Reconstruction of Protein Side-Chain Conformational Free Energy Surfaces From NMR-Derived Methyl Axis Order Parameters

Marimuthu Krishnan*,[†] and Jeremy C. Smith[‡,§]

[†]Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500 032, India

[‡]University of Tennessee/Oak Ridge National Laboratory Center for Molecular Biophysics, Oak Ridge, Tennessee 37831, United States

[§]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, United States

S *Supporting Information*

ABSTRACT: An analytical approach is developed for reconstructing site-specific methyl-bearing protein side-chain conformational energy surfaces from NMR methyl axis order parameters (O_{axis}^2). Application of an enhanced sampling algorithm (adaptive biasing force) to molecular dynamics simulation of a protein, calcium-bound calmodulin, reveals a nonlinear correlation between O_{axis}^2 and the populations of rotamer states of protein side-chains, permitting the rotamer populations to be extracted directly from O_{axis}^2. The analytical approach yields side-chain conformational distributions that are in excellent agreement with those obtained from the enhanced-sampling MD results.
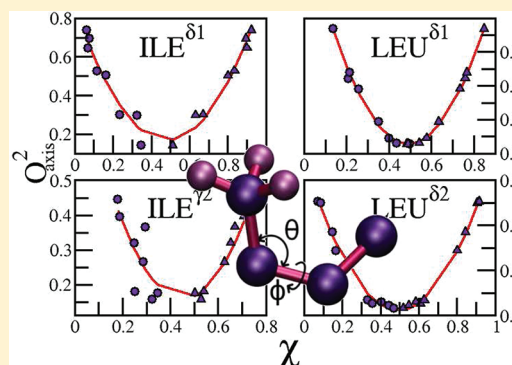
■ INTRODUCTION

Understanding the interplay between the structure, dynamics, and thermodynamics of proteins is essential for comprehending their folding and function.[1−6] Among the experimental approaches to this problem, site-specific nuclear magnetic resonance (NMR) has emerged as a means of determining the contribution of side-chain dynamics to conformational entropy, and, this entropy being an essential component of protein thermodynamics, the approach permits a molecular-level connection between the atomistic dynamics and thermodynamics of proteins.[2−5,7−16] ^{13}C, ^{1}H, and ^{2}H NMR relaxation data of methyl-bearing residues provide information on side-chain motions, while ^{15}N relaxation is used to examine amide backbone fluctuations.[4,6−15,17−19]

The Lipari−Szabo "model-free" formalism is often used to interpret NMR spectral data in terms of fluctuations of bond vectors (amide N−H vectors for backbone dynamics, and C−H vectors of methyl groups for side-chain dynamics), yielding the square of an order parameter, O^2, quantifying the amplitude of motion, and a corresponding relaxation time, τ. The side-chain O^2 and τ quantify the degree of spatial restriction and relaxation time of a given methyl-bearing side-chain.[20,21] The reorientational dynamics of the symmetry axis and the rotation of C−H vectors about the symmetry axis are the two major dynamical contributions that determine the methyl relaxations. Assuming that these two degrees of freedom are uncoupled, O^2 can be written as

$$O^2 = O_{axis}^2 O_{rot}^2 \qquad (1)$$

where O_{rot}^2 is the order parameter for rotation about the methyl symmetry axis, and O_{axis}^2 is that for motion of the symmetry axis.[4,7−15,17,18,22] $O_{rot}^2 = 1/9$ for a methyl group that rotates completely and possesses ideal tetrahedral geometry, and in the analysis of NMR experimental data, this value of O_{rot}^2 is commonly assumed.[2]

A key quantity of interest for connecting NMR-derived order parameters (O_{axis}^2) to thermodynamics is the angular/orientational probability distribution function, $P(\theta,\phi)$, of a given methyl symmetry axis, whose orientation is determined by the polar angles θ and ϕ.[2,16,23−26] A schematic illustration of the polar angles θ and ϕ is shown in Figure 1. The methyl-axis order parameter (O_{axis}^2) associated with a given side-chain is related to $P(\theta,\phi)$ as follows:

$$O_{axis}^2 = \frac{3}{2}\left[\left\langle x^2 \right\rangle^2 + \left\langle y^2 \right\rangle^2 + \left\langle z^2 \right\rangle^2 + 2\left\langle xy \right\rangle^2 + 2\left\langle xz \right\rangle^2 \right.$$
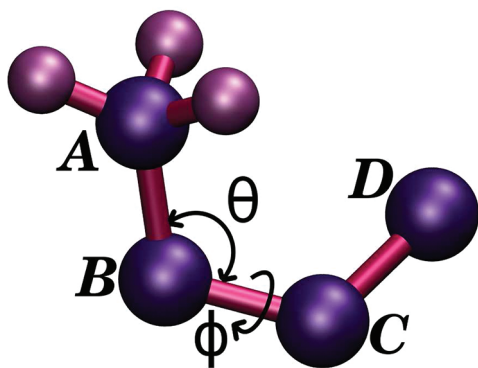$$\left. + 2\left\langle yz \right\rangle^2\right] - \frac{1}{2} \qquad (2)$$

Figure 1. Schematic representation of a methyl group. The atom label A denotes the methyl carbon, while labels B, C, and D represent heavy atoms. The atoms A, B, C, and D are sequentially connected by these three repeating bonds. The dihedral angle, ϕ, is the angle between the planes $A-B-C$ and $B-C-D$. θ is the angle between $B-A$ and $B-C$ bond vectors. The bond vector $B-A$ denotes the symmetry axis of the methyl group. $\pi - \theta$ defines the azimuth angle of the symmetry axis of the methyl group.

where $x = \sin(\pi - \theta)\cos(\phi)$, $y = \sin(\pi - \theta)\sin(\phi)$, and $z = \cos(\pi - \theta)$, and the angular brackets denote averages are computed using $P(\theta,\phi)$.[23,26] In what follows, θ in $P(\theta,\phi)$ is actually $\pi - \theta$.

In many approaches for characterizing the dynamics of protein side-chains, $P(\theta,\phi)$ is assumed to have simple functional forms in determining O_{axis}^2.[10,25,27−30] In another approach, a 3-fold symmetric dihedral potential energy surface (with a barrier height between any two adjacent rotamers of 3 kcal/mol) was used to describe the dihedral fluctuations of the side-chains.[26] Although the rotamer states were assumed to be iso-energetic, they were assumed to be unequally populated, resulting in unrealistic scaling (by the relative population of each rotamer state) of $P(\theta,\phi)$.[26] In a Monte Carlo-based approach, dihedral angles of rotatable side-chain bonds were assumed to be the sole degrees of freedom in proteins, and their fluctuations within rotamer wells were described by a quadratic function of ϕ.[31]

Molecular dynamics (MD) simulation, which employs force-field-based description of interactions between protein atoms, is a useful tool for characterizing fast motions of protein side-chains and thus can be used to provide a microscopic dynamical interpretation of NMR spin-relaxation results.[30,32−37] In an MD-based investigation of NMR relaxation parameters, the time evolution of θ and ϕ of a given methyl-bearing side-chain is monitored to determine the angular distribution function, $P(\theta,\phi)$, directly from the MD trajectories. $P(\theta,\phi)$ thus computed can be used (using eq 2) to determine O_{axis}^2. However, a major bottleneck in the above MD-based approach is an incomplete sampling of the conformational space of the side-chains, preventing the determination of the exact O_{axis}^2. This sampling problem may result in a poor estimation of $P(\theta,\phi)$, which can deviate from the equilibrium distribution, thus leading to the overestimation of O_{axis}^2.[33] Although a long, well-sampled MD trajectory should, in principle, be able to provide accurate estimation of O_{axis}^2, this is often a computationally demanding task.

Here, we examine the possibility of deriving O_{axis}^2 directly from side-chain conformational free energy profiles computed using enhanced sampling algorithms with detailed atomistic interactions described by a molecular mechanics (the

CHARMM) force-field. Since these enhanced-sampling free energy methods are designed to circumvent the sampling problem, the accurate estimation of conformational free energy profiles, $F(\theta,\phi)$, of protein side-chains is possible with a reasonable computing time. Our strategy is to determine $P(\theta,\phi)$ from $F(\theta,\phi)$ using the Boltzmann inversion formula, which is

$$P(\theta, \phi) = \frac{e^{-\beta F(\theta,\phi)}}{\int_0^{2\pi}\int_0^{\pi} e^{-\beta F(\theta,\phi)}\, d\phi\, d\theta} \tag{3}$$

$P(\theta,\phi)$ thus determined can be inserted into eq 2 to compute O_{axis}^2. This enhanced-sampling-based approach not only alleviates the sampling problem but also lead to the establishment of a nonlinear correlation between O_{axis}^2 and the populations of side-chain rotamer states. The nonlinear correlation thus obtained allows the determination of equilibrium conformational distributions of methyl-bearing side-chains of proteins directly from NMR methyl-axis motional parameters.

We apply the above approach to rapidly reconstruct the equilibrium conformational distributions of side-chains of a test protein, calmodulin, from O_{axis}^2. The adaptive biasing force (ABF) enhanced sampling method is used to accurately determine the conformational free energy profiles $F(\theta,\phi)$ of side-chains of CaM. Six independent 10-ns all-atom MD trajectories are also generated to assess the role of sampling in the estimation of O_{axis}^2. It is found that, whereas the individual MD trajectories do not reproduce the equilibrium $P(\theta,\phi)$ distribution, $P(\theta,\phi)$ obtained from an ensemble of multiple MD trajectories does resemble the converged equilibrium distribution, thus indicating that multiple short MD trajectories can be effectively used as an alternative to the enhanced sampling technique to determine the equilibrium $P(\theta,\phi)$ and thereby O_{axis}^2. The side-chain conformational distributions obtained from O_{axis}^2 are in excellent agreement with those obtained from the ABF method.

■ SIMULATION DETAILS

MD simulations of Ca^{2+}-saturated CaM (PDB id: 1CLL) were carried out using NAMD[38] with the CHARMM27 all-atom[39] and TIP3P water[40] force fields. The initial structure of Ca^{2+}-saturated CaM was taken from the crystal structure refined at 1.7 Å using X-ray crystallography (PDB ID: 1CLL).[41]

The model systems were constructed by placing the Ca^{2+}-saturated CaM in a TIP3P water box of dimensions $(59 \times 76 \times 56)$ Å^3. The resulting water molecules that were in hard contact with the protein (protein−water distance < 2.5 Å) were removed. The system was energy minimized to a root-mean-square (rms) energy gradient of 10^{-3} kcal/mol/Å.

The MD simulations were performed in the NPT ensemble at 300 K and 1 atm pressure using a Langevin thermostat and barostat with a damping coefficient of 5 ps^{-1}. The nonbonded pair-interaction potential was truncated at 12 Å and smoothed between 10 Å and 12 Å using a cubic switching function. Periodic boundary conditions were applied. Electrostatic interactions were computed using the particle mesh Ewald (PME) method[42] with a real space cutoff of 13 Å, and the reciprocal space interactions were computed on $60 \times 81 \times 60$ grids using sixth-degree B-splines. The equations of motion were integrated with a time step of 1 fs. The systems were energy minimized and then heated to 300 K followed by 2 ns
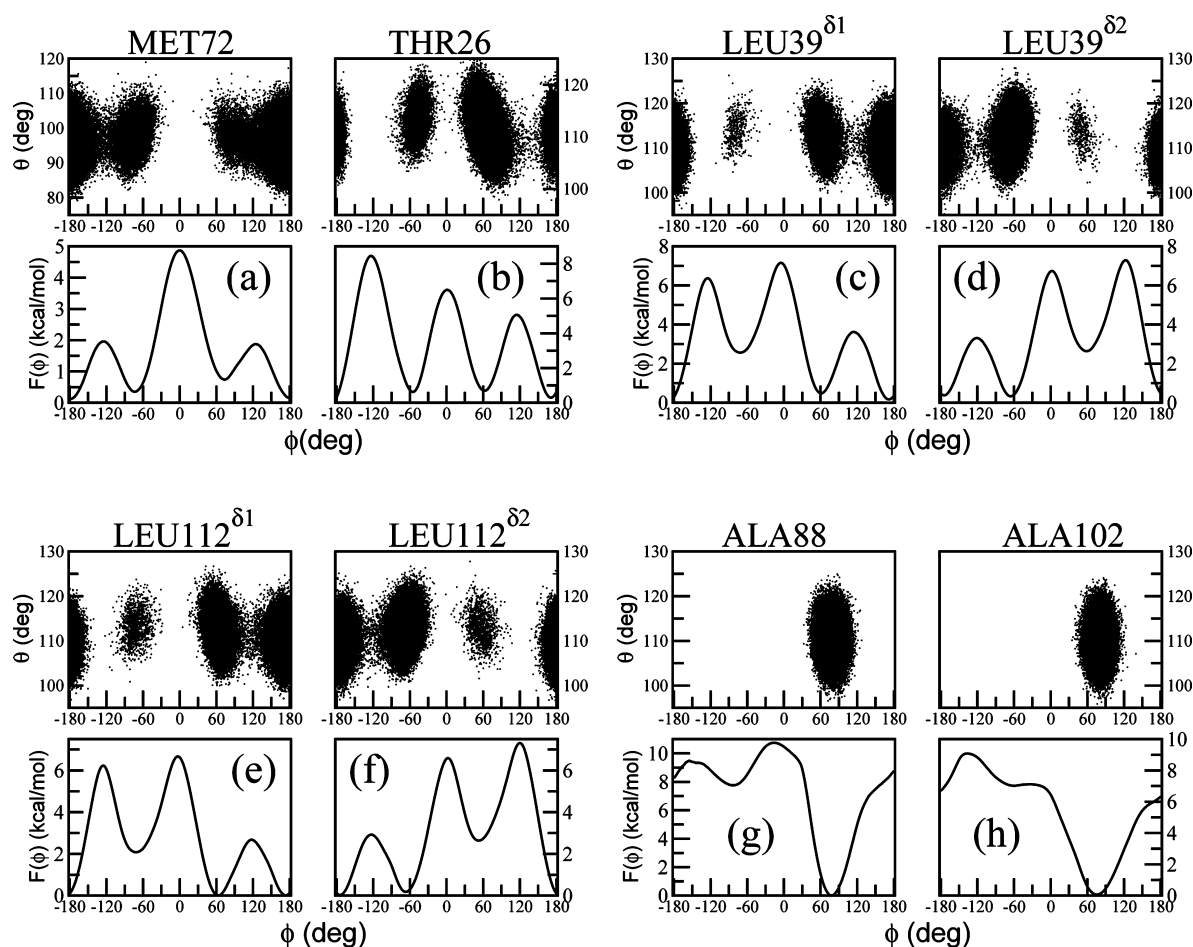
Figure 2. Scatter plots of the polar (θ) and azimuthal (ϕ) angles, which describe the orientation of methyl symmetry axes, obtained from six independent 10-ns MD trajectories (top panels) and the corresponding conformational free energy profiles (computed using the ABF method) associated with the reorientational dynamics of the methyl symmetry axes (bottom panels) shown for representative side-chains of CaM. (a) MET72; (b) THR26; (c) LEU39$^{\delta1}$; (d) LEU39$^{\delta2}$; (e) LEU112$^{\delta1}$; (f) LEU112$^{\delta2}$; (g) ALA88; (h) ALA102.

equilibration and 10 ns of production run. Five further, independent, 10 ns production runs were also performed, starting from the final configuration obtained from the first 10 ns production run and with different initial velocities.

The most common problem in determining the side-chain conformational free energy landscape using MD simulation is that the accessible low-energy regions of the phase space dominate the Boltzmann distribution obtained and the high-energy regions are sampled incompletely. This inadequate sampling of the phase space results in statistical artifacts in conformational free energy profiles along reaction coordinates. To avoid this problem and accurately determine the conformational free energy landscapes of methyl-bearing side-chains of CaM, we use the ABF method implemented in NAMD (Collective Variables Module).[43−45] In the ABF method, the mean force exerted on a selected reaction coordinate is used to accelerate the conformational sampling and to determine the associated free energy profiles by performing thermodynamic integration of the mean force as a function of the reaction coordinate.

The torsional angle, ϕ, of a given methyl-containing side-chain was selected as the reaction coordinate appropriate for the determination the conformational free energy profile, $F(\phi)$, of that side-chain. The mean force as a function of ϕ was accumulated in bins of width 1° from a total simulation time of 60 ns. The average free energy profile obtained from two independent 60 ns ABF calculations carried out at 300 K and 1 atmospheric pressure was used to determine the NMR order parameters of protein side-chains. The convergence of the free energy profiles was examined by monitoring the time evolution of the free energy profiles at regular intervals of 7.2 ns along the trajectory, and the differences in the free energy profiles between two adjacent time windows were used as a measure of convergence of our simulations. Let $F(\phi, t)$ and $F(\phi, t + \Delta t)$ be the values of free energies at ϕ and at two adjacent time windows, t and $t + \Delta t$, respectively. The difference $\Delta F = F(\phi, t + \Delta t) - F(\phi, t)$ was required to be <0.05 kcal/mol at all values of ϕ for all methyl dihedrals (Supporting Information - Figure S3a). In addition, a convergence quantity, ζ, was defined as follows:

$$\zeta = \sum_{\phi=-\pi}^{\phi=\pi} \sqrt{[F(\phi, t + \Delta t) - F(\phi, t)]^2} \tag{4}$$

It was observed that with increasing time, the value of ζ decreases to a value close to zero (Supporting Information - Figure S3b).

RESULTS AND DISCUSSION

First, we examine the side-chain conformational free energy profiles computed using the ABF method. Figure 2 shows the free energy profiles, $F(\phi)$, associated with ϕ for a few representative side-chains of CaM. Since the bend angle, θ, for all proteins is generally restricted to a narrow window around $\theta = 100°$, the free energy profiles, $F(\theta)$, associated with θ can be calculated from the probability distribution of θ obtained from the time evolution of θ ($F(\theta)$ is not shown here). The profiles of MET72, THR26, LEU39$^{\delta 1}$, LEU39$^{\delta 2}$, LEU112$^{\delta 1}$, and LEU112$^{\delta 2}$ consist of three rotamer states at about $\phi = \pm 180°$ (corresponding to the *trans* conformation) and $\phi = \pm 60°$ (corresponding to the *gauche* conformations) separated by three barriers, at $\phi = 0°$ (corresponding to the *cis* state) and $\phi = \pm 120°$ (corresponding to the *eclipsed* conformation). The least stable conformation (with the highest free energy) for all ILE$^{\delta 1}$, MET, VAL$^{\gamma 2}$ residues in CaM is observed at $\phi = 0°$. The free energy profiles of MET residues are symmetric with respect to $\phi = 0°$ while the profiles were nonsymmetric for other residues.

Table 1 presents the range of ΔF_{max}, the difference in free energy between the most and least stable side-chain

Table 1. Range of the Difference in Free Energy between the Most and Least Stable Side-Chain Conformations (ΔF_{max}) and Average Free Energy ($\langle \Delta F_{max} \rangle$) for Different Types of Residues of CaM

residue	ΔF_{max} (kcal/mol)	$\langle \Delta F_{max} \rangle$ (kcal/mol)
MET	3.8−5.2	4.6
LEU$^{\delta 1}$	6.7−7.8	7.1
LEU$^{\delta 2}$	6.9−8.9	7.5
VAL$^{\gamma 1}$	8.2−9.2	8.6
VAL$^{\gamma 2}$	8.3−10.0	9.2
ILE$^{\delta 1}$	7.3−9.7	8.4
ILE$^{\gamma 2}$	8.5−10.0	9.2
THR	8.0−11.1	9.6

conformations, for different types of residues of CaM. The average value of ΔF_{max} for MET residues ($\langle \Delta F_{max} \rangle$) is 4.6 kcal/mol, which is much lower than those of VAL$^{\gamma 2}$, ILE$^{\gamma 2}$, VAL$^{\gamma 1}$, and ILE$^{\delta 1}$. The value of ΔF_{max} is a measure of the extent of conformational space sampled by a given side-chain: the larger the ΔF_{max}, the more conformationally restricted the side-chain. Now the extent of conformational space sampled by a given residue determines the value of O^2_{axis}: a residue with a higher conformational flexibility is expected to have lower O^2_{axis}. Hence, the lower values of ΔF_{max} for MET residues of CaM indicate that the O^2_{axis} values may also be lower for MET residues in CaM, and this observation is consistent with the recent site-specific NMR relaxation studies on calmodulin.[7] Similarly, the values for LEU$^{\delta 1}$ and LEU$^{\delta 2}$ are relatively low compared to other residues, indicating a larger conformational flexibility of LEU residues in CaM. By contrast, the higher values of ΔF_{max} for ILE, THR, and VAL$^{\gamma 2}$ indicate that these residues are more conformationally restricted.

The conformational sampling of ALA residues in proteins merits special consideration. The free energy profiles of ALA residues (shown in Figure 2g,h) consist of only one rotamer state, with ΔF_{max} ranging from 2.3 to 22.5 kcal/mol. The lowest ΔF_{max} value was observed for ALA147, which is thus expected to have a lower O^2_{axis}, which is again consistent with the experimental results.[7] The values of ΔF_{max} for ALA1 (4.8 kcal/

mol), ALA10 (5.6 kcal/mol), and ALA128 (5.9 kcal/mol) of CaM are less than 6.0 kcal/mol, and thus they would be expected to have smaller O^2_{axis} values than ALA73 (22.5 kcal/mol), ALA88 (10.8 kcal/mol), and ALA15 (10.1 kcal/mol), for which ΔF_{max} is greater than 10.0 kcal/mol. Site-specific NMR experiments have revealed such restricted motions within a rotamer state for alanines in CaM.[2,46] This was attributed to the fact that the side-chain dynamics of alanines are governed by the motions of the backbone to which they are directly bonded.

Figure 2 also shows scatter plots of θ/ϕ distributions calculated directly from six independent 10-ns MD trajectories, i.e., without using ABF. The dihedral angle ϕ is distributed over a wide range between $\phi = -180°$ and $\phi = 180°$, while the bend angle θ is restricted to a narrow window between $\theta = 80°$ and $\theta = 125°$. It is also evident from Figure 2 that the θ/ϕ distributions obtained from MD trajectories are consistent with the free energy profiles obtained from the ABF method. For example, θ/ϕ data points for MET72 are clustered around $\phi = \pm 65°$ and $\phi = \pm 180°$, indicating that these are preferred conformations, and these rotamer states exactly coincide with minima on the corresponding free energy surface. Similarly, the absence of data points between $\phi = -60°$ and $\phi = 60°$ is due to the presence of the highest barrier on the corresponding free energy surface. The data on the θ/ϕ scatter plot for ALA residues in CaM are clustered around the single rotamer state, indicating the restricted dynamics of ALA residues of CaM.

The comparison of the free energy profiles with θ/ϕ distributions for each methyl-containing residue allows identification of those side-chains that suffer from poor conformational sampling during MD simulations. Although MD trajectories are useful for deriving methyl symmetry axis order parameters, insufficient sampling results in overestimation of MD-derived order parameters. Figure 3 shows the free energy profile and θ/ϕ distribution for two representative residues the side-chain conformations of which are incompletely sampled during the MD simulations: for ILE125$^{\gamma 2}$ and VAL136$^{\gamma 1}$, the converged free energy profile consists of three stable rotamer states, while the θ/ϕ distribution is clustered
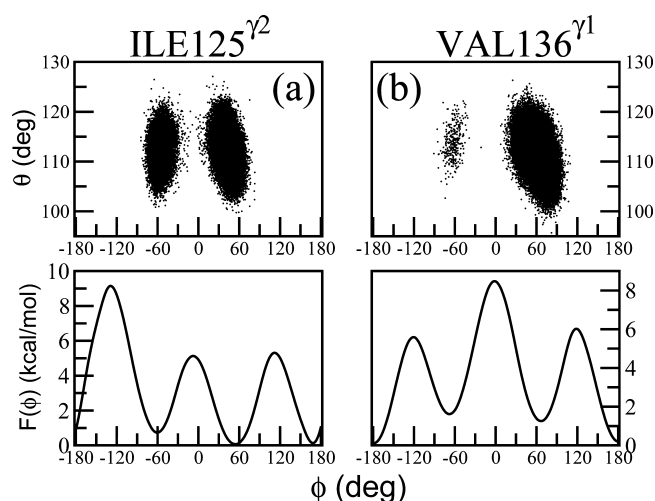

Figure 3. Scatter plots of θ and ϕ obtained from six independent 10-ns MD trajectories (top panels) and the corresponding conformational free energy profiles computed using the ABF method (bottom panels) are shown for two representative side-chains of CaM that exhibit insufficient conformational sampling during MD simulations. (a) ILE125$^{\gamma 2}$; (b) VAL136$^{\gamma 1}$.

around only two rotamer states. The residues that exhibit insufficient sampling were found to have relatively large values of ΔF_{max}.

The probability distribution function, $P(\phi)$ was computed using eq 3 from the side-chain free energy profiles obtained from the ABF method. For all residues except alanines, $P(\phi)$ consists of three peaks of varying heights corresponding to the three stable rotamer states. Let ϕ_1, ϕ_2, and ϕ_3 denote the major (the value of ϕ at which $P(\phi)$ is peaked at its maximum value), intermediate (the value of ϕ at which $P(\phi)$ is peaked at its second maximum), and minor (the value of ϕ at which $P(\phi)$ is peaked at its third maximum) rotamer states. The populations (χ) of the major (χ_1), intermediate (χ_2), and minor (χ_3) rotameric states of any given side-chain can be determined from $P(\phi)$ using the following equation:

$$\chi_i = \int_{\phi_i - \pi/3}^{\phi_i + \pi/3} P(\phi)\, d\phi \tag{5}$$

where i can take any integer value from 1 to 3.

Using eqs 2 and 5, we computed O^2_{axis}, χ_1, and χ_2 for all methyl-containing side-chains of calmodulin, and the results are shown in Figure 4, revealing a systematic correlation between
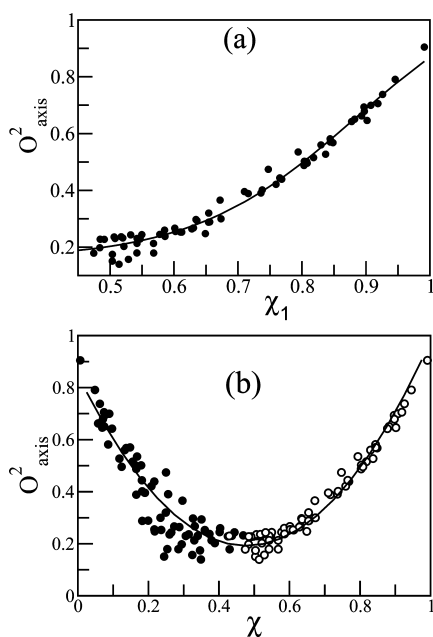

Figure 4. (a) Correlation between the population of major rotamer state (χ_1) and the O^2_{axis}. The solid line is the best fit to the data using eq 6. (b) Relationship between the populations of major (χ_1 - open circles) and intermediate (χ_2 - filled circles) rotamer states and O^2_{axis}. The solid line is the best fit to the data using eq 7.

O^2_{axis} and the populations of rotamer states. Figure 4a shows that O^2_{axis} increases nonlinearly with increasing χ_1; this dependence can be described using a sigmoidal-type function defined as follows:

$$O^2_{axis} = a_0 + \frac{1}{1 + e^{-a_1(\chi_1 - a_2)}} \tag{6}$$

The solid line in Figure 4a shows the best fit to the data using eq 6, obtained with parameters $a_0 = 0.156$, $a_1 = 7.844$, and $a_2 = 0.884$.

Equation 6 has a simple physical meaning. The higher the value of χ_1 (i.e., the larger the barrier separating the major rotamer state from the other states), the more restricted the motion of the side-chain. Also, the parameter a_2 is the value of χ_1 at which the order parameter is equal to $0.5 + a_0$. Hence, those side-chains with $\chi_1 > a_2$ are more sterically hindered (i.e., the dynamics is confined to a single rotamer state), while side-chains with $\chi_1 < a_2$ experience less steric hindrance and thus exhibit higher conformational flexibility. The rate of decay of O^2_{axis} with respect to χ_1 is determined by the parameter a_1. A similar, sigmoidal-type relationship has been derived between the barrier height on a 3-fold symmetric $F(\phi)$, which describes methyl rotation, and the O^2_{rot} order parameter corresponding to methyl rotation.[36,47,48]

The dependence of O^2_{axis} on the populations of the major and intermediate rotamer states is shown Figure 4b. The value of O^2_{axis} decreases from ~0.9 to ~0.15 when χ_2 increases from 0 to ~0.5. $\chi_2 = 0$ corresponds to the side-chain exhibiting restricted motion within a single rotamer state, while $\chi_2 = 0.5$ corresponds to the side-chain conformational transitions resulting in the major and intermediate states being equally populated with $\chi_3 = 0$. The lowest limit of O^2_{axis} is 0.111 and is achieved when all three states are equally populated (i.e., $\chi_1 = \chi_2 = \chi_3 = 1/3$). This O^2_{axis} dependence of χ can be fitted by the following quadratic function (i.e., an asymmetric parabola):

$$O^2_{axis} = b_0(\chi - b_1)(\chi - b_2) + b_3 \tag{7}$$

The solid line in Figure 4b shows the best fit of eq 7 to the data, with corresponding parameters $b_0 = 2.88$, $b_1 = 0.501$, $b_2 = 0.454$, and $b_3 = 0.194$. The function defined in eq 7 describes the variation of O^2_{axis} with χ accurately for methyl groups with $0.35 \leq O^2_{axis} \leq 1.0$. The spread in the data for methyls with $O^2_{axis} < 0.35$ indicates that the populations of rotamer states estimated (using eq 7) from O^2_{axis} for these methyls can have non-negligible uncertainties. There are at least two possible reasons for this spread of data: the dependence of O^2_{axis} on χ is different for different types of residue and/or a given value of O^2_{axis} may correspond to different values of χ_1 and χ_2 for a given type of residue.

To estimate the contributions of the above components, we examine the dependence of O^2_{axis} on χ for different types of residue (Figure 5). This dependence can be fitted by the following quadratic function (i.e., an asymmetric parabola):

$$O^2_{axis} = c_0(\chi - c_1)(\chi - c_2) + 0.5(c_1 + c_2) \tag{8}$$

It is evident from Figure 5 that eq 8 describes the variation of O^2_{axis} with χ well and the fitted parameters are different for different residue types: these parameters are listed in Table 2. The parameter c_0 is related to the curvature ($(\partial^2/\partial \chi^2)O^2_{axis}$) of the asymmetric parabola and c_1 and c_2 are chosen such that $\chi = 0.5(c_1 + c_2)$ corresponds to the minimum of the asymmetric parabola.

The observed correlation between O^2_{axis} and χ suggests that it may be possible to determine the relative populations of rotamer states of protein side-chains directly from experimentally determined O^2_{axis} order parameters. That is, the experimental O^2_{axis} order parameters obtained from high-resolution site-specific NMR experiments may be inserted into eq 8 to determine χ_1, χ_2, and χ_3 ($\chi_3 = 1 - \chi_1 - \chi_2$). For a given value of O^2_{axis}, the two roots of this quadratic equation can be determined, and they correspond to the values of χ_1 (the root with a larger value) and χ_2 (lower root). The values of χ_i
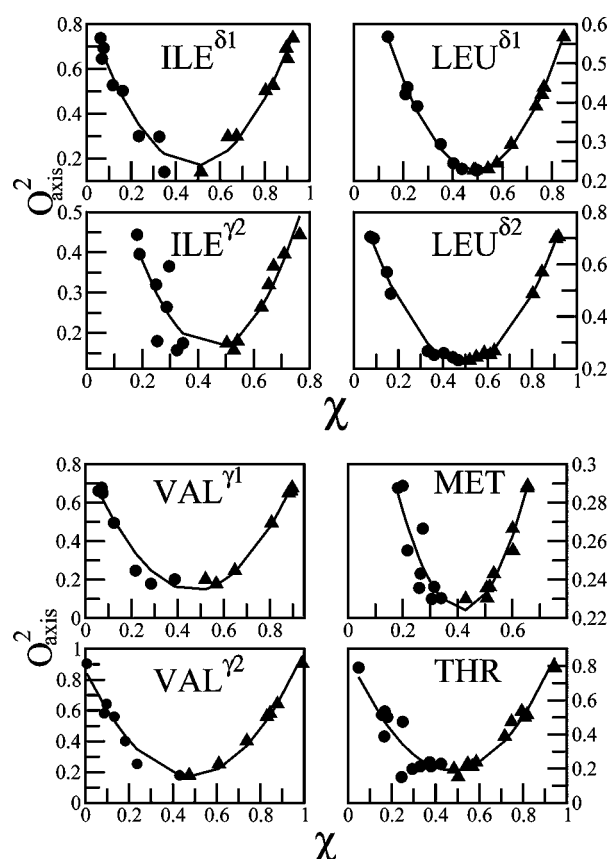
Figure 5. Correlation between the populations of major (χ_1 - filled triangles) and intermediate (χ_2 - filled circles) rotamer states and O^2_{axis} for different types of residues. The solid line is the best fit to the data using eqs 8.

Table 2. Fitted Parameter Values for Eq 8

residue	c_0	c_1	c_2
LEU$^{\delta 1}$	2.692 ± 0.047	0.179 ± 0.002	0.804 ± 0.004
LEU$^{\delta 2}$	2.751 ± 0.059	0.181 ± 0.003	0.807 ± 0.005
VAL$^{\gamma 2}$	2.991 ± 0.167	0.159 ± 0.007	0.801 ± 0.012
VAL$^{\gamma 1}$	3.022 ± 0.171	0.142 ± 0.007	0.802 ± 0.012
ILE$^{\delta 1}$	2.972 ± 0.173	0.157 ± 0.007	0.809 ± 0.011
ILE$^{\gamma 2}$	3.393 ± 0.519	0.159 ± 0.016	0.748 ± 0.020
MET	1.114 ± 0.115	0.001	0.829 ± 0.021
THR	2.992 ± 0.241	0.166 ± 0.010	0.786 ± 0.017

thus obtained can be used to reconstruct the equilibrium conformational distributions, $P(\phi)$, of side-chains using a three-Gaussian distribution function of the following form:

$$P(\phi) = \frac{\chi_1 e^{-(\phi-\phi_1)^2/2\sigma_1^2}}{\sqrt{2\pi\sigma_1^2}} + \frac{\chi_2 e^{-(\phi-\phi_2)^2/2\sigma_2^2}}{\sqrt{2\pi\sigma_2^2}} + \frac{\chi_3 e^{-(\phi-\phi_3)^2/2\sigma_3^2}}{\sqrt{2\pi\sigma_3^2}} \qquad (9)$$

where σ_1, σ_2, and σ_3 are the widths of the Gaussian functions centered at the major (ϕ_1), intermediate (ϕ_2), and minor (ϕ_3) rotamer states.

For a given residue type, it is straightforward to determine χ_i from O^2_{axis} by solving eq 8. However, σ_i and ϕ_i are unknown and

need to be computed for different types of residue. To do this, eq 9 was fitted directly to the data of $P(\phi)$ obtained from the ABF-based MD using eq 3 for each methyl group in CaM. The average values of the widths of the Gaussian functions (σ_i) calculated for each residue type are presented in Table 3. The

Table 3. Average Values of the Widths (in Radians) of the Gaussian Functions Peaked around the Major ($\langle\sigma_1\rangle$), Intermediate ($\langle\sigma_2\rangle$), and Minor ($\langle\sigma_3\rangle$) Rotamer States for Different Types of Residues of CaM

residue	$\langle\sigma_1\rangle$	$\langle\sigma_2\rangle$	$\langle\sigma_3\rangle$
LEU$^{\delta 1}$	0.180 ± 0.007	0.217 ± 0.048	0.230 ± 0.029
LEU$^{\delta 2}$	0.179 ± 0.018	0.218 ± 0.033	0.222 ± 0.022
VAL$^{\gamma 2}$	0.151 ± 0.016	0.152 ± 0.013	0.155 ± 0.021
VAL$^{\gamma 1}$	0.164 ± 0.020	0.160 ± 0.017	0.168 ± 0.014
ILE$^{\delta 1}$	0.213 ± 0.023	0.193 ± 0.048	0.247 ± 0.102
ILE$^{\gamma 2}$	0.167 ± 0.014	0.150 ± 0.025	0.172 ± 0.014
MET	0.280 ± 0.013	0.247 ± 0.010	0.260 ± 0.020
THR	0.154 ± 0.016	0.146 ± 0.014	0.162 ± 0.023

values of $\langle\sigma_i\rangle$ are relatively large for MET, LEU$^{\delta 1}$, and LEU$^{\delta 2}$, indicating higher conformational flexibility, while THR and VAL$^{\gamma 2}$ residues exhibit relatively narrower $P(\phi)$ distributions with lower values of $\langle\sigma_i\rangle$. This observation is consistent with the trend seen in the values of $\langle\Delta F_{max}\rangle$ (see Table 1).

To reconstruct the side-chain conformational distribution the accurate estimation of ϕ_i is essential. However, the computation of the average values of ϕ_i is nontrivial, the reason being that the features of the free energy profiles (such as the rotamer energies and barriers separating rotamer states) vary among residues of a given residue type depending upon the local environment. This variation implies that the populations of rotamer states are not unique for a given residue type but are widely distributed between 0% and 100%, and this observation is consistent with recent NMR results.[49−52] The wide distribution of the populations of rotamer states implies that, for any given residue type, any of the t, g_+, and g_- states can be a major rotamer state.

By examining the free energy profiles, we classify residues based on the rotamer sequence (RS), which is defined based on the values of ϕ_1, ϕ_2, and ϕ_3. For example, "tg_+g_-" is the RS of a residue if t, g_+, and g_- states correspond to the major, intermediate, and minor states, respectively, i.e., the first, second, and third letters in the sequence define the major, intermediate, and minor rotamers, respectively. Let n be the number of residues of a given residue type and n_S be the number of residues of the same residue type having a sequence S. Then, the probability (P_S) that any given type of residue has sequence S is $P_S = n_S/n$. This classification based on the sequence type is useful in determining meaningful values of $\langle\phi_i\rangle$. That is, $\langle\phi_i\rangle$ can be calculated for a specific sequence by averaging over members of a given residue type with same S. The average values of ϕ_i for different RS (S) are presented in Table 4 together with the corresponding values of P_S.

The $trans(t)$ state is the major rotamer state for 70% of the residues, while for 25%, $trans$ is the intermediate state. The values of $\langle\phi_1\rangle$ for all MET, ILE$^{\delta 1}$, and VAL$^{\gamma 2}$ residues are centered around 3.1, while for 67% of THR, 56% of LEU$^{\delta 1}$, and 89% of LEU$^{\delta 2}$, the value of $\langle\phi_1\rangle$ is close to 3.0. For ILE$^{\delta 1}$, tg_-g_+ is the most probable RS with $\langle\phi_1\rangle$ = 3.06, $\langle\phi_2\rangle$ = 5.21, and $\langle\phi_3\rangle$ = 1.20, and these values are consistent with the values $\langle\phi_1\rangle$ = 3.04, $\langle\phi_2\rangle$ = 5.17, and $\langle\phi_1\rangle$ = 1.12 obtained from recent NMR

Table 4. Average Values of ϕ_1, ϕ_2, and ϕ_3 (in Radians) for Different RS Values (S) Obtained from the ABF Calculations Are Presented for Different Types of Residues of CaM[a]

residue	S	P_S	$\langle\phi_1\rangle$	$\langle\phi_2\rangle$	$\langle\phi_3\rangle$
MET	tg_+g_-	6/9	3.14	1.31	5.00
	tg_-g_+	3/9	3.14	5.03	1.28
LEU$^{\delta 2}$	tg_-g_+	6/9	3.20	5.14	1.08
	tg_+g_-	2/9	3.04	1.11	5.13
	g_+tg_-	1/9	1.09	3.03	4.97
LEU$^{\delta 1}$	tg_-g_+	3/9	3.23	5.15	1.00
	tg_+g_-	2/9	3.01	1.08	5.04
	g_+tg_-	3/9	1.07	3.00	4.98
	g_-tg_+	1/9	5.20	3.20	1.20
VAL$^{\gamma 2}$	tg_-g_+	3/7	3.04	5.15	1.20
	tg_+g_-	3/7	3.01	1.16	5.15
	g_+tg_-	1/7	1.08	2.95	5.16
VAL$^{\gamma 1}$	tg_+g_-	2/7	3.07	1.05	5.17
	g_+tg_-	4/7	0.98	2.98	5.24
	g_+g_-t	1/7	0.98	5.28	2.94
ILE$^{\delta 1}$	tg_-g_+	5/8	3.06	5.21	1.20
	tg_+g_-	3/8	3.11	1.16	5.16
ILE$^{\gamma 2}$	tg_+g_-	2/8	3.15	5.13	1.15
	g_+tg_-	3/8	0.95	2.95	5.24
	g_+g_-t	2/8	1.03	5.18	2.85
	g_-g_+t	1/8	5.25	1.05	2.91
THR	tg_-g_+	4/12	3.02	5.31	1.12
	tg_+g_-	4/12	2.93	1.10	5.25
	g_+tg_-	4/12	1.06	2.92	5.29

[a]P_S is the probability of finding a given RS S, and it is expressed as a fraction of residues of a given type that show a given RS S.

experiments.[49] The free energy profiles of ILE$^{\delta 1}$ residues exhibit a least-populated additional rotamer state around $\phi = 1.74$ (Supporting Information - Figure S1), called the "gauche100" state.[49] For the reconstruction of $P(\phi)$ from O^2_{axis}, we consider only the t, g_+, and g_- states.

For most LEU residues, the difference in free energy between the ϕ_1 and ϕ_2 states is less than k_BT (at 300 K), and these states may be considered isoenergetic (Supporting Information - Figure S2). That is, the major and intermediates states are almost equally populated, and thus it is difficult to distinguish them. However, the RS values reported in Table 4 were obtained using a stringent rule based on the difference in free energy between the major and intermediate states: ϕ_1 is the major state if $F(\phi_1) - F(\phi_2) < 0$; otherwise it is an intermediate state. For most LEU residues, it was observed that the free energy of the *trans* state is slightly lower than that at the intermediate *gauche* conformation. The barriers separating the minor state from the other two states are large (greater than 6 kcal/mol), thus providing a rationale for why the minor rotamer states of LEU residues are least (or zero) populated.[51,52] The above RS analysis suggests that for a given residue type and an RS, the average values of σ_i and ϕ_i reported in Table 3 and Table 4 can be used in eq 9 to determine $P(\phi)$, which in turn can be inverted using the Boltzmann formula to derive $F(\phi)$. The values of $\langle\sigma_i\rangle$ and $\langle\phi_i\rangle$ computed here are consistent with the corresponding values obtained from a recently developed X-ray crystallography-based approach (called Ringer), which systematically samples electron density associated with the dihedral angles of proteins and determines the distributions of side-chain rotamers.[53,54]

Figure 6 compares the conformational distributions obtained from O^2_{axis} using eq 9 for representative side-chains of CaM with
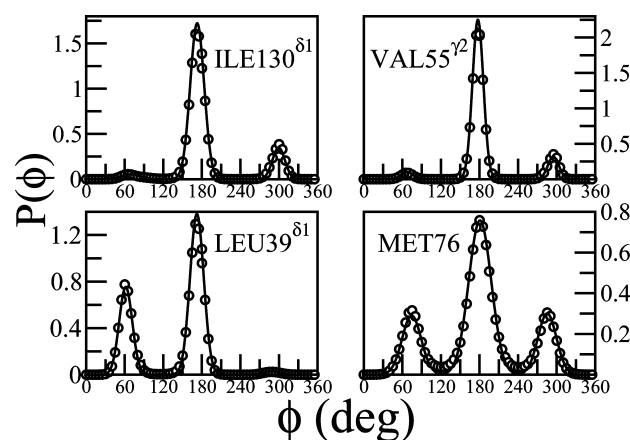

Figure 6. Side-chain conformational distributions reconstructed from O^2_{axis} (solid line) compared with those obtained from the ABF method (open circles) for representative side-chains of CaM.

those obtained from the ABF calculations. It is evident that the side-chain conformational distributions obtained using eq 9 are in excellent agreement with those obtained using ABF.

■ COMPARISON WITH EXPERIMENTAL DATA

The distribution of the generalized order parameters (shown in Figure 7) obtained from the ABF calculation exhibits three
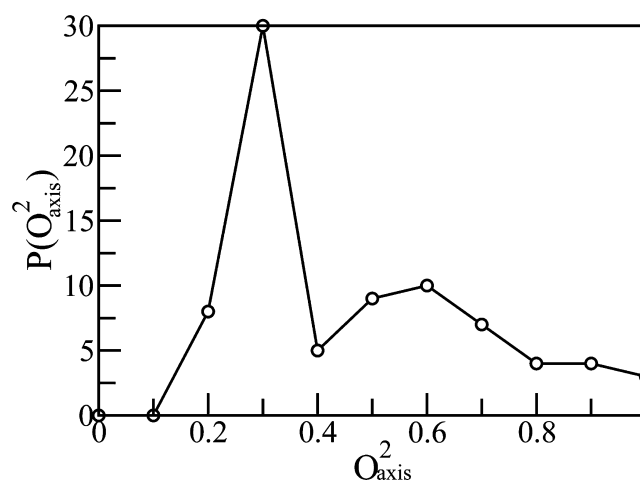

Figure 7. The distribution of the order parameters (O^2_{axis}) obtained from the ABF calculation.

peaks located (in order of decreasing intensity) at $O^2_{axis} \sim 0.3$, $O^2_{axis} \sim 0.6$, and $O^2_{axis} \sim 0.9$. The trimodal distribution of O^2_{axis} and the corresponding peak positions are consistent with recent site-specific NMR experimental investigations, which attribute the three peaks to three classes of side-chain motions: J-class (side-chains with $O^2_{axis} \sim 0.3$), ω-class (side-chains with $O^2_{axis} \sim 0.9$), and α-class (side-chains with $O^2_{axis} \sim 0.6$).[2,46,55] The side-chains in the J-class exhibit frequent rotameric interconversion between different rotamer states on the nanosecond or faster time scale, while members in the ω-class exhibit restricted motion within a rotamer well. The side-chains in the α-class exhibit restricted motion within a rotamer state with occasional jumps between different rotamer states.

Figure 8 compares the calculated and experimental O^2_{axis} for CaM in the unbound form. A striking difference between earlier
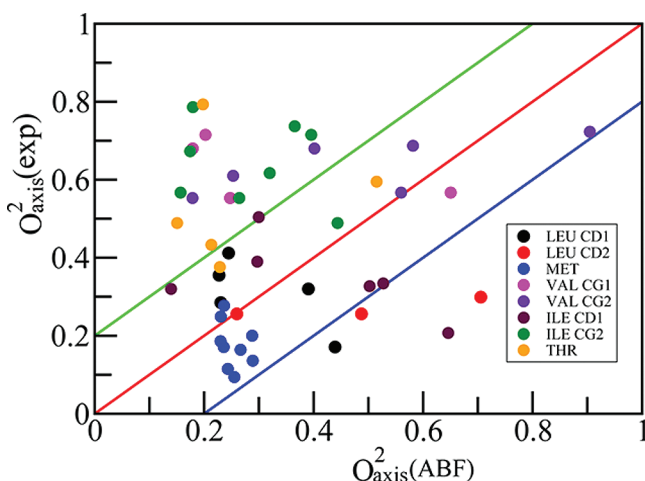

Figure 8. Comparison of the calculated (O^2_{axis}(ABF)) and experimental (O^2_{axis}(exp)) order parameters for CaM in the unbound form. The straight lines (red, O^2_{axis}(exp) = O^2_{axis}(ABF); green, O^2_{axis}(exp) = O^2_{axis}(ABF) + 0.2; blue, O^2_{axis}(exp) = O^2_{axis}(ABF) − 0.2) are shown as guides to the eyes.

MD-based studies of O^2_{axis} and the present study is evident from Figure 8. Although earlier MD-based investigations have better results for the backbone order parameters, the side-chain order parameters were overestimated (the value of O^2_{axis}) for proteins. The discrepancies between the experimental and simulated O^2_{axis} were attributed to inadequate sampling of the conformational space of the side chains. In the present study we observe a negligible number of residues with O^2_{axis} larger than the experimental values, and a significant number of data points are located near the diagonal line (in particular, side-chains of MET, LEU$^{\delta 1}$, and ILE$^{\delta 1}$). It is also evident from Figure 8 that the values of O^2_{axis} for some residues are lower than the corresponding experimental values. In particular, O^2_{axis} of ILE$^{\gamma 2}$, VAL$^{\gamma 1}$, and THR obtained from the ABF calculations are lower than those obtained experimentally. There are at least two possible reasons for these differences. The first reason involves the empirical force field parameters used in the simulation. Although the force field parameters provide good estimates of O^2_{axis} for many residues, such as MET, LEU$^{\delta 1}$, and ILE$^{\delta 1}$, the systematic under-estimation of O^2_{axis} for ILE$^{\gamma 2}$, VAL$^{\gamma 1}$, and THR indicates that there is scope for improvement of these parameters to accurately describe some side-chain dynamics of proteins.[56−59] Site-resolved NMR data-driven force field optimization strategies have been developed in recent years for better description of protein conformational dynamics.[58−60] These approaches perform better in describing the backbone dihedral energy surfaces, and NMR-based force field optimization methods have yet to realize their full potential in describing the side-chain dynamics of proteins. The second reason points toward the time-scale issues. The enhanced-sampling approach developed here is devoid of sampling errors. However, order parameters obtained from NMR experiments are sensitive to picosecond−nanosecond time-scale motions, and some of the calculated jumping between conformational states would be too slow to be picked up by the experimental O^2_{axis}. The relative contributions of these factors to the underestimation of O^2_{axis} will be examined in the future.

The structural data obtained from X-ray crystallography can be used to examine the veracity of the side-chain rotamer distributions generated from O^2_{axis}. The crystal structures (mostly obtained at liquid nitrogen temperatures) derived from X-ray diffraction provide information about most probable conformations of individual side-chains of proteins. For a given residue type, the methyl dihedral angles, ϕ, can be calculated from the atomic coordinates deposited in the Protein Data Bank, and then one can examine whether these values are consistent with the major rotamer states obtained from the ABF calculations.

Figure 9 compares the methyl dihedral angles of the X-ray structure (PDB ID: 1CLL) with $P(\phi)$ obtained from our
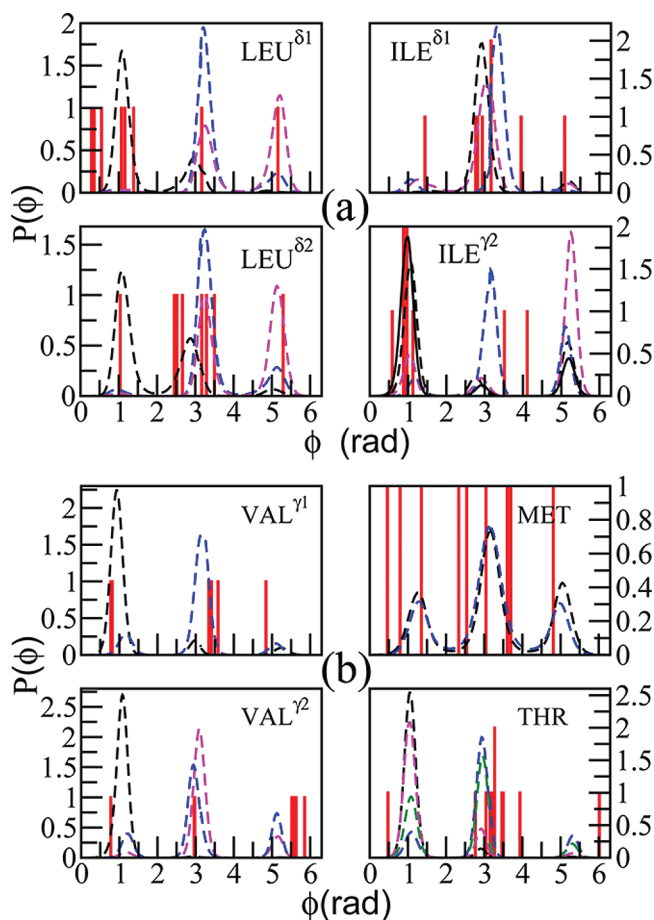

Figure 9. Side-chain conformational distributions reconstructed from O^2_{axis} (dashed line) compared with those obtained from the X-ray structure (red solid line). The calculated distributions are shown for a few representative residues of CaM.

calculations. In the crystal structure, 75% of LEU$^{\delta 2}$ residues are populated near the *trans* state, while the remaining are equally distributed between the g_+ and g_- states. 75% of ILE$^{\gamma 2}$ and 75% of LEU$^{\delta 1}$ are populated in the g_+ state, indicating that the g_+ state is the most stable conformer for these residues. However, the spread in dihedral angles (within the g_+ state) obtained from the X-ray data is more for LEU$^{\delta 1}$ than ILE$^{\gamma 2}$. As discussed before, in the CHARMM-based MD simulations, the g_+ and *trans* states of LEU$^{\delta 1}$ are iso-energetic and thus almost equally populated, whereas the g_+ state of LEU$^{\delta 1}$ is the most highly populated in the X-ray structure. For VAL$^{\gamma 2}$, the g_- state is the most populated in the crystal structure, whereas it is not the

major rotamer state for any of the VAL$^{\gamma2}$ residues in our simulations. Also, for THR residues, the *trans* state is the most highly populated in the crystal, whereas the g_+ state is also almost as highly populated as the *trans* in the simulations. We also notice that for VAL$^{\gamma1}$, the *trans* state calculated from the simulations is slightly shifted to lower values than the X-ray data, and also that the g_+ state seems to be slighly overstabilized in CHARMM-based MD simulations. The above observations suggest that there is some room for improvement of the CHARMM force field parameters in order for them to accurately represent the fine details of the side-chain conformational energy surfaces of proteins.[56−59]

The distributions of side-chain rotamers generated from O^2_{axis} are compared with those from the multiple 10-ns MD simulations in Figure 10 for a few representative side chains.
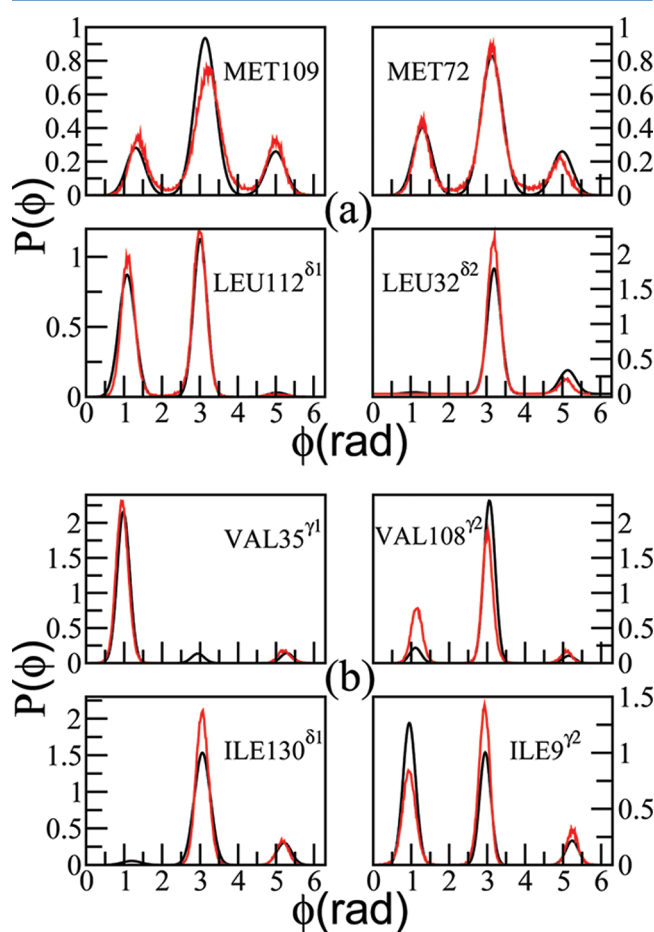

Figure 10. Side-chain conformational distributions reconstructed from O^2_{axis} (black) compared with those obtained from the multiple 10-ns MD simulations (red) for representative side-chains of CaM.

In both cases, the general features (such as peak positions and widths) of the distributions are similar, thus ascertaining the reliability of the O^2_{axis}-based approach to derive the side-chain rotamer distributions. The marginal differences in the peak intensities of distributions of some residues can be attributed to the inadequate sampling of the conformational space of these side-chains in MD simulations.

CONCLUSIONS

An enhanced-sampling-based computational approach, which alleviates the conformational sampling problem, has been developed to accurately determine methyl-containing side-chain conformational free energy profiles directly from NMR-derived methyl symmetry axis order parameters (O^2_{axis}). This approach has permitted the establishment of a nonlinear relationship between O^2_{axis} and the populations of side-chain rotamer states. This relationship permits the determination of the equilibrium conformational distributions of methyl-bearing side-chains of proteins directly from NMR methyl-axis motional parameters. For the test case of calmodulin, the equilibrium side-chain conformational distributions that are rapidly reconstructed using this approach were found to be in excellent agreement with those obtained from the ABF MD calculations.

A major focus of recent high-resolution NMR experiments is on examining how the relative populations of various conformers of side-chains change at different stages of biological processes ranging from protein folding to ligand-induced structural transitions to enzyme catalysis.[49−52] The populations and conformational distributions of side-chain rotamer states determined from O^2_{axis} using the present analytical approach provide insight into conformational free energy landscapes (using the Boltzmann formula (eq 3)) of proteins. It will be of interest to investigate how the rotamer populations are related to thermodynamic quantities (for example, the conformational entropy) and how the rates of transitions between rotamer states are related to free energy barriers that separate these states. This approach of determining $P(\phi)$ from O^2_{axis} may also be useful for examining the reliability of existing force field parameters used in biomolecular simulations and for deriving accurate dihedral and nonbonded force field parameters from the experimental O^2_{axis} order parameters.

A web-based computer program that calculates the side-chain rotamer populations and distribution of rotamer states from O^2_{axis} can be accessed from http://regol.iiit.ac.in/Conformation (Supporting Information - Web-based computer program).

ASSOCIATED CONTENT

Supporting Information
Figures S1 and S2 showing the conformational free energy profiles (computed using the ABF method) associated with the reorientational dynamics of the methyl symmetry axes of ILE$^{\delta1}$ and LEU residues, respectively, are provided. Figure S3 depicting the time evolution of the convergence quantity ζ is also provided. A discussion about a web-based computer program that calculates the side-chain rotamer populations and distribution of rotamer states from O^2_{axis} is included. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: m.krishnan@iiit.ac.in.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Daniel, R. M.; Dunn, R. V.; Finney, J. L.; Smith, J. C. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 69−92.

(2) Igumenova, T. I.; Frederick, K. K.; Wand, A. J. *Chem. Rev.* **2006**, *106*, 1672−99.

(3) Wand, A. J. *Nat. Struct. Biol.* **2001**, *8*, 926−931.

(4) Kay, L. E. *Nat. Struct. Biol.* **1998**, *5*, 513−517.

(5) Forman-Kay, J. D. *Nat. Struct. Biol.* **1999**, *6*, 1086−1087.

(6) Ishima, R.; Torchia, D. A. *Nat. Struct. Biol.* **2000**, *7*, 740−743.

(7) Lee, A. L.; Kinnear, S. A.; Wand, A. J. *Nat. Struct. Biol.* **2000**, *7*, 72−77.

(8) Lee, A. L.; Wand, A. J. *Nature* **2001**, *411*, 501−504.

(9) Tugarinov, V.; Hwang, P.; Ollerenshaw, J.; Kay, L. *J. Am. Chem. Soc.* **2003**, *125*, 10420−10428.

(10) Bremi, T.; Brüschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1997**, *119*, 4272−4284.

(11) Brüschweiler, R. *Curr. Opin. Struc. Biol.* **2003**, *13*, 175−183.

(12) Li, D. W.; Showalter, S. A.; Brüschweiler, R. *J. Phys. Chem. B* **2010**, *114*, 16036−16044.

(13) Law, A. B.; Fuentes, E. J.; Lee, A. L. *J. Am. Chem. Soc.* **2009**, *131*, 6322−6323.

(14) Petit, C. M.; Zhang, J.; Sapienza, P. J.; Fuentes, E. J.; Lee, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18249.

(15) Mauldin, R. V.; Lee, A. L. *Biochemistry* **2010**, *49*, 1606−1615.

(16) Li, Z.; Raychaudhuri, S.; Wand, A. J. *Protein Sci.* **1996**, *5*, 2647−2650.

(17) Nicholson, L. K.; Kay, L. E.; Baldisseri, D. M.; Arango, J.; Young, P. E.; Bax, A.; Torchia, D. A. *Biochemistry* **1992**, *31*, 5253−5263.

(18) Ruschak, A. M.; Kay, L. *J. Biomol. NMR* **2010**, *46*, 75−87.

(19) Ishima, R.; Petkova, P.; Louis, J. M.; Torchia, D. A. *J. Am. Chem. Soc.* **2001**, *123*, 6164−6171.

(20) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

(21) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.

(22) LeMaster, D. M. *J. Am. Chem. Soc.* **1999**, *121*, 1726−1742.

(23) Yang, D.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369−382.

(24) Yang, D.; Mok, Y.; Forman-Kay, J.; Farrow, N.; Kay, L. *J. Mol. Biol.* **1997**, *272*, 790−804.

(25) Akke, M.; Brüeschweiler, R.; Palmer, A. G. III. *J. Am. Chem. Soc.* **1993**, *115*, 9832−9833.

(26) Hu, H.; Hermans, J.; Lee, A. L. *J. Biomol. NMR* **2005**, *32*, 151−162.

(27) Woessner, D. E. *J. Chem. Phys.* **1962**, *36*, 1−4.

(28) London, R.; Avitabile, J. *J. Am. Chem. Soc.* **1978**, *100*, 7159−7165.

(29) Wittebort, R. J.; Szabo, A. *J. Chem. Phys.* **1978**, *69*, 1722−1736.

(30) Prabhu, N. V.; Lee, A. L.; Wand, A. J.; Sharp, K. A. *Biochemistry* **2003**, *42*, 562−570.

(31) DuBay, K. H.; Geissler, P. L. *J. Mol. Biol.* **2009**, *391*, 484−497.

(32) Case, D. *Acc. Chem. Res.* **2002**, *35*, 325−331.

(33) Best, R.; Clarke, J.; Karplus, M. *J. Mol. Biol.* **2005**, *349*, 185−203.

(34) Chatfield, D. C.; Szabo, A.; Brooks, B. R. *J. Am. Chem. Soc.* **1998**, *120*, 5301−5311.

(35) Chatfield, D. C.; Wong, S. E. *J. Phys. Chem. B* **2000**, *104*, 11342−11348.

(36) Krishnan, M.; Kurkal-Siebert, V.; Smith, J. C. *J. Phys. Chem. B* **2008**, *112*, 5522.

(37) Krishnan, M.; Smith, J. C. *J. Am. Chem. Soc.* **2009**, *131*, 10083−10091.

(38) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(39) Mackerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(41) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quiocho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177−1192.

(42) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(43) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9169−9183.

(44) Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904−2914.

(45) Hénin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. *J. Chem. Theory Comput.* **2009**, *6*, 35−47.

(46) Frederick, K.; Kranz, J.; Wand, A. *Biochemistry* **2006**, *45*, 9841−9848.

(47) Baudry, J. *J. Am. Chem. Soc.* **2006**, *128*, 11088−11093.

(48) Baudry, J.; Smith, J. C. *J. Phys. Chem. B* **2005**, *109*, 20572−20578.

(49) Hansen, D. F.; Neudecker, P.; Kay, L. E. *J. Am. Chem. Soc.* **2010**, *132*, 7589−7591.

(50) Hansen, D. F.; Kay, L. E. *J. Am. Chem. Soc.* **2011**, *133*, 8272−8281.

(51) Hansen, D.; Neudecker, P.; Vallurupalli, P.; Mulder, F.; Kay, L. *J. Am. Chem. Soc.* **2010**, *132*, 42−43.

(52) Bouvignies, G.; Vallurupalli, P.; Hansen, D. F.; Correia, B. E.; Lange, O.; Bah, A.; Vernon, R. M.; Dahlquist, F. W.; Baker, D.; Kay, L. E. *Nature* **2011**, *477*, 111−114.

(53) Lang, P. T.; Ng, H. L.; Fraser, J. S.; Corn, J. E.; Echols, N.; Sales, M.; Holton, J. M.; Alber, T. *Protein Sci.* **2010**, *19*, 1420−1431.

(54) Fraser, J. S.; van den Bedem, H.; Samelson, A. J.; Lang, P. T.; Holton, J. M.; Echols, N.; Alber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16247−16252.

(55) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448*, 325−329.

(56) Mackerell, A. D. Jr.; Feig, M.; Brooks, C. L. III. *J. Comput. Chem.* **2004**, *25*, 1400−1415.

(57) MacKerell, A. D. Jr.; Feig, M.; Brooks, C. L. III. *J. Am. Chem. Soc.* **2004**, *126*, 698−699.

(58) Showalter, S. A.; Brüschweiler. *J. Chem. Theory Comput.* **2007**, *3*, 961−975.

(59) Li, D. W.; Brüschweiler. *J. Chem. Theor. Comput.* **2011**, *7*, 1773−1782.

(60) Li, D. W.; Brüschweiler, R. *Angew. Chem., Int. Ed.* **2010**, *122*, 6930−6932.

4133

dx.doi.org/10.1021/jp2104853 | *J. Phys. Chem. B* 2012, 116, 4124−4133