Protein Flexibility in Aqueous and Nonaqueous Solutions

David S. Hartsough and Kenneth M. Merz, Jr.*

Contribution from the Department of Chemistry, 152 Davey Laboratory, Pennsylvania State University, University Park, Pennsylvania 16802. *Received May 26, 1992*

Abstract: Herein we present the results of molecular dynamics simulations of a protein in an organic solvent. Our results show that while the amino acid side chains are more flexible in water, there is only a very small difference in backbone flexibility between protein in chloroform and protein in water. We propose that the major difference between the two solvent environments is the preferred orientations of the amino acid side chains. These results aid in the understanding of experimental observations and offer important insights for protein engineering in nonaqueous media.

Introduction

One of the most exciting developments in enzyme chemistry has been the recent discovery that enzymes are able to retain their catalytic activity in nonaqueous media.[1] Although this notion was initially dispelled by much of the chemical community,[2] the persistent efforts of Klibanov and others have clearly demonstrated this to be true.[3] In addition to retaining catalytic activity, enzymes in organic solvents were also found to exhibit profoundly enhanced thermostability.[4] Futhermore, it has been found that it is possible in organic solutions to "imprint" the protein with an inhibitor and to induce changes in the enzymatic rate and selectivity after removal of the inhibitor.[5] These characteristics of enzymes in organic solutions were explained on the basis of increased rigidity of proteins in nonaqueous media. Klibanov and others have proposed that water acts as a "molecular lubricant", resulting in greater protein flexibility.[6] However, scant information is available at the molecular level to enable one either to prove or to disprove this proposal. In order to accurately assess the role of flexibility in the activity of proteins, information concerning protein flexibility in both aqueous and nonaqueous environments is required.

Computer simulation has proven to be an invaluable tool in the understanding of protein structure and dynamics.[7,8] Since the seminal theoretical work of McCammon, Gelin, and Karplus demonstrating the dynamic nature of protein structures,[8] numerous studies have appeared examining the behavior of proteins in the crystal lattice, in vacuo, and in aqueous solution.[7] These studies have shown that proteins possess a wealth of conformational substates of nearly equal energy.[9] The surrounding medium may, therefore, influence protein flexibility by altering the energetic barriers separating these conformational substates and thereby affect the ease with which the protein may sample these states. Previous computational studies have considered the effect of hydration upon protein flexibility, but these studies have not examined the effects of explicit organic solvents.[10] Instead, hydrated protein was compared to protein in vacuo. However, because it neglects the damping effects of solvent friction and the van der Waals interactions of protein and solvent, protein in vacuo is not an accurate model of protein in nonaqueous solution.[11] A limited number of studies have directly addressed the behavior of small oligopeptides in organic solvents. In some cases these studies have found evidence for changes in both structure and dynamics upon placement of the polypeptide in a nonaqueous environment.[12] However, none of these studies have examined the behavior of polypeptides containing charged/polar residues in hydrophobic, low-dielectric solvents. During the course of this work, a report appeared describing the results of simulations of α-chymotrypsin in low- and high-dielectric environments.[13] These authors found increased flexibility in the high-dielectric environment. However, these studies did not include the effects of explicit organic solvent molecules.

Presented herein are the results of a theoretical investigation of protein dynamics in water and chloroform. Molecular dynamics (MD) simulations of bovine pancreatic trypsin inhibitor (BPTI) in each of these solvents have enabled us to evaluate the influence of solvent upon protein structure and dynamics.

Computational Methods

Starting coordinates for all heavy atoms of BPTI and 73 waters of hydration were obtained from the crystal structure.[14] Counterions were introduced to neutralize the 6+ charge of BPTI by replacing six of the waters of hydration near positively charged groups with chloride ions. The structure was optimized in vacuo using the OPLS[15] parameter set and the AMBER united atom force field in AMBER 4.0.[16] The binding energy of each water of hydration was evaluated, and the 20 most strongly bound waters were retained and the others deleted. This amount of water roughly corresponds to that required for hydration of charged

(1) Klibanov, A. M. *Acc. Chem. Res.* **1990**, *23*, 114.
(2) Dagani, R. *Chem. Eng. News* **1984**, *62*, 1249.
(3) (a) Zaks, A.; Klibanov, A. M. *Science* **1984**, *224*, 1249. (b) Zaks, A.; Klibanov, A. M. *J. Am. Chem. Soc.* **1986**, *108*, 2767. (c) Klibanov, A. M. *TIBS* **1989**, *14*, 141. (d) Margolin, A. L.; Tai, D.-F.; Klibanov, A. M. *J. Am. Chem. Soc.* **1987**, *109*, 7885. (e) Sakurai, T.; Margolin, A. L.; Russell, A. J.; Klibanov, A. M. *J. Am. Chem. Soc.* **1988**, *110*, 7236. (f) Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 3194. (g) Riva, S.; Chopineau, J.; Kieboom, A. P. G.; Klibanov, A. M. *J. Am. Chem. Soc.* **1988**, *110*, 584. (h) Zaks, A.; Klibanov, A. M. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 3192. (i) Wheeler, C. J.; Croteau, R. *Arch. Biochem. Biophys.* **1986**, *248*, 429. (j) Deetz, J. S.; Rozzell, J. D. *TIBTECH* **1988**, *6*, 15. (k) Reslow, M.; Adlercreutz, P.; Mattiasson, B. *Appl. Microbiol. Biotechnol.* **1987**, *26*, 1. (l) Ryu, K.; Dordick, J. *J. Am. Chem. Soc.* **1989**, *111*, 8026. (m) Ståhl, M.; Jeppson-Wistrand, U.; Månsson, B. *J. Am. Chem. Soc.* **1991**, *113*, 9366.
(4) (a) Ayala, G.; de Gómez-Puyou, M. T.; Gómez-Puyou, A.; Darszon, A. *FEBS* **1986**, *203*, 41. (b) Capaldi, R. A.; Malatesta, F.; Darley-Usmar, V. M. *Biochim. Biophys. Acta.* **1983**, *726*, 135-148. (c) Hatefi, Y. *Annu. Rev. Biochem.* **1985**, *54*, 1015.
(5) (a) Russell, A. J.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 11624. (b) Ståhl, M.; Jeppsson-Wistrand, U.; Månsson, M.-O.; Mosbach, K. *J. Am. Chem. Soc.* **1991**, *113*, 9366.
(6) (a) Poole, P. L.; Finney, J. L. *Biopolymers* **1983**, *22*, 255. (b) Poole, P. L.; Finney, J. L. *Int. J. Biol. Macromol.* **1983**, *5*, 308. (c) Poole, P. L.; Finney, J. L. *Comments Mol. Cell. Biophys.* **1984**, *2*, 129. (d) Careri, G.; Gratton, E.; Yang, P.-H.; Rupley, J. A. *Nature* **1980**, *284*, 572.
(7) (a) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: New York, 1987. (b) Brooks, C. L., III; Petit, B. M. *Advances in Chemical Physics Volume LXXI. Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; John Wiley & Sons: New York, 1988.
(8) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585.
(9) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598.
(10) (a) Steinbach, P. J.; Loncharich, R. J.; Brooks, B. R. *Chem. Phys.* **1991**, *158*, 383. (b) Kitao, A.; Hirata, F.; Go, N. *Chem. Phys.* **1991**, *158*, 447.
(11) Levitt, M. Sharon, R. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 7757.
(12) (a) Lautz, J.; Kessler, H.; van Gunsteren, W. F.; Weber, H.-P.; Wegner, R. M. *Biopolymers* **1990**, *29*, 1669. (b) Mierke, D. F.; Kessler, H. *J. Am. Chem. Soc.* **1991**, *113*, 9466. (c) De Loof, H.; Nilsson, L.; Rigler, R. *J. Am. Chem. Soc.* **1992**, *114*, 4028.
(13) Affleck, R.; Haynes, C. A.; Clark, D. S. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5167.
(14) (a) Walter, J.; Huber, R. *J. Mol. Biol.* **1983**, *167*, 911. (b) Wlodawer, A.; Walter, J.; Huber, R.; Sjolin, L. *J. Mol. Biol.* **1984**, *180*, 307. (c) Wlodawer, A.; Diesenhofer, J.; Huber, R. *J. Mol. Biol.* **1987**, *193*, 145.
(15) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(16) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, C.; Weiner, P.; Kollman, P. A. *AMBER 4.0*, University of California, San Francisco, 1991.

0002-7863/92/1514-10113$03.00/0

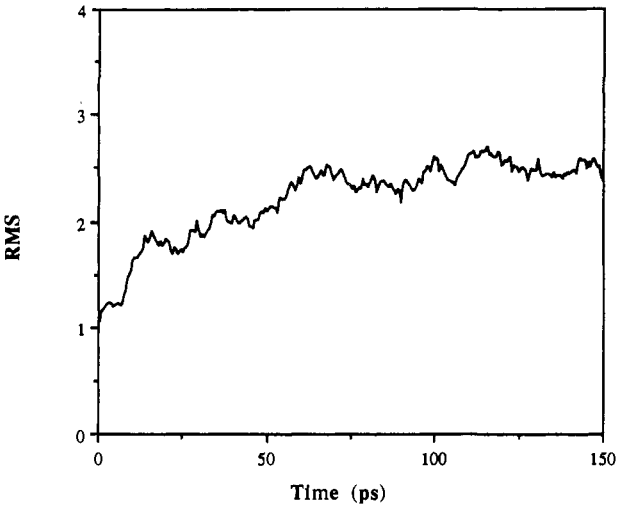

Figure 1. RMS deviations between the instantaneous computed structure in water and the crystal structure for all residues as a function of time. Sampling was begun at 102 ps.

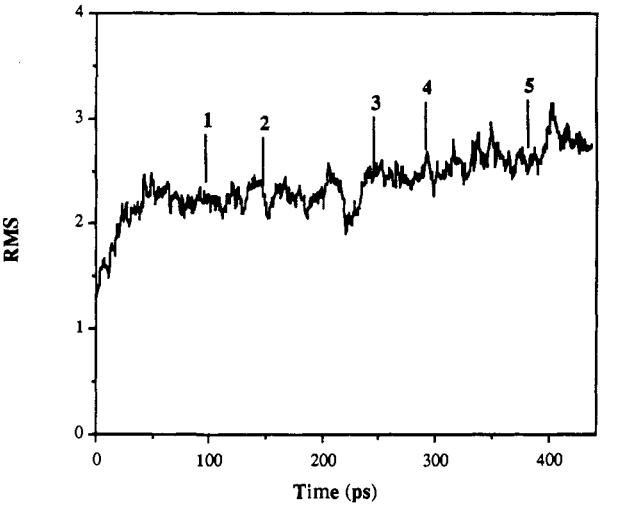

Figure 2. RMS deviations between the instantaneous computed structure in chloroform and the crystal structure for all residues as a function of time. (1) Beginning of sampling period for 298 K. (2) End of sampling period and beginning of temperature ramp to 310 K. (3) Beginning of sampling period for 310 K. (4) End of sampling period for 310 K and start of temperature ramp to 360 K. (5) Beginning of sampling period for 360 K.

groups.[17,18] The optimized structure was immersed in a box of either TIP3P water[20] or chloroform,[21] and all solvent atoms closer than 2.2 or further than 6.0 Å in any Cartesian direction were eliminated, producing solvent boxes with initial dimensions of 45.9 × 35.7 × 41.3 Å containing 1689 and 362 water and chloroform solvent molecules, respectively. Prior to MD, both solvent boxes were minimized with periodic boundary conditions to an RMS gradient <0.5 kcal mol^{-1}. For all MD simulations, solute and solvent were independently coupled to a constant-temperature heat bath. Periodic boundary conditions at constant pressure were enforced, and all bond lengths were constrained using the SHAKE[22] algorithm with a tolerance of 0.0005 Å, allowing a time step of 1.5 fs. A pairlist was generated every 16 steps using a residue-based cutoff distance of 10.0 Å. A 150-ps trajectory of BPTI in water was obtained beginning at 5 K and heating to 310 K over a period of ~7 ps.[23] Data were accumulated over the last 48 ps of this trajectory. A 150-ps trajectory of BPTI in chloroform at 298 K was obtained in a similar manner. The final structure of this trajectory was used as the starting point of a second MD run of 144 ps at 310 K. Again, data were collected over the latter 48 ps of the trajectory. Finally, the end point of this run was used as the initial point of a 144-ps trajectory at 360 K. Coordinates were saved for analysis every 50 steps. Solvent-accesible surface areas were calculated using an algorithm recently developed in this lab.[25]

Results and Discussion

We chose for our simulations the protein bovine pancreatic trypsin inhibitor (BPTI). BPTI is large enough to exhibit the gross structural behavior of larger proteins, yet is still amenable to simulation. The RMS deviations of instantaneous structures from the crystal structure are plotted as a function of time in Figures 1 and 2 for BPTI in water and chloroform, respectively. The RMS deviations from the crystal are typical of those seen when using the OPLS parameter set.[26] Both systems appear to have reached an equilibrium state by ca. 100 ps. Following this initial equilibration period, each system was simulated for an additional 48 ps, during which time data was accumulated for analysis. At the

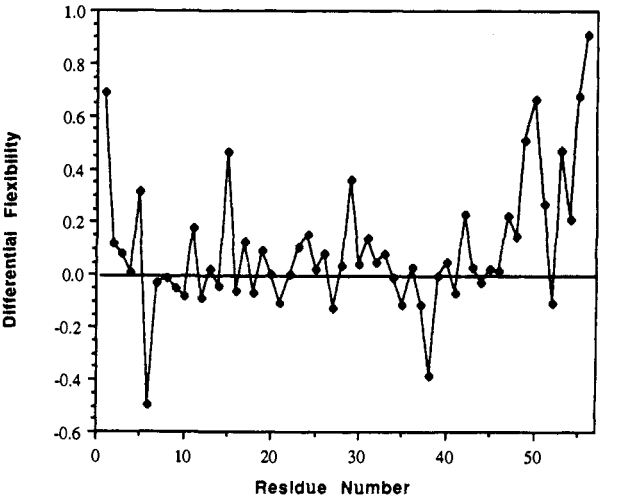

Figure 3. Differential flexibility by residue (water − 310 K chloroform). Average flexibilities were calculated (eq 1) for all heavy atoms of each residue for the 48-ps trajectories in water and chloroform. The flexibility values in chloroform were then subtracted from those in water to yield the differential flexibility.

Table I. Flexibility of BPTI (eq 1)

run	$(RMS_{all})^a$ (Å)	$(RMS_{back})^b$ (Å)
water	0.741 (0.688)[a]	0.531 (0.478)
$CHCl_3$, 298 K	0.528	0.418
$CHCl_3$, 310 K	0.614	0.476
$CHCl_3$, 360 K	0.630	0.512

[a] Value for all heavy atoms. [b] Value for backbone only. [c] In parentheses is value calculated if residues 1, 55, and 56 are not included in the analysis.

(17) Rupley, J. A.; Gratton, E.; Careri, G. *TIBS* **1983**, 18.
(18) These waters are included in an attempt to model the water which is often found to accelerate these reactions.[19]
(19) Kazandjian, R. Z.; Klibanov, A. M. *J. Am. Chem. Soc.* **1985**, *107*, 5448.
(20) (a) Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 335. (b) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(21) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.* **1990**, *94*, 1683.
(22) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.
(23) Due to water molecules entering and exiting the cutoff region, the temperature coupling in AMBER 4.0 does not rigorously produce the desired temperature.[24] When a target temperature of 298 K was used, an average bulk temperature of 310 K was obtained, consisting of 311 K for the solvent and 302 K for the solute. In an attempt to minimize complications arising from these temperature differences, comparisons are made with chloroform solutions which bracket the temperature of the water solution (i.e., 300 and 310 K). Similar difficulties were not encountered in the temperature coupling of the chloroform solutions.
(24) Personal communication with David A. Pearlman.
(25) LeGrand, S. M.; Merz, K. M., Jr. *J. Comput. Chem.*, in press.
(26) Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1990**, *112*, 2773.

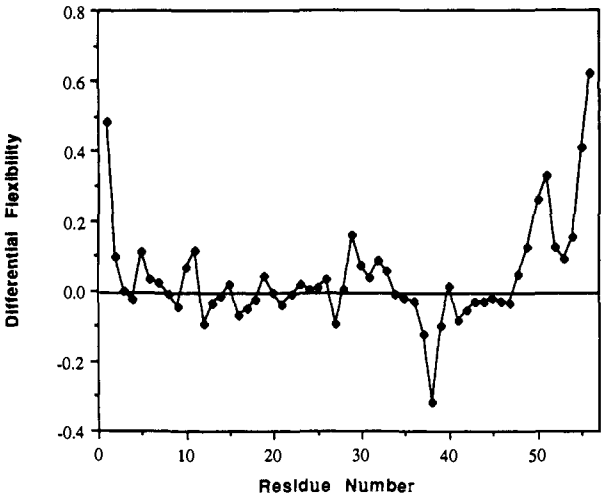

Figure 4. Differential flexibility of backbone by residue (water − 310 K chloroform). Same as Figure 1, except only backbone atoms were examined.

times indicated in Figure 2 (150 and 294 ps), additional energy was added to the system to increase the temperature of the chloroform solution to 310 and 360 K, respectively.

The flexibility/rigidity of a protein is a measure of the amount by which instantaneous structures deviate from a time-averaged structure (eq 1). It is apparent from the results of the simulations

$$\text{flexibility} = \langle(((r_i - \langle r_{ave}\rangle)^2))^{1/2}\rangle \qquad (1)$$

run in chloroform that protein flexibility and solution temperature are intrinsically related (Table I). Although the flexibility does not increase in direct linear proportion to the temperature, as the temperature of the solution is increased, the flexibility of the protein increases as expected. This increase in flexibility is manifested in both the motions of the side-chain atoms and in increased flexibility of the backbone.

BPTI in water is calculated to be more flexible than BPTI in chloroform (Table I). Even when the temperature of the chloroform solution is 50° greater than that of the aqueous solution, BPTI is more flexible in water than it is in chloroform (Table I). Examination of the differential flexibility of each residue reveals that the increased flexibility of BPTI in water is distributed throughout several residues of the protein, with the terminal residues possessing the greatest increase in flexibility (see Figure 3). However, it is clear that much of the increased flexibility of the backbone in water is due to motion of the first residue (Arg1) and several residues near the C-terminus (Gly56, Cys55, Cys51, Asp50) (see Figure 4). If the contributions made by residues 1, 55, and 56 are discounted, the flexibility of the BPTI backbone in water is reduced from 0.531 to 0.478, essentially the same as the value in 310 K chloroform (0.476). Thus, although our results are supportive of increased side-chain flexibility in an aqueous medium, we find a much smaller difference in the flexibility of the backbone.

It is extremely difficult to predict a priori the degree of conformational flexibility that must be lost to give rise to the thermostability and memory effects observed experimentally. Furthermore, the experimental observations may be due to a combination of differential flexibility and other structural or energetic perturbations arising from the nonpolar environment. For this reason, we have examined structural changes that may also be responsible for the experimental results.

Changing the surrounding medium from water to chloroform strongly influences the geometries of the amino acid side chains. In water the side chains of the surface residues are almost fully extended into the polar solvent, while in chloroform the side chains are folded back onto the surface of the protein (Figure 5). Consistent with this observation are the differences in solvent-accesible surface area in the two environments. If the surface residues are either engaged in polar contacts with other side chains or folded back onto the surface of the protein itself, a significantly smaller solvent-accesible surface area would be expected. Using a common probe radius of 1.4 Å to facilitate comparisons, the solvent-accesible surface of BPTI ranges from ca. 4100 to 4300 Å^2 and from 3600 to 3700 Å^2 for BPTI in water and chloroform, respectively, at 310 K. The observed differences in side-chain extension are maintained throughout the course of the simulations. From this evidence, it is apparent that the organic medium is, not

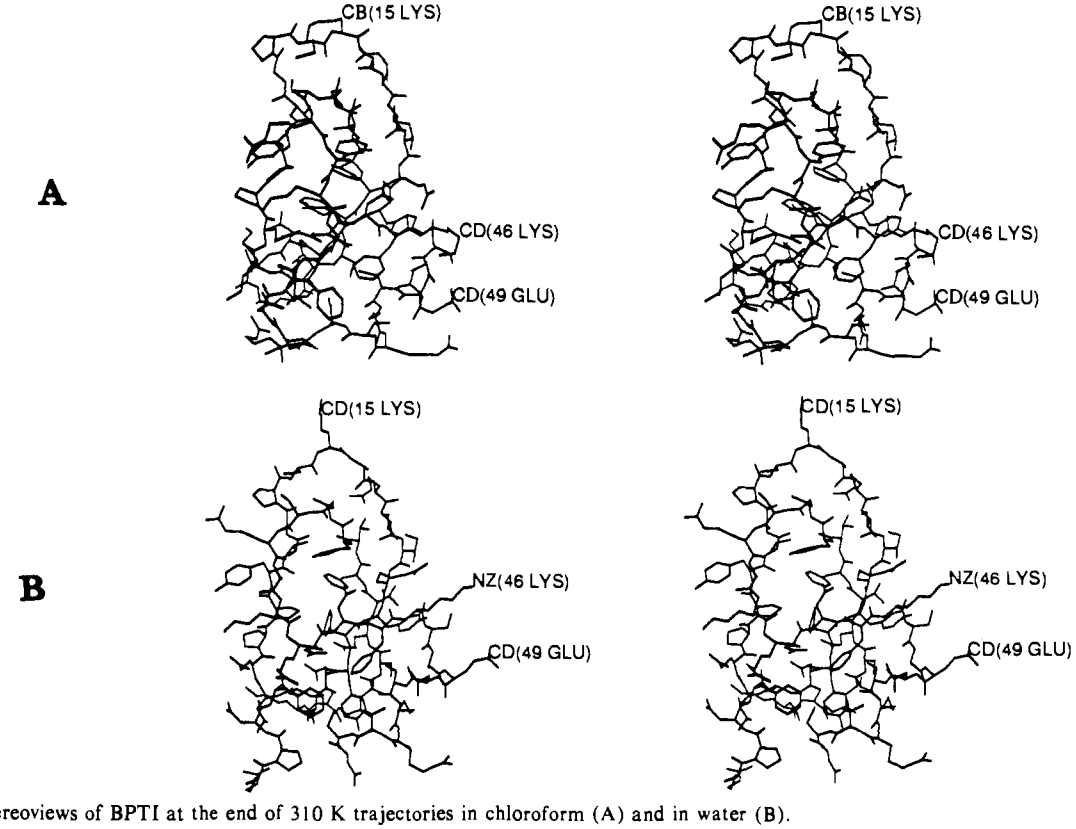

Figure 5. Stereoviews of BPTI at the end of 310 K trajectories in chloroform (A) and in water (B).

surprisingly, unable to stabilize the charged/polar side chains.

These structural changes can be used to rationalize the experimental results. In order for a protein to unfold, various hydrogen-bond and salt bridge contacts must be disrupted. As has been noted previously, these newly exposed polar groups can be readily stabilized by an aqueous environment.[27,28] The ability to form hydrogen bonds with the solvent is, therefore, important for protein unfolding. Clearly this will not be possible in the nonpolar environment. This observation offers an explanation of the noted increase in protein stability in nonpolar solvents.[3,4] Furthermore, these results offer a possible explanation for "molecular imprinting".[5] If the enzyme active site involves polar/charged residues, these side chains would be unable to reorient significantly in the nonpolar environment. Once moved to their "correct" positions by the inhibitor/imprinter, the polar residues in the active site would be unable to return to their old positions, as this would involve exposing these groups to the nonpolar solvent. It should be noted that this does not necessarily imply greater rigidity but only that the side chains occupy different regions of space in the two environments. Although these explanations are consistent with the experimental findings, they do not provide a concrete basis for identifying the source of the observed effects. Clearly, only a detailed analysis of the enzyme reaction path, which takes into account such factors as substrate solvation and possible changes in the identity of the rate-limiting step, will provide the desired insight into these novel processes.

Finally, the geometric changes observed offer a possible explanation for any loss of conformational flexibility in nonaqueous solvents. If the polar side chains are moved close to the surface of the protein, the strength of their electrostatic interactions with other charged/polar groups will increase. These interactions will effectively increase the number of cross-links between different regions of the protein. As these interactions increase in strength, the ability of the side chains to move away from these favored geometries will be hindered, resulting in decreased conformational flexibility.

Conclusions

The results presented provide one of the first molecular-level pictures of a protein in a nonaqueous environment. Our studies have shown that placement of a protein in a nonaqueous environment results in a significant loss of side-chain flexibility and a much smaller decrease in backbone flexibility. In addition to finding evidence for some loss of conformational flexibility, we also observe a pronounced change in side-chain geometries upon moving from an aqueous to a nonaqueous environment.

Acknowledgment. This research was supported by the ONR Grant No. N00014-90-3-4002. We would like to thank the Pittsburgh Supercomputer Center for Cray YMP time. We would also like to thank Dr. J. Tirado-Rives of Yale University for allowing us to use his QRMS program for the calculation of RMS deviations and Dr. Greg Farber of this department for many helpful discussions.

Registry No. BPTI, 9087-70-1; $CHCl_3$, 67-66-3; H_2O, 7732-18-5.

(27) Shirley, B. A.; Stanssens, P.; Hahn, U.; Pace, C. N. *Biochemistry* **1992**, *31*, 725.

(28) Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. *Biochemistry* **1991**, *30*, 9686.