A rigid disulfide-linked nitroxide side chain simplifies the quantitative analysis of PRE data

Nicolas L. Fawzi[1†], Mark R. Fleissner[2†], Nicholas J. Anthis[1†], Támás Kálai[3], Kalman Hideg[3], Wayne L. Hubbell[2]*, G. Marius Clore[1]*

[1]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA.

[2]Jules Stein Eye Institute and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA.

[3]Institute of Organic and Medicinal Chemistry, University of Pécs, H-7624 Pécs, Szigeti str. 12, Hungary.

*Corresponding authors email: G.M.C. (mariusc@mail.nih.gov) and W.L.H. (hubbellw@jsei.ucla.edu)

[†]These authors contributed equally.

Abstract

The measurement of ^{1}H transverse paramagnetic relaxation enhancement (PRE) has been used in biomolecular systems to determine long-range distance restraints and to visualize sparsely-populated transient states. The intrinsic flexibility of most nitroxide and metal-chelating paramagnetic spin-labels, however, complicates the quantitative interpretation of PREs due to delocalization of the paramagnetic center. Here, we present a novel, disulfide-linked nitroxide spin label, R1p, as an alternative to these flexible labels for PRE studies. When introduced at solvent-exposed α-helical positions in two model proteins, calmodulin (CaM) and T4 lysozyme (T4L), EPR measurements show that the R1p side chain exhibits dramatically reduced internal motion compared to the commonly used R1 spin label (generated by reacting cysteine with the spin labeling compound often referred to as MTSL). Further, only a single nitroxide position is necessary to account for the PREs arising from CaM S17R1p, while an ensemble comprising multiple conformations is necessary for those observed for CaM S17R1. Together, these observations suggest that the nitroxide adopts a single, fixed position when R1p is placed at solvent-exposed α-helical positions, greatly simplifying the interpretation of PRE data by removing the need to account for the intrinsic flexibility of the spin label.

Keywords

Paramagnetic relaxation enhancement, site-directed spin-labeling, electron paramagnetic resonance.

Introduction

In recent years, a number of research groups have developed a variety of NMR methods to probe biomolecular structure and dynamics based on the site-specific introduction of paramagnetic species (Clore, 2008, 2011; Clore and Iwahara, 2009; Keizers and Ubbink, 2011; Otting, 2010). The most generally applicable method for introducing paramagnetic species is site-directed spin labeling (SDSL) (Hubbell et al., 2000a; Todd et al., 1989), in which a cysteine is introduced at a site of interest by site-directed mutagenesis, and then modified with a sulfhydryl-specific compound to generate a spin label side chain. Due to the large dipolar interaction between the unpaired electrons of the paramagnetic species and the nuclear spins of the biomolecule, position-specific distance, angular information, or both can be extracted. Among the various paramagnetic NMR techniques, the measurement of transverse paramagnetic relaxation enhancements rates (PREs), particularly of $^{1}H_{N}$ nuclei, has shown broad applicability, allowing a number of important questions in structure and dynamics to be addressed (Fawzi et al., 2010; Iwahara and Clore, 2006; Iwahara et al., 2004a; Iwahara et al., 2006; Takayama and Clore, 2011; Tang et al., 2006; Tang et al., 2008; Tang et al., 2007; Volkov et al., 2010; Volkov et al., 2006). PREs arising from site-directed introduction of nitroxide spin labels or paramagnetic metal ions are large enough to reliably measure distances up to 35 Å from the unpaired electron and to probe transient, sparsely-populated intra- or inter-molecular interactions that occasionally bring the spin label position in close proximity to ^{1}H nuclei of interest (Iwahara et al., 2004b; Iwahara et al., 2007).

Up to this point, however, the motion of the spin label relative to the protein has complicated direct interpretation of both intra- and intermolecular PREs due to the conformational space that can be sampled by the paramagnetic center (Iwahara et al., 2004b). Among the myriad spin labels available, R1 (Figure 1a) is the most widely used for EPR studies, and has also been used successfully in several PRE studies (Bermejo et

al., 2009; Volkov et al., 2006). At solvent-exposed helical sites, the R1 side chain occupies a set of rotamers that are limited from the backbone through the disulfide linkage region due in part due to interactions of the disulfide with the peptide backbone (Fleissner et al., 2009; Guo et al., 2008; Guo et al., 2007; Warshaviak et al., 2011). Transitions between rotamers in this region of the side chain change the spatial position of the nitroxide relative to the protein on the time scale of microseconds (Bridges et al., 2010), though additional rotamer transitions for the terminal rotatable bonds are site-dependent and on a much faster (likely in the ns range) time scale (Chou et al., 2003). NMR studies attempting to quantitatively interpret the PREs arising from R1 (Bermejo et al., 2009), like any other label with internal motions, must take into account the ensemble of nitroxide positions, their interconversion dynamics and the order parameters for the interspin (nitroxide-nucleus) vectors (Clore and Iwahara, 2009; Iwahara et al., 2004b).

To simplify the interpretation of NMR PRE data, an ideal spin label would not perturb the intrinsic structure and dynamics of the protein, and would have a single paramagnetic center that is rigidly attached to the protein. Indeed, several relatively rigid spin labels have been evaluated, but their general use for PRE studies appears to be limited. One such label is TOAC, an amino acid containing a stable nitroxide radical that can be introduced into synthetic peptides and has been used successfully for PRE studies (Lindfors et al., 2011). Although both the C_α and the nitroxide of TOAC are contained within a single piperidine ring, the radical can occupy at least two distinct positions because of ring isomerizations (Crisma et al., 2005; Flippen-Anderson et al., 1996). Moreover, the utility of TOAC for PRE studies is limited because its incorporation into large or isotope labeled proteins is not practical, and also risks perturbing local protein structure and dynamics. Two additional strategies, primarily used for the measurement of pseudo-contact chemical shifts (PCS), rely on multiple functional groups, either engineered or existing, in order to introduce paramagnetic ion chelating moieties. One is CLaNP-5 (Keizers et al., 2007), a large, organic molecule pre-chelated to a lanthanide ion

(614 Da plus a 139 to 175 Da lanthanide ion, compared to 185 Da for R1), that can be introduced via two proximally located, surface cysteine residues. The other is 4-mercaptomethyl-dipicolinic acid (Su et al., 2008), which requires (i) a carboxyl bearing side chain to be located near a single cysteine residue for incorporation and (ii) the addition of lanthanide ions to the protein sample. This labeling strategy is incompatible with metal-binding proteins, and may cause protein aggregation due to non-specific protein-lanthanide binding. As multiple functional groups are necessary to incorporate these labels, this strategy is potentially limited to specific protein topologies, and the covalent attachment of such labels may perturb the structure, dynamics, and stability of the protein under investigation. Thus, a nitroxide label that can be incorporated via a single surface cysteine residue that combines the reliable, sequence-context independent rigidity of a lanthanide-based label with the ease of conjugation and small chemical structure of R1, would be of greater general use for PRE studies.

A suitable candidate may come from a series of R1 analogs, which show more restricted internal motions than R1 at solvent-exposed sites (Columbus et al., 2001) and presumably have fewer allowed conformations due to the presence of a rigid substituent at position 4 of the pyrroline ring. In particular, the 4-phenyl analog exhibits highly restricted motions and adopts a single conformer at a model solvent-exposed helical site (Fleissner, 2007) (PDB code 1ZUR). However, spin labeled proteins bearing this particular derivative are prone to aggregation, probably due to an increase in surface hydrophobicity. To circumvent this problem, we synthesized a reagent (HO-3606) that generates a more hydrophilic, 4-pyridyl derivative of R1 (designated R1p, Figure 1b) and evaluated its general use for PRE studies.

First, we examined the effect of the pyridyl group on the internal flexibility of the spin label side chain by EPR spectroscopy by introducing it into solvent-exposed helical sites in T4 lysozyme (T4L) and human calmodulin (CaM) as model systems. We demonstrate that the EPR spectra of R1p mutants of both proteins are consistent with a

single, highly ordered component, and, most importantly, that the PREs arising from R1p can be accounted for by a single label conformation, greatly simplifying the procedure for using PREs as long range distance restraints or as probes of transient states.

Experimental Methods

Synthesis of the methanethiosulfonate reagent HO-3606

The reagent HO-3606 (radical **5**) was prepared according to the general procedure of Scheme 1. Melting points were determined with a Boetius micro melting point apparatus and are uncorrected. Elemental analyses (C, H, N, S) were performed on Fisons EA 1110 CHNS elemental analyzer. Mass spectra were recorded on a Finnigan Automass-Multi instrument in the EI mode. The IR (Specord 85) spectra were in each case consistent with the assigned structure. Flash column chromatography was performed on Merck Kieselgel 60 (0.040-0.063 mm). Qualitative TLC was carried out on commercially prepared plates (20 x 20 x 0.02 cm) coated with Merck Kieselgel GF_{254}. Compound **1** was prepared according to published procedures (Kalai et al., 1998); 3-pyridineboronic acid and other reagents were purchased from Aldrich.

Synthesis of 3-Formyl-4-(pyridin-3-yl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl radical [2]. To a deoxygenated, stirred solution of aldehyde **1** (988 mg, 4.0 mmol) in dioxane (15 mL), $Pd(PPh_3)_4$ (100 mg, 0.1 mmol) was added in one portion and the mixture was stirred at room temperature (r.t.) for 10 min. Next, 3-pyridineboronic acid (504 mg, 4.1 mmol) was added followed by the addition of aqueous 10% Na_2CO_3 (10 mL, 9.4 mmol), and the mixture was stirred and refluxed for 2h under N_2. After cooling, the solvents were evaporated off *in vacuo*, and the residue was partitioned between $CHCl_3$ (25 mL) and water (10 mL). The organic phase was separated, dried ($MgSO_4$), filtered and evaporated to give aldehyde **2**. Yield: 617 mg (63%) after flash chromatography purification ($CHCl_3/Et_2O$, 2:1), yellow solid, mp 98-99 °C, R_f : 0.21 ($CHCl_3/Et_2O$, 2:1). MS EI m/z (%): 245 (M^+, 100), 230(30), 200 (42), 172(88). IR (nujol): 1670 (C=O), 1640, 1560, 1535 (C=C) cm^{-1}. Composition calculated for $C_{14}H_{17}N_2O_2$: C, 68.55; H, 6.99; N, 11.42. Found: C, 68.50; H, 7.00; N, 11.36.

Synthesis of 3-Hydroxymethyl-4-(pyridin-3-yl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl radical [3]: To a solution of aldehyde **2** (980 mg, 4.0 mmol) in EtOH (15 mL), $NaBH_4$ (190 mg, 5.0 mmol) was added in one portion, and the mixture was stirred at r.t. for 15 min. and then quenched with water (10 mL). The solution was extracted with $CHCl_3$ (2 x 10 mL), the organic phase was dried ($MgSO_4$), filtered and evaporated and the residue was purified by flash column chromatography ($CHCl_3$ / MeOH, 9:1) to offer alcohol **3**. Yield: 830 mg (84%), off-white solid, mp 145-147 °C, R_f : 0.14 ($CHCl_3$/MeOH, 9:1). MS EI m/z (%): 247 (M^+, 100), 232(61), 217 (16), 184(35). IR (nujol): 3150 (OH), 1620, 1560, 1535 (C=C) cm^{-1}. Composition calculated for $C_{14}H_{19}N_2O_2$: C, 67.99; H, 7.74; N, 11.33. Found: C, 67.92; H, 7.71; N, 11.25.

Synthesis of 3-Chloromethyl-4-(pyridin-3-yl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl radical [4]: To a solution of alcohol **3** (1.23 g, 5.0 mmol) in CH_2Cl_2 (15 mL) and Et_3N (555 mg, 5.5 mmol), $MeSO_2Cl$ (630 mg, 5.5 mmol) was added at 0 °C. The mixture was stirred at r.t. for 30 min., then washed with water (10 mL), and the organic phase was separated, dried ($MgSO_4$), filtered and evaporated. The residue was dissolved in dry acetone (40 mL), then LiCl (420 mg, 10 mmol) was added and the mixture stirred and refluxed for 30 min. The solvent was evaporated off *in vacuo*, and the residue was partitioned between EtOAc (30 mL and water (10 mL). The organic phase was separated, dried ($MgSO_4$), filtered, and evaporated, and the crude residue was further purified by flash column chromatography ($CHCl_3$/Et_2O, 2:1) to provide radical **4**. Yield: 689 mg (52%), pale yellow solid, mp 87-88 °C, R_f : 0.31 ($CHCl_3$/Et_2O, 2:1). MS EI m/z (%): 265/267 (M^+, 82/28), 250/252(56/21), 235/237 (11/4), 215(100), 184(17). IR (nujol): 1620, 1560, 1535 (C=C) cm^{-1}. Composition calculated for $C_{14}H_{18}ClN_2O$: C, 63.27; H, 6.83; N, 10.54. Found: C, 63.20; H, 6.85; N, 10.52.

Synthesis of 3-Methanesulfonilthiomethyl -4-(pyridin-3-yl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl Radical [5] (i.e., HO-3606). To a solution of compound **4** (532 mg, 2.0 mmol) in acetone (10 mL), $NaSSO_2CH_3$ (536 mg, 4.0 mmol) in water (5 mL) was added and the mixture was kept at 40 °C. The reaction mixture was monitored by TLC. After consumption of most of the starting material (~30 min), the acetone was evaporated off *in vacuo*, and the residue was partitioned between $CHCl_3$ (30 ml) and brine (10 mL). The organic phase was separated, dried ($MgSO_4$), filtered, and evaporated. The resulting residue was further purified by flash column chromatography (hexane/EtOAc and $CHCl_3/Et_2O$) to give the methanethiosulfonate [5]. Yield: 245 mg (36%), pale yellow solid, mp-125-127 °C, R_f 0.12 ($CHCl_3/Et_2O$, 2:1). MS EI m/z (%): 341 (M^+, 29), 326(10), 346 (39), 215(100), 79(42). IR (nujol): 1615, 1560, 1535 (C=C) cm^{-1}. Composition calculated for $C_{15}H_{21}N_2O_3S_2$: C, 52.76; H, 6.20; N, 8.20; S, 18.78. Found: C, 52.53; H, 6.13; N, 8.05; S, 18.55.

Sample preparation

The genetic construct used for expressing single cysteine mutants of T4 lysozyme (T4L) in the WT* background (C54T, C97A) has been described previously (Mchaourab et al., 1996). The mutants were expressed, purified, and spin labeled with either the common spin label (1-Oxyl-2,2,5,5-tetramethyl-2,5-dihydropyrrol-3-ylmethyl) methanethiosulfonate (sometimes referred to as “MTSL” or “MTSSL”) (Hankovszky et al., 1980) or HO-3606 according to the method described previously (Fleissner et al., 2009).

For experiments on calmodulin (CaM), a pET21a (Novagen) vector encoding the 148-residue human (CaM) protein was created by cloning the gene for CaM from *Xenopus laevis* (whose amino acid sequence is identical to the human protein) into the vector using PCR primers with *BamH1* and *NdeI* restriction sites. A CaM variant for site-

directed spin labeling, CaM S17C was created with the QuikChange Site-Directed Mutagenesis Kit (Stratagene).

CaM S17C with uniform $^{2}H/^{13}C/^{15}N$ labeling was expressed in BL21-CodonPlus (DE3) RIPL cells (Stratagene) at 37°C using standard methods. Cultures of 1 L of minimal M9 medium using $^{2}H_2O$, $^{15}NH_4$,Cl and $^{2}H_7/^{13}C$-glucose as the sole sources of hydrogen, nitrogen and carbon, respectively, were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at A_{600nm} ~0.6, and harvested by centrifugation 6-9 hours later.

Cells were resuspended in ~25 ml of 40 mM Tris pH 7.4, 2 mM EDTA, 2 mM dithiothreitol (DTT), and 1x complete protease inhibitor (Roche), and lysed using a microfluidizer. The lysate was cleared by centrifugation and loaded onto a DEAE ion-exchange column (GE Life Sciences) pre-equilibrated with 20 mM Tris, pH 7.4, 80 mM $(NH_4)_2SO_4$ and 1 mM $MgCl_2$. Protein was eluted with a linear 0-100% gradient over 10 column volumes with a high-salt buffer containing 1.2 M $(NH_4)_2SO_4$. Fractions containing CaM were pooled and loaded onto phenyl-sepharose resin (GE Life Sciences) pre-equilibrated with 50 mM Tris, pH 7.4, 2 mM $CaCl_2$ and 100 mM NaCl. CaM was eluted with the addition of 2 mM EGTA. Protein was stored in the elution buffer plus 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) and 1x complete protease inhibitor.

CaM S17C (~ 200μM) was spin labeled by the addition of a 10-fold molar excess of either "MTSL," HO-3606, or (1-Acetoxy-2,2,5,5-tetramethyl-δ-3-pyrroline-3-methyl) methanethiosulfonate (Toronto Research Chemicals), to generate the spin label side chains R1, R1p, and a diamagnetic analog of R1, respectively. Labeling reactions were allowed to proceed at room temperature in the dark for ~2 hours and tested for completion by liquid chromatography/mass spectrometry. Unreacted spin-label was removed with a HiPrep 26/10 Desalting Column (GE Life Sciences) and exchanged into NMR buffer, comprising 5% $^{2}H_2O$/95% $^{1}H_2O$, 100 mM KCl, 0.1x EDTA-free complete protease inhibitor (Roche), 25 mM HEPES, pH 6.5, and 8 mM $CaCl_2$, or into EPR buffer

containing the same salts and buffering agents with $^{1}H_2O$ only. NMR samples were concentrated to 0.3 mM with Amicon Ultra Centrifugal Filter Units (3 kDa molecular weight cutoff). EPR experiments were performed at a concentration of 0.1 mM. Experiments were performed in the presence of 1.2-fold excess (added prior to centrifugation) of a tightly binding synthetic peptide (KRRWKKNFIAVSAANRFKKIS SSGAL; Anaspec) corresponding to the CaM-binding domain (CBD) of human skeletal muscle myosin-light-chain kinase (skMLCK).

EPR Spectroscopy

For EPR experiments, ~5 μL samples of T4L (~0.5 mM protein in 50 mM MOPS pH 6.8, 25 mM NaCl) or CaM (~0.1 mM protein in NMR buffer without $^{2}H_2O$) were loaded into glass capillaries (0.6 mm x 0.84 mm) sealed at one end. Where indicated, the sample also included 30% w/v sucrose to increase the viscosity of the medium, a 1.2 molar excess of the previously described CBD synthetic peptide from human skMLCK, or both. Spectra were collected at X-band (9 GHz) over 100 G on a Bruker Elexys 580 spectrometer fitted with a high-sensitivity resonator using 20 mW incident microwave power and 1 G field modulation amplitude at 100 kHz. Sample temperature was maintained within a single degree of the reported value using a Bruker ER4131VT temperature controller. EPR spectra were simulated using the MOMD model (Budil et al., 1996), using spatially averaged A and g tensor principal values corresponding to the nitroxide motion relative to the protein (the time-independent, effective Hamiltonian) (Hubbell et al., 2000a).

NMR Spectroscopy

Backbone resonance assignments for peptide bound CaM were transferred from known assignments and verified by triple resonance correlation spectroscopy (Clore and Gronenborn, 1994) using standard HNCO, HNCACB, and CBCA(CO)NH pulse

sequences. Transverse $^{1}H_{N}$-Γ_2 PRE rates were obtained from the differences in the $^{1}H_{N}$-R_2 transverse relaxation rates between the paramagnetic and diamagnetic samples (Clore and Iwahara, 2009) using a 3D HNCO-based pulse scheme similar to a previously described scheme (Hu et al., 2009) except without TROSY. Compared to most PRE experiments, this experiment adds a $^{13}C'$ dimension to resolve overlapping resonances present in the 2D ^{1}H-^{15}N correlation spectrum of CaM. Data were recorded at 27°C on a Bruker 600 MHz spectrometer equipped with a triple resonance *z*-gradient cryoprobe. Two time points (separated by 20 ms) were used for the $^{1}H_{N}$-R_2 measurements (Iwahara et al., 2007). In this case, no significant differences in $^{1}H_{N}$-R_2 were observed for CaM incorporating the diamagnetic control label and without a label (free-cysteine) in NMR buffer containing 5mM DTT.

PRE data analysis

The position, or ensemble of positions, of the nitroxide spin label that best fits the observed $^{1}H_{N}$-Γ_2 PRE rates was computed by comparing predicted and observed $^{1}H_{N}$-Γ_2 PREs using the coordinates of the structure of Ca^{2+}-loaded, peptide-bound CaM in complex with smooth muscle MLCK (PDB 1CDL) (Meador et al., 1992). $^{1}H_{N}$-Γ_2 PREs were back-calculated in Xplor-NIH (Schwieters et al., 2006; Schwieters et al., 2003) with either a single conformer or a five-conformer ensemble for the spin label together with the Solomon-Bloembergen Model Free (SBMF) representation (Iwahara et al., 2004b). The label and immediately adjacent side chain coordinate positions were optimized in torsion angle space by simulated annealing to minimize the difference between observed and calculated $^{1}H_{N}$-G_2 PREs, as previously described (Iwahara et al., 2004b) with the addition of a weighting factor for the PRE force constant equal to the square of the inverse of the expected error in the observed PRE at each residue (Xplor-NIH PRE force constant option sigma). The error in the observed PRE was calculated from the peak intensities at the two time points in the paramagnetic and diamagnetic samples (Iwahara

et al., 2004b) and a minimum error of ±1.5 s^{-1} was used to account for small sample-to-sample variations. In addition to the PRE pseudo-potential, the target function includes stereochemical restraints, a quartic van der Waals repulsion term to prevent atomic overlap between the spin-label and the protein, and a multidimensional torsion angle database potential of mean force (Clore and Kuszewski, 2002). (Note atomic overlap between spin labels in the five-conformer ensemble is allowed since the ensemble represents a distribution of states.) To evaluate the goodness-of-fit, the appropriate reduced χ^2:

$$\chi^2 = \frac{1}{M}\sum_N \left(\frac{\Gamma_2^{observed} - \Gamma_2^{predicted}}{\sigma_{observed}} \right)^2 \qquad [1]$$

was computed. The number of degrees of freedom M is given by:

$$M = N - (3C) \qquad [2]$$

where N is the number of observed PREs, and C, the number of spin label conformers in the fitting ensemble, one or five. C is multiplied by three, for the three translational degrees of freedom involved in fitting each nitroxide position. Because the effective electron relaxation time τ_s, which is equal to the electron spin-lattice relaxation time T_{1e} in the case of ^{1}H-Γ_2 PREs for macromolecules, is much longer than the protein rotational correlation time τ_r (Iwahara et al., 2007), the PRE correlation time τ_c [$=(\tau_r^{-1} + \tau_s^{-1})^{-1}$] for the calculation of PRE rates was assumed to be approximately the same as τ_r (9.9 ns) for CaM (Chang et al., 2003).

Complete Xplor-NIH scripts and documentation to introduce the spin label into a protein of known structure, find the label position(s) that best-fits the intramolecular PRE data, and compute χ^2 are available online at http://nmr.cit.nih.gov/xplor-nih/doc/current.

Results

EPR demonstrates restricted internal motion of R1p relative to R1

To compare the internal flexibility of the R1p side chain to that of R1, the solvent-exposed site D72 of T4L was selected for introduction of the spin labels. This site has previously been used as a reference since contributions of backbone fluctuations to the motion of the nitroxide are minimal (Columbus et al., 2001). The room temperature EPR spectra of the T4L D72R1p mutant and the previously published D72R1 mutant (Mchaourab et al., 1996) are presented in Figure 2a. Both mutants were studied in buffer (Figure 2a, red traces), but were also studied in a 30% sucrose solution (Figure 2a, black traces), which reduces the effect of protein rotational diffusion on the spectrum and allows the motion of the nitroxide relative to the protein to be examined (Mchaourab et al., 1996). Assuming a spherical geometry for T4L, an estimate of its rotational correlation time (τ_R) at room temperature according to the Stokes-Einstein relation is 6 to 8 ns in buffer and 18 to 24 ns in 30% sucrose. Fits to multifrequency EPR spectra of spin labeled T4L yield $\tau_R \approx 10$ ns in buffer and approximately 3 times longer in a viscosity corresponding to 30% sucrose (Zhang et al., 2010). Regardless of the exact values, rotational diffusion within this time regime will strongly influence the EPR spectrum as shown in Figure 2.

In solutions of 30% sucrose, the T4L mutants D72R1 and D72R1p have spectra characterized by well-resolved outer hyperfine extrema with splitting of $2A_{zz}'$ (Figure 2) that correspond to nitroxides with the 2p orbital on nitrogen oriented along the laboratory magnetic field. For T4L D72R1, the overall lineshape arises from a fast anisotropic internal motion of the nitroxide that can be accurately described by an order parameter (***S***) and correlation time (τ_i) on the ns time scale (Columbus, 2001). In the fast motion limit, the time-dependent Hamiltonian describing the motion can be replaced by an effective, time-independent, Hamiltonian with the A and g magnetic interactions tensors averaged over motion within a cone. In this limit, which has been shown to be a

reasonable approximation for T4L D72R1, the value of $2A_{zz}'$ is a quantitative measure of the order parameter, increasing with increased order (Columbus et al., 2001).

As is evident in Figure 2a, the spectra of T4L D72R1p and T4L D72R1 in sucrose are similar except that $2A_{zz}'$ is considerably larger for that of T4L D72R1p (64 G for D72R1p and 53 G for D72R1), indicating a higher order for the internal motion of R1p. This result is consistent with expectations from the crystal structure of the 4-phenyl analog of R1 in which the entire side chain was resolved (see Discussion). Due to this increase in order, the rotational diffusion of the protein, even in 30% sucrose, will contribute to the EPR spectrum of T4L D72R1p by reducing $2A_{zz}'$. To deduce the $2A_{zz}'$ value corresponding to purely internal motion of the side chain, spectra of T4L D72R1p were recorded as a function of sucrose concentration from 15 to 35% wt/wt, and then analyzed (Timofeev and Tsetlin, 1983) to give $2A_{zz}' \approx 70$ G, 5 G less than that in the rigid limit (Kusnetzow et al., 2006). Within the context of the effective Hamiltonian model for motion in a cone (Griffith and Jost, 1976), this result gives an order parameter for the internal motion of R1p of $\boldsymbol{S} = 0.76$, compared to 0.43 for R1 (Columbus et al., 2001). If this simple model has merit, then it should be possible to reproduce the experimental spectra of T4L D72R1p in both buffer and 30% sucrose using the effective Hamiltonian parameters corresponding to $\boldsymbol{S} = 0.76$ and computing the effect of protein rotational diffusion according to the stochastic Louisville method that is widely used for simulating the effect of motion on EPR spectra (Budil et al., 1996). The dashed traces shown in Figure 2 are fits to the experimental spectra using the effective Hamiltonian parameters given in the legend and allowing the isotropic τ_R of the protein to vary. The best-fit τ_R values are 5 and 17 ns for buffer and 30% sucrose, respectively, and both are within the expected range, considering the simple nature of the model and the assumed isotropic diffusion of the protein. Thus, a first-order working model for the overall motion of the R1p sidechain is a fast, highly-ordered internal motion of small amplitude modulated by slow rotational diffusion of the protein. To explore the generality of the

ordered internal motion of R1p, it was placed at several other solvent-exposed helical sites in T4L (Figure 3). In each case, the spectra are very similar to T4L D72R1p, suggesting that the above results will be general.

The spectrum of CaM S17R1 is similar to other non-interacting, solvent exposed helical sites that exhibit fast (~2 ns), weakly ordered ($\boldsymbol{S} < 0.5$) motion. As for T4L, the spectra of CaM S17R1p exhibit reduced mobility compared to those of the analogous R1 mutant (Figure 2b). Remarkably, the spectra of CaM S17R1p are similar to those of T4L D72R1p. Because T4L and CaM are of similar size and the sites in each case are located at the solvent exposed surface of a helix, the internal motions relative to the protein must be comparable.

A single label conformation accounts for the experimental PREs

The ^1H_N-Γ_2 PRE rates for CaM S17C reacted with the R1 side chain label, CaM S17R1, and with the pyridyl label, CaM S17R1p, are presented in Figure 4. In the Ca^{2+}-loaded and peptide-bound state, CaM adopts a rigid structure in which the hydrophobic clefts of the N- and C-terminal domains of CaM close around the target peptide (Ikura et al., 1992; Meador et al., 1992). Since the labels are attached in the same S17C position, the pattern of PREs arising from S17R1 and S17R1p are similar. Quantitatively, however, the effect of the larger conformational dynamics of R1 compared to R1p on the observed PREs can be easily distinguished.

Assuming a single, static conformation of the side chain, CaM S17R1 shows a reasonable fit with a reduced χ^2 of 1.59 (Figure 4, top, blue line) (a value of 1.0 suggesting a perfect fit within the error of the measurement). Though the fit is good overall, the observed PREs at positions 67, 68, and 69 are clearly not well fit, suggesting that the nitroxide position may sample many different states with distinct PREs that cannot be captured by a single structure. By introducing an ensemble of five (equally weighted) conformations, the fit to the data for S17R1 can be improved such that the

predicted PREs are within the experimental error in this region (Figure 4, top, red line). The reduced χ^2 for the five-conformer ensemble drops to 1.22, indicating that the fit is quantitatively superior to the fit using a single conformation, and hence a multiple conformer ensemble is necessary to quantitatively treat the PREs for S17R1.

In contrast, the PREs for CaM S17R1p can be quantitatively accounted for by a single conformation of the label, resulting in a χ^2 of 1.29 (Figure 4, bottom, blue line). This value is considerably smaller than the χ^2 for a single conformer representation of S17R1 and comparable to the χ^2 for the five-conformer ensemble. Unlike in the case of S17R1, fitting an ensemble of five conformations of the label to the observed PREs for S17R1p leaves the fit essentially unchanged (red and blue lines overlap) and does not improve the reduced χ^2, indicating that a single conformation of the R1p spin-label side chain is sufficient to account for the observed PREs.

Discussion

Recent studies have demonstrated that PREs can not only provide sensitive, long-range structural restraints (Iwahara et al., 2004b), but can also allow transient, sparsely-populated structures to be observed (Clore, 2011; Clore and Iwahara, 2009). However, a quantitative description of the PREs arising from most spin labels (e.g., R1) typically requires the use of a ensemble of spin label conformers to represent the positions of the nitroxides due to the inherent conformational flexibility of the spin label side chain. Here we introduce and characterize a new, rigid spin label for proteins that, when placed at solvent-exposed helical positions, can simplify PRE data analysis. In contrast to other rigidly attached, lanthanide labels developed for PCS, only a single cysteine is required to incorporate this new, disulfide-linked spin label, designated R1p; the thiol-specific reagent for generating R1p (HO-3606, scheme 1) is a 4-pyridyl analog of the commonly used MTSL reagent that generates the R1 spin label. Although R1 and R1p differ only by the 4-subsituent, the EPR spectra of R1p at solvent-exposed sites in T4L and CaM are

dramatically different from those for the analogous R1 mutants, exhibiting highly restricted motion relative to the protein (Figure 2). This result is consistent with the crystal structure of the 4-phenyl analog at T4L V131, in which the entire side chain is resolved in a single rotamer (Fleissner, 2007) (PDB accession code 1ZUR). In contrast, two rotamers were observed in the crystal structure of R1 at the same site though the nitroxide rings of neither rotamer was resolved due to disorder (Fleissner et al., 2009). Interestingly, the EPR spectra of the rigidly attached R1p mutants are strikingly similar to mutants bearing the 4-phenyl analog of R1 (Columbus et al., 2001), suggesting that the conformational space of the R1p side chain is also restricted. Indeed, the intramolecular PREs observed for CaM S17R1p are fully accounted for within experimental error by a single nitroxide position, and the fit cannot be improved by introducing an ensemble representation. This is in sharp contrast to the PREs arising from S17R1 where a multiple conformer ensemble is required and yields a significantly better fit than a single conformer representation.

At any surface helical site, the assumption that R1p can be treated as a occupying a single position for PRE studies can be quickly determined by combining a pair of $^{1}H_{N}$-R_2 experiments to compute intramolecular (or intradomain) PREs with the Xplor-NIH scripts we have provided (see Methods). Since quantitative calculation of PREs for a label with a single nitroxide position does not require the representation of a conformational ensemble of labels nor a model free analysis incorporating order-parameters for the nitroxide-proton vectors, we propose that site directed spin-labeling with HO-3606 to form the nitroxide bearing side chain R1p will dramatically simplify the quantitative procedures needed for analyzing intra- and intermolecular PRE data.

Acknowledgments

We thank Vincenzo Venditti and Mengli Cai for helpful suggestions, Dusty Baber and Jinfa Ying for assistance with NMR spectrometers, and Evan Brooks for expert technical assistance. This work was in part supported by funds from the Intramural Program of the NIH, NIDDK and the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.), NIH grants R01EY05216 (to W.L.H.), and the Jules Stein Professorship Endowment (to W.L.H.). The synthesis of new spin label reagents was supported by Hungarian National Research Funds (OTKA K81123).

Figure Legends

Figure 1 Reaction of the widely used spin-labeling compound ("MTSL") and HO-3606 with a protein sulfhydryl group to generate the R1 and R1p side chains, respectively.

Figure 2 EPR spectra of R1 and R1p labeled proteins. **a,** EPR spectra of T4L D72R1 (left) and T4L D72R1p (right) at 298 K. **b,** EPR spectra of CaM S17R1 (left) and CaM S17R1p (right) at 298 K. Spectra shown in red are of the mutants in buffer, and those in black are in buffer containing 30% (w/v) sucrose. Shown below the T4L D72R1p spectra (gray, dashed trace) are simulated spectra according to a modified effective Hamiltonian model using values for the magnetic tensor derived assuming an order parameter of 0.76 and no diffusion axis tilt ($A_{xx} = A_{yy} = 7.07$ G, $A_{zz} = 34.9$ G, $g_{xx} = g_{yy} = 2.00667$, $g_{zz} = 2.00267$). Best-fit values for τ_r were 5.4 ns (buffer) and 17 ns (30% sucrose). All spectra are normalized to the same concentration of nitroxide and have a magnetic field scan width of 100 G. Vertical scaling factors (×2, left side of spectrum) were applied to select spectra for display purposes.

Figure 3 EPR spectra of T4L R1p mutants in a buffered solution containing 30% sucrose at 295 K. Solvent-exposed helical sites in T4L were chosen for incorporation of the R1p spin label. The magnetic field scan width is 100 G.

Figure 4 Comparison of observed and calculated PRE data for Ca^{2+}-loaded, peptide-bound CaM with a spin label conjugated to S17C: **a,** S17R1; **b,** S17R1p. Experimental PRE profiles are shown as circles (error bar, 1 s.d.). PREs too large (> 60 s^{-1}) to be accurately measured are plotted at the top of the charts. PRE profiles back-calculated from the structures of CaM (Ca^{2+}-loaded, peptide-bound) are displayed as solid lines using either a single conformer (blue) or 5 conformer ensemble (red) to represent the spin label. Reduced χ^2 values are given for each of the fits. Note that use of a multiple-

conformer ensemble representation significantly improves the fit to the experimental data in the case of R1 but not R1p.

References

Bermejo GA, Strub MP, Ho C, Tjandra N (2009) Determination of the solution-bound conformation of an amino acid binding protein by NMR paramagnetic relaxation enhancement: Use of a single flexible paramagnetic probe with improved estimation of its sampling space. J Am Chem Soc 131:9532-9537.

Bridges MD, Hideg K, Hubbell WL (2010) Resolving conformational and rotameric exchange in spin-labeled proteins using saturation recovery epr. Applied Magn Reson 37:363-390.

Budil DE, Lee S, Saxena S, Freed JH (1996) Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified levenberg-marquardt algorithm. J Magn Reson Series A 120:155-189.

Chang SL, Szabo A, Tjandra N (2003) Temperature dependence of domain motions of calmodulin probed by NMR relaxation at multiple fields. J Am Chem Soc 125: 11379-11384.

Chou JJ, Case DA, Bax A (2003) Insights into the mobility of methyl-bearing side chains in proteins from $^3J_{CC}$ and $^3J_{CN}$ couplings. J Am Chem Soc 125:8959-8966.

Clore GM (2008) Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement. Mol Biosystems 4:1058-1069.

Clore GM (2011) Exploring sparsely populated states of macromolecules by diamagnetic and paramagnetic nmr relaxation. Protein Sci 20:229-246.

Clore GM, Gronenborn AM (1994) Multidimensional heteronuclear nuclear magnetic resonance of proteins. Meth Enzymol 239:349-363.

Clore GM, Iwahara J (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem Rev 109:4108-4139.

Clore GM, Kuszewski J (2002) χ_1 rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force. J Am Chem Soc 124:2866-2867.

Columbus L (2001) Investigating backbone and side chain dynamics of α-helices in the nanosecond regime with site-directed spin labeling. PhD Dissertation, UCLA, pp 51-58.

Columbus L, Kalai T, Jeko J, Hideg K, Hubbell WL (2001) Molecular motion of spin labeled side chains in α-helices: Analysis by variation of side chain structure. Biochemistry 40:3828-3846.

Crisma M, Deschamps JR, George C, Flippen-Anderson JL, Kaptein B, Broxterman QB, Moretto A, Oancea S, Jost M, Formaggio F, et al. (2005) A topographically and conformationally constrained, spin-labeled, α-amino acid: Crystallographic characterization in peptides. J Pep Res 65:564-579.

Fawzi NL, Doucleff M, Suh JY, Clore GM (2010) Mechanistic details of a protein-protein association pathway revealed by paramagnetic relaxation enhancement titration measurements. Proc Natl Acad Sci USA 107:1379-1384.

Fleissner MR (2007). X-ray structures of nitroxide side chains in proteins: A basis for interpreting distance measurements and dynamic studies by electron paramagnetic resonance. PhD Thesis, UCLA (2007).

Fleissner MR, Cascio D, Hubbell WL (2009) Structural origin of weakly ordered nitroxide motion in spin-labeled proteins. Protein Sci 18:893-908.

Flippen-Anderson JL, George C, Valle G, Valente E, Bianco A, Formaggio F, Crisma M, Toniolo C (1996) Crystallographic characterization of geometry and conformation of TOAC, a nitroxide spin-labelled C α,α-disubstituted glycine, in simple derivatives and model peptides. Int J Pept Protein Res 47:231-238.

Griffith OH, Jost PC (1976). Lipid spin labels in biological membranes. In Spin Labeling: Theory and Applications, L.J. Berliner, ed. (New York, Academic Press), pp. 454-523.

Guo ZF, Cascio D, Hideg K, Hubbell WL (2008) Structural determinants of nitroxide motion in spin-labeled proteins: Solvent-exposed sites in helix B of T4 lysozyme. Protein Sci 17:228-239.

Guo ZF, Cascio D, Hideg K, Kalai T, Hubbell WL (2007) Structural determinants of nitroxide motion in spin-labeled proteins: tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme. Protein Sci 16:1069-1086.

Hankovszky HO, Hideg K, Lex L (1980) Nitroxyls VII. Synthesis and reactions of highly reactive 1-oxyl-2,2,5,5-tetramethyl-2,5-dihydropyrrole-3-ylmethyl sulfonates. Synthesis-Stuttgart 914-916.

Hu K, Doucleff M, Clore GM (2009) Using multiple quantum coherence to increase the ^{15}N resolution in a three-dimensional TROSY HNCO experiment for accurate PRE and RDC measurements. J Magn Reson 200:173-177.

Hubbell WL, Cafiso DS, Altenbach C (2000) Identifying conformational changes with site-directed spin labeling. Nature Struct Biol 7:735-739.

Ikura M, Clore GM, Gronenborn AM, Zhu G, Klee CB, Bax A (1992) Solution structure of a calmodulin-target peptide complex by multidimensional NMR. Science 256:632-638.

Iwahara J, Clore GM (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. Nature 440: 1227-1230.

Iwahara J, Schwieters CD, Clore GM (2004a) Characterization of nonspecific protein-DNA interactions by ^{1}H paramagnetic relaxation enhancement. J Am Chem Soc 126:12800-12808.

Iwahara J, Schwieters CD, Clore GM (2004b) Ensemble approach for nmr structure refinement against ^{1}H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. J Am Chem Soc 126:5879-5896.

Iwahara J, Zweckstetter M, Clore GM (2006) NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. Proc Natl Acad Sci USA 103:15062-15067.

Iwahara J, Tang C, Clore GM (2007) Practical aspects of ^{1}H transverse paramagnetic relaxation enhancement measurements on macromolecules. J Magn Reson 184:185-195.

Kalai T, Balog M, Jeko J, Hideg K (1998) 3-substituted 4-bromo-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxyl radicals as versatile synthons for synthesis of new paramagnetic heterocycles. Synthesis-Stuttgart 1476-1482.

Keizers PHJ, Desreux JF, Overhand M, Ubbink M (2007) Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. J Am Chem Soc 129: 9292-9293.

Keizers PM, Ubbink M (2011) Paramagnetic tagging for protein structure and dynamics analysis. Progr Nucl Magn Reson Spectros 58: 88-96.

Kusnetzow AK, Altenbach C, Hubbell WL (2006) Conformational states and dynamics of rhodopsin in micelles and bilayers. Biochemistry 45: 5538-5550.

Lindfors HE, Venkata BS, Drijfhout JW, Ubbink M (2011) Linker length dependent binding of a focal adhesion kinase derived peptide to the src SH3-SH2 domains. FEBS Lett 585: 601-605.

Mchaourab HS, Lietzow MA, Hideg K, Hubbell WL (1996) Motion of spin-labeled side chains in T4 lysozyme, correlation with protein structure and dynamics. Biochemistry 35:7692-7704.

Meador WE, Means AR, Quiocho FA (1992) Target enzyme recognition by calmodulin - 2.4Å structure of a calmodulin-peptide complex. Science 257:1251-1255.

Otting G (2010) Protein NMR using paramagnetic ions. Ann Rev Biophys 39:387-405.

Schwieters CD, Kuszewski JJ, Clore GM (2006) Using Xplor-NIH for NMR molecular structure determination. Prog Nucl Magn Reson Spec 48:47-62.

Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160:65-73.

Su XC, Man B, Beeren S, Liang H, Simonsen S, Schmitz C, Huber T, Messerle BA, Otting G (2008) A dipicolinic acid tag for rigid lanthanide tagging of proteins and paramagnetic nmr spectroscopy. J Am Chem Soc 130: 10486-10847.

Takayama Y, Clore GM (2011) Intra- and intermolecular translocation of the bi-domin tramnscription factor Oct1 characterized by liquid crystal and paramagnetic NMR. Proc Natl Acad Sci USA.

Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein-protein association. Nature 444: 383-386.

Tang C, Louis JM, Aniana A, Suh JY, Clore GM (2008) Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. Nature 455:693-696.

Tang C, Schwieters CD, Clore GM (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. Nature 449:1078-1082.

Timofeev VP, Tsetlin VI (1983) Analysis of mobility of protein side-chains by spin label technique. Biophys Struct Mech 10: 93-108.

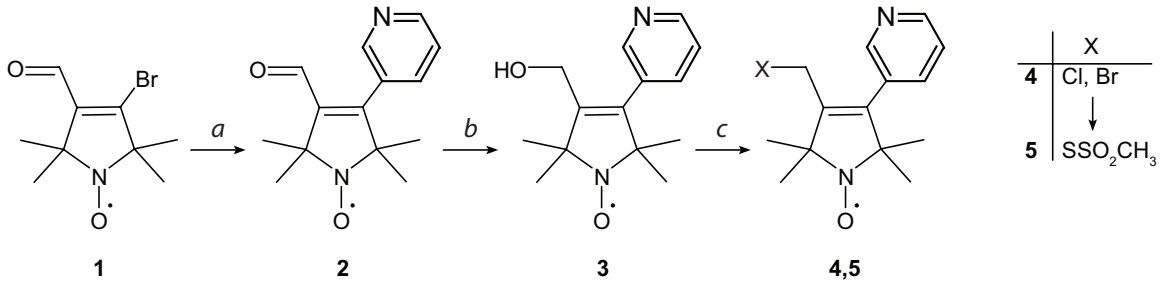
Todd AP, Cong J, Levinthal F, Levinthal C, Hubbell WL (1989) Site-directed mutagenesis of colicin E1 provides specific attachment sites for spin labels whose spectra are sensitive to local conformation. Proteins 6:294-305.

Volkov AN, Ubbink M, van Nuland NAJ (2010) Mapping the encounter state of a transient protein complex by PRE NMR spectroscopy. J Biomol NMR 48:225-236.

Volkov AN, Worrall JAR, Holtzmann E, Ubbink M (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. Proc Natl Acad Sci USA 103:18945-18950.

Warshaviak DT, Serbulea L, Houk KN, Hubbell WL (2011) Conformational analysis of a nitroxide side chain in an α-helix with density functional theory. J Phys Chem B 115:397-405.

Zhang ZW, Fleissner MR, Tipikin DS, Liang ZC, Moscicki JK, Earle KA, Hubbell WL, Freed JH (2010) Multifrequency electron spin resonance study of the dynamics of spin labeled T4 lysozyme. J Phys Chem B 114:5503-5521.



Scheme 1

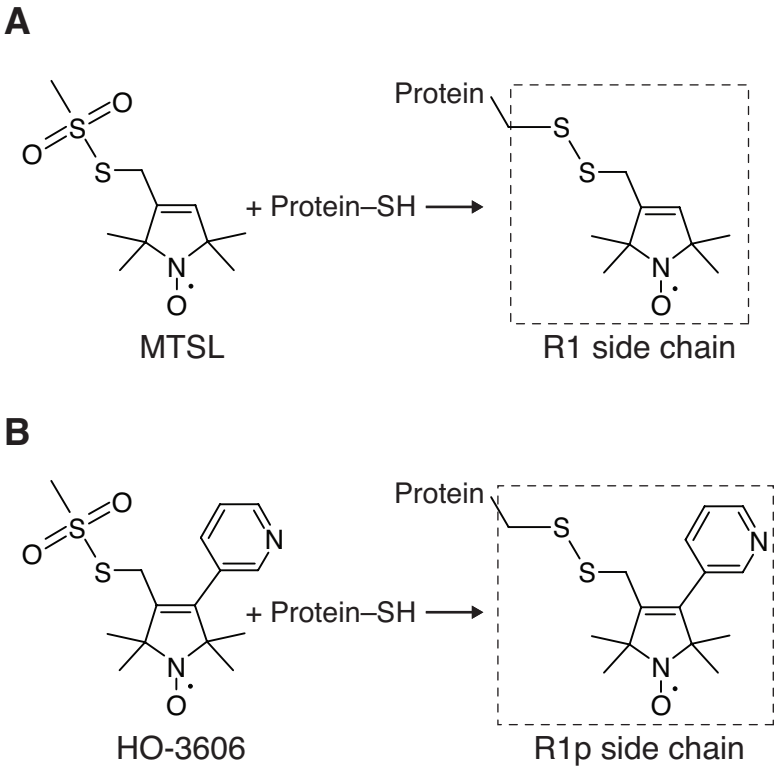

A
Protein
+ Protein–SH
MTSL
R1 side chain
B
Protein
+ Protein–SH
HO-3606
R1p side chain

Fig. 1

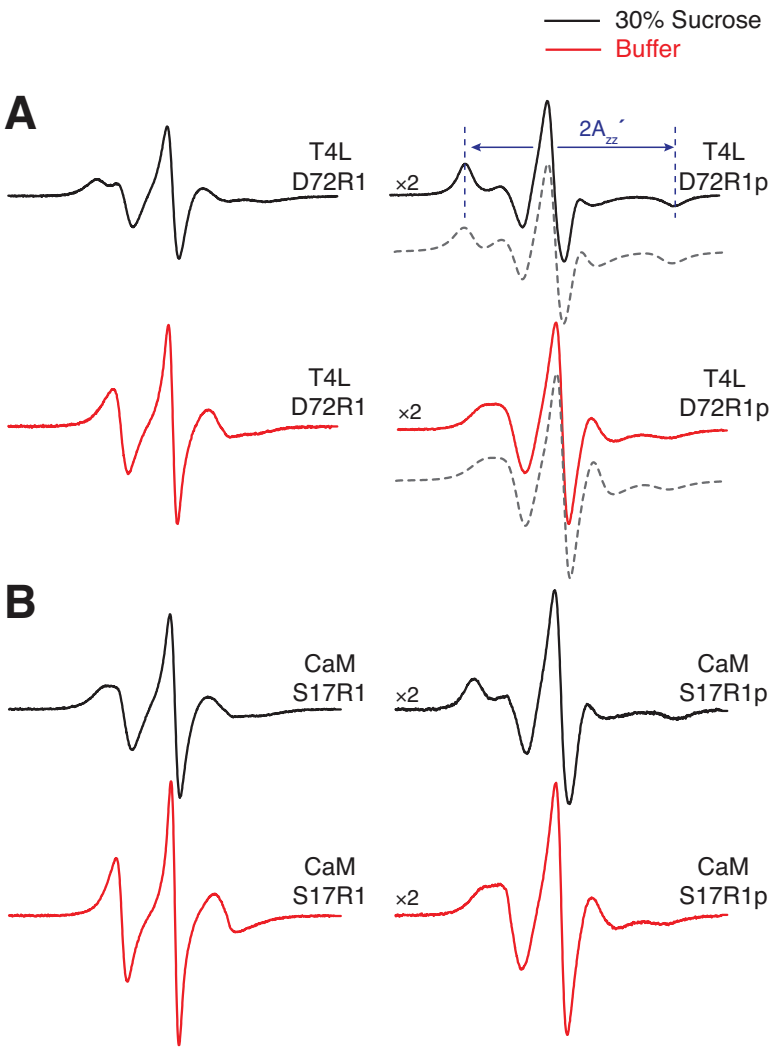

30% Sucrose
Buffer
A
T4L
D72R1
×2
2A_zz´
T4L
D72R1p
T4L
D72R1
×2
T4L
D72R1p
B
CaM
S17R1
×2
CaM
S17R1p
CaM
S17R1
×2
CaM
S17R1p

Fig. 2

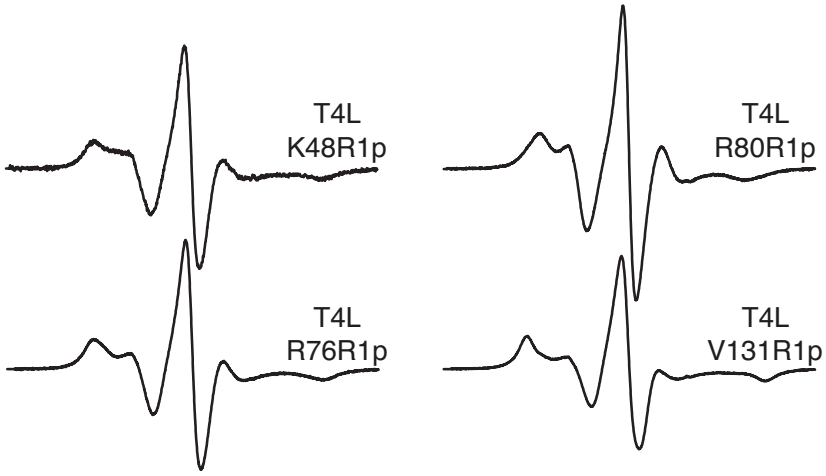


Fig. 3

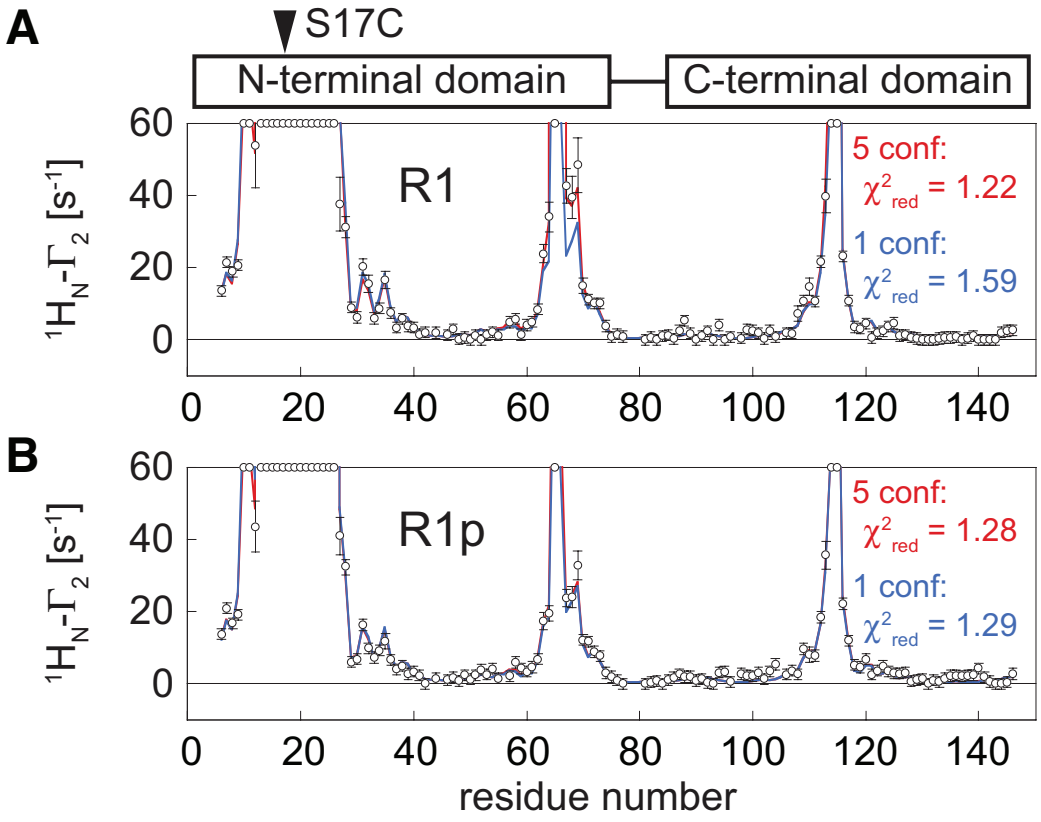

A
S17C
N-terminal domain
C-terminal domain
R1
5 conf:
χ^2_{red} = 1.22
1 conf:
χ^2_{red} = 1.59
^1H_N-Γ_2 [s^{-1}]
B
R1p
5 conf:
χ^2_{red} = 1.28
1 conf:
χ^2_{red} = 1.29
residue number

Fig. 4