SUBTILISIN STRUCTURE AND CATALYSIS IN NON-AQUEOUS MEDIA

by

JENNIFER L. SCHMITKE

B.Sc. Honors, Chemistry
Queen's University, Canada, 1992

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
IN CHEMISTRY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June, 1998

Signature of Author: ___ _____
May 14, 1998

Certified by: _____
Professor Alexander M. Klibanov, Thesis Supervisor

Accepted by: _____
Professor Dietmar Seyferth, Chairman
Departmental Committee on Graduate Students

This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Prof. Lawrence J. Stern
Department of Chemistry — Chairman

Prof. Alexander M. Klibanov
Department of Chemistry — Thesis Supervisor

Prof. Scott C. Virgil
Department of Chemistry

SUBTILISIN STRUCTURE AND CATALYSIS IN NON-AQUEOUS MEDIA

by JENNIFER L. SCHMITKE

Submitted to the Department of Chemistry on May 19, 1998, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Chemistry

ABSTRACT

The X-ray crystal structure of the serine protease subtilisin Carlsberg in anhydrous dioxane has been determined to 2.6 Å resolution. This structure has turned out to be nearly indistinguishable from those previously determined in water and in acetonitrile. Seven enzyme-bound dioxane molecules have been detected, each potentially forming at least one hydrogen bond with a subtilisin hydrogen-bond donor or bound water. Two of the bound dioxane molecules are in the active site region. The other five dioxane molecules are located on the surface of subtilisin at interprotein crystal contacts. The locations of the bound solvent in the dioxane structure are distinct from those in the structures in acetonitrile and in water.

Subtilisin dissolved in aqueous solution is several orders of magnitude more active than the enzyme suspended in anhydrous acetonitrile. In order to ascertain why, we have employed lightly cross-linked crystalline subtilisin as a catalyst in both aqueous and organic media. Since the structure of this crystalline enzyme in anhydrous acetonitrile (and dioxane) is virtually identical to that in water (see above), solvent-induced conformational changes as the cause of the enzymatic activity drop can be ruled out. Titration of the competent active centers of subtilisin has revealed that k_{cat}/K_M is solely responsible for this activity plunge. Quantitative mechanistic analysis of the 7-order-of-magnitude difference in k_{cat}/K_M values between subtilisin dissolved in water and cross-linked subtilisin crystals suspended in anhydrous acetonitrile allows accounting for at least 5.6 orders of magnitude. This drastic decline is due to (i) a marked shift in the activity *vs.* pH profile of the cross-linked crystalline enzyme compared to its soluble counterpart; (ii) different energetics of substrate desolvation; and (iii) very low thermodynamic activity of water in anhydrous acetonitrile resulting in a much more rigid and thus less active enzyme.

The X-ray crystal structures of *trans*-cinnamoyl-subtilisin, an acyl-enzyme covalent intermediate, have been determined to 2.2-Å resolution in anhydrous acetonitrile and in water. The cinnamoyl-subtilisin structures are virtually identical in the two solvents. In addition, their enzyme portions are nearly indistinguishable from the structures of the free enzyme in acetonitrile and in water; thus acylation in either aqueous or nonaqueous solvent causes no appreciable conformational changes. However, the locations of bound solvent molecules in the active site of the acyl- and free enzyme forms in acetonitrile and in water are distinct. Such differences in the active site solvation may contribute to the observed variations in enzymatic activities left unexplained in the aforementioned kinetic study. Notably, prolonged exposure to organic solvent or removal of interstitial solvent from the crystal lattice causes the channels within enzyme

crystals to collapse, leading to a drop in the number of active sites accessible to the substrate.

The crystal structures of subtilisin in 40% acetonitrile and 20% dioxane have been determined to 2.3- and 2.2-Å resolution, respectively, and their solvent binding patterns have been compared to those observed in the neat organic solvents. The structures of the protein in the two aqueous-organic mixtures are the same as in the neat water, acetonitrile, and dioxane. Interestingly, the enzyme-bound organic solvent molecules tend to congregate in the active site. Three of the five bound acetonitrile molecules in the structure of subtilisin in 40% acetonitrile are situated in the enzyme active site, as is the single enzyme-bound dioxane molecule in 20% dioxane. Significantly, these organic solvent molecules are present in the same locations in the enzyme active site in the corresponding neat solvents.

Thesis Supervisor: Dr. Alexander M. Klibanov, Professor of Chemistry

ACKNOWLEDGMENTS

I wish to recognize my thesis advisor, Alexander Klibanov, for his guidance, encouragement, and financial support .

I am also grateful to the past and present members of the Klibanov group, especially Charles Wescott, Henry Costantino, and Tao Ke for helpful discussions and for their friendship.

Most importantly, I would like to thank my parents, Rodney and Mary Schmitke, my sister Rosalind Schmitke, and David Schuster for their constant support, understanding, and patience.

TABLE OF CONTENTS

INTRODUCTION

In recent years, it has been shown that not only do enzymes remain synthetically useful catalysts in organic solvents, instead of their natural aqueous milieu, but they offer many significant advantages in such media as well.* For example, many substrates of interest to synthetic chemists are insoluble in water but readily soluble in organic solvents. Moreover, equilibria which are unfavorable in aqueous environments can be shifted to the desired products in organic solvents. In addition, undesirable hydrolysis of products or other side reactions involving water can be avoided in anhydrous media.

Importantly, enzyme preparations in non-aqueous media are insoluble, allowing the facile separation of dissolved products from the enzyme catalyst suspension. Also, the biocatalysts themselves can be easily recovered for reuse simply by filtration. In addition, enzyme suspensions often show increased stability in non-aqueous media compared to their soluble counterparts in water. Finally, enzymes display remarkable novel properties in organic media, such as solvent-dependent selectivity, including enantioselectivity. Despite all these impressive assets, however, a main drawback limiting the utility of enzymes in anhydrous solvents, a focus of this thesis, is that their catalytic activity in such milieu is drastically diminished relative to that in water.*

Lyophilized powders have been the most widely used form of enzyme suspensions in organic solvents. Recently, though, cross-linked crystalline enzymes emerged as useful catalysts.† This enzyme preparation proffers several benefits, such as increased thermostability and insolubility in aqueous-organic mixtures and even water

---

* See reference (1) in Chapter I and (1) in Chapter II.

† See reference (7) in Chapter II and (15) in Chapter III.

itself. Most importantly, direct structural information via X-ray crystallography can be obtained for these crystalline formulations, while enzymes in lyophilized powders, which are amorphous, are of generally unknown conformations. Thus, molecular structure-activity relationships can, for the first time, be developed for such crystalline enzyme catalysts in organic solvents. In fact, the X-ray crystal structure of such an enzyme catalyst in an anhydrous solvent, namely that of subtilisin Carlsberg in neat acetonitrile, was recently determined and shown to be indistinguishable from that in water.[‡]

At this time, however, several key questions remain unresolved, namely: Is the case of acetonitrile unique, or are the structures of crystalline enzymes in other neat organic solvents also the same as in water?, Are cross-linked crystalline enzymes also much less active in organic solvents compared to water?, If so, why? Is the enzyme mechanism of such a catalyst the same in anhydrous solvents as it is in water? These issues will be addressed in Chapters I through III in this thesis using the model enzyme catalyst of cross-linked crystals of subtilisin Carlsberg.

In Chapter I, the generality of the finding that the structure of crystalline subtilisin does not change upon transition from water to a neat organic solvent will be tested by solving the crystal structure of subtilisin in anhydrous dioxane. Also, the solvent binding patterns of the two organic solvents will be compared.

Knowing that structural changes of the enzyme cannot account for differences in observed activity of cross-linked crystals of subtilisin in water *vs.* anhydrous acetonitrile, in Chapter II the activity differences between the two solvents will be systematically

---

[‡] See reference (2) in Chapter I.

investigated. A mechanistic study of the causes of the detected plunge in enzymatic activity upon transition from water to the organic solvent will then be performed.

In order to shed further light on the issue of the mechanism of subtilisin in different solvents, in Chapter III the crystal structures of an acyl-enzyme intermediate, cinnamoyl-subtilisin, will be determined in water and in anhydrous acetonitrile. To complement the findings in Chapter II, the differences in solvent binding and subsequent displacement by the substrate in the enzyme active site will also be scrutinized. In addition, Chapter III will include an investigation of the consequences of the crystallinity of the catalyst itself to provide insight into improving the utility of such preparations for use in non-aqueous media.

A related subject, involving the determination of enzyme crystal structures in organic solvents for the creation of functional group maps of protein surfaces,[§] will be examined in Chapter IV. The crystal structures of subtilisin in the aqueous-organic (predominantly aqueous) mixtures of 40% acetonitrile and 20% dioxane will be determined. The resulting solvent binding patterns will be compared to those found in the corresponding neat solvents, and the implications for rational drug design will be discussed.

For the convenience of the reader, each chapter contains its own introduction, results and discussion, methods, and reference sections. Note that each chapter has resulted in a publication.

[§] See reference (1-4) in Chapter IV.

I. The Crystal Structure of Subtilisin Carlsberg in Anhydrous Dioxane and Its Comparison with Those in Water and Acetonitrile

A. Introduction

Enzymatic catalysis in organic media has emerged as an important area of research in biotechnology and bioorganic chemistry (1). The structure of enzymes in such milieu should aid in the understanding and optimization of this catalysis. Several years ago, the first such X-ray crystal structure, that of subtilisin Carlsberg in anhydrous acetonitrile was determined (2). It was found that the structure of the lightly cross-linked enzyme crystal in the anhydrous solvent was virtually indistinguishable from that in water (2,3). This conclusion was corroborated by subsequent crystallographic studies of two other enzymes — γ-chymotrypsin in hexane (4,5) and elastase in acetonitrile (6).

To date (2-6) only one nonaqueous solvent has been examined for each crystalline enzyme. A rationale has been tentatively adopted (7,8) that if the enzyme structure does not change upon transition from water to acetonitrile or hexane, then it also should not change upon transition to another solvent because the difference in physicochemical properties between two organic solvents should be less than between water and an organic solvent. However, a direct validation of this reasoning until now has been missing.

In the present study, we have determined the X-ray crystal structure of lightly cross-linked subtilisin in a second, unrelated anhydrous solvent, dioxane. This structure

has turned out to be essentially the same as those in water (3) and in acetonitrile (2). Interestingly, the locations of the subtilisin-bound dioxane and acetonitrile molecules are distinct.

B. Results and Discussion

The structure of subtilisin Carlsberg in anhydrous dioxane determined to 2.6 Å resolution (Table 1.1) includes 65 bound water molecules and seven bound dioxane molecules (Fig. 1.1A). The protein structure is virtually identical to that in water (3) (Fig. 1.1B), with root mean squared (RMS) shifts of the main-chain atoms of 0.33 Å.

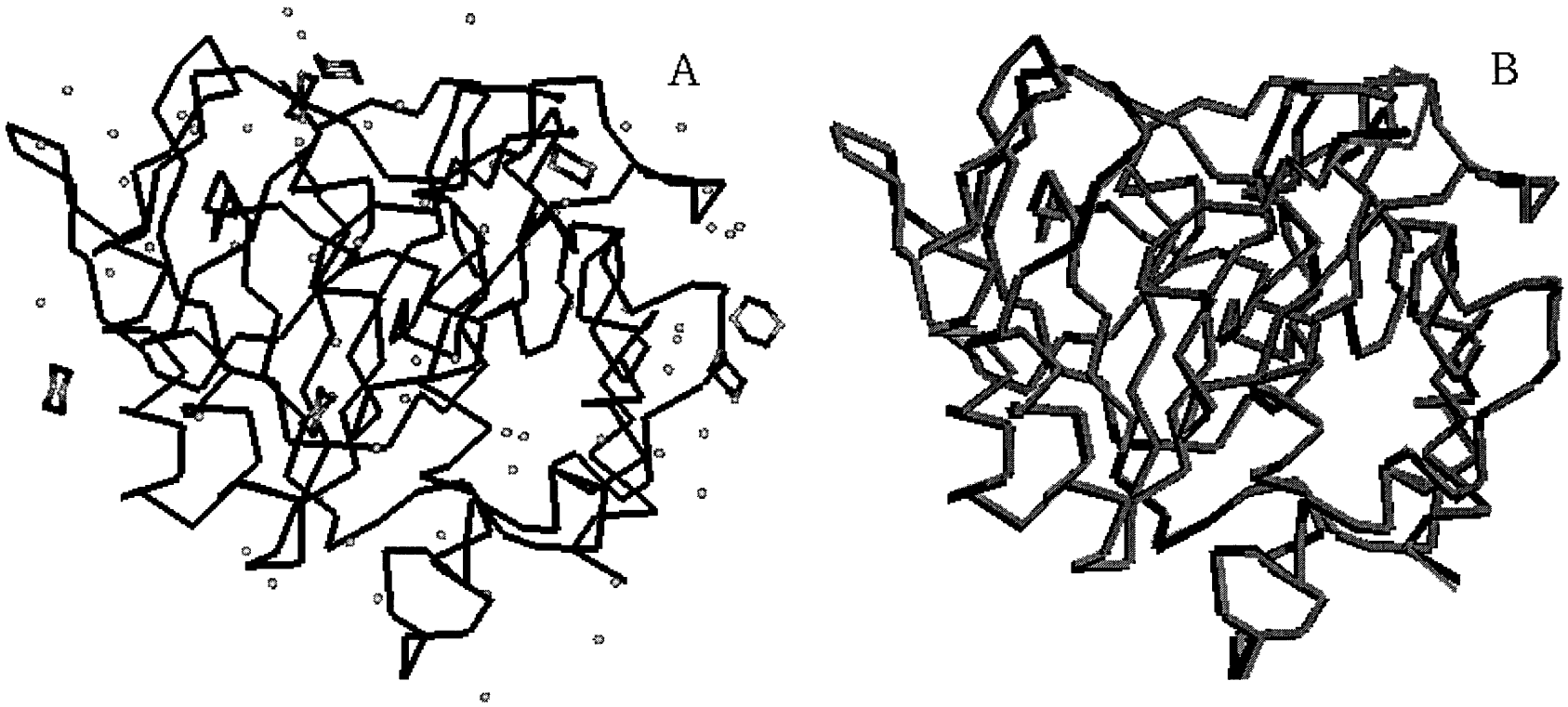


Figure 1.1. (A) C_α trace of the protein structure of subtilisin Carlsberg in anhydrous dioxane (black lines) with bound water (gray balls) and dioxane (balls-and-sticks: C atoms in black and O atoms in gray) molecules. (B) Superimposition of the C_α traces of subtilisin in water (thin gray lines) and in anhydrous dioxane (thicker black lines). The RMS shifts of the C_α atoms and of the main-chain atoms are 0.31 Å and 0.33 Å, respectively. Ser159 and Gly160 are not included in the model of subtilisin in dioxane because of uncertainties in their location (see Materials and Methods); the C_α atoms of Asn158 and Ser161 are shown as balls (upper right).

Not only does the overall global structure of the enzyme not change upon transition from water to anhydrous dioxane, but the local, side-chain structure is also essentially unaffected. Small deviations in the side-chain conformations in dioxane compared to their conformations in water are no greater than those between two previously determined structures in water, namely those of subtilisin in the cross-linked crystal in distilled water (3) and in the non-cross-linked crystal in aqueous buffer (9). In addition, the apparent number of the intramolecular hydrogen bonds of subtilisin in dioxane and in water were found to be similar — 45 and 49, respectively. Figure 1.2A illustrates that both the overall *B*-factor and the *B*-factor trends through the protein are also similar in water, dioxane, and acetonitrile.

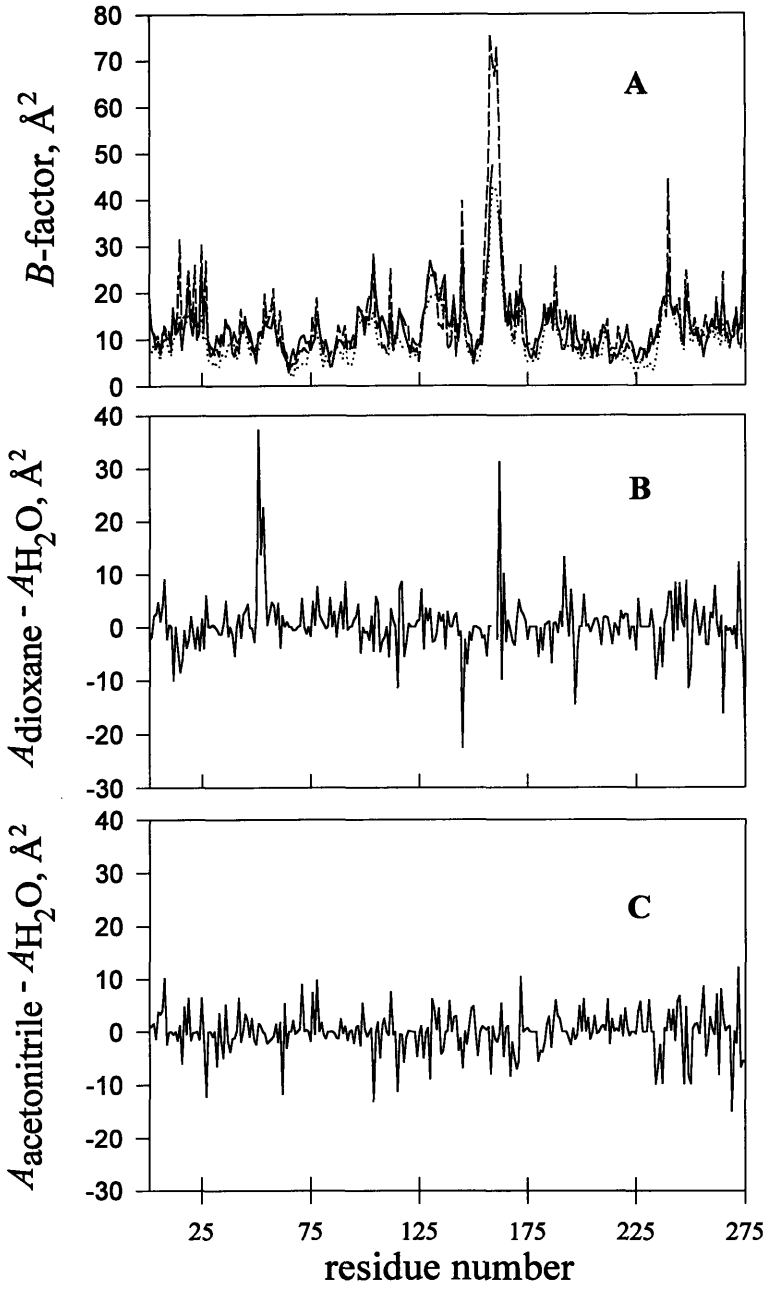


Figure 1.2. (A) Refined crystallographic *B*-factors (residue average) of subtilisin Carlsberg in dioxane (——), water (----), and acetonitrile (····). (B) The difference in the exposed surface areas of subtilisin between dioxane and water, and (C) between acetonitrile and water. As in Fig. 1.1, Ser159 and Gly160 are left out in A and B.

Studies involving molecular dynamics simulations have suggested that proteins in organic solvents should possess a diminished overall exposed surface area relative to their structures in water (10,11). We observed (Fig. 1.2B) no significant change in the overall surface area of subtilisin in dioxane (9100 $Å^2$) compared to water (9100 $Å^2$). Additionally, there is only some 1% difference in the surface area in acetonitrile (9000 $Å^2$) compared to water (Fig. 1.2C). The simulation studies (10,10) also predict that nonpolar residues will be more exposed and polar residues less exposed in organic solvents compared to water. Although some subtilisin residues agree with this prediction, others do not. In general, the differences appear random: the average change in the exposed surface area for a given residue type does not follow the predicted trend for the structures in either dioxane or acetonitrile compared to water (Fig. 1.2 B and C). These inconsistencies could have arisen because the simulation studies (10,10) were performed on a different protein, bovine pancreatic trypsin inhibitor, and in a different, non-hydrogen bonding solvent, chloroform. An alternative possibility is that the enzyme is kinetically trapped in the same conformation as in water, and consequently cannot reach a minimum energy state of the sort predicted by the molecular dynamics simulations.

Since subtilisin exhibits markedly different catalytic activity and other properties in different solvents (7,8,12), its active site structure in solvents is a critical issue in our study. We find that, in addition to the global enzyme structure, this region in particular of the lightly cross-linked crystals of subtilisin in dioxane is essentially the same as in water. The RMS displacement is 0.17 Å for all the atoms within a 10 Å radius of the head nucleophile Ser-221. Hence, the organic solvent does not alter subtilisin's catalytic activity by causing a significant structural change in the active site.

Examination of the solvent bound in the active site region in dioxane compared to that in water (Fig. 1.3A) reveals that the catalytic water molecule, a hallmark of the active site of subtilisin and other serine proteases (13), is also present in dioxane. One can speculate that in a transesterification or another non-hydrolysis reaction catalyzed by proteases in organic solvents (1) this bound water may, in fact, lead to a hydrolysis product in the first turnover. Inspection of Fig. 1.3A also shows that one dioxane molecule replaces two water molecules and one dioxane is located in an unoccupied site relative to the water structure. Additionally, some water molecules are observed in different locations in the two structures. An analogous study carried out for the situation in dioxane compared to acetonitrile (Fig. 1.3B) reveals that in the former a dioxane molecule binds in an analogous position to a bound acetonitrile, while the other dioxane replaces a water molecule, and two waters replace two bound acetonitriles. The distinct solvation of the subtilisin active site in different organic solvents possibly contributes to the altered enzymatic activity and other properties in those media (7,8,12).

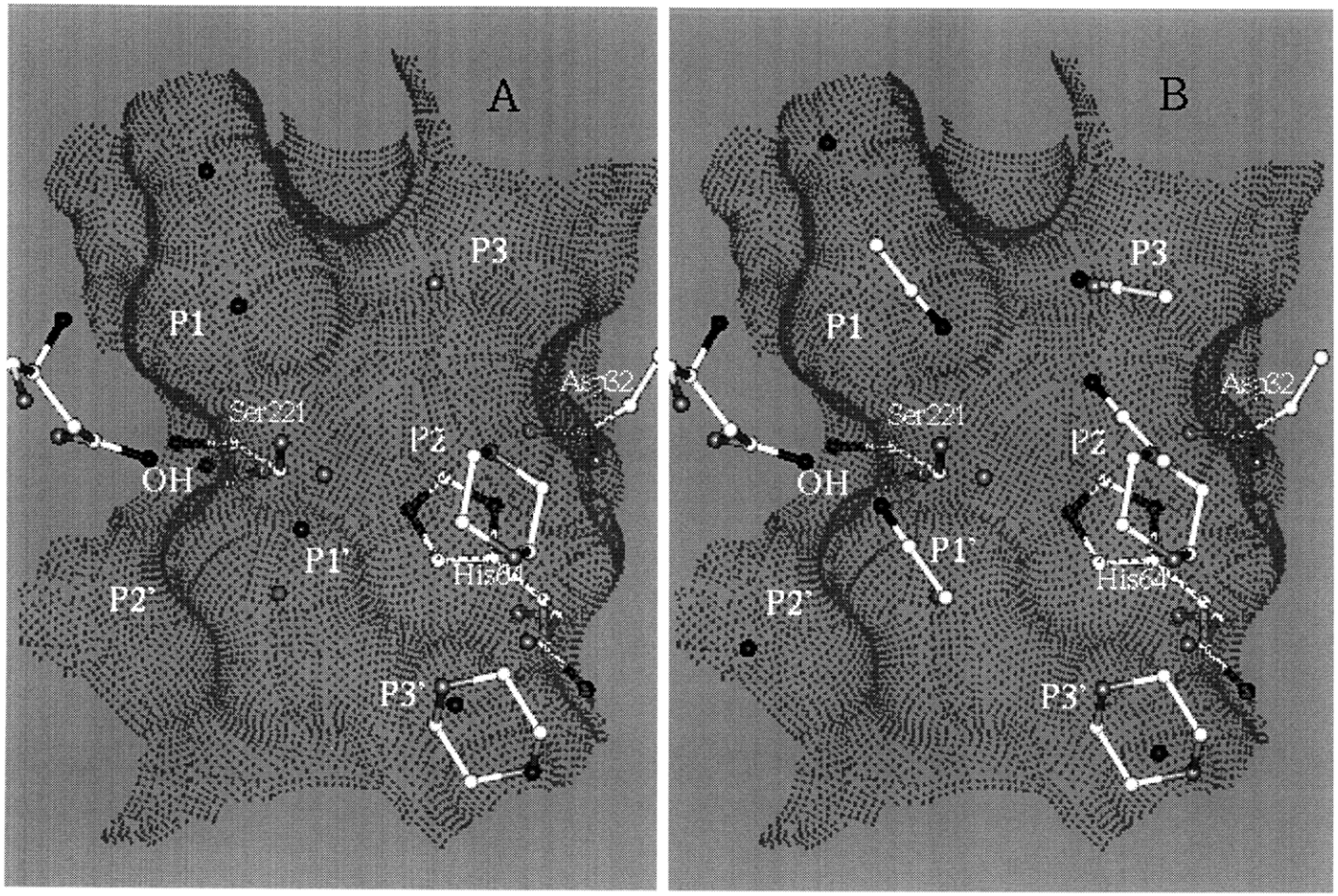


Figure 1.3. Active site of subtilisin Carlsberg in anhydrous dioxane compared with that in water (A) and acetonitrile (B). The catalytic triad, Asn155 of the oxyanion hole (OH), and solvent molecules in the dioxane structure are shown as balls-and-sticks with carbon, oxygen, and nitrogen shown in white, light-gray, and black, respectively. The water molecules in the aqueous and acetonitrile structures are depicted as dark-gray balls in A and B. The acetonitrile molecules in that structure (B) are also shown as linear balls-and-sticks. The surface of the protein (Connolly algorithm (29)) in the active site region in the dioxane structure is portrayed by black dots.

Crystals of subtilisin, even after cross-linking, usually fail to diffract X-rays upon replacement of their interstitial water by most anhydrous solvents (ref. 3 and our unpublished results). The crystal lattice is held together by weak intermolecular interactions between the packed proteins (14). Presumably, most solvents disrupt these interactions, thereby causing the crystal to lose its integrity. We have ruled out the possibility of significant *intramolecular* distortions: not only are the crystals catalytically active in solvents in which they fail to diffract, but, in fact, the activities of non-cross-linked and cross-linked crystals of subtilisin in acetonitrile are the same (2). It is, therefore, important to assess the effect of dioxane on the crystal contact regions. We observed no significant difference in the direct protein-protein crystal contacts between the dioxane and aqueous structures or between the dioxane and acetonitrile structures. It is of possible significance that of seven bound dioxanes, five are involved in crystal contacts with the other two in the active site. Of the five, two dioxanes replace two waters, another two replace one water, and one is in a location unoccupied in the structure in water. Perhaps the subtilisin crystal remains intact in dioxane because the molecules of the latter sustain interactions (Fig. 1.4) maintained by the water that they replace. Likewise, the enzyme crystal should not diffract in solvents which do not preserve these interactions, e.g., toluene and octane (our unpublished results), as well as several others (3).

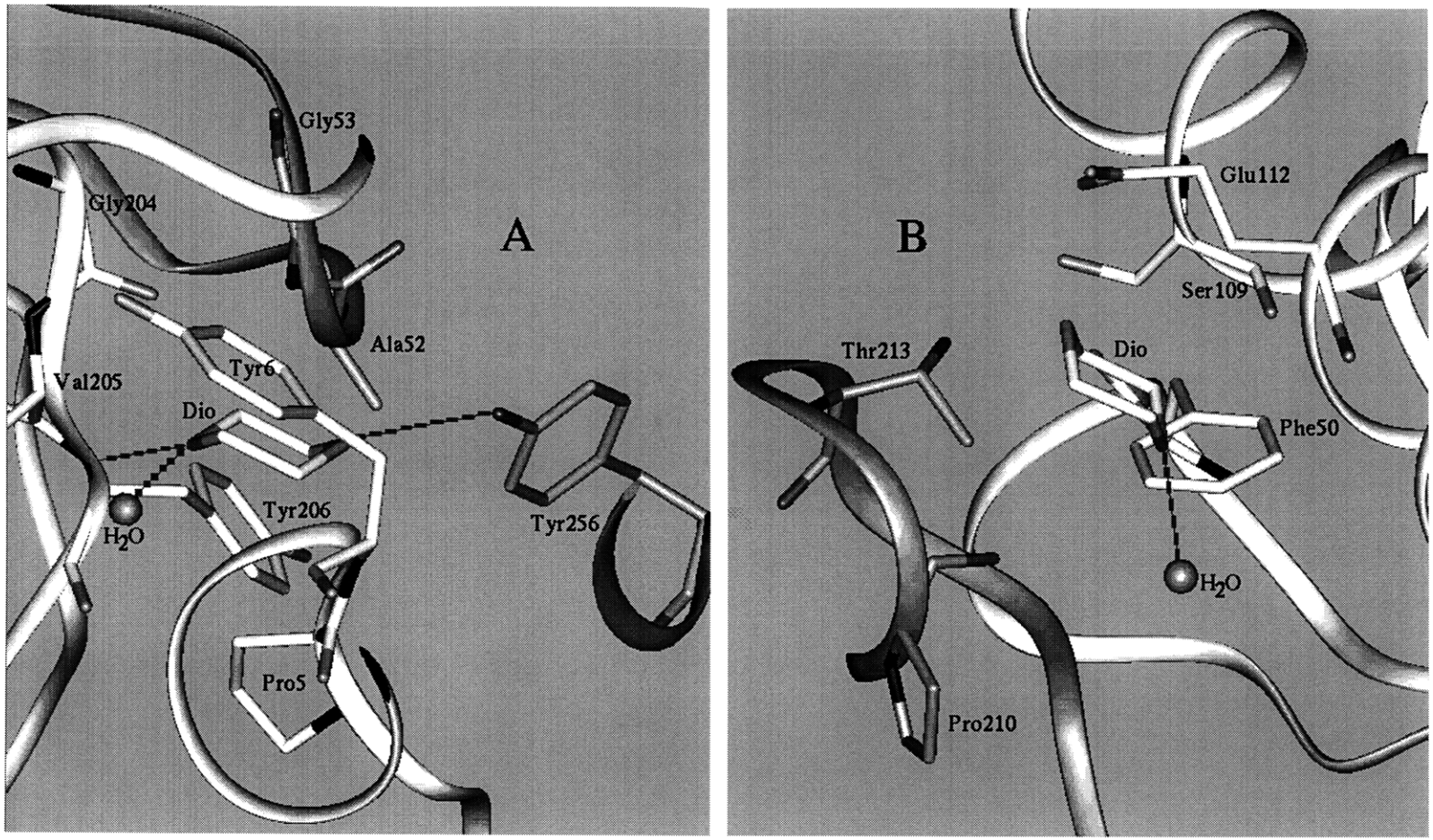


Figure 1.4. Two examples of dioxane (Dio) molecules binding to subtilisin Carlsberg in anhydrous dioxane. (A) A dioxane molecule is seen interacting with Pro5, Tyr6, Gly204, Val205, Tyr206, and a water molecule, and with Ala52 and Gly53 of one neighboring, symmetry-related subtilisin molecule, as well as Tyr256 of another. (All residues depicted are within 3.5 Å of the dioxane.) (B) A dioxane molecule is displayed interacting with Phe50, Ser109, Glu112, and a water molecule, and with Pro210 and Thr213 of a neighboring subtilisin molecule. In both A and B, the central subtilisin's residues and bound solvent molecules mentioned above are displayed as balls-and-sticks where the carbon, oxygen, and nitrogen are white, light-gray, and black, respectively; the ribbon diagram of the subtilisin backbone is also white. The neighboring, symmetry-related subtilisin's residues and ribbons are shown in darker gray colors. The hydrogen bonds, in which the dioxane oxygen atoms could be involved, are shown as gray dashed lines.

There is an interest in determining the potential binding sites in proteins for small organic molecules representing particular functional groups, with the eventual goal of designing tightly binding compounds consisting of a combination of such groups (2,6,15). We therefore examined subtilisin's binding sites for dioxane and for acetonitrile and compared them. When a solvent molecule binds to the protein surface, only a fraction of the former would be in direct contact with the protein, thus leaving a significant percentage of the molecule exposed. The average exposed surface areas of the seven dioxane and twelve acetonitrile molecules are 34% and 51%, respectively. The value for the dioxane molecules must be lower because of their involvement in the crystal contacts where the dioxane molecules tend to be sandwiched between two neighboring subtilisins (Fig. 1.4). Thus it appears that both dioxane and acetonitrile bind at the surface of the enzyme. Meanwhile, only one dioxane molecule binds near an acetonitrile site in the respective structures (Fig. 1.3B). One dioxane molecule replaces two waters, and two replace one water each (example in Fig. 1.3B). In contrast, three acetonitrile molecules replace one water molecule each (two in the active site (Fig. 1.3B)). Since the dioxane and acetonitrile molecules replace *different* water molecules in their respective structures, it follows that the two solvents bind to distinct sites on subtilisin.

The question arises about the driving force of the binding of dioxane to subtilisin. A likely major contributor is hydrogen bonding — all the bound dioxanes can form at least one hydrogen bond with a subtilisin hydrogen-bond donor or an enzyme-bound water. Moreover, all but one of the bound dioxanes can form at least two hydrogen bonds, where each of the dioxane oxygen atoms is involved in at least one hydrogen bond

(Fig. 1.4). A principal force in the acetonitrile binding to the protein surface also may be hydrogen bonding: all but one of the twelve bound acetonitrile molecules can form hydrogen bonds with donor atoms of subtilisin or water (allowing for the ambiguity in acetonitrile orientation in electron density maps) (2).

Since the enzyme-bound dioxane molecules that are not in the active site region of subtilisin are found at crystal contacts, they can form more contacts with the protein than they could have in other areas of subtilisin's surface. Consequently, these dioxane molecules might be more ordered and therefore have more pronounced electron density. Interestingly, four of the seven enzyme-bound dioxane molecules interact with a Tyr residue (Fig. 1.4A), although tyrosines constitute less than 8% of subtilisin's amino acid residues with an exposed surface area above 30%. (On the other hand, only 5 of the 65 bound water molecules in the dioxane structure interact with a Tyr residue.) Of the three dioxanes that do not interact with a Tyr, two are found near other side-chain rings, those of His and Phe (Fig. 1.4B). Thus, six out of seven enzyme-bound dioxanes are located near rings, with a dioxane oxygen apparently interacting with the ring (Fig. 1.4). In contrast, the acetonitrile molecules exhibit no such preference for binding at the crystal contacts (only two of the twelve bind in such regions), nor do they appear to favor association with ring structures.

In summary, in the present study we have demonstrated that the crystal structures of subtilisin in two unrelated organic solvents, dioxane and acetonitrile, are virtually identical to each other and to that in water. In particular, the active site structures in all three solvents are essentially the same. Also, the overall exposed surface area of the protein in the crystal is not substantially affected by the solvent. With respect to the

bound solvent, two dioxane molecules bind in the active site of the enzyme, while the other five are all involved in interprotein crystal contacts. Finally, the dioxane binding sites observed in the enzyme are distinct from those for acetonitrile.

C. Materials and Methods

Crystal preparation. Subtilisin Carlsberg (serine protease from *Bacillus licheniformis*, EC 3.4.21.14) was purchased from Sigma Chemical Co. Crystals were grown from an aqueous 330 mM cacodylate buffer, pH 5.6, saturated with Na_2SO_4, approximately 13% (9). A single crystal (~0.8 × 0.1 × 0.05 mm) was placed in 1 mL of a 10% glutaraldehyde cross-linking solution containing 30 mM cacodylate buffer, pH 7.5, and 10% Na_2SO_4. The glutaraldehyde solution was aged for three days at room temperature before reaction. (Crystal appearance and diffraction limit were dependent on the glutaraldehyde aging protocol.) We presume a relatively low extent of cross-linking: there are only 9 lysine residues in subtilisin and only 5 which could reasonably be involved in an intermolecular cross-linking event (distances below 20 Å). No glutaraldehyde or modified Lys electron density was detected in the electron density maps.

The crystal was incubated in the cross-linking solution for 30 min and washed five times with 2 mL of distilled water. The water was then removed from the crystal and replaced with anhydrous dioxane by washing five times with 2 mL each. The crystal was left in the solvent for 20 min following each wash. Finally, the crystal was mounted in a

0.5-mm diameter quartz capillary for data collection. The water content of similarly treated crystals was found to be 3% (w/w) using a Mettler Karl-Fisher titrimeter (16).

Data collection and reduction. X-ray diffraction data were obtained at ambient temperature (23±2°C) to a nominal resolution of 2.6 Å with an RAXIS II area detector. X-rays were generated with a Rigaku RU200 rotating copper anode source and the Cu K_α radiation was selected with a graphite monochromator. An oscillation range of 1.75° with an exposure time of 35 min was used to collect data over a total of 137° (> 95% of the possible data 14-2.6 Å). The reflections were indexed and the intensities integrated and scaled using DENZO and SCALEPACK of the HKL Package (17). Data collection statistics are presented in Table 1.1. Structure factor magnitudes were calculated from the intensities and truncated using programs in the CCP4 Suite (18).

Table 1.1. Crystal properties and data collection, refinement, and model statistics

Crystal properties		
space group	$P2_12_12_1$	
cell dimensions: [a] $a \times b \times c$, Å	77.0×55.2×53.9	
molecules per asymmetric unit	1	
Data collection statistics	resolution limits	
	14 - 2.6 Å	2.7 - 2.6Å
number of measurements	39403	3891
number of unique reflections	7111	745
completeness, %	95.9	96.2
mean I / σ_I	6.2	4.3
R_{merge}, [b] %	14.2	38.7
Crystallographic refinement statistics	resolution limits	
	6.0 - 2.6Å	2.7 - 2.6Å
number of reflections	6554	745
number of reflections with $F>2\sigma_F$	6307	708
crystallographic *R*-factor, [c] %	16.1	19.3
free *R*-factor, [d] %	24.6	28.0
Model statistics[e]		
total number of non-hydrogen atoms	2017	
number of protein atoms	1910	
number of solvent atoms	107	
RMSD[f] bond length, Å	0.014	
RMSD[f] bond angle, °	1.66	
mean *B*-factor (standard deviation), Å^2		
protein	12.5 (5.9)	
solvent	22.8 (12.3)	
mean real space fit[g] (standard deviation), %		
protein	89.7 (3.8)	
solvent	83.1 (6.6)	

[a] The unit cell dimensions for the cross-linked crystal are 77.2 × 55.8 × 53.7 Å in water (3) and 77.1 × 55.4 × 53.6 Å in acetonitrile (2), and for the non-cross-linked crystal in aqueous buffer are 76.7 × 55.6 × 53.1 Å (9). [b] *R*-factor on the intensity for symmetry-related measurements. [c] *R*-factor on structure factors for reflections used in the refinement (90% of total). [d] *R*-factor on structure factors for reflections randomly omitted from the refinement and used as the test set (10% of total). [e] Only atoms included in the refinement are shown. A Luzzati analysis revealed the average coordinate error to be 0.3 Å. [f] Root mean square deviation from ideal geometric values (Enge and Huber parameter set). [g] Real space fit values calculated from the $|2F_o-F_c|$ electron density maps.

Space group determination. Pseudo-precession pictures were generated revealing two-fold rotation axes on *h*, *k*, and *l*, placing the crystal in the 222 rotation group. Systematic extinctions revealed *h* to be a screw axis, while *k* and *l* remained ambiguous. Thus, the possible space groups at this point were $P222_1$, $P2_12_12$, and $P2_12_12_1$. Molecular replacement translation searches (see below) only yielded a real solution, a peak of 12.8σ, in the $P2_12_12_1$ space group; no peaks were found above 7.0σ for the other space groups. Thus, the space group was $P2_12_12_1$, the same as for the cross-linked crystal in both water and acetonitrile.

Molecular replacement. Initial phases were determined by molecular replacement using the program AMORE (19) in the CCP4 Suite with the 2.75 Å subtilisin structure in aqueous buffer reported by Neidhart & Petsko (9) as the search model. The rotation search using the 10 to 4.0 Å data revealed one unique solution at 9.7σ (standard deviation above the mean rotation function). The translation search in the $P2_12_12_1$ space group yielded a single peak of 12.8σ. Comparison of the molecular replacement solution to that of the original search model showed that the RMS displacement was only 0.001 Å, i.e., no significant crystal rearrangement was observed for the cross-linking and solvent transfer steps. A series of $|2F_{obs}-F_{calc}|$ maps were calculated using phases determined from models which had sequential regions omitted (11% of the total atoms). The omit maps showed clear density for all omitted regions with the exception of residues 158-162 (see below).

Refinement and model building. After molecular replacement the *R*-factor for the search model was 0.30 for all the reflections in the 6.0-2.6 Å resolution shell and the

R-free (20), consisting of reflections that are left out of the refinement (10% of total), was 0.32. The search model was first rebuilt residue by residue into the omitted regions using the omit maps generated as above, adjusting side chain torsion angles and occasionally changing residue rotamers. Approximately 25% of the residues required some adjustment to center them in their electron density. Two rounds of refinement using X-PLOR (21) followed by manual rebuilding using the program O (22) were then performed on the protein structure. The refinement protocol in X-PLOR consisted of: a rigid body refinement (40 cycles using the 10-4.0 Å data) followed by simulated annealing (6.0-2.6 Å), a conventional positional refinement (50 cycles, 6.0-2.6 Å), and a restrained group *B*-factor refinement (20 cycles, 6.0-2.6 Å). The simulated annealing runs involved heating to 3000 K and then cooling in increments of 25 K with 50 steps of molecular dynamics simulations for 0.5 fs at each temperature. At this stage, the *R*-factor and *R*-free were brought down to 0.22 and 0.30, respectively, in the 6.0-2.6 Å resolution shell.

Subtilisin contains a flexible surface loop, Asn158 to Thr162, for which initially very weak density was observed. This entire loop was omitted during the early refinement cycles. Residues of the loop were gradually included in the model in later refinement rounds. However, Ser159 and Gly160 did not have well-defined electron density even in the final maps, and they were therefore not included in the final model. This loop also appears to be very flexible in the cross-linked crystal in water, reflected by high *B*-factors of residues 158-162.

Solvent molecules. Two rounds of manual building in O (22) and refinement with X-PLOR (21) were performed resulting in the addition of the first 35 waters.

Potential waters were located using the $|F_{obs}-F_{calc}|$ electron density map at the 3.0σ contour level (23). Initially, they were only built into density peaks that clearly exhibited the appropriate shape; any peaks that were possibly due to dioxane molecules were excluded. The refinement protocol using X-PLOR was the same as described above. Potential waters were retained if their electron density in subsequent $|2F_{obs}-F_{calc}|$maps (23) persisted after the simulated annealing refinement, and if they fulfilled the following criteria: within 3.4 Å of a subtilisin oxygen or nitrogen atom (or bound water in the second round) with good hydrogen-bonding geometry, *B*-factors less than 45 $Å^2$, and real space fit correlation coefficients above 65%. After inclusion of these 35 waters, the *R*-factor and *R*-free were brought down to 0.286 and 0.167, respectively, for the data in the 6.0-2.6 Å resolution shell. Three rounds of manual building in O and refinement with the program REFMAC (24) were carried out resulting in the inclusion of 30 additional waters, for a total of 65 water molecules in the model. The water identification and manual building was performed as above, but the refinement was done using REFMAC, where a combined positional and a restrained individual *B*-factor refinement was performed using the Maximum Likelihood residual and all the working data, 14-2.6 Å. Criteria for acceptable waters were the same as above. The inclusion of the final 30 waters resulted in an improvement in *R*-free to 0.276 with no change in the *R*-factor (0.167).

Dioxane molecules were introduced into the model by means of four rounds of manual building in O and automatic refinement using X-PLOR. The dioxane model used has a chair conformation as reported by Buschmann *et al.* (25) for the phase I crystal at 279 K. The dioxane parameter and topology files for X-PLOR were generated using

XPLO2D (26). Potential dioxane molecules were located using the $|F_{obs}-F_{calc}|$ map. Electron density was considered to represent a potential dioxane if the peak was above 3.0σ with at least one dioxane oxygen within hydrogen-bonding distance to a hydrogen-bond donor on the protein or water (Fig. 1.5A). Dioxane electron density was recognized by comparison with F_{calc} maps calculated for a theoretical dioxane molecule between the resolution limits of 14-2.6 Å (Fig. 1.5B).

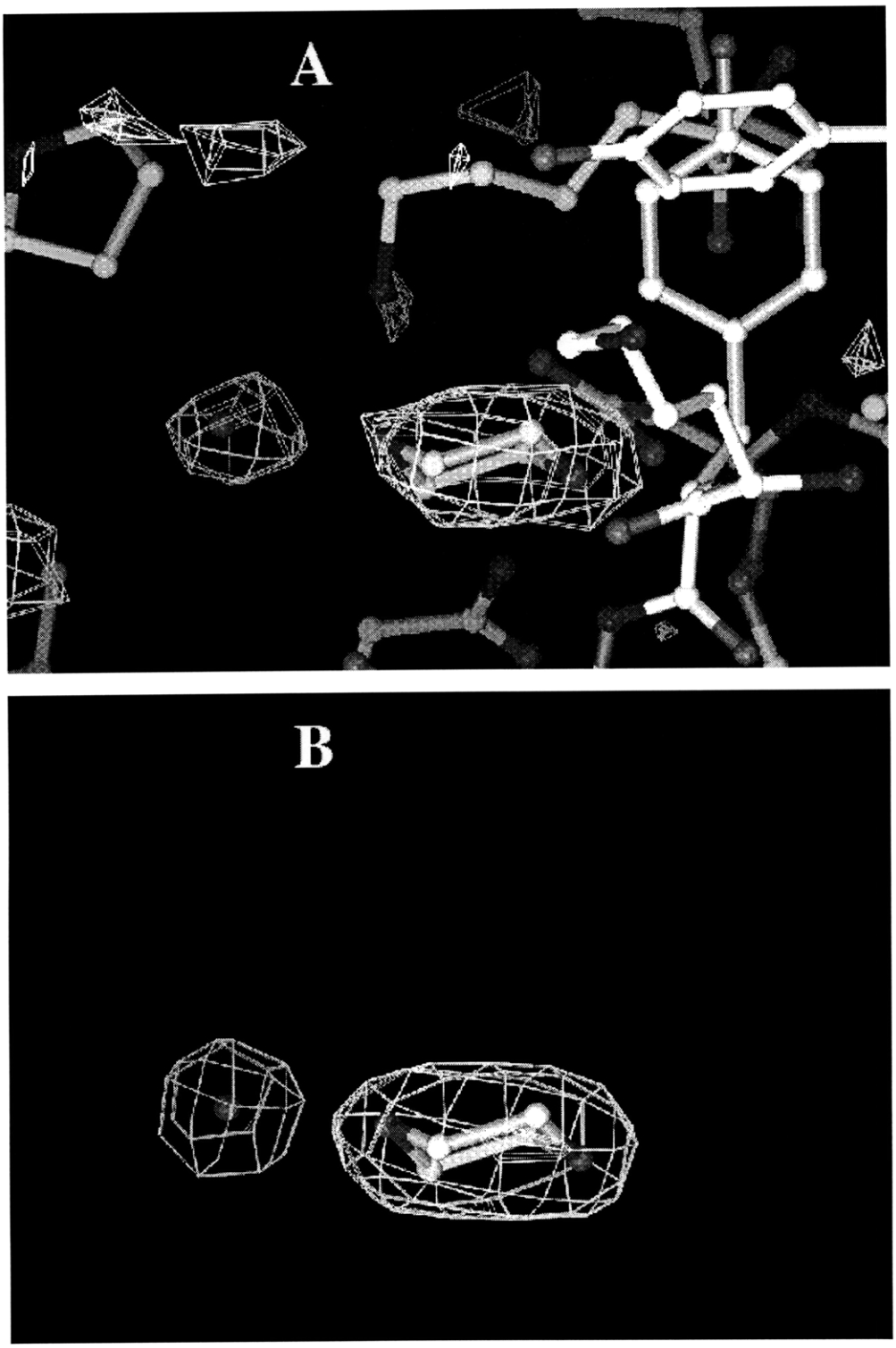


Figure 1.5. (A) Ball-and-stick model of a dioxane molecule and a water molecule binding at a subtilisin crystal contact. Carbon, oxygen, and nitrogen are depicted in white, light gray, and dark gray, respectively. $|F_{obs}-F_{calc}|\phi_{calc}$ map, omitting solvent molecules, shown in white at 3σ contour. (B) The solvent molecules are displayed as in A; shown in white is the F_{calc} theoretical electron density map for the water and dioxane molecules.

A complication that must be noted is that the electron density for a single dioxane molecule at this resolution appears similar to the density for two hydrogen-bonded water molecules (the O-O distance in dioxane is 2.68 Å). In a few cases, the hydrogen-bonding pattern in the model was consistent only with two waters; however, when it was ambiguous, the following criteria were used to distinguish between the possibilities of the density corresponding to one dioxane or two waters. Initially, a real space fit correlation coefficient was calculated for both a dioxane molecule and two waters built into such an electron density peak. A conventional positional refinement, 40 cycles, followed by a restrained individual *B*-factor refinement, 20 cycles, (both using the 6.0-2.6 Å data) was performed using X-PLOR (21) on the model with either one dioxane molecule or two waters. The respective real space fit and *R*-free values were calculated; dioxane molecules were kept only if these parameters, before and after refinement, were superior to those for the two waters. Furthermore, putative dioxane molecules with average *B*-factors greater than 50 $Å^2$ or real space fit correlation coefficients less than 60% were discarded. Seven dioxane molecules were included in the final model. The final *R*-factor and *R*-free for the 6.0-2.6 Å resolution shell were 0.161 and 0.246, respectively (Table 1.1). The coordinates for this structure can be found in the Protein Data Bank (PDB; Brookhaven National Laboratory, entry 1af4).

Structure comparisons. The structure of subtilisin Carlsberg in dioxane was superimposed with the previously determined structures in distilled water (3) and acetonitrile (2) using LSQMAN (27) to minimize the RMS distance of the main chain atoms between the appropriate pair of structures. The RMS distances between subtilisin in dioxane and in distilled water and between dioxane and acetonitrile were 0.33 and 0.36

Å, respectively. The identification of the binding pockets in the enzyme active site was accomplished by superimposition (as above) with the structure of subtilisin bound to the protease inhibitor eglin c (28). Surface area calculations were performed using the Connolly algorithm (29) with a probe radius of 1.4 Å in the InsightII package from Biosym. The percentages of the exposed surface areas of the dioxane and acetonitrile molecules in their respective structures were calculated as follows: the surface area ($Å^2$) of a given bound dioxane or acetonitrile molecule in the structure was calculated and then expressed as a percentage of the surface area ($Å^2$) calculated for the free dioxane or acetonitrile molecule. The crystal contacts, as well as the root mean square deviations from ideality of the model, were determined using the geometric analysis features in X-PLOR (21).

The number of intramolecular hydrogen bonds in subtilisin was determined using the program HBPLUS (McDonald, I., Naylor, D., Jones, D., Thornton, J.M., 1994). Hydrogen bonds were considered to exist between hydrogen-bond donor and acceptor atoms if the distance between them was less than 3.4 Å and the donor-hydrogen-acceptor angle was greater than 90°.

D. References

1 Koskinen, A. M. P.; Klibanov, A. M., Eds. *Enzymatic Reactions in Organic Media*. Blackie: London, 1996.

2. Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klibanov, A. M. *Proc. Natl. Acad. Sci. USA* **1993** *90*, 8653-8657.

3. Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675-681.

4. Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. *Biochemistry* **1994**, *33*, 7326-7336.

5. Yennawar, H. P.; Yennawar, N. H.; Farber, G. K. *J. Am. Chem. Soc.* **1995**, *117*, 577-585.

6. Allen, K. N.; Bellamacina, C. R.; Ding, X.; Jeffery, C. J.; Mattos, C.; Petsko, G.A.; Ringe, D. *J. Phys. Chem.* **1996**, *100*, 2605-2611.

7. Wescott, C. R.; Hidetaka, N.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 10365-10370.

8. Ke, T.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3366-3374.

9. Neidhart, D. J.; Petsko, G. A. *Protein Eng.* **1988**, *2*, 271-276.

10. Hartsough, D. S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 10113-10116.

11. Hartsough, D. S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 6529-6537.

12. Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 1629-1631.

13. Voet, D.; Voet, J. G. *Biochemistry*, 2nd Ed; Wiley: New York, pp. 398-400, 1995.

14. Giegé, R.; Ducruix, A. in *Crystallization of Nucleic Acids and Proteins* Ducruix, A.; Giegé, R. eds. Oxford University Press: New York, 1992, pp. 1-18.

15. Mattos, C.; Ringe, D. *Nature Biotechnol.* **1996**, *14*, 595-599.

16. Laitinen, H. A.; Harris, W. E. *Chemical Analysis*, 2nd Ed. McGraw-Hill: New York, 1975, pp. 361-363.

17. Otwinowski, Z.; Minor, W. *Meth. Enzymol.* **1997**, *276*, 307-326.

18. Collaborative Computational Project, Number 4 *Acta Cryst.* **1994**, *D50*, 760-763.

19. Navaza, J. *Acta Cryst.* **1994**, *D50*, 157-163.

20. Brünger, A.T. *Nature* **1992**, *355*, 472-475.

21. Brünger, A. T.; Kuriyan, J.; Karplus, M. *Science* **1987**, *35*, 458-460.

22. Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Cryst.* **1991**, *A47*, 110-119.

23. McRee, D. E. *Practical Protein Crystallography* Academic Press: San Diego, 1993, pp. 226-227.

24. Pannu, N. S.; Read, R. J. *Acta Cryst.* **1996**, *A52*, 659.

25. Buschmann, J.; Müller, E.; Luger, P. *Acta Cryst.* **1986**, *C42*, 873-876.

26. Kleywegt, G. J. *ESF/CCP4 Newsletter* **1995**, *32*.

27. Kleywegt, G. J. *ESF/CCP4 Newsletter* **1994**, *31*.

28. Bode, W.; Papamokos, E.; Musil, D. *Eur. J. Biochem.* **1987**, *166*, 673-692.

29. Connolly, M. L. *Science* **1983**, *221*, 709-712.

II. The Mechanistic Dissection of the Plunge in Enzymatic Activity upon Transition from Water to Anhydrous Solvents

A. Introduction

Enzymes have been found to possess unique properties in nonaqueous media, e.g., the ability to catalyze transformations virtually impossible in aqueous solution (1) However, a major drawback limiting the utility of enzymes is that their catalytic activity in anhydrous organic solvents is drastically diminished relative to that in water. For example, lyophilized powder of the serine protease subtilisin Carlsberg is some 6 orders of magnitude less active when suspended in acetonitrile than the soluble enzyme is in water (2). A tempting explanation for this drop in activity is that a conformational change inflicted upon the enzyme by lyophilization and/or subsequent placement in the solvent is responsible (3). Recently, however, it was demonstrated that this could be no more than a minor part of the explanation because the crystal structure of cross-linked crystals (CLCs) of subtilisin in anhydrous acetonitrile was found to be indistinguishable from that in water (4,5),* and yet the CLCs are still far less catalytically active in acetonitrile than in water (4).

In the present study, we addressed this issue mechanistically using the aforementioned CLCs of subtilisin as a model system. Since the structure of this crystalline enzyme preparation is the same in acetonitrile and in water (4,5), the solvent-

* It was also found in chapter I of this thesis that the structure of subtilisin also does not change upon transition from acetonitrile to other organic solvents, specifically dioxane (6).

induced conformational change hypothesis, while viable with the amorphous enzyme, may not be invoked. In addition, because CLCs are highly stable and possess other attractive properties in aqueous and aqueous-organic mixtures (7), they may turn out to be the future catalyst of choice in both aqueous and nonaqueous enzymology.

B. Results and Discussion

In order to approach the explanation of the difference in catalytic activity of enzymes in organic solvents *vs.* in water, we selected a model system on which to focus our study, namely, the difference in the catalytic activity of soluble subtilisin in aqueous solution at pH 7.8 for the hydrolysis of N-Ac-L-Phe-OEt and that of the CLCs of subtilisin in anhydrous acetonitrile for the transesterification of this ester with propanol. We found that V_{max}/K_M plummets more than 7 orders of magnitude upon transition from the aqueous to organic system,[†] from 2.1 s^{-1} to 4.5×10^{-8} s^{-1} (1 mg/mL enzyme).

Two factors could contribute to the observed reduction in V_{max}/K_M, namely, a reduction in the catalytic efficiency of the enzyme (k_{cat}/K_M) or in the concentration of catalytically competent enzyme ($[E]_0$). Although the conformation of cross-linked crystalline subtilisin suspended in acetonitrile is virtually identical to that in water (4,5) (indicating that subtilisin in the CLCs is native), it is possible that only a small fraction of the active centers in the CLCs is accessible to the substrate because of hindering

[†] V_{max}/K_M in the two systems can be directly compared as long as they involve the same ester substrate. This is because $(k_{cat}/K_M)_{N\text{-}Ac\text{-}L\text{-}Phe\text{-}OEt}$, the bimolecular rate constant for the reaction of the free enzyme with this ester substrate, is independent of the nucleophile substrate – whether water (in aqueous solution) or an alcohol (in organic solvents) (2).

protein-protein contacts or cross-links within the crystal, resulting in a reduced $[E]_0$. To test for such an effect, the percentage of catalytically competent active centers was measured by the direct titration of Ser 221 of subtilisin's catalytic triad (see Methods for details). The fractions of active enzyme dissolved in aqueous solution and in CLCs suspended in acetonitrile were found to be similar, 63±5% and 33±14%, respectively, thus ruling out the possibility that the 7-order-of-magnitude drop in activity could be due to a decrease in $[E]_0$. When the $[E]_0$ values determined from the active center titrations are used to convert V_{max}/K_M to k_{cat}/K_M, the diminished activity of CLCs in acetonitrile compared to that of the dissolved enzyme in water is unequivocally attributable to decreased catalytic efficiency of the enzyme (Table 2.1).

Table 2.1. Values of $(k_{cat}/K_M)_{N\text{-}Ac\text{-}L\text{-}Phe\text{-}OEt}$ for the subtilisin-catalyzed hydrolysis of N-Ac-L-Phe-OEt or its transesterification with propanol

Form of subtilisin[a]	Solvent	k_{cat}/K_M[b] ($M^{-1}s^{-1}$)
dissolved	water (pH 7.8)	$(9.1\pm1.3)\times10^4$
CLCs	water (pH 7.8)	$(3.4\pm0.5)\times10^3$
CLCs	octane (a_w = 1)	40 ± 20
CLCs	octane (a_w= 0.006)	0.93 ± 0.40
CLCs	acetonitrile (a_w = 0.006)	$(2.3\pm0.5)\times10^{-2}$
CLCs	acetonitrile ($a_w \leq 0.002$)	$(1.1\pm0.4)\times10^{-2}$

[a] All organic solvents contained 100 mM propanol as the nucleophile. [b] The values of the portion of the CLCs that is active subtilisin, used to calculate $[E]_0$, are 57±20% in water, 45±24% in octane, and 33±14% in acetonitrile. These values were in turn used to calculate k_{cat}/K_M from V_{max}/K_M.

When one compares soluble subtilisin in water to CLCs of subtilisin in acetonitrile, two parameters have been changed simultaneously, namely the solvent and the form of the enzyme. In order to determine the contribution of each factor individually, we examined the behavior of the CLCs relative to the soluble enzyme in water, thus changing only the form of the enzyme, while keeping the solvent the same. The k_{cat}/K_M value of the CLCs in water at pH 7.8 was found to be 1.4 orders of magnitude lower than that of soluble subtilisin (Table 2.1). Therefore, the first question to be addressed was why this is so.

One plausible explanation is that the reaction catalyzed by the CLCs is slowed by internal diffusion of the substrate within them. To test this, we co-crystallized active and inactivated (with phenylmethylsulfonyl fluoride) subtilisin at various *f* values (where *f* is the fraction of the active subtilisin in the CLCs) and studied the dependence of the enzymatic activity on *f*. If the reaction is limited by internal diffusion, this dependence should be convex (8). If, however, the reaction is free of diffusional limitations (internal or external), a linear dependence should result (8) As seen in Figure 2.1, the dependence of the activity of the CLCs on *f* is essentially linear, thus ruling out diffusional limitations as the cause of the 27-fold drop in activity of the CLCs in water compared to the soluble enzyme. Note that if the enzymatic activity of the CLCs of subtilisin is not limited by diffusion in water, it certainly will not be limited by diffusion in organic solvents, where the reaction is far slower (Table 2.1) and diffusion faster (because the solvents' viscosities are lower than water's) (8).

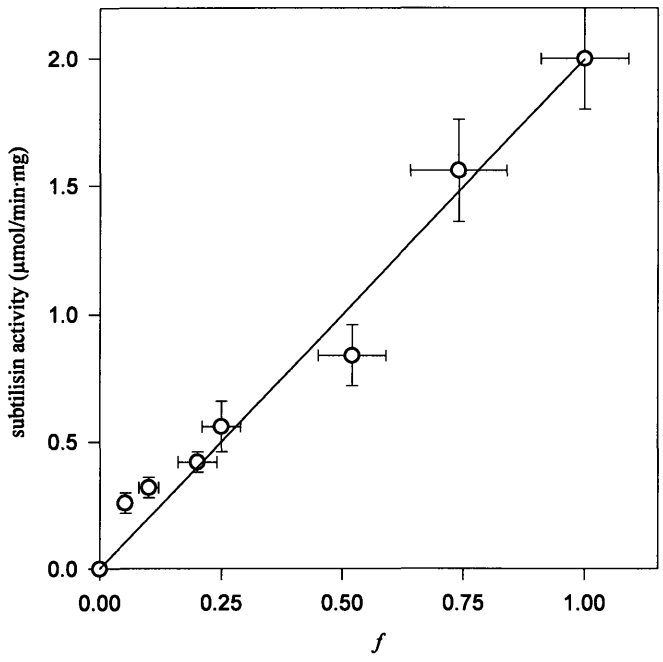


Figure 2.1. Dependence of the catalytic activity of CLCs of subtilisin in the hydrolysis of N-Ac-L-Phe-OEt in water (pH 7.8) on the fraction (*f*) of the active subtilisin in the CLCs. The value of *f* was manipulated by co-crystallizing mixtures having different ratios of active subtilisin and PMSF-inactivated subtilisin.

Another tenable reason for the difference in activity of the CLCs compared to soluble subtilisin in water at pH 7.8 is that the activity *vs.* pH profiles of the two forms of the enzyme are distinct. Indeed, we found (Figure 2.2) that the pH dependence of the activity of the CLCs is markedly shifted (by 3.7 pH units) compared to that of the dissolved enzyme. This shift can readily account for the 1.4-order of magnitude difference in the k_{cat}/K_M values between the cross-linked crystalline and dissolved

subtilisins at pH 7.8.‡ Therefore, the first drop in enzymatic activity that occurs upon transition from soluble subtilisin in aqueous solution at pH 7.8 to the CLCs in dry acetonitrile can be explained by the shift of the activity *vs.* pH profile of the CLCs relative to that of the enzyme in solution.

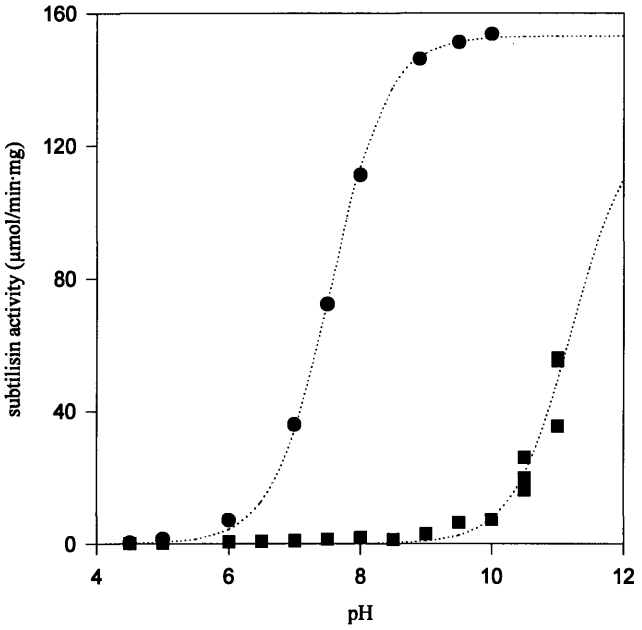


Figure 2.2. The pH dependence of the initial rate (v) of the enzymatic hydrolysis of N-Ac-L-Phe-OEt in water for dissolved subtilisin (●) and the CLCs of subtilisin (■). The data were fitted to the theoretical ionization curve (-------) by nonlinear regression, where $v = \frac{A \times 10^{(pH-pK_a)}}{(1+10^{(pH-pK_a)})}$, where A is a constant.

‡ Although the exact cause of the shift in the activity *vs.* pH profile is not known, at least two possible explanations seem likely. In the crystalline state, the subtilisin molecules are held in a close proximity and fixed orientation to each other (packed in the crystal lattice). The ensuing intermolecular electrostatic interactions could affect the pK_as of the active center's catalytic residues, leading to the observed effect. Alternately, due to the overall positive charge of subtilisin, whose pI in solution is 9.4 (Ottesen, M.; Svendsen, I. *Meth. Enzymol.* **1970**, *19*, 199-215), a partitioning of hydrogen ions away from the crystal could occur. This would lower the effective pH in the crystal relative to that in the bulk solution and result in a shift in the activity *vs.* pH profile.

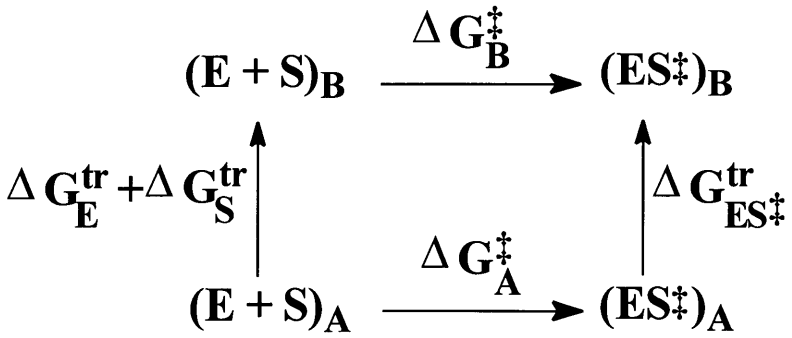
Having accounted for 1.4 orders of magnitude, there is still a 5.6-order of magnitude decrease in activity to be rationalized that transpires upon moving the CLCs from water to anhydrous acetonitrile. Two factors, which have previously been shown to influence the kinetics of enzymes suspended in organic solvents, are likely to be responsible. The first is the difference between the desolvation free energies of the substrate, N-Ac-L-Phe-OEt, in acetonitrile and water. Because the binding energy of a substrate is utilized for enzyme catalysis (9), and this binding of a substrate molecule to an enzyme active center requires the desolvation of the substrate, k_{cat}/K_M depends on the free energy of substrate desolvation (10). The second likely contributor is the dehydration of the CLCs, i.e., their exposure to a much lower thermodynamic water activity (a_w). For lyophilized powders of enzymes in general (11), and of subtilisin in particular (12), increasing the a_w in organic solvents leads to enhanced enzymatic activity, presumably because water acts as a molecular lubricant (13), increasing the conformational mobility of the enzyme. We endeavored to isolate these two factors and elucidate their individual contributions to the decrease in catalytic efficiency of the CLCs upon transition from water to dry acetonitrile.

The two effects can be segregated using the following four-staged stepladder (Table 2.1). Because in anhydrous acetonitrile the a_w is less than 0.002 but not known exactly, the first move is to acetonitrile at a still low but defined a_w (step 1). Salt hydrate pairs can be used in organic solvents (where they are insoluble) to buffer the a_w to a constant value (14). We selected the pair $BaBr_2 \cdot 1H_2O/BaBr_2$ which buffers the a_w to 0.006. The next transition is from acetonitrile with a_w=0.006 to octane with the same a_w to isolate the contribution of the substrate desolvation energy difference (step 2); the

introduction of this intermediate solvent is necessary because it is impossible to attain the same water activity, a_w=1, in acetonitrile (infinitely miscible with water) and in water. The subsequent shift from octane with a_w=0.006 to that with a_w=1 addresses the water activity effect alone (step 3). The final move from octane with a_w=1 to water should again indicate the role of substrate desolvation in these solvents (step 4).

The aforementioned thermodynamic effect, which arises from the difference in the substrate desolvation energy between two solvents and the resulting difference in reactivity toward the enzyme, is depicted by the cycle in Scheme 1 below (10).

Scheme 2.1.



The lower horizontal arrow represents the enzyme (E) reacting with the substrate (S) in solvent A to form the transition state ($ES^{\ddagger}$). In a hypothetical, equivalent path, E and S separately partition from solvent A into solvent B, the transition state is formed there and is transferred from solvent B back into A. In the cycle, $\Delta G^{\ddagger}_{A}$ and $\Delta G^{\ddagger}_{B}$ represent the free energies of activation in solvents A and B, respectively, and ΔG^{tr}_{E}, ΔG^{tr}_{S}, and $\Delta G^{tr}_{ES^{\ddagger}}$ are the free energies of transfer of the enzyme, substrate, and transition state, respectively,

from A to B. Expressing $\Delta G^{\ddagger}_{A}$ as the sum of the energetic terms of this alternative path produces

$$\Delta G^{\ddagger}_{A} = \Delta G^{tr}_{E} + \Delta G^{tr}_{S} + \Delta G^{\ddagger}_{B} - \Delta G^{tr}_{ES\ddagger} \quad (1)$$

Assuming that the entire substrate is removed from the solvent in the transition state, the solvation of E and $ES^{\ddagger}$ is essentially identical, provided that there is no substrate-induced conformational change in the enzyme[§] and that desolvation of the active center can be neglected. Thus, $\Delta G^{tr}_{E} \approx \Delta G^{tr}_{ES\ddagger}$. (If the assumption does not hold, then the effect described below will be less.) Canceling the equal terms yields

$$\Delta G^{\ddagger}_{A} = \Delta G^{\ddagger}_{B} + \Delta G^{tr}_{S} \quad (2)$$

$\Delta G^{\ddagger}_{A}$ is related to $(k_{cat}/K_M)_A$ by (9)

$$\Delta G^{\ddagger}_{A} = -RT\ln\left[\left(\frac{k_{cat}}{K_M}\right)_A\left(\frac{h}{\kappa T}\right)\right] \quad (3)$$

where h, κ, and R are Planck's, the Boltzmann, and the gas constants, respectively, and T is the absolute temperature. The analogous expression can be written for solvent B.

[§] Serine proteases do not act via an induced-fit mechanism (9). The X-ray crystal structures of subtilisin Carlsberg in its native state and that inhibited by eglin c are indistinguishable (4).

ΔG_S^{tr} can in turn be expressed in terms of the thermodynamic activity coefficient of the substrate** (γ):

$$\Delta G_S^{tr} = RT \ln\left(\frac{\gamma_B}{\gamma_A}\right) \tag{4}$$

Substitution of equations 3 and 4 into equation 2 yields:

$$\frac{\left(\frac{k_{cat}}{K_M}\right)_A}{\left(\frac{k_{cat}}{K_M}\right)_B} = \frac{\gamma_A}{\gamma_B} \tag{5}$$

This predictive model, culminating in equation 5, was tested in step 2 of the ladder. In this step, when the situation in octane with a_w=0.006 is compared to acetonitrile with a_w=0.006, the experimentally obtained k_{cat}/K_M ratio of 40±19 (Table 2.1) was similar to the value predicted by the activity coefficient ratio of N-Ac-L-Phe-OEt in these two solvents, 63±6 (Table 2.2). These data support the thermodynamic model that predicts changes in k_{cat}/K_M based on variations in the substrate desolvation energetics.

** The free energy of a solute dissolved in a solvent is described by the equation $G = G° + RT \ln (x\gamma)$, where x and γ are the solute mole fraction and activity coefficient, respectively. Because the solvent is the primary variable in our work, the standard state is chosen as the pure liquid solute. Thus G° is independent of the solvent, and the free energy of transfer of the solute from solvent A to solvent B (ΔG_S^{tr}) is $RT \ln(x_B\gamma_B/x_A\gamma_A)$. If the transfer is made at constant mole fraction (i.e., the volumes of the solvents are directly proportional to their molar volumes), this simplifies to $\Delta G_S^{tr} = RT \ln(\gamma_B/\gamma_A)$. For dilute solutions, the concentration dependence of the activity coefficient can be disregarded. Thus, the variation in x which occurs in our experiments (e.g., due to concentration variation in the measurement of k_{cat}/K_M or differences in the molar volumes of the solvent) is neglected.

Table 2.2. Thermodynamic activity coefficients γ of N-Ac-L-Phe-OEt in various solvent systems

Solvent[a]	$\gamma_{\text{N-Ac-L-Phe-OEt}}$[b]
water (pH 7.8)	$(2.6\pm0.2)\times10^3$
hydrous octane ($a_w = 1$)	$(8.2\pm0.2)\times10^2$
octane ($a_w = 0.006$)	$(8.2\pm0.3)\times10^2$
acetonitrile ($a_w = 0.006$)	13±1
acetonitrile ($a_w \leq 0.002$)	13±1

[a] All organic solvents contained 100 mM propanol. [b] Calculated from the solubility of the ester in the appropriate solvent system (see Methods).

In addition to describing the solvent dependence of the acylation step of subtilisin-catalyzed transesterifications, Scheme 1 is equally applicable to the deacylation step. The model predicts that a difference in the desolvation energy of the nucleophile substrate (as opposed to the ester substrate) between two solvents will also result in a k_{cat}/K_M difference. In agreement with this, when a small activity coefficient ratio was calculated, e.g., for propanol in acetonitrile *vs.* in *tert*-amyl alcohol (2.7), the corresponding experimentally determined k_{cat}/K_M ratio was also small (2.2). When a larger difference in $(k_{cat}/K_M)_{\text{nucleophile}}$ was predicted by the activity coefficient ratio, as is the case for 2-methoxyethanol in cumene *vs.* in acetonitrile (14), the experimental ratio of the specificity constants was also higher (7.6).[††]

[††] The water activity was not buffered in the solvents in which $(k_{cat}/K_M)_{\text{nucleophile}}$ was determined.

On the basis of all these data, one can conclude that the difference in the desolvation energy of N-Ac-L-Phe-OEt in water compared to dry acetonitrile can account, as indicated by $\gamma_{water}/\gamma_{acetonitrile}$ (Table 2.2), for 2.3 orders of magnitude of the difference in k_{cat}/K_M of the CLCs in these solvents. Hence only 3.3 orders of magnitude out of the original 7 are now left to explain.

As discussed earlier, the water activity may have a profound effect on the catalytic efficiency of the CLCs in organic solvents. Indeed, as seen in step 3 of the ladder (Table 2.1), k_{cat}/K_M of the CLCs in octane at a_w=1 is 43-fold higher than that in the same solvent at a_w=0.006. In addition, a doubling in the activity of the CLCs is observed in step 1, i.e., in acetonitrile with a_w=0.006 compared to this solvent with a_w<0.002. As mentioned above, one of the likely explanations for the increase in enzymatic activity with a_w is that water acts as a molecular lubricant and affords a higher conformational flexibility of the enzyme resulting in enhanced catalysis. Another possibility (15) is that the added water increases the dielectric constant (polarity) of the solvent and thus in the enzyme active center, with an ensuing stabilization of the charged transition state and hence an elevation in k_{cat}/K_M.

To distinguish between these two possibilities, we investigated the effects of two selected additives on k_{cat}/K_M in octane. The first, formamide, is a good molecular lubricant due to its high propensity to form hydrogen bonds (16) and the second, N-methylacetamide, while a poorer molecular lubricant (because it can form fewer hydrogen bonds), has a higher dielectric constant than formamide (191 *vs.* 111) (17). If the flexibility hypothesis is correct, formamide should increase the observed catalytic

activity in octane to a greater extent than N-methylacetamide. If the polarity hypothesis is correct, the opposite will be the case.[‡‡] While the addition of formamide (at a=1) raises k_{cat}/K_M by a factor of 13 (from 0.93±0.40 to 12 ±5 $M^{-1}s^{-1}$), no enhancement occurs when N-methylacetamide is added to the same activity – k_{cat}/K_M=0.79±0.30 $M^{-1}s^{-1}$. (The value of $\gamma_{N\text{-}Ac\text{-}L\text{-}Phe\text{-}OEt}$ in octane is unaffected by these additives.) These data support the hypothesis that an elevated a_w leads to a greater conformational mobility of the enzyme that translates into a higher k_{cat}/K_M. Therefore, the depressed enzymatic flexibility of the CLCs in dry acetonitrile (a_w<0.002) can account for 1.9 orders of magnitude of activity lost between water and acetonitrile. (This value is obtained by adding the effects of going from a_w<0.002 to a_w=0.006 (0.28, step 1) and from a_w=0.006 to a_w=1 (1.6, step 3) in Table 2.1.)

Thus only 1.4 out of the original 7 orders of magnitude are now left unexplained. This gap is seen in Table 2.1 as the difference between CLCs in water and CLCs in octane with a_w=1, where the relative substrate desolvation energies (Table 2.2) predict only a 3.2-fold drop in catalytic efficiency as opposed to the 85-fold observed. The exact origin of this disparity is still unclear. One possibility is that the activity *vs.* pH profile in organic solvents is responsible, i.e., the protonation state of the catalytic triad's histidine is different for CLCs suspended in organic solvents than it is in water at pH 7.8. In fact, we did find that the ratio of $(k_{cat}/K_M)_{N\text{-}Ac\text{-}L\text{-}Phe\text{-}OEt}$ for CLCs at pH 7.8 *vs.* pH 4.5 in water to

[‡‡] The dielectric constants for these solvent systems are 2.0, 2.1, 2.2, and 2.6 for pure octane, octane saturated with water, octane saturated with formamide, and octane saturated with N-methylacetamide, respectively. These values were calculated as described by Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486-1493.

be 34-fold; whereas the $(k_{cat}/K_M)_{N\text{-}Ac\text{-}L\text{-}Phe\text{-}OEt}$ ratio in octane for CLCs that were removed from aqueous buffers of pH 7.8 *vs.* pH 4.5 is only 7-fold. Alternatively, the solvent binding in the active center could compete with the substrate and thus inhibit subtilisin in organic solvents. Four organic solvent molecules bind in subtilisin's active center region in the CLCs in acetonitrile (4,5). Also, in the γ-chymotrypsin crystal structure in hexane (a solvent homologous to octane), two hexane molecules bind in the vicinity of the active center (18). This interesting possibility of the existence of specific interactions between the solvent and the enzyme active site is investigated further in Chapter III.

C. Concluding Remarks

Using a systematic approach to explain the 7-order of magnitude drop in catalytic activity of CLCs of subtilisin in anhydrous acetonitrile relative to the enzyme dissolved in aqueous solution, we have been able to account for 5.6 of those by: (i) a shift in the activity *vs.* pH profile of the CLCs in water relative to the aqueous enzyme solution; (ii) the unfavorable desolvation energy of the substrate in the organic solvent relative to water; and (iii) the diminished conformational flexibility of the CLCs in the solvent due to dehydration. Consequently, this study has provided the rationale for avoiding much of this activity loss by: optimizing the pH of the solution from which the CLCs are collected, judiciously selecting the organic solvent to minimize the unfavorable substrate desolvation energetics, and maintaining a high water activity in the system.

D. Materials and Methods

Cross-linked enzyme crystals (CLCs) of subtilisin. Subtilisin Carlsberg (serine protease from *Bacillus licheniformis*, EC 3.4.21.14) was purchased from Sigma Chemical Co. Crystals were grown at 30 °C from an aqueous 330 mM cacodylate buffer, pH 5.6, saturated with Na_2SO_4 (19). The crystals were cross-linked with a 1.5% glutaraldehyde solution (pH 7.5, 30 mM cacodylate buffer, 13% Na_2SO_4), followed by washing with the buffer (containing no glutaraldehyde) and distilled water, and stored in a 20 mM phosphate buffer, pH 7.8, at 10 °C (4,5). The average dimensions of the needle-like CLCs were (100±40)μm × (15±5)μm × (15±5)μm. Note that the size of the CLCs did not change upon transition from water to acetonitrile and octane, as evidenced by microscopic examination. Furthermore, the unit cell dimensions of the subtilisin CLCs in acetonitrile are essentially the same as in water (4,5).

Substrates, products, and solvents. N-Ac-L-Phe-OEt, N-*trans*-cinnamoylimidazole, and phenylmethylsulfonyl fluoride (PMSF) were purchased from Sigma. The organic solvents, including alcohols, utilized were reagent grade and were dried over 3-Å molecular (Linde) sieves prior to use to a water content below 0.01%. Anhydrous $BaBr_2$ was purchased from Aldrich Chemical Co. The mixture of $BaBr_2 \cdot 1H_2O/BaBr_2$ was prepared by placing the anhydrous salt (5 g / 16.8 mmol) over pure water (75.8 mg / 4.2 mmol) in a sealed vessel saturated with water vapor.

One of the product esters, N-acetyl-L-phenylalanine 2-methoxyethyl ester (N-Ac-L-Phe-OEtOMe), used for GC calibration, was synthesized as follows. To a mixture of 2 g of N-Ac-L-Phe (10 mmol) in 45 mL of benzene, 4.5 mL of 2-methoxyethanol (60

mmol) and 4 drops of 95% H_2SO_4 were added. The reaction mixture was refluxed for 3 h and then transferred to a separatory funnel containing 100 mL of ether, where it was washed 5 times with 50 mL of a 5% aqueous solution of $NaHCO_3$. The ether phase was dried over anhydrous $MgSO_4$ and concentrated by rotary evaporation; the product ester crystallized out of the solution. The purity (99%) of the product N-Ac-L-Phe-OEtOMe was verified by GC. N-acetyl-L-phenylalanine propyl ester was similarly synthesized from the free acid and propanol, and its purity (99%) was likewise confirmed.

Phenylmethylsulfonic acid (PMSOH) was synthesized from PMSF. One gram of the latter (6 mmol) was dissolved in 10 mL of propanol, to which 10 mL of 5 M NaOH (50 mmol) was added. The reaction mixture was refluxed overnight. The solution was then cooled, acidified with H_2SO_4, and the resultant PMSOH was extracted into 50 mL of ether. The ether phase was washed thrice with 50 mL of a 10% solution of H_2SO_4, dried over anhydrous $MgSO_4$, and subjected to rotary evaporation. The resultant solid PMSOH was 99% pure, as verified by HPLC.

Active center titration. The percentage of the catalytically competent subtilisin molecules (used to calculate $[E]_0$) of the soluble enzyme preparation was determined by titration with *N-trans*-cinnamoylimidazole (20). The concentration of the catalytically competent subtilisin in the CLCs which is accessible to the substrate in water and in acetonitrile was determined by titration of the active centers with the irreversible serine protease inhibitor PMSF (to avoid multiple turnovers due to long titration times) in two independent experiments, each done in triplicate. CLCs (20 mg/mL) were placed in 2 mL of either water (pH 5.0; 10 mM acetate buffer) or acetonitrile, both containing 1 mM PMSF, and the suspension was shaken at 30 °C and 300 rpm. The disappearance of

PMSF, as well as any PMSOH produced by spontaneous hydrolysis, was monitored by HPLC. The validity of this titration method was verified with a subtilisin solution of an independently determined $[E]_0$.

Kinetic measurements. The k_{cat}/K_M values in water were measured potentiometrically for the subtilisin-catalyzed hydrolysis of N-Ac-L-Phe-OEt (1-12 mM ester; 79 μg/L and 29 mg/L dissolved and suspended CLCs of subtilisin, respectively; pH 7.8; 30 °C; 100 mM KCl). Initial rate data were fitted to the Michaelis-Menten equation using the nonlinear curve fitting function of SigmaPlot (Jandel Scientific). The pH dependence of the soluble subtilisin (1.3-17 μg/L) and of the CLCs (15-190 mg/L) was also measured potentiometrically for the same reaction (1.0 mM ester substrate).

In anhydrous organic solvents, the ester and alcohol k_{cat}/K_M values were determined in the following manner. The CLCs of subtilisin (5 mg) were recovered from the pH 7.8 phosphate buffer, washed with anhydrous acetonitrile (2 × 1 mL), and then by the organic solvent in which the kinetics were studied (3 × 1 mL). For $(k_{cat}/K_M)_{N\text{-}Ac\text{-}L\text{-}Phe\text{-}OEt}$ the appropriate solvent containing a solution of the ester (0.6-6 mM in octane; 10-100 mM in acetonitrile) and 100 mM propanol was prepared immediately before kinetic measurements. In the determination of $(k_{cat}/K_M)_{propanol}$, the solutions of 100 mM N-Ac-L-Phe-OEt and propanol (2.5-100 mM) in *tert*-amyl alcohol and acetonitrile were prepared immediately before use, as was the case for the measurement of $(k_{cat}/K_M)_{2\text{-}methoxyethanol}$ where the solution was comprised of N-Ac-L-Phe-OEt (100 mM in acetonitrile; 50 mM in cumene) and alcohol (50 mM in acetonitrile; 2.5-100 mM in cumene). One milliliter of the appropriate solution was then added to 1 mg of the CLCs. The reaction mixture was

shaken at 30 °C and 300 rpm. Periodically, a 1-μL sample was withdrawn and assayed by GC. All k_{cat}/K_M values, except that for 2-methoxyethanol in acetonitrile, were determined by fitting the kinetic data as described above. The value of $(k_{cat}/K_M)_{2\text{-methoxyethanol}}$ in acetonitrile was calculated via the initial rate (v) ratios of 2-methoxyethanol and propanol and the k_{cat}/K_M value of propanol, where $(k_{cat}/K_M)_{2\text{-methoxyethanol}} = v_{2\text{-methoxyethanol}}/v_{propanol} \times (k_{cat}/K_M)_{propanol}$ since both nucleophiles were present in equal concentrations (10).

The measurement of $(k_{cat}/K_M)_{N\text{-Ac-L-Phe-OEt}}$ in octane and acetonitrile where the a_w was buffered to 0.006 was carried out as follows. The solvent containing the ester (0.6-6 mM in octane; 10-100 mM in acetonitrile) and 100 mM propanol was added to 100 mg of the mixture of $BaBr_2 \cdot 1H_2O/BaBr_2$; the suspension was equilibrated by shaking at 30 °C and 300 rpm for 90 min. One milligram of the CLCs was added, and the resulting mixture was shaken and the reaction monitored as described above. Again the kinetic data were fitted using nonlinear regression to obtain k_{cat}/K_M.

The $(k_{cat}/K_M)_{N\text{-Ac-L-Phe-OEt}}$ values in octane with a_w=0.006 and $a_{additive}$=1 (where the additive was formamide or N-methylacetamide) were determined as follows. The ester (0.6-6 mM) and 100 mM propanol were dissolved in octane containing 100 mg/mL $BaBr_2 \cdot 1H_2O/BaBr_2$ at 30 °C. The additive (containing 100 mM propanol) was added until the appearance of a second phase . The octane phase (1 mL) was then added to 1 mg of the CLCs. Kinetic measurements and k_{cat}/K_M calculations were performed as above.

Finally, the determination of k_{cat}/K_M of N-Ac-L-Phe-OEt in "hydrous" octane (a_w=1) was performed as follows. The ester (0.6-6 mM) and 100 mM propanol were

dissolved in 2 mL of octane, and the solution was allowed to equilibrate with water (containing propanol) through the vapor phase at 30 °C until the activity of water reached unity (three days). One milliliter of the resulting octane solution was withdrawn and added to 1 mg of CLCs. The initial rate of the reaction shaken at 30 °C and 300 rpm was monitored by GC, and the k_{cat}/K_M value was obtained by nonlinear curve fitting of the kinetic data.

Diffusion limitation assay. To test whether the hydrolysis of N-Ac-L-Phe-OEt catalyzed by the CLCs in water at pH 7.8 was limited by diffusion, active and PMSF-inactivated subtilisin were co-crystallized at various *f* values. The dependence of the activity of the resulting partially inactive crystals as a function of *f* was found to be linear (Fig. 2.1).

Inactivation of subtilisin. To a 58 mL subtilisin solution (10 mg/mL, pH 7.8, 20 mM phosphate buffer), 2.0 mL of a PMSF solution (43 mM; 6-fold molar excess) in propanol was added. The mixture was incubated for 4 h at room temperature. The remaining activity of subtilisin, assayed by measuring the initial rate of the enzymatic hydrolysis of 1.0 mM N-Ac-L-Phe-OEt in water (pH 7.8, 100 mM KCl, 30 °C), was found to be less than 0.1%. The excess PMSF was then removed by centrifugation/ultrafiltration (10 °C), followed by a 15-fold dilution with deionized water (thrice); the PMSF-inactivated subtilisin solution was then concentrated to 15 mg/mL.

Co-crystallization of active and PMSF-inactivated subtilisin. Varying volumes of solutions of the PMSF-inactivated and active subtilisins were mixed to achieve a range of values of *f* from 0.05 to 0.75. The crystals were then grown in the same manner as described above. To verify that the active and inactive subtilisin had co-crystallized, i.e.,

a single crystal was comprised of both active and inactive subtilisin, a single crystal was removed from a given batch of crystals, dissolved in aqueous solution (0.63 mg/L, pH 7.8, 100 mM KCl, 30 °C), and its activity was measured in the hydrolysis of 1.0 mM N-Ac-L-Phe-OEt. The activity obtained was compared to that of crystals of completely active subtilisin (f=1), and the f value of the given batch of crystals was calculated.

Activity vs. f studies. The crystals of inactive/active subtilisin, of a given f, were cross-linked as described above. The activity of the CLCs (0.05 mg/mL) was measured in the hydrolysis of 1.0 mM N-Ac-L-Phe-OEt in water (pH 7.8, 100 mM KCl, 30 °C) and plotted against f (Fig. 2.1).

Activity coefficient calculations. The activity coefficients of the nucleophiles propanol and 2-methoxyethanol in organic solvents were calculated using the UNIFAC algorithm (21), while that of N-Ac-L-Phe-OEt was obtained from its solubility in the appropriate solvent. The relationship between solubility and activity coefficient can be described as follows. If the standard state is the pure liquid solute, the activity of the solute is 1, n is the number of moles of solute in solution, n_S is the number of moles of solvent in solution, γ is the solution phase activity coefficient of the solute, and x is the solution phase mole fraction of the solute, then the solute will dissolve in the solvent until $\gamma x = 1$. If the solute is sparingly soluble, then $x = \frac{n}{n_S}$, where $n = SV$ (S is the molar solubility of the solute and V is the system volume) and $n_S = V/V_M$, where V_M is the molar volume of the solvent. Then $x = SV_M$ and $\gamma = \frac{1}{SV_M}$.

The value of γ, which is a function of solute concentration, was calculated at the concentration equaling *S*. At low mole fractions, where solute-solute interactions are negligible, γ becomes constant with respect to concentration. This is true for both Henry's and Raoult's law activity coefficients.

Determining the solubility (S). In anhydrous solvents, an excess of N-Ac-L-Phe-OEt was placed in the solvent containing 100 mM propanol. In hydrous solvents, where the water activity was buffered to 0.006, the excess of ester was placed in 1 mL of solvent containing 100 mM propanol and 100 mg of the $BaBr_2 \cdot 1H_2O/BaBr_2$ mixture. For *S* in water, the ester was placed in a 100 mM KCl solution, pH 7.8. For the octane systems with a_w=0.006 and $a_{additive}$=1, an excess of the ester was added to the octane (100 mM propanol, 100 mg/mL $BaBr_2 \cdot 1H_2O/BaBr_2$, 30 °C) and the additive was added until its solubility was reached. The resulting suspensions were shaken at 30 °C and 300 rpm. In octane with a_w=1, the ester was placed in a vial containing octane (and 100 mM propanol) in contact through the vapor phase with water containing propanol. The system was allowed to equilibrate at 30 °C until the water activity reached unity. Periodically, the ester concentration in each system was assayed by GC; the final concentration was denoted *S*.

E. References

1. Koskinen, A. M. P.; Klibanov, A. M., Eds. *Enzymatic Reactions in Organic Media*; Blackie: London, 1996. Santaniello, E.; Ferraboschi, P.; Grisenti, P. *Enzyme Microb. Technol.* **1993**, *15*, 367-382. Faber, K. *Biotransformations in*

Organic Chemistry; Springer-Verlag: Berlin, 1992. Klibanov, A. M. *Acc. Chem. Res.* **1990**, *23*, 114-120. Dordick, J. S. *Enzyme Microb. Technol.* **1989**, *11*, 194-211. Chen, C. S.; Sih, C. J. *Angew. Chem. Int. Ed. Eng.* **1989**, *28*, 695-707.

2. Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 3194-3201.

3. Desai, U. R.; Osterhout, J. J.; Klibanov, A. M. *J. Am. Chem. Soc.* **1994**, *116*, 9420-9422.

4. Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klibanov, A. M. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8653-8657.

5. Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675-681.

6. Schmitke, J. L.; Stern, L. J.; Klibanov, A. M. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4250-4255.

7. Lalonde, J. J.; Govardhan, C.; Khalaf, N.; Martinez, A. G.; Visuri, K.; Margolin, A. L. *J. Am. Chem. Soc.* **1995**, *117*, 6845-6852. Persichetti, R. A.; St. Clair, N. L.; Griffith, J. P.; Navia, M. A.; Margolin, A. L. *J. Am. Chem. Soc.* **1995**, *117*, 2732-2737. Sobolov, S. B.; Bartoszko-Malik, A.; Oeschger, T. R.; Montelbano, M. M. *Tetrahedron Lett.* **1994**, *35*, 7751-7754. St. Clair, N. L.; Navia, M. A. *J. Am. Chem. Soc.* **1992**, *114*, 7314-7316.

8. Boudart, M; Burwell, R. L. *Techniques in Chemistry,* 3rd ed.; Wiley: New York, 1973; Vol. 6, Chapter 12. Crank, J. *The Mathematics of Diffusion*; Oxford University Press: London, 1957; Chapter 5.

9. Fersht, A. *Enzyme Structure and Mechanism,* 2nd ed.; Freeman: New York, 1985; Chapter 12.

10. Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 1629-1631.

11. Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 8017-8021. Halling, P. J. *Trends Biotechnol.* **1989**, *7*, 50-52. Halling, P. J. *Enzyme Microb. Technol.* **1994**, *16*, 178-206.

12. Yang, Z.; Zacherl, D.; Russell, A. J. *J. Am. Chem. Soc.* **1993**, *115*, 12251-12257.

13. Poole, P. L.; Finney, J. L. *Int. J. Biol. Macromol.* **1983**, *5*, 308-310.

14. Halling, P. J. *Biotechnol. Tech.* **1992**, *6*, 271-276.

15. Xu, Z.; Affleck, R.; Wangikar, P.; Suzawa, V.; Dordick, J. S.; Clark, D. S. *Biotechnol. Bioeng.* **1994**, *43*, 515-520.

16. Kitaguchi, H.; Klibanov, A. M. *J. Am. Chem. Soc.* **1989**, *111*, 9272-9273. Ray, A. *Nature* **1971**, *231*, 313-315.

17. Reichardt, C.; *Solvent and Solvent Effects in Organic Chemistry*, 2nd ed.; VCH: New York, 1988, Appendix Table A-1.

18. Yennawar, N. H.; Yennawar, H. P.; Farber, K.; *Biochemistry* **1994**, *33*, 7326-7336.

19. Niedhart, D. J.; Petsko, G. A. *Protein Eng.* **1988**, *2*, 271-276.

20. Schonbaum, G. R.; Zerner, B.; Bender, M. L. *J. Biol. Chem.* **1961**, *236*, 2930-2935.

21. Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 10362-10363.

III. Comparison of X-ray Crystal Structures of an Acyl-Enzyme Intermediate of Subtilisin Carlsberg Formed in Anhydrous Acetonitrile and in Water

A. Introduction

Enzymes in organic solvents, instead of their natural aqueous milieu, remain synthetically useful catalysts (1). In addition, they display remarkable novel properties, e.g., solvent-dependent selectivity (2). Such solvent dependences of prochiral selectivity (3a) and enantioselectivity (3b) of crystalline enzymes in nonaqueous media have been nearly quantitatively predicted using structure-based molecular modeling and thermodynamic calculations.

Despite recent advances (4 and Chapter II), the question of why enzymatic activity is often much reduced in organic solvents compared to water is far from resolved. Such an understanding, like that concerning the selectivity, requires both structural and mechanistic information. Recently, structures of several enzymes in neat organic solvents, namely those of subtilisin Carlsberg in acetonitrile (5a,b) and dioxane (5f), γ-chymotrypsin in hexane (5c,d), and elastase in acetonitrile (5e), have been elucidated and found to be essentially the same as in water. In order to provide further insight into the enzyme mechanism in these media, though, structures of *reaction intermediates* should be determined. Until now, no such structures have been available for comparison

between water and organic solvents.*

All the aforementioned enzyme structures are of serine proteases, for which the universal reaction intermediate in water is an acyl-enzyme. In organic solvents, however, the formation of such an intermediate in the catalysis by subtilisin is supported by kinetic but not direct structural data (6). Moreover, it is simply presumed that the structure of the enzyme-substrate intermediate formed in different solvents is the same (3). The most direct way to experimentally test this assumption is to determine the corresponding structures.

At this point, three major pertinent questions remain unanswered: (i) Is there an acyl-enzyme intermediate formed in organic solvent?, If so, (ii) Is the structure of the intermediate the same as that formed in water?, and (iii) Are solvent molecules displaced from the active site by the substrate portion of the acyl-enzyme intermediate? In this study, we address these questions, as well as some related issues.

---

* The structures of cinnamoyl-γ-chymotrypsin and similar acyl-enzyme intermediates in water have been reported (Stoddard, B. L.; Bruhnke, J.; Porter, N. A., Ringe, D.; Petsko, G. A. *Biochemistry* **1990**, *29*, 4871-4879). In hexane, a tetrahedral intermediate for the tetrapeptide found in the γ-chymotrypsin active site specificity pocket was suggested (5d). However, said tetrapeptide was present during the crystallization from aqueous solution, and thus the intermediate may have formed there, rather than in hexane. In addition, the inhibitor N-acetyl-D-Trp, soaked into the crystal in hexane, partially occupies the specificity pocket in this structure and makes the interpretation of the electron density maps ambiguous (5d).

B. Results and Discussion

Previously, the X-ray crystal structures of lightly cross-linked subtilisin Carlsberg were solved in dioxane (5f), acetonitrile (5a), and water (5b), and were found to be virtually identical, with a root mean squared displacement (rmsd) of the backbone atoms of 0.3 Å among the different structures. However, to determine whether the enzyme mechanism is the same in organic solvent as in water, one must examine the structure of a reaction intermediate formed in both media, not just the structure of the free enzyme. To this end, in the present study, we set out to solve the structure of an acyl-enzyme reaction intermediate of subtilisin, specifically that of *trans*-cinnamoyl-subtilisin, formed in acetonitrile and in water. Its formation was achieved as follows. Crystals of subtilisin were grown from aqueous solution, lightly cross-linked, and placed in either anhydrous acetonitrile or water, to which *N-trans*-cinnamoylimidazole was subsequently added. The resulting cinnamoyl-subtilisin crystals were then mounted in capillaries, X-ray diffraction data were collected, and the structures were determined to 2.2-A resolution in acetonitrile (Figure 3.1) and in water (Figure 3.2). The structures included the substrate cinnamoyl group, as well as 74 bound water and 12 bound acetonitrile molecules in acetonitrile, and 88 bound water molecules in water.

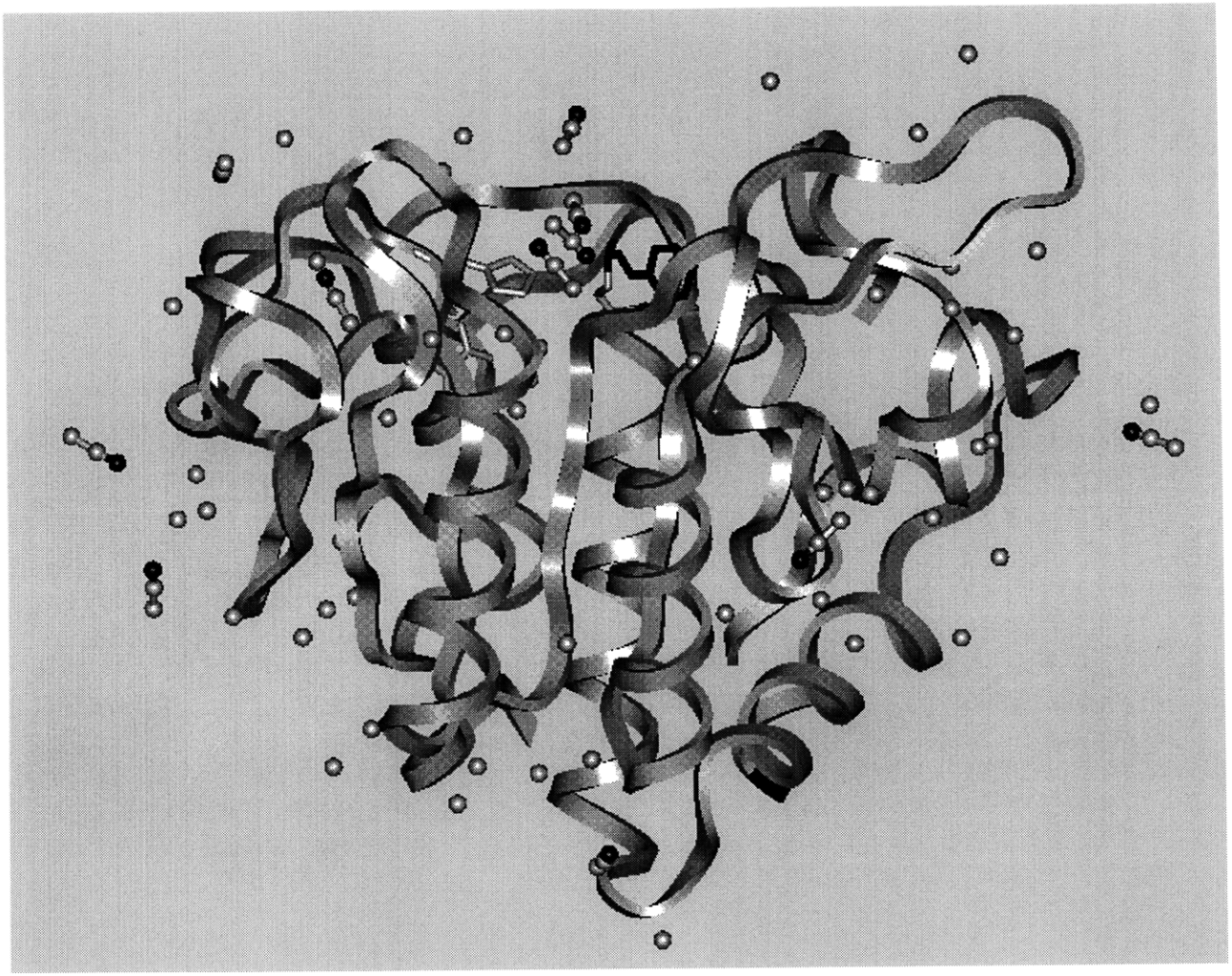


Figure 3.1. Ribbon diagram of the protein structure of *trans*-cinnamoyl-subtilisin in acetonitrile. The catalytic triad (Asp-32, His-64, and Ser-221) is portrayed as sticks. The cinnamoyl group is shown in black. Water molecules and acetonitrile molecules are depicted by balls-and-sticks with the nitrogen atoms of acetonitrile in black.

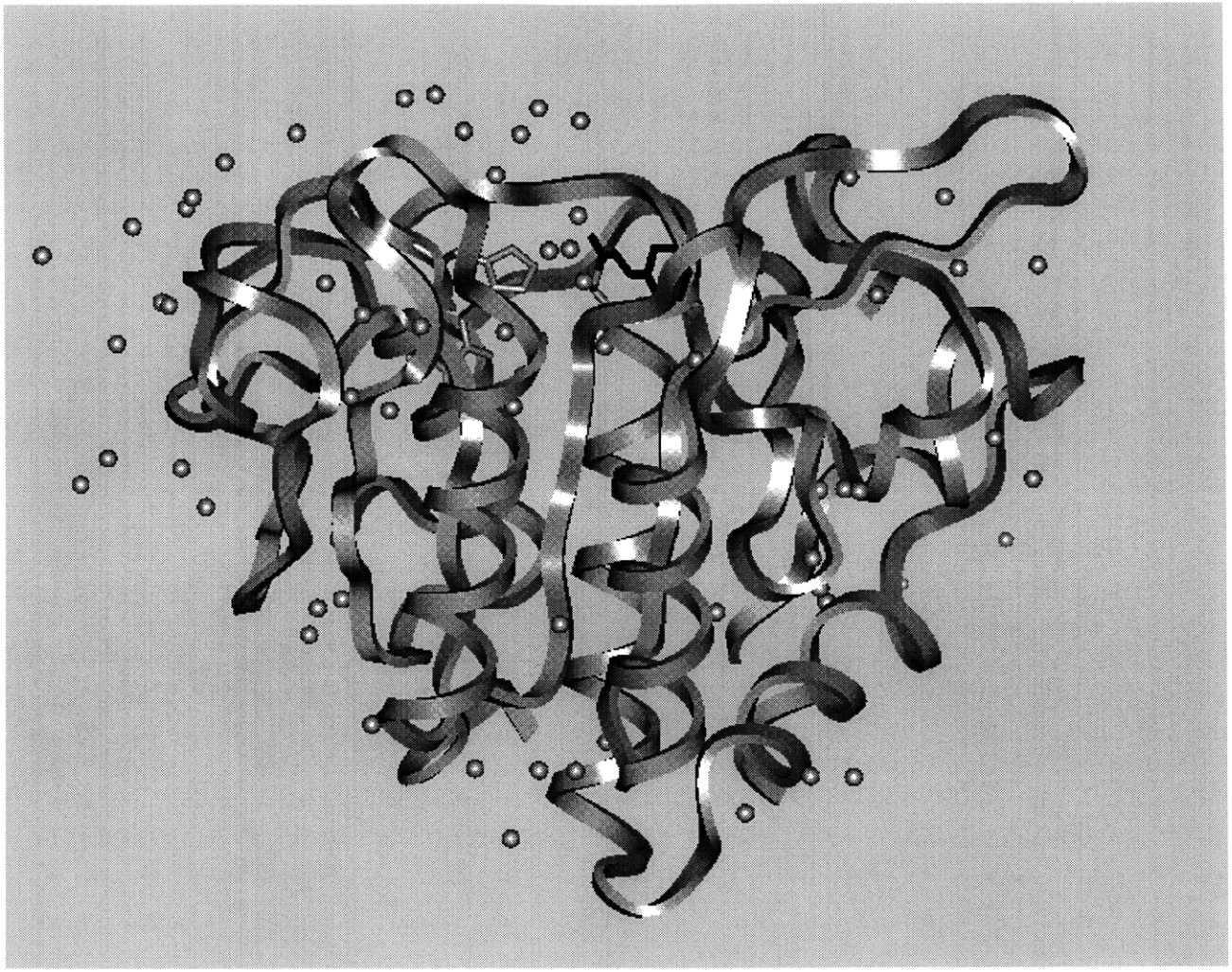


Figure 3.2. Ribbon diagram of the protein structure of *trans*-cinnamoyl-subtilisin in water. The catalytic triad and cinnamoyl group (black) are portrayed as sticks. Water molecules are depicted by gray balls.

The enzyme portion of subtilisin does not change upon formation of the acyl-enzyme intermediate in either acetonitrile (Fig. 3.3A) or water (Fig. 3.3B). In order to estimate the average coordinate error inherent in the acyl-enzyme structures, Luzzati (7) and sigmaA (8) analyses were performed (see Methods for details) and a value of 0.3 Å was estimated by either method for both structures. The rmsd of the backbone atoms between the free and acyl-enzyme structures are 0.30 Å in acetonitrile and 0.25 Å in water; these values are on the order of the average coordinate error inherent in each of the independent structures. In addition, only 3 out of 274 amino acid residues have an average displacement above 0.5 Å between the acyl- and free enzyme structures in acetonitrile (Gly-157 and Ser-159 of a flexible surface loop remote from the active center, and Tyr-171, also on the surface). Similarly, only 5 residues have an average displacement greater than 0.5 Å between the two structures in water (Ala-129 and Ser-130, and Asn-158, Ser-161 and Thr-162 of the flexible surface loop). Figure 3.3C reveals that the structure of the cinnamoyl-subtilisin formed in acetonitrile is the same as that formed in water, with an rmsd of 0.25 Å for the backbone atoms. Likewise, only 4 residues have an average displacement larger than 0.5 Å, namely Gly-157 and Ser-161, located in the flexible surface loop, Ala-274 at the C-terminus, and Ser-101, also on the enzyme surface remote from the active site.

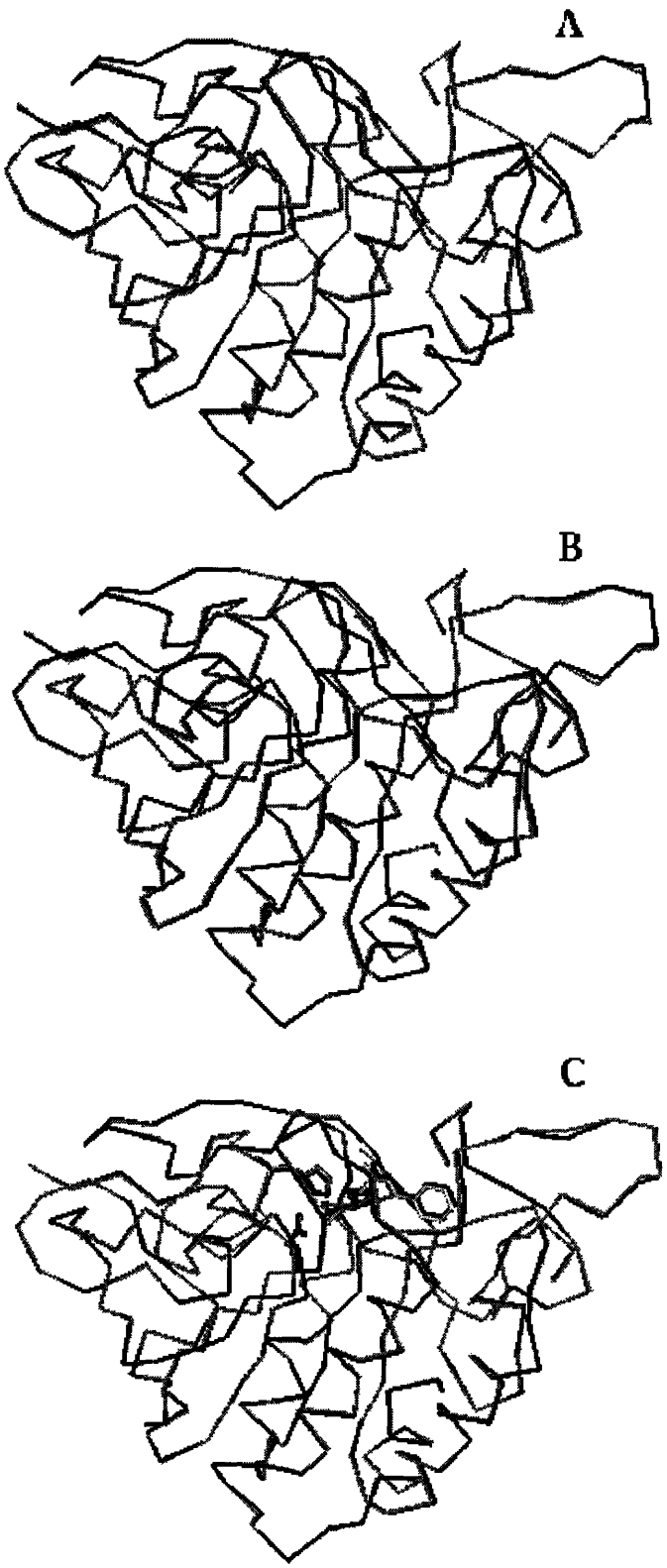


Figure 3.3. (A) Superimposition of the C_α traces of the protein structure of *trans*-cinnamoyl-subtilisin (black) and free subtilisin (gray) in anhydrous acetonitrile. (B) Superimposition of the C_α traces of *trans*-cinnamoyl-subtilisin (black) and free subtilisin (gray) in water. (C) Superimposition of the C_α traces of *trans*-cinnamoyl-subtilisin in anhydrous acetonitrile (black) and water (gray). The catalytic triad and cinnamoyl group are portrayed as sticks.

Because subtilisin exhibits markedly different catalytic activity and other properties in different solvents (2-4,6), the structure of the active site in these solvents is a critical issue. We find that not only is the overall structure of the crystalline acyl-enzyme the same whether formed in acetonitrile or in water, but it can be seen in Fig. 3.3C and by comparing Figures 3.4 A and B that the structure of the active site region, including that of the covalently bound cinnamoyl moiety, is also identical. The rmsd of all atoms within a 10-Å sphere of the active site catalytic triad nucleophile, the Ser-221 Oγ atom, is 0.29 Å in acetonitrile compared to water. Thus, it appears that the solvent does not change the activity of subtilisin by causing a conformational change in the active site or by affecting the mode of binding of the substrate.

As mentioned earlier, structure-based molecular modeling has been successfully used by us to predict the solvent dependence of enzymatic stereoselectivity (3) In so doing, the minimum energy conformation of the tetrahedral intermediate of the enzymatic, e.g., acylation, reaction is determined in vacuum. The assumption is then made that this minimum energy structure will be the same in solvent. We tested this assumption and found that the conformation of the substrate's cinnamoyl portion predicted by molecular modeling is indeed identical to that experimentally determined by X-ray crystallography (depicted in Fig. 3.4). Thus, the structure of the acyl-enzyme, which is the same in acetonitrile and water in our study, can be correctly predicted by molecular modeling.

An analysis of the enzyme-bound solvent molecules in the active site of the acyl- *vs.* free enzyme reveals that in acetonitrile (Figs. 3.4 A and C, respectively), the

cinnamoyl moiety displaces one acetonitrile molecule, that (labeled *b* in Fig. 3.4C) from the P1 specificity pocket. In the free enzyme structure, an additional acetonitrile molecule (*a* in Fig. 3.4C) is bound in the location of the 'catalytic water' molecule†, found in the structure in water (*a* in Fig. 3.4D). This acetonitrile molecule is still present in the acyl-enzyme structure in acetonitrile (*a* in Figs. 3.4 A and C). Note that particular attention was paid to the issue of the inclusion of acetonitrile and water molecules in the acyl-enzyme structure in acetonitrile (see Methods for details).

Examination of the active site situation in water shows that in the acyl-enzyme (Fig. 3.4B), the cinnamoyl group displaces one water from the specificity pocket and another from the oxyanion hole (*b* and *c*, respectively, in Fig. 3.4D). This gives rise to the question: can the dissimilar solvation of the active site in different solvents be a contributor to the distinct subtilisin activities in them (6)? Serine proteases like subtilisin act via a ping-pong mechanism,† where the 'pong' is the deacylation step. In enzymatic transesterification reactions in organic solvents (1), the nucleophile involved in the deacylation must first displace whatever molecule occupies the 'catalytic water' site. In acetonitrile, an acetonitrile molecule resides in this site (*a* in Figs. 3.4 A and C), while in dioxane (5f) the 'catalytic water' molecule is present. Consequently, differences in the enzymatic deacylation rates observed in, e.g., these two solvents may be partially due to the difference in displacement energies of an acetonitrile molecule in acetonitrile and a water molecule in dioxane.

---

† The 'catalytic water' molecule is a hallmark of the active site of subtilisin and other serine proteases which act via a ping-pong mechanism (Voet, D.; Voet, J. G. *Biochemistry*, 2nd Ed; Wiley: New York, pp. 398-400, 1995.)

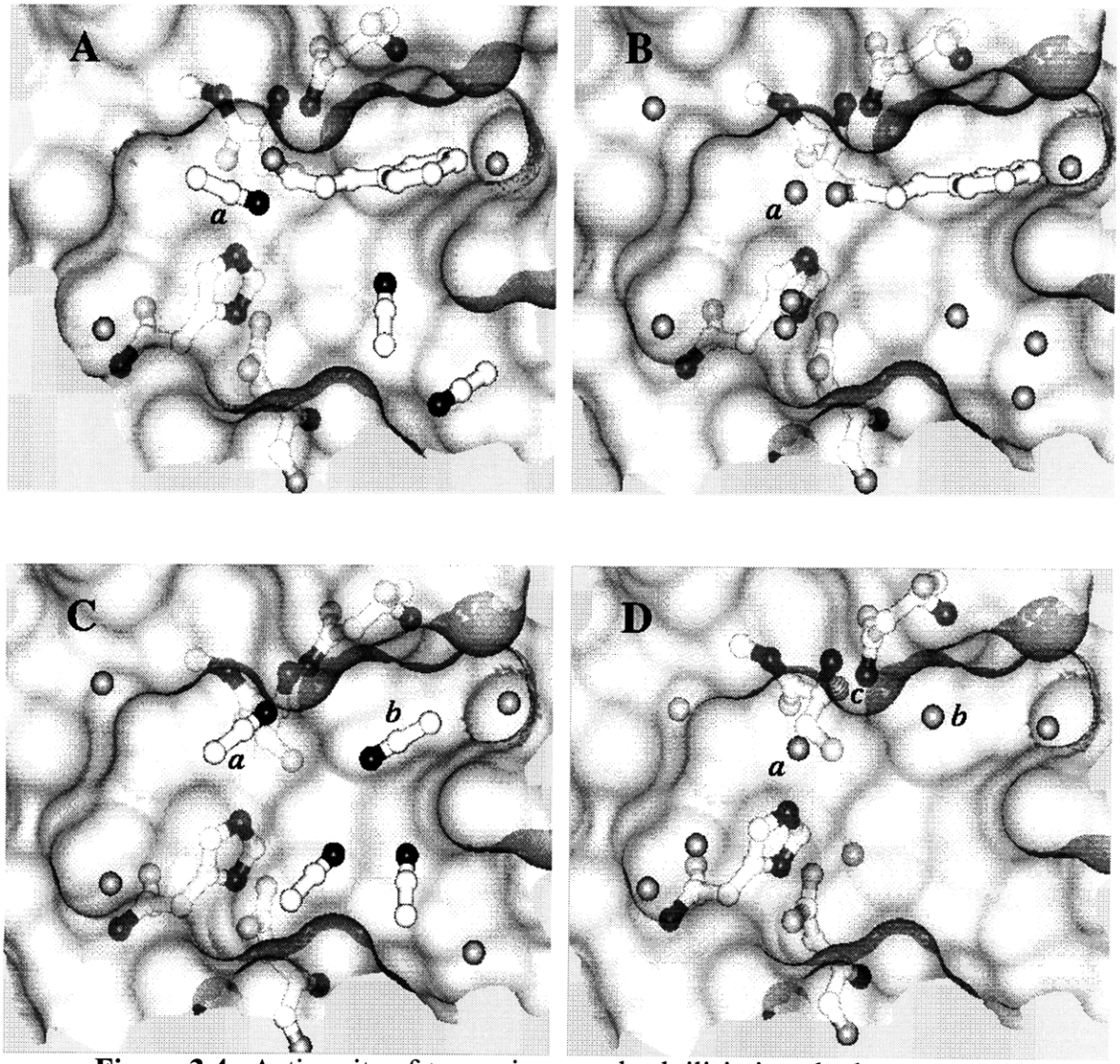


Figure 3.4. Active site of *trans*-cinnamoyl subtilisin in anhydrous acetonitrile (A) and in water (B), and that of the free subtilisin in anhydrous acetonitrile (C) and in water (D). The catalytic triad, Asn-155 of the oxyanion hole, the cinnamoyl group (A and B), and solvent molecules are shown as balls-and-sticks with carbon, oxygen, and nitrogen shown in white, light-gray, and black, respectively. The semitransparent solvent accessible surface of the protein (calculated using the Connolly algorithm () with a probe radius of 1.4 Å in the Insight II package from Biosym) in the active site region is shown in light gray.

The deacylation step is not the only one in which differential solvation of the active site could play a role. Since the cinnamoyl group in the acyl-enzyme displaces water molecules in water and an acetonitrile molecule in this solvent (Fig. 3.4), different

energies are required to displace these molecules to allow substrate binding. In effect, the solvent acts as a competitive inhibitor of the substrate. In an earlier work (Chapter II, 4b) where we attempted to explain the decrease in activity of cross-linked crystals of subtilisin upon transition from water to anhydrous acetonitrile, 1.4 out of some 7 orders of magnitude of activity loss remained unaccounted for. Perhaps the differential energies of displacing enzyme-bound water and acetonitrile molecules from the enzyme active site, not considered in that work, explain this shortfall.

The concept of specific interactions of solvent molecules with the enzyme can be used to rationalize some literature data. For instance, Partridge et al. (10) washed cross-linked crystals of subtilisin with either acetonitrile or propanol, then assayed the enzyme crystals in propanol, and found the resultant enzymatic activity lower in the former case. The acetonitrile introduced by washing could be acting as an inhibitor of the enzyme in propanol. Such specific interactions were also suggested by molecular dynamics studies performed on subtilisins immersed in dimethyl sulfoxide (11a) and dimethyl formamide (DMF) (11b) which predicted that organic solvent molecules would bind to the protein. Moreover, DMF molecules were predicted to bind in the enzyme active site (11b). In particular, molecular dynamics simulations in DMF placed at least one DMF molecule into the active site in both wild-type subtilisin E and a mutant (12) thereof. Interestingly, similar studies of these subtilisins in water showed that the water, but not DMF, molecules did not remain bound in the active site and diffused in and out during the simulations (11b).

In the examination of reactions involving enzyme crystals in organic solvents, it may be important to consider the role of the crystal itself, i.e., the overall physical

presentation of the enzyme. Different pretreatments of cross-linked crystals of subtilisin have been shown to result in distinct enzymatic activities in organic solvent. When Partridge et al. (10) incubated crystals for three days in anhydrous acetonitrile or in air, the result was a 3- and 200-fold drop in activity, respectively, when subsequently assayed in acetonitrile. These data can be explained by the following hypothesis: When either immersed in organic solvents for long times (days) or exposed to conditions that remove interstitial solvent, cross-linked enzyme crystals experience a partial collapse of their crystal lattice and concomitant closure of their solvent channels to the substrates. Consequently, the enzymatic activity drops due to that supramolecular phenomenon rather than due to a specific change on the molecular level.

We tested this hypothesis as follows. First, the accessible active sites of subtilisin in the crystal were determined by titration with PMSF, an irreversible inhibitor of serine proteases (13), as a function of the pretreatment described above. After a 1.7- and 3-day incubation in acetonitrile, the number of accessible active sites dropped to 13% and 2.5%, respectively (see Methods for details). Moreover, when the crystals were exposed to air for 3 days, less than 1% of the initial active sites remained accessible. In a second experiment, we investigated the crystals' ability to uptake methylene blue, an organic dye, as a function of the pretreatment. If the crystals' solvent channels are open (intact), the dye should penetrate the crystal and turn it blue. If they are not open, e.g., have collapsed, the dye will fail to do so. As seen in Figure 3.5A, crystals with no pretreatment, placed in anhydrous acetonitrile containing methylene blue, do in fact turn blue. Conversely, crystals exposed to air overnight remain colorless (Fig. 3.5B). Furthermore, the latter crystals appeared damaged, curled up, and deformed.

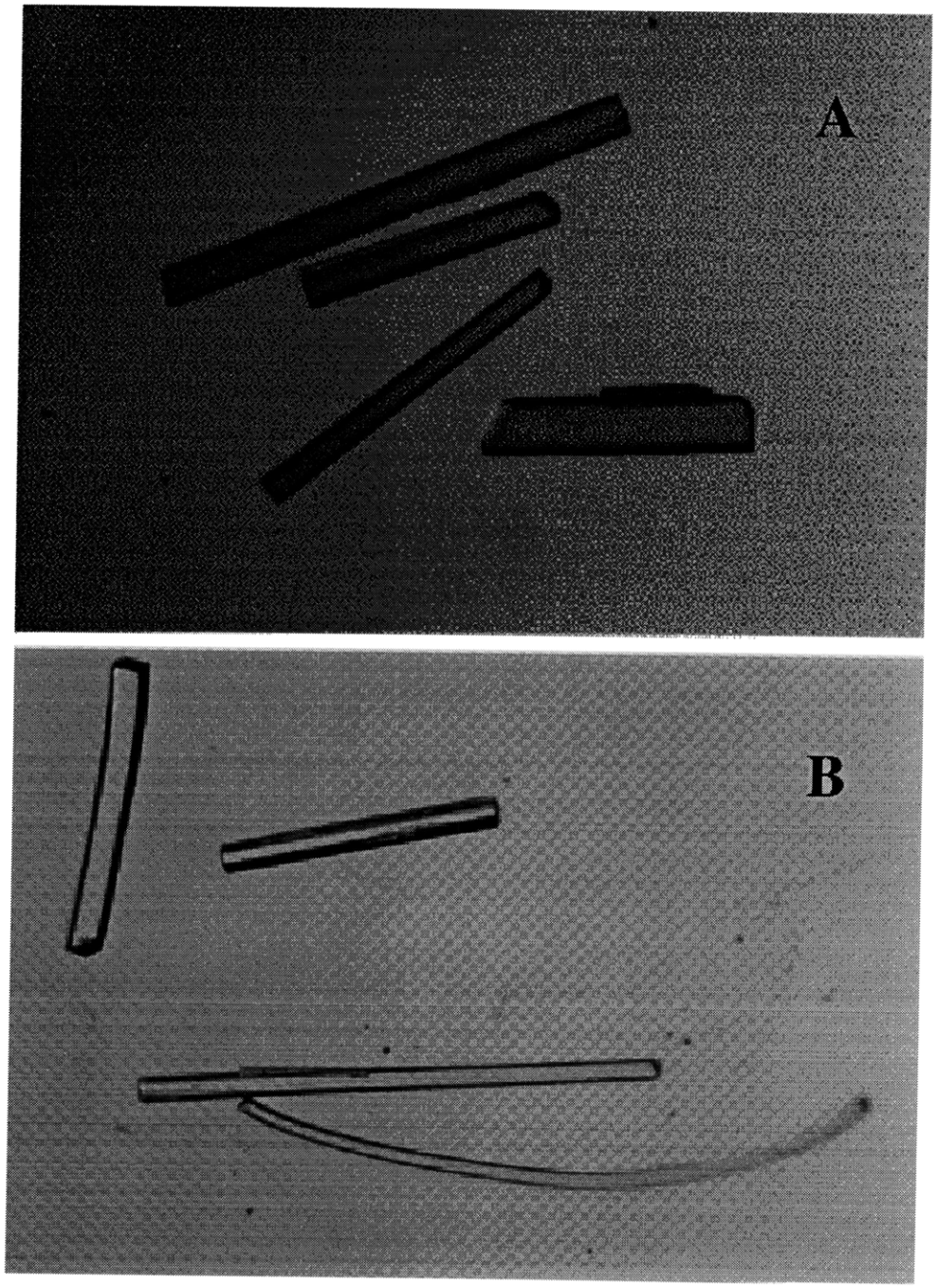


Figure 3.5. (A) Cross-linked crystals of subtilisin in a methylene blue solution (light gray background), where the crystals, having absorbed the dye, are dark gray. (B) Crystals that were exposed to air overnight prior to placement in the dye solution. They do not uptake the dye and remain clear (see Methods for details).

These data add another dimension to the question of catalytically competent active sites.

In contrast to the situation of an enzyme dissolved in water where only the number of

competent active sites is important, their accessibility is also an issue for a crystalline enzyme suspended in organic solvents. A similar predicament may also occur in lyophilized enzyme powders suspended in organic solvents, where prolonged incubation of subtilisin in acetonitrile also led to a reversible loss of activity (14). Thus, not only is structural information on the molecular level of the enzyme important, but so is an understanding of what is occurring at the supramolecular, crystalline, level of the enzyme catalyst.

C. Concluding Remarks

That the structures of the crystalline acyl-enzyme intermediate formed in acetonitrile and in water are the same verifies earlier kinetic evidence (6) that the enzymatic mechanisms in organic solvents and in water are similar. The observed distinct binding of solvent molecules in the enzyme active site between the acyl-enzyme intermediate and the free enzyme structures in different solvents may be responsible for at least some of the heretofore unexplained disparity in the activity of enzymes in organic solvents. This factor should be considered when endeavoring to manipulate enzymatic stereoselectivity by rational selection of the organic solvent (2,3). On a separate note, a decrease in the accessible active sites in crystalline (and amorphous) enzyme preparations due to the collapse of solvent channels causes a reduction in the observed enzymatic activity in organic solvents. Khalaf et al. (15) recently showed that when cross-linked crystals of subtilisin are first treated with certain detergents before drying to a constant water activity, a different catalytic activity in organic solvent ensues. Given our findings herein, the addition of a detergent to the solvent in the crystal lattice may combat the

phenomenon of channel collapse and preserve the uniformity of the crystal network. Thus, carefully designing the conditions to maintain the order of the solvent channels and protein crystal lattice should be explored as a strategy to achieve the full enzymatic potency.

D. Materials and Methods

Crystal preparation and acylation. Subtilisin Carlsberg (protease from *Bacillus licheniformis*, EC 3.4.21.14) was purchased from Sigma Chemical Co. Its crystals were grown from an aqueous 330 mM cacodylate buffer, pH 5.6, saturated with (~13%) Na_2SO_4 (16). Single crystals (~0.8 × 0.1 × 0.08 mm) were placed in 1 mL of a 10% glutaraldehyde cross-linking solution (aged for three days at room temperature before reaction (5f), see Methods in Chapter I) containing 30 mM cacodylate buffer, pH 7.5, and 10% Na_2SO_4. Again, we presume a relatively low extent of cross-linking: there are only 9 lysine residues in subtilisin of which only 5 which could be involved in an intermolecular cross-linking event (distances below 20 Å). No glutaraldehyde or modified Lys electron density was detected in the electron density maps.

The crystals were incubated in the cross-linking solution for 30 min and washed five times with 2 mL of distilled water. When preparing a crystal for acylation in acetonitrile, the water was then removed and replaced with anhydrous acetonitrile by washing five times with 2 mL each. We have shown that crystals similarly washed with organic solvent have a residual water content of 3% (w/w) (5f, see Methods in Chapter I). Conversely, for the acylation in water, the crystal was transferred to a 5 mM acetate

buffer, pH 5.0. The crystal was left in the appropriate solvent system for 5 min following each wash. Crystals were then acylated with *N-trans*-cinnamoylimidazole (NTC) in both water and acetonitrile. In acetonitrile, the crystal was transferred to an anhydrous solution containing 100 mM NTC, followed by mounting in a 0.7-mm glass capillary after the completion of the reaction in less than 5 min. In water, the crystal was transferred to a solution containing 110 μM NTC in cacodylate buffer, pH 5.0. After 1 min the reaction was complete, the pH was lowered to 3.0 with 1 M HCl, and the crystal was mounted in a 0.7-mm glass capillary. The extent of acylation was assayed by measuring the catalytic activity of the subtilisin crystal toward the hydrolysis of the substrate N-acetyl-L-phenylalanine ethyl ester. In water, the disappearance of absorption at 335 nm for the NTC was also monitored (17).

Data collection and reduction. X-ray diffraction data were obtained at ambient temperature (23±1°C) with an RAXIS II area detector. X-rays were generated with a Rigaku RU200 rotating copper anode source, and the Cu K_α radiation was selected and focused with mirrors. In acetonitrile, diffraction data of one crystal were measured over 119° as 2.25° oscillations to a nominal resolution of 2.15 Å. All data were collected within 20 hours of transfer to the solvent. In water, diffraction data of two crystals were collected to a nominal resolution of 2.20 Å with ranges of 2.75° and 2.25° over a total of 63.00° and 79.75°. The appropriate reflections were indexed and the intensities integrated and scaled using DENZO and SCALEPACK of the HKL Package (18). Data collection statistics are presented in Table 3.1. Structure factor magnitudes were calculated from the intensities and truncated using programs in the CCP4 Suite (19).

Table 3.1. Crystal properties, data collection, refinement, and model statistics for *trans*-cinnamoyl-subtilisin in anhydrous acetonitrile and in water

	ACETONITRILE		WATER	
Crystal properties				
space group	$P2_12_12_1$		$P2_12_12_1$	
cell dimensions: $a \times b \times c$, Å	76.4×55.3×52.8		77.0×55.0×53.6	
	resolution limits,[b] Å			
Data collection statistics	14-2.2	2.3-2.2	14-2.2	2.3-2.2
number of measurements	44825	4339	57568	4311
number of unique reflections	10495	1038	10663	1089
completeness, %	91.1	77.3	88.6	72.3
mean I / σ_I	24.5	10.1	14.6	4.9
R_{sym},[a] %	5.3	15.5	10.4	41.5
	resolution limits, Å			
Crystallographic refinement statistics	8.0-2.2	2.3-2.2	8.0-2.2	2.3-2.2
reflections used in refinement ($F>2.0\sigma_F$)	10376	1024	10193	1055
crystallographic *R*-factor,[c] %	19.5	24.2	18.4	25.6
free *R*-factor,[c] %	23.4	26.3	22.5	29.2
Model statistics[d]				
total number of non-hydrogen atoms	2041		2019	
number of protein atoms	1931		1931	
number of solvent atoms	110		88	
rmsd[e] bond length, Å	0.009		0.010	
rmsd[e] bond angle, °	1.3		1.5	
mean *B*-factor (standard deviation), Å^2				
protein	20 (9)		23 (8)	
solvent	43 (17)		35 (11)	
rmsd of structure in acetonitrile *vs.* water, Å				
protein backbone — 0.25	active site atoms — 0.29		cinnamoyl group — 0.24	

[a] $R_{sym} = \sum | I - \langle I \rangle | / \sum I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity of multiple observations of symmetry related reflections. In acetonitrile, the sum is over data from a single crystal. In water, this sum is over data from two largely overlapping data sets taken from two independent crystals. [b] In the acetonitrile data set, the resolution shell of 2.20-2.15 Å was also included in the refinement as described in the Methods. This shell contained 2557 measurements with 722 unique reflections, a completeness of 57.8%, a mean I/σ_I of 9.0, and an R_{sym} of 18.0%. [c] $R = \sum | | F_{obs} | - | F_{calc} | | / \sum | F_{obs} |$, where the free *R*-factor is calculated for a randomly chosen 10% of reflections, and the crystallographic *R*-factor is calculated for the remaining 90% of reflections ($F > 2.0\sigma_F$) used for the refinement. [d] Includes 10 atoms from the covalently bound cinnamoyl group and one Ca^{2+} atom. [e] Root mean square deviation from ideal geometric values (Enge and Huber parameter set).

Refinement and model building. A rigid body refinement, using the 10-4.0 Å data, utilizing the program X-PLOR (20) was performed for both the acetonitrile and water data where the starting model was that of the corresponding cross-linked free enzyme crystal structure previously published (Brookhaven Protein Data Bank coordinate files reference numbers 1scb in acetonitrile and 1sca in water) (5a,b). A series of $|2F_{obs}-F_{calc}|\exp(-i\alpha_{calc})$ maps were prepared using phases from the starting model for each solvent where one-eighth of the total residues was omitted in each map. These maps were examined only in the areas of omitted residues. The models and corresponding electron density maps were examined using the program O (21). Any region or residue for which there was no clear density was not included in the model in the first round of refinement. Subtilisin contains a flexible surface loop, Asn-158 to Thr-162, for which initially very weak density was observed. This entire loop was omitted at this stage. Residues of the loop were gradually included in the model in later refinement rounds. In addition, at this point all the residues within a 6-Å sphere of the active site Ser-221 were omitted from the refinement in order to prevent any active site residues from incorrectly refining into the density of the cinnamoyl group which was not included in the model until the final stages.

In all refinement protocols in acetonitrile, no solvent corrections in the reflection file were used with the 8.0-2.15 Å data. In water, a bulk solvent mask corrected reflection file with the 14-2.2 Å data was used. The data in the 2.3-2.2 Å resolution shell were used in the refinement in water, because the *R*-free (*R*-factor calculated for 10% of the total reflections left out of the refinement) (22), improved in the 14-2.3 Å resolution

shell when these data were included. The first round of automated refinement in X-PLOR, including only the model atoms described above, consisted of a simulated annealing slow-cool followed by a restrained individual *B*-factor refinement. The simulated annealing runs involved an initial positional refinement, 40 cycles, followed by 5 trials of heating to 3000 K and then cooling in increments of 25 K with 50 steps of molecular dynamics simulations for 0.5 fs at each temperature. A restrained individual *B*-factor refinement, 20 cycles, was then performed on the coordinate file (trial) with the lowest *R*-free value (22). At this stage, the *R*-factor and the *R*-free for the structure in acetonitrile were, respectively, 0.31 and 0.33 (8.0-2.15 Å). For the structure in water, the values were, respectively, 0.28 and 0.32 (14-2.2 Å) using the bulk solvent mask correction.

The refinements were followed by manual rebuilding using the program O (21). The models in both water and acetonitrile were first rebuilt residue by residue into the omitted regions using omit maps generated as described above, adjusting side chain torsion angles and occasionally changing residue rotamers. The rebuild in acetonitrile resulted in the inclusion of all the residues in the model except for Gly-160 and Ser-161. The catalytic triad, Asp-32, His-64, and Ser-221, as well as Asn-155 of the oxyanion hole, were still omitted from the model at this stage. In water the rebuild resulted in the inclusion of all residues in the model, except for the loop at Asn-158 to Thr-162; again, the catalytic triad residues and Asn-155 were left out of the model.

The next round of refinement using X-PLOR consisted of a conventional positional refinement, 20 cycles, followed by a restrained individual *B*-factor refinement,

20 cycles, At this stage, the *R*-factor and *R*-free were brought down to 0.25 and 0.29 (8.0-2.15 Å), respectively, in acetonitrile and 0.26 and 0.31 (14-2.2 Å), respectively, in water.

Solvent molecules were included in the models by means of manual building in O (21) and automated refinement in X-PLOR (20). Potential waters were located using the $|F_{obs}-F_{calc}|$ electron density map at the 3.0σ contour level (23). Initially, they were built only into density peaks that clearly exhibited the appropriate shape; any peaks that were possibly due to acetonitrile molecules in acetonitrile were excluded. A conventional positional refinement, 30 cycles, followed by a restrained individual *B*-factor refinement, 25 cycles, was performed on the models in both water and acetonitrile. Potential waters were retained if their electron density in subsequent $|2F_{obs}-F_{calc}|$ maps persisted after the positional refinement, and if they fulfilled the following criteria: oxygen atom within 3.4 Å of a subtilisin oxygen or nitrogen atom (or bound water in the following rounds below) with good hydrogen-bonding geometry, *B*-factors less than 65 $Å^2$, and real space fit correlation coefficients above 70%. The first round of manual building in O and refinement with X-PLOR resulted in the addition of the first 8 water molecules in acetonitrile and 9 water molecules in water. The remaining residues of the surface loop described above were also included in the models in both water and acetonitrile. At this point, the *R*-factor and *R*-free were brought down to 0.23 and 0.28 (8.0-2.15 Å), respectively, in acetonitrile, and 0.21 and 0.27 (14-2.2 Å), respectively, in water. Three additional rounds of manual building in O and refinement with the program X-PLOR were carried out resulting in the inclusion of an additional 53 waters in acetonitrile and 57 waters in water (and all protein atoms except for Ser-221), for a total of 62 water

molecules in the model in acetonitrile and 65 water molecules in the model in water. The water identification, manual building, and refinement were performed as above. Criteria for acceptable waters were the same as above. At this juncture, the *R*-factor and *R*-free were reduced to 0.20 and 0.25 (8.0-2.15 Å), respectively, in acetonitrile, and 0.19 and 0.24 (14-2.2 Å), respectively, in water. Two final such rounds of manual rebuilding in O and refinement in X-PLOR were performed on the model in water, resulting in the inclusion of a total of 88 water molecules and one Ca^{2+} (with an occupancy of 0.98) molecule. The resulting *R*-factor and *R*-free were and 0.18 and 0.23, respectively.

For cinnamoyl-subtilisin in acetonitrile, acetonitrile molecules were introduced into the model by means of two rounds of manual building in O and automatic refinement using X-PLOR. The acetonitrile model used was that of Fitzpatrick et al (5a). The acetonitrile parameter and topology files for X-PLOR were generated with the aid of XPLO2D (24). Potential acetonitrile molecules were located using the $|F_{obs}-F_{calc}|$ map. Electron density was considered to represent a potential acetonitrile molecule if the peak was above 3.0σ with the acetonitrile's nitrogen within hydrogen-bonding distance to a hydrogen-bond donor on the protein or water. Acetonitrile electron density was recognized by comparison with F_{calc} maps calculated for a theoretical acetonitrile molecule between the resolution limits of 14-2.15 Å. A complication that must be noted is that the electron density for a single acetonitrile molecule at this resolution appears similar to the density for two hydrogen-bonded water molecules. In some cases, the hydrogen-bonding pattern in the model was consistent only with two waters; however, when it was ambiguous, the following criteria were used to distinguish between the

possibilities of the density corresponding to one acetonitrile or two waters. Initially, a real space fit correlation coefficient was calculated for both an acetonitrile molecule and two waters molecules built into such an electron density peak. A conventional positional refinement, 30 cycles, followed by a restrained individual *B*-factor refinement, 25 cycles, was performed using X-PLOR on the model with either one acetonitrile molecule or two water molecules. The respective real space fit and *R*-free values were calculated; acetonitrile molecules were kept only if these parameters, before and after refinement, were superior to those for the two waters. Furthermore, putative acetonitrile molecules with average *B*-factors greater than 65 $Å^2$ or real space fit correlation coefficients less than 75% were discarded. Twelve acetonitrile molecules were included in the final model as well as 74 water molecules and one Ca^{2+} molecule. The *R*-factor and *R*-free at this point were 0.20 and 0.24, respectively.

Cinnamoyl group. The *trans*-cinnamoyl group model used was that of the crystal structure of *trans*-cinnamic acid (25). Again, the parameter and topology files for X-PLOR were generated with the aid of XPLO2D. The cinnamoyl group and Ser-221 were now finally included in the models in water and acetonitrile, built in using O. A conventional positional refinement, 20 cycles, followed by an occupancy refinement for the *trans*-cinnamoyl group, 10 cycles, followed by a restrained individual *B*-factor, 25 cycles, were performed in X-PLOR. The occupancy refined to 0.77 and 0.78 for the cinnamoyl residue in acetonitrile and water, respectively. The final values of *R*-factor and *R*-free were 0.20 and 0.23 (8.0-2.15 Å), respectively, in acetonitrile and 0.18 and 0.22 (14-2.2 Å), respectively, in water (Table 3.1).

The coordinates for these structures can be found in the Protein Data Bank (PDB; Brookhaven National Laboratory, entries ..... in acetonitrile and ..... in water).

Model validation. The final models in acetonitrile and water contained no residues that were in disallowed regions of the Ramachandran plot, as determined with the program Procheck (26). The models in acetonitrile and water contained one (Thr-211) and two (Ser-161 and Asn-212) residues, respectively, in generously allowed regions. The root mean squared deviation from ideal values for both structures can be found in Table 3.1. The real space fit correlation coefficients for the models in acetonitrile and water were 0.90 ± 0.05 protein and 0.84 ± 0.10 solvent; and 0.90 ± 0.05 protein and 0.84 ± 0.07 solvent, respectively using the $2F_{obs}$-F_{calc} maps. The final models contained no peaks above 3.5σ in the $|F_{obs}$-$F_{calc}|$ electron density maps. The overall *B*-factors for the models in acetonitrile and water were 21 and 24 $Å^2$, respectively, i.e., comparable to Wilson *B*-factors for the crystal data in acetonitrile and water of 27 and 29 $Å^2$, respectively. The average coordinate error predicted by Luzzati (7) or sigmaA (8) analyses (performed in X-PLOR) for both models were approximately 0.3 Å using the test set of reflections.

Active site titration. The accessible active sites of subtilisin in the crystal that could be titrated with phenylmethylsulfonyl fluoride, PMSF, an irreversible inhibitor of serine proteases, were measured as a function of the pretreatment of the crystal (4b, see Methods in Chapter II). Briefly, cross-linked crystals (10-50 mg/mL) were placed in 1 mL of acetonitrile containing 500 μM PMSF, and the suspension was shaken at 30 °C and 300 rpm. The disappearance of PMSF, as well as any phenylmethylsulfonic acid

produced by spontaneous hydrolysis, was monitored by HPLC. Each measurement was done in triplicate. The pretreatments to which the crystals were subjected were taken from Partridge et al. (10) Pretreated crystals were washed with anhydrous acetonitrile (3 × 1 mL), followed by incubation in either acetonitrile (10 mL) for 1.7 and 3 days or in air for 3 days in a sealed jar. Crystals with no pretreatment were washed with anhydrous acetonitrile (3 × 1 mL), recovered by centrifugation, and used immediately. The fraction of the cross-linked crystal that was catalytically competent and accessible subtilisin in each case was: 44 ± 2%, no pretreatment; 5.5 ± 1.3% and 1.1 ± 0.7% for the 1.7- and 3-day incubation in acetonitrile, respectively; and less than 0.5% after the 3-day incubation in air.

Crystal dye soak. An aqueous solution of methylene blue was purchased from Hampton Research. The crystals with no pretreatment were washed with anhydrous acetonitrile (3 × 1 mL) and immediately placed in a solution of methylene blue in anhydrous acetonitrile (1 mL). The dye was then allowed to soak into the crystal. After three hours, the crystals became quite blue and the photograph represented in Fig. 3.5A was taken. Other crystals were prepared by first washing with anhydrous acetonitrile (3 × 1 mL), followed by placement of the crystals in a sealed jar overnight. The crystals were then placed in a solution of methylene blue identical to that described above. After three hours, no change in the crystal color was detected and the photograph represented in Fig. 3.5B was taken. In addition, no color was observed in these crystals even for incubations longer than one day.

Molecular modeling. The molecular modeling methodology utilized for the *trans*-cinnamoyl-subtilisin was that of Ke et al. (3a). The only exception was that the starting model was that of the covalently bound acyl-intermediate (rather than tetrahedral one).

E. References

1. Chen, C.-S.; Sih, C. J. *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 695-707. Klibanov, A. M. *Acc. Chem. Res.* **1990**, *23*, 114-120. Dordick, J. S. *Biotechnol. Prog.* **1992**, *8*, 259-267. Santaniello, E.; Ferraboschi, P.; Grisenti, P. *Enzyme Microb. Technol.* **1993**, *15*, 367-382. Gutman, A. L.; Shapira, M. *Adv. Biochem. Eng./Biotechnol.* **1995**, *52*, 87-128. Koskinen, A. M. P.; Klibanov, A. M., Eds. *Enzymatic Reactions in Organic Media*; Blackie: London, 1996.
2. Wescott, C. R.; Klibanov, A. M. *Biochim. Biophys. Acta* **1994**, *1206*, 1-9. Carrea, G.; Ottolina, G.; Riva, S. *TIBTECH* **1995**, *13*, 63-70.
3. (a) Ke, T.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3366-3374. (b) Wescott, C. R.; Noritomi, H.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 10365-10370.
4. (a) Halling, P. J. *Enzyme Microb. Technol.* **1994**, *16*, 178-206. (b) Schmitke, J. L.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3360-3365. (c) Margolin, A. L. *TIBTECH* **1996**, *14*, 223-230. (d) Klibanov, A. M. *TIBTECH* **1997**, *15*, 97-101.

5. (a) Fitzpatrick; P. A., Steinmetz; A. C. U.; Ringe, D.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8653-8657. (b) Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675-681. (c) Yennawar, N. H; Yennawar, H. P.; Farber, G. K. *Biochemistry* **1994**, *33*, 7326-7336. (d) Yennawar, H. P.; Yennawar, N. H.; Farber, G. K. *J. Am. Chem. Soc.* **1995**, *117*, 577-585. (e) Allen, K. N.; Bellamacina, C. R.; Ding, X.; Jeffery, C. J.; Mattos, C.; Petsko, G. A.; Ringe, D. *J. Phys. Chem.* **1996**, *100*, 2605-2611. (f) Schmitke, J. L.; Stern, L. J.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4250-4255.

6. Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 3194-3201. Kanerva, L. T.; Klibanov, A. M. *J. Am. Chem. Soc.* **1989**, *111*, 6864-6865. Adams, K. A. H.; Chung, S.; Klibanov, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 9418-9419. Chatterjee, S.; Russell, A. J. *Biotechnol. Bioeng.* **1992**, *40*, 1069-1077.

7. Luzzati, V. *Acta Cryst.* **1952**, *5*, 802-810. Kleywegt, G. L.; Jones, T. A. *Structure* **1994**, *2*, 1241-1258. Kleywegt, G. L.; Brünger, A. T. *Structure* **1996**, *4*, 897-904.

8. Read, R. J. *Acta Cryst.* **1986**, *A42*, 140-149.

9. Connolly, M. L. *Science* **1983**, *221*, 709-712.

10. Partridge, J.; Hutcheon, G. A.; Moore, B. D.; Halling, P. J. *J. Am. Chem. Soc.* **1996**, *118*, 12874-12877.

11. (a) Zheng, Y.; Ornstein, R. L.; *J. Am. Chem. Soc.* **1996**, *118*, 4175-4180. (b) Toba, S.; Mertz, K. M., Jr. *J. Am. Chem. Soc.* **1997**, *119*, 9939-9948.

12. Chen, K.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5618-5622.

13. Moss, D. E.; Fahrney, D. E. *Biochem. Pharmacol.* **1978**, *27*, 2693.

14. Schulze, B.; Klibanov, A. M. *Biotechnol. Bioeng.* **1991**, *38*, 1001-1006.

15. Khalaf, N.; Govardhan, C. P.; Lalonde, J. J.; Persichetti, R. A.; Wang, Y.; Margolin, A. L. *J. Am. Chem. Soc.* **1996**, *118*, 5494-5495.

16. Neidhart, D. J.; Petsko, G. A. *Protein Eng.* **1988**, *2*, 271-276.

17. Schonbaum, G. R.; Zerner, B.; Bender, M. L. *J. Biol. Chem.* **1961**, *236*, 2930-2935.

18. Otwinowski, Z.; Minor, W. *Meth. Enzymol.* **1997**, *276*, 307-326.

19. Collaborative Computational Project, Number 4 *Acta Cryst.* **1994**, *D50*, 760-763.

20. Brünger, A. T.; Kuriyan, J.; Karplus, M. *Science* **1987**, *35*, 458-460.

21. Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Cryst.* **1991**, *A47*, 110-119.

22. Brünger, A. T. *Nature* **1992**, *355*, 472-475.

23. McRee, D. E. *Practical Protein Crystallography*; Academic Press: San Diego, pp. 226-227, 1993.

24. G. J. Kleywegt, *ESF/CCP4 Newsletter* **1995**, *32*.

25. Wierda, D. A.; Feng, T. L.; Barron, A. R. *Acta Cryst.* **1989**, *C45*, 339-341.

26. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283-291.

IV. Organic Solvent Binding to Crystalline Subtilisin in Mostly Aqueous Mixtures *vs.* in Neat Solvents

A. Introduction

The determination of X-ray crystal structures in anhydrous or nearly anhydrous solvents has been advocated as a direct, experimental, functional-group mapping technique (1-4) analogous to the computational multiple-copy simultaneous-search method (MCSS) (5,6). MCSS is used to determine the local energy minima of a functional group in the force field of a protein. Multiple copies of the functional group are distributed in the region of interest of the protein and energy minimization of these copies using, e.g., molecular mechanics yields distinct local minima. The result is a functional group binding map of the protein surface (5,6). In the crystallographic technique, structures of the protein in several neat organic solvents are determined, and the locations of the protein-bound organic solvent molecules, each representing a given functional group, are likewise ascertained (1-4). The ultimate goal is to then construct larger, tighter-binding molecules, 'drug candidates', which incorporate these bound functional groups.

A major obstacle faced by the crystallographic approach is that it is very difficult to determine crystal structures of proteins in neat organic solvents due to crystal contact distortions (7,8). It should be much easier to solve the structure of a lightly cross-linked enzyme crystal placed in a predominantly aqueous solution containing organic solvent. In addition, since the ultimate goal of such rational drug design methods is to identify molecules that bind to proteins *in aqueous solution*, determining the binding sites of

organic solvent molecules in the neat solvent may yield some positions involving such weak associations that they are not relevant in aqueous solution.

The preceding considerations lead to the following questions: (i) Do solvent molecules bind to crystalline enzymes in aqueous-organic, principally aqueous mixtures (which are more relevant media than neat solvents if one intends to design drug molecules that bind to proteins under physiological conditions)?, If so, (ii) Where do these organic solvent molecules bind to the enzyme?, and (iii) Are these solvent molecules bound in the same locations as in neat organic solvent?

To address these questions, we determined the structures of crystals of subtilisin, lightly cross-linked and placed in 40% acetonitrile and 20% dioxane, and compared them to those previously determined in the corresponding neat solvents (1and 8, respectively).

B. Results and Discussion

Previously, the structures of lightly cross-linked crystals of subtilisin were determined in water (7), anhydrous acetonitrile (1), and neat dioxane (Chapter I, (8)). All these structures were found to be essentially identical: the root mean squared distance (rmsd) between any two was approximately 0.3 Å, which is on the order of the error inherent in each of the individual structures. In neat acetonitrile,12 organic solvent molecules were found in the enzyme structure, while 7 dioxane molecules were present in that anhydrous solvent. In the current study, as outlined in the Introduction, we were interested in determining whether organic solvent molecules bind to crystalline subtilisin in aqueous-organic, primarily aqueous mixtures, and, if so, how this binding compared to

that in the neat solvents. We therefore solved the X-ray crystal structures of subtilisin in aqueous solutions containing either 40% acetonitrile or 20% dioxane.

Subtilisin crystals were grown from aqueous solution, lightly cross-linked to render them insoluble and fortify them, and transferred to an anqueous solution of 40% acetonitrile or 20% dioxane; diffraction data were collected, and the structures determined to 2.3- and 2.2- Å, respectively. The structure of subtilisin in these aqueous-organic systems were found to be virtually identical to those in the neat solvents; the rmsd among them is roughly 0.3 Å, i.e., again on the order of the error inherent in each of the individual structures (see Methods for details). In regard to the question of the binding of organic solvent molecules to the enzyme, in the structure of subtilisin in 40% acetonitrile 74 water and 5 acetonitrile molecules were found, while that in dioxane contained 91 water molecules and only a single dioxane molecule.

Note that in neat acetonitrile 12 molecules of the solvent are bound to subtilisin (1). Of the 5 acetonitrile molecules which were present in the structure of subtilisin in 40% acetonitrile, 4 are in the same locations as 4 acetonitrile molecules found in the neat solvent. The fifth acetonitrile molecule in the 40% acetonitrile structure, which was found at a crystal contact, has no counterpart in the neat organic solvent. Perhaps the density was not sufficiently strong in the neat solvent in order to discriminate between water and acetonitrile molecules, and therefore nothing was included in that location. Notably, as mentioned in the literature (3,9), organic solvent molecules located at crystal contacts are most likely an artifact of the crystal itself, thus binding at these locations would not occur in solution.

In regard to the natural question concerning the other 8 acetonitrile molecules observed in the neat solvent we found the following. In 40% acetonitrile, water molecules were found instead in 5 of the acetonitrile molecule locations observed in the neat solvent. No solvent molecules were present in the subtilisin structure in 40% acetonitrile where the remaining three acetonitrile molecules were situated in the anhydrous solvent. Hence, 8 of the 12 acetonitrile molecules present in the structure in the anhydrous solvent apparently bind so weakly that they are not observed in 40% acetonitrile. This suggests the possibility that large designer molecules constructed from such weak binding functional groups will also not bind to the enzyme at relevant concentrations under physiological conditions.

We performed a similar analysis of the organic solvent binding pattern of subtilisin in 20% dioxane *vs.* in the neat solvent. Only a single dioxane molecule was found bound to the enzyme in 20% dioxane, compared to seven in the neat solvent. This lone dioxane molecule, though, is in the same location as one of the dioxane molecules in the structure of subtilisin in the anhydrous solvent. In 20% dioxane, water molecules were found in 3 of the dioxane molecule locations in the neat solvent. Once more, no solvent molecules were present in the 20% dioxane structure where the 3 final dioxane molecules were situated in the neat solvent. In one case, however, namely where the second active site-bound dioxane molecule was found in the neat solvent, a dioxane molecule was included in early models in 20% dioxane but eventually discarded due to a high B-factor. It should be noted that of the 7 dioxane molecules which were present in subtilisin's structure in the anhydrous solvent, all but the 2 in the active site were situated at crystal contacts, and thus, as pointed out above, would doubtless not occur in solution.

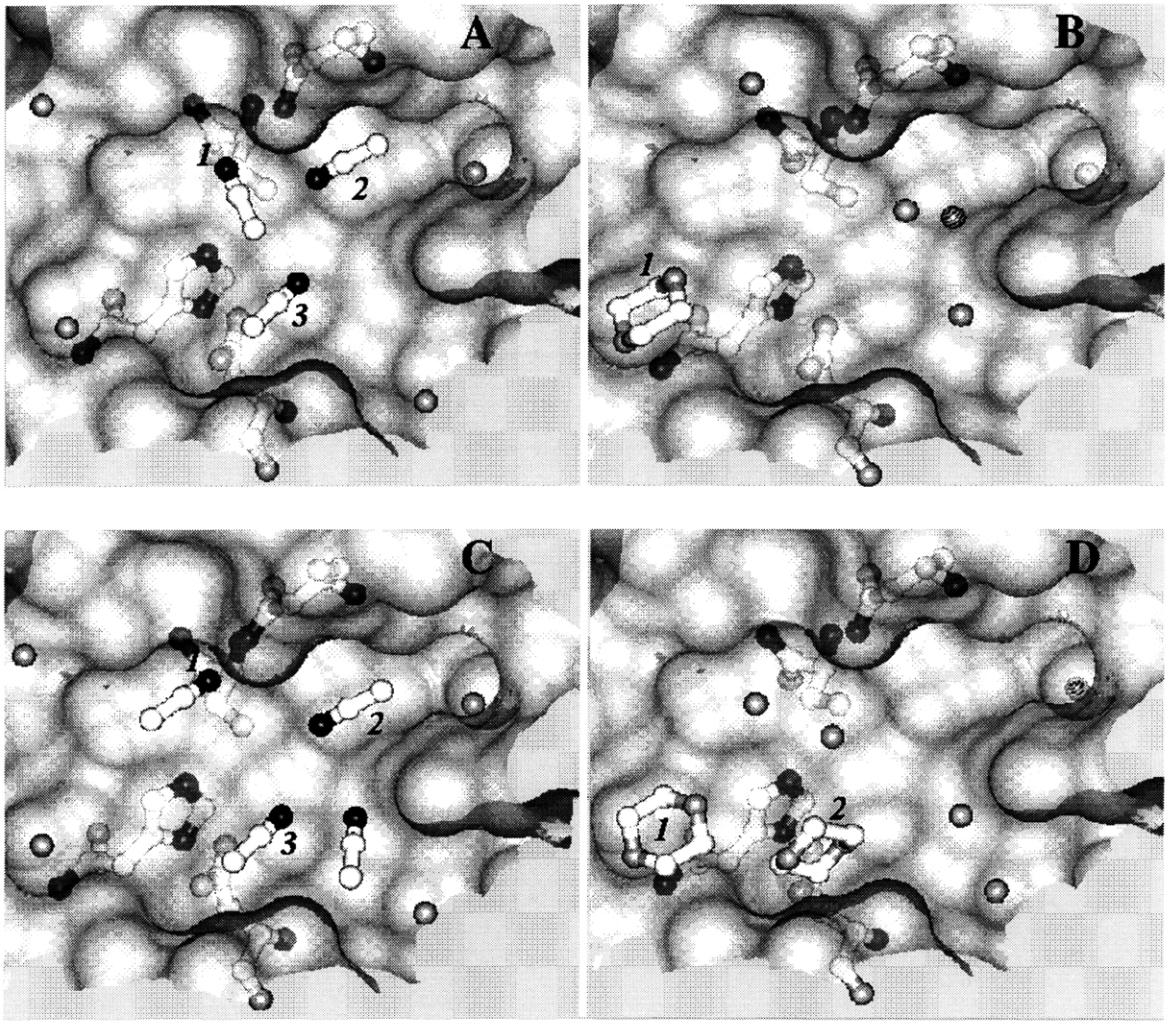


Figure 4.1. Active site of subtilisin in 40% acetonitrile (A), in 20% dioxane (B), in anhydrous acetonitrile (C), and in neat dioxane (D). The catalytic triad, Asn-155 of the oxyanion hole, and solvent molecules are shown as balls-and-sticks with carbon, oxygen, and nitrogen shown in white, gray, and black, respectively. The semitransparent solvent accessible surface of the protein (calculated using the Connolly algorithm () with a probe radius of 1.4 Å in the Insight II package from Biosym) in the active site region is shown in light gray.

Enzyme-bound organic solvent molecules in subtilisin crystals placed in aqueous systems containing organic solvent appear to congregate in the active site. In 40% acetonitrile 3 of the 5 acetonitrile molecules found in the structure are in the active site (labeled *1-3* in Fig. 4.1A), and are, importantly, in the same locations as 3 in the neat

solvent (*1-3* in Fig. 4.1C). For subtilisin in 20% dioxane, the single dioxane molecule present in the structure is also located in the active site (*1* in Fig. 4.1B), again, in the same position as one in the neat solvent (*1* in Fig. 4.1D). Perhaps this observed preferential occupation of the active site by the organic solvent occurs because the enzyme active site is naturally designed to bind substrates.

Evidence for preferential binding of organic solvent molecules to enzyme active sites also exists in the literature. For example, when the enzyme ribonuclease A was crystallized from an aqueous solution containing 8 M formate, two bound formate molecules were found in the structure, both located in the active site of the enzyme (9). Also, solution NMR studies of lysozyme have also shown that in aqueous solution containing molar concentrations of a variety of organic solvents, molecules of the latter bind to the enzyme active site substrate binding cleft, and nowhere else on the protein (11). On a separate note, these data support the recently advanced concept of the occurrence of specific interactions of the solvent with the enzyme active site resulting in organic solvent inhibition of enzymatic catalysis in non-aqueous media (Chapter III, (12).

To summarize, this study presents a preferred method for the elucidation of organic molecule binding sites on enzymes and other (receptor) proteins. Not only is it easier to determine the structures of enzymes in aqueous-organic mixtures than in neat organic solvents, but potential false positives should be minimized in this method as well.

C. Materials and Methods

Crystal preparation. Subtilisin Carlsberg (protease from *Bacillus licheniformis*, EC 3.4.21.14) was purchased from Sigma Chemical Co. Its crystals were grown from an

aqueous 330 mM cacodylate buffer, pH 5.6, saturated with (~13%) Na_2SO_4 ([13]). Single crystals (~1.0 × 0.1 × 0.08 mm) were placed in 1 mL of a 10% glutaraldehyde cross-linking solution (aged for three days at room temperature before reaction, see Methods in Chapter I (8)) containing 30 mM cacodylate buffer, pH 7.5, and 10% Na_2SO_4. Once again, we presume a relatively low extent of cross-linking: there are only 9 lysine residues in subtilisin of which only 5 which could reasonably be involved in an intermolecular cross-linking event (distances below 20 Å). No glutaraldehyde or modified Lys electron density was detected in the electron density maps, but the cross-linking rendered the crystals insoluble in aqueous systems and more robust against distortions caused by organic solvents (Chapter I, (7,8)).

The crystals were incubated in the cross-linking solution for 30 min and washed five times with 2 mL of distilled water. The water was then removed and replaced with either 40% (v/v) acetonitrile or 20% (v/v) dioxane by washing five times with 2 mL each. The crystals were finally mounted in a 0.7-mm glass capillary.

Data collection and reduction. X-ray diffraction data were obtained at ambient temperature (23±2°C) with an RAXIS II area detector. X-rays were generated with a Rigaku RU200 rotating copper anode source, and the Cu K_α radiation was selected and focused with mirrors. In 40% acetonitrile, diffraction data of two crystals were measured over 35.75° and 57.75° as 2.75° oscillations to a nominal resolution of 2.3 Å. In 20% dioxane, diffraction data of two crystals were collected to 2.2 Å with a range of 3.00° over a total of 48.00° and 57.00°. The appropriate reflections were indexed and the intensities integrated and scaled using DENZO and SCALEPACK of the HKL Package

(14). Data collection statistics are presented in Table 4.1. Structure factor magnitudes were calculated from truncated intensities using programs in the CCP4 Suite (15).

Table 4.1. Crystal properties, data collection, refinement, and model statistics for subtilisin in 40% acetonitrile and in 20% dioxane

	40% ACETONITRILE		20% DIOXANE	
Crystal properties				
space group	$P2_12_12_1$		$P2_12_12_1$	
cell dimensions: $a \times b \times c$, Å	79.0×54.8×53.9		77.6×55.2×53.9	
	resolution limits, Å			
Data collection statistics	20-2.3	2.4-2.3	20-2.2	2.3-2.2
number of measurements	34019	3523	44882	3478
number of unique reflections	10215	1098	11408	1195
completeness, %	93.7	85.9	93.8	85.0
mean I / σ_I	10	3.9	9.9	3.9
R_{sym},[a] %	12.6	40.6	12.9	35.7
	resolution limits, Å			
Crystallographic refinement statistics	20-2.3	2.4-2.3	20-2.2	2.3-2.2
reflections used in refinement ($F>2.0\sigma_F$)	9817	1053	10193	1055
crystallographic *R*-factor,[b] %	18.6	28.4	18.7	26.2
free *R*-factor,[b] %	24.7	29.6	24.5	33.8
Model statistics[c]				
total number of non-hydrogen atoms	2010		2017	
number of protein atoms	1921		1921	
number of solvent atoms	89		96	
rmsd[d] bond length, Å	0.007		0.007	
rmsd[d] bond angle, °	1.3		1.2	
mean *B*-factor (standard deviation), Å^2				
protein	27 (8)		22 (7)	
solvent	43 (14)		33 (13)	

[a] $R_{sym} = \Sigma \mid I - \langle I \rangle \mid / \Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity of multiple observations of symmetry related reflections. In both 40% acetonitrile and 20% dioxane, the sum is over data from two overlapping data sets taken from two independent crystals. [b] $R = \Sigma \mid\mid F_{obs} \mid - \mid F_{calc} \mid\mid / \Sigma \mid F_{obs} \mid$, where the free *R*-factor is calculated for a randomly chosen 10% of reflections, and the crystallographic *R*-factor is calculated for the remaining 90% of reflections ($F > 2.0\sigma_F$) used for the refinement. [c] Includes one Ca^{2+} atom. [d] Root mean square deviation from ideal geometric values (Enge and Huber parameter set).

Refinement and model building. A rigid body refinement, using the 10-4.0 Å data, utilizing the program X-PLOR (16) was performed on both the 40% acetonitrile and 20% dioxane data. The starting models used were the previously published cross-linked enzyme crystal structures in the corresponding neat solvents (Brookhaven Protein Data Bank (PDB) coordinate files reference numbers 1scb in acetonitrile and 1af4 in dioxane).

A series of $|2F_{obs}-F_{calc}|\exp(-i\alpha_{calc})$ maps were prepared using phases from the starting model for each solvent system where one-eighth of the total residues was omitted in each map. These maps were examined only in the areas of omitted residues. The models and corresponding electron density maps were examined using the program O (17). Several cycles of manual rebuilding using O, followed by automated refinement with X-PLOR, were carried out. The manual rebuilding, using omit maps generated as described above, involved adjusting side chain torsion angles and occasionally changing residue rotamers.

In all refinement protocols, bulk solvent mask corrected reflection files including the 20-2.3 Å data in 40% acetonitrile and 20-2.2 Å data in 20% dioxane were used. The data in the highest resolution shell in both cases were used in the refinement, because the *R*-free (*R*-factor calculated for 10% of the total reflections left out of the refinement (18)) improved in the lower resolution shells when these data were included. The initial refinement consisted of simulated annealing runs, which involved an initial positional refinement, 40 cycles, followed by 5 trials of heating to 3000 K and then cooling in increments of 25 K with 50 steps of molecular dynamics simulations for 0.5 fs at each temperature. This was followed by a restrained individual *B*-factor refinement, 25 cycles,

performed on the coordinate file (trial) with the lowest *R*-free value (18). Later refinement rounds were comprised of conventional positional refinements, 30-70 cycles, followed by restrained individual *B*-factor refinements, 25 cycles.

Solvent molecules were included in the models as described previously (8) by means of manual building in O and automated refinement (positional and individual *B*-factor) in X-PLOR. Potential water molecules were located using the $|2F_{obs}-F_{calc}|$ electron density map at the 1.5σ contour level. Initially, they were built only into density peaks that clearly exhibited the appropriate shape; any peaks that were possibly due to acetonitrile molecules in 40% acetonitrile, or dioxane molecules in 20% dioxane, were excluded. Water molecules were retained if their electron density in subsequent $|2F_{obs}-F_{calc}|$ maps persisted after the positional refinement, and if they fulfilled the following criteria: oxygen atom within 3.4 Å of a subtilisin oxygen or nitrogen atom, or bound water with good hydrogen-bonding geometry, *B*-factors less than 65 $Å^2$, and real space fit correlation coefficients above 70%.

For subtilisin in 40% acetonitrile, the acetonitrile model used was that found in the PDB file 1scb. For subtilisin in 20% dioxane, the dioxane model used is in 1af4. The acetonitrile and dioxane parameter and topology files for X-PLOR were generated with the aid of the program XPLO2D (19).

In 40% acetonitrile, potential acetonitrile molecules were located using the $|2F_{obs}-F_{calc}|$ map. Electron density was considered to represent a potential acetonitrile molecule if the peak was above 1.5σ with the acetonitrile's nitrogen within hydrogen-bonding distance to a hydrogen-bond donor on the protein or to a bound water molecule.

Acetonitrile electron density was recognized by comparison with F_{calc} maps calculated for a theoretical acetonitrile molecule. A complication that must be noted is that the electron density for a single acetonitrile molecule at this resolution appears similar to the density for two hydrogen-bonded water molecules. In some cases, the hydrogen-bonding pattern in the model was consistent only with two waters; however, when it was ambiguous, the following criteria were used to distinguish between the possibilities of the density corresponding to one acetonitrile or two waters. Real space fit correlation coefficients and *R*-free values were calculated for both an acetonitrile molecule and two waters molecules built into such an electron density peak. Acetonitrile molecules were kept only if these parameters, before and after refinement, were superior to those for the two waters. Furthermore, putative acetonitrile molecules with average *B*-factors greater than 65 $Å^2$ or real space fit correlation coefficients less than 75% were discarded. Five acetonitrile molecules were included in the final model as well as 74 water molecules and one Ca^{2+} molecule. The final *R*-factor and *R*-free values were 0.186 and 0.247, respectively (Table 4.1).

For subtilisin in 20% dioxane, potential dioxane molecules were also located using the $|2F_{obs}-F_{calc}|$ map. Electron density was considered to represent a potential dioxane if the peak was above 1.5σ with at least one dioxane oxygen within hydrogen-bonding distance to a hydrogen-bond donor on the protein or an enzyme-bound water molecule. Dioxane electron density was recognized by comparison with F_{calc} maps calculated for a theoretical dioxane molecule. Dioxane molecules were kept if their electron density in subsequent $|2F_{obs}-F_{calc}|$ maps persisted after refinement, with *B*-factors

less than 65 $Å^2$, and real space fit correlation coefficients above 75%. One dioxane molecule was included in the final model as well as 91 water molecules and one Ca^{2+} molecule. The final *R*-factor and *R*-free values were 0.187 and 0.245, respectively (Table 4.1).

The coordinates for both structures can be found in the Protein Data Bank (PDB; Brookhaven National Laboratory, entries … in 40% acetonitrile and .. in 20% dioxane).

Model validation. The final models in 40% acetonitrile and 20% dioxane contained no residues in disallowed regions of the Ramachandran plot, as determined with the program Procheck (20). The model in 40% acetonitrile contained two residues (Asn-25 and Thr-211) in generously allowed regions. The root mean squared deviation from ideal values for both structures can be found in Table 4.1. The real space fit correlation coefficients for the models in 40% acetonitrile and 20% dioxane were 0.91 ± 0.04, protein and 0.87 ± 0.05, solvent; and 0.90 ± 0.04, protein and 0.87 ± 0.06, solvent, respectively using the $|2F_{obs}-F_{calc}|$ maps. The overall *B*-factors for the models in 40% acetonitrile and 20% dioxane, 28 and 23 $Å^2$, respectively, were comparable to the Wilson *B*-factors for the crystal data of 35 and 27 $Å^2$, respectively. The average coordinate error predicted by Luzzati (21-23) or sigmaA (24) analyses (performed in X-PLOR) for both models was approximately 0.3 Å using the test set of reflections.

D. References

1. Fitzpatrick; P. A., Steinmetz; A. C. U.; Ringe, D.; Klibanov, A. M. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8653-8657.

2. Peisach, E.; Casebier, D.; Gallion, S. L.; Furth, P.; Petsko, G. A.; Hogan, J. C. Jr.; Ringe, D. *Science* **1995**, *296*, 66-69.

3. Allen, K. N.; Bellamacina, C. R.; Ding, X.; Jeffery, C. J.; Mattos, C.; Petsko, G. A.; Ringe, D. *J. Phys. Chem.* **1996**, *100*, 2605-2611.

4. Mattos, C.; Ringe, D. *Nature Biotechnol.* **1996**, *14*, 595-599.

5. Miranker, A.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 29-34.

6. Caflish, A.; Miranker, A.; Karplus, M. *J. Med. Chem.* **1993**, *36*, 2142-2167.

7. Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675-681.

8. Schmitke, J. L.; Stern, L. J.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4250-4255.

9. Joseph-McCarthy, D.; Federov, A. A.; Almo, S. L. *Protein Eng.* **1996**, *9*, 773-780.

10. Connolly, M. L. *Science* **1983**, *221*, 709-712.

11. Liepinsh, E.; Otting, G. *Nature Biotechnol.* **1997**, *15*, 264-268.

12. Schmitke, J. L.; Stern, L. J.; Klibanov, A. M. *J. Am. Chem. Soc., submitted.*

13. Neidhart, D. J.; Petsko, G. A. *Protein Eng.* **1988**, *2*, 271-276.

14. Otwinowski, Z.; Minor, W. *Meth. Enzymol.* **1997**, *276*, 307-326.

15. Collaborative Computational Project, Number 4 *Acta Cryst.* **1994**, *D50*, 760-763.

16. Brünger, A. T.; Kuriyan, J.; Karplus, M. *Science* **1987**, *35*, 458-460.

17. Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Cryst.* **1991**, *A47*, 110-119.

18. Brünger, A. T. *Nature* **1992**, *355*, 472-475.

19. G. J. Kleywegt, *ESF/CCP4 Newsletter* **1995**, *32*.

20. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283-291.

21. Luzzati, V. *Acta Cryst.* **1952**, *5*, 802-810.

22. Kleywegt, G. L.; Jones, T. A. *Structure* **1994**, *2*, 1241-1258.

23. Kleywegt, G. L.; Brünger, A. T. *Structure* **1996**, *4*, 897-904.

24. Read, R. J. *Acta Cryst.* **1986**, *A42*, 140-149.

V. BIOGRAPHICAL NOTE

EDUCATION

1998 Ph.D. in Biological Chemistry, **Massachusetts Institute of Technology**
Thesis title "Subtilisin Structure and Catalysis in Non-Aqueous Media"
Advisor Professor Alexander M. Klibanov

1992 B.Sc. First Class Honors in Chemistry, **Queen's University, Canada**

EXPERIENCE

1992 - 1998 **Graduate Research Assistant,** Massachusetts Institute of Technology

- Investigated the structures of crystalline enzymes and their substrate complexes in organic solvents via protein X-ray crystallography
- Performed molecular modeling of proteins and their complexes using Insight II / Discover, O and related programs, X-PLOR, CCP4 Suite, GRASP
- Determined the causes of reduced enzymatic activity in anhydrous solvents relative to that in water; involved enzyme kinetics: GC, HPLC, UV, pH STAT, protein crystallization and lyophilization
- Conducted UNIX system administration: installed, organized, software
- Directed the research of several undergraduate and summer students
- Managed the users and was part of the team that maintained the X-ray Diffraction Facility: Rigaku RU200 rotating Cu anode generator, mirror optics, RAXIS II area detector
- Played a major role in the planning/design of the new (present) Klibanov Lab
- Participated in the writing of NIH/NSF grant proposals and DOE progress reports
- *Laboratory Instructor:* Taught an undergraduate introductory chemistry lab

1995 - 1998 **Laboratory Safety Officer**. Ensured that the laboratory was in compliance with federal and state laws covering academic labs.

Summer 1991 Performed ab initio studies of sulfate and phosphate systems, Queen's University

Summer 1990 Carried out analytical work consisting of the removal of uranium and thorium from ultra-pure water, National Research Council, Canada

AWARDS

1993	Excellence in Teaching Award, Chemistry Department, **MIT**
1988-1992	Queen's Provincial Scholarship
1988-1992	Canada Scholarship
1990	R.T. Mohan Scholarship in Chemistry

AFFILIATIONS Member of the American Chemical Society

LANGUAGES German and French

PUBLICATIONS

"Selective Hydrogen Bonding as a Mechanism for Differentiation of Sulfate and Phosphate at Biomolecular Receptor Sites, " G.R.J. Thatcher, D.R. Cameron, R. Nagelkerke, J.L. Schmitke, *Chemical Communications* **4**, 386-388 (1992)

"The Mechanistic Dissection of the Plunge in Enzymatic Activity upon Transition from Water to Anhydrous Solvents," J.L. Schmitke, C.R. Wescott, and A.M. Klibanov, *Journal of the American Chemical Society* **118**, 3360-3365 (1996)

"The Crystal Structure of Subtilisin Carlsberg in Anhydrous Dioxane and its Comparison with Those in Water and Acetonitrile," J.L. Schmitke, L.J. Stern, and A.M. Klibanov, *Proceedings of the National Academy of Science USA* **94,** 4250-4255 (1997)

"Comparison of X-ray Crystal Structures of an Acyl-Enzyme Intermediate of Subtilisin Carlsberg Formed in Anhydrous Acetonitrile and in Water," J.L. Schmitke, L.J. Stern, and A.M. Klibanov, *Journal of the American Chemical Society, submitted*

"Computational and Experimental Examination of the Effect of Inorganic Salts on Chymotryptic Enantioselectivity in Water," T. Ke, R. V. Rariy, J.L. Schmitke, and A.M. Klibanov, *Biotechnology and Bioengineering, submitted*

"Organic Solvent Binding to Crystalline Subtilisin in Mostly Aqueous Media and in the Neat Solvents," J.L. Schmitke, L.J. Stern, and A.M. Klibanov, *Biochemical and Biophysical Research Communications, submitted*