent problems associated with RCS evaluations.

In the process of determining D, a conservative choice of resonances was made. Most assigned signals were rejected from the basic set because the RCS calculation revealed a ΔRCS between the paramagnetic form and the diamagnetic form larger than the experimental error in chemical shift determination (i.e., >0.01 ppm). ΔRCS values above this cutoff suggest that the local environment of a proton is poorly described in one of the structures. Alternatively, the environment might be adequately described but susceptible to deformation, either by the change in external conditions and complexation or by crystal-packing effects. In any case, the excessive ΔRCS makes that proton less reliable for the determination of D. A description of the structural difference between $metMbH_2O$ and MbCO can also be obtained through the geometric factor, which contains the structural parameters R and θ, or through R and θ themselves. The RCS calculation provides not only the locations of structural difference in the protein but also the sensitivity of the diamagnetic chemical shift to the local structural change and is taken to be a better criterion.

After applying the RCS structural selection criterion and standard statistical treatment, it is found that the protons in the basic set (Figure 3A) display a significant range of δ_{dip} (–0.6 to +0.6 ppm) and have a δ_{dip} relatively insensitive to their position. The span of δ_{dip} values and the small sensitivity contribute to the precision and accuracy of D for further calculations. The

(45) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311–333.
(46) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647–1651.
(47) Jesson, J. P. In *NMR of Paramagnetic Molecules. Principles and Applications*; La Mar, G. N.; Horrocks, W. DeW., Jr.; Holm, R. H., Eds.; Academic: New York, 1973; pp 1–52.
(48) *NMR of Paramagnetic Molecules. Principles and Applications*; La Mar, G. N., Horrocks, W. DeW., Jr., Holm, R. H., Eds.; Academic: New York, 1973.
(49) Harper, L. V.; Amann, B. T.; Kilfoil Vinson, V.; Berg, J. M. *J. Am. Chem. Soc.* **1993**, *115*, 2577–2580.
(50) Memory, J. D. *J. Magn. Reson.* **1977**, *27*, 241–244.
(51) Williams, G.; Clayden, N. J.; Moore, G. R.; Williams, R. J. P. *J. Mol. Biol.* **1985**, *183*, 447–460.
(52) Emerson, S. D.; La Mar, G. N. *Biochemistry* **1990**, *29*, 1545–1555.
(53) Emerson, S. D.; La Mar, G. N. *Biochemistry* **1990**, *29*, 1556–1566.

correlation coefficient for the linear fit is 0.999, a high value which demonstrates the validity of eq 3 as a model. In addition, the small rms deviation confirms that the structure around the protons in this set is minimally perturbed by the change of ligand or phase, or both.

D obtained with chemical shift information (9.66 ± 0.08 cm^{-1}) agrees well with the value (9.5 cm^{-1}) from the IR studies on polycrystalline metMbH$_2$O at 4 K[41] and the values (9.26 cm^{-1} for single crystal and 9.14 cm^{-1} for frozen solution) obtained through the temperature dependence of the electron spin–lattice relaxation rate.[43] Values ranging between 4 and 10 cm^{-1} have also been reported by others,[14,42,54] although the lower value[54] has been questioned.[43] The chemical shift method reported here yields a reliable *D* on the basis of simple room temperature measurements but requires X-ray coordinates and extensive spectral assignments. Its advantage resides in the self-consistency it provides for proton NMR applications.

The NMR *D* constant can be used to calculate the dipolar shift experienced at any location in the structure under the same conditions. As shown in Figure 3B, δ_{dip} for most of the 104 assigned resonances not included in the basic set is well predicted. This results in part from the stringency of the criterion for constructing the basic set, since many protons rejected from it do have small ΔRCS values, albeit larger than 0.01 ppm. These are expected to be closely predicted if the solid-state structures are appropriate models for the solution structures. Overall, 80% of the data points lie within ±0.05 ppm of the calculated shift. The narrow distribution simplifies the spectral assignment process greatly because it allows one to predict where a particular resonance is likely to be found if its assignment is available in one of the two forms of the protein (diamagnetic or paramagnetic). Large discrepancies should be few; when encountered they first raise the question of assignment correctness and suggest reinspection of the spectra.

A small number of observed δ_{dip} values, for which assignment was carefully confirmed, depart visibly from the calculated values. The protons giving rise to incorrect shifts, or rather the residues bearing these protons, point to regions of structure where the solid-state coordinates are not adequate to describe the solution conformation. Since the sensitivity of δ_{dip} varies with position, the structural interpretation of the chemical shift deviations must be adjusted for each proton and it is necessary to examine the possible origins of discrepancies between observed and calculated values.

The dipolar chemical shift contribution can be affected by three main sources of error: (1) Inacceptable van der Waals contacts and the inherent imprecision of the atomic positions in the metMbH$_2$O X-ray structure may invalidate the calculated dipolar shift. (2) Structural alterations on going from the solid-state (often under high-salt conditions) to the solution state of metMbH$_2$O may also invalidate the calculated dipolar shift. (3) Differences in the solution structure of the diamagnetic and the paramagnetic forms may invalidate the diamagnetic reference contribution and falsify the observed dipolar shift. These three possibilities may of course combine.

The deviating residues (Figures 3B and 4B) can be segregated in two major groups: Ala-94, Ile-101, Lys-102, Phe-138, Arg-139, Lys-140, Asp-141, Ile-142, Ala-143, and Ala-144, residues located at the F–G–H interface; and Val-13, Trp-14, Ala-15, Val-17, Glu-18, Leu-115, and Met 131, residues located at the A–G–H interface. Departure is clearly distinct in the two regions; the first region includes residues close to the iron, with large δ_{dip} and sensitivity to displacement, and the second region, residues remote from the iron, with small δ_{dip} and weak sensitivity. Other residues not located in the above two regions also display significant deviations: Leu-2, Leu-40, Met-55, Ala-57, Glu-59, Ala-74, and Gly-80.

In the refined X-ray structure of metMbH$_2$O[29] a few residues participate in sterically unfavorable contacts. These residues are therefore expected to undergo a structural rearrangement in solution, perhaps with an amplitude larger than the Δx_j tolerance of 0.2 Å. For example, the last three residues at the carboxy terminus fold back so that Gly-153 is positioned close to Ile-101, so close in fact that the distance between the NH of Ile-101 (G2) and one of the C$^\alpha$H's of Gly-153 is only 1.6 Å. For these two protons, the X-ray geometry is unrealistic.[55] The unfavorable contact can be eliminated by displacement of Gly-153, the last residue of Mb, and by local reorganization around Ile-101. Conformational differences among available Mb solid-state structures and a high *B* factor at the C-terminus establish that such a rearrangement is feasible. It is also known that the side chain of Tyr-151 in sperm whale Mb and that of Phe-151 in horse heart Mb adopt relatively different conformations with respect to Tyr-146 and Ala-94 in the metMbH$_2$O form[56,57] and in the des-Fe (iron-free) form of myoglobin.[58] This lends further support to the possibility that the structure near the C-terminus is tolerant of changes and may explain the discrepancy between the predicted and observed dipolar shift for the side chain of Ala-94 (Figure 3B). Ala-94 is in close contact with the H-helix through Phe-146. Ile-101 is also in close contact with the H-helix at Ile-142 and Tyr-146 and lies between three structural elements (G-helix, H-helix, and C-terminus). The poor prediction of dipolar chemical shift for several residues near Ile-101 suggests that perturbations extend through the interface.

In the second group of deviating residues, close van der Waals contacts are also detected but these are not as severe as those in the first group and might not require relaxation outside the 0.2-Å limit. In order to characterize fully the discrepancy between predicted and observed shifts, it is necessary to apply ΔRCS to the diamagnetic reference and to inspect the remaining deviation. When this approach is taken, the data points are not brought within the error limits of the predicted value. Such a result does not allow one to distinguish between a failure of the RCS calculation, which is likely to be less accurate than acceptable for this particular purpose,[39] and structural differences beyond those manifested in the crystal structures. The ambiguity will have to remain until complementary structural data are gathered and we limit the interpretation to the statement that the deviations observed in the second group indicate a region of Mb for which the solid-state coordinates might be inadequate.

The comparison of MbCO and metMbH$_2$O crystal structures shows that, even in the heme binding site, the differences in the two structures are minimal. Hence, in spite of the difficulty in identifying the error source, the poor correlation between calculated and observed shift is probably a result of the structural perturbation by dissolution rather than by complexation. Interestingly, the second site is also located in an interface involving distinct structural elements: the side chains of Val-13 (A11), Val-17 (A15), and Leu-115 (G16) participate in a hydrophobic core between the A and G helices. The side chain of Met-131 from the H-helix points toward this hydrophobic core. Steric incompatibilities involving residues 13, 16, 18, 122, and 123 in the sperm whale protein render reorganization around Val-13 necessary. Precedent for structural variability in this region of Mb is found in horse heart Mb.[56,57] For this protein, the C$^\alpha$ positions in the contact region (residues 15–17 and 118–123) differ markedly from the sperm whale positions in spite of strong sequence homology. The observation is consistent with the potential for rearrangement in solution.

Figure 4B also illustrates that the δ_{dip} values for some of the protons of Leu-2, Leu-40, Met-55, Ala-57, Glu-59, Ala-74, and

(54) Eisenberger, P.; Pershan, P. S. *J. Chem. Phys.* **1966**, *45*, 2832–2835.
(55) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283–437.
(56) Evans, S. V.; Brayer, G. D. *J. Biol. Chem.* **1988**, *263*, 4263–4268.
(57) Evans, S. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *213*, 885–897.
(58) Lecomte, J. T. J.; Cocco, M. J. *Biochemistry* **1990**, *29*, 11057–11067.

Gly-80 are not well predicted. Owing to the low number of assignments in and around these residues, it is not clear yet how and to what extent the surroundings are affected, but most of these residues display suspicious structural features. For example, Leu-2 (NA2) and Leu-40 (C5) have a *B* factor higher than the average. Leu-40 and Met-55 (D5) point at each other and block an opening between the C- and D-helices and the CD corner. Glu-59 (E2) is in contact with Arg-62 (E5) via a water molecule, and its amide hydrogen is coordinated to a sulfate ion bridging a neighboring Mb molecule in the crystal. Gly-80 (EF3) is in a short turn connecting the E-helix and the F-helix and docking against the H-helix. Small perturbations, apparently not directly associated with close contacts, might occur in these regions of myoglobin upon dissolution.

The translation of dipolar shift deviations into corrected coordinates is a cumbersome multivariable fitting problem. Qualitatively, Figure 4B illustrates that the 0.2-Å tolerance is close to adequate to restore agreement in many instances; overall then, the displacements with respect to the X-ray structure are expected to be small. With more complete spectral assignments, a quantitative structural study integrating NOE data, and the incorporation of the geometric information contained in the observed dipolar shift as a constraint for energy minimization programs, it will be feasible to characterize these displacements. Even without a full structural determination, the method can be quite useful. One application is to site-directed mutants: an individual side chain can be oriented with respect to the wild-type matrix simply by adjusting the dihedral angles until all dipolar shifts are simultaneously accounted for. This approach has been used for more difficult systems by La Mar and co-workers[59] and Berg and co-workers.[49] Another application is in the analysis of structural perturbations caused by experimental conditions. For example, when one of the experimental parameters is altered (e.g., ionic strength), the new δ_{dip} can be obtained by using δ_{dia} from the spectrum of the diamagnetic reference recorded under the same conditions. If a conformational change occurs in one or both forms, a deviation from the calculated value will locate the rearrangement. To ascertain the nature of δ_{dip}, the T^2 temperature dependence of the dipolar contribution (eq 3) can be tested. We are currently inspecting the environment of histidine side chains with this procedure.

(59) Rajarathnam, K.; Qin, J.; La Mar, G. N.; Chiu, M. L.; Sligar, S. G. *Biochemistry* **1993**, *32*, 5670–5680.

Conclusions

The zero-field splitting constant for metMbH_2O obtained from NMR chemical shifts and solid-state coordinates is a sensitive parameter useful for structural purposes. If the resonance of a given proton is not found within a tolerated range of the value calculated with the NMR-derived *D*, a critical analysis of the solid-state coordinates is justified at that site. It appears that the X-ray structure of metMbH_2O is not adequate for dipolar shift calculations in more than two regions of the protein. Most observations can be rationalized by flaws in the X-ray structure (bad van der Waals contacts) or peculiarities encountered in the solid-state (bound ions and intermolecular contacts). The fact that shifts are poorly predicted when there is an obvious cause reinforces the validity of the tool for general application.

When appropriate data are collected, the dipolar shift criterion will contribute to a precise description of the regions of Mb that either are not well depicted by the X-ray data or are affected by the external conditions. Furthermore, because the dipolar shift is so sensitive to the geometry, it will provide valuable constraints in the determination of solution structures of metaquomyoglobin and constitute a stringent test for the outcome of energy minimization methods.

Acknowledgment. This work was supported in part by PHS Grant 2 S07 RR07082 (Biomedical Science Support) and in part by Grant 26495-AC4,7 from the American Chemical Society, Petroleum Research Fund. The authors thank Melanie Cocco for help with the MbCO sample preparation, Dr. Klára Ösapay for providing her ring current shift results for comparison to ours, Dr. Gregory Farber for advice on the use of X-ray data, Dr. Christopher Falzone for reading the manuscript, Dr. Bertrand Garcia-Moreno for critical discussions, and the reviewers for useful suggestions.

Supplementary Material Available: A listing of chemical shifts in the metaquo and carbonmonoxy forms and of the NOEs used for spectral assignment (10 pages). Ordering information is given on any current masthead page.