Interdependent Folding of the N- and C-Terminal Domains Defines the Cooperative Folding of α-Lytic Protease†

Erin L. Cunningham‡,§ and David A. Agard*,‡,§,||

The Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, Graduate Group in Biophysics, University of California at San Francisco, San Francisco, California 94143-2240

Received August 8, 2003; Revised Manuscript Received September 18, 2003

ABSTRACT: α-Lytic protease (αLP) serves as an important model in achieving a quantitative and physical understanding of protein folding reactions. Synthesized as a pro-protease, αLP belongs to an interesting class of proteins that require pro regions to facilitate their proper folding. αLP's pro region (Pro) acts as a potent folding catalyst for the protease, accelerating αLP folding to its native conformation nearly 10^{10}-fold. Structural and mutational studies suggested that Pro's considerable foldase activity is directed toward structuring the αLP C-terminal domain (CαLP), a seemingly folding-impaired domain, which is believed to contribute significantly to the high-energy folding and unfolding transition states of αLP. Pro-mediated nucleation of αLP folding within CαLP was hypothesized to subsequently enable the αLP N-terminal domain (NαLP) to dock and fold, completing the formation of native protease. In this paper, we find that ternary folding reactions of Pro and noncovalent NαLP and CαLP domains are unaffected by the order in which the components are added or by the relative concentrations of the αLP domains, indicating that neither discrete CαLP structuring nor docking of the two αLP domains is involved in the folding transition state. Instead, the rate-limiting step of these folding reactions appears to be a slow and concerted rearrangement of the NαLP and CαLP domains to form active protease. This cooperative and interdependent folding of both protease domains defines the large αLP folding barrier and is an apparent extension of the highly cooperative αLP unfolding transition that imparts the protease with remarkable kinetic stability and functional longevity.

Nearly all extracellular bacterial proteases, as well as a number of vacuolar and lysosomal proteases, are synthesized with pro regions that facilitate the proteases' folding to native conformations (*1*−*4*). One of the most thoroughly studied examples of pro-mediated folding is α-lytic protease (αLP,[1] Figure 1) (*5*); a 198 residue digestive enzyme secreted by the Gram-negative bacterium, *Lysobacter enzymogenes*, to degrade other soil microorganisms (*6*, *7*). Evolved to function in harsh, proteolytic environments, αLP resists degradation via its large and highly cooperative unfolding barrier of 26 kcal/mol ($t_{1/2}$ = 1.2 years) that effectively limits unfolding events that would render αLP susceptible to proteolysis (*8*, *9*). This remarkable level of kinetic stability extends αLP's functional lifetime but carries the costly energetic penalty of an extremely high folding barrier (30 kcal/mol) that has necessitated the coevolution of a 166 residue pro region to mediate folding of the protease to its thermodynamically metastable native state (S. M. E. Truhlar, manuscript submitted and ref *8*).

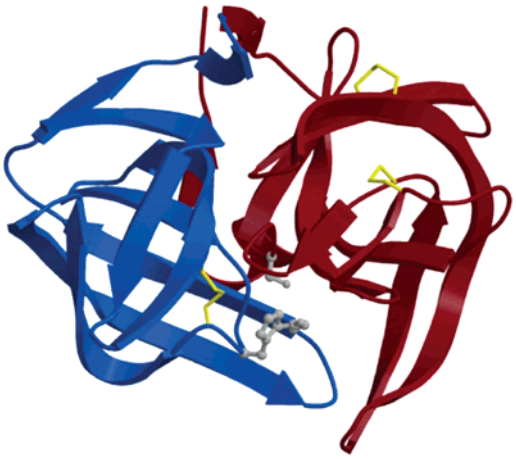

FIGURE 1: Structure of αLP. A ribbon diagram of native αLP displays the two-domain β-barrel topology of the protease, with the N-terminal domain (residues 1−85) and the C-terminal domain (residues 86−198) in blue and red, respectively. The side chains of the catalytic triad are shown in gray and αLP's three intradomain disulfide bonds (17−37, 101−111, and 137−170) are highlighted in yellow. The figure was prepared from the 1.5 Å resolution αLP crystal structure (*25*) using Molscript (*26*) and Raster 3D (*27*).

Alone, αLP folds to a stable, molten-globule folding intermediate (Int) that converts to native protease on the exceed-

† This work was supported by the Howard Hughes Medical Institute.

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

§ The Howard Hughes Medical Institute.

|| Department of Biochemistry and Biophysics.

‡ Graduate Group in 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; TS, αLP folding transition state; NαLP, N-terminal domain of αLP; CαLP, C-terminal domain of αLP; αLP-PSL, αLP with a protease sensitive loop insertion between the two αLP domains; TEV protease, tobacco etch virus protease; CD, circular dichroism; ANS, 8-anilino-naphthalenesulfonic acid.

10.1021/bi035409q CCC: $25.00
Published on Web 00/00/0000 PAGE EST: 7.2

ingly slow time scale of millennia ($t_{1/2}$ = 1800 years); however, in the presence of its N-terminal pro region (Pro), αLP folding is dramatically accelerated 3×10^9-fold, to the order of seconds ($t_{1/2}$ = 23 s) (*8*, *10*). Pro catalyzes αLP folding by preferentially stabilizing the folding transition state by 18.2 kcal/mol, and it also makes the folding reaction efficient by shifting the folding equilibrium in favor of native αLP through tight native-state binding (*8*, *11*, *12*). Structural analysis of this Pro·αLP native-state complex (Pro·Nat) revealed an extensive interface (>4000 Å^2 buried surface area) between the proteins, where Pro surrounds the C-terminal domain of αLP (CαLP), and the Pro C-terminal tail lies in a substrate-like manner in the protease active site (*13*). With only the Pro N- and C-termini contacting the N-terminal domain of the protease (NαLP), Pro binds almost exclusively to CαLP, suggesting that this domain is the focused substrate of Pro foldase activity.

Mutagenesis studies indicate that the structuring of CαLP is an integral part of the high folding barrier, making CαLP a pertinent target for Pro binding and folding catalysis. Screens of libraries of mutagenized αLP reveal that mutations that stabilize the folding transition state all map to CαLP, with mutations in two core residues (R102H and G134S) reducing the folding barrier by 3 kcal/mol (*14*). Indeed, tight packing within the CαLP core appears to play an important role in establishing αLP's large folding and unfolding barriers. Recent ultrahigh-resolution X-ray crystallographic structures of αLP reveal geometric distortions in the CαLP core residue, F181, that are calculated to cost ~4 kcal/mol (C. Fuhrmann, manuscript submitted). Furthermore, the highly buried W147, which interacts extensively with all three of the CαLP core residues discussed, is replaced with smaller residues in most homologues that do not require pro regions for proper folding but is conserved as a large aromatic residue among homologues that possess extremely high folding barriers such as that of αLP.

Similarly, the β-hairpin structural motif found in CαLP is conserved among members of the chymotrypsin superfamily that are synthesized with pro regions but is noticeably absent in homologues that do not need pro regions for folding (*5*). The likely importance of the β-hairpin is also reflected in the strong conservation of the β-strand residues in the pro region with which it pairs in the Pro·Nat interface (*13*). Kinetic analysis of Pro-catalyzed folding reactions using Pro or αLP variants that contain mutations in the β-strand or -hairpin, respectively, identify this interaction as also occurring in the initial binding of the Pro C-terminal domain to Int during formation of the Int·Pro Michaelis complex (*12*). Additional mutational studies suggest that the binding of both the Pro N-terminal and the C-terminal domains acts to arrange the β-hairpin and other key structural elements in CαLP in the folding transition state (TS) (*15*), which in turn is believed to allow NαLP to dock and fold, thereby completing the nascent active site. Activation of protease activity allows efficient intramolecular processing of the Pro-αLP junction. Finally, intermolecular cleavage of Pro by αLP, or other exogenous proteases, leads to its rapid destruction, releasing active protease from the Pro·Nat complex (E. Cunningham, manuscript in press).

In this model of Pro-catalyzed folding, structuring of the individual αLP domains is presumably fairly independent; Pro binding nucleates αLP folding within the folding-impaired CαLP domain, followed by NαLP collision and folding to complete the two β-barrel topology of the native protease. If this general folding scheme is correct, what is the rate-limiting step of the folding reaction? For the mammalian homologues, trypsinogen and chymotrypsinogen, the individual N- and C-terminal domains of these proteins fold independently, with the rate-limiting step being the formation of interdomain disulfide bonds upon collision of the two pre-folded domains (*16*, *17*). Unlike these related enzymes, αLP does not contain interdomain disulfide bonds and is strongly dependent on its pro region for proper folding, suggesting that the folding transition state of αLP may be fundamentally quite different than that of its mammalian counterparts. What then is the physical nature of the rate-limiting folding transition state in Pro-catalyzed αLP folding? Is it the structuring of CαLP, the docking and folding of NαLP, or maybe the proper ordering of the NαLP:CαLP interface, and thus, formation of the protease active site? Furthermore, if CαLP is the substrate for Pro foldase activity, is NαLP folding competent in the absence of Pro?

This paper addresses these unanswered questions by exploring the independent folding behaviors of the individual NαLP and CαLP domains in the absence and presence of Pro. Here, we show that Pro-catalyzed folding of the separated domains reconstitutes active protease, despite the lack of a covalent linkage between NαLP and CαLP, and is insensitive to both the order of addition of the components and the relative concentrations of the αLP domains. Although these ternary (NαLP·CαLP·Pro) complexes appear to form rapidly, structuring of the αLP domains to produce active protease occurs on a much slower time scale. Pro-catalyzed folding of Int is considerably faster, with intact Int demonstrating a greater extent of organized, albeit flexible, structure than the individual, or combined, NαLP and CαLP domains. Last, in sharp contrast to the mammalian serine proteases, αLP folding is defined by the highly interdependent folding of its N- and C-terminal domains and mirrors the extreme cooperativity of the αLP unfolding transition that is vital to αLP's functional longevity.

MATERIALS AND METHODS

αLP-PSL Plasmid Construction. A silent mutation was introduced into the pALP12 plasmid (*18*) using cassette mutagenesis to create a new *Bam*HI restriction site within the αLP gene (pALP12BamHI). The TEV protease recognition sequence (ENLYFQGG) was inserted between αLP residues S83 and S84, which are located within the loop that connects the two αLP domains, by ligating a synthetic oligonucleotide cassette into the pALP12BamHI plasmid digested with *Msc*I and *Bam*HI (pALP12insert). All oligonucleotides were synthesized by the DNA Facility of the Howard Hughes Medical Institute at UCSF, and mutations were verified by sequencing.

Preparation of αLP-PSL and Individual αLP Domains. The αLP-PSL mutant was expressed from the pALP12insert plasmid, where the pro region and αLP-PSL mutant are expressed as a continuous polypeptide, as previously described (*18*, *19*). The protein was purified by S-sepharose ion exchange chromatography as described (*14*, *19*), except that the protein was eluted at pH 9.0.

αLP-PSL fractions were pooled, adjusted to pH 3.0 with 1/5 volume 0.5 M glycine (pH 2.5), and incubated with 0.25

mg/mL pepsin for 6 h at room temperature to proteolyze any residual Pro. Incubation with pepsin also selectively cleaved the αLP-PSL insertion loop after the tyrosine residue, producing the NαLP and CαLP domain fragments without additional incubation with TEV protease. Pepsin-treated αLP-PSL was filtered through a 0.22 μm syringe filter, diluted 3-fold to reduce the salt concentration, and further purified by Mono-S HPLC as described (*19*), except that a 0−250 mM NaCl gradient was used to elute the protein.

Pepsin-cleaved αLP-PSL (αLP-PSL_{cut}) eluted as a single peak discrete from intact αLP-PSL (αLP-PSL_{uncut}). αLP-PSL_{cut} was adjusted to pH 3 with 1/5 volume 0.5 M glycine (pH 2.5) to prevent proteolysis when urea was added to 1.6 M to help dissociate the two αLP domains, which were then separated by reverse phase chromatography. NαLP and CαLP domains eluted as distinct peaks from a semipreparative reverse phase column (Polymer Laboratories, Inc.) using a 5−95% gradient of 95% acetonitrile and 0.1% TFA elution buffer. NαLP fractions were pooled and further purified by a second round of reverse phase chromatography to remove small amounts of CαLP contaminant. Purified NαLP and CαLP fractions were pooled, lyophilized, and stored at 4 °C. Both domains were solubilized in 2 M GdHCl, and aliquots of these denatured stocks were rapidly diluted to ∼30 mM GdHCl to prepare NαLP and CαLP for use in experiments. A portion of HPLC-purified αLP-PSL_{uncut} was also denatured as described (*14*). Sample purity was analyzed using a combination of denaturing SDS−PAGE and MALDI mass spectrometry throughout the preparation.

Pro Region Expression and Purification. Wild-type Pro was purified from *Escherichia coli* strain BL21(DE3)/pLysS as described (*14*) and contains an additional N-terminal proline residue, a cloning artifact that does not affect its behavior (*12*).

Fluorescence Measurements. Tryptophan fluorescence spectra (283 nm excitation) of 1 μM NαLP, CαLP, the combined αLP domains (NαLP + CαLP), Int, native αLP (inactive S143A mutant), and 4 μM Pro in 20 mM potassium succinate (pH 5.6) at 0 °C were recorded, as were ANS fluorescence spectra (380 nm excitation) of the same samples in the presence of 50 μM ANS. Tryptophan fluorescence quenching experiments were performed under the conditions described previously, with either 100 mM I^- (sodium salt) as an external quencher or 100 mM NaCl as a nonquenching control of equivalent ionic strength.

Time-resolved fluorescence measurements of 1 μM NαLP, CαLP, the combined αLP domains (NαLP + CαLP), and Int in 20 mM potassium succinate (pH 5.6) at 0 °C were each recorded at 322 nm, with excitation at 283 nm, upon addition of 16 μM Pro or buffer alone. NαLP and CαLP were also individually preincubated with Pro for 30 min at 0 °C before adding the opposite αLP domain and monitoring fluorescence at 322 nm. Data from these ternary reactions were fit to single exponentials to determine k_{obs} rate constants, whereas Pro-catalyzed Int refolding was fit to a biphasic exponential as previously described (*12*, *14*). Once the fluorescence signal of each sample stabilized, an equilibrium fluorescence measurement was also recorded by determining the average fluorescence plateau value. Plateau data were averaged from 400 to 1200 s after $t = 0$, except for Int + Pro, which was averaged from 800 to 1200 s. All fluorescence measurements were made in a Fluoromax-3 (J. Y. Horiba) connected to an external water bath, except for the fluorescence quenching experiments, which were made in a Fluorolog-3 (J. Y. Horiba), again connected to an external water bath.

Pro-Catalyzed Refolding Assays. The 1.5 μM NαLP and CαLP were preincubated with each other, or with 22.5 μM Pro, for 30 min at 0 °C (ice water) in 20 mM potassium succinate (pH 5.6) prior to addition of the third component. The resulting refolding reactions contained 1 μM NαLP and CαLP and 15 μM Pro in 20 mM potassium succinate (pH 5.6) at 0 °C. Refolding of 1 μM αLP-PSL_{uncut} was also observed under these conditions. For each reaction, production of folded, active αLP was monitored by cleavage of a chromogenic substrate as previously described (*14*).

Data Analysis. Data analysis was performed using Kaleidagraph version 3.6 (Synergy Software) unless otherwise specified.

RESULTS

Production of Separated NαLP and CαLP Domains. Previous in vitro αLP folding studies have utilized unfolded αLP populations with intact disulfide bonds so that the folding kinetics would not be complicated by the slow rate of disulfide bond formation (*8*, *10*, *14*). To produce the individual NαLP and CαLP domains with native αLP disulfide bond pairings (see Figure 1), an αLP mutant containing an inserted protease recognition sequence within the loop connecting the two domains (αLP-protease sensitive loop or αLP-PSL) was expressed with its pro region folding catalyst, purified, and then cleaved using an exogenous protease. In this scheme, αLP-PSL is efficiently folded by the covalently attached Pro, forming the native-state disulfide bonds and producing active protease that is capable of efficiently processing the Pro-αLP junction. Addition of either TEV (tobacco etch virus) protease or pepsin to partially purified αLP-PSL resulted in efficient and highly selective loop cleavage (data not shown), allowing the two individual domains to be separated. However, since pepsin also degrades any residual Pro contaminants, pepsin was employed in subsequent large-scale cleavage reactions.

Pepsin-cleaved αLP-PSL (αLP-PSL_{cut}) was further purified by Mono-S HPLC chromatography and eluted as a single peak that was distinct from that of intact αLP-PSL (αLP-PSL_{uncut}), as shown in Figure 2a. αLP-PSL_{cut} and αLP-PSL_{uncut} display nearly equivalent cleavage activities against chromogenic substrates and are inhibited by Pro with wild-type-like affinities; however, the specific activities of these variants are ∼2-fold less than that of wild-type αLP (data not shown). Having established that the individual αLP domains are capable of functioning together to achieve proteolytic activity without a covalent linkage, the two domains were then purified from one another by reverse phase chromatography (Figure 2b), lyophilized, and resolubilized in 2 M GdHCl.

Characterization of NαLP and CαLP Structure. Aliquots of the denatured NαLP and CαLP stocks were rapidly diluted out of denaturant and analyzed by a variety of spectroscopic methods to determine the extent of structure formed in the individual domains under native conditions. Binding of the hydrophobic dye, ANS, to the separated αLP domains was monitored by fluorescence (Figure 3). ANS fluoresces intensely upon binding to exposed pockets of hydrophobic

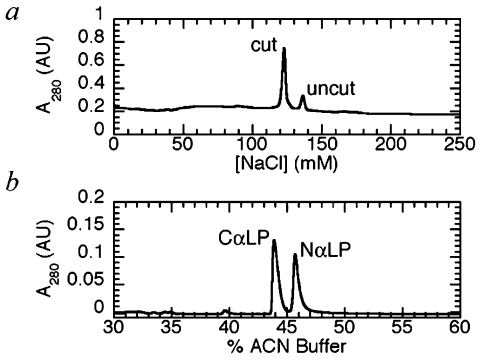

FIGURE 2: Production of separated NαLP and CαLP domains. (a) Mono-S HPLC chromatography of pepsin-treated αLP-PSL. The individual αLP domains remain associated in the αLP-PSL_{cut} species and elute as a single peak (~125 mM NaCl) distinct from that of the remaining uncut material (αLP-PSL_{uncut}). (b) Reverse phase chromatography of αLP-PSL_{cut}. NαLP and CαLP elute as separate peaks between 40 and 50% acetonitrile buffer (see Materials and Methods). The small amount of CαLP contaminant present in the NαLP fractions was removed by a second round of reverse phase chromatography (data not shown).

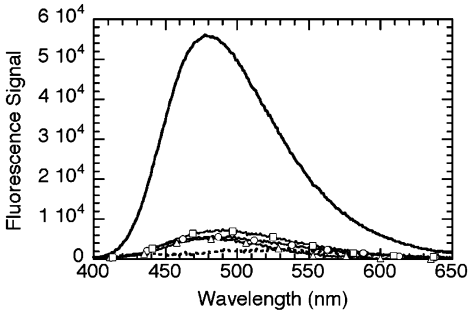

FIGURE 3: ANS binding. The fluorescence of ANS in the presence of the molten-globule-like Int (solid line) is increased and blue-shifted as compared to that of denatured αLP (dashed line). NαLP (open triangles), CαLP (open circles), and the combined domains (open squares) display moderate ANS fluorescence.

surface area that are formed in the presence of organized secondary and tertiary structure, in contrast to its weak fluorescence at longer wavelengths in aqueous solution (*20*). Considered a hallmark of the flexible protein tertiary structure found in molten globules, ANS does not bind effectively to either fully denatured proteins or to the well-packed, and therefore inaccessible, hydrophobic cores of folded proteins. Consistent with these tenets of ANS behavior, denatured αLP shows negligible ANS binding, whereas the molten globule-like Int displays significant ANS binding, with a blue-shift in ANS fluorescence (Figure 3). ANS fluorescence is also shifted to shorter wavelengths in the presence of either NαLP or CαLP, yet the small magnitude of these fluorescence intensities suggests only modest amounts of ANS binding. Combining both αLP domains gives no major enhancement in ANS fluorescence over that of the individual domains. These findings suggest that the separated domains contain a more ordered structure than denatured αLP but significantly less than that found in the intact Int species. Furthermore, circular dichroism (CD) spectra of NαLP and CαLP reveal a combination of random coil and β-strand signatures (data not shown) that are consistent with the idea that the domains, while mostly unfolded, do have some residual secondary structure.

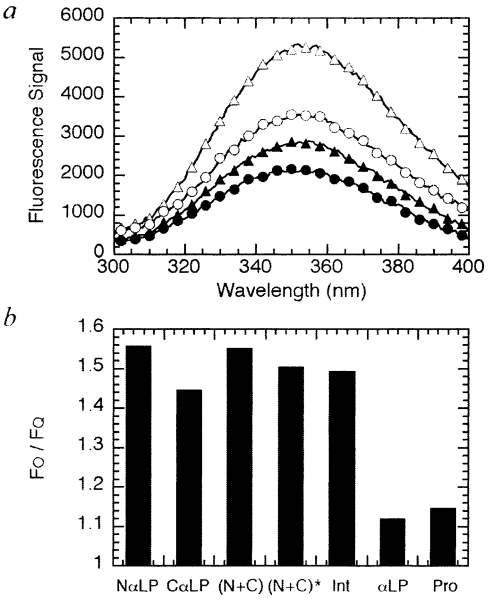

FIGURE 4: Tryptophan fluorescence quenching. (a) NαLP and CαLP fluorescence emission spectra with and without external quencher. In the presence of iodide ion, the fluorescence spectra of NαLP (closed triangles) and CαLP (closed circles) are significantly diminished as compared to spectra recorded in the presence of nonquenching ions (open triangles and open circles, respectively). (b) Comparison of fluorescence quenching ratios. The ratios of nonquenched to quenched fluorescence (F_O/F_Q) at 355 nm are similar for Int (1.49), NαLP (1.56), CαLP (1.45), and the combined domains, whether diluted out of denaturant together ((N + C), 1.55) or combined after dilution ((N + C)*, 1.51). Folded proteins, such as native αLP ($F_O/F_Q = 1.12$) and Pro ($F_O/F_Q = 1.15$), show less sensitivity to the external quencher.

To confirm the formation of partial structure in NαLP and CαLP, we measured the sensitivity of the intrinsic tryptophan fluorescence of each domain to the external quencher (Figure 4a). The fluorescence emission spectra of the single tryptophan residue within each domain were monitored in the presence of the quencher I^- (NaI) and compared to the spectra collected in the presence of nonquenching ions (NaCl). As shown in Figure 4b, folded proteins, such as native αLP and Pro, are rather insensitive to external quencher; however, unfolded proteins, or proteins with fairly plastic tertiary structures, enable the quencher to access the tryptophan residue, thereby decreasing the overall fluorescence intensity. This latter scenario holds true for Int, where the ratio of unquenched to quenched fluorescence intensities (F_O/F_Q) equals 1.49 for this molten globule-like species. Similar quenching ratios of 1.55, 1.45, and 1.55 were found for NαLP, CαLP, and the combined domains, respectively, whereas free tryptophan showed much greater sensitivity to quencher ($F_O/F_Q = 2.27$), suggesting that the tryptophan residues of the individual domains reside in partially protected environments comparable to those found in Int.

Pro-Catalyzed Folding of Separated αLP Domains. Having determined from our initial biochemical characterization of the individual αLP domains that NαLP and CαLP are both well-behaved and marginally structured under native conditions, whether alone or combined, we next sought to evaluate their structure in the presence of the pro region folding catalyst via tryptophan fluorescence. Native αLP demonstrates a substantially elevated and blue-shifted (329 nm maximum) fluorescence signal as compared to the

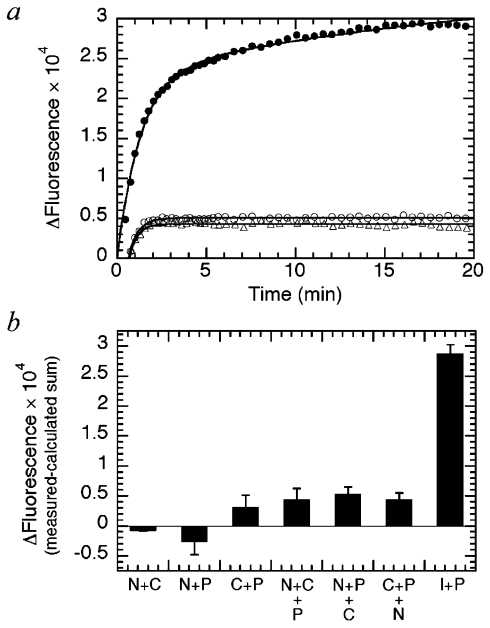

FIGURE 5: αLP fluorescence in the presence of Pro. (a) Time-resolved fluorescence measurements. Pro-catalyzed folding of Int to native αLP (closed circles) results in an increase in fluorescence signal at 322 nm that occurs at a rate ($k_{obs-Int-fluor} = 0.77 \pm 0.03$ min^{-1}) comparable to the refolding rate predicted from a previous kinetic analysis of αLP folding ($k_{obs-calculated} \sim 0.76$ min^{-1}), where the αLP activity was used to monitor folding (*14*). Ternary folding reactions, in which either NαLP + Pro (open circles) or CαLP + Pro (open triangles) were preincubated prior to initiating the reaction, show rapid single-exponential increases in the fluorescence signal ($k_{obs-(NP)+C} = 1.93 \pm 0.27$ min^{-1} and $k_{obs-(CP)+N} = 2.24 \pm 0.16$ min^{-1}). For all data, zero baselines were established by subtracting the sum of the fluorescence of the individual components from the observed fluorescence. Nonzero baseline values are due to slight errors in the calculated fluorescence sums. (b) Equilibrium fluorescence measurements. Average deviations ($n = 3$) of measured, stable fluorescence values from the calculated sums of the individual components are plotted for binary and ternary combinations of NαLP, CαLP, and Pro, as well as an Int + Pro sample (see Materials and Methods).

extremely weak fluorescence spectra of denatured αLP (*8*), Int, and the individual αLP domains that are centered about a 355 nm maximum (see Figure 4a). These spectroscopic differences between the folded and unfolded states of αLP have proven quite useful in measuring αLP unfolding kinetics (*8*) but have not been employed to follow Pro-catalyzed αLP folding because of technical difficulties with this methodology. Specifically, the marginally stable Pro ($\Delta G = 2$ kcal/mol) (*15, 21*) contains two tryptophan residues whose combined fluorescence is highly sensitive to slight temperature fluctuations or any minor environmental changes that shift the folded equilibrium of Pro (*22*). Since Pro is added in >10-fold excess of αLP to approximate first-order folding kinetics, the instability of Pro can easily dominate the observed fluorescence signal. Yet, under carefully temperature-controlled conditions, the significant increase in the fluorescence signal realized upon the Pro-catalyzed folding of Int to Nat can be kinetically evaluated (Figure 5a), yielding an observed folding rate constant comparable to that determined from studies that utilized protease activity to monitor αLP folding (*8*).

Kinetic analysis of the relatively small changes in the fluorescence of binary combinations of Pro with either αLP domain, (NαLP + Pro) and (CαLP + Pro), was impeded

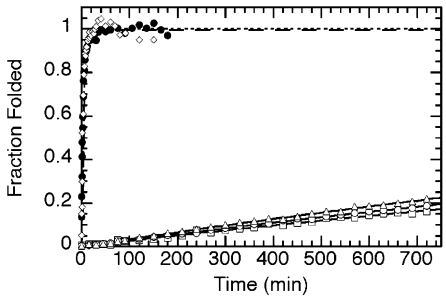

FIGURE 6: Pro-mediated production of active αLP. Binary combinations of NαLP + CαLP (open squares), NαLP + Pro (open circles), and CαLP + Pro (open triangles) were preincubated before adding the third component to initiate the ternary refolding reactions and monitoring αLP folding by protease activity (see Materials and Methods). Linear fits of these data give similar rates of $k_{obs-(NC)+P} = 2.48 \times 10^{-4} \pm 0.05 \times 10^{-4}$ min^{-1}, $k_{obs-(NP)+C} = 2.87 \times 10^{-4} \pm 0.03 \times 10^{-4}$ min^{-1}, and $k_{obs-(CP)+N} = 3.39 \times 10^{-4} \pm 0.03 \times 10^{-4}$ min^{-1}. Pro-catalyzed folding of intact αLP−PSL$_{uncut}$ (open diamonds) and wild-type Int (closed circles) are ∼1300−2000 times faster than the ternary refolding reactions under the same experimental conditions ($k_{obs-αLP-PSLuncut} = 0.40 \pm 0.06$ min^{-1} and $k_{obs-Int} = 0.64 \pm 0.03$ min^{-1}; dashed and dotted lines represents biphasic exponential fits of the αLP-PSL$_{uncut}$ and wild-type Int data (*14*), respectively).

by signal fluctuations, presumably due to the aforementioned sensitivity of Pro fluorescence (data not shown); however, within minutes, these fluorescence signals stabilized, allowing for equilibrium measurements of the samples (Figure 5b). Addition of the omitted αLP domain to the equilibrated binary samples resulted in interpretable time-resolved measurements (Figure 5a) in which each three-component reaction rapidly achieves a stable fluorescence signal that is greater than the sum of its parts and that is reached approximately 2.5-fold faster than the rate observed for Pro-catalyzed Int refolding. Interestingly, the increase in fluorescence signal for these ternary reactions occurs with a rate constant of $k_{obs} \sim 2$ min^{-1}, regardless of which domain is preincubated with Pro.

Equilibrium fluorescence measurements of all binary and ternary samples were taken once the fluorescence signals reached a stable plateau value, and the deviation of these final fluorescence signals from the calculated fluorescence sums of the individual components is illustrated in Figure 5b. All combinations show increased fluorescence over that of the calculated, noninteracting sums, except for the binary samples (NαLP + CαLP) and (NαLP + Pro), suggesting that these two pairings do not interact to induce nativelike structuring of the αLP components. While (CαLP + Pro) does show a modest increase in fluorescence intensity, neither NαLP nor CαLP were observed to compete with Int for Pro binding (data not shown). Yet, due to the concentration limits of NαLP and CαLP, an effect would only have been detectable if the affinity of the individual domains for Pro were stronger than 50 μM (2-fold weaker than that of Int). Thus, any binding of CαLP to Pro occurs with an affinity that is considerably weaker than Pro-αLP binding in the native complex or the folding transition state complex (subnanomolar and femtamolar affinities, respectively).

As the final and ultimate measure of αLP domain folding, ternary reactions of Pro, NαLP, and CαLP were assayed for protease activity as a function of time (Figure 6). While Pro did indeed successfully mediate the folding of the separated domains to reconstitute active protease, the folding rate is

more than 2000 times slower than the rate of Pro-catalyzed Int refolding ($k_{obs-Int} = 0.64 \pm 0.03$ min^{-1}). Preincubation of Pro with either αLP domain was inconsequential to the observed folding rate ($k_{obs} \sim 3.0 \times 10^{-4}$ min^{-1}), congruent with the tryptophan fluorescence studies despite the strikingly disparate time-scales of the two kinetic analyses. Extended preincubation (12 h) of Pro with CαLP and/or addition of NαLP in 5-fold excess of CαLP also had no effect on folding (data not shown). Finally, refolding of intact αLP-PSL$_{uncut}$ ($k_{obs-αLP-PSLuncut} = 0.40 \pm 0.06$ min^{-1}) occurs at a rate similar to that of wild-type Int, indicating that it is the lack of a covalent linkage between the NαLP and the CαLP domains, not the mutations introduced into the interdomain loop, that is responsible for the observed decrease in refolding rate for the ternary refolding reactions.

DISCUSSION

αLP can only fold to its functional native state on a biologically reasonable time scale with the assistance of its pro region folding catalyst. In the absence of Pro, αLP folds to a stable molten-globule folding intermediate that is prevented from converting to the native protease by an extremely large folding barrier (*8*, *10*). Int displays nativelike secondary structure, yet contains little defined tertiary structure, with a hydrodynamic radius that suggests that the two domains of the protease are expanded and separated in comparison to the compact native state (*10*). In this study, characterization of the individual NαLP and CαLP domains reveals these truly separated domains to possess substantially less organized secondary structure than that of intact Int. The domains, whether alone or together, show very modest ANS binding that is only slightly elevated over that of denatured αLP and significantly less than that of Int. This weak ANS binding of the NαLP and CαLP domains suggests that the domains are less structured than in the intact Int species. Consistent with these findings, CD spectra of the noncovalent domains are predominately random coil, even when combined, whereas Int shows a mostly β-sheet signature (*10*).

While structuring of the NαLP and CαLP domains appears minimal, it is enough to partially shield tryptophan residues within the domains from the effects of an external fluorescence quencher. Tryptophan fluorescence spectra of NαLP and CαLP each displayed sensitivity to the external quencher, I^-, analogous to that of Int. Furthermore, whether diluted out of denaturant together, or combined after dilution, stoichiometric mixtures of NαLP and CαLP showed similar levels of sensitivity to I^- as Int and the individual domains. These data indicate that the tryptophan residues located in the separated domains are at least partially buried, such that, even in the isolated domains, they achieve protection from the external quencher that is nearly equivalent to that afforded by the flexible tertiary structure of intact Int. However, although the local environments of these tryptophan residues within the individual NαLP and CαLP domains appear to be comparable to those of intact Int, the overall amount of structuring of these domains is significantly less than that of Int.

The presence of a substantial structure in full-length Int appears to make it significantly more folding competent than the separated NαLP and CαLP domains, with Pro-catalyzed folding of Int to the native, active protease occurring >2000 times faster than that of the divided domains. Mutation of the interdomain loop alone does not appreciably affect refolding, with the rate of catalyzed αLP-PSL$_{uncut}$ refolding only 1.6 times slower than that of wild-type Int. The substantial reduction in refolding rate for the ternary folding reactions therefore appears to be due to the physical separation of the two protease domains. Although considerably slower than the Pro-mediated folding of intact αLP-PSL$_{uncut}$ or wild-type Int, Pro does catalyze folding of NαLP + CαLP mixtures at a rate that is $>10^5$ times faster than the uncatalyzed Int folding reaction. Addition of NαLP in 5-fold molar excess of CαLP had no effect on the observed refolding rate, indicating that bimolecular collision of the two domains is not the rate-limiting step in the ternary folding reaction. Furthermore, our previously proposed model of Pro-catalyzed folding predicted that preincubation of Pro with CαLP would cause structural rearrangements in CαLP, which would in turn facilitate more rapid binding and structuring of NαLP to form active protease than if the order of addition were reversed. However, no such acceleration was observed. Even extended preincubation of CαLP with Pro for as much as 12 h did not result in an increase in the refolding rate, clearly demonstrating that discrete folding of CαLP is not the rate-limiting step of the folding reaction.

Kinetic analysis of the ternary folding reactions, as monitored by increases in tryptophan fluorescence, also produced rates that were insensitive to the order of addition of the three components. Interestingly, while these kinetic fluorescence measurements suggest that the ternary folding reactions of Pro, NαLP, and CαLP rapidly form stable complexes ($k_{obs} \sim 2$ min^{-1}), activity measurements reveal that these complexes only very slowly rearrange to the active form of the protease ($k_{obs} \sim 3.0 \times 10^{-4}$ min^{-1}). Thus, the rate-limiting transition state of the Pro-catalyzed ternary folding reaction appears to involve the simultaneous and interdependent folding of the NαLP and CαLP domains to form native protease.

Together, these studies on the folding behavior of the individual αLP domains provide important insights into the αLP folding mechanism and give rise to the revised model of Pro-mediated αLP folding shown in Figure 7. Contrary to previous assumptions, the N- and C-terminal domains of αLP appear to interact significantly in the intermediate state of the protease (Figure 7a), resulting in substantially more ordered structure in the intact Int species than in the noncovalently linked, but associated, αLP domains. Residual structure present in the separated NαLP and CαLP domains does sequester nonpolar tryptophan residues from solvent to a similar extent as Int but is considerably folding-deficient as compared to Int. Although Pro is known to initiate αLP binding via the C-terminal domain of the protease (Figure 7b) (*12*), CαLP alone is incapable of competing with full-length Int for Pro binding, albeit under concentration-limited assay conditions. Taking into account the sensitivity of this assay, CαLP affinity for Pro is estimated to be at least 2-fold weaker than that of Int ($K_D \geq 50$ μM) and is therefore in marked contrast to the femtomolar binding of the TS·Pro complex. While equilibrium fluorescence measurements suggest that Pro does bind to CαLP, and may in fact induce some nativelike structuring, this putative complex does not lead to substantial folding of the CαLP domain (Figure 7c), and more importantly, folding of isolated CαLP is not the rate-limiting step of the folding reaction. Neither is the

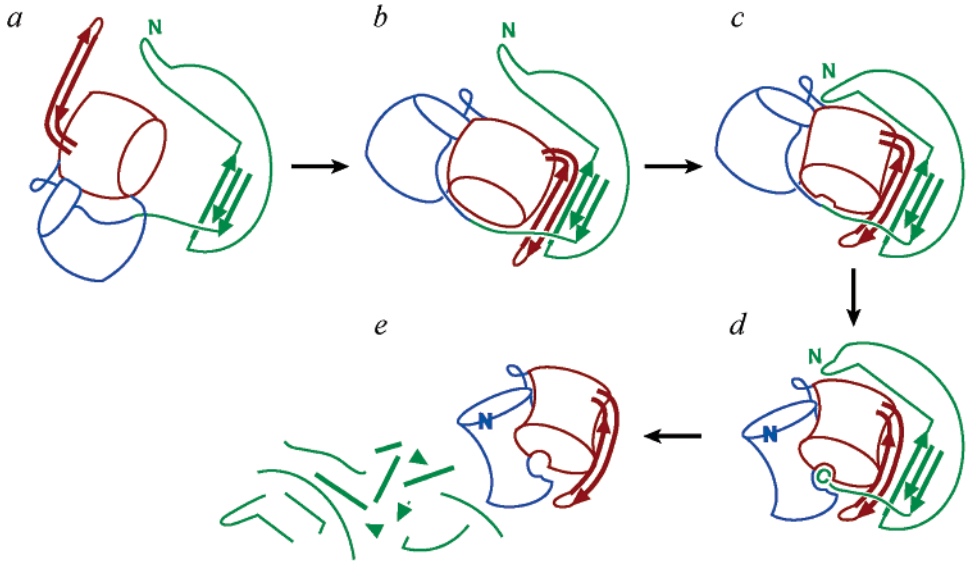

FIGURE 7: Revised model of Pro-catalyzed αLP folding. (a) The pro domain (green) of the Pro-αLP precursor folds separately from the N- and C-terminal domains of αLP (blue and red, respectively), which associate with one another to form substantial local secondary structure. (b) The three-stranded β-sheet from the Pro C-terminal domain pairs with the β-hairpin of CαLP to form a continuous five-stranded β-sheet. (c) Both N- and C-terminal domains of Pro bind to CαLP to help arrange key structural elements, enabling (d) NαLP and CαLP to simultaneously fold. Interdependent folding of the αLP domains completes the protease active site, which can then process the Pro-αLP junction. The new N-terminus of αLP repositions to its native conformation, while the Pro C-terminal tail remains bound to the αLP active site, inhibiting protease activity. (e) Intermolecular cleavage of Pro by αLP, or other exogenous proteases, leads to the degradation of Pro and the release of mature, active αLP.

stepwise docking and folding of NαLP onto CαLP. Furthermore, as previously discussed, the covalently attached NαLP and CαLP domains appear to associate early in the structuring of Int, independent of Pro foldase activity. Instead, this work identifies the interreliant and concerted folding of the NαLP and CαLP domains together to form active protease as the rate-limiting transition state of the Pro-catalyzed αLP folding reaction (Figure 7c,d).

This highly cooperative and simultaneous folding of the two αLP domains is quite different from our original model of sequential αLP domain folding and is also different from the folding mechanisms of related serine proteases. In analogous domain separations of the αLP homologues chymotrypsin and trypsin, zymogen forms of the proteases were selectively cleaved to separate the N- and C-terminal domains of these proteins and to evaluate their individual domain folding behaviors (*16*, *17*). In both cases, the N- and C-terminal domains were shown to fold independently, and once properly folded, to then recognize one another to form productive stable complexes in which interdomain disulfide bonds and the correct active site geometries are formed.

Although the concerted folding of NαLP and CαLP is in direct opposition to the folding mechanisms of these related enzymes, it is entirely consistent with the extreme cooperativity of αLP unfolding (S. M. E. Truhlar, manuscript in press, ref *9*). αLP's large and highly cooperative unfolding barrier limits local and global unfolding events, effectively suppressing proteolytic sensitivity and extending the functional lifetime of the protease. While trypsin, and presumably, chymotrypsin undergo partial unfolding events that allow them to be proteolytically degraded faster than their global unfolding rate (S. M. E. Truhlar, manuscript submitted), αLP is only degraded upon complete unfolding (*9*). The remarkable rigidity of αLP's native state is further reflected in its unusually low crystallographic *B* factors and hydrogen deuterium exchange protection factors of $>10^{10}$ for ~20 amides (*9*), which are among the highest ever measured for a protein (*23*). Importantly, these most slowly exchanging αLP residues are not isolated to a discrete core, as is the case for most proteins (*24*), but are instead spread across both NαLP and CαLP domains. Distribution of this exceedingly stable core across the domain interface may in fact dictate the highly cooperative nature of αLP's folding and unfolding transitions, which in turn determines the functional properties of the native protease.

ACKNOWLEDGMENT

We thank Pinar Erciyas and Caleb Bashor for assistance with protein preparation and John Dueber, Brian Kelch, Dr. Luke Rice, Dr. Andy Shiau, and Stephanie Truhlar for helpful discussions and critical commentary on the manuscript. We thank C. Craik for the use of his fluorometer and Stephanie Truhlar for her assistance in adapting the Craik lab's fluorometer for our experimental needs.

REFERENCES

1. Baker, D., Shiau, A. K., and Agard, D. A. (1993) *Curr. Opin. Cell Biol. 5*, 966−970.
2. Bryan, P. (2002) *Chem. Rev. 102*, 4805−4815.
3. Serkina, A. V., Shevelev, A. B. and Chestukhina, G. G. (2001) *Russ. J. Bioorg. Chem. 27*, 285−305.
4. Shinde, U., and Inouye, M. (1995) *Intramolecular chaperones and protein folding*, R. G. Landes, Austin, TX.
5. Cunningham, E. L., Jaswal, S. S., Sohl, J. L., and Agard, D. A. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 11008−11014.
6. Brayer, G. D., Delbaere, L. T. J., and James, M. N. G. (1979) *J. Mol. Biol. 131*, 743−775.
7. Whitaker, D. R. (1970) *Methods Enzymol. 19*, 599−613.
8. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature 395*, 817−819.
9. Jaswal, S. S., Sohl, J. L., Davis, J. H., and Agard, D. A. (2002) *Nature 415*, 343−346.
10. Baker, D., Sohl, J. L., and Agard, D. A. (1992) *Nature 356*, 263−265.

11. Sohl, J. L., Shiau, A. K., Rader, S. D., Wilk, B., and Agard, D. A. (1997) *Biochemistry 36*, 3894−3902.
12. Peters, R. J., Shiau, A. K., Sohl, J. L., Anderson, D. E., Tang, G., Silen, J. L., and Agard, D. A. (1998) *Biochemistry 37*, 12058−12067.
13. Sauter, N. K., Mau, T., Rader, S. D., and Agard, D. A. (1998) *Nat. Struct. Biol. 5*, 945−950.
14. Derman, A. I., and Agard, D. A. (2000) *Nat. Struct. Biol. 7*, 394−397.
15. Cunningham, E. L., Mau, T., Truhlar, S. M. E., and Agard, D. A. (2002) *Biochemistry 41*, 8860−8867.
16. Higaki, J. N., and Light, A. (1986) *J. Biol. Chem. 261*, 10606−10609.
17. Light, A., and Al-Obeidi, A. M. (1991) *J. Biol. Chem. 266*, 7694−7698.
18. Mace, J. E., Wilk, B. J., and Agard, D. A. (1995) *J. Mol. Biol. 251*, 116−134.
19. Mace, J. E., and Agard, D. A. (1995) *J. Mol. Biol. 254*, 720−736.
20. Royer, C. A. (1995) *Methods Enzymol. 40*, 65−89.
21. Anderson, D. E., Peters, R. J., Wilk, B., and Agard, D. A. (1999) *Biochemistry 38*, 4728−4735.
22. Sohl, J. L. (1997) Ph.D. Thesis, University of California−San Francisco, San Francisco, CA.
23. Huyghues-Despointes, B. M., Scholtz, J. M., and Pace, C. N. (1999) *Nat. Struct. Biol. 6*, 910−912.
24. Li, R., and Woodward, C. (1999) *Protein Sci. 8*, 1571−1590.
25. Rader, S. D., and Agard, D. A. (1997) *Protein Sci. 6*, 1375−1386.
26. Kraulis, P. (1996) *J. Appl. Crystallogr. 24*, 946−950.
27. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
28. Nomenclature: αLP and Pro residue numberings are sequential.

BI035409Q