Environmentally Dependent Conformational Preferences of Peptides

Paul E. Smith, Herb D. Blatt, and B. Montgomery Pettitt*

Department of Chemistry, University of Houston
4800 Calhoun Road, Houston, Texas 77204-5641

Received October 28, 1996
Revised Manuscript Received January 24, 1997

The conformations adopted by peptides and proteins are known to be dependent on their environment.[1−5] The question then arises as to which conformations are more or less stable in different environments. For instance, protein folding involves a highly correlated transition from an unfolded state with a high degree of solvent exposure to a folded state of relatively low solvent exposure where many amino acid residues are surrounded by protein. Hence, understanding the effects of changes in environment on the conformation of amino acid chains could enable the rationalization of certain underlying principles governing the protein folding mechanism.[6,7]

In this paper, we present molecular dynamics simulations and experimental results which compare the different conformational preferences of amino acid chains on changing their environments. We will explicitly consider three different environments, namely, the gas phase, aqueous solution, and the protein environment. To achieve this we have assumed that the backbone conformation (determined by ϕ and ψ) is, to a first approximation, essentially independent of the nature of the amino acid side chain (with the exception of glycine and proline) and independent of the nature of the rest of the chain.[8]

In this spirit, our gas phase model is a four alanine residue peptide whose conformational preferences have been determined by 100 ns stochastic dynamics simulations using continuum solvents with relative permittivities of 1 and 80.[9] Explicit aqueous solvent simulations of the same tetrapeptide have also been obtained over 10 ns for four different permutations of the terminal group charge (NH_3^+ and NH_2 for the N terminus, CO_2^- and CO_2H for the C terminus).[8] All simulations were performed with the latest CHARMM all atom force field.[10] Although the results are force field dependent, it is assumed that the underlying potential surface in the CHARMM model and the limited sampling during 10 ns are sufficiently accurate for this coarse comparison with experiment. The ϕ and ψ distributions obtained correspond to averages over the central residues (residues 2 and 3). In addition, the aqueous environment results have been averaged over the different charge end group combinations. Probability distributions corresponding to a protein environment are traditionally obtained from X-ray crystallography.[11−14] Hence, we have obtained the relevant ϕ and ψ dihedral angle distributions directly from 63 protein crystal structures as deposited in the Brookhaven protein database (a full list of the proteins and further details are given in ref 15).

Scheme 1. Thermodynamic Cycle Involving Changes in Conformation (ξ) and Environment (g = gas, aq = aqueous, and prot = protein)

$$\begin{array}{ccc} W_g(\xi_1) & \longrightarrow & W_g(\xi_2) \\ \downarrow & & \downarrow \\ W_{aq}(\xi_1) & \longrightarrow & W_{aq}(\xi_2) \\ \downarrow & & \downarrow \\ W_{prot}(\xi_1) & \longrightarrow & W_{prot}(\xi_2) \end{array}$$

The relationship between the various population distributions is illustrated in Scheme 1, where the probability distributions are reflected in terms of torsional potentials of mean force (PMF) obtained from the usual Boltzmann relationship.[16] Moving horizontally, one obtains the free energy for rotation $\Delta W_{1\rightarrow 2}$ between two different conformations (ξ_1 and ξ_2; $\xi_i = \phi$ or ψ), while vertical changes correspond to the free energy of solvation $\Delta W_{g\rightarrow aq}$ of a fixed conformation of the backbone on changing the environment. In general, for conformationally flexible molecules, $\Delta W_{solv} = W_{aq}(\xi_2) - W_g(\xi_1)$ corresponds to the free energy of solvation per residue during a conformational change, while $\Delta W_{fold} = W_{prot}(\xi_2) - W_{aq}(\xi_1)$ corresponds to the free energy of protein folding per residue in solution. Here, we are primarily concerned with differences in the PMF's given by $\Delta\Delta W = \Delta W_{1\rightarrow 2,aq} - \Delta W_{1\rightarrow 2,g} = \Delta W_{1,g\rightarrow aq} - \Delta W_{2,g\rightarrow aq}$, i.e., which represent the relative changes in solvation between two different conformations of the peptide backbone.

The probability distributions obtained for the three different environments have been converted to rotational PMFs and are displayed in Figure 1. A quantitative picture of the average effect that the different environments have on the backbone conformations adopted by amino acid chains is apparent. For the ϕ degree of freedom, all methods give minima around −90° and 60°, except for a small shift to −70° observed for the protein environment. Relatively high free energy conformations near −10° and 120° observed in the gas phase are also observed in explicit aqueous and continuum aqueous environments. However, proteins appear to be more tolerant of these conformations. Populations of the two different major conformations (−90°/60°) at 300 K are 97/3, 99/1, 100/0, and 92/8 for the $\epsilon = 1$, $\epsilon = 80$, and explicit aqueous and protein environments, respectively.

The ψ dihedral potentials of mean force display a marked shift from the gas phase potential. In particular, the minimum at 70° disappears on moving to an aqueous or protein environment. The PMFs obtained for the aqueous and protein environments display very similar features with both giving two minima around −60° and 160°, the major differences occurring in the magnitude of the barriers dividing these minima. Populations for the two minima (−60°/160°) at 300 K are 56/44, 50/50, and 55/45 for the $\epsilon = 80$ and explicit aqueous and protein

(1) Brooks, C. L., III; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(2) Sönnichsen, F. D.; van Eyk, J. E.; Hodges, R. S.; Sykes, B. D. *Biochemistry* **1992**, *31*, 8790−8798.
(3) Tayar, N. E.; Mark, A. E.; Vallat, P.; Brunne, R. M.; Testa, B.; van Gunsteren, W. F. *J. Med. Chem.* **1993**, *36*, 3757−3764.
(4) Marlow, G. E.; Perkyns, J. S.; Pettitt, B. M. *Chem. Rev.* **1993**, *93*, 2503−2521.
(5) Smith, P. E.; Pettitt, B. M. *J. Phys. Chem.* **1994**, *98*, 9700−9711.
(6) Timasheff, S. N.; Arakawa, T. Stabilization of Protein Structure by Solvents. In *Protein Structure: A Practical Approach*; Creighton, T. E., Ed.; IRL Press: London, 1989; pp 331−345.
(7) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(8) Blatt, H. D.; Smith, P. E.; Pettitt, B. M. Manuscript in preparation.
(9) Terminal NH_3 and CO_2 groups were neutralized by setting their charges to 0. Stochastic dynamics simulations were performed at 300 K including all interactions and using a 2 fs time step. Friction coefficients were assigned as 2 ps^{-1} for all heavy atoms (carbon, nitrogen, and oxygen) and 0 ps^{-1} for all hydrogens. Configurations were saved every 1 ps for analysis.
(10) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Wiorkiewicz-Kuczera, J.; Karplus, M. *FASEB J.* **1992**, *6*, A143.
(11) Levitt, M.; Warshel, A. *Nature* **1975**, *253*, 694−698.
(12) Kolinski, A.; Skolnick, J. *J. Chem. Phys.* **1992**, *97*, 9412−9426.
(13) Avbelj, F. *Biochemistry* **1992**, *31*, 6290−6297.
(14) Kikuchi, T. *J. Comput. Chem.* **1996**, *17*, 226−237.
(15) Smith, P. E.; Blatt, H. D.; Pettitt, B. M. *Proteins: Struct., Funct., Genet.* **1997**, *27*, 227−234.
(16) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. Proteins. A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons: New York, 1988; Vol. 71.

S0002-7863(96)03752-3 CCC: $14.00

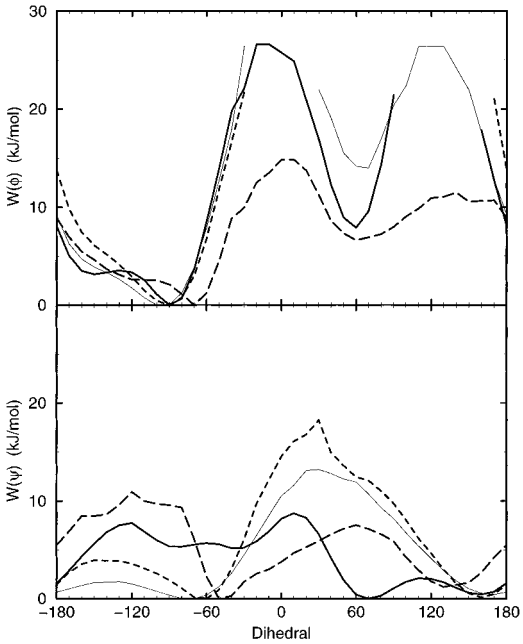

Figure 1. Free energy profiles for the ϕ (top) and ψ (bottom) degrees of freedom as a function of environment: vacuum $\epsilon = 1$ (thick solid), vacuum $\epsilon = 80$ (thin solid), aqueous (thick dotted), and protein (thick dashed).

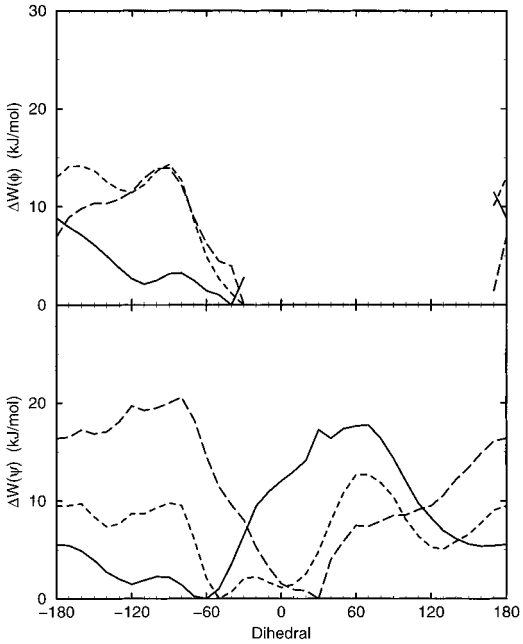

Figure 2. Solvation free energy differences for the ϕ (top) and ψ (bottom) degrees of freedom as a function of change in environment: vacuum to aqueous (thick solid), vacuum to protein (thick dotted), and aqueous to protein (thick dashed).

environments, respectively. Interestingly, the very simple $\epsilon = 80$ continuum model produces results which are in near quantitative agreement with the explicit solvent simulations; the only significant difference being that the continuum model gives somewhat lower barriers to rotation. This is also supported by other theoretical results for small peptides which have suggested that the major contribution to the change in PMF on going from gas phase to aqueous solution is a simple dielectric effect.[17] Hence, the coarse continuum approximation appears to be a satisfactory representation of solvent effects for these coarse comparisons of nonpolar systems while providing a vast increase in computational efficiency.

For our purposes it is more informative to look at differences in effects between the various environments. The data presented in Figure 2 illustrate the changes in the torsional rotational potential energy surfaces on moving from one environment to another. Thus, the influence of the intramolecular potential has been removed in comparison with that in Figure 1. Data for ϕ are incomplete, due to insufficient sampling in the explicit solvent simulations (rate constants for ϕ transitions are $1-2$ orders of magnitude slower than ψ transitions),[8,18] but suggest that a change in environment from the gas phase to aqueous or protein, or a change from aqueous to protein, all tend to favor folded conformations around 0°. This type of environmental effect has also been observed in an integral equation study of a small zwitterionic peptide in salt solution.[19] The change in solvation of the average ψ rotational free energy surface on going from aqueous solution to a protein environment again favors folded conformations in the region of 0°.[20] This suggests that desolvation of random coil like backbone conformations, and their subsequent interaction with a protein matrix, is such that folded configurations are favored by up to 20kJ/mol per ψ degree of freedom on the average. However, by comparing small single-residue results with those for polypeptide chains, we neglected to include explicitly in Scheme 1 the configurational entropy ($-T\Delta S$) lost on folding.[7] This entropy cost is on the order of 5−42 kJ/mol per residue or ϕ/ψ pair,[21,22] with total free energies of folding in ranging from −20 to −40kJ/mol for a medium size protein (100 residues). Therefore, environmental stabilization free energies in the region of 20 kJ/mol per ψ degree of freedom are consistent with these numbers within the previously stated approximate framework. This semiquantitative picture of changes in peptide backbone solvation could prove valuable in attempts to model and understand protein folding pathways.[23,24]

Acknowledgment. This publication was supported in part by the Keck Center and grant number BIR-92-56580 from the National Science Foundation. We also thank the Robert A. Welch Foundation, the National Science Foundation, and the National Institutes of Health for partial support. The Metacenter is acknowledged for computational support.

JA963752V

(17) Lau, W. F.; Pettitt, B. M. *Biopolymers* **1987**, *26*, 1817−1831.
(18) Smith, P. E.; Pettitt, B. M.; Karplus, M. *J. Phys. Chem.* **1993**, *97*, 6907−6913.
(19) Perkyns, J. S.; Pettitt, B. M. *J. Phys. Chem.* **1995**, *99*, 1−2.
(20) Note that in such a double difference the underlying potential surface is removed and only environmental influences remain.
(21) Privalov, P. L. *Adv. Protein Chem.* **1979**, *33*, 167−241.
(22) Creighton, T. E. *Proteins*; Freeman & Co.: New York, 1984.
(23) Chan, H. S.; Dill, K. A. *Ann. Rev. Biophys. Biophys. Chem.* **1991**, *29*, 447−490.
(24) Friesner, R. A.; Gunn, J. R. *Ann. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 315−342.