Quantitative Interpretation of Solvent Paramagnetic Relaxation for Probing Protein–Cosolute Interactions

Yusuke Okuno, Attila Szabo, and G. Marius Clore*

ABSTRACT: Protein–small cosolute molecule interactions are ubiquitous and known to modulate the solubility, stability, and function of many proteins. Characterization of such transient weak interactions at atomic resolution remains challenging. In this work, we develop a simple and practical NMR method for extracting both energetic and dynamic information on protein–cosolute interactions from solvent paramagnetic relaxation enhancement (sPRE) measurements. Our procedure is based on an approximate (non-Lorentzian) spectral density that behaves exactly at both high and low frequencies. This spectral density contains two parameters, one global related to the translational diffusion coefficient of the paramagnetic cosolute, and the other residue specific. These parameters can be readily determined from sPRE data, and then used to calculate analytically a concentration normalized equilibrium average of the interspin distance, \( \langle r^6 \rangle_{\text{norm}} \) and an effective correlation time, \( \tau_\text{c} \), that provide measures of the energetics and dynamics of the interaction at atomic resolution. We compare our approach with existing ones, and demonstrate the utility of our method using experimental \(^1\)H longitudinal and transverse sPRE data recorded on the protein ubiquitin in the presence of two different nitroxyde radical cosolutes, at multiple static magnetic fields. The approach for analyzing sPRE data outlined here provides a powerful tool for deepening our understanding of extremely weak protein–cosolute interactions.

INTRODUCTION

Protein–cosolute interactions play a critical role in protein stability,\(^1\)\(^-\)\(^5\) function,\(^6\)\(^-\)\(^7\) and molecular recognition.\(^8\)\(^-\)\(^7\) In particular, weak interactions between proteins and small cosolutes (such as urea, ATP, trimethylamine N-oxide, amino acids, etc.) are ubiquitous in biochemical processes within cells.\(^8\)\(^,\)\(^9\) For example, one class of small cosolutes known as osmolytes, such as sorbitals and betaines, tend to accumulate in kidney cells to counteract the denaturation effect of urea.\(^10\)\(^,\)\(^11\) Recently, osmolytes have also been shown to modulate protein aggregation processes involved in a number of neurodegenerative diseases.\(^12\)\(^-\)\(^15\)

Despite considerable scientific interest,\(^9\)\(^,\)\(^16\)\(^-\)\(^18\) the mechanistic details of protein–cosolute interactions are not well established at the molecular level, largely due to their extremely weak and transient nature that are invisible to most biophysical techniques. As a result, molecular details have largely been inferred from theoretical\(^19\)\(^-\)\(^24\) and computational\(^25\)\(^-\)\(^29\) studies.

NMR spectroscopy can potentially report on the energetics and time scale of protein–cosolute interactions.\(^30\)\(^-\)\(^36\) In particular, solvent paramagnetic relaxation enhancement (sPRE)\(^37\)\(^-\)\(^42\) can directly probe the interaction between a paramagnetic cosolute and a protein at atomic resolution. A new way of interpreting sPRE data is the subject of the current work.

Even for the simple analytically tractable model where the protein and paramagnetic cosolute are described as hard-spheres with no intermolecular forces between them, analysis and interpretation of the sPRE is difficult and of little practical utility.\(^43\)\(^,\)\(^44\) In the absence of intermolecular interactions other than excluded volume, the transverse sPRE relaxation time can be reasonably well described by computing ensemble average distances, \( \langle r^6 \rangle_{\text{exc}} \) (normalized for cosolute concentration) between the unpaired electron on the cosolute and protons on the protein with the distance of closest approach being limited by the protein molecular surface.\(^41\)\(^,\)\(^42\)\(^,\)\(^43\) The latter is easily calculated directly from atomic coordinates. However, agreement with experiment is poor when preferential protein–cosolute interactions are present\(^42\) (see Figure S1 of the
Supporting Information, SI), which is the main interest of this work.

Here, we propose a simple, but more accurate alternative, based on an approximate non-Lorentzian spectral density function that can be used to fit spectral densities obtained from experimental longitudinal \( \Gamma_1 \) and transverse \( \Gamma_2 \) sPRE rates.\(^\text{36}\) This approach enables one to extract two residue-specific quantities from the experimental sPRE data that characterize paramagnetic cosolute–protein interactions at the atomic level: (i) the cosolute concentration normalized, equilibrium average protein–cosolute interspin distance, \( \langle r^{--6}_\text{norm} \rangle \) and (ii) the effective correlation time \( \tau_c \) which is a measure of the time-scale of the protein–cosolute interaction. \( \langle r^{--6}_\text{norm} \rangle \) is independent of dynamics and is very sensitive to local intermolecular interactions. In particular, the value of \( \langle r^{--6}_\text{norm} \rangle \) will increase in the presence of an attractive potential and decrease in the presence of a repulsive potential. Both \( \langle r^{--6}_\text{norm} \rangle \) and \( \tau_c \) are meaningful even when well-defined binding states are not formed, as in the case of weak protein–cosolute interactions. Further, comparison of \( \langle r^{--6}_\text{exc} \rangle \) values calculated directly from a protein structure,\(^\text{41,42,43}\) with the experimentally determined \( \langle r^{--6}_\text{norm} \rangle \) values enables one to ascertain at the atomic level whether the protein–cosolute interactions are attractive or repulsive.

## THEORY

Relaxation due to direct dipole–dipole interaction between a proton spin on the protein and the electron spin on the paramagnetic cosolute in an isotropic liquid is described by the time-correlation function \( C(t) \) given by:\(^\text{46}\)

\[
C(t) = 4\pi N_\perp \left[ \frac{Y_{2,0}(\Omega_\perp(0))}{r^1(0)} \frac{Y_{2,0}(\Omega_\perp(t))}{r^3(t)} \right]
\]

where \( r \) and \( \Omega_\perp \) denote the length and orientation of a vector, \( \bar{r} \), between the protein and the cosolute spins in the laboratory frame (see Figure 1). \( Y_{2,0}(\Omega_\perp) \) is the spherical harmonic, \( Y_{l,m}(\Omega_\perp) \), with \( l = 2 \) and \( m = 0 \); \( N_\perp \) is the number of cosolute spins; and the brackets \( \langle \rangle \) denote an equilibrium average of all positions and orientations of the protein–cosolute pair.

Equation 1 ignores fluctuations of dipole–dipole interactions arising from electron spin relaxation.

We define the spectral density as the cosine transform of \( C(t) \):

\[
J(\omega) = \int_0^\infty C(t) \cos(\omega t) \, dt
\]

The difference in the longitudinal and transverse relaxation rates of the protein spin in the presence and absence of the paramagnetic cosolute, denoted as \( \Gamma_1 \) and \( \Gamma_2 \),\(^\text{36}\) respectively, are related to \( J(\omega) \) by:\(^\text{46}\)

\[
\Gamma_1 = \frac{3}{10} \left( \frac{\mu_0}{4\pi} \right)^2 \hbar^2 \gamma_H^2 \gamma_e^2 \langle J(\omega \perp) \rangle
\]

and

\[
\Gamma_2 = \frac{1}{5} \left( \frac{\mu_0}{4\pi} \right)^2 \hbar^2 \gamma_H^2 \gamma_e^2 \langle J(0) + \frac{3}{4} J(\omega \perp) \rangle
\]

where \( \mu_0 \) is the vacuum permeability constant divided by \( 2\pi \); \( \hbar \) is Planck’s constant; \( \gamma_H \) and \( \gamma_e \) are the gyromagnetic ratios of the \( ^1\text{H} \) spin of interest in the protein and the electron spin of the cosolute; and \( \omega \perp /2\pi \) is the Larmor frequency of the proton spins at given spectrometer field. It should be pointed out that \( \langle \rangle \) and \( J(\omega \perp) \) neglect terms involving \( \omega \perp \) as the latter is ~660 times larger than \( \omega \perp \) and hence terms containing \( \langle \omega \perp \rangle \) and \( J(\omega \perp \pm \omega_e) \) are negligibly small compared to those with \( J(0) \) or \( J(\omega_e) \). If \( \Gamma_1 \) is measured at \( n \) magnetic fields and \( \Gamma_2 \) at one of them, then \( n + 1 \) spectral densities can be determined: \( J(0), J(\omega_e), \ldots, J(\omega_e) \).

In isotropic solution, the correlation function is identical between the protein and cosolute and individual protons in the protein.

For comparison of \( \langle r^{--6}_\text{exc} \rangle \) values calculated directly from a protein structure,\(^\text{41,42,43}\) with the experimentally determined \( \langle r^{--6}_\text{norm} \rangle \) values enables one to ascertain whether the protein–cosolute interactions are attractive or repulsive.

\[ C(t) = N_\perp \left[ \frac{P_2(\hat{r}(0) \cdot \hat{r}(t))}{r^3(0) r^3(t)} \right] \]

where \( P_2(x) \) is the second order Legendre polynomial, and \( \hat{r} \) is a unit vector between the protein and cosolute spins. This expression for \( C(t) \) depends on the dot product of \( \hat{r}(0) \) and \( \hat{r}(t) \), and is therefore independent of the reference frame.

From eq 5 it follows that the initial value of the correlation function is given, in general, by:

\[
C(0) = N_\perp \langle r^{--6}_\text{exc} \rangle = 4\pi n_\perp \int_0^\infty g(r) \frac{dr}{r^4}
\]

where \( g(r) \) is the pair correlation function: \( g(r) / V \) \( dr \) is the probability of finding the protein and cosolute spins at a distance between \( r \) and \( r + \, dr \) in volume \( V \); and \( n_\perp \) is the number density (concentration) of the paramagnetic cosolute (i.e., \( N_\perp / V \)). In general, \( g(r) \) is related to the potential of the mean force \( U(r) \) by \( g(r) = \exp(-U(r)/k_BT) \) where \( k_B \) is the Boltzmann constant, and \( T \) the temperature. It should be emphasized that eq 6 is valid for arbitrary shapes of the protein and cosolute and for arbitrary intermolecular interactions between them.

It is convenient to define a concentration independent equilibrium average:

\[
\langle r^{--6}_\text{norm} \rangle = \frac{C(0)}{n_\perp} = 4\pi \int_0^\infty \frac{g(r) \, dr}{r^4}
\]

which has units of \( m^{-3} \) instead of \( m^{-6} \). This average is determined not only by the excluded volume interactions between the protein and cosolute but also by additional site-specific intermolecular forces due to electrostatic and hydrophobic interactions. Let us denote the values of this average calculated considering the excluded volume of the protein and cosolute by \( \langle r^{--6}_\text{exc} \rangle \), which, in practice, can be calculated from the protein coordinates.\(^\text{41,42,43}\) Clearly when \( \langle r^{--6}_\text{norm} \rangle \) is larger than \( \langle r^{--6}_\text{exc} \rangle \) local attractive interactions are present, whereas when \( \langle r^{--6}_\text{norm} \rangle \) is smaller than \( \langle r^{--6}_\text{exc} \rangle \) there are local repulsive interactions. Consequently \( \langle r^{--6}_\text{norm} \rangle \) can be used to identify the types of interactions (attractive, no force or repulsive) between the cosolute and individual protons in the protein.
In principle, \( \langle r^{-6}_\text{norm}\rangle \) could be found\(^{46}\) if \( J(\omega) \) were known at all frequencies using the inverse cosine transform of eq 2 at \( t = 0 \),

\[
\langle r^{-6}_\text{norm}\rangle = \frac{2}{n_0^4} \int_0^\infty J(\omega) \, d\omega
\]

(8)

In practice, \( J(\omega) \) can only be determined at a small number of frequencies owing to the limited availability of NMR spectrometer fields.

The simplest approximate spectral density, \( J_{\text{approx}}(\omega) \), is a Lorentzian:

\[
J_{\text{approx}}(\omega) = \frac{J(0)}{1 + (\omega \tau_C)^2}
\]

(9)

If \( \Gamma_1 \) and \( \Gamma_2 \) (and hence \( J(0) \) and \( J(\omega) \)) were measured at a single magnetic field, \( \tau_C \) could be found uniquely from eq 9,

\[
\tau_C = \frac{\sqrt{J(0) - J(\omega)}}{\omega J(\omega)} \approx \frac{J(0)}{\omega^2 J(\omega)}
\]

(10)

and from eq 8 it follows that,

\[
\langle r^{-6}_\text{norm}\rangle = \frac{J(0)}{n_0^4 \tau_C}
\]

(11)

If data at more fields are available, then \( \tau_C \) can be found by least-squares fitting. If only \( \Gamma_2 \) at a single magnetic field is determined for a large number of residues, then the corresponding \( \langle r^{-6}_\text{norm}\rangle \) values can be obtained from eq 11 assuming that \( \Gamma_2 \propto J(0) \) (i.e., \( J(\omega) \ll J(0) \)) and \( \tau_C \) is same for all residues. The latter is not generally valid (see Figure S2). The purpose of this work is to do better without significantly complicating the analysis.

For the sPRE, the Lorentzian spectral density is flawed because in the low frequency limit, as a result of translational diffusion, the spectral density behaves as:

\[
\lim_{\omega \to 0} J(\omega) = J(0) - B \sqrt{\omega} + \ldots
\]

(12)

where \( B \) is a constant. The Lorentzian spectral density, given by eq 9, however, approaches \( J(0) \) as \( \omega \to 0 \). However, the exact \( J(\omega) \) behaves like a Lorentzian in the high frequency limit,

\[
\lim_{\omega \to \infty} J(\omega) = A \omega^{-2} + \ldots
\]

(13)

where \( A \) is a constant.

According to Sholl\(^{49}\) and Fries,\(^{50}\) the constant \( B \) in eq 12 is independent of intermolecular interactions and is given by:

\[
B = \frac{2\sqrt{2} \pi n_0}{9 D_{\text{trans}}^{1/2}}
\]

(14)

where \( D_{\text{trans}} \) is the relative translational diffusion constant of the interacting molecules.

Our approach to the analysis of sPRE data is based on the following ansatz for the spectral density,

\[
J_{\text{approx}}(\omega) = \frac{J(0)}{1 + a \omega + b \sqrt{\omega}^2}
\]

(15)

when \( a \) and \( b \) are set to:

\[
a = \frac{J(0)}{A}
\]

(16)

and,

\[
\tau_C = \frac{\int_0^\infty C(t) \, dt}{\int_0^\infty C(0)} \approx \frac{J(0)}{n_0^4 \langle r^{-6}_\text{norm}\rangle}
\]

(19)

\( \tau_C \) is a measure of the time scale of the fluctuations of the interspin vector \( \mathbf{r} \) which arise from both translational and rotational diffusion and are influenced by both short and long-range intermolecular interactions. The quantitative interpretation of the physical meaning of \( \tau_C \) is model dependent (see SI Figure S3).

## RESULTS AND DISCUSSION

### Validation of the Approximate Spectral Density Function Equation 15 for the Force-Free Model

The simplest theoretical analysis of intermolecular dipolar relaxation, developed by Hwang\(^{44}\) and Ayant\(^{44}\) is based on a “force-free” model which assumes the protein and cosolute are hard-spheres with no intermolecular forces (other than excluded volume) between them. The cosolute and protein spins can be located anywhere in spheres that undergo rotational diffusion independently of one another. The spectral density function for this simple hard-sphere model is given analytically by:

\[
J(\omega) = \frac{\pi n_0}{6 D_{\text{trans}}^2} \sum_{i=0}^m \sum_{j=0}^l \frac{1}{(2k+1)(2l+1)} \left( \frac{a}{d} \right)^{2(l+k)} \left( \mathbf{p}_i \cdot \mathbf{p}_j \right) \text{Re}[j_{k+1} (\epsilon_0)]
\]

(20)

where

\[
j_k(x) = \left( \frac{1}{x^2} \right)^{l+1} \left( \frac{1}{2l+1} - \frac{1}{x^2} \right)^{l+1} \left( 1 + \frac{1}{x^2} \right)^{-1} \left( 1 + \frac{1}{x^2} \right)^{-1} \left( 1 + \frac{1}{x^2} \right)^{-1}
\]

(21)

\[
\tau = d^2 / D_{\text{trans}}
\]

(22)

\[
\frac{1}{\tau_{Q1}} = \tau_{\text{Rot},Q1}
\]

(23)
Figure 2. Validation of the approximate spectral density function, eq 15, for hard-sphere protein−cosolute interactions. (A) and (B) comparison of the exact force-free hard-sphere spectral densities (continuous gray lines) calculated using eq 20 with those calculated using either our approximate (eq 15, panel A) or single Lorentzian (eq 9, panel B) spectral density functions (shown as dashed red lines) at several distances of the 1H spin to the surface of an idealized protein represented by a sphere. The parameters used to generate the exact spectral densities are \( \tau_P = 3.8 \) ns and \( \tau_S = 39 \) ps (corresponding to spheres of radii \( \sim 17 \) Å and \( \sim 3.5 \) Å, respectively). The 1H spin is located within the protein sphere at distances of 1, 2, 3, 4, and 5 Å from the protein surface. The electron spin is located at 1 Å from the surface of the cosolute. The approximate spectral densities obtained using eq 15 superimpose almost exactly on the exact spectral densities calculated using eq 20. The blue circles represent the points from the exact spectral densities used to fit the approximate spectral density functions. (C) and (D) complete cross-validation showing the correlation between experimental and predicted \( \Gamma_1 \) values obtained using eqs 15 and 9, respectively, for a range of parameters: \( \tau_P = 2, 5, 10, \) and 15 ns; \( \tau_S = 5, 25, 50, 100, \) and 250 ps; and the 1H spin located at 2 (red), 4 (green), and 6 (blue) Å from the protein surface. The electron spin is located at the center of the cosolute sphere.

\[
\left\{ \sum_i \left( \Gamma_1^{\text{obs,free}}(i) - \Gamma_1^{\text{calc,free}}(i) \right)^2 / \sum_i \Gamma_1^{\text{obs,free}}(i)^2 \right\}^{1/2}
\]

where the summation with index \( i \) runs over all data points \( i \) at a given field. (E) and (F) ratios of approximate to exact \( \langle r^{-6} \rangle_\text{norm} \) values obtained using eqs 15 and 9, respectively, for 2- and 3-field fits. (G) and (H) correlation plots between exact and approximate \( \langle r^{-6} \rangle_\text{norm} \) values obtained using eqs 15 and 9, respectively, for 2- and 3-field fits.

Figure 3. Experimental backbone amide proton sPRE measured for 0.5 mM 2H/15N-labeled ubiquitin in the presence of 25 mM (A) 3-carboxy-PROXYL or (B) 3-carbamoyl-PROXYL paramagnetic cosolutes. The chemical structures of the paramagnetic cosolute are shown at the top. The middle and bottom panels display the experimental 1HN-\( \Gamma_2 \) and 1HN-\( \Gamma_1 \) profiles, respectively, at 500 (blue), 800 (green), and 900 (red) MHz spectrometer fields. (C) and (D) complete cross-validation showing the correlation between experimental and predicted 1HN-\( \Gamma_1 \) values at 500, 800, and 900 MHz (free data sets) for 3-carboxyl-PROXYL and 3-carbamoyl-PROXYL, respectively, when eq 15 is used to fit the working data sets at the other two fields (800 and 900 MHz, 500 and 900 MHz, and 800 and 900 MHz, respectively).
Large standard deviations in PDB fi, i.e., exact force-free spectra density eq 20 for a range of physically reasonable values of 15, we analyzed a large set of simulated data points using the times of the protein (respectively; are the center-to-spin distances of cosolute and protein, Stoke sample are presented below. For each simulation, we used the calculated directly from the coordinates of ubiquitin (PDB 1D3Z),66 taking into account only the excluded volume with no intermolecular forces between protein and cosolute, are displayed as continuous red lines. The associated with a and 8285

\[ \sigma_K(x) = \frac{\omega \tau}{\tau_{P,k} + \tau_{S,l-k}} \] (24)

\[ \frac{1}{\tau_Q} = 6D_{\text{Rot},Q} \] (25)

where Q is P or S.

To demonstrate the validity of our spectral density given by eq 15, we analyzed a large set of simulated data points using the exact force-free spectra density eq 20 for a range of physically reasonable values of \( \tau_p \) and \( \tau_s \) and the results for a representative sample are presented below. For each simulation, we used the Stoke—Einstein relation,

\[ D_{\text{trans}} = \frac{k_B T}{6 \pi \eta R_p} + \frac{k_B T}{6 \pi \eta R_s} \] (26)

and

\[ D_{\text{Rot},Q} = \frac{k_B T}{8 \pi \eta R_Q} \] (27)

where \( k_B \) is the Boltzmann constant; \( \eta \) is the viscosity of the solvent; and \( R_p \) and \( R_s \) are the hard sphere radii of the protein and cosolute, respectively. We set \( T = 298 \) K and \( \eta = 8.9 \times 10^{-2} \) J s m\(^{-3}\) (equal to the viscosity of water at 298 K). For each value of \( \tau_p \) and \( \tau_s \) \( R_p \) and \( R_s \) were back calculated from eqs 25 and 27 and \( d \) was set to \( R_p + R_s \). Simulated values of \( f(\omega) \) at \( \omega = 0 \), 2\( \pi \) (500 MHz), 2\( \pi \) (800 MHz), and 2\( \pi \) (900 MHz) were generated and fitted to eq 15 optimizing the values of \( a \) and \( D_{\text{trans}} \). Comparisons between representative exact spectral densities obtained from eq 20 with \( \tau_p = 3.8 \) ns and \( \tau_s = 39 \) ps and the fits using the approximate spectral density of eq 15 and the single Lorentzian spectral density of eq 9, are shown in Figure 2A and B, respectively. Equation 15 reproduces the exact spectral density essentially perfectly, whereas clear systematic differences are observed for eq 9. The optimized value of \( D_{\text{trans}} \) obtained from the fit is 7.3 (±0.1) \( \times 10^{-10} \) m\(^2\) s\(^{-1}\) which is close to the exact value of 8.5 \( \times 10^{-10} \) m\(^2\) s\(^{-1}\). The optimized value of \( D_{\text{trans}} \) is slightly underestimated probably due to insufficient number of data points in the low-frequency regime where \( f(\omega) \) is most sensitive to \( D_{\text{trans}} \) (see eq 12). Complete cross-validation in which fitting the \( \Gamma_1 \) values at two fields is used to predict the \( \Gamma_1 \) values at the third field for all three combinations of fields reveal excellent agreement for eq 15 (Figure 2C), but poor agreement for eq 9 (Figure 2D).

Comparison between the exact \( \langle r^n \rangle_{\text{norm}} \) and the approximate \( \langle r^n \rangle_{\text{approx}} \) obtained with the approximate spectral density functions using both a 3-field fit and three different combinations of 2-field fits, shows excellent agreement for eq 15 (Figure 2E), but poor agreement with systematic deviations for eq 9 (Figure 2F). In addition, the correlation between \( \langle r^n \rangle_{\text{norm}} \)
and $\langle r^{-6}_{\text{norm}} \rangle_{\text{approx}}$ for a variety of different $\tau_P$ and $\tau_S$ values is excellent for eq 15 (Figure 2G) but exhibits clear systematic deviations for eq 9 (Figure 2H).

The impact of 1-field and 2-field fitting on the accuracy of the extracted $\langle r^{-6}_{\text{approx}} \rangle_{\text{norm}}$ values is investigated in Figures S4 and S5. In the case of eq 15 there is essentially no difference between 3-field and 2-field fitting, and significant deviations between $\langle r^{-6}_{\text{approx}} \rangle_{\text{norm}}$ and $\langle r^{-6}_{\text{exact}} \rangle_{\text{norm}}$ values only become apparent for 1-field fitting when the $^1H$ spin is at 1 Å from the protein surface. In contrast, eq 9 performs poorly.

The approximate spectral density given by eq 15 is not unique. In the SI we examine two other spectral density functions: eq S6 omits various cross-terms in the denominator of eq. 15; and eq.S7, 55,56 is equivalent to the force-free hard-sphere spectral density given by eq 20 with $l = 0$. However, except in the case of the force-free hard-sphere model with both spins at the center, eq S7 does not have the flexibility to reproduce the amplitude of both the low and high frequency limits given by eqs 12 and 13. For the force-free model, $\langle r^{-6}_{\text{norm}} \rangle_{\text{approx}}$ calculated from eq 15 is more accurate than that obtained from either eqs S6 or S7 (see Figure S6).

As we noted above, eqs 3 and 4 for $\Gamma_1$ and $\Gamma_2$ neglect the contribution from dipole moment fluctuations arising from electron spin relaxation (see eqs S9 and S10). This is valid for nitroxide radicals where the electron relaxation time $T_1 e$ ($\sim 10^{-7}$ s)56−58 is much longer than $\tau_P$, the rotational correlation time of the protein; for cosolutes with metal ions, however, $T_1 e \leq \tau_P$,42,59 and, consequently, electron spin relaxation has to be taken into account in the analysis of intramolecular protein $\Gamma_1$ and $\Gamma_2$ data (i.e., where the PREs originate from a spin label covalently attached to the protein).36,60 What, however, is the impact of $T_1 e$ on sPRE data? We therefore carried out a series of calculations to test how well eq 15 reproduces the exact hard-sphere force-free spectral density when electron relaxation is taken into account by substituting eq S12 instead of eqs 24 into 20.61 The results presented in Figure S7 for $T_1 e$ values ranging from 0.1 to 100 ns, and $\tau_P$ values of 3.8 and 38 ns, corresponding to proteins of about 8 and 80 kDa, respectively, indicate that eq 15 is applicable when $T_1 e \geq 1$ ns, and the $^1H$ spin is within 5 Å from the protein surface.

Figure 5. Mapping sites of preferential protein-cosolute interactions on the surface of ubiquitin. $\langle r^{-6}_{\text{norm}} \rangle_{\text{norm}} - \langle r^{-6}_{\text{norm}} \rangle_{\text{exc}}$ profiles obtained from backbone amide proton sPRE measurements on $^2H/^{15}N$ labeled ubiquitin in the presence of 25 mM (A) 3-carboxy-PROXYL and (B) 3-carbamoyl-PROXYL radicals. The dashed line is drawn at $3 \times 10^{-28}$ m$^{-3}$ and the points corresponding to residues showing significant preferential interactions are colored in purple. Mapping of $\langle r^{-6}_{\text{norm}} \rangle_{\text{norm}} - \langle r^{-6}_{\text{norm}} \rangle_{\text{exc}}$ on the surface of ubiquitin color graded from purple ($\sim 3 \times 10^{-28}$ m$^{-3}$) to white for (C) 3-carboxy-PROXYL and (D) 3-carbamoyl-PROXYL. Two views related by a 180° rotation are shown; the top panels display the molecular surface, and the bottom panels, the corresponding ribbons diagrams with the side chains of residues showing significant preferential interactions in purple. For comparison, the (E) hydrophobicity scale51 (color coded from green to white) and (F) electrostatic potential (graded from blue, positive; white, neutral; and red, negative) are displayed on the molecular surface of ubiquitin. Hydrophobicity and electrostatic potential were calculated using the PyMol Molecular Graphics System (version 2.0 Schrödinger, LLC).
surface (such that $T_{1e}$ is equal to or greater than the correlation time given by eq 19).

**Validation of the Approximate Spectral Density Function Equation 15 Using Experimental NMR Relaxation Dispersion Measurements on Solvents in the Presence of TEMPOL.** To assess the validity of eq 15 for the presence of TEMPOL.

$\Gamma$ is the number of observed $1H$ spins: $1H$ spin specific $\Gamma$ values for $a$ and a single $D_{\text{trans}}$ that determines the values of $b$ using eq 17. The values of $a$ and $b$ determined from the fits are provided in Table S4.

Cross-validation was carried out by predicting the $\Gamma_1$ values at one field from eq 15 using the fits to the data from the two other fields (for all three combinations of fields). The initial values of $\Gamma_1$ were set according to eq 17 with $D_{\text{trans}}$ assumed to be equal to that of the translational diffusion constant of the cosolutes. Excellent agreement between predicted and experimental $\Gamma_1$ values at all fields are obtained for both 3-carboxy-PROXYL and 3-carbamoyl-PROXYL radicals (Figure 3C and D). The optimized $D_{\text{trans}}$ determined from the fits are $7.8(\pm 2.7) \times 10^{-10} m^2 s^{-1}$ and $8.0 (\pm 1.0) \times 10^{-10} m^2 s^{-1}$, respectively, which are close to the translational diffusion constant of the Di-$t$-butyl-nitroxide radical ($D_{\text{trans}} = 7.2 \times 10^{-10} m^2 s^{-1}$).

The experimentally derived $(r^{-6})_{\text{norm}}$ and $T_2$ profiles as a function of residue are shown in Figure 4. For comparison, a plot of $(r^{-6})_{\text{norm}}$ calculated directly from the coordinates of ubiquitin (PDB code 1D3Z) using the PSolPot module in Xplor-NIH that takes into account only excluded volume with no intermolecular forces, is also included. For both radicals there are several regions where the experimentally derived $(r^{-6})_{\text{norm}}$ values are significantly higher than the $(r^{-6})_{\text{norm}}$ values, indicative of preferential sites of interactions of the radicals on the surface of ubiquitin. The average value of $T_2$ is $\sim 1 ns$ with lower values in some of the loops and at the C-termini, consistent with expected lifetimes for weak protein-cosolute interactions.

A more detailed examination of the sites of preferential interaction of ubiquitin with the 3-carboxy- and 3-carbamoyl-PROXYL radicals is provided in Figure 5. The plots of the $(r^{-6})_{\text{norm}} = (r^{-6})_{\text{norm}}$ profiles as a function of residue shown in Figure 5A and B provide a quantitative description of the relative strength of the interaction sites which is color coded on a graded scale from purple to white on the molecular surface of ubiquitin in Figure 5C and D. There are two major patches of preferential interaction, broadly comparable for the two radicals, and involving a single face of ubiquitin: first patch comprising Leu71 and Leu73; and a second patch consisting of Thr12, Lys63. The opposing face displays only a single weak patch (Asp52) for the 3-carboxy-PROXYL radical only. For comparison surface plots of hydrophobicity and electrostatic potential are displayed in Figure 5E and F, respectively. The sites of preferential interactions cluster in hydrophobic areas and involve residues either located in loops, at the end of $\beta$-strands or in the C-terminal tail. However, it is also clear that the majority of hydrophobic regions on the surface do not exhibit any preferential interaction with the cosolute, and no correlation is seen with electrostatics, despite the fact that 3-carbamoyl-PROXYL is neutral while 3-carboxy-PROXYL is negatively charged at neutral pH.

Finally, we note that fitting the sPRE data with our approximate spectral density function (eq 15) at only a single field (900 MHz), fixing the values of $b$ (eq 17) to those calculated using the experimentally determined translational diffusion constant for the Di-$t$-butyl-nitroxide radical ($D_{\text{trans}} = 7.2 \times 10^{-10} m^2 s^{-1}$) yields almost superimposable $(r^{-6})_{\text{norm}}$ values compared to the 3-field fit and good cross-validation of the $\Gamma_1$ data at 500 MHz (free data set) (Figure S9).

**CONCLUSIONS**

Here we have introduced a simple, approximate non-Lorentzian spectral density function in eq 15 that can be used to quantitatively analyze experimental sPRE data at multiple magnetic fields to extract, in a straightforward manner, two parameters at atomic resolution that relate to both the energetics and time-scale of paramagnetic cosolute-protein interactions: namely, the concentration normalized, ensemble average protein-cosolute interspin distance, $(r^{-6})_{\text{norm}}$ and the effective correlation time $T_2$ (a measure of the length of time that the cosolute electron spin interacts with a $1H$ spin on the protein).

Equation 15 faithfully reproduces the exact spectral densities in eq 20 and $(r^{-6})_{\text{norm}}$ values calculated for an idealized force-free model describing the interaction between two spheres, representing a protein and cosolute, for a wide range of parameters. For experimental sPRE data, excellent cross-
validation between experimental and predicted $\Gamma_1$ values is obtained (i.e. predicting the $\Gamma_1$ values at one field using the fits to the experimental $\Gamma_1$ data from two other fields, for all three combinations of fields). This approach is tailored to a system with very weak, short-lived interactions between a protein and cosolute and does not necessitate any specific type of binding mode. By comparing the $(r^6)_{\text{norm}}$ values derived from fitting the experimental sPRE data at multiple fields with eq 15 with the calculated values obtained directly from the three-dimensional protein coordinates assuming only excluded volume interactions, one can identify regions of preferential interaction between the protein and paramagnetic cosolute at atomic resolution. This approach is expected to provide a powerful tool to characterize and understand the influence of paramagnetic cosolutes on protein stability and folding/unfolding equilibria at atomic resolution.

**EXPERIMENTAL SECTION**

Spectral Density Function Calculation and Fitting Procedure. All calculations of spectral density functions and fits were performed using MATLAB (The MathWorks). Since one cannot easily sample sufficient number of points in the low-frequency regime (given the practical limitations of available magnetic fields and the technical difficulties associated with field cycling), it is a good idea to supply initial estimates of $b$ values, given by eq 17, using the value of the relative translational diffusion constant $D_{\text{trans}}$ (which can be assumed to be equal to the translational diffusion coefficient of the cosolute, given that the protein diffuses slowly on account of its much larger size) estimated using the Stokes–Einstein relationship (eq 26). The radius of the cosolute was estimated from its molecular structure.

NMR Experiments. $^2$H/$^{15}$N-labeled ubiquitin was expressed, isotopically labeled and purified as described previously. $^3$-(Carbamoyl)-2,2,5,5-tetramethyl-1-pyrrolidinyloxy (3-carbamoylPROXYL) and 3-(carboxy)-2,2,5,5-tetramethyl-1-pyrrolidinyloxy (3-carboxyPROXYL) paramagnetic cosolutes were purchased from Millipore Sigma. NMR samples consisted of 0.5 mM $^2$H/$^{15}$N-labeled ubiquitin, 10 mM phosphate buffer ($K_2HPO_4/KH_2PO_4$), pH 7.0, and 5% $D_2O/95$% H$_2$O (v/v) in the absence or presence of 25 mM paramagnetic cosolute. All NMR experiments were recorded at 25 °C on 500, 800, and 900 MHz Bruker AVANCE III spectrometers equipped with TCI z-axis gradient cryogenic probes. Longitudinal $^1$H$_S$–$R_1$ and transverse $^1$H$_S$–$R_2$ rates were measured as described previously. $^1$H$_S$–$\Gamma_1$ and $^1$H$_S$–$\Gamma_2$ values are given by the differences in corresponding $^1$H$_S$–$R_1$ and $^1$H$_S$–$R_2$ values obtained in the presence and absence of the paramagnetic cosolute.

Determination of the Uncertainty in $(r^6)_{\text{norm}}$ and $r_2$ Values Derived from the Experimental sPRE Data. The errors in the values of $(r^6)_{\text{norm}}$ and $r_2$ plotted in Figure 4 were estimated by adding random errors to the calculated $f(0)$ and $f(\infty)$ obtained from fitting the experimental $^1$H$_S$–$\Gamma_1$ and $^1$H$_S$–$\Gamma_2$ data to our approximate spectral density function, eq 15, using a Monte Carlo approach. The random errors were generated using a normal distribution with mean 0 and variance of the corresponding fitting errors. $(r^6)_{\text{norm}}$ values were subsequently calculated for a sample size of 100, and the standard deviation of the calculated $(r^6)_{\text{norm}}$ are plotted as error bars.

Units and Constants. $r_2$ is in units of number of paramagnetic cosolute molecules per m$^3$, given by the concentration in molar units multiplied by Avogadro's number times 10$^6$ (i.e., 6.02 x 10$^{23}$). The values of the various constants in eqs 3 and 4 are as follows: $\mu_\nu = 4\pi \times 10^{−15}$ T$^1$ m$^3$ J$^1$; $\lambda_1 = 0.5 \times 10^{−34}$ J s $^1$; $\eta_2 = 1.76 \times 10^3$ rad s$^1$ T$^1$ and $\eta_2 = 1.76 \times 10^3$ rad s$^1$ T$^1$. For eqs 26 and 27, the Boltzmann constant $k_B = 1.38 \times 10^{−23}$ J K$^1$, with the viscosity $\eta$ in units of J s m$^3$ and the distances $R_i$ in units of meters.

Calculation of $(r^6)_{\text{norm}}$ values for ubiquitin. Relative sPRE values between cosolute and amide protons of ubiquitin were calculated from very high precision atomic coordinates determined by NMR (PDB code 1D3Z) using a modified version of the PSolPot module in Xplor-NIH$^{45}$ to take into account that the unpaired electron is not located at the center of the nitroxide radicals. PSolPot calculates the integral of $k/r^6$ over all solvent accessible volume. When $k = 1$, this is equal to $(r^6)_{\text{norm}}$ that is $\langle \tau^6 \rangle_{\text{norm}}$ in the absence of any intermolecular potential other than volume exclusion of the protein. The radical cosolute was represented by a sphere of radius 3.5 Å with the electron spin located 2.0 Å from the center. To take into account that surface side chains may occupy an ensemble of conformations, $(r^6)_{\text{norm}}$ was calculated for all 10 structures in the 1D3Z PDB file.

**REFERENCES**


**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c00747.

Further theoretical considerations and five additional figures (PDF)

**AUTHOR INFORMATION**

Corresponding Author

G. Marius Clore — Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States; orcid.org/0000-0003-3809-1027; Email: mariusc@mail.nih.gov

Authors

Yusuke Okuno — Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

Attila Szabo — Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c00747

Notes

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We thank Drs. Irina Gopich, Theodoros Karamanos, Dennis Torchia, Vitali Tugarinov, and Zhilin Yang for helpful discussions and experimental data analysis. Dr. Charles Schweitzer kindly modified the PsolPot module in Xplor-NIH for the calculation of the predicted “force-free” solvent PRE from atomic coordinates. We thank Drs. James Baber, Dan Garrett, and Jinfang Yang for technical support on the spectrometers, and Dr. Alberto Cucco for the gift of [H$^{2}$/N$^{15}$]-labeled ubiquitin. This work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (to G.M.C. and A.S.).

**REFERENCES**


