# Stability and Activity of Mesophilic Subtilisin E and Its Thermophilic Homolog: Insights from Molecular Dynamics Simulations

**Giorgio Colombo[†,‡] and Kenneth M. Merz, Jr.*[,‡]**

*Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802 and Istituto di Biocatalisi e Riconoscimento Molecolare CNR, via Mario Bianco, 9 20131, Milano, Italy*

*Received February 9, 1999*

**Abstract:** Herein we examine the origin of the high-temperature (350 K) behavior of a thermophilic mutant enzyme (labeled as 5-3H5; see Zhao and Arnold *Prot. Eng.* **1999**, *12*, 47−53) derived from subtilisin E by eight amino acid substitutions. Through the use of molecular dynamics (MD) simulations, we have provided molecular-level insights into how point mutations can affect protein structure and dynamics. From our simulations we observed a reduced rmsd in several key regions, an increased overall flexibility, an increase in the number of hydrogen bonds, and an increase in the number of stabilizing interactions in the thermophilic system. We also show that it is not a necessary requirement that thermophilic enzymes be less flexible than their mesophilic counterparts at low temperatures. However, thermophilic enzymes must retain their three-dimensional structures and flexibility at high temperatures in order to retain activity. Furthermore, we have been able to point out the effects of some of the single substitutions. Even if it is not possible yet to give general rules for rational protein design, we are able to make some predictions on how a protein should be stabilized in order to be thermophilic. In particular, we suggest that a promising strategy toward speeding up the design of thermally stable proteins would be to identify fluxional regions within a protein through the use of MD simulations (or suitable experiments). Presumably these regions allow for autocatalytic reactions to occur and are also involved in allowing water to gain access to the interior of the protein and initiate protein unfolding. These fluxional regions could also adversely affect the positioning of the catalytic machinery, thereby decreasing catalytic efficiency. Thus, once these locations have been identified, "focused" directed evolution studies could be designed that stabilize these "fluxional" regions.

## Introduction

Understanding the molecular-level origins of the thermal stability of proteins is an important issue for both practical and fundamental reasons. From the practical point of view, thermophilic enzymes can have import in the catalysis of reactions that require high temperatures to occur, for example, as in polymerase chain reactions (PCR) applications, or in the detergent industry.[1] From the fundamental point of view, the comparison between mesophilic enzymes and their thermophilic counterparts can help in understanding the molecular basis of thermostability[2] as well as furthering our understanding of the relationship between protein sequence and structure (i.e., protein folding). Hence, by increasing our understanding of the molecular-level origins of protein (thermo)stability, we will be able to begin to rationally engineer novel proteins with characteristics suitable for a particular application.

Homologous enzymes can be found in organisms that live in very different environments, covering a temperature range of about 200 degrees.[2] The three-dimensional structures of many proteins from thermophilic organisms have been solved and published; for example, thermolysin,[3] aldehyde ferredoxin oxido reductase,[4] glutamate dehydrogenase,[5] and rubredoxin.[5] The three-dimensional structures of thermostable mutants obtained by site-directed mutagenesis have also been solved and published, and the effects of mutations on stability have been investigated.[6] Comparison of the structures and sequences from these related proteins has led researchers to suggest many factors that could determine thermal stability in proteins. These factors include amino acid composition,[7] the effect of particular turns and loops,[8] salt bridges,[9] electrostatic interactions,[10] hydrogen bonding,[11] metal ion binding,[12] and the formation of disulfide bridges.[13] However, no "unifying" theory has been widely accepted and the reason for this is the fact that stability reflects

[†] The Pennsylvania State University.

[‡] Istituto di Biocatalisi e Riconoscimento Molecolare CNR.

(1) Adams, M. W. W.; Kelly, R. M. *Chem. Eng. News* **1992**, *73*, 32−42.

(2) Zhao, H.; Arnold, F. H. *Protein Eng.* **1999**, *12*, 47−53.

(3) Matthews, B. W.; Jansonius, J. N.; Colman, P. M.; Schenborn, B. P.; Dupourque, D. *Nature (London), New. Biol.* **1972**, *238*, 37−41.

(4) Chan, M. K.; Mukund, S.; Kletzin, A.; Adams, M. W. W.; Rees, D. C. *Science* **1995**, 1463−1469.

(5) Yip, K. S. P.; Stillman, T. J.; Britton, K. L.; Artymiuk, P. J.; Bocker, P. J.; Sedelnikova, S. E.; Engel, P. C.; Pasquo, A.; Chiavaluce, P.; Consalvi, V.; Scandurra, R.; Rice, D. W. *Structure* **1995**, *3*, 1147−1158.

(6) Bryan, P. N.; Rollence, M. L.; Pantoliano, M. W.; Wood, J.; Finzel, B. C.; Gilliland, G. L.; Howard, A. J.; Poulos, T. L. *Proteins: Struct., Funct., Genet.* **1986**, *1*, 326−334.

(7) Menendez-Arias, L.; Argos, P. *J. Mol. Biol.* **1989**, 397−406.

(8) Hardy, F.; Vriend, G.; van der Vinne, B.; Frigerio, F.; Grandi, G.; Venema, G.; Eijsink, V. G. H. *Protein Eng.* **1994**, *7*, 425−430.

(9) Perutz, M. F.; Raidt, H. *Nature* **1975**, 256−259.

(10) Spassov, V. Z.; Karshikoff, A. D.; Ladenstein, R. *Protein Sci.* **1995**, *4*, 1516−1527.

(11) Day, M. W.; Hsu, B. T.; Joshua-Tor, L.; Park, J. B.; Zhou, Z. H.; Adams, M. W. W.; Rees, D. C. *Protein Sci.* **1992**, *1*, 1494−1507.

(12) Toma, S.; Campagnoli, S.; Margarit, I.; Gianna, R.; Grandi, G.; Bolognesi, M.; De Filippis, V.; Fontana, A. *Biochemistry* **1991**, *30*, 97−106.

(13) Clarke, J.; Fersht, A. R. *Biochemistry* **1993**, *32*, 4322−4329.

a complex interplay of inter- and intramolecular interactions, and it is very difficult to delineate the role each of these disparate interactions play in destabilizing or stabilizing a protein molecule.[14–16] Moreover, homologous sequences generally differ at a substantial fraction of their amino acid positions, thereby making it extremely difficult to determine which mutations either stabilize or destabilize a protein in a systematic manner.

In light of these difficulties, a promising approach toward the elucidation of the molecular "mechanism" of thermostabilization is to study engineered proteins in which the effects of specific amino acid substitutions can be systematically evaluated.[17,18] The use of directed evolution has proven to be particularly useful in this case. This technique involves mutation, recombination, and screening or selection to accumulate the mutations required to achieve significant changes in protein function.[19] A wide variety of protein properties have been improved through this approach such as activity in nonaqueous solvents,[20] enhanced expression levels, enantioselectivity, substrate specificity,[21] and alkaline stability.[22]

The structural aspects of subtilisins have been well characterized[23–25] and, despite considerable differences in amino acid sequences, the secondary structure and the global fold of these proteins are very similar. Furthermore, the catalytic mechanism and thermal stabilities of the subtilisins have been extensively studied from the experimental and theoretical point of view.[25–28] The active site of the subtilisins is composed of the classic catalytic triad Ser 221, His 64 and Asp 32. Subtilisin E is expressed as a pre-pro enzyme, with a pre-sequence of 29 amino acids necessary for secretion and a pro-sequence of 77 residues needed for proper folding. The mature reactive enzyme consists of a single polypeptide chain of 275 amino acids with no disulfide bridges. Zhao and Arnold have used directed evolution to convert *Bacillus subtilis* subtilisin E into an enzyme functionally equivalent to its thermophilic homologue thermitase from *Thermoactinomices vulgaris*. Five generations of random mutagenesis, recombination, and screening were sufficient to create subtilisin 5-3H5, whose half-life at 83 °C (3.5 min) and temperature optimum for activity ($T_{opt}$ = 76 °C) are identical to those of thermitase. The $T_{opt}$ of the evolved enzyme is 17 °C higher and its half-life at 65 °C is >200 times that of wild-type subtilisin E. 5-3H5 is more active toward the hydrolysis of succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide than wild type over the whole range of temperatures from 10 to 90 °C. Thermitase differs from subtilisin E at 157 amino acids. However, only eight amino acid substitutions were sufficient to convert

subtilisin E into an enzyme equally thermostable. These substitutions were the following: Pro 14 Leu, Asn 76 Asp, Asn 118 Ser, Ser 161 Cys, Gly 166 Arg, Asn 181 Asp, Ser 194 Pro, and Asn 218 Ser. Only two of those (Asn 218 Ser and Asn 76 Asp) were found in thermitase.[2]

In this paper we will address the problem of activity and thermostability of the serine protease enzyme subtilisin E and the 5-3H5 mutant.[2] We decided to use molecular dynamics (MD) simulations for the two proteins at two different temperatures (300 and 350 K), to get a more complete understanding, at the molecular level, of the reasons that determine the enhanced thermostability of the mutant. Molecular dynamics (MD) simulations have been used by several authors to address the problem of differential thermostability of homologous proteins,[29] to investigate protein stability vs temperature,[30] and to investigate the thermal and denaturant-catalyzed unfolding of proteins.[31,32]

## Computational Details

Molecular dynamics (MD) simulations were carried out on subtilisin E and its mutant derivative 5-3H5 starting from two different configurations which we labeled SRT (subtilisin room temperature) and HRT (5-3H5 room temperature), respectively. All simulations were performed using the ROAR 1.0[33] module of AMBER 5.0[34] in conjunction with the AMBER94[35] all atom force field and TIP3P water.[36]

The starting coordinates for all the heavy atoms of Subtilisin E were taken from the end point of a previous simulation of this system.[37] The thermostable mutant 5-3H5 differs from the wild-type subtilisin E at eight amino acid positions (Pro 14 Leu, Asn 76 Asp, Asn118 Ser, Ser 161 Cys, Gly 166 Arg, Asn 181 Asp, Ser 194 Pro, and Asn 218 Ser). The amino acid substitutions from subtilisin E to 5-3H5 were carried out using MIDAS.[38] The mutated positions were then minimized using the ROAR 1.0 module of AMBER, while the remainder of the protein was held fixed.

For the 5-3H5 and subtilisin E systems, one and two counterions (chloride ions) were added, respectively, to neutralize the systems. The two proteins were then solvated using TIP3P water molecules, and this resulted in 8350 and 8204 water molecules being added to subtilisin E and 5-3H5, respectively. The two configurations were then minimized for 2000 steps using steepest descent energy minimization for the first 500 steps and conjugated gradient minimization for the remaining part. Periodic boundary conditions were used at this stage. These minimized structures of subtilisin E and 5-3H5 were used as the reference structures for all of our simulations and subsequent analyses.

The minimized configurations were used as starting points for molecular dynamics (MD) simulations. The temperatures of both systems were raised from 0 to 300 K over the first 15 ps of the simulations. The temperature and pressure were maintained at 300 K

(14) Jaenicke, H.; Schurig, H.; Beaucamp, N.; Ostendorp, R. *Adv. Protein Chem.* **1996**, *48*, 181–269.

(15) Vogt, G.; Argos, P. *Folding Des.* **1997**, *2*, S40–S46.

(16) Richards, F. M. *Cell. Mol. Life Sci.* **1997**, *53*, 790–802.

(17) Fersht, A. R.; Serrano, L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 75–83.

(18) Matthews, B. W. *Adv. Protein Chem.* **1995**, *46*, 249–278.

(19) Arnold, F. H. *Acc. Chem. Res.* **1998**, *31*, 125–131.

(20) Chen, K.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5618–5622.

(21) Kuchner, O.; Arnold, F. H. *Trends Biotechnol.* **1997**, *15*, 523–530.

(22) Cunningham, B. C.; Wells, J. A. *Protein Eng.* **1987**, *1*, 319–325.

(23) Markland, J. F. S.; Smith, E. L. *Subtilisins: Primary Structure, Chemical and Physical Properties*; Boyer, P. D., Ed.; Academic Press: New York, 1971; Vol. 3, pp 561–608.

(24) Krait, J., Ed. *Subtilisin: X-ray Structure*; Boyer, P. D., Ed.; Academic Press: New York, 1971; Vol. 3, pp 547–560.

(25) Wells, J. A.; Estell, D. A. *Trends Biochem. Sci.* **1988**, *13*, 291–297.

(26) Seizen, R. J.; de Vos, W. M.; Leunissen, J. A. M.; Dijkstra, B. W. *Protein Eng.* **1991**, *4*, 719–737.

(27) Takagi, H.; Matsuzawa, H.; Ohta, T.; Yamasaki, M.; Inouye, M. *Ann. NY Acad. Sci.* **1992**, 52–59.

(28) Daggett, V.; Schröder, S.; Kollman, P. *J. Am. Chem. Soc.* **1991**, *113*, 8926–8935.

(29) Lazaridis, T.; Lee, I.; Karplus, M. *Protein Sci.* **1997**, *6*, 2589–2605.

(30) Mark, A. E.; van Gusteren, W. F. *Biochemistry* **1992**, *31*, 7745–7748.

(31) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1993**, *32*, 4175–4184.

(32) Tirado-Rives, J.; Orozco, M.; Jorgensen, W. L. *Biochemistry* **1997**, *36*, 7313–7329.

(33) Cheng, A.; Stanton, R. S.; Vincent, J. J.; Damodaran, K. V.; Dixon, S. L.; Hartsough, D. S.; Best, S. A.; Merz, K. M., Jr. *ROAR*, 1.0 ed.; The Pennsylvania State University, 1997.

(34) Case, D. A.; Pearlman, D. A.; Caldwell, J. C.; Cheatham, T. E. I.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M., Jr.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 5.0*; University of California, San Francisco, 1997.

(35) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(37) Toba, S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1997**, *119*, 9939–0048.

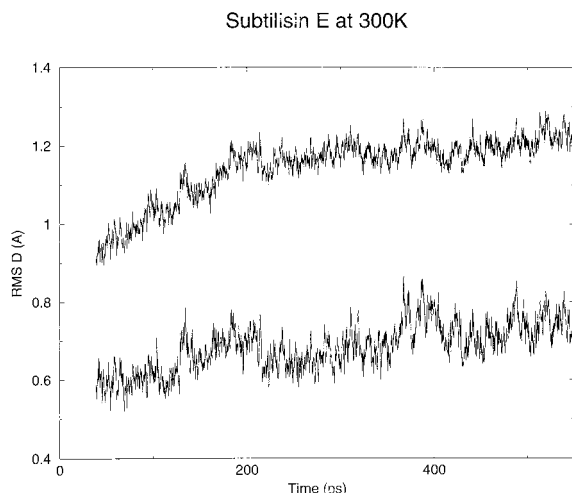(38) Ferrin, T. E. *J. Mol. Graphics* **1988**, *6*, 13–27.

Subtilisin E at 300K



**Figure 1.** rms deviations between the instantaneous computed structures and the crystal structure for all residues as a function of time for Subitlisin E at 300 K.

5-3h5 at 300K



**Figure 2.** rms deviations between the instantaneous computed structures and the crystal structure for all residues as a function of time for 5-3H5 at 300 K.

and 1 atm using the Nóse Hoover Chain (NHC) algorithm for temperature and pressure coupling.[39] Ewald sums and long-range van der Waals corrections were used for the accurate treatment of long-range electrostatic and long-range van der Waals interactions, respectively.[40] A $K_{max}$ value of 8 was used along with a nonbond pairlist cutoff of 10 Å, and the pairlist was updated every 50 steps. In all simulations a 1.5 fs time step was used. In both cases, the bond lengths were constrained to their equilibrium values using the SHAKE[41] algorithm with a tolerance of 0.0005 Å. Periodic boundary conditions were applied in all cases.

Both the subtilisin E and the 5-3H5 simulations at 300 K (labeled SRT and HRT, respectively) covered 550 ps. Starting from the structures obtained after 450 ps, the temperature for both systems was raised to 350 K over a 15 ps time interval. The high-temperature simulations for subtilisin E and 5-3H5 (labeled SHT and HHT, respectively) were carried out for a total time of 1.2 ns. The last 300 ps of coordinates for the four systems (from 250 to 550 ps for SRT and HRT, and from 0.9 to 1.2 ns for SHT and HHT) were saved for analysis every 50 time steps.

## Results and Discussion

**rms Deviations.** The backbone and total root-mean-square deviation (rmsd) (between the instantaneous structures and the starting "crystal" structures) are plotted as a function of time in Figures 1 (SRT) and 3 (SHT) for the native enzyme and in Figures 2 (HRT) and 4 (HHT) for the 5-3H5 system at 300 and 350 K, respectively. Table 1 reports the average rms values for the 4 systems in the time range over which the analysis was carried out. The SRT and the HRT simulations have reached equilibrium after ∼200 ps for both the total and the backbone rms deviations. Their average rms deviation values are very similar, and both of the proteins at 300 K show a very close resemblance to the starting structure. Thus, we conclude that both of these systems are able to preserve structural features which are necessary for activity and stability at 300 K, which is consistent with experimental observation.[2]

After 450 ps at 300 K the temperature was raised to 350 K for both the wild-type subtilisin E and for the thermostable mutant 5-3H5. Through a comparison of the two systems at high temperature, we found that the rmsd for the HHT simulation (Figure 4) stabilizes after ∼200 ps at 350 K, while
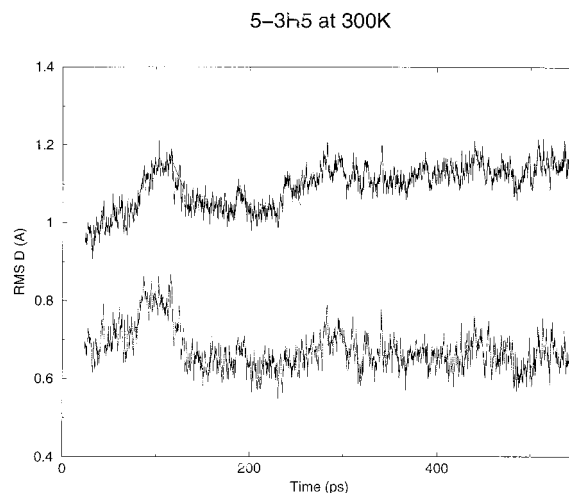
(39) Cheng, A.; Merz, K. M., Jr. *J. Phys. Chem.* **1996**, *100*, 1927−1937.
(40) Smith, P. E.; Pettitt, B. M. *J. Chem. Phys.* **1996**, *105*, 4289−4293.
(41) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.
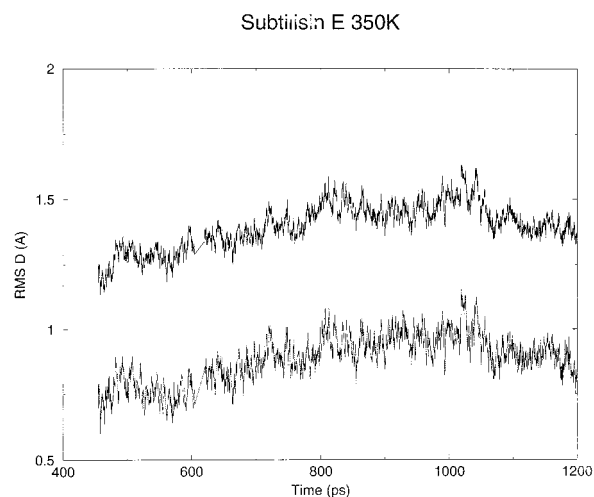
Subtilisin E 350K



**Figure 3.** rms deviations between the instantaneous computed structures and the crystal structure for all residues as a function of time for Subitlisin E at 350 K.
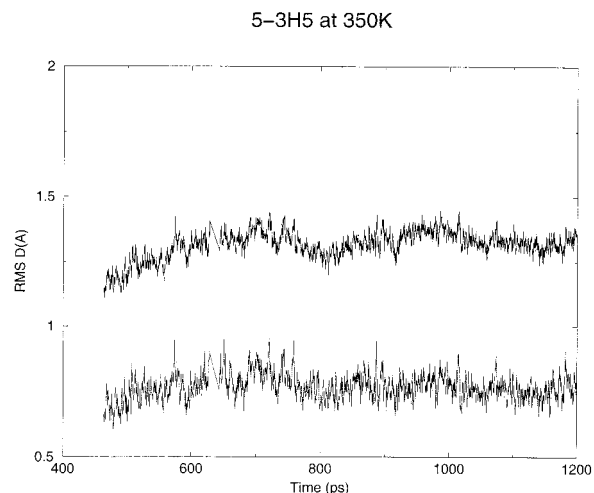
5-3H5 at 350K



**Figure 4.** rms deviations between the instantaneous computed structures and the crystal structure for all residues as a function of time for 5-3H5 at 350 K.

the wild-type system (SHT; see Figure 3 takes longer to stabilize. For both high temperature simulations the final 300 ps of the respective trajectories were equilibrated (see Figures 3 and 4).

**Table 1.** Summary of the Radius of Gyration, rms Deviation, rms Fluctuations, SASA, and Hydrogen-Bonding Data from the Simulations

|  | SRT | SHT | HRT | HHT |
|---|---|---|---|---|
| rms Deviation (Å) | | | | |
| total | 1.19 | 1.43 | 1.12 | 1.33 |
| backbone | 0.71 | 0.93 | 0.66 | 0.76 |
| RMS Fluctuation (Å) | | | | |
| total | 1.52 | 1.89 | 1.50 | 1.99 |
| backbone | 1.41 | 1.76 | 1.40 | 1.87 |
| Radius of Gyration (Å) | | | | |
| average | 16.82 | 16.93 | 16.82 | 16.83 |
| Hydrogen Bonding | | | | |
| >90% | 164 | 139 | 168 | 151 |
| <10% | 609 | 747 | 597 | 739 |
| SASA (Å$^2$) | | | | |
| total | 9999.25 | 10119.9 | 9745.0 | 9775 |
| hydrophobic | 2536 | 2631 | 2705 | 2760 |
| hydrophilic | 7049 | 7014 | 7462 | 7487 |
| Ser 221 | 9.91 | 11.00 | 10.96 | 10.02 |
| His 64 | 41.01 | 45.93 | 44.04 | 45.49 |
| Asp 32 | 0.59 | 1.60 | 0.58 | 1.08 |
| Tyr 217 | 68.65 | 86.23 | 75.07 | 75.43 |
| Glu 156 | 118.86 | 126.23 | 94.85 | 86.77 |

For HHT the increase in the rmsd of the backbone atoms was only 0.1 relative to HRT, while for SHT the increase was 0.2 relative to SRT. On the other hand, the total rmsd for these two systems was essentially identical (difference of only ~0.1), but both the SRT and SHT simulations had higher rmsds than the corresponding HRT and HHT simulations.

Overall, we found that the thermostable mutant protein (HHT) maintained its global three-dimensional structure and had a close correspondence to the crystal structure. The observed rms deviation values for SHT are not typical of an unfolding process, but from Figure 3 we can infer that the native protein is of lower stability at 350 K. We note that we cannot quantify "stability" in terms of thermodynamic quantities ($\Delta G_{unfolding}$, etc.) using MD simulations, but the fact that the native protein, at high temperatures, starts to unfold indicates a reduction in the stability of the system (i.e., movement away from the global minimum at mesophilic temperatures). We also discuss activity below and, again, we cannot quantify activity in terms of thermodynamic quantities ($k_{cat}$, etc.), but we assume that if the protein is not at its global minimum and if the active site is distorted it will be difficult for an enzyme to maintain optimal activity.

Graphical inspection of the structure of the native system showed high mobility of three loop regions situated on the surface of the protein: the loops formed by residues 96−105, 158−162, and 38−45. The results from our visual inspection of the protein are supported by the calculation of the average rmsd on a per residue basis (see Figures 5 and 6). By comparing Figures 5 and 6, we observe a substantial similarity between the SRT and HRT systems. On the other hand, a comparison of SHT and HHT show a striking difference in the patterns between the two graphs. In particular, we notice very high rmsd values (range of 3−5.5) for the residues between 96 and 105 in the high-temperature subtilisin E simulation (SHT). These residues tend to assume conformations that bring the side chains into contact with solvent. This motion also opens up a large pocket close to the active site and exposes the interior of the protein to the penetration of solvent. Most of the secondary structure is retained, but this type of loop movement and solvent penetration into the core of the protein has been proposed as important factors that favor autolysis and in unfolding
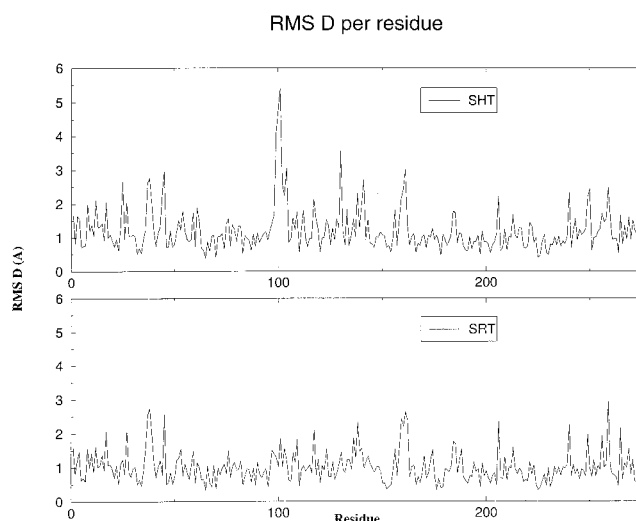


**Figure 5.** rms deviations per residue averaged over the last 300 ps for subtilisin E: top, subtilisin E at 350 K; bottom, subtilisin E at 300 K.
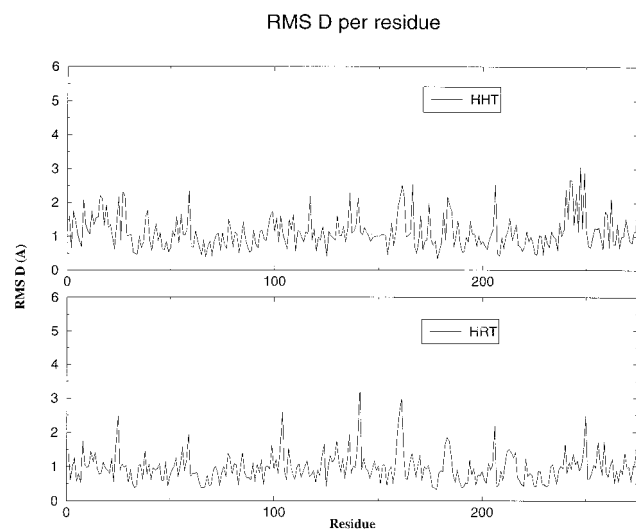


**Figure 6.** rms deviations per residue averaged over the last 300 ps for 5-3H5: top, 5-3H5 at 350 K; bottom, 5-3H5 at 300 K.

initiation.[8,42−45] Both of these phenomena can be correlated with the loss of activity and stability of subtilisin E at 350 K. High rmsd (rmsd = ~3.0) also characterizes the 158−162 loop in SHT, as well. This loop also tends to "open" and thereby exposes its amino acid residues to solvent. These two "high rmsd regions" are not present in the case of the thermostable mutant at 350 K (HHT), though the rmsd of the 158−162 loop is only slightly reduced over SHT (rmsd = ~2.5).

Arnold and co-workers[46] have also noted that loop motions can be correlated with thermostbility in a related system (*p*-nitorbenzyl esterase (pNB)). They observed that an evolutionarily evolved thermostable pNB had decreased loop motions via X-ray crystallographic studies. This observation is strongly supported by our own results on subtilisin E.

(42) Matouschek, A.; Kellis, J. T.; Serrano, L.; Fersht, A. R. *Nature* **1989**, *340*, 122−126.

(43) Serrano, L.; Matouschek, A.; Fersht, A. R. *J. Mol. Biol.* **1992**, *224*, 805−818.

(44) Jackson, S. E.; Fersht, A. R. *Biochemistry* **1991**, *30*, 10436−10443.

(45) Fontana, A.; Fassina, G.; Vita, C.; Dalzoppo, D.; Zamai, M.; Zambonin, M. *Biochemistry* **1986**, *25*, 1847−1851.

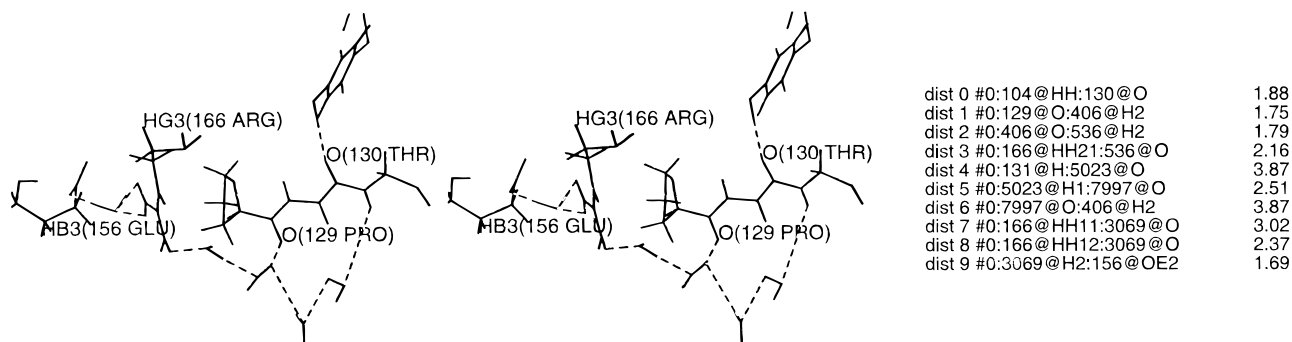(46) Spiller, B.; Gershenson, A.; Arnold, F. H.; Stevens, R. C. 1999, to be submitted for publication.

| dist 0 #0:104@HH:130@O | 1.88 |
| dist 1 #0:129@O:406@H2 | 1.75 |
| dist 2 #0:406@O:536@H2 | 1.79 |
| dist 3 #0:166@HH21:536@O | 2.16 |
| dist 4 #0:131@H:5023@O | 3.87 |
| dist 5 #0:5023@H1:7997@O | 2.51 |
| dist 6 #0:7997@O:406@H2 | 3.87 |
| dist 7 #0:166@HH11:3069@O | 3.02 |
| dist 8 #0:166@HH12:3069@O | 2.37 |
| dist 9 #0:3069@H2:156@OE2 | 1.69 |

**Figure 7.** Three-dimensional representation of the hydrogen-bonding pattern for Ser 118 in 5-3H5.



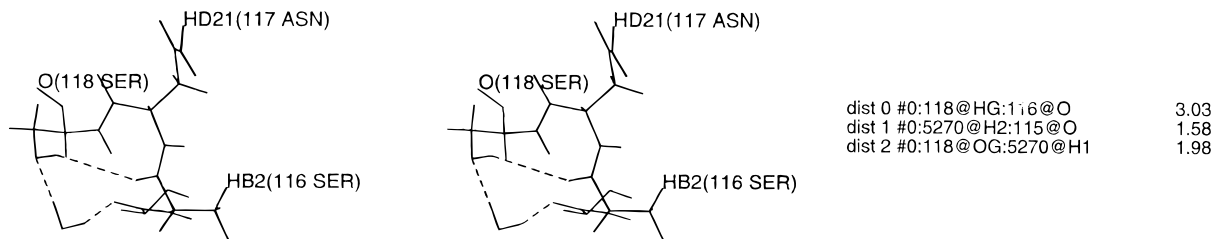| dist 0 #0:118@HG:116@O | 3.03 |
| dist 1 #0:5270@H2:115@O | 1.58 |
| dist 2 #0:118@OG:5270@H1 | 1.98 |

**Figure 8.** Three-dimensional representation of the hydrogen-bonding pattern for residues 156, 166, 130, and 104 in 5-3H5.

In the three-dimensional structure of 5-3H5 we found that the OH group of Tyr 104 forms a strong hydrogen bond with the backbone NH group of Thr 130 (see Figure 7). Moreover, Tyr 104 is situated close to the end of an α-helix at the beginning of which an Asn has been substituted with a Ser (Asn 118 Ser). The side chain OH group of this Ser residue can hydrogen bond directly or via a one-water bridge with the carbonyl of Ser 116 or with the carbonyl of residue 115 via a one- or two-water bridge, thereby stabilizing this α-helix (see Figure 8). The carbonyl group of Pro 129 and the amide group (NH) of Gly 131 are hydrogen-bound, via multiple water molecules, to one of the $NH_2$ groups of the side chain of Arg 166 (which is one of the mutated sites: Gly 166 Arg). The side chain $NH_2$ groups of this residue also hydrogen bond with the side chain oxygens of Asp 156 directly or via a water bridge. Moreover, the distance between the negatively charged side chain of Asp 156 and the positively charged side chain of Arg 166 favors the presence of an electrostatic interaction (see Figure 7). Thus, through a series of hydrogen-bonding interactions, we find that the Asn 118 Ser and Gly 166 Arg point mutations facilitate the stabilization of the 5-3H5 system relative to the native protein by eliminating flexible loops (96−105 and 158−162) that could serve as an autolytic site. Moreover, Arg 166 in the 5-3H5 system is located at the end of a distinct elongated cleft, known as the S1 pocket, which determines substrate specificity. The stabilization due to this residue may also be attributed to a reduction in cavity volume, which could affect the autolytic process by preventing the substrate protein from fitting into that pocket.[2]

The slightly lower rms deviation of the loop formed by residues 158−162 in the thermostable mutant at 350 K can also be attributed to the formation of a hydrogen bond between the side chain OH group of Ser 163 and the side chain carboxylate group of Glu 195. Glu 195 is, in fact, directly connected in 5-3H5 to Pro 194 (Ser 194 Pro). The presence of this mutation appears to stabilize this loop, thereby forcing Glu 195 to have its side chain closer to Ser 163 which results in a longer-lived hydrogen-bond interaction. This enhanced stabilization is not present in the SHT system, where residue 194 is a Ser. From the preceding analysis, we can qualitatively determine how three

point mutations (Asn 188 Ser, Gly 166 Arg, and Ser 194 Pro) provided an enhanced hydrogen-bonding pattern that helped the mutant protein retain its three-dimensional structure at higher temperatures.

Overall, our rms deviation results for the two systems, at both 300 and 350 K, are in good accord with the experimental observation that wild-type subtilisin E has low activity and stability at high temperature. On the other hand, the thermostable mutant 5-3H5 can preserve both of these properties at high temperatures, presumably by maintaining its global three-dimensional structure.

**Protein Flexibility.** The rms fluctuation from average coordinates is a good measure of the flexibility of the system during a given time period. Thus, we decided to calculate the rms fluctuations about the time averaged structure for the last 300 ps of both the 300 and 350 K simulations. From Table 1, as expected, we observe that the total average flexibility of the protein is higher than the average flexibility of the backbone for all simulations. Data at 300 K indicate that the average rms fluctuation for both the wild-type and the mutant system are very similar, thus supporting the experimental observation of comparable reactivities and stabilities at this temperature.[2] The analysis of the data at 350 K shows that the mutant system has higher average flexibility values for both the total and backbone atoms.

To gain further insight into this observation, we compared the rms fluctuations per residue for the different systems. From this analysis, we see that the flexibility characteristics of subtilisin E and 5-3H5 are remarkably similar at 300 K, while significant differences were observed for the high-temperature simulations (see Figures 9 and 10). The rms fluctuations for the catalytic triad (Table 2) are comparable (with the mutant protein generally being higher than the native system) for the 300 K simulations: thus, in both systems the active site retains its overall structure and catalytic ability. The fact that the two enzymes have comparable activity and stability at 300 K is again reflected in their analogous flexibility properties for the active site residues. For the 350 K simulations, the rms flexibility of the active site residues are greater for the mutant than for the native protein, while the rmsds of key residues in the native
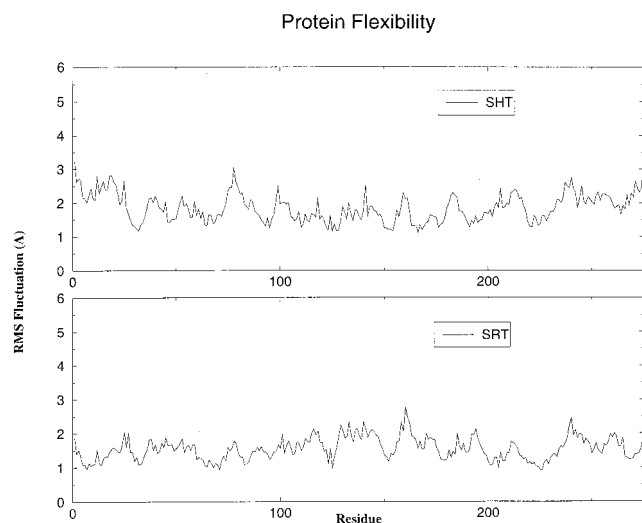
Protein Flexibility



**Figure 9.** rms fluctuations per residue about the time averaged structure for subtilisin E: top, subtilisin E at 350 K; bottom, subtilisin E at 300 K.
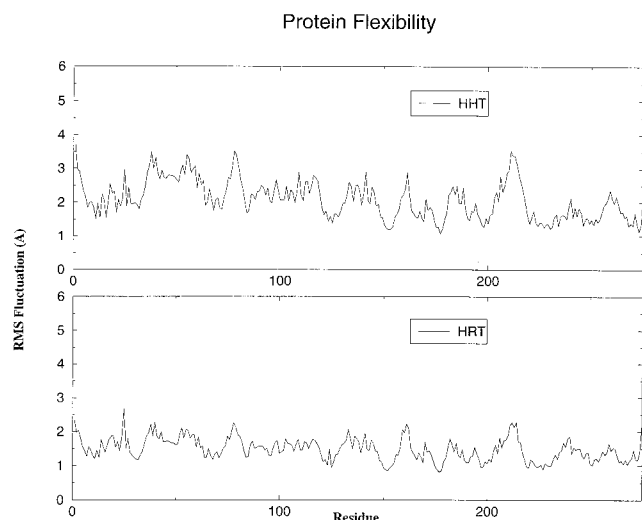
Protein Flexibility



**Figure 10.** rms fluctuations per residue about the time averaged structure for 5-3H5: top, 5-3H5 at 350 K; bottom, 5-3H5 at 300 K.

**Table 2.** Flexibility and rms Deviations of Active Site Residues

|  | SRT | SHT | HRT | HHT |
|---|---|---|---|---|
| rms Deviations of Acitve Site Residues (Å) | | | | |
| Ser 221 | 1.14 | 1.44 | 1.16 | 1.23 |
| His 64 | 0.64 | 0.63 | 0.57 | 0.65 |
| Asp 32 | 0.51 | 0.52 | 0.50 | 0.53 |
| Asn 155 | 0.93 | 0.99 | 0.89 | 0.87 |
| RMS Flexibility of Acitve Site Residues (Å) | | | | |
| Ser 221 | 1.08 | 1.28 | 1.17 | 1.56 |
| His 64 | 1.05 | 1.36 | 1.23 | 1.91 |
| Asp 32 | 1.09 | 1.17 | 1.17 | 1.81 |

system are larger (e.g., Ser 221 and Asn 155). For example, the greater flexibility of Ser 221 could play a key role in dimished catalytic activity because this motion will have to be "frozen" prior to proton transfer and attack at the electrophilic carbonyl carbon of the peptide substrate.

The rms fluctuations per residue for the high-temperature simulations show interesting trends. In particular, on average, 5-3H5 tends to be more flexible at high temperature than the native system (see Table 2). Flexibility implies increased conformational entropy and, therefore, favors increased thermodynamic stability at higher temperatures.[29] The increased

rigidity of wild-type subtilisin E, at high temperature, does not let the protein explore regions of conformational space that restricts the ability of the protein to reach a state in which it could still have high activity (i.e., $k_{cat}$) and stability (i.e., against autocatalytic processes, etc.) at 350 K. On the other hand, 5-3H5, thanks to its increased flexibility, can sample a wider portion of phase space, thus populating a larger number of minimum energy states. This can slow the unfolding process by increasing the activation barrier between the folded (and catalytically active) state and the transition region for unfolding. Indeed, Fersht et al. have correlated the rate of unfolding with thermodynamic stability, through linear free-energy relationships.[42] In their hypothesis, the more stable a protein is thermodynamically (relative to its unfolded state), the higher the barrier for unfolding. This could certainly be the case for 5-3H5, which shows relatively higher rms fluctuations at 350 K. Finally, greater flexibility, at high temperatures, could facilitate the binding and release of substrates and products.

Recently, it was suggested that the adjustment of conformational flexibility is a critical factor in the adaptation of a mesophilic protein to high-temperature conditions.[47] Thus, it was observed in 3-isopropylmalate dehydrogenase (IPMDH) that the thermophilic enzyme (from *Thermus thermophilus*) was less conformationally flexible than the mesophilic enzyme (from *E. coli*) and, hence, less active. At high temperatures, the thermophilic enzyme became as conformationally flexible as the native protein at low temperatures. Given our (and the experimental results of Zhao and Arnold[2]) results. this observation is not general. We find that the mesophilc and thermophilic forms of these enzymes are equally flexible at low temperature (300 K). However, at high temperature (350 K) the native protein unfolds partially and is less flexible, while the mutant enzyme does not unfold but is more conformationally flexible than at 300 K. Thus, the observation of Závodszky et al., while correct for IPMDH, is not true in the present case. Clearly, conformational flexibility is necessary for activity, but thermophilic enzymes need not be inactive at low temperatures (e.g., 300 K).

An important point should be kept in mind: rms fluctuations are calculated from the time-averaged equilibrated structure, which is different from the crystal structure. Thus, a low rms fluctuation does not mean that a specific secondary structure element was conserved. It only provides us with an insight into the flexibility of these elements. Hence, even if rms fluctuation values are low (e.g. SHT case), the secondary structure elements can have undergone major changes, as observed in the previous section, through the calculated rms deviations from the crystal structure.

**Radius of Gyration.** The radius of gyration ($R_{gyr}$) is defined as the mass-weighted rms distance of a collection of atoms from their common center of mass. Hence, this analysis gives us insight into the overall dimensions of the protein. The results of the $R_{gyr}$ calculation performed over the last 300 ps of the four simulations (SRT, HRT, SHT, and HHT) are reported in Table 1.

These results show that the $R_{gyr}$ values for the wild-type subtilisin E and the mutant enzyme at 300 K are essentially the same. However, an increase in $R_{gyr}$ is observed for the wild-type protein at high temperature (SHT) relative to the mutant enzyme. This variation is a further indication of a slow change in the tertiary structure of the native protein that can eventually lead to unfolding and certainly contributes to the decrease in activity and stability that has been observed experimentally.[2]

(47) Závodszky, P.; Kardos, J.; Svingor, A.; Petsko, G. A. *Proc. Natl. Acad. Sci.* **1998**, *95*, 7406−7411.

**Table 3.** Calculated Average Distances

| distance (Å) | SRT | SHT | distance (Å) | HRT | HHT |
|---|---|---|---|---|---|
| 218@OD1−204@NH | 1.95 | 1.99 | 218@OG-204@NH | 2.02 | 2.03 |
| 217@O-205@NH | 2.16 | 2.18 | 217@O-205@NH | 2.18 | 2.20 |
| 219@NH-202@O | 2.00 | 2.04 | 219@NH-202@O | 1.93 | 1.97 |

On the other hand, the analysis of the $R_{gyr}$ behavior of the thermostable mutant is again indicative of the capability of the engineered protein to maintain its tertiary structure even at high temperature.

**Hydrogen-Bonding Interactions.** The criteria used for the determination of the presence or absence of a hydrogen bond was purely geometric.[37] All polar hydrogens were taken into account, and distances to potential hydrogen-bond acceptors were calculated. A distance cutoff of 2.6 Å between the donor and the acceptor and an angle cutoff between 120° and 180° at the hydrogen atom were applied. These criteria are the same as those applied in previous publications by Hartsough and Merz[48] and by Toba and Merz.[37] We have also separated the strong (present for more than 90% of time) from the weak hydrogen bonds (present for <10% of time).

During the simulations the mutant protein always had more stable hydrogen bonds than did the native protein (see Table 1). Moreover, we noticed that almost all of the mutations (e.g., Asn 188 Ser, Gly 166 Arg, and Ser 194 Pro) that increase the hydrogen-bonding interactions were located in the areas of lower flexibility of the 5-3H5 system. Thus, from a hydrogen-bonding perspective our results are consistent with the observation of greater activity and thermostability of the mutant relative to the native enzyme.[2]

The substitution of Asn 218 with Ser has been shown to increase thermal stability in several subtilisins.[2] Asn (or Ser) 218 are components of an antiparallel $\beta$-sheet formed by residues 202−219[6] where residues 218 and Ser 204 are directly opposite one another. In the wild-type enzyme, the side chain carbonyl oxygen of Asn 218 is positioned above the plane of the antiparallel $\beta$-sheet, where it accepts a hydrogen bond from the backbone nitrogen of Ser 204. In the mutated enzyme the Ser 218 side chain hydroxyl group accepts the hydrogen bond from the backbone nitrogen of Ser 204. In this way the overall hydrogen-bonding pattern is maintained in the mutant system; however, the steric signature of the side chain of Ser 218 is smaller than that of the wild-type Asn 218 side chain. Thus, it can fit more readily into the plane of the backbone nitrogen of Ser 204 than can the side chain carbonyl oxygen of Asn 218, thereby giving a more energetically favorable (and stabilizing) interaction. Moreover, the presence of a "smaller" side chain at position 218 permits closer packing of the two $\beta$-strands. Table 3 reports the hydrogen-bonding distances on going from Asn 218 to Ser218.

The substitution of Asn 181 with Asp also contributes toward improving the hydrogen-bonding pattern of the molecule and its interactions with solvent. Asp 181 is located on a loop and is in direct contact with solvent, and clearly, the carboxylate group of Asp 181 is a better H-bond acceptor than the amido group of Asn. Furthermore, Asp 181 is always involved in hydrogen-bonding interactions with the hydroxyl group and the peptide N atom of Ser 183 in the HRT and HHT cases. These interactions were observed to be either direct or mediated by 1−3 water molecules. The side chain of Asp 181 also occasionally hydrogen bonds with the OH group of Ser 182.

Another mutation of note is Asn 76 to Asp. In the thermostable enzyme the side chain group of Asp 76 can accept a

hydrogen bond from Gln 2. Moreover, the carbonyl oxygen of Gln 2 is involved in calcium binding. Once the hydrogen bond to Asp 76 is formed, the conformation of the Gln 2 side chain is locked in a way that favors calcium binding (i.e., preorganization).[49] A higher affinity for calcium ions can be important in improving stability; moreover, this interaction could further stabilize the N-terminus and, thereby, prevent unraveling.

**Solvent Accessible Surface Area.** Solvent accessible surface area (SASA) can be considered to be an indicator of how the different parts of a protein can interact with and be affected by the medium under different conditions. For our analysis we used the algorithm developed by LeGrand and Merz, which uses a probe of 1.4 Å.[50] The SASA values are reported in Table 1. The total SASA for native subtilisin is higher than the mutant one both in the low and in the high-temperature simulations. This is indicative (as shown also by the $R_{gyr}$ calculation) of a more compact structure for the mutated system. Arnold et al. have shown that thermal stability of proteins can be a consequence of reduced surface area and increased packing density of the internal core, which is consistent with our observations.[51]

We also studied how the exposure of hydrophobic and hydrophilic residues changes with temperature. The biggest difference (2536 Å$^2$ versus 2631 Å$^2$) is in the exposure of the apolar residues of subtilisin E at 300 and 350 K. This is indicative of an increase in solvent penetration into the internal core of the protein. This is also considered an important event at the beginning of the unfolding process and is another factor that helps in explaining the reduced activity and stability of the wild-type enzyme at higher temperature.[42−44]

It's interesting to consider the solvent exposure of residues close to mutation sites or those in proximity to active site and substrate-binding pockets. As we have seen in the previous section, the substitution between Asn 218 and Ser 218 can have important consequences. Tyr 217 is directly H-bound to this amino acid, and Tyr 218 and Asp 60 have been shown to affect the burial of His 64 in the PC3 mutant of subtilisin E in DMF.[37] In the present case solvent exposure of Tyr 217 remains constant in the HRT and HHT simulations, while there is a large variation in the SASA for this residue between the SRT and SHT simulations, which arises from Tyr 217 extending into solvent at high temperature. This movement of residue 217 toward the external region of the protein controls the opening of a pocket, near the active site, which controls penetration of solvent into the core of the protein. The fact that no significant SASA change is observed for Tyr 217 in the case of the two mutant protein simulations can be considered a further consequence of how Ser 218 stabilizes the 202−219 $\beta$-sheet region. Thus, the active site region is not influenced by steric changes in its proximity or by solvent penetration into undesired regions, so that reactivity and stability can be retained at the higher temperature.

Another interesting variation in the SASA involves Glu 156. The SASA values for this amino acid are very different between subtilisin E and 5-3H5. In the mutant enzyme this residue can interact through a one- or two-water molecule bridge with Arg 166 (Gly 166 Arg). Moreover, because of this interaction, Glu 156 is kept in place at both 300 and 350 K and its side chain points toward the interior of the protein. In both of the mutant simulations the side chain of Glu 156 interacts, directly or through a two-water molecule bridge, with the amide hydrogens of the side chain of Asn 155. This residue is very important in terms of reactivity because it determines the formation and the

(48) Hartsough, D. S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 10113−10116.

(49) Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 1039−1057.
(50) LeGrand, S. M.; Merz, K. M., Jr. *J. Comput. Chem.* **1993**, *14*, 349−352.
(51) Arnold, F. H. *Curr. Opin. Biotechnol.* **1993**, *4*, 450−455.

**Table 4.**   Calculated Average Distances

| | hydrogen-bonding distances between active site residues Ser-His-Asp (Å) | | | |
|---|---|---|---|---|
| | SRT | SHT | HRT | HHT |
| Ser221 HG-His64 NE2 | 3.66 | 3.76 | 3.58 | 3.61 |
| His64 HD1-Asp32 OD1 | 1.90 | 1.94 | 1.92 | 1.94 |
| His64 HD1-Asp32 OD2 | 2.69 | 2.64 | 2.69 | 2.76 |

position of the oxyanion hole. The presence of the hydrogen bond between one of the amide hydrogens of Asn 155 and Glu 156 automatically "preorganizes"[49] the position of the second amide hydrogen and, thereby, facilitates catalysis. In the wild-type enzyme residue 166 is a Gly, so that this network of hydrogen bonds cannot be formed. This is obviously reflected in the SASA value of Glu 156, which prefers to have its side chain more exposed to solvent rather than in the internal region of the protein as in the mutant case.

**Active Site Analysis.** The active site of subtilisin is formed by the classic catalytic triad of Ser 221, His 64, and Asp 32. We first ran calculations to evaluate average distances for the hydrogen bonds among the three residues. It had been shown in a previous paper by our group[37] and by X-ray crystallographic studies by Matthews et al.[52] that there is a hydrogen bond between Asp 32 and His 64, whereas one is not present between Ser 221 and His 64. Our distance analysis confirms these findings for all four simulations (see Table 4). From these data we can see that the Ser 221 HG- His 64 NE2 distance is shorter in the case of HHT than in the case of SHT. This might affect how efficiently the hydroxyl hydrogen of Ser 221 is transferred to His 64 in the first step of the hydrolytic reaction catalyzed by subtilisin, but we cannot quantitate this effect.

We also calculated the rms deviations per residue for the catalytic triad and for Asn 155, which, we believe, has an influence on activity (see Table 2). Our calculations show that while the rms deviation values for His and Asp are very close to one another during the four simulations, the values for Ser 221 indicate significant differences. In the case of SRT, HRT, and HHT, the deviation from the crystal structure is 1.14−1.23 Å, while in the case of SHT this value increases to 1.44 Å. Ser 221 is linked, through residue 220, to the $\beta$-sheet comprised of residues 202−219. This is stabilized, in the mutant system, by the Asn 218 to Ser substitution, so that the whole region can maintain its three-dimensional structure even at 350 K. The hydrogen-bonding network present (Asn 218 instead of Ser) in the native protein apparently does not allow the favorable orientation to be preserved at high temperature. Recall that, in 5-3H5, Glu 156 and Asn 155 are part of a complicated hydrogen-bonding system that involves the presence of Arg 166. The whole network brings the Asn 155 side chain in a very favorable position to stabilize the reaction intermediate in the oxyanion hole. This does not happen in the case of subtilisin E where residue 166 is a Gly. The net effect is reflected in the rms deviation values for Asn 155. In the HRT and HHT cases the values are lower and the conformation more stable, as opposed to the higher values calculated for the wild-type case.

We were also interested in the effect mutations can have on pockets which are close to the active site. The substitution of Gly 166 with Arg influences substrate specificity and binding and has been shown to have an effect on $K_M$.[2] Residue 166 is located at the bottom of the S1 pocket into which a side chain of a substrate would be placed. The substitution of a residue with no side chain (Gly) with one with a very large (and

(52) Matthews, D. A.; Alden, R. A.; Birktoff, J. J.; Freer, S. T.; Kraut, J. *J. Biol. Chem.* **1977**, *252*, 8875−8883.
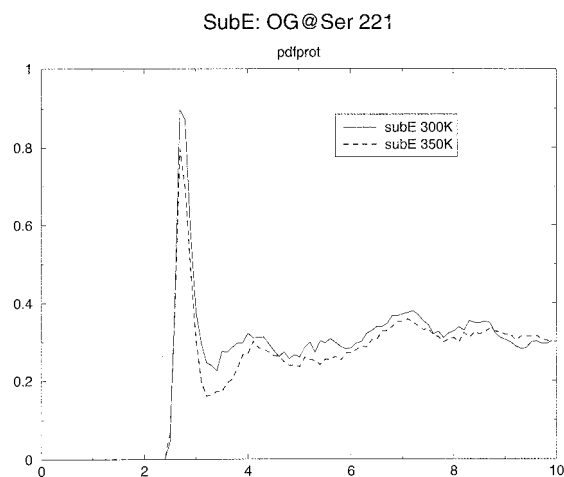


**Figure 11.**   Plot of the pair distribution function for OG@Ser 221 to the oxygens in the surrounding water molecules for subtilisin E.
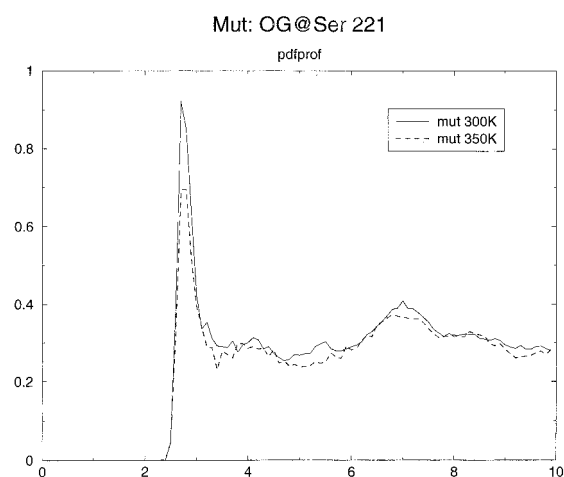


**Figure 12.**   Plot of the pair distribution function for OG@Ser 221 to the oxygens in the surrounding water molecules for 5-3H5.

charged) side chain (Arg) will certainly have a dramatic effect on substrate recognition and release. Furthermore, due to the presence of Asn 218 in the wild-type enzyme, Tyr 217 tends to be exposed to solvent, thereby opening a pocket, at high temperatures, which leads to solvent penetration near His 64 in the active site.

Pair distribution functions (pdfs) allow us to obtain insights into the extent of water penetration and solvent ordering around the atoms of interest in the active site. These atoms were the following:  OG@Ser 221 (Figures 11 and 12 and ND1@His 64 (Figures 13 and 14). Our pdfs for OG@Ser 221 and the oxygens of the surrounding water molecules show a peak for the first solvation shell at 2.7 Å and another broader peak at 6.9 Å for the HRT and HHT systems (see Figure 12). In the case of native subtilisin E, we also observed a peak at 4.1 Å from OG of Ser 221 (see Figure 11). The pdfs for ND1@His 64 in the HRT and HHT (see Figure 14) simulations show a first peak at 3.3 Å where the intensity of the peak for the 350 K system is slightly higher than that at 300 K. A second peak at 4.8 Å and a third at 5.7 Å are also present, but the intensity of the 300 K peaks is greater than that at 350 K. The situation is reversed when we consider the pdfs for ND1@His 64 for subtilisin E. In this case solvent penetrates into the interior of the protein at 350 K due to the repositioning of Tyr 217's side chain. This is evidenced by the greater intensity of the third peak (at 5.7 Å) in the SHT simulation. Thus, we see that solvent can penetrate into the pocket opened by Tyr 217, thereby
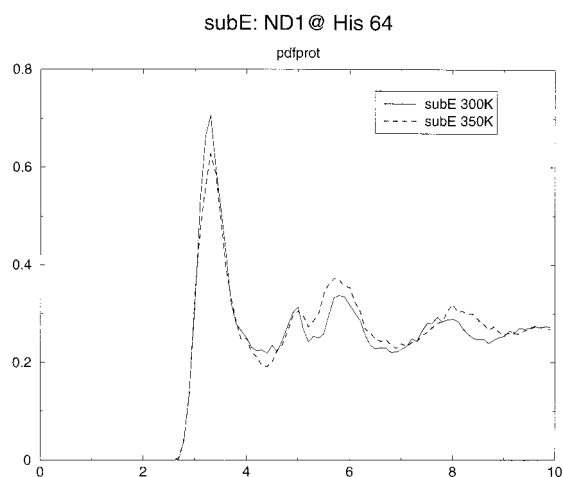
**Figure 13.** Plot of the pair distribution function for ND1@His 221 to the oxygens in the surrounding water molecules for subtilisin E.
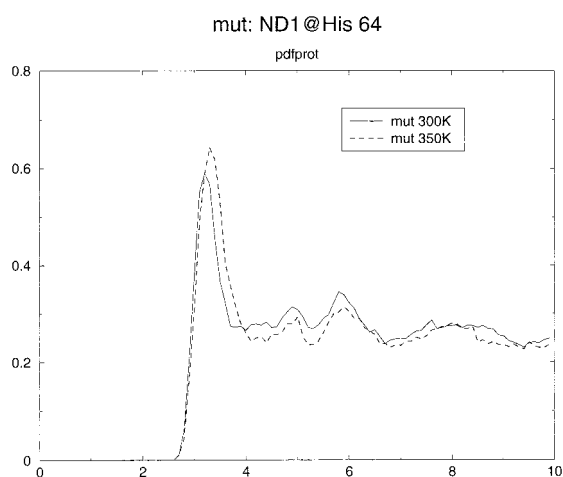


**Figure 14.** Plot of the pair distribution function for ND1@His 221 to the oxygens in the surrounding water molecules for 5-3H5.

destabilizing the protein through unfavorable contacts with hydrophobic interior residues.

## Conclusions

A protein in water has to fulfill several requirements for catalytic activity and stability. First of all, it must fold into its proper active structure that does not alter significantly with time under the reaction conditions (e.g., presence of organic solvents, ionic strength of the solution, temperature, etc.). The enzyme must not be too rigid such that it becomes difficult to bind or to release substrate or product molecules. Moreover, the active site region must maintain a reactive conformation which is not disrupted by loop movements, partial unfolding, or solvent penetration.[53] From the results of our simulations, we find that 5-3H5 has been modified, with respect to native subtilisin E, in such a way that these key characteristics are maintained at higher temperatures. On the other hand, the wild-type enzyme undergoes significant loop movements, reductions in the numbers of hydrogen bonds, etc. that could eventually lead to unfolding.

In particular, our simulations showed how the substitution of uncharged amino acids with charged ones (Asn 76 Asp, Asn

(53) Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250−5254.

181 Asp, Gly 166 Arg) improved the stability of the protein through a better series of hydrogen-bonding interactions. The substitution of the Gly 166 side chain with the much larger Arg side chain has significant steric consequences as well. Position 166 is in close proximity to the active site (so-called S1 pocket), and the presence of Arg's larger side chain could have an influence on substrate binding and release through a reduction in cavity volume. Autolysis could also be slowed by the presence of this larger side chain (i.e., the autolysis "substrate specificity" characteristics have been altered).

We have also demonstrated the importance of the Asn 218 Ser substitution, which has been discussed previously.[2] The presence of Ser 118 in close proximity to an α-helix can afford a stabilizing influence similar to that of a helix-cap. The substitution of Pro 14 with Leu in helix-B increases both the helical propensity and the conformational entropy of the region, because of the rotational freedom of the Leu side chain relative to Pro. The thermostable mutant protein also has a higher flexibility than the wild-type system at 350 K. This increases the entropy content of the folded state and could be an important factor that stabilizes the folded state relative to the unfolding transition state and the unfolded protein. The higher rigidity of subtilisin E does not bring any thermodynamic advantage and could be a factor in the lower stability and activity at high temperature. The $R_{gyr}$ calculations showed how 5-3H5 maintains its global folded three-dimensional structure over all simulation conditions. The increase of $R_{gyr}$ for subtilisin E on going from 300 to 350 K is associated with the loop movements described above that force the system to partially lose its globular and, hence, reactive three-dimensional structure.

We have examined each single mutation in order to try to define rational design rules for proteins, which can be simply implemented. As shown in this paper we were successful in several instances but it is still very difficult to make individual predictions. Nonetheless, a design strategy that could be employed is evidenced by our work. A key observation is the loss of conformational mobility (as evidenced by rmsd) of several loop regions in the mutant subtilisin. We hypothesize that these regions may serve as unfolding initiation sites or they may be regions where autolysis could occur. These regions can be readily identified from MD simulations or X-ray structures[46] of the native protein and, once identified, experimental (i.e., directed evolution) efforts could be focused on stabilizing these regions. Alternatively, all loop regions could be examined experimentally to identify key regions that need to be stabilized. However, it would be interesting to see if a combined theoretical/experimental design process, as envisioned above, could be successful.