PROTEINS: Structure, Function, and Genetics 12:339–344 (1992)

Protease Pro Region Required for Folding Is a Potent Inhibitor of the Mature Enzyme

David Baker, Joy L. Silen, and David A. Agard
Department of Biochemistry and Biophysics and The Howard Hughes Medical Institute, University of California, San Francisco, California 94143-0448

ABSTRACT **α-Lytic protease, an extracellular bacterial serine protease, is synthesized with a large pro region that is required in vivo for the proper folding of the protease domain. To allow detailed mechanistic study, we have reconstituted pro region-dependent folding in vitro. The pro region promotes folding of the protease domain in the absence of other protein factors or exogenous energy sources. Surprisingly, we find that the pro region is a high affinity inhibitor of the mature protease. The pro region also inhibits the closely related *Streptomyces griseus* protease B, but not the more distantly related, yet structurally similar protease, elastase. Based on these data, we suggest a mechanism in which pro region binding reduces the free energy of a late folding transition state having native-like structure.**

Key words: protein folding, pro region, protease inhibition

INTRODUCTION

The folding of most single subunit proteins is thought to proceed through a sequence of well-defined kinetic intermediates.[1] Understanding the mechanism of protein folding will require determination of the structures of these intermediates and the energetics underlying their formation. The high cooperativity of folding greatly complicates the analysis of single steps in the folding pathway. Furthermore, direct structural characterization of folding intermediates is a formidable task as they are present only transiently and then in low concentration.

An unusual feature of the folding of two small serine proteases provides a means to circumvent the high cooperativity of folding. α-Lytic protease, an extracellular serine protease of *Lysobacter enzymogenes*, is synthesized as a larger precursor containing a 33 amino acid signal sequence, a 166 amino acid pro region, and a 198 amino acid protease domain.[2] The pro region is only transiently associated with the catalytic domain and does not form part of the active enzyme. In vivo studies have shown that the pro region is required for the proper folding of the protease domain.[3,4] An evolutionarily unrelated serine protease, subtilisin, also requires a pro region for proper folding.[5,6] This is quite remarkable as it suggests that not only the catalytic mechanism, but also the folding mechanism of these two proteases has arisen by convergent evolution.

The pro region dependence of α-lytic protease folding potentially allows study of a single step in a folding pathway. Previously, we had shown that for α-lytic protease, covalent linkage between the pro region and the protease domain is not required for pro region-assisted folding in vivo.[4] Here we extend this analysis with in vitro studies of the interactions between purified pro region and protease. We find that the pro region is both necessary and sufficient for α-lytic protease folding in vitro, and surprisingly, that the pro region is an extremely high affinity inhibitor of the protease. The latter finding has important implications for the mechanism by which the pro region promotes the folding of the protease domain.

MATERIALS AND METHODS

Construction of a Pro Region-Containing Fusion Protein

To facilitate production and purification of the pro region, a glutathione transferase-pro region fusion was constructed by cloning the pro region-encoding *Eae*1 fragment from the expression plasmid pALP5[3] into the *Sma*1 site of pGEX-2T.[7] The fusion protein (GEXPRO) contains the entire pro region and five additional amino acids (APNSS) at the C terminus introduced during cloning. GEXPRO and GEX proteins were purified from *E. coli* cells harboring the relevant plasmids using glutathione agarose (Sigma) as described.[7]

Reconstitution of Pro Region-Dependent Folding

α-Lytic protease was denatured by treatment with 6 M guanidine HCl, 200mM glycine pH 3.0 for 12 hr at 25°C. The low pH was used to minimize autolysis

Received June 18, 1991; accepted August 5, 1991.
Address reprint requests to Dr. David Agard, Howard Hughes Medical Institute Research Laboratory, Dept. of Biochemistry and Biophysics, Structural Biology Unit, University of California, S-964, Box 0448, San Francisco, CA 94143-0448.
J.L. Silen is now at Syva Corp., MS-2-204, 900 Arastradero, Palo Alto, CA 94304.

during denaturation. Twenty μl portions of the denatured protease solution were mixed with 40 μl of GEX or GEXPRO solutions in 3 M guanidine HCl, 20 mM glycine pH 3. The reactions were further supplemented with 30 μM of the boronic acid protease inhibitor Ac-Pro-boroVal[8] (K_i for α-lytic protease inhibition 3.3 μM, final concentration of α-lytic protease was 1 μM) to prevent autolysis during refolding and then dialyzed against 30 μM Ac-Pro-boroVal, 5mM TrisCl pH 8.0 for 6 hr at 4°C using a microdialysis chamber (Pierce). Ten μl portions of each reaction were then supplemented with bovine trypsin (1 μg, Cooper Biomedical) and incubated at 25°C for 15 min to release pro region inhibition (see below). Trypsin rapidly degrades the pro region and nonnative conformations of α-lytic protease while leaving the folding enzyme intact (data not shown). α-Lytic protease activity was determined from the linear portion of progress curves using succinyl-Ala-Pro-Ala-para-nitroanalide as a chromogenic substrate as described.[3] Little enzymatic activity was recovered when the trypsin treatment step was omitted due to pro region-dependent inhibition. Trypsin did not hydrolyze the para-nitroanilide substrate used in the α-lytic protease activity assay. To relieve Ac-Pro-boroVal inhibition, the treated reactions were diluted 1:100 in 0.1 M Tris pH 8.0 for activity assay.

Protease Inhibition Studies

Inhibition by the pro region was investigated by assaying α-lytic protease activity in the presence of increasing amounts of the GEXPRO fusion protein. α-Lytic protease activity was determined as above except that the assay mix was supplemented with 0.1 mg/ml BSA (Sigma) to stabilize the very low enzyme concentrations required for these studies. BSA at this concentration did not interfere with the enzyme assay. The high affinity of the pro region-protease interaction precluded conventional methods of measuring K_i, which rely on competition by substrate. The detection limit in the α-lytic protease activity assay is ~7 nM and thus it was not possible to carry out the inhibition studies at concentrations bracketing the K_i ($\sim 10^{-10}$ M). Instead, the apparent K_i was estimated by fitting the inhibition data to the equation $v = k_2 S/k_M + S \times (E - K_i - I + ((K_i + E + I)^2 - 4EI)^{1/2})/2$, a standard procedure in studies of tight binding inhibitors.[9] The GEXPRO concentration was initially estimated using the extinction coefficient suggested for GEX fusion proteins by Smith and Johnson,[7] and then refined during the fitting of the inhibition curves to a value ($\epsilon_{280} = 76800\ M^{-1}\ cm^{-1}$) 20% lower than this estimate. Dilute solutions (< 0.5 μM) of the GEX and GEXPRO proteins were stabilized by addition of 1.0 mg/ml BSA. *Streptomyces griseus* protease B (SGPB) activity was assayed as described for α-lytic protease except that the substrate used was succinyl-Ala-Ala-Pro-Phe-para-nitroanalide. The k_{cat}/K_m for SGPB with this substrate is 10,900 ($M^{-1}\ s^{-1}$). The K_i of the pro region for protease B was obtained from the ratio of the slopes of Lineweaver-Burk plots in the presence and absence of the pro region.

RESULTS

Pro Region Dependent Folding In Vitro

To allow detailed mechanistic study, we sought to reconstitute pro region-dependent folding in vitro. Mature α-lytic protease was denatured by treatment with 6 M guanidine-HCl buffered at pH 3.0 to prevent autolysis during unfolding. Due to the absence of reducing agent, the three disulfide bonds in native α-lytic protease presumably remain intact in the denatured protein. The guanidine was removed by dialysis against 5 mM Tris pH 8.0. As denatured α-lytic protease is extremely protease sensitive, any lack of synchrony during refolding could potentially result in degradation of a substantial fraction of the incompletely folded protease molecules. To eliminate this possibility, the reactions were supplemented with a peptide boronic acid protease inhibitor to prevent autolysis during refolding.

Importantly, under these conditions no activity was recovered when the protease was dialyzed alone (Fig. 1) or in the presence of 0.1 mg/ml BSA (data not shown). To examine the effect of the pro region on refolding, the denatured protease was supplemented with increasing amounts of either the α-lytic protease pro region in the form of a glutathione transferase-pro region fusion protein (GEXPRO) or glutathione transferase (GEX) prior to dialysis. Following removal of the guanidine, the samples were treated with trypsin to relieve potential inhibition by the pro region (see below), and then assayed for protease activity. Supplementation with stoichiometric amounts of pro region led to recovery of up to 35% of the activity present prior to denaturation, whereas no activity was recovered in the presence of the GEX control (Fig. 1). Gel electrophoretic analysis showed that intact protease was present in all samples after dialysis (data not shown), ruling out the trivial possibility that the pro region promoted recovery simply by preventing autolysis during refolding. Similar high levels of protease activity were recovered in the presence of GEXPRO when the boronic acid protease inhibitor was omitted. The inhibitor, therefore, does not interfere with pro region-assisted folding. The pro region is thus both necessary and sufficient for the proper folding of denatured mature α-lytic protease containing "native" disulfide bonds. Notably, pro-mediated folding does not require any exogenous energy source such as ATP.

The recovery of only a fraction of the input protease activity is not surprising. Substantial aggregation was observed when the denatured protease

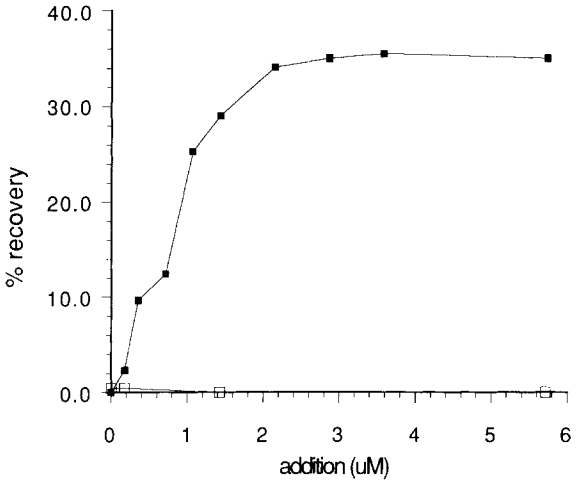

Fig. 1. Pro region dependent folding in vitro. Denatured α-lytic protease (2.9 μM) was allowed to refold in the presence of GEX (open symbols) or GEXPRO (closed symbols) and then assayed for protease activity.

was dialyzed in the absence of the pro region. The level of activity recovery presumably reflects competition between aggregation and folding (interaction with the pro region) during dialysis. As both the pro region and the protease are denatured at the beginning of the reaction, recovery of activity requires that the pro region regain the ability to promote folding at a concentration of guanidine high enough to prevent aggregation of the protease.

α-Lytic Protease Is Inhibited by Its Pro Region

Previously,[4] we postulated that the α-lytic protease pro region might act as a "foldase" by lowering the energy of a rate-limiting kinetic barrier. If this high energy transition state were to occur at a late stage in folding (as in several other well-characterized folding pathways)[10], then the pro region might interact with the mature enzyme. We therefore investigated the effect of the pro region on α-lytic protease activity.

As indicated in Figure 2, GEXPRO was a potent inhibitor of α-lytic protease, whereas GEX alone had no effect. Accurate estimation of the K_i requires that equilibrium binding has been achieved, which is slow for such high affinity inhibitors. The protease was preincubated with either GEXPRO or GEX for 5 min before the addition of substrate; steady-state levels of protease activity were reached within 10 min after substrate addition. Preincubation was necessary to prevent substrate depletion prior to attaining equilibrium. The steady-state level of protease activity was found to be independent of the order of addition of the pro region and the substrate to solutions containing the protease. Taken together with the nonlinearity of the inhibition curves, this strongly suggests that inhibition by the pro region is

TABLE I. Inhibition by the Pro Region is Reversed by Protease Treatment*

Addition	Activity (% max)
1. α-lytic	100
2. trypsin	< 0.1
3. α-lytic + gexpro incubate 5′	< 1
4. α-lytic + gexpro incubate 25′	< 1
5. α-lytic + trypsin incubate 25′	95
6. α-lytic + gexpro incubate 5′, add trypsin, incubate 20′	89

*Ten-μl reactions containing 1 μg bovine trypsin, GEXPRO (2.7 μM), or α-lytic protease (2.2 μM) were prepared as indicated and then assayed for α-lytic protease activity. Trypsin did not cleave the tripeptide substrate used in the activity assay (line 2) nor did it have a significant effect on the activity of α-lytic protease alone (line 5).

reversible (see also Table I). Inhibition data from several experiments similar to the one shown in Figure 2 fit the classical tight binding inhibitor equation (ref 9, see Methods) with an apparent K_i of 0.5–2.0 × 10^{-10}M. This makes the pro region the highest affinity inhibitor known for α-lytic protease (cf ref 8).

α-Lytic protease is thus synthesized with a covalently attached, high affinity inhibitor. A mechanism must exist for relieving the inhibition in vivo. We have found that a wide variety of proteases (data not shown), including bovine trypsin, will rapidly cleave within the pro region and relieve inhibition. As shown in Table I, inhibition of α-lytic protease by the pro region was almost completely reversed by treatment with trypsin (compare rows 1, 3, and 6). This result rules out the possibility that the pro region causes irreversible denaturation of the protease. In the natural host, *Lysobacter,* it is likely that one or more of the many secreted proteases is responsible for relief of pro region inhibition. Inhibition by the pro region may be important in preventing premature and possibility deleterious expression of protease activity in the periplasm.

Specificity of Pro Region Inhibition

To determine the specificity of pro region inhibition, we examined the effect of the α-lytic protease pro region on the activity of the two evolutionarily related serine proteases: the bacterial enzyme *Streptomyces griseus* protease B (SGPB) and the mammalian enzyme pancreatic elastase. The global folding topology of both proteases is quite similar to α-lytic protease, but SGPB has significantly higher amino acid sequence homology to α-lytic protease than does elastase (37% sequence identity vs. 18%).[11] The α-lytic protease pro region strongly inhibited SGPB (K_i = 1.1 nM, Fig. 3), but had no measurable effect on elastase activity (data not shown). The high affinity of the α-lytic protease pro region for SGPB is

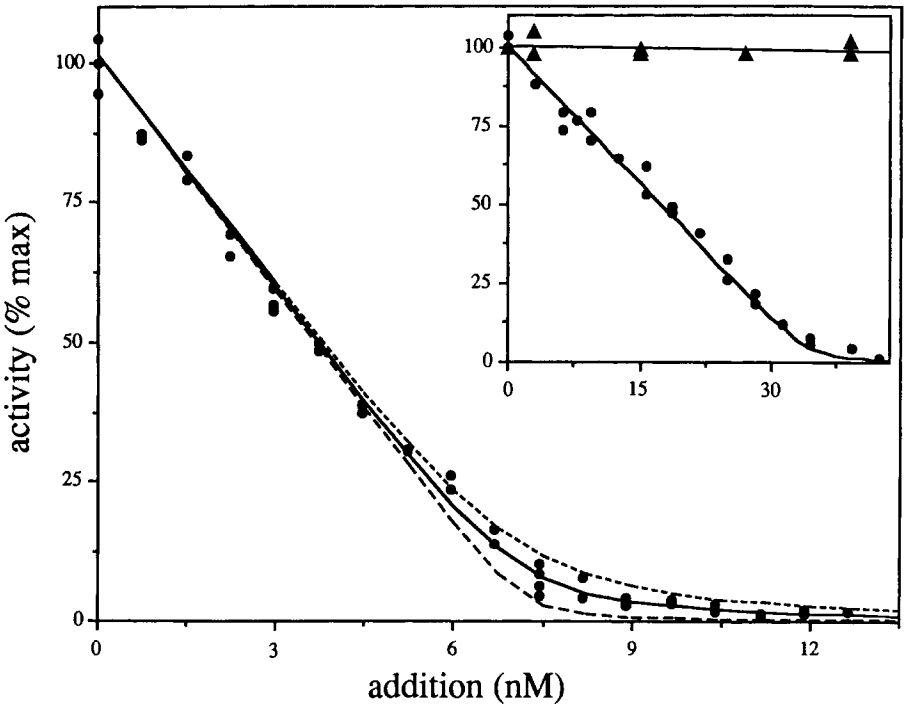

Fig. 2. Inhibition of α-lytic protease by its pro region. α-Lytic protease (main figure 8.5 nM, inset 40 nM) was mixed with the indicated concentrations of GEX (triangles) or GEXPRO (circles) and then assayed for protease activity. Solid lines are plots of the tight binding inhibitor equation (see Methods) with $K_i = 9.0 \times 10^{-11}$ M. For comparison theoretical curves are shown for $K_i = 1.8 \times 10^{-11}$ M (lower dashed line), and $K_i = 1.8 \times 10^{-10}$ M (upper dashed line).

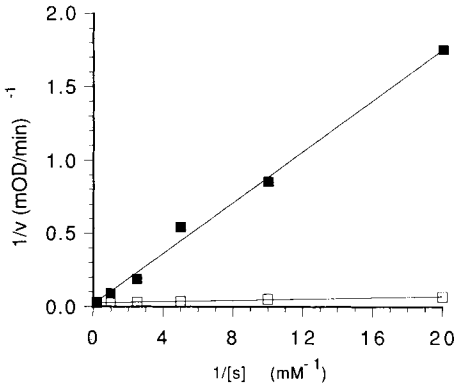

Fig. 3. *S. griseus* protease B is inhibited by the α-lytic protease pro region. SGPB (15 nM) activity was assayed in the presence (closed symbols) and absence (open symbols) of 74 nM GEXPRO with the indicated concentration of substrate. Data points are the average of three separate measurements.

interesting because the two proteases have quite different substrate specificities; whereas both α-lytic protease and elastase preferentially cleave peptide bonds following small hydrophobic amino acids, SGPB cleaves after large hydrophobic residues. It may be possible to identify regions involved in pro region binding by determining the affinity for the pro region of hybrids containing different segments of the three structurally related proteases.

The slightly lower affinity of the pro region for SGPB compared to α-lytic protease makes it possible to distinguish between competitive and non-competitive modes of inhibition. The α-lytic protease pro region behaves as a classical competitive inhibitor of SGPB; the pro region is displaced at high substrate concentrations (Fig. 3). Making a similar discrimination for α-lytic protease inhibition is at present limited by the large difference in the affinity of the protease substrate and the pro region for α-lytic protease and by the limit of detection in the protease activity assay (~7nM protease). These two constraints limit the accuracy of the K_i measurement and complicate determination of the mode of inhibition (competitive vs noncompetitive). However, by analogy with SGPB, it seems very likely that the pro region also competitively inhibits α-lytic protease.

The finding that the α-lytic pro region inhibits SGPB suggests that the *S. Griseus* protease pro regions also function as inhibitors and, by analogy, might play an important role in folding their protease domains. In support of this, recent in vivo experiments have indicated that the SGPB pro region is required for production of active enzyme.[12] The *S. griseus* proteases A and B are synthesized with N terminal pro regions (78 and 76 amino acids), which have detectable sequence homology (23 and 37% sequence identity) to the α-lytic protease pro region.

The highest homology is in the C-terminal 50 residues of the 166 amino acid α-lytic pro region suggesting a possible location for the inhibitory function. Although it is possible that this region also contains the folding functionality, deletion experiments failed to find any functional subdomain within the pro region.[13]

Almost all known protein inhibitors of serine proteases share a common mechanism for inhibiting protease activity in which a peptide bond on the inhibitor surface combines with the enzyme active site in a substrate like manner. Whereas the k_{cat}/K_m for the interaction is large, both k_{cat} and K_m are very small and the inhibitor, instead of being turned over, remains bound at the protease active site.[14] Although binding of substrate and pro region to SGPB and very likely to α-lytic protease are mutually exclusive, we have no evidence that the pro region obeys this "standard mechanism" or even that the pro region binds at the protease activite site. For example, pro-region inhibition could occur via a partial unfolding of the enzyme leading to a destruction of the substrate binding pocket. Although competitive with substrate binding, this would not require the pro region to directly block the active site.

X-ray crystallographic analysis of the complex between pro region and α-lytic protease should clarify the mechanism of inhibition. For this purpose, conditions have been found to release the pro region from the GEXPRO fusion protein by limited proteolysis. Preliminary results with isolated pro region suggest that the presence of the N-terminal glutathione transferase in the fusion protein does not influence either the refolding or the inhibition activities. Although much remains to be learned about the inhibition mechanism, even at this stage our results have important implications for the mechanisms by which the pro region facilitates the folding of the protease domain.

DISCUSSION

Inhibition of proteases by pro regions appears to have arisen independently several times during evolution. Whereas the pro regions of the closest eukaryotic relatives of α-lytic protease, elastase, trypsin, and chymotrypsin, are not protease inhibitors,[15] the unrelated pancreatic carboxypeptidase A is inhibited by its pro region,[16] the pro region of the aspartyl protease precursor pepsinogen lies across the enzyme active site,[17] and the pro region of subtilisin has been recently shown to be a potent subtilisin inhibitor.[18] It is interesting to speculate that the functional similarities between the prokaryotic and eukaryotic protease pro regions may extend beyond protease inhibition; the carboxypeptidase A pro region, like the α-lytic protease pro region, may also play a role in protein folding. Such speculation may be extended to the large number of prokaryotic and eukaryotic proteases, which have pro regions whose function has not yet been determined; many of these pro regions may be protease inhibitors that function in folding.

Our results suggest that the pro region promotes folding in a manner quite distinct from that of the "molecular chaperonins," a ubiquitous family of proteins which also facilitate protein folding reactions. Two of the best-characterized chaperonins, the prokaryotic GroE and the eukaryotic HSP70, are ATPases and appear to utilize the energy of ATP hydrolysis to repeatedly bind and release partially folded protein conformations.[19,20] Sequestration of folding intermediates by chaperonins is thought to promote folding by greatly reducing the rate of competing aggregation reactions. In contrast, no exogenous energy source is required for promotion of α-lytic protease folding by the pro region. Furthermore, whereas the pro region binds the folded protease with subnanomolar affinity, the chaperonins typically have extremely weak affinities for native protein conformations.[21] These differences in affinity have a correlate in differences in specificity: the pro region interacts with α-lytic protease and the closely related *S. griseus* protease B, but not with the more distantly related, but structurally similar, elastase. By contrast, the chaperonins promote the folding of a wide variety of structurally unrelated proteins. The lack of requirement for exogenous energy source and the high affinity for the folded state suggest that the pro region promotes folding through a novel mechanism.

Protein folding involves kinetic competition between reactions off the pathway (e.g., aggregation) and on the pathway leading to the native state. As discussed earlier, the pro region could promote proper folding of α-lytic protease either by destabilizing an off-pathway transition state (as suggested for the chaperonins) or stabilizing an on-pathway transition state.[4] The observation that the pro region interacts strongly with the native enzyme suggests, first, that the pro region directly facilitates the on-pathway reaction and, second, that the rate-limiting folding transition state has a native-like conformation. It should be emphasized that since the folded enzyme is stable upon removal of the pro region, the pro region cannot function simply by shifting the folding equilibrium in favor of the folded state. If the interactions between the pro region and the native enzyme are representative of those between the pro region and the folding transition state, pro region binding could reduce the standard free energy of the transition state by on the order of 14 kcal/mole. The degree of folding rate enhancement depends, in addition, on the degree of interaction between the pro region and the relevant folding intermediate, and thus cannot be estimated at this stage of the analysis.

Structural studies of the pro region-protease complex should shed further light on the mechanism of

pro region-induced folding. Finally, since the pro region appears to function late in folding, it may be possible to trap a folding intermediate and to investigate the conformational changes and energetics involved in its conversion to the native state simply by controlling the time of addition of the pro region to a refolding reaction.

ACKNOWLEDGMENTS

We thank Daniel Santi, Judy White, Roger Bone, and members of the Agard lab for comments on the manuscript and Jonathan Davis for providing us with α-lytic protease. Ac-Pro-boroVal was kindly provided by Charles Kettner of DuPont. The *S. griseus* protease B was a kind gift from Scott Leigh of the Clorox Co. This work has been supported by funding from the Howard Hughes Medical Institute and from an NSF PYI Award (DAA). D.B. is a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation.

REFERENCES

1. Kim, P., Baldwin, R. L. Intermediates in the folding reactions of small proteins, Ann. Rev. Biochem. 59:631–660, 1990.
2. Silen, J. L., McGrath, C. N., Smith, K. R., Agard, D. A. Molecular analysis of the gene encoding α-lytic protease: Evidence for a preproenzyme. Gene 69:237–244, 1988.
3. Silen, J. L., Frank, D., Fujishige, A., Bone. R., Agard, D. A. Analysis of prepro-α-lytic protease expression reveals that the pro region is required for activity. J. Bact. 171:1320–1325, 1989.
4. Silen, J. L., Agard, D. A. The α-lytic protease pro region does not require a physical linkage to activate the protease domain in vivo. Nature 341:462–464, 1989.
5. Ikemura, H., Takagi, H., Inouye, M. Requirement of pro sequence for the production of active subtilisin in Escherichia coli. J. Biol. Chem. 262:7859–7864, 1987.
6. Zhu, X., Ohta, Y., Jordan, F., Inouye, M. Pro sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature 339:483–484, 1989.
7. Smith, D. B., Johnson, K. S. Single step purification of polypeptides expressed in E coli as fusions with glutathione-s-transferase. Gene 67:31–40, 1988.
8. Kettner, C. A., Bone, R., Agard, D. A., Bachovchin, W. B. Kinetic properties of binding of α-lytic protease to peptide boronic acids, Biochemistry 27:7682–7688, 1988.
9. Williams, J. W., Morrison, J. F. The kinetics of reversible tight binding inhibition. Methods in Enzymology 63:437–467, 1979.
10. Creighton, T. E. Towards a better understanding of protein folding pathways. Proc. Natl. Acad. Sci. U. S. A. 85:5082–5086, 1988.
11. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D., James M. N. Refined structure of α-lytic protease at 1.7 A resolution. J. Mol. Biol. 183:479–502, 1985.
12. Couto, J.R., Lo, E., Terzakis, C. The pro region of alkaline protease from Streptomyces griseus is required for expression of active enzyme in *E. coll*, J. Bact. (submitted) (1991).
13. Fujishige, A., Smith, K., Silen, J.L., Agard, D.A. Correct folding of α-lytic protease is required for its extracellular secretion from *E. coli*. J. Cell Biol. (submitted) (1991).
14. Laskowski, M., Kato, I. Protein inhibitors of proteinases. Ann. Rev. Biochem. 49:593–626, 1980.
15. Stroud, R. M., Kossiakoff, A. A., Chambers, J. L. Mechanisms of zymogen activation. Ann. Rev. Biophys. Bioeng. 6:177–193, 1977.
16. SanSegundo, B., Martinez, M. C., Vilanova, M., Cuchillo, C. M., Aviles, F. X. The severed activation segment of porcine pancreatic procarboxypeptidase is a powerful inhibitor of the active enzyme. Biochim. Biophys. Acta 707: 74–80, 1982.
17. James, M. N., Sielecki, A. R. Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 A resolution. Nature 319:33–38, 1986.
18. Ohta, Y., Hojo, H., Aimoto, S., Kobayashi, T., Zhu, X., Inouye, M. Propeptide as an intermolecular chaperone: renaturation of a denatured subtilisin E with a synthetic pro-peptide. Molecular Microbiology 5:1907–1510, 1991.
19. Goloubinoff, P., Christeller, J. T., Gatenby, A. A., Lorimer, G. H. Reconstitution of active dimeric rubulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. Nature 342:884–889, 1989.
20. Flynn, G. C., Chappell, T. G., Rothman, J. E., Peptide binding and release by proteins implicated as catalysts of protein assembly. Science 245:385–390, 1989.
21. Bochkareva, E. S., Lissin, N. M., Girschovich, A. S., Transient association of newly synthesized unfolded proteins with the heat shock Gro EL protein. Nature 336:254–257, 1988.