α-Lytic Protease Precursor: Characterization of a Structured Folding Intermediate†

D. Eric Anderson,‡,§ Reuben J. Peters,‡,|| Barry Wilk, and David A. Agard*

The Howard Hughes Medical Institute and the Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143-0448

Received September 10, 1998; Revised Manuscript Received January 11, 1999

ABSTRACT: The bacterial α-lytic protease (αLP) is synthesized as a precursor containing a large N-terminal pro region (Pro) transiently required for correct folding of the protease [Silen, J. L., and Agard, D. A. (1989) *Nature 341*, 462−464]. Upon folding, the precursor is autocatalyticly cleaved to yield a tight-binding inhibitory complex of the pro region and the fully folded protease (Pro/αLP). An in vitro purification and refolding protocol has been developed for production of the disulfide-bonded precursor. A combination of spectroscopic approaches have been used to compare the structure and stability of the precursor with either the Pro/αLP complex or isolated Pro. The precursor and complex have significant similarities in secondary structure but some differences in tertiary structure, as well as a dramatic difference in stability. Correlations with isolated Pro suggest that the pro region part of the precursor is fully folded and acts to stabilize and structure the αLP region. Precursor folding is shown to be biphasic with the fast phase matching the rate of pro region folding. Further, the rate-limiting step in oxidative folding is formation of the disulfide bonds and autocatalytic processing occurs rapidly thereafter. These studies suggests a model in which the pro region folds first and catalyzes folding of the protease domain, forming the active site and finally causing autocatalytic cleavage of the bond separating pro region and protease. This last processing step is critical as it allows the protease N-terminus to rearrange, providing the majority of net stabilization of the product Pro/αLP complex.

α-Lytic protease (αLP)[1,2] is a 198 residue extracellular serine protease of the trypsin family that is secreted by the soil bacterium *Lysobacter enzymogenes*. αLP is synthesized as a pre-pro-enzyme and targeted to the periplasmic space as a soluble precursor with a large (166 residue) N-terminal pro region (Pro). Upon folding, the Pro-αLP precursor is autocatalyticly processed (*2*) to produce a tight binding noncovalent inhibitory complex [$K_i = 0.3$ nM (*3*)] of Pro and native αLP (Pro/αLP). Active enzyme is subsequently released through degradation of the proteolytically sensitive Pro. Production of active, secreted αLP is completely dependent upon the presence of the Pro region, supplied either as part of the precursor (*4*) or in trans as a separate polypeptide chain (*5*).

In vitro refolding experiments using denatured mature αLP (with its three disulfide bonds intact) also show an absolute requirement of Pro for folding (*6*). In the absence of Pro, denatured αLP folds to an inactive monomeric intermediate (Int) that has all the hallmarks of a stable molten globule: greatly expanded size, significant secondary structure but little or no tertiary structure, and the ability to bind ANS [unpublished results (*6*)]. Int is stable for weeks but readily folds to native, active αLP upon addition of Pro. Kinetic analysis of Int folding ($t_{1/2} \approx 2000$ years) and αLP unfolding ($t_{1/2} \approx 1$ year) demonstrates that Int and the native state are not in thermodynamic equilibrium on any practical time scale (*7*). Surprisingly, the well-behaved native αLP is thermodynamically less stable than either Int or the fully unfolded molecule (*7*). The pro region functions both as a folding catalyst to increase the rate of Int folding by a factor of 3×10^9 and also to shift the thermodynamic equilibrium in favor of native αLP(*7*).

By contrast, little is known about the folding mechanism or the structural and energetic characteristics of the intact pro-protease precursor. It has been shown that cleavage of the covalent bond separating the protease from the pro region was critically important for producing active, secreted protease in vivo (*4*). In particular, in vivo data suggested a strong link between completion of folding and the ability to be secreted through the *Escherichia coli* outer membrane (*2*). Thus, the precursor represents a true intermediate along the vivo folding pathway that may have properties quite distinct from the Pro/αLP complexes previously studied.

Here we present the characterization of the structural and energetic properties of the precursor to determine its role in αLP folding. In addition, we have examined the kinetics of precursor folding and processing. At physiological pH, salt,

† 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.

‡ These authors contributed equally to this study.

§ Current address: ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, MD 21702-1201.

|| Current address: Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340.

[1] Nomenclature: α-Lytic protease residue numbering is by homology to chymotrypsin (*1*). Mutations such as the active site mutant Ser 195 → Ala are denoted SA195.

[2] Abbreviations: αLP, α-lytic protease; Pro, the pro region of αLP; Int, the αLP folding intermediate; CD, circular dichroism; ANS, 8-anilino-1-naphtha-lensulfonic acid; SEC, size-exclusion chromatography; PDI, protein disulfide isomerase.

10.1021/bi982165e CCC: $18.00
Published on Web 03/26/1999

and temperature, the αLP precursor appears to be mostly unfolded. The use of osmolytes and low temperature was required to fully stabilize precursor structure. Under these conditions, precursor exposes excess hydrophobic surface area and contains comparable, but not identical, secondary and tertiary structural organization when compared to the Pro/αLP complex. Furthermore, the precursor folds in a sequential two-step process which seems to correspond to folding of the pro region followed by pro-region-mediated folding of the αLP region. We propose a structural model for the precursor where the pro region is fully folded while the αLP region depends on folded pro region for its structure. Once folded, the protease region cleaves the precursor to form the significantly more stable Pro/αLP complex. This model is discussed in light of a study of the Pro-Subtilisin BPN′ precursor protein (*8*) and our accumulated understanding of αLP folding.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions. The *E. coli* K12 strain DH5α (Focus, 1986, 8:2, 9) was used for cloning experiments, and *E. coli* B strain BL21(DE3) (Novagen, Madison, WI) was used for expression. Growth, induction, and expression were carried out using standard conditions or as published in the pET System Manual fourth edition (Novagen, Inc., Madison WI), using TB media.

Plasmid Construction. Expression constructs for His-tag αLP precursor and Pro region were made using the pET-28a(+) vector (Novagen, Madison, WI). An initial construct, pKK, had been made that introduced a PCR-amplified α-lytic protease pro region (*9*). A *Bam*HI/*Blp*1 DNA fragment excised from the previously constructed pP_{A1}phALP14SA195 (*9*) containing the 3′ end of the Pro region and the complete inactive α-lytic protease gene was ligated into pKK DNA. This converted it into a plasmid, pPALP-S195A, encoding the inactive full-length α-lytic protease precursor used in this paper. A similar strategy was used to construct a plasmid, pPALP, encoding the active protease. All constructs were confirmed by sequencing.

Precursor Preparation. His-tag precursors (pPALP-S195A and pPALP) were expressed as inclusion bodies in BL21-(DE3), inducing at 0.8 OD and expressing for 3−5 h at 37 °C. Inclusion bodies were purified as described (*9*). The inclusion bodies were dissolved in 5 M urea and partially purified using an *S*-Sepharose column under denaturing conditions (5 mM MES, 5 mM Tris, 15 mM NaCl, pH 6.4, and fresh 4 M urea). The peak fractions were adjusted to pH 8.0 with 1 M Tris base and purified using a Ni-NTA superflow IMAC column following the standard protocol (Qiagen). Peak fractions were precipitated by adding 5 vol of cold 100% ethanol. The pellets were washed and then redissolved in 0.1 M Tris (pH 8.0) and 5 M urea. The purified precursor was reduced with 10 mM DTT and then glutathione conjugates formed by the addition of oxidized glutathione to 0.1 M.

To ensure that refolding was performed at very low precursor concentrations, the reaction was carried out by very slowly (0.1−0.2 mL/min) diluting <50 mL of the denatured glutathione conjugated material into 3 L of refolding buffer (15−20% sucrose or glycerol, 0.1 M Tris, 3 mM cysteine, and 200 mM methionine, pH 6.4) at low temperature with constant stirring. After dilution, the folding reaction was allowed to incubate in the cold room for 2−4 days. The solution was next clarified by filtration and adjusted to pH 6.0 by slowly adding 10% cold acetic acid. A batch binding *S*-Sepharose step was used to concentrate the oxidized precursor. The resulting slurry was poured into a column and the precursor eluted using a 0.15 to 1 M NaCl gradient in 10 mM MES, pH 6.4 buffer.

Monomers (∼20% of total material) could be cleanly isolated from the multimers by performing gel filtration using a 7.5 mm × 60 cm Toyo TSK-GS3000 HPLC gel filtration column in 5 M urea, 20 mM MES, 100 mM Tris, 400 mM NaCl, and 20 mM methionine. Denatured monomer was refolded at 5 °C by either slow dialysis against 20% sucrose, 10 mM MES, 20 mM methionine, and 3 mM EDTA, pH 6.0, or by using a SFM-4 stopped flow system (Biologic) operating preperatively with a 1−20 dilution. Folded precursor was concentrated by batch binding to *S*-Sepharose.

Folding Assay for Wild-Type Precursor. Early optimization of the folding protocol for precursor involved tracking the development of αLP activity from folding of wild-type precursor into refolding buffer. Time points were obtained by withdrawing an aliquot from the folding reaction, incubating at room temperature with 1/10th of the volume of 2 mg/mL trypsin (Worthington Biochemical). After a minimum of 5 min of incubation this solution was assayed for the presence of αLP as described previously (*6*) using the chromagenic substrate Suc-Ala-Pro-Ala-pNA (Peptide Institute).

His-Tagged Pro Region. His-tagged Pro region inclusion bodies were prepared as described (*9*). After ultracentrifugation and solubilization, the supernatant was applied directly to a Ni-NTA column and the protein purified as described above for the precursor. The eluted His-tagged pro region was then ethanol precipitated and redissolved in 4 M urea/100 mM Tris/10 mM MOPS, pH 6.5, and dialyzed against several liters of 5 mM MOPS, pH 6.5, precipitated protein was removed by centrifugation, and the refolded Pro region was used without further purification.

Stoichiometric Complex of Pro/αLP. Comparison of the precursor to the noncovalent complex required that an accurate stoichiometric complex be formed between αLP and His-tagged Pro region. To reduce proteolysis of Pro, αLP was treated with excess AEBSF (Sigma) at pH 8 until the assayed activity was reduced at least 10000-fold. His-tagged Pro was then added in approximately 2-fold excess, and the sample was incubated on ice for 30 min. The complex solution (2 mL) was applied to an 80 mL G-25 M gel filtration column equilibrated with 15 mM Aces/15 mM MES, pH 6.0. The sample was concentrated using a 4 mL *S*-Sepharose column preequilibrated with the same buffer and eluted in 1 M NaCl. The complex was further concentrated to 200 μM using a Centriprep-10 (Amicon) and the buffer exchanged using a 4 mL gel filtration spin column equilibrated with buffers containing 50 nM Pro region. Experiments were performed immediately as gradual decay of the complex occurs due to residual proteolytic activity.

CD Measurements. Circular dichroism measurements were made using a J-715 spectroplarimeter (Jasco) with a temperature controlled cuvette holder and a 1 mm path length cuvette. For direct comparisons between the complex and precursor the buffer generally used for this study (10 mM

sodium phosphate, 15% glycerol, 0.1 M NaF, pH 6.4) allowed the precursor to be fully stabilized but limited measurements below 195 nm. Comparisons were also made in the absence of glycerol in the same buffer. Far-UV circular dichroism spectra were analyzed (10) only using circular dichroism data between 205 and 260 nm due to the necessity of using glycerol to stabilize the precursor material.

Free-Energy Estimate. Chemical denaturation curves were obtained using freshly dissolved urea, 10 mM MES (pH 6.4), 0.1 M NaCl, at 5 °C, either with or without 15% weight/volume glycerol (addition of 15% glycerol allowed determination of the folded baseline from 0 to 1 M urea). The data were fit to a two-state transition using a linear free-energy method (*11*) and the software KaleidaGraph (Abelbeck Software). Urea concentrations were determined by refractive index measurement (*11*).

Fluorescence. Fluorescence measurements were made using an SLM-8100 spectrofluorometer and a 5 mm square fluorescence cuvette (no. 4, NSG) at a protein concentration of 0.1 OD/mL, using the salt/glycerol buffer described at a temperature of 5 °C. ANS binding experiments were carried out with the addition of 50 μM ANS and fluorescence emissions were measured between 400 and 600 nm during excitation at 380 nm.

Fluorescence-Monitored Stopped-Flow. Stopped-flow experiments were performed using a Biologic SM-4 stopped flow mixer attached to a Jasco-715 spectropolarimeter with the CD instrument's PMT, placed behind a filter that eliminated light below the wavelength of 320 nm, mounted at a right angle to the incident beam (set at 280 nm, 4 nm slit width) so that it detected fluorescence. Preparations of Pro and oxidized SA195 precursor were made up in 4 M urea with 50 mM NaCl and 5 mM MES (pH 6.4) at a concentration of 4 mg/mL. Refolding was initiated by mixing with 50 mM NaCl, 5 mM MES (pH 6.4), 10% glycerol at a ratio of 1:10. The quenched-flow apparatus, prefiltered and degassed samples, and buffers were maintained at a temperature of 10 °C. Data analysis was done using the program KaleidaGraph with single and double exponential equations.

Processing Time Course. After purification, denatured His-tagged wild-type precursor was exchanged into a buffer containing 5 M urea, 20 mM Tris (pH 8.3), and 5 mM DTT using a G-25M (Pharmacia) desalting column. This material was diluted into 50 parts of a redox refolding buffer (50 mM Tris, 0.3 mM oxidized glutathione, 3 mM reduced glutathione, 50 mM NaCl) either with or without excess (250 μM) PDI and all experiments were performed on ice. At intervals, two aliquots were withdrawn and treated separately: a tenth volume of 0.2 M iodoacetamide was added to one aliquot and incubated on ice until the end of the time course. This prevents further disulfide bond formation and allows properly oxidatively folded precursor proteins to complete processing. The second aliquot was made 0.1 mg/mL in trypsin and incubated on ice for 5 min to remove Pro region from the Pro/αLP noncovalent complex (*6*) and to destroy all unfolded precursor. αLP activity provides an accurate measure of the formation of active mature αLP as a function of time. After the end of the time course, all the iodoacetamide treated aliquots were treated with trypsin in a similar way and αLP activity quantitated. In the absence of trypsinization there was no detectable αLP activity due to Pro inhibition.

RESULTS

Production and Folding of αLP Precursor. Previously, we had tried to express αLP precursor using the signal sequence-directed secretion expression system developed for production of active αLP (*2*). Because precursor processing is normally autolytic (*4*) and almost certainly intramolecular (*9*), production of precursor required either disabling the catalytic machinery (mutating the active site Ser 195 → Ala) or expression at elevated temperatures. Unfortunately, this led to insertion of either wild-type or SA195 precursor into the bacterial outer membrane (*2*), from which purification was extremely challenging. A small amount of reduced and denatured wild-type precursor was partially purified and refolded under redox conditions to give rise to a detectable amount of α-lytic protease activity (*2*). Although not practical for preparative-scale precursor production, these early experiments demonstrated that precursor was autocatalytically processed and that wild-type α-lytic protease could be obtained from a redox refolding reaction. Therefore, we chose to express the precursor as inclusion bodies even though this would require both refolding and formation of the correct disulfide bonds. To further simplify purification, a His_6 tail was added to the N-terminus of the precursor.

Initial refolding experiments utilized wild-type precursor because the processed active protease is easily assayed (after removal of the inhibitory pro region) which facilitated monitoring of folding reaction efficiency. These experiments suggested an approximate pH optimum, a sensitivity to temperature, as well as a preference for glutathione-conjugated material over reduced material in a refolding reaction similar to that described for propapain (*12*). Without optimizing mixing and postfolding purification, recoveries of over 20% could be obtained from the wild-type precursor.

Unfortunately, the inactive precursor variant (SA195) was refolded at much lower efficiency; requiring further optimization of conditions. We suspect that the rapid self-processing of wild-type precursor to the more stable Pro/αLP complex is responsible for the significantly higher yields. SDS−PAGE analysis of the inactive folded precursor indicated that disulfide bonds could be formed even in the presence of a reducing agent (3 mM cysteine), indicating that there is some structural stabilization leading to formation of the correct disulfide bonds. However, significant amounts of high molecular weight disulfide bond cross-linked multimers were also observed. Addition of protein disulfide isomerase did not reduce multimer formation. By contrast, the use of 20% glycerol or sucrose coupled with a mixing protocol designed to minimize local protein concentrations (very slow addition of sample to a large stirred volume) significantly improved oxidative refolding efficiency. The overall efficiency was approximately 20%, with a final yield of folded monomeric precursor of 8 mg/L. A large number of approaches for separating folded monomeric precursor from the multimeric precursors were examined; the only successful method used high-pressure gel filtration under denaturing conditions. Refolding the disulfide intact precursor is quite efficient, with >90% yields.

The SA195 Precursor is Marginally Stable. Equilibrium ultracentrifugation was performed, in the presence of 15% glycerol/10 mM P_i/100 mM NaF, pH 6.4 at 4 °C to maintain the precursor (20 μM) in a soluble form (data not shown).

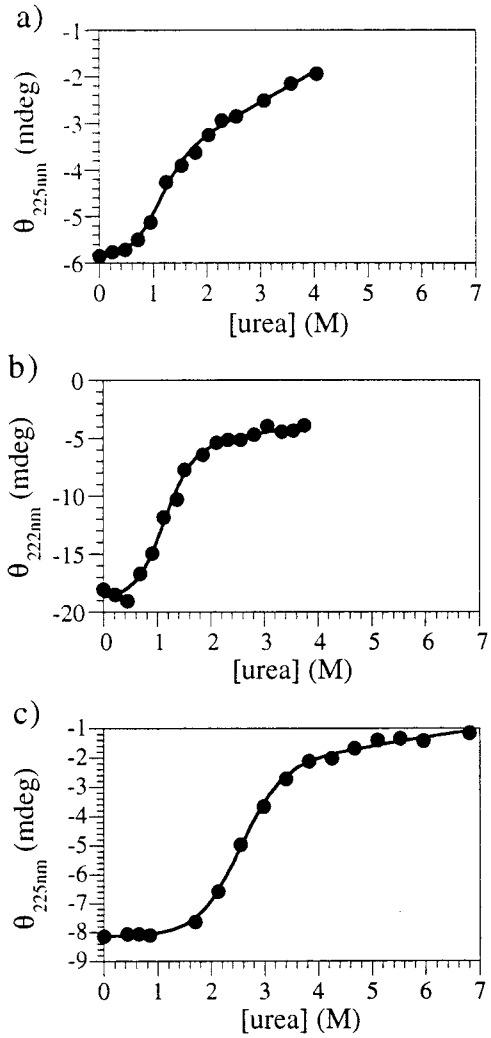

FIGURE 1: Urea denaturation curves. (a) Precursor. (b) Pro. (c) Precursor in the presence of 15% glycerol. Shown are nonlinear fits of the data to the linear free energy model, using a horizontal folded baseline and linear unfolded baseline.

Under these conditions, the calculated molecular weight (39 500 Da) closely approximates that expected (37 900 Da) for monomeric precursor. Therefore, under conditions approximating those used throughout this study, precursor is monomeric.

Precursor stability was measured by monitoring circular dichroism (CD) at 225 nm at 5 °C in the presence of varying amounts of urea (Figure 1a). This curve does not contain a folded baseline at low concentrations of denaturant, even at low temperature, and so it is uncertain if the protein is fully folded. A nonlinear fit of the data using the linear free-energy model with a horizontal baseline for the folded protein and a linear unfolded baseline (Figure 1a), results in fit values of $\Delta G = -2.2$ (± 0.3) kcal/mol and an m value of 2.1 (± 0.3) kcal/mol/mol denaturant. Clearly, the precursor is, at best, marginally stable under these conditions.

Intriguingly, the stability and m value of the precursor agree extremely well with those determined for the isolated Pro protein [Figure 1b; $\Delta G = -2.3$ (± 0.3) kcal/mol, m value $= 2.1$ (± 0.3) kcal/mol/mol denaturant]. These results suggest that the pro region in the precursor may be fully folded and account for the observed early cooperative unfolding transition. Further, the noncooperative loss of precursor 2° structure at higher denaturant concentrations mirrors the behavior of the Int state of mature αLP under similar conditions (*7*). Presumably, the loss of folded pro region allows the αLP region to behave like the Int state of mature protease. Therefore, we propose that the pro region in precursor is fully and independently folded.

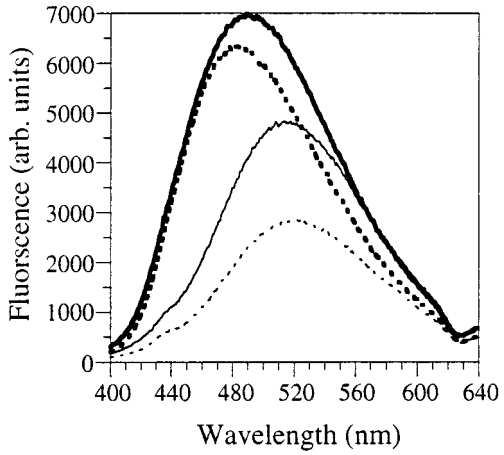

FIGURE 2: ANS binding. The fluorescence of ANS in buffer (thin lines) relative to its fluorescence with precursor present (thick lines) both with (solid lines) and without (dashed lines) 15% glycerol.

In the presence of 15% glycerol, precursor exhibits a folded baseline and undergoes an unfolding transition at higher denaturant concentrations, indicating that it is significantly more stable under these conditions (Figure 1c). A nonlinear fit of these data results in fit values of -3.3 (± 0.2) kcal/mol and 1.3 (± 0.1) kcal/mol/mol for the ΔG and m value, respectively. Clearly, the intact precursor is fully folded in the presence of osmolyte and unfolds in a single, cooperative transition.

The hydrophobic dye ANS is commonly used to indicate the presence of accessible hydrophobic surface area and is often considered the signature of flexible protein tertiary structure. ANS fluoresces only weakly with a maximum near 520 nm in aqueous solutions, but in a hydrophobic environment such as the interior of proteins, it fluoresces much more intensely with a blue-shifted maximum (*13*). Precursor, both in the absence and presence of osmolyte, clearly binds ANS (Figure 2). Which suggests that the hydrophobic core of the precursor is accessible to ANS and is not as well packed as native αLP (which does not bind ANS). As expected, ANS binding to the precursor is enhanced in the absence of glycerol (Figure 2), indicative of an even less well-packed structure. ANS binding is dependent on the presence of organized 2° and 3° structure since denatured precursor does not bind ANS (data not shown). Importantly, ANS fluorescence is enhanced even more by binding to the αLP Int state (J. Nephew, personal communication) than by binding to the Pro-αLP precursor. This indicates that the αLP domain of the precursor is more highly structured than is the Int state of mature αLP. Nonetheless, these results clearly demonstrate that precursor contains accessible hydrophobic surfaces and suggests that at least parts of the precursor have flexible tertiary structure.

Comparison of Precursor and Pro/αLP Complex. The far-UV circular dichroism spectra of αLP precursor and Pro/αLP complex in the absence of osmolytes are similar in overall shape and intensity (Figure 3a). However, relative

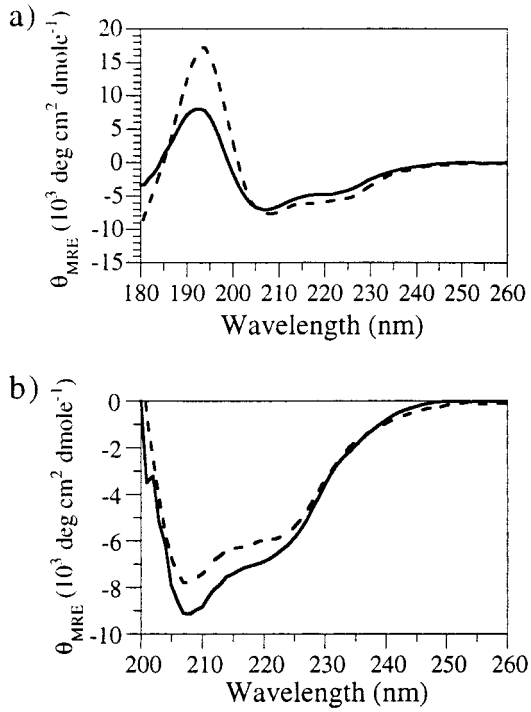

FIGURE 3: Far-UV CD of precursor relative to the Pro/αLP complex. (a) Spectra in the absence of glycerol. (b) Spectra in the presence of 15% glycerol. The Pro/αLP complex spectra is indicated by a dashed line while the precursor spectra is indicated by a solid line.

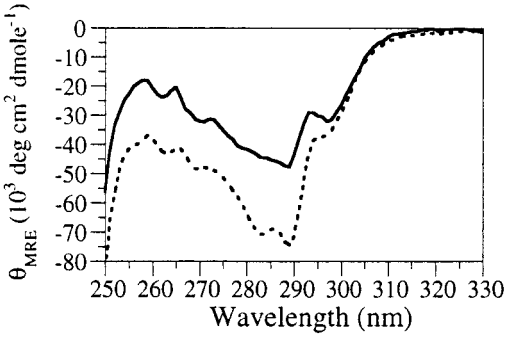

FIGURE 4: Near-UV CD of precursor relative to the Pro/αLP complex in the presence of 15% glycerol. The Pro/αLP complex spectra is indicated by a dashed line while the precursor spectra is indicated by a solid line.

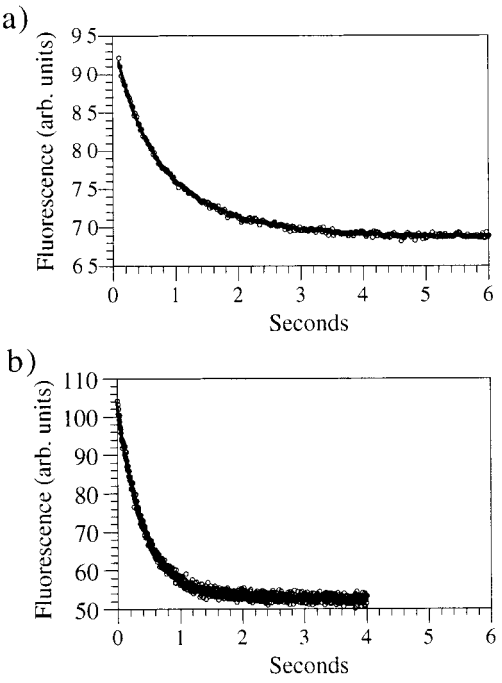

FIGURE 5: Fluorescence monitored refolding after 10-fold dilution out of 4 M urea. (a) Precursor kinetic data fit to a double exponential equation by the shown curve. (b) Pro kinetic data fit to a single-exponential equation by the shown curve.

differences are apparent, with a dramatic difference at 190 nm, as well as a uniform relative reduction in precursor optical activity between 200 and 230 nm. In the presence of 15% glycerol the difference between 200 and 230 nm is reversed and precursor shows more dichroism than the complex (Figure 3b). Due to the absence of short wavelength data in the presence of glycerol, spectral deconvolution was limited to the three general classes of secondary structure: α-helix, *β*-sheet, and disordered (*10*). Furthermore, the precision of the secondary structure composition determinations are limited, and are best used for purposes of comparison. The results (complex, 17% helix; 53% sheet; 32% disordered; precursor, 21% helix; 45% sheet; 32% disordered) indicate that there are no large differences in secondary structure between the precursor and the Pro/αLP complex. However, there does appear to be 4% more α-helix and 8% less *β*-sheet in the precursor, which may reflect structural alterations in the αLP N-terminus that occur after processing.

While the near-UV aromatic CD indicates similarities between the precursor and the complex, it also reveals some significant differences (Figure 4). The spectra share many features, strongly suggesting that precursor folds to a state very similar to the structure of the complex. However, the absence of the peak at 285 nm in the precursor spectrum is informative. This peak in the complex spectra most likely originates from αLP since it is observed in the αLP spectra and not that of Pro (the aromatic CD spectra of isolated Pro and αLP are negative and are essentially additive upon complex formation) (*9*). The observed spectral difference indicates that the local environments of the αLP tryptophans must be different in the precursor relative to the complex. Importantly, the data also clearly demonstrates that the precursor αLP region has more tertiary organization than Int, whose CD spectra has essentially no near-UV signal (personal communication, J. Sohl). Thus, the spectra indicate that the precursor has tertiary structure very similar to that of the complex, but also suggests that the αLP region has a somewhat altered conformation or exhibits some tertiary flexibility. In support of this, preliminary studies using a fluorescent quenching reagent suggest that aromatic residues are more solvent accessible in the precursor than in the complex.

Precursor Folding Shows Biphasic Kinetics. Fluorescence-monitored refolding of the disulfide bond-intact α-lytic precursor is biphasic (Figure 5a), which was clear from examination of the residual errors from the fit to single and double exponential equations. The biphasic fit indicates the presence of two kinetic components having roughly equal magnitudes, and time constants of 0.4 and 1.1 s. Intriguingly, the isolated pro region folds in a single step with a rate

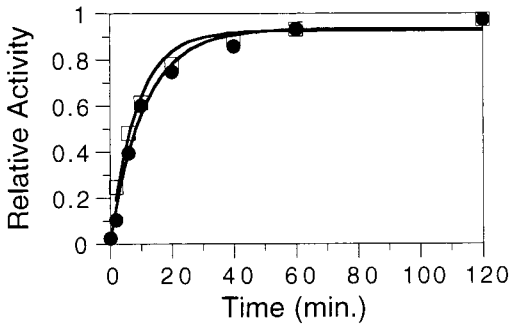

FIGURE 6: α-Lytic protease activity generated by redox folding and processing of wild-type precursor. The ability of wild-type precursor to generate active protease was monitored after 50-fold dilution out of 5 M urea, either by immediately stopping folding reactions (solid circles) or by quenching disulfide formation and letting processing complete (open squares), as described in the methods.

(0.4 s) exactly matching that of the fast phase of precursor folding (Figure 5b). Further, the roughly equal amplitudes (40%, 60%) of these phases also correlates well with the presence of two tryptophans in both the Pro and αLP domains. The observed parallel in folding rates and the correlation between the amplitudes and distribution of tryptophans strongly suggest that the biphasic nature of precursor folding arises from a two step process in which initial folding of Pro is followed by Pro-mediated folding of the αLP domain, generating the final precursor structure.

Kinetics of Disulfide Bond Formation. The formation of native disulfide bonds is a chemically definable moment in protein folding because free cysteines can be readily trapped by reaction with irreversible reagents. We followed the formation of disulfide bonds during folding of the αLP precursor and its subsequent processing to form the mature Pro/αLP complex. Of central interest was the relative timing of disulfide bond formation and the completion of folding and processing of the αLP domain. Figure 6 shows the kinetics of formation of active αLP from a redox refolding reaction of wild-type αLP precursor in which time points were either immediately stopped with trypsin treatment (destroying Pro and all unfolded precursor) or disulfide bond formation quenched and trypsin treatment delayed to allow folding and processing to complete. Within the expected precision of the manual-mixing experiment (one minute), the two preparations appear to yield very similar αLP activity at all times. Control experiments did not detect any inhibition of αLP activity caused by iodoacetamide and the residual urea which is present in the refolding reaction. Since the amounts of αLP produced in the two reactions are similar, it is unlikely that partially oxidized forms of the precursor are competent for processing; this would have lead to differences in either the level of αLP produced from the alkylation reactions and/or disruption of its enzymatic activity by attached alkyl groups. This experiment strongly suggests that autoprocessing and subsequent structural rearrangements occur very rapidly compared to disulfide bond formation. The addition of protein disulfide isomerase to the reaction increases the rate of Pro/αLP formation 3-fold (data not shown), this effect further suggests that disulfide bond formation, and not processing, is rate limiting for complex formation.

DISCUSSION

The folding of αLP is completely dependent on the presence of its pro region both in vivo (*5*) and in refolding reactions in vitro (*6*). Previous in vitro refolding studies, using separate Pro and αLP domains, led to an understanding of the role of the pro region and the discovery of the large barriers to αLP folding and unfolding. However, these studies provided little insight into the in vivo folding reaction that necessarily proceeds via a Pro-αLP precursor. Once folded, the pro-enzyme precursor is subsequently intramolecularly processed to yield the Pro/αLP complex. Proteolytic degradation of Pro then yields active, mature enzyme. Here, we have begun the analysis of the properties and folding pathway of the αLP precursor.

Equilibrium unfolding of the precursor seems to proceed by the following mechanism:

$$\text{Folded} \leftrightarrow \text{Pro}_{\text{U}} - \alpha\text{LP}_{\text{Int}} \leftrightarrow \text{Unfolded}$$

While in the presence of glycerol, precursor unfolding is fully cooperative; in its absence, there is a cooperative transition at low denaturant concentrations followed by a gradual and linear loss of residual structure as denaturant is further increased. The early cooperative transition fits a two-state model with a net stability of $\Delta G = 2.2$ kcal/mol. This stability as well as the *m* value precisely matches that observed for the isolated Pro. The gradual and linear loss of secondary structure at high denaturant concentrations mimics the behavior observed with the αLP molten globule Int state (*7*). Thus, we propose that, in the absence of glycerol, equilibrium unfolding necessarily proceeds through an intermediate where Pro is unfolded while the attached αLP region behaves like the Int state of mature αLP. The addition of glycerol must sufficiently stabilizes both the Pro and αLP regions such that they cooperatively and fully unfold. Significantly, the precursor αLP region is more stable than Int; this enhanced stability must derive from interactions with the attached pro region.

The marginal stability of the precursor contrasts sharply with the highly stable Pro/αLP complex [$\Delta G \approx 10$ kcal/mol (*7*)]. This dramatic difference in stability is the direct result of the cleavage of the peptide bond linking Pro with αLP. This cleavage, and the resultant enhancement in stability, plays a critical role in driving the folding reaction to completion under physiological conditions where the precursor is largely unfolded. Biochemical and structural studies have demonstrated that the C-terminus of the pro region is precisely positioned in a substrate-like manner in the protease active site (*9*, *14*). From this, it follows that in the precursor, the αLP N-terminus, which is covalently linked to the pro region C-terminus, is also positioned in the active site. With a wild-type protease under permissive conditions, we envision that, as soon as the precursor folds, the active site is constituted and the precursor is cleaved to yield the more stable Pro/αLP complex. Thus, the lifetime of the physiologically unstable precursor is minimized. This fits well with our in vivo and in vitro observations: in vivo precursor becomes embedded in the outer membrane, while the Pro/αLP complex efficiently translocates through the outer membrane into the media (*2*), and in vitro refolding of inactive precursor has lower yields and is much more

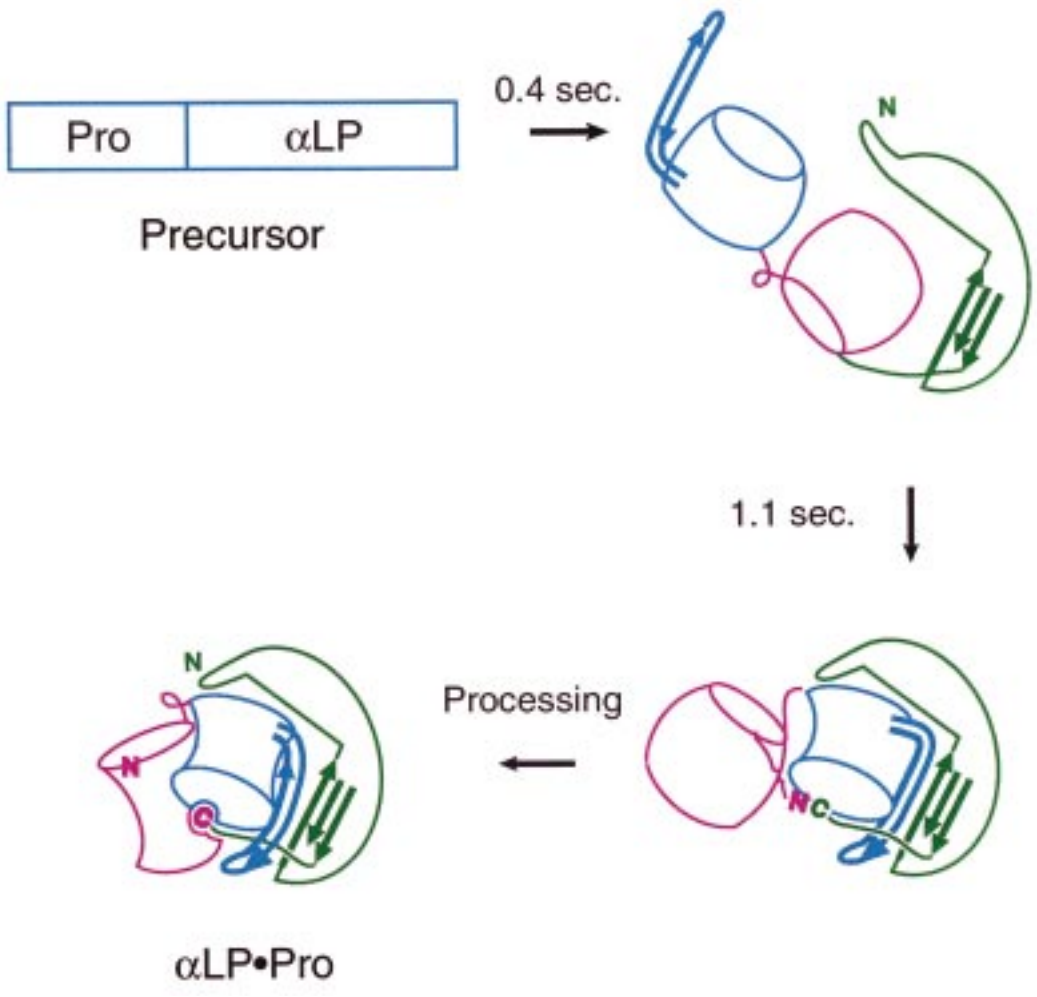

FIGURE 7: The proposed structural and folding model for precursor.

prone to multimerization than production of the active enzyme.

In addition to having a lower stability, the precursor appears to have other properties intermediate between those of a normal folded protein and a molten globule. The precursor shows a ~2× enhancement in ANS fluorescence (Figure 2 vs ~50-fold for a true molten globule like the αLP Int state; J. Nephew, personal communication) and the buried tryptophans are sensitive to the addition of polar fluorescence quenchers (data not shown). The incomplete folding of the precursor must be a direct result of the covalent linkage between the Pro and αLP domains.

While the C-terminus of the Pro region is precisely located in the protease active site in the Pro/αLP complex (*9*, *14*), the αLP N-terminus lies approximately 24 Å distant in both the native αLP (*1*) and the Pro/αLP complex (*14*). Since these residues are covalently joined in the precursor, a significant change must occur in the structure and location of the αLP N-terminus upon precursor cleavage. The precursor configuration of the αLP N-terminus could be destabilizing, or the cleaved conformation could be stabilizing. Support for the latter model comes from the observation that the αLP N-terminus forms the second β-strand of a β-hairpin, which cannot exist in the absence of the N-terminal β-strand. The loss of the β-hairpin is consistent with the CD data indicating that the precursor has more α-helix and less β-sheet than the complex. It is interesting to speculate that the increased α-helix content in the precursor arises from a helical conformation of the αLP N-terminus as it connects to the Pro C-terminus. Such a model would be reminiscent of proteolytic processing in the serpins, where cleavage converts an α-helix into a β-strand which then inserts into a preexisting β-sheet (*16*).

We know that the structure of Pro does not change upon complex formation with αLP (*14*). It thus seems likely that the structure of Pro is also unchanged in the precursor. We propose that the precursor is comprised of fully folded pro region while the αLP region requires folded Pro for its structure and exhibits flexible tertiary organization. Moreover, because Pro makes quite extensive (buried surface area $\approx$ 4000 Å^2) and virtually exclusively contacts with the αLP C-terminal domain in the Pro/αLP complex (*14*), the αLP C-terminal domain is likely to be most stabilized by Pro in the precursor. Therefore, we suggest that it is the αLP N-terminal domain that is least structured and most molten globule-like in the precursor. The loss of structure in the N-terminal domain is presumed to result from the loss of the N-terminal β-strand and other constraints dictated by the covalent linkage.

We further propose a sequential domain-folding model where the pro region necessarily folds first and then acts to fold the attached protease domain. The proposal is supported by the precise equivalence of folding rates for isolated Pro and the fast phase of precursor folding. The equal distribution

of tryptophan residues between Pro and αLP also matches the approximately equal amplitudes observed for the two phases of precursor folding. Indeed, the dependence of αLP region structure on folded pro region almost dictates that the pro region fold first.

Previous studies have led us postulate that the "folding defect" in αLP lies within the αLP C-terminal domain and that the extended β-hairpin in this domain, αLP residues 118−130, plays a central role in creating the high barriers to folding and unfolding (*14*). Chymotrypsin family members synthesized with pro regions all have this conserved β-hairpin, while it is absent or replaced by alternative structures in family members without pro regions. On the basis of this, we suggested that a major function of the pro region is to bind to the β-hairpin, forming an intermolecular five-stranded β-sheet, and then, with the assistance of the Pro C-terminus, to correctly position the β-hairpin, resulting in a folded αLP C-terminal domain (*14*). We believe a similar model is appropriate for precursor folding (Figure 7). First, the pro region folds, which then interacts with the β-hairpin and stabilizes structure in the αLP C-terminal domain. As the αLP C-terminal domain becomes structured, this creates an interaction surface to stabilize the formation of structure in the αLP N-terminal domain. Either strain or loss of the N-terminal β-strand prevents either the αLP N-terminal domain by itself or both αLP domains from completely folding until processing occurs. While the covalent linkage between the Pro and αLP greatly destabilizes the precursor relative to the complex, it also leads to a 30-fold increase in folding rate relative to Pro catalyzed folding of Int state mature αLP (*7*). Since binding of the Pro C-terminus takes place in the rate-limiting folding transition state (*3*), the Pro-αLP linkage could increase folding rates by optimally positioning the Pro C-terminus.

Although not evolutionarily related, subtilisins also require a pro region for efficient folding. Fersht and co-workers have characterized the properties of a catalytic serine to alanine variant of pro-subtilisin which is unable to self-process (*8*). While we are able to compare the αLP precursor with the Pro/αLP complex as well as isolated Pro, owing to the instability of the isolated subtilisin pro region, it was only possible to make comparisons between the subtilisin precursor and native subtilisin. Nevertheless, they present a model of pro-subtilisin generally similar to ours; near- and far- UV circular dichroism spectra of subtilisin precursor are shown to be generally similar to native subtilisin. Further, they also observe an equilibrium folding intermediate and demonstrate that all tertiary structure formation occurs during folding of this intermediate. The nature of this folding intermediate is suggested to be similar to the subtilisin Int state, similar to our proposed intermediate. The two precursors also share a lack of stability, both require the addition of a cosolute for stabilization, which is unusual for biologically active proteins. One reason that these two precursors might lack significant equilibrium stability is that they are transient conformations that self-process to generate highly stable protease/inhibitor complexes. It is only necessary that the proteolytically active precursor conformation be kinetically accessible in order for processing to occur.

The large change in stability observed upon processing of the αLP precursor suggests that processing of other precursor molecules, such as viral fusion glycoproteins, may play an important role in creating new and highly evolved energetic landscapes in these systems. In particular, cleavage of the peptide chain can drive the formation of metastable states such as observed with αLP (*7*) and the neutral pH form of the influenza virus fusion glycoprotein hemagglutinin (*15*).

ACKNOWLEDGMENT

We especially wish to thank Dr. Alan Derman for his critical comments on the manuscript, and Sheila Jaswal for assistance with the ANS experiments.

REFERENCES

1. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D., and James, M. N. G. (1985) *J. Mol. Biol. 184*, 479.
2. Fujishige, A., Smith, K. R., Silen, J. L., and Agard, D. A. (1992) *J. Cell Biol. 118*, 33−42.
3. 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.
4. Silen, J. L., Frank, D., Fujishige, A., Bone, R., and Agard, D. A. (1989) *J. Bacteriol. 171*, 1320−1325.
5. Silen, J. L., and Agard, D. A. (1989) *Nature 341*, 462−464.
6. Baker, D., Sohl, J. L., and Agard, D. A. (1992) *Nature 356*, 263−265.
7. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature 395*, 817−819.
8. Eder, J., Rheinnecker, M., and Fersht, A. R. (1993) *J. Mol. Biol. 233*, 292−304.
9. Sohl, J. L., Shiau, A. K., Rader, S. D., Wilk, B., and Agard, D. A. (1997) *Biochemistry 36*, 3894−3902.
10. Sreerama, N., and Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
11. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
12. Taylor, M. A., Pratt, K. A., Revell, D. F., Baker, K. C., Sumner, I. G., and Goodenough, P. W. (1992) *Protein Eng. 5*, 455−459.
13. Royer, C. A. (1995) *Methods Mol. Biol. 40*, 65−89.
14. Sauter, N., Mau, I. T., Rader, S., and Agard, D. A. (1998) *Nat. Struct. Biol. 5*, 945−950.
15. Carr, C. M., Chaudhry, C., and Kim, P. S. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 14306−14313.
16. Goldsmith, E. J., and Mottonen, J. (1994) *Structure 2*, 241−244.

BI982165E