Supplementary Information

Microscopic insights into the NMR relaxation based protein conformational entropy meter

Vignesh Kasinath, Kim A. Sharp and A. Joshua Wand

Graduate group in Biochemistry and Molecular Biophysics and the Johnson Research Foundation and Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, United States of America

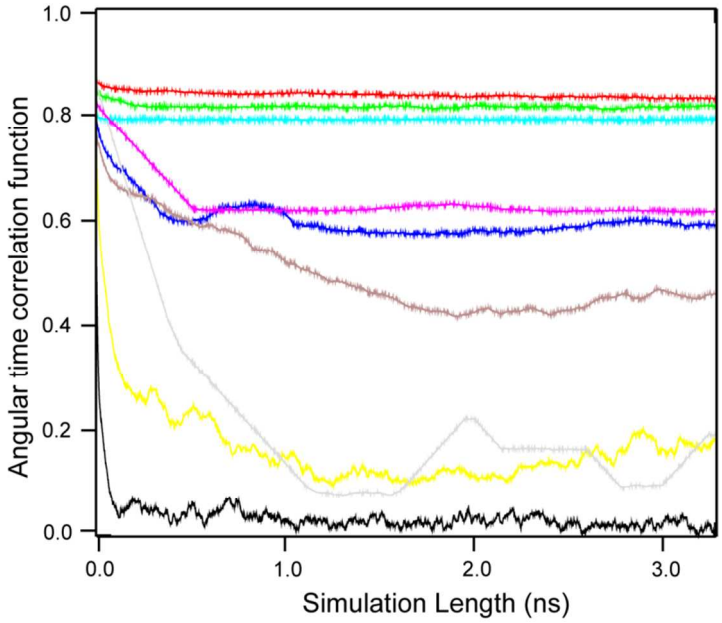

Figure S1. Convergence of methyl side-chain order parameters from MD simulations. Shown above are angular time correlation functions $C(t)=\langle P_2[\mu(t)(t+\tau)]\rangle$, where $\mu(t)$ is the instantaneous unit axis vector and $P_2(x)$ is the second degree Legendre polynomial, for a variety of residues (rigid to dynamic) from the highly dynamic α_3D and the highly rigid Hen Egg White Lysozyme (HEWL). Adequately sampling is indicated even on this short simulation time scale and that the excellent decay of autocorrelation functions suggest an overall satisfactory convergence of side-chain order parameters. The order parameters calculated from the plateau value of the autocorrelation functions above correlate well with the order parameter values calculated from the entire simulation length, using Chatfield et al. formulation [Eq. 1 main text]. The different residues of α_3D shown above are: (Red) - 16 THR, (Green) - 21 LEU, (Blue) - 42 LEU, (Yellow) - 56 LEU, and (Black) - 1 MET. The different residues of HEWL shown above are: (Brown) - 44 MET, (Grey) - 43 THR, (Cyan) - 56 LEU, and (Magenta) - 58 ILE.

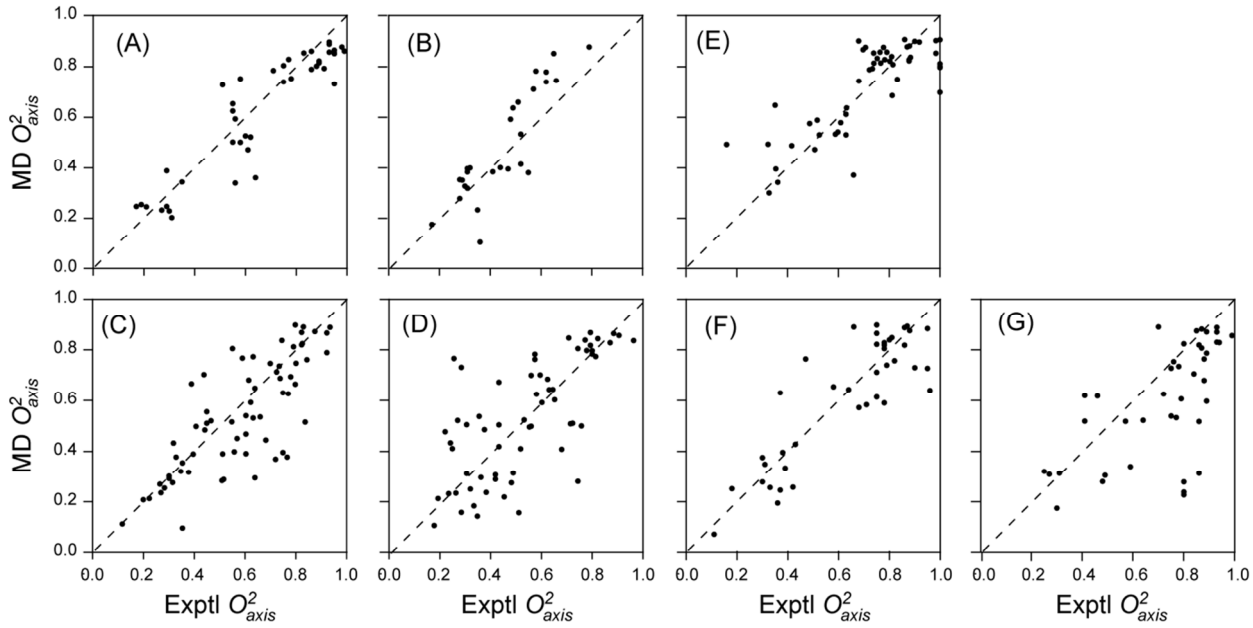

Figure S2. Correlation of methyl side-chain order parameters calculated from MD simulations versus those from NMR-relaxation based experiments. Each of the panels (a-g) above represent the different proteins used: (A) Ubiquitin (<r.m.s.d.> = 0.104) (B) α_3D (<r.m.s.d.> = 0.116) (C Calmodulin-smMLCKp (<r.m.s.d.> = 0.146) (D) Calmodulin-nNOSp (<r.m.s.d.> = 0.172) (E) Hen egg white lysozyme (<r.m.s.d.> = 0.124) (F) Adipocyte lipid binding protein (<r.m.s.d.> = 0.125) (G) Cytochrome c_2 (<r.m.s.d.> = 0.223). All the proteins studied exhibit a reasonable linear correlation with observable site-to-site variation. The parameters of the best fitted line are summarized in Table 2 (main text).

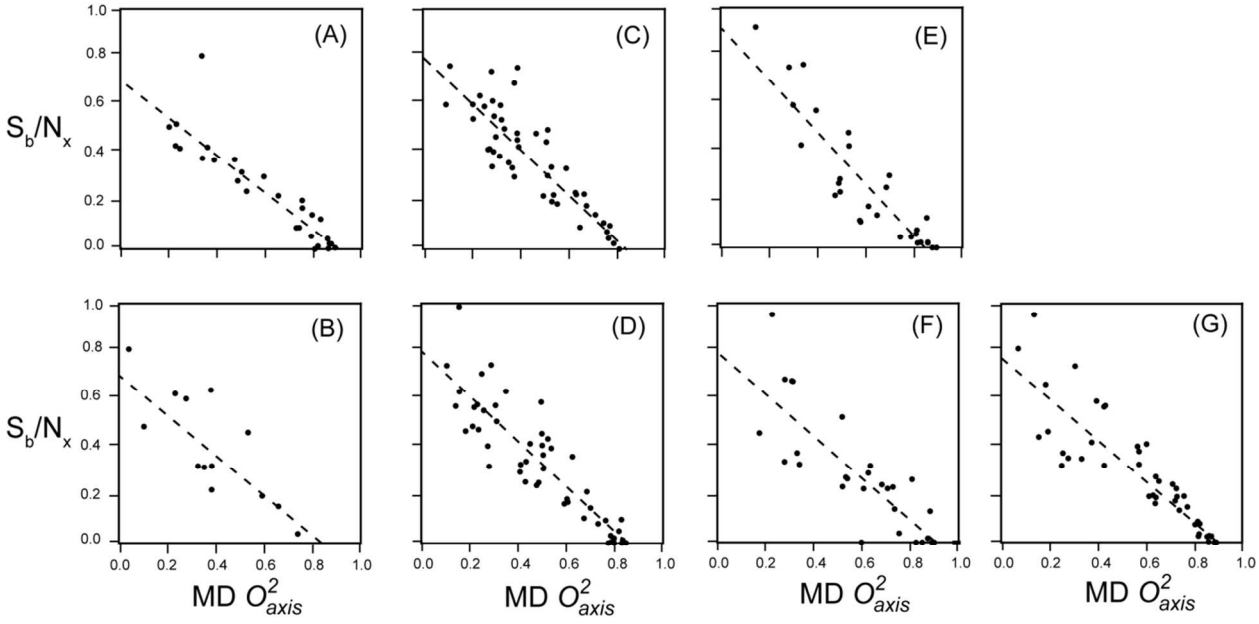

Figure S3. Correlation of rotamer entropy normalized by number of side-chain torsion angle to O2axis derived from MD simulations for methyl residues. Each of the panels (a-g) above represent the different proteins used: (A) Ubiquitin (B) α_3D (C) Calmodulin-smMLCKp (D) Calmodulin-nNOSp (E) Hen egg white

lysozyme (F) Adipocyte lipid binding protein (g) Cytochrome c_2 .The parameters of the best fitted line are summarized in Table S1.

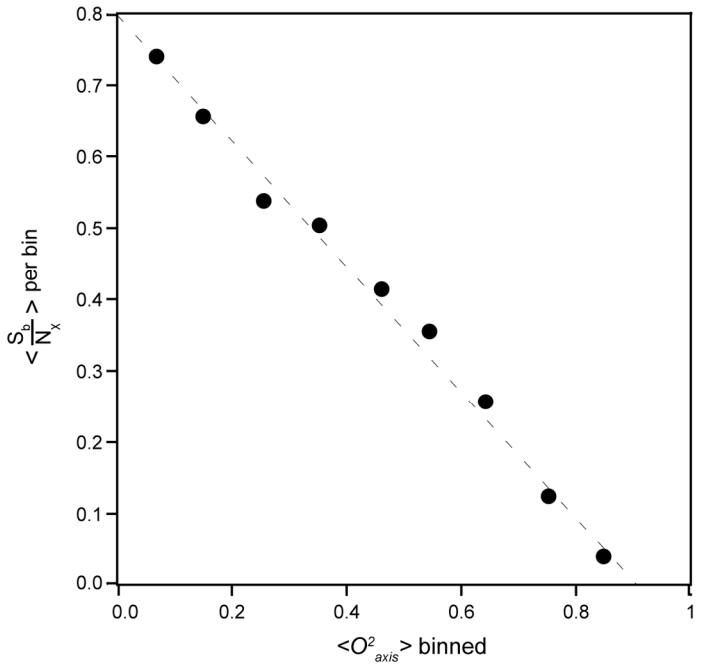

Figure S4. Rotamer entropy and side-chain methyl dynamics are transforms of each other over the entire range of order parameter. The linear relationship between the average O^2_{axis} in every $\Delta\ O^2_{axis}$ = 0.1 sized bins to the average rotamer entropy (normalized by the number of side-chain torsion angles) of the corresponding residues is high over the entire range of methyl side-chain. Data from all seven proteins simulated were used here resulting in a best fit linear line with slope of -0.88 and a R^2 of 0.99 (identical to Fig. 2 main text).

Table S1. Correlation of normalized rotamer entropy to methyl side-chain order parameters from MD simulations.

Protein[a]	slope	R^2	methyl s.c. entropy $\langle S_b/N_\chi \rangle$ [b]
ALBP	-0.84	0.79	0.276
α_3D	-0.88	0.64	0.399
Cyt c_2	-0.88	0.68	0.382
CaM-smMLCKp	-0.92	0.74	0.381
CaM-nNOSp	-0.95	0.80	0.335
HEWL	-1.10	0.81	0.257
Ubiquitin	-0.78	0.82	0.236

[a] Abbreviations: ALBP, adipocyte lipid binding protein; Cyt c_2, cytochrome c_2; CaM-smMLCKp, calcium-saturated calmodulin in complex with a peptide corresponding to the myosin light chain kinase calmodulin-binding domain; CaM-nNOSp, calcium-saturated calmodulin in complex with a peptide corresponding to the neuronal nitric oxide synthase calmodulin-binding domain; HEWL, hen egg white lysozyme

[b] Average rotamer entropy normalized by the respective number of side-chain torsion angles for methyl groups

Table S2. Dynamical proxy of methyl group is an excellent indicator of both methyl side-chain and total protein rotamer entropy.

Protein	methyl s.c. entropy[b] $(S_b/k_B)/N_\chi$	total sc. entropy[c] $(S_b/k_B)/N_\chi$	$\langle O^2_{axis}\rangle_{MD}$	1D s.c. entropy[d] (S_b/k_B)	MIST 2nd order[e] (S_b/k_B)	MIST 3rd order[f] (S_b/k_B)
ALBP	0.303 ± 0.020	0.433 ± 0.014	0.619 ± 0.014	101.73	89.34	79.37
α_3D	0.404 ± 0.009	0.495 ± 0.023	0.571 ± 0.008	82.94	76.85	76.75
Cyt c_2	0.261 ± 0.010	0.415 ± 0.008	0.670 ± 0.021	81.92	76.81	64.52
CaM-smMLCKp	0.395 ± 0.011	0.484 ± 0.019	0.562 ± 0.016	145.73	129.48	127.78
CaM-nNOSp	0.357 ± 0.027	0.503 ± 0.024	0.561 ± 0.010	150.70	131.71	122.71
HEWL	0.233 ± 0.029	0.399 ± 0.026	0.699 ± 0.003	80.51	75.96	67.66
Ubiquitin	0.286 ± 0.022	0.425 ± 0.026	0.629 ± 0.011	64.56	52.75	48.58

[a] Abbreviations: ALBP, adipocyte lipid binding protein; Cyt c_2, cytochrome c_2; CaM:MLCK, calcium-saturated calmodulin in complex with a peptide corresponding to the myosin light chain kinase calmodulin-binding domain; CaM:NOS, calcium-saturated calmodulin in complex with a peptide corresponding to the neuronal nitric oxide synthase calmodulin-binding domain; HEWL, hen egg white lysozyme

[b] Methyl rotamer entropy of protein normalized by the total number of unique connected methyl side-chain χ angles

[c] Total rotamer entropy of protein normalized by the total number of unique side-chain χ angles

[d] Total rotamer entropy of protein accounting for only intra-residue correlations

[e] Total rotamer entropy of protein accounting for intra-residue and 2nd order inter-residue correlations

[f] Total rotamer entropy of protein accounting for intra-residue and up to 3rd order inter-residue correlations