Engineering the Independent Folding of the Subtilisin BPN′ Prodomain: Analysis of Two-State Folding versus Protein Stability†

Sergei Ruvinov,‡ Lan Wang,‡ Biao Ruan,‡ Orna Almog,§ Gary L. Gilliland,§ Edward Eisenstein,§ and Philip N. Bryan*,‡

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, and The National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received February 20, 1997; Revised Manuscript Received June 18, 1997⊗

ABSTRACT: In complex with subtilisin BPN′, the 77 amino acid prodomain folds into a stable compact structure comprising a four-stranded antiparallel β-sheet and two three-turn α-helices. When isolated from subtilisin, the prodomain is 97% unfolded even under optimal folding conditions. Traditionally, to study stable proteins, denaturing cosolvents or temperatures are used to shift the equilibrium from folded to unfolded. Here we manipulate the folding equilibrium of the unstable prodomain by introducing stabilizing mutations generated by design. By sequentially introducing three stabilizing mutations into the prodomain we are able to shift the equilibrium for independent folding from 97% unfolded to 65% folded. Spectroscopic and thermodynamic analysis of the folding reaction was carried out to assess the effect of stability on two-state behavior and the denatured state. The denatured states of single and combination mutants are not discernably different in spite of a range of $\Delta G_{\text{unfolding}}$ from −2.1 to 0.4 kcal/mol. Conclusions about the nature of the denatured state of the prodomain are based on CD spectral data and calorimetric data. Two state folding is observed for a combination mutant of marginal stability ($\Delta G = 0$). Evidence for its two-state folding is based on the observed additivity of individual mutations to the overall $\Delta G_{\text{unfolding}}$ and the conformity of $\Delta G_{\text{unfolding}}$ vs T to two-state assumptions as embodied in the Gibbs−Helmholz equation. We believe our success in stabilizing the two-state folding reaction of the prodomain originates from the selection of mutations with improved ability to fold subtilisin rather than selection for increase in secondary structure content. The fact that a small number of mutations can stabilize the independent folding of the prodomain implies that most of the folding information already exists in the wild-type amino acid sequence in spite of the fact that the unfolded state predominates.

In order to understand protein stability and the folding process, it is essential to understand the equilibrium that exists between the native state and the denatured state under conditions which strongly favor folding (Dill & Shortle, 1991). It is known that certain denatured states are preferentially populated under native conditions and that these denatured stated are critical in guiding the folding process (Bai & Englander, 1996; Buckler et al., 1995; Orban et al., 1995). Paradoxically, increasing stability of particular denatured states can undermine two-state unfolding behavior and decrease the overall $\Delta G_{\text{unfolding}}$. This paradox has important implications for protein design, because in order to achieve a stable, unique protein structure a free energy barrier must separate native and denatured forms. The goal of this paper is to study the origin of the free energy barrier that results in two-state unfolding behavior.

The folding equilibrium is difficult to study under native conditions because the denatured state is infrequently populated. In order to examine the folding equilibrium spectroscopically and calorimetrically under conditions that strongly favor folding, we have used the 77 amino acid prodomain of subtilisin BPN′. Biosynthesis of subtilisin is dependent on the prodomain, which is eventually cleaved from the N-terminus of subtilisin to create the 275 amino acid mature form of the enzyme (Ikemura et al., 1987). The prodomain has been shown to promote the in vitro folding of subilisin in a bimolecular folding reaction resulting in a tight complex between the two (Strausberg et al., 1993). The isolated prodomain is 97% unfolded even under optimal folding conditions (0.1 M KP_i, pH 7.0, 20 °C) (Bryan et al., 1995). When complexed with subtilisin, the prodomain folds into a stable compact structure comprising a four-stranded antiparallel β-sheet and two three-turn α-helices (Gallagher et al., 1995).

Traditionally, to study stable proteins, denaturing cosolvents or temperatures are used to shift the equilibrium from folded to unfolded. Here we manipulate the folding equilibrium of the unstable prodomain by introducing stabilizing mutations generated by design. By sequentially introducing three stabilizing mutations we are able to shift the apparent equilibrium for folding from 97% unfolded to 65% folded. Mutants in this range are useful because both the native and denatured states are populated under conditions that strongly favor folding. Spectroscopic and thermodynamic analysis of the folding reaction was carried out to assess the effect of stability on two-state behavior and the denatured state. This analysis reveals that two-state behavior can be manifest even in unstable proteins and that increasing stability does not necessarily result in observable changes in the spectral and thermodynamic properties of the denatured state.

† This work was supported by NIH Grant GM42560.
* Corresponding author.
‡ University of Maryland Biotechnology Institute.
§ The National Institute of Standards and Technology.
⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1997.

S0006-2960(97)00395-4 CCC: $14.00

MATERIALS AND METHODS

Cloning and Expression of the Prodomain of Subtilisin. The prodomain region of the subtilisin BPN′ gene was subcloned using the polymerase chain reaction in an Eppendorf MicroCycler according to conditions outlined in the GeneAmp PCR reagent kit. Oligonucleotides were synthesized which amplified the coding sequence for the 77 prodomain. Digestion of the amplified product with the appropriate restriction enzymes allowed a precise excision of the DNA to the ATG initiation codon of the *Escherichia coli* expression plasmid pT7-7, a generous gift from Stanley Tabor of Harvard Medical School. pT7-7 contains the bacteriophage T7 RNA polymerase promoter ϕ10 and the translation start site for gene 10. The resulting plasmid, denoted pP1, was used to transform the *E. coli* production strain BL21(DE3). Fermentations were run in a 1.5 L BioFlo Model fermenter (Strausberg et al., 1993). The expression strategy for the prodomain is identical to that described in detail for high-level production of the 56 amino acid B-domain of protein G (Alexander et al., 1992).

Mutagenesis of the cloned prodomain gene was performed according to the oligonucleotide-directed *in vitro* mutagenesis system, version 2 (Amersham International plc). Single-strand plasmid DNA was sequenced with Sequenase (U.S. Biochemical Corp.).

Kinetics of Catalyzed Subtilisin Folding. Refolding of subtilisin was followed by monitoring changes in tryptophan fluorescence (excitation $\lambda = 300$ nm, emission $\lambda = 345$ nm) using a Spex FluoroMax spectrofluorometer for manual mixing experiments and a KinTek stopped-flow apparatus, Model SF2001, for rapid kinetic measurements as described (Strausberg et al., 1993). The subtilisin, Sbt-15,[1] used in this study was an engineered version that was catalytically impaired by mutating the active-site serine 221 to cysteine and was modified to eliminate the high-affinity calcium site A. Details of the construction and characterization of this mutant are described in Strausberg et al. (1993).

Circular Dichroism. Circular dichroism (CD) measurements were performed with a Jasco spectropolarimeter, Model J-720, using water-jacketed quartz cells with path lengths 1 cm, 1 mm, or 0.1 mm depending on prodomain concentrations, which ranged from 5 to 800 μM. Temperature control was provided by a Neslab RTE-110 circulating water bath interfaced with a MTP-6 temperature programmer. Far-UV wavelength scans were recorded at 25 °C from 250 to 190 nm. In many cases, however, due to high solution absorbance the lower wavelength of determination was limited to 200 nm or even to 205 nm. Typically, the averages for five CD spectra were presented. The ellipticity results were expressed as mean residue ellipticity, $[\theta]$, in degrees centimeter^2 decimole^{-1}.

Temperature-induced unfolding of the prodomain was performed in the temperature range between −5 and 90 °C in 1 cm, 1 mm, and 0.1 mm cuvettes. Ellipticities at 222 nm were continuously monitored at a scanning rate of 1 °C/min. Reversibility of the denaturations was confirmed by comparing the CD spectra at 25 °C before and after cooling or heating. In all cases the reversibilities were essentially complete.

Analytical Ultracentrifugation. Sedimentation equilibrium experiments using wild-type and mutant prodomains were performed with a Beckman XL-A Optima analytical ultracentrifuge equipped with a four-hole, An-55 rotor. All experiments were performed at 25 °C at rotor speed of 40 000 rpm. The concentration distribution of prodomains at sedimentation equilibrium was acquired in a step mode as an average of 25 measurements of absorbance at 275 nm at each radial position, with nominal spacing of 0.001 cm between radial positions. Equilibrium was verified by comparing consecutive scans separated by time intervals from 40 min to 2 h. Lyophilized prodomain samples were dissolved in the corresponding buffer, as indicated in the Results section. Partial specific volumes (cubic centimeters per gram) of 0.734 for the wild type, 0.735 for the 40L-57E mutant, and 0.740 for mut-3 prodomain, respectively, were calculated on the basis of their amino acid compositions. In a typical experiment, equilibrium sedimentation data were collected after 16−24 h for three initial prodomain concentrations varying between ∼35 and ∼110 μM, ranging in A_{275} from 0.2 to 0.6.

Equilibrium constants for the dissociation of dimeric prodomain to monomers were estimated by using a modified version of IGOR software (Wavemetrics, Lake Oswego, OR) running on a Macintosh computer (Brooks et al., 1993). Data were first analyzed according to a possible monomer−dimer, monomer−dimer−trimer, and monomer−dimer−tetramer equilibrium. However, for the higher association models, the highest association constant value was found to be at least 100-fold lower than the monomer−dimer association constant values. Moreover, since the monomer−dimer association constants values were found to be similar for the three different prodomain concentrations examined, the data from the three concentration distributions were analyzed globally in terms of a monomer−dimer model.

The pH dependence of dimerization was performed in a three-component buffer mixture giving virtually constant ionic strength $I = 0.1$ M. The concentrations of individual components were the following: acetic acid 0.05 M, MES 0.05 M, and Tris 0.1 M. This system allows one to work in the pH range from 3.5 to 9.2 (Ellis & Morrison, 1982).

Differential Scanning Calorimetry. Differential scanning calorimetry (DSC) measurements were performed with a Hart 7707 DSC heat conduction scanning microcalorimeter interfaced with an IBM personal computer (Alexander et al., 1992; Schwarz & Kirchhoff, 1988). The temperature was increased from −10 to 95 °C at a scan rate of 30 °C/h. The solution mass of all protein and control solutions was near 0.70 g/ampoule. The scans were done in 100 mM NaOAc, pH 5.0. Samples to be scanned were prepared by rehydrating lyophilized protein in the appropriate buffer and dialyzing against the same buffer. The concentration of the dialyzed protein was determined by UV absorbance using 1 mg/mL = $A_{275\text{nm}}$ of 0.67. The number of nanomoles of protein ranged from 400 to 800, corresponding to 5−10 mg/mL.

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) experiments were performed with a

[1] Abbreviations: A_{280}, absorbance at 280 nm; CD, circular dichroism; DSC, differential scanning calorimetry: Gu-HCl, guanidine hydrochloride; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; ITC, isothermal titration calorimetry; K_a, association constant for prodomain binding; Mes, 2-*N*-morpholinoethanesulfonic acid; mre, mean residue ellipticity; mut-3, prodomain with the mutations Q32E, Q40L, and K57E; [P], prodomain concentration; P_i, phosphate; pro-wt, wild-type prodomain of subtilisin BPN′; [S], subtilisin concentration; Sbt-15, subtilisin BPN′ with the mutations Δ75−83, N218S, and S221C; TFE, trifluoroethanol; Tris, tris(hydroxymethyl)aminomethane.

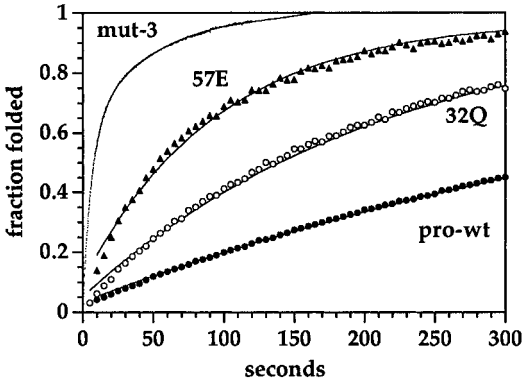

FIGURE 1: Effect of prodomain mutations on the kinetics of subtilisin folding. Denatured Sbt-15 (1 μM) and 5 μM prodomain were mixed in 5 mM KP_i and 30 mM Tris, pH 7.5, at 25 °C. The reaction was followed by the increase in tryptophan fluorescence at 345 nm, which occurs upon folding of subtilisin into the prodomain−subtilisin complex. The data for pro-wt, 57E, and 32Q are fit to single, exponential equations to determine pseudo-first-order rate constants for folding (solid lines). Sbt-15 folding with mut-3 is faster than with single mutants but is not a single-exponential phase. No fit is shown for mut-3.

Microcal Omega titration calorimeter as described (Strausberg et al., 1993; Wang et al., 1995; Wiseman et al., 1989). Binding constants were determined by fitting the heat per injection, $dQ/d[P]_{total}$, to

$$dQ/d[P]_{total} = \Delta H[^1/_2 + [1 - (1 + r)/2 - X_r/2]/ [X_r - 2X_r(1 - r) + 1 + r^2]^{1/2}] \quad (1)$$

where $1/r = [S]_{total}K_a$ and $X_r = [P]_{total}/[S]_{total}$ (Wiseman et al., 1989). $[S]_{total}$ was 25 μM in all experiments. Each binding constant and enthalpy determination was based on at least two titration runs at each temperature. Titration runs were performed until no heat was produced by further addition of prodomain. The subtilisin used in these experiments is a mutant with the catalytic serine 221 converted to alanine. The mutant denoted Sbt-110 is also engineered to be highly stable (Strausberg et al., 1995; Strausberg, unpublished results). Sbt-110 remains fully native throughout the temperature range of 11−45 °C at pH 5.0.

RESULTS

Generation of Stable Folds. On the basis of the X-ray crystal structure of the prodomain−subtilisin complex (Bryan et al., 1995; Gallagher et al., 1995), we have designed mutations that increase the independent stability of the prodomain. The prodomain folds into a single compact domain with a four-stranded antiparallel β-sheet and two three-turn α-helices when complexed with subtilisin. Stabilizing mutations were identified in three different areas of the structure of prodomain: α-helix 23−32 (E32Q),[2] β-strands 35−51 (Q40L), and α-helix 53−61 (K57E). These amino acid positions were selected because they do not contact subtilisin in the complex and because they appear to be in regions of the structure that are not well packed in the wild-type prodomain.

The mutations in the two α-helices were introduced to improve upon what appeared to be unfavorable electrostatic interactions in the folded state. In the first α-helix an ϵ-oxygen of E32 is separated by only 4 Å from a δ-oxygen of D28. The E32Q mutation was introduced to reduce electrostatic repulsion. In the second helix, three surface lysines (K54, K57, and K61) are located on consecutive α-helical turns and thus the charged ϵ-amino groups are separated from one another by only 6−7 Å. To reduce electrostatic repulsion, the K57E mutation was introduced. The Q40L mutation was introduced to improve hydrophobic packing between β-strand 35−51 and α-helix 23−32.

Two double mutants and a triple mutant were also constructed. The mutant combining Q40L and K57E into a single molecule is denoted 40L-57E. The mutant combining E32Q and K57E into a single molecule is denoted 32Q-57E. A mutant combining E32Q, Q40L, and K57E into a single molecule is denoted mut-3. The structure of mut-3 in complex with subtilisin was determined by X-ray crystallography (Almog et al., manuscript in preparation). No changes in α-carbon backbone are observed to result from the mutations on comparison of mut-3 to pro-wt (Gallagher et al., 1995). Potentially the K57E mutation could have resulted in an intrahelical salt bridge with either K54 or K61. It appears, however, that the ϵ-oxygen of E57 reamains centered ∼7 Å from ϵ-amino groups K54 and K61. Substitution of the partially buried Q40 with leucine resulted in hydrophobic contacts of its Cδ groups with Cβ of A23 (3.3 Å) and Cγ1 of V37 (3.5 Å).

Effect of Prodomain Mutations on the Kinetics of Subtilisin Folding. Our initial screen for determining whether a mutation stabilizes the native state was to measure the rate at which that prodomain mutant catalyzes subtilisin folding. The bimolecular folding reaction of subtilisin and the prodomain can be followed by an increase in tryptophan fluorescence of 1.7-fold due to changes in the environments of the three tryptophans in subtilisin upon its folding and binding of the prodomain. The prodomain does not contain tryptophan residues and thus has no intrinsic fluorescence at 345 nm upon excitation at 300 nm. Therefore, fluorescence increases observed at 345 nm are due solely to the conversion of unfolded subtilisin to the folded complex.

The folding reaction was followed in 30 mM Tris and 5 mM KP_i, pH 7.5, using $[S_u] = 1$ μM and $[P] = 5$ μM for each of the mutant prodomains. Under these conditions a single cycle of subtilisin folding is measured. The product is a 1:1 complex of Sbt-15 and prodomain. The reactions with the pro-wt and single mutants were pseudo-first-order and the folding curves were fit with a single-exponential equation to determine a $k_{observed}$ (Figure 1). All of the single mutations increased the rate of subtilisin folding. The reactions with the 40L-57E and mut-3 were the fastest but the reactions were not first-order. The kinetics of catalyzed folding are analyzed in detail by Wang et al. (manuscript in preparation).

Since none of the mutations in the prodomain directly contacts subtilisin in the complex, their effects on the folding of subtilisin are linked to whether or not they stabilize a conformation of the prodomain that promotes subtilisin folding. This screen allowed us to detect beneficial mutations with small effects on the folding equilibrium. For example, the 32Q mutant folds subtilisin 2.4 times faster than does pro-wt. The effect of the 32Q mutation is difficult to detect by CD, however, because the fraction of folded prodomain increases by only a few percent. The screen also

[2] A shorthand for denoting amino acid substitutions employs the single-letter amino acid code as follows: E32Q denotes the change of glutamate 32 to glutamine.

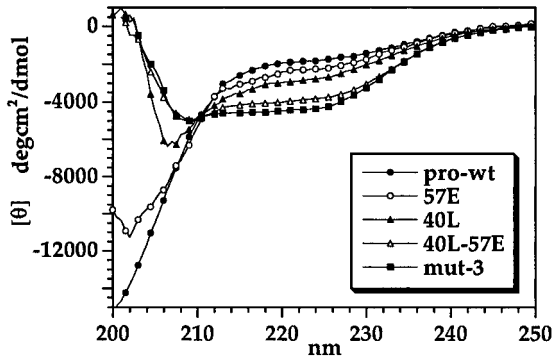

FIGURE 2: CD profiles of the prodomain mutants. Mean residue ellipticity (deg cm^2 dmol^{-1}) is plotted versus wavelength for pro-wt, 57E, 40L, 40L-57E, and mut-3. Spectra were measured in 0.01 M KP_i, pH 7.0, using a 1 cm cylindrical cuvette at 25 °C with [P] = 10 μM.

enabled us to distinguish mutations which increase compact but nonnative structure from those that stabilize the native conformation.

CD Profiles of the Prodomain Mutants. Circular dichroic spectra in the far uv region were used to determine the apparent equilibrium constants for folding the mutant prodomains. Figure 2 shows the CD spectrum of pro-wt, which is typical of a largely random coil structure with a minimum ellipticity at 198 nm (Goodman & Kim, 1989; Merutka et al., 1991). The mean residue ellipticity at 222 nm is −2000 deg cm^2 dmol^{-1}. CD spectra of isolated pro mutants 57E, 40L, 40L-57E, and mut-3 are shown in Figure 2. The mutants with the highest percentage of regular secondary structure are the combination mutants 40L-57E and mut-3. The mean residue ellipticity of 40L-57E at 222 nm is −3900 deg cm^2 dmol^{-1} and for mut-3 is −4400 deg cm^2 dmol^{-1}. If increasing amounts of the stabilizing cosolvent ethylene glycol are added to the wild type or any of the pro mutants, their ellipticity at 222 nm converges on −5100 deg cm^2 dmol^{-1} as the ethylene glycol concentration approaches 50%. At 50% EG, the CD spectrum of the prodomain resembles the difference spectrum of the prodomain−subtilisin complex minus subtilisin (Bryan et al., 1995). The difference in ellipticity at 222 nm of the prodomain−subtilisin complex minus subtilisin is equal to −5100 deg cm^2 dmol^{-1}. This value is taken to be the ellipticity of the native prodomain at 25 °C.

Dimerization of the Prodomain: Sedimentation Equilibrium Studies. The ellipticity of pro-wt is independent of concentration. The ellipticity at 222 nm of 40L-57E and mut-3 increases noticably in the concentration range between 5 and 800 μM, however. Concentration dependence of ellipticity is often an indication of association and/or aggregation. The sedimentation equilibrium method is widely used for accurate determination of molecular homogeneity and associations of proteins. We applied this method to the pro-wt and mutants in order to quantitatively characterize the tendency of the prodomain to associate and to evaluate the pH and ionic strength dependences of its association propensity.

An initial demonstration of reversible equilibrium was obtained by centrifuging several different initial concentrations of the prodomain to the state of equilibrium at a single rotor speed and obtaining the same equilibrium constant for all distributions. Typically, three concentrations ranging from 0.2 to 1 mg/mL of the prodomain solutions were centrifuged to equilibrium. Pro-wt in 100 mM potassium phosphate buffer, pH 7.0, behaves as a homogeneous population of monomers with average molecular mass of 8500 Da, which is close to its analytical molecular mass. Data for 40L-57E and mut-3 fit well to a monomer−dimer equilibrium. Differences between the fitted curves and the data were small (root mean square deviations ∼0.005) and were distributed almost randomly around zero. The three initial concentrations of the prodomain yielded a dissociation constant $K_d = 120 \pm 22$ μM for 40L-57E. The K_d for mut-3 was 70 ± 35 μM in 100 mM KP_i buffer, pH 7.0.

Table 1: Dissociation Constants for Dimerization of the 40L-57E Prodomain vs pH and Ionic Strength

	I^b						
pH[a]	0.02	0.05	0.1	0.2	0.25	0.4	0.5
5.0			8400				
5.4			3713				
5.7			2840				
6.0			1269				
6.4			423				
7.0	15	34	69	185	248	372	486
7.4			73				
7.9			78				
8.3			46				
8.8			54				
9.2			60				

[a] Measurements were made at 25 °C in a three-component buffer system giving a virtually constant ionic strength of 0.1 M. The composition of the buffer system was 0.05 M acetic acid, 0.05 M MES, and 0.1 M Tris. [b] Ionic strength was varied by adding NaCl to 20 mM HEPES, pH 7.0. Dissociation constants are given in units of micromolar.

The influence of salt concentration and pH on the monomer−dimer equilibrium was investigated using 40L-57E (Table 1). The values of the dissociation constant, K_d, increase progressively upon increasing the salt concentration, from $K_d = 15$ μM at 20 mM NaCl to $K_d = 486$ μM at 500 mM NaCl, indicating that the dimers of the mutant prodomain dissociate parallel to the increase of salt concentration. The values of K_d increase progressively upon decreasing the pH. In 100 mM NaOAc, pH 5.0, 40L-57E exhibits no significant dimerization.

The present results demonstrate that the structured mutant prodomains exist at neutral pH in a monomer−dimer equilibrium that is highly sensitive to both pH and salt concentration. Only the more stable prodomains, 40L-57E and mut-3, displayed any tendency to dimerize.

Effects of Mutations of the Equilibrium Constant for Prodomain Folding. In 10 mM KP_i, pH 7.0, the fraction of folded material is dependent on prodomain concentration because of the tendency of the folded state to dimerize. In order to determine an intrinsic folding equilibrium constant (K_f) for each mutant, the CD data were analyzed in terms of a three-state system

$$P_u \overset{K_f}{\rightleftharpoons} P_f \overset{K_a}{\rightleftharpoons} {}^1/_2[P_f]_2$$

where K_f is the equilibrium constant for folding, K_a is the equilibrium constant for dimerization (association constant), and P_u, P_f, and $[P_f]_2$ are molar concentrations of unfolded monomer, folded monomer, and folded dimer, respectively. The amount of folded prodomain (folded monomer + folded dimer) is determined from CD experiments (Figure 2). This analysis requires that the ellipticity of the prodomain in the

Table 2: Effects of Prodomain Mutations on Its Folding Equilibrium[a]

	P_f/P_u	$\Delta G_{folding}$ (kcal/mol)	$\Delta\Delta G$ (kcal/mol)	ΔG_{sum} (kcal/mol)
wt	0.03	2.1		
40L	0.18	1.0	1.1	
57K	0.15	1.1	1.0	
32Q-57E	0.3	0.8	1.3	
40L-57E	1.0	0.0	2.1	2.1
mut-3	1.8	−0.4	2.5	2.4

[a] Intrinsic equilibrium constants for the folding of monomers (P_f/P_u) were calculated as described in the text. ΔG_{sum} for 40L-57E and mut-3 were determined by adding the $\Delta\Delta G$ values for single and double mutants.

dimer is the same as the ellipticity of the folded monomer at 222 nm. High similarity between monomer and dimer structures is indicated by the existence of a isodichroic point at 210 nm, which is independent of prodomain concentration, and also the observation that the difference spectrum of prodomain−subtilisin complex minus subtilisin is very similar to the spectrum of the folded dimer.

The equilibrium between all monomers ($P_u + P_f$) and dimers ($[P_f]_2$) can be expressed in terms of the dimerization constant (K_a):

$$P_u + P_f \overset{K_a}{\rightleftharpoons} [P_f]_2$$

$$K_a = [P_f]_2/(P_u + P_f)^2$$

Since K_a is known from sedimentation experiments, the fraction of monomer ($P_f + P_u$) is determined by using the quadratic equation:

$$(P_f + P_u) = [(8K_aP_{total} + 1)^{1/2} - 1]/4K_a$$

At any given concentration of total prodomain, such as $[P]_{total} = 10\ \mu M$ in Figure 2, P_u for each mutant can be determined from its ellipticity at 222 nm and total monomer ($P_f + P_u$) can be determined from the dimerization constant for each mutant. P_f/P_u, the intrinsic equilibrium constant for folding independent of dimerization, is reported for each mutant in Table 2.

We performed similar analysis at 100 mM KP_i, pH 7.0, and 100 mM NaOAc, pH 5.0, for 40L-57E. The intrinsic equilibrium constant for folding is insensitive to pH (between 5 and 7) and to ionic strength between 0.01 and 0.5 M.

In a two-state system, the energetic consequences of a localized alteration can be measured by its effect on the overall free energy of the unfolding reaction. In a multistate system, one would expect these mutations to influence their local structure but not necessarily stabilize the entire native fold. The stabilizing mutations are in three different areas of the structure of prodomain: the β-strands 35−51, α-helix 23−32, and α-helix 53−61. In each case the stabilizing effects of mutations appears to accrue additively. There is not any indication from CD that the denatured state changes as a function of mutation. A clear isodichroic point is observed for pro-wt and this family of mutants, indicating that no populated states exist with nonnative secondary structure (Figure 2). Non-native structure can be induced by TFE. In fact, the wild-type prodomain can be driven into a highly α-helical state by 50% TFE. The relevance of this phenomenon to the folding equilibrium without cosolvent is debatable.

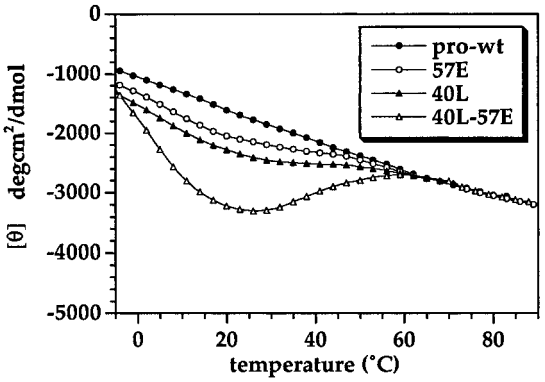

FIGURE 3: Temperature unfolding profile for mutant prodomains at pH 5. Mean residue ellipticity (degrees centimeter^2 decimole^{-1}) at 222 nm is plotted vs temperature in the range from −5 to 90 °C. All temperature profiles were recorded using a 1 mm cylindrical cuvette with a protein concentration of 50 μM in 100 mM sodium acetate buffer, pH 5.0.

Analysis of the Temperature Unfolding Profile. (A) Circular Dichroism. To avoid dimerization, analysis of prodomain unfolding vs temperature was carried out in 100 mM NaOAc, pH 5.0, at [P] = 50 μM. On the basis of the equilibrium sedimentation data, L40-K57 is 99% monomeric under these conditions. The equilibrium constant for folding in 100 mM NaOAc, pH 5.0, is the same as the intrinsic equilibrium constant for folding in 10 mM KP_i, pH 7.0 (Table 2).

The ellipticity of pro-wt at 222 nm decreases linearly with temperature with $\Delta mre = -27$ deg cm^2 dmol^{-1} °C^{-1}. This temperature dependence is typical of random polypeptides and small proteins in high Gu-HCl (Goodman et al., 1989; Merutka et al., 1991), indicating that the pro-wt is unfolded at all temperatures at pH 5.

Temperature unfolding profiles of pro-wt, 40L, 57E, and 40L-57E at pH 5 are presented in Figure 3. The unfolding reactions are completely reversible after scanning to 90 °C. Pro-wt is used as the baseline for the unfolded state. 40L-57E in the presence of 50% ethylene glycol is used for the baseline corresponding to fully folded protein. The fraction native is determined by subtracting unfolded baseline from the experimental CD signal and then dividing by the total CD difference between 100% folded and 0% folded at that temperature.

Stabilized prodomain mutants acquire maximum structure at ∼20 °C, based on ellipticity at 222 nm. 40L-57E has 50% native structure at this temperature. All mutants undergo two unfolding transitions, one induced by decreasing temperature below 20 °C and one induced by increasing temperature above 20 °C. As the temperature is raised or lowered the ellipticity of the mutants converges and then coincides with the wt prodomain ellipticity, indicating that the heat- and cold-denatured states of the prodomain mutants are unfolded.

(B) Differential Scanning Calorimetry. The amount of excess heat absorbed by a protein sample as the temperature T is increased through a transition from the folded to unfolded state at constant pressure provides a direct measurement of the ΔH of unfolding (Privalov & Khechinashvili, 1974). Calorimetric data were obtained for 40L-57E in 100 mM NaOAc, pH 5.0. A decrease in heat capacity occurs as

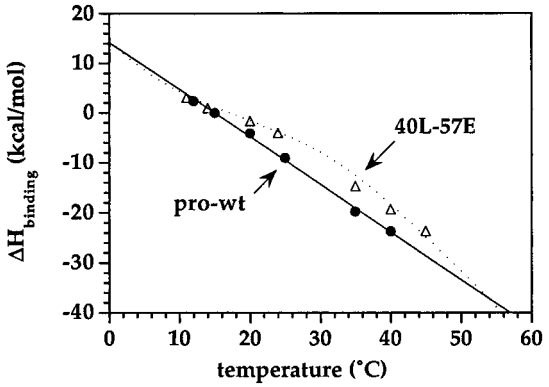

FIGURE 4: $\Delta H_{binding}$ vs temperature. A plot of $\Delta H_{binding}$ vs temperature for pro-wt is linear over the range of 11−40 °C and is fit to the equation $\Delta H = \Delta H_0 + \Delta C_p(T - T_0)$, where $\Delta C_p = -1.0$ kcal deg^{-1} mol^{-1}). A plot of $\Delta H_{binding}$ vs temperature for 40L-57E is fit to the equation $\Delta H = \Delta H_0 + (-1.0$ kcal deg^{-1} mol^{-1})[fraction folded]$(T - T_0)$.

the protein folds upon scanning from −10 to 20 °C, followed by an increase in heat capacity as the protein unfolds upon scanning from 20 to 60 °C. The magnitude of the heat capacity change (ΔC_p) is on the order of 1 kcal deg^{-1} mol^{-1} but is not well determined due to the difficulty in determining a baseline in the low temperature range. Little excess heat is evolved in the folding and unfolding reactions, indicating the ΔH is small over the temperature range where the folding and unfolding transitions are occurring. While DSC results are qualitatively similar to CD melting results, the small ΔH makes the folding reaction difficult to follow by DSC. For a more accurate measure of ΔH, isothermal calorimetry (ITC) was used.

Effect of Prodomain Mutations on the Thermodynamics of Its Binding to Subtilisin. Titration calorimetry was used to determine the enthalpies of binding ($\Delta H_{binding}$) and the association constants (K_as) of pro-wt and 40L-57E for folded subtilisin. The catalytically inactive subtilisin Sbt-110 was used in the titrations. Titrations were performed in 100 mM NaOAc, pH 5.0, to eliminate dimerization of 40L-57E as described in Materials and Methods.

$$P + S_f \overset{K_a}{\rightleftharpoons} P_f\text{-}S_f$$

(A) Enthalpy of Binding. Titrations of pro-wt and 40L-57E were performed over the temperature range of 11−45 °C to determine the temperature dependence of $\Delta H_{binding}$. $\Delta H_{binding}$ is equal to 0 at 16−18 °C for both prodomains and then decreases as T is increased. The decrease in $\Delta H_{binding}$ with increasing temperature results from the change in heat capacity (ΔC_p) upon binding prodomain to subtilisin, according to

$$\Delta H = \Delta H_0 + \Delta C_p(T - T_0)$$

For pro-wt, $\Delta H_{binding}$ was plotted vs temperature and fit to a straight line of slope (ΔC_p) equal to −1 kcal deg^{-1} mol^{-1} (Figure 4). ΔC_p has been shown to be correlated with the decreased exposure of hydrophobic surface area upon folding or in this case binding (Privalov & Gill, 1988; Livingstone et al., 1991).

For 40L-57E, $\Delta H_{binding}$ is not linear with temperature between 11 and 45 °C, indicating that ΔC_p varies with temperature. Presumably, ΔC_p changes between 11 and 45 °C because the fraction of folded 40L-57E is changing. The temperature dependence of $\Delta H_{binding}$ for 40L-57E can be accounted for by calculating $\Delta C_{pbinding}$ as a function of the fraction of folded 40L-57E (Figure 4). $\Delta C_{pbinding}$ is predicted to be greater than −1 kcal deg^{-1} mol^{-1} between 0 and 50 °C, where a significant fraction of 40L-57E is folded. This is consistent with the observation that the $\Delta H_{binding}$ vs T curve is flatter for 40L-57E than for pro-wt (Figure 4). As the fraction of folded 40L-57E approached 0 (below 0 or above 50 °C), ΔC_p is predicted to approach −1 kcal/mol and the $\Delta H_{binding}$ vs T curve becomes parallel to the line for pro-wt and offset by ≤1 kcal/mol. The L40 and E57 mutations thus result in a change in the folding enthalpy of ±1 kcal/mol.

Table 3: Titration Calorimetry of Subtilisin Sbt-110 with Pro-wt and 40L-57E[a]

	T (°C)	K_a (M^{-1})	$\Delta H_{binding}$ (kcal/mol)
wild type	11.3		3.1
	11.7		2.3
	14.8		−0.05
	19.6		−4.2
	24.8		−9.1
	34.5	4×10^6	−18.9
	40.0	3×10^6	−23.7
40L-57E	11.1		3.0
	14.3		0.9
	19.6		−1.65
	24.4		−4.1
	34.5		−14.7
	40.0	$\geq 5 \times 10^7$	−19.3
	45.0	$\geq 5 \times 10^7$	−23.6

[a] Binding constant (K_a) and binding enthalpy ($\Delta H_{binding}$) were determined using nonlinear least-squares minimization of the titration data to eq 1 in the text (Wiseman et al., 1989). The stoichiometry of binding is 1. Measurements for each experimental condition were performed in duplicate at each temperature. K_as at the lower temperatures cannot be determined as accurately due to the decreasing enthalpy of binding.

(B) Free Energy of Binding. The titration calorimeter is sensitive to changes in K_a under conditions at which the product K_a[S] is between 1 and 1000 (Wiseman et al., 1989). At temperatures > 35 °C, $\Delta H_{binding}$ is < −15 kcal/mol for both prodomains, allowing determination of K_as (Table 3). At the lower temperatures the smaller amount of heat produced per titration, coupled with tighter binding, makes direct determination of K_a less accurate. Titration curves of Sbt-110 with pro-wt and L40- E57 at 40 °C are shown in Figure 5. The data points correspond to the negative heat of binding associated with each addition of prodomain to Sbt-110. Pro-wt is bound to Sbt-110 with a K_a of 3×10^6 M^{-1}. L40-E57 is bound with a K_a of $\geq 5 \times 10^7$ M^{-1}, ≥15 times more tightly than pro-wt (Table 3). Since mutations are introduced in regions of the prodomain that do not directly contact subtilisin upon formation of the complex, their effects on binding to subtilisin are linked to whether or not they stabilize the native conformation. Therefore, mutations that stabilize independent folding of the prodomain into its native conformation should increase binding to subtilisin. The equilibrium for binding the prodomain to subtilisin is

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

where K_f is the equilibrium constant for folding the pro-

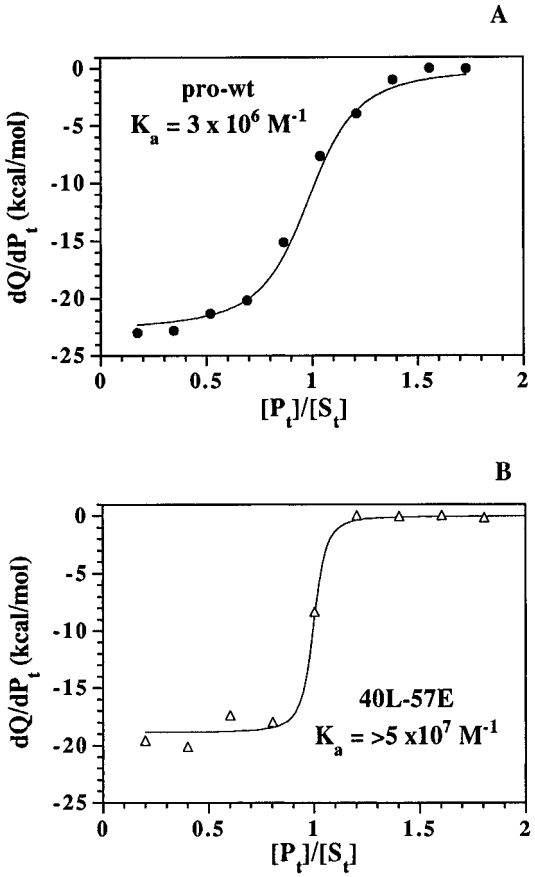

FIGURE 5: Titration of subtilisin Sbt-110 with pro-wt and 40L-57E. The heats of binding for successive additions of prodomain are plotted vs the ratio of [P]/[Sbt-110]. Binding curves were determined by nonlinear least-squares minimization method using eq 1 in the text (Wiseman et al., 1989). Temperature was 40 °C.

domain and K_P is the association constant of folded prodomain for subtilisin. The observed binding constant is then $K_{(P+P_u)} = [K_f/(1 + K_f)]K_P$.

As the fraction of folded prodomain approaches 1, the observed association constant approaches its maximum, K_P. The 15−20-fold difference in the fraction of folded protein between 40L-57E and pro-wt observed by CD (Table 2) is consistent with the ≥15-fold difference in binding constant.

Combining Calorimetric and CD Data. If a protein unfolds in a two-state manner, then the temperature dependence of the unfolding reaction will be determined by the thermodynamic state functions ΔH, ΔS, and ΔC_p according to the Gibbs−Helmholz equation:

$$\Delta G = \Delta H_0 - T\Delta S_0 + \Delta C_p[T - T_0 - T \ln (T/T_0)] \quad (2)$$

where ΔH_0 and ΔS_0 are the enthalpy and entropy of unfolding evaluated at a reference temperature T_0, and ΔG is the free energy of unfolding at a temperature T (Becktel & Schellman, 1987; Brandts, 1964; Pace & Tanford, 1968; Privalov, 1979).

The ellipticity data from CD experiements were converted to $\Delta G_{unfolding}$ (Figure 6). The maximum $\Delta G_{unfolding}$ (0 kcal/mol) is observed at 20 °C. 40L-57E unfolds more if the temperature is either raised or lowered from 20 °C (293 K). According to the Gibbs equation, ΔS =0 at the maximum $\Delta G_{unfolding}$ and stabilization is entirely enthalpic with $\Delta G_{293} = \Delta H_{293}$ (Becktel & Schellman, 1987). The predictions of the Gibbs equation are reasonable in light of calorimetric

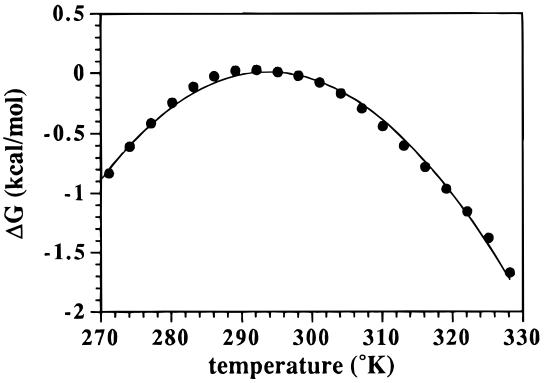

FIGURE 6: ΔG as a function of temperature for the unfolding of 40L-57E. The temperature unfolding profile measured by far-UV CD for 40L-57E (Figure 3) was converted to an apparent $\Delta G_{unfolding}$ (solid circles). The solid line is a theoretical curve calculated using the Gibbs−Helmholtz equation: $\Delta G_{unfolding} = \Delta H_0 - T\Delta S_0 + \Delta C_p[T - T_0 - T \ln (T/T_0)]$, where $T_0 = 293$ K, $\Delta H_0 = 0$ kcal/mol, $\Delta S_0 = 0$ cal deg^{-1} mol^{-1} and $\Delta C_p = 0.9$ kcal deg^{-1} mol^{-1}.

results from ITC, which show that $\Delta H_{binding}$ passes through 0 between 16 and 18 °C (289 and 291 K).

The curvature of the free energy profile ($\Delta G_{unfolding}$ vs T) is determined by the increase in heat capacity of the system when the protein unfolds (Privalov, 1979; Becktel & Schellman, 1987). A theoretical stability curve was calculated using eq 2 with $\Delta H_{293} = 0$, $\Delta S_{293} = 0$, and $\Delta C_p = 0.9$ kcal deg^{-1} mol^{-1}. The calculated $\Delta G_{unfolding}$ vs T curve closely coincides with the experimental $\Delta G_{unfolding}$ determined from CD melting. The temperature of maximum stability of the 40L-57E (∼20 °C) is typical of many globular proteins and generally is attributed to the large contribution of hydrophobic burial to the stabilization of the folded state. The ΔH_{293} for stable proteins averages 200 cal/mol per residue (Makhatadze & Privalov, 1995). $\Delta H_{unfolding}$ at 293−298 K is generally interpreted as the enthaply of stabilization due to nonhydrophobic contributions (Baldwin, 1986). The fact that $\Delta H_{293} = 0$ for 40L-57E helps account for its marginal stability.

DISCUSSION

Studies of the stabilized prodomain mutants have revealed two surprising features about its unfolding reaction. First, cooperative, two-state folding is observed for 40L-57E in spite of its marginal stability ($\Delta G_{unfolding} = 0$ kcal/mol at 20 °C). Evidence for two-state folding is based on the observed additivity of individual mutations to the overall $\Delta G_{unfolding}$ and the conformity of $\Delta G_{unfolding}$ vs T to two-state assumptions as embodied in the Gibbs−Helmholz equation. In fact, single mutants and the pro-wt may be two-state reactions even though $\Delta G_{unfolding}$ is negative. Second, the denatured states of all mutants are not discernibly different in spite of the differences in $\Delta G_{unfolding}$. This finding was surprising in light of results with staphylococcal nuclease, which clearly show that the denatured state tends to become more structured as $\Delta G_{unfolding}$ increases (Shortle, 1996). Conclusions about the nature of the denatured state of the prodomain are based on CD spectral data and calorimetric data. CD measurements show a clear isodichroic point for all mutants (Figure 2), and the temperature dependence of ellipticity at 222 nm suggests that the cold- and heat-denatured states of the prodomain are largely unfolded. Calorimetric determination of $\Delta H_{binding}$ shows that the mutations cause only minor

changes in ΔH. There is a great deal of data relating specific structural changes in the native state with changes in $\Delta G_{unfolding}$ [e.g., Matthews (1987) and Shirley et al. (1992)]. In the cases where calorimetric analysis is done, it usually turns out that relatively small changes in ΔG are the result of large but compensating changes in ΔH and ΔS [e.g., Hu et al. (1992)]. Often even conservative mutations with small effects on ΔG result in $\Delta H_{folding}$ changes >10 kcal/mol. Large changes in ΔH and ΔS may arise from the fact that mutations affect the distribution of denatured states as well as affecting the structure of the native state (Shortle et al., 1988). In the case of the 40L-57E mutant, the magnitude of the ΔH change is similar to the ΔG change.

To produce a stable, unique protein fold there must be an energetic barrier between the native and denatured states. A prevalent view in the protein design field is that to produce a unique native state, in addition to creating favorable native interactions, competing compact folds must be uncoded (Betz et al., 1993; Lattman & Rose, 1993). Avoiding stable misfolds, however, has proven difficult. De novo designed folds often have the intented secondary structure but generally are not cooperative two-state systems because intermediates exist with a continuum of stabilities in between the native and unfolded states. The code for α-helical structure is highly degenerate (Kamtekar et al., 1993); hence it is easier to design mutations that increase helical structure than those that stabilize a specific tertiary fold. Design is difficult, therefore, because one must be able to predict stability of the native state relative to the stability of partial folds.

We believe our success in stabilizing the native state of the prodomain without stabilizing intermediates originates from the ability to select for its function in facilitating subtilisin folding rather than simply selecting for an increase in secondary structure content. Since the ability of the prodomain to catalyze subtilisin folding depends upon the correct tertiary fold, we were able to evaluate mutations initially in terms of function. This relationship is analyzed in detail by Wang et al. (manuscript in preparation). It is probable that some mutations in the prodomain would increase the content of structure while decreasing the probability of the active conformation; thus, assays which detect only increases in structure could be misleading. The fact that a small number of mutations can stabilize the independent folding of the prodomain implies that most of the folding information already exists in the wild-type amino acid sequence in spite of the fact that the unfolded state predominates. On the mutational path to a stable fold, we coevolved structure and function simultaneously, thus avoiding missteps into unproductive regions of mutational space.

ACKNOWLEDGMENT

We thank Joel Hoskins for synthesizing the oligonucleotides used in site-directed mutagenesis and DNA sequencing and Patrick Alexander and John Moult for useful discussion. The identification of commercial equipment and materials in this paper does not imply recommendation or endorsement by the National Institute of Standards and Technology.

REFERENCES

Alexander, P., Fahnestock, S., Lee, T., Orban, J., & Bryan, P. (1992) *Biochemistry 31*, 3597−3603.
Bai, Y., & Englander, S. W. (1996) *Proteins: Struct., Funct., Genet. 24*, 145−151.
Baldwin, R. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8069−8072.
Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859−1877.
Betz, S. F., Raleigh, D. P., & DeGrado, W. F. (1993) *Curr. Opin. Struct. Biol. 3*,
Brandts, J. F. (1964) *J. Am. Chem. Soc. 86*, 4291−4301.
Brooks, I. S., Soneson, K. K., & Hensley, P. (1993) *Biophys. J. 64*, A244.
Bryan, P., Wang, L., Hoskins, J., Ruvinov, S., Strausberg, S., Alexander, P., Almog, O., Gilliland, G., & Gallagher, T. D. (1995) *Biochemistry 34*, 10310−10318.
Buckler, D. R., Haas, E., & Scheraga, H. A. (1995) *Biochemistry 34*, 15965−15978.
Dill, K. A., & Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.
Ellis, K. J., & Morrison, J. F. (1982) *Methods Enzymol. 87*, 405−426.
Gallagher, T. D., Gilliland, G., Wang, L., & Bryan, P. (1995) *Structure 3*, 907−914.
Goodman, E. M., & Kim, P. S. (1989) *Biochemistry 28*, 4343−4347.
Hu, C.-Q., Kitamura, S., Tanaka, A., & Sturtevant, J. M. (1992) *Biochemistry 31*, 1643−1647.
Ikemura, H., Takagi, H., & Inouye, M. (1987) *J. Biol. Chem. 262*, 7859−7864.
Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M., & Hecht, M. H. (1993) *Science 262*, 1680−1685.
Lattman, E. E., & Rose, G. D. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 439−441.
Livingstone, J. R., Spolar, R. S., & Record, T. M. (1991) *Biochemistry 30*, 4237−4244.
Makhatadze, G. I., & Privalov, P. L. (1995) *Adv. Protein Chem. 47*, 307−425.
Matthews, B. W. (1987) *Biochemistry 26*, 6885−6888.
Merutka, G., Shalongo, W., & Stellwagen, E. (1991) *Biochemistry 30*, 4225−4248.
Orban, J., Alexander, P., Bryan, P., & Khare, D. (1995) *Biochemistry 34*, 15291−15300.
Pace, C. N., & Tanford, C. (1968) *Biochemistry 7*, 198−208.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167−241.
Privalov, P. L., & Gill, S. J. (1988) *Adv. Protein Chem. 39*, 191−234.
Privalov, P. L., & Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665−684.
Schwarz, F. P., & Kirchhoff, W. H. (1988) *Thermochim. Acta 128*, 267−295.
Shirley, B. A., Stanssens, P., Ulrich, H., & Pace, C. N. (1992) *Biochemistry 31*, 725−732.
Shortle, D. (1996) *Protein Sci. 5*, 991−1000.
Shortle, D., Meeker, A. K., & Freire, E. (1988) *Biochemistry 27*, 4761−4768.
Strausberg, S., Alexander, P., Wang, L., Schwarz, F., & Bryan, P. (1993) *Biochemistry 32*, 8112−8119.
Strausberg, S., Alexander, P., Gallagher, D. T., Gilliland, G., Barnett, B. L., & Bryan, P. (1995) *Bio/Technology 13*, 669−673.
Wang, L., Ruvinov, S., Strausberg, S., Gallagher, T. D., Gilliland, G., & Bryan, P. (1995) *Biochemistry 15*, 415−15,420.
Wiseman, T., Williston, S., Brandts, J. F., & Lin, L.-N. (1989) *Anal. Biochem. 179*, 131−137.

BI9703958