J|A|C|S
COMMUNICATIONS

Published on Web 02/27/2002

χ_1 Rotamer Populations and Angles of Mobile Surface Side Chains Are Accurately Predicted by a Torsion Angle Database Potential of Mean Force

G. Marius Clore*,† and John Kuszewski§

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0510, and Computational Bioscience and Engineering Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892-5624

Received December 8, 2001

Conformational plasticity of surface side chains, manifested by rotamer averaging that is fast on the chemical shift time scale, can play an important role in protein−protein recognition, permitting optimal intermolecular contacts to be formed.[1] This is critical in cases where a single protein employs the same interaction surface to recognize a variety of different proteins.[1] A recent study has shown that equilibrium χ_1 side chain rotamer populations and angles can be probed in solution by the measurement of $^{13}C\beta-^{1}H\beta$ residual dipolar couplings (RDC) in multiple alignment media.[2] Using the small 63-residue B1 domain of protein L, it was shown that while the data for all internal and some surface side chains could be accounted for by a single χ_1 rotamer with angles very close to those observed in the X-ray structure,[3] the data for 11 surface side chains could only be accounted for by multiple χ_1 rotamers.[2] The crystal structure has three independent molecules in the asymmetric unit, and some of these mobile side chains are also found to have different χ_1 rotamers in the three molecules.[3] ϕ/ψ backbone and χ_1 side chain torsion angles tend to be correlated,[4,5] mainly due to steric interactions between side chain and backbone atoms. For example, the χ_1 rotamer for valine is almost invariably trans in an α-helix. In this contribution, we show that the equilibrium χ_1 angle and rotamer populations observed for the mobile surface side chains of protein L in solution can be well predicted from the backbone coordinates of the crystal structure by means of a torsion angle database potential of mean force (TADBP)[5] combined with the application of conjoined rigid body/torsion angle dynamics.[6]

The TADBP has been described previously[5] and consists of a set of multidimensional potential surfaces derived from high-resolution crystal structures describing various torsion angle correlations in two (ϕ/ψ, χ_1/χ_2, χ_2/χ_3, χ_3/χ_4), three ($\phi/\psi/\chi_1$, $\chi_1/\chi_2/\chi_3$, $\chi_2/\chi_3/\chi_4$), and four ($\phi/\psi/\chi_1/\chi_2$, $\chi_1/\chi_2/\chi_3/\chi_4$) dimensions. The aim of the TADBP in the context of an NMR structure determination is to bias sampling during simulated annealing refinement to conformations that are likely to be energetically possible, as defined by those conformations that are known to be physically realizable from high-resolution crystal structures.[5] The TADBP used in the present work differs from that described previously[5b] in two regards: (1) the potential is derived from a larger database of crystal structures, specifically 518 structures (only residues with thermal B-factors <25 Å^2 are considered) solved at a resolution ≤ 1.8 Å and an R-factor of $<25\%$, with $<30\%$ identity among any pairs of sequences;[7] (2) the raw potentials are fitted by a sum of multidimensional quartic functions rather than a sum of multidimensional Gaussians.[5b] While the overall shape of the quartic function is very similar to that of an equivalent Gaussian, the quartic function has the advantage of not requiring the evaluation of an exponential function, which is computationally expensive.

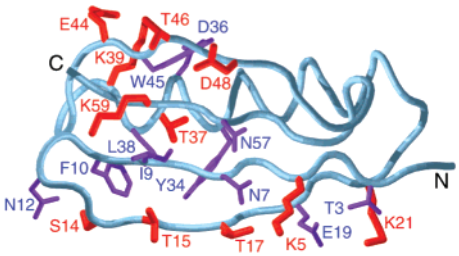

Figure 1. Ribbon diagram[10] of protein L[3] with the 11 mobile surface side chains in red and spatially adjacent surface side chains in blue.

A ribbon diagram of protein L is shown in Figure 1 illustrating the position of the 11 mobile surface side chains, as well as that of spatially adjacent surface side chains. Because there are some significant differences in the ϕ/ψ backbone torsion angles associated with the mobile surface side chains among the three molecules in the crystal structure, calculations were performed on all three molecules. Simulated annealing, with the program XPLOR-NIH,[8] was carried out as follows: the backbone, internal side chains, and surface side chains, other than the 11 mobile ones or those adjacent to them, were held rigid (i.e., treated as a rigid body). The surface side chains adjacent to the mobile ones were allowed torsion angle degrees of freedom limited to within ±20° of those observed in the crystal structure by torsion angle square-well potentials.[9] The 11 mobile surface side chains were given full torsional degrees of freedom. The target function for simulated annealing in torsion angle space[6] consists of only three terms: a quartic van der Waals repulsion term,[9] a square-well torsion angle restraint term[9] with a force constant of 200 kcal·mol^{-1}·rad^{-2} applied only to the surface side chains adjacent to the 11 mobile ones throughout the calculations, and the TADBP. The rationale for permitting some movement for the surface side chains adjacent to the mobile ones is to ensure that conformational space for the latter is fully and effectively sampled and not significantly affected by steric interactions with neighboring side chains. No electrostatic or H-bonding terms are present in the target function. The simulated annealing protocol is as follows: 8 ps of dynamics at 3000 K with the van der Waals repulsion potential and TADBP turned off to randomize the torsion angles of the mobile surface side chains; 8 ps of dynamics at 3000 K with the van der Waals repulsion term still turned off but the force constant, k_{db}, for the TADBP set to 1; 48 ps (240 cycles of 0.2 ps each) during which time the system is slowly cooled from 3000 to 12.5 K, the force constant and radius scale factor for the van der Waals repulsive term are increased from 0.004 to 4 kcal·mol^{-1}·rad^{-2} and decreased from 0.9 to 0.8,

† National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.
§ Computational Bioscience and Engineering Laboratory, Center for Information Technology, National Institutes of Health.

10.1021/ja017712p CCC: $22.00

Table 1. Comparison of χ_1 Rotamer Populations and Angles for Mobile Surface Side Chains Determined by Simulated Annealing Using the Torsion Angle Database Potential of Mean Force and Experimentally from Dipolar Couplings[2], Together with the χ_1 Values Observed in the X-ray Structure[3]

	simulated annealing and torsion angle database potential			$^{13}C\beta$–$^{1}H\beta$ dipolar couplings and Monte Carlo fitting			X-ray χ_1		
residue	% $g^+(\chi_1)$	% $t(\chi_1)$	% $g^-(\chi_1)$	% $g^+(\chi_1)$	% $t(\chi_1)$	% $g^-(\chi_1)$	mol1	mol2	mol3
Lys5	0	49 (−178 ± 6°)	51 (−59 ± 7°)	11 ± 6 (71 ± 13°)	49 ± 4 (−178 ± 5°)	40 ± 4 (−66 ± 8°)	179°	178°	179°
Ser14	32 (68 ± 17°)	36 (167 ± 14°)	32 (−74 ± 14°)	60 ± 4 (65 ± 6°)	20 ± 7 (165 ± 17°	20 ± 4 (−52 ± 13°)	56°	53°	−56°
Thr15	39 (73 ± 10°)	28 (174 ± 5°)	33 (−58 ± 10°)	21 ± 11 (54 ± 6°)	66 ± 18 (−166 ± 6°	13 ± 9 (−53 ± 10°)	56°	53°	−56°
Thr17	36 (64 ± 4°)	24 (174 ± 4°)	40 (−64 ± 8°)	33 ± 8 (59 ± 6°)	11 ± 8 (−169 ± 6°)	56 ± 1 (−56 ± 7°)	−64°	66°	−62°
Lys21	2 (66 ± 6°)	39 (−176 ± 7°)	59 (−65 ± 5°)	14 ± 1 (71 ± 11°)	19 ± 1 (−176 ± 12°	67 ± 1 (−63 ± 4°)	−79°	−176°	−63°
Thr37	67 (59 ± 1°)	0	33 (−55 ± 5°)	60 ± 18 (74 ± 5°)	11 ± 7 (169 ± 6°)	30 ± 24 (−67 ± 7°)	70°	71°	72°
Lys39	0.3 (77 ± 0.3°)	51 (177 ± 6°)	47.7 (−72 ± 2°)	24 ± 3 (75 ± 10°)	31 ± 3 (178 ± 1°)	45 ± 2 (−61 ± 6°)	−67°	−63°	−66°
Glu44	9 (63 ± 6°)	45 (−177 ± 2°)	46 (−67 ± 7°)	58 ± 4 (77 ± 3°)	34 ± 4 (−171 ± 7°	7 ± 5 (−74 ± 22°)	−53°	−61°	−66°
Thr46	50 (67 ± 5°)	29 (176 ± 5°)	21 (−57 ± 2°)	74 ± 8 (56 ± 5°)	12 ± 3 (−168 ± 7°)	14 ± 10 (−60 ± 9°)	56°	60°	−170°
Asp48	0	72 (−178 ± 9°)	28 (−77 ± 5°)	1 ± 1 (64 ± 10°)	47 ± 1 (−167 ± 7°)	52 ± 1 (−69 ± 5°)	−72°	−79°	179°
Lys59	0.1 (62°)	51.6 (−178 ± 1°)	48.3 (−64 ± 4°)	5 ± 2 (69 ± 12°)	46 ± 1 (−163 ± 6°)	49 ± 2 (−52 ± 4°)	171°	−57°	−58°

respectively, and k_{db} is increased from 1 to 2; and finally 500 steps of minimization with the final values of the force constants. Starting from each molecule in the crystal structure, 300 structures were calculated, and the results reported are obtained from the distribution of torsion angles observed in all 900 simulated annealing structures.

The results are summarized in Table 1 and compared to both those determined by Monte Carlo analysis of the measured RDCs in solution[2] and those observed in the three molecules of the crystal structure.[3] Two different rotamers are observed for 7 out of the 11 mobile residues in the crystal structure. There is, however, no example in the crystal structure of a residue where all three rotamers are observed. The agreement in the equilibrium values of the χ_1 angles is excellent. The average difference in the equilibrium χ_1 angle values obtained by simulated annealing and from the RDCs is only 9 ± 6°.

In general, there is also good agreement between the χ_1 rotamer populations predicted by simulation and those obtained from the RDC data. This is particularly satisfying since the derivation of χ_1 rotamer populations and angles from RDCs is not a straightforward and direct procedure, and hence entails a number of uncertainties which, in all likelihood, are much larger than the standard deviations obtained from the Monte Carlo analysis. Thus, Lys5, Asp48, and Lys59 are almost exclusively a mixture of t and g^- χ_1 rotamers. The χ_1 angles of Lys21 and Lys39 are also predicted to be almost exclusively t and g^- by the TADBP, although the presence of some some minor g^+ χ_1 rotamer, is deduced from the RDCs. The theoretical and experimental distributions of χ_1 rotamers observed for Thr17 ($g^- > g^+ > t$), Thr37 ($g^+ > g^- \gg t$) and Thr46 ($g^+ \gg t \approx g$) are essentially the same.

In our calculations all three χ_1 rotamers of Ser14 are roughly equally populated. The RDC data, as well as the crystal structure, however, indicate that the major χ_1 rotamer for Ser14 is g^+. This is probably due to the formation of a transient H-bond between the side chains of Asn12 and Ser14 in the g^+ χ_1 conformation (cf. Figure 1), an effect not modeled in our calculations. All three χ_1 rotamers of Thr15 are roughly equally populated in the simulations, with the g^+ rotamer in slight excess, whereas the t rotamer is the major species derived from the RDCs. Since the χ_1 angles of Thr15 are only in the g^+ and g^- rotamers in the crystal structure, the significance of the disparity for Thr15 between our calculations and the RDC-derived χ_1 rotamers is uncertain.

There is one clear anomaly: Glu44. In the simulations the populations of t and g^- χ_1 rotamers for Glu44 are approximately equal, with only a small fraction (∼10%) in the g^+ rotamer; in the case of the χ_1 distribution derived from the RDCs, the g^+ and g^- rotamers are interchanged such that g^+ is the major species (∼60%), followed by t (∼35%) with only a small proportion (<10%) of g^-. The results from the RDCs are a little surprising in this regard and may be a consequence of the small number of Monte Carlo iterations with statistical significance.[2] First, a survey of the crystallographic protein database[4a,5] indicates that the g^+ χ_1 conformation for a glutamate is rarely populated for the combination of ϕ/ψ angles observed in the crystal structure ($\phi \approx -70°$, $\psi \approx 153-157°$); second, Glu44 is in the g^- χ_1 rotamer for all three molecules in the crystal structure;[3] third, if Glu44 would be anomalous in solution, one would predict the existence of some potential nonbonded interaction (electrostatic or H-bonding) that would stabilize the g^+ χ_1 conformation, but none is apparent from the crystal structure.

In conclusion, we have shown that equilibirum side chain χ_1 angles and rotamer distributions can be predicted from static crystal structures using a torsion angle database potential of mean force in conjunction with rigid body/torsion angle dynamics. Thus, one can conclude that for a mobile surface side chain exhibiting rotamer averaging, the equilibrium distribution of rotamers will largely be determined by its backbone ϕ/ψ angles. That distribution may be modulated to a small extent by other factors, for example, the formation of transient hydrogen bonds with neighboring side chains.

Supporting Information Available: Description of the quartic form of the torsion angle database potential of mean force (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wang, G.; Louis, J. M.; Sondej, M.; Seok, Y.-J.; Peterkofsky, A.; Clore, G. M. *EMBO J.* **2000**, *19*, 5635−5649.
(2) Mittermaier, A.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 6892−6903.
(3) O'Neill, J. W.; Kim, D. E.; Baker, D.; Zhang, K. Y. J. *Acta Crystallogr.* **2001**, *D57*, 480−487.
(4) (a) Dunbrack, R. L.; Karplus, M. *J. Mol. Biol.* **1993**, *230*, 543−574. (b) Vriend, G.; Sander, C. *J. Appl. Crystallogr.* **1993**, *26*, 47−60. (c) Xiang, Z.; Honig, B. *J. Mol. Biol.* **2001**, *311*, 421−430.
(5) (a) Kuszewski, J.; Gronenborn, A. M.; Clore, G. M. *Protein Sci.* **1996**, *5*, 1067−1080. (b) Kuszewski, J.; Clore, G. M. *J. Magn. Reson.* **2000**, *146*, 249−254.
(6) Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2000**, *152*, 288−302.
(7) Hobohm, U.; Scharf, M.; Schneider, R.; Sander, C. *Protein Sci.* **1992**, *1*, 409−417.
(8) Clore, G. M.; Kuszewski, J.; Schwieters, C. D.; Tjandra, N. *XPLOR-NIH* version 1.1.2, **2001**, http://nmr.cit.nih.gov/xplor-nih.
(9) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett* **1988**, *229*, 317−324.
(10) Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2001**, *149*, 239−244.

JA017712P