KOR. JOUR. MICROBIOL December, 1992 p. 507~513 Vol. 30, No. 6

Cloning, Sequencing and Expression of an Extracellular Protease Gene from *Serratia marcescens* RH1 in *Escherichia coli*

Lee, Seung-Hwan, Jeong-Min Kim, Young-Tae Kwon, Young-Hee Kho† and Hyune-Mo Rho*

Department of Molecular Biology, Seoul National University, Seoul 151-742, Korea
†Genetic Engineering Research Institute, KIST, Taejon 305-606, Korea

Serratia marcescens RH1 isolated from soil samples produced large amount of extracellular proteases. One of the genes encoding an extracellular protease from *S. marcescens* RH1 was cloned in *Escherichia coli* by shot gun cloning method. The cloned protease, SSP, was stably expressed by its own promoter and excreted into the extracellular medium from *E. coli* host cells. The nucleotide sequence of the protease gene (*ssp*) showed a single open reading frame (ORF) of 3,135 nucleotides corresponding to 1,045 amino acids (112 kDa). The nucleotide and deduced amino acid sequence of SSP showed high overall homology (88%) to one of the *S. marcescens* protease (27), but low homology to other serine protease families. The optimal pH and temperature of the enzyme were pH 9.0 and 45°C, respectively. The activity of the protease was inhibited by phenylmethylsulfonyl fluoride (PMSF), which suggests that the enzyme is a serine protease.

KEY WORDS □ *Serratia marcescens* RH1, extracellular serine protease, cloning, DNA sequencing, promoter

Microbial proteases have been widely used in industrial fields, such as food, leather, detergent and pharmaceuticals. These enzymes are produced from various microorganisms, including genus *Bacillus* (25, 26, 28), *Serratia* (23, 27), *Pseudomonas* (11, 17), and *Vibrio* (6). Some of these proteases are also well recognized at least suspected as pathogenic factors in many diseases such as inflammatory diseases (15, 16, 19, 21).

Serratia marcescens, a Gram-negative bacterium which belongs to the family *Enterobacteriaceae*, secretes a number of extracellular proteins, a nuclease, chitinases, lipases and proteases (2, 3, 8, 9, 14). Of these, extracellular protease was shown to be related to pathogenesis of *S. marcescens* in experimental pneumonia in guinea pigs and mice (19) and to insect pathogenicity (10). Analysis of a variety of different *Serratia* strains revealed that four or more different proteases are produced from these strains (7, 12, 21). The major extracellular protease gene which is widely used as anti-inflammatory agent has been cloned from *S. marcescens* E-15 (23). The major protease was a metalloprotease with a molecular weight of 51 kDa, and its mature enzyme followed a presequence of 16 amino acids. One of the minor protease gene was also cloned and sequenced (27). The minor protease was a serine protease with a molecular weight of 41 kDa, and its precursor protein was composed of a typical N-terminal signal sequence, a mature enzyme domain and a large C-terminal domain.

In this report, the extracellular protease-producing bacteria was isolated from soil samples and identified *S. marcescens*, named as *S. marcescens* RH1. One of the genes encoding an extracellular protease from *S. marcescens* RH1 was cloned and sequenced in *E. coli*. The cloned protease (SSP) was stably expressed by its own promoter and excreted into the culture medium from *E. coli* host cells. The biochemical properties of the enzyme were also investigated.

*To whom correspondence should be addressed

MATERIALS AND METHODS

Bacterial strains and growth media

E. coli HB101 (F^-, r^-, m^-, recAB, lacY1) was used for transformation and amplification of plasmid DNA. JM83 (ara^-, Δ(lac-proAB), rpsL, thi, φ80, lacZΔDM15)) was used for transfor-

mation with plasmid pUC19 and selection of colonies containing recombinant plasmid DNA. JM101 (supE, thi, Δ(lac-proAB), F$^-$(lacI^q, lacZΔDM15)) was used for the selection of colonies containing recombinant DNA on X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), IPTG (isopropylthio-β-D-galactoside) plates and transfection with M13 phage. *E. coli* transformants were grown aerobically in Luria broth (LB) containing 100 μg/ml ampicilin at 37°C.

Screening and identification of protease producing strains

The collected samples from soil were overlayed on LB agar plate containing 1.0% skim milk by dilution with distilled water. The halo forming colonies were streaked on new media generation to generation. The isolates showing high proteolytic activity were subjected to identification using Gram-staining method, microscopic analysis and API 20E kit (Analytab Products Inc.).

Recombinant DNA techniques

Chromosomal DNA of *S. marcescens* RH1 was isolated according to the method of Mamur (20). Purified chromosomal DNA was digested with *Pst*I and ligated into the corresponding site of pUC19, followed by transformation into *E. coli* JM83 (4). The protease-positive clone was identified by a clear zone around the colony on LB agar containing 1.0% skim milk. The recombinant plasmid pRH1 containing the 8.0 kb DNA fragment encoding the extracellular protease was subcloned by a serial deletion with *Bal*31 exonuclease and restriction enzymes.

The protease activity was measured by the method of Yanagida *et al.* (27). The 2.5 ml of prewarmed 1% casein in 50 mM phosphate buffer (pH 8.0) and 0.5 ml of enzyme solution were mixed and incubated at 37°C for 10 min. The reaction was stopped by adding stopping mixture containing 0.11 M trichloroacetic acid, 0.22 M acetic acid and 0.33 M sodium acetate, and centrifuged. The supernatant was subject to optical density determination at A_{280}. One unit of the protease activity was defined as an amount of enzyme which catalyzed an increase in absorbance of 0.1 per 10 min by the result of tyrosine liberation.

RESULTS

Screening and identification of *S. marcescens* RH1

The isolated microorganism was a motile, rod-shaped, Gram-negative bacterium by microscopic analysis and Gram-staining. The strain was identified as *S. marcescens* by API 20E kit, and named as *S. marcescens* RH1. (data not shown).

Cloning and subcloning of the *ssp* gene from *S. marcescens* RH1

The *Pst*I fragments of chromosomal DNA of *S. marcescens* RH1 were inserted into the pUC19 and transformed into *E. coli* JM83. Among 15,000 transformants, one colony showed a clear zone on the LB plate containing 1.0% skim milk. The clear zone of the transformant suggested the extracellular production of the protease cloned in *E. coli* JM83. The recombinant plasmid isolated from the clone (1), designated pRH1, contained a 8.0 kb DNA fragment and was analyzed by the restriction enzymes (Fig. 1). Southern blot analysis of the chromosomal DNA of *S. marcescens* RH1 using ^{32}P-labeled 1.5 kb *Bam*H1-*Pst*I fragment of pRH1 as a probe showed that the cloned fragment was derived from the parental strain (data not shown). The initial recombinant plasmid pRH1 was subcloned to a minimal length for the expression of SSP by *Bal*31 nuclease and restriction enzymes (Fig. 1). The plasmid pRH2 oppositely oriented to *LacZ* promoter, but the plasmid pRH1 did not. The two plasmid produced nearly the same amount of SSP, suggesting that *ssp* was expressed by its own promoter. The plsmid pRH4 which had opposite orientation to pRH5 showed 2.5 times higher proteolytic activity than pRH5 unlike the case of pRH1 and pRH2. This result shows that *LacZ* promoter also function for the expression of SSP. The production of SSP by the plasmid pRH5 was much lower than pRH2, suggesting that some of the regulatory element of the promoter of pRH5 may be deleted during subcloning.

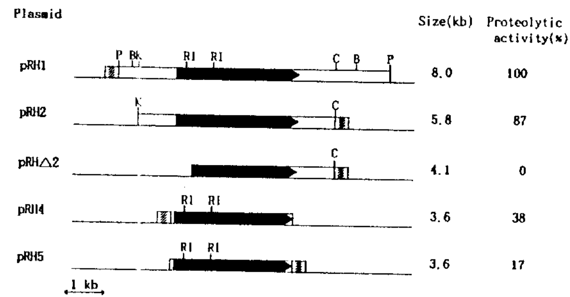

Fig. 1. *Subcloning of the DNA encoding the ssp gene.* The plsmid pRH2 was shortened by *Kpn*I and *Cla*I digestion to make pRH2. The plasmid pRHΔ2 was obtained by *Bal*31 nuclease treament following *Kpn*I digestion of pRH2. The plasmid pRH4 and pRH5 was obtained by *Bal*31 nuclease of *Bgl*II fragment of pRH1. The hatched region represents *LacZ* promoter. The blackened and open region represent the coding and its flanking sequence of the *ssp* gene, respectively. Abbreviation: B, *Bgl*II; C, *Cla*I; K, *Kpn*I; P, *Pst*I; RI, *Eco*RI.

Nucleotide sequence of the *ssp* gene

The nucleotide sequence of the *ssp* gene contained a single open reading frame of 3,135 base pairs directing a polypeptide composed of

```
                         -189 TGC CTG CAG GTC GAC TCT AGA GGA TCC CCG ACA TCT TCA TGT -148

 -147 CCG CAC GCC CGA ATT TTT ATC ACT CAG CCT GTT TTA TGC GGC CCG CGT CGC GAG GCA TGC GTT  -85

  -84 ATC TGT CGT TTC ACG ATC ATA GGA GGT TTC AGC GCG GTT AAA ACG TCA GTT TTC TTC TGG GTA  -22

  -21 TAA ATA AAC ATG GAG CGT AAA ATG ATA CTT AAT AAA AAA TTG AAG TTA GCG TAT TGC GTG TTT   42
                   SD             MET Ile Leu Asn Lys Lys Leu Lys Leu Ala Tyr Cys Val Phe

   43 CTG GGT TGT TAT GGC TTG TCT CTT CAT TCA TCT CTT GCC GCT TAT CGG GAA CCA GGA CAA TTA  105
      Leu Gly Cys Tyr Gly Leu Ser Leu His Ser Ser Leu Ala Ala Tyr Arg Glu Pro Gly Gln Leu

  106 GGT TCG CCC GAC AGT TGG AAA AAC GCA GAG TTT AAT CGC CAA TGG GGG CTT GAG GCT ATT TCC  168
      Gly Ser Pro Asp Ser Trp Lys Asn Ala Glu Phe Asn Arg Gln Trp Gly Leu Glu Ala Ile Ser

  169 GCC GAG TTC GCC TAT GCC AGA GCG TAT ACC GGA AAG GGC GTG ACG ATT GGC GTT ATT GAT GAT  231
      Ala Glu Phe Ala Tyr Ala Arg Ala Tyr Thr Gly Lys Gly Val Thr Ile Gly Val Ile Asp Asp

  232 GCC ATT CTC TCT CAT CCG GAA TTC GCC GGT AAA CTG ACG CGG CTG GAT AAC GGC AGC TAT AAT  294
      Ala Ile Leu Ser His Pro Glu Phe Ala Gly Lys Leu Thr Arg Leu Asp Asn Gly Ser Tyr Asn

  295 TTC TCG TAT GAT AAA CAA GAT AAT ATG TCA TTC GGC ACG CAT GGC ACG CAT GTC GCC GGT ATT  357
      Phe Ser Tyr Asp Lys Gln Asp Asn MET Ser Phe Gly Thr His Gly Thr His Val Ala Gly Ile

  358 GCA GCC GCC AAA AGA GAC GGC TCC GGT ATG CAT GGC GTC GCT TAC GAC GCG GAT ATT ATC GGC  420
      Ala Ala Ala Lys Arg Asp Gly Ser Gly MET His Gly Val Ala Tyr Asp Ala Asp Ile Ile Gly

  421 ACC AAA TTG AAT GAT TAC GGT AAT CGC AAC GGC CGT GAA GAG CTG ATT CAG AGC GCG GCT CGC  483
      Thr Lys Leu Asn Asp Tyr Gly Asn Arg Asn Gly Arg Glu Glu Leu Ile Gln Ser Ala Ala Arg

  484 GTC ATC AAT AAC AGC TGG GGG ATC CGC CCC GAT ATT CGG CGA GAC GCC AAA GGC GAT ATT ATC  546
      Val Ile Asn Asn Ser Trp Gly Ile Arg Pro Asp Ile Arg Arg Asp Ala Lys Gly Asp Ile Ile

  547 TGG TTG CCG AAC GGC AGG CCG GAC TAC GTG GCG TGG GTG AAA ACC GAC GTC ATC AAT GAA GTG  609
      Trp Leu Pro Asn Gly Arg Pro Asp Tyr Val Ala Trp Val Lys Thr Asp Val Ile Asn Glu Val

  610 ATG CGC AAT AAG TCC AAC CTG GAA TGG GGC AGC GAA CAA CCC GTG CCC ACC GGC GGG CAC AGC  672
      MET Arg Asn Lys Ser Asn Leu Glu Trp Gly Ser Glu Gln Pro Val Pro Thr Gly Gly His Ser

  673 GCC ATG GCG ACG CTG CTG CGG GCC GCC AAG CAT GGC AAG CTG ATC GTC TTC TCG GCA GGC AAT  735
      Ala MET Ala Thr Leu Leu Arg Ala Ala Lys His Gly Lys Leu Ile Val Phe Ser Ala Gly Asn

  736 TAC AAT AAC TAC AAT ATT CCT GAG GCG CAA AAG TCA CTT CCC TAT GCT TTC CCG GAG GTA TTA  798
      Tyr Asn Asn Tyr Asn Ile Pro Glu Ala Gln Lys Ser Leu Pro Tyr Ala Phe Pro Glu Val Leu

  799 AAC AAT TAT CTG ATC GTG ACC AAT CTG AGT AAT AAC GAT AAG TTA AGC GTT TCC TCA ACC AGC  861
      Asn Asn Tyr Leu Ile Val Thr Asn Leu Ser Asn Asn Asp Lys Leu Ser Val Ser Ser Thr Ser

  862 TGC GGA CAT ACG GCC AGT TTT TTA GCG TGT CAG CCG GGC AGC AGT ATT TAC AGC AGC GTA GGG  924
      Cys Gly His Thr Ala Ser Phe Leu Ala Cys Gln Pro Gly Ser Ser Ile Tyr Ser Ser Val Gly

  925 GAG CTG GTT TCC AAT ACC GGC GGG GCG GTG AAT CGC GAA GCG TAT AAT AAG GGC GAG CTG ACG  987
      Glu Leu Val Ser Asn Thr Gly Gly Ala Val Asn Arg Glu Ala Tyr Asn Lys Gly Glu Leu Thr

  988 GTT AAA CCC GAT TAC GGA AAT ATG TCG GGC ACC TCA ATG GCG CCG GAC GTC ACC GGA TTT GCC 1050
      Val Lys Pro Asp Tyr Gly Asn MET Ser Gly Thr Ser MET Ala Pro Asp Val Thr Gly Phe Ala

 1051 GCC GTA TTG ATG CAG CGC TTC CCT TAT ATG AGC GCC GCT CAA ATA TCC GCG GTC ATC AAA ACC 1113
      Ala Val Leu MET Gln Arg Phe Pro Tyr MET Ser Ala Ala Gln Ile Ser Ala Val Ile Lys Thr

 1114 ACC GCC ACC GAT TTA GGC GAG GTC GGC ATT GAT CAT TTA TTC GGC TGG GGC CGC GTC AAT TTA 1176
      Thr Ala Thr Asp Leu Gly Glu Val Gly Ile Asp His Leu Phe Gly Trp Gly Arg Val Asn Leu

 1177 CGC GAT GCG ATC AAC GGG CCG AAG ATG TTT ATC ACC CAA GAG GAT ATT CCG CAG GAA TTC TAT 1239
      Arg Asp Ala Ile Asn Gly Pro Lys MET Phe Ile Thr Gln Glu Asp Ile Pro Gln Glu Phe Tyr

 1240 GTG CCG GGT TCC TAC AGT GAA AAA CAG TTT GTG GTG AAT ATC CCC GGT CTT GGG AAT ATC GTC 1302
      Val Pro Gly Ser Tyr Ser Glu Lys Gln Phe Val Val Asn Ile Pro Gly Leu Gly Asn Ile Val

 1303 GAA GCC GGA ACC CCG GTT GAG CGG CGT TGC ACG TCG GGC GAA TGC GAC TTC GAT TCG TGG AGC 1365
      Glu Ala Gly Thr Pro Val Glu Arg Arg Cys Thr Ser Gly Glu Cys Asp Phe Asp Ser Trp Ser

 1366 AAT GAC ATC CGC GGC CAC GGC GGG TTG ACC AAA ACG GGC GCC GGT ACG CTG GCC GTG TTG GGC 1428
      Asn Asp Ile Arg Gly His Gly Gly Leu Thr Lys Thr Gly Ala Gly Thr Leu Ala Val Leu Gly

 1429 AAT AAT ACT TAC AGC GGT GAC ACC TGG GTG AAA CAG GGC GTA CTG GCG TAC AAC GGT TCG GTG 1491
      Asn Asn Thr Tyr Ser Gly Asp Thr Trp Val, Lys Gln Gly Val Leu Ala Tyr Asn Gly Ser Val

 1492 GCG TCC AAT GTC TAT ATC GAA AAT AGC GGC ACC GTG GCC GGC GAC CGG ACG GTG GGG GCG TTC 1554
      Ala Ser Asn Val Tyr Ile Glu Asn Ser Gly Thr Val Ala Gly Asp Arg Thr Val Gly Ala Phe

 1555 AGA GCG GTT CGG GGC TGT GAA CAC GGT GAC GCG GGA AAT GGC TAC GGT ACG CTG CAC GTG TTG 1617
      Arg Ala Val Arg Gly Cys Glu His Gly Asp Ala Gly Asn Gly Tyr Gly Thr Leu His Val Leu

 1618 CTA GAC GCC GTG TTT GAT CGC GGC TCG CAG TAT AAC GTG GAA TTG GCC GAC AAA GGC CGC AGC 1680
      Leu Asp Ala Val Phe Asp Arg Gly Ser Gln Tyr Asn Val Glu Leu Ala Asp Lys Gly Arg Ser

 1681 GAT AAG TTG GCC GCG CGG CGC GCT TTT CTC AAC GGC GGC AGT ATG AAT GTC AGC CTG GAT CGC 1743
      Asp Lys Leu Ala Ala Arg Arg Ala Phe Leu Asn Gly Gly Ser MET Asn Val Ser Leu Asp Arg

 1744 AGC CAA AAA CTG ATG TCG CAG AAT GAG GCG GAA CTG CTG GTG GGC AAC AAC TAT ACC ATC CTG 1806
      Ser Gln Lys Leu MET Ser Gln Asn Glu Ala Glu Leu Leu Val Gly Asn Asn Tyr Thr Ile Leu
```

```
1807 ACC ACG CTG GAC GGC GTT ACC GGT AGA TTC GAT AAC GCT AAC CCC TCG TAT CCG TTT GTT AAA 1869
     Thr Thr Leu Asp Gly Val Thr Gly Arg Phe Asp Asn Ala Asn Pro Ser Tyr Pro Phe Val Lys
1870 GTG GCG CTG GAT TAT CGG GGC AAT GAC ACC GGG CTG GGC ATC ACT AAA ACC GAC GCC ACG TTT 1932
     Val Ala Leu Asp Tyr Arg Gly Asn Asp Thr Gly Leu Gly Ile Thr Lys Thr Asp Ala Thr Phe
1933 GAC AGC CTG GCC AGC ACT GAG AAC GAC AAA GAG GTA GCC CGC GCG GTG GAA ACG CTG AAC GCG 1995
     Asp Ser Leu Ala Ser Thr Glu Asn Asp Lys Glu Val Ala Arg Ala Val Glu Thr Leu Asn Ala
1996 ACG GAG CCG GTC ACG GAA ACG GCC AAA CGC AGC GTG AGT ATC CCC GCG AGC GAA GAG GCC AAT 2058
     Thr Glu Pro Val Thr Glu Thr Ala Lys Arg Ser Val Ser Ile Pro Ala Ser Glu Glu Ala Asn
2059 CTG CAG CAA AGC GAT GCG GGC GCC GCG CAA GCC GTG AAC GAA GAG GCG AGC ATC GTG GCG GGG 2121
     Leu Gln Gln Ser Asp Ala Gly Ala Ala Gln Ala Val Asn Glu Glu Ala Ser Ile Val Ala Gly
2122 CAT CCC ATC TAC GAA AGC TTT CTG GGC TTC ACC ACC GCC AGA GAA TTG CAA CAG GCG ACC CGT 2184
     His Pro Ile Tyr Glu Ser Phe Leu Gly Phe Thr Thr Ala Arg Glu Leu Gln Gln Ala Thr Arg
2185 CAA CTG TCC GGC CAG ATC CAC GCG GAT ATG GCG TCC GCC CAG ATT AAC GAA AGC CGT TAT CTG 2247
     Gln Leu Ser Gly Gln Ile His Ala Asp MET Ala Ser Ala Gln Ile Asn Glu Ser Arg Tyr Leu
2248 CGC GAT ACC GCC ACC GAG CGG TTG CGC CAG GCC GAT GGC CGC CGC ACC GCT TCC GAT ATC AAA 2310
     Arg Asp Thr Ala Thr Glu Arg Leu Arg Gln Ala Asp Gly Arg Arg Thr Ala Ser Asp Ile Lys
2311 GCG GAT GAT AAT GGC GCC TGG GCG AAA TTG CTG GGC AAC TGG GGG CAT GCT TCC GGC AAC GAC 2373
     Ala Asp Asp Asn Gly Ala Trp Ala Lys Leu Leu Gly Asn Trp Gly His Ala Ser Gly Asn Asp
2374 AAC GCT ACC GGT TAC CAG ACA TCC ACC TAT GGC GTG CTG TTG GGT CTG GAC AGC GAA CTG TTT 2436
     Asn Ala Thr Gly Tyr Gln Thr Ser Thr Tyr Gly Val Leu Leu Gly Leu Asp Ser Glu Leu Phe
2437 GAC GAC GGC CGG CTG GGC GTG ATG ACC GGG TAT ACC CGC ACG TCG CTG GTA GGC GGT CTA CAG 2499
     Asp Asp Gly Arg Leu Gly Val MET Thr Gly Tyr Thr Arg Thr Ser Leu Val Gly Gly Leu Gln
2500 TCA GTA GTC CAC AGC GAC ACT ACA CAT CTG GGG CTG TAC GGC GAC AAA CGC TTC GGC GCG TTG 2562
     Ser Val Val His Ser Asp Thr Thr His Leu Gly Leu Tyr Gly Asp Lys Arg Phe Gly Ala Leu
2563 GCG CTG CCA GCG GGC GGC ACC TAT ACC TGG CAT CGC ATC GAC ACG TCG CGC TCG GTA AAC TAC 2625
     Ala Leu Pro Ala Gly Gly Thr Tyr Thr Trp His Arg Ile Asp Thr Ser Arg Ser Val Asn Tyr
2626 GGC GCG CAG GCG GAT CGC GAA AAG GCC CGC TAT AAC GCG CGC ACC GGT CAG CTG TTT ATC GAA 2688
     Gly Ala Gln Ala Asp Arg Glu Lys Ala Arg Tyr Asn Ala Arg Thr Gly Gln Leu Phe Ile Glu
2689 AGC GGC TAC GAT TGG AGC AAC GAC GTG GTC AAT CTT GAG CCG TTC GCC AAC CTG GCG TAC ACC 2751
     Ser Gly Tyr Asp Trp Ser Asn Asp Val Val Asn Leu Glu Pro Phe Ala Asn Leu Ala Tyr Thr
2752 CAC TAT CGC AAC GAG GGG ATC AAC GAG CAA GGC GGG GCG GCG GCG CTG CGC GGC GAT AAG CAA 2814
     His Tyr Arg Asn Glu Gly Ile Asn Glu Gln Gly Gly Ala Ala Ala Leu Arg Gly Asp Lys Gln

2815 AGT CAG TCC GCC ACC GCT TCG ACG CTG GGC CTG CGC GCC GAT ACG CAA TGG CAG ACC GAC AGC 2877
     Ser Gln Ser Ala Thr Ala Ser Thr Leu Gly Leu Arg Ala Asp Thr Gln Trp Gln Thr Asp Ser
2878 GTG GCG ATC GCC CTG CCG GGC GAG CTG GGT TGG CAA CAT CAG TAC GGC AAG CTG GAG CGT AAA 2940
     Val Ala Ile Ala Leu Pro Gly Glu Leu Gly Trp Gln His Gln Tyr Gly Lys Leu Glu Arg Lys

2941 ACA CAG CTG ATG TTC AAA CGC AGC GAT GTC GCG TTC GAC GTG AAC AGC GTC CCT GTT TCT CGC 3003
     Thr Gln Leu MET Phe Lys Arg Ser Asp Val Ala Phe Asp Val Asn Ser Val Pro Val Ser Arg
3004 GAT GGG GCC ATT CTG AAA GCG GGC GTC GAT GTA TCG ATT AAC AAA AAC GTC GTC CTG TCC CTT 3066
     Asp Gly Ala Ile Leu Lys Ala Gly Val Asp Val Ser Ile Asn Lys Asn Val Val Leu Ser Leu
3067 GGG TAC GGC GGG CAG CTG TCG TCC AAC CAC CAG GAC AAC AGC GTC AAC GCC GGC CTG ACC TGG 3129
     Gly Tyr Gly Gly Gln Leu Ser Ser Asn His Gln Asp Asn Ser Val Asn Ala Gly Leu Thr Trp
3130 CGG TTC TGA TAG CCG TTG ATC CGT TTC ACG CTA CTC AGC CCC CAT CGC AAC GAT GGG GTT TTT 3192
     Arg Phe TER                                     ---------->    <----------
3193 TAT TGG GCC GCG GAA GAC AGG CCG GGC GGC GGT GTC TTC ACA GGA TAA AGG CGC GTT AGA ACA 3255
3256 GGA TGT TTC TAA TTC TTT TGG TTA GAA AAA TTG ATG ATA ATC ATC TGA AGT GAC TAA AAG GAT 3318
3319 TTA GGC CGA AAC ATG GAA TTT CTC TTC ATT TTA TTA TTT TTT ACC TTA TAC ATA TCA ATT GCC 3381
3382 TGT GAG TAT CAC CG
```

Fig. 2. *The nucleotide and deduced amino acid sequences of ssp.*
The nucleotide sequence is numbered from the first base of the translation start codon. The deduced amino acid sequence (three letter codon) is shown below the nucleotide sequence of the coding region. The active site of SSP are boxed. The SD sequence is underlined. The putative transcrition terminator is indicated by arrows. EMBL accession Number; The accession number for the sequence reported in this paper is X59719.

1,045 amino acid residues with a calculated molecular mass of 112,000 daltons (Fig. 2). A putative ribosome binding site (RBS), GGAG, was found 6 bp upstream of the translation initiation codon. A 19-bp palindromic sequence was located 31 to 51 bases downstream from a termination codon (TGA). This sequence is thought to be a rho-independent transcription terminator. The G

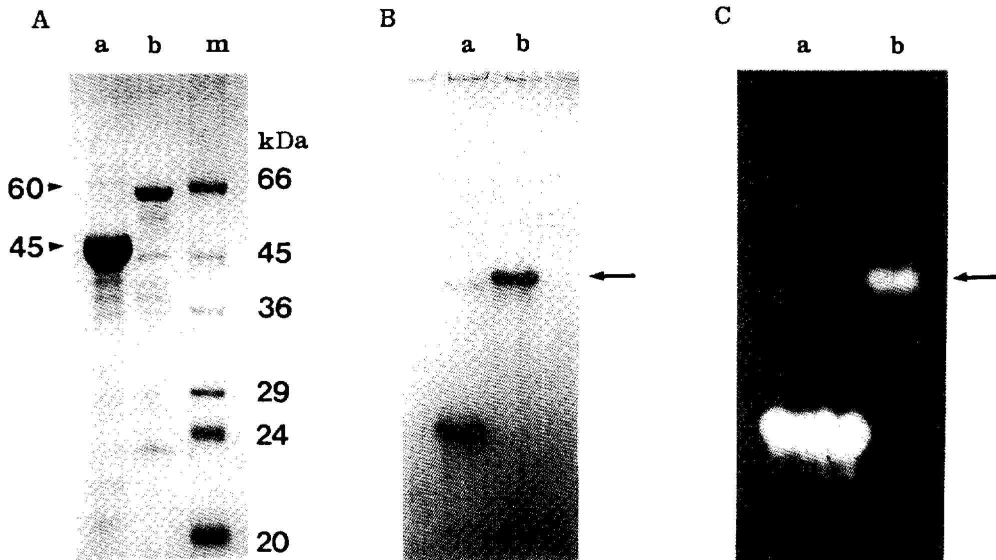

Fig. 3. *Characterization of SSP from E. coli [pRH1].*
(A) SDS-PAGE of the concentrates of the culture media. a; *Serratia marcescens* RH1, b; *E. coli* ···pRH1], m; size marker. (B) The Coomassie blue staining after native-PAGE of the same sample. a; *Serratia marcescens* RH1. b; *E. coli* [pRH1]. (C) The activity staining after native-PAGE of the same sample. a; *S. marcescens* RH1, b; *E. coli* [pRH1]. SSP is indicated by arrow.

+C content of the coding region of the *ssp* gene was 57.3%, which was similar to that of the chromosomal DNA of *S. marcescens* (56 to 59%). The G+C content of the 5′ noncoding region (46.7%) was lower than that of the coding region.

Deduced amino acid sequence of SSP

The deduced amino acid sequence of SSP was compared with other proteases using EMBL databases. The predicted amino acid sequence of SSP showed 88% homology with the serine protease from *S. marcescens* IFO-3046 (27), but relatively low homology (15% to 20%) to other serine proteases. But SSP showed high homology to these proteases over the regions comprising the active sites. Three essential residues for catalytic function of SSP, Asp^{49}, His^{85} and Ser^{314}, could be assigned according to the homology with other proteases. Comparison of the nucleotide and amino acid sequence of SSP showed a typical N-terminal signal sequence (nucleotide 1 to 81), and a mature domain (n.t. 82 to 1221) followed by pro-region (n.t. 1222 to 3135).

Excretion of the cloned *Serratia* protease from *E. coli* JM83

When *E. coli* JM83 harboring pRH1 was cultured in LB aerobically at 37°C, the cells produced the protease into the medium in parallel with their growth. Almost all the protease activity was found in the extracellular fraction of the cells at the early-stationary phase, while most of the β-lactamase (27) and β-galactosidase (27) activities were localized in the periplasm and the cytoplasm, respectively (data not shown).

Characterization of the cloned SSP

The culture medium of *E. coli* [pRH1] was concentrated with ammonium sulfate, and applied to SDS-PAGE (Fig. 3A). The major band of about 60 kDa was detected, but it was different from the band of 45 kDa which is thought to be the major metalloprotease of *S. marcescens* (23). To confirm that the major band is truely a protease, the activity staining using nondissociating PAGE was performed. As shown in Fig. 3, the major band in *E. coli* sample truely had a proteolytic activity, and was distinctive from the major metalloprotease from *S. marcescens.* To characterize SSP more detail, the concentrate of SSP was purified using DE-52 column followed by CM-52 column (data not shown). The purified SSP was completely inhibited by 2 mM PMSF and DFP (di-isopropylfluorophosphate), but not by 5 mM EDTA, suggesting SSP is a serine protease. The optimal pH and temperature of the enzyme were pH 9.0 and 45°C, respectively, suggesting SSP is a slightly alkaline protease (Fig. 4).

DISCUSSION

We have cloned one of the serine protease genes from *S. marcescens* RH1 isolated from soil samples. The ORF of the *ssp* gene was composed of 3,135 bp encoding 1,045 amino acids (112 kDa).

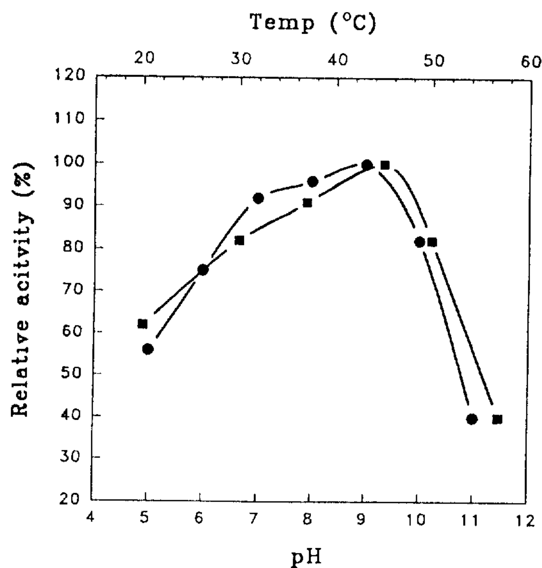

Fig. 4. *Optimum pH (●—●) and optimum temperature (■—■) of the protease activity.*
Effect of pH was examined in 50 mM CH_3COONa-CH_3COOH buffer (pH 5.0~pH 6.0), 50 mM Na_2HPO_4-NaH_2PO_4 (pH 7.0~pH 9.0) and 50 mM Na_2CO_3-$NaHCO_3$ (pH 10.0~pH 11.0) at 37°C. Effect of temperature was examined in 50 mM Na_2HPO_4-NaH_2PO_4 buffer (pH 7.5).

The deduced amino acid sequence of SSP showed relatively high homology (88%) to the other serine protease from *S. marcescens* (27). Comparison of the two proteases showed that the precursor of SSP was composed of a N-terminal signal sequence (3.5 kDa), mature domain (41.0 kDa) and a large C-terminal pro-peptide (67.0 kDa). This structure of SSP was interesting in that such a bulky C-terminal pro-peptide as well as an N-terminal signal peptide is cleaved off during the excretion process (22). The predicted molecular weight (45 kDa) of SSP was quite different from that (60 kDa) on SDS-PAGE (Fig. 3A). This is thought to be due to the tendency of retardation of the proteases on SDS-PAGE. This phenomenon was also observed by Yanagida *et al.* (27).

The *ssp* gene was expressed by its own promoter unlike the metalloprotease from *S. marcescens* E-15 (23) and the serine protease from *S. marcescens* IFO-3046 (27). It was expected that the promoter sequence of SSP be similar to the consensus sequence of *E. coli* promoters, but no such similarity was found at upstream of the *ssp* gene (13). The plasmid pRH5 derived from the initial plasmid pRH1 showed much lower activity than the plasmid pRH2 (Fig. 1).

Acknowledgements

This research was in part supported by grants from the ministry of Science and Technology and from the ministry of Health and Social Affairs.

REFERENCES

1. **Birnboim, H.C. and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.
2. **Braun, V. and G. Schmitz.** 1980. Excretion of a protease by *Serratia marcescens*. *Arch. Microbiol.* **124**, 55-61.
3. **Bromke, B.J. and J.M. Hammel.** 1979. Regulation of extracellular protease formation by *Serratia marcescens*. *Can. J. Microbiol.* **25**, 47-52.
4. **Cohen, S.N., S.L. Niemela, and R.H. Miller.** 1973. Nonchromosomal antibiotics resistance in bacteria: Genetic transformation of *E. coli* by R-factor DNA. *Proc. Nat'l. Acad. Sci. USA.* **69**, 2110-2114.
5. **Dale, R.S.M., B.A. McClure, and J.P. Houchins.** 1985. A rapid single stranded cloning strategy for production of a sequential series of overlapping clones for use in DNA sequencing. *Plasmid.* **13**, 31-40.
6. **Deane, S.M., F.T. Robb, S.M. Robb, and D.R. Woods.** 1989. Nucleotide sequence of the *Vibrio alginolyticus* calcium-dependent, detergent-resistant alkaline serine exoprotease A. *Gene.* **76**, 281-288.
7. **Decedue, C.J., E.A. Broussard, A. Larsan, and H. D. Braymer.** 1979. Purification and characterization of the extracellular proteinase of *Serratia marcescens*. *Biochemica et Biophysica Acta.* **569**, 293-301.
8. **Eaves, G. and C.D. Jefferies.** 1962. Isolation and characterization of an extracellular nuclease of *Serratia marcescens*. *J. Bacteriol.* **85**, 273-278.
9. **Gikskov, M., L. Olsen, and S. Molin.** 1988. Cloning and expression in *E. coli* of the gene for extracellular phospholipase A1 from *Serratia liquefaciens*. *J. Bacteriol.* **170**, 5855-5862.
10. **Flyg, C. and K.G. Xanthopoulos.** 1983. Insect pathogenic properties of *Serratia marcescens*. Passive and active resistance to insect immunity studied with protease-deficient and phage-resistant mutants. *J. Gen. Microbiol.* **129**, 453-464.
11. **Fukushima, J., S. Yamamoto, K. Morihara, Y. Atsumi, H. Takeuchi, S. Kawamoto, and K. Okuda.** 1989. Structural gene and complete amino acid sequence of *Pseudomonas aeruginosa* IFO 3455 elastase. *J. Bacteriol.* **171**, 1698-1704.
12. **Grimmont, P.A.D. and F. Grimmont.** 1978. The genus *Serratia*. *Ann. Rev. Microbiol.* **32**, 221-248.
13. **Hawley, D.K. and W.R. McClure.** 1983. Compilation and analysis of *E. coli* promoter DNA sequences. *Nucleic Acids Res.* **11**, 2237-2255.

14. **Heller, K.** 1979. Lypolytic activity copurified with the outer membrane of *Serratia marcescens*. *J. Bacteriol.* **140**, 1120-1122.
15. **Holder, I.A.** 1983. Experimental studies of the pathogenesis of infections due to *Pseudomonas aeruginosa* effect of treatment with protease inhibitors. *Rev. Infect. Dis.* **5**, 5914-5921.
16. **Kessler, E., A. Spierer, and S. Blumberg.** 1983. Specific inhibition of *Pseudomonas aeruginosa* elastase injected intracorneally in rabbit eyes. *Invest. Ophthalmol. Vis. Sci.* **24**, 1093-1097.
17. **Kessler, E. and M. Safrin.** 1988. Synthesis, processing, and transport of *Pseudomonas aeruginosa* elastase. *J. Bacteriol.* **170**, 5241-5247.
18. **Lammeli, U.K.** 1970. Clevage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London).* **227**, 680-685.
19. **Lyerly, D. and A.S. Kreger.** 1983. Importance of Serratia protease in the pathogenesis of experimental *Serratia marcescens pneumonia. Infec. and Immun.* **40**, 113-119.
20. **Mamur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganism. *J. Mol. Biol.* **3**, 208-218.
21. **Matsumoto, D., H. Maeda, K. Takada, R. Kamata, and R. Okamura.** 1984. Purification and characterization of four proteases from a clinical isolate of *Serratia marcescens kums* 3958. *J. Bacteriol.* **157**, 225-232.
22. **Miyazaki, H., N. Yanagida, S. Horinouchi, and T. Beppu.** 1989. Chracterization of the prcursor of *Serratia marcescens* serine protease and COOH-terminal processing of the precursor during its excretion through the outermembrane of *E. coli*. *J. Bacteriol.* **171**, 6555-6562.
23. **Nakahama, K., K. Yoshimura, R. Marumoto, K. Mosakazu, I.S. Lee, T. Hase, and H. Matsubara.** 1986. Cloning and sequencing of *Serratia* protease gene. *Nucleic Acids. Res.* **14**, 5843-5855.
24. **Sanger, F., S. Nicklen, and A.R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. *Proc. Nat'l. Acad. Sci. U.S.A.* **74**, 5463-5467.
25. **Takagi, M., T. Imanaka, and S. Aida.** 1985. Nucleotide sequence and promoter region for the neutral protease gene from *Bacillus stearothermophilus*. *J. Bacteriol.* **163**, 824-831.
26. **Vasantha, N., L.D. Thompson, C. Rhodes, C. Banner, J. Nagle, and D. Filpula.** 1984. Genes for alkaline protease and neutral proteases from *Bacillus amyloliquefaciens* contains a large open reading frame between the regions coding for signal sequence and mature protein. *J. Bacteriol.* **159**, 811-819.
27. **Yanagida. N., T. Uozumi, and T. Beppu.** 1986. Specific excretion of *Serratia marcescens* protease through the outermembrane of *E. coli*. *J. Bacteriol.* **166**, 937-944.
28. **Yang, M.Y., E. Ferrari, and D.J. Henner.** 1984. Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an *in vitro* derived deletion mutation. *J. Bacteriol.* **160**, 15-21.

(Received October 26, 1992)
(Accepted November 20, 1992)

초 록: *Serratia marcescens* RH1의 분비성 단백질 분해효소 유전자의 클로닝, 염기서열 결정 및 대장균 내에서의 발현

이승환 · 김정민 · 권용태 · 고영희[†] **· 노현모** (서울대학교 자연과학대학 분자생물학과, [†]한국과학기술원 부설 유전공학연구소)

한국의 토양표본으로부터 다량의 세포외 단백질 분해효소를 분비하는 균주를 선별, 동정하여 *Serratia marcescens* RH1이라 명명하였다. 동 균주로부터 단백질 분해효소 유전자(*ssp*)를 shot gun cloning 방법을 이용하여 대장균에 클로닝한뒤 그 염기서열을 결정하였다. *ssp* 유전자는 3,135 개의 염기로 이루어진 open reading frame(ORF)를 가졌으며 112 kDa의 단백질을 coding 하였다. *ssp* 유전자의 염기서열 및 아미노산 서열을 분석한 결과, 이 유전자는 N-말단의 signal peptide(3.5 kDa), 성숙형 단백질(41.0 kDa), 그리고 C-말단의 거대 pro-peptide(67.0 kDa)로 이루어져 있었다. SSP는 대장균에서 자신의 promoter를 이용하여 대량 발현되었으며, 액체 배양중 활성효소가 배양액으로 분비되었다. *ssp* 유전자를 가지는 대장균 배양액을 SDS-PAGE 및 native-PAGE에 적용시킨 결과, SSP는 모균인 *S. marcescens* RH1의 주요 단백질 분해효소(45 kDa)와는 다른, 미량 단백질 분해효소로 밝혀졌다. 또한 대장균 배양액으로부터 SSP를 정제하여 그 생화학적인 특성을 연구한 결과, SSP는 단백질 분해효소로 밝혀졌으며, 최적 pH는 pH 9.0, 최적온도는 45°C였다.