

Conformational dynamics and distribution of nitroxide spin labels

Journal Article**Author(s):**

Jeschke, Gunnar

Publication date:

2013-07

Permanent link:

<https://doi.org/10.3929/ethz-a-010784332>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

Progress in Nuclear Magnetic Resonance Spectroscopy 72, <https://doi.org/10.1016/j.pnmrs.2013.03.001>

Journal: Progress in Nuclear Magnetic Resonance Spectroscopy

Conformational dynamics and distribution of nitroxide spin labels

Gunnar Jeschke

ETH Zürich, Lab. Phys. Chem., Wolfgang-Pauli-Strasse 10, 8093 Zürich, Switzerland

e-mail: gjeschke@ethz.ch

Tel.: +41 44 632 5702

Fax: +41 44 633 1448

Keywords:

paramagnetic relaxation enhancement; dynamic nuclear polarization; electron paramagnetic resonance; distance measurements; molecular dynamics

Journal: Progress in Nuclear Magnetic Resonance Spectroscopy

Number of pages: 53

Number of figures: 9

Number of tables: 3

Edited by GB and DGG

Abstract

Long-range distance measurements based on paramagnetic relaxation enhancement (PRE) in NMR, quantification of surface water dynamics near biomacromolecules by Overhauser dynamic nuclear polarization (DNP) and sensitivity enhancement by solid-state DNP all depend on introducing paramagnetic species into an otherwise diamagnetic NMR sample. The species can be introduced by site-directed spin labeling, which offers precise control for positioning the label in the sequence of a biopolymer. However, internal flexibility of the spin label gives rise to dynamic processes that potentially influence PRE and DNP behavior and leads to a spatial distribution of the electron spin even in solid samples. Internal dynamics of spin labels and their static conformational distributions have been studied mainly by electron paramagnetic resonance spectroscopy and molecular dynamics simulations, with a large body of results for the most widely applied methanethiosulfonate spin label MTSL. These results are critically discussed in a unifying picture based on rotameric states of the group that carries the spin label. Deficiencies in our current understanding of dynamics and conformations of spin labeled groups and of their influence on NMR observables are highlighted and directions for further research suggested.

Table of contents:

1. Introduction

2. Choice of spin labels and labeling sites

2.1 Labeling strategy and common spin labels

2.2 Choice of label and conformational variability

2.3 Choice of the labeling site

3. Label dynamics

3.1. Processes and interactions influencing NMR and EPR response

3.2 Molecular dynamics computations

3.3 Manifestation in continuous-wave EPR lineshapes

3.3.1 Simplified models of the dynamics of spin labels

3.3.2 Simulation of CW EPR spectra from MD trajectories

3.4 Dynamics of spin labels and relaxation

4. Conformational distribution

4.1 Conformational distribution in solids *versus* conformational dynamics in fluids

4.2 Rotamer library approach

4.3 Conformations observed in crystal structures

4.4 Effects on measurements of distance distributions

5. Conclusions and Perspectives

1. Introduction

Nitroxide free radicals [1] can be attached as spin labels [2] to proteins, nucleic acids, peptides, or synthetic macromolecules. Such labeling schemes allow the targeted introduction of an electron spin into otherwise diamagnetic systems. This electron spin can be observed either directly by electron paramagnetic resonance (EPR) or indirectly by NMR spectroscopy via its hyperfine interactions with nuclear spins. Since the magnetic moment of an electron spin is 658 times larger than the one of a proton, such experiments can provide access to longer distances than NMR experiments on systems that only contain nuclear spins, for instance by measuring the paramagnetic relaxation enhancement (PRE) for soluble proteins in liquid phase [3,4,5], including detergent-solubilized membrane proteins [6], and intrinsically disordered proteins [7]. PRE measurements have also been performed in the solid state [8]. Furthermore, the large magnetic moment leads to a proportionally large Boltzmann polarization of electron spin transitions compared to nuclear spin transitions. Polarization transfer induced by microwave irradiation and mediated by the hyperfine coupling, called dynamic nuclear polarization (DNP) [9], can be used to enhance sensitivity of NMR experiments [10,11]. In addition, the larger magnetic moment of the electron spin leads to higher frequencies at comparable magnetic fields and to larger anisotropies of the interaction of the electron spin with the magnetic field or with other spins than can be found in spin systems containing only nuclei. This in turn causes a strong sensitivity of paramagnetic systems to motions on time scales where NMR of diamagnetic systems is less sensitive. For instance, liquid-state DNP induced by the Overhauser effect is sensitive on the time scale of translational diffusion of water [12].

Experiments that depend on the introduction of electron spins into otherwise diamagnetic systems can strongly benefit from site-directed spin labeling (SDSL), where the nitroxide spin label is introduced at a specified position in a protein [13,14] or DNA molecule [15]. Early EPR

[16] and NMR [17] studies have indicated that perturbations of structure due to the label are usually small and rather local, a finding that has been largely confirmed by later work. Such SDSL techniques are widely applied in PRE studies. They have also been used in recent liquid-state DNP work [18,19], and may become attractive for solid-state DNP [20]. Once a protein or nucleic acid is spin labeled, further information can be obtained from CW EPR experiments [14] or distance measurements between spin labels by pulsed EPR techniques [21]. Constraints obtained from NMR and EPR experiments may complement each other in cases where constraints from NMR alone are insufficient for a full determination of the structure of a protein [22] or protein complex [23]. The spectroscopic signature in SDSL experiments is determined not only by inherent properties of the system under investigation, but also by specific properties of the spin label itself. In particular, the average site of the electron spin, which is approximately located at the center of the nitroxide N-O bond, is connected to the macromolecule by a linker that is generally flexible. Differences between the free energies of different linker conformations are usually smaller than the thermal energy. Hence, the site of the electron spin is distributed over space to an extent that depends on the type and conformational distribution of the linker. This distribution must be taken into account in the interpretation of distance measurements between spin labels in the solid state [24] and for a quantitative analysis of solid-state PRE measurements [25]. Furthermore, liquid-state PRE and DNP efficiency depend on the spectral density functions of dynamic processes that modulate the hyperfine couplings between the electron and nuclear spins. Conformational dynamics of the label contributes to this spectral density function and thus cannot be neglected.

Current knowledge of the conformational dynamics of spin labels originates mainly from continuous-wave (CW) EPR studies and molecular dynamics (MD) simulations that were aimed at explaining the observed spectral lineshapes. This article is devoted to a critical review of the results of such studies and an assessment of their importance for the interpretation of NMR experiments. Conformational dynamics of labels in liquid phase is linked to their static

distribution in solid phase. Direct information on preferred conformations can be obtained from diffraction of crystallized spin-labeled proteins. Indirect information can be derived from label-to-label distance distributions measured in proteins with known structures. Since such measurements are less demanding than protein crystallization, they have become very popular in the past decade, so that a large number of distances are known experimentally. The combined analysis of conformations of spin labels in protein crystals and label-to-label distances is the second main topic of this article. The majority of experimental results discussed in this review stems from EPR experiments. Care is taken to highlight the importance of these results for the interpretation and quantitative analysis of NMR experiments.

This review is structured as follows. Section 2 discusses how the distribution and dynamics of conformations can be influenced by the choice of the spin label and the labeling site. Section 3 is devoted to theoretical, computational, and experimental approaches to the dynamics of labels. First, motional processes and interactions between the macromolecule and the label are considered since they influence the spatial dynamics of the electron spin and determine the spectral response. Second, findings from MD simulations and limitations of this approach are discussed. Third, we examine how dynamics of the macromolecule and the label influence CW EPR lineshapes. Fourth, the relation between spatial dynamics of the electron spin and the relaxation times of the electron spin itself and of nearby nuclear spins is clarified and the implications for the interpretation of PRE and Overhauser DNP data are reviewed. Section 4 focuses on the distribution of conformations of spin labels in solid samples. The “rotamer library approach” is presented as a computationally inexpensive way of predicting the distribution of conformations of spin labels from a structural model of the unlabeled macromolecule, and alternative approaches are mentioned. The crystal structures of spin-labeled proteins tell us about preferences of the label for certain rotameric states and label-to-label distances in proteins with known structures can discriminate between different models of the distribution of rotameric states and different approaches for modeling conformational

distributions. The review ends with a summary of the most important findings and open questions.

2. Choice of spin labels and labeling sites

2.1 Labeling strategy and common spin labels

Spin labels can be incorporated at specific sites into macromolecules either during step-by-step synthesis of the macromolecule or by post-synthetic modification of specific residues. The former strategy is usually applied in solid-state [26] or solution-state [27] synthesis of peptides and for oligonucleotides [15] and allows a broad variation of labeling chemistry [28]. In this scenario, one often aims for a rigid coupling of the label to the backbone of the macromolecule and for the least possible conformational ambiguity of the label. For peptides, this can be achieved by incorporating the unnatural amino acid 2,2,6,6-tetramethylpiperidine-N-oxy-4-amino-4-carboxylic acid (TOAC, Fig. 1) [26,27]. While TOAC is a good substitute for the rare amino acid α -aminoisobutyric acid that features in antibiotic peptaibols [29], it is an achiral amino acid with a tetrasubstituted C^α atom and unusual preferences for a limited range of backbone dihedral angles. Thus, TOAC is liable to modify the secondary structure, for it is a helicogenic amino acid. Recently, the chiral unnatural amino acid 4-(3,3,5,5-tetramethyl-2,6-dioxo-4-oxypiperazin-1-yl)-L-phenylglycine (TOPP, Scheme1) was introduced to overcome this problem [30]. The TOPP label allows for rotations about the C^α - C^β bond; however, this bond is collinear with the nitroxide N-O bond and the distance between the C^α atom and the site of the electron spin lies within a narrow range.

The rigid and conformationally unambiguous DOXYL labels (Fig. 1) can be introduced during synthesis of lipids and steroids. Some members of this class of structures are commercially available.

Post-synthetic modification is the method of choice for proteins that are expressed in living cells. The standard route is generation of a mutant of the protein devoid of any of the native cysteine residues, and the introduction of one or several cysteine residues by directed point mutations [13,14]. Sometimes buried native cysteine residues need not be removed since, unlike the intentionally introduced surface cysteines, they react with labels much more slowly or not at all. Nitroxide labels suited for this approach carry a thiol-reactive group, such as a maleimido group [2], the methanethiosulfonate group [31] (MTSL, Fig. 1), or the iodoacetamido group (IA-PROXYL, Fig. 1). By a wide margin, MTSL is the most widely used cysteine-specific spin label, mainly for its high selectivity to thiol groups, high reactivity under mild conditions, and small propensity to distort the protein structure. The diamagnetic analogue dMTSL [32] can be used to check the structural integrity, for instance by testing for chemical shift changes. Furthermore, dMTSL can be used to obtain diamagnetic reference values for transverse relaxation times T_2 in PRE measurements [3]. The main disadvantage of MTSL is the labile character of the disulfide bond that is formed on attachment to the thiol group of cysteine. If a protein requires slightly reducing conditions, or if a labeled protein is to be studied in a living cell, maleimido [33] or iodoacetamido labels are preferable. Note that under more strongly reducing conditions, nitroxides are converted to hydroxylamines, a reaction that proceeds in living cells on a time scale of a few minutes to a few hours, depending on the type of nitroxides. If a mutant without cysteines does not fold or is not functional, nitroxide labels can also be introduced after protein synthesis if an unnatural amino acid is introduced that is suitable for labeling [34].

2.2 Choice of label and conformational variability

It might appear that complications due to internal dynamics and conformational distributions of spin labels are best avoided by selecting a rigid, conformationally unambiguous label. This would call for a short linker, ideally without any bonds that can undergo rotations.

Unfortunately, such labels also have the highest propensity for distorting the structure and function of the biomacromolecule under investigation. Furthermore, shortening the linker requires compromises on the selectivity or reactivity of the thiol-reactive group, which make such labels less versatile than MTSL. Research on restricting the mobility of spin labels has therefore focused on modifications of MTSL. Based on the hypothesis that the dynamics of MTSL attached to surface-exposed labeling sites in α -helices is strongly dominated by fluctuations in the dihedral angles χ_4 and χ_5 (Fig. 2), substituents were introduced that were expected to reduce the variability of χ_5 [35]. Indeed this led to a slowdown of motions of spin labels, as manifested in CW EPR spectra. Examples of commercially available spin labels with reduced mobility of this type are Br-MTSL (Fig. 1) and the analogous CH₃-MTSL. The same strategy using a pyridine group instead of a methyl or bromide group was applied to simplify the interpretation of PRE data [36].

The mobility of spin labels can also be strongly reduced by two-point attachment to the protein, as is possible with the bisfunctionalized MTSL analogue bisMTSL (Fig. 1). This label can assume energetically relaxed conformations when two cysteine sites are introduced at residues i and $i+3$ or $i+4$ in an α -helix, at positions i and $i+1$ in a β -strand or at the nearest position in an adjacent β -strand [37]. Labeling with bisMTSL can potentially result in the misleading attachment of two distinct spin labels to two adjacent reactive thiol groups in a side reaction, which does however appear to be much slower.

Despite the commercial availability of some conformationally restricted MTSL analogues over the last few years, the vast majority of SDSL studies are still performed with MTSL. While this requires that substantial linker dynamics and conformational distributions are taken into account (*vide infra*), it also decreases the probability of producing dysfunctional or structurally distorted labeled proteins. More often than not, generating mutants and expressing, purifying and labeling the proteins are the most tedious parts of the study. In this situation it makes

sense to opt for the label that is likely to introduce the least complications in these steps.

Hence, understanding conformational preferences and dynamics of MTSL is likely to remain a key to the quantitative analysis of SDSL by NMR and EPR in many future studies.

2.3 Choice of the labeling site

First and foremost, labeling should not significantly alter the structure or impair the function of the protein. Known functional residues and binding sites should not be labeled. Structural distortions are more likely when mutating residues that can potentially bind metals, such as histidine, or charged residues. Conservative labeling thus focuses on residues with non-coordinating, uncharged, hydrophobic side groups. The MTSL side group is slightly more apolar than leucine, and its hydrophobicity lies between methionine and tryptophan [38]. Very conservative labeling schemes should also avoid mutations of residues with aromatic side groups, which might be involved in π stacking or π -cation interactions. When possible, mutations of Gly and Pro residues should also be avoided, since their preferences for backbone dihedral angles, as expressed in terms of Ramachandran plots, differ from other native residues and MTSL. The scan of conservative sites implemented in the MMM software [39,40,41] considers Cys, Ile, Leu, Met, Ser, Thr, and Val residues as appropriate labeling sites. When conformational changes that may be relevant for protein function are studied with SDSL techniques, it is good practice to perform functional assays with both cysteine mutants and spin labeled proteins. In NMR studies, agreement of key 2D spectra between the wild-type and dMTSL-labeled protein can be used to confirm the structural integrity of the labeled protein.

Second, a high labeling yield is desired. This point is of particular importance for PRE measurements, where quantitative analysis can fail if labeling is incomplete, and for distance measurements with CW EPR techniques [42], where the presence of a fraction of unlabeled proteins increases the uncertainty of distances, in particular towards the upper end of the range of distances. If a protein is labeled in its folded state, high labeling yield can be expected

only for surface-exposed residues, although there is no guarantee that buried cysteines are fully protected against spin labeling. In general, labeling of buried residues should be avoided, as such labeling is more likely to distort the structure or modify protein dynamics than labeling of surface-exposed residues.

Even at surface-exposed sites, the labeling yield may be lowered by steric hindrance from neighboring side groups. In MMM such a hindrance is indicated by a small value of the partition function and a small number of allowed rotamers [40,41], as will be further explained in Section 4.1. The 'Select site pairs for fitting' function in the EPR menu of MMM sorts possible labeling sites into favorable and unfavorable sites based on the number of allowed rotamers and gives a warning if a potential labeling site is a known binding site. Fig. 3 shows the distribution of spin label conformations (using transparent stick models) and midpoint positions of N-O bonds (the blue and red spheres have radii proportional to the populations) for the largely unrestricted surface sites at residues 72 and 131 of T4 lysozyme as predicted by MMM.

Third, depending on the experiment to be performed, one may be interested in a narrow spatial distribution and in restricted conformational dynamics of the label. For a given label, this can be achieved by selecting sites with moderate sterical hindrance by neighboring side groups. Note however that, at any currently accessible level of theory and approximation, the prediction of the spatial distribution and conformational dynamics is fraught with significant errors (see Sections 3 and 4). In this situation, predictions are most accurate at sites without sterical hindrance. In particular, when studying conformational changes of the protein backbone with EPR [14,21] or PRE [43] measurements, interpretation of the data may be compromised if the conformational distribution of the labeled side group strongly depends on the conformation of the protein backbone. This is much more likely for narrow than for broad

conformational distributions. In the case of PRE measurements, the problem can be detected by careful analysis of the local PRE pattern [43].

For PRE studies on α -helical membrane proteins, it has been suggested that one spin label per α -helix is sufficient to generate a sufficient number of long-range constraints to determine structures with an accuracy of about 1.5 Å [44].

3. Label dynamics

3.1. Processes and interactions influencing NMR and EPR responses

NMR and EPR spectra are influenced by the orientation of the molecules relative to the magnetic field and by distances between spins. Likewise, relaxation times are influenced by fluctuations of the orientation and spin-spin distances on time scales close to the inverse of the resonance frequency of the observed spin (for both longitudinal and transverse relaxation) or at frequencies close to zero (only for transverse relaxation). The orientation of the whole protein in solution changes with global tumbling, which for soluble proteins is often approximated by isotropic Brownian rotational diffusion. Assuming the Stokes-Einstein law and neglecting the hydration shell, one expects the rotational correlation time of a globular protein in water at 20°C in nanoseconds to be 0.3 times the molecular weight in kDa. Actual rotational correlation times $\tau_{r,prot}$ are about twice as long. At typical molecular weights of proteins accessible by NMR techniques, one has $\tau_{r,prot}$ between 1 and 15 ns, which lies in the relevant range for inducing both longitudinal and transverse relaxation. Usually, only global tumbling of the protein is considered in PRE computations for nitroxide spin labels [5], although internal dynamics of spin-labeled side groups occurs on the same time scale (Fig. 4). For detergent-solubilized membrane proteins the rotational correlation time for global tumbling depends on the radius of the micelles. For membrane proteins reconstituted in liposomes, global tumbling is much slower than all other processes discussed below.

Proteins and, in some cases, nucleic acids, do not move as rigid bodies. Local backbone motions at the spin-labeled site can also lead to a reorientation of the molecular frame of the spin label, and affect spin-spin vectors and spin-spin distances. Small-amplitude backbone librations occur with characteristic times between 2 and 10 ps. However, the distribution of local backbone relaxation times is rather broad and may extend into the nanosecond range due to collective segmental motions [45]. Large-scale conformational changes of the backbone occur on microsecond to millisecond time scales. While global tumbling of proteins can be characterized more easily with other techniques, SDSL approaches are of interest to characterize backbone motion.

The characterization of backbone motions is complicated by internal motions of the spin-labeled side groups, which influences the same magnetic parameters and occurs on similar time scales (Fig. 4). Vibrational modes, corresponding to oscillations of bond lengths and bond angles, are fast small-scale motions that do not significantly affect EPR and NMR experiments. Vibrational averaging may influence spin Hamiltonian parameters, such as the g tensor or hyperfine tensor, which can be accounted for by proper averaging of the parameters along classical or *ab initio* MD trajectories [46]. The main influence on CW EPR spectra and on electron as well as nuclear spin relaxation times originates from variations of dihedral angles. For the spin label MTSL, the relevant angles χ_1 to χ_5 are shown in Fig. 2.

Rotations about dihedral angles are characterized by potentials with two or three minima. If only the four atoms required to define a dihedral angle are considered, the potentials for angles χ_1 , χ_2 , and χ_4 have energy minima at -60° (m or g^-), 60° (p or g^+), and 180° (t), while the potentials for χ_3 and χ_5 have minima at -90° (m) and $+90^\circ$ (p). Combinations of these values lead to 108 canonical rotamers. Distributions of dihedral angles observed in MD computations for MTSL attached to the central cysteine residue of a pentadecaglycine helix in explicit water indeed roughly agree with the expected minima, with small deviations of the optimum

dihedral angles resulting from steric hindrance by atoms that are remote from the considered dihedral angle [47]. Corresponding plots for MTSL attached to a single cysteine residue, based on data from [41], are shown in Fig. 5.

Internal motions of the label on time scales relevant for NMR and EPR spectroscopy can thus be understood in terms of moderate fluctuations of dihedral angles for a given rotameric state (libration) on the one hand, and in terms of transitions between different rotameric states on the other hand. As can be seen in Fig. 5, librations have typical amplitudes of 10-30°, although excursions up to 45° are observed [48]. These librations occur on a 1-100 ps time scale.

Transitions between rotameric states happen on time scales between about 1 ns and 100 ns, depending on which dihedral angles change from one canonical value to another. Transitions between the *p* and *m* states of χ_3 are particularly slow and are observed only rarely in MD trajectories up to 100 ns at ambient temperature [47-49]. Sampling problems were also encountered in an MD study where distance distributions between spin labels were directly taken from MD trajectories and compared to experimental distributions [50]. For some of the trajectories, good agreement between simulation and experiment was observed, while other trajectories were discarded because the agreement was substantially worse. The enhanced 'umbrella' sampling method was used to estimate the *m:p* ratio for χ_3 with an uncertainty of about 4% in MD computations for residues in T4 lysozyme labeled with MTSL [49]. The ratio was found to be 27:73 for residue 72, and 55:45 for residue 131, indicating that subtle differences in the environment at these two surface-exposed helical sites can shift this ratio significantly. Note however that for otherwise spatially unrestricted labels, the *p* and *m* groups of χ_3 rotamers sample very similar parts of space (Fig. 3).

In MD simulations using current standard force fields, transitions between different states of χ_1 and χ_2 are somewhat slower (a few nanoseconds to tens of nanoseconds [49]) than those between different states of χ_4 and χ_5 . The frequency of transitions between χ_5 states depends

on the other dihedral angles, in particular on χ_4 , and can be very high for some conformations due to an almost flat χ_5 potential [51]. Such a coupling between the dynamics of χ_4 and χ_5 dihedral angles is included in the CHARMM parametrization of MTSL by CMAP correction [52]. The weak restriction for χ_5 in MTSL explains why substitutions such as CH₃-MTSL and Br-MTSL have such a strong influence of CW EPR line shapes.

Transitions between rotameric states correspond to jumps of the mean position of the electron spin between the blue and red spheres in Fig. 3. Jumps between spheres of the same color are much more frequent. Obviously, such transitions can lead to substantial changes in the orientation and length of the vector between the electron and nuclear spins for nuclei in the sensitive range of PRE measurements (up to 25 Å for nitroxides). Collisions with water molecules can occur in the new rotameric state while they were remote from the electron spin in the old rotameric state. Thus, neglecting rotamer transitions in quantitative computations of PRE or liquid-state Overhauser DNP may not be allowed in general. An overview of the relevant time scales for protein, internal label, and solvent motions is given in Fig. 4.

3.2 Molecular dynamics computations

The estimates of rotamer transition rates given in Section 3.1 and the distributions of dihedral angles shown in Fig. 5 are based on MD computations. As such data cannot be directly obtained by experiment, all assumptions involved in these computations have to be considered carefully before conclusions can be drawn. First, transitions between the *p* and *m* states of χ_3 are too infrequent in MD simulations at ambient temperature to obtain good statistics [49,52]. Frequent transitions are observed at a simulation temperature of 600 K [47]; however the populations of the rotameric states also depend on the temperature. Since each rotameric state is composed of a multitude of microstates, it is not in general permissible to compute relative rotamer energies from populations at one temperature and recompute populations at another temperature from the Boltzmann distribution [41]. In particular,

relative energies of canonical rotamers at zero temperature, as they can be obtained from *ab initio* quantum chemical computations, are not sufficient for predicting rotamer populations at higher temperatures (see Section 4.4).

The sampling problem implies that rotamer populations derived for MTSL from most currently available MD trajectories cannot be considered to be quantitatively correct with respect to χ_3 , and possibly not even with respect to the other dihedral angles. General use of enhanced sampling techniques may be required to overcome this problem. In a recent study, rotamer distributions obtained from enhanced sampling MD were found to be in excellent agreement with those reconstructed from NMR-derived order parameters for Ile, Val, Leu, and Met side groups [53].

The accuracy of the molecular force fields underlying MD computations is also a limiting factor. An early study that compared enhanced (umbrella) sampling MD with NMR-derived correlation times and with 3J coupling constants based on the Karplus relationship found only a qualitative agreement, with 3J coupling constant predictions from MD being poor for two out of the four Phe residues in antamanide [54]. More recent work with local elevation-biased enhanced sampling showed that experimental 3J coupling constants can be used to detect deficiencies in force field parameters [55]. Note however that the Karplus relationship was found to be only a rough approximation for the dependence of $^3J_{\alpha\beta}$ on the dihedral angle χ_1 [56]. Conclusions about the internal dynamics of spin labels were also found to depend on the force field [57].

The accuracy of force fields may be compromised not only because of insufficient parameter quality, but also because of their functional form that may exclude some interactions which significantly influence rotamer populations and transition rates. As will be discussed in Section 4.2, crystal structures of MTSL attached to surface-exposed helical sites [58,59] suggest a significant interaction between the S^δ atom (labeled S1 in Fig. 2) and the C^α atom of the same

residue, possibly mediated by the C^α hydrogen. Such interactions are expected to influence the distributions of χ_1 and χ_2 and to lower their transition rates. However, no such interactions were apparent in standard *ab initio* quantum chemical parametrization of CHARMM force field parameters for MTSL [52]. As remarked upon earlier [47], sulfur atoms are highly polarizable, so that spherically symmetric Lennard-Jones potentials in standard force fields may not be sufficient to model van-der-Waals interactions, a point that has been reiterated recently [60]. This would suggest the use of polarizable force fields, such as AMOEBA [61]. Note however that more elaborate computations and lower transition rates for χ_1 and χ_2 rotamers would aggravate the sampling problem.

Another problem of MD simulations arises from the fact that current water models in molecular force fields do not reproduce water viscosity. The most commonly used TIP3P water model underestimates viscosity by a factor of 2.87 [62]. This is expected to lead to a scaling of the diffusion coefficients of small solutes and of molecular moieties by about the same factor. This was approximately corrected for in some simulations by scaling the MD time step by a factor of 2.5 [63]. Note that such correction does not fully recover the true molecular relaxation times, since not all dynamic processes are slowed down in proportion to the viscosity of water.

The possible deficiencies of MD simulations of spin label dynamics raise the question of experimental tests. CW EPR spectra of nitroxide spin labels are sensitive to motions on the time scales of greatest interest that concern collective backbone motions, transitions between rotameric states, and global tumbling of the biomacromolecule. This dependence and ways for simulating such spectra are described in the next Section.

3.3 Manifestation in continuous-wave EPR lineshapes

3.3.1 Simplified models of the dynamics of spin labels

Lineshapes in CW EPR spectra are affected by the reorientation of nitroxide labels with respect to the magnetic field because of two anisotropic interactions, the field-independent anisotropy of the ^{14}N hyperfine coupling of the N-O group, and the field-dependent anisotropy of the electron Zeeman interaction described by g anisotropy. The hyperfine tensor has approximately axial symmetry with typical principal values $(A_x, A_y, A_z) \approx (0.6, 0.6, 3.3) \text{ mT} \approx (17, 17, 92) \text{ MHz}$, whereas the g tensor is non-axial with an asymmetry parameter $\eta \approx 1$ and typical principal values $(g_x, g_y, g_z) \approx (2.0090, 2.0059, 2.0027)$. For $(g_x - g_z) \mu_B B_0/h < 2A_z$ the total width of the spectrum is determined by the hyperfine splitting between the $m_l = +1$ and $m_l = -1$ manifolds of the electron spin levels, which applies at fields below $B_0 \approx 1.94 \text{ T}$. This is the case at X-band frequencies around 9.6 GHz ($B_0 \approx 0.34 \text{ T}$) and at Q-band frequencies around 35 GHz ($B_0 \approx 1.1 \text{ T}$). In particular, at X-band frequencies, where most studies were done, hyperfine anisotropy dominates, although g anisotropy is not negligible. At Q-band frequencies, the two anisotropies are comparable, whereas at W-band frequencies of about 94 GHz (3.55 T, ^1H frequencies around 140 MHz) or higher, the g anisotropy dominates.

The CW EPR lineshape can be understood as arising from multi-site exchange between different orientations. Since up to Q band frequencies, the anisotropy of the resonance frequencies is approximately $\Delta\nu \approx 140 \text{ MHz}$, the spectra are most sensitive to motions with rotational correlation times $\tau_r \approx (2)^{1/2}/(\pi\Delta\nu) \approx 3 \text{ ns}$. At a frequency of 263 GHz ($B_0 \approx 9.4 \text{ T}$, 400 MHz ^1H NMR frequency), where $\Delta\nu \approx 842 \text{ MHz}$ is dominated by g anisotropy, the rotational correlation time that has the most pronounced effect on CW EPR spectra is reduced to $\tau_r \approx 0.53 \text{ ns}$. The sensitivity to global tumbling and to slow rotamer exchange is thus reduced at high frequencies, whereas the sensitivity to libration increases. For rotational correlation times larger than about 100 ns, a powder CW EPR lineshape with a slightly reduced total width is observed. For rotational correlation times larger than about 1 μs the CW EPR lineshape is insensitive to motions. Such slower motions can still be observed by CW EPR when the

electron spin transitions are saturated and the signal is detected 90° out-of-phase at the second harmonic of the modulation frequency with large modulation amplitude (0.5 mT) . This saturation-transfer EPR technique is sensitive to motions on time scales up to the longitudinal relaxation time of the electron spins, which can exceed 1 μ s [64]. If exchange between groups of rotameric states is slow on the CW EPR time scale, whereas exchange between states within each group is faster than or matches the EPR time scale, the EPR spectrum shows a weighted superposition of the spectra of individual groups of rotamers. Each group of rotamers corresponds to a basin on the potential energy hypersurface, which is explored on the EPR time scale, while transitions between basins are negligible. The weighting factors are the populations in each basin.

The overlap of time scales of slow rotamer exchange and of collective segmental backbone motions (Fig. 5) creates difficulties in reliably detecting the latter type of motions with SDSL approaches. It has been argued that the two types of motion can be distinguished by measuring the CW EPR spectrum as a function of pressure up to very high pressures around 4000 bar [65]. Pressure leads to changes of the conformational equilibrium constant K depending on the molar volumes of the conformations. If compressibility is the same for the two states, as would be expected for two different rotameric states of a label, the dependence of $\ln[K(p)/K(0)]$ on pressure p should be linear. This was indeed observed for MTSL for the surface-exposed residue 44 in T4 lysozyme, whereas a non-linear dependence was found for the buried residues 46 and 118 [65].

Computation of CW EPR lineshapes for exchange between multiple orientations can be performed most efficiently by numerically solving the stochastic Liouville equation (SLE) with the superoperator $\Gamma + i \mathcal{L}$, where Γ is a diffusion superoperator and \mathcal{L} the commutator superoperator of the spin Hamiltonian [66,67]. The solution involves inversion of large matrices, which is efficiently done by combining the Lanczos algorithm for tridiagonalization

with a conjugate gradients algorithm. The Liouville superoperator matrix \mathcal{L} for the spin Hamiltonian is set up in terms of Wigner rotation matrices, such that the integral over all relative orientations Ω_{LG} between the laboratory frame L and the g tensor frame G of the radical (Fig. 6) can be expressed in terms of Wigner 3-J symbols [66]. We shall use the notation of Liang and co-workers [68]. Unless relaxation broadens the lines, the CW EPR lineshape depends only the principal values of the g - and hyperfine tensors and the time-dependent Euler angles $\Omega_{LG}(t) = (\alpha_{LG}(t), \beta_{LG}(t), \gamma_{LG}(t))$. To relate this transformation to a model of molecular dynamics, it can be decomposed into a series of transformations [68]. Depending on the symmetry of individual dynamic processes, only two instead of three Euler angles may be required for certain transformations Ω_{ij} .

In the simplest case the label is rigidly attached to a globular molecule that tumbles isotropically. In this case, no subtransformations are required and the stochastic dynamics $\Omega_{LG}(t)$ can be described as isotropic Brownian rotational diffusion with a single rotational correlation time τ_r . If the label is rigidly attached to a prolate (rodlike) or oblate molecule, Brownian rotational diffusion becomes anisotropic. Usually it is sufficient to consider an axially symmetric rotational diffusion tensor with eigenvalues $R_{||}^\circ$ for rotations about the symmetry axis of the molecule, and R_{\perp}° for rotations about perpendicular axes. Since the symmetry axis of the whole molecule usually does not coincide with the z_G axis of the g tensor, a molecular diffusion frame (M frame) needs to be defined (Fig. 6). Note that the M frame is fixed with respect to the nitroxide moiety. Dynamics is described by anisotropic Brownian rotational diffusion encoded in $\Omega_{LM}(t)$. A time-independent transformation with Euler angles Ω_{MG} relates the M frame to the G frame.

The next step of complexity arises when an orienting potential is introduced, which defines a C' frame where the $z_{C'}$ axis lies along the director of the potential. Such an orienting potential can be related to the laboratory frame by a time-independent transformation Ω_{LC} if the

orientation is induced by the magnetic field or if the reorientation of the director of the local orienting potential due to global tumbling is slow on the EPR time scale. The latter situation is assumed in the model of microscopic order and macroscopic disorder (MOMD) [67], where the director orientation varies across the sample but can be considered to be fixed in a local environment on time scales up to at least 1 μ s. The orienting potential is expanded into spherical harmonics [67]. In early work on spin-labeled proteins it was assumed that this applies even for small proteins at X-band frequencies in highly viscous solutions prepared by adding 30% sucrose to water, which results in rotational diffusion rates of about $6 \cdot 10^6$ s⁻¹ for a protein of the size of T4 lysozyme (18.7 kDa). Based on later theoretical work it was suggested that such rotational diffusion rates, corresponding to $\tau_r = 28$ ns, are not sufficiently slow for applications of the MOMD model at X-band frequencies [63]. Note however that for residue 131 in T4 lysozyme at a largely unrestricted α -helical surface site, CW EPR spectra obtained in 30% sucrose solution and in a suspension of microcrystals are virtually identical [59], suggesting that global tumbling of the protein in the sucrose solution is negligible on the time scale of internal motions of the spin label. For large proteins in viscous solution or for membrane proteins in liposomes, the MOMD model is expected to be applicable in any case. The MOMD model is also applicable to DOXYL-labeled lipid molecules in lipid bilayers [69]. In the MOMD model, the local diffusion frame M of the nitroxide may move with respect to this C' frame, necessitating a time-dependent transformation Ω_{CM} . In the context of spin-labeled proteins, the C' frame is fixed with respect to the protein (Fig. 6) and is related to the potential that the spin label experiences from neighboring side groups, the backbone, and the mean linker direction.

If global tumbling of the protein cannot be neglected, the description must be extended to the model of slowly relaxing local structures (SRLS), where the protein and spin label are treated as two coupled rotators [70,71]. This necessitates the introduction of a global diffusion frame of

the protein (C frame in Fig. 6). The time-dependent transformation Ω_{LC} from the laboratory frame to the C frame describes global tumbling of the protein, which is usually treated as isotropic Brownian rotational diffusion. If protein tumbling is anisotropic, another time-independent transformation $\Omega_{CC'}$ between the protein-fixed global diffusion frame C and the protein-fixed frame C' of the orienting potential needs to be introduced. In the SMLS model, Ω_{CM} describes coupling between motions of the protein and the label.

Programs for simulating and fitting spectra with the MOMD [67] and SMLS [72] models and simpler models are available on the ACERT homepage [73]. The MOMD model and the simpler models are also supported by the EasySpin simulation toolbox [74].

Good simultaneous fits were obtained with the SMLS model for CW EPR spectra measured in aqueous solution in the temperature range between 2 and 37.5°C on T4 lysozyme labeled at residues 72 and 131 [68]. In both cases, two-component models were required at lower temperatures, i.e., there are at least two subpopulations in slow exchange on the EPR time scale. For residue 131 a single component was sufficient at 37.5°C, and for residue 72 at both 22 and 37.5°C. Although the dynamic and magnetic parameters were fitted simultaneously for both frequencies, the populations of the two components at lower temperatures were allowed to differ between the 9 and 250 GHz spectra, implying that a two-component SMLS model does not fully account for the complexity of the dynamics of spin labels at these surface-exposed sites. Furthermore, in free fits different global tumbling rates of the protein were found for the two labeling sites, although fits that constrained the global tumbling rate to be the same for both sites were still good and the other dynamics parameters did not change significantly when τ_r was changed.

In later work the same two mutants of T4 lysozyme were studied at four EPR frequencies (9, 95, 170, and 240 GHz) in the presence and absence of the high molecular weight sucrose polymer Ficoll as a viscosity-enhancing agent [75]. When using four frequencies, spectra at 2

and 12°C could be fitted for residue 72 only by assuming three components, one of which was immobile and is clearly revealed only at the three higher frequencies. Introducing the third component improved the agreement of the global tumbling rates, although some difference between MTSL at sites 72 and 131 persisted. Likewise, for residue 131 three components were now required for fitting. The orienting potential, quantified by an order parameter S_{20} was found to be larger for residue 72 than for residue 131, which was attributed to a larger backbone flexibility of residue 131. In this study, CH₃-MTSL was also included. For this label, order parameters were similar at the two sites. At both sites, the number of components could be reduced to two. After addition of Ficoll, slightly larger order parameters and a slow-down of local diffusion of the spin labels were observed in some cases. Altogether the study using MTSL suggested an immobile group of rotamers interacting with the protein surface and two distinct groups of more or less mobile rotamers with relatively low ordering.

3.3.2 Simulation of CW EPR spectra from MD trajectories

The most efficient way of simulating a CW EPR spectrum from a trajectory $\Omega_{LG}(t)$ that describes the dynamics of the g -tensor frame with respect to the laboratory frame consists of the computation of the free induction decay (FID), subsequent Fourier transformation, and pseudomodulation [76] to compute the derivative-like lineshapes. The first implementation sampled a Brownian dynamics trajectory at intervals of 0.3 ns to provide a Nyquist band corresponding to a 100 mT field sweep, assuming a trajectory length of 1 μ s with 2^{14} data points to obtain 30 μ T resolution [77]. At that time, MD trajectories at atomic resolution of such lengths could not be obtained. At the time of writing this review they can only be obtained with dedicated massively parallel computers [78]. It was realized early that shorter MD trajectories at atomic resolution can be used to construct an ordering potential, and to compute longer trajectories by Brownian dynamics simulations based on such ordering potentials [79]. Initially a 6 ns MD trajectory was considered sufficient and a potential in terms

of the Euler angles $\Omega_{LG} = (\alpha_{LG}, \beta_{LG}, \gamma_{LG})$ was constructed with 6° angular resolution. Recently it was recognized that the required MD trajectory length depends on decay of the re-orientational auto-correlation function of the spin label [48]. From a 40 ns trajectory, effective auto-correlation times of about 0.7 ns were derived. Note however that the sampling problem of MD still persists: MD trajectories extending over a few tens of nanoseconds cannot be expected to properly describe relative populations of rotamer subgroups in slow exchange with each other. However, if the relative populations in the individual basins of the energy hypersurface were known, it would be sufficient to run individual MD trajectories in each basin with a length of 5-50 ns, depending on intrabasin auto-correlation functions. Assuming that exchange between basins is slow on the EPR time scale, a weighted superposition of the spectra of all basins would then provide a good approximation of the total spectrum. The spectra for individual basins should correspond to the dynamic components found in SRLS simulations. This simple idea was formalized by constructing Markov models (*vide infra*) [49,80].

Two further problems arise. First, atomic coordinates in MD simulations are typically propagated at intervals of 1 fs and sampled at intervals of 1 ps. The latter interval would still lead to unnecessary computational expense for FID simulations, yet sampling at intervals of 0.3 ns may be too slow compared to the auto-correlation times of re-orientation. In particular, resonances between fluctuations of Ω_{LG} and the electron spin Larmor frequency may cause longitudinal relaxation. However, it has been argued that transverse relaxation is much faster than this contribution to longitudinal relaxation. Hence, the contribution of longitudinal relaxation to the linewidths is negligible and the prediction of the FID's should be good when MD trajectories are averaged over intervals of 1 ns at X band ($B_0 = 0.34$ T) and shorter intervals at higher frequencies, for instance 1/8 ns at $B_0 = 8.92$ T [63].

Second, global tumbling of the protein may have a slower auto-correlation function than intrabasin spin label dynamics. It would be wasteful to run long MD trajectories at atomic resolution merely to obtain estimates for the rotational diffusion tensor that describes global tumbling, in particular given the fact that poor reproduction of solvent viscosity in current water models compromises the quality of such estimates. Hence, it is common practice to remove global tumbling of the protein from the MD trajectories and to reintroduce it in FID computations by a Brownian dynamics approach [81]. An efficient numerical integrator for such Brownian dynamics, which also allows for simulating anisotropic Brownian diffusion in an external potential, has been developed [63].

Above it was suggested to identify slowly exchanging subgroups of rotamers and to construct the overall spectrum from weighted spectra of the subgroups. This simple idea runs into difficulties, since the time scale of transitions between rotamer subgroups covers a broad range that overlaps with the time scale where the EPR lineshape is sensitive to spin label dynamics. In MD trajectories, transitions occur stochastically. Hence, an individual trajectory may sample several groups of rotamers, even if transitions between those groups are slow on the EPR time scale. Still it is attractive to describe spin label dynamics in terms of transitions between macrostates, where each macrostate consists of several conformations (microstates). Such a concept reduces the complex dynamics of the MD trajectory to a simpler model that is easier to interpret and can be related to the rotameric states. A systematic identification of the macrostates is possible in the framework of Markov models, assuming that the spin label does not have any conformational memory on the time scale required for simulating EPR spectra [80]. Given the macrostates, their equilibrium populations, and a state-to-state transition probability matrix, arbitrarily long stochastic trajectories can be simulated with relatively little computational effort.

Construction of the macrostates takes into account that, for a Markov model to be applicable, their internal structure must relax on a time scale shorter than the interval τ between two successive observations [80]. This requires identification of groups of microstates with a limited spatial range so that these groups can be lumped together into a single macrostate. Proper accounting for probability densities introduces problems that can be solved using a hidden Markov model, where a vector of order parameters is observed instead of the states of the Markov chain. These order parameters can be identified either with the dihedral angles χ_1 - χ_5 (for MTSL, see Fig. 2) or with the Euler angles $\Omega_{LG} = (\alpha_{LG}, \beta_{LG}, \gamma_{LG})$ [80]. The latter choice is related to the derivation of a potential energy function from the MD trajectory in terms of Euler angles [79]. Starting from the 108 canonical rotamers of MTSL and 12 random non-rotameric conformations, the model was simplified by lumping states. A model with 14 rotameric states was found to capture most of the dynamics and to provide reasonable simulations of multi-frequency EPR spectra, although the quality of the simulations increased significantly upon extending the model to 23 states [80].

Later work using this Markov model of MTSL attached to a polyalanine helix provided a set of the 18 most strongly populated rotamers and their populations, taking into account the uncertainty of the χ_3 distribution due to insufficient sampling [52]. The same approach was then applied to residues 72 and 131 in T4 lysozyme [49]. At site 131, spectra simulated for the p and m states of χ_3 were nearly identical and a good agreement of spectra obtained at 22°C and $B_0 = 0.33, 3.4$ and 6.1 T was achieved irrespective of the sampling problem. In contrast, at site 72 the spectra for the p and m states of χ_3 were significantly different. In particular at the two higher fields, a mixture of 27% m and 73% p states, as suggested by enhanced (umbrella) sampling, provided better agreement than the p component alone, which gave a good fit only for the spectrum at $B_0 = 0.33$ T.

3.4 Dynamics of spin labels and relaxation

Internal conformational dynamics of the spin label can induce longitudinal relaxation of the electron spin if the time scale matches the electron Larmor frequency. This happens for correlation times between 20 ps (X band frequencies) and 0.6 ps (263 GHz corresponding to 400 MHz ^1H frequency), which are encountered for libration in surface-exposed spin labels. Internal dynamics of spin labels can also cause PRE of nearby nuclear spins, both directly and indirectly. The direct effect arises from modulations of the amplitude of the dipole-dipole contribution to the hyperfine interaction as well as from reorientation of the electron spin-nuclear spin vector. The indirect effect arises from the influence of the dynamics of the spin label on the longitudinal relaxation of the electron spin, since longitudinal electron relaxation also modulates the hyperfine field at the nucleus. For nitroxide spin labels the indirect effect is usually negligible, since typical electron T_1 values are much longer than the inverse Larmor frequency of the nuclear spins [5]. The direct effect is also generally neglected in PRE studies, where global protein tumbling is considered to make the only relevant contribution, albeit with less justification. At a proton frequency of 400 MHz, the effect is maximal at correlation times around 0.4 ns, which almost matches the auto-correlation time of the internal reorientation of spin labels of about 0.7 ns found by Oganessian [48].

Such a neglect of significant dynamic modes had led to the prediction that Overhauser DNP in solutions would level off rapidly towards high fields [82]. Recent experiments have shown that, although the DNP efficiency indeed decreases with increasing fields, this effect is not as steep as expected [83,84]. These results were rationalized by MD simulations of the nitroxide molecule TEMPOL in TIP3P water [85]. The long tail of the correlation function was extrapolated by a $\tau^{-3/2}$ scaling, as appropriate for an analytical force-free model of translational diffusion. Such a model should apply at long times where the water molecules have diffused to sites remote from the nitroxide radical. This force-free model contains only two parameters, the constant $D_{\text{ff}} = D_{\text{NO}} + D_{\text{w}}$ of the relative translational diffusion of the nitroxide radical (D_{NO}) and water molecules (D_{w}) and the distance of closest approach d_{ff} . Although the MD

simulations nicely reproduced the experimentally known diffusion coefficient D_{NO} of TEMPOL (while the water diffusion constant was adjusted by manipulating the friction coefficient of the thermostat), the distance of closest approach and the correlation time predicted by the force-free model were found to vary quite strongly with frequency, with d_{ff} decreasing from 3.44 Å at an EPR frequency of 9.6 GHz to 2.31 Å at 360 GHz [85]. A fit of the dipolar correlation functions extracted from the MD trajectories required three exponentially decaying functions that were tentatively assigned to fast thermal librations of water molecules with time constants of roughly 0.4, 4, and 30 ps. These time-scales were attributed to, hydrogen-bonding of water to the N-O group, the making and breaking of such hydrogen bonds, and large-scale translational diffusion of water. The authors pointed out that the processes can actually be more complex [85]. Taken together, the two fastest processes contribute more than 70% to the total correlation function, which explains why DNP enhancements drop less steeply with increasing frequency than would be expected if the modulation of the hyperfine coupling were only due to translational diffusion.

A similar study for MTSL attached to a protein would be of interest in order to interpret recent Overhauser DNP results for spin-labeled membrane proteins [18,19], but this has not yet been carried out. However, diffusion constants of water in the vicinity of MTSL spin labels and native amino acids have been investigated in an earlier MD study [86]. Diffusion coefficients of water along the axis of the spin label and perpendicular to it were found to be virtually the same. In general, diffusion coefficients are smaller near a protein side group than in the bulk, and particularly small for water molecules trapped in cavities, such as the ones observable in crystal structures. For such water molecules residence times up to 200 ps have been found [86].

In PRE studies, the flexibility of a spin label does not only lead to internal motion at time scales relevant for inducing nuclear spin relaxation, but also to a distribution of the electron spin in

space. PRE scales with r^{-6} as a function of distance r , so that apparent mean distances $\langle r^{-6} \rangle^{-1/6}$ averaged over the spatial distribution of the electron spin are not the true distances, but rather distances to a point that is closer to the nuclear spin than to the center of gravity of the distribution [25]. This necessitates ensemble fitting with a multi-conformational representation of the label. The required number of conformations of the spin label depends on the spatial extent of the electron spin distribution and on the precision that is required for nuclei that may come closest to the electron spin, since for these nuclei the errors are largest. In a study of interdomain movement of glutamine binding protein, it was found that estimating the sampling space of MTSL is a non-trivial task [43]. A significant improvement in structure determination of the protein-substrate complex was obtained when the sampling space was pre-estimated from only intradomain PREs and the known structure of the *apo* protein. Note that the prediction of the conformational distribution of a spin label, as discussed in Section 4.2, is equivalent to the determination of this sampling space, which could be applied in an iterative way in structural modeling. Restriction of the sampling space was attempted [36] based on the χ_4 - χ_5 model for MTSL derivatives and the finding that substitution of the hydrogen at C2 (see Fig. 2) by a bulky group reduces spin label mobility [35]. Indeed, after introducing a 3-pyridyl group at C2, experimental PRE measurements in calmodulin could be satisfyingly reproduced by assuming only a single conformation of the spin label [36]. Note that a good fit of back-calculated PRE measurements is not a guarantee for good accuracy either of the determined distances [25] or of the motional model.

Magnetic resonance relaxation rates can also be directly simulated from MD trajectories without first extracting correlation functions [87]. Compared to the use of simplified motional models, this reduces the risk of missing significant dynamic processes. The relevant equation of motion is

$$d\Delta\rho/dt = -i[\mathbf{H},\Delta\rho] - i[\mathbf{H}_\omega,\rho] \quad , \quad (1)$$

where ρ is the spin density operator and $\Delta\rho = \rho - \rho_{\text{eq}}$ the difference with respect to the density operator ρ_{eq} in thermal equilibrium. The Hamiltonian \mathbf{H} includes the stationary Hamiltonian (mean field) as well as the fluctuating Hamiltonian that drives spin-lattice relaxation, whereas \mathbf{H}_ω describes a coherent driving Hamiltonian due to microwave or radiofrequency irradiation. Eq. (1) is conveniently expressed in terms of spherical tensor operators in Liouville space with superoperators \mathcal{L} and \mathcal{L}_ω corresponding to \mathbf{H} and \mathbf{H}_ω , respectively [88]. After computing the time average of the Liouville superoperator, the fluctuating Liouville-space Hamiltonian \mathcal{L}_f that drives relaxation can be obtained as the difference of the total Liouvillian from its time average. Following the ideas of Redfield, \mathcal{L}_f can be transformed into an interaction frame and a correlation matrix $C(\tau)$ can be computed as the time average of the product of the fluctuations of the Liouvillian in the interaction frame at times 0 and τ . Integration of this correlation matrix from time $t' = 0$ to infinity provides the relaxation matrix R . An approximation to R can be computed as a double sum over M discrete trajectories, each with N time steps Δt [87]. Although the method was initially applied to longitudinal relaxation induced by isotropic Brownian rotational diffusion with time steps of 100 ps for a total time span of 1 μs , processing of atomistic MD trajectories by the same approach appears feasible.

4. Conformational distribution

4.1 Conformational distribution in solids versus conformational dynamics in fluids

Many NMR studies and virtually all pulsed EPR studies on spin-labeled biomacromolecules are performed on solid samples, which may be either microcrystalline or glassy. Internal motions of the spin label, as discussed in the preceding sections, is largely frozen below the glass transition temperature, except for small-angle harmonic librations with correlation times in the picosecond range and below [89]. On the other hand, such motions can occur above the glass transition temperature and proceed on time scales that are much faster than the time required for freezing by immersion of the sample in liquid nitrogen (~ 1 s), and also faster than

the freezing time of about 100 μ s of freeze-quench experiments. Hence, it is commonly assumed that during freezing ergodic conditions are preserved down to the vicinity of the glass transition temperature, where motions over several decades of correlation times cease within a temperature interval of a few Kelvin [90]. The conformational distribution of macromolecules in glasses is thus expected to correspond to the conformational space explored by macromolecular dynamics at the glass transition temperature. This expectation is in qualitative agreement with force constants measured via the width of distance distributions between rigid TOAC labels in a coiled-coil leucine zipper motif [91] and quantitatively consistent with end-to-end distance distributions in oligo(phenyleneethynylene) [92].

This simple picture is slightly modified by experiments that can detect differences in label-to-label distance distributions in T4 lysozyme samples frozen on time scales of seconds or hundreds of microseconds [93]. At low cryoprotectant content (10% glycerol), enhanced local protein concentrations lead to faster background decays in DEER experiments for slow freezing, but not for fast freezing. This problem can be avoided by increasing the cryoprotectant concentration (30% glycerol). More significantly, in some cases distance distributions were broader for fast freezing than for slow freezing. This result was interpreted as an indication that the freeze quench was too fast for the proteins and spin labels to reach conformational equilibrium down to the glass transition temperature. Note however that properties of glasses generally depend on the cooling rates [90] and that the observed broadening of the distance distributions might also arise from large internal strains in the rapidly cooled sample.

The glass transition of protein surface water occurs near 170 K [94]. The glass transition of liposome preparations is believed to depend only on the solvent/cryoprotectant combination and is expected to lie in the range between 220 and 240 K [95]. Hence, solvent-exposed

residues at the surface of membrane proteins may experience the same glass transition temperature of 170 K as in soluble proteins.

Solid-state NMR studies are preferentially performed on microcrystalline samples to avoid inhomogeneous line broadening due to chemical shift dispersion caused by heterogeneous local environments [96]. Besides small-angle libration on Terahertz frequency scales, which also occurs in crystals, substantial side group dynamics or conformational distributions may occur in protein crystals, which are often not very densely packed. This is further discussed in Section 4.3 on the basis of experimental evidence for spin-label side groups.

The use of cryoprotectants also raises the question whether solvent composition influences the internal conformational distribution of spin labels. A recent DFT study has revealed that relative rotamer energies strongly change between computations *in vacuo* and more sophisticated solvation models [97]. Preferential interactions of the label with glycerol and poly(ethylene glycol) were suggested on the basis of changes in CW EPR lineshapes in a study of a water-exposed site in the flap region of HIV-1 protease [98]. In a study on the β -barrel outer membrane protein OmpA, a strong influence of the membrane-mimetic environment on the conformational distribution of spin labels was found by analyzing both CW EPR lineshapes and distance measurements across the β -barrel [99]. Significant differences have also been found between distance measurements on the ABC transporter MsbA in detergent micelles and in liposomes [100], although it is not clear for this relatively flexible protein whether the differences are due to different conformations of the spin label or to different protein backbone conformations. Conformational changes of ligand binding loops induced by poly(ethylene glycol) solutes have been observed in the outer membrane transporters OmpA and FecA and, in general, enhanced ordering and reduced dynamics of whole proteins due to interactions with solutes has been suggested [101]. It has also been suggested that in lipid bilayers, conformations of spin labels that lie largely parallel to layer are disfavored, while

conformations perpendicular to the layer, and thus parallel to the lipid chain director, are favored [102,103]. For the interpretation of spin label dynamics in terms of a cone model, this effect has been parametrized by a probability function $f = 1 - \sin\theta$, where θ is the angle between the membrane normal and a mean side group direction. To the best of our knowledge, no experimental evidence has been provided for this model.

4.2 Rotamer library approach

Assuming that the conformational distribution of a spin label corresponds to the conformational space explored by side group dynamics at the glass transition temperature, this distribution could be obtained from MD simulations at 175 K. Compared to MD simulations at ambient temperature, this would aggravate the sampling problem. Furthermore the computational effort would be forbidding in large studies with many labeling sites and, in particular, for *in silico* site scans that can help to select optimal labeling sites. Given limitations in the accuracy of current MD approaches, the question arises as to whether simplified modeling of the conformational distribution of the label could achieve comparable accuracy with much less computational effort. For native side groups, good accuracy can be achieved with rotamer library approaches [104].

In rotamer library modeling, geometries and relative energies of rotameric states are pre-computed once and for all for a mean protein environment. To predict side group conformations for a given backbone geometry, interactions between a rotamer and its environment are computed with a simplified attraction-repulsion potential. For native side groups, relative energies of the rotamers can be estimated from the frequency of their occurrence in known protein structures. Because of the large size of the PDB and the relatively small number of native side groups, backbone-dependent relative rotamer energies and their standard deviations could be derived for a grid of backbone dihedral angles ϕ and ψ [104].

Although several crystal structures of spin-labeled proteins have been published (Section 4.3) the experimental data are still far too sparse to parametrize a rotamer library for MTSL in the same fashion. Relative energies for the rotamers of unrestricted MTSL and IA-PROXYL side groups were obtained instead by running MD simulations over a total length of 100 ns sampling in 1 ps intervals [41]. To reduce the sampling problem, MD trajectories for MTSL with implicit water solvation were obtained in slices of 200 ps at target temperatures of 175 or 298 K that were interspersed with 1000 ps intervals at 600 K and 200 ps intervals for thermalization. This procedure leads to sufficiently frequent transitions between otherwise slowly exchanging basins, but annealing is too fast to obtain accurate relative populations for the basins. The 100'000 trajectory frames were projected onto approximately 200 rotamers. The relative energies of these rotamers for unrestricted spin labels were obtained by Boltzmann inversion of the frequency of their occurrence in the trajectories. As is seen in Fig. 5, rotamer distributions extracted from these MD simulations differ from those resulting from more elaborate and sophisticated MD simulations on a spin-labeled poly(alanine) α -helix. For χ_1 and χ_2 these differences are probably largely caused by steric hindrance between the spin label and the α -helical backbone, which is accounted for in the rotamer library approach by the interaction potential between the spin label and the protein (*vide infra*). However, the difference in the χ_4 rotamer distribution is most likely caused by neglecting the CMAP correction for coupling between χ_4 and χ_5 dihedral angle dynamics [52] in the less sophisticated MD simulations used to generate the rotamer library [41].

To compute rotamer populations at a given labeling site, relative rotamer energies of the unrestricted spin label must be augmented by the energy of the interaction of the rotamer with its environment. Information on the spin label environment can be derived from a crystal or NMR structure or from a structural model of the protein. This environment is approximated as rigid, which provides good computational efficiency, but also leads to unrealistically large

repulsive energies when the interaction energy is computed by standard pairwise atom-atom Lennard-Jones potentials. This problem can be solved by artificially reducing the van-der-Waals radius of the atoms by a so-called '*forgive factor*'. In practice, forgive factors in the range from 0.35 to 0.6 provide suitable accuracy. Electrostatic interactions and hydrogen bonding are neglected. The total energy of a rotamer is the sum of all Lennard-Jones energies and the relative energy of the free spin label. The partition function can be computed from the total energies of all rotamers and provides a measure for the steric hindrance at the labeling site considered. Relative rotamer populations follow from the Boltzmann distribution. The assumption of a rigid environment also implies that native side groups do not repack to accommodate the spin label, which for core residues is justified by x-ray crystal studies on the incorporation of unnatural amino acids [105]. Changes in rotameric states of native amino acids are rarely seen in the core of a protein, where the states are determined by optimized packing.

This approach is implemented in the open source software package MMM and allows one to compute a rotamer distribution at a given labeling site within a few seconds on a laptop . Such computational efficiency allows one to scan all possible labeling sites even in large proteins [40] for a systematic comparison with experimental data as discussed in Section 4.4. Rotamer distributions obtained with MMM for residues 72 and 131 in T4 lysozyme are visualized in Fig. 3.

The rotamericity in crystal structures, defined as the fraction of native side groups with dihedral angles in a range of $\pm 20^\circ$ around a canonical value, varies between 70 and 95% for native amino acids. It has been suggested that outliers occur systematically [106]. For spin labels, a similar behavior is suggested by the dihedral angles observed in reliable crystal structures (blue lines in Fig. 5, see also Section 4.3). Based on this consideration, it was suggested to use a fine-grained search space for spin label conformations [107]. For native

amino acids, this problem has been addressed by a sophisticated approach that provides well-packed, low-energy structures within reasonable computation times [108]. Recently, two alternative approaches with fine-grained sampling were introduced for modeling the conformational distribution of spin labels. In PRONOX spin label conformations are created by scanning a user-defined grid of the dihedral angles, without considering a potential energy function of the χ_i [109]. However, dihedral angle values for χ_1 and χ_2 that are frequently observed in crystal structures (Section 4.3) can be favored as an option. Conformations are rejected when clashes with atoms in the environment occur, defined as an approach of two atoms closer than 0.75 times the sum of their van-der-Waals radii. This yields larger optimal forgive factors compared to the rotamer library approach because of a finer sampling of the conformational space. PRONOX is available on a server [110].

The second alternative approach, known as MtsslWizard, is based on similar sampling of conformational space without taking into account potential energy functions [107]. Clashes are defined as an approach of any pair of atoms to a distance closer than a predefined value (the default is 3.4 Å, but it can be reduced to 2.6 Å). A user-defined number of clashes may be allowed. To search for hydrophobic pockets, MtsslWizard includes an optional 'snuggly fit' feature that constructs a rough attraction potential by counting non-clashing atoms in the environment with a distance of up to 4.5 Å to an atom on the spin label. Indiscriminate use of this feature is not recommended. MtsslWizard is available as a PyMol plugin [111].

Comparisons between the three approaches are deferred to Section 4.4. Note that the conformational ensemble obtained with such approaches is physically unrealistic, as the torsion potentials are completely ignored. Nevertheless spatial sampling of the electron spin position by such accessible volume approaches may be good [114], as has been observed for a similar problem encountered for chromophores in Förster resonance energy transfer (FRET) measurements [115].

4.3 Conformations observed in crystal structures

Within the past few years several soluble proteins (T4 lysozyme, Spa15, GB1, CylR2, cytochrome P450cam) and membrane proteins (LeuT, BtuB, KcsA) with attached spin labels have been crystallized. Table 1 gives an overview of the structures that are sufficiently well resolved to infer the conformations of the spin labels. This table is based on all structures containing the ligand MTN that were found in the PDB as of January 31st 2013. This ligand is the fragment of MTSL that binds to cysteine; a code that has replaced its name as a spin-labeled residue R1A in earlier versions of the PDB (for the numbering of atoms in the MTN ligand, see Fig. 2). NMR structures were ignored, since PRE constraints do not allow one to fully specify the conformation of spin labels. Note that conformations given in the NMR structure PDB:2KSQ obtained with X-PLOR-NIH feature an unrealistic bend of the N-O group with respect to the five-membered ring that has been corrected in later versions (X-PLOR NIH 2.18 and X-PLOR 3.85). NMR structures obtained with these later versions still show some tendency of χ_3 angles to be near $\pm 160^\circ$ that are not observed in crystal structures or MD simulations.

For all crystal structures, spin label conformations provided in the PDB files were checked against the electron densities provided by the Uppsala Electron-Density Server [116]. This excluded structures PDB:2W8H and PDB:2OU9 which do not display electron density because of a mismatch of more than 5% between predicted and reported R-values. In the former case, this may be due to crystal twinning, which was considered in refinement of the crystal structure [117]. Structure PDB:2IFX was excluded because of insufficient resolution (3.56 Å). In all other structures the model of the spin label was considered reliable up to the first discontinuity in electron density or up to the last atom that had at least 50% occupancy. Models for a given side group were disregarded altogether if significant unexplained electron density was found near that side group. For these reasons, some modeled spin labels within

the accepted protein structures are not listed in Table 1. From the first structures of T4 lysozyme that is singly labeled at residues 75 and 119 and doubly labeled at residues 65/80 [118], residues 75 and 119 are included in Table 2, although these structures are not found in the PDB. In this case agreement with electron density can be inferred from figures in the paper. Residue 80 was excluded because of poor electron density and residue 65 because of crystal packing interactions. Some spin labels are reported with less resolved atoms than are actually reported in the PDB file. These precautions were taken to ensure that all listed rotameric states have been observed with sufficient significance. Often electron density is poorly defined for spin label atoms that are remote from the backbone. Such a disorder of remote atoms had also been observed for other non-native amino acids [105] and may indicate significant dynamics with respect to χ_4 and χ_5 in microcrystals.

The strong preference for χ_1/χ_2 *mm* and *tp* rotamers observed for surface-exposed helical and loop sites in T4 lysozyme structures PDB:2IGC, PDB:2OU8, PDB:2NTH [119], PDB:2Q9D, PDB:2Q9E [58], and PDB:1ZYT, PDB:2CUU, PDB:3G3V, PDB:3G3X, and PDB:3G3W [59] is also observed in other soluble proteins (CylR2, PDB:2XIU, [120]; GB1, PDB:3V3X [121]) and, in one instance, at a buried site (Spa15, PDB:2XGA, [122]). For exposed residue Asn-8 in a β strand of GB1 (PDB:3V3X) non-rotameric χ_2 angles of -120.4° and 162° have been found [121]. As seen in Fig. 5, χ_2 features a rather broad distribution around the canonical values in MD simulations. Non-rotameric angles, defined here as dihedral angles deviating by more than 30° from their canonical value are also observed for χ_2 in loop site Ser-48 of cytochrome P450cam (PDB:4EK1, [123]), but in no instance for χ_1 or χ_3 .

Except for Asn-8 in chain B of PDB:3V3X and Trp-371 in the outer membrane protein BtuB (PDB:3RGM, [124]) the first two dihedral angles χ_1 and χ_2 are always resolved, although sometimes two rotameric states coexist. The existence of more than two rotameric states with respect to χ_1 and χ_2 cannot be safely excluded for Thr-151 in T4 lysozyme at 100 K (PDB:3G3X),

Val-10 in BtuB at 90 K (PDB:3M8B), and Thr-156 in BtuB where more 40-50% of the electron density of the spin label is unaccounted for.

At solvent-exposed loop site Val-10 and lipid-exposed β -strand site Trp-371 of BtuB [124] as well as at solvent-exposed loop site Ser-48 in cytochrome P450cam [123] and the solvent- or lipid-headgroup exposed loop site Tyr-82 in KcsA, there is clear evidence for p rotamers with respect to χ_1 that are sterically forbidden at α -helical sites. Biasing of a rotamer library towards the states observed in α -helical sites (see Section 4.4) may thus not be beneficial for simulations in loops and β -strands. In contrast, the two lipid-exposed α -helical sites that have been studied in the secondary inner membrane transporter LeuT [125], Phe-177 (PDB:3MPN) and Ile-204 (PDB:2MPQ) show the same mm rotamers that are frequently observed at solvent-exposed sites in soluble proteins as well as the relatively short C^α - S^δ distances that have been interpreted in terms of a non-classical hydrogen bond. Such C^α - S^δ distances of 3.5 Å or shorter are found in 19 out of 36 resolved spin label side group conformations. Note that a DFT study including dispersion correction and solvation also found a preference for those rotamers that can form a close C^α - S^δ contact [97].

In half of the reported spin label conformations, the complete side group is clearly resolved, although this includes one case with two equally populated alternate conformers and two further cases where 13 and 30% of the electron density is unaccounted for. Note that well resolved unique conformations in x-ray crystal structures at 90-120 K are no indication for frozen side group dynamics in solution at ambient temperature. For instance, rotamer exchange has been observed by NMR for thioredoxin residues [126] despite a 100% occupation of a single rotamer in the crystal structure [127]. Taking this into account the crystal structures suggest that for MTSL coexistence of several rotamers at ambient temperature is the rule rather than an exception. Such conformational distributions can be detected by distance distribution measurements between two spin labels in the same protein.

4.4 Effects on measurements of distance distributions

Within the past decade, distance distribution measurements for pairs of spin labels with pulsed dipolar spectroscopy methods have been performed for a number of proteins with known structures [60,100,128- 140]. From these studies, a few were selected where mean distances were given and where the duration of the dipolar evolution was sufficient [21] for these mean distances to be reliable. For cytochrome P450cam the most probable distances [141] were interpreted as mean distances of the major protein conformation. For T4 lysozyme, mean distances measured by CW EPR that were included in the data of a previous study [107] were also included in the data used for preparing Table 2. A file specifying the full set of site pairs with references and remarks is included with the MMM 2013 software package [39].

Assuming that the solution structure of these proteins is very close to the crystal structure, these data allow one to test predictions of conformational distributions of spin labels. In a first series of tests, mean square deviations and mean deviations between predicted and experimental distances were determined for 104 pairs of labeling sites in 8 proteins with the three available software packages MMM [41], PRONOX [109], and MtsslWizard [107] (Tables 2,3). For ten pairs of sites, predictions by all three programs deviated from the experimental results to an extent that was hardly consistent with the assumption that the protein structure remains unchanged in solution. Two of these pairs of sites involve residue 60 in T4 lysozyme (PDB:2LZM, pairs 60/90 and 60/94 in Table 3). Hence, the two remaining distances involving residue 60 in T4 lysozyme were also removed from the test set. Data for the 12 pairs of residues excluded in the second run are listed in Table 3. Note that for three out of five outliers in MsbA (PDB:3B60) good predictions can be obtained by using the 'snuggly fit' feature of MtsslWizard (see R_{SF} column in Table 3). This result indicates that for these solvent-exposed or buried residues in a membrane protein the attractive part of the potential may be important.

The attractive potential is partially suppressed in MMM by scaling of van-der-Waals-radii with the forgive factor and is not at all present in PRONOX or in the default mode of MtsslWizard.

Results for the remaining 92 site pairs are given in Table 2. Data for PRONOX were computed in the low clash mode (scaling of van-der-Waals radii to 40%), since this gave slightly better results for the total set and strongly improved results for cytochrome P450cam (PDB:2MBW) and for an archaeal homolog of the eukaryotic RNA polymerase II (PDB:1GO3). In MtsslWizard, 'tight' conditions were chosen for sites where they returned at least 10 feasible conformations, while 'loose' conditions were used for other sites. Data for MMM were computed with the ambient-temperature rotamer library, since on average this improved results compared to the 175 K library, as had been pointed out before [107]. In the bacteriorhodopsin structure (PDB:1C3W) lipid molecules were removed for processing with MMM and MtsslWizard, whereas PRONOX gave the same results with or without the lipid molecules.

For all test data, the three programs predict mean distances with comparable accuracy, corresponding to a mean square deviation between 3 and 3.5 Å, as was also found for the smaller test data set that included only T4 lysozyme (PDB:2LZM) and core-histone octamer (PDB:1TZY) [107]. The true accuracy is probably slightly better, since not all cases where labeling causes changes in the backbone structure may have been recognized, in particular for the relatively flexible ABC transporter MsbA (PDB:3B60). Note also that for this protein the DEER measurements were performed with ADP/vanadate ligation [100,133], whereas a crystal structure of sufficient quality is available only with AMPPNP ligation. For T4 lysozyme, cytochrome P450cam, the RNA polymerase II homolog, and the lipopolysaccharide transport protein LptA (PDB:2R19) all predictions have a root mean square accuracy of 2.5 Å or better.

These deviations can be compared to the standard deviation of predicted mean distances for 40 structural models each in NMR ensembles of the regulatory domain of human cardiac troponin C in the calcium-free state (PDB:1SPY) and bound to cardiac troponin I (PDB:1MXL).

For residue pair 35/84 standard deviations of 1.3 and 2.1 Å, respectively, are found. These uncertainties arising from variation in the NMR structural models are almost as large as the mean square deviations of predictions based on crystal structures from experimental distances. In contrast, the standard deviation of 8 predicted mean distances between pairs of adjacent protomers in the octameric Wza complex (PDB:2W8I) is only in the range from 0.1 to 0.4 Å for labels at residues 58, 78, 143, and 335, although the protomers are not related by crystal symmetry.

The results can also be compared to an earlier study where mean distances were predicted with a more elaborate Monte Carlo approach, resulting in a mean deviation of 3 Å [24]. A detailed study for a smaller number of pairs of sites concluded that MD and Monte Carlo simulations are not superior and possibly even slightly inferior to the simpler rotamer library approach [60]. Better performance of MD simulations was claimed in another study with a smaller number of site pairs, but in this case only those MD trajectories were selected that agreed well with experimental distance distribution data [50]. For the 16 distances in native histone octamer (PDB:1TZY), predictions from MD trajectories [134] have a root mean square deviation of 4.2 Å [107], slightly worse than with any of the simpler prediction approaches (see Table 2). Although the distribution of spin label conformations in MD simulations is certainly physically more realistic than in rotamer library or even accessible volume approaches, this advantage does not appear to result in a better accuracy of the predictions, most likely because of sampling problems.

In conjunction with the rotamer library approach in MMM, the experimental data also allow one to test predictions of the distribution of χ_1 , χ_2 rotamer groups that are based on rotamers observed in crystal structures (Table 1, R_{xray} in Table 2, blue lines in Fig. 5), or in the MD simulation of R1 at the central residue in a poly(alanine) α -helix ($R_{\text{MD}\alpha}$ in Table 2, red bars in Fig. 5) [52], or on relative rotamer energies computed by DFT and an elaborate solvation

model (R_{DFT} in Table 2, green bars in Fig. 5) [97]. In the case of DFT, relative populations of χ_1 , χ_2 , χ_3 groups of rotamers were computed from the energies by the Boltzmann distribution at 298 K. In this case only the *tpm* and *mmp* rotamer groups contribute to any significant extent, with a strong preference for the *tpm* group. In the Boltzmann distribution at the protein glass transition temperature of 175 K, the *tpm* group is almost exclusively populated. To test these predictions, relative populations of rotamer groups in the ambient-temperature library for MTSL in MMM were rescaled accordingly. Note that these populations correspond to the spin label that does not interact with the protein, which leads to double counting of interactions with an α -helical structure in the $R_{\text{MD}\alpha}$ set. The forgive factors for scaling van-der-Waals radii were separately optimized for each library with rescaled populations and are given as footnotes to Table 2.

Similar population rescaling leads to an improved agreement of rotamer library simulations with experimental results in a study on Rpo4/7 stalk module of the *M. jannaschii* RNA polymerase (PDB:1GO3) [60]. We find that population rescaling does not lead to significant improvement in accuracy when taking into account the whole set of test data. Accuracy for the five distances measured in the RNA polymerase varies in the range 1.2 to 2.5 Å root mean square deviation for different prediction algorithms and is strongly sensitive to parameter changes, for instance to the forgive factor in rotamer library simulations. Consistent loss of accuracy is observed with population rescaling according to DFT results. This suggests that entropic contributions, which are missing in these computations, are important, as was already suggested in an early study on rotamer distribution and dynamics of native side groups [54]. Note also that the Boltzmann distribution computed from DFT energies does not agree either with the rotamer distribution observed in x-ray crystal structures or with the one observed in MD simulations (Fig. 4).

For soluble proteins with only α -helices and loops, improvements of accuracy are seen in both the R_{xray} and $R_{\text{MD}\alpha}$ results, except for the cytoplasmic domain of erythrocyte membrane band 3 (PDB:1HYN), where prediction quality slightly worsens although α -helical sites also strongly dominate. Such improvements are not seen for the RNA polymerase II (PDB:1GO3) and the periplasmic lipopolysaccharide transport protein LptA (PDB:2R19) where several β -strand sites were labeled.

Such comparison could potentially answer the question whether the χ_1 , χ_2 rotamers preferentially observed in x-ray crystal structures (mm , tp , and weakly tm) are almost exclusively populated in solution or whether the states mt and tt that are sterically allowed and often encountered in MD simulations, are also present. However, the data presented in Table 2 are inconclusive with respect to this question.

Predictions of spin label conformational distributions can provide a complete distance distribution rather than only the mean distance. In a rotamer library approach, a comparison of distance distributions can provide insight into the influence of relative populations of rotamer groups on prediction quality. For instance, relative populations of the p and m rotamer groups with respect to χ_3 are not well known because of MD sampling problems. Fig. 3 suggests that spatial sampling of these two rotamer groups is very similar. This is borne out by comparison of the simulated distance distributions shown in Fig. 7 for MTSL at residues 72 and 131 in T4 lysozyme (PDB:2LZM) taking into account all rotamers (black line), only χ_3 m rotamers (blue line), and only χ_3 p rotamers (red line). The small differences seen between these simulations would be very hard to detect reliably in experiments. In contrast, there are significant differences between χ_1 , χ_2 rotamer group populations seen in the MD simulations of the unrestricted spin label underlying the original ambient temperature library (black line) and in MD simulations for a spin label attached to an α -helix [52] (green line).

Such simulated distance distributions can also be compared to the experimental distribution for this doubly labeled protein. Furthermore, from the distance distribution the primary DEER data can be simulated by fitting only modulation depth and a decay constant of the intermolecular background function. If the measurements are performed at relatively low concentration, background decay is weak and modulation depth fitting corresponds only to vertical scaling. Such comparisons between experimental and predicted distance distribution and primary DEER data are shown in Fig. 8a and b, respectively. The measurements were performed with high-power Q-band DEER [142] in order to obtain high signal-to-noise ratio at sufficiently long dipolar evolution times. Distance distributions are normalized to their integral to accentuate differences in distribution width. All predicted distributions are somewhat broader than the experimental one. For this particular example the best predictions come from the rotamer library approach using the low-temperature library (solid purple line) or the ambient temperature library biased towards rotamers observed in crystal structures (solid blue line). With PRONOX, default biasing towards χ_1 , χ_2 values observed in crystal structures (solid red line) leads to a better prediction than a pure accessible volume approach (dotted red line). MtsslWizard provides predictions with a quality intermediate between the two PRONOX predictions and with not much difference between the tight (dotted green line) and loose (solid green line) modes.

These results show that biasing towards favorable dihedral angles χ_1 , χ_2 can cause changes that are comparable to typical deviations of mean distances and distribution widths between prediction and experiment. However, from the discussion of mean distance predictions it should be clear that comparison of the approaches for a single site pair cannot provide a general judgment on relative performance. This is also borne out by the distance distributions shown in Fig. 9 that were measured between pairs of residues at α -helical sites in rhodopsin [131] or predicted based on structure PDB:1GZM. Site pair 252/326 from the original work was excluded from this comparison, as site 326 is a terminal residue whose backbone conformation

may be more strongly influenced by spin labeling. Overall prediction quality is quite good, except for site pair 63/241. The data show that differences between different predictions are comparable to the differences between individual predictions and experiments and that none of the three approaches is clearly superior. This conclusion does not change by χ_1 , χ_2 biasing in the rotamer library approach (data not shown). Taken together all comparisons of long-range distance and distance distribution measurements with predictions of conformational distributions of spin labels suggest that an improvement of prediction quality will require simultaneous improvement of several approximations made in existing approaches.

5. Conclusions and Perspectives

Spin label dynamics on time scales between picoseconds and microseconds influences spectral responses in NMR and EPR experiments. The situation is complex due to an overlap of time scales of internal motions of spin labels and motions of the labeled biomacromolecules themselves. Quantitative predictions are further complicated by the limited accuracy of molecular force fields, sampling problems in MD computations, and possibly even shortcomings of the functional form of the most widely used molecular force fields.

Nevertheless, in recent work the agreement between experimental and predicted EPR spectral lineshapes is rather good, indicating that internal motions on the relevant time scales between 10 ps and 10 ns are reasonably well captured by MD simulations. Internal dynamics of spin labels on the same time scales are also expected to affect PRE and Overhauser DNP.

In the solid state, internal dynamics of spin labels can be neglected, except for small-angle librations and vibrations that hardly influence NMR and EPR spectra. Yet, PRE and EPR distance measurements in solids, and some phenomena in solid-state DNP, depend on the conformational distributions of the spin labeled side groups, which are hard to predict. Mean

distances between two spin labels can be predicted with an accuracy of about 3 Å (root mean square deviation), corresponding to an uncertainty of about 2.2 Å for the position of a single label. Surprisingly, the accuracy is about the same for rotamer library approaches, accessible volume methods, more elaborate Monte Carlo conformational searches, or MD approaches.

Within the past decade, the understanding of conformational dynamics and the distribution of nitroxide labels has greatly improved and an impressive body of high-quality experimental data has been obtained for proteins with known structures, in some cases using complementary experimental approaches for the same labeling sites on the same protein. Deficiencies in current theoretical descriptions have been identified. Further work along these directions may not only lead to a more quantitative analysis of PRE, Overhauser DNP, CW EPR lineshapes, and pulsed EPR data, but may also shed some light on general aspects of the dynamics of protein side groups and on molecular force fields.

Acknowledgments

I thank Remmia Mathews for providing the T4 lysozyme sample, Enrica Bordignon for the Q-band DEER measurements, Ian Haworth for help with PRONOX, Gregor Hagelueken for providing all MtsslWizard simulation results and recognizing the improvements with the 'snuggly fit' mode for three sites in MsbA, as well as Wayne Hubbell and Christian Altenbach for providing the experimental data for rhodopsin. Helpful discussions by Enrica Bordignon and Wayne Hubbell are gratefully acknowledged. This work was supported by SNF grant 200020_14441. It was edited by Geoffrey Bodenhausen and David G. Gadian.

References

- [1] G. I. Likhtenstein, J. Yamauchi, S. Nakatsuji, A. I. Smirnov, R. Tamura Nitroxides – Applications in Chemistry, Biomedicine, and Materials Science, VCH-Wiley, Weinheim, 2008.
- [2] O. H. Griffith, H. M. McConnell, Proc. Natl. Acad. Sci. U. S. A. 55 (1966) 8-11.
- [3] L. K. Tamm, B. Y. Liang, Prog. Nucl. Magn. Reson. Spectrosc. 48 (2006) 201-210.

- [4] G. M. Clore, C. Tang, J. Iwahara, *Curr. Opin. Struct. Biol.* 17 (2007) 603-616.
- [5] G. M. Clore, J. Iwahara, *Chem. Rev.* 109 (2009) 4108-4139.
- [6] R. C. Page, S. Lee, J. D. Moore, S. J. Opella, T. A. Cross, *Protein Sci.* 18 (2009) 134-146.
- [7] J. R. Allison, P. Varnai, C. M. Dobson, M. Vendruscolo, *J. Am. Chem. Soc.* 131 (2009) 18314-18326.
- [8] S. Wang, R. A. Munro, S. Y. Kim, K. H. Jung, L. S. Brown, V. Ladizhansky, *J. Am. Chem. Soc.* 134 (2012) 16995-8.
- [9] A. Abragam, M. Goldman, *Rep. Prog. Phys.* 41 (1978) 395-467.
- [10] T. Maly, G. T. Debelouchina, V. S. Bajaj, K. N. Hu, C. G. Joo, M. L. Mak-Jurkauskas, J. R. Sirigiri, P. C. A. van der Wel, J. Herzfeld, R. J. Temkin, R. G. Griffin, *J. Chem. Phys.* 128 (2008) 052211.
- [11] C. Griesinger, M. Bennati, H. M. Vieth, C. Luchinat, G. Parigi, P. Hofer, F. Engelke, S. J. Glaser, V. Denysenkov, T. F. Prisner, *Prog. Nucl. Magn. Reson. Spectrosc.* 64 (2012) 4-28.
- [12] M. D. Lingwood, S. I. Han, *Annu. Rep. NMR Spectrosc.* 73 (2011) 83-126.
- [13] C. Altenbach, T. Marti, H. G. Khorana, W. L. Hubbell, *Science* 248 (1990) 1088-92.
- [14] W. L. Hubbell, D. S. Cafiso, C. Altenbach, *Nat. Struct. Biol.* 7 (2000) 735-9.
- [15] O. Schiemann, N. Piton, J. Plackmeyer, Bode BE, T. F. Prisner, J. W. Engels, *Nat. Protoc.* 2 (2007) 904-923.
- [16] H. S. Mchaourab, M. A. Lietzow, K. Hideg, W. L. Hubbell, *Biochemistry* 35 (1996) 7692-704.
- [17] J. L. Battiste, G. Wagner, *Biochemistry* 39 (2000) 5355-65.
- [18] A. Doll, E. Bordignon, B. Joseph, R. Tschaggelar, G. Jeschke, *J. Magn. Reson.* 222 (2012) 34-43.
- [19] S. Hussain, J. M. Franck, S. Han, *Angew. Chem. Int. Ed. Engl.* 52 (2013) 1953-8.
- [20] V. Vitzthum, F. Borcard, S. Jannin, M. Morin, P. Mieville, M. A. Caporini, A. Sienkiewicz, S. Gerber-Lemaire, G. Bodenhausen, *ChemPhysChem* 12 (2011) 2929-32.
- [21] G. Jeschke, *Annu. Rev. Phys. Chem.* 63 (2012) 419-46.
- [22] R. Ward, M. Zoltner, L. Beer, H. El Mkami, I. R. Henderson, T. Palmer, D. G. Norman, *Structure* 17 (2009) 1187-1194.
- [23] Y. H. Yang, T. A. Ramelot, R. M. McCarrick, S. S. Ni, E. A. Feldmann, J. R. Cort, H. A. Wang, C. Ciccocanti, M. Jiang, H. Janjua, T. B. Acton, R. Xiao, J. K. Everett, G. T. Montelione, M. A. Kennedy, *J. Am. Chem. Soc.* 132 (2010) 11910-11913.
- [24] K. Sale, L. K. Song, Y. S. Liu, E. Perozo, P. Fajer, *J. Am. Chem. Soc.* 127 (2005) 9334-9335.
- [25] J. Iwahara, C. D. Schwieters, G. M. Clore, *J. Am. Chem. Soc.* 126 (2004) 5879-96.
- [26] R. Marchetto, S. Schreier, C. R. Nakaie, *J. Am. Chem. Soc.* 115 (1993) 11042-11043.
- [27] C. Toniolo, E. Valente, F. Formaggio, M. Crisma, G. Pilloni, C. Corvaja, A. Toffoletti, G. V. Martinez, M. P. Hanson, G. L. Millhauser, a. I. et, *J. Pept. Sci.* 1 (1995) 45-57.
- [28] S. A. Shelke, S. T. Sigurdsson, *Eur. J. Org. Chem.* (2012) 2291-2301.
- [29] A. D. Milov, Y. D. Tsvetkov, F. Formaggio, M. Crisma, C. Toniolo, J. Raap, *J. Am. Chem. Soc.* 123 (2001) 3784-3789.
- [30] S. Stoller, G. Sicoli, T. Y. Baranova, M. Bennati, U. Diederichsen, *Angew. Chem. Int. Edit.* 50 (2011) 9743-9746.
- [31] L. J. Berliner, J. Grunwald, H. O. Hankovszky, K. Hideg, *Anal. Biochem.* 119 (1982) 450-5.
- [32] C. Altenbach, K. J. Oh, R. J. Trabanino, K. Hideg, W. L. Hubbell, *Biochemistry* 40 (2001) 15471-82.

- [33] R. Igarashi, T. Sakai, H. Hara, T. Tenno, T. Tanaka, H. Tochio, M. Shirakawa, *J. Am. Chem. Soc.* 132 (2010) 8228-9.
- [34] M. R. Fleissner, E. M. Brustad, T. Kalai, C. Altenbach, D. Cascio, F. B. Peters, K. Hideg, S. Peucker, P. G. Schultz, W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 21637-42.
- [35] L. Columbus, T. Kalai, J. Jeko, K. Hideg, W. L. Hubbell, *Biochemistry* 40 (2001) 3828-46.
- [36] N. L. Fawzi, M. R. Fleissner, N. J. Anthis, T. Kalai, K. Hideg, W. L. Hubbell, G. M. Clore, *J. Biomol. NMR* 51 (2011) 105-114.
- [37] M. R. Fleissner, M. D. Bridges, E. K. Brooks, D. Cascio, T. Kalai, K. Hideg, W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 16241-16246.
- [38] Y. G. Yu, T. E. Thorgeirsson, Y. K. Shin, *Biochemistry* 33 (1994) 14221-6.
- [39] MMM 2013 software package, www.epr.ethz.ch/software/index, accessed Feb 11th 2013.
- [40] Y. Polyhach, G. Jeschke, *Spectr. Int. J.* 24 (2010) 651-659.
- [41] Y. Polyhach, E. Bordignon, G. Jeschke, *Phys. Chem. Chem. Phys.* 13 (2010) 2356-66.
- [42] M. D. Rabenstein, Y. K. Shin, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 8239-43.
- [43] G. A. Bermejo, M. P. Strub, C. Ho, N. Tjandra, *J. Am. Chem. Soc.* 131 (2009) 9532-9537.
- [44] D. Gottstein, S. Reckel, V. Dotsch, P. Guntert, *Structure* 20 (2012) 1019-1027.
- [45] O. B. Okan, A. R. Atilgan, C. Atilgan, *Biophys. J.* 97 (2009) 2080-2088.
- [46] V. Barone, P. Cimino, A. Pedone, *Magn. Reson. Chem.* 48 (2010) S11-22.
- [47] C. Beier, H. J. Steinhoff, *Biophys. J.* 91 (2006) 2647-2664.
- [48] V. S. Oganessian, *Phys. Chem. Chem. Phys.* 13 (2011) 4724-4737.
- [49] D. Sezer, J. H. Freed, B. Roux, *J. Am. Chem. Soc.* 131 (2009) 2597-605.
- [50] J. L. Sarver, J. E. Townsend, G. Rajapakse, L. Jen-Jacobson, S. Saxena, *J. Phys. Chem. B.* 116 (2012) 4024-4033.
- [51] F. Tombolato, A. Ferrarini, J. H. Freed, *J. Phys. Chem. B.* 110 (2006) 26248-26259.
- [52] D. Sezer, J. H. Freed, B. Roux, *J. Phys. Chem. B.* 112 (2008) 5755-5767.
- [53] M. Krishnan, J. C. Smith, *J. Phys. Chem. B.* 116 (2012) 4124-33.
- [54] T. C. Beutler, T. Bremi, R. R. Ernst, W. F. vanGunsteren, *J. Phys. Chem.* 100 (1996) 2637-2645.
- [55] J. R. Allison, W. F. van Gunsteren, *ChemPhysChem* 10 (2009) 3213-3228.
- [56] D. Steiner, J. R. Allison, A. P. Eichenberger, W. F. van Gunsteren, *J. Biomol. NMR.* 53 (2012) 223-246.
- [57] I. Stoica, *J. Mol. Model.* 11 (2005) 210-225.
- [58] Z. F. Guo, D. Cascio, K. Hideg, W. L. Hubbell, *Protein Sci.* 17 (2008) 228-239.
- [59] M. R. Fleissner, D. Cascio, W. L. Hubbell, *Protein Sci.* 18 (2009) 893-908.
- [60] D. Klose, J. P. Klare, D. Grohmann, C. W. M. Kay, F. Werner, H. J. Steinhoff, *PLoS One*, 7 (2012) e39492.
- [61] J. W. Ponder, C. J. Wu, P. Y. Ren, Pande VS, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, T. Head-Gordon, *J. Phys. Chem. B.* 114 (2010) 2549-2564.
- [62] I. C. Yeh, G. Hummer, *J. Am. Chem. Soc.* 124 (2002) 6563-8.
- [63] D. Sezer, J. H. Freed, B. Roux, *J. Chem. Phys. (USA)* 128 (2008) 165106-1-16.
- [64] D. D. Thomas, L. R. Dalton, Hyde JS, *J. Chem. Phys.* 65 (1976) 3006-3024.
- [65] J. McCoy, W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 1331-6.
- [66] D. J. Schneider, J. H. Freed, *Biol. Magn. Reson.* 8 (1989) 1-76.

- [67] D. E. Budil, S. Lee, S. Saxena, J. H. Freed, *J. Magn. Reson. A.* 120 (1996) 155-189.
- [68] Z. C. Liang, Y. Lou, J. H. Freed, L. Columbus, W. L. Hubbell, *J. Phys. Chem. B.* 108 (2004) 17649-17659.
- [69] A. K. Smith, J. H. Freed, *Chem. Phys. Lipids.* 165 (2012) 348-361.
- [70] A. Polimeno, J. H. Freed, *J. Phys. Chem.* 99 (1995) 10995-11006.
- [71] E. Meirovitch, Y. E. Shapiro, A. Polimeno, J. H. Freed, *Prog. Nucl. Magn. Reson. Spectrosc.* 56 (2010) 360-405.
- [72] Z. C. Liang, J. H. Freed, *J. Phys. Chem. B.* 103 (1999) 6384-6396.
- [73] ACERT homepage, www.acert.cornell.edu/, accessed Feb 11th 2013
- [74] EasySpin homepage, <http://www.easyspin.org/>, accessed Feb 11th 2013.
- [75] Z. W. Zhang, M. R. Fleissner, D. S. Tipikin, Z. C. Liang, J. K. Moscicki, K. A. Earle, W. L. Hubbell, J. H. Freed, *J. Phys. Chem. B.* 114 (2010) 5503-5521.
- [76] Hyde JS, M. Pasenkiewicz-Gierula, A. Lesmanowicz, W. E. Antholine, *Appl. Magn. Reson.* 1 (1990) 483-496.
- [77] B. H. Robinson, L. J. Slutsky, F. P. Auteri, *J. Chem. Phys.* 96 (1992) 2609-2616.
- [78] S. Piana, K. Lindorff-Larsen, D. E. Shaw, *Proc. Natl. Acad. Sci U. S. A.* 109 (2012) 17845-17850.
- [79] H. J. Steinhoff, W. L. Hubbell, *Biophys. J.* 71 (1996) 2201-2212.
- [80] D. Sezer, J. H. Freed, B. Roux, *J. Phys. Chem. B.* 112 (2008) 11014-11027.
- [81] S. C. DeSensi, D. P. Rangel, A. H. Beth, T. P. Lybrand, E. J. Hustedt, *Biophys. J.* 94 (2008) 3798-3809.
- [82] K. H. Hausser, *Z. Phys.* 183 (1965) 265-273.
- [83] P. Hofer, G. Parigi, C. Luchinat, P. Carl, G. Guthausen, M. Reese, T. Carlomagno, C. Griesinger, M. Bennati, *J. Am. Chem. Soc.* 130 (2008) 3254-5.
- [84] M. J. Prandolini, V. P. Denysenkov, M. Gafurov, S. Lyubenova, B. Endeward, M. Bennati, T. F. Prisner, *Appl. Magn. Reson.* 34 (2008) 399-407.
- [85] D. Sezer, M. J. Prandolini, T. F. Prisner, *Phys. Chem. Chem. Phys.* 11 (2009) 6626-6637.
- [86] I. Stoica, *J. Biomol. Struct. Dyn.* 21 (2004) 745-760.
- [87] D. P. Rangel, P. C. Baveye, B. H. Robinson, *J. Phys. Chem. B.* 116 (2012) 6233-49.
- [88] R. D. Nielsen, B. H. Robinson, *Concepts Magn. Reson. Part. A.* 28A (2006) 270-290.
- [89] S. A. Dzuba, E. P. Kirilina, E. S. Salnikov, *J. Chem. Phys.* 125 (2006) .
- [90] J. L. Barrat, J. Baschnagel, A. Lyulin, *Soft Matter.* 6 (2010) 3430-3446.
- [91] S. V. Gulla, G. Sharma, P. Borbat, J. H. Freed, H. Ghimire, M. R. Benedikt, N. L. Holt, G. A. Lorigan, K. Rege, C. Mavroidis, D. E. Budil, *J. Am. Chem. Soc.* 131 (2009) 5374-5.
- [92] G. Jeschke, M. Sajid, M. Schulte, N. Ramezani, A. Volkov, H. Zimmermann, A. Godt, *J. Am. Chem. Soc.* 132 (2010) 10107-10117.
- [93] E. R. Georgieva, A. S. Roy, V. M. Grigoryants, P. P. Borbat, K. A. Earle, C. P. Scholes, J. H. Freed, *J. Magn. Reson.* 216 (2012) 69-77.
- [94] W. Doster, *Biochim. Biophys. Acta* 1804 (2010) 3-14.
- [95] E. C. van Winden, W. Zhang, D. J. Crommelin, *Pharm. Res.* 14 (1997) 1151-60.
- [96] W. T. Franks, A. H. Linden, B. Kunert, B. J. van Rossum, H. Oschkinat, *Eur. J. Cell. Biol.* 91 (2012) 340-348.
- [97] D. T. Warshaviak, L. Serbulea, K. N. Houk, W. L. Hubbell, *J. Phys. Chem. B.* 115 (2011) 397-405.

- [98] L. Galiano, M. E. Blackburn, A. M. Veloro, M. Bonora, G. E. Fanucci, *J. Phys. Chem. B.* 113 (2009) 1673-1680.
- [99] R. H. F. Jimenez, D. M. Freed, D. S. Cafiso, *J. Phys. Chem. B.* 115 (2011) 14822-14830.
- [100] P. P. Borbat, K. Surendhran, M. Bortolus, P. Zou, J. H. Freed, H. S. Mchaourab, *PLoS Biol.* 5 (2007) e271.
- [101] M. Kim, Q. Xu, D. Murray, D. S. Cafiso, *Biochemistry* 47 (2008) 670-679.
- [102] J. Strancar, A. Kavalenka, P. Zihnerl, D. Stopar, M. A. Hemminga, *J. Magn. Reson.* 197 (2009) 245-248.
- [103] J. Strancar, A. Kavalenka, I. Urbancic, A. Ljubetic, M. A. Hemminga, *Eur. Biophys. J. Biophys. Lett.* 39 (2010) 499-511.
- [104] G. G. Krivov, M. V. Shapovalov, R. L. Dunbrack Jr, *Proteins* 77 (2009) 778-95.
- [105] R. Wynn, P. C. Harkins, F. M. Richards, R. O. Fox, *Protein Sci.* 6 (1997) 1621-6.
- [106] H. Schrauber, F. Eisenhaber, P. Argos, *J. Mol. Biol.* 230 (1993) 592-612.
- [107] G. Hagelueken, R. Ward, J. H. Naismith, O. Schiemann, *Appl. Magn. Reson.* 42 (2012) 377-391.
- [108] A. Fahmy, G. Wagner, *Biophys. J.* 101 (2011) 1690-1698.
- [109] M. M. Hatmal, Y. Y. Li, Hegde BG, Hegde PB, C. C. Jao, R. Langen, I. S. Haworth, *Biopolymers* 97 (2012) 35-44.
- [110] PRONOX server,
<http://rockscluster.hsc.usc.edu/research/software/pronox/pronox.html>, accessed Feb 11th 2013
- [111] MtsslWizard page of PyMol Wiki, www.pymolwiki.org/index.php/mtsslWizard, accessed Feb 18th 2013.
- [114] Q. Cai, A. K. Kusnetzow, K. Hideg, E. A. Price, I. S. Haworth, P. Z. Qin, *Biophys. J.* 93 (2007) 2110-7.
- [115] S. Kalinin, T. Peulen, S. Sindbert, P. J. Rothwell, S. Berger, T. Restle, R. S. Goody, H. Gohlke, C. A. Seidel, *Nat. Methods.* 9 (2012) 1218-25.
- [116] G. J. Kleywegt, M. R. Harris, J. Y. Zou, T. C. Taylor, A. Wahlby, T. A. Jones, *Acta Crystallogr. D. Biol. Crystallogr.* 60 (2004) 2240-9.
- [117] G. Hagelueken, W. J. Ingledew, H. Huang, B. Petrovic-Stojanovska, C. Whitfield, H. EIMkami, O. Schiemann, J. H. Naismith, *Angew. Chem. Int. Ed.* 48 (2009) 2904-2906.
- [118] R. Langen, K. J. Oh, D. Cascio, W. L. Hubbell, *Biochemistry* 39 (2000) 8396-405.
- [119] Z. F. Guo, D. Cascio, K. Hideg, T. Kalai, W. L. Hubbell, *Protein Sci.* 16 (2007) 1069-1086.
- [120] T. Gruene, M. K. Cho, I. Karyagina, H. Y. Kim, C. Grosse, K. Giller, M. Zweckstetter, S. Becker, *J. Biomol. NMR.* 49 (2011) 111-119.
- [121] T. F. Cunningham, M. S. McGoff, I. Sengupta, C. P. Jaroniec, W. S. Horne, S. Saxena, *Biochemistry* 51 (2012) 6350-6359.
- [122] J. E. D. Lillington, J. E. Lovett, S. Johnson, P. Roversi, C. R. Timmel, S. M. Lea, *J. Mol. Biol.* 405 (2011) 427-435.
- [123] S. Stoll, Y. T. Lee, M. Zhang, R. F. Wilson, R. D. Britt, D. B. Goodin, *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 12888-12893.
- [124] D. M. Freed, A. K. Khan, P. S. Horanyi, D. S. Cafiso, *Biochemistry* 50 (2011) 8792-803.
- [125] B. M. Kroncke, P. S. Horanyi, L. Columbus, *Biochemistry* 49 (2010) 10045-60.
- [126] D. M. LeMaster, *J. Am. Chem. Soc.* 121 (1999) 1726-1742.
- [127] S. K. Katti, D. M. LeMaster, H. Eklund, *J. Mol. Biol.* 212 (1990) 167-84.

- [128] P. P. Borbat, H. S. Mchaourab, J. H. Freed, *J. Am. Chem. Soc.* 124 (2002) 5304-5314.
- [129] Z. Zhou, S. C. DeSensi, R. A. Stein, S. Brandon, M. Dixit, E. J. McArdle, E. M. Warren, H. K. Kroh, L. K. Song, C. E. Cobb, E. J. Hustedt, A. H. Beth, *Biochemistry* 44 (2005) 15115-15128.
- [130] N. Alexander, M. Bortolus, A. Al-Mestarihi, H. Mchaourab, J. Meilerl, *Structure* 16 (2008) 181-195.
- [131] C. Altenbach, A. K. Kusnetzow, O. P. Ernst, K. P. Hofmann, W. L. Hubbell, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 7439-7444.
- [132] L. Song, M. Larion, J. Chamoun, M. Bonora, P. G. Fajer, *Eur. Biophys. J.* 39 (2009) 711-9.
- [133] P. Zou, M. Bortolus, H. S. Mchaourab, *J. Mol. Biol.* 393 (2009) 586-597.
- [134] A. Bowman, R. Ward, H. El-Mkami, T. Owen-Hughes, D. G. Norman, *Nucleic Acids Res.* 38 (2010) 695-707.
- [135] K. Kazmier, N. S. Alexander, J. Meiler, H. S. Mchaourab, *J. Struct. Biol.* 173 (2011) 549-557.
- [136] J. A. Merten, K. M. Schultz, C. S. Klug, *Protein Sci.* 21 (2011) 211-8.
- [137] V. Krishnamani, Hegde BG, R. Langen, J. K. Lanyi, *Biochemistry* 51 (2012) 1051-1060.
- [138] A. Mittal, S. Bohm, M. G. Grutter, E. Bordignon, M. A. Seeger, *J. Biol. Chem.* 287 (2012) 20395-20406.
- [139] D. M. Freed, P. S. Horanyi, M. C. Wiener, D. S. Cafiso, *Biophys. J.* 99 (2010) 1604-10.
- [140] H. Raghuraman, J. F. Cordero-Morales, V. Jogini, A. C. Pan, A. Kollwe, B. Roux, E. Perozo, *Structure* 20 (2012) 1332-42.
- [141] C. J. Lopez, S. Oga, W. L. Hubbell, *Biochemistry* 51 (2012) 6568-6583.
- [142] Y. Polyhach, E. Bordignon, R. Tschaggelar, S. Gandra, A. Godt, G. Jeschke, *Phys. Chem. Chem. Phys.* 14 (2012) 10762-73.

Glossary

CW continuous wave

DNP dynamic nuclear polarization

EPR electron paramagnetic resonance

FID free induction decay

MD molecular dynamics

MOMD microscopic order, macroscopic disorder

MTSL methanethiosulfonate spin label

PDB protein data bank

PRE paramagnetic relaxation enhancement

SDSL site-directed spin labeling

SLE stochastic Liouville equation

SRLS slowly relaxing local structure

Tables

Table 1: Conformations of the spin-labeled residue R1 formed by reaction of the MTSL label with a cysteine residue (Fig. 2) in crystal structures of proteins with their PDB identifiers, resolution, site address in the format (chain)residue.location, native residue, secondary structure type, solvent exposure, resolved R1 atoms, occupancy, dihedral angles χ_1 - χ_3 , C^α - S^δ distance, rotamer type, temperature, and reference.

PDB ID	Res. (Å)	Site	Native	Sec.	Exposure	Resolved	Occ.	χ_1	χ_2	χ_3	C^α - S^δ (Å)	Rotamer	T (K)	Ref.
2Q9D	1.40	(A)41	Ala	α	exposed	all	1.0	-174.8	56.9	86.4	3.6	<i>tp</i>	100	[58]
2Q9E	2.10	(A)44	Ser	α	exposed	all	1.0	-83.1	-57.5	-95.3	3.4	<i>mmm</i>	100	[58]
2Q9E	2.10	(B)44	Ser	α	exposed	all	1.0	-85.4	-55.4	-95.9	3.3	<i>mmm</i>	100	[58]
2Q9E	2.10	(C)44	Ser	α	exposed	S1	1.0	173.0	-96.0	n.a.	3.9	<i>tm</i>	100	[58]
n.a.	1.5	75	Val	α	exposed	all	n.a.	-73	173	n.a.	n.a.	<i>mt</i>	100	[117]
1ZYT	1.70	(A)82	Ala	l	exposed	S1,C4,C3	0.8	-68.4	-56.0	101.9	3.4	<i>mmp</i>	100	[59]
2IGC	1.40	(A)115	Thr	α (t)	exposed	all	1.0	-81.4	-57.5	-92.2	3.7	<i>mmm</i>	100	[118]
2OU8	1.80	(A)115.A	Thr	α (t)	exposed	S1,C4	0.5	-88.3	-28.7	n.a.	3.6	<i>mm</i>	298	[118]

2OU8	1.80	(A)115.B	Thr	$\alpha(t)$	exposed	S1,C4	0.5	166.3	-58.2	n.a.	3.7	<i>tm</i>	298	[118]
2NTH	1.80	(A)118	Leu	α	buried	all	1.0	-104.7	33.0	88.2	3.5	<i>mpp</i>	100	[118]
n.a.	1.43	119.A	Arg	α	exposed	S1	n.a.	-50	-50	n.a.	n.a.	<i>mm</i>	100	[117]
n.a.	1.43	119.B	Arg	α	exposed	S1,C4	n.a.	175	54	n.a.	n.a.	<i>tp</i>	100	[117]
2CUU	1.75	(A)131.A	Val	α	exposed	S1,C4	0.5	175.3	80.0	n.a.	3.7	<i>tp</i>	100	[59]
2CUU	1.75	(A)131.B	Val	α	exposed	S1,C4	0.5	-68.7	-60.2	n.a.	3.4	<i>mm</i>	100	[59]
3G3V	2.10	(A)131.A	Val	α	exposed	S1	0.5	175.0	83.1	n.a.	3.7	<i>tp</i>	291	[59]
3G3V	2.10	(A)131.B	Val	α	exposed	S1	0.5	-74.5	-57.4	n.a.	3.5	<i>mm</i>	291	[59]
3G3X	1.80	(A)151	Thr	$\alpha(t)$	exposed	S1	0.5	-83.2	-72.1	n.a.	3.5	<i>mm</i>	100	[59]
3G3W	2.30	(A)151	Thr	$\alpha(t)$	exposed	S1	0.7	-82.2	-72.1	n.a.	3.5	<i>mm</i>	291	[59]
2XGA	2.30	(A)19	Cys	I	buried	all	1.0	-76.8	-61.0	-111.7	3.4	<i>mm</i>	120	[121]
2XGA	2.30	(B)19	Cys	I	buried	all	1.0	-84.0	-62.6	-114.7	3.5	<i>mm</i>	120	[121]
2XIU	1.50	(A)55	Thr	I	exposed	all	0.70	-62.5	-61.2	-87.4	3.5	<i>mmm</i>	100	[119]
2XIU	1.50	(B)55	Thr	I	exposed	all	0.87	-73.9	-53.1	-96.7	3.4	<i>mmm</i>	100	[119]
3M8B	2.44	(A)10	Val	I	exposed	S1	0.6	56.4	69.1	n.a.	3.5	<i>pp</i>	90	[138]

3M8D	2.44	(A)10	Val	l	exposed	S1	0.75	49.1	59.8	n.a.	3.6	<i>pp</i>	90	[138]
3RGM	2.60	(A)156	Thr	β	lipid	S1	0.6	176.3	-83.0	n.a.	3.7	<i>tm</i>	90	[123]
3RGN	2.30	(A)371	Trp	β	lipid	SG	0.65	46.9	n.a.	n.a.	3.6	<i>p</i>	90	[123]
3MPN	2.25	(A)177	Phe	α	lipid	S1	1.0	-69.0	-56.6	n.a.	3.3	<i>mm</i>	100	[124]
2MPQ	2.25	(A)204	Ile	α	lipid	all	1.0	-69.2	-58.9	-87.2	3.5	<i>mmm</i>	100	[124]
3STZ	2.50	(C)82	Tyr	l	exposed	S1,C4,C3	0.8	58.8	-88.8	90.7	4.1	<i>pmp</i>	n.a.	[139]
3V3X	2.00	(A)8	Asn	β	exposed	all	1.0	-56.6	-120.4	-87.5	4.2	<i>m?m</i>	100	[120]
3V3X	2.00	(C)8	Asn	β	exposed	all	1.0	-56.3	-74.3	-76.3	3.6	<i>mmm</i>	100	[120]
3V3X	2.00	(D)8.A	Asn	β	exposed	all	0.5	-59.3	162.0	72.3	4.3	<i>mtp</i>	100	[120]
3V3X	2.00	(D)8.B	Asn	β	exposed	all	0.5	-63.1	-57.6	-82.5	3.4	<i>mmm</i>	100	[120]
3V3X	2.00	(B)28	Lys	α	exposed	S1,C4,C3	1.0	-61.0	-63.1	111.7	3.5	<i>mmp</i>	100	[120]
3V3X	2.00	(C)28	Lys	α	exposed	all	1.0	-59.5	-61.4	95.6	3.3	<i>mmp</i>	100	[120]
3V3X	2.00	(D)28	Lys	α	exposed	S1	1.0	-58.2	-66.4	n.a.	3.5	<i>mm</i>	100	[120]
4EK1	1.97	(A)48	Ser	l	exposed	all	1.0	72.3	-106.9	93.2	4.1	<i>p?p</i>	100	[122]
4EK1	1.97	(B)48	Ser	l	exposed	all	1.0	75.0	-118.6	101.9	4.1	<i>p?p</i>	100	[122]

4EK1	1.97	(A)190	Ser	I	buried	all	1.0	-71.0	-79.7	-92.2	3.7	<i>mmm</i>	100	[122]
------	------	--------	-----	---	--------	-----	-----	-------	-------	-------	-----	------------	-----	-------

Table 2

Root mean square deviation and mean deviation (in parentheses) in Å of different predictions of mean label-to-label distances in proteins with known structures from experimental values. N is the number of experimental distances observed in a given protein.

PDB ID (N)	MMM	Pronox	MtsslWizard	R_{xray}^a	$R_{MD\alpha}^b$	R_{DFT}^c
2LZM (32)	2.4 (0.1)	2.4 (-1.0)	2.8 (0.9)	1.7 (0.4)	2.0 (0.3)	3.0 (1.1)
1TZY (16)	3.9 (1.3)	3.4 (-1.6)	3.2 (0.7)	3.3 (0.2)	3.6 (1.1)	4.8 (-0.7)
1HYN (6)	3.8 (-0.4)	2.8 (-0.3)	2.9 (0.2)	4.0 (-1.1)	4.1 (0.2)	4.2 (-2.0)
2MBW (4)	2.5 (-1.2)	1.7 (1.1)	2.8 (2.6)	1.6 (-0.8)	1.3 (-0.2)	2.5 (-2.2)
1C3W ^d (5)	4.3 (-2.4)	3.5 (0.8)	5.0 (2.9)	4.5 (-0.4)	4.4 (-1.3)	5.2 (-0.9)
1GO3 (6)	2.4 (-1.0)	1.6 (-0.3)	1.8 (0.6)	2.5 (1.3)	1.2 (0.5)	3.3 (-0.2)
2R19 (4)	2.3 (-0.2)	2.2 (-0.1)	2.6 (0.8)	2.3 (-0.4)	2.3 (-0.3)	2.5 (-0.1)
3B60 (19)	3.8 (-0.3)	4.8 (-0.8)	4.8 (1.4)	4.6 (0.2)	3.9 (0.0)	5.2 (0.1)
all (92)	3.2 (-0.1)	3.3 (-0.6)	3.5 (1.1)	3.2 (0.1)	3.1 (0.3)	4.1 (0.0)

^a MMM, relative rotamer populations adapted to the frequency of their occurrence in x-ray structures, forgive factor 0.40, ^b MMM, relative rotamer populations adapted to their frequency in MD trajectories for a labeled α -helix [52], forgive factor 0.40, ^c MMM, relative rotamer populations with Boltzmann populations according to their relative energies predicted by DFT computations [97], forgive factor 0.65, ^d lipid molecules removed for MMM and MtsslWizard.

Table 3

Systematic deviations of predicted mean label-to-label distances (Å) from experimental values in proteins with known structures.

PDB ID	site 1	site 2	exp.	MMM	Pronox	MtsslWizard	R _{xray} ^a	R _{MDα} ^b	R _{SF} ^c	Ref.
2LZM	(A)60	(A)90	37.8	44.6	45.5	45	42.9	43.7	n.a.	[134]
2LZM	(A)60	(A)94	25.5	32.1	33.5	31	28.2	30.8	n.a.	[134]
2LZM	(A)60	(A)109	35.2	35.5	39.6	39	36.6	36.1	n.a.	[134]
2LZM	(A)60	(A)154	34.1	38.1	39.6	38	37.4	37.7	n.a.	[134]
1HYN	(P)341	(Q)341	31.9	38.0	41.7	37.1	38.1	37.9	n.a.	[128]
1C3W	(A)201	(A)226	37	43.5	45.7	44.6	42.1	42.8	n.a.	[136]
2R19	(A)164	(B)164	40	48.9	51.2	48.7	51.8	49.1	n.a.	[135]
3B60	(A)61	(B)61	40	49.8	53.5	52.1	55.0	51.7	41.0	[100]
3B60	(A)539	(B)539	27	15.2	18.9	14.0	19.9	18.9	14.4	[100]
3B60	(A)116	(B)116	26	17.5	17.4	13.2	19.6	17.3	27.3	[132]
3B60	(A)143	(B)143	26	37.4	40.9	42.8	38.6	36.1	32.3	[132]
3B60	(A)162	(B)162	51	59.6	63.8	61.7	61.8	60.3	48.6	[132]

^a MMM, relative rotamer populations adapted to the frequency of their occurrence in x-ray structures, forgive factor 0.40, ^b MMM, relative rotamer populations adapted to their frequency in MD trajectories for a labeled α -helix [52], forgive factor 0.40, ^c MtsslWizard using the 'snuggly fit' mode.

Figure captions

Figure 1

Structures of spin labels used in proteins and lipids. The TOAC and TOPP labels can be introduced in peptide synthesis, while DOXYL can be introduced during synthesis of a lipid or detergent. The other labels react with free cysteine residues under mild conditions.

Figure 2

Structure of residue R1 formed by reaction of the MTSL label (PDB ligand code MTN) with a cysteine residue. Side group dihedral angles χ_1 - χ_5 are marked in red and the atom identifiers in PDB files are given as green labels, with O, C, N, CA, CB, and SG belonging to the cysteine residue and other atoms to the MTN ligand. The x and y axis of the nitroxide g tensor frame are shown in blue, with the z axis being perpendicular to the plane and pointing to the reader.

Figure 3

Ribbon model of T4 lysozyme (green), based on structure PDB:2LZM, with simulations of R1 rotamers (transparent stick models) by the MMM software [39-41]. Red ($\chi_3 > 0$) and blue ($\chi_3 < 0$) spheres denote loci of the N-O bond midpoint of the MTN ligand, their volume being proportional to the predicted populations. Visualized with MMM.

Figure 4

Time scales of the most important dynamic processes in a protein (top) and of the internal motions of a spin label (bottom).

Figure 5

Distribution of dihedral angles in the spin-labeled residue R1 formed by reaction of the MTSL label with a cysteine residue (Fig. 2) predicted by MD of the unrestrained residue (black lines)

[41], MD of a poly(alanine) α -helix labeled at its central residue (red bars) [52], DFT computations of the energies of solvated rotameric states (green bars) [97], and observed in crystal structures of proteins labeled with MTSL (blue lines, see Table 1).

Figure 6

Frames and transformations for stochastic Liouville simulations of spin label dynamics. Spectral response and relaxation times are fully described by the time-dependent set of Euler angles $\Omega_{LG}(t)$ denoting the transformation between laboratory frame and the nitroxide-fixed g tensor frame. The most elaborate slowly relaxing local structure (SRLS) model involves intermediary transformations (i) $\Omega_{LC}(t)$ to a protein-fixed diffusion frame, (ii) $\Omega_{CC'}$ to a protein-fixed director frame, where an ordering potential for the label is defined, (iii) $\Omega_{CM}(t)$ to a nitroxide-fixed diffusion frame, and (iv) Ω_{MG} to the nitroxide-fixed g tensor frame. Shortcuts can be taken to describe anisotropic rotational diffusion without or with an ordering potential (using ($\Omega_{LM}(t)$) or (Ω_{LC}) respectively), the latter choice corresponding to the MOMD model. Adapted from [68].

Figure 7

Effect of biasing of the rotamer distribution on the predicted distance distribution for spin-labeled residues 72 and 131 in T4 lysozyme (MMM simulations based on structure PDB:2LZM). The distance distributions are normalized with respect to their integral. Black line: Original ambient temperature rotamer library. Blue line: Only rotamers with $\chi_3 < 0$. Red line: Only rotamers with $\chi_3 > 0$. Green line: Populations of non-interacting rotamers rescaled to the distribution observed in MD simulations of an MTSL-labeled poly(alanine) α helix [52].

Figure 8

(a) Comparison between predicted (colored curves) and experimental (black curves) distance distributions, and (b) primary Q-band DEER data for T4 lysozyme with spin labels at residues 72

and 131. The predictions were computed by MMM [39-41] based on structure PDB:2LZM with the original ambient temperature rotamer library (dotted blue lines), with a library biased to match the rotamer distribution observed in x-ray structures (solid blue lines), and the low-temperature (175 K) rotamer library (solid purple lines), with PRONOX [109,110] without (dotted red lines) and with (solid red lines) bias favoring dihedral angles χ_1 and χ_2 observed in x-ray structures, and with MtsslWizard [107,111] with loose (solid green lines) and tight (dotted green lines) settings. Distance distributions are normalized with respect to their integral. Superposition with the primary data is not shown for the worst-fitting cases of the two PRONOX and MtsslWizard predictions.

Figure 9

Comparison between predicted (colored curves) and experimental distance distributions (black curves) in rhodopsin, based on data from Alexander et al. [130] and the structure PDB:1GZM. Predictions were made with the ambient temperature rotamer library of MMM (blue) [39-41], with PRONOX [109,110] in the low-clash mode (red) and with MtsslWizard [107,111] in the recommended mode (green). Distance distributions are normalized with respect to the maximum probability density.

Figures

Figure 1

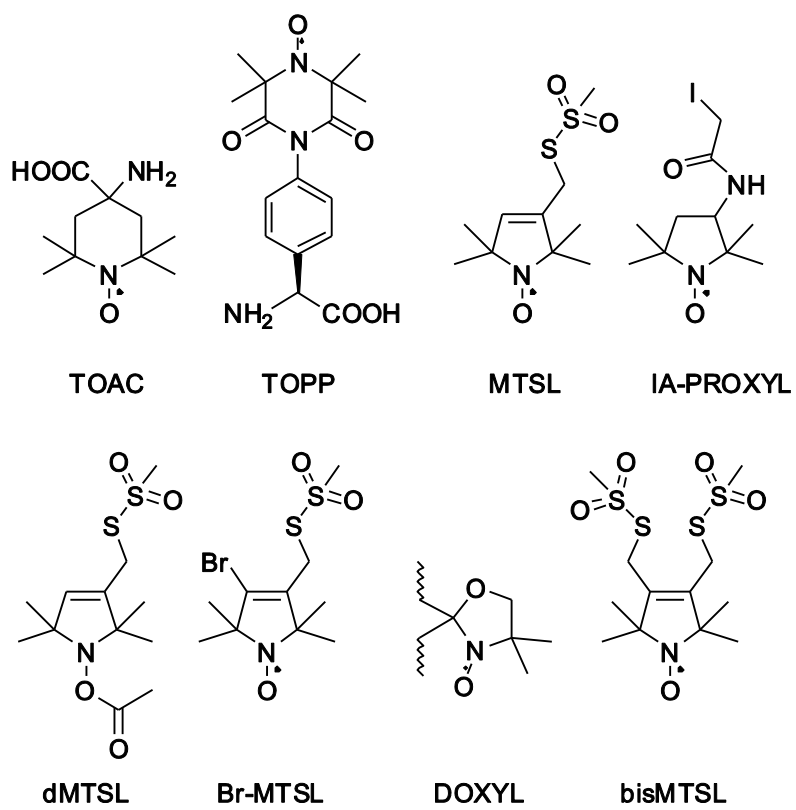


Figure 2

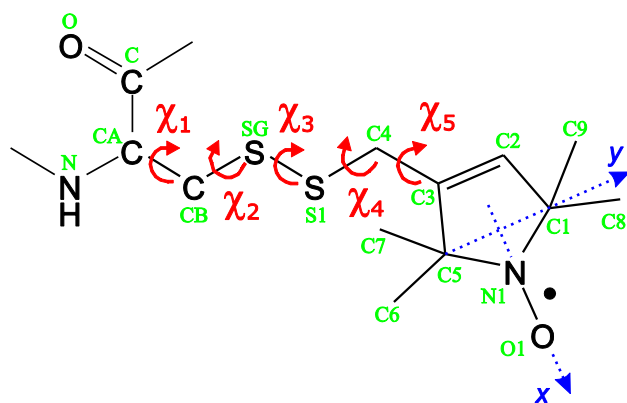


Figure 3

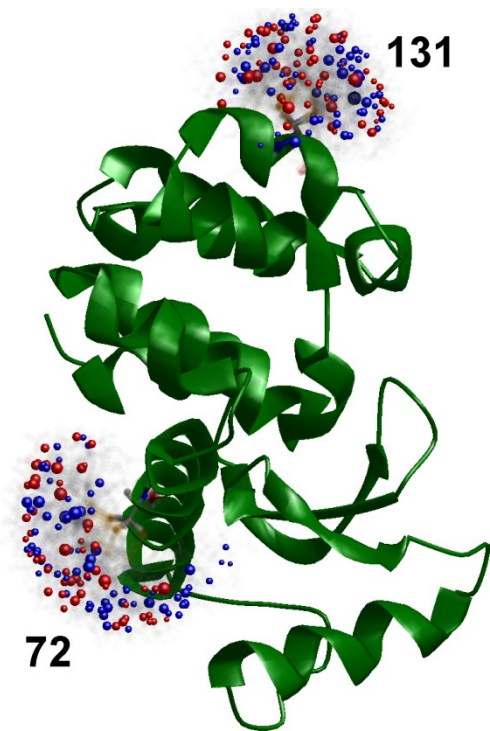


Figure 4

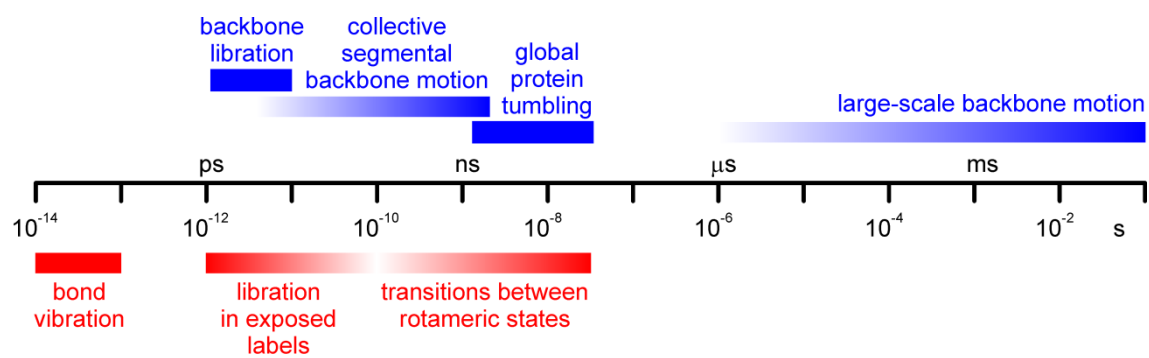


Figure 5

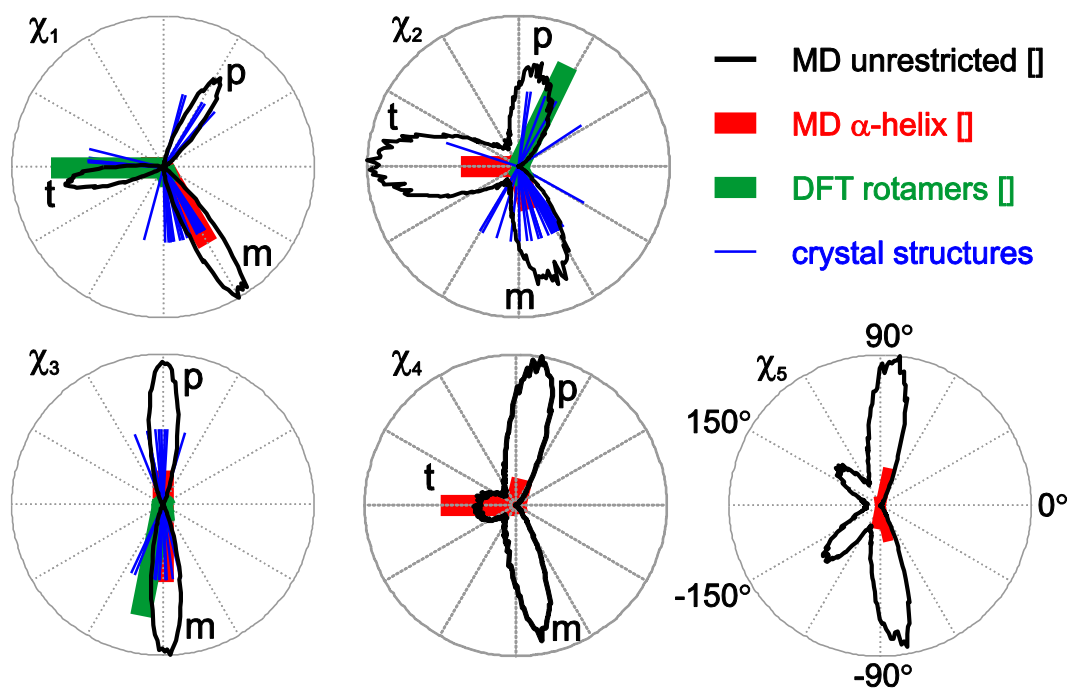


Figure 6

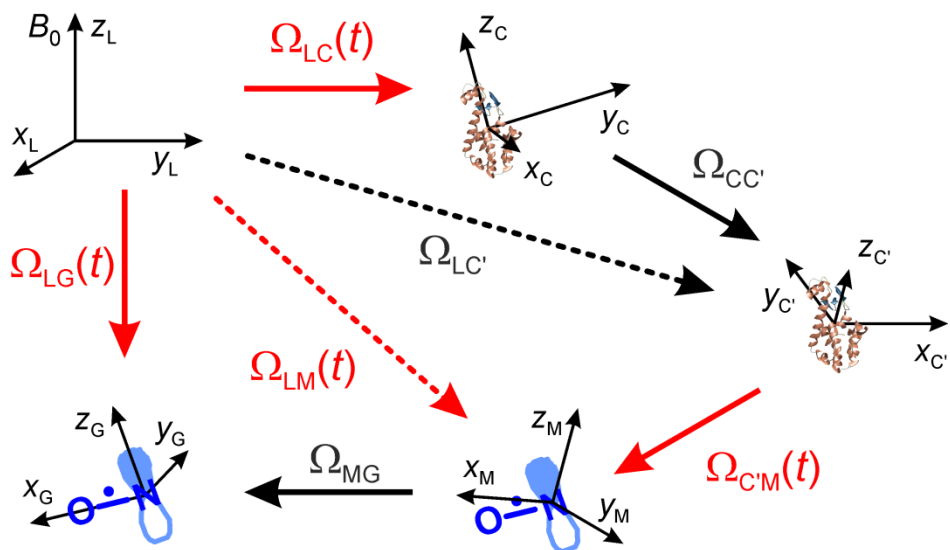


Figure 7

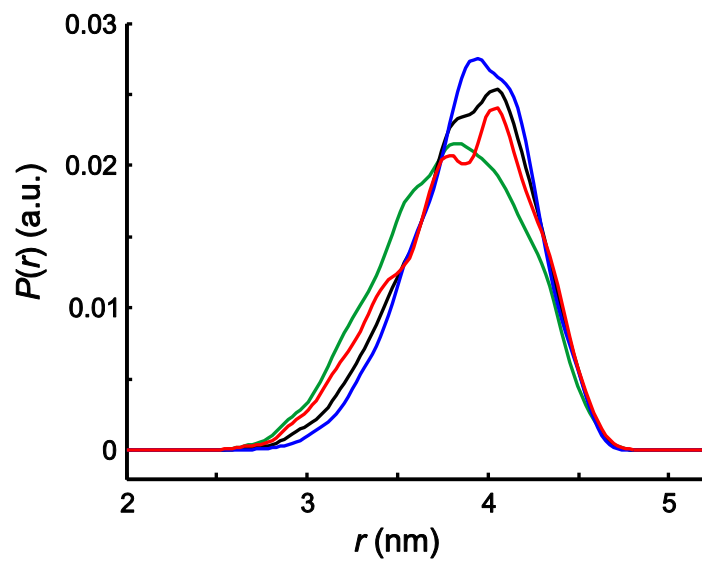


Figure 8

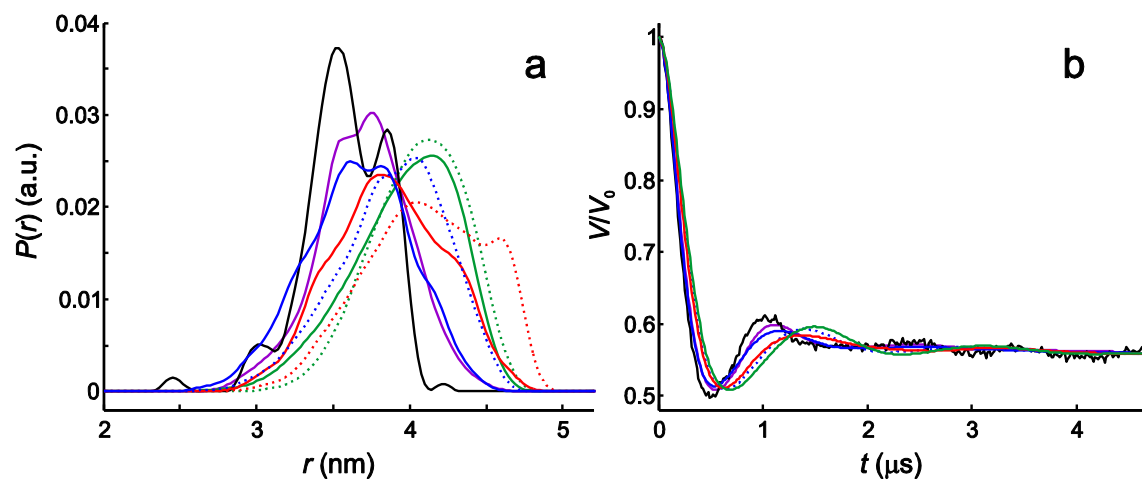


Figure 9

