Pro Region C-Terminus:Protease Active Site Interactions Are Critical in Catalyzing the Folding of α-Lytic Protease†

Reuben J. Peters,‡ Andrew K. Shiau,‡ Julie L. Sohl,§ D. Eric Anderson,‡ Gale Tang,|| Joy L. Silen,⊥ and David A. Agard*,‡

The Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, Graduate Group in Biophysics, and Department of Physiology and Biochemistry, University of California, San Francisco, California 94143, and Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, California 94304

Received April 20, 1998; Revised Manuscript Received June 26, 1998

ABSTRACT: α-Lytic protease is encoded with a large (166 amino acid) N-terminal pro region that is required transiently both in vivo and in vitro for the correct folding of the protease domain [Silen, J. L., and Agard, D. A. (1989) *Nature 341*, 462−464; Baker, D., et al. (1992) *Nature 356*, 263−265]. The pro region also acts as a potent inhibitor of the mature enzyme [Baker, D., et al. (1992) *Proteins: Struct., Funct., Genet. 12*, 339−344]. This inhibition is mediated through direct steric occlusion of the active site by the C-terminal residues of the pro region [Sohl, J. L., et al. (1997) *Biochemistry 36*, 3894−3904]. Through mutagenesis and structure−function analyses we have begun to investigate the mechanism by which the pro region acts as a single turnover catalyst to facilitate folding of the mature protease. Of central interest has been mapping the interface between the pro region and the protease and identifying interactions critical for stabilizing the rate-limiting folding transition state. Progressive C-terminal deletions of the pro region were found to have drastic effects on the rate at which the pro region folds the protease but surprisingly little effect on inhibition of protease activity. The observed kinetic data strongly support a model in which the detailed interactions between the pro region C-terminus and the protease are remarkably similar to those of known substrate/inhibitor complexes. Further, mutation of two protease residues near the active site have significant effects on stabilization of the folding transition state (k_{cat}) or in binding to the folding intermediate (K_M). Our results suggest a model for the α-lytic protease pro region-mediated folding reaction that may be generally applicable to other pro region-dependent folding reactions.

α-Lytic protease (αLP)[1,2] is a 198-residue chymotrypsin-like serine protease secreted to the extracellular environment by the soil bacterium *Lysobacter enzymogenes*. The protease is encoded with a 166-residue N-terminal pro region (Pro) whose coexpression (either in cis as a contiguous peptide or in trans as a separate polypeptide) is transiently required for production of active protease in vivo (*3*). Similarly, in vitro experiments show that denatured mature protease requires Pro to refold. Chemically denatured αLP rapidly diluted into buffer folds to an inactive state that is stable for weeks yet rapidly folds to the native state upon addition of Pro (*4*). This inactive or intermediate state (Int) exhibits many of the characteristics of a stable molten globule: significant secondary structure, no detectable tertiary structure, and a greatly expanded hydrodynamic radius relative to the native state (*4*). Remarkably, the inability of αLP to fold in the absence of Pro indicates the presence of a barrier of >27 kcal/mol that effectively blocks folding of Int (*4*, *5*). The pro region functions to facilitate folding by directly catalyzing the folding reaction (folding is accelerated by a factor of $>10^7$) and by stabilizing the native state ($K_i \sim 10^{-10}$ M; *6*). Once the protease has folded, the pro region is rapidly destroyed by proteolysis and the protease assumes a kinetically trapped native state. Energetic analysis of the folding reaction indicates that Pro binds the folding transition state ~1000-fold more tightly than it binds the native state (*4*, *5*), indicating that it has evolved to be optimally complementary to the folding transition state.

This paper is an initial examination of the mechanism by which αLP is folded by its pro region. Of central interest is probing the different interactions between the protease domain and the pro region that occur throughout the folding

† This work was supported by the Howard Hughes Medical Institute. A.K.S. was a Howard Hughes Medical Institute predoctoral fellow, and J.L.S. was a National Science Foundation predoctoral fellow.

* Corresponding author. Tel: (415) 476-2521. Fax: (415) 476-1902. E-mail: agard@msg.ucsf.edu.

‡ The Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, UCSF.

§ Graduate Group in Biophysics, UCSF.

|| Department of Physiology and Biochemistry, UCSF.

⊥ Affymax Research Institute.

[1] Nomenclature: α-Lytic protease residue numbering is by homology to chymotrypsin (*1*). Mutations such as Val 167 → Ile are denoted V167I. The P and S nomenclature (*2*) is used to denote substrate residues and binding sites. P_1 is the substrate residue before the sissile bond; P_2, P_3, etc. extend toward the N-terminus. S_1, S_2, etc. are the corresponding binding sites on the enzyme.

[2] Abbreviations: αLP, α-lytic protease; V167I, αLP with Val167 replaced by Ile; N170K, αLP with Asn170 replaced by Lys; Pro, the pro region of αLP; ProWt, wild-type pro region; Pro-1 to Pro-4, pro region mutants with one to four amino acids deleted from the C-terminus; ProQ-3A, pro region mutant with the Gln three residues from the C-terminus (Q-3) replaced by Ala; CD, circular dichroism; Int, folding intermediate state of αLP; Int‡, refolding transition state of αLP; RMSD, root-mean-square deviation.

S0006-2960(98)00883-6 CCC: $15.00
Published on Web 08/14/1998

reaction. By direct analogy with enzyme kinetics, we might expect unique interactions to occur in the formation of the initial binding (or Michaelis) complex with Int (Int·Pro), the folding transition state complex ($Int^{‡}$·Pro), and the inhibitory complex with the native enzyme (Nat·Pro). These key functional states can be differentiated through kinetic analysis of the folding reaction, which demonstrates Michaelis−Menten behavior. Further, interactions specific to various states of the αLP·Pro complex throughout the folding reaction can be mapped through mutagenesis of both the pro region and the protease. Of particular importance are mutations or deletions that specifically affect the $Int^{‡}$·Pro complex. Such residues must be involved, either directly or indirectly, in catalyzing the folding reaction and are thus of great mechanistic interest.

The C-terminal residues of Pro and the residues near the active site of αLP are particularly interesting areas for investigation of the interactions between Pro and αLP. It has recently been determined that Pro inhibition of αLP is mediated by steric occlusion of the αLP substrate binding cleft by the C-terminal residues of Pro, demonstrating that such interactions exist in the native state complex (*7*). In vivo studies also suggested that C-terminal deletions of the pro region could impair production of active protease (*31*). In this study we demonstrate that Pro region C-terminal deletions and mutations in αLP surface residues near the active site specifically affect the refolding reaction. On the basis of these results, we suggest a mechanistic model of how pro regions catalyze refolding.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. Wild-type αLP was produced in *L. enzymogenes* 495 (ATCC 29487) as described (*8*). Mutagenesis of αLP was carried out using the plasmid pALPE (*7*). Mutant αLP was produced in *Escherichia coli* bacterial strain MH1 carrying the appropriate pALP5-derived plasmid as previously described (*9*). Wild-type and mutant Pro regions were expressed with a T7 expression system, using bacterial strain BL21(DE3)pLysS (*10*), as described previously (*7*).

Cassette Mutagenesis. Mutagenesis of targeted codons was carried out by the synthesis of a "doped" cassette (a *Kpn*I/*Stu*I restriction fragment) containing the codons for residues V167, T168, and N170, which were mutated. The corresponding oligonucleotides were synthesized with a mixture consisting of 85% of the naturally occurring nucleotide and 15% of the remaining dNTP's in the reaction steps for the chosen codons but otherwise contained the wild-type sequence. The "doped" cassette was ligated into pALPE, replacing the wild-type fragment, thus generating a library of αLP mutants at the chosen codons. Given the level of "doping", ∼25% of the colonies generated by transformation of this library were expected to be wild type or silent mutations; in reality only ∼10% of the ∼1000 colonies screened exhibited wild-type activity levels.

In Vivo Characterization of α*-Lytic Protease Expression and Activity.* Detection and characterization of αLP activity in vivo was carried out using a plate assay to screen the colonies for protease activity (*12*). As a secondary screen, more exact activity measurements were made using liquid cultures. Cultures were inoculated with a single colony, and

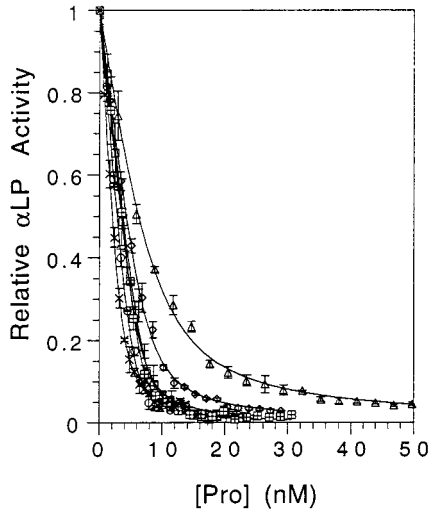


FIGURE 1: Mutant inhibition studies. The pro region was added to native protease in the indicated amounts and bound 3−5 min at 25 °C. Chromogenic substrate was added and the rate of catalytic activity determined. Activity was normalized to the activity in the absence of the pro region. K_i was calculated from fitting the data to the classic tight-binding inhibitor equation (see Materials and Methods); the fit to the data for each is shown. Each data point was determined in triplicate and is shown as the average with the standard deviation error. The variants are represented as follows: ProWt (open circles), Pro-1 (open squares), Pro-2 (open diamonds), Pro-3 (open triangles), ProQ-3A (closed circles), V167I (×), and N170K (crosses).

αLP was expressed via phoA induction as described (*9*). One milliliter aliquots were taken, and supernatant was harvested by spinning out the cells and assayed for protease activity. Initial velocity measurements were taken using succinyl-Ala-Pro-Ala-*p*-nitroanilide (Peptide Institute, Inc.) as a chromogenic substrate, as described (*9*). The amount of mature αLP present was determined by ELISA and compared to a wild-type standard of known concentration as described (*11*).

Production and Purification of α*-Lytic Protease.* Wild-type αLP was purified from cultures of *L. enzymogenes* 495 (ATCC 29487). Mutant αLP was purified from *E. coli* strain MH1 carrying pALP5-derived plasmids containing the mutant version of αLP. Large-scale preparations of the mutant proteases were carried out using a 12 L fermentor to grow 10 L cultures for phoA-directed expression as described (*9*). It is interesting to note here that V167I and N170K mutant αLPs were produced much better at 12 °C rather than the 18 °C normally used for wild-type production. Protease was purified from culture filtrates using both low-flow and HPLC ion-exchange chromatography as described (*12*).

Production and Purification of the Pro Region. The wild-type pro region and all C-terminal deletion mutants were purified from *E. coli* strain BL21(DE3)/pLysS transformed with the appropriate plasmid as described (*7*).

Protease Inhibition Studies. Inhibition by the pro region was investigated by assaying αLP activity in the presence of increasing amounts of purified pro region, as described by Baker et al. (*6*), except that the purified pro region concentrations were initially estimated using a linear extrapolation of the initial portion of the inhibition curve and then refined during the fitting of inhibition curves to values within 20% of the initial estimates. The apparent K_i was estimated by fitting the inhibition data, using the Macintosh

program KaleidaGraph (Abelbeck Software), to the tight-binding inhibitor equation:

$$v = \frac{k_{cat}[S]}{K_M + [S]} \times \frac{[\alpha LP] - [Pro] - K_i + \sqrt{(K_i + [\alpha LP] + [Pro])^2 - 4[\alpha LP][Pro]}}{2}$$

also as described (*6*). The reported apparent K_i of the pro region for the active protease in the context of each of the mutants we describe here (except Pro-4) is the average of at least three experiments such as the ones shown in Figure 1.

The large reduction in inhibition constant for Pro-4 allowed the use of Lineweaver−Burk plots to determine its K_i. αLP (4 nM) was incubated in the presence and absence of 40 nM Pro-4, and the rate of substrate hydrolysis was determined in the presence of varying amounts of substrate; again all data points were collected in triplicate. The data in the absence of Pro were fit to the Michaelis−Menten equation; in the presence of Pro the data were fit to a modified Michaelis−Menten equation reflecting the presence of inhibitor as follows:

kinetic mechanism

$$\alpha LP{\cdot}Pro \underset{+Pro}{\overset{K_i}{\rightleftharpoons}} \alpha LP \underset{+substrate}{\overset{K_M}{\rightleftharpoons}} \alpha LP{\cdot}substrate \xrightarrow{k_{cat}} \alpha LP + products$$

rate equation

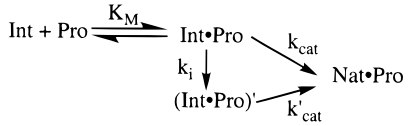
$$v = \frac{k_{cat}[\alpha LP][Pro]}{[Pro] + K_M(1 + [\alpha LP]/K_i)}; \qquad K_M = \frac{k_{-1} + k_{cat}}{k_1}$$

Surface Plasmon Resonance Binding Experiments. Binding experiments were carried out using a BiaCore (Pharmacia BioSensor). Int and native αLP, in 20 mM succinate at pH 6, were immobilized using standard EDC/NHS cross-linking chemistry. Pro was dialyzed into HBS (10 mM HEPES, 150 mM NaCl, 2.5 mM EDTA, pH 7.4) and P20 surfactant (Pharmacia BioSensor) added to 0.0005% prior to use. P20 surfactant (0.0005%) was added to degassed and 0.2 μm filtered HBS and used as the running buffer for all manipulations of the system. Injections of 10 mM HCl were used to regenerate the binding surfaces between Pro injections.

Reconstitution of Pro Region-Dependent Folding. α-Lytic protease was concentrated by ultrafiltration through a 10 kDa MWCO membrane to ∼50 mg/mL, then denatured by addition of 3 volume equivalents of 8 M guanidine hydrochloride and, 0.67 M glycine, pH 3.0, allowed to incubate for at least 12 h at 25 °C, and reconcentrated by ultrafiltration (10 kDa MWCO) to ∼50 mg/mL. The stable folding intermediate (Int) was made by a 1:200 dilution of this concentrated denatured αLP into 20 mM $NaPO_4$, pH 7.2, as described (*4*).

Wild-type refolding data are from Sohl et al. (*5*); mutant refolding reactions were initiated by the addition of the indicated amounts of the pro region to 2 μM Int (except below 20 μM pro region where we maintained the pro region in a 10-fold molar excess over Int in order to maintain pseudo-first-order reaction conditions) and incubated at 4 °C. The high affinity of Pro for native state αLP (the product of the folding reaction) constrains Pro-dependent folding to be single turnover; excess Pro was used so that effects due to Pro depletion could be ignored. The folding reactions were followed by determining the amount of αLP activity at various times throughout the reaction as described (*4*). This procedure was modified so that 5−25 μL aliquots of the folding reaction were diluted (1:10 v/v) into 1 mg/mL bovine trypsin, in assay buffer at room temperature, to rapidly degrade the high concentrations of the pro region before aggregation could occur at the higher temperatures.

The observed pro-dependent folding reaction progress curves were biphasic. Careful analysis of the data indicated the following kinetic mechanism (see Results):



where K_M is the disassociation constant for the Michaelis complex (and represents an approximation to the "true" Michaelis constant, see below), k_i is the isomerization rate of the Int·Pro complex to a slower folding form (Int·Pro)′, and k_{cat} and k'_{cat} are the folding rates of Int·Pro and (Int·Pro)′, respectively. All of the progress curve data for each variant (refolding time courses at different Pro concentrations) were simultaneously fit using FitSim (*13*). In this way a single set of K_M, k_{cat}, k_i, and k'_{cat} values could be determined for each variant. Importantly, formation of the Int·Pro complex was assumed to be in rapid equilibrium and was only constrained by the complex disassociation constant, K_M. Surface plasmon resonance binding experiments indicate that the on and off rates for the Int·Pro complex are sufficiently fast to make this assumption valid. As a consequence, the dissociation constant K_d is essentially identical to what the standard K_M measurement would be for this reaction, justifying our use of K_M instead of K_d. Initial values for kinetic rates and constants were estimated from initial fitting of the data to a biphasic exponential equation. All the fits exhibited goodness of fit (R^2) values of >0.99.

Circular Dichroism Examination of the Pro Region. Pro region stability was examined by circular dichroism in the far-UV range with an AVIV 60DS spectropolarimeter. Far-UV spectra were taken at 2 °C with 2.5 μM solutions of the wild-type pro region and Pro-4 in 5 mM MES, pH 7.0, with scans averaged over four repeats with 5 s readings every 0.5 nm. Thermal denaturation curves were recorded with 8 s readings every 1 °C after 1 min of equilibration at that temperature, using the same buffer conditions as above. The observed denaturation was fully reversible.

Crystallography and Structure Determination. Crystals of the αLP active site proximal surface residue mutant V167I were grown by serial seeding using wild-type αLP microcrystals as described (*18*). Crystals grew over a 7-week period in drops with three to six crystals. Data for the V167I αLP mutant were collected from a single crystal at room temperature with an R-axis II image plate detector using X-rays from a Rigaku 18 kW generator. Data were reduced using software provided by Rigaku. The mutant protein crystallized isomorphously with the wild-type enzyme in the space group $P3_121$, with unit cell dimensions $a = b = 66.2$ Å, $c = 80.2$ Å, $\alpha = \beta = 90°$, and $\gamma = 120°$.

Initial difference Fourier maps were computed using structure factors and phases calculated from the refined

Table 1: Affinity and Rate Constants Determined for Each of the Mutants[a]

variant	K_i (nM)	k_{cat} (min^{-1})	k_i (min^{-1})	k'_{cat} (min^{-1})	K_M (μM)	relative k_{cat}/K_M	$\Delta\Delta G_{Int^{\ddagger}\cdot Pro}$ (kcal/mol)
wild type	0.32	1.78 ± 0.09	1.18 ± 0.1	0.25 ± 0.01	23 ± 3	1.0	NA
Pro-1	0.36	0.15 ± 0.02	0.24 ± 0.05	0.04 ± 0.002	19 ± 4	0.10	1.4
Pro-2	0.79	0.062 ± 0.007	0.14 ± 0.03	0.02 ± 0.001	24 ± 5	0.033	2.0
Pro-3	1.7	0.0026 ± 0.0002	NA[b]	NA[b]	5 ± 1	0.0067	3.0
Pro-4	99	$<10^{-6}$	NA[b]	NA[b]	~ 7[c]	$<10^{-6}$	>8.0
Q-3A	0.17	0.72 ± 0.06	0.59 ± 0.09	0.12 ± 0.01	34 ± 5	0.27	0.8
V167I	0.58	0.20 ± 0.01	0.03 ± 0.01	0.03 ± 0.01	64 ± 7	0.040	1.9
N170K	0.44	0.11 ± 0.02	0.06 ± 0.02	0.03 ± 0.002	54 ± 13	0.026	2.1

[a] k_{cat} and K_M were determined by fitting the refolding time courses to the αLP refolding kinetic mechanism (see Materials and Methods). Inhibition constants were determined using the tight-binding analysis for all Pro mutations except Pro-4, which was weak enough to allow use of Lineweaver−Burk analysis (see Materials and Methods). [b] NA = not applicable. [c] The affinity of the Int·Pro-4 complex was measured by surface plasmon resonance binding experiments.

coordinates of wild-type αLP (*14*). The map was inspected and the V167 → I side chain replacement performed using the interactive graphics package FRODO, after which the coordinates and *B*-factors were refined using X-PLOR (*15*). Omit maps were used to assay the correctness of the modeled structure. Finally, the presence and position of solvent molecules were examined, a total of 18 water molecules were removed, and 5 new water molecules were placed in the final structure. The final refined coordinates for the structure reported here have been deposited in the Brookhaven Protein Data Bank (code 1BOQ). Superposition and RMSD calculations were done using the MidasPlus program from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH Grant RR-01081). Figures were produced using the programs RasMol (Roger Sayle, Galxo Welcome Research and Development) and Bobscript, a modified version of Molscript (*16*) developed by Rob Esnouf (Oxford).

RESULTS

Pro Region C-Terminal Deletion Mutants. To examine their importance to the folding reaction, the C-terminal residues of Pro were progressively deleted in a series of Pro region constructs (Pro-1 → Pro-4), beginning with full-length ProWt (...GlyLysLeuGlnThrThr) and ending with the deletion of four residues from the C-terminus (...GlyLys). All the Pro variants were cloned into a T7 expression system, expressed, and purified as described previously (*7*).

To determine the effect of these deletions on the affinity of the modified Pro for native αLP, the K_i of each variant for wild-type αLP was measured. As shown in Figure 1 (see also Table 1), the series of Pro C-terminal deletions had surprisingly little effect on inhibition of αLP. Full-length ProWt inhibits αLP with a K_i of 0.3 nM; deleting three residues results in only a 5-fold reduction (Pro-3; $K_i = 1.7$ nM); deleting another residue, Pro-4, does significantly decrease the K_i ~300-fold to 99 nM (Table 1). The reduced affinity of Pro-4 was sufficiently great that it was no longer necessary to use a tight binding formalism (see Materials and Methods). Because the substrate was equivalent in length to the four residues removed in Pro-4, surface plasmon resonance binding experiments were used to confirm that this large drop in affinity of Pro-4 for native αLP was not an artifact due to increased substrate accessibility (data not shown).

Refolding rates were determined by measuring the amount of αLP activity present after Int was incubated with Pro for varying times. In the refolding reaction, Pro acts as a single turnover catalyst because Pro binds so tightly to native αLP, which is the product of the refolding reaction. To minimize any consequences of Pro depletion during the refolding reaction, Pro is always present in great excess (see Materials and Methods). The resulting data were clearly biphasic (Figure 2a), indicating a rate-limiting isomerization of some species or alternative parallel reaction pathways. However, as was previously observed (*5*), only the fast phase exhibited a systematic dependence on Pro concentration (Figure 2b,c).

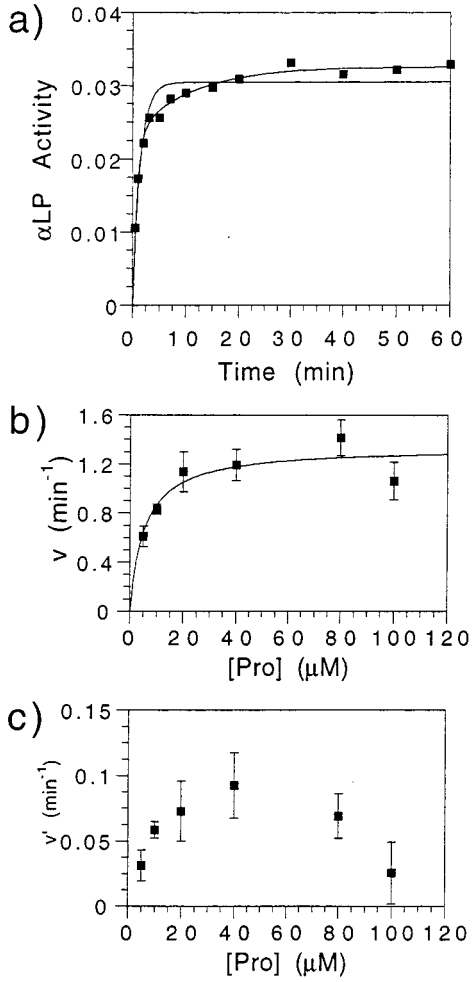
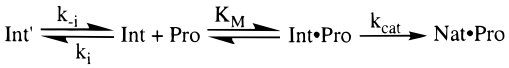


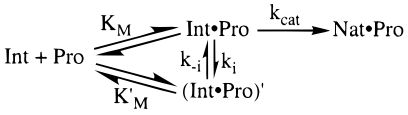
FIGURE 2: Biphasic nature of the pro region-dependent refolding of αLP. (a) A typical refolding progress curve with the best single and double exponential fits shown; (b) the fast-phase rate dependence on [Pro] with the fit to the single turnover Michaelis−Menten equation shown; (c) the slow phase rate versus [Pro].

This suggests one of the following kinetic schemes:



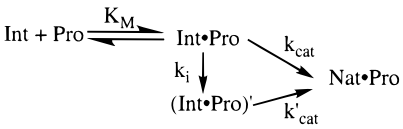
(a)

or



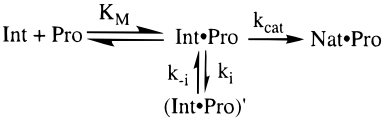
(b)

or



(c)

or



(d)

In all schemes, the fast-phase rate is interpreted as the exponential rate of a single turnover Michaelis−Menten reaction, as indicated by the good fit of the Pro concentration-dependent fast-phase rates to the appropriate equation (Figure 2b). These schemes differ with respect to the origins of the slow phase. In scheme a, Int can exist in two forms: Int which can productively interact with Pro and Int′ which cannot. If scheme a was correct, the rate and amplitude of the slow phase should be independent of Pro region mutations since the slow kinetic phase should only be a property of Int. Initial fitting of the refolding data determined that the slow-phase rates and amplitudes did vary with different Pro regions (data not shown), demonstrating that scheme a cannot be correct. In scheme b, Int·Pro binding results in formation of either folding-competent or -incompetent complexes, where the folding-incompetent complex slowly converts to the folding-competent complex either directly or indirectly through a release/rebind cycle. However, the requirement that the slow-phase rates be independent of Pro concentration leads to unreasonable assumptions: K_M' must be much smaller than the concentrations examined, yet must still have an on rate similar to that for formation of the productive complex. Otherwise, the kinetics would not appear biphasic. This seems an unlikely combination, and thus scheme b was also discarded. In scheme c, the Int·Pro complex can fold directly or isomerize to a folding-incompetent state (Int·Pro)′ that slowly folds to the native state with a relatively slower rate (k'_{cat}). Alternatively, the (Int·Pro)′ complex could fold via reisomerization to Int·Pro as in scheme d. These two reaction mechanisms cannot be distinguished by our kinetic experiments which only monitor the formation of Nat·Pro. Understanding the meaning of the different rates is simplified in scheme c: the slow-phase rate is k'_{cat} and the relative amplitudes between the fast and slow phases are determined by the relative rates of k_{cat} and k_i. While we have done the data fits both ways, all of the data presented here have been analyzed using scheme c.

In practice, the relevant rate and equilibrium constants (K_M, k_{cat}, k_i, and k'_{cat}) were extracted by fitting simulations of the kinetic mechanism to the refolding time courses using the computer program FitSim (*13*). To focus on the mechanism of Pro-mediated refolding of αLP, only the K_M and k_{cat} constants are referred to in all following discussions of the refolding reaction, although both the isomerization and slow-folding rates are given in Table 1.

The series of Pro C-terminal deletions had essentially no effect on formation of the Michaelis complex (Int·Pro) but exhibited progressively greater effects on the rate at which refolding of wild-type αLP Int was catalyzed (Figure 3 and Table 1). ProWt exhibits a k_{cat} of 1.78 min^{-1} and a K_M of 23 μM; with Pro-1 there is essentially no change in K_M but an ∼12-fold reduction in k_{cat} (Table 1). Pro-2 also shows no change in affinity for Int but exhibits a significant ∼29-fold reduction in k_{cat} (Table 1). The refolding mediated by Pro-3 is extremely slow and in fact is not detectably biphasic (data not shown). Indeed, k_{cat} for Pro-3 is much slower than any of the folding or isomerization rates found for other Int·Pro complexes (Table 1). Once k_{cat} is slower than the isomerization rate, (Int·Pro)′ would build up and all productive folding would be via k_{cat}'. Alternatively using scheme d, the rate of folding would be sufficiently slow for complex isomerization to be under preequilibrium conditions. The data demonstrate that Pro-3 binds Int quite well, in fact 4-fold better than ProWt; however, its k_{cat} is reduced >600-fold (Table 1). No refolding of αLP Int could be detected in the presence of Pro-4, even after 2 days in the presence of ∼200 μM Pro-4 (data not shown). Surface plasmon resonance binding experiments indicate that Pro-4 does binds Int in the micromolar range ($K_d \sim 7$ μM), with little change in the on and off rates relative to ProWt (data not shown). Therefore, while the Int·Pro-4 complex does form, from the detection limits of the αLP activity assay, any refolding catalyzed by Pro-4 must occur at a rate of $<10^{-6}$ min^{-1}. Remarkably, although deletion of the C-terminal residues of Pro had little or no effect on Pro affinity for wild-type Int, these shortened Pro constructs were greatly impaired in their ability to catalyze the refolding of αLP.

To determine if the effects of the Pro deletion series were due to the loss of the interactions of the deleted C-terminal residues or were the consequence of a global effect on Pro structure or stability, the amount of secondary structure and thermal stability were compared for ProWt and the most extreme mutation, Pro-4. The near-UV circular dichroism (CD) spectra of ProWt and Pro-4 were identical (Figure 4a), as were their thermal melting behavior, as followed by CD (Figure 4b). The identical behavior of these molecules indicates that they contain equivalent amounts of secondary structure and that their structural stability, and presumably tertiary structure, is also equivalent. Therefore, deletion of Pro C-terminal residues has no effect on the amount of secondary structure and thermal stability of Pro. Thus, the effects of these Pro C-terminal deletion mutants are due to the loss of these residues and not loss of Pro region structure or stability.

Sequence Specificity Requirements at the −3 Position of the Pro Region. Because initial analysis of the series of Pro C-terminal deletions suggested that deletion of three residues abolished the ability of Pro to refold Int, the side chain requirements at this position were investigated. A series of substitutions (Asn, Glu, Lys, Ala) for the glutamine residue at the −3 position were constructed in the context of full-length Pro. Surprisingly, initial analysis of these Pro mutants

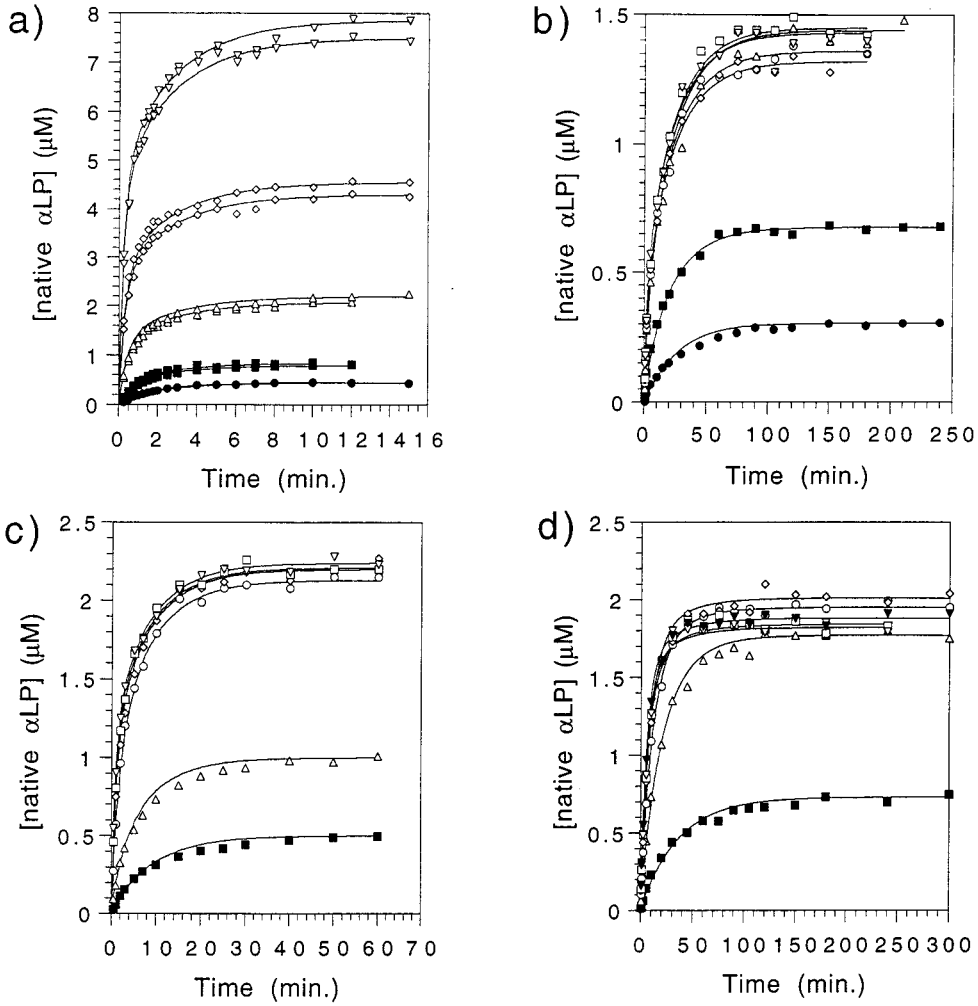


FIGURE 3: Kinetics of the pro region-dependent refolding of the mutants. The pro region was added at the indicated concentrations in at least a 10-fold molar excess over the intermediate state, and the reaction was incubated at 4 °C. Refolding was followed by the amount of recoverable protease activity. Shown are fits of the observed refolding at various concentrations of the pro region to the αLP refolding mechanism by FitSim (*13*). The various concentrations of Pro in the refolding data are represented as follows: 5 μM (filled circles), 10 μM (filled squares), 20 μM (open upright triangles), 40 μM (open circles), 60 μM (open diamonds), 80 μM (open squares), 100 μM (open inverted triangles). (a) ProWt refolding wild-type αLP; (b) Pro-1 refolding wild-type αLP; (c) ProQ-3A refolding wild-type αLP; (d) ProWt refolding V167I αLP.

indicated that they had a minimal effect on αLP refolding (D. E. Anderson, unpublished results). The alanine substitution (ProQ-3A) was selected as an interesting nonconservative substitution and fully characterized.

The ProQ-3A substitution exhibited minimal effects on all aspects of the refolding reaction (Figures 1 and 3c and Table 1). Interestingly, binding to the native state improves by a factor of 1.9 while transition state and Michaelis complex stabilization are reduced by factors of 2.5 and 1.5, respectively. Although these changes are marginal, the preferred binding conformation for ProQ-3A appears to be shifted toward the native state and away from optimally recognizing the αLP refolding transition state or Int, as demonstrated by the 3.7-fold decrease in k_{cat}/K_M.

α-Lytic Protease Mutants Which Alter Pro Region Binding. In a first attempt to map the interface between αLP and Pro, the surface of αLP was examined for likely regions of interaction. Knowing that Pro acts as a competitive inhibitor of αLP (*6*, *7*), we had suspected that there would be direct interactions between Pro and the substrate binding pocket, as later confirmed by biochemical analysis (*7*). On the basis of surface topology, we looked for other potential regions of Pro interaction and chose a surface loop ~15 Å from the active site (residues 167−170) for initial mutagenesis efforts (Figure 5). This site has a cluster of three residues whose side chains were fully solvent exposed and did not interact with one another. Our intent was to target sites where mutations would have a negligible effect on the stability and function of native αLP, yet might perturb Pro interactions during refolding. Using doped-cassette mutagenesis a minilibrary of variants was created at three positions (V167, T168, N170) where random amino acids were introduced in a background biased toward the native sequence. This library was screened for mutants that showed a decrease in αLP production using a simple plate activity assay (*12*). Unfortunately, the plate assay could not discriminate between a decrease in protease level due to impaired folding by Pro and a decrease in proteolytic activity per se with no change in the amount of active αLP being produced. Thus, mutants identified by the plate assay were further characterized to determine the cause of the reduced proteolytic activity. The results, summarized in Table 2, indicate that the V167I, V167L, and N170K mutations are good candidates for residues involved in Pro−αLP interactions as they predomi-

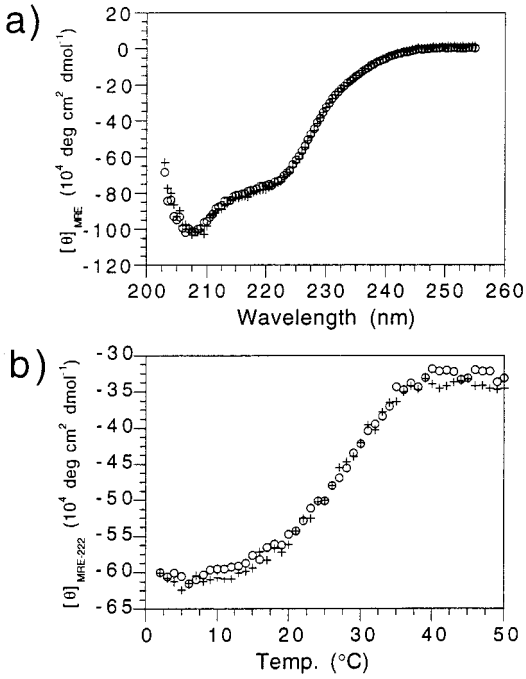


FIGURE 4: CD analysis of the stability and amount of secondary structure of Pro-4 compared to ProWt. (a) Far-UV spectra; (b) thermal denaturation curves. ProWt (open circles) and Pro-4 (crosses).

FIGURE 5: α-Lytic protease crystal structure. Ribbon drawing of wild-type αLP. Shown is a ribbon representation of the α-carbon backbone with the two orthogonal β-barrel domains shown in blue and green with ball-and-stick representations of the side chains of the catalytic triad in red and the stretch of surface residues that were mutated in yellow.

nantly affect production levels and not specific activity. To facilitate further characterization, the mutants V167I and N170K were subcloned into a phosphate depletion induction expression system, expressed, and purified (see Materials and Methods).

Mutation of the active site proximal surface residues Val167 and Asn170 minimally affect their ability to be inhibited by wild-type Pro (Figure 1 and Table 1). Pro inhibition of V167I was reduced 2-fold, while the K_i for N170K was essentially unchanged. By contrast, these active site proximal surface mutants significantly affect both the affinity of the Int state of these variants for ProWt and the rate at which these variant Int state αLPs can be refolded (Figure 3 and Table 1). V167I is refolded at a rate 9-fold less than that of wild-type αLP and also exhibits a 3-fold decrease in affinity for Pro in the Int state. N170K is refolded at a rate 16-fold slower and exhibits an 2-fold decrease in affinity for Pro in the Int state. Therefore, mutation of these surface residues near the active site affect k_{cat} and K_M but do not have a significant effect on K_i.

Table 2: In Vivo Characterization of α-Lytic Protease Active Site Proximal Surface Mutations[a]

mutant	relative activity (%)	relative expression (%)
V167L	1.5	1.2
V167I	4.5	1.5
V167H	ND[b]	0.02
T168S[c]	ND[b]	ND[b]
N170K	1.1	1.1
N170S	0.6	ND[b]

[a] Colonies exhibiting reduced protease activity were selected from a plate screen, and mutants were identified by sequencing. Activity was determined by a spectroscopic assay and compared to a representative wild-type culture. Expression levels were determined by ELISA and also compared to a representative wild-type culture. [b] ND = not determined. [c] T168S exhibited near wild-type activity in the plate assay. Given the conservative nature of the mutation, we decided not to pursue this particular mutant.

To show that the effects of these surface substitutions in αLP are a consequence of altered interactions and not due to alterations in αLP structure, the X-ray crystal structure was solved for V167I and compared to that of wild-type αLP (2ALP; *1*). The two crystal structures are superimposable with an α-carbon RMSD of 0.14 Å, demonstrating that at least this active site proximal surface substitution does not appreciably distort the αLP native state structure.

DISCUSSION

The αLP pro region functions as a powerful single-turnover catalyst to facilitate the folding of the molten globule-like folding intermediate state (Int) of the protease. To better understand how the pro region stabilizes the αLP refolding transition state, we have analyzed the functional consequences of mutations in both the pro region and within the protease.

Interpretation of the Pro-mediated refolding of αLP was initially complicated by the biphasic nature of refolding progress curves. However, this biphasic behavior is readily understood to result from a second process in which Int·Pro isomerizes into a slower folding form (see Results). The resulting refolding kinetic mechanism can be used to extract the relevant rate and equilibrium constants from the refolding time courses.

The ability of the Pro region to stabilize three kinetically distinct species along the folding pathway can be analyzed using the αLP refolding kinetic mechanism (see Results). The kinetically observable species correspond to formation of (i) the initial Michaelis complex (Int·Pro, K_M), (ii) the refolding transition state complex (Int‡·Pro, k_{cat}/K_M), and (iii) the product complex (Nat·Pro, K_i). Thus, the effects of mutations in either αLP or Pro can be separated into effects on three separate states of the refolding complex. Importantly, structural characterization of several mutants demonstrates that the changes observed in the folding reaction

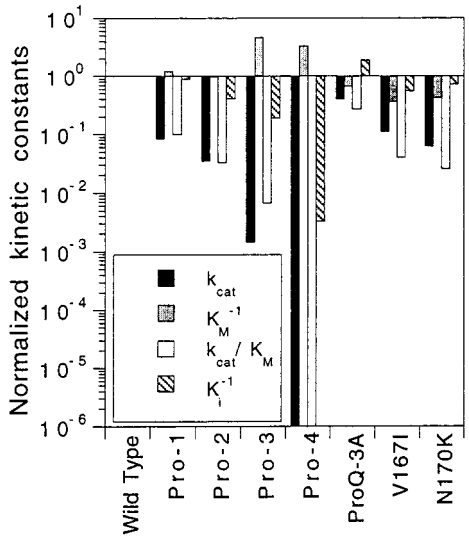


FIGURE 6: Graphical comparisons of the changes in the binding constants and rate of refolding for the mutants. The rate constants are shown normalized to the wild-type folding reaction, and the binding constants were inverted and also normalized to the inverted wild-type binding constants. The fold change is shown on a log scale.

result from the loss or disruption of specific interactions and do not arise from structural perturbations or instabilities induced by the mutations. A graphical representation of the effects of the mutants studied here on the three states is given in Figure 6. Deletion of Pro C-terminal residues has very little effect on formation of either the Michaelis complex or the native state inhibitory complex. However, there is a drastic reduction in the rate at which these Pro deletion variants mediate the refolding of αLP. Therefore, the Pro C-terminal residues must be directly or indirectly required to specifically stabilize the refolding transition state of αLP.

On the basis of previous biochemical studies (*7*), it is known that in the inhibitory native state complex the C-terminal residues of Pro lie in the αLP substrate binding cleft in a substrate-like manner. Therefore, our detailed knowledge of substrate binding within the active site can be used to help interpret the observed effects of pro region mutagenesis. For simplicity of comparison of the C-terminal residues of Pro with peptide substrate positions, we proposed the following alignment (*7*):

αLP substrate binding sites:	S4 S3 S2 S1
Pro Region Sequence:	...LysLeuGlnThrThr
Substrate Residues:	P4 P3 P2 P1

From crystallographic analysis of substrate analogues, such as peptide boronic acids (*17*), bound in the αLP substrate binding cleft (*18*, *19*), we know that a combination of three substrate backbone hydrogen bonds, P1 side chain packing into the S1 specificity pocket, and the hydrophobic packing of P2 into the αLP S2 subsite stabilizes substrate binding (schematized in Figure 7). While biochemical studies indicate that the Pro C-terminus binds in a substrate-like manner (*7*), surprisingly, the data presented here show that these interactions do not contribute to the stability of the αLP·Pro inhibitory complex. Instead, the dominant effect of Pro C-terminal deletions is on the refolding reaction.

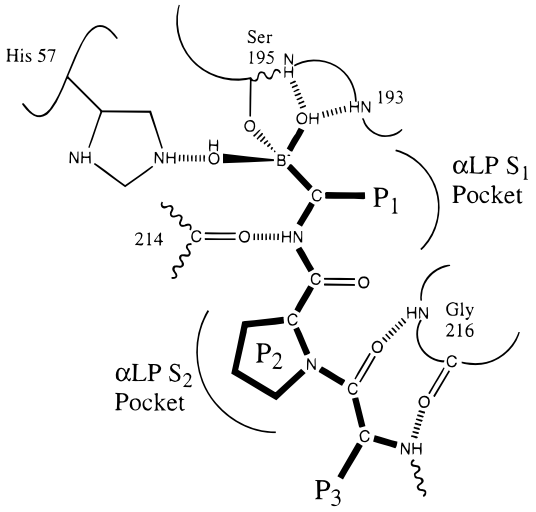


FIGURE 7: Schematic diagram of the binding of a generalized boronic acid peptide inhibitor (in bold) to the substrate binding cleft of αLP. Figure adapted from Bone et al. (*18*).

Intriguingly, the magnitude of these effects can be understood in terms of the loss of substrate-like interactions which must be forming during the folding transition state.

Specifically, the observed data indicate a loss of 1.4 kcal/mol in folding transition state stabilization at the P1 site upon deletion of the Pro C-terminal Thr. This effect can be due to either side chain or main chain interactions, or both. However, the minimal effects of varying the Pro C-terminal residue on production of active αLP in vivo (*20*) suggest that the backbone hydrogen bond at P1 is more important for stabilizing the refolding transition state complex. At the P2 position, substrates do not contribute any backbone hydrogen bonds, but the P2 side chain does pack into a hydrophobic pocket. We suggest that the corresponding threonine at Pro-2 packs in an analogous manner and is responsible for the relatively smaller 0.6 kcal/mol $\Delta\Delta G$ after deletion of a second Pro C-terminal residue (Pro-1 versus Pro-2). The nearly equivalent effects on k_{cat}/K_M of deleting the residue at Pro-3 (Pro-3 vs Pro-2; $\Delta\Delta G = 1.0$ kcal/mol) compared to just removing the majority of the side chain (ProQ-3A; $\Delta\Delta G = 0.8$ kcal/mol) suggest that side chain interactions and not the backbone hydrogen bonds are most important for transition state stabilization at this position. Alternatively, the ProQ-3A mutation could alter the ability of the Pro C-terminus to fit into the substrate binding pocket. Finally, deletion of the leucine at the Pro −4 position has such a dramatic effect (>5 kcal/mol; Pro-4 versus Pro-3) on Pro-mediated refolding that both side chain and backbone interactions must be involved. Interestingly, while the P4 alanine side chain found in the available complex structures (*19*) is observed pointing into solution, there is a hydrophobic surface patch nearby (*18*) with which the leucine side chain could interact.

Substitution of αLP surface residues near the active site has multiple consequences. While these substitutions have little effect on formation of the native state inhibitory complex, there are clear effects on formation of the Michaelis complex and stabilization of the refolding transition state. On the basis of the effects on K_M, we conclude that the side chains of V167I and N170K participate in forming the Michaelis complex. Further, both mutants also affect k_{cat},

and so these side chains are likely to form additional interactions with Pro in the refolding transition state. These results emphasize the importance of interactions between the nascent αLP substrate binding cleft and Pro in the refolding reaction, specifically in stabilizing the αLP refolding transition state.

The remarkable parallelism between interactions which stabilize substrates for hydrolysis by αLP (Figure 7) and interactions with the Pro C-terminus which facilitate refolding suggests that the pro region C-terminus precisely mimics substrate binding modes. Further, the data indicate that such substrate-like interactions between the Pro C-terminus and the nascent protease substrate binding cleft are critical for stabilizing the refolding transition state. Because such similar interactions must be present in the inhibitory complex, even though they are unimportant for K_i, we propose that this portion of the structure is strained in Nat•Pro but not in $Int^{‡}$•Pro. This, then, may be the basis for the ∼4.4 kcal/mol difference in Pro stabilization of the folding transition state compared to the native state (*4*, *5*).

These results suggest a model where interactions between Pro and the protease in the area around residues 167−170 occur early in the folding reaction as evidenced by effects of mutations on K_M. Subsequently, interactions between Pro C-terminal residues and the nascent substrate binding cleft play a central role in stabilization of the folding transition state. It is important to note that interactions outside of the protease active site must also be required to catalyze αLP folding, as expected given the extremely large stabilization energies observed (*4*, *5*). Incubation of αLP Int with a substantial excess of a very tight binding small peptide mimic inhibitor (suc-Ala-Ala-Pro-boroVal, K_i = 6.4 nM; *17*) does not promote folding of the protease (David Baker, personal communication). This clearly demonstrates that active site stabilization is necessary but not sufficient for catalysis of αLP folding. The high affinity of Pro for native αLP implies that at least some of the required interactions for catalyzing folding are conserved in the native state inhibitory complex. A combination of structural studies and further mutagenesis should provide insight into these other critical interactions.

Many other extracellular proteases are encoded with pro regions, and a number of them have been shown to be inhibitors of their associated proteases (*21*−*25*). In the case of the evolutionarily unrelated serine protease, subtilisin, the role of its pro region has been extensively studied. The folding of subtilisin is highly dependent on the presence of its pro region, either in vivo or in vitro (*23*, *26*). Its pro region also inhibits subtilisin activity by steric occlusion of the protease active site, as demonstrated by the crystal structure of the complex (*27*). However, if the calcium binding site of subtilisin is deleted, then the protease will slowly fold on its own (*28*), suggesting that the refolding barrier for subtilisin may be lower than that for αLP.

Deletion and mutation of C-terminal residues of the subtilisin E pro region affect the ability of the pro region both to catalyze refolding and to inhibit protease activity to very similar degrees (*29*). Similarly, deletion of two or five residues from the C-terminus of the subtilisin BPN′ pro region reduces both catalysis of folding and binding to native subtilisin by equivalent amounts (*30*). In both cases, deletion of pro region C-terminal residues have large effects on the refolding reaction. Given that the C-terminus of the subtilisin pro region has also been shown to bind in the protease active site in the native state complex (*27*), these results suggest that interactions of the subtilisin pro region C-terminus and the protease active site are also important in folding catalysis. This implies that there may be fundamental similarities in the mechanisms of pro region-mediated refolding for subtilisin and αLP. However, there are also differences; the structures of αLP and subtilisin are unrelated, the pro region of subtilisin is much smaller (77 vs 166 residues), and C-terminal deletions in the subtilisin pro region do not preferentially affect refolding rates but instead exhibit equivalent effects on both the refolding reaction and inhibition. One possibility is that the subtilisin pro region is not as optimally tuned to stabilize the folding transition state for its protease. This could be a consequence of the lower thermodynamic stability of the subtilisin pro regions which, by themselves, are almost entirely unfolded (*23*). This contrasts with the substantially larger αLP pro region which is fully folded in solution (Figure 4b). Even if the interactions are not specific to the refolding transition state, it is apparent that deletion of C-terminal residues from the subtilisin pro region greatly inhibit its ability to carry out refolding.

It is clear from both the results presented here and those published for subtilisin that interactions between the enzyme active site and its pro region C-terminal residues are critical in catalyzing folding of these proteases. Thus, we suggest that such pro region C-terminal/active site interactions will be crucial in folding of proteases with amino-terminal pro regions and that an equivalent pro region N-terminal/active site interaction will be crucial in the folding of proteases with carboxyl-terminal pro regions.

ACKNOWLEDGMENT

The authors thank Drs. Stephen Rader and Earl Rutenber for assistance with solving the crystal structure of the V167I αLP mutant, Dr. Kazuo Harada for assistance with the CD experiments, Dr. Alan Frankel for use of the CD spectrometer, Dr. Arthur Weiss for use of the BiaCore, and Dr. Alan Derman for critically reading the manuscript.

REFERENCES

1. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D., and James, M. N. G. (1985) *J. Mol. Biol. 184*, 479.
2. Schechter, I., and Berger, A. (1968) *Biochem. Biophys. Res. Commun. 27*.
3. Silen, J. L., and Agard, D. A. (1989) *Nature 341*, 462−464.
4. Baker, D., Sohl, J. L., and Agard, D. A. (1992) *Nature 356*, 263−265.
5. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature* (in press).
6. Baker, D., Silen, J. L., and Agard, D. A. (1992) *Proteins 12*, 339−344.
7. Sohl, J. L., Shiau, A. K., Rader, S. D., Wilk, B., and Agard, D. A. (1997) *Biochemistry 36*, 3894−3902.
8. Whitaker, D. R. (1970) *Methods Enzymol. 19*, 599−613.
9. Silen, J. L., Frank, D., Fujishige, A., Bone, R., and Agard, D. A. (1989) *J. Bacteriol. 171*, 1320−1325.
10. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
11. Mace, J. E., Wilk, B. J., and Agard, D. A. (1995) *J. Mol. Biol. 251*, 116−134.

12. Mace, J. E., and Agard, D. A. (1995) *J. Mol. Biol. 254*, 720−736.
13. Zimmerle, C. T., and Frieden, C. (1989) *Biochem. J. 258*, 381−387.
14. Brayer, G. D., Delbaere, L. T. J., and James, M. N. G. (1979) *J. Mol. Biol. 131*, 743−775.
15. Brunger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science 235*, 458−460.
16. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
17. Kettner, C. A., Bone, R., Agard, D. A., and Bachovchin, W. W. (1988) *Biochemistry 27*, 7682−7688.
18. Bone, R., Shenvi, A. B., Kettner, C. A., and Agard, D. A. (1987) *Biochemistry 26*, 7609−7614.
19. Bone, R., Frank, D., Kettner, C. A., and Agard, D. A. (1989) *Biochemistry 28*, 7600−7609.
20. Reidhaar-Olson, J. F., and Agard, D. A. (manuscript in preparation).
21. Winther, J. R., and Sorensen, P. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9330−9334.
22. Segundo, B. S., Martinez, M. C., Vilanova, M., Cuchillo, C. M., and Aviles, F. X. (1982) *Biochim. Biophys. Acta 707*, 74−80.
23. Strausberg, S., Alexander, P., Wang, L., Schwarz, F., and Bryan, P. (1993) *Biochemistry 32*, 8112−8119.
24. Fusek, M., Mares, M., Vagner, J., Voburka, Z., and Baudys, M. (1991) *FEBS Lett. 287*, 160−162.
25. Fox, T., de Miguel, E., Mort, J. S., and Storer, A. C. (1992) *Biochemistry 31*, 12571−12576.
26. Ohta, Y., et al. (1991) *Mol. Microbiol. 5*, 1507−1510.
27. Gallagher, T., Gilliland, G., Wang, L., and Bryan, P. (1995) *Structure 3*, 907−914.
28. Bryan, P., et al. (1992) *Biochemistry 31*, 4937−4945.
29. Li, Y., Hu, Z., Jordan, F., and Inouye, M. (1995) *J. Biol. Chem. 270*, 25127−25132.
30. Wang, L., et al. (1995) *Biochemistry 34*, 15415−15420.
31. Fujishige, A., Smith, K. R., Silen, J. L., and Agard, D. A. (1992) *J. Cell Biol. 118*, 33−42.

BI980883V