Energetics of Folding Subtilisin BPN′ †

Philip Bryan,* Patrick Alexander, Susan Strausberg, Fredrick Schwarz, Wang Lan, Gary Gilliland, and D. Travis Gallagher

Maryland Biotechnology Institute, Center for Advanced Research in Biotechnology, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received January 27, 1992; Revised Manuscript Received April 1, 1992

ABSTRACT: Subtilisin is an unusual example of a monomeric protein with a substantial kinetic barrier to folding and unfolding. Here we document for the first time the in vitro folding of the mature form of subtilisin. Subtilisin was modified by site-directed mutagenesis to be proteolytically inactive, allowing the impediments to folding to be systematically examined. First, the thermodynamics and kinetics of calcium binding to the high-affinity calcium A-site have been measured by microcalorimetry and fluorescence spectroscopy. Binding is an enthalpically driven process with an association constant (K_a) equal to 7×10^6 M^{-1}. Furthermore, the kinetic barrier to calcium removal from the A-site (23 kcal/mol) is substantially larger than the standard free energy of binding (9.3 kcal/mol). The kinetics of calcium dissociation from subtilisin (e.g., in excess EDTA) are accordingly very slow ($t_{1/2}$ = 1.3 h at 25 °C). Second, to measure the kinetics of subtilisin folding independent of calcium binding, the high-affinity calcium binding site was deleted from the protein. At low ionic strength (I = 0.01) refolding of this mutant requires several days. The folding rate is accelerated almost 100-fold by a 10-fold increase in ionic strength, indicating that part of the free energy of activation may be electrostatic. At relatively high ionic strength (I = 0.5) refolding of the mutant subtilisin is complete in less than 1 h at 25 °C. We suggest that part of the electrostatic contribution to the activation free energy for folding subtilisin is related to the highly charged region of the protein comprising the weak ion binding site (site B). The coalescing of the charged amino acids to form this ion binding site may present an electrostatic barrier to folding unless the charges can be shielded as they are brought into proximity during folding. A high kinetic barrier between the folded and unfolded states may be an essential characteristic of extracellular, microbial proteases to prevent autodegradation. Since this kinetic barrier is also an impediment to folding, the pro sequence may function in vivo to catalyze the folding reaction.

Subtilisin BPN′ is a 275 amino acid serine protease secreted by *Bacillus amyloliquefaciens*. It is an enzyme of considerable industrial importance and has been the subject of numerous protein engineering studies (Siezen et al., 1991; Bryan, 1992; Wells & Estell, 1988). We have previously used several strategies to increase the stability of subtilisin to thermal denaturation by assuming simple thermodynamic models to approximate the unfolding transition (Pantoliano et al., 1987, 1988, 1989; Rollence et al., 1988). One complication to this approach is the fact that subtilisin is initially secreted as a proenzyme and the presence of the 77 amino acid pro sequence is important for in vivo production of mature subtilisin (Wells et al., 1983; Vasantha et al., 1984; Ikemura et al., 1987; Zhu et al., 1989). In fact, in vitro refolding of mature subtilisin has not been demonstrated previously, even though earlier calorimetric data could be fit with a two-state model, indicating that the unfolding reaction is fundamentally reversible (Bryan et al., 1986; Pantoliano et al., 1989).

Since refolding of mature subtilisin is a very inefficient process under many conditions, we have tried to dissect the impediments to the refolding of mature subtilisin to answer the following questions: Is the mature folded form of subtilisin, in fact, in a global free energy minimum at neutral pH and 25 °C? If so, what are the energetic barriers that slow refolding?

The subtilisin BPN′ structure has been highly refined (R = 0.14) to 1.3-Å resolution and has revealed structural details for two ion binding sites (Finzel et al., 1986; Pantoliano et al., 1988; McPhalen & James, 1988). One of these (site A) binds Ca^{2+} with high affinity and is located near the N-terminus, while the other (site B) binds calcium and other cations much more weakly and is located about 32 Å away (Figure 1). Structural evidence for two calcium binding sites was also reported by Bode et al. (1987) for the homologous enzyme, subtilisin Carlsberg.

The unfolding reaction of subtilisin can be divided into two parts as follows:

$$N(Ca) \rightleftharpoons N \rightleftharpoons U$$

where N(Ca) is the native form of subtilisin with calcium bound to the high-affinity calcium binding site A (Finzel et al., 1986; Pantoliano et al., 1988; McPhalen & James, 1988), N is the folded protein without calcium bound, and U is the unfolded protein. Subtilisin is a relatively stable protein whose stability is in large part mediated by the high-affinity calcium site (Voordouw et al., 1976; Pantoliano et al., 1988). The melting temperature of subtilisin at pH 8.0 in the presence of micromolar concentrations of calcium is ~75 °C and ~56 °C in the presence of excess EDTA[1] (Takehashi & Sturtevant, 1981; Bryan et al., 1986a).

We have constructed a series of mutant subtilisins which have allowed the two parts of the folding process to be exam-

† This work was supported by NIH Grant GM42560.

* Corresponding author.

[1] Abbreviations: C_{ex}, excess specific heat as measured by calorimetry; C_{max}, maximum excess specific heat; ΔH_{cal}, calorimetric enthalpy for unfolding; ΔH_{vH}, van't Hoff enthalpy for unfolding; DSC, differential scanning calorimetry: EDTA, disodium salt of ethylenediaminetetraacetic acid; Gu-HCl, guanidine hydrochloride; I, ionic strength; [P], protein concentration, rms, root mean square; sAAPFna, succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide; Tris, tris(hydroxymethyl)aminomethane; $t_{1/2}$, half-life for a kinetic experiment; wt, wild type.

0006-2960/92/0431-4937$03.00/0

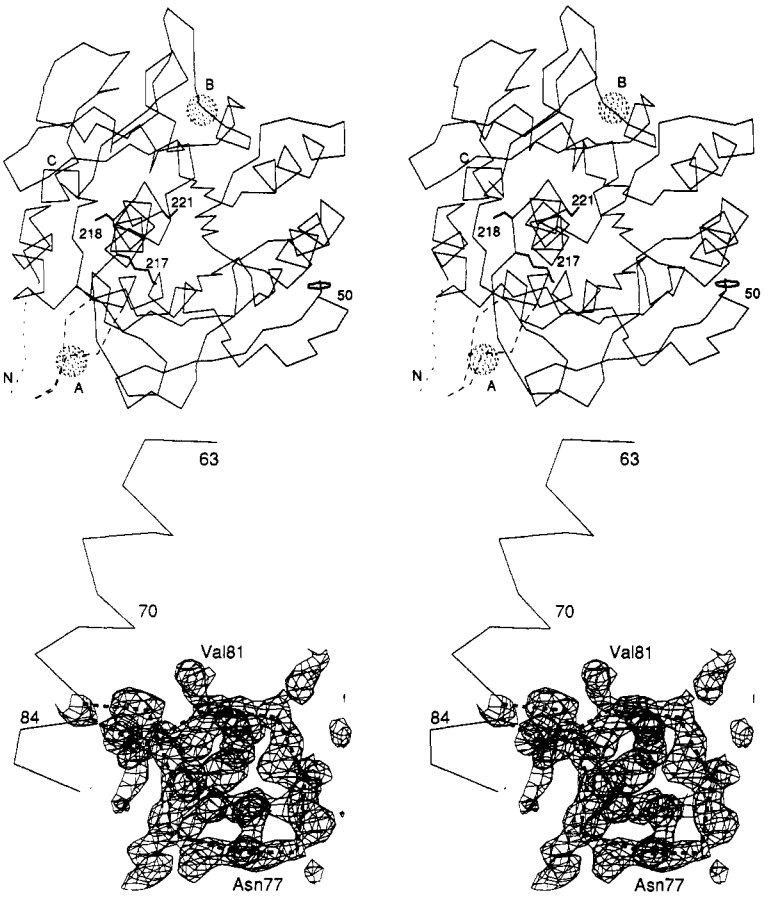

FIGURE 1: X-ray crystal structure of S15 subtilisin. (A, top) α-Carbon plot showing the positions of mutations as noted. The numbering of wild-type subtilisin is kept. Dotted spheres show the position of calcium at the weak ion binding site (B-site) and the former position of the A-site. The A-site loop (dashed line) is absent in this mutant. N- and C-termini are indicated. The N-terminus is disordered (dotted line). (B, bottom) Close-up view of the A-site deletion. The loop from S12 subtilisin is shown as a dotted line with the continuous helix of S15. Superimposed is the 3σ difference electron density (FO12–FO15, phases from S15) showing the deleted A-site loop.

ined independently. First, it was necessary to eliminate proteolytic activity, in order to prevent autodegradation from occurring during the unfolding and refolding reactions. This was accomplished by converting the active-site serine 221 to cysteine.[2] This mutation has little effect on the thermal denaturation temperature of subtilisin but reduces peptidase activity by a factor of $\sim 3 \times 10^4$ (Abrahmsen et al., 1991). This mutant has become a vehicle through which the folding questions can be addressed without complications of proteolysis. The S221C mutant is designated S1.[3]

Second, previous calorimetric studies of the calcium-free (apoenzyme) form of subtilisin indicated that it is of marginal stability at 25 °C with a ΔG of unfolding of <5 kcal/mol (Pantoliano et al., 1989). Because calcium is such an integral part of the subtilisin structure, the apoenzyme might be thought of as a folding intermediate. Therefore, to make the relatively unstable apoenzyme easier to produce and purify, we made versions of S1 with three or four mutations (M50F, Q206I, Y217K, and N218S) which cumulatively increase the free energy of unfolding by 3.0 kcal/mol and increase the thermal denaturation temperature of the apoenzyme by 11.5 °C (Pantoliano et al., 1989). The mutant containing the M50F, Q206I, Y217K, N218S, and S221C mutations is designated S-11, and the mutant containing the M50F, Y217K, N218S, and S221C mutations is designated S12.[4]

Third, to study refolding of subtilisin independent of calcium binding, we have made a deletion to remove any potential for calcium binding to site A. Four of the six ligands to the calcium are provided by a loop composed of amino acids 75–83. This loop is an interruption in an α-helix involving amino acids 63–85 (McPhalen & James, 1988). We have used site-directed mutagenesis to delete amino acids 75–83, which creates an uninterrupted helix and abolishes the calcium binding potential at this site (Figure 1). This mutant (S15)[4] offers two major experimental advantages for characterizing the N $\rightleftharpoons$ U reaction. First, the removal of all calcium from S12 subtilisin requires several days of dialysis at 25 °C. Thus this time-consuming step is eliminated. Second, keeping calcium from rebinding to subtilisin is difficult without including metal chelating agents. The presence of the chelator EDTA is undesirable, however, since it binds to the protein and seems to affect the refolding reaction.

Materials and Methods

Cloning and Expression. The subtilisin gene from *B. amyloliquefaciens* (subtilisin BPN′) has been cloned, sequenced, and expressed at high levels from its natural promoter sequences in *Bacillus subtilis* (Wells et al., 1983; Vasantha

[2] The S221A mutant was originally constructed for this purpose. The mature form of this mutant was heterogeneous on its N-terminus, however, presumably due to some incorrect processing of the proenzyme.

[3] A shorthand for denoting amino acid substitutions employs the single-letter amino acid code as follows: Y217K denotes the change of Tyr 217 to Lys, and the variant enzyme containing this single change is denoted subtilisin Y217K.

[4] The specific activities of S11, S12, and S15 against the synthetic substrate, sAAPFna, are the same (SA = 0.0025 unit/mg at 25 °C, pH 8.0). These measurements were performed on protein freshly purified on a mercury affinity column.

et al., 1984). All mutant genes were recloned into a pUB110-based expression plasmid and used to transform *B. subtilis*. The *B. subtilis* strain used as the host contains a chromosomal deletion of its subtilisin gene and therefore produces no background wild-type (wt) activity (Fahnestock & Fisher, 1987). Oligonucleotide mutagenesis was carried out as previously described (Zoller & Smith, 1983; Bryan et al., 1986b). S221C was expressed in a 1.5 l New Brunswick fermentor at a level of ~100 mg of the correctly processed mature form per liter. The addition of wild-type subtilisin to promote production of the mature form of S221C subtilisin was not required in our *Bacillus* host strain as was the case for Abrahmsen et al. (1991).

Protein Purification and Characterization. Wild-type subtilisin BPN′ and the variant enzymes were purified and verified for homogeneity essentially as described (Bryan et al., 1986b; Pantoliano et al., 1987, 1988). In some cases the C221 mutant subtilisins were repurified on a sulfhydryl-specific mercury affinity column (Affi-Gel 501, Bio-Rad). Assays of peptidase activity were performed by monitoring the hydrolysis of sAAPFna as described by DelMar et al. (1979). The [P] was determined using $P^{0.1\%} = 1.17$ at 280 nm (Pantoliano et al., 1989). For variants which contain the Y217K change, the $P^{0.1\%}$ at 280 nm was calculated to be 1.15 (or 0.96 × wt), based on the loss of one Tyr residue (Pantoliano et al., 1989).

N-Terminal Analysis. The first five amino acids of subtilisin S15 were determined by sequential Edman degradation and HPLC analysis. This revealed that 100% of the material had the amino acid sequence expected from the DNA sequence of the gene and that processing of the propeptide was at the same position as for the wild-type enzyme.

Preparation of Aposubtilisin. S11 and S12 subtilisins contain an equal molar amount of tightly bound calcium after purification. X-ray crystallography has shown this calcium to be bound to the A-site (Finzel et al., 1986; Pantoliano et al., 1988; McPhalen & James, 1988). Complete removal of calcium from subtilisin is very slow, requiring 24 h of dialysis against EDTA at 25 °C to remove all calcium from the protein and then an additional 48 h of dialysis in high salt (Brown et al., 1977) at 4 °C to remove all EDTA from the protein. To prepare the calcium-free form of subtilisins S11 and S12, 20 mg of lyophilized protein was dissolved in 5 mL of 10 mM EDTA and 10 mM Tris-HCl, pH 7.5, and dialyzed against the same buffer for 24 h at 25 °C. In order to remove EDTA, which binds to subtilisin at low ionic strength, the protein was then dialyzed against 2× 2 L of 0.9 M NaCl and 10 mM Tris-HCl, pH 7.5, at 4 °C for a total of 24 h and then against 3× 2 L of 2.5 mM Tris-HCl, pH 7.5, at 4 °C for a total of 24 h. Chelex 100 was added to all buffers not containing EDTA. When versions of C221 subtilisin not containing stabilizing amino acid substitutions were used, up to 50% of the protein precipitated during this procedure. It is essential to use pure native apoenzyme in titration experiments so that spurious heat produced by precipitation upon the addition of calcium does not interfere with the measurement of the heat of binding.

To ensure that preparations of aposubtilisin were not contaminated with calcium or EDTA, samples were checked by titration with calcium in the presence of Quin2 prior to performing titration calorimetry.

Titration Calorimetry Measurements. The calorimetric titrations were performed with a Microcal Omega titration calorimeter as described in detail by Wiseman et al. (1989). The titration calorimeter consists of a matched reference cell containing the buffer and a solution cell (1.374 mL) containing the protein solution. Microliter aliquots of the ligand solution are added to the solution cell through a rotating stirrer syringe operated with a plunger driven by a stepping motor. After a stable baseline was achieved at a given temperature, the automated injections were initiated and the accompanying heat change per injection was determined by a thermocouple sensor between the cells. During each injection, a sharp exothermic peak appeared which returned to the baseline prior to the next injection occurring 4 min later. The area of each peak represents the amount of heat accompanying binding of the added ligand to the protein. The total heat, Q, was then fit by a nonlinear least squares minimization method (Wiseman et al., 1989) to the total ligand concentration, $[Ca]_{total}$, according to the equation:

$$dQ/d[Ca]_{total} = \Delta H[1/2 + (1 - (1 + r)/2 - X_r/2)/(X_r - 2X_r(1 - r) + 1 + r^2)^{1/2}] \quad (1)$$

where $1/r = [P]_{total} \times K_a$ and $X_r = [Ca]_{total}/[P]_{total}$. The protein concentrations ranged from 30 to 100 μM while the concentration of the calcium solutions were about 20 times the protein concentrations. Each binding constant and enthalpy were based on several titration runs at different concentrations. Titration runs were performed until the titration peaks were close to the baseline.

In Vitro Refolding of S15 Subtilisin. For refolding studies subtilisin was maintained as a stock solution in 2.5 mM Tris-HCl, pH 7.5, and 50 mM KCl at a concentration of ~100 μM. The protein was denatured by diluting the stock solution into 5 M Gu-HCl, pH 7.5, or, in most cases, 25 mM H_3PO_4 or HCl, pH 1.8–2.0. The final concentration of protein was 0.1–5 μM. S15 was completely denatured in less than 30 s by these conditions. S12 required ~60 min to become fully denatured. Acid-denatured protein was then neutralized to pH 7.5 by the addition of Tris base (if denatured in HCl) or 5 M NaOH (if denatured in H_3PO_4). Refolding was initiated by the addition of KCl, NaCl, or $CaCl_2$ to the desired concentration. For example, KCl was added from a stock solution of 4 M to a final concentration of 0.1–1.5 M with rapid stirring. In most cases renaturation was carried out at 25 °C. The rate of renaturation was determined spectrophotometrically by UV absorption from the increase in extinction at λ = 286 and from the increase in intrinsic tyrosine and tyrptophan fluorescence (excitation λ = 282, emission λ = 347) or by CD from the increase in negative ellipticity at λ = 222 nm.

X-ray Crystallography. Large single-crystal growth and X-ray diffraction data collection were performed essentially as previously reported (Bryan et al., 1986a; Pantoliano et al., 1988, 1989) except that it was not necessary to inactivate the S221C variants with diisopropyl fluorophosphate (DFP) in order to obtain suitable crystals. The starting model for S12 was made from the hyperstable subtilisin BPN′ mutant 8350 (Protein Data Bank entry 1SO1.pdb). The S12 structure was refined and then modified to provide the starting model for S15.

Data sets with about 20 000 reflections between 8.0- and 1.8-Å resolution were used to refine both models using restrained least squares techniques (Hendrickson & Konnert, 1980). Initial difference maps for S15, phased by a version of S12 with the entire site A region omitted, clearly showed continuous density representing the uninterrupted helix, permitting an initial S15 model to be constructed and refinement begun. Each mutant was refined from $R \sim 0.30$ to $R \sim 0.18$ in about 80 cycles, interspersed with calculations of electron density maps and manual adjustments using the graphics modeling program FRODO (Jones, 1978).

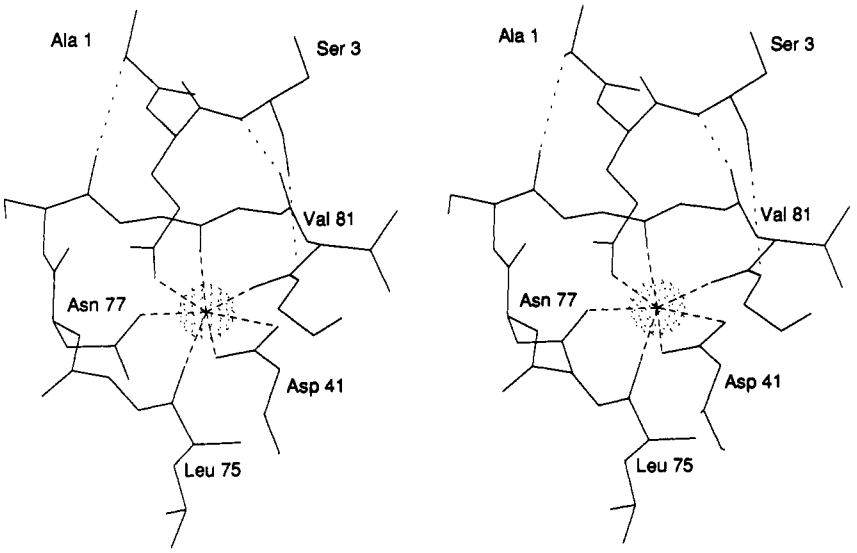

FIGURE 2: X-ray crystal structure of the calcium A-site region of S12 subtilisin. Calcium is shown as a dotted sphere with one-half the van der Waals radius. Dashed lines are coordination bonds, while dotted lines represent hydrogen bonds under 3.2 Å.

RESULTS

Comparison of S12 and S15 Subtilisins

X-ray Crystallography. Except for the region of the deleted calcium binding loop, the structures of S12 and S15 are very similar, with a rms deviation of 0.18 Å between 262 α-carbons. The N-terminus of S12 (as in the wild type) lies beside the site A loop, furnishing one calcium coordination ligand, the side-chain oxygen of Q2. In S15 the loop is gone, leaving residues 1–4 disordered. In S12 (as in the wild type) the site A loop occurs as an interruption in the last turn of a 14-residue α-helix; in S15 this helix is uninterrupted and shows normal helical geometry over its entire length. Diffuse difference density and higher temperature factors indicate some disorder in the newly exposed residues adjacent to the deletion. A thorough description and comparison of the refined structures are in preparation.

Structure of the Calcium A Site of S12 Subtilisin. Calcium at site A is coordinated by five carbonyl oxygen ligands and one aspartic acid. Four of the carbonyl oxygen ligands to the calcium are provided by a loop composed of amino acids 75–83 (Figure 2). The geometry of the ligands is that of a pentagonal bipyramid whose axis runs through the carbonyls of 75 and 79. On one side of the loop is the bidentate carboxylate (D41), while on the other side is the N-terminus of the protein and the side chain of Q2. The seven coordination distances range from 2.3 to 2.6 Å, the shortest being to the aspartyl carboxylate. Three hydrogen bonds link the N-terminal segment to loop residues 78–82 in parallel-β arrangement.

Differential Scanning Calorimetry. The stability properties of S12 and S15 have been studied using DSC. The Δ75–83 mutant (S15) is very similar in melting temperature to the apoenzyme of S12. The DSC profiles of apo-S12 and S15 are shown in Figure 3. The temperature of maximum heat capacity is 63.0 °C for S15 and 63.5 °C for apo-S12 at pH 9.63. The DSC experiments were carried out at high pH to avoid aggregation during the denaturation process. The amount of excess heat absorbed by a protein sample as the temperature is increased through a transition from the folded to unfolded state at constant pressure provides a direct measurement of the ΔH of unfolding (Privalov & Potekhin, 1986). ΔH_{cal} of unfolding for apo-S12 and S15 is about 140 kcal/mol. Above pH 10.0, the unfolding transition for S15 fits a two-state model reasonably well, consistent with equilibrium thermodynamics

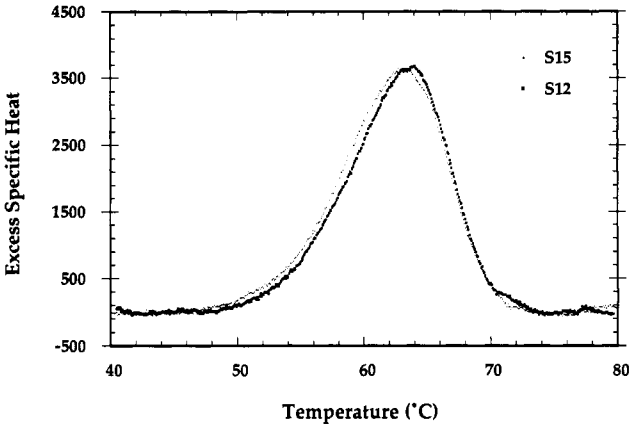

FIGURE 3: Differential scanning calorimetry. The calorimetric scans apo-S12 (T_m = 63.5 °C) and S15 (T_m = 63.0 °C) are shown. Measurements were performed with a Hart 7707 DSC heat conduction scanning microcalorimeter as described (Pantoliano et al., 1989). Sample conditions were 50 mM glycine, pH 9.63. The scan rate was 0.5 °C/min. Excess heat capacity is in units of microjoules per degree. The calorimeter ampules contained 1.78 mg of protein.

as expressed in the van't Hoff equation, $d \ln K/dT = \Delta H_{vH}/(RT^2)$, with ΔH_{vH}, the van't Hoff enthalpy, or apparent enthalpy, approximately equal to the calorimetric, ΔH_{cal}, or true enthalpy. At pH 9.63, however, the melting profile for both proteins is asymmetric, indicating that the unfolding is not a pure two-state process. A more complete analysis of these proteins by DSC is in progress.

Characterization of Calcium Binding to the A-Site

Titration Calorimetry. The binding of calcium to the S11 and S12 subtilisins has been measured by titration calorimetry, allowing both the binding constant and the enthalpy of binding to be determined (Wiseman et al., 1989; Schwarz et al., 1991).

The S11 and S12 subtilisin mutants were used in titration experiments, as explained under Materials and Methods, because production of the wild-type apoenzyme is impossible due to its proteolytic activity and low stability. Titrations of S11 and S12 were performed at protein concentrations [P] = 30 and 100 μM. Titration of the S11 apoenzyme with calcium at 25 °C is shown in Figure 4. The data points correspond to the negative heat of calcium binding associated with each titration of calcium. The titration calorimeter is sensitive to changes in K_a under conditions at which the product of K_a ×

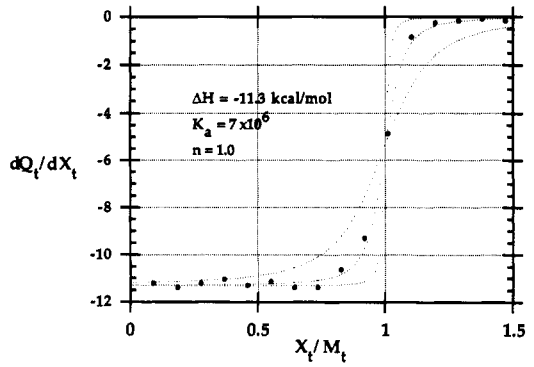

FIGURE 4: Titration calorimetry of subtilisin S11. The heat of calcium binding for successive additions of calcium is plotted vs the ratio of [Ca]/[P]. The data are best fit by a calculated binding curve assuming a binding constant of 7×10^6 and ΔH equal to −11.3 kcal/mol using eq 1 from the text. For comparison, calculated curves assuming $K_a = 1 \times 10^6$ and 1×10^8 are also shown. In this titration [P] = 100 μM. The temperature was 25 °C.

Table I: Titration Calorimetry of the Calcium A-Site in Subtilisin Mutants S11 and S12[a]

		parameters calculated from fit		
mutant	[P] (μM)	n	K_a	ΔH
S11	100	0.98 ± 0.01	7.8 ± 0.2 × 10^6	−11.3 ± 0.1
S11	33	0.9 ± 0.3	6.8 ± 1.5 × 10^6	−10.9 ± 0.2
S12	100	0.99 ± 0.01	6.4 ± 0.2 × 10^6	−11.8 ± 0.5

[a] Binding parameters for the stoichiometric ratio (n), binding constant (K_a), and binding enthalpy (ΔH) were determined from deconvolution using nonlinear least squares minimization (Wiseman et al., 1989). Measurements for each experimental condition were performed in duplicate at 25 °C.

[P] is between 1 and 1000 (Wiseman et al., 1989). Since the K_a of subtilisin for Ca is about 1×10^7 M^{-1}, these protein conncentrations result in values of $K_a \times$ [P] = 300 and 1000. At lower protein concentrations the amount of heat produced per titration is difficult to measure accurately. The results of fitting the titrations of S11 and S12 to a calculated curve are summarized in Table I.

The average values obtained are similar for S11 and S12: $\Delta H = \sim -11$ kcal/mol; $K_a = 7 \times 10^6$ M^{-1} and a stoichiometry of binding of one calcium site per molecule. The weak site B does not bind calcium at concentrations below the millimolar range and therefore does not interefere with measurement of binding to the A site. The standard free energy of binding at 25 °C is 9.3 kcal/mol. The binding of calcium is therefore primarily enthalpically driven with only a small net loss in entropy [$\Delta S_{binding} = -6.7$ cal/(deg mol)].

Measuring Kinetics of Calcium Dissociation. The dissociation of calcium from subtilisin is a slow process. To measure this rate, we used the fluorescent calcium chelator Quin 2. Quin 2 binds calcium with a K_a of 1.8×10^8 at pH 7.5 (Linse et al., 1987). The fluorescence of Quin 2 at 495 nm increases by ~6-fold when bound to calcium (Bryant, 1985). Subtilisin S11 or S12 as isolated contains one calcium ion per molecule. When mixed with an excess of Quin 2, the kinetics of calcium release from the protein can be followed from the increase in fluorescence at 495 nm. The reaction is assumed to follow the pathway N(Ca) ⇌ N + Ca + Quin 2 ⇌ Quin(Ca). The dissociation of calcium from subtilisin is very slow relative to calcium binding by Quin 2, such that the change in fluorescence of Quin 2 is equal to the rate of calcium dissociation from subtilisin. The inital release of calcium from S11 follows simple first-order kinetics (Figure 5A).

Temperature Dependence of Calcium Dissociation. The first-order rate constant (k) for calcium dissociation was

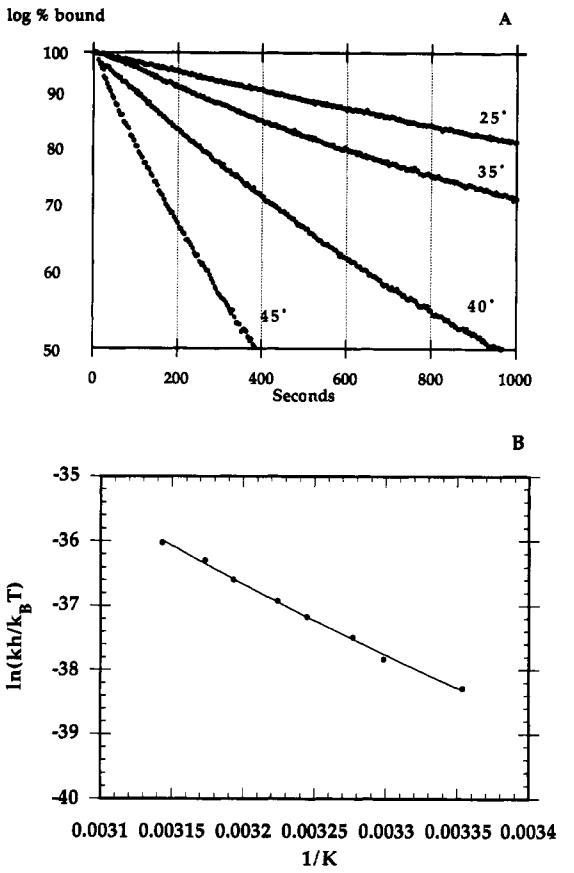

FIGURE 5: Kinetics of calcium dissociation from subtilisin S11 as a function of temperature. 1 μM subtilisin S11 was added to 10 μM Quin 2 at time = 0. Calcium dissociates from subtilisin and binds to Quin 2 until a new equilibrium is achieved. The rate of calcium dissociation is followed by the increase in fluorescence of Quin 2 when it binds to calcium. (A) The log of the percent of the protein bound to calcium plotted vs time. The kinetics of dissociation at four temperatures are shown. The dissociation follows first-order kinetics for the first 25% of the reaction. This is well before equilibrium is approached; thus reassociation of calcium can be neglected. (B) Temperature dependence of the rate of calcium dissociation from S11 subtilisin in the presence of excess Quin 2, pH 7.4, over the temperature range 25–45 °C. The natural log of the equilibrium constant for the transition state (calculated from the Eyring equation) is plotted vs the reciprocal of the absolute temperature. The line is fit according to eq 3 in the text with $T_0 = 298$ K.

measured from 20 to 45 °C. The plot of ln k vs $1/T$(K) is roughly linear. We have fit our data using transition-state theory according to the Eyring equation:

$$\Delta G^{\ddagger} = -RT \ln K^{\ddagger} = -RT \ln (kh/k_B T) \qquad (2)$$

where k_B is the Boltzmann constant, h is Planck's constant, and k is the first-order rate constant for folding. A graph of ln $(hk/k_B T)$ vs $1/T$ is shown in Figure 5B. Fitting the data to the equation (Chen et al., 1989):

$$\ln K^{\ddagger} = A + B(T_0/T) + C \ln (T_0/T) \qquad (3)$$

where $A = [-\Delta C_p^{\ddagger} + \Delta S^{\ddagger}(T_0)]/R$, $B = -A - \Delta G^{\ddagger}(T_0)/RT_0$, and $C = -\Delta C_p^{\ddagger}/R$ yields $\Delta G^{\ddagger} = 22.7$ kcal/mol, $\Delta C_p^{\ddagger} = -0.2$ kcal/(deg mol), $\Delta S^{\ddagger} = -10$ cal/(deg mol), and $\Delta H^{\ddagger} = 19.7$ kcal/mol at a reference temperature of 25 °C. A possible slight curvature of the plot would be due to a change in heat capacity associated with formation of the transition state [$\Delta C_p^{\ddagger} = 0.2$ kcal/(deg mol)]. ΔC_p for protein folding has been shown to be closely correlated with the change in exposure of hydrophobic groups to water (Privalov & Gill, 1988; Livingstone et al., 1991). In terms of heat capacity, the transition state therefore appears similar to that of the native protein. The values for $\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}$ obtained from Figure 5B indicate that

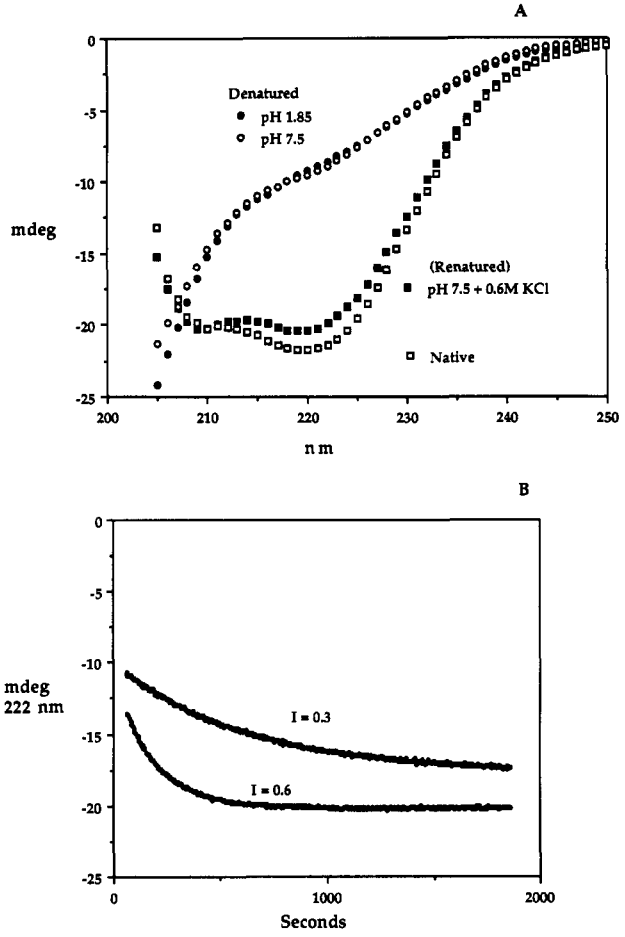

FIGURE 6: Analysis of subtilisin refolding monitored by CD. (A) CD spectra are shown as noted for (1) S15 in 25 mM H_3PO_4 at pH 1.85, (2) S15 denatured at pH 1.85 and the neutralized to pH 7.5 by the addition of NaOH, (3) S15 denatured at pH 1.85 and neutralized to pH 7.5 30 min after the addition of KCl to 0.6 M, and (4) native S15 subtilisin. The protein concentration of all samples was 1 μM. (B) Kinetics of refolding of S15. Samples were denatured at pH 1.85 and then the pH was adjusted to 7.5. At time 0, KCl was added to the denatured protein. Recovery of the native structure was followed at 222 nm at KCl concentrations of 0.3 and 0.6 M. The 0.6 M sample after 30 min of refolding was then used to record the corresponding spectrum in part A. Spectra were obtained in a Jasco 720 spectropolarimeter using a water-jacketed cell of 1 cm path length thermostated to 25 °C.

the transition state is enthalpically less favorable than the calcium-bound form with only a small change in entropy.

Characterization of the Refolding Reaction

Spectroscopic Analysis of S15 Refolding. The refolding reaction of S15 has been followed by three spectroscopic methods: circular dichroism, changes in tyrosine and tryptophan fluorescence, and changes in extinction at 286 nm. To measure refolding, the protein is first denatured at pH 1.85 by adding phosphoric acid or hydrochloric acid. S15 is completely denatured within several seconds at this pH. S12 is also completely denatured at pH 1.85 but requires about 60 min to completely unfold. If the pH of the S15 solution is quickly raised to 7.5 at relatively low ionic strength (I = 0.05), no significant refolding can be measured over the course of 1 h at 25 °C. This can best be seen by comparing the circular dichroism spectra of S15 at pH 1.85 and then the same solution neutralized to pH 7.5 (Figure 6A). The spectra are essentially identical and indicate mostly random structures. At the relatively low concentrations of protein used in the spectroscopic experiments (0.1–5.0 μM), there is no evidence of protein association or aggregation detectable by an increase in light scattering. Surprisingly, denatured S15 does not appear to collapse into a partially structured compact intermediate upon neutralization. The substantially random unfolded form of the protein appears to be kinetically locked from the native form of the protein and from any partially folded intermediates.

The native state of subtilisin is known to be stabilized by binding calcium or a monovalent cation at the low-affinity site B (Pantoliano et al., 1988). The melting temperature of S15, measured by CD, is 55 °C in 2.5 mM Tris-HCl, pH 8.0, and increases to 70 °C in the presence of 0.6 M KCl. We hypothesized that unfavorable electrostatic interactions contributing to the high activation barrier for folding S15 might be lessened by raising the ionic strength of the refolding reaction. It was found that ~90% refolding to the native form occurred within 30 min if KCl was added to the unfolded protein to a concentration of 0.6 M at pH 7.5 (Figure 6).[5] The rate of refolding is not affected by varying the protein concentration over the range of 0.1–5 μM, indicating that association of unfolded protein molecules does not appear to influence the rate of folding.

To verify that the native state of the protein had been achieved after refolding, fluorescence spectroscopy and thermal denaturation profiles were also used. Recovery of activity was not a useful criterion because the specific activity of C221 subtilisin is so low that it could not be accurately measured from the concentration of protein used in refolding experiments.[4] The fluorescence spectra of tyrosine and tryptophan for native and refolded S15 are virtually identical, as are the T_m's and shape of the denaturation profiles of native and refolded S15 subtilisin.

The folding reaction can be fit to a single first-order rate constant (Figure 6B and Figure 7A), equal to 2.8×10^{-3}/s at I = 0.6 at 25 °C (the time required for half of the protein to fold, $t_{1/2}$ = 16.6 min). If the ionic strength of the folding reaction is varied, then the log of the rate constant increases linearly with the log of ionic strength (Figure 7). Increasing the ionic strength by 10-fold results in ~90-fold increase in the rate of folding. A 90-fold increase in rate at 25 °C would correspond to a decrease in activation energy for folding of 2.65 kcal/mol. Thus at an ionic strength of 1.0, $t_{1/2}$ = 1.5 min, at I = 0.1, $t_{1/2}$ = 2.1 h, and at I = 0.01, one would predict $t_{1/2}$ = 7.6 days. Because of this strong dependence of folding rate on ionic strength, folding of subtilisin is sufficiently slow at 25 °C that we initially assumed it was not occurring at all in many previous experiments at low ionic strength.

Examination of Possible Salting Out Effects. We initially assumed that KCl accelerated folding by preferentially binding to and stabilizing the folded and transition states of S15 relative to the unfolded state. Another possibility, however, is that, at high concentration, salt competes with the protein for water such that the unfolded form of the protein is "salted out" of solution, resulting in the protein being driven to the folded form (Arakawa et al., 1990). This occurs because the native form has less area exposed to solvent than the unfolded form. We have tested whether the salting out effect contributes significantly to the folding rate by monitoring folding in the presence of a number of salts and cosolvents known to be salting out agents including $MgSO_4$, Na_2SO_4, NaCl, $MgCl_2$, poly(ethylene glycol) (PEG 3000), glucose, and gly-

[5] Exposure of S15 to the high-intensity 222-nm light beam of the Jasco 720 for 30 min appears to decrease the amount of renatured material by about 10%. Quantitative renaturation can be achieved if the sample is allowed to renature outside of the light path.

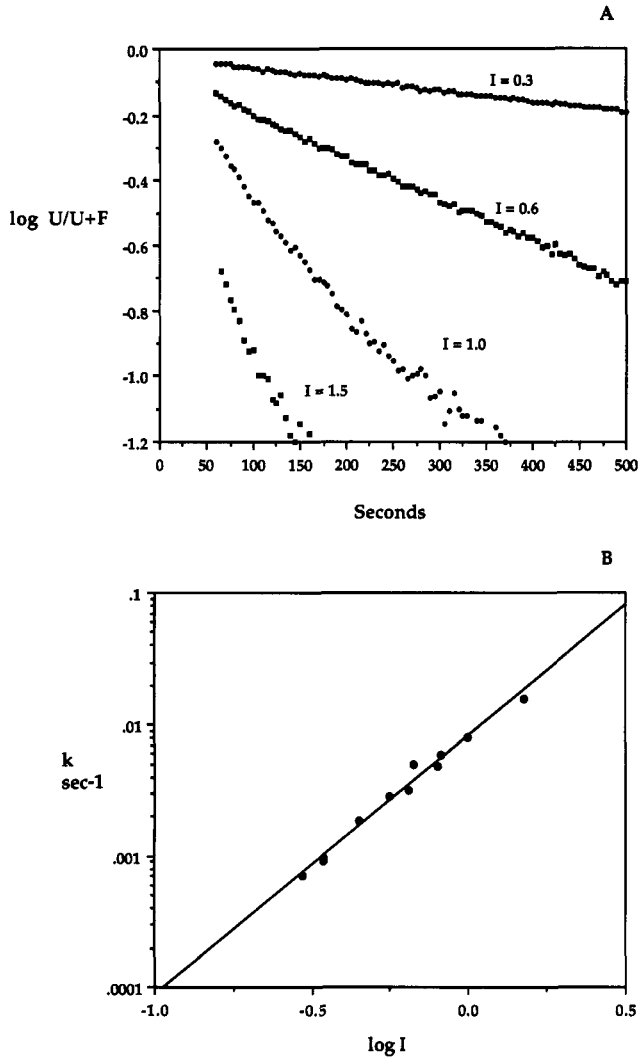

FIGURE 7: Kinetics of refolding of S15 as a function of ionic strength. (A) The log of the percent unfolded protein plotted vs time. The kinetics of refolding are shown at four ionic strengths. The amount of refolding was determined by CD from the increase in negative ellipticity at 222 nm. 100% folding is determined from the signal at 222 nm for native S15 at the same concentration and 0% folding is determined from the signal for acid-denatured S15. The refolding approximately follows first-order kinetics for the first 90% of the reaction. Refolding was carried out at 25 °C. (B) The log of first-order rate constants for refolding obtained by CD or fluorescence measurements at 25 °C plotted as a function of the log of ionic strength. Ionic strength was varied from $I = 0.25$ to $I = 1.5$. The rate of refolding increases linearly with log I. A 10-fold increase in I results in an ~90-fold increase in the refolding rate.

cerol. Some of these, such as $MgSO_4$, Na_2SO_4, NaCl, glucose, and glycerol, are known to increase protein stability because of their preferential destabilization of the unfolded state (Arakawa et al., 1990). Adding 10–50% glycerol, glucose, or PEG to denatured S15 at pH 7.5 did not result in significant refolding of S15. In contrast, all of the salts accelerated folding, but at the same ionic strength $MgSO_4$ and $MgCl_2$ were less effective than NaCl, KCl, or $CaCl_2$. The ability of a cosolvent to accelerate folding, therefore, was not correlated with its effectiveness as a salting out agent. The ability of a salt to accelerate folding seems related to the nature of the specific cation, with Na, K, and Ca similarly effective at the same ionic strength.

Temperature Dependence of the Refolding Rate. The kinetics of refolding S15 subtilisin were measured over the range 20–40 °C in the presence of 0.6 M KCl. The extent of refolding was measured by the change in tyrosine and tryptophan fluorescence vs time. The folding reaction can be fit to a single relaxation time at each temperature, but a plot of ln k vs $1/T$ is not linear, indicating that the folding reaction cannot be described in terms of simple Arrhenius theory. This behavior is not unexpected and has been observed previously for the folding reaction of other proteins (Pohl, 1968; Chen et al., 1989; Jackson & Fersht, 1991). As pointed out by Chen et al., a change in heat capacity going from the unfolded state to the transition state is expected since the overall ΔC_p of folding is large (3.3 kcal/mol). This would explain the observed curvature in an Arrhenius plot.

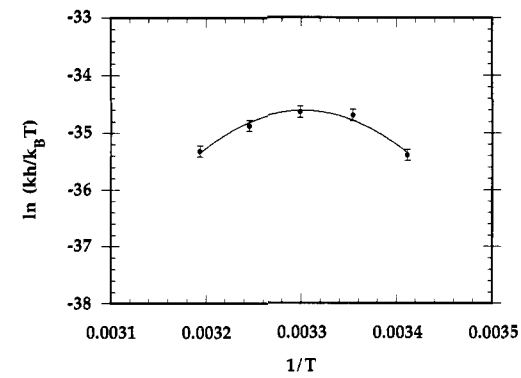

FIGURE 8: Temperature dependence of the refolding rate of S15 subtilisin in 0.6 M KCl and 23 mM potassium phosphate, pH 7.3. The natural log of the equilibrium constant for the transition state (calculated from the Eyring equation) is plotted vs the reciprocal of the absolute temperature. The line is fit according to eq 3 in the text with $T_0 = 298$ K.

We have fit our data using transition-state theory according to eq 2. A graph of ln (hk/k_BT) vs $1/T$ is shown in Figure 8. Fitting the data to eq 3 (Chen et al., 1989) yields $\Delta G'^{\ddagger} = 20.6$ kcal/mol, $\Delta C_p'^{\ddagger} = -2.7$ kcal/(deg mol), $\Delta S'^{\ddagger} = -26$ cal/(deg mol), and $\Delta H'^{\ddagger} = 12.8$ kcal/mol at a reference temperature of 25 °C. The prime symbol denotes that parameters are for subtilisin folding in the presence of 0.6 M KCl. This treatment of the kinetic data is based on vibrational theory in which the frequency of breakdown of the transition state to product is equal to $\kappa kBT/h$, where κ is a transmission factor assumed to be equal to 1. It is doubtful that this theoretical treatment is applicable to protein folding reactions in all respects; thus thermodynamic parameters determined for the transition state for S15 folding should be viewed circumspectly. Figure 8 shows that the rate of folding reached a maximum around 30 °C. The pronounced curvature is due to the large $\Delta C_p'^{\ddagger}$ [-2.7 kcal/(deg mol)]. The change in heat capacity associated with formation of the transition state is about 80% of the total heat capacity change for folding. As noted above, the ΔC_p for protein folding is closely correlated with the change in exposure of hydrophobic groups to water; thus the transition state appears to be relatively compact and more similar to the fully folded state than to the unfolded state. Similar observations were made for the transition states of a T4 lysozyme mutant in 3 M Gu-HCl (Chen et al., 1989) and of chymotrypsin inhibitor 2 (Jackson & Fersht, 1991). The values for $\Delta S'^{\ddagger}$ and $\Delta H'^{\ddagger}$ obtained from Figure 8, though prone to errors, indicate that the transition state is both entropically and enthalpically less favorable at 25 °C than the unfolded state. At an ionic strength of 0.01, the rate of folding extrapolates to about 1×10^{-6}/s, which would correspond to $\Delta G^{\ddagger} = \sim 26$ kcal/mol at 25 °C.

Discussion

Typical of extracellular proteases, subtilisin BPN′ is a robust protein (Matthews et al., 1974; Voordouw et al., 1976;

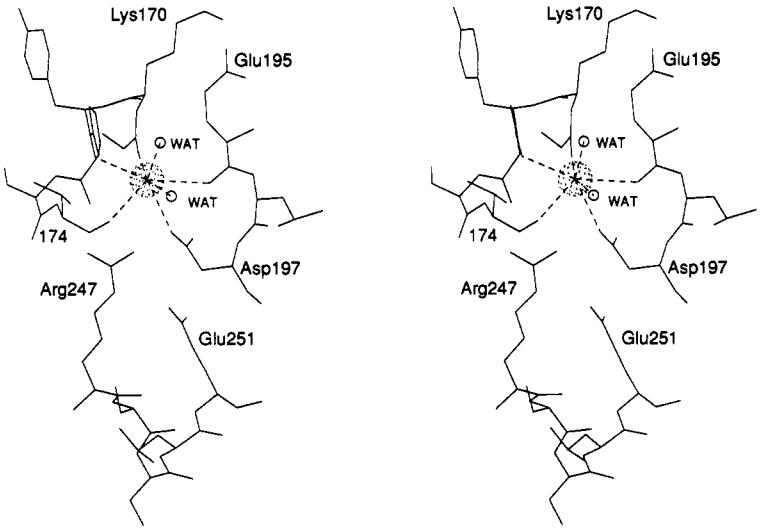

FIGURE 9: X-ray crystal structure of the weak ion binding region of S15 subtilisin. Coordination bonds are shown as dashed lines. Note the preponderance of charged amino acids.

Frommel & Höhne, 1981; Betzel, 1990; Gros et al., 1991). It is an unusual example of a monomeric protein with a substantial kinetic barrier to folding and unfolding (Sanchez-Ruiz et al., 1988). The high thermodynamic and kinetic stability of subtilisin is in large part mediated by the high-affinity calcium site A. Calcium binding to site A contributes about 5 kcal/mol to the free energy of unfolding at 1 mM calcium and 25 °C. Furthermore, the kinetics of calcium dissociation are very slow at 25 °C due to a free energy of activation for calcium dissociation of ~23 kcal/mol, which explains in part the high kinetic stability.

The structure of the calcium A-site in subtilisin is fundamentally different from the EF hand calcium binding sites of calbindin, parvalbumin, calmodulin, α-lactoalbumin, or troponin C. All of the protein ligands to calcium in an EF hand structure are on a single loop of about 12 amino acids connecting two α-helices, and three or four of the side-chain ligands to the calcium are aspartic acid or glutamic acid. By contrast, the protein ligands of the A-site come from three noncontiguous regions of subtilisin, and only one of the ligands (D41) carries a formal charge. The K_a of the A-site ($\sim 10^7$ M^{-1}) is comparable to that of EF hand structures which range from 10^5 to 10^9 M^{-1}. Also, the binding of calcium to both types of site is enthalpically driven. A major difference in the A-site, however, is that the on and off rates for calcium are three to four *orders of magnitude* slower. This difference may reflect that EF hand proteins are typically part of a calcium messenger system and are designed to be responsive to changes in calcium concentration in order to trigger various biological processes. In subtilisin, calcium is an integral part of the structure, and its association or dissociation probably requires significant but transient disruption in surrounding protein–protein interactions. This disruption in structure would explain the high activation energy and slow kinetics of calcium binding and dissociation. For example, breaking main-chain hydrogen bonds between the N-terminal region and the loop region (Figure 2) would allow the relatively buried calcium a passageway into or out of the protein. This is only one of several plausible models for the process of calcium dissociation, however. Calcium binding sites are common features of extracellular proteases probably because of their large contribution to both thermodynamic and kinetic stability (Matthews et al., 1974; Voordouw et al., 1976; Frommel & Höhne, 1981; Betzel, 1990; Gros et al., 1991).

The S15 mutant, which lacks the calcium A-site, is accordingly no longer very stable by a thermodynamic definition. Surprisingly, S15 still appears to be quite stable by a kinetic definition, having a significant activation barrier between the native and unfolded form, independent of the A-site. This robust design of subtilisin is probably necessitated by the rigors of the extracellular environment into which subtilisin is secreted, which by virtue of its own presence is protease-filled. Accordingly, high activation barriers to unfolding may be essential to lock the native conformation and prevent transient unfolding and proteolysis. The disadvantage of a kinetically stable native state is that folding may also be slow. We estimate that the free energy of activation for refolding S15 subtilisin is ~26 kcal/mol at an ionic strength of 0.01 and 25 °C. One reason that subtilisin is initially synthesized in a pro form may be to catalyze folding (Ikemura et al., 1987; Zhu et al., 1989). The precise mechanism by which this is accomplished is under investigation.

Many proteins are difficult to fold in vitro. One common obstacle to proper folding is aggregation during the folding process. Initially we had assumed that this was the reason that mature subtilisin was difficult to fold. Unfolded, mature subtilisin does aggregate at neutral pH at concentrations greater than 1 mg/mL. At lower protein concentrations, as reported here, the primary impediment to folding seems to be the high free energy of activation for the process. We have shown that one way this barrier can be reduced is by the addition of salt to the folding reaction.

We suggest that, in the S15 mutant, one component of the high-energy barrier to folding is related to the cation binding site located around Asp 197 in subtilisin BPN′ (Figure 9) (Pantoliano et al., 1988). This weak ion binding site (site B) has an association constant for calcium of ~30 mM and in the absence of high calcium concentrations will bind sodium or potassium. We have observed that the melting temperature of S15 subtilisin is increased by ~15 °C by 0.6 M KCl. Ribonuclease T1 is also known to be stabilized by the preferential binding of monovalent and divalent cations to the native state (Pace & Grimesly, 1988). The region of subtilisin around site B has a high charge density with two Glu, one Asp, one Arg, and one Lys within an 8-Å radius. Particularly striking is the close juxtaposition of Glu 251 and Asp 197. Kinetic analysis of S15 indicates that the transition state for folding is relatively compact and nativelike in its thermody-

namic properties. The coalescing of the charged amino acids to form a rudimentary B-site in the transition state may present an electrostatic barrier to folding unless the charges can be shielded as they are brought into proximity. Cations therefore would be expected to stabilize both the transition state for folding and the folded state. Kinetic analysis of mutations in the B-site region will be useful in assessing its role in the transition state.

This work should be relevant to other problems of in vitro protein folding in that it demonstrates how efficient folding can be achieved by modest reductions in the activation energy of the folding reaction. In this case, the activation energy for subtilisin folding can be reduced by high ionic strength. In a subsequent report, we will describe how destabilizing the unfolded state through the introduction of a disulfide bond into S15 subtilisin results in more than a 100-fold increase in the refolding rate.

ACKNOWLEDGMENTS

We thank Joel Hoskins for synthesizing the oligonucleotides used in site-directed mutagenesis and DNA sequencing and Kathryn Fisher for advice on gene expression and protein production. We acknowledge Michael Pantoliano, John Moult, Maurice Krauss, Matthew Mauro, and Edward Eisenstein for advice and helpful discussion. We also thank one of the referees for helpful comments regarding the interpretation of binding free energies.

REFERENCES

Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kosiakoff, A., & Wells, J. A. (1991) *Biochemistry 30*, 4151–4159.

Arakawa, T., Bhat, R., & Timasheff, S. N. (1990) *Biochemistry 29*, 1914–1923.

Betzel, Ch., Teplyakov, A. V., Harutyunyan, E. H., Sänger, W., & Wilson, K. S. (1990) *Protein Eng. 3*, 161–172.

Bode, W., Papamokos, E., & Musil, D. (1987) *Eur. J. Biochem. 166*, 673–692.

Brown, R. D., Brewer, C. F., & Koenig, S. H. (1977) *Biochemistry 16*, 3883–3896.

Bryan, P. N. (1992) *Pharmaceutical Biotechnology* (Botchard, R. T., Ed.) Vol. 3, Plenum Press, New York, NY (in press).

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

Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H. Y., & Poulos, T. (1986b) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3743–3745.

Bryant, D. T. W. (1985) *Biochem. J. 226*, 613–616.

Chen, B., Baase, W. A., & Schellman, J. A. (1989) *Biochemistry 28*, 691–699.

DelMar, E., Largman, C., Brodrick, J., & Geokas, M. (1979) *Anal. Biochem. 99*, 316–320.

Fahnestock, S. R., & Fisher, K. E. (1987) *Appl. Environ. Microbiol. 53*, 379–384.

Finzel, B. C., Howard, A. J., & Pantoliano, M. W. (1986) *J. Cell. Biochem., Suppl. 10A*, 272.

Gros, P., Kalk, K. H., & Hol, W. G. J. (1991) *J. Biol. Chem. 266*, 2953–2961.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ranseshan, S., & Venkatesan, K., Eds.) pp 13.01–13.23, Indian Institute of Science, Bangalore.

Ikemura, H., Takagi, H., & Inouye, M. (1987) *J. Biol. Chem. 262*, 7859–7864.

Jackson, S. E., & Fersht, A. R. (1991) *Biochemistry 30*, 10436–10443.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Linse, S., Brodlin, P., Drakenberg, T., Thulin, E., Sellers, P., Elmdén, K., Grundström, T., & Forsén, S. (1987) *Biochemistry 26*, 6723–6753.

Livingstone, J. R., Spolar, R. S., & Record, T. M. (1991) *Biochemistry 30*, 4237–4244.

Matthews, B. W., Weaver, L. H., & Kester, W. R. (1974) *J. Biol. Chem. 249*, 8030–8044.

McPhalen, C. A., & James, M. N. G. (1988) *Biochemistry 27*, 6582–6598.

Pace, C. N., & Grimesly, G. R. (1988) *Biochemistry 27*, 3242–3246.

Pantoliano, M. W., Ladner, R. C., Bryan, P. N., Rollence, M. L., Wood, J. F., & Poulos, T. L. (1987) *Biochemistry 26*, 2077–2082.

Pantoliano, M. W., Whitlow, M., Wood, J. F., Rollence, M. L., Finzel, B. C., Gilliland, G., Poulos, T. L., & Bryan, P. N. (1988) *Biochemistry 27*, 8311–8317.

Pantoliano, M. W., Whitlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L., & Bryan, P. N. (1989) *Biochemistry 28*, 7205–7213.

Pohl, F. M. (1968) *Eur. J. Biochem. 7*, 146–152.

Privalov, P. L., & Potekhin, S. A. (1986) *Methods Enzymol. 131*, 4–15.

Privalov, P. L., & Gill, S. J. (1988) *Adv. Protein Chem. 39*, 191–234.

Rollence, M. L., Filpula, D., Pantoliano, M. W., & Bryan, P. N. (1988) *CRC Crit. Rev. Biotechnol. 8*, 217–224.

Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L., Cortijo, M., & Mateo, P. L. (1988) *Biochemistry 27*, 1648–1652.

Schwarz, F. P., Puri, K. D., & Surolia, A. (1991) *J. Biol. Chem. 266*, 24344–24350.

Siezen, R. L., de Vos, W. M., Leunissen, J. A. M., & Dijkstra, B. W. (1991) *Protein Eng. 4*, 719–737.

Takahashi, K., & Sturtevant, J. M. (1981) *Biochemistry 20*, 6185–6190.

Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J., & Filpula, D. (1984) *J. Bacteriol. 159*, 811–819.

Voordouw, G., Milo, C., & Roche, R. S. (1976) *Biochemistry 15*, 3716–3724.

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

Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., & Chen, E. Y. (1983) *Nucleic Acids Res. 11*, 7911–7925.

Wiseman, T., Williston, S., Brandts, J. F., & Lin, L.-N. (1989) *Anal. Biochem. 179*, 131–137.

Zhu, X., Ohta, Y., Jordan, F., & Inouye, M. (1989) *Nature 339*, 483–484.

Zoller, M. J., & Smith, M. (1983) *Methods Enzymol. 100*, 468–500.