Oxygen Accessibility to Ribonuclease A: Quantitative Interpretation of Nuclear Spin Relaxation Induced by a Freely Diffusing Paramagnet†

Ching-Ling Teng, Brian Hinderliter, and Robert G. Bryant*

The Biophysics Program and the Chemistry Department, University of Virginia, Charlottesville, Virginia 22904-4319

Received: May 20, 2005; In Final Form: July 30, 2005

The nuclear spin relaxation induced by a freely diffusing paramagnetic center provides a direct measure of intermolecular accessibility. A number of factors are involved in a quantitative interpretation of relaxation data including excluded volume effects, solvation differences, and the details of the electron spin relaxation in the paramagnetic center. In the case where the electron relaxation time is short compared with correlation times describing the electron−nuclear coupling, the nuclear spin relaxation rates may be related to the effective local concentration of the paramagnetic center at different locations about the solute of interest. The local concentrations may in turn be related to differences in the local free energies of interaction between the diffusing paramagnet and the cosolute. We demonstrate this approach for the case of ribonuclease A and deduce surface free energy differences for a large number of protein proton sites. We find that the oxygen accessibility is poorly represented by hard-sphere models such as computed solvent or steric accessibility. There is a distribution of local intermolecular interactions with a width of the order of *RT* that dominates the report of the intermolecular exploration of the protein by this simple solute.

Intermolecular accessibility is key to molecular recognition and associative phenomena that are critical for reactivity, catalysis, or regulation. In the case of a macromolecule such as a protein, there may be a very large number of contributions that add together to provide the final free energy change that is measured for the binding of a small solute to the protein. It is generally difficult to get a detailed understanding of all of the small energetic contributions to the free energy that characterizes intermolecular couplings; however, a reasonably detailed view of some of the contributions may be gained by exploiting intermolecular effects in spectroscopy where the short range interactions of one molecule affect the spectral properties of another. Nuclear magnetic resonance spectroscopy of soluble proteins provides excellent spectroscopic and structural resolution for the study of intermolecular interactions using relaxation spectroscopy.[1−14] The electron−nuclear dipole−dipole contribution to nuclear spin relaxation depends inversely on the sixth power of the intermoment distance. When the dipolar coupling is modulated by translational diffusion, the effective distance dependence decreases because the total intermoment coupling requires an integration of effects over a volume accessible to the interacting spin pairs; however, the dependence is still a sufficiently strong function of distance that the changes in the nuclear spin relaxation rate constants for solute protons report the effective accessibility of the freely diffusing paramagnetic cosolute to these molecular positions.[15−18] Although there have been a number of qualitative discussions of the application of this effect to measure molecular accessibility, a quantitative interpretative platform for this potentially very powerful technique has not been presented. In this paper, we present a foundation that provides the basis for quantitative interpretation of nuclear spin relaxation rate constant contributions from freely diffusing paramagnetic centers in terms of differences in the local intermolecular potential or free energy. We demonstrate the approach using ribonuclease A and deduce free energy differences for the exploration of this protein by the cosolute oxygen. We find that the oxygen accessibility is poorly represented by hard-sphere models such as computed solvent or steric accessibility and that there is a distribution of local intermolecular free energies with a width of the order of *RT*. The oxygen affinity is generally higher in hydrophobic regions, but there is not a simple correlation with elementary measures of individual amino acid hydrophobicity.

Background and Theory

In the absence of long-lived specific binding interactions, the electron−nuclear dipole−dipole coupling between a freely diffusing electron spin and a nuclear spin is modulated by the relative translational diffusion of the spin-bearing molecules. In the present experiments, oxygen is the diffusing paramagnet. If the paramagnetic oxygen were to bind to the protein at some position with full occupancy, the correlation time for the coupling would be the oxygen spin−lattice relaxation time, and at van der Waals contact with a proton, the relaxation rate constant for the proton would be approximately 6200 s^{-1}.[19] The relaxation rate constants that we measure are generally less than 10 s^{-1}. No protein proton spin−lattice rate constant or line broadening approaches 6200 s^{-1}. Thus, what oxygen binding interactions there are at the protein interface involve very low occupancies. If one adopted the view that the oxygen binds significantly, then the binding probability would be generally of the order of 0.1% or less. We therefore adopt the analytical strategy that the oxygen explores the protein−water interface by diffusive motions and employ models for the paramagnetic contributions to the proton spin−lattice relaxation rate constants based on the relative diffusive motion of the interacting spins.

† Part of the special issue "Donald G. Truhlar Festschrift".

* Corresponding author. Address: Chemistry Department, University of Virginia, P.O. Box 400319, Charlottesville, VA 22904-4319. E-mail: rgb4g@virginia.edu. Phone: 434-924-1494.

10.1021/jp0526593 CCC: $33.50
Published on Web 09/10/2005

The intermolecular relaxation models assume that the point dipole approximation for the intermoment coupling is valid, and we know of no evidence to question this assumption, particularly for intermolecular couplings where covalent connectivity is absent.[20]

The analytical models for the spin−lattice relaxation rate constants when the length as well as the orientation of the intermoment vector may change are generally very different from those for intramolecular dipolar couplings modulated by rotational diffusion. There are two major differences: the spectral density functions that enter the relaxation equation are not Lorentzian so that the magnetic relaxation dispersion profile is not Lorentzian but may be considerably broader.[21] The dependence of the paramagnetic contribution to the relaxation rate constant on distance appears to be weaker because one integrates over the range of intermoment distances, which may include the extremes of the sample volume. Analytical expressions for the translational contribution to the paramagnetic relaxation rate constant have been developed by Ayant, Hwang, Freed, and others[16−18,22−24] on the basis of the relative diffusion of the interacting particles. For the applications of interest here, there are two limits for the electron spin relaxation time, T_{1e}, of the freely diffusing paramagnetic center: (1) T_{1e} is long compared with the translational correlation time or the rotational correlation time of the target molecule of interest. (2) T_{1e} is short compared with translational diffusion correlation times. The first case is relevant to nitroxide centers and some metal centers that have been used most often in the past;[2,4,6,7,9−11,21,25−28] the second is important to the present discussion where we use dioxygen as the paramagnetic relaxation agent.[2,3,29,30] The paramagnetic contribution to the kth proton in solution originating from a dipolar coupling to the oxygen center is given by Freed[16]

$$\frac{1}{T_{1k}} = \frac{32\pi}{405}\gamma_I^2\gamma_S^2\hbar^2 S(S+1)P\frac{N_A}{1000}\frac{[C]}{bD}\{j_2(\omega_S-\omega_I) + 3j_1(\omega_I) + 6j_2(\omega_S+\omega_I)\} \quad (1)$$

where

$$j_j(\omega) = j_j^a(\omega) + j_j^b(\omega) \quad (2)$$

$$j_j^a(\omega) = Re\left[1/(1+s_j\tau_j)\frac{1+\frac{\sigma_j}{4}}{1+\sigma_j+\frac{4\sigma_j^2}{9}+\frac{\sigma_j^3}{9}}\right] \quad (3a)$$

$$j_j^b(\omega) = \tau/\tau_j\, Re\left[1/(1+s_j\tau_j)\frac{^9/_4 - (5\sigma_j^2/4) + (3\sigma_j^3/4) - (\sigma_j^4/4)}{1-(\sigma_j^2/9)-(2\sigma_j^4/81)-(\sigma_j^6/81)}\right] \quad (3b)$$

$$s_j = \left(i\omega + \frac{1}{T_{jS}}\right), \quad j = 1, 2 \quad (3c)$$

$$\sigma_j = s_j\tau/(1+s_j\tau_j) \quad (3d)$$

$$\tau = b^2/D \quad (3e)$$

$$D = D_{O_2} + D_{protein} \quad (3f)$$

$$\tau_j = \frac{\langle r^2\rangle}{6D_{O_2}} \quad (3g)$$

$[C]$ is the concentration of the paramagnetic molecule, while S is the electron spin, 1 in this case, ω is the Larmor frequency for the nuclear spin, I, or the electron spin, S, P is a factor discussed below that equals 1 for a hydrogen atom, b is the distance of closest approach between the electron and nuclear spin, D is the relative diffusion constant, and $\langle r^2\rangle$ is the mean-square jump length of the small solute. In the present case, the motion of the protein is slow so that the effective diffusion constant becomes that for the paramagnet, oxygen. T_{jS} is the electron spin−lattice or transverse relaxation time for $j = 1$ or 2, respectively, and the other symbols have their usual meaning.[16]

The electron relaxation rate constant for dioxygen is 7.5 ps and is not sensitive to viscosity.[31] The translational correlation time for water computed from eq 3g is also 7.5 ps; however, estimates of the translational mobility at the protein interface are a factor of 3 or more slower.[21] Thus, the fluctuations in the electron−nuclear coupling at the macromolecular interface will be dominated by the rapid electron spin relaxation. Because the correlation time for the electron−nuclear coupling in oxygen solutions is short and minimally sensitive to oxygen rotational or translational dynamics, the correlation time for the electron−nuclear coupling is practically constant and the paramagnetic contribution to the spin relaxation rate constant at each proton is proportional to $P[C]$.

The relaxation of small molecules is somewhat different from that of large molecules. In the case of small molecules, the diffusion jump length is often approximated as the distance of closest approach, which is close to the molecular diameter of the small molecule. However, in the case of macromolecules such as proteins, a proton may be sterically excluded from contact with the small molecule or buried by the protein fold. In this case, the diffusion jump length and the distance of closest approach may be different. Thus, eq 1 suggests that the relaxation rate constant is proportional to $1/b$; however, for small molecules, the mean jump length is approximately the same as the distance of closest approach and $r^2 \cong b^2$.[32] Substitution of $D = b^2/6\tau_j$ into eq 1 yields the steeper distance dependence of b^{-3}, which is appropriate for small solutes or for surface protein protons. However, for an arbitrary protein proton, the distance of closest approach of the diffusing paramagnetic center may be very different from one proton to another, although the relative protein−oxygen diffusion constant is the same. Therefore, $\langle r^2\rangle \neq b^2$ and the effective dependence of the nuclear relaxation rate constant on distance is more complex.[16] We point this out for generality: this manuscript focuses on surface protons, where the distance of closest approach, b, is practically the same for all surface protons and this complication is unimportant.

Figure 1 was computed on the basis of eqs 1−3 with the assumption of spherical symmetry of a hydrogen atom, the correlation time set at 7.5 ps, and the translational diffusion constant equal to that of the solvent water, 2×10^{-5} cm^2 s^{-1}. The calculation is truncated at distances below the sum of the van der Waals radii of the hydrogen and oxygen atoms. The dashed line indicates the relaxation rate that would be observed if the distance of closest approach to the hydrogen atom is

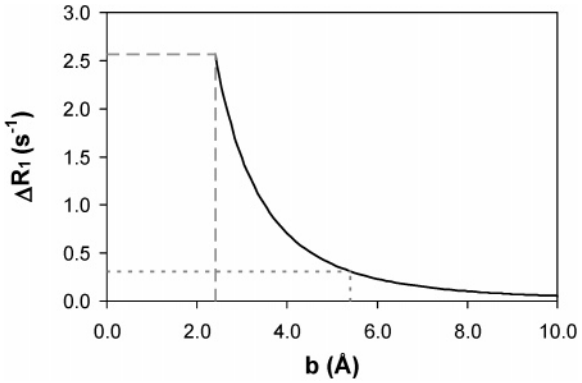

Figure 1. Oxygen induced relaxation rate constant contribution as a function of distance from the proton assuming an oxygen concentration of 13.3 mM, a translational diffusion constant of 2×10^{-5} cm^2 s^{-1}, and a correlation time of 7.5 ps for the electron−nuclear coupling. The plateau is established at van der Waals contact, the dashed line represents the relaxation rate at a distance of 3.0 Å further away, that is, the distance an intervening water molecule would add.

increased by the presence of an intervening water molecule; the relaxation rate constant falls by nearly a factor of 10 if water is in the way so that *b* increases by 3 Å. The strong dependence of the proton relaxation rate on distance is supported by early calculations which indicate that, for purely diffusional contributions, approximately 90% of the relaxation occurs within 10 Å of the detected proton.[21] Because the correlation time is the T_{1e} value for oxygen, the relaxation rate constant is not significantly sensitive to changes in the translational mobility of the oxygen in the vicinity of the protein or other solute molecule. Therefore, the paramagnetic contribution to the proton relaxation rate is an effective measure of the time-average proximity of oxygen to the measured proton or the local concentration, which is generally found to be larger in the vicinity of the protein protons than implied by the equilibrium concentration of oxygen in water. The volume associated with defining the effective local oxygen concentration is implied by Figure 1. For the surface proton, the majority of the paramagnetic relaxation contribution derives from the first several angstroms and more than 90% of the paramagnetic relaxation contribution derives from the oxygen within 10 Å of the detected proton as noted. Thus, the effective volume associated with the local concentration is defined by a radial distance of approximately 2 or 3 water molecule diameters. For the spectral effects observed to be in the fast chemical exchange limit, the time scale for the averaging process must be short compared to the proton relaxation time when the oxygen is adjacent to the proton, which is approximately 0.2 ms. All present evidence indicates that the mean lifetimes for the oxygen interactions are short compared with this value.

The parameter *P* includes factors that may change the accessibility to a particular nuclear spin including steric factors and any biases in the effective local concentration created by the local intermolecular energetic profile. This factor may be divided into two parts: a steric factor and an equilibrium constant that accounts for locally different concentrations of the diffusing paramagnet created by intermolecular interactions. Thus, we write for the *k*th proton site

$$P_k = f_k K_k \qquad K_k = [O_{2,\mathrm{local}}]/[O_{2,\mathrm{bulk}}] \tag{4}$$

where f_k is a geometric factor that accounts for the local bonding pattern in the molecule that limits the uniform distance of closest approach to the detected nuclear spin.

f_k will be computed in detail, but it is useful to consider a proton located on a planar surface to deduce the magnitude of the factors. In this case, the oxygen or other relaxation agent may approach the detected proton from only one side of the plane; thus, one anticipates a reduction in the relaxation rate by a factor of about 2. In a macromolecule such as ribonuclease A (RNase A), the proton on the surface may have van der Waals contact with the oxygen from one side but the distance of closest approach on the opposite side may be several tens of angstroms away on the other side of the protein. Because of the short electron relaxation time for oxygen, the paramagnetic relaxation rate constant and the steric factors may be computed using a lattice model. The protein is placed in a box of lattice points, the dimensions of which are defined by the largest dimensions of the protein in each coordinate plus the van der Waals diameter of the oxygen molecule. The box is then enlarged until the lattice sums computed do not change in the fourth place as computed for a hydrogen atom. The total paramagnetic contribution to the relaxation rate constant is then a sum of contributions from each lattice point weighted by the volume element appropriate for the grid size, 0.025 Å, setting to zero any point that is inside a protein atom or inside the molecular surface defined by the van der Waals contact between oxygen and the atoms of the protein. This calculation agrees well with the paramagnetic contribution to proton spin−lattice relaxation rates measured in small molecules and incorporates the several aspects of excluded volume and steric factors associated with the anisotropy in the distances of closest approach between the protein protons and the oxygen molecule.[32] The calculation also establishes the relaxation rate constants appropriate for the force-free or hard-sphere limit for the intermolecular potential. The strategy of the approach developed here is then to compare measured relaxation rate constants with those computed from the lattice model to deduce local concentrations of the oxygen. The geometric factors, f_k, are computed on the basis of the structural coordinates of the protein deduced from the X-ray structure or NMR structure. The remaining unknown in eq 4 is then K_k, which provides the free energy difference between the bulk solution and the region immediate to the detected proton spin through the classic relation

$$\Delta G_k^\circ = -RT \ln K_k = -RT \ln\left(\frac{R_{1\mathrm{meas}}^{\mathrm{para}}}{R_{1\mathrm{calcd}}^{\mathrm{para}}}\right)_k \tag{5}$$

An alternative way to define the equilibrium constant K_k is in terms of a specific independent binding constant

$$K_k' = \frac{[PrO_2]_k}{[Pr_k][O_2]} \tag{6}$$

where $[O_2]$ is the equilibrium solubility of oxygen in water at this pressure. The concentration of oxygen in the *k*th site is then

$$[PrO_2]_k = \frac{K_k'[O_2]Pr_T}{1 + K_k'[O_2]} \rightarrow K_k'[O_2]Pr_T \tag{7}$$

where Pr_T is the total protein concentration and the arrow yields the limit of small K_k' and low oxygen concentration, which is appropriate for the present cases. Comparison of eq 7 with eq 4 shows that $K_k = K_k'Pr_T$, and for comparison of the free energy differences on the surface of the protein, the two approaches are equivalent because the total protein concentration divides to unity.

Computation

Protein proton accessible surface areas are computed on the basis of the ACCESS program (v. 2.2.4) from the Center of Structural Biology at Yale University (http://www.csb.yale.edu/download/download_descrip.html),[33] modified by David LeMaster.[2] The distance of closest approach, *b*, for the intermolecular electron−nuclear interaction of each protein proton is defined by computing the shortest distance between a proton and the calculated Connolly protein molecular surface defined by a surface probe with a radius of 1.4 Å (Molecular Simulations Inc., San Diego, CA). The lattice sums were computed using a Fortran program written in this laboratory using a lattice spacing of 0.025 Å in each coordinate based on the coordinates retrieved from the Protein Data Bank (7RSA).[34] The results are summarized in the Supporting Information.

NMR Spectroscopy

All NMR data were acquired using a 500 MHz Varian Unity Plus spectrometer. Both T_1-weighted gradient double-quantum-filtered correlation spectroscopy (gDQF-COSY) spectra and one-dimensional proton (^{1}H) spectra were acquired for each sample. gDQF-COSY spectra were recorded in the phase-sensitive mode using quadrature detection in the directly detected dimension (D2) and hypercomplex detection in the indirectly detected dimension (D1). The carrier frequency was centered at the HOD resonance frequency which is also used as the internal reference at 4.75 ppm. The RNase A spectra were acquired at 30 °C with spectral widths of 5000 Hz in both dimensions. Each acquisition consisted of 16 transients, containing 1024 complex points in D2 and 256 complex points in D1, corresponding to a total acquisition time of ∼3 h for every 1 s of saturation recovery delay.

The proton T_1-weighed pulse sequence was edited by adding a ^{1}H $\pi/2$ pulse (saturation pulse), followed by a saturation recovery delay before a gDQF-COSY pulse sequence.[35,36] Two additional gradient pulses were added immediately before and after the saturation pulse to spoil any remaining coherences. The delay table was constructed according to the response of the average peak intensity of the protein protons in a one-dimensional inversion recovery experiment for each sample. The spectra were typically measured with a saturation recovery delay list of 0.105, 0.233, 0.357, 0.511, 0.693, 0.911, 1.20, 1.61, 2.30, 3.0, 4.0, 5.0, and 6.0 s for the diamagnetic sample and 0.05, 0.1, 0.156, 0.244, 0.304, 0.402, 0.528, 0.706, 1.01, 1.31, 2.0, 3.0, 4.0, and 5.0 s for the paramagnetic sample.

The time domain data were first filtered with a $\pi/2$ phase-shifted sine-bell window function in both dimensions and then zero-filled to 2048 points in D2 and to 1024 points in D1 prior to Fourier transformation. Absolute values of cross-peak intensities were used to simplify the effects of antiphase components. All data were processed using VNMR 6.1B (Varian Inc., Palo Alto, CA) and Felix 2000 NMR software (Molecular Simulations Inc., San Diego, CA). ^{1}H chemical shift assignments for RNase A were used as reported at 30 °C and pH 3.2[37] and at 35 °C and pH 4.0.[38]

The peak intensities in a COSY spectrum were measured by integrating the peak magnitude over the peak area. Negative multiplet components of a normal COSY cross-peak were inverted to positive values so that the integration of the peak magnitude over the peak area gives nonzero values of peak intensities. However, peak overlap is not unusual, which may result in composite peak intensities that could interfere with the determination of the relaxation parameters unique to each proton in the protein structure. In such a case, peak areas were defined to include only multiplet components that did not overlap. Spectral peaks where clear isolation was not possible using this strategy were not used. Variations in relaxation rates as a result of the choice of peak area were small compared to the standard error from the data fitting routine.

A total of 463 cross-peaks, including some slowly exchanging amide proton cross-peaks, were unambiguously assigned in the 2D COSY spectra. Since scalar coupling occurs between nuclei that are separated by one to three chemical bonds, a specific proton may be coupled to two distinct protons, resulting in two cross-peaks with the same relaxation character in the COSY spectrum. For example, if protons A, B, and C are coupled to one another, six cross-peaks at (ωA, ωB), (ωA, ωC), and (ωB, ωC) and symmetrically (ωB, ωA), (ωC, ωA), and (ωB, ωC) can be observed in the spectrum. However, both cross-peaks at (ωA, ωB) and (ωA, ωC) characterize the relaxation parameters unique to proton A. The resonance frequency in the second dimension does not encode the initial population of proton B or C and merely shows connectivity as a result of scalar coupling.

Ideally, relaxation parameters characterized by cross-peaks at (ωA, ωB) and (ωA, ωC), for example, should give the same ΔR_1 value for proton A; however, fitting to an exponential is subject to error that is propagated by the necessity of subtracting two relaxation rate constants, that is, that for the paramagnetic oxygenated sample and that for the deoxygenated sample under the same pressure of nitrogen.

$$\Delta R_1 = \Delta R_1^{O_2} - \Delta R_1^{N_2} \tag{8}$$

In the case where the values of ΔR_1 for proton A as measured by different cross-peaks at (ωA, ωB) and (ωA, ωC) differed slightly, the cross-peak with a greater signal-to-noise ratio, a smaller residual in the fit to a single exponential, or more complete isolation from other cross-peaks was used to characterize the relaxation for proton A. For most of the cross-peaks that carry redundant information, the differences in the relaxation rates are small and comparable to the standard errors of the fit (∼10%). Some cross-peaks have larger errors because of rapid transverse relaxation or longitudinal relaxation.

In the COSY spectra, 300 out of 463 cross-peaks were uniquely resolved out of a total of 787 protons in RNase A, excluding amide protons. The relaxation parameters for amide protons were not characterized because both the proton relaxation and the hydrogen/deuterium exchange mechanism may affect the peak intensities. Nevertheless, the number of reporter sites within a protein from a COSY spectrum is large; most of the residues provide more than two reporter protons. Some residues, such as 8F, 115Y, and 120F, provide five resolvable reporter protons, and all reporter protons of the same residue render unique relaxation rate constants for measurements of oxygen accessibility at that site.

The apparent proton spin−lattice relaxation rate, R_1, characterized by a cross-peak was determined by fitting the peak intensities to a single exponential (eq 2) using the nonlinear least-squares fitting routine (Levenberg−Marquardt method) in the Mathematica 3.0 program (Wolfram Research, Inc., Champaign, IL). The cross-peak intensity data were fit to the equation

$$I(t_i;R_1,I_0) = I_0 \times (1 - e^{-t_i \times R_1}) + B \tag{9}$$

where I_0 is the intensity at $t = \infty$, R_1 is the protein proton spin−lattice relaxation rate constant, and B is the background intensity. If the intensity I_0 is normalized to 1, then we fit the only parameter R_1; B is experimentally measured by integrating the

noise density over the width of the COSY peak. Assuming a normal distribution of error between the data and the fit,

$$D(t_i) - I(t_i;R_1,I_0) \sim N(0,\sigma) \tag{10}$$

To find the maximal likelihood, the parameter R_1 is adjusted to minimize χ^2 defined as

$$\chi^2(R_1) = \sum_{i=1}^{m}(D(t_i) - I(t_i;R_1))^2 \tag{11}$$

R_1 is found from the condition

$$\frac{\partial\chi^2}{\partial R_1} = 0 \tag{12}$$

Substituting eq 9 into eq 12 yields

$$\sum_{i=1}^{m}((D(t_i) - I(t_i;R_1)) \times t_i e^{-t_i \times R_1} = 0 \tag{13}$$

We note that in eq 13 the difference between the experimental measurement $D(t_i)$ and the hypothesis $I(t_i;R_1)$ is weighted by a factor of $t_i e^{-t_i \times R_1}$. In the range of observed protein proton spin−lattice relaxation rates, the initial points of the data collected have a much stronger influence in determining the parameter R_1 than the later points. Therefore, the values of R_1 extracted are not very sensitive to the effects of spin diffusion, which appear at long times.

Results and Discussion

Comparisons with Hard-Sphere Models. If only steric factors are important in intermolecular exploration, the measured paramagnetic contributions to the proton relaxation rate should be proportional to the accessible surface area of each proton. For the relaxation rates computed from lattice sums of the dipolar couplings, this is the case, as shown in Figure 2A. The correlation is good ($r^2 = 0.98$ and $p = 0$); the scatter derives from the discreteness of the lattice and the adjacent steric constraints that do not affect the surface area in the same way as the relaxation rates which are weighted as the inverse sixth power of the distance. Experimental relaxation rate constants do not scale simply with accessible surface area based on the crystal structure, as shown in Figure 2B ($r^2 = 0.049$ and $p = 0.403$). Thus, factors other than exposed surface area are important, but we note that the relaxation rate constants do not correlate simply with other often-used metrics such as hydrophobicity indices of each amino acid, charge, or Debye−Waller factors.[15]

A classic and major difficulty in local measures of spin−lattice relaxation rate constants in macromolecules is the contribution from spin diffusion or proton−proton dipolar coupling. Although initial decay constants are little affected, the effects at long times may be significant. In this work, we minimize χ^2 to find the spin−lattice relaxation rate, R_1, in eq 9, which is predominately determined by the initial points of the data sets (details in the Supporting Information) shown in Figure 3. Indeed, at long times, deviations from exponential decay are apparent. There are several ways to minimize the effects of spin diffusion; one is to measure T_2. However, the dynamic range of the relaxation experiment is compressed because protein T_2 values are fairly short, which requires significantly higher concentrations of paramagnet to develop a reliable measurement of the paramagnetic contributions. An alternative and technically challenging approach is to execute a magnetic field cycle on the sample so that the sample resides at a magnetic field where the cross-relaxation rate is zero during the relaxation period. For the surface protons, the paramagnetic contributions to the individual protons will be underestimated if spin diffusion is significant because the interior protons will generally be less well relaxed by the paramagnet and thus slow the surface proton relaxation rate through the spin coupling. A possible consequence is that the local concentrations and thus the free energy differences deduced by this approach will be systematically somewhat underestimated. Thus, in the present work, we will focus discussion on the surface residues where the paramagnetic contributions are the largest and local motions minimize the efficiency of proton−proton coupling. We note that the approach we present is, however, general and provides an interesting map of the energy landscape of the protein−water interface.

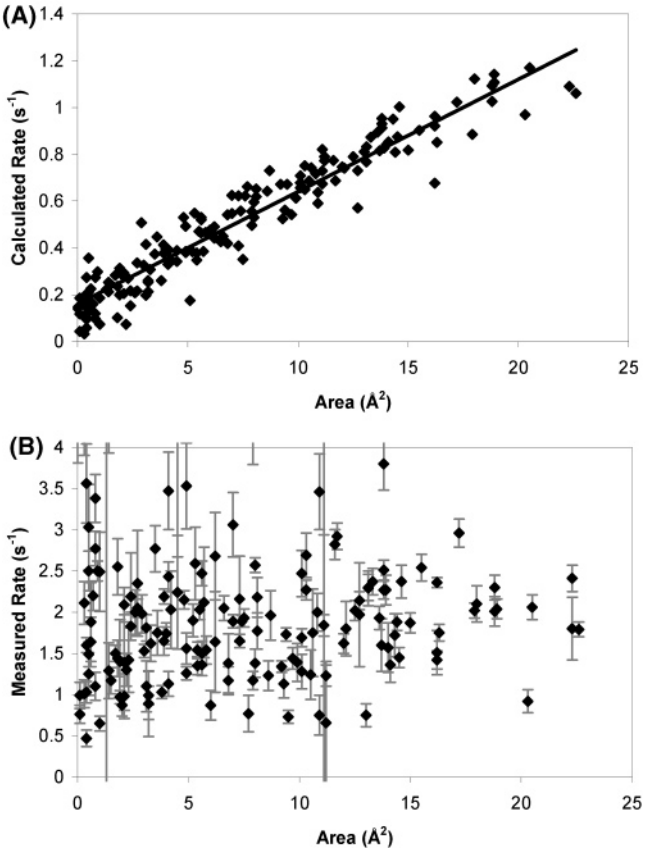

Figure 2. (A) Computed proton spin−lattice relaxation rate constant as a function of exposed surface area based on the Lee and Richards algorithm[33] for ribonuclease A. (B) Experimental proton spin−lattice relaxation rate constant as a function of exposed surface area of proton computed from the Lee and Richards algorithm.

In summary, hard-sphere models fail to account for the magnitude of the spin−lattice relaxation rate constants or their distribution. The relaxation rates computed using the lattice sum approach provide a useful reference that includes the geometric factors created by the details of the folded protein structure as well as the steric factors of the chemical bonding at each proton site but is otherwise *force free*. Substitution of eq 4 into eq 1 provides the basis for analyzing the experimental relaxation rate constants at each site in terms of the effective local concentration of the paramagnet at each site taking the computed rate as the reference. We note that, in nonaqueous solutions of small solutes, work in this laboratory has shown that the computed

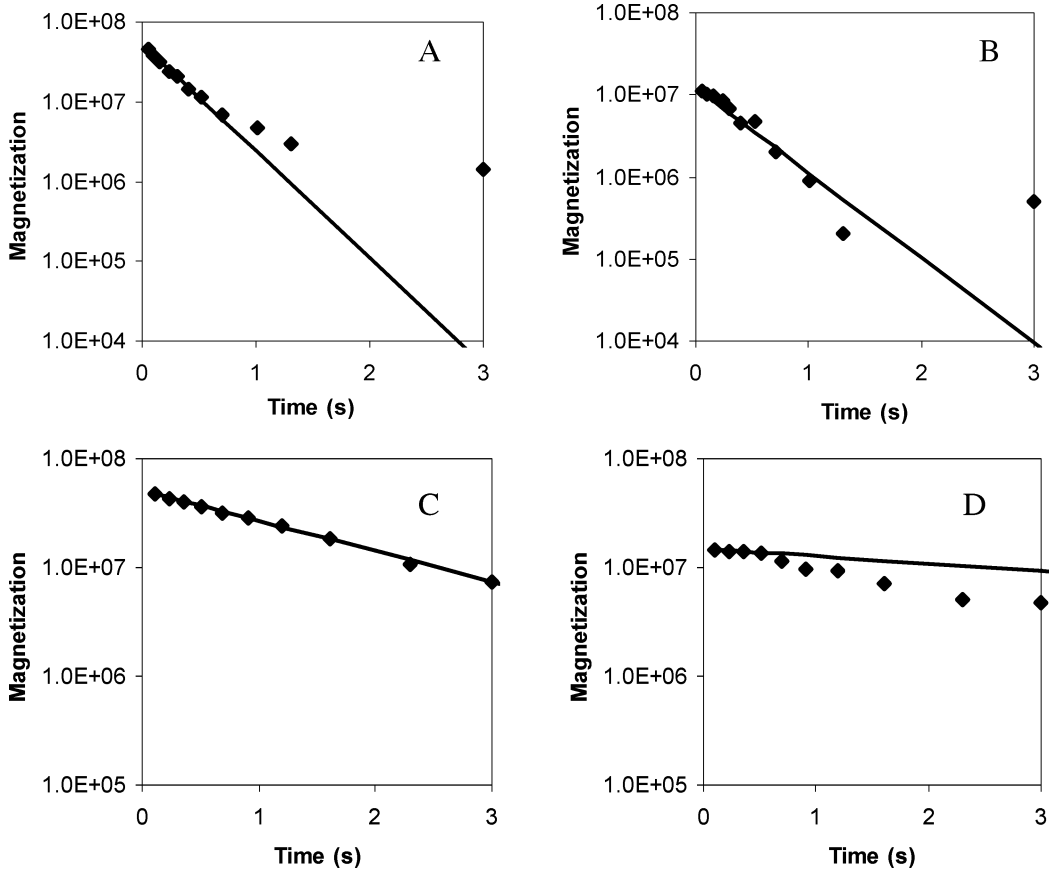

Figure 3. Semilogarithmic graphs of the magnetization decay vs time in the COSY experiments for protons of ribonuclease A. The magnetization is in arbitrary units. (A and B) (34HB2:HB1) An example of the relaxation curves of a surface proton in the presence of oxygen and nitrogen. (C and D) (17HG1:HB) An example of the relaxation curves of an interior proton (depth 4.0 Å) in the presence of oxygen and nitrogen.

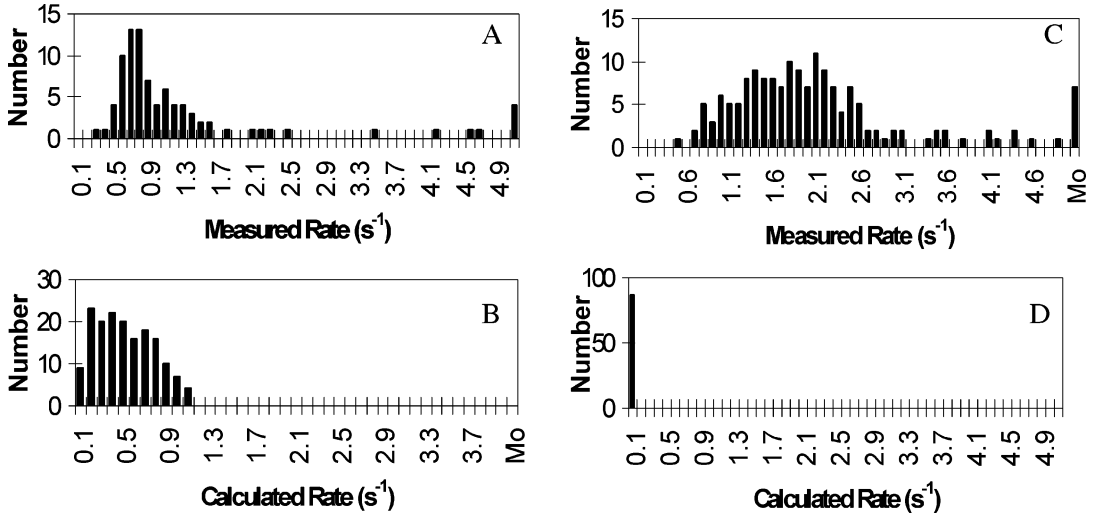

Figure 4. (A) Distribution of the measured paramagnetic relaxation rate constant contributions from oxygen for 165 resolved surface protons of ribonuclease A. (B) Distribution of the computed paramagnetic relaxation rate constant contributions from oxygen for the surface protons of ribonuclease A using a lattice model with a 0.025 Å grid spacing. (C) Distribution of the measured paramagnetic relaxation rate constant contributions induced by oxygen for interior protons in ribonuclease A. (D) Distribution of the calculated relaxation rate constants for interior protons in ribonuclease A using a lattice model with a 0.025 Å grid spacing.

relaxation rates agree with the measured relaxation rates within experimental error after the geometrical factors have been included.[32]

Relaxation Rate Distributions. Protons that are within 0.1 Å of the protein surface defined by the Lee and Richards criterion[33] are considered as surface protons, while protons that are buried more than the short axis of the oxygen molecule or 1.3 Å deeper than the surface are considered as interior. The measured and computed paramagnetic contributions to the surface proton relaxation rate constants are summarized in parts A and B of Figure 4, respectively, based on the high resolution crystal structure (7RSA);[34] the interior proton data are summarized in Figure 4C and D. The measured relaxation rate constants are generally larger than expected on the basis of the calculation assuming completely free diffusion. For the interior protons, the difference is dramatic and affected significantly

by specific interactions where the oxygen affinity is larger than that associated with diffusional contact which have been described in detail.[19] The computed interior proton spin−lattice relaxation rates at large effective intermoment distances or distances of closest approach implied by the Lee and Richards contact surface are not consistent with the measured interior proton relaxation rates. These differences may arise from several factors in addition to weak but specific binding to interior sites: The oxygen may penetrate the protein interior, which may be made more accessible by structural fluctuations. The surface proton relaxation rates, which are generally larger, may raise the interior proton relaxation rates by interproton dipole−dipole couplings, that is, spin diffusion. The magnitude of this effect is difficult to model quantitatively because the assumption of a uniform global correlation time for the protein and solving the coupled differential equations predict interior paramagnetic relaxation rates that are considerably larger than those observed. We note that the spin diffusion problem may be experimentally defeated using highly deuterated proteins and by exploiting nitrogen-15 or carbon-13 labeled proteins. However, these approaches may reduce the number of spins available to report molecular accessibility. We focus quantitative interpretation on the surface protons where the effects of spin diffusion are the smallest; the effect of spin diffusion is to reduce the surface proton relaxation rate constants and the associated free energy differences.

The surface proton relaxation rate constants are generally larger than the computed relaxation rate constants. If an oxygen molecule were bound in van der Waals contact at a specific proton site for the duration of the relaxation time measurement, the relaxation rate constant would be 6200 s^{-1}.[39,40] Weak oxygen binding interactions have been discussed in detail;[19] those that have been detected are weak with an effective occupancy of the order of 0.1−0.5% in the strongest cases. For surface protein protons, such weak binding interactions have the same net effect of increased local oxygen concentration.[15] The surface proton relaxation distribution includes some protons with small relaxation rate constants. There are not many of these sites, which may result in part from the protein structural fluctuations changing the effective exposure to the diffusing paramagnet. The small calculated relaxation rate constants tend to be localized in the concave hydrophobic regions of the protein surface. However, these regions also provide a favorable opportunity for the hydrophobic oxygen molecule to escape from the water, which may cancel the apparent reduction in accessibility implied by the X-ray diffraction data that is the basis for the calculated reference rate. As expected, no large rates are computed for the surface protons.

On average, the fact that the measured surface relaxation rate constants are larger than the computed relaxation rate constants demonstrates that the local oxygen concentration at the surface is larger than that implied by the solubility of oxygen in water by approximately a factor of 3. The calculated relaxation rates for surface protons are plotted against the measured relaxation rates in Figure 5. Unlike the case of small molecule in nonaqueous solvents, the agreement is very poor. Therefore, factors other than steric constraints must dominate the measurement. The reference state implicit in eq 5 is the bulk solution with the assumption that the oxygen accessibility to the protein protons is determined by hard-sphere potentials and the geometric constraints implied by the structure of the folded protein. The free energy deduced from the ratio of relaxation rate constant differences includes contributions to the increased oxygen contact from fluctuations in the protein structure which

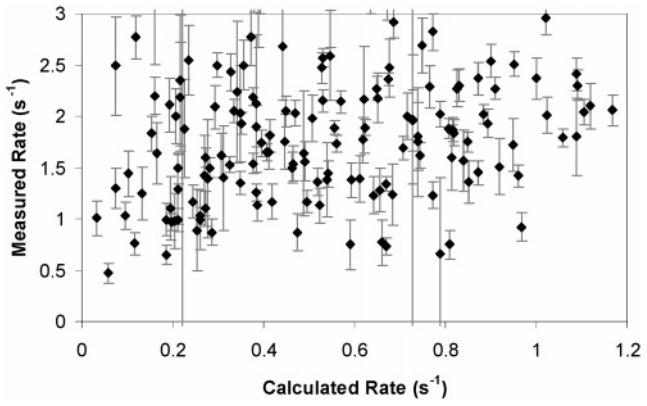

Figure 5. Measured oxygen induced relaxation rate constant contributions plotted against the relaxation rate constants computed from a lattice model using a 0.025 Å grid spacing for surface protons of ribonuclease A.

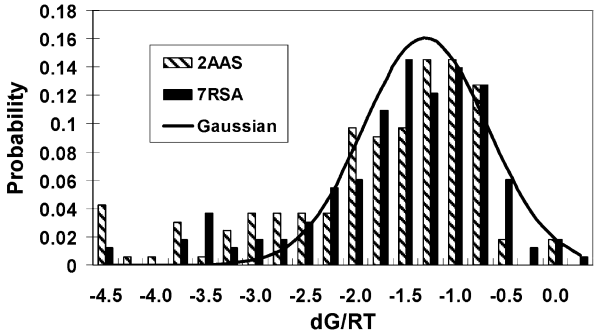

Figure 6. Distribution of free energy values deduced for each site on the basis of eq 5 for surface protons of ribonuclease A. The coordinates reported for a high resolution crystal structure (7RSA) were used[34] for one calculation, and the average of the free energies for each of 11 coordinate sets reported on the basis of the NMR solution structure (2AAS) were used for the other.[46]

may increase the apparent accessibility. The values provided by eq 5 are then the free energy differences for oxygen between the bulk solution and the vicinity of each detected surface proton and include effects of structural fluctuations and spin diffusion that act in opposite ways. The measurement error contributes an average uncertainty of approximately 10% to the free energy deduced from the ratio of relaxation rate contributions.

Figure 6 shows the distribution of surface proton free energy differences reported in units of *RT* using the crystal structure 7RSA[34] and a subset of the 32 NMR structures 2AAS[41] as reference coordinates. The solid line is a Gaussian fit to the distribution that accounts for the shape well except for several points in the region of high affinity. The distribution corresponding to the solid line is centered at $-1.26RT$ with a full width of $1.2RT$, which corresponds to a factor of 3.5 in the average effective local oxygen concentration from one side of the distribution to the other. The strongest interactions are not well determined because of the uncertainty in the relaxation rates related to the selection of delay times in the data acquisition; however, that the oxygen induced relaxation at these protons is not in doubt on the basis of both the line widths and the rapid decay of longitudinal magnetization.

The measured relaxation rate constants that enter the construction of Figure 6 represent averages over the dynamics of the protein; however, the reference state employed was either the static X-ray structure or a selection of NMR structures. The trajectories of the protons are complex, and Figure 6 includes the effects of the dynamical averaging over the acquisition time of the experiment in the magnitude of the free energy difference

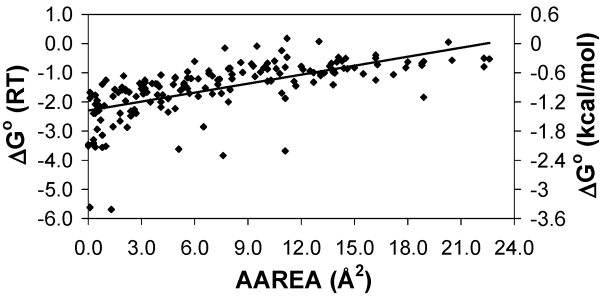

Figure 7. Free energy values deduced from eq 5 vs the accessible area of each surface proton for ribonuclease A. No correction has been made for local surface motions, although a strategy for first-order corrections has been presented elsewhere.[15]

relative to that for the static structure. One may approach the magnitude of the dynamical correction for the local motion of the protons by considering a harmonic displacement of the proton in the CH bonds using the temperature dependence of the B-factors as a guide.[19] An alternative is to use different structures in the set reported on the basis of the solution NMR determinations and compare the free energy differences deduced from these with those deduced from the crystal structure. We have done this for 11 of the coordinate sets reported on the basis of the NMR structure.[41] Although the computed relaxation rate for a proton that is partially buried by neighbors at the interface changes from one coordinate set to another with a standard deviation of the order of 0.25*RT*, the width of the distribution is 4 times larger than this variation. Computing the free energy distribution from the average of 11 of the NMR coordinate sets yields a similar free energy distribution, as shown in Figure 6; the results based on the NMR coordinates as a reference change little with the addition of more than five coordinate sets. The problem of an appropriate reference structure is of greater concern for the protein protons than for a small molecule solution. However, Figure 6 shows that although different reference coordinate systems may cause individual proton energy differences to change, the distribution of free energy differences is relatively insensitive to the coordinate choices and Figure 6 provides a useful characterization of the distribution of intermolecular free energy differences.

The values of the free energy associated with each resolved surface proton site correlate with the accessible surface area of the proton on the basis of the X-ray structure (7RSA coordinates), as shown in Figure 7. This correlation coefficient is 0.628, and p is practically 0. The correlation is still poor but much better than that of Figure 2C because the steric factors that affect accessibility are incorporated in the lattice sum calculations that serve as the reference. The largest absolute deviations occur below the line or to large negative free energies that may reflect the effects of weak binding interactions. These deviations are large for protons when the computed surface area approaches zero and the free energy differences increase, particularly in regions of the protein that are hydrophobic and concave in the molecular surfaces. We note that it is unlikely that such a simple correlation should be very good because the surface area metric does not directly include contributions from adjacent positions and side chains that may modulate the local free energy significantly.

The surface proton data are summarized differently in Figure 8, which shows the free energy difference represented on a two-color scale ranging from red to white to blue. White represents the center of the distribution in Figure 6. Red corresponds to the largest negative free energies relative to the mean, and blue, the largest positive free energies relative to the mean. The largest

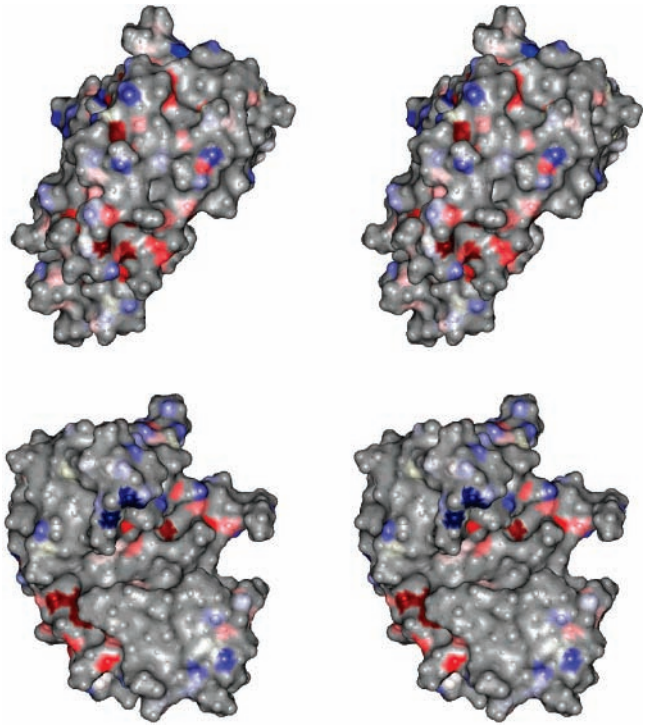

Figure 8. Stereopair representations of the surface of ribonuclease A defined with a 1.4 Å probe at two different orientations. The values of the intermolecular free energy deduced from eq 5 are represented on a two-color scale; the most red color corresponds to free energies equal to 1.26*RT* larger than the mean of the distribution shown in Figure 6, white, the mean of that distribution, which is still a negative free energy relative to the hard-sphere model, and the most blue, 1.26*RT* times more positive free energy than the mean of the distribution in Figure 6.

negative free energy regions correspond to concave regions of contact with the molecular surface defined by a 1.4 Å probe size. In addition, these regions often have favorable side chain interactions that provide the hydrophobic oxygen with an alternative hydrophobic escape from the water where the entropic price has already been paid by the protein. The smallest interactions are found in the relatively convex regions of the contour surface, which is consistent with weaker hydrophobic contributions to the free energy.[42−44]

Figure 1 shows that the presence of an intervening water molecule with high occupancy would reduce the oxygen relaxation rate contribution dramatically on the basis of the hard-sphere collision model, which is the reference for measurements. However, the distribution of relaxation rate contributions from oxygen for the surface protons shows practically no dramatic reduction in accessibility to the paramagnet implied by Figure 1; there are practically no large positive free energy values deduced. Therefore, while there is a distribution of effective interaction strengths of the oxygen with different regions of the protein surface that undoubtedly reflect differences in local solvation effects, the positions identified in the crystal structure with strong localization of water do not appear strongly in these experiments that probe the water occupancy indirectly by the accessibility to oxygen. For surface proton positions where one water molecule is reported to be localized in the crystal structure within 3 Å of the observed proton, the range of relaxation rate constants is 0.47 ± 0.1 to 7.4 ± 1.35 s^{-1}. When more than one water molecule is identified in the crystal structure within 3 Å of the observed proton site, the relaxation rate constants are all smaller than 3 s^{-1}, which is consistent with modest steric exclusion of the oxygen from these protons by water, yet water

molecule induced oxygen exclusion is far from complete. The surface free energy differences may also be modulated by binding of counterions, which may be quite large as hydrated species. We note that magnetically silent ion binding sites may be identified on the basis of paramagnetic effects of explorer species, although they are not of clear importance in the present case.[45]

Conclusion

The use of freely diffusible paramagnetic centers provides a useful characterization of the free energies that govern how one molecule interacts with another on a local scale determined by the spectral and spatial resolution provided by the proton NMR spectrum. The basic idea is similar to spectroscopic quenching experiments, but the localization depends on the strong distance dependence of a dipole−dipole coupling. The present work advances an analytical strategy that provides intermolecular free energy differences sensed at hundreds of places on a large molecule such as a protein. This work has focused on the surface sites interacting with the freely diffusing oxygen molecule, which is found to contact essentially all surface sites more frequently than predicted from a hard-sphere force-free model taken as the reference. Correlations between the observed paramagnetic contributions to the proton spin−lattice relaxation rate constants and obvious physical parameters such as the exposed surface area of an observed proton are largely absent because the factors that regulate local oxygen concentrations are not uniform. The local oxygen concentrations are generally higher than those predicted by force-free models by a factor of the order of 3; however, the local interactions differ considerably in water. The distribution of free energies that characterize the local oxygen−protein interaction is centered at $-1.25RT$ (-3.1 kJ/mol), and the nearly Gaussian distribution has a full width of $1.24RT$ (3.1 kJ/mol). Although no strong oxygen binding sites are found, there are several positions both outside and inside the protein that have magnetic coupling to oxygen that is significantly greater than the average. Nevertheless, even these sites are populated only at a level of 0.2% or less. The energy of any one of these interactions is small; however, the several hundred characterized represent a fraction of the total associated with the intermolecular interaction. The local heterogeneity detected provides a useful view of how the energetic landscape varies as a small molecule explores a protein in water.

Acknowledgment. This work was supported by the National Institutes of Health, GM34541, EB002054, and the University of Virginia. We gratefully acknowledge helpful discussions with Professors David Cafiso, Rodney Biltonen, and Jack Freed and Drs. Griselda Hernandez and David LeMaster.

Supporting Information Available: Table showing for each residue detected the sequence number, amino acid, observed proton, coupled proton, relaxation rate change induced by the oxygen, standard error, intermoment distance, accessible surface area, and steric accessibility factor. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Pintacuda, G.; Otting, G. *J. Am. Chem. Soc.* **2002**, *124*, 372.
(2) Hernandez, G.; Teng, C. L.; Bryant, R. G.; LeMaster, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 4463.
(3) Teng, C. L.; Bryant, R. G. *J. Am. Chem. Soc.* **2000**, *122*, 2667.
(4) Niccolai, N.; Rossi, C.; Valensin, G.; Mascagni, P.; Gibbons, W. A. *J. Phys. Chem.* **1984**, *88*, 5689.
(5) Zhou, N.; Mascagni, P.; Gibbons, W.; Niccolai, N.; Rossi, C.; Wyssbrod, H. *J. Chem. Soc., Perkin Trans. 2* **1985**, 581.
(6) Niccolai, N.; Lampariello, R.; Bovalini, L.; Rustici, M.; Mascagni, P.; Martelli, P. *Biophys. Chem.* **1990**, *38*, 155.
(7) Esposito, G.; Lesk, A. M.; Molinari, H.; Motta, A.; Niccolai, N.; Pastore, A. *J. Mol. Biol.* **1992**, *224*, 659.
(8) Esposito, G.; Lesk, A. M.; Molinari, H.; Motta, A.; Niccolai, N.; Pastore, A. *Biopolymers* **1993**, *33*, 839.
(9) Molinari, H.; Esposito, G.; Ragona, L.; Pegna, M.; Niccolai, N.; Brunne, R. M.; Lesk, A. M.; Zetta, L. *Biophys. J.* **1997**, *73*, 382.
(10) Niccolai, N.; Ciutti, A.; Spiga, O.; Scarselli, M.; Bernini, A.; Bracci, L.; Di Maro, D.; Dalvit, C.; Molinari, H.; Esposito, G.; Temussi, P. A. *J. Biol. Chem.* **2001**, *276*, 42455.
(11) Niccolai, N.; Spadaccini, R.; Scarselli, M.; Bernini, A.; Crescenzi, O.; Spiga, O.; Ciutti, A.; Di Maro, D.; Bracci, L.; Dalvit, C.; Temussi, P. A. *Protein Sci.* **2001**, *10*, 1498.
(12) Luchette, P. A.; Prosser, R. S.; Sanders, C. R. *J. Am. Chem. Soc.* **2002**, *124*, 1778.
(13) Prosser, R. S.; Luchette, P. A.; Westerman, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9967.
(14) Prosser, R. S.; Luchette, P. A.; Westerman, P. W.; Rozek, A.; Hancock, R. E. *Biophys. J.* **2001**, *80*, 1406.
(15) Teng, C. L. Mapping Molecular Accessibility. Ph.D. Thesis, University of Virginia, 2002.
(16) Freed, J. H. *J. Chem. Phys.* **1978**, *68*, 4034.
(17) Ayant, Y.; Belorizky, E.; Alizon, J.; Gallice, J. *J. Phys. I* **1975**, *36*, 991.
(18) Hwang, L.; Freed, J. H. *J. Chem. Phys.* **1975**, *63*, 4017.
(19) Teng, C. L.; Bryant, R. G. *Biophys. J.* **2004**, *86*, 1713.
(20) Kowalewski, J.; Laaksonen, A.; Nordenskiold, L. *J. Chem. Phys.* **1981**, *74*, 2927.
(21) Polnaszek, C. F.; Bryant, R. G. *J. Chem. Phys.* **1984**, *81*, 4038.
(22) Fries, P.; Belorizky, E. *J. Phys. I* **1978**, *39*, 1263.
(23) Attard, P.; Patey, G. N. *J. Chem. Phys.* **1990**, *92*, 4970.
(24) Fries, P.; Hansen, J. P. *Mol. Phys.* **1983**, *48*, 891.
(25) Dinesen, T. R.; Bryant, R. G. *J. Magn. Reson.* **1998**, *132*, 19.
(26) Dinesen, T. R. J.; Seymour, J.; McGowan, L.; Wagner, S.; Bryant, R. G. *J. Phys. Chem. A* **1999**, *103*, 782.
(27) Hodges, M. W.; Cafiso, D. S.; Polnaszek, C. F.; Lester, C. C.; Bryant, R. G. *Biophys. J.* **1997**, *73*, 2575.
(28) Likhtenshtein, G. I.; Adin, I.; Novoselsky, A.; Shames, A.; Vaisbuch, I.; Glaser, R. *Biophys. J.* **1999**, *77*, 443.
(29) Prosser, R. S.; Luchette, P. A.; Westerman, P. W.; Rozek, A.; Hancock, R. E. *Biophys. J.* **2001**, *80*, 1406.
(30) Prosser, R. S.; Luchette, P. A.; Westerman, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9967.
(31) Teng, C. L.; Hong, H.; Kiihne, S.; Bryant, R. G. *J. Magn. Reson.* **2001**, *148*, 31.
(32) Teng, C.-L.; Martini, S.; Bryant, R. G. *J. Am. Chem. Soc.* **2004**, *126*, 15253.
(33) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379.
(34) Wlodawer, A.; Svensson, L. A.; Sjolin, L.; Gilliland, G. L. *Biochemistry* **1988**, *27*, 2705.
(35) Piantini, U.; Sorensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 6800.
(36) Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479.
(37) Robertson, A. D.; Purisima, E. O.; Eastman, M. A.; Scheraga, H. A. *Biochemistry* **1989**, *28*, 5930.
(38) Rico, M.; Bruix, M.; Santoro, J.; Gonzalez, C.; Neira, J. L.; Nieto, J. L.; Herranz, J. *Eur. J. Biochem.* **1989**, *183*, 623.
(39) Bloembergen, N.; Morgan, L. O. *J. Chem. Phys.* **1961**, *34*, 842.
(40) Solomon, I. *Phys. Rev.* **1955**, *99*, 559.
(41) Santoro, J.; Gonzalez, C.; Bruix, M.; Neira, J. L.; Nieto, J. L.; Herranz, J.; Rico, M. *J. Mol. Biol.* **1993**, *229*, 722.
(42) Cheng, Y. K.; Rossky, P. J. *Nature* **1998**, *392*, 696.
(43) Cheng, Y. K.; Rossky, P. J. *Biopolymers* **1999**, *50*, 742.
(44) Cheng, Y. K.; Sheu, W. S.; Rossky, P. J. *Biophys. J.* **1999**, *76*, 1734.
(45) Diakova, G.; Igbani, M.; Nagatani, G.; Teng, C. L.; Bryant, R. G., unpublished.
(46) Santoro, J.; Gonzalez, C.; Bruix, M.; Neira, J. L.; Nieto, J. L.; Herranz, J.; Rico, M. *J. Mol. Biol.* **1993**, *229*, 722.