Solvation and Dynamics of Chymotrypsin in Hexane

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

Contribution from the Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

Received January 17, 1996[⊗]

Abstract: We have investigated the serine protease γ-chymotrypsin (γ-CT) in three different solvation environments using molecular dynamics simulations. These solvation environments include the following: (1) γ-CT taken from the crystal structure of Yennawar *et al*. (*Biochemistry* **1994**, *33*, 7326−7336) with seven surface bound hexane molecules and 50 bound "essential" water molecules all immersed in 1109 hexane molecules (simulation labeled CT); (2) γ-CT taken from Yennawar *et al*. and solvated with 50 "essential" water molecules and immersed in 1107 hexane molecules (simulation labeled CTWAT); and (3) γ-CT taken from Yennawar *et al*. and solvated with a monolayer of 444 water molecules and immersed in 931 hexane molecules (simulation labeled CTMONO). From these trajectories we found that the placement of bound water molecules and the amount of hydration of the protein in the simulated structure had an effect on the protein flexibility as indicated by changes in the RMS deviation. The radius of gyration value was similar for the three systems, indicating no significant unfolding or denaturation. Hydrophobic residues were found to have increased solvent accessible surface area (SASA), while hydrophilic residues experienced a decrease in SASA in the CT and CTWAT simulations. No hexane diffusion into the protein interior was found, with the exception of one bound hexane site in the CT simulation. The secondary structure analysis of the active site indicates that the active site structure was retained in all three simulations. We also found that intramolecular forces (*i.e*., hydrogen bonding) that stabilize proteins are stronger in the CT and CTWAT systems, as shown by the increase in the number of stable hydrogen bonds found. Net ion pair interactions and reduced ratio of surface area to volume of the protein also contributed to the stability of the protein in anhydrous organic media.

Introduction

The suspension or dissolution of enzymes into non-aqueous solvents affords many advantages (as well as some disadvantages) over the use of proteins in aqueous solution. For example, Klibanov and co-workers have demonstrated that proteins have increased thermostability,[1] molecular pH memory,[2] and altered substrate specificity[3] when they are placed in an anhydrous organic solvent. Furthermore, the ability of enzymes to catalyze reactions that are either thermodynamically or kinetically impossible in aqueous solution has also been observed.[4−8] This has led to the possibility of widespread use of enzymes as highly specific catalysts in non-aqueous solvents in the synthesis of organic compounds.[9]

The protease family of enzymes has long generated considerable interest due to their role in peptide degradation and possibly peptide synthesis.[10−12] Hence, the catalytic activities of serine proteases have been extensively studied.[13] Several molecular dynamics (MD) studies have been reported on serine proteases in aqueous media.[14−16] However, nothing is known about the dynamics of serine proteases in non-aqueous solvents. Moreover, only a few studies on MD simulations of proteins in non-aqueous solvation have been reported. For example, van Gunsteren and Karplus[17] examined bovine pancreatic trypsin inhibitor (BPTI) in a non-polar solvent while Hartsough and Merz[18,19] compared the structure and dynamics of BPTI in water and chloroform. A simulation of neuropeptide galanin in water

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1996.

(1) Zaks, A.; Klibanov, A. M. Enzymatic Catalysis in Organic Media at 100 °C. *Science* **1984**, *224*, 1249−1251.

(2) Klibanov, A. M. Enzymes That Work in Organic Solvents. *CHEMTECH* **1986**, *63*, 354−359.

(3) Zaks, A.; Klibanov, A. M. Substrate Specificity of Enzymes in Organic Solvents vs Water Is Reversed. *J. Am. Chem. Soc*. **1986**, *108*, 2767−2768.

(4) Zaks, A.; Klibanov, A. M. Enzyme-Catalyzed Processes in Organic Solvents. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3192−3196.

(5) Kuhl, P.; Halling, P. J.; Jakubke, H. D. Chymotrypsin Suspended in Organic Solvents with Salt Hydrates Is a Good Catalyst for Peptide Synthesis from Mainly Undissolved Reactants. *Tetrahedron Lett*. **1990**, *31*, 5213−5216.

(6) Stahl, M.; Mansson, M. O.; Mosbach, K. The Synthesis of a D-Amino Acid Ester in an Organic Media with α-Chymotrypsin Modified by a Bio-Imprinting Procedure. *Biotech. Lett*. **1990**, *12*, 161.

(7) West, J. B.; Hennen, W. J.; Lalonde, J. L.; Bibbs, J. A.; Zhong, Z.; Meyer, E. F., Jr.; Wong, C. H. Enzymes as Synthetic Catalysts: Mechanistic and Active-Site Considerations of Natural and Modified Chymotrypsin. *J. Am. Chem. Soc*. **1990**, *112*, 5313−5320.

(8) Kasche, V.; Michaelis, G.; Galunsky, B. Binding of Organic Solvent Molecules influences the P′1 P′2 Stereo and Sequence-Specificity of α-Chymotrypsin in Kinetically Controlled Peptide Synthesis. *Biotech. Lett*. **1991**, *13*, 75.

(9) Faber, K. Biotransformations in Organic Chemistry. In *Biocatalyts in Non-Conventional Media*; Tramper, J., Vermüe, M. H., Beeftink, H. H., Eds.; Elsevier Science: Amsterdam; Springer-Verlag: Berlin, New York, 1992.

(10) Kullman, W. *Enzymatic Peptide Synthesis*; CRC Press: Boca Raton, FL, 1987.

(11) Schellenberger, V.; Jakubke, H. D. Protease-Catalyzed Kinetically Controlled Peptide Synthesis. *Agnew. Chem., Int. Ed. Engl*. **1991**, *30*, 1437−1449.

(12) Wong, C. H.; Wang, K. T. New Developments in Enzymatic Peptide Synthesis. *Experientia* **1991**, *47*, 1123.

(13) Polgar, L. In *Hydrolytic Enzymes*; Neuberger, A., Brocklehurst, K., Eds.; Elsevier Science Publishers B.V. (Biomedical Division): Amsterdam, 1987; Chapter 3, pp 159−200.

(14) Yu, H.-A.; Karplus, M .; Nakagawa, S.; Umeyama, H. Active Site Dynamics of Acyl-Chymotrypsin. *PROTEINS: Structure, Function and Genetics* **1993**, *16*, 172−194.

(15) Gerig, J. T. Dynamics at the Active Site of *N*,*N*-Bis(2-fluorophenyl)-carbamoyl α-Chymotrypsin. *Magn. Reson. Chem*. **1990**, *28*, 47.

(16) Bemis, G. W.; Carlson-Golab, G.; Katzenellenbogen, J. A. A Molecular Dynamics Study of the Stability of Chymotrypsin Acyl Enzymes. *J. Am. Chem. Soc.* **1992**, *114*, 570−578.

(17) van Gunsteren, W. F.; Karplus, M. Protein Dynamics in Solution and in a Crystalline Environment: A Molecular Dynamics Study. *Biochemistry* **1982**, *21*, 2259−2274.

(18) Hartsough, D. S.; Merz, K. M. J. Protein Flexibility in Aqueous and Nonaqueous Solutions. *J. Am. Chem. Soc.* **1992**, *114*, 10113−10116.

(19) Hartsough, D. S.; Merz, K. M. J. Protein Dynamics and Solvation in Aqueous and Nonaqueous Environments. *J. Am. Chem. Soc*. **1993**, *115*, 6529−6537.

S0002-7863(96)00153-9 CCC: $12.00

and 2,2,2-trifluoroethanol was carried out by De Loof *et al*.,[20] while Lautz *et al*.[21] compared cyclosporin A in water and a nonpolar solvent. Hult *et al*.[22] have recently reported an MD study of a lipase in vacuum, water, and methyl hexanoate. In general, significant changes in the structure of the protein in an organic solvent *versus* water were observed.

Experimental studies have shown that proteins, in order to retain their activity in anhydrous solvents, require some water molecules to be present. These water molecules can be described as "essential water", and it has been suggested that they function as molecular lubricants for the protein.[23] However, an MD simulation has shown that the water hydration shell around the protein myoglobin in water is not a uniform monolayer, but a patchwork of water clusters, covering some charged groups extensively while leaving other regions of the protein surface poorly solvated.[24] In addition, a recent NMR study of chymotrypsin in hexane has shown that water molecules are likely to be localized in the most polar regions and the majority of the enzyme surface will be in direct contact with hexane molecules.[25] Thus, water may not play a global role as a "molecular lubricant" in either aqueous or non-aqueous solvents, but may only facilitate molecular motion in localized regions on the protein surface.

In this study, we report our findings on the structure and dynamics of the serine protease γ-chymotrypsin (γ-CT) in hexane using MD simulations. Unlike membrane associated proteases, γ-CT has been designed to function most efficiently in an aqueous solution. A so-called "catalytic triad" consisting of Ser-His-Asp is involved in the hydrolysis of peptide bonds on the C-terminal side of large hydrophobic residues. γ-CT has been suspended in nearly anhydrous solvents and has been found to be catalytically active.[26] Several crystal structures of chymotrypsin have been reported and recently a structural study of a γ-CT crystal soaked in hexane has been solved.[27] Furthermore, recent X-ray crystallographic studies of γ-CT have provided an alternate view[28] of enzymatic memory.[23,29] The structure of cross-linked subtilisin crystals in non-aqueous solvents has also been solved.[30,31]

Hexane is a suitable non-aqueous solvent choice for enzymatic reactions since it is thought that it does not strip essential water from enzyme surfaces and, therefore, allows enzymes to retain their catalytic activity. Hence, γ-CT in hexane is an ideal model system to explore the effect of non-aqueous solvation on protein structure, function, and dynamics. A molecular-level understanding regarding how proteins are affected by anhydrous organic solvents would further our ability to employ proteins as catalytic agents in organic solvents. Moreover, by comparing the origin of the enhanced thermal stability of proteins in non-aqueous solvents to that of proteins obtained from thermophilic organisms[32,33] we will be able to further our understanding of the origin of protein stability in extreme environments.

Table 1. Summary of the Setup of Starting Configurations[a,b]

molecule type	CT	CTWAT	CTMONO
X-ray water	50	50[c]	50
added water	0	0	394
X-ray hexane	7	0	0
added hexane	1109	1107	931

[a] All protein coordinates are from the hexane crystal (see ref 27). [b] Three chloride ions were added as neutralizing counterions in all cases. [c] The 50 water locations used were taken from the native crystal structure and placed onto the hexane crystal structure.

Computational Methods

Molecular Dynamics (MD) simulations were carried out on γ-CT using three different starting configurations which we have labeled CT, CTWAT, and CTMONO. The details regarding the starting conditions for the three simulations presented herein are given in Table 1. All the simulations were performed using a parallelized version[34] of the MINMD module of AMBER 4.0[35] simulation package in conjunction with the OPLS force field.[36] TIP3P[37] water and a united atom hexane model[38] were used in these simulations. The starting coordinates for all heavy atoms were obtained from the crystal structures of γ-CT determined by Farber and co-workers in hexane.[27]

In all three of the starting configurations, the most tightly bound 50 water molecules (as determined from experimental *B* values) were retained. This amount of water roughly corresponds to that required for hydration of charged groups and has been found to participate in the catalytic activity.[23,39] Three chloride ions were introduced near positively charged groups as counterions to neutralize the protein charge. γ-CT consists of three peptide segments (1−13, 16−146, and 151−245) linked together by disulfide linkages. In the X-ray structure, residues 1−10, 16−146, and 151−245 were observed[27] and it was only these residues that were used in the MD simulations. Furthermore, the pentapeptide, which is present in the active site of γ-CT,[27] was removed in the simulations. The initial configuration for the CT simulation also included the seven hexane molecules that were observed crystallographically.[27] The initial positions for the water molecules in the simulation labeled CTWAT were obtained from the native crystal (*i.e.*, the crystal was not exposed to hexane prior to structure

(20) 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.

(21) 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.

(22) Norin, M.; Edholm, O.; Haeffner, F.; Hult, K. Molecular Dynamics Simulations of an Enzyme Surrounded by Vacuum, Water, or a Hydrophobic Solvent. *Biophys. J.* **1994**, *67*, 548−559.

(23) Klibanov, A. M. Enzymatic Catalysis in Anhydrous Organic Solvents. *TIBS* **1989**, *14*, 141−144.

(24) Steinbach, P. J.; Brooks, B. R. Protein Hydration Elucidated by Molecular Dynamics Simulation. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9135−9139.

(25) Parker, M. C.; Moore, B. D.; Blacker, A. J. Measuring Enzyme Hydration in Non-polar Organic Solvents Using NMR. *Biotech. Bioeng.* **1995**, *46*, 452.

(26) Paradkar, V. M.; Dordick, J. S. Aqueous-like Activity of α-Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents. *J. Am. Chem. Soc.* **1994**, *116*, 5009−5010.

(27) Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. X-ray Crystal Structure of γ-Chymotrypsin in Hexane. *Biochemistry* **1994**, *33*, 7326−7336.

(28) Yennawar, H. P.; Tennawar, N. H.; Farber, G. K. A Structural Explanation for Enzyme Memory in Nonaqueous Solvents. *J. Am. Chem. Soc.* **1995**, *117*, 577−585.

(29) Klibanov, A. M. What is Remembered and Why? *Nature (London)* **1995**, *374*, 596.

(30) Fitzpatrick, P. A.; Steinmetz, A. C.; Ringe, D.; Klibanov, A. M. Enzyme Crystal Structure in a Neat Organic Solvent. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8653−8657.

(31) Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. X-Ray Crystal Structure of Cross-Linked Subtilisin Carlsberg in Water *vs.* Acetonitrile. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675−681.

(32) Rees, D. C.; Chan, M. K.; Mukund, S.; Arnulf, K.; Adams, M. W. W. Structure of a Hyperthermophilic Tungstopterin Enzyme, Aldehyde Ferredoxin Oxidoreductase. *Science* **1995**, *267*, 1463−1469.

(33) Goldman, A. How to Make My Blood Boil. *Structure* **1995**, *3*, 1277−1279.

(34) Vincent, J. J.; Merz, J., K. M. A Highly Portable Parallel Implementation of AMBER 4 Using the Message Passing Interface Standard. *J. Comput. Chem.* **1995**, *16*, 1420−1427.

(35) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. (1991) AMBER 4.0, University of California, San Francisco.

(36) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins: Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.

(37) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for the Simulation of Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(38) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *J. Am. Chem. Soc.* **1984**, *106*, 6638−6646.

(39) Rupley, J. A.; Gratton, E.; Careri, G. Water and Globular Proteins. *TIBS* **1983**, *8*, 18−22.

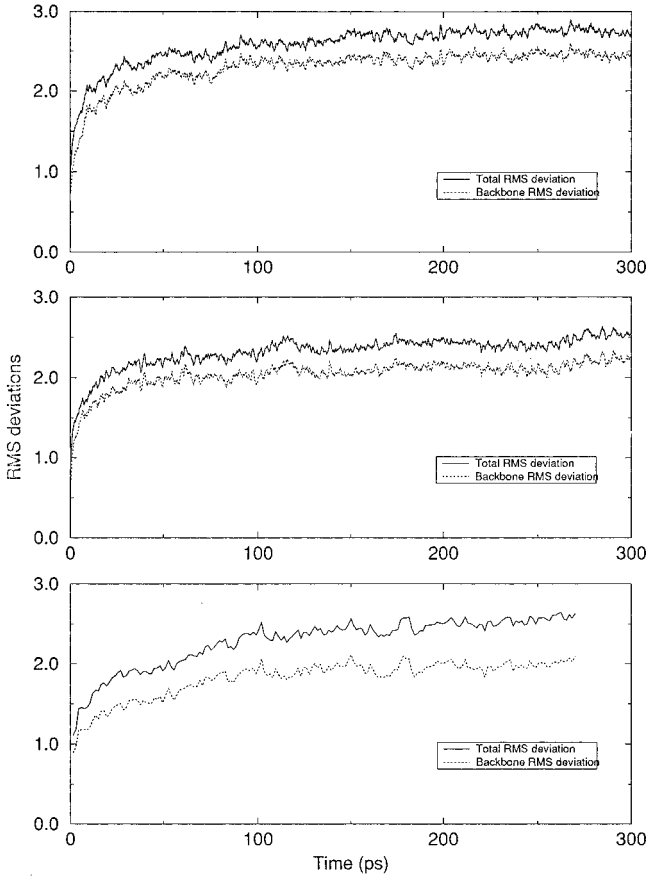

Figure 1. RMS deviation between instantaneous computed structure and the crystal structure for all residues as a function of time. The top panel (a) is for the simulation labeled CT, the middle panel (b) is for CTWAT, and the bottom panel (c) is for CTMONO.

determination) structure of γ-CT[27] and the crystallographically observed hexane molecules were deleted. In the CTMONO simulation these hexane molecules were again ignored and a (mono)layer of water molecules was placed around the protein. The resulting three starting configurations were then solvated in hexane (see Table 1 for total amounts) and subjected to energy minimization for 2000 steps with periodic boundary conditions. In all cases all of the bond lengths were constrained to their equilibrium values using the SHAKE[40] algorithm with 0.0005-Å tolerance.

The minimized systems were equilibrated by gradually increasing the system temperature from 0 to 300 K over a period of ~30 ps by coupling to a temperature bath.[41] A time step of 1.5 fs as well as a constant pressure of 1 atm[41] was used in all the simulations. Both CT and CTWAT simulations were performed for 300 ps (150 ps equilibration and 150 ps sampling), while CTMONO dynamics was run for 270 ps (150 ps equilibration and 120 ps sampling). Coordinates were saved every 50 steps for CT and CTWAT, and every 1000 steps for CTMONO for analysis. All analyses were carried out for the last 150 ps for CT and CTWAT, and the last 120 ps for CTMONO.

Results and Discussion

Root-Mean-Square Deviations. The root-mean-squared (RMS) deviations of the instantaneous structures from the starting crystal structure are plotted as a function of time in Figure 1. The RMS deviations from the crystal are typical of those seen in protein simulations using the OPLS parameter set.[42] The three systems appear to have reached an equilibrium state with respect to RMS deviation by ca. 150 ps for both backbone and total RMS deviations (see Figure 1). The average

Table 2. Summary of the RMS Deviation, Radius of Gyration, SASA, Hydrophobic Exposed SASA, Solvent Diffusion, Average Distances, and Hydrogen Bonding Data from the Simulations

analysis type	CT	CTWAT	CTMONO
RMS Deviation (A)			
total	2.7	2.4	2.5
backbone	2.4	2.1	2.0
Radius of Gyration (A)			
minimized structure	18.0	17.5	17.7
MD average	17.8	17.1	17.3
SASA (10^3 Å^2)			
crystal structure	9.2	9.2	9.2
MD[a] average	8.7	8.7	9.2
Hydrophobic Residues SASA/Total SASA (% exposed residues)			
crystal structure	34.3	34.3	34.3
MD average	39.8	39.1	38.3
Solvent Diffusion (10^5 cm^2/s)			
water	0.4	0.5	3.0
Average Distances[b]			
NE2-HOG	4.2 (2.0)	3.1 (2.0)	4.1 (2.0)
HND-OD1	1.8 (1.8)	1.8 (1.8)	1.7 (1.8)
HND-OD2	2.5 (2.5)	2.5 (2.5)	2.8 (2.5)
Hydrogen Bonding			
crystal (total)	183	183	183
min. structure (total)	216	193	203
total	570	561	605
<10%[c]	216	224	294
>90%[d]	139	141	93

[a] Using the structure obtained from the end of the simulation. [b] Value given in parentheses is from the crystal structure. [c] Number of hydrogen bonds present <10% of the time. [d] Number of hydrogen bonds present >90% of the time.

RMS deviations for the three systems after 150 ps (*i.e.*, the instantaneous RMS was averaged from 150 to 300 ps) are reported in Table 2.

Both CTWAT and CTMONO have lower RMS deviations compared to the CT system, indicating that they have a closer resemblance to the starting crystal structure. In our earlier work on BPTI[18,19] we found that the RMS deviation from the crystal was on average higher for the system containing only water. If we assume that the CTMONO simulation is a better representation of γ-CT in an aqueous environment than CT and CTWAT (a reasonable assumption even though water molecules still tend to cluster in the CTMONO case (see discussion below)), we observe the opposite trend from our earlier work on BPTI.[18,19] Namely, the aqueous simulation experiences on average a smaller RMS deviation from the crystal than do the anhydrous hexane simulations. This point is very clear if we examine the RMS deviation of the backbone atoms only. This is an interesting observation given that the starting structure we used in our simulations came from a crystal structure in which hexane molecules were co-crystalized with γ-CT. Hence, we might expect that the simulations that were designed to represent an anhydrous organic solvent environment would not significantly deviate from the observed crystal structure (this is especially true for the CT run). Thus, one interpretation of our results suggests that the X-ray structure may not give as accurate a picture of the structure of γ-CT in a non-aqueous solvent as one would expect. This is not surprising since it is known that there can be structural differences observed between NMR and X-ray structures for the same protein molecule. Moreover, the X-ray structure data have not demonstrated definitively that the solvent channels between the protein molecules are filled with water or hexane molecules. Our results seem to best support

(40) van Gunsteren, W. F.; Berendsen, H. J. C. Algorithm for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311.

(41) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(42) 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.

the former over the latter. Our results are not definitive in this regard and it is clear that further experimental and simulation studies are needed to clarify these issues. Finally, if we consider RMS deviation to be indicative of protein mobility, we conclude that γ-CT exhibits higher mobility in anhydrous nonpolar organic media than in water. This is in agreement with experimental NMR results that have shown that the flexibility of cytochrome *c* is clearly enhanced in organic medium.[43]

Another interesting aspect of the RMS data has to do with the dependence of the results on the starting configuration. The only change on going from CT to CTWAT was the placement of the 50 "essential" water molecules. In the CT simulation the water coordinates were taken from the "hexane" crystal structure[27] by selecting the water molecules with the lowest *B* factors. In order to test the dependence of our simulated results on the starting configuration we created CTWAT where we placed the 50 essential water molecules based on the lowest *B* factors from the "native" (*i.e.*, in the absence of hexane) crystal structure. The importance of the correct placement of buried water molecules in a simulation model for locally correct structure has been reported.[44] Given the seemingly modest change between CT and CTWAT, we observe that the RMS deviation between the two runs can be quite different over the 300 ps time scale used in the simulations. This leads us to conclude that the placement of the essential water molecules is important for protein mobility and we note that once they are positioned it is very difficult for the water structure to rearrange along the protein surface due to the presence of the low dielectric solvent.

Analyzing the RMS deviations on a per residue basis, we observe thst the first ten-residue peptide of γ-CT experiences the highest RMS deviation for all three runs (see Figure 2). Higher RMS deviations are further observed at the beginning and the end of each peptide segment of the protein (see gaps at 10−16 and 146−151 in Figure 2). In Figure 2a the next largest series of peaks we observe (located at residue number 36−39, 74−78, 125−129, and 202−205) for the CT simulation correspond to loop regions exposed to the solvent. For the CTWAT simulation (see Figure 2b) we again observe significant peaks at the termini of the peptide fragments as well as at 59−63, 202−206, and 215−219, which again correspond to surface loops. Interestingly, we find in the CTMONO simulation that the terminal regions deviate significantly from the crystal structure, but the remaining large peaks do not all correspond to surface loops. The large peaks in CTMONO are located at 60−65 (β-sheet buried region), 70−72 (semiburied region adjacent to a surface loop), 96−99 (buried loop region), 110−114 (surface loop), 126−129 (surface loop), 176−178 (semiburied region adjacent to a surface loop), and 202−206 (surface loop). Thus, the major contributor to the RMS deviation of γ-CT in hexane arises at the peptide termini as well at surface exposed loops. In the monolayer simulation the terminal regions do deviate significantly from the crystal structure, but the surface loops are no longer the sole regions that experience large RMS deviations. The reason for this behavior is not entirely clear, but the surface loops are known to be flexible regions within proteins and when a protein is placed into hexane the surface region undergoes substantial changes in response to the low dielectric solvent. Since the surface loops are flexible these regions can undergo larger rearrangements to satisfy the hydrogen bonding deficit present at the surface due to the presence of the nonpolar solvent. In the monolayer simulation

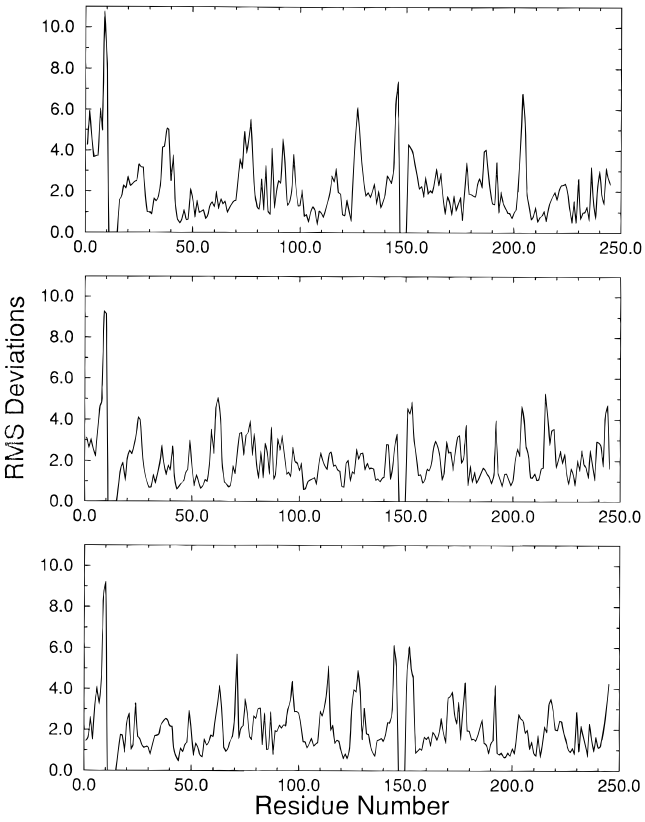

Figure 2. RMS deviation per residue. The top panel (a) is for the simulation labeled CT, the middle panel (b) is for CTWAT, and the bottom panel (c) is for CTMONO.

we do not entirely observe this behavior since there is sufficient water present to satisfy the hydrogen bond interactions of the surface residues. Similar effects of solvent on the conformation of surface exposed loops, which do not have extensive internal hydrogen bonds, have also been observed in other MD simulations.[22,45]

The difference in RMS behavior between predominantly aqueous and non-aqueous environments is quite interesting in light of recent experimental observations of Arnold and coworkers.[46−48] Utilizing a random mutagenesis scheme Chen and Arnold[46] have generated a mutant of subtilisin E that is almost as active in 60% DMF as the native enzyme is in water. The mutant enzyme has had ten positions altered all of which are found in surface loops near the active site of the enzyme. Given that we observe the greatest RMS deviations to occur at surface loops in a non-aqueous solvent we can postulate that the site-directed mutants of Chen and Arnold[46] are stabilizing these loop regions of the enzyme in order to enhance the activity of the enzyme. This is an interesting possibility and it would be useful to see if altering the sequence of loop regions has an effect on the activity of other enzymes in non-aqueous solvents.

The active site residues show lower RMS deviation than the protein's average (see Figure 3). An MD simulation of acylchymotrypsin in water by Nakagawa *et al*. also found the active site triad residues to have smaller fluctuations than the rest of

(43) Wu, J.; Gorenstein, D. G. Structure and Dynamics of Cytochrome *c* in Nonaqueous Solvents by 2D NH-Exchange NMR Spectroscopy. *J. Am. Chem. Soc.* **1993**, *115*, 6843−6850.

(44) Sreenivasan, U.; Axelsen, P. H. Burried Water in Homologous Serine Proteases. *Biochemistry* **1992**, *31*, 12785−12791.

(45) Åqvist, J.; Tapia, O. Molecular Dynamics Simulation of the Solution Structure of the C-Terminal Fragment of L7/L12 Ribosomal Protein. *Biopolymers* **1990**, *30*, 205−209.

(46) Chen, K.; Arnold, F. H. Tuning the Activity of an Enzyme for Unusual Environments: Sequential Random Mutagenesis of Subtilisin E for Catalysis in Dimethylformamide. *Proc. Natl. Acad. Sci.* **1993**, *90*, 5618−5622.

(47) Arnold, F. H. Protein Engineering for Unusual Environments. *Curr. Opin. Biotechnol.* **1993**, *4*, 450−455.

(48) Arnold, F. H. Engineering Enzymes for Nonnatural Environments. *FASEB J.* **1993**, *7*, 744−749.

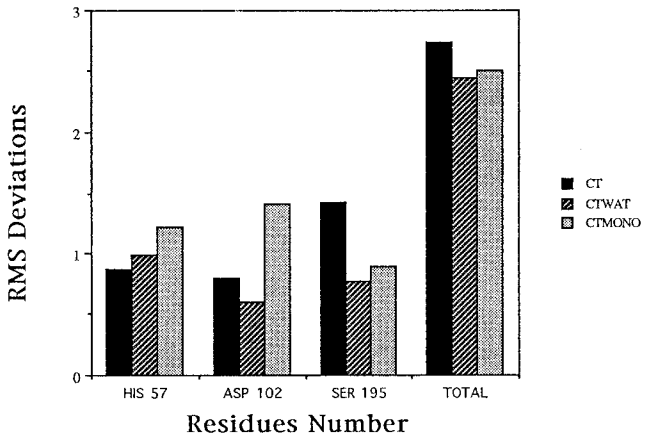

Figure 3. RMS deviations of the active-site residues.

the protein.[14] This is also observed when CT is in an anhydrous non-aqueous environment. Furthermore, this calculated observation is consistent with the experimentally observed structural integrity of the catalytic triad in the serine protease α-lytic protease.[49] Hence, the observed reduced flexibility of the active site side residues in both water and hexane allows the protein to remain catalytically active in both aqueous and non-aqueous environments.

Radius of Gyration. The radius of gyration (r_{gyr}) is defined as the mass-weighted RMS distance of a collection of atoms from their common center of gravity. This provides insight into the relative size of a protein molecule. Assuming that γ-CT is a roughly spherical protein, the expected radius of gyration should fall within 17−18 Å.[50] The radius of gyration values for the three systems were calculated and averaged over the sampling phase of the MD simulations and the results are given in Table 2. The results obtained for all three systems fall within the expected range, with the CT simulation having the largest r_{gyr} and CTWAT having the smallest. This observation is in contrast with the trend observed in our earlier BPTI simulations where we found a general reduction in the r_{gyr} of the protein in the non-aqueous solvent relative to the same protein in water.[18,19] Assuming that the CTMONO simulation is representative of an aqueous environment we observe the opposite trend with respect to the CT run while the CTWAT run follows the trend observed in the BPTI simulations. These observations taken together indicate that the starting configurations for the MD simulations are critical in determining future r_{gyr} behavior and assuming that a protein in a non-aqueous solvent suffers an overall decrease in r_{gyr} is not borne out. Nonetheless, the small differences among the r_{gyr} values suggest that conformational changes are in general modest and that the protein backbone structure is largely unaffected. We still observe that the major motions that occur when a protein is placed in a nonpolar non-aqueous solvent are due to the movements of the side chains, with the hydrophobic side chains orienting out toward the solvent and the hydrophilic residues folding inward toward the protein's interior.

Surface Area, Salt Bridges, and Protein Thermostability. Solvent accessible surface area (SASA) is another indicator of how the surrounding medium affects protein structure and dynamics. In general, the stability of a protein can be increased by minimizing the ratio of the surface area to volume, which simultaneously reduces the unfavorable surface energy and increases the attractive interior packing interactions. A relatively simple indicator to measure the efficiency of packing is given by the fraction of atoms in a protein with no accessible surface area.[32]

In our simulations, the calculated SASA values, A_o, were calculated for the three systems using XPLOR[51] with a probe radius of 1.4 Å. The total SASA from the X-ray structure is slightly higher than the total SASA for both the CT and CTWAT systems (Table 2). The protein, when placed in a non-aqueous media, repositions its solvent exposed side chains, which results in a reduced SASA. The total SASA for both the CT and CTWAT systems have the same decrease in SASA value relative to the X-ray structure. However, the CTMONO system experienced an increase in total SASA when compared to the crystal structure. This is not surprising since the polar groups on the protein surface do not bury themselves into the protein matrix to avoid the low dielectric solvent. Hence, the SASA is larger than that observed for the CT and CTWAT runs.

By comparing a hyperthermophilic enzyme to other proteins, Rees *et al.*[32] have also found that the relatively small solvent-exposed surface area, and a relatively large number of both ion pairs and buried atoms, may contribute to the extreme thermostability of aldehyde ferredoxin oxidoreductase (AOR). Applying the same formula used by Rees *et al.*, the expected surface area, A_c, for chymotrypsin was determined to be 9593 Å. We consider an atom to be buried only if 100% of the surface area is inaccessible to the solvent. Similarly, the contribution of electrostatic interactions toward the thermal stability of γ-CT in hexane was also assessed. The number of ionic interactions per residue was defined as the difference between the number of interactions between oppositely charged side chains and the number of interactions between side chains of like charge, divided by the total number of residues in the protein. Interactions were defined as occurring between polar (oxygen or nitrogen) atoms of potentially ionizable side chains (Arg, Asp, Glu, His, and Lys) that were positioned <4 Å from one another.[32,52]

In terms of these parameters, a low value of A_o/A_c and a high value for the number of buried atoms corresponds to a minimization of the ratio of surface area to volume of a protein. In our simulations, we found that the fraction of atoms buried for the three systems is nearly identical to the crystal structure (see 3GCH in Figure 4, middle). However, while the value of A_o/A_c for CTMONO is nearly identical to the crystal structure, both CT and CTWAT exhibited a decreased value for A_o/A_c (see Figure 4, top). The lower value of A_o/A_c for CT and CTWAT is possibly due to the decreased amount of water molecules present in both simulation systems, while the number of water molecules in CTMONO simulation is closer to the number of water molecules present in the crystal. This further demonstrates that the extent of hydration is a plausible factor in affecting the thermal stability of an enzyme in organic media. γ-CT has clearly shown that it has a preference for the reduced ratio of surface area to volume in anhydrous (CT and CTWAT) media versus the nearly anhydrous (CTMONO) environment. Moreover, experimental studies have shown that suspended γ-CT retains approximately 50 or fewer water molecules and also has increased thermal stability compared to its aqueous counterparts.[23] Our calculations further show that both CT and CTWAT have a relatively higher number of salt bridges per residue compared to the CTMONO and the crystal structure. The trend observed in the CT and CTWAT system mimics AOR while the CTWAT system mimics the crystal structure. Thus, the increased net ion pair per residue which might contribute[33]

(49) Burke, P. A.; Smith, S. O.; Bachovchin, W. W.; Klibanov, A. M. Demonstration of Structural Integrity of an Enzyme in Organic Solvents by Solid-State NMR. *J. Am. Chem. Soc.* **1989**, *111*, 8290−8291.

(50) Creighton, T. E. *Proteins: Structures and Molecular Properties*; W. H. Freeman and Company: New York, 1993.

(51) Brunger, A. T.; Kuriyan, J.; Karplus, M. Crystallographic R Factor Refinement by Molecular Dynamics. *Science* **1987**, *235*, 458−460.

(52) Barlow, D. J.; Thornton, J. M. Ion-Pairs in Proteins. *J. Mol. Biol.* **1983**, *168*, 867−885.

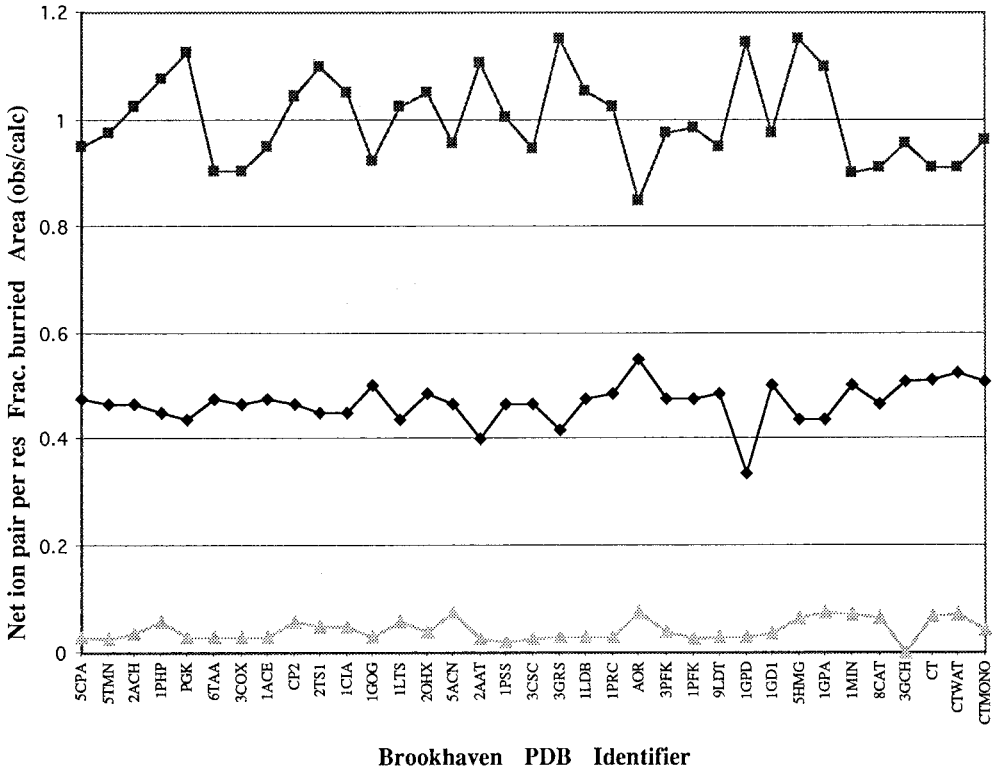

Figure 4. Variation in the net number of ionic pairs per residue (bottom panel), fraction of buried atoms (middle panel), and relative surface area (top panel) for crystalline chymotrypsin (PDB ID 3GCH), CT, CTWAT, CTMONO, and a series of monomeric and oligomeric proteins as determined by Rees *et al.*[32]

to the enhanced thermostability observed in AOR[32] is also observed for chymotrypsin in anhydrous hexane.

The salt bridge between amino terminus (Ile 16) and the side chain of Asp 194 has been shown to be buried[53] in the structure of the active enzyme, whereas in the inactive enzyme the amino group is thought to be exposed.[54,55] Our simulations show the presence of the salt bridge and the amino terminus of Ile 16 as being buried for the three systems. This result is another example of chymotrypsin's ability to retain its active conformation in organic solvent and, hence, its catalytic activity.

We further observed a marked increase in the exposed surface area of the hydrophobic residues, while the large decrease occurs at polar residues (Table 2). Both the CT and CTWAT systems experienced about 5% increase while CTMONO experiences a 4% increase for the SASA of the hydrophobic residues. This is in accord with the expectation of "like dissolves like". The hydrophobic side chains of CT reorient themselves on the surface and become more exposed to hexane, while the charged residues fold inward. However, applying an arbitrary cutoff value of 40 Å^2 in the change in surface area per residue, Lys and Thr, which are hydrophilic residues, are observed to have experienced both an increase and a decrease in SASA. In the case of Lys, it is probably due to its long hydrocarbon chain which gives hydrophobic characteristics to the side chain, despite it being considered a polar residue. Interestingly, several Lys residues are also found to undergo a large decrease in SASA in all of the runs, which is again the result of the burial of the Lys side chain into the protein matrix. Hence, whether Lys experiences a decrease or increase in SASA depends on the local interactions present around the side chain. For Thr one can imagine that the hydroxyl group could become buried into the protein matrix, while the methyl group could be exposed to the nonpolar solvent. This could result in a net increase in the SASA for Thr, but the net result again probably depends on the local environment. The CTMONO run has roughly the same number of hydrophobic and hydrophilic residues that experienced a significant increase in SASA. This is presumably due to the closer resemblance of the CTMONO system to an aqueous system rather than a non-aqueous system.

It is tempting to assign all of the enhanced thermal stability for proteins in non-aqueous solvents to the parameters used to characterize thermal stability in aqueous solvents.[56] However, there are clearly other factors present in the case of proteins in non-aqueous solvents. Clearly, the low dielectric solvent is enhancing the packing of the protein and is facilitating the formation of extensive salt-bridge interactions. Thus, the free energy of the folded state of the protein is being lowered, but the free energy of the unfolded state in a non-aqueous solvent is likely to be very different from that in aqueous solvent. While only a 10−20 kcal/mol free energy barrier separates a folded and an unfolded protein in aqueous solution, it is very likely that this is not the case for proteins in non-aqueous solvents. If one imagines a protein that is unfolding in a non-aqueous solvent it becomes very clear that numerous salt bridges have to be broken as well as having to place charged groups into a nonpolar environment. These unfavorable interactions result in highly unfavorable free energies, which suggests that the protein unfolding transition state in a nonpolar solvent would have a significantly large free energy. What the free energy of the final unfolded state would be is unclear, but it is possible that by forming "inside-out" (*i.e.*, hydrophobic exterior and hydrophilic interior) the free energy for this state may not differ greatly from the folded state. Other scenarios are possible, so in the

(53) Cohen, G. H.; Silverton, E. W.; Davies, D. R. Refined Crystal Structure of γ-Chymotrypsin at 1.9 Å Resolution. Comparison With Other Pancreatic Serine Proteases. *J. Mol. Biol.* **1981**, *148*, 449−479.

(54) Fersht, A. R. Conformational Equilibria and the Salt Bridge in Chymotrypsin. *Cold Spring Harbor Symp. Quantum Biol.* **1971**, *36*, 71−73.

(55) Fersht, A. R. Conformational Equilibria in α- and δ-Chymotrypsin. The Energetics and importance of Salt Bridge. *J. Mol. Biol.* **1972**, *64*, 497−509.

(56) Wescott, C. R.; Klibanov, A. M. The Solvent Dependence of Enzyme Specificity. *Biochim. Biophys. Acta* **1994**, *1206*, 1−9.

absence of concrete information, our proposed structure of the unfolded state in a non-aqueous solvent is, at this point, speculation. Nonetheless, it is clear that in non-aqueous solvents it is very likely that the folded structure is significantly stabilized, while the unfolding transition state is destabilized relative to the situation in water. It is the combination of these factors that results in the enhanced thermal stability of proteins in non-aqueous solvents. While most attention has been focused on the stabilization of the folded protein structure from thermophilic organisms, it is also possible that the unfolding transition state is destabilized in these systems.

Solvent Diffusion. The self-diffusion coefficients of water and hexane were calculated from the mean squared displacements after equilibration. The hexane diffusion was separated into bound hexane, that is hexane trapped in the protein, and the bulk or "free" hexane. Time intervals, δt, of up to 15 ps were used and the results averaged over all starting times. The values of the diffusion coefficients were obtained from the slope of the mean squared displacements *vs* δt over the δt interval from 5 to 15 ps (Table 2). To assess the effect of the protein on 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 molecules 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.

In the CT system, the bound water molecules have a low diffusion coefficient of 0.5×10^{-5} cm^2/s compared to the neat water value of 3.8×10^{-5} cm^2/s for TIP3P[37] water. This reduction of water mobility near the protein surface is due to water−protein interactions as well as a truncation of the space in which the water molecule can move (*i.e.*, due to the presence of the protein molecule, water molecules cannot diffuse into this space as opposed to a free water molecule that can diffuse in all directions). For the same system, the seven bound hexane molecules have a diffusion value of 3.89×10^{-5} cm^2/s. This is significantly lower than the united atom hexane[38] diffusion value of 7.2×10^{-5} cm^2/s. In addition, six of the seven initially "bound" hexane molecules were found to have diffused off of the protein surface at the end of the dynamics run. Only one bound hexane (Hex 238) is still attached to the protein, surrounded by several large hydrophobic residues forming a stable hydrophobic pocket on the protein surface. The self-diffusion value for hexane 238 was calculated to be a very low value of 0.262×10^{-5} cm^2/s, and supports the entrapment of hexane 238 in this hydrophobic pocket. Omitting hexane 238, the six total bound hexane gave a higher diffusion value of 4.48×10^{-5} cm^2/s (Figure 5). Interestingly, the site for hexane 238 was again observed in a subsequent chymotrypsin crystal grown and then exposed to hexane and 2-propanol mixtures by Farber and co-workers.[28] However, the remaining hexane sites were not observed, suggesting that these other pockets, in general, have a lower affinity for hexane and that in a hexane environment these bound hexane molecules can exchange with the bulk hexane solvent.

In the CTWAT and CTMONO systems, no hexane molecules were present in the hydrophobic pocket that was found in the CT system[27] and no hexane molecules were found in other interior regions of the protein. We postulate that the trapped hexane molecules observed in the crystal structure[27] were present because the large number of water molecules present in the crystals forced the hexane molecules to partition onto the surface of the protein. Unlike the crystal environment, where we presumably have a few hexane molecules surrounded by large amounts of water, we have a situation where a few water

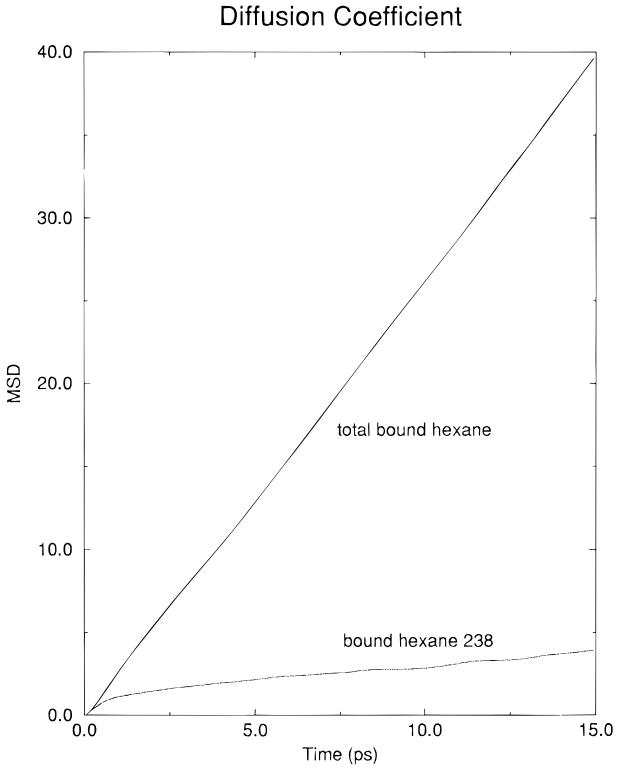

Figure 5. Solvent diffusion coefficient values for bound hexane molecules in the CT system. Only one hexane molecule (hexane 238) remains bound to the protein.

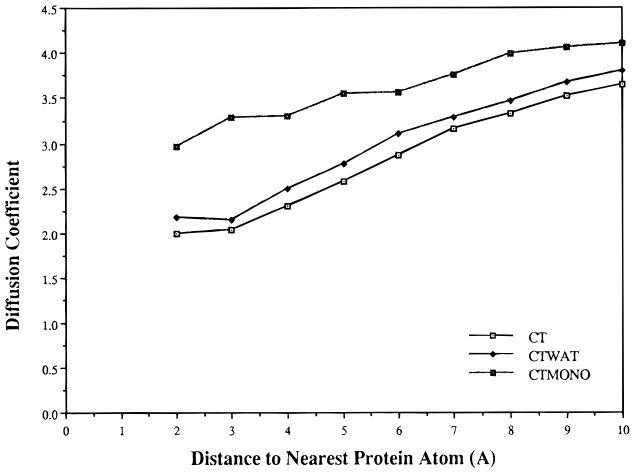

Figure 6. Plot of the calculated diffusion coefficients for solvent molecules as a function of starting distance from protein.

molecules are surrounded by large quantities of the nonpolar solvent, hexane. Thus, there is not enough of a "hydrophobic" driving force in the CTWAT and CTMONO simulations to force the hexane molecules to diffuse into the hydrophobic pockets that were found in the crystal structure.

Although no hexane molecules were found in the protein's interior for the CTWAT and CTMONO systems, hydrophobic contacts were observed between hexane molecules near the protein surface and hydrophobic side chains in all three systems. Many hydrophobic side chains were extended into the hexane environment forming many surface "clefts" for hexane molecules to transiently attach themselves to the protein surface. Hydrophilic residues, on the other hand, tended to fold back onto the surface of the protein in order to minimize surface contacts. Furthermore, the solvent mobility is strongly correlated with protein-to-solvent distance (Figure 6). Hexane molecules in CTMONO have a higher diffusion coefficient near the protein surface when compared to both CT and CTWAT.

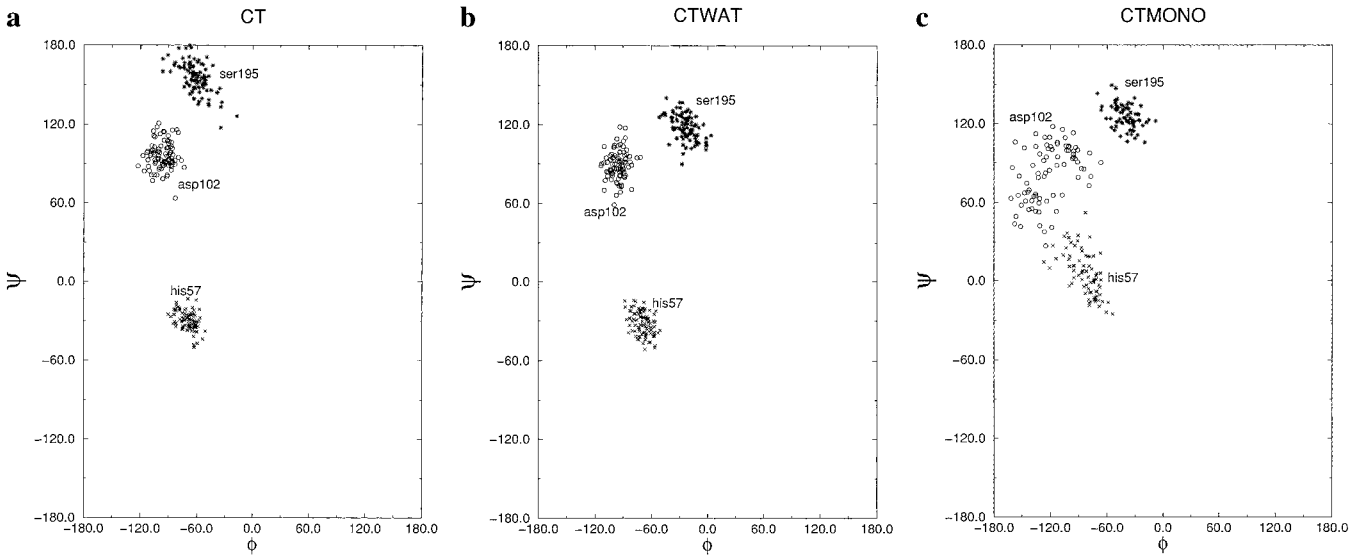

Figure 7. Ramachandran plot of catalytic triad residues in the active site for the three systems: (a) CT, (b) CTWAT, and (c) is CTMONO.

This higher diffusion coefficient is due to the "shielding" effect (*e.g.*, reduction in surface irregularities, *etc.*) provided by the monolayer of water present on the protein surface in the CTMONO system. Moreover, the hexane molecules cannot form stable interactions with the water molecules surrounding the protein itself, which further allows the hexane molecules to freely diffuse. Although the calculated hexane diffusion value near the protein surface is higher for CTMONO, the overall trend is similar to that seen in CT and CTWAT. The effect on the nonpolar solvent from the protein and water diminishes as distance increases.

For the CTMONO system, the self-diffusion value for the water monolayer is 2.99×10^{-5} cm^2/s. This value is lower than the diffusion value of neat TIP3P[37] water (3.8×10^{-5} cm^2/s). Nonetheless, the monolayer of water resembles the mobility of a free water environment and appears to not be tightly bound to the protein surface. The slight reduction of water mobility in the aqueous monolayer (*i.e.*, near the protein surface) has been attributed to both polar water−protein interactions and the formation of clathrate-like structures around the nonpolar groups.[57,58] However, other studies have also shown that water at the surface of a protein is not very different from bulk water.[59,60] In our CTMONO simulation, we observed significant rearrangements of the monolayer of water molecules. The water molecules clustered around charged hydrophilic residues, while leaving the hydrophobic residues exposed to the solvent. This process might have resulted in few water molecules being truly immobilized on the protein surface. From this observation, we agree with Klibanov and co-workers that hexane does not strip the water layer[1,61] nor does it immobilize the water molecules at the protein/solvent interface. However, it does cause some rearrangements of the water structure on the protein surface which are favorable. This clustering of water molecules on the surface of the protein is consistent with experimental results which show that the formation of a complete monolayer of water over the protein surface is thermodynamically unfavorable in a nonpolar environment.[25] The net result of this observation is the preferential solvation of the hydrophobic regions of the protein surface by the nonpolar solvent.[25]

Secondary Structure Analysis of the Active Site. Secondary structure analysis was carried out and averaged over the trajectories. Active site residues were examined for their values of ψ and ϕ with respect to time. The Ramachandran plots of the active site for the three systems were calculated and positional similarities were observed for the active site catalytic triad residues (Figure 7). All the active site residues fall within allowed regions and they are in close agreement with the active site values obtained from the crystal structure. The fluctuations in position are within the scatter that is usually associated with protein fluctuations. This indicates that the active site of chymotrypsin remains relatively unchanged, which is in accord with structural measurements of active enzymes in most anhydrous solvents.[27,28,30,31,49,56] Interestingly, the scatter for His 57 and Asp 102 observed in the CT and CTWAT simulations is less than that observed in the CTMONO simulation. Thus, the fluctuation of the active site residues in the "aqueous" simulation was greater than that observed in the non-aqueous simulation. This observation corroborates the idea that proteins in non-aqueous solvents are less flexible, but if the flexibility of the active site residues is critical for catalytic activity this might partially explain the reduced activity of enzymes in nonpolar non-aqueous solvents.

Examination of the catalytic triad in the γ-CT active site demonstrates that our simulations give results that are consistent with the NMR and crystallographic results of serine proteases in aqueous solvent.[62] From our simulations, we found that the Nδ@His-57 hydrogen interacts with the carbonyl group of Asp-102, forming a stable hydrogen bond, which is also thought to be present in the crystal structure. However, the hydrogen bond that was present in the beginning of our simulation (*i.e.*, from the crystal structure) between the hydrogen on the Oγ@Ser-195 and Nϵ@His-57 has been broken (see distances in Table 2). This probably results from the removal of the peptide from the active site of the protein, in our simulations, that was present in the crystal structure. Steitz *et al.*[62] found that the hydrogen bonds between His-57 and Asp-102 do exist but the hydrogen bond between Ser-195 and His-57 does not exist in the enzyme without the substrate in the active site. Their results, derived

(57) Brooks, C. L., III; Karplus, M. Solvent Effects on Protein Motion and Protein Effects on Solvent Motion. Dynamics of the Active Site Region of Lysozyme. *J. Mol. Biol.* **1989**, *208*, 159−181.

(58) Rossky, P. J.; Karplus, M. Solvation. A Molecular Dynamics Study of a Dipeptide in Water. *J. Am. Chem. Soc.* **1979**, *101*, 1913−1937.

(59) Van Belle, D.; Prévost, M.; Wodak, S. J. Electrostatic Properties of Solvated Proteins: A Microscopic Analysis Based on Computer Simulations. *Chem. Scr.* **1989**, *29A*, 181−189.

(60) Ahlström, P.; Teleman, O.; Jönsson, B. Molecular Dynamics Simulation of Interfacial Water Structure and Dynamics in a Parvalbumin Solution. *J. Am. Chem. Soc.* **1988**, *110*, 4198−4203.

(61) Zaks, A.; Klibanov, A. M. The Effect of Water on Enzyme Action in Organic Media. *J. Biol. Chem.* **1988**, *263*, 8017−8021.

(62) Steitz, T. A.; R. G., S. Crystallographic and NMR Studies of the Serine Proteases. *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 419−444.

from an aqueous study, and our results from a non-aqueous study show the ability of the enzyme to conserve the preferred active site conformation in both aqueous and organic medium and, therefore, explain the ability of enzymes to function in non-aqueous solvents.

A study of the active site within a 5-Å radius of the Oγ@Ser-195 "pocket" reveals some interesting features. In CTMONO, despite the protein being solvated in hexane, the hydrophillicity of the active site is conserved. There were no hexane molecules present in the substrate binding "pocket". Instead, there are between 2 and 4 water molecules residing in this "pocket" at any time throughout the simulation. The Oγ@Ser-195 hydrogen tends to form a hydrogen bond with the oxygen of a water molecule nearby. Occasionally, the hydrogen on the same water molecule also forms a hydrogen bond with Nϵ@His-57. This arrangement allows the protein to maintain the basic structural features of the catalytic triad (Asp-His-Ser). In the CTWAT system, there were no hexane or water molecules observed throughout the simulations. Instead, the Oγ@Ser-195 hydrogen forms a hydrogen bond with Oγ@Ser-205. Similarly, in the CT system the Oγ@Ser-195 hydrogen hydrogen bonds with O@Phe-36. One of the "essential" water molecules that was present in the beginning of the CT simulation is also present in the "pocket" periodically throughout the CT simulations but does not participate in the hydrogen bond formation with Oγ@Ser-195. There are also between 1 and 2 hexane molecules present in the CT substrate binding "pocket".

Hydrogen Bonding Interactions. Hydrogen bonding analyses were carried out on the coordinate sets saved from the dynamics trajectories after equilibration. The criteria used to determine 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. A distance cutoff of 2.6 Å between the donor and the acceptor and an angle cutoff between 120° and 180° at the hydrogen atom were applied. The criteria used are similar to those used by Tirado-Rives and Jorgensen.[42]

We have broken down the hydrogen bond characteristics into those present >90% of the time and those present <10% of the time. Hydrogen bonds present for more than 90% of the time can be considered to be very stable, while those hydrogen bonds present for less than 10% of the time can be considered weak or non-existent (Table 2). It can be clearly seen that although CTMONO has the highest number of total hydrogen bonds, it also has the least number of stable hydrogen bonds. On the other hand, the CT and CTWAT systems have significantly higher numbers of stable hydrogen bonds than the CTMONO system. This clearly indicates the preference for formation of more intramolecular hydrogen bonds within a protein when placed in an organic environment. This has also been noted in other MD studies of proteins in organic solvents.[18,19,22]

Conclusions

The results presented provide molecular-level insights into the structure and dynamics of γ-CT in an hexane environment. Our studies have shown that both hydration and the placement of "essential" water molecules affect the flexibility of the protein when placed in a non-aqueous environment. As hydration increases, the structural similarities of the protein to the crystal structure increases and placement of bound water molecules affect the "local" mobility of the protein, mainly the surface loops. Chen and Arnold[46] have observed that mutagenesis at the surface loops of subtilisin E allows this enzyme, when suspended in a non-aqueous solvent (DMF in this case), to retain its catalytic activity for a longer time. Our results suggest that these mutations may stabilize the conformationally flexible surface loops, thereby stabilizing the protein. Further studies aimed at understanding this observation would be interesting.

We find that the protein total solvent accessible surface area (SASA) is reduced when placed in hexane, with hydrophobic residues experiencing the biggest increse in SASA while hydrophilic residues undergo the greatest decrease in SASA. Increasing the number of water molecules in the simulation results in an overall increase in SASA due to swelling of the protein. The initial placement of bound water does not appear to have any appreciable effect on the total SASA. The monolayer of water covering the protein is also not uniform, with the hydrophilic residues receiving preferential hydration and leaving the hydrophobic residues more exposed to the nonpolar solvent.

The protein maintained its spherical nature in hexane, showing no significant unfolding or denaturation. The secondary structure for the γ-CT active site is also relatively unchanged and the active site side chain fluctuations are less than the average fluctuations of the protein. We further found that the hexane molecules did not diffuse into the hydrophobic interior of the protein indicating that although the protein remains stable in hexane, it remains unsolvated or impermeable to the organic solvent (*i.e.*, suspended in hexane). Our simulations show that only one of the seven observed hexane molecules remained bound during the course of the CT simulation. Hence, we have to conclude that the X-ray structure is not representative of the solution phase conditions simulated herein. Possibly the experimentally observed hexane molecules are held in place by crystal packing forces or by the presence of water molecules in the interstitial spaces within the crystal. The placement of bound water molecules does not have a major effect on the protein intramolecular interactions, but the number of water molecules in a non-aqueous environment affects the formation of more intramolecular hydrogen bonds. As water hydration increases, the number of stable intramolecular hydrogen bonds decreases indicating the role that aqueous solvent plays in facilitating fluctuations in the protein structure.

From our simulations, we note that the increase in the stability of the folded protein in organic media is due to the increase in intramolecular hydrogen bonds, net ion pair interactions, and reduced ratio of surface area to volume of the protein. These same criteria have been used to explain the origin of thermal stability of protein obtained from thermophilic organisms.[32] The relatively hydrophobic protein interior and the hydrophilic active site appear to be unaffected by the presence of non-aqueous solvent. The protein remains suspended, yet stable, if the minimum hydration required is provided.

Acknowledgments. We thank the Office of Naval Research for support of this research through Grant No. N00014-90-3-4002. The Pittsburgh Supercomputer Center and the Cornell Theory Center are acknowledged for generous allocations of supercomputer time through a MetaCenter grant. We thank Dr. Greg Farber for helpful discussions. We also thank Professor D. Rees for much of the data contained in Figure 4.[32]

JA960153F