Volume 14 Number 14 1986 Nucleic Acids Research

Cloning and sequencing of *Serratia* protease gene

Kazuo Nakahama, Koji Yoshimura, Ryuji Marumoto, Masakazu Kikuchi, In Sook Lee[1], Toshiharu Hase[1] and Hiroshi Matsubara[1]

Biotechnology Laboratories, Central Research Division, Takeda Chemical Industries, Ltd., Yodogawa-ku, Osaka 532 and [1]Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan

Received 17 June 1986; Accepted 30 June 1986

ABSTRACT

The gene encoding an extracellular metalloproteinase from Serratia sp. E-15 has been cloned, and its complete nucleotide sequence determined. The amino acid sequence deduced from the nucleotide sequence reveals that the mature protein of the Serratia protease consists of 470 amino acids with a molecular weight of 50,632. The G+C content of the coding region for the mature protein is 58 %; this high G+C content is due to a marked preference for G+C bases at the third position of the codons. The gene codes for a short pro-peptide preceding the mature protein. The Serratia protease gene was expressed in Escherichia coli and Serratia marcescens; the former produced the Serratia protease in the cells and the latter in the culture medium. Three zinc ligands and an active site of the Serratia protease were predicted by comparing the structure of the enzyme with those of thermolysin and Bacillus subtilis neutral protease.

INTRODUCTION

In culture medium Serratia sp. E-15 produces a potent protease (1), which is widely used as an anti-inflammatory agent. The enzyme has a zinc atom, which is essential for proteolytic activity, and its substrate specificity is somewhat similar to that of thermolysin (EC 3.4.24.4) produced by Bacillus thermoproteolyticus (1-3). The molecular weight of the Serratia protease has been estimated to be 45,000-60,000 by various physical methods (1,4), however, the complete amino acid sequence has not been determined. Lee et al. reported the amino acid sequences of the NH_2-terminal half (278 residues) and the COOH-terminal peptide (13 residues) of the enzyme and proposed possible zinc ligands (two histidine residues) on the basis of structural similarity with thermolysin (5,6).

Schmitz and Braun (7) also purified an exoprotease of Serratia marcescens ATCC 25419 and determined its NH_2-terminal (10 residues) and COOH-terminal (4 residues) amino acid sequences, which are almost identical with the sequences of the Serratia protease reported by Lee et al. (5,6).

It has been difficult to determine the complete amino acid sequence of the Serratia protease because of the large molecular weight. We have attempted to determine the amino acid sequence from the nucleotide sequence of the Serratia protease gene. The complete nucleotide sequences of only two Serratia genes, the glutamine

amidotransferase gene (8) and the chitinase gene (9) from S. marcescens, have been reported. In this paper, we describe the cloning and DNA sequencing of the protease gene from Serratia sp. E-15, and the deduced amino acid sequence of the enzyme. We also predict three zinc ligands and an active site of the enzyme by comparing the enzyme with thermolysin (10,11) and Bacillus subtilis protease (12,13).

MATERIALS AND METHODS

Materials

Restriction enzymes were purchased from New England Biolabs, Takara Shuzo, or Nippon Gene. T4 DNA ligase and Escherichia coli DNA polymerase I were from New England Biolabs. T4 polynucleotide kinase was from Takara Shuzo. [γ-^{32}P]ATP (3000 Ci/mmol) and [α-^{32}P]dCTP (>400 Ci/ mmol) were purchased from Amersham. M13 sequencing kit was from Takara Shuzo. Oligonucleotides were synthesized by the phosphoamidite method (14) using DNA synthesizers (Genet A-II, Nippon Zeon; 380A DNA Synthesizer, Applied Biosystems). Authentic Serratia protease and rabbit anti-Serratia protease antiserum were provided by Dr. K. Miyata of our Central Research Division.

Bacterial strains, plasmids, and media

Serratia sp. E-15 was obtained from the Institute for Fermentation, Osaka. E. coli JM103 (15) and E. coli DH1 (16) were used as hosts. S. marcescens NC-4 was provided by Dr. H. Ono of our Central Reasearch Division. Plasmid pUC12 and M13 phage vectors were purchased from P-L Biochemicals. Bacteria were grown in L-broth (17) or brain heart infusion (Difco).

Isolation of DNA

Chromosomal DNA of Serratia sp. E-15 was prepared as described by Lovett and Keggins (18). Plasmid DNA was isolated by the alkaline extraction procedure (19).

Transformation

E. coli JM103 and E. coli DH1 were transformed by the $CaCl_2$/RbCl procedure (16). In the case of E. coli JM103, cells were grown on L-agar containing ampicillin, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside and isopropyl-β-D-thiogalactopyranoside, and white colonies were used for colony hybridization. S. marcescens NC-4 was transformed by the $CaCl_2$ procedure (16).

Colony hybridization

Colony hybridization was carried out as described by Maniatis et al. (16). Transformants were transferred onto nitrocellulose filters on L-agar plates and grown overnight at 37°C. The colonies on the filters were lyzed, and the liberated DNA was fixed to the filters by baking. The synthetic oligonucleotides were labeled by phosphorylation with [γ-^{32}P]ATP and T4 polynucleotide kinase. The DNA fragment was labeled by nick translation with [α-^{32}P]dCTP and E. coli DNA polymerase I. Filters were

hybridized with labeled oligonucleotides overnight at 37°C in 6 x SSC-0.01 M EDTA-5 x Denhardt's solution-0.5 % SDS-100 μg/ml denaturated salmon sperm DNA and washed with 2 x SSC-0.1 % SDS at 37°C. Hybridization and washing conditions for the labeled fragment were as previously described (16).

DNA sequencing

DNA sequencing was carried out by means of the M13 cloning protocols of Messing et al. (15) and the dideoxy method of Sanger et al. (20). All experiments were performed according to the manuals of the sequencing kit supplied by Bethesda Research Laboratories and Takara Shuzo. Reactions for the sequencing were carried out at 37°C.

Western blotting analysis

Samples were subjected to electrophoresis on a 7.5 % polyacrylamide gel in the presence of 0.1 % SDS, according to the method of Laemmli (21). Western blotting was done by the method of Burnette (22). Serratia protease was detected using purified rabbit anti-Serratia protease antibody and goat anti-rabbit immunogloblin G-horseradish peroxidase conjugate (Bio-Rad Laboratories).

Assay of protease activity

Protease activity was assayed as described by Miyata et al. (1).

RESULTS

Cloning of Serratia protease gene

On the basis of the published amino acid sequences of the Serratia protease at positions 143-147 and 42-51 (6), the 14-mer oligonucleotide probes dTC(T/C)TG(A/G/T/C)CC(T/C)TGCCA and the 29-mer oligonucleotide probes dACTTTGTA(T/G)CCGTCCCA(T/G)GTTTGGTTTTC, respectively, were chemically synthesized (14). Chromosomal DNA of Serratia sp. E-15 was digested with BamHI and ligated into the BamHI site of pUC12. Out of 700 colonies, one colony hybridized with both probes. The plasmid from the positive colony was designated pTSP20. The restriction map of the cloned Serratia protease gene in pTSP20 is shown in Figs. 1 and 2A. As the result of the DNA sequencing, the insert in pTSP20 was found to lack the sequence encoding the COOH-terminal region of the mature protein of the enzyme. A further cloning of the gene was carried out using the ^{32}P-labeled 0.9 kb PstI-BamHI fragment from pTSP20 as a probe (shown in Fig. 2). The chromosomal DNA of Serratia sp. E-15 was digested with HindIII and ligated into the the HindIII site of pBR322. Out of about 6,000 colonies, 4 colonies hybridized to the probe. The plasmid isolated from one of the positive colonies was designated pTSP21 (Fig. 1). The DNA sequencing indicated that the insert in pTSP21 contained the sequence encoding the whole mature protein as shown in Fig. 2B.

The chromosomal DNAs of Serratia sp. E-15 and E. coli were digested with BamHI,

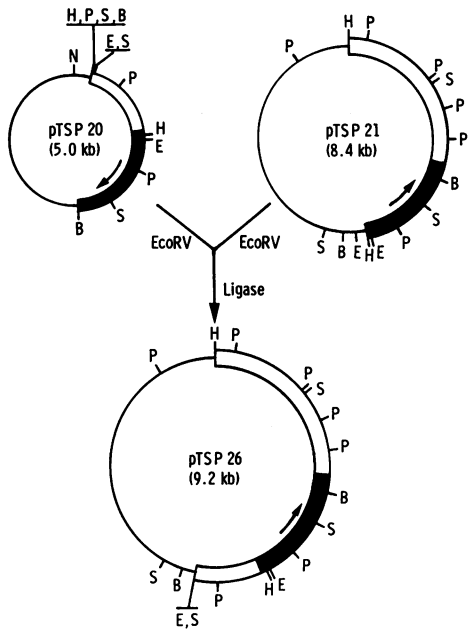

Fig. 1. Construction of the Serratia protease expression plasmid pTSP26 from pTSP20 and pTSP21. The thin lines represent the vectors. The boxes indicate the DNA fragments derived from Serratia sp. E-15, and the shaded regions show the positions of the Serratia protease gene. The arrows indicate the direction of the gene. Restriction sites are abbreviated as follows: H, HindIII; P, PstI; S, SalI; B, BamHI; E, EcoRV; N, NdeI.

electrophoresed on an agarose gel, and transferred onto a nitrocellulose filter. One fragment (2.3 kb) of the Serratia chromosomal DNA hybridized with the ^{32}P-labeled 0.9 kb PstI-BamHI fragment from pTSP20 whereas E. coli chromosomal DNA did not (data not shown). This result shows that the 0.9 kb PstI-BamHI fragment came from the Serratia chromosomal DNA.

Nucleotide sequence of the Serratia protease gene

The nucleotide sequences of two fragments from pTSP20 and pTSP21 were determined according to the strategy shown in Fig. 2. The two fragments overlap; one (pTSP20) lacks the 3' region of the Serratia protease gene and the other (pTSP21) lacks the 5' noncoding region. The nucleotide sequence determined from the two fragments is shown in Fig. 3. It contains a long open reading frame of 1,509 nucleotides extending from positions 728 to 2,236. Lee et al. determined the amino acid sequences of the NH_2-terminal 278 residues and COOH-terminal 13 residues of the Serratia protease (6). One part of the amino acid sequence deduced from the open reading frame almost completely coincides with the sequence of the NH_2-terminal 278 residues. There are some

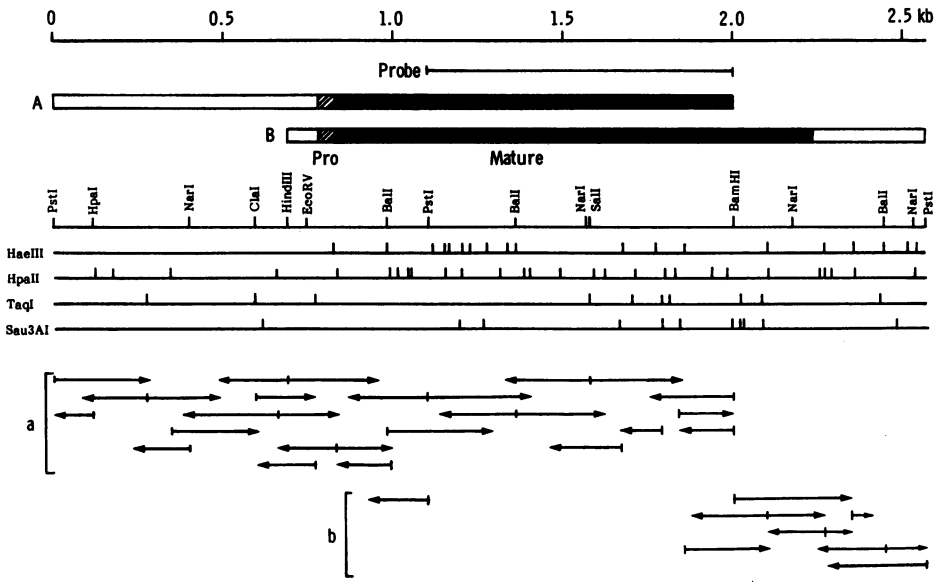

Fig. 2. Restriction map and sequencing strategy of the Serratia protease gene. The boxes indicate the DNA fragments from pTSP20 (A) and pTSP21 (B). The cross-hatched and shaded regions represent the pro-peptide and mature coding regions, respectively. The arrows indicate the direction and extent of sequencing from restriction sites of DNA fragments from pTSP20 (a) and pTSP21 (b). The horizontal line above the box (A) shows the 0.9 kb PstI-BamHI fragment used as the probe.

differences between the two amino acid sequences: Asp is replaced by Asn at positions 20 and 47, Gly-Asp is replaced by Asn-Gly at positions 25-26. The other part of the deduced amino acid sequence completely coincides with the sequence of the COOH-terminal 13 residues and is followed by a termination codon (TAA). It is clear that the sequence (1,410 nucleotides) encoding from the NH_2-terminal 278 residues to the COOH-terminal 13 residues corresponds to the coding region for the mature protein of Serratia protease. This protein contains 470 amino acids and has a calculated molecular weight of 50,632. Table I shows the amino acid composition of the mature protein calculated from the deduced amino acid sequence and that found by amino acid analysis (6). All values are identical within the experimental errors of amino acid analysis. The enzyme contains no cysteine like many other extracellular proteins of bacteria (23).

There are two ATG codons at nucleotides 728 and 779 preceded by sequences resembling the Shine-Dalgarno (SD) sequence for the long open reading frame. The two possible initiation codons must generate the translation products with molecular weights of 54.3 kd (kilodaltons) and 52.4 kd, which are larger than the mature protein (50.6 kd) by 3.7 kd and 1.8 kd, respectively. The experiment for expressing the Serratia protease

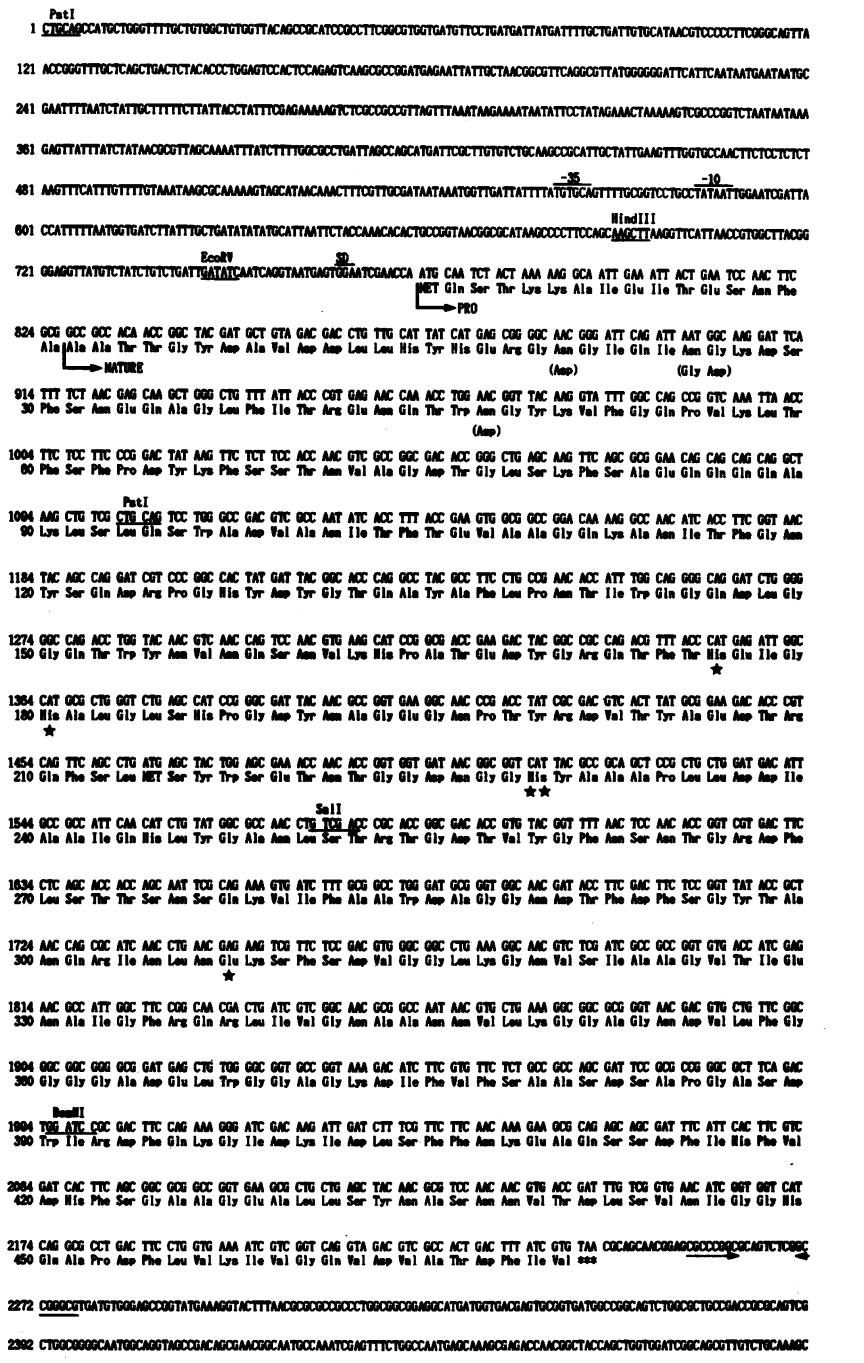
PstI
-35
-10
HindIII
EcoRV
SD
MET Gln Ser Thr Lys Lys Ala Ile Glu Ile Thr Glu Ser Asn Phe
PRO
Ala Ala Ala Thr Thr Gly Tyr Asp Ala Val Asp Asp Leu Leu His Tyr His Glu Arg Gly Asn Gly Ile Gln Ile Asn Gly Lys Asp Ser
MATURE
(Asp)
(Gly Asp)
(Asp)
PstI
SalI
BamHI

Table I. Amino acid composition of *Serratia* protease

Amino acid	Residue	
	Deduced from DNA sequence	Amino acid analysis[a)]
Ala	47	49
Arg	12	11
Asn	36	
Asp	37	79
Cys	0	0
Gln	26	
Glu	15	39
Gly	56	62
His	12	12
Ile	23	22
Leu	28	28
Lys	17	17
Met	1	1
Phe	30	31
Pro	10	10
Ser	35	35
Thr	32	29
Trp	8	ND[b)]
Tyr	19	17
Val	26	27

a) All values are adjusted from those reported by Lee *et al.* (6) on the basis of the molecular weight, 50,632. b) not determined.

gene in *E. coli* showed that the gene product was larger than the mature protein by 1 kd as described later. From these results, the most reasonable initiation codon for the open reading frame is the methionine residue at nucleotide 779. It is suggested that the mature protein coding region is preceded by the coding region for the pro-peptide consisting of 16 amino acids. The coding region for the signal peptide is not found in the determined nucleotide sequence.

Upstream from the long open reading frame, there are putative -35 (TGTGCA) and -10 (TATAAT) sequences; the spacer between the two sequences is 16 nucleotides. The

Fig. 3. Nucleotide sequence of the *Serratia* protease gene and deduced amino acid sequence of the enzyme. The nucleotide sequence is numbered from the 5' end of the sequenced fragment. Numbering of the amino acids starts at the NH_2-terminus of the mature protein of the *Serratia* protease. The putative -35 and -10 sequences and potential SD sequence are overlined. The putative transcription terminator is indicated by arrows. Only the differences in the amino acid sequence reported by Lee *et al.* (6) are given below the deduced amino acid sequence. PRO is the pro-peptide and MATURE is the mature protein. The probable zinc ligands and the probable active site of the enzyme are indicated by single asterisks (*) and double asterisk (**), respectively.

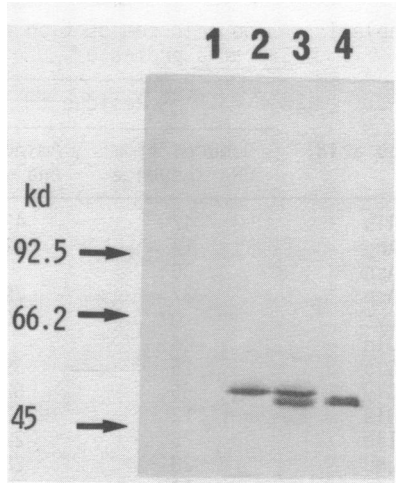

Fig. 4. Western blotting analysis of the product by E. coli. Molecular weights were calculated from the mobilities of the following standards: phosphorylase B (Mr=92,500), bovine serum albumin (66,200) and ovalbumin (45,000). Lane 1, sample from E. coli JM103/pBR322; lane 2, sample from E. coli JM103/pTSP26; lane 3, a mixture of the sample from E. coli JM103/pTSP26 and authentic Serratia protease; lane 4, the authentic Serratia protease.

open reading frame is followed by a palindromic sequence, which appears to be a transcription terminator (Fig. 3).

Another open reading frame is present downstream from the putative transcription terminator (from nucleotide 2,338), although the sequence downstream from the PstI site has not yet been determined.

Expression of the Serratia protease gene

The cloned insert in pTSP20 lacks the 3' region (about 240 bp) of the protease gene, and the one in pTSP21 lacks the promoter region. A Serratia protease expression plasmid, pTSP26, was constructed by isolating the 1.0 kb EcoRV fragment containing the promoter region of the gene from pTSP20 and ligating it with the 8.2 kb EcoRV fragment from pTSP21 (Fig. 1).

E. coli harboring pTSP26 showed no detectable zone of clearing on an agar plate containing 3 % skim milk, while Serratia sp. E-15 created a large clearing zone.

E. coli JM103/pTSP26 and E. coli JM103/pBR322 were grown in brain heart infusion medium for 24 hr at 28°C on a rotary shaker. The cultures were centrifuged, and the cells washed twice with 50 mM Tris-HCl (pH 8.0)-50 mM NaCl and frozen. The frozen cells were suspended in 1/10 original volume of the same buffer and disrupted with a Kaijo Denki sonic oscillator for 2 min at 2 A. The treated cell suspensions were centrifuged to give extracts. The proteins in the extracts were precipitated by

ammonium sulfate and acetone as described by Miyata et al. (1) and analyzed by the Western blotting. The sample from E. coli JM103/pTSP26 showed an immuno-reactive protein with a molecular weight of 51 kd while authentic Serratia protease migrated to the position corresponding to a molecular weight of 50 kd. The immuno-reactive protein was not detected in the sample from E. coli JM103/pBR322 (Fig. 4).

S. marcescens NC-4/pTSP26 and S. marcescens NC-4 were grown as just described, and the culture supernatants were assayed for protease activity. The protease activity of S. marcescens NC-4/pTSP26 was more than twice as high as that of S. marcescens NC-4. The protease thus produced was identical with the authentic Serratia protease, as examined by the Western blotting.

DISCUSSION

Lee et al. determined the amino acid sequences of the NH_2-terminal half and COOH-terminal peptide of the Serratia protease (5,6), which is widely used for medical purposes. As described above, we have cloned the Serratia protease gene and determined its complete nucleotide sequence. The amino acid sequence deduced from the nucleotide sequence is almost identical to the partial sequence determined by Lee et al. (5,6). There are some differences between the two determinants described above. Most of the differences are between Asn and Asp, and seem to be due to an artifact during the Edman degradation of the protein.

The Serratia protease has a zinc atom, which is essential for proteolytic activity, and its substrate specificity is somewhat similar to that of thermolysin (1-3). In thermolysin, the zinc ligands are His 142, His 146, and Glu 166, and the active site is His 231 (11). Glu 143 and Asp 226 are also thought to participate in the catalysis (24). Lee et al. have compared the amino acid sequence of the NH_2-terminal half of the Serratia protease with the sequence of thermolysin, and showed that a region containing two histidine residues (at positions 176 and 180) in the Serratia protease is similar to the zinc binding region in thermolysin (5,6). The amino acid sequence of the neutral protease from B. subtilis (Bacillus amyloliquefaciens) has also been shown to be similar to that of thermolysin; the three zinc ligands present in thermolysin are also present in the neutral protease (His 143, His 147, and Glu 167) (25). The zinc ligands and the active site of the Serratia protease were searched by comparing the structure of the enzyme with the structures of thermolysin (10,11) and B. subtilis neutral protease (12,13). The region from residue 176 to 182 in Serratia protease is similar to the regions containing two zinc ligands in thermolysin (residues 142-148) and in B. subtilis neutral protease (residues 143-149) (Fig. 5A). The region from residue 305 to 315 in Serratia protease is very similar to the regions containing another zinc ligand (Glu) in thermolysin (residues 164-173) and B. subtilis neutral protease (residues 165-174), although a residue, Lys, is inserted in Serratia protease (Fig. 5B). The region from residue 224 to 230 in Serratia protease is

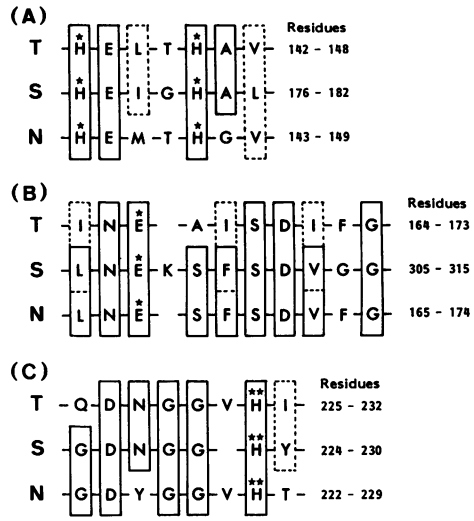

Fig. 5. Comparison of the regions containing zinc ligands and active sites of Serratia protease, thermolysin, and B. subtilis neutral protease. Serratia protease, thermolysin, and B. subtilis neutral protease are represented by S, T, and N, respectively. The zinc ligands and the active sites are indicated by single asterisks (*) and double asterisks (**), respectively.

similar to the active site region in thermolysin (residues 225-232) and to the region from residue 222 to 229 in B. subtilis neutral protease (Fig. 5C). The similarities of these regions suggest that the zinc ligands of Serratia protease are His 176, His 180, and Glu 307 and that the active site is His 229. However, the relative positions of the zinc ligand (Glu) and the active site (His) in Serratia protease is different from those in thermolysin and B. subtilis neutral protease, indicating that the similarities in the zinc binding and active site regions between Serratia protease and the other two enzymes might be derived by a convergent evolution. The three zinc ligands of thermolysin have been shown to occur in helical regions by X-ray analysis (11). The three zinc ligands of the Serratia protease also occur in helical regions, as predicted by the method of Chou and Fasman (26). X-ray analysis of the Serratia protease is now in progress (4), and the three-dimensional structure of the enzyme will be determined in near future.

The nucleotide sequence of the Serratia protease gene shows that the mature protein is preceded by a peptide consisting of 33 or 16 amino acids. The expression of the Serratia protease gene in E. coli suggests that the peptide consists of 16 amino acids. However, the possibility still remains that the peptide consists of 33 amino acids and is processed to the shorter peptide whose molecular weight is 1 kd. The peptide is different from the signal peptides of secretory proteins because the ratio of hydrophobic

Table II. Codon usage of the mature protein

Phe	TTT	8	Ser	TCT	3	Tyr	TAT	7	Cys	TGT	0
	TTC	22		TCC	9		TAC	12		TGC	0
Leu	TTA	1		TCA	2	Term	TAA	0	Term	TGA	0
	TTG	2		TCG	7		TAG	0	Trp	TGG	8
	CTT	1	Pro	CCT	1	His	CAT	9	Arg	CGT	4
	CTC	1		CCC	1		CAC	3		CGC	5
	CTA	0		CCA	0	Gln	CAA	5		CGA	1
	CTG	23		CCG	8		CAG	21		CGG	2
Ile	ATT	10	Thr	ACT	2	Asn	AAT	4	Ser	AGT	0
	ATC	13		ACC	28		AAC	32		AGC	14
	ATA	0		ACA	1	Lys	AAA	8	Arg	AGA	0
Met	ATG	1		ACG	1		AAG	9		AGG	0
Val	GTT	0	Ala	GCT	6	Asp	GAT	16	Gly	GGT	19
	GTC	10		GCC	24		GAC	21		GGC	29
	GTA	3		GCA	1	Glu	GAA	8		GGA	1
	GTG	13		GCG	16		GAG	7		GGG	7

amino acids in the peptide is much lower than those in the signal peptides, even if the peptide consists of 33 amino acids. The role of the pro-peptide is unclear. Alkaline and neutral proteases of B. amyloliquefaciens, B. subtilis, and Bacillus licheniformis have pro-peptides between signal peptides and mature proteins (12,13,27-30). The pro-peptides are thought to regulate the protease activity of the enzymes and to be removed during secretion. The pro-peptide of Serratia protease may also regulate protease activity, although it is much smaller than those of the Bacillus proteases described above. The cleavage seems to occur between an alanine and the alanine that is the NH_2-terminal residue of the mature protein. The nucleotide sequence of the Serratia protease gene indicates that the enzyme seems to have no signal peptide. E. coli hemolysin is released extracellularly without cleavage of a signal peptide, although the mechanism of secretion remains unclear (31,32). The Serratia protease may be released into the culture medium by a mechanism similar to that of hemolysin. S. marcescens harboring pTSP26 produced the Serratia protease in the culture medium, however, the Serratia protease expressed in E. coli harboring pTSP26 was detected in the cells but not in the culture medium. Four cistrons are necessary for the E. coli hemolysin phenotype (32). It seems likely that the Serratia protease also requires several cistrons for the secretion process, and that pTSP26 lacks some parts of the cistrons. The open reading frame downstream from the mature protein coding region may be one of the cistrons.

The G+C content of chromosomal DNA from S. marcescens (59 %) is higher than that of E. coli chromosomal DNA (51 %) (33). The G+C content of the sequenced 2,570 base fragment is 54 %. Within the fragment, the coding region for the mature protein of the Serratia protease (58 %) is much more G+C-rich than the 5' noncoding region (42 %). Table II shows the codon usage of the coding region for the mature protein. The G+C contents at the first, second, and third positions of the codons are 57, 42, and 74 %, respectively. It is clear that the high G+C content of the coding region is due to a

marked preference for G+C bases at the third position of the codons. This pattern has been found in the glutamine amidotransferase gene of S. marcescens (8).

The Serratia protease has been produced in industrial scale and widely used in medical purposes. We have obtained the gene by molecular cloning. The gene will be useful for improving enzyme production and modifying the enzyme by using recombinant DNA techniques in the future.

ACKNOWLEDGEMENTS

We thank Drs. Y. Sugino, A. Kakinuma, and M. Kida for encouragement, Drs. K. Tomoda and K. Miyata for helpful discussions. We also thank Mrs. K. Ishimaru for technical assistance.

REFERENCES

1. Miyata, K., Maejima, K., Tomoda, K. and Isono, M. (1970) Agric. Biol. Chem. 34, 310-318.
2. Miyata, K., Tomoda, K. and Isono, M. (1970) Agric. Biol. Chem. 34, 1457-1462.
3. Miyata, K., Tomoda, K. and Isono, M. (1971) Agric. Biol. Chem. 35, 460-467.
4. Katsuya, Y., Hamada, K., Hata, Y., Tanaka, N., Sato, M., Katsube, Y., Kakiuchi, K. and Miyata, K. (1985) J. Biochem. 98, 1139-1142.
5. Lee, I. S., Wakabayashi, S., Miyata, K., Tomoda, K., Yoneda, M., Kangawa, K., Minamino, N., Matsuo, H. and Matsubara, H. (1984) J. Biochem. 96, 1409-1418.
6. Lee, I. S., Wakabayashi, S., Matsubara, H., Miyata, K. and Tomoda, K. (1985) Fed. Proc. 44, 1057.
7. Schmitz, G. and Braun, V. (1985) J. Bacteriol. 161, 1002-1009.
8. Kaplan, J. B., Merkel W. K. and Nichols, B. P. (1985) J. Mol. Biol. 183, 327-340.
9. Jones, J. D. G., Grady, K. L., Suslow, T. V. and Bedbrook, J. R. (1986) EMBO J. 5, 467-473.
10. Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A. and Neurath, H. (1972) Nature New Biol. 238, 35-37.
11. Matthews, B. W., Jansonius, J. N., Colman, P. M., Schoenborn, B. P. and Dupourque, D. (1972) Nature New Biol. 238, 37-41.
12. Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J. and Filpula, D. (1984) J. Bacteriol. 159, 811-819.
13. Yang, M. Y., Ferrari, E. and Henner, D. J. (1984) J. Bacteriol. 160, 15-21.
14. Beaucagl, S. L. and Carutheri, M. H. (1981) Tetrahedron Lett. 22, 1859-1862.
15. Messing, J., Crea, R. and Seeburg, P. H. (1981) Nucleic Acids Res. 9, 309-321.
16. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, New York.
17. Lennox, E. S. (1955) Virology 1, 190-206.
18. Lovett, P. S. and Keggins, K. M. (1979) Meth. Enzymol. 68, 342-357.
19. Birnboim, H. C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523.
20. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
21. Laemmli, U. K. (1970) Nature 227, 680-685.
22. Burnette, W. N. (1981) Anal. Biochem. 112, 195-203.
23. Pollock, M. R. and Richmond, M. H. (1962) Nature 194, 446-449.
24. Weaver, L. H., Kester, W. R. and Matthews, B. W. (1977) J. Mol. Biol. 114, 119-132.

25. Levy, P. L., Pangburn, M. K., Burstein, Y., Ericsson, L. H., Neurath, H. and Walsh, K. A. (1975) Proc. Natl. Acad. Sci. USA 72, 4341-4345.
26. Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222-245.
27. Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A. and Chen, E. Y. (1983) Nucleic Acids Res. 11, 7911-7925.
28. Shimada, H., Honjo, M., Mita, I., Nakayama, A., Akaoka, A., Manabe, K. and Furutani, Y. (1985) J. Biotechnol. 2, 75-85.
29. Yoshimura, K., Miyazaki, T., Nakahama, K. and Kikuchi, M. (1985) J. Takeda Res. Lab. 44, 42-50.
30. Jacobs, M., Eliasson, M., Uhlen, M. and Flock, J.-I. (1985) Nucleic Acids Res. 13, 8913-8926.
31. Felmlee, T., Pellett, S., Lee, E.-Y. and Welch, R. A. (1985) J. Bacteriol. 163, 88-93.
32. Felmlee, T., Pellett, S. and Welch, R. A. (1985) J. Bacteriol. 163, 94-105.
33. Sober, H. A. (1970) Editor of Handbook of Biochemistry: Selected Data for Molecular Biology, 2nd edn. pp. H83-H85, CRC Press, West Palm Beach, Florida.