Journal of Environmental Biotechnology
Vol. 7, No. 2, 99–102, 2007

Original paper (regular paper)

Cloning and Characterization of A Gene Encoding Algicidal Serine Protease from *Pseudoalteromonas* sp. Strain A28

Daiki Kohno[1], Yuichi Sakiyama[1], Sun-Og Lee[2], Noboru Takiguchi[1], Atsushi Mitsutani[3], Hirotaka Kitaguchi[3] and Junichi Kato[1]*

[1] *Department of Molecular Biotechnology, Hiroshima University, Higashi-Hiroshima, Hiroshima 739–8530, Japan*

[2] *Bio-Control Research Team, Korea Research Institute of Chemical Technology, Yusong PO Box 107, Taejon 305–600, Republic of Korea*

[3] *Department of Marine Biotechnology, Fukuyama University, Fukuyama, Hiroshima 729-0292, Japan*

* TEL: +81–(0)82–424–7757 FAX: +81–(0)82–424–7047

* E-mail: jun@hiroshima-u.ac.jp

(Received; 25 October, 2007/Accepted; 2 November, 2007)

The gene (*espI*) encoding an algicidal extracellular serine protease was isolated from *Pseudoalteromonas* sp. strain A28. DNA sequence analysis revealed that EspI is synthesized as a 711-amino-acid prepropeptide and is further processed to the 50-kDa matured protease after cleavage of a 148-amino-acid N-terminal domain and 6.7-kDa C-terminal domain. The Conserved Domain Search program found that EspI belongs to the subtilisin-like serine protease family. The program also found two PPC domains in the C-terminal region. The PPC domain is normally found at the C-terminus of secreted bacterial peptidases and they are not present in the active peptidase. However, the molecular size of purified EspI predicts that it contains one and half PPC domains in the matured EspI protease.

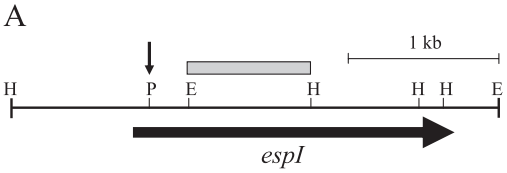
Key words: red tide algae, serine protease, alga-lysing bacterium, *Pseudoalteromonas*

Skeletonema costatum is the marine diatom which occasionally produces blooms in the Ariake sea, Fukuoka, Japan[6]. Its undesirable blooms cause severe damage to aquacultured laver. *Pseudoalteromonas* sp. strain A28, which was isolated from the Ariake Sea, is capable of lysing *S. costatum*[2]. The culture supernatant of strain A28 shows algicidal activity, suggesting that strain A28 produces an extracellular algicidal compound(s)[3]. Genetic analysis revealed an involvement of extracellular proteases in algicidal activity of strain A28. We identified two serine proteases, designated serine protease I (EspI) and serine protease II (EspII), as the major extracellular proteases. EspI is a monomeric protease with molecular weight of 50 kDa, whereas EspII is a heterodimer of 48-kDa and 33-kDa subunits. We demonstrated that EspI was responsible for the algicidal activity but not EspII[3]. Interestingly, some commercially available proteases, including trypsin, pepsin, subtilisin, and pronase, did not show algicidal activity. These results suggest that some properties of EspI, for example, substrate specificity or affinity to algal cells, are required for its algicidal activity. To further investigate molecular mechanism of algicidal activity of EspI, we cloned and characterized the *espI* gene encoding EspI from strain A28 in this study.

In the previous study, the N-terminal amino acid sequence of purified EspI was determined to be Ala-Thr-Pro-Asn-Asp-Pro[3]. To design an oligonucleotide probe, we determined the N-terminal amino acid sequence again. Amino acid sequencing enabled determination of the first 15 amino acid residues (Ala-Thr-Pro-Asn-Asp-Pro-Arg-Tyr-Asp-Gln-Trp-His-Tyr-Tyr-Glu). A degenerated 23-mer oligonucleotide (5'-TA(C/T)GA(C/T)CA(A/G)TGGCA(C/T)TA(C/T)TA(C/T)GA-3') derived from the N-terminal sequence, residue 8–15 was synthesized as a hybridization probe. Southern blot analysis[7] detected a 2.1-kb *Eco*RI fragment of the *Pseudoalteromonas* sp. strain A28 genome. A size-fractionated library of *Eco*RI fragments of strain A28 genomic DNA (1.5–2.5 kb in size) was created with a pUC118 vector[13] and *Escherichia coli* MV1184[13]. Plasmids DNAs, isolated from recombinant colonies, were screened with the oligonucleotide probe. One positive plasmid, designated as pPA28-1, was obtained. Restriction endonuclease analysis showed that pPA28-1 contained a 2.1-kb *Eco*RI fragment of the strain A28 genomic DNA. Nucleotide sequence analysis of the 2.1-kb *Eco*RI fragment found part of an open reading frame (*espI*). To obtain the 5' region of *espI*, Southern blot analysis with a 0.8-kb *Eco*RI-*Hin*dIII fragment as probe (Fig. 1) was performed and the 2.0-kb *Hin*dIII fragment hybridized to the 0.8-kb probe was cloned into pUC118. DNA sequencing found that the 2.0-kb *Hin*dIII fragment contained the 5' region of *espI*. The *espI* gene of 2133 bp could code for a protein of 711 amino acid residues

having a calculated molecular weight of 73.2 kDa. The N-terminal amino acid sequence of purified EspI was located within the 711-amino acid polypeptide starting at Ala-149 and is in perfect agreement with all 15 sequence amino acids (Fig. 1). To confirm that the cloned *espI* gene really encodes EspI, an *espI* mutant of strain A28 was constructed by inactivating the chromosomal *espI* gene. The *espI* gene on pPA28-1 was disrupted by inserting a kanamycin resistant gene cassette[12] into the *Pst*I site within the *espI* gene (Fig. 1). The resulting plasmid was introduced into strain A28 by electroporation[2] and kanamycin resistant transformants were selected. The insertion mutation of the chromosomal *espI* gene was confirmed by Southern blot analysis. The culture supernatant of the *espI*- mutant was examined for proteolytic and algicidal activities as described previously[3]. It showed ca. 70% decreased proteolytic and algicidal activities compared to that of the parent strain (Fig. 2), confirming that the cloned *espI* gene encodes algicidal serine protease I.

A computer-assisted sequence similarity search revealed that EspI showed 74% identity with *Pseudolateromonas piscicida* strain O-7 (formerly named *Alteromonas* sp. strain O-7) serine protease AprI[11]. *P. piscicida* stain O-7 was isolated from a sediment sample at the Sagami Bay of Japan as a chitin-degrading bacterium[9]. There is no information about algicidal activity of *P. piscicida* strain O-7 and AprI. It is predicted that the AprI protein is synthesized as a preproprotein with a 40-amino-acid signal peptide, a



B

```
                                                                                                  CCCTTCGCGAAAAATATCGTGGAGATTTACAATAACAATCTAGGGATCAATTATGACAACTAACAACAACTTGAAGAAGTGCGCAATTGCACTGAGCTTGAC   102
                                                                                                                                    M  T  T  N  N  N  L  K  K  C  A  I  A  L  S  L  T    17
TGCGTTGTTTGGGACGACGGCAGCAATGGCTACTCCAAACCAACTAATGTCACCATCTATGCAGGAAACTGCAGCGAAATTACAAGGCCAAGAAGGCTTTGG   204
  A  L  F  G  T  T  A  A  M  A  T  P  N  Q  L  M  S  P  S  M  Q  E  T  A  A  K  L  Q  G  Q  E  G  F  G    51
TACACAATTCATCATCAAGTATAAAAACAACAGTGATGAAATGATGACAATGTCGGCAGCAGAACAAATGCCGACTATGATGAACAAAAAAGCAAAAGGCTT   306
  T  Q  F  I  I  K  Y  K  N  N  S  D  E  M  M  T  M  S  A  A  E  Q  M  P  T  M  M  N  K  K  A  K  G  F    85
TGTTAAAAACTTCACCAGTAAAAAAGGTAAAGTAAAAGCACAATATGTTCGTGCAATGGCGATGAGTAATCACCACGTAATGCGTGCTGATAAAAAACTTTC   408
  V  K  N  F  T  S  K  K  G  K  V  K  A  Q  Y  V  R  A  M  A  M  S  N  H  H  V  M  R  A  D  K  K  L  S   119
TGCTCAAGAAGCACAAGAATTCATGCAAGAGATGGTTGCATCAGGTAACGTTGAATACATCGAAATCGACCAAATGTTAAAGCCATTTGCAACGCCTAACGA   510
  A  Q  E  A  Q  E  F  M  Q  E  M  V  A  S  G  N  V  E  Y  I  E  I  D  Q  M  L  K  P  F  A  T  P  N  D   153
CCCACGTTATGACGACCAATGGCACTACTACGAGCAAGCGGGTGGACTTAACCTTCCTACTGCGTGGGATACAGCAACGGGTAGCGGTGTGGTTGTTGCGGT   612
  P  R  Y  D  D  Q  W  H  Y  Y  E  Q  A  G  G  L  N  L  P  T  A  W  D  T  A  T  G  S  G  V  V  V  A  V   187
ACTTGATACAGGTTACCGCCCACATATTGATTTAAATGCCAACATCCTTCCTGGCTATGACATGATCTCTAACGTGTCTGTAGCTAACGATGGGGGTGGTCG   714
  L  D  T  G  Y  R  P  H  I  D  L  N  A  N  I  L  P  G  Y  D  M  I  S  N  V  S  V  A  N  D  G  G  G  R   221
TGATAGCGATGCACGCGATCCGGGTGATGCAATCAGCGCTAACGAATGTGGTTACACACATGGTGCGCAAAGCTCAAGCTGGCACGGTACGCACGTTGCGGG   816
  D  S  D  A  R  D  P  G  D  A  I  S  A  N  E  C  G  Y  T  H  G  A  Q  S  S  S  W  H  G  T  H  V  A  G   255
TACTGTTGCGGCAGTAACTAACAACGGTGAAGGTGTTGCCGGTGTTGCATACGATGCGAAAGTAGTTCCTGTACGTGTACTTGGTAAATGTGGTGGTTTAAC   918
  T  V  A  A  V  T  N  N  G  E  G  V  A  G  V  A  Y  D  A  K  V  V  P  V  R  V  L  G  K  C  G  G  L  T   289
TTCTGATATCGCTGACGGTATTATCTGGGCTTCAGGTGGTTCGGTTTCTGGCGTACCAGCAAACGCTAACCCTGCTGATGTTATCAACATGAGTTTAGGTGG  1020
  S  D  I  A  D  G  I  I  W  A  S  G  G  S  V  S  G  V  P  A  N  A  N  P  A  D  V  I  N  M  S  L  G  G   323
TAGCGGTTCATGTAGCTCTACAACACAAAGCGCTATCAACCAAGCACGAAACAACGGCACTGTGATCGTAATCGCTGCGGGTAACGATAACGATAACTCTGC  1122
  S  G  S  C  S  S  T  T  Q  S  A  I  N  Q  A  R  N  N  G  T  V  I  V  I  A  A  G  N  D  N  D  N  S  A   357
AAACTACAACCCAGGTAACTGTAATGGTGTTGTAAACGTTGCATCTGTTGGTCGTAACGGTGGTCGTGCTTATTACTCAAACTACGGTAGCAACATTGATGT  1224
  N  Y  N  P  G  N  C  N  G  V  V  N  V  A  S  V  G  R  N  G  G  R  A  Y  Y  S  N  Y  G  S  N  I  D  V   391
TGCAGCACCGGGTGGCGCACAAAGCTTTGCGAATGACTCTGAAGGTGTTTTATCAACACACAACTCAGGCTCTACAAGCCCAAGCAGCGATTCATATCACTA  1326
  A  A  P  G  G  A  Q  S  F  A  N  D  S  E  G  V  L  S  T  H  N  S  G  S  T  S  P  S  S  D  S  Y  H  Y   425
CTCGCAAGGTACATCGATGGCAGCACCACATGTTGCGGGTGTAGCAGCGCTAATCAAGCAAGCTAAACCATCTGCAACGCCAGATGAAATCGAAAGCATCTT  1428
  S  Q  G  T  S  M  A  A  P  H  V  A  G  V  A  A  L  I  K  Q  A  K  P  S  A  T  P  D  E  I  E  S  I  L   459
AAAATCAACAACTCGTTCTTTCCCTGCAACATGTACAAGCTGTGGTACAGGTATTGTTGATGCAGCAGCCGCTGTGGCAGCAGCATCAGGCACAACACCGCC  1530
  K  S  T  T  R  S  F  P  A  T  C  T  S  C  G  T  G  I  V  D  A  A  A  A  V  A  A  A  S  G  T  T  P  P   493
ACCACCAACAGGCAATGTGTTAGAAGATGGTCAAGCCCTAACAGGTTTAAGTGGCTCTGCTAGCAGCCAAACTTTCTACACAATGGAAGTGCCAACAGGTGC  1632
  P  P  T  G  N  V  L  E  D  G  Q  A  L  T  G  L  S  G  S  A  S  S  Q  T  F  Y  T  M  E  V  P  T  G  A   527
AACTAACGTAACCTTCACAATGAGTGGCGGTACAGGTGATGCTGACCTTTACGTTCGTGCAGGTAGCGCACCAACAACGTCAACCTATGATTGTCGCCCGTA  1734
PPC1
  T  N  V  T  F  T  M  S  G  G  T  G  D  A  D  L  Y  V  R  A  G  S  A  P  T  T  S  T  Y  D  C  R  P  Y   561
TAAAGGCGGTAACAACGAAGTATGTTCTATCGATAACCCAACAGCAGGCACTTACCATGTAATGCTCAATGGTTACTCTGCATACTCAGGTGTTAGCCTTGT  1836
  K  G  G  N  N  E  V  C  S  I  D  N  P  T  A  G  T  Y  H  V  M  L  N  G  Y  S  A  Y  S  G  V  S  L  V   595
GGGTGATTTAACAACTTCTGGCGGTGGTTCAGGCTCAGGTTCACCACAAGCAGGTGGCGGCACAGTGACTGATATCTCTGCAAGTGCTGGTCAATGGAAACA  1938
PPC2
  G  D  L  T  T  S  G  G  G  S  G  S  G  S  P  Q  A  G  G  G  T  V  T  D  I  S  A  S  A  G  Q  W  K  H   629
TTACACACTAGAAGTACCAGCAGGTATGGCAAGCTTCACTGTTACGACGTCTGGTGGTTCAGGTGATGCTGACTTATTCGTTAAGTTTGGTAGCCAACCAAC  2040
  Y  T  L  E  V  P  A  G  M  A  S  F  T  V  T  T  S  G  G  S  G  D  A  D  L  F  V  K  F  G  S  Q  P  T   663
GTCATCTAGCTATGACTGTCGTCCGTACAAAAACGGTAACGCAGAAACATGTACATTCAGTAATCCACAAGCAGGTACGTGGCACTTAAGCTTAAATGCTTA  2142
  S  S  S  Y  D  C  R  P  Y  K  N  G  N  A  E  T  C  T  F  S  N  P  Q  A  G  T  W  H  L  S  L  N  A  Y   697
CCGTACTTTCTCTGGTGTAACATTGGATGCACAGTATCAACCTTAATAGTTGATCTCCCCAGTATCCCCCTGGGACTTTTCAGCCGGACTTACGTCCGGCTT  2244
  R  T  F  S  G  V  T  L  D  A  Q  Y  Q  P                                                              711
```

Fig. 1. Restriction map of the *espI* region of *Pseudoalteromonas* sp. strain A28 (A) and the nucleotide sequence of the cloned *espI* gene and the deduced amino acid sequence of EspI (B).

(A) The location and orientation of *espI* are indicated by a horizontal arrow. The gray bar indicates a DNA fragment used as a probe for the gene walking experiment. Plasmid pPA28-2 was constructed by inserting a kanamycin resistant gene cassette into the *Pst*I sited indicated by the vertical arrow. Restriction sites: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I. (B) The 15-amino acid N-terminal amino acid residues of purified EspI are underlined. PPC1 and PPC2 are illustrated by solid boxes. Potential catalytic triad residues, 217-Asp, 249-His, and 430-Ser, are indicated by gray boxes. Nucleotide sequence of *espI* has been deposited in the GSDB, EMBL, DDBJ, and NCBI nucleotide sequence databases under the accession number AB365799.

110-amino-acid N-terminal propeptide, a 346-mino-acid mature protease, and a 219-amino-acid C-terminal propeptide which is processed into the 32.4-kDa mature protease[11]. In the case of *Pseudoalteromonas* sp. strain 28 EspI, the N-terminal domain with a similar size (148 amino acid residues) to that of strain O-7 AprI is removed during the maturation. However, since the purified EspI protease is 50 kDa[3], only a 6.7-kDa C-terminal domain (ca. 60 amino acid residues) is expected to be removed during the processing of EspI.

The Conserved Domain Search program[5] of the National Center for Biotechnology Information found a subtilase (subtilisin-like serine protease) family motif in EspI. Potential catalytic triad residues, Asp, His, and Ser, which are common in subtilisin-like serine proteases[8], are conserved at the positions corresponding to 217-Asp, 249-His, and 430-Ser (Fig. 1). The Conserved Domain Search program also found two PPC (bacterial pre-peptidase C-terminal) domains, designated PPC1 (residues 513–597) and PPC2 (residues 624–708) in the C-terminal region of EspI (Fig. 1). The PPC domain (Conserved Domain Database accession number: pfam04151 [the National Center for Biotechnology Information]) is normally found at the C-terminus of many secreted bacterial peptidases and they are removed during the maturation of the active peptidases[14]. Another *P. piscicida* strain O-7 extracellular serine protease AprII[10] and the *Mixococcus xanthus* metal protease[1] as well as EspI possess two PPC domains in their C-terminal regions, whereas there is only one PPC domain in many other bacterial proteases. In *P. piscicida* strain O-7 AprI, two PPC domains are cleaved off during maturation, while the molecular weight (50 kDa) of the matured EspI protease predicts that matured EspI still has whole PPC1 and half PPC2 in its C-terminal region. EmpI is an algicidal metal protease produced by strain A28[4]. Expectedly, the EspI serine protease did not show significant similarity to 38-kDa matured EmpI metal protease, however, the C-terminal region (residues 476–711) of EspI was 50.4% identical to that of EmpI. Sequence analysis revealed that EmpI also possesses two PPC domains in the C-terminal region and that these PPC domains are removed from the matured EmpI protease. It was shown that EmpI was 6-fold less algicidal to *S. costatum* than EspI on the basis of the mass of proteins[4]. Thus, the PPC domains in matured EspI might have a role in algicidal activity. We are now investigating this possibility using obtained *espI* gene.

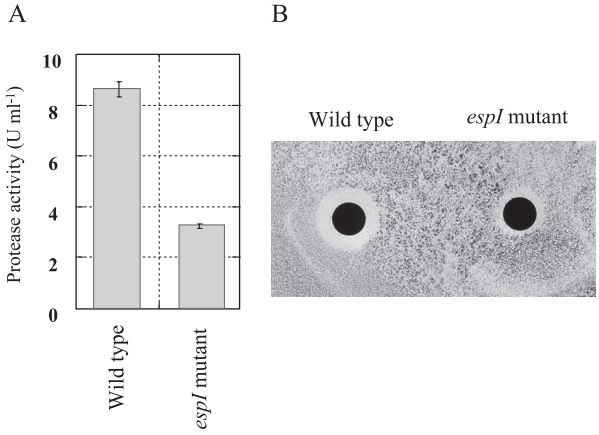


Fig. 2. Protease activity (A) and algicidal activity (B) in culture supernatants of strain A28 and its *espI*- mutant.
(A) Protease activity was measured by using azocasein as a substrate as described by Lee, et al3). One unit protease activity was defined as the amount of enzyme that caused incremental change of one absorbance unit per hour. Vertical bars represent the standard deviations of measurements done in four separate experiments. (B) Algicidal activity was detected as described previously3). Zones of clearing around paper disks indicate the lysis of the diatom *S. costatum* strain NIES-324. Plates were photographed after 48 h of incubation at 20°C.

References

1) Kearns, D.B., P.J. Bonner, D.R. Smith, and L.J. Shimkets. 2002. An extracellular matrix-associated zinc metalloprotease is required for dilauroyl phosphatidylethanolamine chemotactic excitation in *Myxococcus xanthus*. J. Bacteriol. 184: 1678–1684.
2) Kato, J., J. Amie, Y. Murata, A. Kuroda, A. Mitsutani, and H. Ohtake. 1998. Development of a genetic transformation system for an alga-lysing bacterium. Appl. Environ. Microbiol. 64: 2061–2064.
3) Lee, S.-O., J. Kato, N. Takiguchi, A. Kuroda, T. Ikeda, A. Mitsutani, and H. Ohtake. 2000. Involvement of an extracellular protease in algicidal activity of the marine bacterium *Pseudomonas* sp. strain A28. Appl. Environ. Microbiol. 66: 4334–4339.
4) Lee, S.-O., J. Kato, K. Nakashima, A. Kuroda, T. Ikeda, N. Takiguchi, and H. Ohtake. 2002. Cloning and characterization of extracellular metal protease gene of the algicidal marine bacterium *Pseudoalteromonas* sp. strain A28. Biosci. Biotechnol. Biochem. 66: 1366–1369.
5) Marchler-Bauer, A., J.B. Anderson, M.K. Derbyshire, C. DeWeese-Scott, N.R. Gonzales, M. Gwadz, L. Hao, S. He, D.I. Hurwitz, J.D. Jackson, Z. Ke, D. Krylov, C.J. Lanczycki, C.A. Liebert, C. Liu, F. Lu, S. Lu, G.H. Marchler, M. Mullokandov, J.S. Song, N. Thanki, R.A. Yamashita, J.J. Yin, D. Zhang, and S.H. Bryant. 2007. CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res. 35: D237–240.
6) Mitsutani, A., T. Takase, M. Kirita, and Y. Ishida. 1992. Lysis of *Skeletonema costatum* by *Cytophaga* sp. isolated from the coastal waters of the Ariake Sea. Nihon Suisan Gakkai-shi. 58: 2159–2169.
7) Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8) Siezen, R.J., W.M. de Vos, and B.W. Dijkstra. 1991. Homology modeling and protein engineering strategy of subtilinases, the family of subtilisin-like serine proteases. Protein Eng. 4: 719–737.
9) Tsujibo, H., Y. Toshida, C. Imada, Y. Okami, K. Miyamoto, and Y. Inamori. 1991. Isolation and characterization of a chitin degrading marine bacterium belonging to the genus *Alteromonas*. Nippon Suisan Gakkaishi 57: 2127–2131.
10) Tsujibo, H., K. Miyamoto, K. Tanaka, M. Kawai, K. Tainaka, C. Imada, Y. Okami, and Y. Inamori. 1993. Cloning and sequence of an alkaline serine protease-encoding gene from the marine bacterium *Alteromonas* sp. strain O-7. Gene 136: 247–251.

11) Tsujibo, H., K. Miyamoto, K. Tanaka, Y. Kaidzu, C. Imada, Y. Okami, and Y. Inamori. 1996. Cloning and sequence analysis of a protease-encoding gene from the marine bacterium *Alteromonas* sp. strain O-7. Biosci. Biotechnol. Biochem. 60: 1284–1288.
12) Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19: 259–268.
13) Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. Methods Enzymol. 153: 3–11.
14) Yeats, C., S. Bentley, and A. Bateman. 2003. New knowledge from old: in silico discovery of novel protein domains in *Streptomyces coelicolor*. BMC Microbiol. 3: 3 (electronic publication).