Reprinted from Nature, Vol. 341, No. 6241, pp. 462-464, 5th October, 1989

The α-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*

Joy L. Silen & David A. Agard

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

α-LYTIC protease, an extracellular serine protease of *Lysobacter enzymogenes 495*, is synthesized as a pre-pro-protein[1]. Previously it has been shown that when expressed in *Escherichia coli*, the protein is autocatalytically processed in the periplasmic space, and that the functional protease domain accumulates extracellularly[2]. Engineered proteins lacking the 166 amino-acid pro-region were enzymatically inactive and remained cell-associated[2]. By independently expressing the pro- and protease domains *in vivo*, evidence is provided here that direct covalent linkage is not required for production of active protease. We postulate that the pro-region acts as a template to promote the folding of the protease domain into an active configuration. Our results, combined with recent experiments on the evolutionarily unrelated subtilisin E (ref. 3), suggest that the ability of the pro-region of these bacterial proteases to facilitate folding of their protease domains is not a curiosity of a single system, but may reflect a general property of extracellular bacterial serine proteases.

On the basis of denaturation and renaturation experiments on a wide variety of small proteins[4-6], it seems that small, single-subunit, globular proteins fold spontaneously to reach a minimum energy state. For example, Anfinsen demonstrated that ribonuclease A could be unfolded and then refolded with full restoration of activity[4]. Unexpectedly, two small bacterial serine proteases, subtilisin (J. Wells, personal communication) and α-lytic protease (J. Richards, personal communication) could not be refolded. Experiments in both systems[2,7] indicated that the pro-region, which is covalently attached to, but only transiently associated with the protease domain, is necessary for correct folding of these enzymes. Several different models can be proposed for the role of the pro-region in folding the protease domain (Fig. 1). In models 1 and 2, the properly folded protease is considered to be at an energy minimum, but is kinetically inaccessible. That is, there are one or more non-productive folding pathways leading to improperly folded and possibly aggregated protein that have a lower transition-state energy and higher ground-state energy than the productive pathway. The pro-region would then facilitate production of active protease, either by stabilizing the transition state for the correct pathway (Fig. 1, model 1) or by destabilizing the non-productive pathways (Fig. 1, model 2). Unfortunately, it is not yet possible to distinguish between these two models as this requires a detailed kinetic analysis using purified components. Model 3 (Fig. 1*a*, *b*), however, which depends on the energy liberated from cleavage of the precursor, can be tested. In this model, the functional protease is not at an energy minimum (the misfolded forms on the non-productive pathway have lower energies), but is kinetically trapped. In such a model, the pro-region would act as in model 1 to lower the kinetic barrier, but in addition, the energy released by cleaving the precursor is required to shift the overall equilibrium in favour of the productive pathway. Once formed, the active high-energy protease would be unstable in the presence of functional pro-domain (Fig. 1*b*). To avoid unfolding active protease and subsequent accumulation in the lower-energy misfolded state, the pro-region must be physically separated from the protease or inactivated.

To test if the physical linkage and its subsequent cleavage plays a direct part in the folding process (model 3), it was necessary to independently express the pro- and protease domains within the periplasmic space. Expression systems with signal sequences were employed because the protease has three disulphide bonds and thus requires an oxidizing environment to fold properly. The expression construct, pCOMP5 (Fig. 2), uses the *phoA* promoter and signal sequence fused to the protease domain[2] and the *lpp/lac* promoter and *ompA* signal sequence[8] fused to the pro-region. This construct allows induction of the pro-region by isopropyl-β-D-thiogalactoside (IPTG) and the protease domain by phosphate starvation. Only the dual conditions of phosphate starvation and IPTG induction would allow complementation and, therefore, accumulation of the active enzyme.

The results of the induction experiment are shown in Fig. 3. As expected, independent induction of either the protease or pro-domains results in little or no activity. The low level of proteolytic activity seen in the low-phosphate control culture, in which only the protease domain is induced, probably results from the basal expression of the pro-region by the *lpp/lac* promoter[8]. In the cultures containing low phosphate plus IPTG, in which both promoters are induced, α-lytic protease activity

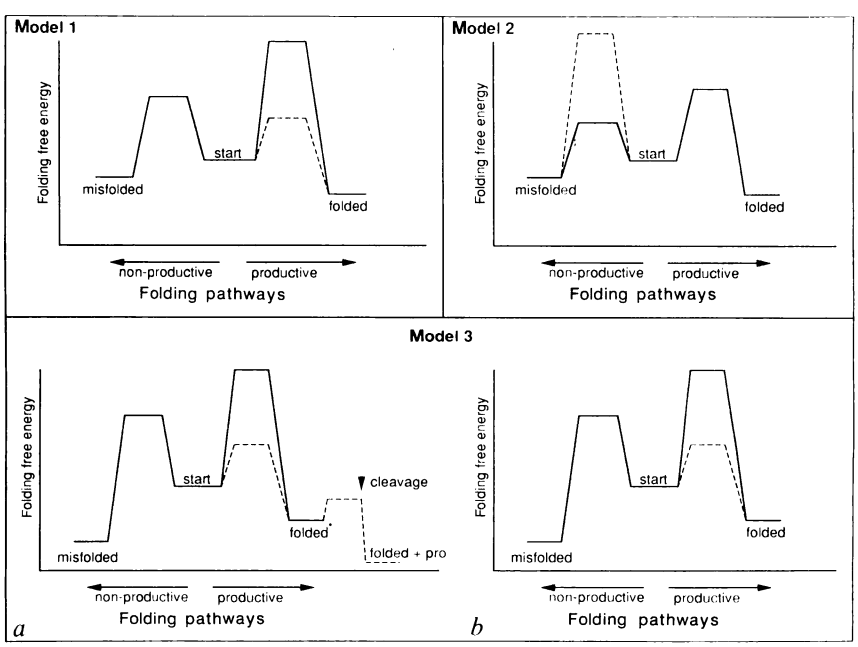

FIG. 1 Three models for the role of the pro-region in folding the protease domain. In all cases the folding process is considered to be a competition between a productive folding pathway that leads to the correctly folded protease, and a non-productive pathway that leads to misfolded and/or aggregated inactive protein. The free energy of folding from the point of view of the protease domain is shown in the absence (solid lines) and presence (dashed lines) of the pro-region. In model 1, the pro-region is seen to stabilize the transition state for folding, whereas in model 2 the pro-region destabilizes the transition state for the competing non-productive pathway. Models 1 and 2 are independent of the covalent attachment of the pro-region and thus, for simplicity, the energetics of the pro-region cleavage are not included. Model 3, depicting a kinetically trapped, higher energy folded state, is shown with either a covalently attached pro-region (panel *a*) or with a non-linked pro-region (panel *b*). The energy of bond cleavage can be used to shift the overall folding equilibrium in favour of the folded configuration only if the pro-region is covalently attached (*a* versus *b*). In model 3 there must be a folded (or mostly folded) intermediate that is still covalently linked to the pro-domain (folded* in panel *a*), which after cleavage results in the final, active form of the protease.

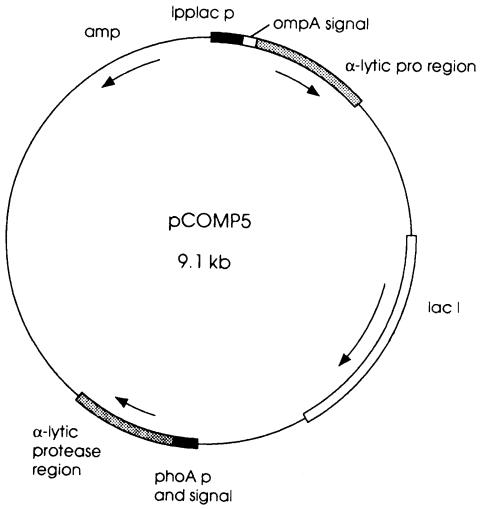

FIG. 2 The pCOMP5 plasmid (9.1 kilobases). This pIN-III-ompA2 derivative[8] places the pro-region of α-lytic protease under the control of the IPTG-inducible *lpp/lac* promoter and downstream of the *ompA* signal sequence. The protease domain of the α-lytic protease gene is fused to the *phoA* promoter and signal sequence; its expression depends on depletion of phosphate in the medium.
METHODS. The pro-region was originally isolated as a filled-in *Eae*I fragment from pDA13 (ref. 1) and cloned into a modified pDBR2 cut with *Sma*I and *Eco*RI (blunted). This construct (pPRO1) established *phoA*-promoter controlled expression and placed an in-frame stop codon six amino acids from the end of the pro-region. A filled-in *Eae*I-*Bal*I fragment from pPRO1 was cloned into the filled-in *Eco*RI site of pINIII-ompA2, which placed it in frame with the *ompA* signal sequence (pPRO3). The coding region for the protease domain of α-lytic protease was added to pPRO3 at a unique blunted *Sty*I site by transferring in a filled-in *Not*I-*Nar*I fragment from pMALP4 (a derivative pMALP2 (ref. 2). Orientation of the protease-encoding fragment was established by restriction enzyme analysis. Constructs were transformed into *E. coli* strain HB101 (*rec*A$^-$) for expression.

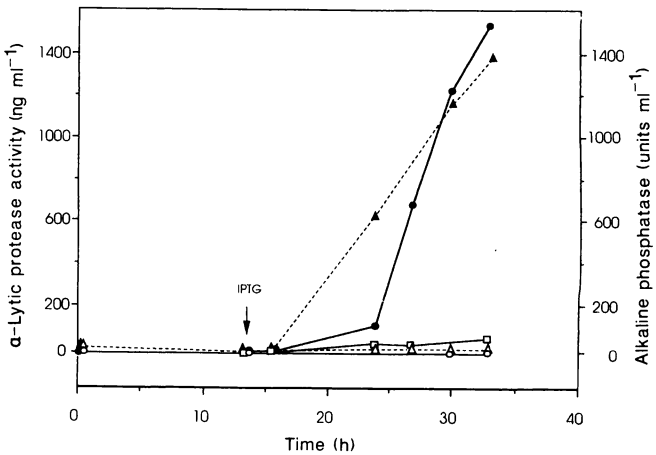

FIG. 3 Time course of α-Lytic protease and alkaline phosphatase production in cultures of HB101/pCOMP5. As the exact time of phosphate depletion is uncertain, levels of chromosomal alkaline phosphatase (--▲--) were monitored to follow the time course of the induction of the *phoA* promoter. Alkaline phosphatase levels under non-inducing conditions (high phosphate plus IPTG) remained at baseline levels (--△--). α-Lytic protease production was also monitored as a function of time: significant levels of activity are seen only when both the pro- and protease regions are co-expressed (compare low phosphate plus IPTG (—●—) with phosphate plus IPTG (—○—) or low phosphate alone (—□—)).
METHODS. Cells grown overnight in high-phosphate medium were concentrated, resuspended in low-phosphate medium, and diluted 1:100 in 250 ml low phosphate (0.1 mM K_2HPO_4/KH_2PO_4, pH 7.2) MOPS medium, or directly diluted 1:100 in 250 ml of high phosphate (20 mM K_2HPO_4/KH_2PO_4, pH 7.2) MOPS medium to an identical optical density at 600 nm ($OD_{600}=0.04$). Four 50 ml aliquots of the low-phosphate culture were taken into 500 ml shaker flasks for duplicates of the low-phosphate and low-phosphate-plus-IPTG cultures; two 50 ml aliquots of the high-phosphate culture were used for the high phosphate duplicates. Cultures were grown with shaking (200 r.p.m.) at 22 °C. After 14 h ($OD_{600}=1.0$), IPTG (2 mM final concentration) was added to the duplicate high-phosphate cultures and to one pair of the low-phosphate cultures. Levels of alkaline phosphatase (both from the cells and the medium), and α-lytic protease (medium only) were determined as previously described[2], using *p*-nitrophenyl phosphate and suc-Ala-Pro-Ala-*p*-nitroanalide, respectively, as substrates (α-lytic protease purified from *Lysobacter enzymogenes* has 188 units mg^{-1}; 1 unit hydrolyses 1 μmol substrate per min). Data shown is the average of the duplicate cultures, except for the low-phosphate culture alkaline phosphatase induction, which is the average of all four low-phosphate cultures, plus and minus IPTG.

reaches levels more than 30 times that of the low-phosphate control culture. As the accumulation of active α-lytic protease occurs after the addition of IPTG and follows the induction of cellular alkaline phosphatase, it is clear that co-induction is required for α-lytic protease production.

These experiments, and the *in vitro* experiments with subtilisin[4], demonstrate the independence of the folding process from the direct linkage between the pro- and protease domains, thereby eliminating the kinetic trapping model (Fig. 1, model 3). By uncoupling the irreversibility of bond cleavage from the folding process, the thermodynamics of folding of the protease domain can be more accurately considered, and more easily investigated experimentally. Because active α-lytic protease displays remarkable temperature stability[9], resistance to autolysis and a high degree of structural rigidity[10], there is probably a high-energy barrier close to the folded state that provides the active enzyme with its high stability, as in model 1. Apparently, this barrier is too high to be traversed unaided. Thus, the pro-region can be thought of as a template that preferentially binds to, and stabilizes, the transition state for folding, just as an enzyme binds and stabilizes the transition state for catalysis of a reaction.

It is quite remarkable that these two classes of serine protease[5], which are unrelated in structure and sequence, have, through convergent evolution, developed both a spatially identical Asp-His-Ser catalytic triad and what appears to be a similar folding mechanism. Molecular analysis of the genes for a wide variety of secreted proteases from both Gram-positive and Gram-negative bacteria[11-18] reveals that these enzymes are also synthesized in precursor form. The widespread use of pro-regions by extracellular bacterial serine proteases indicates that these domains must have an extremely important function.

The folding activity of these pro-regions is reminiscent of the recently discovered class of proteins, called chaperonins, which facilitate the folding and assembly of multi-subunit proteins in chloroplasts[19], mitochondria[20] and of bacteriophage in *E. coli*[19].

Although the chaperonins are capable of aiding the folding of a number of proteins, these extracellular bacterial serine proteases seem to have evolved their own specialized folding aids which are covalently attached. Further examination of the structure and mechanism of the α-lytic protease pro-region may provide new insights into the function of other more generalized folding and assembly systems. □

Received 10 May; accepted 14 August 1989.

1. Silen, J. L., McGrath, C. N., Smith, K. R. & Agard, D. A. *Gene* **69,** 237-244 (1988).
2. Silen, J. L., Frank, D., Fujishige, A., Bone, R. & Agard, D. A. *J. Bact.* **171,** 1320-1325 (1989).
3. Zhu, X., Ohta, Y., Jordan, F. & Inouye, M. *Nature* **339,** 483-484 (1989).
4. Anfinsen, C. B. *Proc. natn. Acad. Sci. U.S.A.* **47,** 1308-1314 (1961).
5. Creighton, T. E. *J. phys. Chem.* **89,** 2452-2459 (1985).
6. Ikai, A., Fish, W. & Tanford, C. J. *J. molec. Biol.* **145,** 265-280 (1973).
7. Ikemura, H., Takagi, H. & Inouye, M. *J. biol. Chem.* **262,** 7859-7864 (1987).
8. Ghrayeb, J. *et al. EMBO J.* **10,** 2437-2422 (1984).
9. Kaplan, H., Symonds, V. B., Dugas, H. & Whitaker, D. R. *Can. J. Biochem.* **48,** 649-658 (1970).
10. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D. & James, M. N. G. *J. molec. Biol.* **183,** 479-502 (1985).
11. Jacobs, M., Eliasson, M., Uhlen, M. & Flock, J. I. *Nucleic Acids Res.* **13,** 8914-8926 (1985).
12. Shimada, H. *et al. Biotechnology* **2,** 75-85 (1985).
13. Stahl, M. L. & Ferrari, E. *J. Bact.* **158,** 411-418 (1984).
14. Takagi, M., Imanaka, T. & Aiba, S. *J. Bact.* **163,** 824-831 (1985).
15. Vasantha, N. *et al. J. Bact.* **159,** 811-819 (1984).
16. Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A. & Chen, E. Y. *Nucleic Acids Res.* **11,** 7911-7925 (1983).
17. Yang, M. Y., Ferrari, E. & Henner, D. J. *J. Bact.* **160,** 15-21 (1984).
18. Henderson, G., Krygsman, P., Liu, C. J., Davey, C. C. & Malek, L. T. *J. Bact.* **169,** 3778-3784 (1987).
19. Hemmingsen, S. *et al. Nature* **333,** 330-334 (1988).
20. McMullin, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **8,** 371-380 (1988).

ACKNOWLEDGMENTS. We thank Roger Bone, Amy Fujishige and Hiten Madhani for comments and M. Inouye for providing the pIN-III-ompA2 expression plasmid. This work was supported by the NSF Presidential Young Investigator Program and the Howard Hughes Medical Institute.

Printed in Great Britain by Turnergraphic Limited, Basingstoke, Hampshire