J|A|C|S ARTICLES

Published on Web 08/26/2010

Structure-Independent Analysis of the Breadth of the Positional Distribution of Disordered Groups in Macromolecules from Order Parameters for Long, Variable-Length Vectors Using NMR Paramagnetic Relaxation Enhancement

Junji Iwahara*,† and G. Marius Clore*,‡

Sealy Center for Structural Biology and Molecular Biophysics, Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, Texas 77555-0647, and Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received June 2, 2010; E-mail: j.iwahara@utmb.edu; mariusc@mail.nih.gov

Abstract: Quantitative information regarding structurally disordered groups is crucial for a complete understanding of the relationship between structure, dynamics, and function in biological macromolecules. Experimental analysis, however, of the positional distribution of disordered groups in the macromolecular frame is extremely difficult. While NMR order parameters, S^2, for fixed-length bond vectors such as N−H and C−H are commonly used for investigations of conformational dynamics of macromolecules, these order parameters provide only angular information about internal motions and are totally insensitive to translational motions. Although analysis of S^2 for bond vectors permits identification of disordered groups in macromolecules, this type of order parameter cannot provide any information about the distribution radii of disordered groups. Here we describe an NMR approach to directly determine the distribution radius of a disordered group independent of any structural knowledge. This approach makes use of order parameters for long, variable-length vectors (including proton−paramagnetic center and proton−proton vectors) between a disordered group and a rigid portion of the macromolecule. We demonstrate the application of this formalism to paramagnetic relaxation enhancement vectors. In addition, the potential utility of the same formalism to ^{1}H−^{1}H cross-relaxation rates is considered as an alternative approach for analyzing the breadth of the positional distribution of disordered groups.

Introduction

Highly flexible or structurally disordered regions in macromolecules, such as surface-exposed loops and disordered tails, often play a key role in function, for example by increasing the "capture radius" [1,2] in macromolecular association.[3−5] Experimental analysis, however, of the positional distribution of disordered groups in the molecular frame is challenging: time-resolved fluorescence energy transfer provides the distance distribution between two fluorophores[6] but does not yield the distribution radius for a disordered group since both extrinsic fluorophores are attached to the macromolecule through mobile linkers; small-angle X-ray scattering provides radial information at low resolution for the entire macromolecule[7] but is not suitable for probing local structural disorder; crystallographic *B* factors can, in principle, provide the mean-square displacement for each atom but are either unavailable or unreliable for structurally disordered regions owing to very weak electron densities;[8] finally, NMR order parameters (S^2) for bond vectors (e.g., N−H and C−H) provide local dynamic and angular information on internal motions but are completely insensitive to translational motions.[9] These issues have hindered experimental studies on the positional distributions of disordered groups in macromolecules, necessary to understand the interplay between structure, dynamics, and function. Here we describe a formalism to directly determine the distribution radius of a disordered group by NMR independent of any structural knowledge. This approach makes use of order parameters for long, variable-length vectors (including proton−paramagnetic center and proton−proton vectors) between a disordered group

† University of Texas Medical Branch.
‡ National Institute of Diabetes and Digestive and Kidney Diseases.

(1) Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8868–8873.
(2) Levy, Y.; Cho, S. S.; Onuchic, J. N.; Wolynes, P. G. *J. Mol. Biol.* **2005**, *346*, 1121–1145.
(3) Spolar, R. S.; Record, M. T., Jr. *Science* **1994**, *263*, 777–784.
(4) Wright, P. E.; Dyson, H. J. *Curr. Opin. Struct. Biol.* **2009**, *19*, 31–38.
(5) Gsponer, J.; Babu, M. M. *Prog. Biophys. Mol. Biol.* **2009**, *99*, 94–103.
(6) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3rd ed.; Springer: New York, 2006.
(7) Bernado, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007**, *129*, 5656–5664.
(8) Kuriyan, J.; Petsko, G. A.; Levy, R. M.; Karplus, M. *J. Mol. Biol.* **1986**, *190*, 227–254.
(9) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.

10.1021/ja1048187

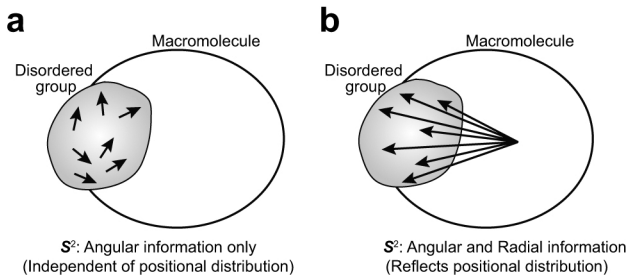

Figure 1. Different characteristics of the order parameters $\mathbf{S}^2$. $\mathbf{S}^2$ for (a) a short, fixed-length vector (e.g., N−H or C−H bond) in a disordered group versus (b) a long, variable-length vector (e.g., PRE or NOE interaction vector) between an atom in a disordered group and another atom in a rigid portion of the macromolecule. Each arrow represents a different state of the same vector in the ensemble.

and a rigid portion of the macromolecule. An application to paramagnetic relaxation enhancement vectors is demonstrated.

Theory

The order parameter $\boldsymbol{S}^2$ for a vector whose length varies due to dynamics includes both angular and radial components and can therefore provide information on the positional distribution. A general expression of $\boldsymbol{S}^2$ that is applicable to both fixed- and variable-length vectors is given by[9−11]

$$\boldsymbol{S}^2 = (4\pi/5)\langle r^{-6}\rangle^{-1} \sum_{m=-2}^{2} |\langle Y_2^m(\theta,\phi)/r^3\rangle|^2 \qquad (1)$$

where $\langle ...\rangle$ represents an ensemble average; r is the distance between the two atoms; Y_2^m are spherical harmonics given by $Y_2^0 = (5/16\pi)^{1/2}(3\cos^2\theta - 1)$, $Y_2^{\pm 1} = \mp(15/8\pi)^{1/2}(\sin\theta\cos\theta)\exp(\pm i\phi)$, and $Y_2^{\pm 2} = (15/32\pi)^{1/2}(\sin^2\theta)\exp(\pm 2i\phi)$; and θ and ϕ are polar angles for the vector within the molecular frame. For a fixed-length vector, the two contributions from r in eq 1 cancel out, and only the angular component remains in $\boldsymbol{S}^2$. Consequently, order parameters for bond vectors located in a disordered group cannot provide information about the breadth of its positional distribution, although they can indicate that the group is indeed disordered (Figure 1a). On the other hand, $\boldsymbol{S}^2$ for variable-length vectors probed, for example, by paramagnetic relaxation enhancement (PRE) and nuclear Overhauser enhancement (NOE) measurements contain, in principle, both angular and radial components. If the vector is between an atom in a disordered region and another atom in a rigid portion of the molecule, the order parameter for the vector should reflect the positional distribution of the disordered group (Figure 1b). We take advantage of this feature to analyze the positional distribution of the disordered group.

Uniform Distribution in a Sphere Model. To interpret the internal motions of variable-length vectors, we consider a model where the two atoms, Atom 1 and Atom 2, of the vector in the molecular frame are independently and uniformly distributed in two separate spheres with radii R_P and R_N, respectively (Figure 2a). We refer to this as the "uniform distribution in a sphere" model (or "Sphere" model for short).

We first consider the special case of $R_N = 0$; that is, Atom 2 is fixed at a point in the molecular frame. We define the following additional parameters: R_P, the distance between Atom 1 and the center O of its distribution; θ, the Atom 1−Atom 2−O angle; β, the Atom 1−O−Atom 2 angle; r, the distance between Atom 1 and Atom 2; and R, the distance between Atom 2 and the center of the spherical distribution of Atom 1. The radius R_P of the sphere describing the distribution of Atom 1 is assumed to be smaller than the center-to-center distance R. Since $r = (R^2 + r_P^2 - 2Rr_P\cos\beta)^{1/2}$, the ensemble $\langle r^{-6}\rangle$ average for the surface of a sphere with radius r_P is given by

$$\langle r^{-6}\rangle = \frac{1}{4\pi r_P^2}\int_0^{\pi}(R^2 + r_P^2 - 2Rr_P\cos\beta)^{-3}\,2\pi r_P^2\sin\beta\,d\beta = \frac{R^2 + r_P^2}{(R^2 - r_P^2)^4} \qquad (2)$$

Using this result, the ensemble average $\langle r^{-6}\rangle$ for a uniform distribution in a sphere with a radius R_P is calculated to be

$$\langle r^{-6}\rangle = \frac{1}{(4/3)\pi R_P^3}\int_0^{R_P}\frac{R^2 + r_P^2}{(R^2 - r_P^2)^4}4\pi r_P^2\,dr_P = (R^2 - R_P^2)^{-3} \qquad (3)$$

Thus, the effective distance for the "Sphere" model with $R_N = 0$ is given by

$$\langle r^{-6}\rangle^{-1/6} = R\{1 - (R_P/R)^2\}^{1/2} \qquad (4)$$

where R is the distance between the centers of the two spheres.

Because of the symmetry of the model, the ensemble average $|\langle Y_2^m(\theta,\varphi)/r^3\rangle|^2$ in eq 1 is nonzero only for $m = 0$. Hence,

$$\boldsymbol{S}^2 = (4\pi/5)\langle r^{-6}\rangle^{-1}|\langle Y_2^0(\theta,\phi)/r^3\rangle|^2 = \langle r^{-6}\rangle^{-1}\langle P_2(\cos\theta)/r^3\rangle^2 \qquad (5)$$

where the Legendre polynomial $P_2(x)$ is given by $(3x^2 - 1)/2$. Since $\cos\theta = (R - r_P\cos\beta)/r$, the ensemble average $\langle P_2(\cos\theta)/r^3\rangle$ for a uniform distribution within a sphere of radius R_P is given by

$$\begin{aligned}\langle P_2(\cos\theta)/r^3\rangle &= \frac{1}{(4/3\pi)R_P^3}\int_0^{R_P}\left\{\int_0^{\pi}P_2\left(\frac{R - r_P\cos\beta}{\sqrt{R^2 + r_P^2 - 2Rr_P\cos\beta}}\right)\times (R^2 + r_P^2 - 2Rr_P\cos\beta)^{-3/2}\frac{\sin\beta}{2}d\beta\right\}4\pi r_P^2\,dr_P \\ &= \frac{1}{R^3}\end{aligned} \qquad (6)$$

$\boldsymbol{S}^2$ is obtained directly from eqs 3−6:

$$\boldsymbol{S}^2 = \{1 - (R_P/R)^2\}^3 \qquad (7)$$

It is important to note that $\boldsymbol{S}^2$ for this model provides purely radial information with respect to the spatial distribution of Atom 1. Equations 4 and 7 then permit the determination of the distribution radius R_P from $\langle r^{-6}\rangle^{-1/6}$ and $\boldsymbol{S}^2$ as

$$R_P = \langle r^{-6}\rangle^{-1/6}\sqrt{(\boldsymbol{S}^2)^{-1/3} - 1} \qquad (8)$$

If the distribution of Atom 1 is ellipsoidal, then the apparent R_P from this relation approximately corresponds to the average length of the semiprincipal axes (see below).

(10) Olejniczak, E. T.; Dobson, C. M.; Karplus, M.; Levy, R. M. *J. Am. Chem. Soc.* **1984**, *106*, 1923–1930.
(11) Brüschweiler, R.; Roux, B.; Blackledge, M.; Griesinger, C.; Karplus, M.; Ernst, R. R. *J. Am. Chem. Soc.* **1992**, *114*, 2289–2302.

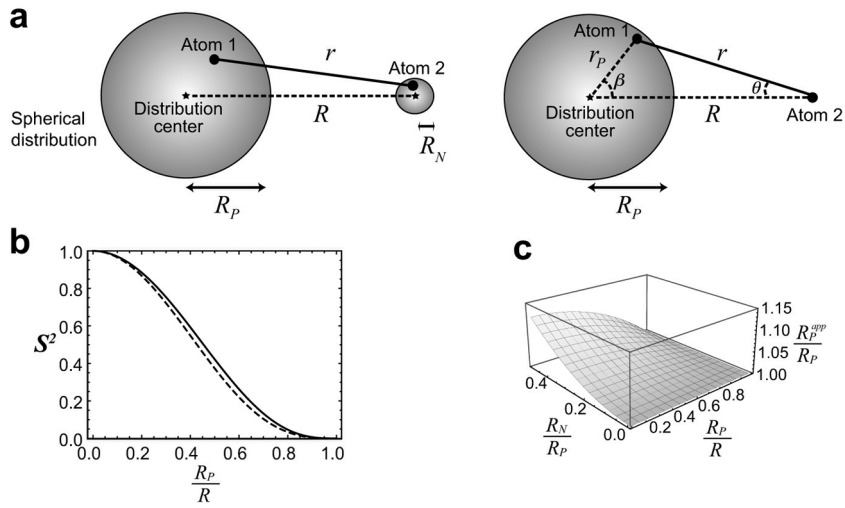

Figure 2. "Sphere" model for the internal motion of a variable-length vector in a macromolecular system. (a) Diagram depicting the model in which Atom 1 and Atom 2 are uniformly distributed in two separate spheres. On the left, Atom 2 is distributed in a sphere of radius R_N; on the right, the case with $R_N = 0$ is shown. (b) Relationship between the order parameter S^2 for the variable-length vector and the ratio R_P/R for the "Sphere" model (solid line, $R_N = 0$; dotted line, $R_N/R_P = 0.4$). (c) Impact of neglecting R_N in the determination of R_P. The ratio of the apparent R_P (obtained from eqs 4, 7, and 8) to the true value of R_P as a function of R_P/R and R_N/R_P is plotted.

Although $R_N = 0$ was assumed for the above derivation, the use of eqs 4, 7, and 8 to estimate R_P for $R_N \neq 0$ is still valid, providing R_N is substantially smaller than R_P (Figure 2b,c). For $R_N \neq 0$, $\langle r^{-6}\rangle^{-1/6}$ and S^2 are approximated by $R\{1 - (R_P/R)^2\}^{1/2}\{1 - (R_N/R)^2\}^{1/2}$ and $\{1 - (R_P/R)^2\}^3\{1 - (R_N/R)^2\}^3$, respectively, providing the internal motions of the two atoms are uncorrelated. Figure 2b shows the relationship between S^2 and R_P/R. The solid line in this figure corresponds to a case with $R_N = 0$, whereas the dotted line corresponds to a case with $R_N \neq 0$ and $R_N/R_P = 0.4$; these curves are very close to one another because of the relatively small scaling introduced by the dynamics of Atom 2. Figure 2c shows the ratio of the apparent R_P obtained from eqs 4 and 7 to the true R_P as a function of R_P/R and R_N/R_P. Although the apparent value of R_P is always larger than the true value of R_P, the error is less than 10% for $R_N/R_P < 0.4$. Thus, eqs 4, 7, and 8 permit the accurate estimation of R_P if R_N is substantially smaller than R_P.

Considerations on the Range of Applicability of the Formalism. Integration of eqs 2, 3, and 6 requires that $R_P < R$. When the condition $R_P \geq R$ occurs in the model shown in Figure 2a, the distance between the nucleus and paramagnetic center can be zero, and r^{-6} would be infinity. Hence, a theoretical treatment is not feasible unless a lower limit for r is given. Such a case can be better dealt with using the "pair correlation functions" originally proposed by Freed et al.[12−14] Indeed, their theory takes the lower limit for r into consideration. However, if $R_P \geq R$, the lower limit of r governs the PRE data, and the calculation of the distribution radius from experimental PRE data is essentially impossible.

Considerations for Nonspherical Distributions. As described above, the distribution radius R_P can be determined from S^2 and $\langle r^{-6}\rangle^{-1/6}$ with eq 8, assuming that the distribution of Atom 1 is spherical. What is the meaning of an apparent R_P (R_P^{app}) if we apply eq 8 for a nonspherical distribution? This question can readily be answered by using an N-state discrete jump model whose spatial distribution is arbitrary. For a discrete lattice model, S^2 and $\langle r^{-6}\rangle^{-1/6}$ are given by[11]

$$S^2 = \left(\sum_{i=1}^{N} p_i r_i^{-6}\right)^{-1} \sum_{j=1}^{N}\sum_{i=1}^{N} p_i p_j P_2(\cos\chi_{ij}) r_i^{-3} r_j^{-3} \qquad (9)$$

$$\langle r^{-6}\rangle^{-1/6} = \left(\sum_{i=1}^{N} p_i r_i^{-6}\right)^{-1/6} \qquad (10)$$

where r represents a vector between Atom 1 and Atom 2; p_i is the probability for the vector r_i; and χ_{ij} is the angle between vectors r_i and r_j. Here we consider the following two types of nonspherical distributions: an ellipsoidal distribution and a bimodal distribution. Other nonspherical/nonuniform distributions can be considered in the same manner.

Numerical calculations of S^2 and $\langle r^{-6}\rangle^{-1/6}$ were carried out using uniformly distributed grid points in an oblate or prolate axially symmetric ellipsoid, and R_P^{app} was calculated with eq 8. As demonstrated in Figure 3, the numerical calculations indicate that the apparent value of R_P is close to the average length of the three semiprincipal axes (i.e., $2R_L/3 + R_S/3$ and $R_L/3 + 2R_S/3$ for oblate and prolate ellipsoids, respectively).

We also considered the bimodal distribution shown in Figure 4a. For this model, Atom 1 is uniformly distributed in two spheres, a and b, in contact with each other (radii R_a and R_b, respectively) with different probabilities (p_a and $1 - p_a$, respectively). The case with $R_b/R_a = 3$ and $R = (R_a + R_b)/2$ was examined. When $p_a = 0.3$ (i.e., the probability for the smaller sphere is low), R_P^{app} for this model is always smaller than $R_a + R_b$ by ~20%, regardless of the position of Atom 2 (green points in Figure 4b). When $p_a = 0.7$ (i.e., the probability for the smaller sphere is high), R_P^{app} can be larger or smaller than $R_a + R_b$, depending on the location of Atom 2, and the deviation from $R_a + R_b$ is as large as 50% (brown points in Figure 4b). In such a case, accurate estimation of the breadth of the distribution from PRE order parameters and effective distances requires sampling of PRE vectors for nuclei at various locations.

(12) Hwang, L. P.; Freed, J. H. *J. Chem. Phys.* **1975**, *63*, 4017–4025.
(13) Freed, J. H. *J. Chem. Phys.* **1978**, *68*, 4034–4037.
(14) Xue, Y.; Podkorytov, I. S.; D.K., R.; Benjamin, N.; Sun, H.; Skrynikov, N. R. *Protein Sci.* **2009**, *18*, 1401–1424.

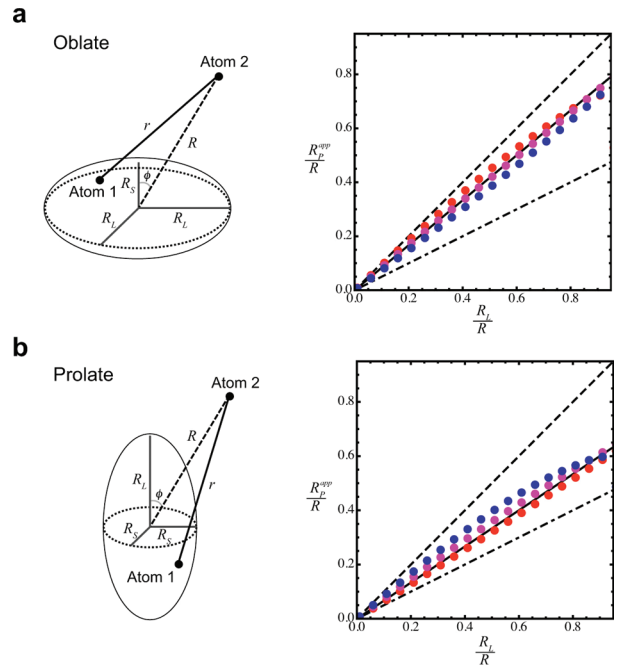

Figure 3. Calculated value of the apparent distribution radius, R_P^{app}, for oblate (a) and prolate (b) axially symmetric ellipsoids. Five parameters are involved: the lengths of the axes R_L and R_S ($R_L > R_S$), the angle ϕ between the symmetry axis and the center−Atom 2 vector, the distance R between the center of an ellipsoid and Atom 2, and the distance r between Atom 2 and an arbitrary position in the ellipsoid. The order parameters S^2 and effective distances $\langle r^{-6}\rangle^{-1/6}$ were calculated using 3581 uniformly distributed grid points for an oblate ellipsoid and 3591 grid points for a prolate ellipsoid. Graphs display apparent R_P/R values (red points, $\phi = 90°$; magenta, $\phi = 45°$; and blue, $\phi = 0°$) as a function of the R_L/R ratio for oblate and prolate ellipsoids with an anisotropy of $R_L/R_S = 2.0$. Dotted and dashed lines correspond to R_L/R and R_S/R, respectively. Solid lines represent $(2R_L/3 + R_S/3)/R$ for an oblate ellipsoid and $(R_L/3 + 2R_S/3)/R$ for a prolate ellipsoid.

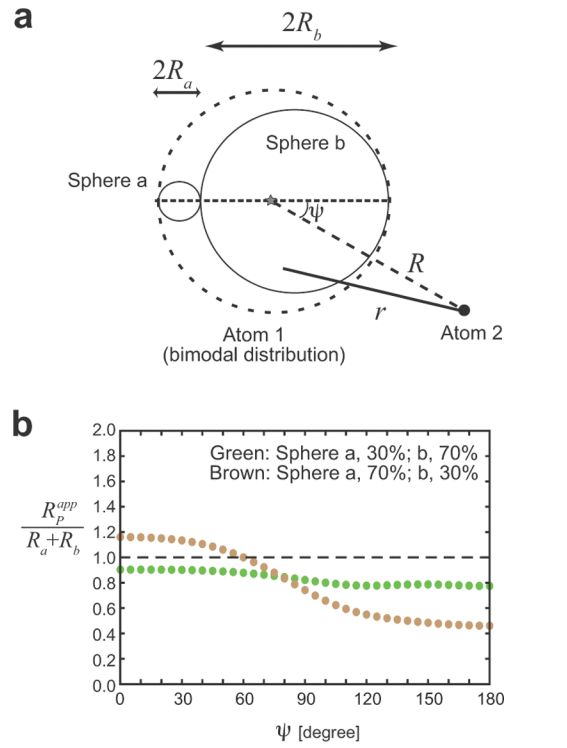

Figure 4. Calculated value of the apparent distribution radius, R_P^{app}, for a bimodal distribution. (a) Model considered to represent the bimodal distribution. In this model, Atom 1 is distributed in two spheres, a and b, that are in contact each other (radii R_a and R_b, respectively), whereas Atom 2 is fixed at a point defined by a distance R to the center of the symmetry axis for the bimodal distribution and an angle ψ between the vector of R and the symmetric axis of the distribution. The distribution of Atom 1 is uniform within each of the two spheres, but their probabilities are different. (b) Ratio of R_P^{app} to $R_a + R_b$ as a function of the angle ψ for a case with $R = (R_a + R_b)/2$ and $R_b/R_a = 3$. Plots were obtained with 3034 grid points in the bimodal distribution. The probability of populating sphere a was 30% for the green points and 70% for the brown points.

Determination of the Distribution Radius from PRE Rates. The above formalism can readily be applied to PRE interaction vectors. The PRE rate Γ is the relaxation rate arising from magnetic dipole−dipole interactions between a nucleus and the unpaired electrons of a paramagnetic center and is measured by taking the difference in the relaxation rates for the paramagnetic and diamagnetic states. We refer to the PRE rates for longitudinal and transverse magnetizations as Γ_1 and Γ_2, respectively. For paramagnetic systems with an isotropic electronic g-tensor (such as Mn^{2+}, Gd^{3+}, and nitroxide), the PRE rates are given by

$$\Gamma_1 = \frac{2}{5}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H^2 g^2 \mu_B^2 s(s+1) J(\omega_H) \qquad (11)$$

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H^2 g^2 \mu_B^2 s(s+1)\{4J(0) + 3J(\omega_H)\} \qquad (12)$$

where s is the electron spin quantum number; g, the electron g-factor; γ_H, the ^{1}H gyromagnetic ratio; μ_0, the permeability of a vacuum; μ_B, the Bohr magneton; and $\omega_H/2\pi$, the Larmor frequency of the proton.[15,16] The spectral density function $J(\omega)$ for a PRE interaction vector undergoing internal motions is given by

$$J(\omega) = \langle r^{-6}\rangle\left\{\frac{S^2\tau_c}{1+\omega^2\tau_c^2} + \frac{(1-S^2)\tau_t}{1+\omega^2\tau_t^2}\right\} \qquad (13)$$

where S^2 is the order parameter for the PRE interaction vector (e.g., a Mn^{2+}-^{1}H vector), and r is the length of the PRE interaction vector (which is variable due to internal motion).[17,18] The correlation times are given by $\tau_c = (\tau_r^{-1} + \tau_s^{-1})^{-1}$ and $\tau_t = (\tau_c^{-1} + \tau_i^{-1})^{-1}$, where τ_r and τ_i are the correlation times for molecular rotation and internal motions, respectively, and τ_s is the electron relaxation time. For ^{1}H nuclei in macromolecules, Γ_1 and Γ_2 are dominated by $J(\omega_H)$ (where $\omega_H/2\pi$ is the ^{1}H frequency) and $J(0)$, respectively. These characteristics make it possible to directly determine values of S^2 and $\langle r^{-6}\rangle^{-1/6}$ for PRE interaction vectors in a structure-independent manner from PRE Γ_1 and Γ_2 rates at multiple magnetic fields by minimization of the χ^2 function, defined as

(15) Bloembergen, N.; Morgan, L. O. *J. Chem. Phys.* **1961**, *34*, 842–850.
(16) Solomon, I. *Phys. Rev.* **1955**, *99*, 559–565.
(17) Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879–5896.
(18) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108–4139.

$$\chi^2 = \sum_{j}^{\text{fields}} \sum_{k}^{\text{nuclei}} \left\{ \frac{(\Gamma_{1,(j,k)}^{\text{calc}} - \Gamma_{1,(j,k)}^{\text{obs}})^2}{\sigma_{1,(j,k)}^2} + \frac{(\Gamma_{2,(j,k)}^{\text{calc}} - \Gamma_{2,(j,k)}^{\text{obs}})^2}{\sigma_{2,(j,k)}^2} \right\} \quad (14)$$

where the superscripts "calc" and "obs" denote calculated and observed values; σ represents the experimental error for each PRE rate; and the indices j and k indicate different magnetic fields and ^{1}H nuclei, respectively. For a system where the paramagnetic center (corresponding to Atom 1 in Figure 2a) is located in a structurally disordered group and the observed nuclei are located in a rigid portion of the macromolecule, the distribution radius R_P can be calculated from $\langle r^{-6}\rangle^{-1/6}$ and S^2 using eq 8. Simultaneously, eqs 4 and 7 yield the distance R, given by

$$R = \langle r^{-6}\rangle^{-1/6}(S^2)^{-1/6} \quad (15)$$

thereby providing paramagnetic center−proton distance restraints for structure calculations using a single average position for the paramagnetic center without the need to represent the mobile paramagnetic group by an ensemble of states.[17]

Practical Considerations. From eqs 11−14, it can be seen that there are up to four unknown parameters per paramagnetic center−proton vector (τ_c, τ_i, S^2, and $\langle r^{-6}\rangle$) but only two experimental observables. Although Γ_2 is field independent, Γ_1 varies with field, and hence, to ensure that the system is not underdetermined, it is essential to measure Γ_1 at several magnetic fields over the frequency range characteristic of high-resolution NMR (500−1000 MHz in ^{1}H). The ratio of Γ_1 rates at two magnetic fields is less than the square of the ratio of the magnetic fields; therefore, considering experimental uncertainties in the measurements, it is advisable to use as wide a range of magnetic fields as is practically feasible. In the current work, fields ranging from 500 to 800 MHz were employed.

We first consider the case where the paramagnetic label is located in a structurally disordered group and PREs are measured to protons in rigid regions of the proteins. The structurally disordered group can include not only the paramagnetic label itself but also neighboring residues if the label is attached to a residue located within a highly mobile region of the protein, such as a flexible loop. In the latter instance, the distribution radius of the paramagnetic center will reflect not only the intrinsic motion of the paramagnetic label but also the motion of the adjacent protein residues. (Note that mobile and rigid portions of the protein can readily be ascertained using conventional ^{15}N relaxation measurements.) To accurately determine the distribution radius of the paramagnetic center, Γ_1 and Γ_2 rates for as many paramagnetic center−proton vectors should be measured. The PRE data can then be fit using two approaches. In the first, the data are fit using eqs 11−14, with the overall correlation time, τ_c, and the correlation time for internal motions, τ_i, treated as global parameters, and S^2 and $\langle r^{-6}\rangle$ optimized for each PRE vector to yield corresponding values of R_P (eq 8) and R (eq 15). (A single value of τ_i is justified in this case since the dynamics of the PRE vectors are governed by the positional dynamics of the paramagnetic center.) From Figures 3 and 4b, it can be seen that the apparent value of R_P for a nonspherical distribution depends on the location of the observed proton. Consequently, it is easy to ascertain if the paramagnetic center is distributed in a highly nonspherical space, as the distribution of apparent R_P values would be much wider than a Gaussian distribution, with a width corresponding to the experimental uncertainties of R_P. An alternative fitting procedure involves treating τ_c, τ_i, and R_P as global parameters and varying only R for each PRE vector. In this instance, the spectral density function using the χ^2 function (eq 14) can be readily derived from eqs 4, 7, and 13, and is given by

$$J(\omega) = R^{-6}\left[\frac{\tau_c}{1 + (\omega\tau_c)^2} + \left\{\left(1 - \frac{R_P^2}{R^2}\right)^{-3} - 1\right\}\frac{\tau_t}{1 + (\omega\tau_t)^2}\right] \quad (16)$$

If the PRE data can be well fit using the second procedure, and the distances R obtained using the two alternate fitting procedures are the same within experimental uncertainty, one can conclude that the spatial distribution of the paramagnetic center is essentially spherical (i.e., the deviations from a spherical distribution are small).

Next, we consider the case where the paramagnetic center is rigid and PREs are measured to protons in a disordered region(s). This situation is more complex since each paramagnetic center−proton vector will be characterized by its own internal correlation time τ_i, and hence τ_i can no longer be treated as a global parameter. (That being said, it would not be unreasonable to assume that τ_i values for adjacent residues would be correlated; hence, in a mobile loop, for example, the variation of τ_i as a function of residue position in the loop would not be expected to be random.) Likewise, each proton will be characterized by its individual distribution radius R_P. For this case, τ_c would still be treated as a global parameter, and the values of τ_i, R, and R_P (using eq 16) or τ_i, S^2, and $\langle r^{-6}\rangle$ (using eq 13) would be optimized for each PRE vector. Given the number of unknowns, Γ_1 rates would have to be measured at a minimum of three magnetic fields. Since the motional amplitudes of immediately adjacent residues will be correlated, the values of both R and R_P would be expected to change in small increments from the backbone of one residue to the next in the disordered region. Thus, in a disordered loop, maximal amplitudes would be predicted for the backbone of residues in the center of the loop becoming increasingly smaller as the residue approaches the rigid portion of the protein. Similarly, within a given residue, increased motional amplitudes would be expected as one advances toward the end of the side chain. Deviations from this behavior would be expected only when the conformational space sampled becomes significantly nonspherical. To a first approximation, it seems likely that the assumption of a spherical distribution will hold in most cases, with the exception of motions involving a transition between two distinct conformational states (i.e., a two-state jump).

Experimental Section

Preparation of the SRY/DNA-EDTA-Mn^{2+}(Ca^{2+}) Complexes. U-[^{2}H/^{15}N]-labeled SRY was expressed and purified as described previously,[19] except that minimal medium containing D_2O was used for deuteration. The 14-bp DNA duplex containing dT-EDTA chelating either Mn^{2+} or Ca^{2+} was prepared as described previously.[17] The SRY/DNA-EDTA-Mn^{2+}(or Ca^{2+}) complexes were prepared by mixing the protein and DNA at a molar ratio of 1:1.2. The complexes were washed to remove excess Mn^{2+} (or Ca^{2+}) as described.[20,21] All buffers in this study were treated with Chelex-20 resin to remove all metal contamination. The NMR samples

(19) Murphy, E. C.; Zhurkin, V. B.; Louis, J. M.; Cornilescu, G.; Clore, G. M. *J. Mol. Biol.* **2001**, *312*, 481–499.
(20) Iwahara, J.; Clore, G. M. *Nature* **2006**, *440*, 1227–1230.
(21) Iwahara, J.; Anderson, D. E.; Murphy, E. C.; Clore, G. M. *J. Am. Chem. Soc.* **2003**, *125*, 6634–6635.

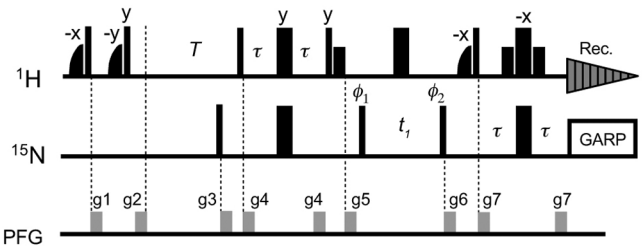

Figure 5. Pulse sequence for the saturation recovery experiment for measurement of PRE ^1H_N-Γ_1 rates for ^2H-labeled proteins. The initial two pairs of shaped and rectangular ^1H pulses along with pulse field gradients kill all ^1H_N magnetization, which recovers during the delay T and is then detected using a water-flip-back $^1H-^{15}N$ HSQC pulse scheme.[28] Two-dimensional spectra with different values of T are recorded in an interleaved manner. Phases are along x unless indicated otherwise. Thin and bold lines represent rectangular 90° and 180° pulses, respectively. Small bold ^1H pulses represent water-selective rectangular 90° pulses (1.4 ms). Half-bell shapes represent ^1H half-Gaussian pulses selective to water (2.1 ms). The delay τ is 2.3 ms. Phase cycling is as follows: $\phi_1 = [x,-x]$; $\phi_2 = [2x,2(-x)]$; receiver $= [x,2(-x),x]$. Quadrature detection for the t_1 dimension was carried out with phase incrementation of ϕ_1 in a States-TPPI manner.

comprised 450-μL solutions of 0.3 mM SRY/DNA complex chelating either Mn^{2+} or Ca^{2+}, 20 mM Tris·HCl (pH 6.8), 20 mM NaCl, and 7% D_2O, sealed into Norell 5-mm NMR tubes. The low concentration for the protein/DNA complexes was chosen to avoid undesired so-called "solvent PRE" effects that arise from random molecular collisions.[22]

NMR Experiments. NMR experiments were carried out using Bruker Avance spectrometers. Spectra were recorded at 308 K with cryogenic probes. PRE ^1H_N-Γ_2 rates for the SRY/DNA-EDTA-Mn^{2+} complex were measured with a two-time-point approach as described.[17,22] Errors in PRE Γ_2 rates were estimated as described previously.[22] For the ^1H_N-Γ_1 measurements, we used the two-dimensional $^1H-^{15}N$ correlation experiment with a saturation recovery period prior to the HSQC scheme described in the following section. The Ca^{2+}-chelated complex was used as the diamagnetic control for all PRE measurements. To identify ^1H nuclei that undergo rapid ($k_{ex} > 0.2\ s^{-1}$) hydrogen exchange with water molecules,[23] a 3D ^{15}N-edited ROESY spectrum[24] was measured with a ^1H spin-lock of 30 ms duration at a radiofrequency strength of 4 kHz. ^{15}N T_1 and T_2 relaxation times were measured for the SRY/DNA-EDTA-Ca^{2+} complex as described by Farrow et al.[25] The rotational correlation time τ_r and the anisotropy of the rotational diffusion tensor were determined from the ^{15}N T_1/T_2 ratios.[26]

Measurement of PRE ^1H_N-Γ_1 Rates. The measurement of PRE Γ_1 rates at multiple magnetic fields is particularly important to obtain values for the order parameters S^2 of the PRE interaction vectors. We measured the PRE Γ_1 rates for the SRY/DNA-EDTA-Mn^{2+} complex using the saturation recovery $^1H-^{15}N$ HSQC-based experiment shown in Figure 5. Although the inversion recovery method for ^1H longitudinal relaxation measurements generally offers higher precision than the saturation recovery method,[27] the saturation recovery method permits more rapid measurement of ^1H longitudinal relaxation rates, since no magnetization recovery is required before the initial ^1H pulse. For the present case involving a deuterated protein where ^1H_N-T_1 relaxation times are very long, this feature is essential to permit completion of the experiment within a reasonable time frame. Delays for the saturation recovery were 0.005, 0.20, 0.60, 0.005, 0.20, 0.60, 1.20, 2.00, 3.00, 4.00, and 5.00 s at 600 MHz; and 0.005, 0.20, 0.60, 1.20, 2.00, 3.00, 4.20, and 5.60 s at 800 MHz. A phenomenological longitudinal relaxation rate R_1 is obtained by fitting the signal intensities to $I(T) = I(\infty)\exp(-R_1T) + I(\infty)$. The PRE ^1H_N-Γ_1 rates were determined as $R_{1,para} - R_{1,dia}$, where $R_{1,para}$ and $R_{1,dia}$ are the longitudinal relaxation rates for paramagnetic and diamagnetic states, respectively. Errors in Γ_1 rates were estimated as $[(\sigma R_{1,para})^2 + (\sigma R_{1,dia})^2]^{1/2}$, where $\sigma R_{1,para}$ and $\sigma R_{1,dia}$ are the errors in the ^1H R_1 rates for the paramagnetic and diamagnetic states, respectively. Data for protons that exhibit non-single-exponential behavior due to rapid hydrogen exchange (or possibly due to $^1H-^1H$ cross relaxation) were excluded; consequently, the total number of Γ_1 data points used in the analysis is less than the number of Γ_2 data points.

χ^2 Minimization. Nonlinear least-squares fitting was carried out using the Levenberg−Marquardt algorithm, and errors in the calculated parameters were estimated using the Monte Carlo method.

Structure-Based Calculation of the Positional Distribution of EDTA-Mn^{2+} in the SRY/DNA-EDTA-Mn^{2+} Complex. The distribution of the EDTA-Mn^{2+} group was also examined by high-temperature simulated annealing calculations carried out using the Xplor-NIH molecular structure determination package.[29] The target function for simulated annealing comprised only terms for covalent geometry (bond, angles, impropers) and van der Waals interactions (in the form of a quartic van der Waals repulsion term). No pseudoenergy terms for experimental data were employed. Topology and parameters for the EDTA-Mn^{2+} group were set as described in the Supporting Information of ref 17. The conformation of the linker between the EDTA-Mn^{2+} group and the thymine base was randomized during the high-temperature stage (3000 K), followed by slow cooling to 25 K. In this calculation, only the dT-EDTA-Mn^{2+} group (from the linker onward) was allowed to move, while the other parts of the complex were kept fixed. A total of 400 different conformations of dT-EDTA-Mn^{2+} were computed.

Results and Discussion

SRY/DNA Complex and PRE Rates. As a model system to test the method described in the Theory section, we determined the distribution radius of the EDTA-Mn^{2+} paramagnetic group covalently attached to a thymine base[21,30] in the SRY/DNA complex (Figure 6a) from PRE rates without recourse to any structural information. The EDTA-Mn^{2+} group is covalently attached to a thymine base in this complex through a flexible linker consisting of five rotatable bonds (Figure 6a) and is therefore disordered with respect to the core of the protein/DNA complex. A previous study on the same SRY/DNA-EDTA-Mn^{2+} complex system showed that the PRE rates for this system are fully consistent with the structure of the complex.[17] Although it is known that low-population intermediates in macromolecular association can make PRE rates inconsistent with the structures of specific complexes,[20,31−33] such effects of intermediates are negligible for the SRY/DNA complex. Thus, this complex represents an ideal model system to test the new method.

(22) Iwahara, J.; Tang, C.; Clore, G. M. *J. Magn. Reson.* **2007**, *184*, 185–195.
(23) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 627–638.
(24) Clore, G. M.; Bax, A.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 5671–5676.
(25) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984–6003.
(26) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562–12566.
(27) Becker, E. D.; Ferretti, J. A.; Gupta, R. K.; Weiss, G. H. *J. Magn. Reson.* **1980**, *37*, 381–394.
(28) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593–12594.
(29) Schwieters, C. D.; Kuszewski, J. J.; Clore, G. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47–62.
(30) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968–972.
(31) Tang, C.; Iwahara, J.; Clore, G. M. *Nature* **2006**, *444*, 383–386.
(32) Volkov, A. N.; Worrall, J. A.; Holtzmann, E.; Ubbink, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18945–18950.
(33) Suh, J. Y.; Tang, C.; Clore, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 12954–12955.

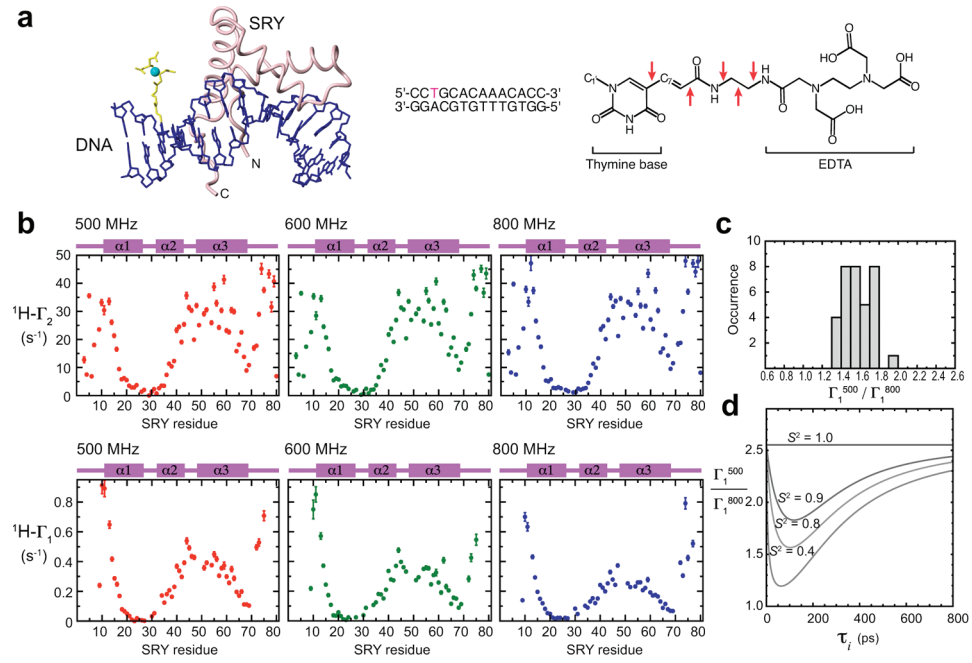

Figure 6. PRE measurements for the SRY/DNA complex at multiple fields. (a) Structure of the SRY/DNA-EDTA-Mn^{2+} complex.[17] The EDTA-Mn^{2+} group (yellow/cyan) is tethered to a thymine base (red in the sequence) in the 14-bp DNA duplex. The chemical structure of T-EDTA is also shown (asterisks, coordination sites; arrows, rotatable bonds in the linker). (b) PRE ^1H_N-Γ_1 and Γ_2 rates measured at three magnetic fields. (c) Histogram of the ratio of PRE ^1H_N-Γ_1 rates measured at 500 and 800 MHz. (d) Theoretical dependence of $\Gamma_1^{500}/\Gamma_1^{800}$ ratio on the order parameter S^2 and the internal motion correlation time τ_i for a Mn^{2+}-^1H vector.

PRE ^1H_N-Γ_1 and ^1H_N-Γ_2 rates were measured for U-[^2H/^{15}N]-labeled SRY bound specifically to a 14-bp DNA duplex conjugated to EDTA-Mn^{2+} (Figure 6b) at three high magnetic fields: 11.7, 14.0, and 18.7 T (^{1}H-frequencies, 500, 600, and 800 MHz, respectively). Deuterated SRY was employed to obtain higher precision for the PRE measurements and to minimize the effect of $^1H-^1H$ cross relaxation on the PRE Γ_1 rates.[34] Figure 6b show the measured ^1H_N-Γ_1 and ^1H_N-Γ_2 rates for SRY backbone amide groups in the complex. Since the PRE is proportional to $\langle r^{-6}\rangle$, residues that are close to Mn^{2+} exhibit large values of ^1H_N-Γ_1 and ^1H_N-Γ_2.

The PRE ^1H_N-Γ_2 rates display negligible field dependence. At the three different magnetic fields, most nuclei exhibit ^1H_N-Γ_2 values that differ only within a factor of 2 of the experimental errors. As described previously,[22] this result indicates that the contribution of Curie spin relaxation to Γ_2 is negligible and that the field dependence of τ_c is minimal in this magnetic field range. The field-independent nature of τ_c is reasonable because, for EDTA-Mn^{2+}, the electron relaxation due to collisional modulation of the zero-field-splitting (ZFS) tensor that can cause field dependence of τ_s^{-1} should occur with a rate of 10^4-10^5 s^{-1} at 11.7−18.7 T, judging from the results of a previous low-field NMR study.[35] Although the actual electron relaxation rate τ_s^{-1} can be substantially larger than this estimate due to the contribution from vibrational modulation of the Mn^{2+} ZFS tensor,[36] the electron relaxation due to this mechanism is essentially field-independent; therefore, the PRE correlation time τ_c, defined as $(\tau_r^{-1} + \tau_s^{-1})^{-1}$, should be virtually field-independent for the present case.

In contrast, the PRE ^1H_N-Γ_1 rates display significant field dependence. Figure 6c shows a histogram of Γ_1(500 MHz)/Γ_1(800 MHz) ratios for residues with $\Gamma_1 > 0.1$ s^{-1} at both fields: the average ratio is 1.70 ± 0.15. If PRE vectors are static with $S^2 = 1.0$, the Γ_1(500 MHz)/Γ_1(800 MHz) ratios for the present case (with $\omega_H\tau_c \gg 1$) should be 2.56, which is considerably larger than the observed values. Even if the contribution from collisional modulation of the ZFS tensor[35,37] is non-negligible in τ_s^{-1}, the field dependence of τ_s^{-1} should make the ratio larger rather than smaller. Dynamics of PRE vectors, however, on the pico- to nanosecond time scale with $S^2 < 1.0$ makes the ratio considerably smaller than 2.56 (Figure 6d). Thus, the observed Γ_1(500 MHz)/Γ_1(800 MHz) ratios clearly indicate that the PRE vectors are dynamic in this system.

Although it is known that the unpaired electron spin density delocalized to the $2p_z$ orbital can render the $\langle r^{-6}\rangle$ dependence of the PRE rate invalid for redox-active systems such as electron-transfer proteins,[38−40] eqs 11−13 were experimentally validated for the SRY/DNA-EDTA-Mn^{2+} complex in our previous study.[17] Thus, these equations were used for subsequent analysis.

Order Parameters and Effective Distances for PRE Vectors. The order parameters S^2 and effective distances $\langle r^{-6}\rangle^{-1/6}$ for the PRE vectors were determined via nonlinear least-squares fitting (cf. eq 14). For the calculation, we used a total of 132 PRE rates for 22 ^1H_N nuclei that are located in structurally rigid portions of the complex (as evidenced by ^{15}N relaxation) and satisfy the following four criteria: (i) $\Gamma_1 > 0.1$ s^{-1}; (ii) $\Gamma_2 > 15$ s^{-1}; (iii) slow hydrogen exchange with water with a rate $k_{ex}^{water} < \sim 0.2$ s^{-1}; and (iv) single-exponential recovery of magnetiza-

(34) Hernandez, G.; Teng, C. L.; Bryant, R. G.; LeMaster, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 4463–4472.
(35) Kruk, D.; Kowalewski, J. *J. Biol. Inorg. Chem.* **2003**, *8*, 512–518.
(36) Miller, J. C.; Sharp, R. R. *J. Phys. Chem. A* **2000**, *104*, 4889–4895.
(37) Rubinstein, M.; Baram, A.; Luz, Z. *Mol. Phys.* **1971**, *20*, 67–80.
(38) Gottlieb, H. P. W.; Barfield, M.; Doddrell, D. M. J. *J. Chem. Phys.* **1977**, *67*, 3785–3794.
(39) Mispelter, J.; Momenteau, M.; Lhoste, J.-M. In *NMR of Paramagnetic Molecules*; Berliner, L. J., Reuben, J., Eds.; Plenum Press: New York, 1993; Vol. 12, pp 299−355.
(40) Ma, L.; Jorgensen, A.-M. M.; Sorensen, G. O.; Ulstrup, J.; Led, J. J. *J. Am. Chem. Soc.* **2000**, *122*, 9473–9485.

Table 1. Models Used for Nonlinear Least-Squares Optimization[a] and Their Statistics

$J(\omega)$	global[b]	individual[c]	degrees of freedom	reduced χ^2 [d]	AIC[e]
fitting 1 (eq 13)	τ_c, τ_i	S^2, $\langle r^{-6}\rangle^{-1/6}$	86	2.7	326.6
fitting 2 (eq 16)	τ_c, τ_i, R_P	R	107	12.6	1401.2
fitting 3 (eq 17)	τ_c	$\langle r^{-6}\rangle^{-1/6}$	109	31.8	3511.4

[a] Minimization of the χ^2 function defined by eq 14. PRE Γ_1 and Γ_2 data measured for 22 different Mn^{2+}-^1H vectors at three magnetic fields were used. Total number of square terms is 132 (= 22 × 2 × 3). [b] Globally optimized parameter(s). [c] Parameter(s) individually optimized for each PRE vector. [d] χ^2 per degree of freedom. [e] Akaike's information criterion; AIC = χ^2 + 2k, where k is the number of optimized parameters.

tion in the saturation recovery experiments for Γ_1 measurements. The latter two conditions are important for accurate measurement of Γ_1 rates. Condition (iii) was verified using a 3D ^{15}N-separated ROE spectrum, in which H_2O-exchange cross peaks appear with opposite sign to ROE cross-peaks, thereby permitting permit identification of H_N atoms that undergo rapid hydrogen exchange.[24] Three fitting calculations were performed, and the results are summarized in Table 1.

First, using eqs 11−14, the order parameter S^2 and effective distance $\langle r^{-6}\rangle^{-1/6}$ were optimized for each PRE vector, while the PRE correlation time τ_c (constrained by the relationship $\tau_c \leq \tau_r$, which always holds from the definition of τ_c) and the internal motion correlation time τ_i were treated as global parameters. The final value of the reduced χ^2 (i.e., χ^2 per degree of freedom) is 2.5, indicating a reasonably good fit (Table 1, fitting procedure 1). This is also clear from the correlation plot of observed versus calculated Γ_1 (Figure 7a) and Γ_2 (Figure 7b) rates. The rotational correlation time τ_r determined from ^{15}N T_1 and T_2 relaxation times is 8.5 ns (with a diffusion anisotropy of 1.1). The values of τ_c and τ_i obtained from the PRE data are 7.2 ± 0.9 and 0.20 ± 0.02 ns, respectively. The values of S^2 range from 0.65 to 0.84, and the effective distances $\langle r^{-6}\rangle^{-1/6}$ range from 20.8 to 24.2 Å (Table 2). A histogram of

Table 2. Values of the Order Parameter S^2 and the Effective Distance $\langle r^{-6}\rangle^{-1/6}$ for Mn^{2+}-^1H_N Vectors Determined by Nonlinear Least-Squares Fitting, and Values of the Center Distances R and the Distribution Radii R_P Calculated from S^2 and $\langle r^{-6}\rangle^{-1/6}$ Using Fitting Procedure 1[a]

^1H_N[b]	S^2	$\langle r^{-6}\rangle^{-1/6}$ (Å)	R (Å)	R_P (Å)
I13	0.65 ± 0.04	20.6 ± 0.3	22.1 ± 0.5	8.0 ± 0.7
V14	0.68 ± 0.04	21.9 ± 0.3	23.3 ± 0.5	8.1 ± 0.7
W15	0.73 ± 0.04	22.9 ± 0.4	24.1 ± 0.6	7.6 ± 0.8
S16	0.74 ± 0.04	23.8 ± 0.4	25.0 ± 0.6	7.7 ± 0.8
Y41	0.73 ± 0.04	22.3 ± 0.3	23.5 ± 0.5	7.4 ± 0.7
W43	0.71 ± 0.04	21.9 ± 0.3	23.2 ± 0.5	7.7 ± 0.7
K44	0.71 ± 0.04	20.8 ± 0.3	22.0 ± 0.5	7.2 ± 0.7
M45	0.72 ± 0.04	21.3 ± 0.3	22.5 ± 0.5	7.3 ± 0.7
L46	0.75 ± 0.04	21.5 ± 0.3	22.5 ± 0.5	6.9 ± 0.7
W52	0.78 ± 0.03	21.6 ± 0.3	22.5 ± 0.5	6.4 ± 0.7
F54	0.73 ± 0.04	21.9 ± 0.3	23.1 ± 0.5	7.3 ± 0.7
F55	0.78 ± 0.04	21.0 ± 0.3	21.9 ± 0.5	6.1 ± 0.7
Q56	0.76 ± 0.04	21.3 ± 0.3	22.3 ± 0.5	6.7 ± 0.7
E57	0.76 ± 0.04	22.4 ± 0.4	23.5 ± 0.5	7.0 ± 0.7
A58	0.78 ± 0.03	21.6 ± 0.4	22.6 ± 0.5	6.4 ± 0.7
Q59	0.82 ± 0.03	20.8 ± 0.3	21.5 ± 0.5	5.5 ± 0.7
K60	0.77 ± 0.03	22.3 ± 0.4	23.3 ± 0.5	6.7 ± 0.7
L61	0.78 ± 0.04	22.8 ± 0.4	23.8 ± 0.6	6.6 ± 0.7
A63	0.83 ± 0.03	22.0 ± 0.4	22.7 ± 0.5	5.6 ± 0.7
M64	0.73 ± 0.04	24.1 ± 0.4	25.4 ± 0.6	8.0 ± 0.8
H65	0.84 ± 0.03	24.2 ± 0.4	24.9 ± 0.6	5.9 ± 0.8
R66	0.84 ± 0.03	23.3 ± 0.4	24.0 ± 0.6	5.8 ± 0.7

[a] Values were obtained via minimization of the χ^2 function defined by eq 14 together with eq 13. The distribution radii (R_P) and center distances (R) were calculated using eqs 8 and 15, respectively. The optimized values of the correlation times τ_c and τ_i are 7.2 ± 0.9 and 0.20 ± 0.02 ns, respectively. [b] Amide protons that satisfied the four criteria described in the main text. Numbering of SRY residues is according to Murphy et al.[19]

the distribution radius, R_P, for the EDTA-Mn^{2+} group, calculated from S^2 and $\langle r^{-6}\rangle^{-1/6}$ (using eq 8) for each Mn^{2+}-^1H vector, is shown in Figure 8a. The average value of R_P is 6.9 ± 0.8 Å. Considering that a typical backbone root-mean-square (rms) deviation in molecular dynamics simulations of proteins is 1−2 Å,[41] the error in R_P arising from neglecting R_N is estimated to

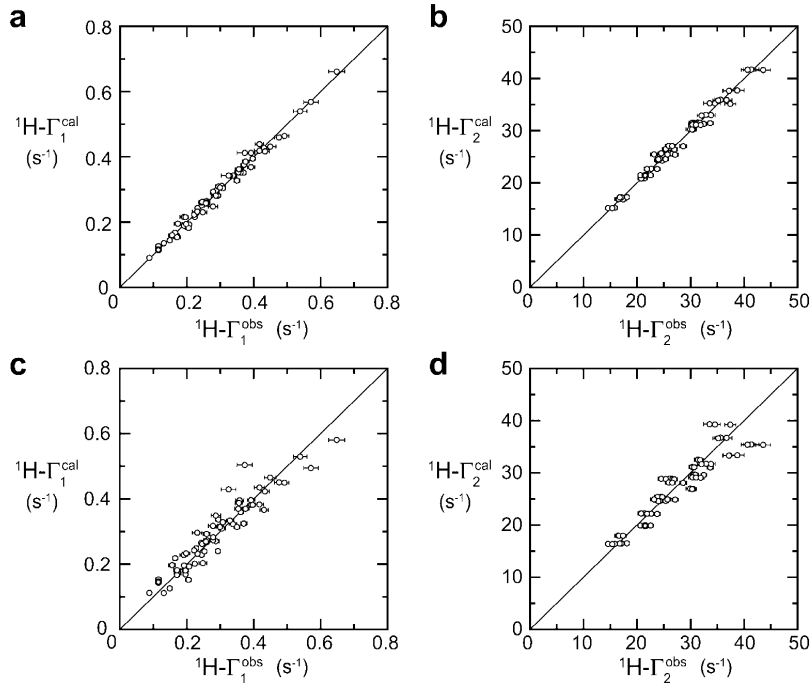

Figure 7. Correlation between calculated and observed PRE rates: (a,b) fitting procedure 1 and (c,d) fitting procedure 2. Statistics are given in Table 1.

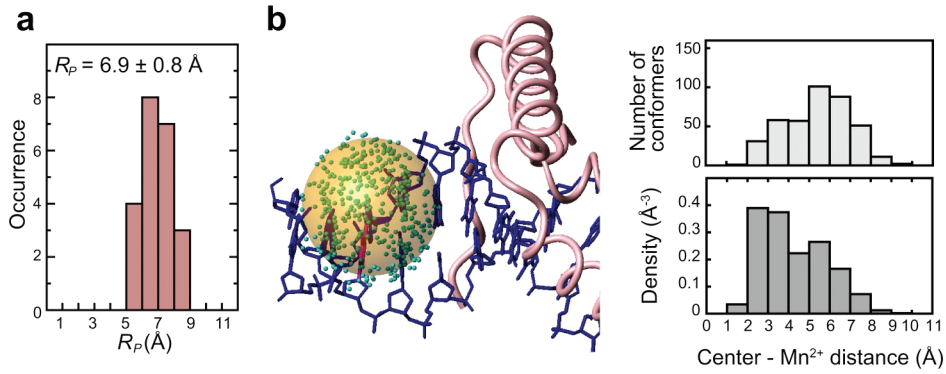

Figure 8. Distribution radius R_P of the EDTA-Mn^{2+} group in the SRY/DNA complex. (a) Histogram of R_P calculated from S^2 and $\langle r^{-6}\rangle^{-1/6}$ for 22 Mn^{2+}-^{1}H vectors in the SRY/DNA-EDTA-Mn^{2+} complex. (b) Mn^{2+} distribution obtained by structure-based simulated annealing calculations. In the left panel of (b), the positions of Mn^{2+} for 400 different conformers are shown in green. The transparent orange sphere, with its origin placed at the center of the distribution, is drawn with a radius of 6.9 Å, corresponding to the average value of R_P determined by the structure-independent method. In the right-hand panel of (b), a histogram of the distance from the center of the distribution to individual Mn^{2+} positions (top) and the density (bottom) of the Mn^{2+} for the 400 conformers (obtained by dividing the number of conformers by the volume of a shell with 1 Å thickness) is shown.

be less than 3% from Figure 2c, which is well below the experimental uncertainty.

We also carried out a second fitting procedure in which a single, global value of R_P was optimized in the nonlinear least-squares fit using eq 16. The resulting reduced χ^2 value is 12.8 (Table 1, fitting procedure 2), indicating that the quality of the fit is not quite as good as that obtained using fitting procedure 1 (eq 13). This is borne out from the correlation plot of observed versus calculated Γ_1 and Γ_2 rates (Figure 7c,d, respectively), although the correlation is still good, except for a couple of outliers in the Γ_1 correlation plot (Figure 7c). The value of Akaike's information criterion (AIC), however, for fitting procedure 1 is far lower than that for fitting procedure 2, implying that the positional distribution of the EDTA-Mn^{2+} is not perfectly symmetric (Table 1). The global value of R_P obtained using fitting procedure 2 is 7.4 ± 0.6 Å, which is still within one standard deviation of the mean value of 6.9 ± 0.8 Å for R_P obtained using fitting procedure 1.

The values of the center distances R obtained for fitting procedure 2 are listed in Table 3. The rms difference between the values of R obtained using fitting procedures 1 and 2 is only 0.7 Å, which represents a difference of less than 4% in the values obtained using the two fitting procedures. This indicates that robust and accurate values of the distances R from ^{1}H nuclei to the center average position of the paramagnetic center (in this instance Mn^{2+}) can be obtained using the formalism and procedures described here. This has important practical implications, since these distances can be used directly in structure calculations without the need to represent the mobile paramagnetic group by an ensemble of states and refine directly against the PRE data.

As a control, we also carried out a fit using the spectral density function,

$$J(\omega) = \langle r^{-6}\rangle \frac{\tau_c}{1 + \omega^2\tau_c^2} \qquad (17)$$

which corresponds to the case where $S^2 = 1$; that is, no internal motion for the paramagnetic label. The resulting reduced χ^2 has a value of 31.8 (Table 1, fitting procedure 3), indicating, as expected, that eq 17 is inappropriate in the present instance owing to dynamics of the Mn^{2+}-^{1}H vectors (i.e., $S^2 \neq 1$).

Table 3. Values of the Center Distances R Calculated Using Fitting Procedure 2 in Table 1[a]

^1H_N	R (Å)
I13	22.3 ± 0.5
V14	23.5 ± 0.5
W15	24.5 ± 0.5
S16	25.4 ± 0.6
Y41	24.0 ± 0.5
W43	23.5 ± 0.5
K44	22.6 ± 0.5
M45	23.0 ± 0.5
L46	23.2 ± 0.5
W52	23.4 ± 0.5
F54	23.6 ± 0.5
F55	22.9 ± 0.5
Q56	23.1 ± 0.5
E57	24.1 ± 0.5
A58	23.5 ± 0.5
Q59	22.7 ± 0.5
K60	24.1 ± 0.5
L61	24.6 ± 0.6
A63	23.8 ± 0.5
M64	25.8 ± 0.6
H65	25.8 ± 0.6
R66	25.0 ± 0.6

[a] Values were obtained via minimization of the χ^2 function defined by eq 14 together with eq 16 (fitting procedure 2 in Table 1). Values of the correlation times τ_c and τ_i were calculated to be 7.7 ± 0.8 and 0.21 ± 0.01 ns, respectively. The distribution radius R_P, defined as a global parameter, was calculated to be 7.4 ± 0.6 Å.

The value of R_P obtained from the structure-independent method is fully consistent with the distribution obtained from structure-based high-temperature simulated annealing calculations incorporating covalent geometry and van der Waals terms, allowing the EDTA-Mn^{2+} group to move while keeping the rest of the complex fixed (Figure 8b,c). This result suggests that the real distribution of the disordered EDTA-Mn^{2+} group in the molecular frame of the SRY/DNA complex is as wide as that predicted from structure-based computations. Equally importantly, the reasonable value of the distribution radius obtained for this model system provides evidence that the structure-independent method works.

Considerations for Obtaining Information about Distribution Radii of Disordered Groups from ^{1}H−^{1}H Cross-Relaxation Rates. Here we briefly consider how and under what circumstances one might obtain information about the distribution radii of

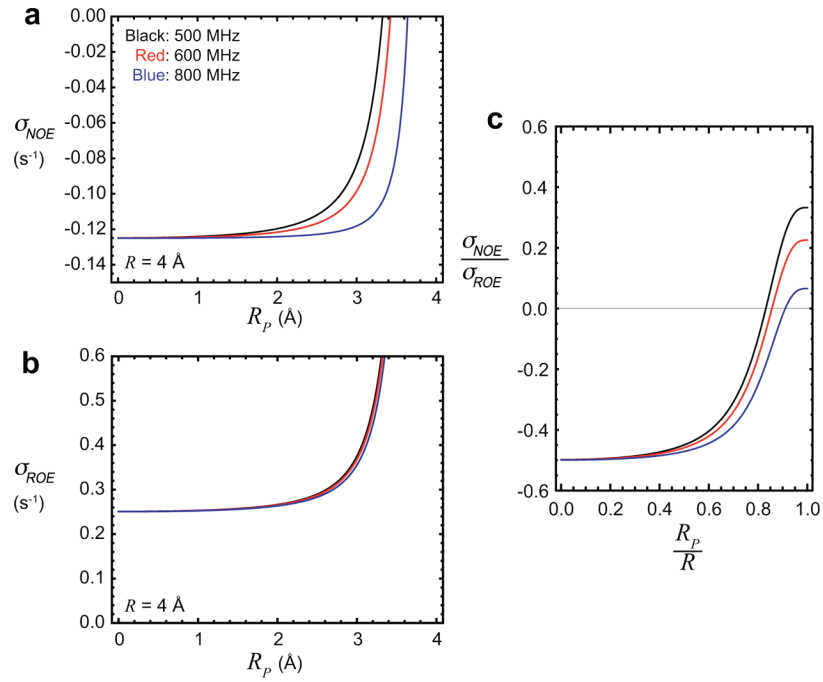

Figure 9. Effect of the distribution radius on ^{1}H−^{1}H cross-relaxation rates. (a) ^{1}H−^{1}H NOE and (b) ^{1}H−^{1}H ROE cross-relaxation rates as a function of the distribution radius R_P for the "Sphere" model with $R = 4$ Å and $R_N = 0$ Å. The curves were obtained using eqs 18−20, together with eqs 4 and 7. (c) Ratio of NOE to ROE cross-relaxation rates as a function of R_P/R. For all panels, calculations at ^{1}H frequencies of 500, 600, and 800 MHz (shown in black, red, and blue, respectively) were carried out with correlations times of $\tau_r = 9$ ns and $\tau_i = 0.2$ ns.

disordered groups in macromolecules from ^{1}H−^{1}H cross-relaxation rates using the formalism presented in this paper. The cross-relaxation rates for the ^{1}H−^{1}H NOE (σ_{NOE}) and ROE (σ_{ROE}) are given by[42]

$$\sigma_{NOE} = \frac{1}{40\pi^2}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H^4 h^2 \{-J_{HH}(0) + 6J_{HH}(2\omega_H)\} \quad (18)$$

$$\sigma_{ROE} = \frac{1}{40\pi^2}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H^4 h^2 \{2J_{HH}(0) + 3J_{HH}(\omega_H)\} \quad (19)$$

where the spectral density function $J_{HH}(\omega)$ is given by[10,11]

$$J_{HH}(\omega) = \langle r^{-6}\rangle \left\{ \frac{S^2\tau_r}{1 + \omega^2\tau_r^2} + \frac{(1 - S^2)\tau_e}{1 + \omega^2\tau_e^2} \right\} \quad (20)$$

and $\tau_e = (\tau_r^{-1} + \tau_i^{-1})^{-1}$. Note that eqs 13 and 20 differ in that the correlation times in eq 13 contain the electron correlation time τ_s. Using eqs 4 and 7, together with eqs 18−20, one can consider the relationship between the distribution radius R_P and cross-relaxation rates for the "Sphere" model.

For ^{1}H−^{1}H vectors between protons in relatively rigid portions of a macromolecule, both NOE and ROE cross-relaxation rates are dominated by the spectral density function $J_{HH}(0) \approx \langle r^{-6}\rangle S^2\tau_r$. Given that $R^{-6} = \langle r^{-6}\rangle S^2$ for the "Sphere" model (cf. eqs 4 and 7), ^{1}H−^{1}H cross-relaxation rates are independent of the distribution radius R_P under conditions where the $J_{HH}(0)$ term predominates. In this limit, it is impossible to determine R_P from ^{1}H−^{1}H cross-relaxation rates. From the perspective of structural analysis, the relationship $J_{HH}(0) \approx R^{-6}\tau_r$ in this limit is extremely useful because it means that the center distance R can be determined without any influence of dynamics.

For ^{1}H−^{1}H vectors, however, between a disordered group and a rigid portion of the macromolecule, the $J_{HH}(2\omega_H)$ term in eq 18 can be non-negligible, and the field dependence of NOE cross-relaxation can be quite significant. As shown in Figure 9, this occurs when R_P and R are comparable. In such a case, the ^{1}H−^{1}H cross-relaxation rates σ_{NOE} and σ_{ROE} are dependent on R_P and, therefore, permit one to derive information about the positional distribution of the disordered group. For a ^{1}H−^{1}H pair that exhibits a field-dependent σ_{NOE}, a χ^2 function containing square terms for σ_{NOE} and σ_{ROE} (analogous to eq 14) can be used to determine the order parameter S^2 and effective distance $\langle r^{-6}\rangle^{-1/6}$, and hence R_P. Since the overall rotational correlation time τ_r in eq 20 can readily be obtained from other data (e.g., ^{15}N T_1 and T_2 relaxation data), σ_{NOE} and σ_{ROE} rates at only two magnetic fields are sufficient to determine τ_i, S^2, and $\langle r^{-6}\rangle^{-1/6}$ (and R_P from the latter two). If the value of τ_i can also be obtained from other data (for example, for a ^{1}H−^{1}H vector between a disordered arginine guanido group and a rigid backbone amide proton, ^{15}Nε relaxation data can provide the correlation time for the internal motion of the ^{1}Hϵ-^{1}H$_N$ vector), σ_{NOE} and σ_{ROE} rates at a single magnetic field should be sufficient to determine R_P. For such a case, the $\sigma_{NOE}/\sigma_{ROE}$ ratio provides a very useful indicator (Figure 9c): if the ratio is significantly larger than −0.5, the distribution radius R_P can be determined using this approach.

Although spin-diffusion effects in both NOE and ROE experiments and undesirable TOCSY transfer in the ROE experiment may potentially hinder accurate measurement of ^{1}H−^{1}H cross-relaxation rates, these problems can readily be overcome for ^{1}H−^{1}H pairs between labile protons (e.g., backbone and side-chain N−H protons) by using perdeuterated proteins in which the replacement of nonexchangeable protons by deuterons removes the spin-diffusion and TOCSY transfer pathways. Indeed, accurate measurement of ^{1}H−^{1}H cross-

relaxation rates on a highly deuterated protein has recently been demonstrated.[43]

Concluding Remarks

Although the proof-of-principle application presented here relates to PRE vectors in which the paramagnetic center is located in a disordered group and the observed ^{1}H nuclei are in rigid portions of the molecule, the opposite situation, in which the paramagnetic center is at a fixed point and the ^{1}H nuclei are located in disordered groups, can also be analyzed to determine the distribution radius of disordered or highly mobile regions of a macromolecule. From this perspective, future applications to disordered groups on engineered proteins with an immobilized paramagnetic group[44−46] (M. R. Feissner and W. L. Hubbell, PDB code 1ZUR) or on metal-binding proteins are of interest. In addition, large-scale interdomain motions can also be analyzed using a highly mobile paramagnetic label, as in the present case, under conditions where the distribution radius of the paramagnetic label is much smaller than that of the interdomain motion. Such a situation would apply when two domains are connected by highly flexible linkers. The same is true with regard to the analysis of the distribution radius of a disordered loop containing the paramagnetic label. Finally, the formalism presented here and described by eqs 4 and 7 can be incorporated into equations for NOE/ROE cross-relaxation rates, and simulations predict that it is also possible, in principle, to obtain information about the distribution radii of disordered groups such as surface side chains from cross-relaxation rates for ^{1}H−^{1}H pairs between a disordered group and a rigid portion of the macromolecule.

In addition to its utility in probing distribution radii of disordered groups, the method presented here can also be used to provide paramagnetic center−proton distances for structure calculations. Previously, we had introduced a multiconformer representation of the paramagnetic group together with a theoretical framework and computational strategy to directly refine against PREs arising from a flexible paramagnetic group.[17] Since the method presented here provides the center distances R from the center of a spherical distribution of the paramagnetic center to the observed nuclei (cf. eq 15, and Tables 2 and 3), PRE-based structure calculations could be carried out using a single-conformer representation of the paramagnetic group together with a standard pseudopotential term for the distance restraints[29,47] and no special treatment for the paramagnetic group, providing the paramagnetic center is distributed approximately uniformly within a sphere. (For PREs involving methyl groups, the conventional "center-average" for NOE-derived interproton distance restraints would be used.[29,47]) In most circumstances, a square-well potential would be employed with a range corresponding to, e.g., the 5−95% confidence limits for the values of R determined from the fits to the Γ_1 and Γ_2 data (cf. eqs 11, 12, and 14 combined with eqs 13 and 15 or eq 16). This approach can be advantageous since it renders structure calculations simpler, faster, and more straightforward. In addition, the convergence properties of distance restraints in simulated annealing calculations are superior to those of PRE Γ_2 restraints. However, there are also some disadvantages: first, extraction of the center distances R requires a far larger number of data sets at multiple magnetic fields and is therefore time-consuming (in the present case, the total experimental time for the Γ_1 measurements at three magnetic field strengths was ∼15 days); second, measurement of accurate Γ_1 data is more difficult than Γ_2, and further Γ_1 is much more sensitive to internal motions, cross-correlation effects, and exchange with water than Γ_2;[17] third, the center distances R are strictly only valid for spherical distributions (anisotropic distributions could be handled by simply increasing the error bounds on the distance restraints, but in the case of bimodal distributions the estimates of R may contain significant errors).

In conclusion, we have demonstrated the NMR strategy and formalism required to obtain the distribution radii of disordered groups in macromolecules in a structure-independent manner using PRE data, which has been very difficult to achieve by any other biophysical experimental method. This new approach can provide new insights into interdomain dynamics and local disorder of proteins, which are important for understanding the role of dynamics in protein function. In addition, the same approach can be used to obtain distance restraints from the center of the distribution of a mobile paramagnetic group to protons of a macromolecule. Although the present application was demonstrated for the relatively small (∼18 kDa) SRY-DNA complex, recent work using methyl TROSY experiments to measure PREs for the 180 kDa proteoasome complex[48] indicates that the method should be applicable to systems larger than 100 kDa.

Acknowledgment. We thank Dr. Attila Szabo for stimulating discussions and for originally pointing out to us that $\langle r^{-6}\rangle \boldsymbol{S}^2 = R^{-6}$ to order R^{-10} for any spherical distribution centered on R. This work was supported by Grant H-1683 from the Welch Foundation (to J.I.), Grant MCB-0920238 from the National Science Foundation (to J.I.), the intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH (to G.M.C.), and the AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.).

Note Added after ASAP Publication. The production errors in the TOC graphic in the version published August 26, 2010, have been corrected in the version posted September 3, 2010.

JA1048187

(41) Fan, H.; Mark, A. E.; Zhu, J.; Honig, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6760–6764.
(42) Brüschweiler, R.; Case, D. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 27–58.
(43) Vögeli, B.; Segawa, T. F.; Leitz, D.; Sobol, A.; Choutko, A.; Trzesniak, D.; van Gunsteren, W.; Riek, R. *J. Am. Chem. Soc.* **2009**, *131*, 17215–17225.
(44) Keizers, P. H.; Desreux, J. F.; Overhand, M.; Ubbink, M. *J. Am. Chem. Soc.* **2007**, *129*, 9292–9293.
(45) Su, X. C.; Man, B.; Beeren, S.; Liang, H.; Simonsen, S.; Schmitz, C.; Huber, T.; Messerle, B. A.; Otting, G. *J. Am. Chem. Soc.* **2008**, *130*, 10486–10487.
(46) Su, X. C.; Liang, H.; Loscha, K. V.; Otting, G. *J. Am. Chem. Soc.* **2009**, *131*, 10352–10353.
(47) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65–73.
(48) Religa, T. L.; Sprangers, R.; Kay, L. E. *Science* , *328*, 98–102.