NMR Investigations of *Clostridium pasteurianum* Rubredoxin. Origin of Hyperfine ^{1}H, ^{2}H, ^{13}C, and ^{15}N NMR Chemical Shifts in Iron−Sulfur Proteins As Determined by Comparison of Experimental Data with Hybrid Density Functional Calculations†

Steven J. Wilkens,‡ Bin Xia,‡ Frank Weinhold,‡,§ John L. Markley,‡,⊥ and William M. Westler*,⊥

Contribution from the National Magnetic Resonance Facility at Madison, Department of Biochemistry, and Department of Chemistry, University of Wisconsin−Madison, 420 Henry Mall, Madison, Wisconsin 53706

Received October 6, 1997

Abstract: High-level, all-electron, density functional calculations on a 104-atom model (B3LYP/6-311G** level) have been used, in conjunction with high-resolution X-ray structural data, to predict the remarkable paramagnetic contact shifts recently measured for ^{1}H, ^{2}H, ^{13}C, and ^{15}N nuclei in *Clostridium pasteurianum* rubredoxin. Three published X-ray structures for the Fe(III) rubredoxin from *C. pasteurianum* were employed to construct a 104-atom model for the iron center that included all atoms shown to have strong electronic interactions with the Fe. Each of these models served as a starting point for quantum mechanical calculations at level B3LYP/6-311G**, which, in turn, yielded calculated values for Fermi contact spin densities. The results indicate that the experimental hyperfine shifts are dominated by Fermi contact interactions: calculated Fermi contact spin densities were found to correlate linearly with isotropic hyperfine ^{1}H, ^{2}H, ^{13}C, and ^{15}N NMR chemical shifts determined for Fe(III) rubredoxin. At the current level of hyperfine peak assignments (all signals assigned to residue and atom types; some assigned to sequence specifically), comparisons were made between experimental shifts and those calculated from the structural model derived from each of the three X-ray structures. For Fe(III) rubredoxin, the R^2 values for the correlation between the calculated spin densities and experimental chemical shifts ranged, depending on the model, from 0.93 to 0.96 for 12 experimental ^{2}H signals and from 0.85 to 0.96 for 12 experimental ^{15}N signals. The correlation with experiment was improved by performing partial geometry optimizations at B3LYP/3-21G* on two of the 104-atom models. In particular, the R^2 correlation with experimental ^{15}N chemical shifts was improved from 0.85 to 0.94 upon optimizing the positions of the nitrogen bound protons reported for one of the X-ray structures. A small increase in the correlation with experiment was also found after optimizing the positions of the α- and β-protons of the cysteines of another model. The consistent overall agreement between calculation and experiment supports the validity and usefulness of combining quantum chemical methods with NMR spectroscopy and X-ray crystallography for the testing and refinement of molecular structures. Significantly poorer correlations with experiment were obtained when hypothetical Fe(II) models, derived from two of the X-ray structures of Fe(III) rubredoxin, were used as the basis for the spin density calculations; this suggests that the protein undergoes subtle structural changes upon reduction.

Introduction

Rubredoxins represent the simplest type of iron−sulfur proteins, in that they contain a single iron ion ligated by sulfur atoms from a surrounding cage of four cysteine residues. A simplified model of the iron−sulfur center is shown in Figure 1. Whereas the rubredoxin from *Pseudomonas oleovorans* has been shown to be involved in the hydroxylation of alkanes,[1−3] the biological function of the rubredoxin from *Clostridium pasteurianum* remains to be discovered. The X-ray structures of several rubredoxins have been reported,[4−9] including two independent structures of *C. pasteurianum* rubredoxin at resolutions as high as 0.9 Å.[9] In addition, rubredoxins have been investigated extensively by NMR,[10−16] EPR,[17,18] Mossbauer,[19,20] and MCD[21] spectroscopy.

† Coordinates for the model structures, the input data for the Gaussian 94 calculations, and the calculated values for the Fermi contact spin densities have been deposited at BioMagResBank (URL: http://www.bmrb.wisc.edu) under the following accession numbers (in parentheses): oxidized rubredoxin models (4118); reduced rubredoxin models (4119). Except as noted above, all calculations were performed at B3LYP/6-311G**.

‡ Graduate Program in Biophysics.

§ Department of Chemistry.

⊥ Department of Biochemistry.

(1) Eggink, G.; Lageveen, R. G.; Altenburg, B.; Witholt, B. *J. Biol. Chem.* **1987**, *262*, 17712−17718.

(2) Eggink, G.; Engel, H.; Variend, G.; Terpstra, P.; Witholt, B. *J. Mol. Biol.* **1990**, *212*, 135−142.

(3) Kok, M.; Oldenhuis, R.; van der Linden, M. P. G.; Raatjes, P.; Kingma, J.; van Lelyveld, P. H.; Witholt, B. *J. Biol. Chem.* **1989**, *264*, 5435−5441.

(4) Watenpaugh, K. D.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1979**, *131*, 509−522.

(5) Watenpaugh, K. D.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1980**, *138*, 615−633.

(6) Frey, M.; Sieker, L. C.; Payan, F.; Haser, R.; Bruschi, M.; Pepe, G.; LeGall, J. *J. Mol. Biol.* **1987**, *197*, 525−541.

(7) Stenkamp, R. E.; Sieker, L. C.; Jensen, L. H. *Proteins Struct. Funct. Genet.* **1990**, *8*, 352−364.

(8) Adman, E. T.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1991**, *217*, 337−352.

(9) Dauter, Z.; Wilson, K. S.; Sieker, L. C.; Moulis, J. M.; Meyer, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8836−8840.

S0002-7863(97)03489-6 CCC: $15.00
Published on Web 05/20/1998

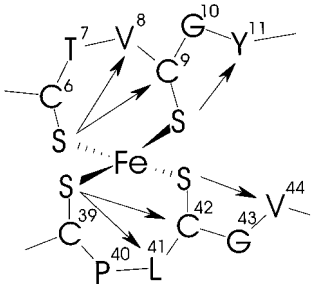

Figure 1. Two-dimensional representation of the covalent (lines) and H-bonding (arrows) network for the residues included in the simplified structural models of the iron−sulfur site in rubredoxin. The S^{γ} atoms of the four cysteines are denoted by "S"; the other letters are conventional single-letter amino acid designations.

Rubredoxins provide attractive targets for detailed theoretical investigations, owing to the simplicity of their iron environment (relative to other iron−sulfur proteins) and the wealth of experimental data available for comparison. Previous theoretical studies on iron−sulfur proteins have focused primarily on calculations of the charge distribution and redox potential of the small synthetic rubredoxin analogues[22−25] studied by Holm and co-workers.[26,27] Our interest lies in studying the remarkably large chemical shifts that have been measured for ^{1}H, ^{2}H, ^{13}C, and ^{15}N nuclei near the iron center, reflecting conjugative interactions with the paramagnetic center.[16] These shifts are evidently an extremely sensitive probe of protein structure and function, and a more complete theoretical understanding of their electronic origin could considerably advance the utility of NMR investigations of this class of proteins.

The *C. pasteurianum* rubredoxin apoprotein has been overproduced in *Escherichia coli* and reconstituted *in vitro*[16] to obtain the recombinant holoprotein, which is abbreviated here as "Rdx". Rdx samples were labeled uniformly with ^{13}C and/or ^{15}N or selectively with ^{2}H, ^{13}C, and ^{15}N, leading to full assignments for the diamagnetic ^{1}H, ^{13}C, and ^{15}N signals[28,29] and to partial assignments for paramagnetically shifted ^{1}H, ^{2}H, ^{13}C, and ^{15}N resonances[16] in both the Fe(III) and Fe(II) forms of the protein. This wealth of NMR data, exhibiting sensitive dependence on minor changes in local structure, provide an excellent framework for testing the utility of quantum mechanical calculations to predict realistic NMR spectra from independently determined structural models.

In paramagnetic molecules, the spatial distribution of electrons of "up" (α) and "down" (β) spin is generally different, leading to nonzero spin densities distributed throughout the system. For atomic orbitals with zero orbital angular momentum (s orbitals), the characteristic nonvanishing wave function amplitude ("cusp") at the origin leads to nonzero spin density (a net imbalance between α and β spin) at the position of the nucleus. Positive spin density at the nucleus causes a shift of the NMR signal of that nucleus to higher frequency (downfield), while negative spin density shifts the peak to lower frequency (upfield). Atomic orbitals originating from neighboring nuclei can also contribute spin density at the position of the nucleus of interest. This isotropic shift mechanism, termed the Fermi contact interaction,[30] is the focus of the present study.

Alternative approaches have been developed for utilizing the structural information inherent in electron−nuclear (hyperfine) interactions in paramagnetic biological macromolecules. Several recent investigations have exploited through-space metal-centered dipolar interactions (pseudocontact interactions) including studies of the docking of cytochrome b_5 with cytochrome *c*,[31] refinement of the solution structure of horse heart ferricytochrome *c* by the addition of paramagnetic dipolar constraints,[32] and investigation of conformational changes in horseradish peroxidase induced by substrate binding.[33]

Much work has gone into developing techniques for calculating Fermi contact coupling constants based on *ab initio*[34−37] and density functional methods,[34,37−43] and the latter have proven particularly efficient for treating transition metal species. Because the isotropic shifts experienced by NMR-active nuclei in the proximity of the iron center of rubredoxin are dominated by the Fermi contact interaction, and since high-resolution structural models are available for this protein, rubredoxin qualifies as an excellent candidate for theoretical investigation. Previous theoretical studies of the contact interaction in smaller species have noted the profound dependence on molecular geometry.[42,43] This sensitivity to geometry suggests the utility of combining Fermi contact shifted NMR data and quantum

(10) Phillips, W. D.; Poe, M.; Weiher, J. F.; McDonald, C. C. *Nature* **1970**, *227*, 574−577.
(11) Krishnamoorthi, R.; Markley, J. L.; Cusanovich, M. A.; Przysiecki, C. *Biochemistry* **1986**, *25*, 50−54.
(12) Werth, M. T.; Kurtz, D. M., Jr.; Moura, I.; LeGall, J. *J. Am. Chem. Soc.* **1987**, *109*, 273−275.
(13) Blake, P. R.; Park, J. B.; Bryant, F. O.; Aono, S.; Magnuson, J. K.; Eccleston, E.; Howard, J. B.; Summers, M. F.; Adams, M. W. W. *Biochemistry* **1991**, *30*, 10885.
(14) Blake, P. R.; Park, J. B.; Zhou, Z. H.; Hare, D. R.; Adams, M. W. W.; Summers, M. F. *Protein Sci.* **1992**, *1*, 1508.
(15) Blake, P. R.; Park, J.-B.; Adams, M. W. W.; Summers, M. F. *J. Am. Chem. Soc.* **1992**, *114*, 4931−4933.
(16) Xia, B.; Westler, W. M.; Cheng, H.; Meyer, J.; Moulis, J. M.; Markley, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 5347−5350.
(17) Peisach, J.; Blumberg, W. E.; Lode, E. T.; Coon, M. J. *J. Biol. Chem.* **1971**, *246*, 5877−5881.
(18) Lode, E. T.; Coon, M. J. *J. Biol. Chem.* **1971**, *246*, 791.
(19) Rao, K. K.; Evans, M. C. W.; Cammack, R.; Hall, D. O.; Thompson, C. L.; Jackson, P. J.; Johnson, C. E. *Biochem. J.* **1972**, *129*, 1063−1070.
(20) Schulz, C.; Debrunner, P. G. *J. Phys.* **1976**, *37*, C6-153-C6-158.
(21) Bennett, D. E.; Johnson, M. K. *Biochim. Biophys. Acta* **1987**, *911*, 71.
(22) Bair, R. A.; Goddard, W. A., III. *J. Am. Chem. Soc.* **1978**, *100*, 5669−5676.
(23) Noodleman, L.; Norman Jr., J. G.; Osborne, J. H.; Aizman, A.; Case, D. A. *J. Am. Chem. Soc.* **1985**, *107*, 3418−3426.
(24) Moulis, J. M.; Chen, J. L.; Noodleman, L.; Bashford, D.; Case, D. A. *J. Am. Chem. Soc.* **1994**, *116*, 11898−11914.
(25) Koerner, J. B.; Ichiye, T. *J. Phys. Chem. B* **1997**, *101*, 33633-3643.
(26) Holm, R. H.; Ibers, J. A. In *Iron−Sulfur Proteins*; Lovenberg, W., Ed.; Academic Press: New York, 1977; p 205.
(27) Holm, R. H. *Adv. Inorg. Chem.* **1992**, *38*, 1.
(28) Volkman, B. F.; Prantner, A. M.; Wilkens, S. J.; Xia, B.; Markley, J. L. *J. Biomol. NMR* **1997**, *10*, 409−410.
(29) Prantner, A. M.; Volkman, B. F.; Wilkens, S. J.; Xia, B.; Markley, J. L. *J. Biomol. NMR* **1997**, *10*, 411−412.
(30) Bertini, I.; Luchinat, C. In *NMR of Paramagnetic Molecules in Biological Systems*; Benjamin Cummings: Menlo Park, CA, 1986.
(31) Guiles, R. D.; Sarma, S.; Digate, R. J.; Banville, D.; Basus, V. J.; Kuntz, I. D.; Waskell, L. *Nat. Struct. Biol.* **1996**, *3*, 333−339.
(32) Qi, P. X. R.; Beckman, R. A.; Wand, A. J. *Biochemistry* **1996**, *35*, 5(38), 12275−12286.
(33) de Ropp, J. S.; Mandal, P.; Brauer, S. L.; La Mar, G. N. *J. Am. Chem. Soc.* **1997**, *119*, 4732−4739.
(34) Lim, M. H.; Worthington, S. E.; Dulles, F. J.; Cramer, C. J. *ACS Symp. Ser.* **1996**, *629*, 402−422.
(35) Cramer, C. J.; Lim, M. H. *J. Phys. Chem.* **1994**, *98*, 5024−5033.
(36) Barone, V.; Adamo, C.; Grand, A.; Subra, R. *Chem. Phys. Lett.* **1995**, *246*, 53−58.
(37) Cohen, M. J.; Chong, D. P. *Chem. Phys. Lett.* **1995**, *234*, 405−412.
(38) Malkina, O. L.; Salahub, D. R.; Malkin, V. G. *J. Chem. Phys.* **1996**, *105*, 8793−8800.
(39) Barone, V.; Adamo, C.; Russo, N. *Chem. Phys. Lett.* **1993**, *212*, 5−11.
(40) Adamo, C.; Barone, V.; Fortunelli, A. *J. Chem. Phys.* **1995**, *102*, 384−393.
(41) O'Malley, P. J.; Collins, S. J. *Chem. Phys. Lett.* **1996**, *259*, 296−300.

Table 1. Regression Analysis of Experimental Chemical Shifts vs Calculated $\rho_{\alpha\beta}$ (B3LYP/6-311G** 104-Atom Model). $\rho_{\alpha\beta}$ Calculated at B3LYP/6-311G** from 104-Atom Models

calculation			^{2}H			^{15}N		
104-atom model used in the calculation and its iron oxidation state	Fe(III) rubredoxin structure (reference)[a]	level of the calculation	R^2	slope (std dev)/ 10^5 ppm au^{-1}	intercept (std dev)/ppm	R^2	slope (std dev)/ 10^5 ppm au^{-1}	intercept (std dev)/ppm
5RXN-M, Fe(III)	5RXN (4, 5)	B3LYP/6-311G**	0.94	2.49 (0.19)	43.6 (32.9)	0.93	2.37 (0.20)	162.4 (24.9)
5RXN-G$_{\alpha\beta}$, Fe(III)[b]	5RXN (4, 5)	B3LYP/6-311G**	0.95	2.24 (0.16)	42.6 (30.9)	0.94	2.41 (0.20)	157.0 (24.4)
1IRO-M, Fe(III)	1IRO (9)	B3LYP/6-311G**	0.96	2.64 (0.17)	31.1 (28.0)	0.85	2.86 (0.38)	148.0 (39.2)
1IRO-G$_{HN}$, Fe(III)[c]	1IRO (9)	B3LYP/6-311G**	0.95	2.57 (0.19)	39.6 (31.9)	0.94	2.46 (0.20)	1336 (25.2)
4RXN-M, Fe(III)	4RXN (4, 5)	B3LYP/6-311G**	0.95	2.61 (0.20)	48.3 (31.0)	0.96	2.38 (0.15)	144.8 (18.7)
4RXN-M, Fe(III)	4RXN (4, 5)	B3LYP/3-21G**	0.97	3.62 (0.21)	22.7 (25.9)	0.85	1.28 (0.17)	309.1 (27.1)
4RXN-M, Fe(II)[d]	4RXN (4, 5)	B3LYP/6-311G**	0.93	2.21 (0.19)	21.6 (13.4)	0.75	1.61 (0.29)	129.5 (32.1)
5RXN-M, Fe(II)[d]	5RXN (4, 5)	B3LYP/6-311G**	0.94	2.13 (0.17)	19.3 (12.6)	0.72	1.51 (0.30)	137.8 (33.6)

[a] Protein Data Bank accession code for the X-ray structure used in generating the simplified structural model for the iron center. [b] Geometry-optimized cysteine α and β protons. [c] Geometry optimized backbone amide protons. [d] Model Fe(III) geometry used to calculate chemical shifts for Fe(II) rubredoxin for comparison with experimental NMR data (see the text).

chemical methods to refine protein structures, particularly in cases for which NMR techniques relying on NOE or scalar coupling are ineffective (e.g., due to relaxation effects caused by electron−nuclear interactions). The development of such a combined theoretical/experimental methodology for iron−sulfur proteins is the goal of the present work.

Methods

Structural Models. The input structures for the quantum chemical calculations were derived from three high-resolution X-ray crystal structures of oxidized *C. pasteurianum* rubredoxin. The resolution in all three was sufficiently high that hydrogens were included in the structural models deposited in the Protein Data Bank (URL: http://www.pdb.bnl.gov).[44,45] Two structures were derived from data collected at 1.2 Å resolution: in one model (5RXN), structural constraints had been imposed during refinement, whereas in the other (4RXN), no restraints were imposed during the refinement process.[4,5] The third structure (1IRO) was derived from data at 1.1 Å resolution which had been refined using the SHELXL-93 program.[9] Each of these coordinate sets was truncated in an identical fashion to produce the three, 104-atom, initial models of oxidized rubredoxin used for the calculations: 5RXN-M, 4RXN-M, and 1IRO-M (Table 1).

The X-ray structure of the corresponding reduced Fe(II) *C. pasteurianum* rubredoxin has been solved at 0.9 Å resolution, and the structures of the oxidized and reduced protein are reported to be highly similar,[9] but the coordinates are not yet available for the latter. In the absence of an experimentally derived structure for reduced rubredoxin, the 4RXN-M and 5RXN-M geometries (see Table 1 for nomenclature of rubredoxin models) for Fe(III) rubredoxin were therefore taken as provisional structural models for the reduced species (Table 1).

The goal in selecting a tractable model structure for the calculations was to make the model as simple as possible, while retaining enough atoms to produce a realistic electronic environment at the atoms of interest, including all nuclei known be characterized by large, hyperfine NMR shifts. Figure 1 shows a simple two-dimensional picture of the residues included in the chosen model, with their covalent and hydrogen bond connectivities. Coordinates for the amino acid residues and iron atom for each model were extracted from the respective PDB file from the Protein Data Bank. The model consisted of the iron center and two hexapeptide chains, Cys^6-Thr^7-Val^8-Cys^9-Gly^{10}-Tyr^{11} and Cys^{39}-Pro^{40}-Leu^{41}-Cys^{42}-Gly^{43}-Val^{44}, each containing two cysteines ligated to the iron. A PDB file was generated containing the coordinates of the residues of interest plus Thr^5 and Val^{38}. InsightII (Molecular Simulations) was employed to convert Thr^7, Val^8, and Leu^{41} into glycine residues and to truncate residues 5 and 38 to formyl groups (by converting their C$^\alpha$ atoms to hydrogens) in order to retain a realistic electronic environment for the N-terminus of each peptide chain. The C-terminal residues of each chain, Val^{44} and Tyr^{11}, were truncated to *N*-methyl groups by converting their C′ and C$^\beta$ atoms to hydrogens. To simplify the model further, the C$^\beta$ and C$^\gamma$ atoms of Pro^{40} were converted to hydrogens to yield an *N*-methylglycine residue. Once completed, the simplified models contained a total of 104 atoms each. The models were then converted to **Z**-matrix format using Xmol (Network Computing Services, Inc., 1993) and visualized using Molden.[46]

Basis Sets and Computational Strategy. Computational efficiency was a particular concern with the 104-atom models. For systems of this size, perturbative and multiconfigurational correlation methods were impractical. Therefore, the spin-unrestricted B3LYP[47] hybrid density functional approach (as implemented in Gaussian 94[48,49]) was employed for all calculations, allowing the effects of spin polarization and electron correlation to be included at a reasonable computational cost. Our results support previous conclusions that the B3LYP hybrid density functional approach offers a good combination of accuracy and computational efficiency,[37] especially when used in combination with the 6-311G** basis set.[34]

Using a locally dense methodology that places a greater number of basis functions on atoms of particular interest and fewer basis functions on those atoms of lesser importance may prove to be a cost saving measure. Such a technique was not attempted here, because we were investigating the utility of combining quantum chemical methods in conjunction with X-ray crystallographic data in order to gain insight into hyperfine interactions measured by NMR. Moreover, since electron density is delocalized through several covalent and hydrogen bonds, it seemed likely that the use of a well-balanced basis set is important in achieving an accurate description of the overall delocalization pattern.

Achieving convergence on such a large model in a timely manner required special measures. Two factors contributed to difficulties in

(42) Ellinger, Y.; Pauzat, F.; Barone, V.; Douady, J.; Subra, R. *J. Chem. Phys.* **1980**, *72*, 6390−6397.
(43) Barone, V.; Grand, A.; Minichino, C.; Subra, R. *J. Chem. Phys.* **1993**, *99*, 6787−6798.
(44) Abola, E. E.; Bernstein, F. C.; Bryant, S. H.; Koetzle, T. F.; Weng, J. In *Crystallographic Databases-Information Content, Software Systems, Scientific Applications*; Allen, F. H., Bergerhoff, G., Sievers, R., Eds.; Data Commission of the International Union of Crystallography: Bonn/Cambridge/Chester, 1987; pp 107−132.
(45) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, Jr. E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.
(46) Schaftenaar, G. *QCPE* **1992**, *12*.
(47) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(48) *Gaussian 94 (Revision D.4)*, Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. J., Cheeseman, J. R., Keith, T. A., Petersson, G. A., Montgomery, J. A., Raghavachari, K., Al-Laham, M. A., Zakrzewski, V. G., Ortiz, J. V., Foresman, J. B., Cioslowski, J., Stefanov, B. B., Nanayakkara, A., Challacombe, M., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, J. L., Replogle, E. S., Gomperts, R., Martin, R. L., Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. P., Head-Gordon, M., Gonzalez, C., Pople, J. A., Gaussian, Inc., Pittsburgh, PA, 1995.
(49) The B3LYP functional employed in Gaussian 94 is slightly different from the one given in ref 48. The form used is in Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

convergence. The first obstacle to be overcome was a near degeneracy in HOMO and LUMO energies, which led to oscillatory convergence behavior. Such oscillations often can be damped by shifting the virtual orbitals up in energy so as to break the near degeneracy. We found a virtual orbital shift of 100 mhartrees sufficient to damp oscillations in the convergence algorithm. The second factor contributing to difficulty in wave function convergence was the poor initial semiempirical guess by Gaussian 94 at the form of the final wave function. To minimize the number of SCF iterations performed when using the 6-311G** basis set (1380 basis functions), we bootstrapped from calculations performed at lower levels of theory. This was done by reading in a converged B3LYP/3-21G* wave function as the initial guess for a B3LYP/3-21G** calculation, which was in turn read in as an initial guess for the final calculation done at B3LYP/6-311G**. Therefore, at each step in the calculation, hybrid density functional wave functions were used as initial guesses instead of vastly inferior semiempirical guesses. The final calculation of 20 SCF iterations required approximately 500 CPU hours on a Silicon Graphics Onyx workstation (six 195 MHz R10000 CPUs).

Theory. The total observed chemical shift can be factored into contributions from Fermi contact, diamagnetic, metal-centered dipolar (metal-centered pseudocontact), and ligand-centered dipolar (ligand-centered pseudocontact) terms as shown below:[30]

$$\delta^{\mathrm{tot}} = \delta^{\mathrm{con}} + \delta^{\mathrm{dia}} + \delta^{\mathrm{dip}}_{\mathrm{MC}} + \delta^{\mathrm{dip}}_{\mathrm{LC}} \tag{1}$$

For systems containing high-spin iron, the Fermi contact term δ^{con} usually is much larger than the others.[50,51] This separation in scale of coupling strengths greatly reduces the complexity involved in analyzing the hyperfine interaction by allowing one to neglect both the metal- and ligand-centered pseudocontact terms. Furthermore, Kurland and McGarvey have shown that the zero-field splitting parameter D typically has a negligible effect on the Fermi contact coupling constant.[50] Our results indicate that these approximations hold adequately for oxidized rubredoxin.

The chemical shift δ^{con} that results from a nonzero Fermi contact contribution is proportional to A_{c}, the scalar Fermi contact coupling constant, and S_{c}, the thermal average of the total electron spin S in the direction of the magnetic field.

$$\delta^{\mathrm{con}} = \left(\frac{\Delta\nu}{\nu_0}\right)^{\mathrm{con}} = -\left(\frac{\Delta B}{B_0}\right)^{\mathrm{con}} = -\left(\frac{A_{\mathrm{c}}}{\hbar\gamma_{\mathrm{n}}B_0}\right)S_{\mathrm{c}} \tag{2}$$

S_{c} is also referred to as the *Curie spin*[52] and is defined in the high-temperature approximation by

$$S_{\mathrm{c}} = -S(S+1)\frac{g_{\mathrm{e}}\mu_{\mathrm{B}}B_0}{3kT} \tag{3}$$

In these equations, $\hbar$ is Planck's constant divided by 2π, γ_{n} is the nuclear gyromagnetic ratio, B_0 is the applied magnetic field, g_{e} is the free electron g-factor, μ_{B} is the Bohr magneton, k is Boltzmann's constant, and T is the absolute temperature.[53]

The hyperfine coupling constant A_{c} is proportional to the spin state S and the difference in the α and β spin densities at the nucleus

$$A_{\mathrm{c}} = \frac{\mu_0}{3S}\hbar\gamma_{\mathrm{n}}g_{\mathrm{e}}\mu_{\mathrm{B}}\sum_{i}^{\mathrm{occ}}[|\psi^{-}{}_i(0)|^2 - |\psi_i^{+(0)}|^2] \tag{4}$$

where i is an index that runs over all occupied molecular orbitals and is evaluated at the position (the origin or "0" of a local coordinate system) of a nucleus of interest. The *contact spin density* $\rho_{\alpha\beta}$ is defined as

$$\rho_{\alpha\beta} = \sum_{i}^{\mathrm{occ}}[|\psi_i^{-}(0)|^2 - |\psi_i^{+}(0)|^2] \tag{5}$$

Substitution of A_{c} and S_{c} into eq 2 yields the desired formula for the contact NMR chemical shift[30]

$$\delta^{\mathrm{con}} = \frac{\mu_0\mu_{\mathrm{B}}{}^2g_{\mathrm{e}}{}^2(S+1)}{9kT}\rho_{\alpha\beta} \tag{6}$$

expressed in terms of spin-up (ψ_i^{+}) and spin-down (ψ_i^{-}) molecular orbitals which can be obtained from quantum chemical calculations.

Equations 1 and 6 can be written in the form

$$\delta^{\mathrm{tot}} = m\rho_{\alpha\beta} + \delta^{\mathrm{res}} \tag{7}$$

where m, the coefficient of the theoretical spin density expression $\rho_{\alpha\beta}$ (in atomic units), is a collection of physical constants

$$m = \frac{\mu_0\mu_{\mathrm{B}}{}^2g_{\mathrm{e}}{}^2(S+1)}{9kT} \tag{8}$$

and δ^{res} represents the residual contribution of noncontact terms ($\delta^{\mathrm{dia}} + \delta^{\mathrm{dip}}_{\mathrm{MC}} + \delta^{\mathrm{dip}}_{\mathrm{LC}}$) in eq 1. At $T = 308$ K, the theoretical value of the coefficient m for oxidized rubredoxin ($S = 5/2$) is

$$m^{\mathrm{theor}}_{\mathrm{OX}} = 2.6743 \times 10^5 \text{ ppm} \tag{9}$$

and that for reduced rubredoxin ($S = 2$) is

$$m^{\mathrm{theor}}_{\mathrm{RE}} = 2.2923 \times 10^5 \text{ ppm} \tag{10}$$

Equations 9 and 10 show that a very small amount of net spin density on a given nucleus will result in a large observable chemical shift for both oxidized and reduced rubredoxin. For example, for an ordinary hydrogenic 1s orbital with amplitude $1/\pi^{1/2}$ (in atomic units) at the nuclear cusp, an imbalance as small as 1% in the density of α and β spins should lead to contact chemical shifts on the order of

$$\delta^{\mathrm{con}} = \frac{(2.67 \times 10^5)}{3.14}(0.01) \approx 850 \text{ ppm} \tag{11}$$

for an isotropic $S = 5/2$ system at 308 K.

It is known that contracted Gaussian basis functions, although suitable for describing the valence region, cannot accurately describe the nuclear cusp.[54] Thus, evaluations of δ^{con} based directly on eq 6 are likely to incur a systematic overall error due to the incorrect amplitudes of ψ_i^{+} and ψ_i^{-} at the nuclear cusp. This absolute error is ordinarily of little chemical consequence, since the molecular orbitals are expected to show the proper *variations* of cusp amplitude with changes in chemical environment. However, this systematic error of the Gaussian basis set results in a change in the "apparent" value of the coefficient m in eq 8, with respect to its ideal theoretical value. We can therefore account for this systematic difference by expressing δ^{tot} as shown in eq 7 and finally compare m_{exp} (obtained as the slope in a direct linear regression fit with experimental hyperfine chemical shifts) with the "ideal" values from eqs 9 and 10. Similarly, the apparent $\delta^{\mathrm{res}}_{\mathrm{exp}}$ obtained as the intercept in such a regression can be compared with the ideal diamagnetic shifts in order to assess the importance of neglected dipolar contributions and other systematic error. Data from the regression analysis of each model structure are given in Table 1.

Results and Discussion

Hyperfine-Shifted ^{2}H and ^{15}N Signals of Fe(III) Rubredoxin. Figure 2 shows linear regression fits of the experimental ^{2}H and ^{15}N chemical shifts of oxidized rubredoxin versus the

(50) Kurland, R. J.; McGarvey, B. R. *J. Magn. Reson.* **1970**, *2*, 286−301.

(51) Shporer, M.; Ron, G.; Loewenstein, A.; Navon, G. *Inorg. Chem.* **1965**, *4*, 358−361.

(52) Gueron, M. *J. Magn. Reson.* **1975**, *19*, 58−66.

(53) Standard values for the physical constants can be found at http://physics.nist.gov/PhysRefData/contents.html.

(54) Hehre, W. J.; Radon, L.; Schleyer, P.v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986; p 19.

Table 2. Assignments for the Hyperfine Shifted ^{2}H and ^{15}N Peaks from Oxidized *C. pasteurianum* Rubredoxin[a]

^{2}H chemical shift assignments[b]				^{15}N chemical shift assignments[c,d]			
NMR data	structural model for the iron center			NMR data	structural model for the iron center		
^{2}H shifts/ppm	4RXN-M	5RXN-M 5RXN-$G_{\alpha\beta}$	1IRO-M 1IRO-G_{HN}	15N shifts/ppm	4RXN-M	5RXN-M	1IRO-M 1IRO-G_{HN}
−11.5	C^{39} H^{α}	C^{6} H^{α}	C^{39} H^{α}	−25.1	C^{42}	C^{9}	C^{42}
−11.5	C^{6} H^{α}	C^{39} H^{α}	C^{6} H^{α}	−68.8	C^{9}	C^{42}	C^{9}
150.4	C^{42} H^{α}	C^{9} H^{α}	C^{9} H^{α}	305.9	$\underline{G^{43}}$	G^{10}	G^{43}
177.3	C^{9} H^{α}	C^{42} H^{α}	C^{42} H^{α}	319.5	$\underline{G^{10}}$	G^{43}	$\underline{G^{10}}$
376.9	C^{39} $H^{\beta 2}$	C^{39} $H^{\beta 2}$	C^{39} $H^{\beta 2}$	336.6	$\underline{V^{44}}$	$\underline{V^{44}}$	$\underline{V^{44}}$
406.7	C^{6} $H^{\beta 2}$	C^{6} $H^{\beta 2}$	C^{6} $H^{\beta 2}$	351.0	$\underline{T^{7}}$	$\underline{T^{7}}$	$\underline{T^{7}}$
437.0	C^{39} $H^{\beta 3}$	C^{39} $H^{\beta 3}$	C^{39} $H^{\beta 3}$	398.2	$\underline{P^{40}}$	$\underline{P^{40}}$	$\underline{P^{40}}$
478.5	C^{6} $H^{\beta 3}$	C^{6} $H^{\beta 3}$	C^{6} $H^{\beta 3}$	508.9	$\underline{V^{8}}$	$\underline{V^{8}}$	$\underline{V^{8}}$
520.6	C^{9} $H^{\beta 3}$	C^{9} $H^{\beta 3}$	C^{9} $H^{\beta 3}$	521.3	$\underline{L^{41}}$	$\underline{L^{41}}$	$\underline{L^{41}}$
557.4	C^{42} $H^{\beta 3}$	C^{42} $H^{\beta 3}$	C^{42} $H^{\beta 3}$	571.1	$\underline{C^{6}}$	C^{6}	$\underline{C^{39}}$
771.7	C^{9} $H^{\beta 2}$	C^{9} $H^{\beta 2}$	C^{9} $H^{\beta 2}$	594.4	C^{39}	C^{39}	$\underline{C^{6}}$
771.7	C^{42} $H^{\beta 2}$	C^{42} $H^{\beta 2}$	C^{42} $H^{\beta 2}$	606.0	Y^{11}	Y^{11}	$\underline{Y^{11}}$

[a] Assignments take into account prior information about residue type, atom type, and assignment and provide maximal agreement with chemical shift calculations (B3LYP/6-311G** 104-atom model). [b] All correlations are in agreement with prior assignment of ^{1}H and ^{2}H peaks of cysteines to H^{α}, $H^{\beta 2}$, and $H^{\beta 3}$ positions (but not to specific cysteine residues) by selective labeling. [c] Underlined residue designations indicate correlations that were altered in order to be consistent with experiments involving selectively labeled samples. [d] Only the 5RXN-M and 5RXN-$G_{\alpha\beta}$ models were affected by the V^{44} backbone ^{15}N resonance assignment.

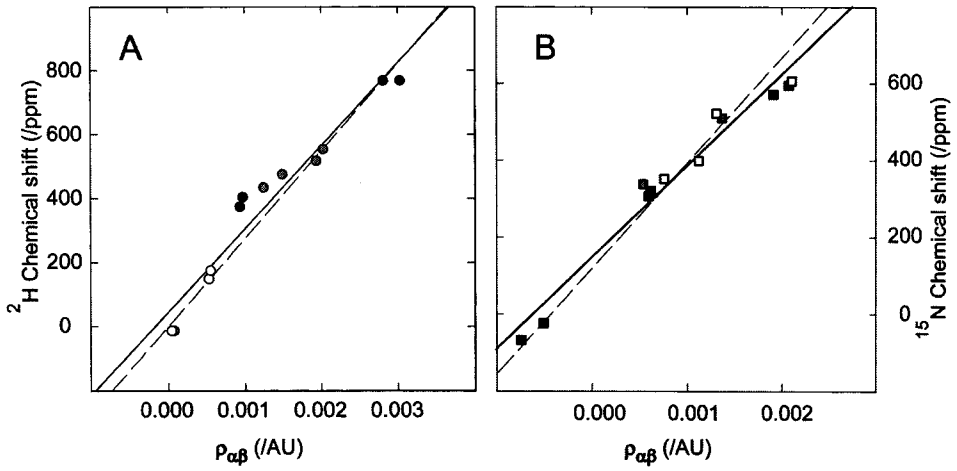

Figure 2. Linear regression fits of experimental chemical shifts of the rubredoxin from *C. pasteurianum* (in ppm) versus $\rho_{\alpha\beta}$ (au) calculated from a 104-atom simplified model of the active site of the iron−sulfur protein (4RXN-M) used as input to Gaussian 94 with the iron as Fe(III). Atom positions in this model were taken from one of the X-ray structures of *C. pasteurianum* Fe(III) rubredoxin (4RXN).[4,5] The solid lines indicate best fits to the data, and the dashed lines represent the theoretical slopes and intercepts. (A) Cysteine α- and β-^{2}H chemical shifts assigned experimentally to the H^{α} (black circles), $H^{\beta 2}$ (open circles), and $H^{\beta 3}$ positions (gray circles). The best fit is in agreement with these assignments. (B) Backbone ^{15}N chemical shifts. The best fit shown takes into account prior residue-specific assignments to T^7, Y^{11}, P^{40}, and L^{41} (black squares) with V^{44} and V^8 shown separately (gray squares), and identifications by residue type for the remaining ^{15}N NMR signals (open squares). The imposition of this information forced less optimal correlations for the $\rho_{\alpha\beta}$ values for T^7, V^8, G^{10}, L^{41}, G^{43}, and V^{44} (Table 2) with the experimental chemical shifts (and hence degraded the R^2 value for the fit).

calculated $\rho_{\alpha\beta}$ (for the 4RXN-M structure). The experimental ^{2}H and ^{15}N chemical shifts were ranked in ascending order and plotted against the ascending order of calculated $\rho_{\alpha\beta}$. Theoretical chemical shifts from the contact spin densities were not calculated in this correlation to avoid the introduction of assumptions. All ^{15}N resonances have been identified experimentally by amino acid type, and a few have been assigned to individual residues in the sequence (Xia, B., et al., manuscript to be published elsewhere). In addition, for oxidized rubredoxin resonance assignments to V^8 and V^{44} have been made by reference to the ^{15}N NMR spectrum of a chemically synthesized rubredoxin containing ^{15}N only in the amide of V^{44} (unpublished data). The cysteine ^{2}H signals have been classified as H^{α}, $H^{\beta 2}$, or $H^{\beta 3}$ on the basis of selective labeling of the protein.[16] All known information about residue type, atom type, and assignment was imposed prior to ranking the calculated $\rho_{\alpha\beta}$ with experimental data. In no case did this procedure alter the ordering of the cysteine H^{α}, $H^{\beta 2}$, or $H^{\beta 3}$ chemical shifts, but in some cases, it changed the order of correlation between experimental ^{15}N chemical shifts and those calculated from certain models (resulting in lower correlation coefficients).

The solid line in each of the graphs corresponds to a linear regression fit of the data points, while the dashed line represents a theoretical line with a slope (267 430 ppm) obtained from eq 7. The *y*-intercept for each plot in Figure 2 represents the diamagnetic chemical shift (zero Fermi contact contribution) with respect to the reference compound for the given nuclear species, in all cases DSS, with indirect referencing for nuclei other than ^{1}H.[55] For Figure 2A, the theoretical *y*-intercept for the dotted line was obtained from a weighted average of H^{α}

and H^{β} chemical shifts for cysteine (3.6 ppm). The theoretical *y*-intercept for the dotted lines in Figure 2B was obtained from a weighted average of the average diamagnetic ^{15}N chemical shifts for the amino acid types used in the model (118.7 ppm). Values for the diamagnetic shifts were obtained from BioMagResBank (URL: http://www.bmrb.wisc.edu/).[56] Fit parameters and chemical shift assignments derived from ranking $\rho_{\alpha\beta}$ are summarized in Tables 1 and 2 for all model structures. Results for individual types of nuclei are discussed below.

Cysteine ^2H Resonances of Fe(III) Rubredoxin. Figure 2A shows a linear regression fit of the experimental cysteine $^2H^{\alpha}$ and $^2H^{\beta}$ chemical shifts vs the calculated $\rho_{\alpha\beta}$ for the 4RXN-M structure. Similar results were obtained with the other two models (5RXN-M and 1IRO-M). The pairings of ^2H NMR peaks with hydrogen atoms in the structure that yielded optimal linear regressions were in full agreement with assignments derived from selective labeling with [$^2H^{\beta2,\beta3}$]cysteine and chiral labeling with [$^2H^{\beta2}$]cysteine. All three models were consistent with the same assignments for the cysteine $H^{\beta2}$ and $H^{\beta3}$ resonances, but not for the cysteine H^{α} resonances. These latter differences are present despite the minor geometrical differences between the models (0.063 Å RMS deviation between 1IRO-M, 4RXN-M, and 5RXN-M structures at the H^{α}, $H^{\beta2}$, and $H^{\beta3}$ positions). For all three model structures, the slopes are systematically lower (by 5−10%) than predicted by eq 8, and the *y*-intercepts are also significantly larger than the average diamagnetic shift of 3.6 ppm.

Because all atoms of the model are in close proximity to the iron center, the metal-centered pseudocontact contribution to the chemical shift could be significant. Since the orientations of the magnetic susceptibility and **g** tensors are unknown, an exact value for the metal-centered pseudocontact shift cannot be determined. Although the point-dipole approximation breaks down for nuclei close to the iron, this approach can be used to estimate the maximum value for the metal-centered pseudocontact contribution to the chemical shift.[30]

$$\delta_{MC}^{dip} = \frac{\mu_0}{4\pi}\frac{35\mu_B{}^2(g_{||}{}^2 - g_{\perp}{}^2)}{36kT}\frac{3\cos^2\theta - 1}{r^3}\left[1 - \frac{32(g_{||}{}^2 + {}^1/_2 g_{\perp}{}^2)}{15(g_{||}{}^2 - g_{\perp}{}^2)}\frac{D}{kT}\right] \quad (12)$$

For oxidized rubredoxin, $D = 2.53$ K,[21] and for a closely homologous rubredoxin, $g_{\perp} = (g_x + g_y)/2 = 2.01$ and $g_{||} = g_z = 2.09$.[57] Taking $T = 308$ K, the maximum chemical shifts (at θ equal to 0) are found to be approximately 12 ppm for the α proton positions ($r = 4.8$ Å) and 43 ppm for the β positions ($r = 3.1$ Å). These contributions, which are small relative to the Fermi contact interaction, have magnitudes comparable to the uncertainties in the intercepts at $\rho_{\alpha\beta} = 0$ (Table 1).

Contributions to observed chemical shifts from ligand-centered dipolar interaction usually are assumed to be negligibly small for one-electron atoms, because nearly all the electron density resides in 1s orbitals, which have no orbital angular momentum. Dipolar fields originating from net electron spin density on surrounding nuclei can also make contributions to the chemical shift of the protons, although their *net* contribution is likely to be small.

In summary, contributions to the chemical shift from sources other than Fermi contact are expected to be small. The relative significance of such effects may be larger for chemical shifts of hydrogens in α positions than in β positions. The calculations reported here indicate that the sensitivity of $\rho_{\alpha\beta}$ to small structural perturbations is larger than that expected from mechanisms other than Fermi contact; thus, slight errors in the placement of the protons during the X-ray refinement process may be a greater source of disagreement with experiment than the neglect of other contributions to the chemical shift.

^{15}N Resonances of Fe(III) Rubredoxin. The linear regression fit of experimental ^{15}N chemical shifts vs calculated $\rho_{\alpha\beta}$ for the 4RXN-M structure is shown in Figure 2B. The chemical shift assignments for the three model structures are listed in Table 2, and the fitted parameters are given in Table 1. It should be noted the greater number of missed assignments for ^{15}N resonances compared to ^2H is likely to be primarily due to the greater number of accumulated errors in nitrogen positions in comparison to the cysteine hydrogens. Models derived from the two coordinate sets deposited by Watenpaugh and coworkers (4RXN and 5RXN) yielded calculated shifts that agreed closely with experiment, whereas that from the independently determined structure (1IRO) yielded chemical shifts that showed significantly poorer correlation ($R^2 = 0.85$) with experiment. The reason for the poor correlation of the one model with experiment is discussed below; discussion here is restricted to results with the 4RXN-M ($R^2 = 0.96$) and 5RXN-M ($R^2 = 0.93$) structures.

Interestingly, the optimal assignment of ^{15}N chemical shifts was different for two of the structural models for Fe(III) rubredoxin (4RXN and 5RXN): rankings were swapped for two residue pairs: Gly^{43} and Gly^{10}, and Cys^{42} and Cys^{9} (Table 2). Thus, the independent assignment of these signals is of critical interest, since it will allow further refinement of the correlations. In addition, the calculated spin density for Val^8 is less than that of Val^{44} for the 5RXN-M structure in disagreement with experiment. Imposing the correct order on the valine assignments before fitting lowered the R^2 slightly (from 0.94 to 0.93). We are in the process of synthesizing rubredoxin peptides with individual glycine and cysteine labels to resolve the remaining uncertainties in the resonance assignments.

Effects of Optimizing the Positions of the Cysteine α- and β-Hydrogen Atoms. In the refinement of the X-ray structures of Fe(III) rubredoxin used in this study, the authors had positioned each hydrogen on the basis of the coordinates of the heavy-atom framework. They then employed the resulting structural model in calculating the structure factors. To investigate possible errors introduced by this method, the bond angles and distances for the α- and β-hydrogens of the four cysteines in the 5RXN-M simplified model were optimized at B3LYP/3-21G* while the remaining atoms in the model cluster were held fixed. This procedure led to very small changes in the coordinates for the hydrogen atoms in the resulting model, 5RXN-$G_{\alpha\beta}$. The $\rho_{\alpha\beta}$ values calculated from model 5RXN-$G_{\alpha\beta}$ (with optimized positions for cysteine α- and β-hydrogens) correlated only slightly better with the experimental ^2H NMR chemical shifts than those from the unoptimized model (Table 1), and the refinement had only small effects on values for the fitted slope and intercept (Table 1). Interestingly, the quality of the ^{15}N fit improved slightly after optimizing the cysteine α- and β-hydrogens.

(55) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135−140.

(56) Ulrich, E. L.; Argentar, D. R.; Manabat, N. C.; Ioannidis, Y. E.; Livny, M.; Markley, J. L. *FASEB J.* **1997**, *11*, 1553.

(57) Hearshen, D. O.; Hagen, W. R.; Sands, R. H.; Grande, H. J.; Crespi, H. L.; Gunsalus, I. C.; Dunham, W. R. *J. Magn. Reson.* **1986**, *69*, 440−459.

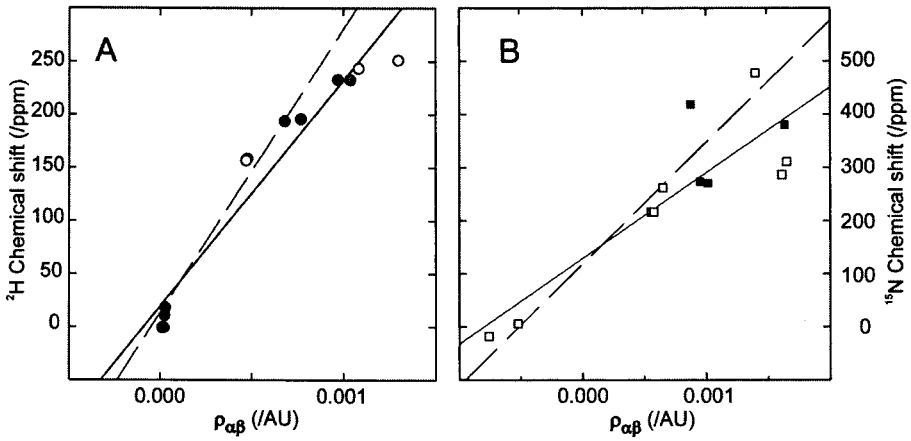

Figure 3. Linear regression fits of experimental chemical shifts of the rubredoxin from *C. pasteurianum* (ppm) versus $\rho_{\alpha\beta}$ (au) calculated from a 104-atom simplified model of the active site of the iron−sulfur protein (4RXN-M) used as input to Gaussian 94 with the iron as Fe(II). Atom positions in this model were taken from one of the X-ray structures of *C. pasteurianum* Fe(III) rubredoxin (4RXN).[4,5] The solid lines indicate the fits to the data, and the dashed lines represent the theoretical slopes and intercepts. (A) Cysteine α- and β-^{2}H chemical shifts assigned experimentally to the H^{α} (black circles), $H^{\beta 2}$ (open circles), and $H^{\beta 3}$ positions (gray circles). The best fit is in agreement with these assignments. (B) Backbone ^{15}N chemical shifts. The best fit shown takes into account prior residue-specific assignments to T^7, Y^{11}, P^{40}, and L^{41} (black squares) and identifications by residue type for the remaining ^{15}N NMR signals (open squares). The imposition of this information forced less optimal correlations for the $\rho_{\alpha\beta}$ values for C^6, V^8, Y^{11}, C^{39}, P^{40}, and L^{41} (Table 3) with the experimental chemical shifts (and hence degraded the R^2 value for the fit).

Effect of Optimizing the Positions of the Backbone Amide Hydrogens in the 1IRO Structural Model. Attempts were made to identify the reason results calculated from the model derived from the 1IRO X-ray structure gave much poorer agreement with experiment than those derived from the 4RXN and 5RXN structures (Table 1). The three structures were superimposed on each other in Insight II (Molecular Simulations), and the RMS deviations between various atomic coordinates and bond distances were calculated. Although the coordinates for the backbone heavy atoms were quite similar (RMS deviation for nitrogen of 0.042 Å), the 1IRO structure showed a systematic difference in the backbone amide proton to nitrogen bond distances. This distance for all such protons in the 1IRO structure was approximately 0.86 Å, considerably smaller than the average value of 1.00 Å found in the 4RXN and 5RXN structures. Consequently, the bond lengths and angles of the eleven backbone amide protons (excluding Pro^{40} which has no amide proton) of all three simplified models were optimized at B3LYP/3-21G* while the coordinates of the other atoms were held fixed. Not surprisingly, the bond lengths of the optimized protons converged to an average length of 1.02 Å in close agreement with the 4RXN and 5RXN structures, while the bond angles changed little (RMS deviation in position of 0.109 Å). The structural model resulting from this refinement of model 1IRO-M model is denoted 1IRO-G_{HN}.

Calculated $\rho_{\alpha\beta}$ values derived from model 1IRO-G_{HN} correlated significantly better with experimental ^{15}N chemical shifts ($R^2 = 0.94$) than values derived from the nonoptimized 1IRO-M model ($R^2 = 0.85$). This result indicates that these calculations are capable of detecting and rectifying errors in the molecular structures of paramagnetic molecules.

Hyperfine-Shifted ^{2}H and ^{15}N Signals of Fe(II) Rubredoxin. No high-resolution structure of reduced *C. pasteurianum* rubredoxin is currently available. As a first approximation, the models that gave the best fit to the experimental NMR data for Fe(III) rubredoxin (5RXN and 4RXN) were used in calculating $\rho_{\alpha\beta}$ for Fe(II) rubredoxin. Although the correlation with ^{2}H chemical shifts (Figure 3A) was nearly as good as that obtained for Fe(III) rubredoxin, that for ^{15}N chemical shifts (Figure 3B) was clearly worse. The optimal ranking of ^{2}H chemical shifts also was in agreement with the atom type identifications derived from selective labeling.[16] We interpret this as indicating that the positions of the cysteine ligand side chain atoms do not change significantly when Fe(III) rubredoxin is reduced; on the other hand, the poor fit of the ^{15}N data suggests the existence of conformational differences or other small structural variations between the oxidized and reduced rubredoxin that affect the positions of backbone nitrogens. We consider it highly desirable that these calculations be repeated when coordinates for the reduced X-ray crystal structure are made available.

Table 3. Assignments for the Hyperfine Shifted ^{2}H and ^{15}N Peaks from Reduced *C. pasteurianum* Rubredoxin[a]

^{2}H chemical shift assignments[b]			^{15}N chemical shift assignments[c]		
NMR data	structural model for the iron center		NMR data	structural model for the iron center	
^{2}H shift/ppm	4RXN-M	5RXN-M	15N shift/ppm	4RXN-M	5RXN-M
−0.4	C^{42} H^{α}	C^{42} H^{α}	−17.3	C^{42}	C^{9}
−0.4	C^{9} H^{α}	C^{9} H^{α}	6.5	C^{9}	C^{42}
10.9	C^{39} H^{α}	C^{6} H^{α}	217.5	G^{43}	G^{10}
18.6	C^{6} H^{α}	C^{39} H^{α}	217.5	G^{10}	G^{43}
157.0	C^{6} $H^{\beta 2}$	C^{39} $H^{\beta 2}$	262.8	V^{44}	V^{8}
158.6	C^{39} $H^{\beta 2}$	C^{6} $H^{\beta 2}$	271.4	P^{40}	P^{40}
193.4	C^{39} $H^{\beta 3}$	C^{39} $H^{\beta 3}$	274.1	T^{7}	T^{7}
196.0	C^{6} $H^{\beta 3}$	C^{6} $H^{\beta 3}$	287.4	C^{6}	C^{39}
233.0	C^{9} $H^{\beta 3}$	C^{9} $H^{\beta 3}$	311.9	C^{39}	C^{6}
233.0	C^{42} $H^{\beta 3}$	C^{42} $H^{\beta 3}$	381.6	Y^{11}	Y^{11}
243.7	C^{9} $H^{\beta 2}$	C^{9} $H^{\beta 2}$	419.6	L^{41}	L^{41}
251.3	C^{42} $H^{\beta 2}$	C^{42} $H^{\beta 2}$	477.7	V^{8}	V^{44}

[a] Assignments take into account prior information about residue type, atom type, and assignment and provide maximal agreement with chemical shifts calculations (B3LYP/6-311G** 104-atom model). [b] All correlations are in agreement with prior assignment of ^{1}H and ^{2}H peaks of cysteines to H^{α}, $H^{\beta 2}$, and $H^{\beta 3}$ positions (but not to specific cysteine residues) by selective labeling. [c] Underlined residue designations indicate assignments that were altered in order to be consistent with experiments involving selectively labeled samples.

Hyperfine-Shifted ^{13}C Resonances of Fe(II) and Fe(III) Rubredoxin. Table 4 contains the predicted chemical shifts for both oxidized and reduced cysteine carbons. Of these, only signals from the $^{13}C^{\beta}$ nuclei in reduced Fe(II) rubredoxin have

Table 4. Predicted Hyperfine Contact Shifts for Cysteine Carbons in Oxidized and Reduced *C. pasteurianum* Rubredoxin[a]

	oxidized			reduced[b]		
residue	C^{α} shift/ ppm	C^{β} shift/ ppm	C′ shift/ ppm	C^{α} shift/ ppm	C^{β} shift/ ppm[c]	C′ shift/ ppm
cysteine 6	2496	338	−41	1767	288 (327)	−23
cysteine 9	1225	1033	−60	647	−928 (−939)	−22
cysteine 39	2181	305	−37	1549	−127 (−36)	−30
cysteine 42	1386	1228	−66	709	−761 (−755)	−19

[a] Contact shifts (ppm) for oxidized rubredoxin are 267 430 times the $\rho_{\alpha\beta}$ obtained from B3LYP/6-311G**. Shifts for reduced rubredoxin used 229 230 as the multiplier. [b] Reduced Fe(II) $\rho_{\alpha\beta}$ values were calculated for the oxidized Fe(III) structure (5RXN[4,5]) by changing the oxidation state in the input to Gaussian 94. [c] Shifts in parentheses are experimentally observed values. Assignments to specific cysteine residues shown were based on a simple ranking of the shifts.

been resolved. The correlations of chemical shifts with residues shown in Table 4 will be tested by residue selective labeling. New strategies will be used in attempts to locate the other ^{13}C resonances.

Transmission of Fermi Contact Interactions across Hydrogen Bonds. The role of hydrogen bonds to iron ligated sulfurs as electron delocalization pathways is illustrated by the nitrogen of Tyr^{11}, which is hydrogen bonded to the sulfur of Cys^{9}. This nitrogen exhibits the largest calculated $\rho_{\alpha\beta}$ and the shortest hydrogen bond to sulfur. Since the nitrogen of Tyr^{11} is eight covalent bonds from the iron center, the mechanism for generating the large net α electron spin density at this atom clearly must be delocalization through the amide hydrogen bond to sulfur. Similar effects were observed in NMR studies of cadmium- and mercury-substituted *Pyrococcus furiosus* rubredoxins, in which it was found that backbone amide protons hydrogen bonded to cysteine sulfurs exhibited scalar coupling to the metal nucleus (^{113}Cd and ^{199}Hg).[58] Strong evidence suggests that the scalar coupling is electron orbital mediated.

On structural grounds, the residues that give rise to hyperfine shifted nitrogen signals can be separated into four groups. The first group consists of cysteines or residues that immediately follow cysteine whose H^N is not hydrogen bonded to S^{γ} or O of cysteine (Cys^{6}, Thr^{7}, Cys^{39}, and Pro^{40}). The second consists of residues that immediately follow cysteine in the protein sequence and whose H^N is hydrogen bonded to O of a cysteine (Gly^{10} and Gly^{43}). The third group consists of cysteines whose H^N is hydrogen bonded to S^{γ} of another cysteine (Cys^{9} and Cys^{42}). The fourth group consists of residues other than cysteine whose H^N is hydrogen bonded to S^{γ} of a cysteine (Val^{8}, Tyr^{11}, Leu^{41}, and Val^{44}). A clear symmetry exists in these groupings as shown in Figure 1.

The fourth group (that consisting of Val^{8}, Tyr^{11}, Leu^{41}, and Val^{44}) is of particular interest because their backbone nitrogens exhibit large isotope shifts to lower frequency (upfield) when the nitrogen-bound protons are replaced by deuterons.[59] For these residues a strong linear relationship is found for $\rho_{\alpha\beta}$ and the distance between the hydrogen-bonded proton and the sulfur (Figure 4). Because an analogous relationship does not exist between $\rho_{\alpha\beta}$ and the sulfur-to-nitrogen distance, the orientation of the nitrogen-to-proton bond vector with respect to the sulfur appears to have minimal effect on the amount of orbital overlap. The electron density is being delocalized from a lone pair orbital on the sulfur to the nitrogen by way of the hydrogen in the hydrogen bond. The insensitivity to bond vector orientation is postulated to be caused by rough spherical symmetry in the sulfur electron donor and electron accepting nitrogen-to-proton antibonding σ^* orbitals over the range of angles presented in the crystal structures.

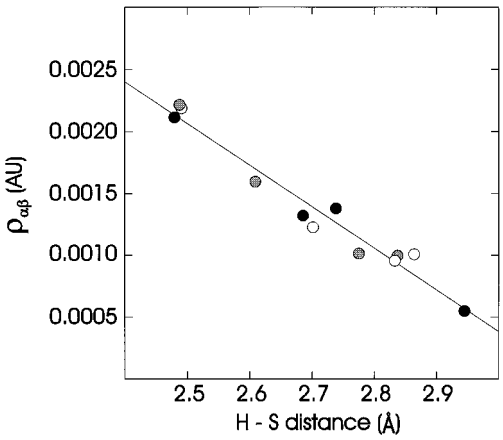

Figure 4. Relationship between the calculated values of $\rho_{\alpha\beta}$ (au) for V^{8}, Y^{11}, L^{41}, and V^{44} and the backbone amide proton to cysteine sulfur bond distances in the structural models used in the calculations (Å): 4RXN-M (black circles), 5RXN-M (open circles), and 1IRO-G_{HN} (gray circles).

Conclusions

In this paper we have shown how Fermi contact spin densities in an iron sulfur protein can be calculated from a simplified structural model derived from a protein X-ray structure by means of an efficient, hybrid density functional treatment employing a large basis set. Chemical shifts calculated directly from these results, with no empirical adjustment, gave remarkably good agreement with experiment. It is encouraging that the calculations made a strong distinction between cysteine nitrogens shifted to higher frequency (downfield) and to lower frequency (upfield). The model structures yielded negative $\rho_{\alpha\beta}$ for the nitrogens of both Cys^{9} and Cys^{42}, suggesting that the shifts to lower frequencies (upfield) are not due to pseudocontact shifts. Because the metal site contains a large excess of *positive* spin density arising from the five unpaired α electrons formally on the iron atom, the finding of *negative* spin density at these nitrogen nuclei is counterintuitive. The excellent agreement between the experimental chemical shifts from Fe(II) rubredoxin in solution and those calculated from models based on the X-ray structures indicates that the structure of the region around the metal center is very similar in solution and in the crystals.

As stated above, it is likely that the geometry of the iron-coordinated cysteines is very similar in Fe(III) and Fe(II) rubredoxin. A comparison of X-ray crystal structures of oxidized and reduced *P. furiosus* rubredoxin indicated that the most significant structural change occurring in close proximity to the iron center was an increase in iron-to-sulfur bond distances by an average of 0.04 Å.[60] This comparison also showed a decrease in the amide proton-to-sulfur H-bond distances by an average 0.09 Å upon reduction, while a molecular dynamics study predicted that many of the structural differences between oxidized and reduced forms of rubredoxin are localized in the region surrounding the Fe−S center.[61] This is in agreement

(58) Blake, P. R.; Lee, B.; Summers, M. F.; Adams, M. W. W.; Park, J. B.; Zhou, Z. H.; Bax, A. *J. Biomol. NMR* **1992**, *2*, 527.
(59) Xia, B.; Wilkens, S. J.; Westler, W. M.; Markley, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 4893−4894.
(60) Day, M. W.; Hsu, B. T.; Joshua-Tor, L.; Park, J.-B.; Zhou, Z. H.; Adams, M. W. W.; Rees, D. C. *Protein Sci.* **1992**, *1*, 1494−1507.
(61) Shenoy, V. S.; Ichiye, T. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 152−160.

with results from a recent NMR study of the diamagnetic region of rubredoxin: the largest chemical shift differences between Fe(II) Rdx and Fe(III) Rdx were found to occur at residues adjacent to the hyperfine shifted CXXCGX regions.[28,29] Individual changes for atoms in the cluster may be small. However, small changes coupled together can have a large impact on the electronic structure and thus the observed isotropic shift.

Likely sources of residual differences between calculation and experiment include incomplete assignments (to be remedied by additional experimentation), the use of oversimplified basis sets (to be approached through more extensive computations), errors in the atomic coordinates of the starting models (to be evaluated through additional refinement of X-ray structures), and the neglect of contributions to the chemical shift other than the random coil diamagnetic shifts and Fermi contact paramagnetic shifts.

Deviations between calculated spin densities and experimental NMR data in rubredoxin could result from the neglect of motions in the calculations. Atom positions determined by single-crystal X-ray diffraction analysis represent time-averaged positions. It is conceivable that the time average of spin densities over the range of vibrationally accessible atom positions would be significantly different from that calculated from a single average position. Such differences would enter into the uncertainties as a systematic error.

Inadequacies in the basis set are likely to be the largest source of error in the calculation. At the B3LYP/3-21G** level the R^2 for backbone nitrogens (0.85) is significantly worse than that (0.96) obtained at the higher 6-311G** level (Table 1). Moreover, the slope for the ^{15}N data in the 3-21G** basis level is less than half the theoretical value.

Unlike pure density functional methods, the hybrid density functional methods such as B3LYP are well-behaved over a wide range of chemical systems for calculating Fermi contact spin densities.[34] However, like pure density functional methods, hybrid density functionals do not obey the variation principle as do pure *ab initio* methods. Thus, one cannot guarantee that increasing the number of basis functions will result in an improvement of the quality of the wave function and Fermi contact spin densities. It is highly probable that improvements in the basis set would result in greater agreement with experiment, but it is difficult to judge when these improvements would become large enough to justify the additional computation time.

Acknowledgment. This work was supported by a grant (GM35976) from the National Institute of General Medical Sciences, National Institutes of Health. Calculations utilized the Silicon Graphics Power Challenge Array at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, and the Silicon Graphics Onyx at the National Magnetic Resonance Facility at Madison (NMRFAM) which is supported by a grant (RR02307) from the Biomedical Research Technology Program of the National Center for Research Resources, National Institutes of Health. S.J.W. was supported in part by a traineeship from an NIH Molecular Biophysics Training Grant (GM08293).

JA973489D