THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Structural Refinement from Restrained-Ensemble Simulations Based on EPR/DEER Data: Application to T4 Lysozyme

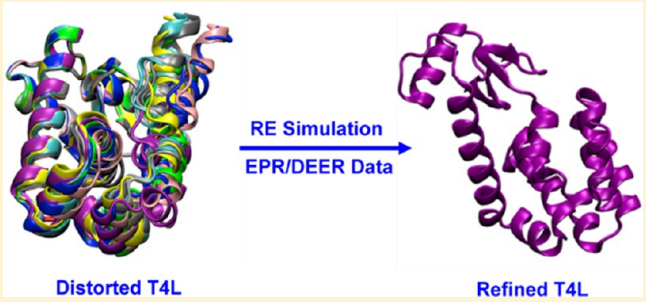
Shahidul M. Islam,[†] Richard A. Stein,[‡] Hassane S. Mchaourab,[‡] and Benoît Roux*[,†]

[†]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, United States

[‡]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee 37232, United States

Supporting Information

ABSTRACT: DEER (double electron−electron resonance) is a powerful pulsed ESR (electron spin resonance) technique allowing the determination of distance histograms between pairs of nitroxide spin-labels linked to a protein in a native-like solution environment. However, exploiting the huge amount of information provided by ESR/DEER histograms to refine structural models is extremely challenging. In this study, a restrained ensemble (RE) molecular dynamics (MD) simulation methodology is developed to address this issue. In RE simulation, the spin−spin distance distribution histograms calculated from a multiple-copy MD simulation are enforced, via a global ensemble-based energy restraint, to match those obtained from ESR/DEER experiments. The RE simulation is applied to 51 ESR/DEER distance histogram data from spin-labels inserted at 37 different positions in T4 lysozyme (T4L). The rotamer population distribution along the five dihedral angles connecting the nitroxide ring to the protein backbone is determined and shown to be consistent with available information from X-ray crystallography. For the purpose of structural refinement, the concept of a simplified nitroxide dummy spin-label is designed and parametrized on the basis of these all-atom RE simulations with explicit solvent. It is demonstrated that RE simulations with the dummy nitroxide spin-labels imposing the ESR/DEER experimental distance distribution data are able to systematically correct and refine a series of distorted T4L structures, while simple harmonic distance restraints are unsuccessful. This computationally efficient approach allows experimental restraints from DEER experiments to be incorporated into RE simulations for efficient structural refinement.



INTRODUCTION

Structural information is critical to understand the function of proteins. While X-ray crystallography is the best technique to obtain high-resolution structural information, it is also necessary to examine the protein in its native environment, free from the constraints imposed by the crystal lattice, to achieve a complete picture of its function. Two important spectroscopic approaches occupy a central role in these efforts, nuclear magnetic resonance (NMR) and electron spin resonance (ESR). One advantage of NMR is that it can report directly on the protein conformation through the nuclei in the system. In contrast, ESR requires the introduction of spectroscopic probes into the system via site-directed spin-labeling (SDSL) techniques, with the possible drawback that these could introduce unwanted perturbations. On the other hand, NMR investigations are limited by the size of the protein system, whereas the strong signal from the spin-label unpaired electron confers a sensitivity advantage to ESR even in the case of extremely large macromolecular complexes. Of particular interest, DEER (double electron−electron resonance) is a powerful ESR technique that makes it possible to measure the distance histogram between a pair of spin-labels inserted in a macromolecule.[1−3] While the experimental distance histograms between pairs of spin-labels contain a lot of information, designing a structural refinement strategy able to exploit the huge amount of data available from ESR/DEER is confronted with a number of challenges.

Any probe-based spectroscopic technique requires one to translate experimental measurements into valid information about the protein structure. In the case of ESR spectroscopy, such an operation is complicated by the internal dynamics of the spin-labels, rendering the extraction of detailed structural and dynamic information of the protein difficult. The most commonly used nitroxide spin-label used in ESR is MTSSL (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl methanethiosulfonate), which is typically linked to a cysteine residue in the protein through a disulfide bond (Figure 1A). The MTSSL moiety has five dihedral bonds, resulting in a highly flexible side chain once it is linked to a protein. A reliable characterization of the dynamical properties of the spin-label is essential to fully exploit the available structural information from ESR/DEER spectra. The well-characterized soluble protein, T4 lysozyme (T4L, Figure 1B), has provided an important model system in the characterization of the conformational properties of MTSSL. Structural and spectroscopic analyses revealed general properties

Received: November 28, 2012
Revised: March 13, 2013
Published: March 19, 2013

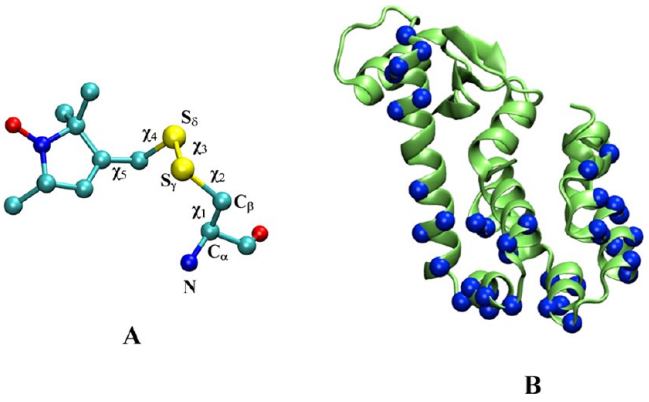


Figure 1. (A) Side chain resulting from linking MTSSL (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl methanethiosulfonate) to a cysteine through a disulfide bond. (B) Cartoon representation of T4 lysozyme with the 37 spin-labeled sites (colored in blue).

regarding the spin-label motion in proteins, captured by the "χ_4/χ_5 model" for solvent-exposed helix surface (SEHS) sites in T4L.[4] According to this model, internal motion of the spin-label side chain is essentially restricted to rotations about the dihedrals χ_4 and χ_5 and the remaining dihedrals are effectively "locked" on the ESR time scale. The χ_3 disulfide torsion is opposed by a large energy barrier,[5] while the χ_1 and χ_2 torsions are hindered by the formation of a hydrogen bond between the sulfur atom of the spin-label and the backbone amide[6] and C_α.[7] Such sulfur−backbone contacts are indeed observed in a number of X-ray crystal structures of T4L with spin-labels[8−15] in support of the χ_4/χ_5 model. Despite these efforts, there remains a paucity of data on the population of accessible rotamers of the MTSSL spin-label.

Computations can provide a "virtual route" to link the atomic structures to the experimental ESR observations. In principle, quantum mechanical ab initio methods offer the most accurate approaches to characterize the energy and conformations of the spin-labels.[16−18] However, these methods are generally too onerous to treat large protein systems and account for thermal fluctuations. Molecular dynamics (MD) simulations based on molecular mechanical force fields offer a realistic alternative strategy to calculate ESR observables.[19,20] For example, MD simulations were able to quantitatively reproduce multifrequency spectra of spin-labels attached to T4 lysozyme with remarkable accuracy.[21] Furthermore, the distance distributions from ESR/DEER experiments have been used to ascertain the validity and accuracy of MD simulations and structural models.[2,3,22−27] Computational modeling methods, such as the Multiscale Modeling of Macromolecular systems (MMM) software package of Yevhen Polyhach and Gunnar Jeschke,[28,29] and the PRONOX algorithm of Hatmal et al.,[30] have been developed to determine the interlabel distances distributions based on the analysis of spin-label rotamers. Attempts were also made to build structural models based on elastic network model and MD simulations that satisfies ESR/DEER distance restraints.[23] In spite of these efforts, there remains a need to develop a robust and effective computational method for making best use of ESR/DEER data in structural refinement.

Translating the ESR/DEER distance histograms into constraints that can be utilized in structural refinement is not straightforward. All the previous computational simulation studies[22−26] and modeling methods[28,30] use the ESR/DEER distance histogram in postanalysis to assess the correctness of models that were generated independently from the experimental data. In other words, none of the existing methods "drive" the structural model toward a 3D conformation that satisfies the ESR/DEER data. The standard approach to incorporate experimental knowledge of any average property, $\langle q(\mathbf{x})\rangle = Q$, consists in simulating the system in the presence of an artificial energy restraint, $k(q(\mathbf{x}) - Q)^2$, that biases the configuration $\mathbf{x}$ of the system toward the desired outcome.[31] For example, this is how NOE (nuclear Overhauser effect) distances are imposed in NMR refinement. This approach, however, is not applicable in the case of ESR/DEER data; the measurements report distance histograms and trying to ascribe a single average distance between any pair of atoms (e.g., backbone C_α) is not a valid approximation. A more sophisticated strategy is based on the notion of "restrained-ensemble" (RE) molecular dynamics (MD) simulations introduced by Vendruscolo and co-workers,[32−36] and exploited by Im and co-workers to determine the structure of membrane-bound peptides on the basis of solid-state NMR observables.[37−40] Such restrained-ensemble MD simulation scheme consists in carrying out parallel MD simulations of N replicas of the basic system in the presence of a biasing potential that enforces the ensemble average of a given property toward its known experimental value. Recently, the formal equivalence between the results from restrained-ensemble MD simulation scheme and the maximum entropy method for biasing thermodynamics ensembles[41] was established.[42] This analysis led to the development of a novel method to incorporate ESR/DEER distance histogram data into multiple-copy restrained-ensemble MD simulations.[43] In the RE simulations, distributions of the spin−spin distances are constrained by the spin pair distance histograms obtained from ESR/DEER experiments. The methodology based on RE simulations is designed to enforce a given set of real-space distance histograms for the purpose of structural refinement, although in practice, such distance histograms are obtained by postprocessing the measured time-dependent ESR/DEER signal.[29] The construction of distance histograms from DEER measurements is a challenging inverse problem that necessitates the utilization of specialized regulation procedures. While our main focus is to enable structural refinement based on experimental data, it is important to keep in mind that the distance histograms can be sensitive to the treatment of the inverse problem.

In the present study, detailed RE simulations based on ESR/DEER data for 51 spin-label pairs inserted at 37 positions in T4L are used to provide a rich source of detailed information about the rotameric states of the MTSSL spin-labels. Using this unprecedented new information from RE simulations and ESR/DEER data, a simplified nitroxide dummy spin-label is then parametrized for the purpose of structural refinement. Finally, the usefulness of the method is illustrated by showing that RE simulations with the dummy nitroxide spin-labels imposing the ESR/DEER experimental distance distribution data are able to systematically correct a series of distorted T4L structures.

METHODS

(a). ESR Experiments. Nitroxide spin-labels were inserted at 37 different positions in T4 lysozyme (T4L), for a total of 51 different pairs for DEER experiments. The positions of the spin-labels in the T4L structure and their relationship to secondary structure elements are summarized in Table 1. All double mutants of T4L used in this study are shown in the Supporting Information, Table S1. Cysteine mutants generated for this study were introduced into the cysteine-free T4L DNA by site-directed

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

Table 1. Spin-Label Positions in T4 Lysozyme, Number of Spin−Spin Distance Distributions and Spin-Label Rotameric States Obtained from RE Simulation

secondary structure	no. of spin-labels	no. of ESR/DEER histograms	labeled residues	χ_1/χ_2 rotamer	
				$\chi_3 = +90°$ (p)	$\chi_3 = -90°$ (m)
turn (59)	1	1	59	*tp*	–
H1 (60−80)	9	25	60	–	*mm*
			61	–	*mm*
			62	*mm*	*mm*
			64	*mt*	–
			65[b]	–	*mm*
			72[b]	*mm, mt*	–
			75	*mt*	–
			76[b]	–	*mm, tm*
			79	*mt*	–
H2 (82−90)	6	8	82[b]	–	*mm*
			83	–	*mm*
			85	–	*mm*
			86	*mt*	–
			89	–	*mm*
			90	–	*tt*
H3 (93−106)	2	6	93	–	*mm*
			94	*mt*	–
H4 (108−113)	3	6	108	–	*mt, mm*
			109	–	*mm*
			112	*mt*	–
H5 (115−123)	5	13	115[b]	*mt*	–
			116	*mt*	–
			119[b]	*mt*	–
			122	*mt*	–
			123	*pp*[a]	*pt, pp*[a]
H6 (126−134)	5	15	127	–	*mm*
			128	*mt*	–
			131[b]	–	*mm*
			132	*mt*	*mm*
			134	*mt*	–
turn (135)	1	2	135	–	*mm*
H7 (137−141)	1	1	140	–	*mm*
H8 (143−155)	3	13	151[b]	*mt*	*mm*
			154	*mt*	–
			155	–	*mm*
turn (159)	1	1	159	*mt*	–

[a]These rotameric states were populated by less than 30% [b]A crystallographic X-ray structure is available for a spin-label at this position (Table 5).

mutagenesis.[3] Protein expression, purification, and labeling with MTSSL were carried out as previously described.[1,3,6] DEER measurements were preformed on a Bruker 580 pulsed ESR spectrometer operating at either X- (9.6 GHz) or Q- (33.4 GHz) band using a standard four-pulse protocol.[44] Data was collected with the samples at 83 K with 23% (v/v) glycerol as cryoprotectant. Analysis of the DEER data to determine the distance distributions was carried in DeerAnalysis2010 or DeerAnalysis2011.[45] The data was fit with Tikhonov regularization and L-curve determination of the optimal regularization parameter.[46] The distance distributions were subsequently converted to 1 Å bins and normalized for use in the RE simulations using a script implemented in MATLAB.

(b). Computational Details. Molecular dynamic simulations of spin-labeled T4L were carried out with the CHARMM[47] and NAMD[48] programs, using the all-atom CHARMM27 protein force field[49] with the CMAP corrections.[50] The force field parameters for the nitroxide spin-label developed by Sezer et al.,[19] which were found to provide very accurate multifrequency ESR spectrum in previous studies,[19−21] were employed for the description of the nitroxide spin-labels. The crystal structure of native T4L (2LZM)[51] was used to construct the initial geometry of the T4L system with its spin-labels. Since 37 sites in total were experimentally labeled, it is not possible to introduce all of the nitroxide spin-labels simultaneously into a single protein without extensive steric clashes. To avoid this problem, four copies of the protein (A, B, C, D) were included in the simulation cell (Figure 2) and the 37 spin-labels were distributed among the four copies

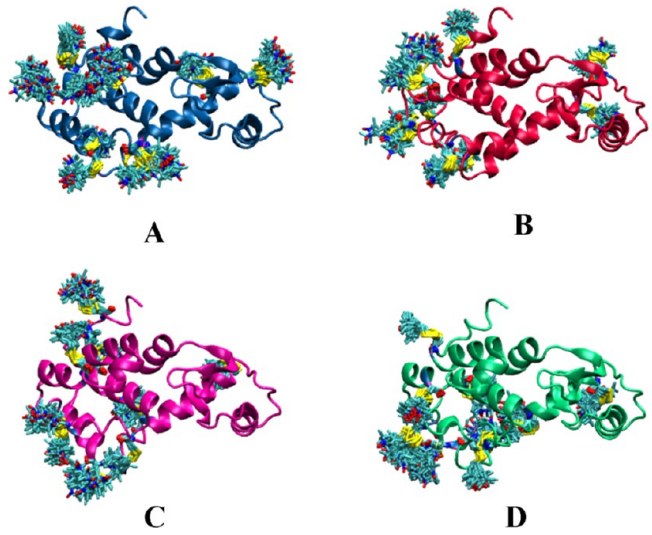


Figure 2. Simulation setup consisting of 37 spin-labeled side chains distributed in four T4L systems. The distance between the center of mass of each T4L is 80 Å from each other; thereby, they are far enough not to have any interaction. Similarly, the spin-label side chains were also kept at least four residues apart with a distance of about 10−15 Å from each other; therefore, the interaction between the spin-labels are also minimum. The copies of the spin-labels were generated with CHARMM using the command REPLICA.

to ensure that they do not interact with each other. Each copy comprises 9, 9, 9, and 10 spin-labels, respectively. The labeled residues are as follows: [59, 64, 76, 83, 89, 108, 128, 135, 140, 154] in A, [60, 65, 85, 90, 115, 123, 127, 131, 151] in B, [62, 72, 79, 86, 93, 112, 119, 132, 155] in C, and [61, 75, 82, 94, 109, 116, 122, 134, 159] in D. All spin-labels are separated by at least a distance of 10−15 Å. To ensure that the four proteins do not interact with each other, the center of mass of each adjacent T4L was separated by 80 Å. The whole system thus constructed has a total of 11 349 atoms, and is solvated by 41 547 TIP3P water molecules within a 160 × 60 × 130 Å rectangular box. The salt concentration was maintained at 0.15 mM/mol by adding 44 potassium and 117 chloride ions. For all the simulations, a weak positional harmonic restraint with a force constant of 1 (kcal/mol)/Å^2 was applied to the protein backbone to avoid any large displacement. Three different simulation methodologies were utilized in this study: conventional molecular dynamics (MD), locally enhanced sampling (LES),[52−54] and restrained-ensemble (RE) simulations.[43]

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

Conventional MD Simulation Protocol. Prior to LES and RE simulations, the simulation box was minimized and equilibrated and a 140 ns MD simulation was performed with NAMD.[48] Both equilibration and production simulations were performed under NPT conditions where the temperature was kept at 300 K and the pressure at 1 atm to remain consistent with experimental conditions. The simulations were performed with Langevin[55] thermostat to control the temperature of the simulation box. A collision frequency, γ, of 5.0 ps^{-1} was used for the Langevin thermostat. A dielectric constant of 1.0 was used during the simulation. Bonds involving hydrogen atoms in water were constrained to their equilibrium values using the SHAKE[56] algorithm. Periodic boundary conditions (PBC) were imposed and a cutoff of 10 Å was used for nonbonded interactions with a switch distance of 8 Å and pairlist distance of 12 Å. Long-range electrostatic behavior was controlled with the particle mesh Ewald (PME) method.[57,58] First, the conjugate gradient energy minimization was performed for 1000 steps to eliminate initial bad contacts between various atoms. This was then followed by a short equilibration run of 5 ns. Then the production MD simulation was performed for 140 ns from which the ensemble-averaged properties of the spin-labels were calculated. Even with such a long trajectory, it is possible that these ensemble average properties are weighted to the starting rotamers due to the lower probability of the interconversion of the disulfide bond.[5]

Locally Enhanced Sampling Simulation Protocol. The locally enhanced sampling (LES) method[52−54,59] aims at increasing the sampling efficacy of conformational space of selected groups of atoms relative to conventional MD simulations. LES consists in introducing N noninteracting copies (replicas) of selected groups of atoms, which then interact with the rest of the system with forces that are scaled down by a factor of $1/N$. According to an approximate mean-field argument, the statistical distribution of the configurations resulting from LES is expected to be representative of the single-copy conventional MD. To try improving the sampling efficacy of the nitroxide spin-labels conformations using the LES method, all 37 labeled residues were replicated 25 times including their backbone atoms. The multiple copies were generated using the PSFGEN module of VMD (Visual Molecular Dynamics)[60] visualization program. The temperature of the replicated atoms was scaled by a factor of $1/N$ and the masses of the enhanced atoms were also reduced by the same factor to restore the normal velocity distribution. The temperature of the rest of the molecule was kept at 300 K. The last frame from the 140 ns MD simulation was used to start the LES simulation; therefore, the system is already well equilibrated. The simulation was performed with NAMD[48] keeping the backbone atoms of the T4L restrained to their X-ray structure by applying positional restraints with a harmonic force constant of 1 (kcal/mol)/Å^2. The production LES simulation was performed for 7 ns after an initial 1000 steps of minimization and 1 ns of equilibration. All the other simulation parameters were kept the same to those used in the conventional MD simulation.

Restrained Ensemble Simulation Protocol. The restrained ensemble (RE) method was recently developed to exploit the information from experimental distance histograms obtained from ESR/DEER experiments.[43] The RE ensemble as implemented here builds on the framework provided by the multiple-copy LES simulation method. As in LES, an ensemble of N noninteracting replicas was created for each spin-label side chain of the basis system (Figure 2). The multiple replicas yield a total of N^2 distances for each pair of spin-labels. Based on these distances, an energy restraint was implemented in CHARMM[47] so that at every step of the simulation the histogram obtained from the N^2 spin−spin distances would be forced to match the experimental distance histogram obtained from the DEER experiment. The possibility of restraining the simulations from the time-dependent DEER signal rather than the postprocessed distance histograms was discussed previously.[43] To avoid discontinuities between the bins, a smooth Gaussian of width σ is used in the construction of the histograms. Let $\bar{h}^{ij}(n)$ be the ensemble-averaged "instantaneous" histogram produced by the multiple replica and $H^{ij}(n)$ be the target experimental histogram. We introduce the energy restraint on all the spin-label i−j pairs, in the system

$$U_{\mathrm{RE}} = \frac{1}{2}K \sum_{\{\mathrm{pair}\,ij\}} \sum_{\{\mathrm{bin}\,n\}} (\bar{h}^{ij}(n) - H^{ij}(n))^2 \tag{1}$$

where K is a large harmonic force constant used to enforce the restraint. The ensemble-average histogram $\bar{h}^{ij}(n)$ is calculated as

$$\bar{h}^{ij}(n) = \frac{1}{N^2} \sum_{s=1}^{N} \sum_{s'=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{(n\Delta r - |\mathbf{r}_i^s - \mathbf{r}_j^{s'}|)}{2\sigma^2}\right] \tag{2}$$

where $|\mathbf{r}_i^s - \mathbf{r}_j^{s'}|$ is the distance between the spin-label i and spin-label j. Both the experimental and the calculated histograms are recorded with the same bin size Δr of 1 Å, and are normalized to 1 (the distributions $\bar{h}^{ij}(n)$ and $H^{ij}(n)$ have dimension of inverse length, and the units of the force constant K is (kcal/mol)/Å^2). For details about the implementation of the RE method, see ref 43.

All the spin-labels were replicated 25 times using the command REPLICA in CHARMM. Interactions between the replicas and the rest of the system were scaled by 1/25 using the BLOCK command; interactions within each replica were not scaled. A harmonic force constant K of 10 000 (kcal/mol)/Å^2 was used to impose the energy restraint to match the calculated distributions $\bar{h}^{ij}(n)$ with those of the experiment $H^{ij}(n)$, and the natural spread σ of the Gaussian was set to 1.7 Å. A total of 51 spin−spin distance histograms obtained from ESR/DEER experiments were used simultaneously in the RE simulations of the system. Large displacements of the protein backbone atoms were prevented by applying positional restraints with a harmonic force constant of 1 (kcal/mol)/Å^2 relative to the X-ray structure. At each time step, the distance histogram between any selected pair of spin-labels was calculated from Cartesian coordinates of the four proteins translated to a unique reference frame. For example, there is an experimental DEER histogram for spin-labels at positions 62 and 109 even though they are located in protein C and D, respectively. To calculate the instantaneous ensemble-average histogram of the pair $\bar{h}^{62,109}(n)$, the C and D protein systems were translated to a unique reference frame and then the distance between the spin-label at position 62 and the spin-label at position 109 was calculated.

To reduce the size of the simulated system, only the water molecules within 35 Å from the center of mass of each T4L system were kept, for a total of 19 479. A spherical half-harmonic containing restraint with a force constant of 0.5 (kcal/mol)/Å^2 was used to keep the waters near the proteins. Both equilibration and production simulations were performed under NVT conditions where the temperature for both replica spin-label atoms and normal atoms were kept at 300 K. The rest of the simulation parameters were kept the same to those used in the

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

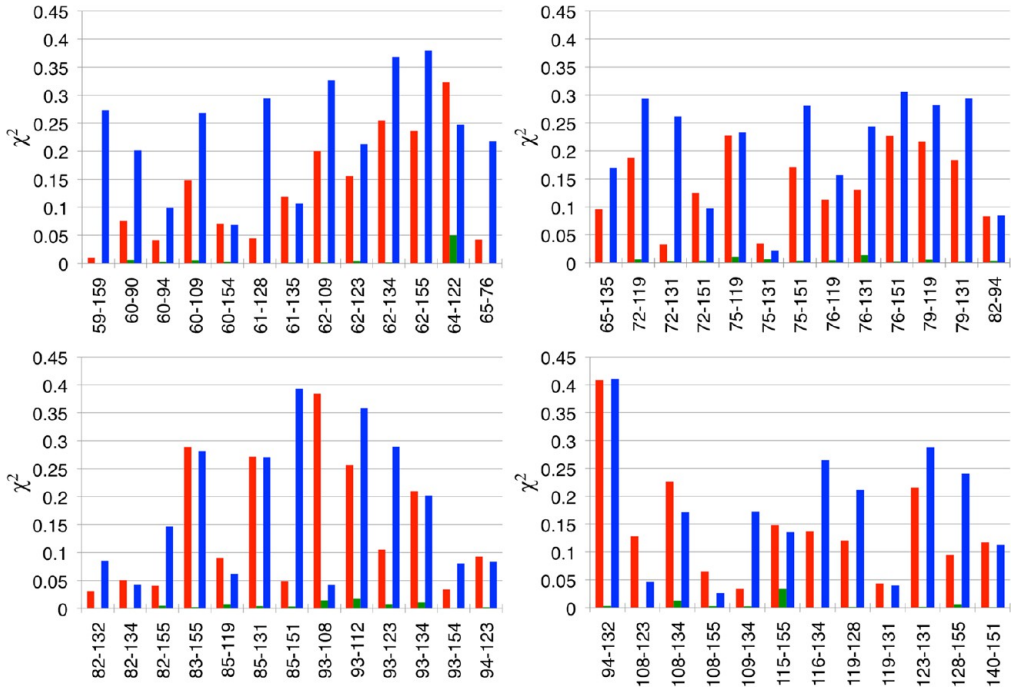


Figure 3. Difference in spin−spin distance histograms between experiment and conventional MD (red line), locally enhanced sampling (blue line), and restrained ensemble (green line) simulations.

conventional single-copy MD and LES simulations. To increase the efficiency, five realizations of the RE ensemble were simulated simultaneously. They were initiated from the frames at 60, 80, 100, 120, and 140 ns of the single-copy MD simulation after energy minimization and equilibration for 1 ns; 8 ns RE simulation was performed for each of the five systems, providing a total of 40 ns RE simulation.

Data Analysis Protocol. Spin pair distances and dihedral angles were collected every 10 ps from 140 ns conventional single-copy MD simulation. Thus, a total of 14 000 data points were obtained for each spin pair distance and dihedral angle. These data were used to calculate the spin pair distance distributions and the rotamer population distributions along five dihedral angles of a particular spin-label. A total of 518 000 data points for every dihedral angle were obtained from 37 spin-labels. This enormous amount of data was used to calculate the global rotameric states of the spin-label side chain. For both LES and RE simulations, a total of 625 spin pair distances were obtained from a single simulation step since each spin-label is replicated to 25 copies. The spin pair distances were collected every 10 ps from the last 1 ns of the trajectory providing a total of 625 000 distances from the simulations, which were then used to calculate the distance distributions. Data for the dihedral angles were collected every 10 ps from the 7 ns LES and the 40 ns RE simulations. Since for every spin-label there are 25 copies, a total of 17 500 data points were obtained for every spin-label yielding a total of 647 500 data points from LES simulation. Following the same procedure, the 40 ns RE simulation provided a total of 987 500 data points for every spin-label and a total of 3 695 375 data points from all spin-labels. These data were used to calculate the dihedral angle distributions. The potential of mean force (PMF) or the change in free energy along the dihedral angles was also calculated by integrating the rotamer population distributions.

RESULTS AND DISCUSSION

As expected, the 51 spin-label distance histograms obtained from the multiple-copy restrained-ensemble (RE) simulation are very similar to those obtained from ESR/DEER data. It is, however, worth emphasizing that satisfying all the distance histograms simultaneously did not generate any unacceptable distortions of the spin-labels, which are attached to protein backbone restrained near the X-ray structure. This is strongly suggestive that the set of 51 distance histograms generated from the postprocessing of the time-dependent DEER signal is broadly consistent with the X-ray structure of T4L and the chemically acceptable range of conformational flexibility of the MTSSL spin-label. In contrast, the distance histograms obtained from conventional MD and the locally enhanced sampling (LES) differ considerably from experiment (see Figure S6 in the Supporting Information for all 51 distance distributions). It is possible to rapidly quantify differences in the 51 distance distributions obtain from MD, LES and RE simulation from the total mean-square deviation relative to the DEER data

$$\chi^2 = \sum_{\{\text{pair}ij\}} \sum_{\{\text{bin}n\}} (\bar{h}^{ij}(n) - H^{ij}(n))^2 \quad (3)$$

As shown in Figure 3, the values of χ^2 for most of the spin−spin distance histograms are large for MD and LES, while they are very small for RE. In most cases, χ^2 from LES is slightly larger than from MD, which is perhaps explained by the approximate nature of the multiple-copy mean-field argument underlying the method. Because the protein backbone was restrained to remain near the X-ray structure, these differences can be directly traced back to the conformations of the flexible spin-labels themselves. Clearly, the unbiased sampling from MD and LES produces spin-label conformations that are in fairly substantial disagreement with experimental data. The discrepancy may, in part, reflect inaccuracies in the molecular mechanical force field used in the simulations. However, it is also important to recall that the

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

samples for ESR/DEER experiments are fast-frozen, and the conformational probability of the flexible spin-labels would reflect the proper Boltzmann distribution at room temperature only if the freezing process was infinitely fast (quenching). On the other hand, if the freezing is slow on the time scale of some molecular motions, the conformational distribution will partly drift toward lower temperature population distributions during the process (annealing). Ultimately, it is likely that some slow degrees of freedom are truly quenched while fast degrees of freedom such as side-chain rotamers are partly annealed to local energy conformations.[61] The implication is that the information from the distance histograms imposed via the RE simulations significantly restricts the conformation of the spin-labels, which provides a unique opportunity to determine the conformational propensity of MTSSL nitroxide spin-labels in multiple positions of a soluble protein under the conditions of ESR/DEER experiments.

(a). Analysis of the T4L Simulations and Rotameric Distributions of MTSSL. The conformation of the MTSSL nitroxide spin-label is completely characterized by the five dihedral angles, χ_1, χ_2, χ_3, χ_4, and χ_5, along the flexible side-chain $C_\alpha-C_\beta-S_\gamma-S_\delta-C_\eta-C_\zeta$ (Figure 1A). The distribution about the χ_1, χ_2, χ_3, χ_4, and χ_5 of all the spin-labels obtained from MD, LES, and RE simulations is shown in Figure 4. For the sake of simplicity, it will be useful to adopt the following convention in the following discussion. The dihedral angles χ_1, χ_2, and χ_4 can adopt 3-fold conformations, +60° (or *gauche*$^+$), 180°(or *trans*), and −60° (or *gauche*$^-$), which will be denoted by *p*, *t*, and *m*, respectively. According to the RE simulations, the *m* rotamer for χ_1 is the most populated, while the rotamers from the conventional MD and LES simulations are more broadly distributed. Similarly, the *m* rotamer for χ_2 is also the most populated according to the RE simulations, with again a broader distribution from the conventional MD and LES simulations. The distributions of the dihedral angle χ_3 of the slowly converting $S_\gamma-S_\delta$ disulfide bond are similar for all three simulations. Essentially, χ_3 can adopt only two stable positions, around +90° and −90° (=270°), which will be denoted as *p* and *m*, respectively. The last dihedral angle, χ_5, between the chain and the ring is broadly distributed, with two positions slightly more populated around +90° and −90°, which will be denoted as *p* and *m*, respectively. The *p* and *m* states were visited for 22 spin-labels in the RE simulation, but only for 11 spin-labels in the conventional MD simulation (Figure S7 in the Supporting Information). Lastly, the rotamer distributions for χ_4 and χ_5 obtained from all simulation methods are very similar (Figure 4), suggesting that considerable fluctuations around these two dihedrals are consistent with the ESR/DEER data.

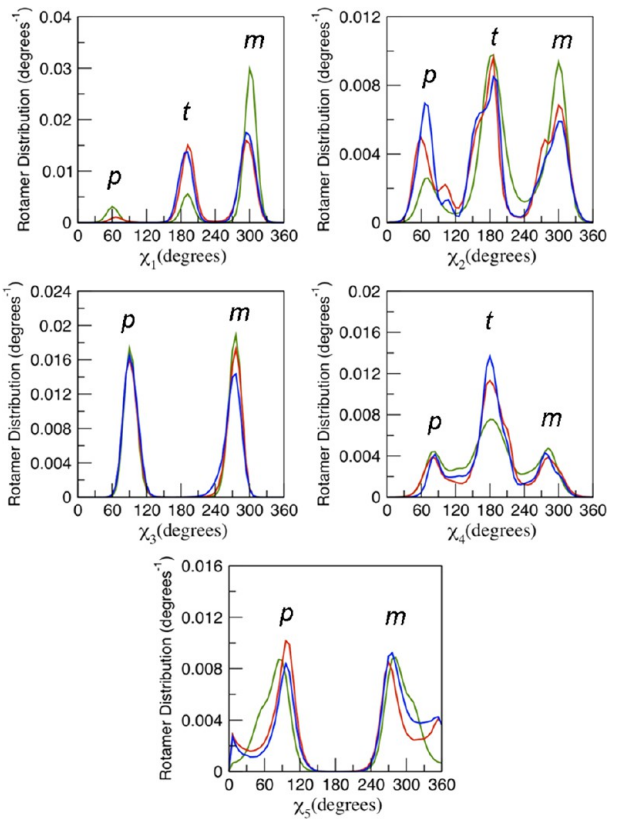


Figure 4. Rotamer population distributions of the dihedral angles χ_1, χ_2, χ_3, χ_4, and χ_5 of the MTSSL spin-label obtained from the conventional MD (red line), locally enhanced sampling (blue line), and restrained ensemble (green line) simulations. All the data from all the spin-labels attached to T4L were combined together to produce the distributions.

The joint 2D distributions along χ_1 and χ_2 for prescribed values of χ_3 obtained from the MD, LES and RE simulations are shown in Figure 5. The populations of the χ_1/χ_2 rotameric states are given in Table 2. According to all simulation methods, it is clear that the χ_1/χ_2 rotamer distribution is strongly affected by the χ_3 dihedral angle of the $S_\gamma-S_\delta$ bond. This is consistent with previous studies that have highlighted the correlations between the various dihedrals.[19,28] Only two predominant $\chi_1/\chi_2/\chi_3$ joint rotameric states are observed with the RE simulation, *mtp* (33%) and *mmm* (32%), whereas several are observed in the LES and MD simulations. This shows that sampling of the conformational space while restraining the spin-label pair via the ESR/DEER distance histograms leads to a considerable reduction in the observed rotameric states. The 2D PMF calculated from the distributions reveal that the free energy barrier between the *mt* and *mm* states is only about 1−2 kcal/mol (Figure S8 in the Supporting Information), suggesting the possibility of rapid interconversion rates for these dihedral angles. The 2D rotamer population distribution along χ_4 and χ_5 shows that they are only weakly correlated. Regardless of the simulation methods, the distributions along χ_4 and χ_5 are very similar and fairly independent of the value of χ_3 (Figure S9 in the Supporting Information). The populations of the χ_4/χ_5 rotameric states are given in Table 3. The results are consistent with the notion that the χ_4 and χ_5 dihedrals are very mobile, an underlying hypothesis of the so-called χ_4/χ_5 model,[4,7] which assumes that the internal motion of the spin-label is largely limited to rotations about the last two dihedrals. Nonetheless, as observed previously,[21] the χ_4 and χ_5 dihedrals display some intrinsic rotameric preferences. Overall, the LES method appears to produce distributions that cover somewhat less rotameric space than the conventional MD simulation. This might indicate that there are differences in the amount of sampling. Alternatively, it is possible that the results are adversely affected by the mean-field approximation that underlies the LES method.[52,59] Nevertheless, the differences are not sufficiently large to cause concerns, which is reassuring given that the same mean-field approximation is also used in the multiple-copy RE ensemble implementation based on ESR/DEER distance histograms.[43]

It is of interest to examine how the rotamer distribution at specific positions departs from the average trend. The results are summarized in Table 4. Due to its particular importance, it is helpful to first focus on the dihedral around the $S_\gamma-S_\delta$ bond, χ_3.

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

Table 2. Rotamer Populations (%) along χ_1 and χ_2 of the Spin-Label Attached to T4 Lysozyme at $\chi_3 \approx 90°$ and 270° Obtained from Conventional MD (MD), Locally Enhanced Sampling (LES), and Restrained Ensemble (RE) Simulations

χ_3	χ_1/χ_2	MD	LES	RE	χ_3	χ_1/χ_2	MD	LES	RE
	pp	1.10	0.00	1.51		*pp*	0.03	0.00	0.66
	tp	13.33	10.00	4.17		*tp*	5.04	12.95	3.89
	mp	1.20	0.09	0.90		*mp*	2.58	2.66	0.16
	pt	0.72	0.00	1.18		*pt*	1.25	0.02	4.23
p (+90°)	*tt*	11.31	8.35	1.10	*m* (270°)	*tt*	7.71	5.80	4.16
	mt	12.20	22.09	33.00		*mt*	8.44	7.20	6.13
	pm	0.0	0.00	0.01		*pm*	0.02	0.00	0.34
	tm	0.12	0.04	0.19		*tm*	6.74	5.27	1.10
	mm	9.25	10.89	5.63		*mm*	18.96	14.66	31.64
	total	49.23	51.46	47.48		total	50.77	48.56	52.31

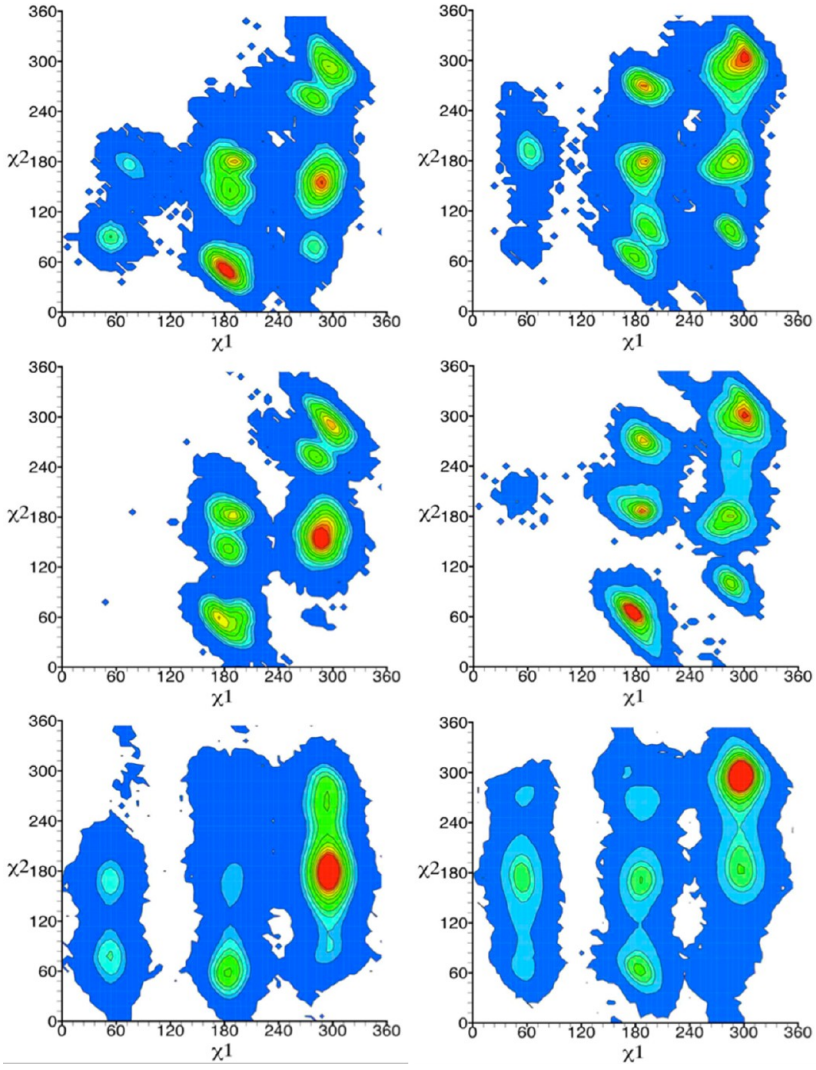


Figure 5. Rotamer population distribution along χ_1 and χ_2 of spin-label attached to T4 lysozyme at $\chi_3 \approx 90°$ (left panel) and 270° (right panel) obtained from conventional MD (upper panel), locally enhanced sampling (middle panel), and restrained ensemble (lower panel) simulations. The units of the angles χ_1 and χ_2 are in degrees. Population densities are indicated with contour lines with densities increasing from blue to red color.

Isomerization of the $S_\gamma-S_\delta$ bond is opposed by a large energy barrier[5] and is a relatively slow process on the time scale of ESR.[21] Furthermore, the *m*:*p* equilibrium ratio was previously shown to be a critical determinant of the line shape of CW-ESR spectra.[21] As observed in Table 4, χ^3 shows a preference for either +90° (*p*) or −90° (*m*) in most cases. Spin-labels at positions 59, 64, 72, 75, 79, 86, 112, 115, 116, 122, 128, 134, 154, and 159 mainly prefer +90° (*p*) with population more than 70%, while spin-labels at positions 60, 61, 65, 76, 83, 85, 89, 90, 93, 108, 109, 123, 131, 135, 140, and 155 mainly prefer −90° (*m*). Only 7 spin-labels (positions 62, 82, 94, 119, 127, 132, and 151) show a significant occupancy of both rotamers. When χ_3 is +90° (*p*), the preferred χ_1/χ_2 rotameric state is typically *mt* (Table 1 and Figure 6). This is observed at 16 positions along the α-helices, a number

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

Table 3. Rotamer Populations (%) along χ_4 and χ_5 of the Spin-Label Attached to T4 Lysozyme at $\chi_3 \approx 90°$ and 270° Obtained from Conventional MD (MD), Locally Enhanced Sampling (LES), and Restrained Ensemble (RE) Simulations

χ_3	χ_4/χ_5	MD	LES	RE	χ_3	χ_4/χ_5	MD	LES	RE
	pp	1.62	1.11	4.13		*pp*	1.33	0.27	5.59
	tp	16.44	15.65	12.65		*tp*	15.26	8.92	14.63
	mp	8.06	10.09	6.21		*mp*	5.89	2.46	5.31
	pt	0.62	0.47	0.07		*pt*	0.09	0.07	0.11
p (+90°)	*tt*	0.61	0.68	0.25	*m* (270°)	*tt*	0.72	0.36	0.30
	mt	0.20	0.18	0.09		*mt*	0.33	0.11	0.10
	pm	6.10	8.26	3.79		*pm*	7.71	5.45	8.20
	tm	12.71	9.71	14.16		*tm*	17.78	30.20	13.48
	mm	2.85	5.30	6.31		*mm*	1.68	0.70	4.60
	total	49.21	51.45	47.67		total	50.79	48.54	52.33

Table 4. Total Rotamer Populations (%) of χ_1 and χ_2 at $\chi_3 \approx 90°$ and 270° of the Spin-Label Attached to the T4 Lysozyme Obtained from Restrained Ensemble (RE) Simulations

residue	*p* (+90°)	*m* (270°)	residue	*p* (+90°)	*m* (270°)	residue	*p* (+90°)	*m* (270°)
59	92.78	7.22	86	89.23	10.77	127	39.42	60.58
60	17.86	82.14	89	1.20	98.80	128	82.71	17.29
61	4.74	95.26	90	8.65	91.35	131	7.18	92.82
62	41.57	58.43	93	4.22	95.78	132	59.82	40.18
64	87.14	12.86	94	68.26	31.74	134	84.17	15.83
65	18.79	81.21	108	6.75	93.25	135	5.73	94.27
72	81.82	18.18	109	7.48	92.52	140	25.25	74.75
75	99.63	0.37	112	98.30	1.70	151	49.47	50.53
76	7.16	92.84	115	100.00	0.00	154	73.75	26.25
79	84.13	15.87	116	96.52	3.48	155	6.53	93.47
82	41.28	58.72	119	63.04	36.96	159	93.39	6.61
83	4.47	95.53	122	78.40	21.60			
85	2.98	97.02	123	29.98	70.02			

of them being near the C-terminus (79, 94, 112, 122, 132, 134, 154). When χ_3 is −90° (*m*), the preferred χ_1/χ_2 rotameric state is typically *mm*. This is observed at 18 positions located along α-helices, a number of them being near the N-terminus (60, 61, 82, 83, 93, 108, 109, 127, 151). A few spin-labels (72, 108, 132, 151) also display both the *mm* and *mt* rotamers. The marked preference of χ_1/χ_2 for the *mm* and *mt* rotamers is consistent with the energy surface of the spin-label,[19] suggesting that the surrounding residues do not lead to large effects. Returning to Table 1, it is clear that spin-labels inserted within α-helices largely prefer the $\chi_1/\chi_2/\chi_3$ rotameric states *mmm* or *mtp* on average. The spin-labels that display a wider variability are either near the ends of an helix, in a loop, or buried. For example, spin-labels at positions 59, 90, and 123 are found to be in the unusual $\chi_1/\chi_2/\chi_3$ rotameric state *tpp*, *ttm*, and *ppp*, respectively. Position 59 is located in a loop while position 90 is buried within the solvent inaccessible protein core, and position 123 is the last residue of helix H5. Finally, regarding the χ_4/χ_5 rotameric state, the majority are either *tp* or *tm*. Only two positions are in the rotameric states *mm* (position 109) or *tp* (position 115) with a population greater than 50%. The predominance of these *tp* and *tm* rotamers is consistent with the initial parametrization of the spin-label.[19] This suggests that these rotamers are predominant unless there are steric clashes with neighboring residues.

(b). Comparison with Available Rotamer Data from X-ray Crystallography. Comparison with the available information from X-ray crystallographic structures of MTSSL spin-labels inserted into T4L[9−11] offers a powerful route to examine the validity of the rotamers populations extracted from the RE simulations. However, in several X-ray structures, the electron density along the flexible side-chain $C_\alpha-C_\beta-S_\gamma-S_\delta-C_\eta-C_\zeta$ is missing beyond the S_δ atom, leaving the position and the orientation of the remaining chain and nitroxide ring undefined. In the small number of X-ray structures in which the C_η is resolved, it is observed that the dihedral angle χ_3 of the $S_\gamma-S_\delta$ bond adopts either values of +90° or −90°. Transitions between these two well-identified states are opposed by a large energy barrier and occur infrequently according to computational analysis.[21,49] For this reason, comparison with X-ray data is mainly limited to χ_1/χ_2, corresponding to the region of the side chain nearest to the backbone $C_\alpha-C_\beta-S_\gamma-S_\delta$. The result of a comparison for spin-labels at nine positions (65, 72, 75, 76, 82, 115, 119, 131, 151) is given in Table 5.

The χ_1/χ_2 rotamers extracted from the RE simulations are broadly consistent with the information from the X-ray crystal structures for seven positions: a *mm* rotamer is observed at 72, 82, 115, 119, 131, and 151, a *mt* rotamer is observed at 75, and a *tm* rotamer is observed at 76. Analysis of the RE simulations suggests that the free energy difference between *mm* and *mt* is small. In the case of position 131, analysis of multifrequency ESR spectra using MD simulations concluded that the rotamers *mm*, *mt*, *tp*, and *tt* were predominant.[21] All four rotamers are observed in the RE simulations, and the first one is observed in the X-ray structure. In the case of position 72, there is no X-ray structure, but there is indirect information from analysis of multifrequency ESR spectra using MD simulations indicating that the *mm* and *tp* rotamers are the most stable,[21] which is consistent with the results from RE simulations. The clearest disagreement concerns the spin-label at position 65, as the rotamer is *tp* in the X-ray structure, whereas it is *mm* in the RE simulations. However, this

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

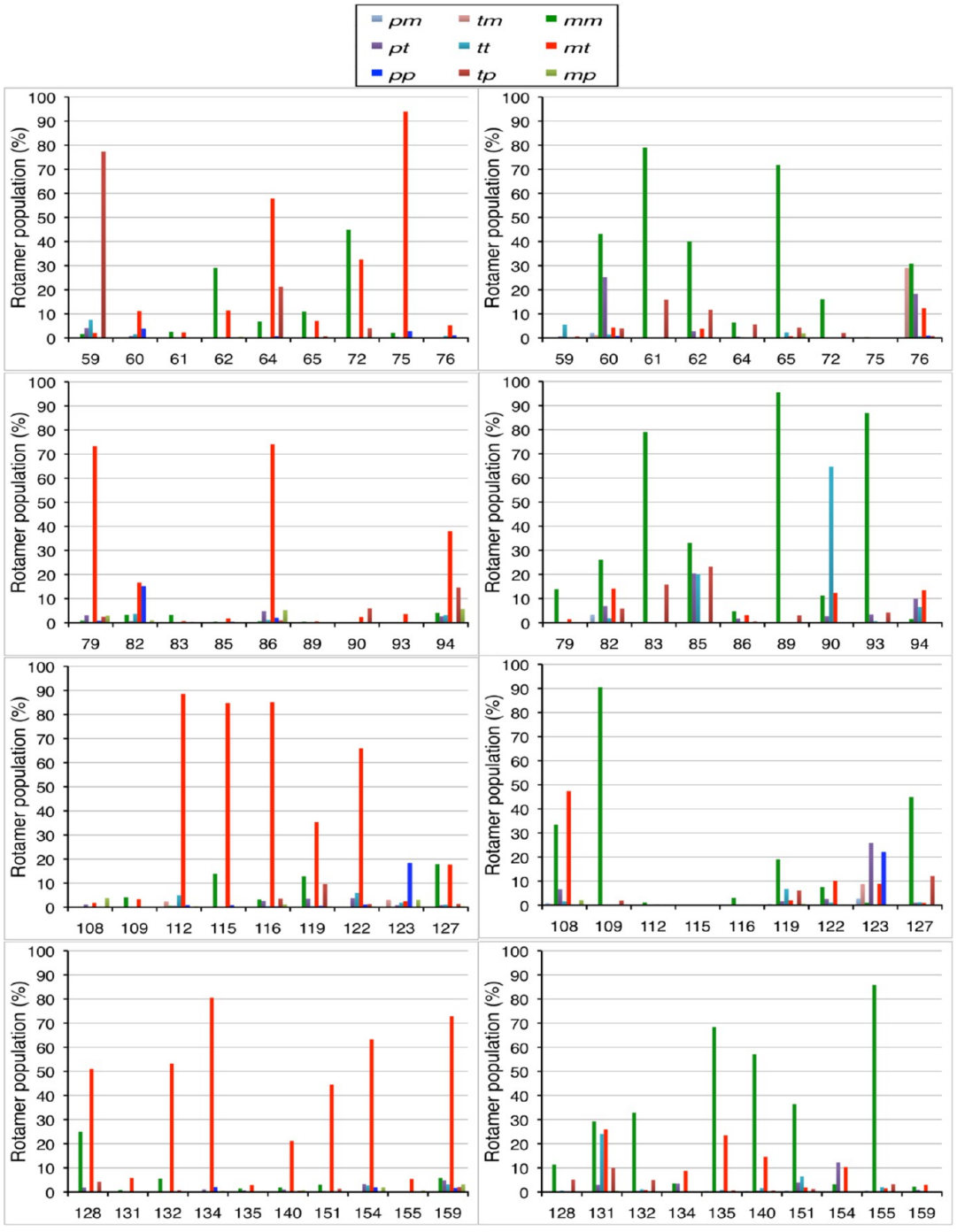


Figure 6. Population of χ_1/χ_2 rotamers when χ_3 is in the *p* state around +90° (left panel) or when χ_3 is in the *m* state around −90° (right panel) for the 37 individual spin-labels attached to T4L obtained from RE simulations.

position is problematic because the spin-label is involved in crystal contacts.[15]

For the dihedral angle χ_1, the most frequently observed rotamer in X-ray structures is *m*, although the *t* rotamer is also observed at a few positions. The *m* rotamer is also predicted to be the most stable energetically according to computational studies.[18,19] It has been argued that the *m* and *t* rotameric states of χ_1 might be stabilized by an interaction between the sulfur S_δ atom of the spin-label with the backbone nitrogen N and oxygen O atoms,[18] though further analysis indicates that such interactions are very weak.[19] For the dihedral angle χ_2, the predominant rotamer observed in X-ray structures is *m*, which is also frequently observed in the RE simulations but absent from the conventional MD or the LES simulations. Thus, there is closer agreement between the conformations from X-ray structures and those extracted from ESR/DEER experiments using RE simulations than with those obtained from unbiased MD or LES simulations. It is likely that these differences reflect the fact that the rate at which the experimental samples are frozen is slow on the molecular time scale, thereby leading to a partial annealing of some side chains' rotameric conformations to local minimum energy conformations. It is noteworthy that the nine sites (65, 72, 75, 76, 82, 115, 119, 131, 151), which are in the helix surface sites in T4L prefer either the *mm* or *mt* rotamer for χ_1/χ_2.

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

Table 5. Comparison of χ_1 and χ_2 Rotamers of the Spin-Labels Obtained from X-ray Structures and RE Simulations

residue	PDB ID	experimental rotamer	RE simulations
65	3K2R[a,b,c]	*tp*	*mm* (83%)
72	MD[d]	*mm*, *tp*	*mm* (64%), *mt* (32%)
75	Fleissner[a]	*mt*	*mt* (93%)
76	3K2R[c]	*tm*	*tm* (30%), *mm* (30%)
82	1ZYT[a,e]	*mm*	*mm* (30%), *mt* (30%)
115	2IGC, 2OU8[a,f]	*mm*, *tp*	*mt* (84%), *mm* (15%)
119	3L2X[f]	*mm*	*mm* (35%), *mt* (35%)
131	2CUU, 3G3V[e,g]	*mm*, *tp*	*mm* (30%), *mt* (30%), *tt* (25%), *tp* (10%)
	MD[d]	*mm*, *tp*, *mt*, *tt*	
151	3G3X[e]	*mm*	*mt* (45%), *mm* (40%)

[a]Fleissner et al.[13] [b]Langen et al.[15] [c]Toledo et al.[14] [d]Sezer et al.[21] [e]Fleissner et al.[11,12] [f]Guo et al.[10] [g]Guo et al.[9]

(c). Simplified Representation of the Spin-Label for Structural Refinement. An important outcome of the present analysis of RE simulations is a detailed characterization of the rotamer population for MTSSL spin-labels bound to T4L based on ESR/DEER data. This information is critically important for the interpretation of ESR experiments, particularly given the general paucity of data about the conformation of spin-label. Nevertheless, these results were obtained by applying an extensive computational framework with RE simulations that may seem somewhat cumbersome if one's sole purpose is to utilize ESR/DEER data for structural refinement. This framework is not without some drawbacks. For instance, because the replicas of the spin-labels are represented with all atomic details, fairly long RE simulations may be required to allow for the conformational transitions of the five dihedrals χ_1 to χ_5 in order to adequately sample all the accessible rotameric states. Furthermore, introducing multiple copies of the protein was necessary in the present application to avoid steric clashes between neighboring spin-labels.

Although all these technical difficulties can be surmounted, it is tempting to try to find ways to further simplify the computational framework without compromising accuracy. A key observation guiding our efforts to design a simplified method is that structural refinement based on ESR/DEER data with RE simulations ought to be concerned only with translating the information from histograms of interatomic distances between the multiple copies of the oxygen of the nitroxide spin-labels into backbone constraints. In the context of structural refinement, the detailed population of accessible rotameric states represents an excess of information that is used only indirectly. In fact, the detailed molecular representation of the spin-labels is needed because it serves to accurately position the oxygen atom relative to the protein backbone. Therefore, one potential route toward a simplified framework is to take a shortcut and discard the rest of the spin-label and retain only the nitroxide oxygen as a dummy ON particle linked to the backbone. Such a simplified representation avoids the burdensome task of accounting for a large ensemble of Boltzmann-weighted spin-label rotamers. Yet, it is perfectly adequate for the purpose of translating the information from ESR/DEER data into the protein structure—as long as one has the ability to determine the position of the ON atom accurately relative to the backbone *without* modeling the molecular structure of the spin-label itself. The position of this dummy ON particle relative to the N, C_α, and C_β of the backbone can be parametrized via a few simple energy terms (bond, angle dihedral) to best reproduce the statistical distribution of the spin-label side chains deduced from the all-atom RE simulations analysis (Figure 7A). As shown in Figure 7B, the configurations

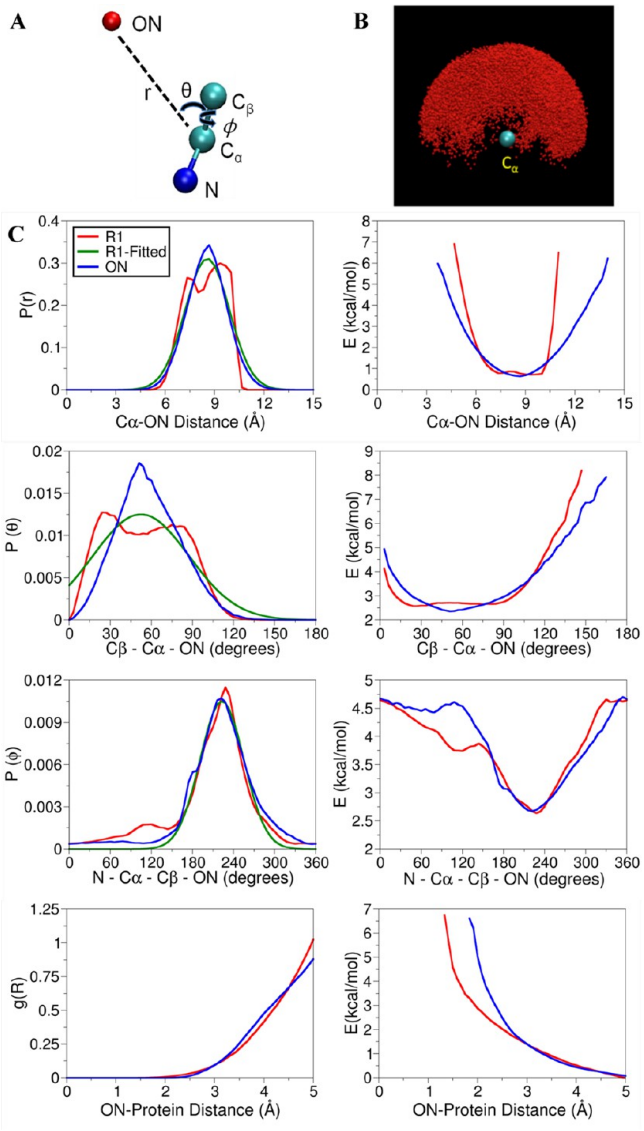


Figure 7. (A) Cartoon representation of the spin-label side chain, R1, replaced with a single dummy nitroxide atom, which is attached to the C_α atom of the protein backbone. Three valuables, C_α–ON distance, C_β–C_α–ON angle, and ON–C_β–C_α–N torsion are used to parametrize the dummy atom. (B) Dynamics of nitroxide oxygen of spin-labels around the C_α atom obtain from RE simulation. (C) Comparison of distribution and potential of mean force of the distance, angle, dihedral angle, and nonbond interactions obtained from MD simulations with R1 and dummy nitroxide atoms.

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740–4754

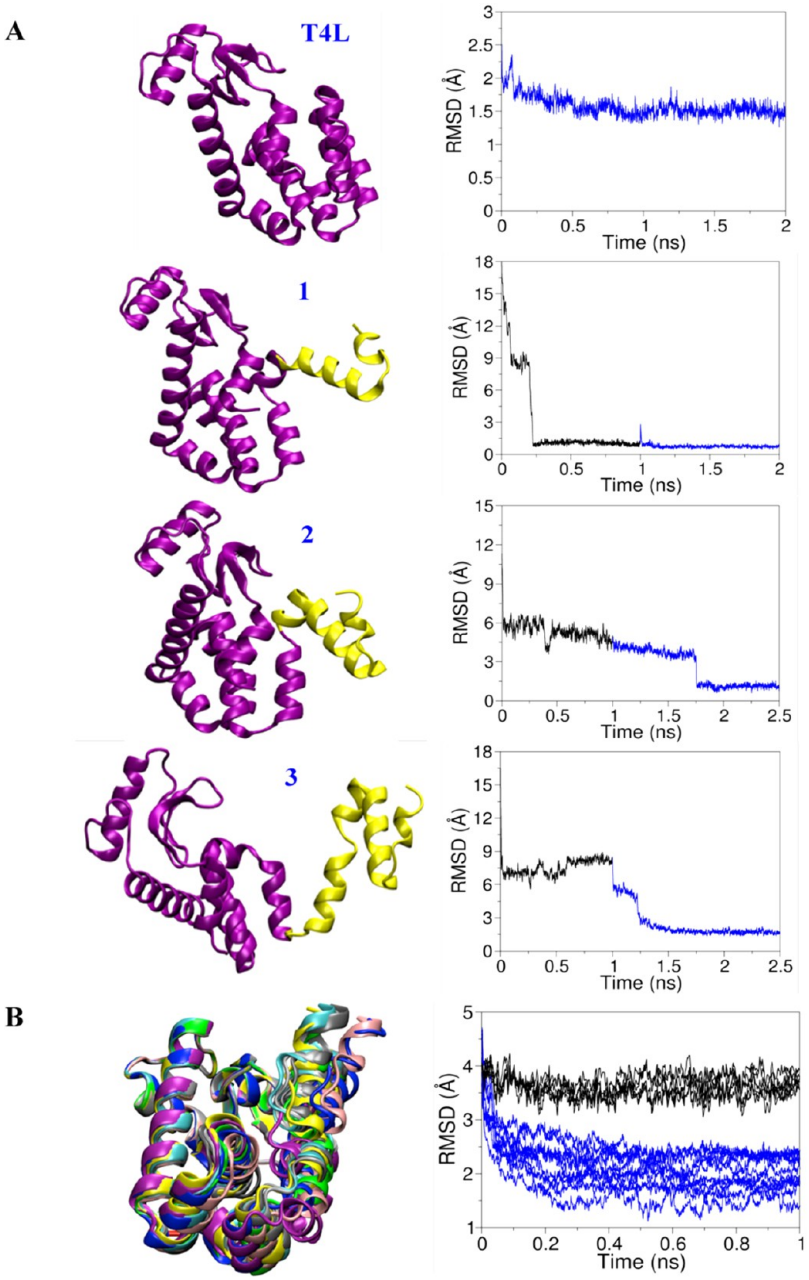


Figure 8. Cartoon representation of X-ray and distorted structures of T4L and the respective root-mean-square deviation of the distorted backbone atoms with respect to the X-ray structure obtained from NOE and RE simulations colored in black and blue, respectively. (A) The X-ray structure is denoted by T4L and the distorted structures are denoted by 1, 2, and 3 depending on the number of distorted helices. The distorted and undistorted residues of the T4L structures are colored in yellow and purple, respective. (B) Ten distorted T4L structures, with a backbone atom rmsd of about 4 Å from the X-ray crystal structure, are refined using the NOE and RE simulations colored in black and blue, respectively. Each distorted structure is represented with a particular color scheme. Three black traces from NOE (black) going down to rmsd of 3.1, 2.8, and 2.2 Å, respectively, are not shown for the sake of clarity.

extracted from the RE simulations reveal that the oxygen atom of the nitroxide spin-label is distributed within a half-sphere around the C_α atom with respect to the N, C_ω, and C_β atoms of the labeled residue. Interestingly, the distribution is almost uniform and does not seem to be strongly correlated with any given rotameric states.

The parametrization of the dummy nitroxide atom was carried out by determining the force constant for the C_α–ON bond (r), the C_β–C_α–ON angle (θ), and the N–C_α–C_β–ON dihedral angel (ϕ) (Figure 7A). The spatial distribution of the nitroxide oxygen atom is mapped in terms of the C_α–ON bond (r), the C_β–C_α–ON angle (θ), and the N–C_α–C_β–ON dihedral angle (ϕ), under the assumption that the values of three variables are uncorrelated (Figure S12 in the Supporting Information). The distributions over r, θ, and ϕ extracted from the RE simulation are shown in Figure 7C (red lines). The probability of finding the ON is maximum at a distance of about 8.0 Å from the C_α and the width of the distribution is about 5 Å. The distribution for θ is quite broad going from 0° to 110° and there appears to be two shallow humps at about 25° and 85°. There is one predominant rotameric state for the dihedral angle ϕ positioned around 180–280°, peaking around 240°. The probability distribution

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740–4754

functions are then modeled on the basis of three simple energy terms, $V_r = k_r(r - r_0)^2$, $V_\theta = k_\theta(\theta - \theta_0)^2$, and $V_\phi = k_\phi(1 + \cos(n\phi - \phi_0))$, where k_r, k_θ, and k_ϕ are the force constants for the distance, the angle, and dihedral potentials and n represents the dihedral multiplicity. In addition, a Lennard-Jones 6-12 potential was used to account for the excluded-volume interactions between the ON particle and the rest of the protein. The radial distribution (R) of all nonbond interactions involving ON atom of the spin-label was calculated in a shell of 5 Å radius and the distribution was converted to the corresponding Boltzmann energy. The result, shown in Figure 7C, indicates that the probability of finding an ON atom is almost zero between 0 and 2.5 Å from any protein atom, which is also evident from the large energy at this region, while the change in energy between 4 and 5 Å is very small (<0.5 kcal/mol). Optimized Lennard-Jones parameters are R_{min} = 4 Å and E_{min} = −0.05 kcal/mol. Nonbonded interactions between the ON dummy atoms and the water molecules are switched off by using the keyword NBFIX in the parameter file, which can be read by both the NAMD[48] and CHARMM[47] program packages. The optimal force constants for k_r, k_θ, and k_ϕ are 0.5 (kcal/mol)/Å^2, 1.0 (kcal/mol)/rad^2, and 1.9 (kcal/mol)/rad^2, respectively. A multiplicity of n = 1 is used for the N−C_α−C_β−ON dihedral angle.

Having parametrized the ON dummy atom representation, we performed 5 ns MD simulation of T4L labeled with ON dummy atom at 37 sites. The calculated ON-ON distance distributions are slightly broader than the experimental ESR/DEER histograms, although the results are somewhat improved compared with the conventional MD simulations carried out with the detailed spin-label models with explicit solvent (Figures S6 and S13 in the Supporting Information). The average distances obtained from MD simulation and experimental spin−spin distance distributions were compared using a linear regression analysis. The correlation between averaged measured and simulated distances is 0.91. These results were compared with those obtained from the Multiscale Modeling of Macromolecular systems (MMM) software package developed by Yevhen Polyhach and Gunnar Jeschke to analyze ESR/DEER data.[28] With MMM, the modeling is carried out on the basis of a precalculated library of ∼200 rotamers which are then inserted at individual sites on a fixed T4L protein backbone. Configurations with clashes are discarded and a Boltzmann-weighted probability distribution is produced for each inserted spin-label, and finally an average pair distance is calculated for each pair of spin-labels. The correlation coefficient of the average distances obtained from MMM with ESR/DEER data is 0.85, which is slightly smaller than the correlation coefficient of 0.91 obtained from MD simulations with the simplified dummy ON spin-label described above.

(d). Illustrative Structural Refinement with Artificially Distorted Conformations of T4L. The simplified dummy ON spin-label model, together with the RE simulation method, can be used to refine protein structures. As a first illustrative test, the T4L structure was refined with the RE simulation from its X-ray structure. Dummy ON atoms were attached to the 37 sites of the spin-labels in T4L and were replicated 25 times for the RE simulations. A 2 ns RE simulation was performed starting from the X-ray crystal structure equilibrated from MD (in vacuum and with explicit solvent). The root-mean-square deviation (rmsd) over the trajectory calculated with respect to the original X-ray structure shows that the rmsd flattens out after 400 ps and remains almost constant to a value of about 1.4 Å (see Figure 8A). The interspin distance distributions are very similar to those obtained from experiment and with those obtained from RE simulation performed with the whole spin-labels (Figure S18 in the Supporting Information). As a second test, the ability of the method to correct artificially distorted structures of T4L was examined. Figure 8 shows some of the distorted structures that were refined with the RE simulations. A total of eight distorted structures were created by altering the dihedral angles in the turns or loops that are connecting pairs of helices, starting from the C-terminal helix of T4L (Figure 8). Some of the distorted models have a backbone atom rmsd of more than 15 Å relative to the X-ray structure. In the refinement, an initial stage of MD simulation was first carried out using a flat-bottom harmonic distance restraints as commonly used in NMR refinement based on NOE relaxation data. The 51 average distances obtained from ESR/DEER data were used. After several nanoseconds of simulation, the rmsd of the backbone atoms of all distorted T4L with respect to the X-ray structure flattens out. The distorted structures of T4L improved for all the structures. However, the rmsd for seven distorted structures remained large, in the range of 5−17 Å, from the X-ray structure. Only one distorted model reached a value of 1.7 Å rmsd (Figure 8). To see if the models could be further improved, RE simulations were then performed starting from the final configuration obtained from the previous simulations. The rmsd of the backbone atoms of most of the models improved considerably, demonstrating that the RE method with the simplified dummy ON spin-label is able to further refine these structures (Figure 8). However, the backbone rmsd for some distorted models remained larger than 6 Å (Figures S19 and S20 in the Supporting Information). This mainly shows the limitation of a sampling algorithm that is based on straight MD simulations, even in the presence of the RE biasing. There are, however, alternative and more powerful computational strategies that have the ability to refine such highly distorted protein structures, and improve the overall rmsd down to ∼4 Å.[27,62] Therefore, as a final illustrative test, we sought to examine the ability of the RE method to refine 10 structures that are only moderately distorted to ∼4 Å rmsd (Figure 8B). In comparison, MD simulations were also performed with a NOE-like flat-bottom harmonic distance restraints. The average backbone atom rmsd, obtained from the last frame of the RE simulations, were found to be about 1.8 Å with respect to the X-ray structure. In contrast, many of the models refined with the NOE-like distance restraint remained at 3.5 Å rmsd (three models refined to 2.0−3.5 Å rmsd). These results demonstrate the usefulness of RE simulation method with the simplified representation of the spin-label as a dummy ON particle for structural refinement.

CONCLUSION

An extensive characterization of the rotamer population of the spin-labels attached to T4 lysozyme (T4L) has been carried out using a novel multiple-copy restrained-ensemble (RE) simulation method to incorporate ESR/DEER distance histogram data. In RE simulations, a global energy restraint forces the simulation spin pair distances collected from multiple copies of the system to match with the experimental distance histograms obtained from ESR/DEER. Here, 51 distance distributions for 37 nitroxide spin-labeled T4L were constrained to the experimental distributions. In contrast, results obtained from single-copy MD and unrestrained LES simulations deviate significantly from the experiments and therefore the conformational preferences of the spin-labels could be very different with these simulations when compared to experiment. In most cases, the rotamers along χ_1

dx.doi.org/10.1021/jp311723a | J. Phys. Chem. B 2013, 117, 4740−4754

and χ_2 at $\chi_3 \approx 90°$ and 270° are predominantly *mt* and *mm*, respectively. On the other hand, multiple rotameric states are found for χ_4 and χ_5 at $\chi_3 \approx 90°$ and 270°. Overall, these findings are in good agreement with the available information from X-ray crystallographic structures,[9−15,21] ESR line-shape analysis[6,21,63] and ab initio studies.[16,18] A few spin-label side chains prefer other rotameric states or multiple rotamers depending on their positions in T4L. On the basis of the rotameric preferences, we classified the spin-labels broadly into two groups: (1) sites in the helices will most likely prefer either *mm* or *mt* or both *mm* and *mt* rotamers and (2) sites on the solvent-exposed loop structures, or sites completely buried in the solvent inaccessible hydrophobic cores, will prefer variable rotameric states.

To develop an efficient strategy aimed at protein structural refinement based on ESR/DEER distance histograms, we designed and parametrized a simplified dummy nitroxide atom "ON" reproducing the statistical distribution of the spin-label side chains deduced from the all-atom RE simulations analysis. The 51 average spin pair distances calculated from these dummy ON spin-label are found in excellent agreement with those obtained from experiment, with a correlation coefficient of 0.91. This result demonstrates that the simplified ON dummy representation is an effective approach to accurately calculate distances between spin-labels at two sites on a protein. Artificially distorted structures of T4L were successfully refined with the final structures deviating from the X-ray structures by an rms deviation of only a few angstroms. The extent of distortion of the initial structure will dictate the refinement process. Therefore, it is important to obtain an X-ray or NMR derived initial structure that is close to the target structure. In most cases, short several nanosecond RE simulations were sufficient to refine the distorted T4L structures. The setup with the dummy spin-label is available in the PDB manipulation step in CHARMM-GUI (www.charmm-gui.org).[64] Current efforts with the RE simulation method are aimed at determining the various conformational states of ion channels and membrane transporters using ESR/DEER data.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

T4L double mutants, simulation setup, distance distributions, time dependence for χ_3 dihedral angle, rotamer population distribution along χ_4 and χ_5, 2D-PMFs along $\chi_1\chi_2$ and $\chi_4\chi_5$, correlation of the $C_\beta-C_\alpha-ON$ angle and the $N-C_\alpha-C_\beta-ON$ dihedral angle, force field parameter and topology for dummy nitroxide atom, correlation of inter label experimental and MD average distances, correlation of inter label experimental and MMM average distances, structures used in refinement. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: roux@uchicago.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was carried out in the context of the Membrane Protein Structural Dynamics Consortium funded by grant U54-GM087519 from the National Institute of Health (NIH). We are grateful to XSEDE for computer time. We also thank Dr. Hanane Koteiche for assistance with generation of the double mutants of T4L.

■ REFERENCES

(1) Borbat, P. P.; Mchaourab, H. S.; Freed, J. H. Electron Spin Resonance in Studies of Membranes and Proteins. *J. Am. Chem. Soc.* **2002**, *124*, 5304−5314.

(2) Bhatnagar, J.; Freed, J. H.; Crane, B. R. Rigid Body Refinement of Protein Complexes with Long-Range Distance Restraints from Pulsed Dipolar ESR. *Methods Enzymol.* **2007**, *423*, 117−133.

(3) Kazmier, K.; Alexander, N. S.; Meiler, J.; Mchaourab, H. S. Algorithm for Selection of Optimized EPR Distance Restraints for De Novo Protein Structure Determination. *J. Struct. Biol.* **2011**, *173*, 549−557.

(4) Columbus, L.; Kalai, T.; Jeko, J.; Hideg, K.; Hubbell, W. L. Molecular Motion of Spin Labeled Side Chains in Alpha-Helices: Analysis by Variation of Side Chain Structure. *Biochemistry* **2001**, *40*, 3828−3846.

(5) Jiao, D.; Barfield, M.; Combariza, J. E.; Hruby, V. J. Ab Initio Molecular Orbital Studies of the Rotational Barriers and the S-33 and C-13 Chemical Shieldings for Dimethyl Disulfide. *J. Am. Chem. Soc.* **1992**, *114*, 3639−3643.

(6) Mchaourab, H. S.; Lietzow, M. A.; Hideg, K.; Hubbell, W. L. Motion of Spin-Labeled Side Chains in T4 Lysozyme, Correlation with Protein Structure and Dynamics. *Biochemistry* **1996**, *35*, 7692−7704.

(7) Columbus, L.; Hubbell, W. L. A New Spin on Protein Dynamics. *Trends. Biochem. Sci.* **2002**, *27*, 288−295.

(8) Barnes, J. P.; Liang, Z. C.; Mchaourab, H. S.; Freed, J. H.; Hubbell, W. L. A Multifrequency Electron Spin Resonance Study of T4 Lysozyme Dynamics. *Biophys. J.* **1999**, *76*, 3298−3306.

(9) Guo, Z. F.; Cascio, D.; Hideg, K.; Hubbell, W. L. Structural Determinants of Nitroxide Motion in Spin-Labeled Proteins: Solvent-Exposed Sites in Helix B of T4 Lysozyme. *Protein Sci.* **2008**, *17*, 228−239.

(10) Guo, Z. F.; Cascio, D.; Hideg, K.; Kalai, T.; Hubbell, W. L. Structural Determinants of Nitroxide Motion in Spin-Labeled Proteins: Tertiary Contact and Solvent-Inaccessible Sites in Helix G of T4 Lysozyme. *Protein Sci.* **2007**, *16*, 1069−1086.

(11) Fleissner, M. R.; Bridges, M. D.; Brooks, E. K.; Cascio, D.; Kalai, T.; Hideg, K.; Hubbell, W. L. Structure and Dynamics of a Conformationally Constrained Nitroxide Side Chain and Applications in EPR Apectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16241−16246.

(12) Fleissner, M. R.; Cascio, D.; Hubbell, W. L. Structural Origin of Weakly Ordered Nitroxide Motion in Spin-Labeled Proteins. *Protein Sci.* **2009**, *18*, 893−908.

(13) Fleissner, M. R. Ph.D. Thesis, University of California, Los Angeles, 2007.

(14) Toledo Warshaviak, D., Cascio, D., Khramtsov, V. V., Hubbell, W. L. Crystal Structure of Spin Labeled T4 Lysozyme Mutant K65V1/R76V1 In *PDB Data Bank*.

(15) Langen, R.; Oh, K. J.; Cascio, D.; Hubbell, W. L. Crystal Structures of Spin Labeled T4 Lysozyme Mutants: Implications for the Interpretation of EPR Spectra in Terms of Structure. *Biochemistry* **2000**, *39*, 8396−8406.

(16) Tombolato, F.; Ferrarini, A.; Freed, J. H. Dynamics of the Nitroxide Side Chain in Spin-Labeled Proteins. *J. Phys. Chem. B* **2006**, *110*, 26248−26259.

(17) Tombolato, F.; Ferrarini, A.; Freed, J. H. Modeling the Effects of Structure and Dynamics of the Nitroxide Side Chain on the ESR Spectra of Spin-Labeled Proteins. *J. Phys. Chem. B* **2006**, *110*, 26260−26271.

(18) Warshaviak, D. T.; Serbulea, L.; Houk, K. N.; Hubbell, W. L. Conformational Analysis of a Nitroxide Side Chain in an Alpha-Helix with Density Functional Theory. *J. Phys. Chem. B* **2011**, *115*, 397−405.

(19) Sezer, D.; Freed, J. H.; Roux, B. Parametrization, Molecular Dynamics Simulation, and Calculation of Electron Spin Resonance Spectra of a Nitroxide Spin Label on a Polyalanine Alpha-Helix. *J. Phys. Chem. B* **2008**, *112*, 5755−5767.

(20) Sezer, D.; Freed, J. H.; Roux, B. Using Markov Models to Simulate Electron Spin Resonance Spectra from Molecular Dynamics Trajectories. *J. Phys. Chem. B* **2008**, *112*, 11014−11027.
(21) Sezer, D.; Freed, J. H.; Roux, B. Multifrequency Electron Spin Resonance Spectra of a Spin-Labeled Protein Calculated from Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2009**, *131*, 2597−2605.
(22) Polyhach, Y.; Godt, A.; Bauer, C.; Jeschke, G. Spin Pair Geometry Revealed by High-Field DEER in the Presence of Conformational Distributions. *J. Magn. Reson.* **2007**, *185*, 118−129.
(23) Tikhonova, I. G.; Best, R. B.; Engel, S.; Gershengorn, M. C.; Hummer, G.; Costanzi, S. Atomistic Insights into Rhodopsin Activation from a Dynamic Model. *J. Am. Chem. Soc.* **2008**, *130*, 10141−10149.
(24) Ding, F.; Layten, M.; Simmerling, C. Solution Structure of HIV-1 Protease Flaps Probed by Comparison of Molecular Dynamics Simulation Ensembles and EPR Experiments. *J. Am. Chem. Soc.* **2008**, *130*, 7184−7185.
(25) Sale, K.; Song, L. K.; Liu, Y. S.; Perozo, E.; Fajer, P. Explicit Treatment of Spin Labels in Modeling of Distance Constraints from Dipolar EPR and DEER. *J. Am. Chem. Soc.* **2005**, *127*, 9334−9335.
(26) Boura, E.; Rozycki, B.; Herrick, D. Z.; Chung, H. S.; Vecer, J.; Eaton, W. A.; Cafiso, D. S.; Hummer, G.; Hurley, J. H. Solution Structure of the ESCRT-I Complex by Small-Angle X-ray Scattering, EPR, and FRET Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 9437−9442.
(27) Hirst, S. J.; Alexander, N.; Mchaourab, H. S.; Meiler, J. RosettaEPR: An Integrated Tool for Protein Structure Determination from Sparse EPR Data. *J. Struct. Biol.* **2011**, *173*, 506−514.
(28) Polyhach, Y.; Bordignon, E.; Jeschke, G. Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2356−2366.
(29) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419−446.
(30) Hatmal, M. M.; Li, Y.; Hegde, B. G.; Hegde, P. B.; Jao, C. C.; Langen, R.; Haworth, I. S. Computer Modeling of Nitroxide Spin Labels on Proteins. *Biopolymers* **2012**, *97*, 35−44.
(31) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905−921.
(32) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. Mapping Long-Range Interactions in Alpha-Synuclein Using Spin-Label NMR and Ensemble Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2005**, *127*, 476−477.
(33) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous Determination of Protein Structure and Dynamics. *Nature* **2005**, *433*, 128−132.
(34) Lindorff-Larsen, K.; Best, R. B.; Vendruscolo, M. Interpreting Dynamically-Averaged Scalar Couplings in Proteins. *J. Biomol. NMR* **2005**, *32*, 273−280.
(35) Best, R. B.; Lindorff-Larsen, K.; DePristo, M. A.; Vendruscolo, M. Relation Between Native Ensembles and Experimental Structures of Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10901−10906.
(36) Lindorff-Larsen, K.; Rogen, P.; Paci, E.; Vendruscolo, M.; Dobson, C. M. Protein Folding and the Organization of the Protein Topology Universe. *Trends. Biochem. Sci.* **2005**, *30*, 13−19.
(37) Lee, J.; Chen, J. H.; Brooks, C. L.; Im, W. P. Application of Solid-State NMR Restraint Potentials in Membrane Protein Modeling. *J. Magn. Reson.* **2008**, *193*, 68−76.
(38) Jo, S.; Im, W. Transmembrane Helix Orientation and Dynamics: Insights from Ensemble Dynamics with Solid-State NMR Observables. *Biophys. J.* **2011**, *100*, 2913−2921.
(39) Kim, T.; Jo, S.; Im, W. Solid-State NMR Ensemble Dynamics as a Mediator Between Experiment and Simulation. *Biophys. J.* **2011**, *100*, 2922−2928.
(40) Im, W.; Jo, S.; Kim, T. An Ensemble Dynamics Approach to Decipher Solid-State NMR Observables of Membrane Proteins. *Biochim. Biophys. Acta* **2012**, *1818*, 252−262.
(41) Jaynes, E. T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620−630.
(42) Roux, B.; Weare, J. On the Statistical Equivalence of Restrained-Ensemble Simulations With the Maximum Entropy Method. *J. Chem. Phys.* **2013**, *138*, 084107.
(43) Roux, B.; Islam, S. M. Restrained-Ensemble Molecular Dynamics Simulations Based on Distance Histograms from Double Electron-Electron Resonance Spectroscopy. *J. Phys. Chem. B* **2013**, DOI: 10.1021/jp3110369.
(44) Jeschke, G. Distance Measurements in the Nanometer Range by Pulse EPR. *Chemphyschem* **2002**, *3*, 927−932.
(45) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. DeerAnalysis2006 - A Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006**, *30*, 473−499.
(46) Chiang, Y. W.; Borbat, P. P.; Freed, J. H. The Determination of Pair Distance Distributions by Pulsed ESR using Tikhonov Regularization. *J. Magn. Reson.* **2005**, *172*, 279−295.
(47) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(48) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(49) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−1616.
(50) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(51) Weaver, L. H.; Matthews, B. W. Structure of Bacteriophage-T4 Lysozyme Refined at 1.7 a Resolution. *J. Mol. Biol.* **1987**, *193*, 189−199.
(52) Roitberg, A.; Elber, R. Modeling Side-Chains in Peptides and Proteins - Application of the Locally Enhanced Sampling and the Simulated Annealing Methods to Find Minimum Energy Conformations. *J. Chem. Phys.* **1991**, *95*, 9277−9287.
(53) Simmerling, C.; Lee, M. R.; Ortiz, A. R.; Kolinski, A.; Skolnick, J.; Kollman, P. A. Combining MONSSTER and LES/PME to Predict Protein Structure from Amino Acid Sequence: Application to the Small Protein CMTI-1. *J. Am. Chem. Soc.* **2000**, *122*, 8392−8402.
(54) Simmerling, C.; Fox, T.; Kollman, P. A. Use of Locally Enhanced Sampling in Free Energy Calculations: Testing and Application to the Alpha → Beta Anomerization of Glucose. *J. Am. Chem. Soc.* **1998**, *120*, 5771−5782.
(55) Adelman, S. A.; Doll, J. D. Generalized Langevin Equation Approach for Atom-Solid-Surface Scattering - General Formulation for Classical Scattering Off Harmonic Solids. *J. Chem. Phys.* **1976**, *64*, 2375−2388.
(56) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(57) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - An N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(58) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(59) Elber, R.; Karplus, M. Enhanced Sampling in Molecular-Dynamics - Use of the Time-Dependent Hartree Approximation for a Simulation of Carbon-Monoxide Diffusion through Myoglobin. *J. Am. Chem. Soc.* **1990**, *112*, 9161−9175.
(60) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.
(61) Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H. Effect of Freezing Conditions on Distances and Their Distributions Derived from Double Electron Electron Resonance (DEER): A Study of Doubly-Spin-Labeled T4 Lysozyme. *J. Magn. Reson.* **2012**, *216*, 69−77.
(62) Brunger, A. T. Version 1.2 of the Crystallography and NMR System. *Nat. Protoc.* **2007**, *2*, 2728−2733.
(63) Mchaourab, H. S.; Kalai, T.; Hideg, K.; Hubbell, W. L. Motion of Spin-Labeled Side Chains in T4 Lysozyme: Effect of Side Chain Structure. *Biochemistry* **1999**, *38*, 2947−2955.
(64) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.