Protein Dynamics and Solvation in Aqueous and Nonaqueous Environments

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 March 1, 1993

Abstract: Herein we report the results of molecular dynamics simulations of a protein in an organic solution. We present a detailed analysis of the dynamics and structure of both protein and solvent. These results are compared and contrasted with corresponding results from simulations of protein in water. We find that placement of protein in an organic medium results in a dramatically increased hydrogen bond network. We propose that it is this increase in intramolecular hydrogen bond interactions that is responsible for the experimentally observed properties of proteins in organic solvents. Furthermore, on the basis of our results we propose that creation of salt bridges will significantly enhance the thermostability of enzymes in nonaqueous solvents.

Introduction

The unique ability of enzymes to accelerate chemical reactions has long been of interest to chemists and biochemists alike. Synthetic chemists have recognized for years the potential value of enzymes as catalysts for a wide variety of organic transformations. However, these efforts have often been stymied by the very environment in which the enzymes occur naturally—aqueous solution. In many instances substrates are unable to withstand exposure to water, while in other cases, for example esterification, the products of the reaction are not stable in the aqueous environment. In an effort to circumvent these difficulties many different strategies have been employed. One of the most promising approaches has been the removal of all but the "essential" water from the enzyme and placement of the enzyme in an organic solvent.[1-14] In this medium the enzyme is able to fulfill its catalytic function without the deleterious effects of an aqueous environment.

In addition to remaining catalytically active, enzymes in organic media also exhibit many useful properties. One of the most interesting observations has been the dramatic increase in thermostability of an enzyme in an organic solution relative to an aqueous environment.[1] Other unique properties of enzymes in organic media include altered stereoselectivity and dramatic rate enhancement when the enzyme is pretreated with an inhibitor.[7,9] Many of these properties have been attributed to the "rigidifying environment of an organic solvent".[7] The proposition that water acts as a "molecular lubricant" loosening up the protein structure has been presented by others as well.[15-18] However, only limited experimental data exist to support this assertion, and it does not provide molecular level insight into the protein–solvent interaction.[15,19]

Computational simulations afford an excellent opportunity to examine, at the molecular level, protein behavior in organic solvents. Simulations have been done examining the effect of hydration upon protein function and dynamics, yet these studies have not directly addressed the effect of an organic solvent.[20,21] Previous work examining the effects of an organic environment has been limited to molecular dynamics simulations of small nonpolar oligopeptides in organic solvents,[22,23] or has not included explicit solvent,[24] which has been shown to be very important for

(1) Kazandjian, R. Z.; Klibanov, A. M. Regioselective Oxidation of Phenols Catalysed by Polyphenol Oxidase in Chloroform. *J. Am. Chem. Soc.* **1985**, *107*, 5448–5450.
(2) Klibanov, A. M. Enzymatic Catalysis in Anhydrous Organic Solvents. *TIBS* **1989**, *14*, 141–144.
(3) Klibanov, A. M. Asymmetric Transformations Catalyzed by Enzymes in Organic Solvents. *Acc. Chem. Res.* **1990**, *23*, 114–120.
(4) Margolin, A. L.; Tai, D.-F.; Klibanov, A. M. Incorporation of D-Amino Acids into Peptides via Enzymatic Condensation in Organic Solvents. *J. Am. Chem. Soc.* **1987**, *109*, 7885–7887.
(5) Reslow, M.; Adlercreutz, P.; Mattiasson, B. Organic Solvents for Bioorganic Synthesis 1. Optimization of Parameters for a Chymotrypsin Catalyzed Process. *Appl. Microbiol. Biotechnol.* **1987**, *26*, 1–8.
(6) Riva, S.; Chopineau, J.; Kieboom, A. P. G.; Klibanov, A. M. Protease Catalyzed Regioselective Esterification of Sugars and Related Compounds in Anhydrous Dimethylformamide. *J. Am. Chem. Soc.* **1988**, *110*, 584–589.
(7) Russell, A. J.; Klibanov, A. M. Inhibitor-Induced Enzyme Activation in Organic Solvents. *J. Biol. Chem.* **1988**, *263*, 11624–11626.
(8) Sakurai, T.; Margolin, A. L.; Russell, A.; Klibanov, A. M. Control of Enzyme Enantioselectivity by the Reaction Medium. *J. Am. Chem. Soc.* **1988**, *110*, 7236–7237.
(9) Ståhl, M.; Jeppson-Wistrand, U.; Månsson, B.; Mosbach, K. Induced Stereoselectivity and Substrate Selectivity of Bio-imprinted α-Chymotrypsin in Anhydrous Organic Media. *J. Am. Chem. Soc.* **1991**, *113*, 9366–9368.
(10) Wheeler, C. J.; Croteau, R. Terpene Cyclase Catalysis in Organic Solvent/Minimal Water Media. Demonstration and Optimization of (+)-α-Pinene Cyclase Activity. *Arch. Biochem. Biophys.* **1986**, *248*, 429–434.
(11) Zaks, A.; Klibanov, A. M. Enzymatic Catalysis in Organic Media at 100 °C. *Science* **1984**, *224*, 1249–1251.
(12) Zaks, A.; Klibanov, A. M. Enzyme-Catalyzed Processes in Organic Solvents. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3192–3196.
(13) Zaks, A.; Klibanov, A. M. Substrate Specificity of Enzymes in Organic Solvents vs Water Is Reversed. *J. Am. Chem. Soc.* **1986**, *108*, 2767–2768.
(14) Zaks, A.; Klibanov, A. M. Enzymatic Catalysis in Nonaqueous Solvents. *J. Biol. Chem.* **1988**, *263*, 3194–3201.
(15) Poole, P. L.; Finney, J. L. Sequential Hydration of a Dry Globular Protein. *Biopolymers* **1983**, *22*, 255–260.
(16) Poole, P. L.; Finney, J. L. Hydration-Induced Conformational and Flexibility Changes in Lysozyme at Low Water Content. *Int. J. Biol. Macromol.* **1983**, *5*, 308–310.
(17) Rupley, J. A.; Gratton, E.; Careri, G. Water and Globular Proteins. *TIBS* **1983**, *8*, 18–22.
(18) Ayala, G.; de Gómez-Puyou, M. T.; Gómez-Puyou, A.; Darszon, A. Thermostability of Membrane Enzymes in Organic Solvents. *FEBS* **1986**, *203*, 41–43.
(19) Guinn, R. M.; Skerker, P. S.; Kavanaugh, P.; Clark, D. S. Activity and Flexibility of Alcohol Dehydrogenase in Organic Solvents. *Biotechnol. Bioeng.* **1990**, *37*, 303–308.
(20) Kitao, A.; Hirata, F.; Go, N. The Effects of Solvent on the Conformation and the Collective Motions of Protein: Normal Mode Analysis and Molecular Dynamics Simulations of Melittin in Water and Vacuum. *Chem. Phys.* **1991**, *158*, 447–472.
(21) Steinbach, P. J.; Loncharich, R. J.; Brooks, B. R. The Effects of Environment and Hydration on Protein Dynamics: A Simulation Study of Myoglobin. *Chem. Phys.* **1991**, *158*, 383–394.
(22) De Loof, H.; Nilsson, L.; Rigler, R. Molecular Dynamics Simulation of Galanin in Aqueous and Nonaqueous Solution. *J. Am. Chem. Soc.* **1992**, *114*, 4028–4035.
(23) Lautz, J.; Kessler, H.; van Gunsteren, W. F.; Weber, H.-P.; Wenger, R. M. On the Dependence of Molecular Conformation on the Type of Solvent Environment: A Molecular Dynamics Study of Cyclosporin A. *Biopolymers* **1990**, *29*, 1669–1687.

0002-7863/93/1515-6529$04.00/0

accurate simulation of protein dynamics in aqueous solutions.[25] In an attempt to examine the effects of solvent environment upon protein structure and dynamics, we have chosen to carry out simulations of a small protein in chloroform and aqueous solution using explicit solvent models. The protein used in these simulations was bovine pancreatic trypsin inhibitor (BPTI). This protein is small enough to be amenable to computational simulation, yet large enough to exhibit the secondary and tertiary structural elements characteristic of protein structure.

In a preliminary report of this research, we primarily addressed the issue of protein flexibility and how it was affected by solvent.[26] We found a significant difference in flexibility between protein in aqueous and nonaqueous media, in agreement with experimental findings and other theoretical studies. However, the difference in flexibility was not evenly spread throughout the protein. For the backbone in particular, the differences in flexibility are limited to a few residues near the C and N termini. If these residues are left out of the analysis, essentially no difference in flexibility is seen between BPTI in water and chloroform. For the side chains, the differences in flexibility are more evenly distributed throughout the protein. In the present report, we expand our analysis and examine in detail the effect of solvent environment on various aspects of protein structure and dynamics. We also present an analysis of solvent structure and dynamics for both the aqueous and non-aqueous environments.

Computational Methods

Starting coordinates for all heavy atoms of BPTI and 73 waters of hydration were obtained from the crystal structure.[27,28] 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 with use of the OPLS[29] parameter set and the AMBER united atom force field in AMBER 4.0.[30] 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 groups and which has been found to accelerate some reactions.[1,17] The optimized structure was immersed in a box of either TIP3P water[31,32] or chloroform,[33] and all solvent atoms closer than 2.2 Å or further than 6.0 Å in any Cartesian direction from the protein were eliminated, producing a solvent box with initial dimensions of 45.9 × 35.7 × 41.3 Å containing 1689 water or 362 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 were enforced at constant pressure, and all bond lengths are constrained with use of the SHAKE[34] algorithm with a tolerance of 0.005 Å, allowing a time step of 1.5 fs. A pairlist was generated every 16 steps by 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.[35] 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 144-ps trajectory at 310 K. Again, data were collected over the latter 48 ps of the trajectory. Following preliminary analysis, an additional 24 ps of dynamics were carried out at 310 K (see text).

Coordinates were saved for analysis every 50 steps. Time dependent flexibility analysis was done with coordinates saved every 0.600 ps. Instantaneous values of the flexibility calculated with use of the longer time between coordinate sets were nearly indistinguishable from those calculated with use of coordinates saved every 0.075 ps. All other analyses were carried out on the full 640 coordinate sets saved during the 48 ps sampling period. Solvent accesible surface areas were calculated with use of an algorithm recently developed in this lab.[36] Solvent accesible surface areas were calculated in both solvent environments with use of a common probe radius of 1.4 Å.

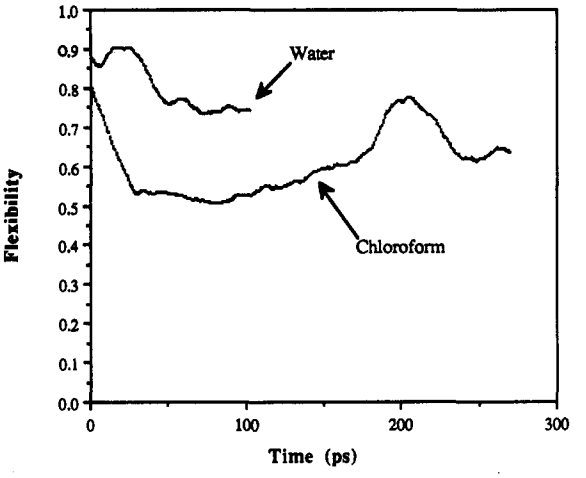

Figure 1. Plot of the flexibility of BPTI (calculated from eq 1) as a function of time for water (upper curve) and chloroform (lower curve). Time on x-axis corresponds to time at which 48 ps of sampling begins (e.g., the point at $t = 50$ represents the flexibility of BPTI calculated with data from 50 to 98 ps). Analysis was done on structures saved every 0.6 ps. Values reported previously correspond to $t = 102$ (water) and 102 and 246 ps (chloroform).

Results and Discussion: Protein Structure and Dynamics

Flexibility. The primary focus of our preliminary report of this research was protein flexibility and how it is affected by different solvent environments. We reported single values of protein flexibility determined by eq 1, where averaging was done over a 48-ps period. At that time we did not address the convergence properties of the flexibility of BPTI.

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

Shown in Figure 1 are plots of protein flexibility as a function of time. The time indicated on the x-axis corresponds to the time at which the 48 ps of structure averaging was begun. Simulations in both chloroform and water show the protein to be rapidly changing shape at early times. However, as BPTI begins to

(24) Affleck, R.; Haynes, C. A.; Clark, D. S. Solvent Dielectric Effects on Protein Dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5167–5170.
(25) Levitt, M.; Sharon, R. Accurate Simulation of Protein Dynamics in Solution. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7557–7561.
(26) Hartsough, D. S.; Merz, K. M., Jr. Protein Flexibility in Aqueous and Nonaqueous Solutions. *J. Am. Chem. Soc.* **1992**, *114*, 10113–10116.
(27) Wlodawer, A.; Walter, J.; Huber, R.; Sjolin, L. Structure of Bovine Pancreatic Trypsin Inhibitor Results of Joint Neutron and X-ray Refinement of Crystal Form II. *J. Mol. Biol.* **1984**, *180*, 301–329.
(28) Wlodawer, A.; Diesenhofer, J.; Huber, R. Comparison of Two Highly Refined Structures of Bovine Pancreatic Trypsin Inhibitor. *J. Mol. Biol.* **1987**, *193*, 145–156.
(29) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimization for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
(30) 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.
(31) Jorgensen, W. L. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers. Application to Liquid Water. *J. Am. Chem. Soc.* **1981**, *103*, 335–340.
(32) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
(33) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. Relative Partition Coefficients for Organic Solvents from Fluid Simulations. *J. Phys. Chem.* **1990**, *94*, 1683–1686.
(34) van Gunsteren, W. F.; Berendsen, H. J. C. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311–1327.
(35) Due to water molecules entering and exiting the cutoff region, the temperature coupling in AMBER 4.0 does not rigorously produce the desired temperature (personal communication from David A. Pearlman). When a target temperature of 298 K was used, an average 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.
(36) LeGrand, S. M.; Merz, K. M., Jr. Rapid Approximation to Molecular Surface Area via the Use of Boolean Logic and Look-Up Tables. *J. Comput. Chem.* **1993**, *14*, 349–352.

Table I. Radius of Gyration and SASA[a]

simulation conditions	$\langle r_{gyr} \rangle$ (Å)	SASA × 10^{-3} (Å^2)
chloroform, 300 K	10.88	3.64
chloroform, 310 K	10.90	3.65
water, 310 K	11.50	4.18

[a] Solvent accessible surface area.

achieve equilibrium with its surroundings, the flexibility moves toward a stable value.

It is clear from Figure 1 that the flexibility of BPTI in water has converged to roughly 0.75 after *ca.* 80 ps. While this value is higher than that observed in a previous simulation of BPTI, it is not inconsistent with experimental observations.[25] In the simulation run in chloroform, the flexibility has also reached a stable value (0.50–0.55) at 70–100 ps. At a time of 102 ps, the temperature of the system was increased from 300 to 310 K. After an equilibration period the flexibility again appears to have converged to 0.60–0.65 at 245 ps. Because of the somewhat abrupt manner in which the flexibility descended during the second equilibration period, an additional 24 ps of sampling was carried out at 310 °C to ensure that convergence had in fact been achieved (data not reported previously). These additional data from 246 to 270 ps demonstrate that the flexibility has reached a stable value of 0.60–0.65.

These new data and this analysis only serve to reinforce the conclusions presented previously. At almost all points, the flexibility of BPTI in water is greater than that in chloroform. Furthermore, for the brief period during which flexibility is greater in chloroform than in water, it is only marginally larger and this is during an equilibration period for BPTI in chloroform; therefore the data are not representative of a protein in equilibrium with its surroundings. Again, it is very difficult to determine whether the observed loss of conformational flexibility is sufficient to give rise to the experimentally observed effects.

Radius of Gyration. The radius of gyration (r_{gyr}) provides insight into the relative "size" of a protein. For a simple protein, assuming roughly spherical shape, changes in the volume and the surface area should be approximately proportional to changes in $(r_{gyr})^3$ and $(r_{gyr})^2$, respectively. Instantaneous values of r_{gyr} were calculated and averaged for BPTI in water and in chloroform (Table I). The value calculated for the aqueous simulation is consistent with the experimental structure and with the results of a previous simulation of BPTI in water.[25,27,28] It is evident from these results that the protein is "smaller" and more compact in chloroform than it is in the aqueous environment. The observed changes in r_{gyr} correspond to a 10% decrease in surface area and a 15% reduction in volume of the protein. The predicted change in surface area is in rough accord with the results of numerical calculations of surface area (vide infra).

It is tempting to ascribe the changes in r_{gyr} to the folding back of the charged side chains onto the protein surface (vide infra).[26] However, in their study of hydrated myoglobin vs myoglobin *in vacuo*, Steinbach *et al.* showed that the majority of the changes in r_{gyr} can be accounted for by a general swelling of the protein backbone upon hydration.[37] One can of course argue that this swelling is due to the charged side chains pulling the backbone with them as they are extended into the polar solvent.

A correlation between the distance of atoms from the center of mass (r_{gyr}) and the RMS fluctuations in the positions of these atoms has been observed previously.[38,39] Frauenfelder and Petsko noted that in crystals of sperm whale myoglobin, the mean-square displacement increased with distance from the center of gravity.[38] Kuczera and co-workers also observed that large dihedral fluctuations could be correlated with accessible surface area which, as noted above, indirectly implicates the radius of gyration.[39] Thus, the observed changes in the radius of gyration lend support to the claim that changing the surrounding medium causes changes in protein flexibility.

Surface Area. Solvent accessible surface area (SASA) provides another indicator of how the surrounding medium affects protein structure and dynamics. In both water and chloroform SASA appears to be relatively stable throughout the sampling period. Furthermore, SASA seems to be little affected by the addition of thermal energy in the chloroform simulation. Throughout all sampling periods SASA is significantly higher in water than in chloroform (Table I). The difference roughly corresponds to the difference expected based upon the radii of gyration (vida supra).

Examined on a residue-by-residue basis, the difference in SASA is not uniformly distributed throughout the protein (Figure 2). As might have been expected on the basis of solvation arguments, the largest increases in SASA, upon moving BPTI from chloroform to water, are observed for charged and polar side chains. Of the nine greatest observed increases in SASA, eight are either for residues bearing charged side chains (ARG, LYS, GLU, and ASP) or for the charged N and C terminal residues. The lone exception is TYR10, which also shows a significantly larger value of SASA in water relative to chloroform.

A substantially smaller number of residues show meaningful increases in SASA upon moving BPTI from water to chloroform. As expected, the largest increase is seen for a neutral hydrophobic residue, LEU6. Surprisingly, the second largest increase is observed for ARG39. Although it is tempting to rationalize the second observation by noting that ARG39 is in extremely close contact with one of the six chloride counterions included in the simulation, it must be pointed out that several other positively charged residues also enjoy close contacts with counterions. The other positively charged residues do not exhibit similar behavior. In fact, LYS41 is in contact with the same chloride ion as ARG39 and it shows a slight decrease in SASA upon moving BPTI from water to chloroform. Given the intricate network of salt bridges and hydrogen bonds that exists on the surface of BPTI in chloroform, no single factor can be offered to explain the anomalous behavior of ARG39. However, the general trend is clear, that residues showing the largest decrease in SASA upon transfer from water to chloroform are those bearing charged side chains.

Although a general correlation is observed between differences in SASA of BPTI and differences in flexibility of BPTI in water and chloroform, the same trend is not readily apparent on a residue basis. While attempts to correlate differences in these two properties may yield an identifiable trend to the optimistic observer, it must be noted that the correlation coefficient is a meager 0.20.

Hydrogen Bonding Interactions. Hydrogen bond analyses were carried out on the structures saved during the sampling periods of the dynamics trajectories. The analysis was not carried out for the chloroform simulation run at 300 K. The criteria for evaluating the presence or absence of a hydrogen bond were solely geometric. All polar hydrogens were identified and distances to potential hydrogen bond acceptors were calculated. If the hydrogen atom–acceptor atom distance was less than 2.6 Å and the angle between the hydrogen bond donor–hydrogen atom bond and the hydrogen bond acceptor–hydrogen atom bond was between 120° and 180°, then a hydrogen bond was considered to exist (criteria used are similar to those used by Tirado-Rives and Jorgensen[40]). The percentage of time during which the hydrogen bond criteria were met was then evaluated. The percentages for

(37) Steinbach, P. J.; Hodoscek, M.; Brooks, B. R. Modeling Solvation Effects on Protein Dynamics. *Biophys. J.* **1993**, *64*, A183.

(38) Frauenfelder, H.; Petsko, G. A. Structural Dynamics of Liganded Myoglobin. *Biophys. J.* **1980**, *32*, 465–483.

(39) Kuczera, K.; Kuriyan, J.; Karplus, M. Temperature Dependence of the Structure and Dynamics of Myoglobin: A Simulation Approach. *J. Mol. Biol.* **1990**, *213*, 351–373.

(40) Tirado-Rives, J.; Jorgensen, W. L. Molecular Dynamics of Protein with the OPLS Potential Functions. Simulation of the Third Domain of Silver Pheasant Ovomucoid in Water. *J. Am. Chem. Soc.* **1990**, *112*, 2773–2781.

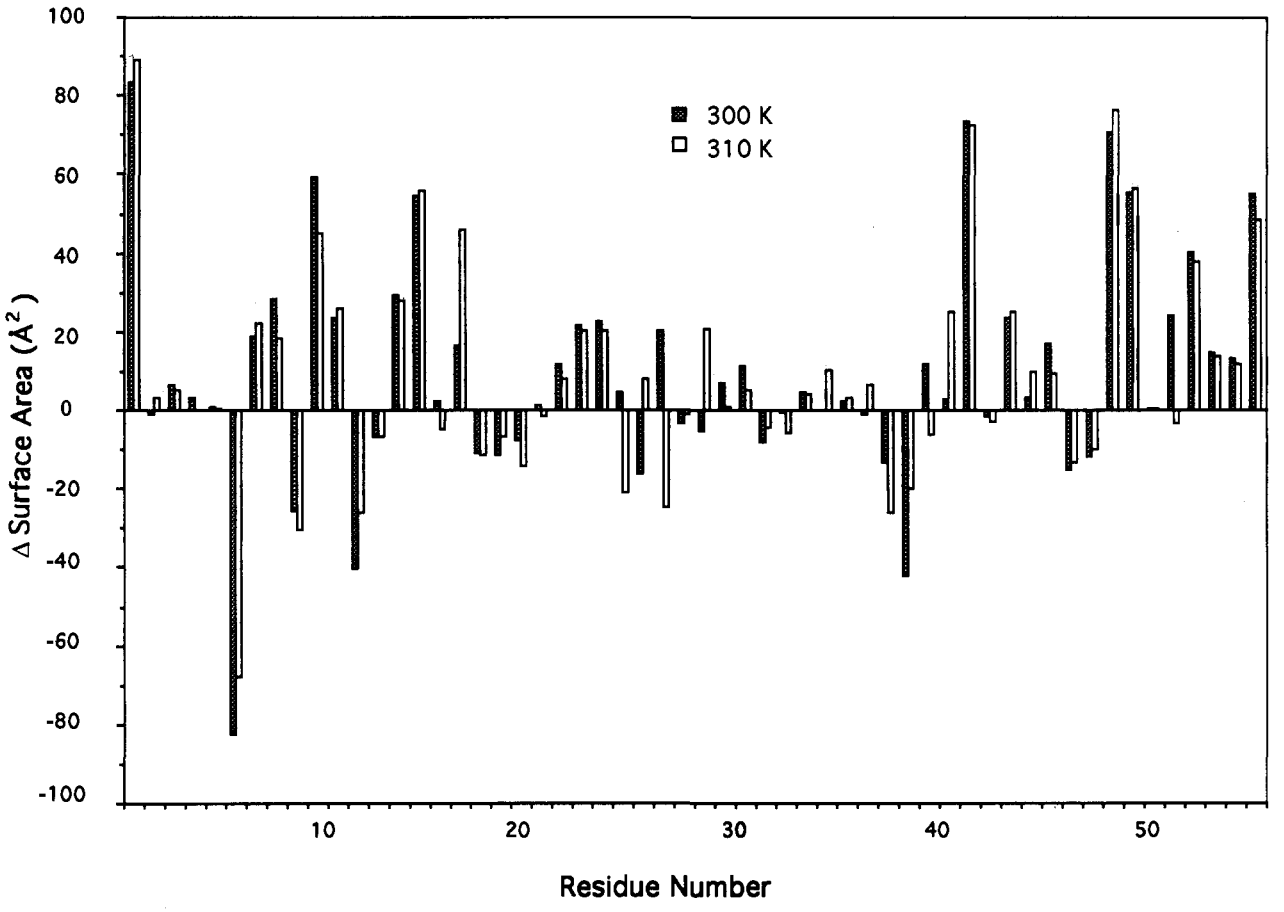

Figure 2. Plot of the differential solvent accesible surface area ($water_{SASA}$ – $chloroform_{SASA}$) as a function of residue number.

all hydrogen bonds between a given pair of residues were then gathered into a single value. Because multiple hydrogen bond contacts exist between some residue pairs, hydrogen bond percentages greater than 100% were found. These values were rounded down to 100%.

The results of the hydrogen bond analyses are displayed in Figure 3 along with a similar analysis of the crystal structure. With few exceptions, all residue pairs which are hydrogen bonded in the crystal structure are also hydrogen bonded in the dynamics simulations. In both water and chloroform, a greater degree of hydrogen bonding is seen than is observed in the crystal structure. In the crystal analysis, the ill defined left to right upwardly diagonal series of points corresponds to the α-helical regions of BPTI. The much more prominent left to right downwardly diagonal series of points corresponds to the β-sheet secondary structure. In water, the hydrogen bonding pattern of the C-terminal α-helical region appears to be disrupted relative to the crystal. Unlike the prior simulation of BPTI in water carried out by Levitt and Sharon, many hydrogen bond contacts not observed in the experimental crystal structure were observed in the water simulation.[25]

A much greater degree of hydrogen bonding was observed in the simulations run in chloroform. Only 5 pairs of residues hydrogen bonded in the crystal are not observed to be bonded in the chloroform simulation. This is much less than the 10 pairs of bonded residues absent in the water simulation. In addition to the hydrogen bonds present in the crystal structure, many new hydrogen bond contacts are observed in the dynamics simulation run in chloroform. The entire surface of the protein is cross-linked by hydrogen bonds between charged side chains and other side chains or backbone atoms. A particularly striking example of this cross-linking network is shown in Figure 4A. For comparison, the same portion of the protein from the simulation run in water is shown in Figure 4B. It is readily apparent that many more intramolecular hydrogen bonding/electrostatic interactions are present in chloroform than in water.

Of particular importance is the fact that the hydrogen bond linkages are binding together regions of the protein very distant from each other in the primary sequence. Because of these linkages, large-scale protein motions will be significantly inhibited. Although the time scales required for such motions are well beyond our capacity to simulate, the structures observed assure that they will be damped by placement of the protein in a nonaqueous environment. The increased hydrogen bond network may thus be the key to many of the remarkable experimental properties of proteins in organic solutions.

Secondary Structure. Secondary structure analysis was carried out on each of the saved instantaneous structures. Each residue was examined and assigned as α–helix, β–sheet, or neither based solely on the values of ϕ and ψ dihedral angles. The geometries of adjacent residues were not considered in making the assignments of secondary structure. The structural requirements for assignment as an α–helical residue were $-100° < \phi < -30°$ and $-80° < \psi < -5°$. To be assigned as β-sheet the criteria used were $-169° < \phi < -89°$ and $83° > \psi > 165°$.[41] All residues that satisfied neither set of criteria were assigned as random coil. The results of this analysis are collected in Table II.

The vast majority of the secondary structural characteristics of the BPTI crystal structure are retained in the simulations run in both water and chloroform. In water, as noted already, the C-terminal α-helical region shows some evidence of unraveling. Although many of the residues are in a helical conformation >60% of the time, three residues in this region (50, 51, and 53) adopt helical structures in <40% of the saved structures. The regions of β-sheet structure are greatly expanded in the water

(41) Daggett, V.; Levitt, M. A Molecular Dynamics Simulation of the C-terminal Fragment of L7/L12 Ribosomal Protein in Solution. *Chem. Phys.* **1991**, *158*, 501–512.

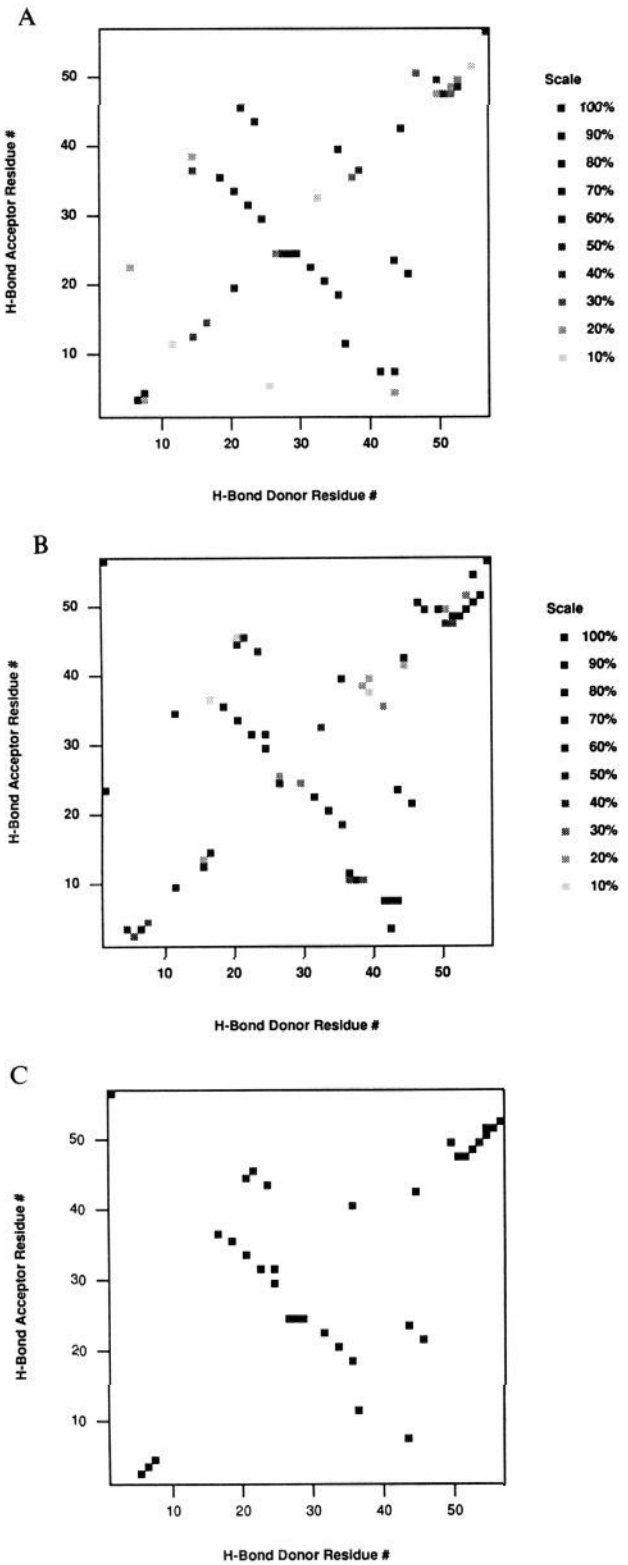

Figure 3. Plot of the frequency of intramolecular hydrogen bond interactions of BPTI in water (A) and at 310 K for chloroform (B) and crystal (C). Hydrogens were added in standard positions to the heavy atoms of the crystal with use of the EDIT module of AMBER 4.0.

simulations relative to the crystal structure. Except for a three-residue segment showing helical structure (25, 26 and 27), almost all residues in the segment from 14 to 35 possess β-sheet type geometries a significant amount of the time.

Much of the secondary structure observed in the crystal is also retained in the simulations run in chloroform. Unlike the structures observed in water, there is little evidence of unwinding of the C-terminal helix for BPTI in chloroform. Several of the residues in this segment possess helical geometries >70% of the time. While the short N-terminal helical region also remains largely intact in chloroform, the three-residue helical segment observed in both the crystal and the water simulations (25, 26 and 27) has completely vanished in the chloroform simulation. The long stretch of β-sheet residues (14–35), however, remains for the most part intact.

Another means of evaluating secondary structural elements is via Ramachandran plots. Ramachandran plots were generated for the average structures of BPTI in each of the solvent environments and for the experimentally determined crystal structure (Figure 5). Consistent with the prior anlaysis of secondary structure, these plots show that structural elements are well-retained in both solvents. No major deviations are observed for the plots in the vastly different aqueous and nonaqueous solvent environments.

Solvent Structure and Dynamics. In much the same way that the solvent environment effects the structure and dynamics of the protein, it is reasonable to expect the protein to affect the structure and dynamics of the solvent. In an attempt to probe the effects of the protein on the solvent, we have examined the diffusion of the solvent and ions and calculated radial distributions of solvent molecules around charged and neutral side chains.

Diffusion. The self-diffusion coefficients of water and chloroform were calculated from the squared displacements during the 48 ps sampling interval by using the Einstein relation. All possible time intervals, Δt, up to 24 ps were used and the results averaged over all starting times. The values of the diffusion coefficient were obtained from the slope of the mean-squared displacements vs Δt over the Δt interval from 10 to 20 ps. To assess the effect of the protein on the solvent dynamics, the solvent was partitioned into shells of 1 Å thickness. The distance from solvent to the nearest protein atom at the beginning of the interval Δt was calculated and the solvent molecule assigned to the corresponding shell. Because solvent molecules can drift from one shell to another during the sampling period, some uncertainty is introduced into the diffusion coefficients.

As anticipated, solvent mobility is strongly correlated with protein-to-solvent distance (Figure 6). The calculated values of the self-diffusion coefficient of water tend to be higher than those seen in previous simulations,[25,40,42] but the overall trend is similar to that seen in previous investigations. This reduction of water mobility near the protein surface has been attributed to both polar water–protein interactions and the formation of clathrate-like structure around nonpolar groups.[43,44] Somewhat unexpected is that the protein affects solvent diffusion in a roughly equal manner for both water and chloroform. In water, the two solvent shells nearest to the protein have diffusion coefficients that are 63% of the values found in the "bulk" solvent. In chloroform, the corresponding ratio is only 54%. It is difficult to identify any specific protein–solvent interaction that would result in drastically reduced mobility for the chloroform solvent. Although dipole–dipole and dipole–charge interactions between chloroform and protein will certainly exist, they should be approximately 50% weaker than the corresponding interactions in water. A less specific mechanism for this result is a generalized effect due to the protein inhibiting solvent motion in certain directions by simply blocking off regions of space. Water, however, does clearly engage in some specific contacts with the protein surface. The waters included in the simulations run in a chloroform box are located

(42) Wong, C. F.; McCammon, J. A. Computer Simulation and the Design of New Biological Molecules. *Isr. J. Chem.* **1986**, *27*, 211–215.
(43) Rossky, P. J.; Karplus, M. Solvation. A Molecular Dynamics Study of a Dipeptide in Water. *J. Am. Chem. Soc.* **1979**, *101*, 1913–1937.
(44) Brooks, C. L., III; Karplus, M. Solvent Effects on Protein Motion and Proein Effects on Solvent Motion. Dynamics of the Active Site Region of Lysozyme. *J. Mol. Biol.* **1989**, *208*, 159–181.

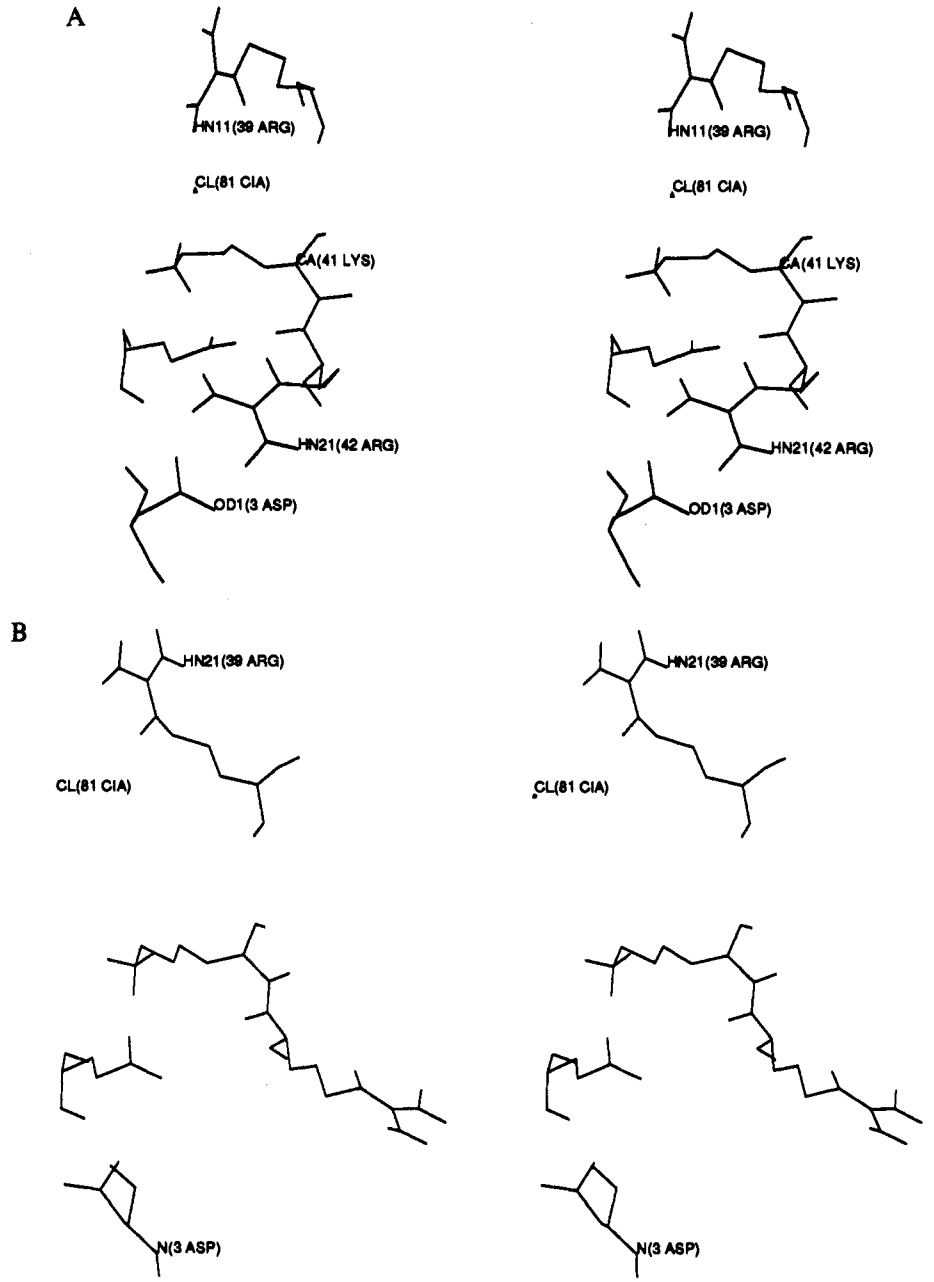

Figure 4. Stereoplots of the network of hydrogen bond/electrostatic interactions for residues ASP3, ARG42, GLU7, LYS41, chloride counterion, and ARG39: chloroform (A) and water (B). Solvent and all other atoms are ommitted for clarity.

primarily around charged groups due to favorable charge–partial charge interactions. The diffusion coefficient for these water molecules is only 9% of the value found for the "bulk" water cited previously (Table III). All of the waters in chloroform remain tightly bound to the protein throughout the trajectory.

Not surprisingly, ion mobility is also decreased in chloroform relative to aqueous solution. The mobility of the chloride counterions in chloroform is 56% of that found in water. This is reasonable given the ability of the ions to be solvated by water, while they are constrained to locations near the protein due to weak electrostatic interactions with chloroform.

Solvent Radial Distribution. Histograms of protein atom to solvent distance corrected for the volume of the radial shell were plotted for several residues of BPTI. The residues chosen displayed a large solvent accessible surface area in one or both of the solvent environments or a large difference in accessible surface between the two environments. Plots of this type were generated for both charged and uncharged residues (ARG1@Cζ, ARG1@protonated N-terminus, ASP3@Cγ, LEU6@Cγ, GLU7@Cδ, LYS15@Nζ, ARG17@Cζ, ILE18@Cδ, ILE19@Cδ, LYS26@Nζ, LEU29@Cγ, ARG39@Cζ, ARG42@Cζ, GLU49@Cδ, ASP50@Cγ, GLY56@deprotonated C-terminus).

In the simulation run in water, structured solvent was observed around the charged side chains. A sharp peak corresponding to the first solvation shell was evident for both positively and negatively charged groups. Also as expected on the basis of previous simulations of protein in water,[44,45] disordered solvent was found around hydrophobic side chains. In only two cases (LEU29 and LEU6) can any gross structure be detected around the hydrophobic groups.

In only a few cases does chloroform show distinct structure around a protein atom. One of the cases where clear solvent structure is observed is around ASP3@Cγ (see Figure 7). In this instance a clear first solvation shell is evident for chloroform. Some degree of structure is also observed around the other negatively charged residues examined. In almost all other cases little or no solvent structure is observed. One contributing factor

(45) Kitchen, D. B.; Reed, L. H.; Levy, R. M. Molecular Dynamics Simulation of Solvated Protein at High Pressure. *Biochemistry* **1992**, *31*, 10083–10093.

Table II. Secondary Structure Analysis

		chloroform, 300 K		chloroform, 310 K		water	
residue	crystal	% α-helix	% β-sheet	% α-helix	% β-sheet	% α-helix	% β-sheet
2		0.0	0.0	0.0	0.0	0.0	1.2
3	α	96.7	0.0	100.0	0.0	79.7	0.0
4	α	52.1	0.0	51.3	0.0	99.5	0.0
5	α	81.0	0.0	58.2	0.0	49.9	0.0
6		17.3	0.5	0.0	0.0	46.5	0.0
7		0.0	15.9	0.0	10.9	0.0	21.5
8		0.0	0.2	0.0	0.5	0.0	0.3
9		0.0	0.2	0.0	0.8	0.0	0.5
10	β	0.0	4.1	0.0	6.9	0.0	24.2
11	α	61.6	0.0	0.0	68.6	88.3	0.0
12		0.0	0.0	0.0	0.0	0.0	0.0
13	α	6.7	0.0	94.7	0.0	28.4	0.0
14		0.0	21.7	0.0	16.9	0.0	23.2
15		0.0	0.2	0.0	0.2	0.0	5.1
16		0.0	50.4	0.0	28.7	0.0	63.5
17		0.0	8.3	0.0	20.6	0.0	30.9
18		0.0	30.1	0.0	24.2	0.0	34.6
19		0.0	35.3	0.0	54.3	0.0	17.9
20	β	0.0	3.9	0.0	5.0	0.0	46.5
21	β	0.0	75.7	0.0	47.7	0.0	80.2
22	β	0.0	79.6	0.0	28.9	0.0	81.9
23		0.0	11.9	0.0	14.5	0.0	39.5
24	β	0.0	39.5	0.0	71.5	0.0	24.3
25	α	0.0	1.1	0.0	0.3	74.9	0.0
26	α	0.0	0.0	0.0	0.0	96.7	0.0
27	α	12.0	0.0	0.0	0.2	86.6	0.0
28		0.0	0.0	0.0	0.0	0.0	0.0
29		0.0	97.7	0.0	87.1	0.0	96.3
30		0.0	14.8	0.0	43.4	0.0	24.3
31		0.0	78.5	0.0	93.3	0.0	69.0
32		0.0	13.6	0.0	5.0	0.0	31.7
33	β	0.0	60.1	0.0	57.4	0.0	77.7
34		0.0	12.3	0.0	10.0	0.0	22.0
35	β	0.0	10.3	0.0	52.6	0.0	33.2
36	α	45.9	0.0	83.8	0.0	11.2	0.6
37		0.0	0.0	0.0	0.0	0.0	0.0
38	β	0.0	72.4	61.3	0.0	0.0	53.8
39		0.0	0.0	0.0	40.3	0.0	0.0
40		0.0	91.9	0.0	13.4	0.0	93.1
41		0.0	4.5	0.0	43.5	0.0	10.1
42	α	32.4	0.0	56.9	0.0	91.4	0.0
43		0.0	2.7	0.0	0.2	0.0	0.3
44	β	0.0	6.2	0.0	24.2	0.0	52.7
45	β	0.0	22.9	0.0	43.7	0.0	59.6
46	α	72.9	0.0	76.9	0.0	89.5	0.0
47	β	0.0	2.0	0.0	0.6	0.0	10.6
48	α	99.8	0.0	98.8	0.0	67.6	0.0
49	α	56.3	0.0	36.5	0.0	83.5	0.0
50	α	98.1	0.0	98.6	0.0	33.1	11.5
51	α	95.6	0.0	88.3	0.0	34.9	0.0
52	α	8.6	0.0	48.7	0.0	94.7	0.0
53	α	63.7	0.0	73.5	0.0	0.0	3.3
54	α	98.4	0.0	96.6	0.0	25.9	0.0
55		0.0	93.0	0.0	87.7	5.8	5.9

to this lack of solvent structure around charged groups is the presence of water molecules, which are better able to solvate the charge centers. Further stabilizing positively charged groups are the counterions which remain in close contact with the positively charged residues. Finally, the formation of salt bridges by the side chains with other regions of the protein surface and the burial of charged groups away from the solvent environment effectively reduce the opportunities for regions of ordered solvent in chloroform. As was observed in water, predominantly disordered solvent is seen around nonpolar hydrophobic residues.

Summary

The results presented clearly demonstrate the profound effect solvent environment can have on protein structure and dynamics. In a polar aqueous environment protein–solvent interactions can successfully compete with protein–protein interactions. This is primarily true because of the high degree of charge separation present in the water molecule. The large partial charges on the hydrogens and the oxygen of water present the opportunity for large magnitude electrostatic interactions with the protein. By contrast, there is much less charge separation in the chloroform molecule and the advantageous protein–solvent interactions are of a smaller magnitude. Accordingly, in a nonaqueous environment the protein must seek out other sources of electrostatic stabilization.

In chloroform the primary sources of electrostatic stabilization available to charged side chains are the chloride counterions and other functional groups on the surface of the protein. The reduced mobility of the chloride ions is indicative of how devoid of favorable solvent–charge interactions the chloroform–protein system is. In water, however, both chloride and protein are afforded more favorable interactions with the solvent than with each other. The conversion of protein–solvent interactions into protein–protein interactions has a substantial effect on protein properties.

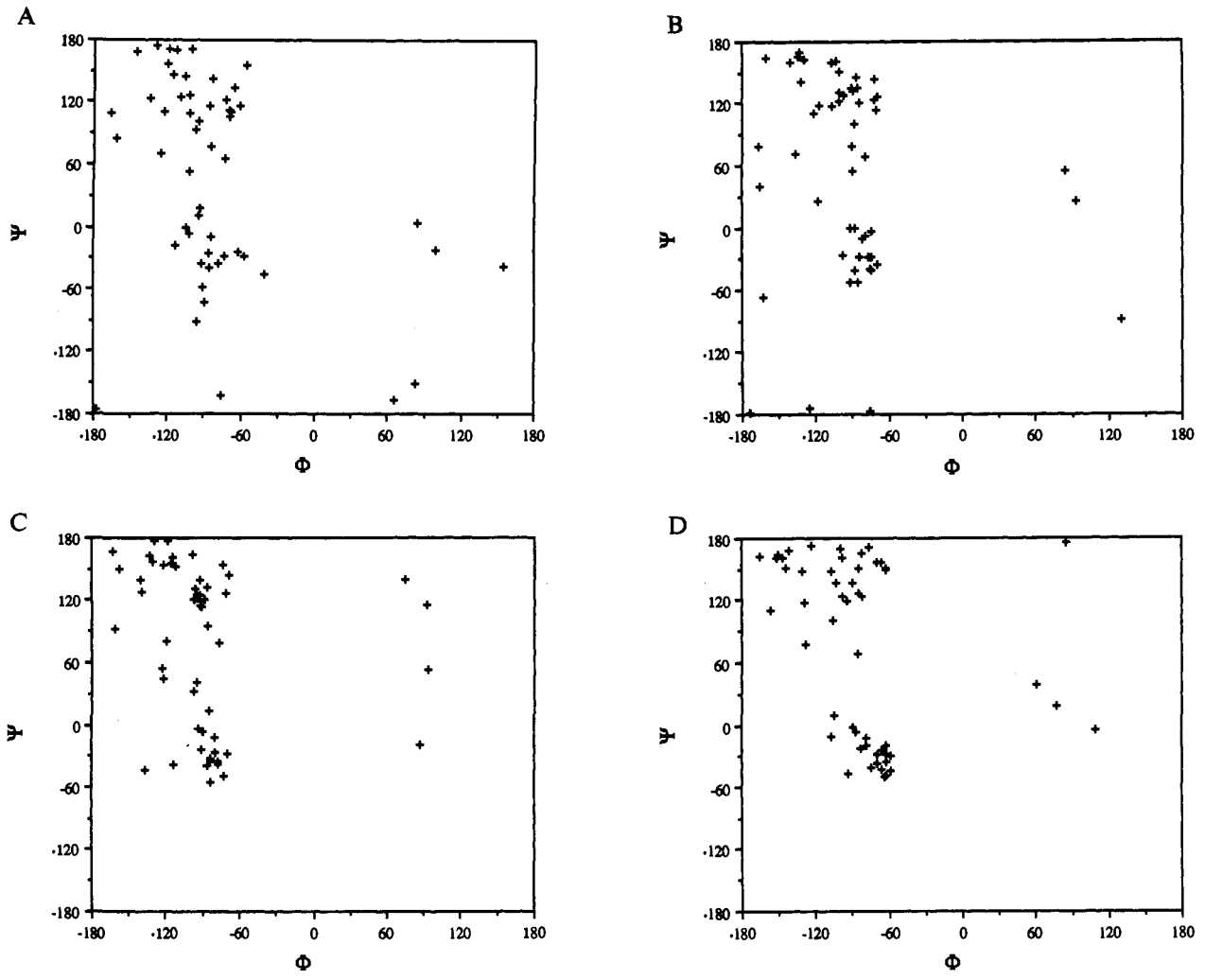

Figure 5. Ramachandran maps for the MD average structures of BPTI at 300 K for chloroform (A) and at 310 K for chloroform (B), water (C), and the experimental crystal structure (D).

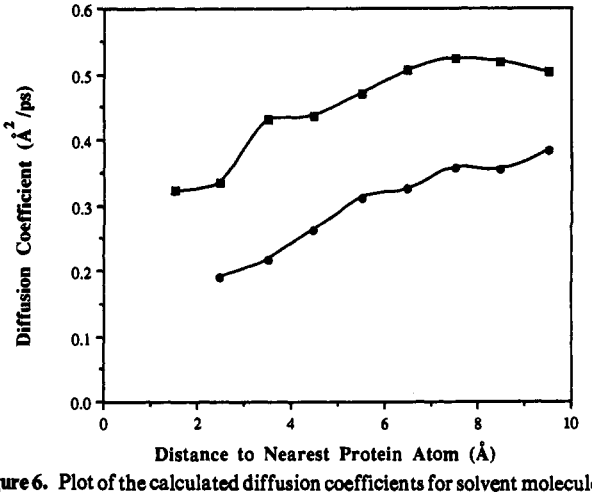

Figure 6. Plot of the calculated diffusion coefficients for solvent molecules as a function of starting distance from protein: water (■); 310 K, chloroform (●).

When protein–solvent contacts become protein–protein contacts, the protein decreases in size and increases in stability. The increase in thermal stability can be directly attributed to the cross-linking effect of the protein–protein interaction. Although these cross-links do not provide the same stability as would be engendered by chemically bonded cross-links, the strength of the electrostatic interactions will be substantial and not easily replaced by the markedly weaker protein–nonpolar solvent interactions. No compelling evidence was found indicating strong preferential

Table III. Diffusion Coefficient

diffusing molecule	diffusion coeff (Å²/ps)
water in chloroform	0.050
chloride in chloroform	0.046
chloride in water	0.083

interactions between nonpolar side chains and the low dielectric solvent environment. Although some distortions of protein structure are observed, as indicated by the increased solvent accessible surface area of nonpolar groups in chloroform, the energetic consequences of these nonpolar side chain–low dielectric solvent contacts will be much less than the protein–solvent interactions of charged groups in water.

The increased network of protein–protein interactions (primarily electrostatic and hydrogen bonding in nature) should reduce the ability of the protein to unfold or undergo significant hinging motions. While processes of this type occur on time scales much too long for computational simulation, the results of the short time scale simulations provide compelling evidence that such motions will be inhibited by a nonaqueous environment. On the shorter time scale a global loss of flexibility is observed, yet it is limited to the terminal residues. However, one must question the functional significance of such short term motions occurring during the picosecond regime examined. Instead, it is almost certainly the case that any experimental manifestations of enzyme rigidity are due to the increased hydrogen bonding network we observe in organic media.[2,7,9,11] The consequences of this intricate cross-linking network of hydrogen bonds are manifested in several

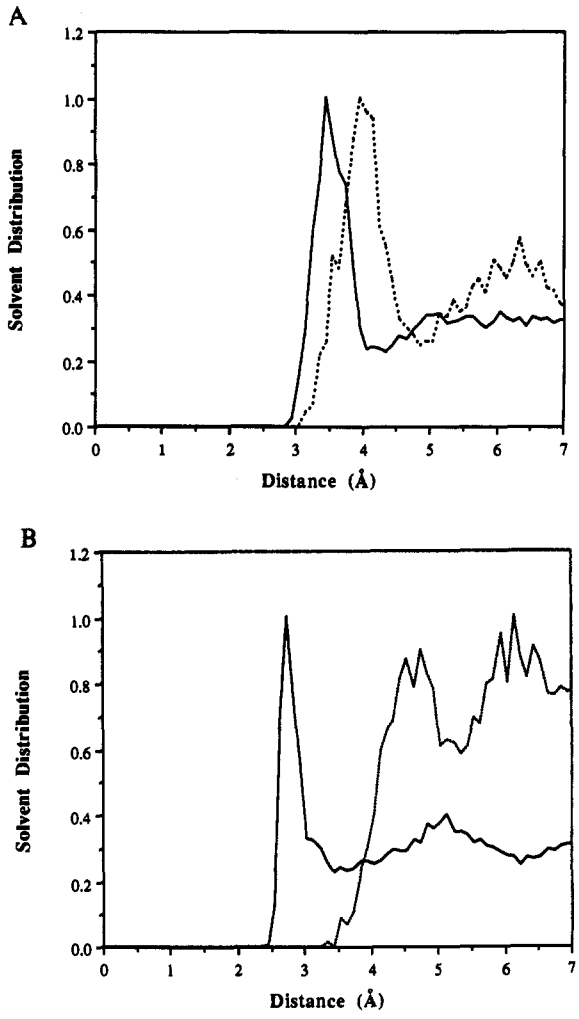

Figure 7. Histograms of solvent density around selected protein atoms. Plots normalized to a maximum of 1.00. Solvent distribution around Cδ of ASP3 (A) at 310 K for chloroform (dashed line) and water (solid line). Solvent distribution around Nζ of LYS26 (B) at 310 K for chloroform (dashed line) and water (solid line).

physically meaningful characteristics we have examined and substantiate the conclusions presented.

Use of these principles should enable design of proteins with even greater thermostability in nonaqueous environments. Previously it was suggested that protein stability may be increased in nonpolar solvents by mutation of polar surface residues into hydrophobic groups.[46-49] While this change would certainly aid in the partitioning of the protein into the nonaqueous phase, it is not clear that such a mutation would increase the stability of the folded state relative to the unfolded state. Furthermore, it has been suggested that mutations which result in the creation of salt bridges on the protein surface will destabilize the folded state in a nonaqueous environment.[47] Instead, we propose that in almost all cases creation of salt bridges will preferentially stabilize the folded state. Specifically, if a pair of nonpolar residues, which are distant in the primary sequence yet spatially proximate in the native state, are mutated to a positive/negative ion pair, a substantial increase in stability of the folded state should result. In the fully solvent exposed unfolded state, each of the charged residues will be relatively unstable in the nonpolar medium. Conversely, in the folded state an energetically favorable salt bridge will be formed by the ion pair. This new interaction will provide not only additional thermodynamic stability for the folded state but also an increase in the kinetic barrier to thermally induced unfolding. Computational simulations are currently underway to test these ideas, and we eagerly anticipate the results of experimental studies designed to test this proposal.

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 Mr. Clayton Springer, formerly of this department, for allowing us to use his modified version of the EDIT module of AMBER for setting up the nonaqueous simulations. Finally, we would like to acknowledge the many enlightening discussions with Dr. Gregory Farber of this department.

(46) Arnold, F. H. Protein Design for Nonaqueous Solvents. *Protein Eng.* **1988**, *2*, 21–25.
(47) Arnold, F. H. Engineering Enzymes for Nonaqueous Solvents. *TIBTECH* **1990**, *8*, 244–255.
(48) Martinez, P.; Arnold, F. H. Surface Charge Substitutions Increase the Stability of α-Lytic Protease in Organic Solvents. *J. Am. Chem. Soc.* **1991**, *113*, 6336–6337.
(49) Chen, K.; Robinson, A. C.; Van Dam, M. E.; Martinez, P.; Economou, C.; Arnold, F. H. Enzyme Engineering for Nonaqueous Solvents. II. Additive Effects of Mutations on the Stability and Activity of Subtilisin E in Polar Organic Media. *Biotechnol. Prog.* **1991**, *7*, 125–129.