The Pro Region N-Terminal Domain Provides Specific Interactions Required for Catalysis of α-Lytic Protease Folding†

Erin L. Cunningham,‡ Ted Mau,‡ Stephanie M. E. Truhlar,§ and David A. Agard*,‡,||

The Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, Graduate Group in Biophysics, and Program in Chemistry and Chemical Biology, University of California at San Francisco, San Francisco, California 94143-0448

Received March 18, 2002; Revised Manuscript Received May 14, 2002

ABSTRACT: The extracellular bacterial protease, α-lytic protease (αLP), is synthesized with a large, two-domain pro region (Pro) that catalyzes the folding of the protease to its native conformation. In the absence of its Pro folding catalyst, αLP encounters a very large folding barrier ($\Delta G = 30$ kcal mol^{-1}) that effectively prevents the protease from folding ($t_{1/2}$ of folding = 1800 years). Although homology data, mutational studies, and structural analysis of the Pro·αLP complex suggested that the Pro C-terminal domain (Pro C-domain) serves as the minimum "foldase" unit responsible for folding catalysis, we find that the Pro N-terminal domain (Pro N-domain) is absolutely required for αLP folding. Detailed kinetic analysis of Pro N-domain point mutants and a complete N-domain deletion reveal that the Pro N-domain both provides direct interactions with αLP that stabilize the folding transition state and confers stability to the Pro C-domain. The Pro N- and C-domains make conflicting demands upon native αLP binding that are alleviated in the optimized interface of the folding transition state complex. From these studies, it appears that the extremely high αLP folding barrier necessitates the presence of both the Pro domains; however, αLP homologues with less demanding folding barriers may not require both domains, thus possibly explaining the wide variation in the pro region size of related pro-proteases.

α-Lytic protease (αLP)[1,2] is a chymotrypsin-like serine protease secreted by the soil bacterium *Lysobacter enzymogenes* as a means of providing nutrients to the host (*1*, *2*). In vivo, αLP is synthesized with a large 166-residue N-terminal pro region (Pro) that is required, either in *cis*, as the naturally occurring peptide, or in *trans*, as a separate polypeptide, for the proper folding of the 198-residue protease to its native state (Nat) (*3*, *4*). This Pro folding requirement can be reconstituted in vitro with denatured αLP. When diluted out of denaturant, αLP folds to a stable molten globule-like intermediate (Int) that is unable to convert to native protease on a biologically relevant time scale due to the presence of a large folding barrier ($t_{1/2}$ of folding = 1800 years) (*5*, *6*). Similarly, the native state is kinetically trapped and prevented from unfolding to the more thermodynamically stable Int and unfolded states by a substantial barrier to unfolding ($t_{1/2}$ of unfolding = 1.2 years). Pro facilitates Int folding in two ways: first, by stabilizing the folding transition state, accelerating folding by nearly 10^{10}-fold, and second, by binding tightly to the native state of the protease, thereby shifting the thermodynamic equilibrium in favor of the Pro·Nat complex. The eventual degradation of Pro then leads to the release of mature, active protease. Recent work demonstrates that the highly cooperative unfolding barrier significantly limits both global and local unfolding events, resulting in a dramatic enhancement in resistance to proteolysis, which in turn leads to the evolutionary advantage of an increased functional lifetime (*6*, *7*). The energetic penalty for such extreme cooperativity is a thermodynamically unstable native state that necessitates the co-evolution of a transient pro region to mediate folding.

Pro-catalyzed folding of αLP follows Michaelis−Menten-like kinetics in which Pro forms an initial binding (or Michaelis) complex with Int (Pro·Int), a folding transition state complex (Pro·Int‡), and an inhibitory complex with the native protease (Pro·Nat), as assessed by K_M, k_{cat}, and K_i values, respectively (*8*). Kinetic analysis of successive truncations (one to four amino acids) of the Pro C-terminus shows these residues to be critical for stabilizing the folding transition state, while having only minor effects on initial and native state binding (*8*). Biochemical and structural studies of the Pro·Nat complex reveal that these residues of the Pro C-terminal tail lie in a substrate-like manner in the protease active site (Figure 1a) (*9*, *10*). In addition to this active site binding of the C-terminal tail, a three-stranded β-sheet from the Pro C-terminal domain (Pro C-domain) pairs with a β-hairpin from the protease to form a continuous five-stranded β-sheet (*10*). Mutations in this β-sheet interface

† This work was supported by the Howard Hughes Medical Institute. S.M.E.T. was a National Science Foundation predoctoral fellow.

* To whom correspondence should be addressed. Telephone: (415) 476-2521. Fax: (415) 476-1902. E-mail: agard@msg.ucsf.edu.

‡ Graduate Group in Biophysics.

§ Program in Chemistry and Chemical Biology.

|| Department of Biochemistry and Biophysics.

[1] Abbreviations: αLP, α-lytic protease; S143A, αLP with active site Ser143 (sequential numbering) replaced with Ala; Int, folding intermediate state of αLP; Nat, native state of αLP; Pro, wild-type pro region; Pro·Nat, complex of Pro with native αLP; Y26F, Pro with Tyr26 replaced with Phe; E30A, Pro with Glu30 replaced with Ala; Y26F/E30A, Pro with both Y26F and E30A mutations; N-domain, N-terminal domain; C-domain, C-terminal domain; SGPB, *S. griseus* protease B; CD, circular dichroism.

[2] αLP and Pro residue numberings are sequential.

10.1021/bi020214o CCC: $22.00
Published on Web 06/20/2002

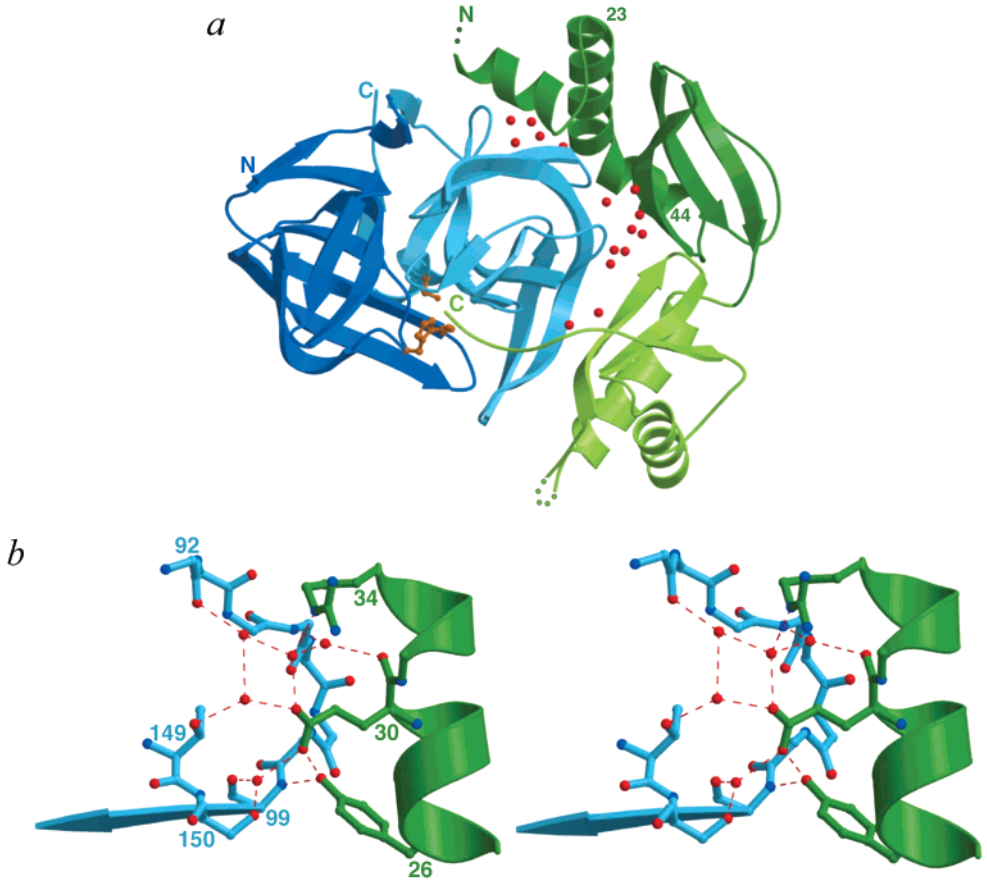

FIGURE 1: Structure of the Pro·Nat complex. (a) Ribbon diagram of the Pro·Nat complex structure (*10*). The αLP N- and C-domains are colored dark and light blue, respectively, with the side chains of the catalytic triad shown in orange. The N-domain (dark green) and C-domain (light green) of Pro bind the αLP C-domain, with the Pro C-terminal tail inserted into the protease active site. Waters in the interface, those that mediate Pro−αLP interactions in lieu of direct protein−protein interactions, are depicted as red spheres. The positions of these ordered waters are based upon the structure of the Pro-3·Nat complex that was determined to a higher resolution than the wild-type complex (*10*). (b) Detailed stereoview of the hydrated αLP·Pro N-domain interface, rotated ~180° about the horizontal axis from the view in panel a. Pro residues Y26, E30, and R34 from the helix (residues 23−44) interact with turn residues 92−99 (only backbone atoms displayed), 149, and 150 of αLP (sequential numbering) via an extensive network of mostly water-mediated hydrogen bonds (dashed red lines).

affect both K_M and k_{cat}, indicating that the Pro C-domain is involved in initial binding of Int, as well as in stabilization of the folding transition state (*8*). The likely importance of this β-hairpin to the αLP folding barrier is reflected in its selective conservation among related proteases in the chymotrypsin superfamily. The β-hairpin is a common structural motif in all bacterial homologues synthesized with pro regions, but is noticeably absent in other related bacterial, viral, and mammalian proteases that do not require pro regions for proper folding. Furthermore, the Pro β-strand that pairs with the hairpin loop is conserved in all homologous pro regions, suggesting that the hairpin and its interactions with the Pro C-domain are important in structuring the protease.

While it is apparent that the extensive contacts between the Pro C-domain and αLP contribute significantly to the Pro-catalyzed folding reaction, these interactions only account for approximately 58% of the >4000 Å^2 of surface area buried in the Pro·Nat complex (*10*). The Pro N-terminal domain (Pro N-domain) also donates a large protease binding surface (1684 Å^2) which, in combination with the Pro C-domain, effectively surrounds one domain of the protease (Figure 1a). However, in contrast to Pro C-domain binding of αLP, the interface between the Pro N-domain and αLP displays less surface complementarity and is marked by a high degree of hydration (*10*, *11*). Primary sequence alignments with homologous pro-proteases show the N-domain to be more variable in size than the more conserved C-domain (*12*). Furthermore, the closely related αLP homologues, *Streptomyces griseus* proteases A and B, are synthesized with smaller pro regions that appear to lack the N-domain entirely and instead correspond to a single domain comparable to αLP's Pro C-domain.

These findings seem to identify the Pro C-domain as the minimum functional "foldase" unit. The Pro C-domain both initially recognizes Int by binding the β-hairpin and continues to provide important interactions with the hairpin and the substrate binding pocket in the folding transition state complex. Our current model of pro-mediated folding holds that Pro C-domain binding to αLP in the folding transition state serves to position the protease β-hairpin, leading to the structuring of the αLP C-domain and, ultimately, the complete attainment of the protease's native fold. What then is the function of the substantial Pro N-domain in αLP folding?

This paper explores the role of the Pro N-domain in the catalysis of αLP folding through a detailed kinetic analysis of several Pro N-domain point mutants and a complete N-domain deletion. Here we show that Pro N-domain mutants designed to disrupt the binding interface with αLP

primarily impact stabilization of the folding transition state and exert their effect in a cooperative manner. Deletion of the entire N-domain to create a mimic of the homologous single-domain pro regions greatly destabilizes the Pro C-domain and virtually abolishes Pro-catalyzed folding of αLP, but has a minimal effect on native state binding. The implications of these results on the mechanism of Pro-catalyzed αLP folding and on pro region domain requirements for related pro-proteases are discussed.

MATERIALS AND METHODS

Pro Region N-Domain Mutants and C-Domain Construct. Point mutations were introduced into the pT7XmaPro plasmid (*9*) using Chameleon (Strategene, La Jolla, CA) to insert single-amino acid substitutions Y26F and E30A into the N-domain of the Pro gene (*11*). QuikChange (Strategene) was then used to add the Y26F mutation to the E30A construct, combining the mutations to form the Y26F/E30A double mutant.

To create a C-domain Pro truncation mutant (CPro), the pT7XmaPro plasmid was digested with *Nde*I and partially digested with *Bam*HI to remove a 370 bp fragment of the Pro gene containing the Pro N-domain. Two overlapping oligonucleotide cassettes were ligated into the resulting *Nde*I- and *Bam*HI-digested pT7XmaPro to build a construct of the Pro C-domain (S87−T166) that included an N-terminal His tag sequence (MGSSHHHHHHSSG) to aid in purification. All oligonucleotides were synthesized by the DNA Facility of the Howard Hughes Medical Institute at the University of California at San Francisco, and all mutations were verified by double-stranded sequencing by the same facility.

Production and Purification of the Pro Region. Wild-type and Pro N-domain mutants were purified from *Escherichia coli* strain BL21(DE3)/pLysS as described previously (*13*). The wild-type pro region that was used contains an additional N-terminal proline residue, a cloning artifact that does not affect its behavior (*8*).

The His-tagged CPro was also expressed in *E. coli* strain BL21(DE3)/pLysS (*13*). Initial purification of CPro with Ni−NTA resin followed the standard denaturing purification protocol for insoluble proteins [Qiagen, QIAexpressionist (2nd ed., 1992)], with the addition of 250 μM benzamidine and 100 μM PMSF to the initial resuspension/solubilization buffer and the use of an imidazole-based alternative elution procedure. Partially purified CPro was dialyzed against 10 mM sodium phosphate (pH 7.2) at 4 °C in dialysis tubing with a 2000 Da molecular mass cutoff to lower the denaturant concentration to 0.1 M urea, then flash-frozen in a dry ice/ethanol bath, and stored at −20 °C. Thawed aliquots of CPro were filtered through 0.2 μm sterilizing filter units and further purified by gel chromatography on a HiLoad 16/60 Superdex 200 preparation grade column with a 100 mM NaCl, 10 mM sodium phosphate (pH 7.2) running buffer. Purified CPro was desalted and concentrated in a 3000 Da molecular mass cutoff Centriprep apparatus (Millipore, Bedford, MA). Subsequent CPro preparations were dialyzed to 1 M urea after Ni−NTA partial purification and were maintained at this higher denaturant concentration throughout the remaining steps of the purification described above. The increase in urea concentration allowed a higher Cpro concentration to be reached in the stock solutions but otherwise had no effect.

Production, Purification, and Denaturation of α*-Lytic Protease*. Wild-type αLP was produced in liquid cultures of *L. enzymogenes* 495 (ATCC 29487) as described previously (*14*). After 6 days of growth, αLP was purified from culture supernatants by ion exchange chromatography and HPLC (*15*), and a portion was denatured as described previously (*13*).

Circular Dichroism Studies. Pro region stability was determined as a function of both urea concentration and temperature by monitoring the circular dichroism at 222 nm in a Jasco 715 spectropolarimeter. Pro (5 μM) in 10 mM sodium phosphate (pH 7.2) was titrated with 8 M urea at 5 °C using a 1 cm path length cuvette, and the resulting denaturation curves were fit to a two-state transition using a nonlinear free energy method (*16*). Urea concentrations were determined by refractive index measurements (*16*). Thermal denaturation curves of 1 μM Pro samples were recorded with 4 s readings every 0.2 °C after equilibration for 30 s at that temperature. Chemical and thermal denaturations of wild-type Pro were repeated with more data points and less experimental error than in previous studies and were consistent with these earlier reports (*8*, *17*). All denaturations were fully reversible for both wild-type Pro and the Pro N-domain mutants.

Stability measurements could not be obtained for CPro as it was unstructured at 5 °C in the absence of urea. However, CD spectra of 10 μM CPro with and without the addition of an equimolar amount of the inactive αLP variant, S143A (*6*), were recorded using a 10 mm path length cuvette, as were the spectra of S143A alone and buffer alone.

Inhibition Measurements. The level of inhibition of αLP by Pro N-domain mutants was determined by assaying αLP activity in the presence of increasing amounts of the purified pro region as previously described, except that purified pro region concentrations were initially estimated by absorbance measurements at 280 nm (*8*). The apparent K_i values were extracted from fits of the inhibition data to the tight-binding inhibitor equation (*18*) using Kaleidagraph version 3.08d (Synergy Software).

Due to the instability of CPro, the standard inhibition assay described above was modified to measure the level of CPro inhibition of αLP; however, this had no effect on wild-type Pro inhibition of αLP (data not shown). The protease was first preincubated with varied amounts of CPro in 30 mM urea, 0.1 mg/mL BSA, 20 mM potassium succinate (pH 5.6) for 5 min at 25 °C. Protease activity was then assayed with 4 mM succinyl-Ala-Ala-Pro-Ala-*p*-nitroanilide substrate in 30 mM urea, 0.05 mg/mL BSA, 50 mM Tris (pH 8.0) at 25 °C. The resulting inhibition data were fit to a competitive inhibitor variation of the Michaelis−Menten equation.

Pro-Catalyzed Folding. Catalyzed folding by Pro N-domain mutants at pH 5.6 was performed as described previously (*13*), maintaining a 16-fold molar excess of Pro. Folding curves were generated for a range of Pro concentrations (typically 5−100 μM), except for Y26F, which was limited to a smaller range of 5−25 μM due to aggregation at higher concentrations. Y26F/E30A-mediated folding of αLP was monophasic and was fit with a single exponential, while both Y26F and E30A folding reactions were biphasic and were fit with a five-parameter double exponential. The amplitudes of these biphasic curves were fixed at the

consensus ratio of 60:40 (fast phase:slow phase) and refit with a three-parameter double exponential, and the resulting fast phase magnitudes were plotted against Pro concentration to determine k_{cat} and K_M values from a fit to a variant of the Michaelis−Menten equation previously described (*8*, *13*). Rate constants from the monophasic Y26F/E30A curves were fit directly to this equation. All data analysis and fits were performed using Kaleidagraph version 3.08d (Synergy Software).

The slow rate of CPro-mediated αLP folding required the use of the very sensitive thiobenzyl ester substrate assay, originally utilized to measure the uncatalyzed folding rate (*6*, *13*). Int (2.5 μM) in 30 mM urea, 20 mM potassium succinate (pH 5.6) was incubated at 0 °C (ice water) in the absence and presence of 13.5 μM Cpro (limited by CPro solubility in the highly concentrated stocks). At each time point, 100 μL aliquots of the folding reaction were treated with 10 μg of pepsin at pH 2.5 for 30 min to digest Int (and Cpro) while leaving native αLP intact. Folded αLP was then assayed as described previously by following absorbance at 324 nm (*13*). The data were corrected for substrate autolysis, and the concentration of native αLP was determined from a standard curve.

Figure Preparation. Figures of the Pro·Nat complex were prepared using Molscript (*19*) and Raster 3D (*20*).

RESULTS

Pro Region N-Domain Mutants. To examine the importance of the Pro N-domain in the catalyzed folding of αLP, Pro mutations Y26F and E30A were introduced to disrupt N-domain binding to the protease. From the Pro·Nat complex structure (Figure 1a,b), it was observed that Y26 makes one of only two direct hydrogen bonds to αLP in the N-domain interface (*10*). The hydrogen bonding potential of this residue is preserved in homologous pro region N-domains, and the hydrogen bond geometry is maintained between two crystal forms of the Pro·Nat complex, which show significantly different main chain positions for Pro residues 23−44, by adoption of an alternate Y26 rotomer. The side chain of E30, a conserved residue in homologous pro region N-domains, further anchors this hydrogen bonding network by making one hydrogen bond with Y26 and three hydrogen bonds with ordered solvent molecules in the interface. The point mutations Y26F and E30A and the double mutant, Y26F/E30A, were made to remove hydrogen bonds from the N-domain interface, therefore disturbing binding to the protease.

To assess the effect of these mutations on native state binding, the K_i of each Pro variant for wild-type αLP was determined. All three mutants demonstrated tight-binding inhibition of the protease (Figure 2 and Table 1), with K_i values within error of the K_i of 0.32 ± 0.064 nM for wild-type Pro (*8*). These results are surprising as the mutations were designed to remove favorable interactions observed in the structure of the Pro·Nat complex, yet they result in no detectable loss in affinity for native αLP.

The mutant pro regions did display altered αLP folding kinetics, as shown in Figure 3. Progress curves of Pro-assisted αLP folding (see Materials and Methods) reveal a successive decrease in folding rates, with E30A, Y26F, and Y26F/E30A slowing pro-catalyzed folding by factors of approximately 2-, 10-, and 100-fold, respectively. Both Y26F and E30A exhibited biphasic folding kinetics as is observed for wild-type Pro and numerous other Pro mutants (*6*, *8*, *13*). The biphasic nature of these refolding curves arises from a parallel pathway in which the Pro·Int complex isomerizes to an alternate conformation [(Pro·Int)′] that only slowly converts to the folded product [(k_{cat})′] (*8*). This isomerization competes with catalyzed folding and accounts for the slow phase of the folding reaction, which unlike the Pro-dependent fast phase (k_{cat}) is invariant with Pro concentration. Y26F/E30A-catalyzed folding is sufficiently slow to make the second phase undetectable and therefore appears to be monophasic, consistent with other more debilitated Pro mutants previously characterized (*8*, *13*).

Full enzymatic profiles of the catalyzed folding reactions were generated using a range of Pro mutant concentrations, and the rate constants for the monophasic progress curves (k_{obs}) and for the fast phases of the biphasic progress curves ($k_{obs,fast}$) were fit to a variant of the Michaelis−Menten equation to extract K_M and k_{cat} values (Figure 4a−c) (*8*). A comparison of the mutant K_M values to that of wild-type Pro showed Y26F and E30A to have essentially no effect on the initial binding of Pro to Int. Instead, the folding defects of these single-point mutations were contained entirely within

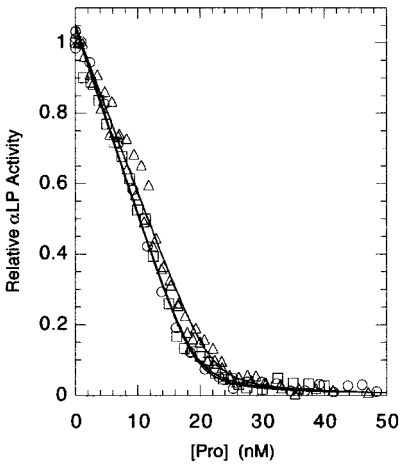

FIGURE 2: Inhibition of αLP by Pro N-domain mutants. Native αLP (22 nM) was incubated with varying amounts of Y26F (○), E30A (△), or Y26F/E30A (□) pro regions for 2−3 min at 25 °C and then assayed for activity. Activity was normalized to the activity in the absence of Pro, and the data were fit to the tight-binding inhibitor equation (see Materials and Methods). Solid lines represent curve fits with K_i values summarized in Table 1.

Table 1: Affinity and Kinetic Constants[a]

	K_i (nM)	k_{cat} (min^{-1})	K_M (μM)	relative k_{cat}/K_M
wild type	0.32 ± 0.064[b]	1.9 ± 0.17[c]	24 ± 6.1[c]	1.0
Y26F	0.28 ± 0.082	0.20 ± 0.048	23 ± 9.1	0.11
E30A	0.35 ± 0.097	0.92 ± 0.064	27 ± 5.1	0.43
Y26F/E30A	0.30 ± 0.091	0.031 ± 0.0020	36 ± 5.4	0.011
CPro	2.9 ± 0.19	NA	NA	NA

[a] Inhibition constants were determined using the tight-binding inhibitor analysis for all Pro variants except CPro, which was weak enough to allow use of competitive inhibition Michaelis−Menten analysis (see Materials and Methods and Figures 2 and 7). k_{cat} and K_M values were determined from catalyzed refolding experiments (see Materials and Methods and Figure 4a−c). NA means not applicable. [b] K_i value for wild-type Pro taken from Peters et al. 1998. [c] Kinetic constants for wild-type Pro taken from Derman et al. 2000.

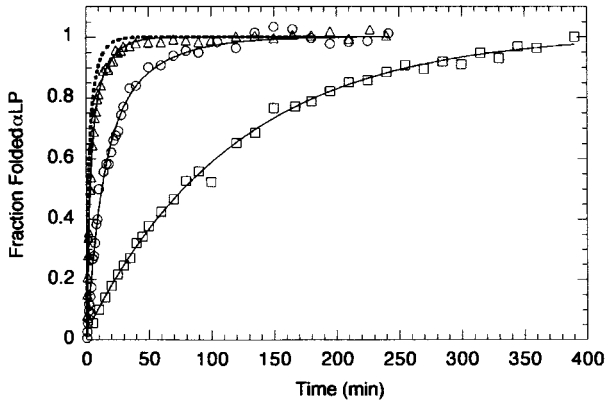

FIGURE 3: αLP folding catalyzed by Pro N-domain mutants. The folding of 0.94 μM Int was catalyzed by 15 μM pro region, and the folding rate constants for the reactions were calculated by fitting the resulting progress curves. E30A (△) and Y26F (○) curves were fit to biphasic exponentials, while the fit of the Y26F/E30A (□) curve was monophasic. The fast phase rate constants for E30A ($k_{obs,fast} = 0.36 \pm 0.009$ min^{-1}) and Y26F ($k_{obs,fast} = 0.083 \pm 0.004$ min^{-1}) correspond to catalyzed folding, and they were approximately 2 and 10 times slower than wild-type Pro (shown as dotted line for comparison) (*13*), respectively. Y26F/E30A's catalyzed folding rate ($k_{obs} = 0.0084 \pm 0.0003$ min^{-1}) was nearly 100 times slower than that of wild-type Pro.

Table 2: Stability Measurements

	T_m (°C)	$\Delta G_{unfolding}$ (kcal/mol)
wild type	30.4 ± 0.2	2.0 ± 0.10
Y26F	29.3 ± 1.0	1.56 ± 0.02
E30A	29.4 ± 0.8	1.65 ± 0.03
Y26F/E30A	29.5 ± 0.5	1.37 ± 0.02

k_{cat} (Table 1). Y26F weakened stabilization of the folding transition state by 10-fold, while E30A had a lesser 2-fold effect. The combination of these mutations in the Y26F/E30A double mutant disabled catalysis further, producing a 60-fold loss in affinity for the folding transition state. In addition to its cooperative effect on folding rate, Y26F/E30A also slightly hindered the formation of the Pro·Int Michaelis complex, as indicated by a 1.5-fold increase in K_M over that of wild-type Pro. Such a subtle change in K_M may be the result of indirect solvent effects, rather than the loss of a direct interaction.

To establish if the effects of these mutations on folding catalysis were simply a consequence of changes in Pro stability, thermal and chemical denaturations of the Pro N-domain mutants were monitored by circular dichroism (CD). All of the mutants exhibited T_m values similar to the T_m of wild-type Pro (30.4 ± 0.2 °C), suggesting little change in Pro stability for these point mutations (Table 2). Nonlinear fits of urea denaturation data (Figure 5) to the linear free energy model resulted in $\Delta G_{unfolding}$ values for the mutants that were a bit lower than that of wild-type Pro ($\Delta G_{unfolding} = 2.0 \pm 0.1$ kcal/mol); however, these minor changes in stabilities were insufficient to account for the observed effects on folding (Table 2).

These results, combined with the detailed analysis of the folding kinetics, indicate that disruption of the hydrogen bonding potentials of Y26 and E30 primarily alters Pro binding to the folding transition state as compared to binding in the initial and native state complexes. The Pro N- and C-domains are therefore both involved in stabilizing the high-

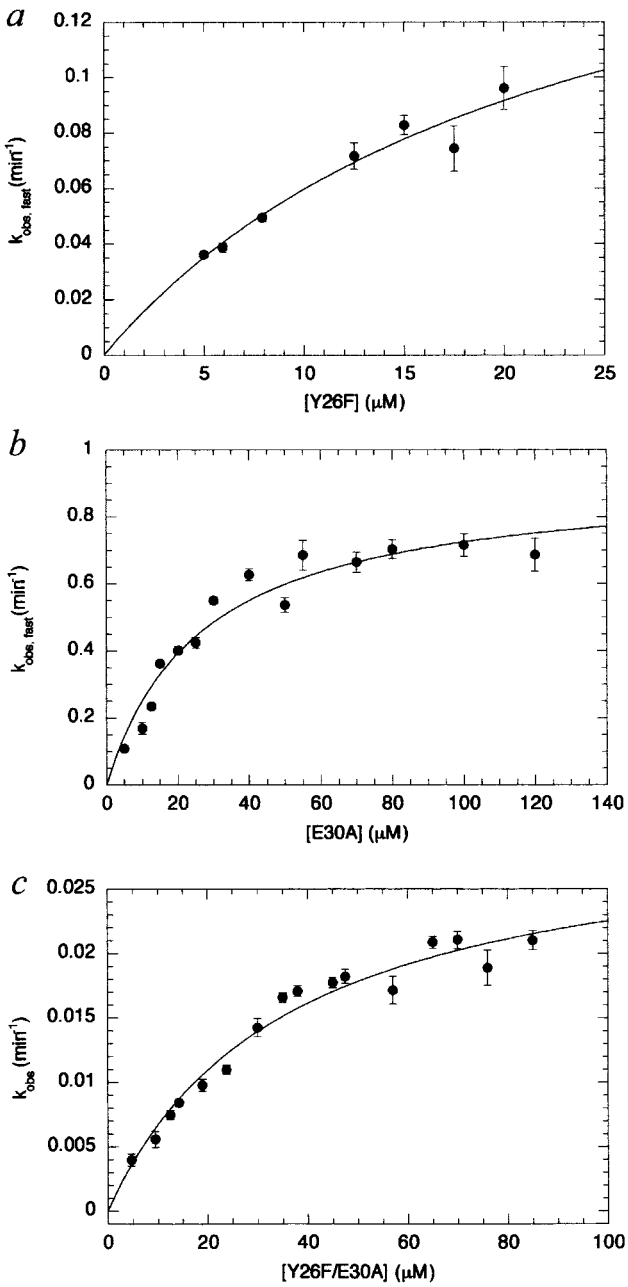

FIGURE 4: Kinetic analysis of Pro-catalyzed folding. Catalyzed folding rate constants are plotted as a function of Pro mutant concentration for αLP folding reactions catalyzed by (a) Y26F, (b) E30A, and (c) Y26F/E30A. Each data point represents the fit from an individual folding reaction like those shown in Figure 3. Enzymatic constants, k_{cat} and K_M, were determined by fitting the data to a variant of the Michaelis−Menten equation (see Materials and Methods) and are summarized in Table 1.

energy transition state of the αLP folding reaction. However, at most, the N-domain mutants combined to slow folding by less than 2 orders of magnitude, a relatively modest effect in the context of the total 3×10^9-fold acceleration in Pro-catalyzed folding. If the Pro N-domain only provides a small portion of the binding energy involved in transition state stabilization, it may not be a necessary component of folding catalysis, as suggested by the single-domain Pro homologues.

N-Domain Deletion Mutant. To determine the overall contribution of the Pro N-domain to αLP folding, we designed a construct consisting of just the Pro C-domain (CPro) to mimic the single-domain pro regions of *S. griseus*

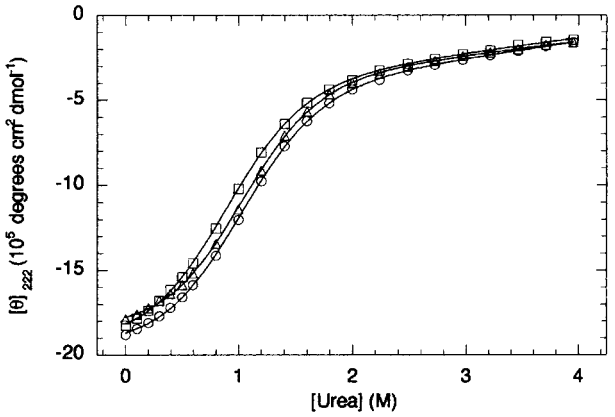

FIGURE 5: Urea denaturation curves for Y26F (○), E30A (△), and Y26F/E30A (□). Nonlinear fits of the data to the linear free energy model, using horizontal folded baselines and linear unfolded baselines, are displayed and were used to determine the $\Delta G_{unfolding}$ values shown in Table 2.

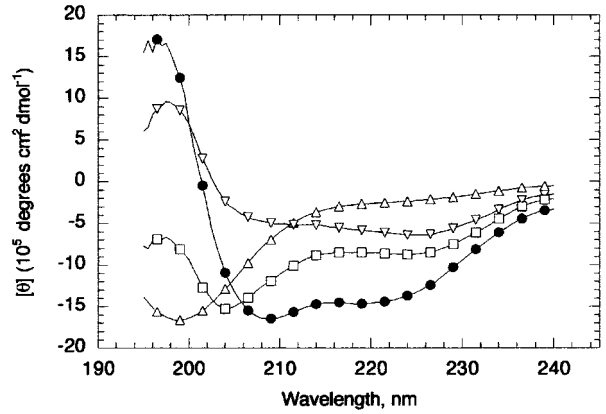

FIGURE 6: CD analysis of CPro structure. The spectrum of CPro alone (△) shows a random coil signature, while the spectrum of native, inactive αLP (▽) displays predominately β-sheet secondary structure. The sum of these two individual spectra (□) is quite different from the spectrum of the preincubated CPro/αLP sample (●), which shows increased helical content similar to that seen for the wild-type Pro·Nat complex structure, demonstrating that CPro becomes structured upon formation of a complex with αLP. Every fifth data point is represented by a symbol to simplify the spectra.

proteases A and B. CPro corresponds to residues S87−T166 of full-length Pro, with an N-terminal His tag added for purification purposes. Initial characterization of CPro by far-UV CD revealed the protein to be a random coil under conditions in which wild-type Pro (*8*) and the Pro N-domain mutants exhibit folded spectra. While no change in secondary structure was observed with the addition of 15% glycerol (data not shown), CPro did become structured upon binding to native αLP (Figure 6). Incubation of CPro with an inactive αLP variant (used to avoid complications due to proteolysis) resulted in a spectrum with helical content greater than that of the predominantly β-sheet spectrum of αLP alone (*6*, *9*). This spectrum differed substantially from the sum of the two individual protein spectra and was consistent with the structuring of CPro upon complex formation with αLP, as the CD signature of native Pro is largely helical, alone or in complex with αLP.

CPro binding to the native protease was further quantified by CPro's ability to inhibit native αLP (Figure 7). CPro competitively inhibited protease activity with a surprisingly tight K_i of 2.9 ± 0.19 nM. This strong inhibition, only 9-fold

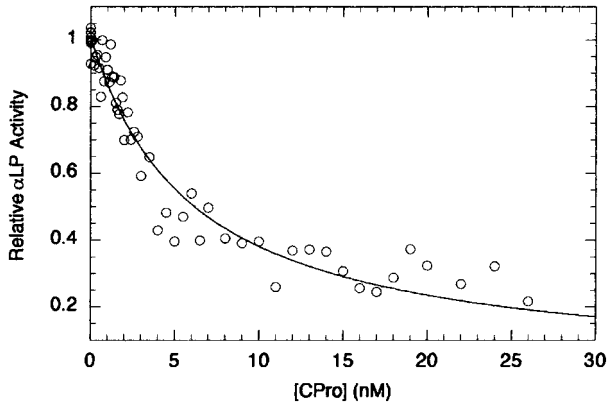

FIGURE 7: Inhibition of wild-type αLP by CPro. Native αLP (0.25 nM) was incubated with increasing amounts of CPro and then assayed for activity (see Materials and Methods). The data were fit to a competitive inhibition variation of the Michaelis−Menten equation, giving a K_i of 2.9 ± 0.19 nM.

weaker than that of wild-type Pro, is remarkable in light of the missing N-domain interactions and the additional loss in binding energy due to CPro instability.

Removal of the N-domain did affect CPro-catalyzed folding, as the production of native αLP could not be detected by standard assay conditions (see Materials and Methods). Monitoring CPro folding of αLP using a more sensitive thiobenzyl ester substrate also failed to measure an acceleration in the folding rate of the protease over that of an uncatalyzed folding reaction control (data not shown). When the detection limits of the assay conditions are considered, these results correspond to a dramatic reduction in the level of catalyzed folding of at least 7 orders of magnitude as compared to that of wild-type Pro. The absence of any measurable rate enhancement prevented a more detailed kinetic analysis that is needed to determine whether k_{cat}, K_M, or both were affected.

DISCUSSION

Previous structural and mutational studies support a model of Pro-catalyzed folding of αLP in which the Pro C-domain initially recognizes and binds the extended β-hairpin of the protease Int state to form a continuous five-stranded β-sheet (*8*, *10*, *12*). Next, binding of the Pro C-terminal tail to the nascent active site positions the β-hairpin and assists in the structuring of the αLP C-domain. The αLP N-domain can then dock and fold against the αLP C-domain to complete the protease active site and the packing of the native state core. This, together with homology data, suggested that the Pro C-domain served as the core folding catalytic unit and raised the question of the function of the Pro N-domain.

Experiments described herein reveal that specific interactions between αLP and the Pro N-domain are also required for efficient folding of the protease. Removing the hydrogen bonding capability of either Pro residue Y26 or E30, via the Y26F or E30A mutation, led to a decrease in the rate of folding catalysis. The Y26F/E30A double mutant showed an even greater loss in folding ability, indicating that these residues contribute cooperatively to the stabilization of the folding transition state. In contrast, these mutations did not significantly alter initial binding to Int, demonstrating that these Pro N-domain residues are not involved in formation of the Michaelis complex.

Deletion of the Pro N-domain produced a completely destabilized protein, Cpro, that possessed only random coil secondary structure under native conditions. The unstructured CPro was incapable of accelerating the αLP folding rate above that of the uncatalyzed reaction, resulting in a catastrophic, at least 10^7-fold, decrease in the rate of catalyzed folding. While the majority of this folding defect is due to the loss of specific N-domain interactions with the protease, CPro instability also debilitates folding by diverting binding energy normally used for stabilizing the transition state toward stabilizing the structure of the CPro catalyst instead. Studies of the full-length Pro-αLP precursor indicate that Pro prefolds before initiating folding of the protease (*17*). A destabilized Pro would be deficient in its ability to maintain a folded scaffold necessary to initiate binding and efficiently fold the protease. Previous in vivo Pro deletion studies found that removal of as few as three amino acids could prevent proper secretion of folded, active αLP to the medium (*21*). While many of these deletions disrupted interactions between Pro and αLP, other seemingly more benign deletions demonstrated similar behavior. These latter deletions may exert their effects by creating instability within Pro, although those experiments could not distinguish between folding and secretion defects.

The role of the Pro N-domain in αLP folding is therefore understood both to impart stability to the Pro C-domain and to contribute direct and significant interactions in the transition state complex. Together, the Pro N- and C-terminal domains position key structural elements in the αLP C-domain, such as the β-hairpin, leading to the structuring of that domain and, finally, the complete folding of the protease to its native state.

Although the Pro N-domain residues highlighted in this study make well-defined interactions with the native state as seen in the structure of the Pro·Nat complex (*10*) and result in significant folding defects when mutated, these mutated residues have only minimal effects on binding to the native state (K_i). This feature is even more pronounced in the complete deletion of the Pro N-domain. Similar mutations, those that alter transition state binding without affecting native state binding, have also been identified via progressive deletions of the Pro C-terminus (*8*). Removal of the last four amino acids results in the near abolition of k_{cat} (reduced by $\sim 10^7$-fold) with a much smaller ~300-fold decrease in K_i.

These seemingly inconsistent results can be resolved by understanding that Pro is not optimized to bind the αLP native state. As a folding catalyst, Pro has evolved to preferentially stabilize the folding transition state to accelerate the αLP folding reaction (the transition state complex is 4.6 kcal/mol tighter than the native state complex). Nonoptimal binding of Pro to native αLP induces strain in the protease, as demonstrated by the increased rate of hydrogen exchange of backbone amides in the protease upon addition of Pro (unpublished results of J. Sohl), and possibly plays an important role in the release and, therefore, subsequent destruction of Pro. Native state binding may create a balance of antagonistic binding between the Pro N- and C-domains, with each domain pulling against the other in an attempt to improve its interface with the protease. In this way, weakening the N-domain interface would be compensated by improved C-domain binding, and vice versa, resulting in no net effect on native state binding. However, these binding defects would be fully realized in the transition state complex.

The most dramatic example of this phenomenon is seen in the binding of CPro to native αLP. The lack of detectable (by CD), folded CPro structure, even in the presence of glycerol, signifies that less than 5% of the CPro population is structured under native conditions, corresponding to a loss in global stability of at least 3.6 kcal/mol compared to wild-type Pro. This alone is sufficient to explain the approximately 10-fold decrease in CPro affinity for native αLP. Incredibly, no additional penalty is observed for removing the substantial N-domain binding interface. We hypothesize that in the absence of the Pro N-domain, Pro C-domain binding is unimpeded, allowing optimization of the Pro C-domain−protease interface and thereby accounting for the unexpectedly high binding energy in the Cpro·Nat inhibitory complex. Without the constraints of Pro N-domain binding, the Pro C-domain may form interactions with the protease that are inaccessible in the full-length Pro·Nat complex, such as filling the solvent-packed gap between αLP and the Pro hinge region (Figure 1a). Similar interactions are also likely to be made in the tighter transition state complex (*10*); thus, the CPro·Nat complex could be an important structural model for the Pro-catalyzed folding transition state.

Despite this proposed transition state-like binding believed to be responsible for the nearly wild-type inhibition of native αLP by CPro, catalysis of αLP folding by CPro was virtually nonexistent. In contrast, some αLP family members, such as *S. griseus* protease B (SGPB), are folded quite well by their own substantially smaller pro regions (*22*). The question of how SGPB achieves a native fold so similar to that of αLP with only "half" of a folding catalyst, whereas CPro is inactive, remains. We propose that while stabilizing the extremely high αLP folding barrier dictates the presence of both Pro N- and C-domains to fold the protease, it is possible that αLP homologues with much smaller pro regions can be efficiently folded as they possess lower, less demanding folding barriers. Furthermore, expression of these small pro regions and their cognate proteases as single polypeptides in vivo effectively supplies the single-domain pro regions at infinite concentration, thus compensating for any deficiency in K_M for these smaller folding catalysts. Future studies that characterize the folding barriers of αLP homologues and the stability and catalytic efficiency of their associated pro regions and that look for relationships between metastability and kinetic stability will provide critical insights into the determinants and evolution of kinetic folding barriers.

ACKNOWLEDGMENT

We thank Peter Chien for his work on testing initial Y26F and E30A refolding conditions, Dr. Sheila Jaswal for providing the S143A inactive αLP variant, Dr. Alan Derman for helpful discussions, and Dr. Luke Rice for his assistance in preparing Figure 1.

REFERENCES

1. Brayer, G. D., Delbaere, L. T. J., and James, M. N. G. (1979) Molecular structure of the alpha-lytic protease from myxobacter 495 at 2 Å resolution, *J. Mol. Biol. 131*, 743−775.
2. Whitaker, D. R. (1970) *Methods Enzymol. 19*, 599−613.
3. Silen, J. L., Frank, D., Fujishige, A., Bone, R., and Agard, D. A. (1989) Analysis of prepro-α-lytic protease expression in *Escheri-*

chia coli reveals that the pro region is required for activity, *J. Bacteriol. 171*, 1320−1325.
4. Silen, J. L., and Agard, D. A. (1989) The α-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*, *Nature 341*, 462−464.
5. Baker, D., Sohl, J. L., and Agard, D. A. (1992) A protein-folding reaction under kinetic control, *Nature 356*, 263−265.
6. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) Unfolded conformations of alpha-lytic protease are more stable than its native state, *Nature 395*, 817−819.
7. Jaswal, S. S., Sohl, J. L., Davis, J. H., and Agard, D. A. (2002) Energetic landscape of α-lytic protease optimizes longevity through kinetic stability, *Nature 415*, 343−346.
8. Peters, R. J., Shiau, A. K., Sohl, J. L., Anderson, D. E., Tang, G., Silen, J. L., and Agard, D. A. (1998) Pro region C-terminus: Protease active site interactions are critical in catalyzing the folding of alpha-lytic protease, *Biochemistry 37*, 12058−12067.
9. Sohl, J. L., Shiau, A. K., Rader, S. D., Wilk, B., and Agard, D. A. (1997) Pro region inhibition by steric occlusion of the α-lytic protease active site, *Biochemistry 36*, 3894−3902.
10. Sauter, N. K., Mau, T., Rader, S. D., and Agard, D. A. (1998) Structure of alpha-lytic protease complexed with its pro region, *Nat. Struct. Biol. 5*, 945−950.
11. Mau, I.-F. T. (1998) Ph.D. Thesis, University of California at San Francisco, San Francisco.
12. Cunningham, E. L., Jaswal, S. S., Sohl, J. L., and Agard, D. A. (1999) Kinetic stability as a mechanism for protease longevity, *Proc. Natl. Acad. Sci. U.S.A. 96*, 11008−11014.
13. Derman, A. I., and Agard, D. A. (2000) Two energetically disparate folding pathways of α-lytic protease share a single transition state, *Nat. Struct. Biol. 7*, 394−397.
14. Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R., and Richards, J. H. (1973) Carbon nuclear magnetic resonance studies of the histidine residue in alpha-lytic protease. Implications for the catalytic mechanism of serine proteases, *Biochemistry 12*, 4732−4743.
15. Mace, J. E., and Agard, D. A. (1995) Kinetic and structural characterization of muations of glycine 216 in α-lytic protease: A new target for engineering substrate specificity, *J. Mol. Biol. 254*, 720−736.
16. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol. 131*, 266−280.
17. Anderson, D. E., Peters, R. J., Wilk, B., and Agard, D. (1999) Alpha-lytic protease precursor: Characterization of a structured folding intermediate, *Biochemistry 38*, 4728−4735.
18. Baker, D., Silen, J. L., and Agard, D. A. (1992) Protease pro region required for folding is a potent inhibitor of the mature enzyme, *Proteins 12*, 339−344.
19. Kraulis, P. (1996) Molscript: A program to produce both detailed and schematic plots of macromolecular structures, *J. Appl. Crystallogr. 24*, 946−950.
20. Merritt, E. A., and Bacon, D. J. (1997) Raster 3D: Photorealistic molecular graphics, *Methods Enzymol. 277*, 505−524.
21. Fujishige, A., Smith, K. R., Silen, J. L., and Agard, D. A. (1992) Correct folding of α-lytic protease is required for its extracellular secretion from *Escherichia coli*, *J. Cell Biol. 118*, 33−42.
22. Baardsnes, J., Sidhu, S., MacLeod, A., Elliott, J., Morden, D., Watson, J., and Borgford, T. (1998) *Streptomyces griseus* protease B: Secretion correlates with the length of the pro peptide, *J. Bacteriol. 180*, 3241−3244.

BI020214O