J|A|C|S
COMMUNICATIONS

Published on Web 06/02/2004

The Origin of Protein Sidechain Order Parameter Distributions

Robert B. Best,[†] Jane Clarke,[†] and Martin Karplus*[,‡,§]

MRC Centre for Protein Engineering, Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, U.K., Department of Chemistry and Biological Chemistry, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, and Laboratoire de Chimie Biophysique, ISIS/ULP, 8 allée Gaspard Monge, BP 70028, F-67083 Strasbourg Cedex, France

Received February 18, 2004; E-mail: marci@tammy.harvard.edu

The dynamics of proteins are known to play an important part in their function.[1] Nuclear magnetic resonance experiments are one of the most valuable sources of information concerning dynamics. In addition to standard experiments for probing backbone dynamics, methods have been developed recently for studying the dynamics of side chain methyl groups using deuterium spin relaxation,[2] and these methods have now been applied to a number of proteins. Order parameters (S^2) derived from these experiments characterize the amplitude of angular motion of each methyl group axis on a scale from 1 (perfectly rigid) to 0 (dynamic).[3] In a study of the temperature dependence of calmodulin dynamics, Lee and Wand observed three distinct peaks in the distribution of methyl order parameters.[4,5] This result was taken as evidence for the existence of three "classes of motion" and related to a hierarchical energy landscape model; it was also used to suggest an alternative origin for the protein "glass transition".[4] The importance of these suggestions for understanding the functional motions in proteins has led us to reinvestigate the existence of the proposed classes. A larger set of proteins shows that there is a broad distribution of order parameters without a clear segregation into classes. Moreover, molecular dynamics simulations provide a physical explanation of the heterogeneity in the observed distribution of the side chain dynamics; the results indicate that the observations do not require a hierarchical landscape model.

The combined distribution of the side chain order parameters for all residue types in 18 proteins for which NMR or crystal structures are available (see legend to Figure 1) is very broad and not unimodal (Figure 1a), although three distinct classes are not evident. The pooling of data from different kinds of methyl groups would be an obvious explanation for this complexity (since, for example, leucine would tend to have lower order parameters than alanine). However, it turns out that the separate distributions for most types of side chains are multimodal as well. For example, the distributions of Val, Thr, and Ile γ methyl order parameters have at least two peaks, while those of Ile and Leu δ methyl groups apparently have three peaks. Ala, by virtue of being essentially part of the backbone has a single peak at high S^2, while Met has a single peak at low-order parameter values; this is probably due to its three rotatable side chain bonds and its tendency to be closer to the protein surface than the other types of methyl-bearing side chains. We note here that although the experimental data were collected at slightly different temperatures, the change in order parameters over this range of temperatures is less than 0.05;[5] restriction to data collected at 303 K does not affect the results.

To probe further the origin of these distributions, molecular dynamics simulations[1] of each of the above 18 proteins were run with CHARMM (Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187−217) for 5 ns at 300 K with a widely used implicit solvent model.[6] This simulation length is sufficient to sample the motion on a picosecond time scale described by the order parameters, and the results do not change substantially on extending the simulations to 30 ns for selected proteins.

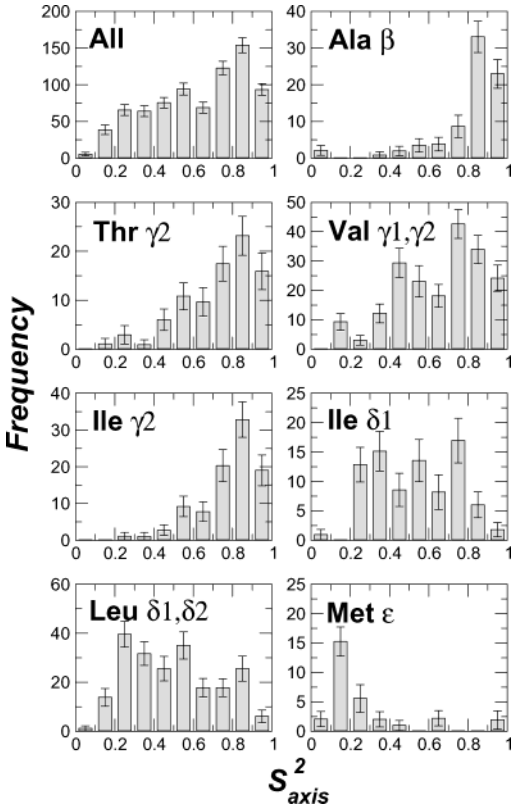

Figure 1. Distributions of experimental methyl axis order parameters for 18 proteins with published side chain dynamics data. Distributions are shown for each type of methyl group and for the pooled distribution of all methyl groups. The proteins used were the third fnIII domain from tenascin (1TEN),[8] the tenth fnIII domain from fibronectin (1FNF),[8] adipocyte lipid binding protein (1LIB),[9] muscle fatty acid binding protein (pdb: 1HMT),[9] ubiquitin (1UBQ),[10] phospholipase Cγ1 SH2 domain (2PLD),[2] holo-calmodulin (3CLN),[11] SAP SH2 (1D1Z),[12] oxidized flavodoxin (1FLV),[13] HIV protease (1HWR),[14] cytochrome *c*2 (1C2R),[15] A3D (2A3D),[16] Cdc42Hs (1AN0),[17] Fyn SH3 (1SHF),[18] mouse urinary protein (1JV4),[19] Syp SH2 (1AYA),[20] eglin c (1EGL),[21] and troponin C (pdb: 5TNC).[22] Error bars were obtained by constructing 100 data sets by a bootstrap Monte Carlo procedure and calculating the variance within each histogram bin.[23]

A satisfactory correlation between order parameters calculated from the simulations[7] and the experimental data is obtained for each protein (linear correlation coefficients 0.4−0.7); this level of agreement is similar to that obtained in long simulations in explicit solvent.[7,24] The distributions obtained for the methyl groups of Val

[†] MRC Centre for Protein Engineering.
[‡] Harvard University.
[§] Laboratoire de Chimie Biophysique.

10.1021/ja049078w CCC: $27.50

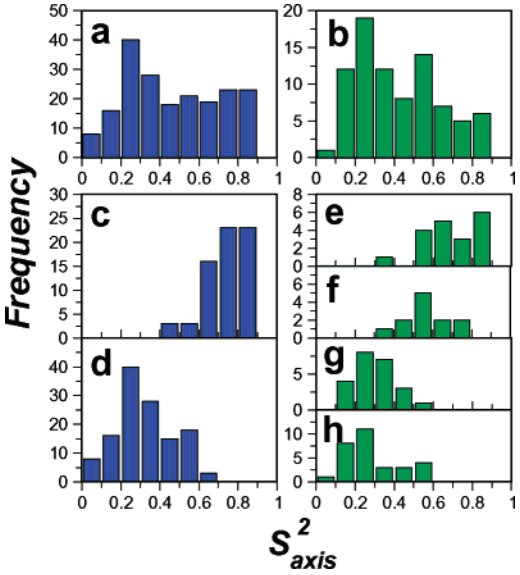

Figure 2. Breakdown of order parameter distributions according to side chain rotamer occupancy. The frequency corresponds to the number of cases with a given value, and Val γ and Ile δ are colored blue and green, respectively. Top panels show the distribution of simulated order parameters for (a) Val γ and (b) Ile δ. These distributions can be decomposed into contributions from a class of methyls making rotameric transitions and a class that does not. The distribution for Val is broken down into the contribution from residues whose χ_1 dihedrals spend greater than 90% of the time in a single rotamer (c) and the remainder (d). The Ile distribution is divided according to side chains in which both (e), only χ_1 (f), only χ_2 (g), or neither χ_1 and χ_2 (h) has greater than 90% preference for one rotamer.

and Ile are shown in Figure 2a,b, respectively; these are shifted to slightly lower values than experiment, most likely due to the implicit solvent model, which permits more flexibility. Very similar results were obtained from simulations of two of these proteins (TNfn3 and FNfn10) in explicit solvent (Supporting Information), without the small shift to lower-order parameter values. This approach was not applied to all proteins because of the computational cost.

Analysis of the simulations identifies the dynamic features of the side chain motions which give rise to the discrete nature of the order parameter distributions. Some of the side chain dihedral angles make rotameric transitions and others do not (because of the confining potential arising from the surrounding protein). Figure 2c,d shows that the calculated bimodal order parameter distribution of Val is the sum of two unimodal distributions: a narrow distribution with high-order parameter values for side chains which fluctuate within a single rotamer well and a broader distribution with lower-order parameters for those making χ_1 rotameric transitions. Slight differences in the environment of each methyl group and contributions from backbone motion account for the variation within these categories. Likewise, the order parameters of Ile δ-methyl groups (Figure 2e−h) can be classified according to whether neither, one, or both of the χ_1 and χ_2 dihedral angles make transitions. The order parameter distributions of the other side chain methyl groups have a corresponding origin, with the γ-methyls of Ile and Thr being analogous to Val, and the δ-methyls of Leu being analogous to those of Ile (Supporting Information).

The present analysis is supported by earlier experimental and simulation studies showing the importance of jumps between rotamers for the interpretation of vicinal coupling constants[25] and of dipolar couplings.[26] The influence of rotameric transitions on order parameters has been recognized but not related to their distributions.[5,26] Also, molecular dynamics simulations have shown that the so-called protein "glass transition" involves the freezing out of the transitions of protein side chains between rotamers,[27] due in part to the solvent.[28] The "glass transition" inferred from the disappearance of the smaller values from the order parameter distribution when the temperature is decreased[4] is therefore in accord with the existing theory and does not require a novel interpretation. In summary, by a combination of experiment and simulations we have obtained a simple explanation for the observed protein side chain order parameter distributions, which does not invoke a hierarchical energy landscape. The conclusions described here will aid in the interpretation of future studies of the internal motions of proteins by NMR and other techniques.

Acknowledgment. R.B.B. was supported by the Cambridge Commonwealth Trust, and J.C. was supported by a Wellcome Trust Senior Research Fellowship. The work at Harvard was supported in part by a grant from the National Institutes of Health. We thank A.J. Wand for helpful comments on the manuscript.

Supporting Information Available: Distributions of order parameters for Ile γ2, Thr γ2, and Ile δ1 from simulations, separated according to side chain rotamer averaging and distributions obtained from explicit solvent simulations. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646−652.
(2) Muhandiram, D. R.; Yamazaki, T.; Sykes, B. D.; Kay, L. E. *J. Am. Chem. Soc.* **1995**, *117*, 11536−11544.
(3) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(4) Lee, A. L.; Wand, A. J. *Nature* **2001**, *411*, 501−504.
(5) Lee, A. L.; Sharp, K. A.; Kranz, J. K.; Song, X.-J.; Wand, A. J. *Biochemistry* **2002**, *41*, 13814−13825.
(6) Lazaridis, T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 133−152.
(7) Chatfield, D. C.; Szabo, A.; Brooks, B. R. *J. Am. Chem. Soc.* **1998**, *120*, 5301−5311.
(8) Best, R. B.; Rutherford, T. J.; Freund, S. M. V.; Clarke, J. *Biochemistry* **2004**, *43*, 1145−1155.
(9) Constantine, K. L.; Friedrichs, M. S.; Wittekind, M.; Jamil, H.; Chu, C.-H.; Parker, R. A.; Goldfarb, V.; Mueller, L.; Farmer, B. T. *Biochemistry* **1998**, *37*, 7965−7980.
(10) Lee, A. L.; Flynn, P. F.; Wand, A. J. *J. Am. Chem. Soc.* **1999**, *121*, 2891−2902.
(11) Lee, A. L.; Kinnear, S. A.; Wand, A. J. *Nat. Struct. Biol.* **2000**, *7*, 72−77.
(12) Finerty, P. J.; Muhandiram, R.; Forman-Kay, J. D. *J. Mol. Biol.* **2002**, *322*, 605−620.
(13) Liu, W.; Flynn, P. F.; Fuentes, E. J.; Krantz, J. K.; McCormick, M.; Wand, A. J. *Biochemistry* **2001**, *40*, 14744−14753.
(14) Ishima, R.; Petkova, A. P.; Louis, J. M.; Torchia, D. A. *J. Am. Chem. Soc.* **2001**, *123*, 6164−6171.
(15) Flynn, P. F.; Urbauer, R. J. B.; Zhang, H.; Lee, A. L.; Wand, A. J. *Biochemistry* **2001**, *40*, 6559−6569.
(16) Walsh, S. T. R.; Lee, A. L.; Degrado, W. F.; Wand, A. J. *Biochemistry* **2001**, *40*, 9560−9569.
(17) Loh, A. P.; Pawley, N.; Nicholson, L. K.; Oswald, R. E. *Biochemistry* **2001**, *40*, 4590−4600.
(18) Mittermaier, A.; Davidson, A. R.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 9004−9005.
(19) Chaykovski, M. M.; Bae, L. C.; Cheng, M.-C.; Murray, J. H.; Tortolani, K. E.; Zhang, R.; Seshadri, K.; Findlay, J. H. B. C.; Hsieh, S.-Y.; Kalverda, A. P.; Homans, S. W.; Brown, J. M. *J. Am. Chem. Soc.* **2003**, *125*, 15767−15771.
(20) Kay, L. E.; Muhandiram, D. R.; Wolf, G.; Shoelson, S. E.; Forman-Kay, J. D. *Nat. Struct. Biol.* **1998**, *5*, 156−163.
(21) Hu, H.; Clarkson, M. W.; Hermans, J.; Lee, A. L. *Biochemistry* **2003**, *42*, 13856−13868.
(22) Gagné, S. M.; Tsuda, S.; Spyracopoulos, L.; Kay, L. E.; Sykes, B. D. *J. Mol. Biol.* **1998**, *278*, 667−686.
(23) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C*, 2nd ed.; Cambridge University Press: Cambridge, 1992.
(24) Prabhu, N. V.; Lee, A. L.; Wand, A. J.; Sharp, K. A. *Biochemistry* **2003**, *42*, 562−570.
(25) Hoch, J. C.; Dobson, C. M.; Karplus, M. *Biochemistry* **1985**, *24*, 3831−3841.
(26) Chou, J. J.; Case, D. A.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 8959−8966.
(27) Steinbach, P. J.; Brooks, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9135−9139.
(28) Vitkup, D.; Ringe, D.; Petsko, G. A.; Karplus, M. *Nat. Struct. Biol.* **2000**, *7*, 34−38.

JA049078W