Article

A QM/MM Approach to Interpreting ^{67}Zn Solid-State NMR Data in Zinc Proteins

Andrew S. Lipton, Robert W. Heck, Greg R. Staeheli, Marat Valiev, Wibe A. De Jong, and Paul D. Ellis

J. Am. Chem. Soc., **2008**, 130 (19), 6224-6230 • DOI: 10.1021/ja711240t • Publication Date (Web): 15 April 2008

J|A|C|S
ARTICLES

Published on Web 04/15/2008

A QM/MM Approach to Interpreting ^{67}Zn Solid-State NMR Data in Zinc Proteins

Andrew S. Lipton,[†] Robert W. Heck,[†] Greg R. Staeheli,[†] Marat Valiev,[‡] Wibe A. De Jong,[‡] and Paul D. Ellis*,[†]

Biological Sciences Division and the Environmental Molecular Sciences Laboratory, Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington 99352

Received December 19, 2007; E-mail: paul.ellis@pnl.gov

Abstract: We present here a ^{67}Zn solid-state NMR investigation of Zn^{2+} substituted rubredoxin. The sample has been prepared as both a dry powder and a frozen solution to determine the effects of static disorder on the NMR line shape. Low-temperature experiments have been performed at multiple fields to determine the relative contributions to the NMR line shape from the electric field gradient and the anisotropic shielding tensors. Finally we present the theoretical interpretation of the experimental results utilizing a combined quantum mechanical molecular mechanics (QM/MM) approach. Theory predicts a sizable contribution from anisotropic shielding as compared with previously examined model systems. This is in good agreement with the experimental data.

Introduction

Historically, cysteine coordination to zinc was thought to be primarily a structural motif. Recently, it has been discovered that there is a chemical role for zinc-dependent thiolate methylation proteins, including methionine synthase,[1,2] farnesyl transferase,[3] and the DNA repair protein Ada[4] among others. In an attempt to determine the factors that make a zinc coordinated with four cysteines either structural or reactive we have begun an investigation into characterizing Zn^{2+} in these environments. Anything done to differentially modulate the chemical properties (ionic character, susceptibility to nucleophilic attack, etc.) of the zinc-bound cysteines should be reflected in the electronic properties of the zinc.

Due to its closed shell d^{10} configuration, zinc is not amenable to most forms of spectroscopy. However, ^{67}Zn is an NMR active isotope with a resonance frequency 1/16th that of ^{1}H and a nuclear spin of $^{5}/_{2}$. These qualities are typically undesirable as they result in broad quadrupole line shapes and low sensitivity. We have recently developed a method by which solid-state ^{67}Zn NMR spectroscopy could be performed on systems with intrinsically low sensitivity,[5,6] such as small molecules at natural isotopic abundance (4% for ^{67}Zn)[7–9] or in a protein (enriched to >88% ^{67}Zn).[10,11] The premise is to overcome the dilution (typically <1% by weight) of the spin in the system of interest by reducing the temperature of the system to cryogenic temperatures to enhance both the Boltzmann factor and reduce the ambient noise of the NMR probe in conjunction with performing a CP/QCPMG experiment.[5] Having acquired the data and extracted the relevant parameters from the NMR line shape through simulation, the next step is interpretation. The principal observable in a solid-state NMR experiment on a quadrupolar nuclide is the quadrupole coupling constant, C_q.[12] This coupling constant is directly proportional to the electric field gradient at the nuclide of interest and is given by

$$C_q = q_{zz}\left[\frac{e^2}{a_0^3 h}\right]Q \tag{1}$$

$$= q_{zz} \cdot 35.245 \text{ MHz (for } ^{67}\text{Zn)} \tag{2}$$

Here Q is the quadrupole moment of the nucleus in question, q_{zz} is the zz element of the field gradient tensor, and the atomic constants (e, a_0, and h) have their usual meanings. The experimentally derived quadrupole coupling constant can therefore be compared to that predicted by an *ab initio* molecular orbital calculation.

We have investigated a series of model complexes with similar coordination to zinc utilizing these methods.[13] In that

[†] Biological Sciences Division.
[‡] Environmental Molecular Sciences Laboratory.

(1) Gonzalez, J. C.; Peariso, K.; Penner-Hahn, J. E.; Matthews, R. G. *Biochemistry* **1996**, *35*, 12228–12234.
(2) Goulding, C. W.; Matthews, R. G. *Biochemistry* **1997**, *36*, 15749–15757.
(3) Huang, C.-C.; Casey, P. J.; Fierke, C. A. *J. Biol. Chem.* **1997**, *272*, 20–23.
(4) Myers, L. C.; Terranova, M. P.; Nash, H. M.; Markus, M. A.; Verdine, G. L. *Biochemistry* **1992**, *31*, 4541–4547.
(5) Lipton, A. S.; Sears, J. A.; Ellis, P. D. *J. Magn. Reson.* **2001**, *151*, 48–59.
(6) Lipton, A. S.; Heck, R. W.; Sears, J. A.; Ellis, P. D. *J. Magn. Reson.* **2004**, *168*, 66–74.
(7) Larsen, F. H.; Lipton, A. S.; Jakobsen, H. J.; Nielsen, N. C.; Ellis, P. D. *J. Am. Chem. Soc.* **1999**, *121*, 3783–3784.
(8) Lipton, A. S.; Wright, T. A.; Bowman, M. K.; Reger, D. L.; Ellis, P. D. *J. Am. Chem. Soc.* **2002**, *124*, 5850–5860.
(9) Lipton, A. S.; Bergquist, C.; Parkin, G.; Ellis, P. D. *J. Am. Chem. Soc.* **2003**, *125*, 3768–3772.
(10) Lipton, A. S.; Buchko, G. W.; Sears, J. A.; Kennedy, M. A.; Ellis, P. D. *J. Am. Chem. Soc.* **2001**, *123*, 992–993.
(11) Lipton, A. S.; Heck, R. W.; Ellis, P. D. *J. Am. Chem. Soc.* **2004**, *126*, 4735–4739.
(12) Cohen, M. H.; Reif, M. G. *Solid-State Phys.* **1957**, *5*.
(13) Lipton, A. S.; Ellis, P. D. *J. Am. Chem. Soc.* **2007**, *129*, 9192–9200.

10.1021/ja711240t CCC: $40.75

study we determined that the average Zn−S bond distance is not a factor in the electric field gradient. However, there is high sensitivity to the nature of the Zn−S bonds. In that study we found good agreement between experiment and field gradients predicted with density functional theory utilizing a local density approximation and a triple-ζ level basis set. For these and other calculations[11] an improvement of the level of agreement for both C_q and the asymmetry parameter, η_q, the model often had to be increased in complexity to include hydrogen-bonding partners and/or nearby charged entities. A typical modeling strategy for a metal center in a protein has traditionally been to only include the first ligation sphere with the metal, and even then the residues are trimmed to a representation of the side chains (i.e., methyl imidazole for histidine).[11]

These calculations can be made more realistic by also including the protein environment represented by classical point charges similar to the self-consistent charge field perturbation (SC-CFP) approach taken by Oldfield and co-workers.[14] The SC-CFP involves treating the target molecule as quantum atoms, while the next set of the repeating monomer units are reduced to point charges. Similar in spirit to SC-CFP, a combined quantum mechanics molecular mechanics (QM/MM)[15–19] approach describes the desired active-site region quantum mechanically with the rest of the protein represented by a classical force field. This allows one to perform an efficient optimization of the entire protein complex that can then be followed by accurate molecular property calculations. We demonstrate here this strategy for utilizing QM/MM calculations as a means for predicting NMR parameters as applied to Zn^{2+}-substituted rubredoxin (Rd). This system was chosen for examination as an example of a Zn^{2+} in a nonreactive or structural protein site.

Experimental Section

Preparation of *Pyrococcus furiosis* (Pf) ^{67}Zn-Rubredoxin. *Escherichia coli* strain BL21 (DE3) containing pT7-7 plasmid[20] with the wild-type Rd gene was obtained from the laboratory of M. Eidness.[21] This expression system was cultured in M9 minimal media at 37 °C until OD_{600} reached 0.8, then protein expression was induced with the addition of 0.4 mM IPTG and 10 μM ^{67}Zn[22] acetate followed by further incubation at 37 °C for 4 h and then 16 °C overnight. Cells were harvested by centrifugation at 6000*g* for 10 min. The pelleted cells were resuspended in PBS buffer and then lysed by three passages through a French Pressure cell. Cell debris was pelleted by centrifugation at 30000*g* for 30 min, and the clear supernatant was collected. This supernatant was heated until solution temperature reached 80 °C and then gently mixed and reheated to 80 °C again for 5 min. Denatured proteins were removed by centrifugation at 30,000*g* for 30 min. One unit of DNase I was added per 10 mL of clear supernatant, $MgCl_2$ was added to 1 mM, and the solution was dialysed 48 h at 4 °C in PBS plus 1 mM $MgCl_2$.

The Rd was further purified by anion-exchange chromatography on a Mono-Q (GE Healthcare) column eluted with a 0–1 M gradient of NaCl in 50 mM Tris, pH 8.5. Rd-containing fractions were dialyzed against water, concentrated by ultrafiltration in a stirred Amicon cell, and then freeze-dried to remove excess water. Freeze-dried Rd was rehydrated inside of a sealed container containing water at 25% of the Rd protein mass for 12 h and then packed into a 5 mm × 20 mm glass NMR tube. In 150 μL of 30% glycerol were mixed 50 mg of ^{67}Zn-Rd and 20 mg of Co-carbonic anhydrase;[11] the mixture was then loaded into a 5 mm × 20 mm NMR tube and frozen at −40 °C.

Solid-State ^{67}Zn NMR. All zinc chemical shifts are referenced with respect to 1 M $Zn(NO_3)_{2(aq)}$ (measured at ambient temperature). The ^{67}Zn spectra were acquired at 10 K utilizing a Varian NMR System (VNMRS) spectrometer with a wide-bore (89 mm) Oxford Instruments magnet operating at 11.7 T (500 MHz for ^{1}H and 31.297 MHz for ^{67}Zn) and a Varian Unity^{Inova} spectrometer with a medium-bore (63 mm) Oxford Instruments magnet operating at 18.8 T (800 MHz for ^{1}H and 50.048 MHz for ^{67}Zn). To obtain cryogenic temperatures (10 K) in the 11.7 and 18.8 T magnets an Oxford Instruments continuous flow cryostat was utilized. The cryostat is top-loaded into the bore of the magnet, and a home-built NMR probe is then inserted into the sample space of the cryostat.[5,6] The pulse sequences used were a combination of cross polarization (CP)[23] with signal detection using either a single Hahn echo or a quadrupole Carr−Purcell−Meiboom−Gill (QCPMG) echo train.[7,24] The spectra acquired at the lower field required a stepped offset frequency and reconstruction using a sky projection.[25] The chosen offsets were 10 kHz steps made relative to the 0.0 ppm shift of 1 M $Zn(NO_3)_2$, for a total of six experiments (transmitter offset = 30, 20, 10, 0, −10, −20 kHz).

The spectra were analyzed using the SIMPSON program.[26] Simulations of the NMR spectra were performed on a Beowulf cluster at PNNL (composed of 40-Verari Dual Socket, Dual Core Intel 5140 2.33 GHz Xeon nodes, 24-Racksaver[27] Dual Pentium IV 2.4 GHz Xeon nodes, and 6-Racksaver Dual Pentium III 1.26 GHz nodes) running the Rocks clustering software and utilizing a gigabit Ethernet connection. The errors in the extracted parameters were estimated from a comparison of the resulting fits at each individual field with the simultaneous fit at both magnetic fields. A simultaneous fit consists of simulating line shapes at both field strengths with identical parameters, summing the individual rmsd's for a combined residual for use in the simplex algorithm, thereby optimizing parameters considering the consequences of all the data at once rather than independently. For experiments that were acquired with a single transmitter frequency the simulations were performed with finite pulse widths. Those experiments that required combining stepped frequencies[8] were simulated with ideal pulses followed by an exact calculation utilizing combined simulations with stepped offsets and finite pulses. We define the shielding anisotropy as $\Delta\sigma = \delta_{33} - \frac{1}{2}(\delta_{11} + \delta_{22})$, where the elements of the shielding tensor are ordered as $|\delta_{33} - \delta_{iso}| \geq |\delta_{11} - \delta_{iso}| \geq |\delta_{22} - \delta_{iso}|$ and δ_{iso} is the trace of the tensor δ.

Ab Initio Molecular Orbital Calculations. Geometry optimizations and calculations of the electric field gradient (EFG) tensors were performed using the QM/MM module[19] of the NWChem computational chemistry package developed at Pacific Northwest National Laboratory (PNNL).[28,29] An X-ray geometry of reduced

(14) Zhang, Y.; Mukherjee, S.; Oldfield, E. *J. Am. Chem. Soc.* **2005**, *127*, 2370–2371.
(15) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.
(16) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.
(17) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley: New York, 1997.
(18) Gao, J. In *Reviews In Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, NY, 1995; Vol. 7, pp119–185.
(19) Valiev, M.; Yang, J.; Adams, J. A.; Taylor, S. S.; Weare, J. H. *J. Phys. Chem. B* **2007**, *111*, 13455–13464.
(20) Tabor, S. In *Current Protocols in Molecular Biology*; Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Eds.; Greene Publishing and Wiley-Interscience: New York, NY, 1990, pp 16.2.1–1.2.11.
(21) Eidness, M. K.; Richie, K. A.; Burden, A. E.; Kurtz, D. M.; Scott, R. A. *Biochemistry* **1997**, *36*, 10406–10413.
(22) Zn-67 metal purchased from Cambridge Isotope Laboratories, Inc. of Andover, MA at >88% enrichment.
(23) Pines, A.; Gibby, M.; G., Waugh; J. S, *J. Chem. Phys.* **1972**, *56*, 1776.
(24) Larsen, F. H.; Jakobsen, H. J.; Ellis, P. D.; Nielsen, N. C. *J. Phys. Chem. A* **1997**, *101*, 8597–8606.
(25) Lipton, A. S.; Wright, T. A.; Bowman, M. K.; Reger, D. L.; Ellis, P. D. *J. Am. Chem. Soc.* **2002**, *124*, 5850–5860.
(26) Bak, M.; Rasmussen, J. T.; Nielsen, N. C. *J. Magn. Reson.* **2000**, *147*, 296–330.
(27) Racksaver is now Verari Systems.

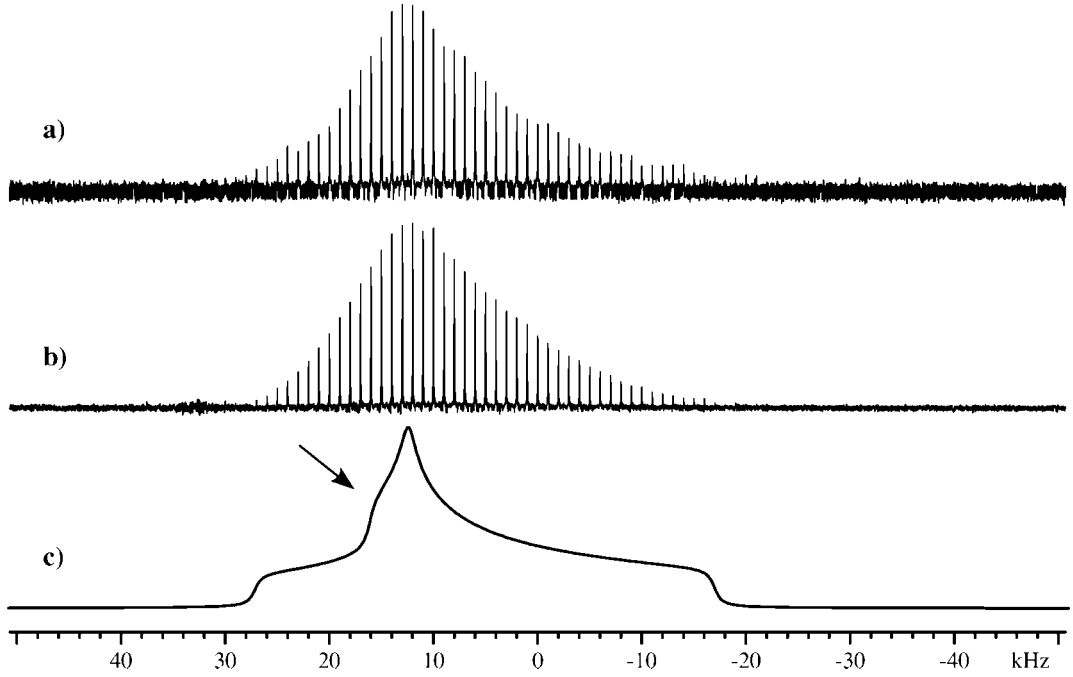

Figure 1. Experimental sky projections and simulated ^{67}Zn NMR data of *Pf*-Rd at 11.7 T where (a) the sample was prepared as a lyophilized powder and (b) the sample was in a frozen 30% glycerol/water solution, and (c) is an ideal powder simulation with $C_q = 4.8$ MHz and $\eta_q = 1$.

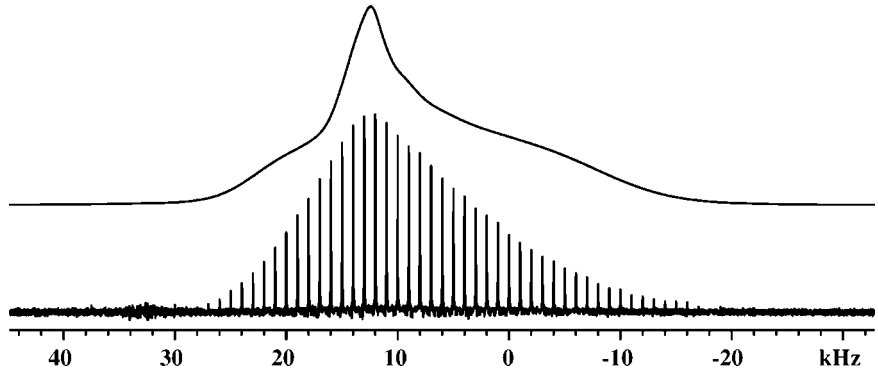

Figure 2. Comparison of the ^{67}Zn NMR spectrum of *Pf*-Rd in a frozen solution at 11.7 T which was acquired at 10 K with CP and QCPMG detection for 1024 transients, a π pulse of 15.0 μs, a ^{1}H $\pi/2$ pulse of 6.0 μs, and a 30 s recycle delay with a simulation using a Gaussian distribution of C_q's centered around 4.2 MHz.

Pf-Rd (1BRF) from the Protein Data Bank (PDB) served as a starting point for geometry optimizations. For the quantum atoms density functional theory (DFT) was used with the exchange-correlation functional defined as the local spin density approximation (LSDA or LDA) utilizing Slater's local spin density exchange,[30] the Vosko, Wilk, and Nusair (VWN) V local spin density correlation functional,[31] with an Ahlrichs double-ζ basis set with polarization (pAVDZ).[32] The LDA functional was chosen for the optimization as it has been shown to give a more reliable geometry for systems with zinc and sulfur interactions.[33] The optimization procedure iteratively cycles through the quantum region, the classical solute, and finally the solvent region until the total energy converges to within 1.0e^{-4} Hartrees. Subsequent property calculations utilized an Ahlrich's triple-ζ basis set including polarization functions (pATZV).[34]The calculations were carried out on the following systems: the Beowulf cluster mentioned above (using only the Xeon nodes) or the 11.8 TFlop Hewlett-Packard system (980 dual Intel 1.5 GHz Itanium-2 processors) present in the Environmental Molecular Sciences Laboratory running a version of Linux based on Red Hat Linux Advanced Server with a QSNet^{II}/Elan-4 interconnect from Quadrics.

Results and Discussion

Figure 1a illustrates the first attempt at a solid-state NMR experiment on rubredoxin with enriched ^{67}Zn. This line shape appears disordered when compared with a simulation of an ideal quadrupole line shape with a C_q of 4.8 MHz and an η_q of 1 (Figure 1c). This sample was prepared using a method typical

(28) Bylaska, E. J. et al.; *NWChem*, v5.0, ed; Pacific Northwest National Laboratory, Richland, WA, 2006.
(29) Kendall, R. A.; Aprà, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis, M.; Fann, G. I.; Harrison, R. J.; Ju, J.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. *Comput. Phys. Commun.* **2000**, *128*, 260–283.
(30) Slater, J. C. *The Self-Consistent Field for Molecules and Solids*; McGraw-Hill: New York, NY, 1974; Vol. 4.
(31) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
(32) Schafer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

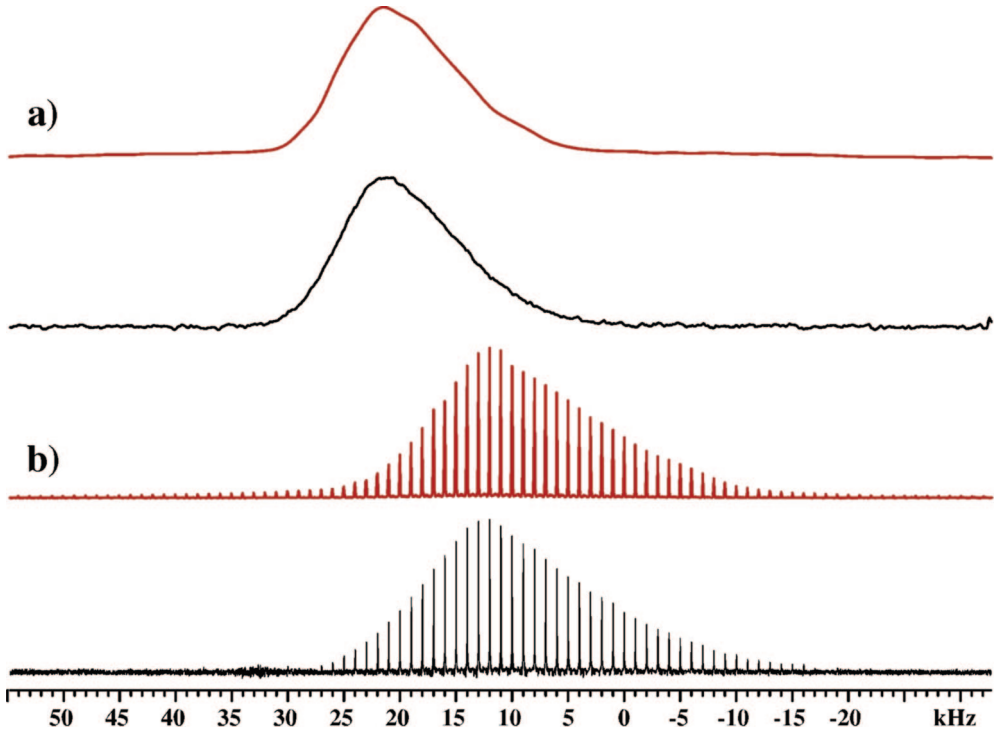

Figure 3. Experimental and simulated ^{67}Zn NMR data at (a) 18.8 T which was acquired at 10 K with CP and a single echo for 512 transients, a π pulse of 15.0 μs, a ^{1}H $\pi/2$ pulse of 6.5 μs, and a 2 min recycle delay and (b) 11.7 T which was acquired as specified in Figure 2. Each simulation (in red) utilized finite pulses and the respective parameters listed in Table 1 for each field strength.

Table 1. Extracted ^{67}Zn Line Shape Parameters for *Pf*-Rd

magnetic field (T)	C_q (MHz)	η_q	δ_{iso} (ppm)[a]	$\Delta\sigma$ (ppm)	η_σ	α	β	γ
11.7	3.92	0.84	429	246	0.77	139	45	201
18.8	4.50	0.89	454	199	0.64	77	61	197
11.7/18.8[b]	4.35	0.84	429/444	221	0.62	77	33	189
errors	0.5	0.05	10	25	0.1	32	15	10

[a] Chemical shifts referenced to 1 M $Zn(NO_3)_{2(aq)}$. [b] Fit against data from both fields simultaneously.

for protein preparation in the solid state, lyophilization followed by rehydration by vapor diffusion. Note the apparent absence of the "wings" of the spectrum as well as the shoulder to higher frequency (indicated in Figure 1). We know from EXAFS that lyophilization induces a disordering of Zn−S bonds and preparing the sample as a frozen glycerol solution can restore that structural order.[35] It is presumed that a disordering of the bonding around the metal center induces a disorder of the magnetic resonance parameters and that removing structural disorder results in tightening up the NMR parameters.

Preparing the sample as a frozen solution with 30% glycerol as a cryoprotectant did indeed improve the disorder of the sample as revealed in the proton NMR spectrum (data not shown) which shows a dipolar splitting instead of the featureless Gaussian line of the lyophilized material. However, the ^{67}Zn NMR spectrum does not show any more features than were observed previously (Figure 1b). This, however, does not detract from this preparation's ability to minimize structural disorder induced by lyophilization as well as the ability a frozen solution offers to freeze a snapshot of a solution equilibrium (such as a pK_a or reactive intermediate). This is where operating at cryogenic temperatures proves useful for other than sensitivity reasons, i.e. freezing dynamical processes. It should also be noted that the X-ray crystallography done on several rubredoxins was also performed at reduced temperatures (103–123 K).

Figure 2 compares the frozen solution preparation with an estimate of what a spectrum with a disordered C_q would look like. The simulated spectrum assumes a disorder in C_q as a Gaussian distribution with a σ value of 0.1 MHz and centered around a value of 4.2 MHz. What is apparent from the figure is that the disorder does remove the sharp features at the edges of the line shape but does not broaden the central portion of the line shape enough to match experiment. To match the width of the center of the line shape would require an estimated C_q that would be too large for the overall width of the powder pattern.

Another explanation for the appearance of the line shape is the presence of anisotropic shielding (CSA, $\Delta\sigma$). This is not unheard of for zinc with four coordinating sulfurs, as we have recently reported for several model complexes.[13] The CSA's previously reported ranged from 22 to 76 ppm for the complexes where shielding was reported. We have also previously shown a zinc spectrum for the DNA repair protein XPA, which also has the metal coordinated by four cysteines, where there was also evidence for shielding anisotropy in the line shape.[10] The unusual aspect about the CSA for rubredoxin is the magnitude.

(33) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2001**, *105*, 10709–10714.
(34) Schafer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.
(35) Buchko, G. W.; Hess, N. J.; Kennedy, M. A. *Carcinogenesis* **2000**, *21*, 1051–1057.

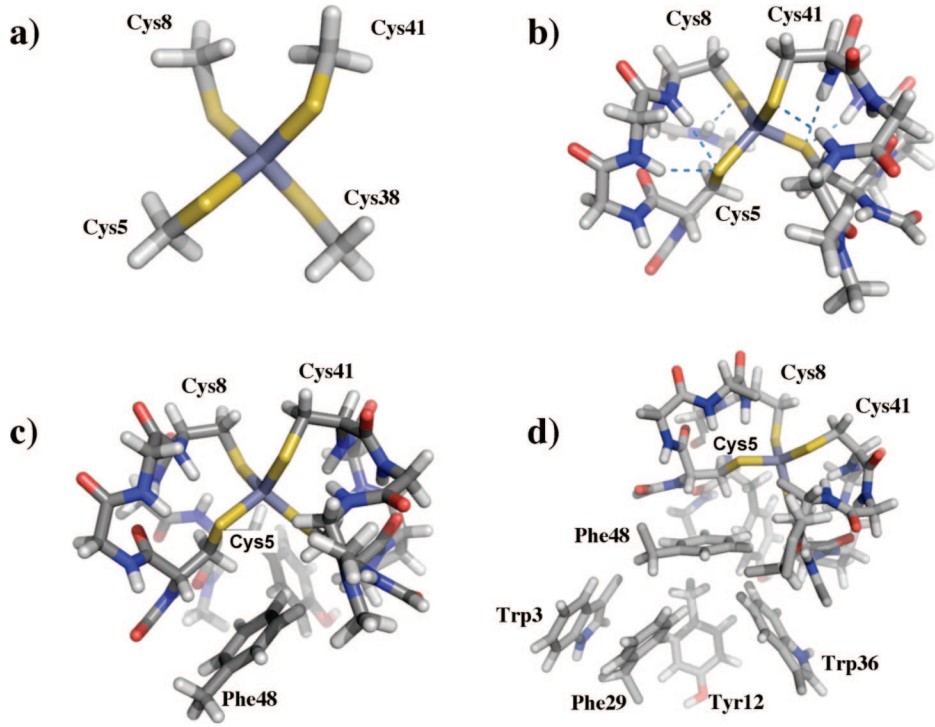

Figure 4. Optimized geometries of the quantum regions of each model (models 1–4 correspond to structures a−d, respectively) of the metal binding site of *Pf*-Rd drawn with MacPyMol 0.99.

Table 2. N−H···S Hydrogen-Bonding Distances near the Fe(II) Core of *Pf*-Rd

donor	acceptor	distance (Å)
Lys6	Cys5	3.71
Ile7	Cys5	3.60
Cys8	Cys5	3.69
Gly9	Cys8	3.59
Tyr10	Cys8	3.51
Pro39	Cys38	3.56
Ile40	Cys38	3.53
Cys41	Cys38	3.62
Gly42	Cys41	3.62
Ala43	Cys41	3.52

To fit the line shape adequately at all magnetic fields requires a $\Delta\sigma$ of nearly 250 ppm with a C_q of 4.5 MHz. The experimental line shapes acquired at 11.75 and 18.8 T with their respective simulations are depicted in Figure 3. One can see from the figure that these are not unreasonable values and that with an extra line broadening mechanism (such as structural disorder) these subtle features can blur into the seemingly featureless spectra that are observed. The extracted quadrupolar and shielding tensor information is listed in Table 1 with the relative orientation between the two principal axis systems (PAS).

Knowing the NMR parameters for this center is not enough; the question is what gives rise to them. If they can be predicted by molecular theory, then we will have gained an understanding about the electronic environment around the metal. This will aid in determining what makes a structural site different from a catalytic one when the primary coordination spheres are similar. The simplest model is a zinc with four coordinated -SH's; however, that is too simple for the present discussion. One could extend the model to four -SMe's or even complete cysteines; however, that is still not adequate to differentiate between sites. In the absence of anything to distinguish the ligands or constrain the geometry, theory will predict equidistant, identical Zn−S bonds. These simple models ignore potential hydrogen-bonding partners as well as any electronic/electrostatic effects from the remainder of the protein. Nearby charges can also be important as we have shown in the case of Zn-$[SPh]_4[NMe_4]_2$ where the field gradient was not accurately predicted until counterions from the next nearest neighbors were taken into consideration.[13] These same considerations were observed by Zhang et al. when they calculated electric field gradients for $Zn(OAc)_2 \cdot 2H_2O$ (among others) using the SC-CFP approach.[14] A QM/MM approach is similar, except that the charges are predefined by the force fields chosen, i.e. Amber[36] or CHARMM.[37]

As our starting geometry for the calculations we have chosen an X-ray structure of Pf-rubredoxin (1BRF)[38] available in the protein databank. The iron−sulfur core has the same geometrical patterns as a Zn Rd from another organism (*Clostridium pasteurianum*) where the X-ray structure was also solved at high precision.[39] Indeed the average M−S distance is almost identical between the reduced iron form (2.33Å) and the zinc-substituted (2.34Å) *Pf*-Rd as determined by EXAFS.[40] After changing the iron to zinc, the first step is to determine what to include in the quantum region. As a first pass the -SCH_2 side chains of the directly coordinated cysteines; 5, 8, 38, and 41, are included as quantum with linker hydrogens added to fill the carbon valence. The converged geometry (shown in Figure 4a) is then used in an electric field gradient calculation with a larger Ahlrich's

(36) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27–85.
(37) Mackerell, A. D., Jr *J. Comput. Chem.* **2004**, *25*, 1584–1604.
(38) Bau, R.; Rees, D. C.; Kurtz, D. M.; Scott, R. A.; Huang, H. S.; Adams, M. W. W.; Eidsness, M. K. *J. Biol. Inorg. Chem.* **1998**, *3*, 484–493.
(39) Dauter, Z.; Wilson, K. S.; Sieker, L. C.; Moulis, J. M.; Meyer, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8836–8840.
(40) George, G. N.; Pickering, I. J.; Prince, R. C.; Zhou, Z. H.; Adams, M. W. W. *J. Biol. Inorg. Chem.* **1996**, *1*, 226–230.

Table 3. Optimized Zn−S Distances and Predicted NMR Parameters

model	Zn−S_{Cys5} (Å)	Zn−S_{Cys8} (Å)	Zn−S_{Cys38} (Å)	Zn−S_{Cys41} (Å)	C_q (MHz)	η_q	$\Delta\sigma$ (ppm)	η^σ	α (deg)	β (deg)	γ (deg)
exp					4.50	0.89	198.6	0.64	77.0	60.7	196.7
1	2.32	2.30	2.32	2.29	−4.09	0.76	−257.8	0.95	88.7	81.1	104.7
2	2.32	2.26	2.34	2.28	−7.89	0.81	−101.1	0.55	81.6	77.3	130.7
3	2.34	2.28	2.33	2.28	5.40	0.59	−194.8	0.92	330.1	91.7	94.1
4	2.31	2.28	2.34	2.28	−6.32	0.80	857.0	0.53	15.4	17.6	252.9

Table 4. Basis Set Dependence of Predicted NMR Parameters of *Pf*-Rd Using Model 3

basis set	C_q (MHz)	η_q	total QMMM energy (Hartrees)	no. of basis functions	Zn functions
ATZV + 1p	5.478	0.686	−7255.13757	2042	6s4p3d
AVTZ + 1p	5.362	0.865	−7255.17945	2149	8s6p3d
AVTZ + 2p2d	6.668	0.494	−7255.15680	2162	8s7p5d
AVTZ + 2p2d2f	6.085	0.635	−7255.19596	2176	8s7p5d2f
cc-pVTZ	6.648	0.572	−7255.43534	4055	7s6p4d2f1g
aug-cc-pVTZ[a]	6.703	0.549	−7255.32265	2434	8s7p5d3f2g

[a] Basis set for HCNO atoms is pAVTZ with aug-cc-pVTZ-NR on Zn and S only.

triple-ζ basis set with polarization (pATZV)[34] to give a value close to experiment at −4.09 MHz with a shielding anisotropy of −257 ppm.

This preliminary model, however, does not help in determining if hydrogen-bonding interactions or other nearby residues affect the Zn^{2+} C_q. From the X-ray structure of Fe−Rd we know that there are several hydrogen-bonding partners from neighboring backbone N−H's to each residue. These interactions are tabulated in Table 2 for the reduced iron form of *Pf*-Rd. For the second model we have included these atoms as well as the C_α and H_α from these residues and the carbonyl from the preceding residue to better approximate the aspects of the N−H's when they are treated as quantum. A depiction of the optimized quantum region is shown in Figure 4b, and the N−H···S hydrogen bonds are shown by a dashed line.

The third model is identical to model 2, but includes in the quantum region the aromatic rings of Tyr10 and Phe48 as well. These two residue side chains are added as the plane of the rings are within 4 Å of the sulfurs of Cys38 and Cys5, respectively. These two cysteines happen to be the "internal" ones as well as those with a higher degree of hydrogen bonding (3 NH···S instead of the 2 that the "surface" cysteines have). The resulting optimized Zn−S bonds have been summarized in Table 3 with the respective predicted NMR parameters for each of the models. For comparison the experimental values of C_q and $\Delta\sigma$ are also tabulated with their respective asymmetry parameters. The optimized quantum region is shown in Figure 4c, where Cys5 and Phe48 are in the foreground, while Cys38 and Tyr10 are behind the structure in this orientation. From the orientations it is easy to envision a lone pair of the sulfur atom interacting with the π orbitals of the aromatic ring. This type of stabilizing interaction has been observed before between an oxygen lone pair and an aromatic ring system.[41] One can see from the data in Table 3 that the inclusion of these two aromatic rings brings the predicted value of C_q back in the right direction; however, the value of the asymmetry parameter is drifting away from experiment. This implies that, although the model is approaching a satisfactory level, it is not quite completely describing the electrostatics around the metal center.

The final model we investigated includes four more aromatic residues. These include Tyr12, Phe29, and Trp36 that reside below Phe48 (opposite from Cys5) in orientations almost orthogonal to the plane of Phe48's ring as well as Trp3, which stacks with Phe29. This is shown in Figure 4d where one can see this aromatic/hydrophobic core forming below the Zn−S center. It is interesting to note that all these aromatic residues (except Trp3) are conserved among all rubredoxins[42] and are most likely implicated in its function. The inclusion of these residues results in an increase in C_q; however, the asymmetry of the predicted field gradient is back in the range of the experimentally observed value.

The final set of calculations performed was a test of the basis set dependence on the EFG. We compare in Table 4 the predicted values for model 3 (Figure 4c) utilizing several triple-ζ basis sets including the Schafer, Horn, and Ahlrichs valence triple-ζ (AVTZ)[32] with varying amounts of polarization (additional functions optimized with g98opt) and the correlation consistent basis sets of Peterson and co-workers (cc-pVTZ-NR and aug-cc-pVTZ-NR).[43,44] From Table 4 it is apparent that the predicted value of C_q is not going to reach the experimental value by increasing the complexity of the basis functions. In fact it is the two minimal basis sets that come closest to the target with the AVTZ basis set reaching good agreement with η_q as well. However the predicted C_q's do seem to be converging, so perhaps this is an indication of the model being incomplete as we have already seen a significant effect by adding the additional aromatic residues (model 4).

Conclusions

We have shown in this work that for Zn^{2+}-substituted *Pf*-rubredoxin, where the metal is coordinated by four cysteines, the resulting ^{67}Zn NMR line shape is consistent with contributions from both quadrupolar and anisotropic shielding tensors. The apparent disorder in the NMR line shape does not result from the sample preparation as frozen glycerol solutions give a result similar to lyophilization. Most of the deviation from an ideal quadrupole line shape can be rationalized with a sizable shielding anisotropy (compared with previously measured model compounds). This increase in the shielding contribution is predicted by molecular theory. The computational approach used for this protein system was a combination of quantum mechanics and molecular mechanics or QM/MM. The predicted EFG is

(41) Egli, M.; Sarkhel, S. *Acc. Chem. Res.* **2007**, *40*, 197–205.

(42) 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*, 10885–10895.

(43) Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283–296.

(44) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.

sensitive to the choice of the quantum region as the sulfur of the cysteines is polarizable due to hydrogen-bonding interactions as well as from aromatic side chains. It is not sufficient to simply include these residues as force fields as these types of interactions can be missed. The gauge of how complete the model is depends not simply on the agreement of a single parameter such as the EFG, but the asymmetry must also be included as well as any anisotropic shielding interactions. Finally, the level of basis set needed to converge on these experimental values of EFG and CSA does not seem to get too large, as the Ahlrichs triple-ζ basis sets utilized for the model complexes[13] seem satisfactory for the protein systems as well.

Acknowledgment. We gratefully acknowledge the generous gift of the rubredoxin clone from Prof. M. Eidness of the University of Georgia. This work was supported by grants from the National Institutes of Health (Federal Grants EB-002050 and EB-003893 for the purchase of computer nodes). G.R.S. gratefully acknowledges support from the SULI program funded through the Department of Energy's Office of Science. This research was carried out in the Environmental Molecular Sciences Laboratory (a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research) located at Pacific Northwest National Laboratory and operated for DOE by Battelle. Computations were performed in part using the Molecular Science Computing Facility in EMSL. NWChem Version 5.0, as developed and distributed by Pacific Northwest National Laboratory, P.O. Box 999, Richland, Washington 99352 U.S.A., and funded by the U.S. Department of Energy, was used to obtain some of these results.

Supporting Information Available: Complete ref 28. This material is available free of charge via the Internet at http://pubs.acs.org.

JA711240T