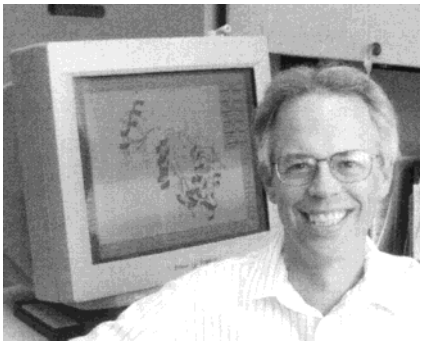
Prodomains and Protein Folding Catalysis†

Philip N. Bryan*

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

Received May 30, 2002

Contents



Philip Bryan obtained a Ph.D. in 1979 with Donald Olins at the Biology Division of Oak Ridge National Laboratory where he used biochemical and biophysical approaches to characterize nucleosome structure. His postdoctoral work with Max Birnstiel at the University of Zurich and William Folk at the University of Michigan investigated the role of chromatin structure on the expression of developmentally regulated genes. Bryan joined Genex Corporation in Rockville, Maryland, in 1984 to use genetic techniques to engineer proteins of increased stability. From 1989 to the present, he has applied of genetic, biochemical, and biophysical methods to the study of protein folding and stability at the University of Maryland Biotechnology Institute's Center for Advanced Research in Biotechnology where he is a professor.

I. Secreted Proteases and Prodomains

Secreted proteases are synthesized as inactive zymogen precursors to tightly regulate the timing of protease activation.[1] Frequently, the zymogen precursor consists of N-terminal amino acids attached to the mature protease sequence. A number of these N-terminal extensions (prodomains) are large enough to fold independently and have been shown to be required for folding of the mature protease into its native conformation. Over time, it has become apparent that the role of the prodomain in protease folding has implications beyond their specific function in regulating the timing of protease activation.[2−9] The major goal of this review is to describe what general inferences about protein folding can be made from prodomain catalyzed folding. General questions include whether native protein conformation is always thermodynamically determined, what causes high kinetic barriers between unfolded and native states, and what do prodomains do to reduce these kinetic barriers.

Many prodomains have been implicated in the folding process.[10] The review will focus on the two proteases for which the most mechanistic information is available: subtilisin (SBT) and α-lytic protease (ALP). Both are serine proteases secreted from bacteria but are not evolutionarily related.[11,12] Each has eukaryotic homologues. ALP is a structural homologue of chymotrypsin and trypsin, and SBT is homologous in sequence and presumably structure to prohormone convertases.[13]

II. Identification of Prosequences

Sequencing of the SBT gene from *Bacillus amyloliquefaciens* in the early 1980s revealed that the primary translation product is a pre-pro-protein.[14,15] A 30-amino acid presequence serves as a signal peptide for protein secretion across the cell membrane and is hydrolyzed by a signal peptidase.[16] A 77-amino acid sequence, termed a propeptide, was found between the signal sequence and the 275-amino acid mature SBT sequence. The prosequence was shown to be removed autocatalytically and was suggested to function as an inhibitor of the mature enzyme to delay its activation until after secretion from Bacillus.[17] In 1987, it was demonstrated that the propeptide is required for the secretion of active SBT in a heterologous host (*Escherichia coli*), and it was suggested that the propeptide guides the folding process.[18]

A role for the propeptide in folding was confirmed subsequently when it was shown to promote the refolding of the mature SBT sequence in vitro in an intermolecular process.[19] In this experiment, the 275-amino acid mature sequence was produced in E. coli in inactive form and then dialyzed against 6 M

† Supported by NIH Grant GM42560.
* To whom correspondence should be addressed. Tel: 301-738-6220. FAX: 301-738-6255. bryan@umbi.umd.edu.

10.1021/cr010190b CCC: $39.75
Published on Web 11/09/2002

guanidine hydrochloride (Gu-HCl) to ensure completely unfolding. When the Gu-HCl then was removed from the mature protein by dialysis, no activity was recovered. However, if the propeptide sequence was added to the renaturing dialysis, then SBT activity was recovered in proportion to the concentration of added propeptide. The experiment was actually a little more complicated in that the refolding catalyst was actually proSBT with a mutation of the active site amino acid aspartic acid 32 to asparagine (D32N). The D32N proSBT is inactive and therefore unable to auto-process.[20] It would be shown subsequently that only the 77-amino acid prodomain was needed to catalyze folding of mature SBT.[21] This paper also made two other key observations. The first was that the 77-amino acid prodomain is a competitive inhibitor of the active enzyme (K_i of 5.4×10^{-7} M), and the second was that the entire prosequence is required for strong inhibition. The first observation showed that the propeptide binds strongly to the native conformation of SBT, and the second was a strong indication that the propeptide was actually a domain that required structure beyond direct interactions with the substrate binding pockets of SBT for its function.

About five years after the cloning of SBT, the gene for ALP was cloned from the gram negative bacterium *Lysobacter enzymogenes*.[22] Sequencing revealed that ALP was also synthesized as a proenzyme. In this case, the prodomain (166 amino acids) was almost as large as the mature protease (198 amino acids). A year later, it was demonstrated that the ALP prodomain was required to produce active ALP in vivo and further that the 166-amino acid prodomain could promote the production of active ALP even when expressed on a separate plasmid. On this basis, it was proposed that the prodomain provides a template for folding the mature ALP into the active conformation.[23]

In 1991, the metalloprotease, carboxypeptide Y from yeast, was discovered to be a third example of a protease whose in vitro folding required the participation of a prodomain.[24] It was suggested that the prodomain might help overcome energy barriers to a productive folding pathway. Other examples of prodomain mediated folding have been found in all four mechanistic families of proteases: serine proteases,[25−30] aspartic proteases,[31−33] metalloproteases,[34−38] and cysteine proteases.[39]

III. Prodomains as Folding Catalysts

In 1992, Baker and Agard performed a series of experiments that began directly to address the fundamental issues about protein folding raised by earlier experiments.[40,41] They performed an in vitro characterization of the following bimolecular folding reaction:

$$U \leftrightarrow I + P \leftrightarrow I{-}P \leftrightarrow N{-}P \leftrightarrow N + P$$

where U is the Gu-HCl denatured form of ALP, I is a partially folded but kinetically trapped intermediate, P is the 166-amino acid prodomain, and N is native ALP. They made the following observations.

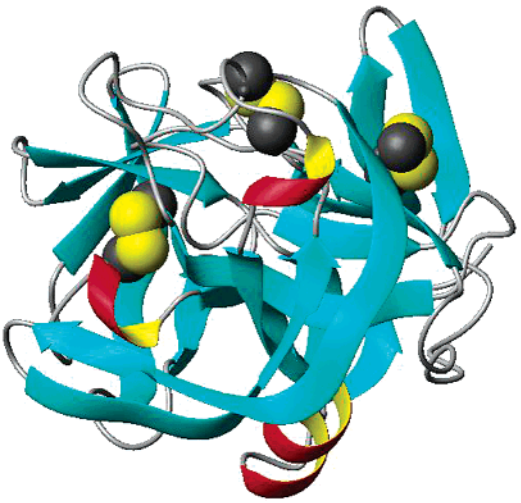


Figure 1. Ribbon drawing depicting the α-carbon backbone of α-lytic protease. Positions of the three disulfide cross-links are shown with space-filling spheres.

Mature ALP denatured in 6 M Gu-HCl and then diluted into native conditions rapidly collapses into a ensemble of conformations (termed "I") with an average hydrodynamic radius about midway between the native and Gu-HCl denatured forms and with almost the same amount of regular secondary structure (β-sheet) as native ALP. Because no reducing agent was added in the experiment, U, I, and N forms of ALP have all three native disulfide bonds intact (17−37, 101−111, and 137−170) (Figure 1). Despite the preorganization provided by the disulfides, conversion of the I state to N was undetectable ($k_{folding} < 10^{-9}$ s^{-1}). When P (0.04 μM) was added to an excess of I in a bimolecular reaction, a single turnover of folding to N was observed at a rate of 0.016 s^{-1}. The product of the folding reaction was a tight complex between P and N with an K_i of about 10^{-10} M[40], which is tight enough to effectively eliminate multiple rounds of ALP folding by a single P molecule. The quantitative assessment of native ALP formation was made by subsequently digesting away P with trypsin to release active ALP. The results showed that the complex I−P converts to N−P at a rate at least 10^7 fold faster than I spontaneously converts to N. The results clearly show that a high kinetic barrier exists between the I and N states, and this barrier prevents equilibrium between the states to be attained on a biological time scale. Hence, kinetics rather than thermodynamics determines the native state of ALP.

The activation barrier between the I and N states was reported as ~30 kcal/mol.[41,42] This seems to have been calculated by comparing the upper limit of folding rate (10^{-9} s^{-1}) with an absolute reaction rate calculated using Eyring theory. In Eyring theory, unimolecular bond breaking is limited by the frequency of molecular vibration (~10^{13} s^{-1}). Although an absolute reaction rate is sometimes used as a hypothetical reference state for protein folding reactions,[43] its relevance is frequently questioned.

In 1993, Eder and Fersht performed characterizations of both the bimolecular folding reaction of SBT and its prodomain[44] and the unimolecular reaction.[45] In the characterization of the bimolecular reaction, they denatured mature SBT (275 amino acids) in 6

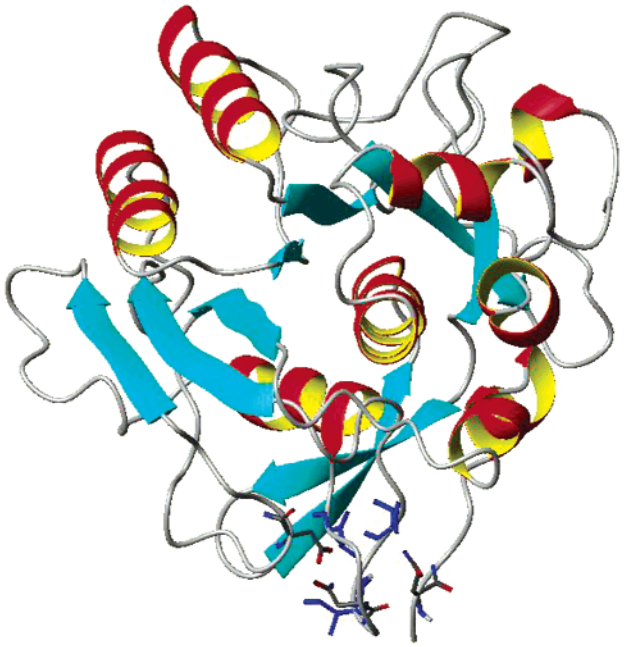


Figure 2. Ribbon drawing depicting the α-carbon backbone of subtilisin. Calcium binding sites A is depicted with only side chain ligands shown.

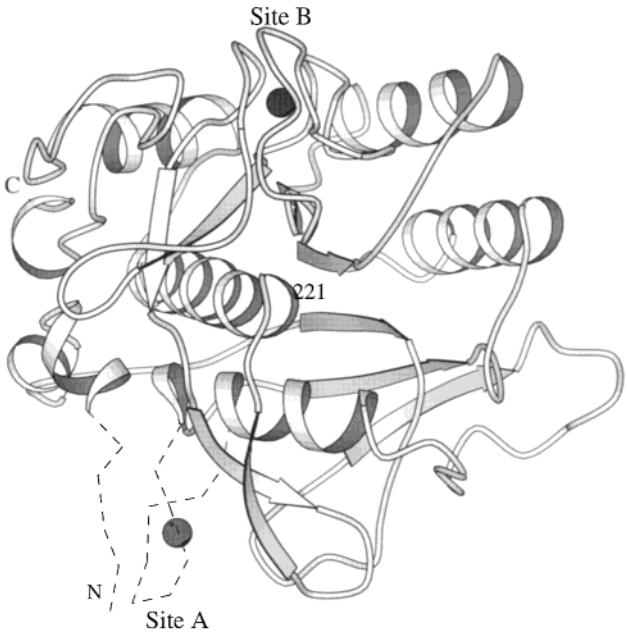


Figure 3. Ribbon drawing depicting the α-carbon backbone of subtilisin. Cation binding sites A and B are shown as gray spheres. The position of the active site serine is numbered (221). The calcium binding loop of site A (amino acids 75−83) and the N-terminal amino acids (1−5) are shown by the dashed lines.

M HCl, pH 1.8 and then either dialyzed or diluted the protein into refolding buffer around neutral pH. The refolding was actually carried out on a mutant SBT in which the active site serine 221 was converted to alanine. The S221A mutation reduces proteolytic activity of SBT by about 10^6-fold[46] and thus eliminates the problem of autoproteolysis during the refolding process. Upon return to native conditions in high salt (e.g., 1.0 M KCl), mature SBT slowly folds into a state reminiscent of the I state of ALP: hydrodynamic radius intermediate between native and unfolded and a high percentage of secondary structure. Remarkably, the intermediate state binds a stoichiometric amount of calcium with an affinity ≤ 1 μM. Native SBT has a calcium-specific, high affinity binding site (site A, $K_d \sim 100$ nM).[47] In native SBT, calcium at site A is coordinated by five carbonyl oxygen ligands and one aspartic acid (Figure 2). Four of the carbonyl oxygen ligands to the calcium are provided by a loop comprising amino acids 75−83 (Figure 3). The geometry of the ligands is that of a pentagonal bipyramid whose axis runs through the carbonyls of 75 and 79. The other ligands in the loop are the delta oxygen of N77 and the carbonyl oxygen of V81. 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-beta arrangement. A less-specific second ion binding (site B, $K_d \sim 10$ μM[48] is located 32 Å from site A in a shallow crevice between two segments of polypeptide chain near the surface of the molecule (Figure 3). The intermediate binds calcium more tightly than site B and which raises the possibility that site A is already formed in the intermediate. This would imply that substantial nativelike tertiary structure is present in the intermediate.

Adding the isolated prodomain (5 μM) to the intermediate (5 μM) resulted in slow refolding to a native complex. After 8 days of incubation at 4 °C, about 50% of the intermediate had been converted to the native form.[44] This would correspond to a rate of 0.2 M^{-1} s^{-1}. Since the diffusion rate would be expected to exceed 10^8 M^{-1} s^{-1}, catalyzed folding is extremely sluggish in the bimolecular reaction. No conversion of the intermediate of the native form was detected in the absence of added prodomain.[45] This experiment has a detection limit of ~1%, which would mean that uncatalyzed folding of mature SBT occurs at a rate of $< 10^{-8}$ s^{-1}.

Eder and Fersht also analyzed the unimolecular folding of the S221A mutant of proSBT.[45] Since proSBT is rapidly autoprocessed to the mature form in Bacillus, the active site mutation S221A mutation was used to trap the unprocessed form of SBT. This allowed expression in *E. coli* and purification from inclusion bodies. Upon dilution from 5 M Gu-HCl, proSBT was found to fold at a rate of 0.0047 s^{-1}. The stability of unprocessed proSBT is quite low compared to the process complex and requires high salt (e.g., 0.5 M ammonium sulfate) to maintain nativelike spectral properties. Under lower salt conditions (e.g., 0.2 M ammonium sulfate) and at protein concentrations of 0.5−1.0 mg/mL, proSBT forms multimers.[49] As with the folding intermediate of mature SBT, proSBT binds stoichiometric calcium and also binds the potent SBT inhibitor CI2 with reduced affinity compared to native SBT.[45] Putting these results into the context of the structure of the bimolecular prodomain SBT complex, gives the following picture.

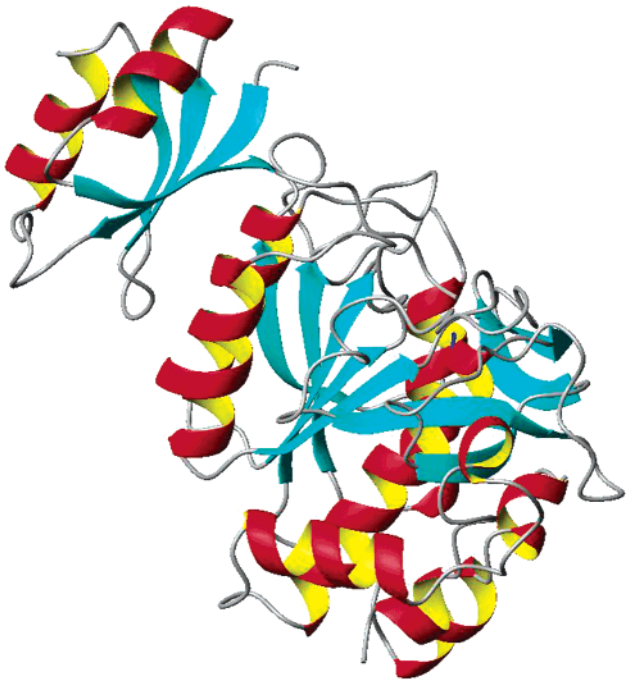


Figure 4. Ribbon drawing depicting the α-carbon backbone of subtilisin in complex with its prodomain.

In the bimolecular complex, the C-terminus of the prodomain binds in the active site cleft of SBT in a substrate-like manner with Y77 occupying the S1 subsite of SBT[50] (Figure 4). If Y77 occupies the same position in unprocessed proSBT then residues 1−10 of mature SBT would be displaced from their native positions. This hypothetical conformer can be modeled by minor torsional reorganizations involving residues 1−10 residues with Ala1 and Gln2 of mature SBT occupying the S1′ and S2′ substrate sites. A major consequence for this "active site" conformer is that the calcium site A cannot fully form until after processing because one of the calcium ligands (Gln2) cannot simultaneously bind in the S2′ subsite and bind to calcium in the A-site. If on the other hand, pro-SBT is a "calcium site" conformer with the N-terminus bound to calcium at site A, then the prodomain will be dislocated from its binding surface on SBT, extended into solvent and available to bind to another SBT molecule intermolecularly. This dichotomous situation would explain the low stability of proSBT, its diminished affinity for CI2, its calcium affinity, and its tendency to form multimers. For more details on unimolecular folding, the reader is referred to refs 51−58.

The unimolecular folding reaction of ALP has some similarity to that of proSBT but also important differences. This study was also carried out on a mutant with the active site serine changed to alanine to prevent autoprocessing.[59] The folded proALP is marginally stable, and 15% glycerol is required to maintain solubility. As with SBT, the C-terminus of the prodomain binds tightly in the active site of ALP in the bimolecular complex[60] (Figure 5). Thus, in the unprocessed molecule, the N-terminus of the active site conformer would be displaced by 24 Å from its position in native ALP. Unlike SBT, ALP does not have the complications of metal binding, but does have three native disulfide bonds. The kinetics of

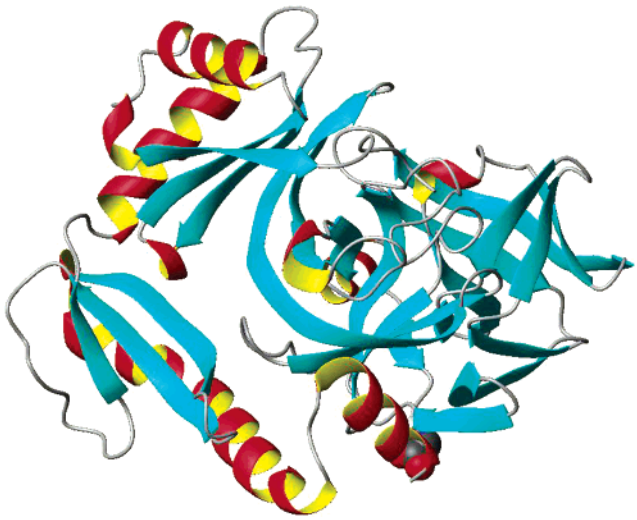


Figure 5. Ribbon drawing depicting the α-carbon backbone of α-lytic protease in complex with its prodomain.

folding when the disulfide bonds are preformed is biphasic but rapid ($k_1 = 1.1$ s^{-1} and $k_1 = 0.4$ s^{-1}). The faster rate corresponds to the rate of prodomain folding and the slower rate appears to correspond to prodomain mediated ALP folding. This rate is ∼30 times faster than the folding step in the bimolecular reaction. This contrasts with SBT in which the folding step in bimolecular folding is >1000 times slower than unimolecular folding. The rate of proALP folding from the reduced state is limited by the rate of disulfide exchange reactions.[59]

IV. Dissecting the Bimolecular Folding Reaction

A. Unimolecular vs Bimolecular Folding

Clearly, the biological folding reaction is unimolecular, and unraveling this reaction will provide crucial insight into how protease activation is regulated through folding and processing. The major goal of this review, however, is to describe what general inferences about protein folding can be made from procatalyzed folding. Most of the research on the energetics of protein folding has concentrated on relatively small, monomeric, globular proteins, which readily unfold and refold. The primary forces driving protein folding as well as the basic features of influencing folding kinetics have been actively explored in these systems. For example, the role of contact order in influencing folding rates of small proteins (<110 amino acids) with two state folding reactions has been an important recent development.[61−63] For larger proteins (>110), in which intermediates usually accumulate in the folding reaction,[64] determinants of folding rates have remained controversial. Debate continues concerning whether some intermediates are on the pathway and accelerate folding or whether they are off the pathway and hinder folding.[65−69] Folding of proteins the size of mature SBT and ALP frequently occurs within seconds. While this is not always true, the reasons for slow or inefficient folding are usually apparent. Two factors with a great impact on folding rate are disulfide exchange reactions and aggregation of

unfolded states. Both of these factors do not influence the reactions, which will be discussed subsequently. Other common impediments to rapid folding such as slow conformational isomerization due to prolines and stable, misfolded conformations may play a role.

In regard to general questions about protein folding, analysis of the bimolecular folding reaction has been particularly informative. Part of the reason for this is practical. Unlike the unimolecular reaction, the product of bimolecular folding is a stable complex between the native protease and the folded prodomain. Atomic resolution structures of this complex has been determined for both SBT and ALP.[50,60,70] The other reason is conceptual. Consider the general folding reaction:

$$U \leftrightarrow N$$

To characterize the reaction, one would need to determine microscopic rate constants in the reaction and determine the structure of the reactant, product, and any populated intermediates in the pathway. This has been the goal of procatalyzed folding studies as well, with the added element that one can study the folding reaction ± prodomain. This is done to answer two basic questions: Why do these proteins fold slowly without the prodomain, and what does the prodomain do to accelerate the folding rate?

The reaction of prodomain with unfolded protease is an intriguing case of complementation. The folded protease is stable in a practical sense. That is, the native conformation persists under favorable solvent conditions for a long period on an experimentalist's time-scale (e.g., >weeks). Yet, once denatured and returned to these same favorable conditions, it remains in an unfolded, conformation for a long period of time. When the isolated prodomain is added, folding occurs spontaneously.

B. What Is the Stability of the Mature Protease?

The most basic folding experiment is determining the free energy of the folding reaction. This is difficult with ALP and SBT because of the impracticality of establishing equilibrium under conditions in which the ratio of folded and unfolded states can be accurately measured. The free energy of a two-state reaction can sometimes be determined, however, by (1) measuring a folding rate at a standard state; (2) measuring unfolding rates as a function of denaturant; (3) extrapolating the unfolding rate to the standard state. From the ratio of the rates of folding and unfolding at the standard state, the free energy of folding is determined. This was the approach taken with ALP.[71] To minimize the influence of autodigestion during the folding reaction, the measurements were made at pH 5.0. Proteolytic activity of serine proteases decreases as the pH is decreased below the pK_a of the active site histidine (p$K_a \sim 7$). A pH around 5 is commonly used as a compromise between low activity and preserving some conformational stability.[72] The refolding rate of ALP was determined by incubating denatured protein in native conditions and removing aliquots at intervals. The amount of active ALP was determined by measuring peptidase activity using a synthetic substrate. Over the course of days, small but linear increases in activity were detected. Quantitation in the experiment is tricky because the accumulation of active ALP causes digestion of the unfolded protein, but the initial rate of the folding reaction is 6×10^{-12} s^{-1} at 25 °C. This was compared to the unfolding rate which was determined as a function of [Gu-HCl]. The rate of unfolding without Gu-HCl was estimated by linear extrapolation to be 1×10^{-7} s^{-1} at 25 °C. Since $\Delta G_{\text{unfolding}} = -RT \ln(k_{\text{unfolding}}/k_{\text{folding}})$ in a two-state system, the simplest interpretation of the unfolding and refolding rates would mean that $\Delta G_{\text{unfolding}}$ is about −6 kcal/mol at 25 °C. Even given the uncertainties in determining the folding and unfolding rates, it is convincing that the unfolded state of ALP is more stable than the folded state under these conditions. Further, this observation is made with the three disulfide bonds of ALP intact in all forms U, I, and N.

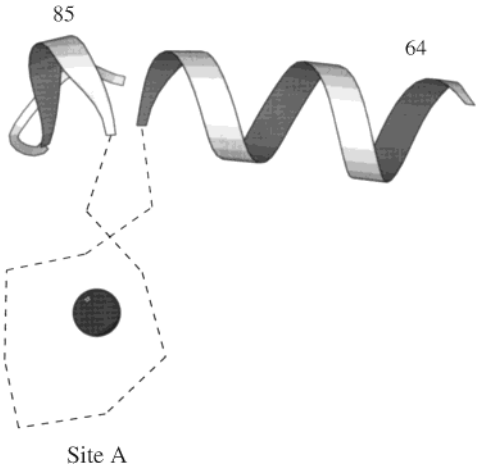


Figure 6. Calcium binding loop of subtilisin. Region of deletion (75−83) is shown by the dashed line.

Measurements of the free energy of SBT folding are also more convoluted than one would like. The most straightforward estimate of the free energy of folding was determined by denaturing and renaturing SBT immobilized on agarose beads to prevent autodigestion.[73] Quantitative renaturation was achieved in 24 h in the presence of 2 M potassium salts, indicating that the folding rate under these conditions must be $> 10^{-5}$ s^{-1}. The unfolding rate is not known under these conditions, but the $\Delta G_{\text{unfolding}}$ must be greater than 2 kcal/mol since recovery of native activity approaches 100%. It is entirely possible, however, that SBT, like ALP, is thermodynamically unstable at lower [salt] or at pH 5.

An additional element in the thermodynamics of SBT folding is calcium binding to site A. A mutant SBT lacking site A was designed to circumvent the additional complexity of calcium binding.[47] The calcium-binding loop is formed from a nine-amino acid bubble in the last turn of a 14-residue α-helix involving amino acids 63−85[74](Figure 6). Deleting amino acids 75−83 creates an uninterrupted helix and abolishes the calcium binding potential at site A.[47,75,76] The X-ray structure has shown that except for the region of the deleted calcium-binding loop, the structures of the mutant and wild-type protein are

remarkably similar considering the size of the deletion. The structures of SBT with and without the deletion superimpose with an rms difference between 261 Cα positions of 0.17 Å. The N-terminus of the wild-type protein lies beside the site A loop, furnishing one calcium coordination ligand, the side chain oxygen of Q2. In Δ75−83 SBT, the loop is gone, leaving residues 1−4 disordered, but the helix is uninterrupted and shows normal helical geometry over its entire length. The folding rate of Δ75−83 BPN′ is orders of magnitude faster than BPN′. In fact, the folding rate of Δ75−83 BPN′ (0.003 s^{-1}, 0.1 M KPi, pH 7.0, 25 °C) is as fast as the unimolecular folding of proSBT under similar conditions.[45,47] H−D exchange data indicate core protons of Δ75−83 BPN′ exchange at rates in the range of $year^{-1}$ in 0.1 M KPi, pH 7.0, 25 °C. Assuming the exchange of the core protons occurs only through global unfolding, then the ratio of unfolding and refolding rates would mean that $\Delta G_{unfolding} \sim 7$ kcal/mol in 0.1 M KPi, pH 7.0 and 25 °C (unpublished data). Since Δ75−83 BPN′ folds $>10^4$-times faster than BPN′ in 0.1 M KPi, pH 7.0, and the unfolding rates of the apo form of BPN′ is ∼ 10-times faster than Δ75−83 BPN′,[77,78] Δ75−83 BPN′ is about 7 kcal/mol greater at 25 °C than apo BPN′. This would indicate that apo BPN′ is near the margin of thermodynamic stability under these conditions.

Despite whether ALP or SBT is thermodynamically stable under a given set of conditions, its slow folding reaction is not due simply to instability of the folded state. If this were true the folding reaction could be thermodynamically driven by adding a ligand that binds tightly and preferentially to the native state. Thus, adding a high affinity inhibitor should result in spontaneous folding. This is not the case. For example, the streptomyces SBT inhibitor protein (SSI) has no detectable influence on the folding rate of SBT even though it binds to SBT more tightly than the prodomain does.[79]

A good example in which a prodomain does cure a thermodynamic problem is the folding reaction of neurophysin.[80] Neurophysin is the 95-amino acid carrier protein for the peptide hormones oxytocin and vasopressin. Isolated neurophysin folds to about 25% in the absence of the hormone ligand. The folding reaction can be driven to 100%, dependent on hormone concentration. The folding rate is not affected by the presence of ligand, however, only the unfolding rate. Hence, hormone−neurophysin complex formation is a thermodynamic sink and drives neurophysin folding to completion. Thus, for ALP and SBT thermodynamic instability, by itself, does not explain the kinetic isolation of folded and unfolded states.

C. How Does the Prodomain Catalyze Folding?

1. Mechanistic Model

To define mechanistically how the prodomain participates in folding, one would like to measure the rates of all steps in the bimolecular folding reaction:

$$P + U \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P{-}I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} P{-}N \underset{k_{on}}{\overset{k_{off}}{\rightleftharpoons}} N + P$$

where P is prodomain, U is unfolded, N is native, P−I is a collision complex of a partially folded intermediate and prodomain, and P−N is the complex of native and prodomain. One would also like to understand the structure of all populated species in this reaction. The approach is to try to describe what is known about the indivdual steps in the reaction and then use this information to understand the overall reaction.

2. Analysis of Prodomain Binding to the Folded Protease

$$P + N \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} P{-}N$$

The most straightforward step to analyze is the binding of the prodomain to the native protease. As mentioned earlier, the SBT prodomain is a competitive inhibitor of the active enzyme (K_i of 5.4×10^{-7}).[21] Binding of ALP to its prodomain is even tighter ($K_i \sim 10^{-10}$).[40] The high affinity of the prodomain for the native conformation suggested that its role in folding might be to stabilize nativelike structures in the transition state for folding.[23] What those nativelike features might be came to light when a high-resolution structure of the prodomain SBT complex was determined.[50]

In the complex, the prodomain folds into a single compact domain with an antiparallel four-stranded β-sheet and two three turn α-helices (Figure 4).[50,70] The β-sheet of the prodomain packs tightly against the two parallel surface α-helices of SBT (residues 104−116 and 133−144). Proresidues E69 and D71 form helix caps for the N-termini of the two SBT helices. In another charge dipole interaction, the carboxylate of E112 of SBT accepts H-bonds from the peptide nitrogens of proresidues 42, 43, and 44. The C-terminal residues 72−77 extend out from the central part of the prodomain and bind in a substrate-like manner along SBT's active site cleft. Residues Y77, A76, H75, and A74 of the prodomain occupy subsites S1 to S4 of SBT, respectively. Almost all of the prodomain's contacts with SBT are made with residues 100−144.

The complex of ALP with its prodomain has also been determined.[60] The 166-amino acid ALP proregion wraps around the C-terminal half (85−198) of mature ALP. The contact interface is more than 4000 $Å^2$, almost four-times that of SBT and its prodomain (Figure 5). Most of the contacts to ALP are made with the C-terminal half (amino acids 65−166) of the prodomain,[5] although both N- and C-terminal portions of the prodomain are required for folding of ALP.[81] Amino acids 160−166 of the prodomain are inserted like a substrate into the active site, reminiscent of the SBT complex. The edges of a three-stranded β-sheet in the C-terminal part of pro and a β-hairpin (118−130) of ALP abut and form and five stranded β-sheet in the complex.

For both ALP and SBT, the structure of the native protease is almost identical in free and complexed states. The shape complimentarity between the prodomains and the native protease is consistent with their properties as inhibitors. The more extensive contact interface of the larger ALP prodomain is also consistent with its higher affinity binding. Not sur-

prisingly, substitution and deletion mutations of prodomain amino acids involved in substrate-like interactions with the protease weaken binding to the native protease.[42,82,83] Also mutations at other regions of the binding interfaces weaken binding.[84,85]

There are some general features of the C-terminal half of the ALP prodomain which are topologically similar to the SBT prodomain.[5] Specifically, β-strands 2, 3, 4 and the α-helices connecting strands 1−3 and 2−4 in the SBT prodomain have structural counterparts in the ALP prodomain. Apart from the general similarity of the substrate-binding cleft interactions outlined above, however, there is little similarity in the interactions of proSBT and proALP at their binding interfaces with the mature proteases.

A final point of comparison is that the ALP prodomain has no similarity to its eukaryotic homologues trypsin and chymotrypsin.[1] The SBT prodomain on the other hand has high structural homology to the prohormone convertase prodomain, although little sequence similarity.[86,87] Surprisingly, the SBT prodomain also has the same basic fold as the prodomains of eukaryotic carboxypeptidases.[88,89] This is despite the lack of any structural similarity between the mature regions of SBT and the carboxypeptidases.

A feature of the isolated prodomains that influences binding properties is independent stability. The ALP prodomain is fully folded at 4 °C, independent of binding to ALP.[85] In contrast, the isolated SBT prodomain is essentially unfolded at all temperatures,[79] even though it is a compact globular structured in complex with SBT. This situation suggests that there is thermodynamic linkage between the SBT prodomain stability and its binding affinity. This linkage can be established by examining the effects of mutations which do not directly affect contacts with the protease but which affect the stability of prodomain. Although the stability of the isolated prodomain is low, it has been measured by titration with stabilizing cosolvents and found to have a $\Delta G_{\text{unfolding}} \sim -2$ kcal/mol at 25 °C.[90] This propensity to fold can be altered by mutations. For example, the mutation I30T in the hydrophobic core of the SBT prodomain weakens binding to SBT by ~100-fold.[84] (This prodomain mutation is reported to cause subtilisin to fold in vivo into an active, alternative conformation with very nativelike properties.[91,92] Detailed structural analysis of the alternative conformation and measurement of its spontaneous rate of conversion to the native form of subtilisin will be crucial to understanding this very surprising phenomenon.)

A number of mutations have been described that increase the independent stability of the SBT prodomain with concomitant increases in binding affinity to native SBT.[93−97] Sequentially introducing stabilizing mutations into the prodomain shifted the equilibrium for independent folding from ~97% unfolded to >98% folded. As the independent stability of the prodomain increased, the binding affinity for SBT increased concomitantly. The most stable prodomain mutant bound about ~30 times more tightly than wild-type prodomain.[98] Since these mutations were introduced in regions of the prodomain that did not directly contact with SBT, their effects on binding to SBT were linked to their stabilizing effects on the prodomain. The linked equilibria for prodomain folding and binding are

$$P_u \overset{K_f}{\rightleftharpoons} P + S \overset{K_P}{\rightleftharpoons} P{-}S$$

where K_f is the equilibrium constant for folding the prodomain and K_P is the association constant of folded prodomain for SBT. The observed binding constant is expected to be $K_{(P+Pu)} = [K_f/(1 + K_f)]K_P$. The data show that as the fraction of folded prodomain approaches one, the observed association constant approaches its maximum, $\sim 10^{10}$ M^{-1}.[93]

The kinetics of the prodomain binding to SBT also have been determined by measuring the rate of fluorescence changes that occurs upon binding. If the reaction is carried out with a 10-fold or greater excess of P, then one observes a pseudo-first-order kinetic process with a rate constant equal to $k_{on}[P] + k_{off}$. The k_{on} for binding the prodomain to folded SBT is $\sim 10^6$ M^{-1} s^{-1}.[99] This value is fairly insensitive to mutations that change the stability of the prodomain.

3. Analysis of Folding Steps

$$P + U \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P{-}I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} P{-}N$$

As described earlier, denatured ALP diluted into native conditions rapidly collapses into a ensemble of conformations with an average hydrodynamic radius about midway between the native and Gu-HCl denatured forms and with almost the same amount of regular secondary structure as native ALP.[41] These experiments are carried out such that all three native disulfides are intact in all ALP species. This kinetically trapped form is called "I" by Baker and Agard. For the sake of common nomenclature with SBT, it will be referred to as "U" in the following discussion of bimolecular folding kinetics. In the kinetic mechanism above, P is prodomain, P−I is an initial complex between prodomain and partially folded ALP, and P−N is the folded complex of prodomain and native ALP.

The kinetics of the conversion of U + P to P−N were determined as a function of [P] ranging from 5 to 100 μM.[85] When the concentration of prodomain is high relative to ALP, the kinetics of a single turnover of folding are pseudo first order. The observed rate of folding can be explained by a rapid equilibrium model where k_{-1} is large relative to k_2. In this model, the equilibrium constant, K_1, is equal to [P−I]/[U][P]. Since $[P{-}I]/[ALP_{total}] = K_1[P]/(1 + K_1[P])$, the observed rate of formation of P−N would be equal to

$$\{k_2K_1[P]/(1 + K_1[P])\} + k_{-2}$$

The concentration dependence of the rate follows a binding hyperbola, with $K_1 = 23$ μM. The maximum rate of the reaction (k_2) is 0.03 s^{-1}. It appears from the plots that k_{-2} is very small compared with k_2. The kinetics described above only pertain to the major (and faster) kinetic process.

Analysis of mutations of the prodomain indicate interactions at or near the active site are involved in stabilizing the transition state between the initial complex and the native complex.[42,85] In contrast, these interactions have less effect on the stability of the initial complex (P−I). For example, successive deletions of the last four amino acids of the prodomain decrease k_2 on average > 10-fold for each lost amino acid. Effects of K_1 are not detectable until three or more amino acids are deleted. Recall that these amino acids bind to ALP in a substrate-like manner.

Similar kinetic analysis of the bimolecular foldng of SBT is difficult because the reaction occurs at a rate of only ~0.2 s^{-1} M^{-1} of prodomain.[44] The slow time scale of in vitro folding makes detailed thermodynamic and kinetic analysis of the process problematic. For this reason, most of the mechanistic analysis of biomolecular SBT folding has been done on a mutant that lacks the calcium binding loop of site A and which folds more rapidly than wild-type SBT (see above). Further, the inactive (S221A) version of the loop-deleted variant was employed because unfolded forms of SBT and prodomain are very sensitive to proteolysis.

If the SBT mutant is denatured and quickly returned to native conditions (e.g., 0.1 M Kpi, pH 7.0), it has little regular secondary structure initially, unlike ALP. The disordered "U" state of mutant SBT will slowly refold to the native form in a single-exponential process (k_{fold} = 0.0015 s^{-1}) even in the absence of the prodomain.[79]

Using this mutant, catalyzed bimolecular folding was analyzed analogously to the ALP experiments described above but with somewhat different results. The rate of SBT folding increases linearly as a function of [P] for [P] ≤ 100 μM. The absence of curvature in the plot implies that the formation of the initial complex, P−I, is the limiting step in the reaction up to [P] = 100 μM. Recall that for ALP that the initial complex reached saturation by [P] ~ 20 μM. These results indicated that K_1 for SBT is at least 10 times weaker.

These results at first glance seem to suggest that the mechanisms of prodomain catalyzed folding of ALP and SBT differ in some fundamental way. Closer examination reveals, however, that the mechanisms are more similar than they first appear. Some of the differences can be explained by examination of the independent stabilities of the two prodomains. Recall that the SBT prodomain is largely unfolded at 25 °C ($\Delta G_{unfolding}$ = −2 kcal/mol). In comparison, the ALP prodomain is about 30−40% unfolded at 25 °C ($\Delta G_{unfolding}$ ~ 0.3 kcal/mol.[85] Further, the analysis of ALP folding usually is carried out at 4 °C, where its prodomain is essentially fully folded. The affect of prodomain stability on catalyzed folding can be seen by examining mutants of the SBT prodomain which stabilize its independent folding. These mutants were selected based on examination of the X-ray structure of the prodomain SBT complex and none of the mutations directly contact SBT. By sequentially introducing stabilizing mutations into the prodomain, the equilibrium for folding the prodomain was shifted from 97% unfolded to >98% folded. As the prodomain is stabilized, the folding reaction becomes faster and distinctly biphasic. Using a fully folded prodomain mutant the kinetics of folding become quite similar to ALP folding.[99] That is, at micromolar concentrations of prodomain, the complex P−I reaches saturation and the reaction is limited by the folding of P−I to P−N. In the case of SBT, it was possible to demonstrate that proline isomerization is rate limiting in this folding step.

There are 13 proline residues in SBT. All but one (Pro168) exist as trans isomers in the native structure.[74] The peptide bond between proline and its preceding amino acid (Xaa−Pro bonds) exist as a mixture of cis and trans isomers in solution unless structural constraints, such as in folded proteins, stabilize one of the two isomers. In the absence of ordered structure, the trans isomer is favored slightly over the cis isomer. The isomerization trans ⇔ cis is an intrinsically slow reaction with rates at 0.1 to 0.01 s^{-1} at 25 °C. Upon denaturation of SBT, the structural constraints are removed and the trans and cis isomers of all 13 prolyl peptide bonds gradually come to equilibrium in the unfolded state, resulting in 2^{13} different proline isomer combinations. To study the catalyzed folding rate of SBT in the absence of prolyl peptide bond isomerization, folding kinetics were measured after a short denaturation time.[100] The SBT mutant used in these studies can be denatured in acid in < 1 s. The rapid denaturation time minimizes the amount of prolyl peptide bond isomerization occurring during the time required to unfold. In the absence of proline isomerization, the observed rate of folding increases as the concentration of prodomain increases, and follows a linear relationship with no intermediate detectable in the course of the reaction. This occurs because the formation of P−N from P−I is faster than the formation of initial collision complex P−I from P and U. The mechanism of folding reaction for denatured SBT with native proline isomers is as follows:

$$P + U \underset{< 0.05\ s^{-1}}{\overset{9 \times 10^4\ M^{-1}\ s^{-1}}{\rightleftharpoons}} P{-}I \underset{<0.01\ s^{-1}}{\overset{> 3\ s^{-1}}{\rightleftharpoons}} P{-}N$$

where U is denatured SBT with all prolyl peptide bond isomers the same as in the native structure. According to the mechanism proposed above, the pseudo-first-order rate constant would reach its maximum when [P] is high enough that the value k_1[P] > (k_2 + k_{-2}). That point has not been reached at [P] = 20 μM. Thus, in the presence of an independently folded prodomain mutant, the rate of SBT folding into a complex becomes typical of in vitro folding rates for many small globular proteins. That is, independent of proline isomerization, folding is largely complete within a few seconds.[101]

In comparison, the rate of ALP folding plateaus at 0.03 s^{-1}, which is in the range of proline isomerization. ALP has three trans and one cis proline. It is not known how proline isomerization is involved in the kinetics of the P−I to P−N transition, however.

4. Model for Catalyzed Folding

The simplest model of catalyzed folding is one in which the prodomain accelerates folding by stabiliz-

ing a substructure on the folding pathway which does not form frequently in the absence of prodomain binding. Once the substructure forms, it acts as a folding nucleus with subsequent folding propagating into other regions. The nature of this substructure can be guessed at based the structures of the folded complexes. In the bimolecular SBT complex, the prodomain binds on SBT's two parallel surface α-helices and supplies caps to the N-termini of the two helices. In the intermediate P−I, the prodomain may stabilize these two helices relative to other unfolded states. Subsequent folding may propagate from this folding nucleus into N- and C-terminal regions of SBT. In the absence of the prodomain, these helices may not have sufficient independent stability to initiate folding very frequently. In ALP, the prodomain may stabilize the formation of a long β-hairpin (118−130) in the C-terminal half of ALP. The formation of this hairpin may be limiting in the absence of the prodomain.[60,102]

V. Why Are Prodomains Required for Folding?

For SBT and ALP folding, there is detailed knowledge of kinetics of the bimolecular reaction, detailed knowledge of the structure of the folded complex, and reasonable inferences about some structural features of the intermediate complex. It is also clear that there is a large kinetic barrier between unfolded and folded forms of the protease. The existence of such high barriers challenges many widely accepted ideas about protein folding, namely, thermodynamic determination of native structure and the sufficiency of thermodynamic stability to determine a pathway. The question remains as to why folding is so slow without the prodomain. With calcium-free SBT, the problem is not that the native state is unstable relative to the intermediate or the other unfolded states, as has been shown for ALP.[42,71] Uncatalyzed folding of calcium-free SBT is slow apparently because of low stability of intermediates.[82,90,99] An efficient folding pathway implies that productive intermediates are significantly more stable than the surrounding landscape of unfolded and misfolded conformations. Native folding intermediates of mature calcium-free SBT may have similar or lower stability than unfolded or misfolded states. This would result in a large entropic barrier to folding since most of the native structure would have to be formed before a significant free energy well is reached in conformational space.[103] This is surprising since a folding pathway seems to be a consequence of a stable native state for many proteins. Native state H−D exchange experiments for several proteins have shown that partially folded states exist with significantly lower energy than the globally unfolded state.[104−106] These partially folded states are important intermediates in the folding pathway of these proteins. In the case of a broad specificity proteinase such as SBT or ALP, however, there is likely a selection against stable intermediate states. Assuming that partially folded intermediates are good substrates for autolysis, the half-life of active protease would be determined by the equilibrium between the proteinase-resistant native state and the lowest energy, proteinase-labile intermediate.[3,47,102,107] Thus, to avoid autolysis, partially folded conformations must be of much higher energy than the native state so that excursions between the two states are rare. In both ALP[107] and SBT (Sari, Orban, and Bryan, in preparation), native state H−D exchange experiments show that amides protons in all elements of secondary structure exchange with solvent deuterons at a rate in the range of year^{-1}.

In ALP, there are 103 amide protons with protection factors $>10^4$ and 31 amide proton with protection factors $>10^9$. The most protected protons occur throughout both domains of ALP. In comparison, there are 131 protons in SBT with protection factors $\geq 10^5$. The 49 slowest exchangers have protection factors $\geq 10^9$. These strongly protected residues occur throughout the main structural elements of SBT, except for the short N-terminal α-helices, A and B, and the β-strands, 8 and 9 (Sari, Orban, and Bryan, in preparation). The large cooperative folding core may create a folding problem, however, because most of the tertiary structure must be acquired before a free energy well is encountered in conformational space. The enormous loss of conformational entropy before that energy well occurs would result in a large transition state barrier to folding.[96,107] In this model, the transition state of uncatalyzed folding is of high energy because of its low conformational entropy, not because unfavorable protein−protein or protein−solvent rearrangements are required in folding. The prodomain decreases the entropic barrier by pushing the transition state back toward a less folded form of the proteinase. Thus, much less conformational entropy is lost in the transition state. Once over the barrier, folding of P−I to P−N is rapid.

In order for the proteinase to be resistant to autoproteolysis, a large barrier must be created between the native state and any partially folded states after folding is completed. This energetic barrier is created by the proteolysis of the prodomain. Once folding to the complex, P−N, has been achieved, the lowest energy intermediate becomes the dissociated species, P + N. Dissociation of the complex results in proteolysis of the prodomain, leaving a large barrier between the native proteinase and other intermediate states.

Unfortunately, detailed understanding about the nature of the kinetic barrier between the unfolded and folded forms of the mature enzyme is still lacking and probably will not be clear until there is more information about the structures of the trapped unfolded state. There is circumstantial evidence, however, that the large kinetic barrier between the compact intermediate states of ALP and wild-type SBT may involve more than a high cooperativity.

The role of calcium site A in the kinetic stability of SBT is well-documented.[108] SBT with the calcium site A occupied unfolds 1000 times more slowly than the apo-form of SBT[48] and the activation energy for calcium dissociation from the folded state is 23 kcal/mol. The effect of calcium binding on the thermodynamics of SBT unfolding is quite unusual, however. The native state binds calcium with an affinity of $\sim 10^7$ M^{-1}. If calcium binding were exclusive to the native state then the contribution of calcium binding

to the unfolding free energy would be

$$\Delta G_{\text{binding}} = -RT \ln(1 + K_a [\text{Ca}])$$

For example, the free energy of unfolding would be increased by 5.5 kcal/mol in 1 mM calcium. The kinetically accessible intermediate form of SBT binds calcium with an affinity $\geq 10^6$ M^{-1}, however.[44] This means that the net contribution of calcium binding to the thermodynamics of unfolding would be minor ($\leq$1.4 kcal/mol). The fact that SBT unfolding is greatly slowed by calcium binding means that the folding rate must be slowed to an almost equivalent extent. This situation seems counterintuitive. If the calcium binding site in the intermediate is similar to that in the native form, then considerable native-like structure is preorganized in the intermediate. This is consistent with the observation that the intermediate has almost as much secondary structure content as the native state. Just the loss of chain entropy in the unfolded state from coordinating the calcium ligands should favor facile folding. Instead calcium binding seems to impose a high kinetic barrier to both folding and unfolding. Conceptually, calcium binding to site A appears to be like locking a dead bolt. A dead bolt is usually thrown with the door in the closed position, locking it closed. It could be thrown with the door opened, however, thereby locking the door open. Thus, a dead bolt alters the kinetics but not the thermodynamics of the door opening−closing reaction. One can speculate that the preorganization of the calcium site in the intermediate imposes topological restrictions on SBT which great hinder folding. Understanding exactly what these restrictions are will require more detailed structural characterization of the intermediate state.

This general model is supported by fact that deleting the calcium binding loop accelerates the folding reaction by $>10^4$-fold. The magnitude of the kinetic effect is unprecedented among mutations affecting folding rate. Further, the mechanism of pro-SBT maturation is consistent with the dead bolt model. The covalent attachment of the prodomain to SBT prevents the final folding of the calcium A-site region, until after the active site is sufficiently formed to cleave the prodomain from the mature enzyme, ensuring that the dead bolt is thrown with the protein folded.[50,57]

ALP has no metal binding sites, but, oddly, the three disulfide bonds could play an analogous role to the SBT calcium site A. An unstrained disulfide cross-link should stabilize a protein by decreasing the entropic cost of folding. The loss of conformational entropy in a polymer due to a cross-link has been estimated by calculating the probability that the ends of a polymer will simultaneously occur in the same volume element. According to statistical mechanics, the three disulfide bonds in ALP should destabilize the unfolded state by ~10 kcal/mol at 25 °C.[109,110] Since the life span of a protease depends on the kinetics rather than the thermodynamics of unfolding, however, disulfide bonds are a surprising feature. The effects of cross-links on the stability of the unfolded state would not generally be manifested in the activation energy of the unfolding reaction, because transition states for unfolding reactions usually are compact, with only slightly larger heat capacity than the native state. In bimolecular folding experiments on ALP, the disulfide bonds are present in both the I and N states of the protein and thus do not contribute to the thermodynamic measurements.[42,71] It is possible, however, that the cross-links impose topological constraints which slow both folding and unfolding as proposed for calcium binding in SBT. No work has been reported on folding of mature ALP from the reduced state to either support of refute this hypothesis. It should be noted that the fungal subtilase Proteinase K has no calcium site A but has two disulfide bonds (27−118 and 174−247).[12]

Whether the high kinetic stability of ALP and SBT is a result of biological imperative or a result of uncoupling evolution of the mature enzyme from the primary sequence of the mature enzyme is not known. It has been suggested that broad specificity proteases face peculiar challenges in surviving their own enzymatic activity and hence this may have necessitated the evolution of high kinetic stability.[102,107] For example, ALP, with broad specificity, has an absolute requirement for catalyzed folding, while the more sequence specific enkaryotic homologues trypsin and chymotrypsin are capable of in vitro folding without a prodomain. SBT and enkaryotic homologues, the prohormone convertases provide a counter example, however. PCs have a strong preference for cutting sequences of the form RX(R/K)R,[111,112] yet they appear to have a calcium site A[12] and a prodomain structurally homologous to the bacterial counterpart.[86,87] In vitro and in vivo folding of PCs appears to require the prodomain. Whatever its causes, the phenomenon of catalyzed folding provides opportunity to study some properties of protein design and stability not evident in the traditionally studied proteins.

VI. References

(1) Khan, A. R.; James, M. N. *Protein Sci.* **1998**, *7*, 815.
(2) Agard, D. A. *Science* **1993**, *260*, 1903.
(3) Baker, D.; Shiau, A. K.; Agard, D. A. *Curr. Opin. Cell Biol.* **1993**, *5*, 966.
(4) Baker, D.; Agard, D. *Biochemistry* **1994**, *33*, 7505.
(5) Baker, D. *Nat. Struct. Biol.* **1998**, *5*, 1021.
(6) Inouye, M. *Enzyme* **1991**, *45*, 314.
(7) Shinde, U.; Li, Y.; Chatterjee, S.; Inouye, M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6924.
(8) Shinde, U.; Inouye, M. *Trends Biochem. Sci.* **1993**, *18*, 442.
(9) Shinde, U.; Inouye, M. *Adv. Exp. Med. Biol.* **1996**, *379*, 147.
(10) Shinde, U.; Inouye, M. *Intramolecular Chaperones and Protein Folding*; R. G. Landes Company: Austin, 1995.
(11) James, M. N.; Delbaere, L. T.; Brayer, G. D. *Can. J. Biochem.* **1978**, *56*, 396.
(12) Siezen, R. J.; Leunissen, J. A. *Protein Sci.* **1997**, *6*, 501.
(13) Steiner, D.; Smeekens, S. P.; Ohagi, S.; Chan, S. J. *J. Mol. Biol.* **1992**, *267*, 23435.
(14) Wells, J. A.; Ferrari, E.; Henner, D. J.; Estell, D. A.; Chen, E. Y. *Nucleic Acids Res.* **1983**, *11*, 7911.
(15) Vasantha, N.; Thompson, L. D.; Rhodes, C.; Banner, C.; Nagle, J.; Filpula, D. *J. Bacteriol.* **1984**, *159*, 811.
(16) Wong, S.; Doi, R. *J. Biol. Chem.* **1986**, *261*, 10176.
(17) Power, S. D.; Adams, R. M.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3096.
(18) Ikemura, H.; Takagi, H.; Inouye, M. *J. Biol. Chem.* **1987**, *262*, 7859.
(19) Zhu, X.; Ohta, Y.; Jordan, F.; Inouye, M. *Nature* **1989**, *339*, 483.
(20) Ikemura, H.; Inouye, M. *J. Biol. Chem.* **1988**, *263*, 12959.
(21) Ohta, Y.; Hojo, H.; Aimoto, S.; Kobayashi, T.; Zhu, X.; Jordan, F.; Inouye, M. *Mol. Microbiol.* **1991**, *5*, 1507.
(22) Silen, J. L.; McGrath, C. N.; Smith, K. R.; Agard, D. A. *Gene* **1988**, *69*, 237.

(23) Silen, J. L.; Agard, D. A. *Nature* **1989**, *341*, 462.
(24) Winther, J. R.; Sorensen, P. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9330.
(25) Zhou, Y.; Lindberg, I. *J. Biol. Chem.* **1993**, *268*, 5615.
(26) Baier, K.; Nicklisch, S.; Maldener, I.; Lockau, W. *Eur. J. Biochem.* **1996**, *241*, 750.
(27) Fabre, E.; Nicaud, J. M.; Lopez, M. C.; Gaillardin, C. *J. Biol. Chem.* **1991**, *266*, 3782.
(28) Fabre, E.; Tharaud, C.; Gaillardin, C. *J. Biol. Chem.* **1992**, *267*, 15049.
(29) Chang, Y. C.; Kadokura, H.; Yoda, K.; Yamasaki, M. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 463.
(30) Baardsnes, J.; Sidhu, S.; MacLeod, A.; Elliott, J.; Morden, D.; Watson, J.; Borgford, T. *J. Bacteriol.* **1998**, *180*, 3241.
(31) van den Hazel, H. B.; Kielland-Brandt, M. C.; Winther, J. R. *J. Biol. Chem.* **1993**, *268*, 18002.
(32) Cawley, N. X.; Olsen, V.; Zhang, C. F.; Chen, H. C.; Tan, M.; Loh, Y. P. *J. Biol. Chem.* **1998**, *273*, 584.
(33) Fukuda, R.; Horiuchi, H.; Ohta, A.; Takagi, M. *J. Biol. Chem.* **1994**, *269*, 9556.
(34) Nirasawa, S.; Nakajima, Y.; Zhang, Z. Z.; Yoshida, M.; Hayashi, K. *Biochem. J.* **1999**, *341* (Pt 1), 25.
(35) Marie-Claire, C.; Ruffet, E.; Beaumont, A.; Roques, B. P. *J. Mol. Biol.* **1999**, *285*, 1911.
(36) Cao, J.; Hymowitz, M.; Conner, C.; Bahou, W. F.; Zucker, S. *J. Biol. Chem.* **2000**, *275*, 29648.
(37) Ventura, S.; Villegas, V.; Sterner, J.; Larson, J.; Vendrell, J.; Hershberger, C. L.; Aviles, F. X. *J. Biol. Chem.* **1999**, *274*, 19925.
(38) Wetmore, D. R.; Hardman, K. D. *Biochemistry* **1996**, *35*, 6549.
(39) Yamamoto, Y.; Watabe, S.; Kageyama, T.; Takahashi, S. Y. *Arch. Insect Biochem. Physiol.* **1999**, *42*, 167.
(40) Baker, D.; Silen, J. L.; Agard, D. A. *Proteins* **1992**, *12*, 339.
(41) Baker, D.; Sohl, J. L.; Agard, D. A. *Nature* **1992**, *356*, 263.
(42) Derman, A. I.; Agard, D. A. *Nat. Struct. Biol.* **2000**, *7*, 394.
(43) Chen, B.-L.; Baase, W. A.; Schellman, J. A. *Biochemistry* **1989**, *28*, 691.
(44) Eder, J.; Rheinnecker, M.; Fersht, A. R. *Biochemistry* **1993**, *32*, 18.
(45) Eder, J.; Rheinnecker, M.; Fersht, A. R. *J. Mol. Biol.* **1993**, *233*, 293.
(46) Carter, P.; Wells, J. A. *Nature* **1988**, *332*, 564.
(47) Bryan, P.; Alexander, P.; Strausberg, S.; Schwarz, F.; Wang, L.; Gilliland, G.; Gallagher, D. T. *Biochemistry* **1992**, *31*, 4937.
(48) Alexander, P. A.; Ruan, B.; Bryan, P. N. *Biochemistry* **2001**, *40*, 10634.
(49) Hu, Z.; Haghjoo, K.; Jordan, F. *J. Biol. Chem.* **1996**, *271*, 3375.
(50) Gallagher, T. D.; Gilliland, G.; Wang, L.; Bryan, P. *Structure* **1995**, *3*, 907.
(51) Hu, Z.; Zhu, X.; Jordan, F.; Inouye, M. *Biochemistry* **1994**, *33*, 562.
(52) Volkov, A.; Jordan, F. *J. Mol. Biol.* **1996**, *262*, 595.
(53) Li, Y.; Inouye, M. *J. Biol. Chem.* **1994**, *269*, 4169.
(54) Li, Y.; Inouye, M. *J. Mol. Biol.* **1996**, *262*, 591.
(55) Shinde, U.; Inouye, M. *J. Mol. Biol.* **1995**, *247*, 390.
(56) Shinde, U.; Inouye, M. *J. Mol. Biol.* **1995**, *252*, 25.
(57) Yabuta, Y.; Subbian, E.; Takagi, H.; Shinde, U.; Inouye, M. *J. Biochem. (Tokyo)* **2002**, *131*, 31.
(58) Yabuta, Y.; Takagi, H.; Inouye, M.; Shinde, U. *J. Biol. Chem.* **2001**, *276*, 44427.
(59) Anderson, D. E.; Peters, R. J.; Wilk, B.; Agard, D. A. *Biochemistry* **1999**, *38*, 4728.
(60) Sauter, N. K.; Mau, T.; Rader, S. D.; Agard, D. A. *Nat. Struct. Biol.* **1998**, *5*, 945.
(61) Plaxco, K. W.; Simons, K. T.; Ruczinski, I.; Baker, D. *Biochemistry* **2000**, *39*, 11177.
(62) Baker, D. *Nature* **2000**, *405*, 39.
(63) Goldenberg, D. P. *Nat. Struct. Biol.* **1999**, *6*, 987.
(64) Roder, H.; Colon, W. *Curr. Opin. Struct. Biol.* **1997**, *7*, 15.
(65) Baldwin, R. L. *Fold Des.* **1996**, *1*, R1.
(66) Baldwin, R. L.; Rose, G. D. *Trends Biochem. Sci.* **1999**, *24*, 77.
(67) Baldwin, R. L.; Rose, G. D. *Trends Biochem. Sci.* **1999**, *24*, 26.
(68) Englander, S. W. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213.
(69) Sosnick, T. R.; Mayne, L.; Hiller, R.; Englander, S. W. *Nat. Struct. Biol.* **1994**, *1*, 149.
(70) Jain, S. C.; Shinde, U.; Li, Y.; Inouye, M.; Berman, H. M. *J. Mol. Biol.* **1998**, *284*, 137.
(71) Sohl, J. L.; Jaswal, S. S.; Agard, D. A. *Nature* **1998**, *395*, 817.
(72) Ottesen, M.; Svendsen, I. *Methods Enzymol.* **1970**, *19*, 199.
(73) Hayashi, T.; Matsubara, M.; Nohara, D.; Kojima, S.; Miura, K.; Sakai, T. *FEBS Lett.* **1994**, *350*, 109.
(74) McPhalen, C. A.; James, M. N. G. *Biochemistry* **1988**, *27*, 6582.
(75) Gallagher, T. D.; Bryan, P.; Gilliland, G. *Proteins: Struct. Funct. Genet.* **1993**, *16*, 205.
(76) Almog, O.; Gallagher, T.; Tordova, M.; Hoskins, J.; Bryan, P.; Gilliland, G. L. *Proteins* **1998**, *31*, 21.
(77) Bryan, P. N. *Biochim. Biophys. Acta* **2000**, *1543*, 203.
(78) Alexander, P. A.; Ruan, B.; Strausberg, S. L.; Bryan, P. N. *Biochemistry* **2001**, *40*, 10640.
(79) Strausberg, S.; Alexander, P.; Wang, L.; Schwarz, F.; Bryan, P. *Biochemistry* **1993**, *32*, 8112.
(80) Deeb, R.; Breslow, E. *Biochemistry* **1996**, *35*, 864.
(81) Cunningham, E. L.; Mau, T.; Truhlar, S. M.; Agard, D. A. *Biochemistry* **2002**, *41*, 8860.
(82) Wang, L.; Ruvinov, S.; Strausberg, S.; Gallagher, T. D.; Gilliland, G.; Bryan, P. *Biochemistry* **1995**, 15.
(83) Sohl, J. L.; Shiau, A. K.; Rader, S. D.; Wilk, B. J.; Agard, D. A. *Biochemistry* **1997**, *36*, 3894.
(84) Li, Y.; Hu, Z.; Jordan, F.; Inouye, M. *J. Biol. Chem.* **1995**, *270*, 25127.
(85) Peters, R. J.; Shiau, A. K.; Sohl, J. L.; Anderson, D. E.; Tang, G.; Silen, J. L.; Agard, D. A. *Biochemistry* **1998**, *37*, 12058.
(86) Tangrea, M. A.; Alexander, P.; Bryan, P. N.; Eisenstein, E.; Toedt, J.; Orban, J. *Biochemistry* **2001**, *40*, 5488.
(87) Tangrea, M. A.; Bryan, P. N.; Sari, N.; Orban, J. *J. Mol. Biol.* **2002**, *320*, 801.
(88) Guasch, A.; Coll, M.; Aviles, F. X.; Huber, R. *J. Mol. Biol.* **1992**, *224*, 141.
(89) Coll, M.; Guasch, A.; Aviles, F. X.; Huber, R. *EMBO J.* **1991**, *10*, 1.
(90) Bryan, P.; Wang, L.; Hoskins, J.; Ruvinov, S.; Strausberg, S.; Alexander, P.; Almog, O.; Gilliland, G.; Gallagher, T. D. *Biochemistry* **1995**, *34*, 10310.
(91) Shinde, U. P.; Liu, J. J.; Inouye, M. *Nature* **1997**, *389*, 520.
(92) Shinde, U.; Fu, X.; Inouye, M. *J. Biol. Chem.* **1999**, *274*, 15615.
(93) Ruvinov, S.; Wang, L.; Ruan, B.; Almog, O.; Gilliland, G.; Eisenstein, E.; Bryan, P. *Biochemistry* **1997**, *36*, 10414.
(94) Kojima, S.; Minagawa, T.; Miura, K. *FEBS Lett.* **1997**, *411*, 128.
(95) Kojima, S.; Minagawa, T.; Miura, K. *J. Mol. Biol.* **1998**, *277*, 1007.
(96) Ruan, B.; Hoskins, J.; Bryan, P. N. *Biochemistry* **1999**, *38*, 8562.
(97) Kojima, S.; Yanai, H.; Miura, K. *J. Biochem. (Tokyo)* **2001**, *130*, 471.
(98) Ruan, B.; Hoskins, J.; Wang, L.; Bryan, P. N. *Protein Sci.* **1998**, *7*, 2345.
(99) Wang, L.; Ruan, B.; Ruvinov, S.; Bryan, P. N. *Biochemistry* **1998**, *37*, 3165.
(100) Ruan, B. Ph.D. Dissertation, University of Maryland, College Park, 1998.
(101) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953.
(102) Cunningham, E. L.; Jaswal, S. S.; Sohl, J. L.; Agard, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11008.
(103) Levinthal, C. *J. Chim. Phys.* **1968**, *65*, 44.
(104) Bai, Y.; Sosnick, T. R.; Mayne, L.; Englander, S. W. *Science* **1995**, *269*, 192.
(105) Chamberlain, A. K.; Handel, T. M.; Marqusee, S. *Nat. Struct. Biol.* **1996**, *3*, 782.
(106) Chung, E. W.; Nettleton, E. J.; Morgan, C. J.; Gross, M.; Miranker, A.; Radford, S. E.; Dobson, C. M.; Robinson, C. V. *Protein Sci.* **1997**, *6*, 1316.
(107) Jaswal, S. S.; Sohl, J. L.; Davis, J. H.; Agard, D. A. *Nature* **2002**, *415*, 343.
(108) Voordouw, G.; Milo, C.; Roche, R. S. *Biochemistry* **1976**, *15*, 3716.
(109) Poland, D. C.; Scheraga, H. A. *Biopolymers* **1965**, *3*, 379.
(110) Pace, C. N.; Grimsley, G. R.; Thomson, J. A.; Barnett, B. J. *J. Biol. Chem.* **1988**, *263*, 11820.
(111) Bresnahan, P. A.; Leduc, R.; Thomas, L.; Thorner, J.; Gibson, H. L.; Brake, A. J.; Barr, P. J.; Thomas, G. *J. Cell Biol.* **1990**, *111*, 2851.
(112) Hosaka, M.; Nagahama, M.; Kim, W. S.; Watanabe, T.; Hatsuzawa, K.; Ikemizu, J.; Murakami, K.; Nakayama, K. *J. Biol. Chem.* **1991**, *266*, 12127.

CR010190B