Volume 244, number 1, 31-33 FEB 06767 February 1989

Location of enzymatic and DNA-binding domains on *E. coli* protease La

Michael E. Baker

Department of Medicine, M-023, University of California, San Diego, La Jolla, CA 92093, USA

Received 14 November 1988

Escherichia coli protease La is an ATP-dependent enzyme that has a DNA-binding site. The locations of the enzymatic and DNA-binding sites are not known. We report that a 75-residue segment at the carboxy-terminus of the protease La is similar to part of *Bacillus licheniformis* β-lactamase, a serine enzyme. The comparison score is 8.2 standard deviations higher than that obtained with 10000 comparisons of randomized sequences of these segments. The probability of obtaining such a score by chance is 1.2×10^{-16}. We also find that a 107-residue segment in the amino-terminus half of protease La is similar to part of the sopB protein, a DNA-binding protein of the plasmid F of *E. coli*. The comparison score for these segments is 8 standard deviations ($P = 6 \times 10^{-16}$). These strong amino acid sequence similarities suggest the locations of the catalytic serine and the DNA-binding domains of protease La.

Protease La; ATP-dependent protease; DNA-binding protease

1. INTRODUCTION

In *Escherichia coli,* as in eukaryotic cells, metabolic energy is required for both the rapid degradation of highly abnormal polypeptides and the increased breakdown of normal cell proteins during starvation [1]. Protease La has been identified as one of the enzymes that is important in the hydrolysis of abnormal proteins in *E. coli* [2-4]. Protease La hydrolyzes denatured proteins and ATP in a coupled manner [3-5]; its proteolytic and ATPase activities are stimulated by DNA [6], and, consistent with one of its presumed functions, protease La synthesis increases when cells accumulate abnormal proteins [7,8]. The recent cloning and sequencing of the gene for protease La [2] provide an opportunity to use site-specific mutagenesis techniques to establish how its binding sites for abnormal proteins, ATP, and DNA interact in proteolysis of abnormal proteins. First, of course, one needs to know the location of protease, DNA, and ATP binding sites. Our interest in steroid receptors, which bind synthetic protease inhibitors and substrates [9,10], as well as ATP and DNA, led us to compare the sequence of protease La with steroid receptors, enzymes, DNA-binding proteins, and other proteins in the database seeking clues to the evolution of proteases and steroid receptors. As reported here, we found that protease La contains a domain similar to *Bacillus licheniformis* β-lactamase, a serine enzyme [11,12], and to the DNA-binding sopB protein [13], a trans-acting protein [14] in *E. coli* plasmid F, that is important in plasmid stability (*sop,* stability of plasmid). Based on these sequence similarities and the common actions of these proteins, we suggest the location of catalytic serine and the DNA-binding domain on protease La.

Correspondence address: M.E. Baker, Department of Medicine, M-023, University of California, San Diego, La Jolla, CA 92093, USA

2. RESULTS

The sequence of protease La was searched against the National Biomedical Research Foundation protein sequence database with the FASTA program [15]. Figs 1 and 2 show the segments on *B.*

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

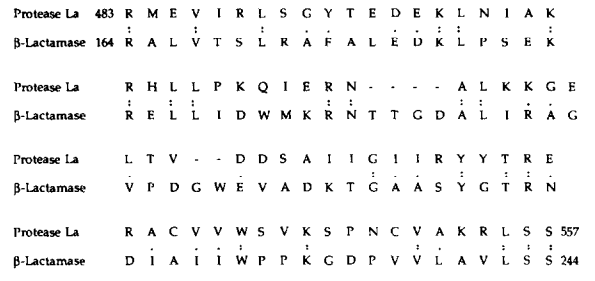

Fig.1. Comparison of *E. coli* protease La with *B. licheniformis* β-lactamase. (:) Identities; (.) conservative replacements. Out of 75 possible matches there are 23 identities (31%) and 15 conservative replacements (20%). The ALIGN analysis of these segments gives a score that is 8.2 standard deviations higher than that obtained with 10000 comparisons of randomized sequences of these segments. The probability of obtaining such a score by chance is 1.2×10^{-16}.

licheniformis β-lactamase and *E. coli* sopB protein that are similar to parts of protease La. The similarities were quantified using the ALIGN program [16], with a penalty of 8 for each gap. The comparison score of residues 483–557 of protease La and residues 164–244 of β-lactamase is 8.2 standard deviations higher than that obtained with 10000 comparisons of randomized comparisons of these proteins. The probability of obtaining such a score by chance is 1.2×10^{-16}. The comparison score for residues 41–148 of protease La and residues 86–192 of sopB protein is 8 standard deviations ($P=6\times10^{-16}$). Although computer analyses cannot distinguish between protease La and either β-lactamase or sopB protein being derived from a common ancestor or if they are examples of convergent evolution, scores of this magnitude suggest some common functions for these segments, especially if there is evidence for common biological actions, which is clearly the case for these proteins.

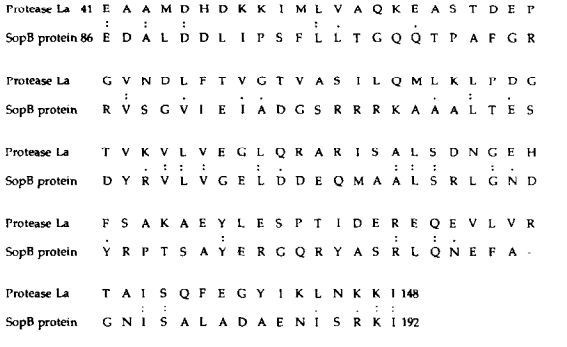

Fig.2. Comparison of protease La with sopB protein. (:) Identities; (.) conservative replacements. Out of 107 possible matches there are 21 identities (20%) and 14 conservative replacements (13%). The ALIGN comparison score is 8 standard deviations higher than that obtained with 10000 comparisons of randomized sequences of these segments. The probability of obtaining such a score by chance is 6×10^{-16}.

3. DISCUSSION

3.1. *Location of the catalytic serine on La protease*

The catalytic nucleophile in protease La has not yet been identified. Because protease La is inhibited by diisopropyl fluorophosphate (DFP) [17,18] as well as halomethyl ketones [18], it has been suggested that protease La is a serine protease. Lack of similarity of protease La to either mammalian or bacterial serine proteases has left this conclusion unresolved. The similarity to β-lactamase that we report here provides additional support for the notion that protease La is a serine protease because β-lactamases use a serine residue for their catalytic action and may share a common ancestor with serine proteases [19].

To locate the catalytic serine on protease La, we used two properties of β-lactamase. (i) The catalytic Ser-52 (codon TCG) is 112 residues towards the amino-terminus from the segment of β-lactamase shown in fig.1. (ii) The catalytic serine is part of a Ser-X-X-Lys motif that is conserved in all β-lactamases [12,19], even though other parts of their sequences are substantially different. Ser-367 (codon TCC) of protease La is about the right distance (116 residues) on the amino-terminus side of the segments shown in fig.1, and it is part of a Ser-X-X-Lys-Ala segment that has three identities with the catalytic site of *B. licheniformis* β-lactamase [12,19], suggesting that Ser-367 is the catalytic residue in protease La.

Chin et al. [2] proposed that residues 346–425 of protease La contain the ATP-binding site on the basis of the sequence and secondary structure of the ATP-binding site on various kinases, including adenylate kinase. An ALIGN analysis of residues 7–149 of rabbit adenylate kinase with residues 347–503 of protease La yielded a score of 6 standard deviations ($P=10^{9}$), supporting their conclu-

sion. Lys-361 of protease La aligns with adenylate kinase's Lys-21, a highly conserved residue that binds ATP. This places the proposed ATP-binding site in overlap with the proposed protease active site, supporting the notion of a novel ATP-dependent mechanism for protease La, involving Ser-367 and possibly Lys-361 and Lys-370.

3.2. *DNA-binding site*

The sopB protein is a trans-acting protein that stabilizes partition of plasmid F in *E. coli* [12,13]. The similarity between the sequences of protease La and sopB and the fact that the latter is a DNA-binding protein suggest that the amino-terminus part of protease La contains the DNA-binding site. The function of the DNA-binding domain on protease La is not known. Its similarity to sopB leaves open the possibility that protease La has a role in gene transcription.

Protease La appears to have arisen by combination of genes coding for proteins with proteolytic, phosphotransferase, and DNA-binding activity. The overlap of the proposed ATP and proteolytic sites raises the possibility that the phosphotransferase and protease and protease activities are derived from a common ancestor. Recently, Brenner [19] suggested that a variety of enzymes, including β-lactamases, serine proteases, phosphatases, and lipases from diverse prokaryoties and eukaryotes, descended from a common cysteine protease ancestor. After the sequences of other homologs of protease La are determined, it may be possible to uncover other evolutionary relationships between this protein family and other enzymes.

Acknowledgements: I thank Darrell Fanestil, Alfred Goldberg, and Donald Helinski for review of an earlier version of this manuscript.

REFERENCES

[1] Hershko, A. and Ciechanover, A. (1982) Annu. Rev. Biochem. 51, 335–364.
[2] Chin, D.T., Goff, S.A., Webster, T., Smith, T. and Goldberg, A.L. (1988) J. Biol. Chem. 263, 11718–11728.
[3] Charette, M.F., Henderson, G.W. and Markovitz, A. (1981) Proc. Natl. Acad. Sci. USA 78, 4728–4732.
[4] Chung, C.H. and Goldberg, A.L. (1981) Proc. Natl. Acad. Sci. USA 78, 4931–4935.
[5] Menon, A.S., Waxman, L. and Goldberg, A.L. (1987) J. Biol. Chem. 262, 722–726.
[6] Chung, C.H. and Goldberg, A.L. (1982) Proc. Natl. Acad. Sci. USA 79, 795–799.
[7] Goff, S.A. and Goldberg, A.L. (1985) Cell 41, 587–595.
[8] Goff, S.A. and Goldberg, A.L. (1987) J. Biol. Chem. 262, 4508–4515.
[9] Baker, M.E., Vaughn, D. and Fanestil, D.D. (1978) J. Supra Mol. Struct. 9, 421–426.
[10] Baker, M.E. and Terry, L. (1983) Steroids 42, 593–602.
[11] Neugenbauer, K., Sprengl, R. and Schaller, H. (1981) Nucleic Acids Res. 9 2577–2588.
[12] Joris, B., Ghuysen, J.-M., Dive, G., Renard, A., Dideberg, O., Charlier, P., Frere, J.-M., Kelly, J.A., Boyington, J.C., Moew, P.C. and Knox, J.R. (1988) Biochem. J. 250, 313–324.
[13] Mori, H., Kondo, A., Ohshima, A., Ogura, T. and Hiraga, S. (1986) J. Mol. Biol. 192, 1–15.
[14] Ogura, T. and Hiraga, S. (1983) Cell 32, 351–360.
[15] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[16] Dayhoff, M.O., Barker, W.C. and Hunt, L.T. (1983) Methods Enzymol. 91, 524–545.
[17] Waxman, L. and Goldberg, A.L. (1982) Proc. Natl. Acad. Sci. USA 79, 4883–4887.
[18] Waxman, L. and Goldberg, A.L. (1985) J. Biol. Chem. 260, 12022–12028.
[19] Brenner, S. (1988) Nature 334, 528–530.