The Early Stage of a β-Hairpin Folding: The Role of Water Molecules

Feng-Yin Li[a]* (李豐穎), Yuming Chen[b] (陳昱名),
Zhi-Qiang You[b] (游志強) and Chung-Yuan Mou[b] (牟中原)
[a]*Department of Applied Chemistry, National Chi Nan University, Nantou, Taiwan, R.O.C.*
[b]*Department of Chemistry, National Taiwan University, Taipei, Taiwan, R.O.C.*

Water has been shown to play a crucial role in stability and catalytical function. Our purpose is to understand the salvation behavior of water molecules during the folding process, in particular, a deca-peptide β-hairpin folding. We performed molecular dynamics simulation by employing AMBER6.0 with an explicit water model. The results indicate that water molecules surrounding a biomolecule can be categorized into at least three shells.

INTRODUCTION

In the past few years, there has been a resurgence of interest in biomolecular hydration and an increasing awareness of the importance of hydration in biomolecular stability and function.[1] There is a general consensus[2] about the fact that the hydrophobic effect is the dominant driving force of protein folding. In contrast to hydrophilic solvation, which is enthalpic in nature, hydrophobic effect is dominated by entropic interaction and involved with a large negative heat capacity. Nonetheless, the physical properties of liquid water underlying the phenomenon of hydrophobicity and giving rise to the characteristic behavior of the change of the thermodynamic functions upon salvation of a protein are not yet completely understood, despite the extensive studies devoted to their investigations.

It is believed that the hydrophobic effect is closely related to the sizes and shapes of the solutes since the geometries of hydrogen-bonded networks play an important role in the properties of water. Due to Stillinger's 1973 paper,[3] structure differences between the solvation of a small hydrophobic species and extended hydrophobic surface have been qualitatively understood. When the solutes are small enough, water can reorganize near them without sacrificing hydrogen bonds. This hydrophobic interaction acts weakly and on small length scales only. On the other hand, close to a large hydrophobic object, the persistence of a hydrogen bond network is geometrically impossible and the nature of hydrophobicity changes. Recently a qualitative theory developed by Lum, Chandler and Weeks (LCW)[4] shows that the crossover from the small to large length scale regimes occurs for surfaces extending roughly one nanometer, the characteristic length scale of protein structure. Many of the effects that can be explained by this analysis have been observed experimentally with studies of forces between macroscopic surfaces in water.[5,6] Current theoretical models for protein folding[7,8] have not yet accounted for the multifaceted nature of hydrophobicity.

MODEL AND METHOD

It is surmised by Huang and Chandler[9] that the complicated phenomena found in protein folding might be due to the change in length scale of hydrophobic surface from unfolded conformation to folded structure. In an effort to better understand the role of hydrophobicity during early folding dynamics, we launched an extensive molecular dynamic simulation of a ten-residue peptide with sequence Ac-VVVV^{D}PGVVVV-NH_2 ($V_4{}^{D}PGV_4$, hereafter) surrounded by 1714 water molecules. We chose this deca-peptide because of its simplicity, and it also has been reported previously to fold as a stable β-hairpin structure (Fig. 1).[10] The method we implemented is molecular dynamics (MD) simulations using AMBER6.0[11] with AMBER96 force field[12] to solve Hamilton's equations of motion for the peptide models coupled with an explicit TIP3P water model.[13] All the simulations were done under an NPT ensemble with the temperature and pressure controlled at physiological conditions (i.e., 300 K and 1 atm). The equations of motion were integrated using the Verlet algorithm[14] with a time step of 2 fs. The electrostatic interactions were calculated with Particle Mesh Ewald summation.[15,16] Bonds involving hydrogen were constrained using the SHAKE algorithm.[17] We started the simulation

Dedicated to the celebration of the seventieth anniversary of Chemical Society the located in Taipei.

from an extended conformation with an equilibration phase of 30 ps with periodic boundary conditions imposed by a nearest image convention in a rectangular box. The density of the system reached 0.99 g/mL within 24 ps, and remained so for the remaining of the trajectory. The simulation then was conducted for 10 ns (5 million integration steps) with the conformation of the β-hairpin and all surrounding water molecules saved every 1ps for further structural analysis. In some time intervals that we are interested in, we saved at 0.01 ps intervals for a detailed dynamic analysis.

We analyzed the peptide conformation by employing radial gyration of the hairpin and the distances among the α-hydrogen atoms of the Valine residues. As a simple measure of the complicated folding process, the R_γ gauges the overall shape of the hairpin in the folding process. In order to monitor the similarity between any instantaneous conformation and the native structure, we found that the distance of two α-hydrogen atoms belonging to a specific residue pair can be a good index, which is referred to as distance of Val α-H.pair (DVHP). There are four such residue pairs that are Val_1~Val_{10}, Val_2~Val_9, Val_3~Val_8, and Val_4~Val_7 and the relative DVHP's in the native structure are 2.5, 7.3, 2.5, and 7.3 Å, respectively.

In an attempt to understand the complicated process that happens during the early folding of the hairpin, we decided to focus on the water molecules surrounding the hairpin instead of the hairpin itself. Hydrogen bonds are responsible

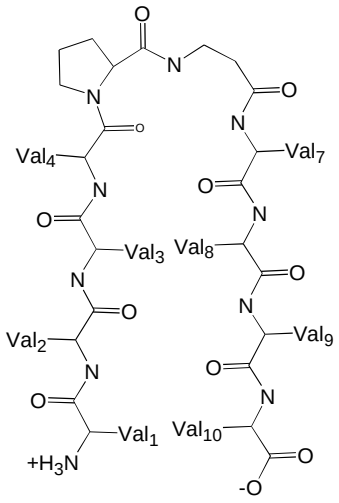

Fig. 1, V4DPGV4

Fig. 1. Native structure of $V_4^DPGV_4$. The native contact is defined as an α-hydrogen pair between two specific residues. There are 4 native contacts that are Val_1~Val_{10}, Val_2~Val_9, Val_3~Val_8 and Val_4~Val_7.

for many of water's peculiar properties. We will analyze the hydrogen bond statistics around the $V_4^DPGV_4$. We adapted a geometric definition of the water-water hydrogen bonds; according to that, a hydrogen bond between a water pair, water and $V_4^DPGV_4$, or within $V_4^DPGV_4$ itself is defined as the following: If the *X*-*X* distance is less than 3.4 Å and the *X* -H~ *X* angle is larger than $2\pi/3$, where *X* can be either an oxygen or a nitrogen atom. The salvation coordinate number (SCN) of a particular water molecule is defined as the number of water molecules with their oxygen atoms lying within a 3.4 Å sphere formed by its oxygen atom. In order to facilitate the analysis of water behavior during the folding process, we employed the widely used definition of salvation shells. Water is considered to belong to the first salvation shell if the water's oxygen is within 4.0 Å to any heavy atoms of the peptide.[18,19] This definition is based on the biological water study by Zewail et al.[20] In this study, the results showed that the average rotational reorientation time for water molecules within a 4- Å distance shell from a protein surface is much slower than that of a 14- Å shell. The water molecules of the second salvation shell are defined as their oxygen atoms are within 3.4 Å from any oxygen atoms of the first shell water molecules. With this definition, the second shell water molecules can be considered as the solvation shell of the first shell water molecules. The rest of the water molecules are categorized as the third shell. From Zewail's study,[20] their results also indicated that water molecules, residing beyond 7 Å from the protein's surface, are considered as bulk type, and the water molecules of the third shell in our definition are designed to behave like their bulk water. In order to correlate the positions of water molecules with the hairpin, we define a direction vector from the oxygen atom of a water molecule to its nearest heavy atom of the hairpin. The orientation of any water molecule with the hairpin can be described by two angles (see Fig. 2 for details). The direction vector is defined as the line connecting the oxygen atom of the related water molecule to the nearest heavy atom of the hairpin. The first angle is the angle between its direction vector and its dipole vector, which is also called the Z-axis or C2-axis. The other one is the angle between this direction vector and the normal vector of the plane containing the water molecule, which is called the Y-axis or C-axis.

As to the orientational relaxation, we calculate the rotational autocorrelation function of water molecules, $C_2^{\alpha}(t) = \langle P_2[\alpha(t)\alpha(0)]\rangle$, where P_2 is the Legendre polynomial of rank *2* and α is the unit vector of certain molecular direction towards the peptide. In particular, we calculate three different α vectors of $C_2^{\alpha}(t)$, which are the molecular dipole vector (also referred to as the Z-axis or C2-axis), the normal vector of wa-

ter's plane (also called the Y-axis or C-axis) and the H-H vector (or X-axis). At short times, the decay of $C_2^o(t)$ is generally non-exponential because of inertial and non-Markovian effects. At long times, when these effects are not important and the relaxation is diffusional, $C_2^o(t)$ decays exponentially with the time constant τ_2^o.

RESULT AND DISCUSSION

Several snapshots of the folding process are shown in Fig. 3, and it is clear that the two strands of the hairpin undergo close and open conformation change. As shown in Fig. 4(a), the time development of the R_γ during the course of the simulation indicates the hairpin undergoes a fast conformation change between collapsing and expansion. The R_γ of the starting confirmation of the hairpin is 7.5 Å. The hairpin quickly collapses to 5.7 Å at 150 ps. The R_γ remains about 6.0 Å until 2000 ps and the conformation of the hairpin undergoes a rapid collapsing-expansion process in the remaining 8000 ps. The time evolution of 4 DVHP's values shown in Fig. 4(b) also indicates that the hairpin reaches only a molten globular state during the entire simulation. We then analyzed the structure properties of water molecules, such as salvation coordination number (SCN), water-water hydrogen bond numbers (nHB) and water orientation according to the salvation shells.

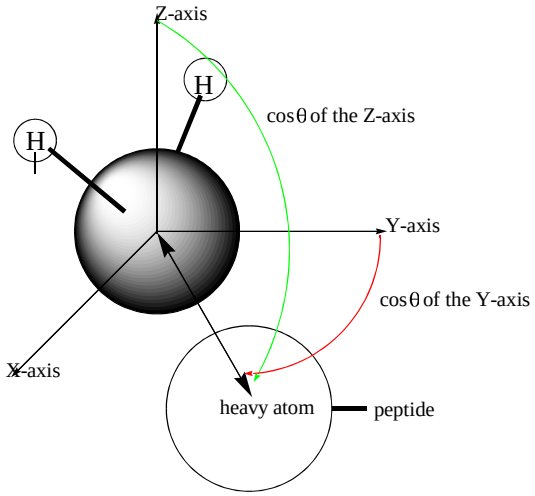

Fig. 2. Schematic representation of axes definition and angle definition. The Z axis is the same as the dipole direction; the X axis is parallel with the connecting line between two hydrogen atoms and the Y axis is perpendicular to the water plane.

Salvation and Hydrogen-bond Statistics

In Table 1, we summarize the average number of water molecules along with the average number of these water molecules close to the side chain carbons (the hydrophobic surface in our case), the main chain carbons, the carbonyl oxygen atoms and amino nitrogen atoms of the hairpin during the 10 ns trajectory (see Fig. 5 for details). These numbers fluctuated a lot during the simulation. From Table 1, it can be found that most of the water molecules are populated around the hydrophobic surface in all three different shells. The distribution maximum and the average of SCN along with the distribution maximum and the average of the hydrogen-bonding statistics among the three different shell water molecules are shown in Table 2. We find that the salvation and hydrogen-bond statistics of the second and third shell are very similar to those of bulk water, and only the first shell shows slightly different statistics. Compared with the distribution maximum and the average value of SCN among the bulk water, which are 5 and 4.8, respectively, it is shown that the first shell water molecules have almost one neighboring water molecule less than bulk water molecules. In the analysis of the hydrogen bond, we found that the distribution and the average value of hydrogen bond number in the first shell water molecules are almost equal to those of the bulk water. If we consider only the water-water hydrogen bond, the distribution maximum and the average value are 3 and 3.32, respectively. Taking the

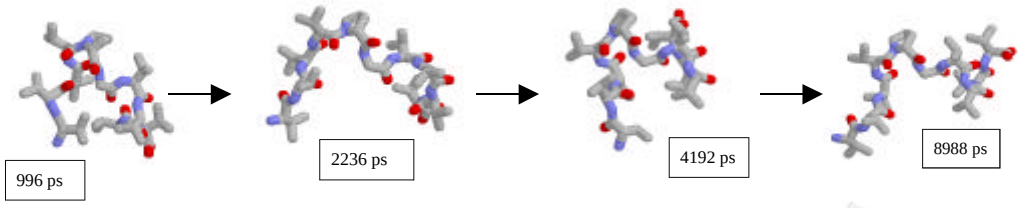

Fig. 3. Snapshots of the conformation along the 10-ns trajectory. The gray tubes are carbon atoms; the red tubes represent oxygen atoms; and the blue tubes symbolize nitrogen atoms. The number inside the box is the time along the trajectory when the snapshot is taken. It is clear that the two strands of the hairpin undergo close and open conformation change.

Table 1. The Salvation Statistics among the Three Salvation Shells

	Average	Close to side-chain carbon	Close to main-chain carbon	Close to carbonyl oxygen	Close to amino nitrogen
First shell	71.2	39.2	2.0	22.1	7.9
Second shell	105.2	75.8	1.8	21.1	6.5
Third shell	155.8	115.2	1.2	28.2	10.9

The columns show the average number of water molecules and the numbers of those water molecules close to the four different heavy atom groups of the hairpin.

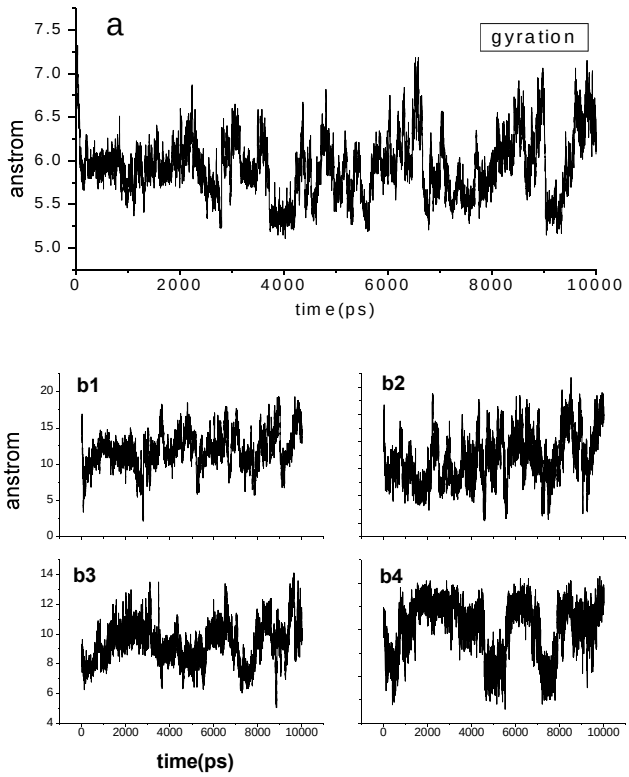

Fig. 4. The gyration and DVHP plots of the 10-ns trajectory. (a) The x-axis of the gyration plot is the time with ps as its unit and the y-axis is the gyration whose unit is $Å^2$. It is obvious that the hairpin does not reach its native structure during the 10-ns trajectory. (b) The plot of DVHP distances during the 10-ns trajectory . The x-axis is the time with ps as its unit and the y-axis is the DVHP distance whose unit is Å. The b1 line is the distance of the α-hydrogen pair between Val_1 and Val_{10}; the b2 line is that between Val_2 and Val_9 ; the b3 line is that Val_3 and Val_8 ; and the b4 line is that between Val_4 and Val_7.

water-peptide hydrogen bond into account, these two values become 4 and 3.66, respectively. If we focus on the water molecules around the hydrophobic surfaces, these two values are also 4 and 3.66, respectively. From Table 2, we find the salvation and hydrogen-bond statistics; the second and third shells are not much different from bulk water.

Structural Orientation Analysis

In the orientation analysis of the water molecules among the three shells, we examine how the hairpin influences the water molecules orientation during its folding pro-

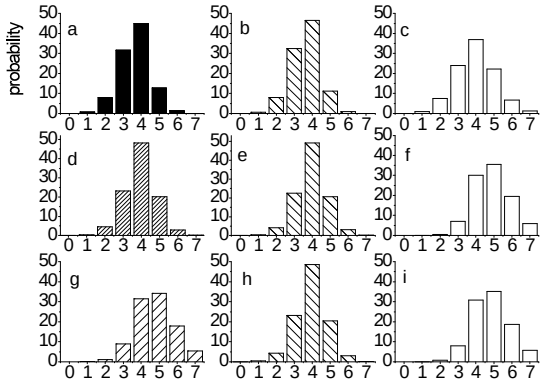

Fig. 5. The salvation and hydrogen bond statistics. The distribution of hydrogen bond numbers in first shell waters, (a) total and (b) adjacent to side-chain carbons. (c) The distribution of the SCN in first shell waters. (d) The distribution of hydrogen bond numbers in bulk water. (e) The distribution of hydrogen bond numbers in second shell waters which are near side-chain carbons. (f) The distribution of SCN in second shell waters. (g) The distribution of SCN in bulk water. (h) The distribution of hydrogen bond numbers in third shell waters which are near side-chain carbons. (i) The distribution of SCN in third shell waters.

Table 2. Salvation Coordinate Number and Hydrogen-bond Statistics

	Max SCN	Average SCN	Max HB	Average HB
First shell	4	4	4	3.32
Second shell	5	4.9	4	3.96
Third shell	5	4.8	4	3.94
Bulk water	5	4.8	4	3.93

The columns indicate the maximum distribution of SCN, the average value of SCN, the maximum distribution of hydrogen bonds and the average number of hydrogen bonds among the three different salvation shells and bulk water.

cess (see Fig. 6 for details). Because the water molecules populate mostly around hydrophobic surfaces, we will examine if these water molecules exist have a certain orientation preference with respect to the hairpin. Figure ? compares the orientation distribution of the water molecules among the first shell along the Z-axis (dipole vector) and Y-axis. Due to the influence of a water-peptide hydrogen bond, the O-H vectors of the water molecules, adjacent to the carbonyl oxygen atoms of the hairpin, mostly point towards the hairpin. For the same reason, the vectors of the water molecules, adjacent to the amide nitrogen atoms of the hairpin, also point in the opposite direction with respect to those of carbonyl oxygen atoms. Even without the influence of a water-peptide hydrogen bond, the water molecules around the hydrophobic surface (adjacent to the side chain carbons) still show a certain orientation preference. The maximum of the Z-axis distribution of these water molecules is found between the $1/2\pi$~ $2/3\pi$, and the maximum of the Y-axis distribution is almost zero. It means that the dipole vectors of the water molecules around the hydrophobic surfaces point away from the hydrophobic surfaces, and the plane containing each of these water molecules is mostly parallel to the related hydrophobic surfaces. Judging from the hydrogen-bonding statistics, the first shell water molecules might form cavities around the related hydrophobic surfaces through a water-water hydrogen bond network. As to the second and third shells, we only calculated the orientation distribution of Z- and Y-axis for those water molecules close to hydrophobic surfaces (see Fig. 7 for details). We find that the maximum of the Z-axis distribution of these water molecules lies between the $1/2\pi$~$2/3\pi$, but the absolute value of the maximum is smaller than that found in the first shell. The distribution of the Y-axis distribution is very close to an even distribution. There exists a slight preference for angles smaller than $1/4\pi$. It means that the second shell water molecules near the hydrophobic surface align their hydrogen atoms toward the first shell waters. As to the orientation distribution of the third shell water molecules, we find that it is almost evenly distributed for all angles (see Fig. 7 for

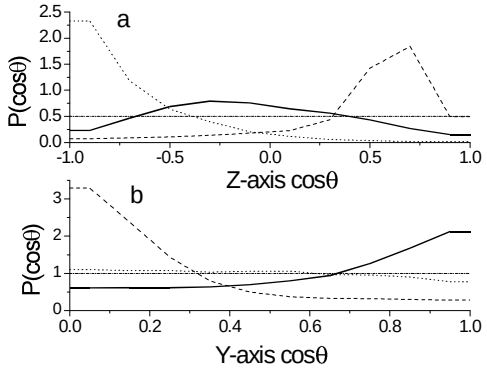

Fig. 6. The distribution of the cosθ of (a) the Z-axis and (b) Y-axis of the first shell waters. The dotted, dashed and solid green lines represent the distribution of the water adjacent to the nitrogen, oxygen and side-chain carbon, respectively. The dot-dash lines indicate the even distribution.

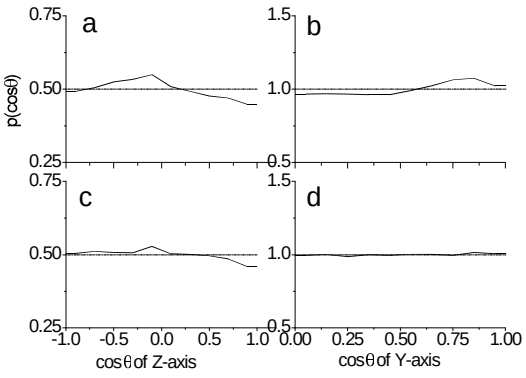

Fig. 7. The distribution of the cosθ of (a) the Z-axis and (b) Y-axis of the second shell water molecules near the side-chain carbon. The distribution of the cosθ of (c) the Z-axis and (d) Y-axis of the third shell waters which are near the side-chain carbon. The dot-dash lines shows the even distribution.

details). It means that the third water molecules show no preferred orientation towards the hairpin.

Orientation Dynamic Analysis

From the results of rotational auto-correlation function (Fig. 8a), we find that the water molecules around hydrophobic surfaces relax significantly more slowly than bulk water (Fig. 8d). All of the three axes of the water molecules show similar trends. Like those found in bulk water, the Z-axis and X-axis of the first shell water molecules, i.e. the axes involved in hydrogen-bonding, relax significantly more slowly than the Y-axis, which is unrelated to hydrogen-bonding. A similar trend also appears in the case of the second shell water molecules except the relaxation is slightly faster than that found in the first shell. We find the rotational auto-correlation functions of the third shell water molecules are not much different from those of the first shell ones.

CONCLUSION

The water molecules in the first shell are influenced by the peptide much more than those in the other shell. If the water molecules in the first shell are adjacent to the polar group, they will tend to form the water-peptide hydrogen bond. But if the first shell water molecules are around a hydrophobic surface, they will tend to form the water-water hydrogen bond and might form a cavity to accommodate the hydrophobic surfaces. The first shell water molecules can reflect the properties of the local structure of the hairpin with the same number of hydrogen bonds as in bulk water. We suggest that the hydrophobic effect comes from the restriction of the water orientation, and we also find that the reorientation of the water molecules is under the influence of the peptide, and the range of this influence goes beyond the immediate surrounding water. Our results indicate that the second shell water molecules are still under the influence of the conformation of the hairpin but it is not the case of those water molecules in the third shell, which behave like bulk water. In order to understand the salvation behavior of the water around biomolecules, the amount of water should be able to extend to the third shell. This finding is very close to the results of an X-ray scattering study of the water structure around myoglobin.[21]

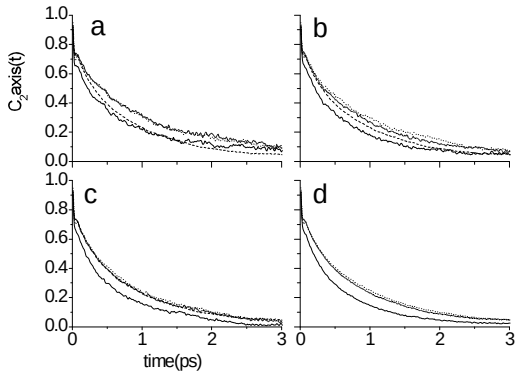

Fig. 8. The auto-correlation function of the waters (a) adjacent to side-chain carbons in the first shell, the waters (b) near side-chain carbon in the second shell and (c) third shell. (d) The auto-correlation function of three axes of bulk water. The bold-solid green, dotted blue and solid brown lines represent the Z-axis, X-axis and Y-axis, respectively. Dashed line is the $C_2Z(t)$ of bulk water.

Received July 3, 2002.

Key Words

Protein folding; β-Hairpin; Hydration.

REFERENCES

1. Israelachvili, J.; Wennerstrom, H. *Nature* **1996**, *379*, 219.
2. Dill, K. A. *Biochemistry* **1990**, *29*, 7133.
3. Stillinger, F. H. *J. Solution Chem.* **1973**, *2*, 141.2
4. Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem.* **1999**, *103*, 4570.
5. Christenson, H. K.; Claesson, P. M. *Science* **1988**, *239*, 390.
6. Ducker, W. A.; Xu, Zhenghe; Israelachvili, J. N. *Langmuir* **1994**, *10*, 3279.
7. Bryngelson, J. D.; Onuchic, J. N.; Socci, N. S.; Wolynes, P. G. *Proteins* **1995**, *21*, 167.
8. Shakhnovich, E. I. *Curr. Opin. Struc. Biol.* **1997**, *7*, 29.
9. Huang, D. M.; Chandler, D. *J. Phys. Chem. B* **2002**, *106*, 2047.
10. Wang, H.; Varady, J.; Ng, L.; Sung, S. S. *Proteins: Structure, Function, and Genetics* **1999**, *37*, 325.
11. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Koll-man, P. A. AMBER 6, University of California, San Francisco, 1999.
12. Kollman, P. A.; Dixon, R. W.; Cornell, W. D.; Chipot, T.; Fox, C.; Pohorille, A. Computer Simulations of Biological Systems, Dordrecht, The Netherlands: ESCOM, 1997.
13. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey,

R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
14. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University: Oxford, 1987.
15. Ewald, P. *Ann. Phys.* **1921**, *64*, 253.
16. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
17. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
18. Cheng,Y. K.; Rossky, P. J. *Nature* **1998**, *392*, 696.
19. Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980**, *102*, 4287.
20. Pal, S. K.; Peon, J.; Zewail, A. H. *Proc. Natl. Acad. Sci.* **2002**, *99*, 1763.
21. Dorbez-Sridi, R.; Cortes, R.; Mayer, E.; Pin, S. *J. Chem. Phys.* **2002**, *116*, 7269.