

Cold-Active Serine Alkaline Protease from the Psychrotrophic Bacterium *Shewanella* Strain Ac10: Gene Cloning and Enzyme Purification and Characterization

LJUDMILA KULAKOVA, ANDREY GALKIN, TATSUO KURIHARA,
TOHRU YOSHIMURA, AND NOBUYOSHI ESAKI*

Institute for Chemical Research, Kyoto University, Uji, Kyoto-Fu 611, Japan

Received 22 June 1998/Accepted 12 October 1998

The gene encoding serine alkaline protease (SapSh) of the psychrotrophic bacterium *Shewanella* strain Ac10 was cloned in *Escherichia coli*. The amino acid sequence deduced from the 2,442-bp nucleotide sequence revealed that the protein was 814 amino acids long and had an estimated molecular weight of 85,113. SapSh exhibited sequence similarities with members of the subtilisin family of proteases, and there was a high level of conservation in the regions around a putative catalytic triad consisting of Asp-30, His-65, and Ser-369. The amino acid sequence contained the following regions which were assigned on the basis of homology to previously described sequences: a signal peptide (26 residues), a propeptide (117 residues), and an extension up to the C terminus (about 250 residues). Another feature of SapSh is the fact that the space between His-65 and Ser-369 is approximately 150 residues longer than the corresponding spaces in other proteases belonging to the subtilisin family. SapSh was purified to homogeneity from the culture supernatant of *E. coli* recombinant cells by affinity chromatography with a bacitracin-Sepharose column. The recombinant SapSh (rSapSh) was found to have a molecular weight of about 44,000 and to be highly active in the alkaline region (optimum pH, around 9.0) when azocasein and synthetic peptides were used as substrates. rSapSh was characterized by its high levels of activity at low temperatures; it was five times more active than subtilisin Carlsberg at temperatures ranging from 5 to 15°C. The activation energy for hydrolysis of azocasein by rSapSh was much lower than the activation energy for hydrolysis of azocasein by the subtilisin. However, rSapSh was far less stable than the subtilisin.

Cold-adapted microorganisms, which include psychrophiles and psychrotrophs, are known to produce enzymes with high levels of activity at low temperatures; these enzymes are called cold-active enzymes (8, 15). Homologous counterparts of the cold-active enzymes are produced by mesophilic or thermophilic microorganisms but are less active at low temperatures (10).

In order to obtain high catalytic efficiency, cold-active enzymes probably have evolved to have high conformational flexibility, although stability has been sacrificed. It is thought that the flexible structures of these enzymes are based on weakened noncovalent interactions, such as salt bridges, hydrogen bonding, hydrophobic interactions, and aromatic-aromatic interactions. α -Amylase from an Antarctic psychrophile, *Alteromonas haloplanktis* A23, has been shown to contain fewer surface salt bridges, to have fewer polar interactions, and to have a lower proline content than porcine pancreatic α -amylase (9). However, the molecular basis of the high levels of activity of cold-active enzymes remains unclear. The fine tertiary structures of cold-active enzymes need to be compared with the fine tertiary structures of their counterparts that exhibit less cold activity. Moreover, it would be effective to study a set of enzymes that exhibit high levels of sequence similarity in order to simplify the comparison. The subtilisin family is probably one of the best models for studying structure-function relationships because of the three-dimensional structures and the primary struc-

tures of various enzymes belonging to this family have been clarified.

We found a psychrotroph of a *Shewanella* species that exhibited a high level of protease activity at low temperatures and cloned a subtilisinlike protease gene from this bacterium. In this paper we describe the characteristics of a cold-active recombinant subtilisinlike protease.

MATERIALS AND METHODS

Fatty acid composition. *Shewanella* strain Ac10 was cultured at 4°C in a medium (pH 7.2) containing 1.5% Polypeptone, 0.1% yeast extract, 0.1% glycerol, 0.2% K_2HPO_4 , 0.1% KH_2PO_4 , 0.01% $MgSO_4 \cdot 7H_2O$, and 3% NaCl. Cells (about 2.7 g, wet weight) were freeze-dried and suspended in 5 ml of 5% (wt/vol) NaOH in 50% aqueous methanol, and the mixture was then incubated at 100°C for 15 min. The saponified material was cooled and acidified to pH 2 by adding concentrated HCl. A 4-ml portion of a 14% (vol/vol) boron trifluoride solution in methanol was added, and the mixture was heated at 100°C for 5 min. Then 10 ml of a saturated sodium chloride solution was added to the mixture, and fatty acid methyl esters were extracted twice with an equal volume of $CHCl_3$ -*n*-hexane (1:4, vol/vol). The combined extracts were evaporated under a stream of nitrogen gas to a volume of about 0.1 ml. The fatty acid methyl esters were analyzed with a Shimadzu model GC-14A gas chromatograph equipped with a flame ionization detector and a type HR-101 capillary column. The temperature of the column was kept at 120°C for the first 5 min; then it was increased to 220°C at a rate of 3°C/min and then kept at 220°C for 10 min. The injector and detector temperatures were 250 and 280°C, respectively. The fatty acid methyl esters were identified by comparison with authentic samples purchased from Sigma.

DNA manipulation. The standard protocols of Sambrook et al. (26) were used for DNA manipulation. A genomic library of *Shewanella* strain Ac10 was prepared with *Bam*HI and pUC118 in *Escherichia coli* TG1 [*F'* *traD36 proAB lacI^q Δ lacZ M15 Δ (lac-pro) thi hsdR ara*]. PCR were performed with a thermal cycler (Perkin-Elmer Cetus) by using 0.02- to 0.1-ml reaction mixtures containing deoxynucleoside triphosphates at concentrations of 0.02 to 0.2 mM, 20 to 100 pmol of primers, 10 to 200 ng of *Shewanella* strain Ac10 genomic DNA, 0.01 ml of 10 \times reaction buffer, and 2.5 U of exTaq or LATaq DNA polymerase (Takara). The program used for the PCR was as follows: up to 45 cycles consisting of denaturation at 95°C for 1 min, annealing at 30 to 45°C for 2 min, and

* Corresponding author. Mailing address: Institute for Chemical Research, Kyoto University, Uji, Kyoto-Fu 611-0011, Japan. Phone: 81-774-38-3240. Fax: 81-774-38-3248. E-mail: esaki@scl.kyoto-u.ac.jp.

TABLE 1. Effect of growth temperature on the cellular fatty acid composition of *Shewanella* strain Ac10

Temp (°C)	% of the following fatty acids:							
	12:0	14:0	16:0	16:1t	18:0	18:1	20:5	Unknown
4	3.7	5.2	1.1	51.0	4.3	22.0	12.1	0.6
15	3.1	8.9	4.9	47.1	3.9	21.5	9.8	0.8
25	4.1	12.5	7.5	43.9	4.1	21.0	5.1	1.8

extension at 72°C for 2 min. Colony hybridization was carried out by the digoxigenin labeling method performed with a kit supplied by Boehringer Mannheim. The oligonucleotides used for 16S ribosomal DNA (rDNA) amplification were prepared as described previously (33). DNA sequencing was performed by using an Applied Biosystems model 377B automated DNA sequencer and a dye-labeled terminator sequencing kit (Applied Biosystems). The sequence data resulted in 1,488 usable bases for the 16S rDNA sequences. Sequences were aligned by using various previously described sequences obtained from the Ribosomal Database Project (19), GenBank, and EMBL databases and the MEGALIGN program of DNASTAR (DNASTAR Inc.). The GenCANS-RDP system obtained from the internet site <http://diana.uthct.edu/gencans.html> (34) was used for classification.

Assays. Proteolytic activity was determined by using *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (AAPF) as the substrate in 50 mM Tris-HCl (pH 9.0) containing 2 mM CaCl₂. One unit of enzyme activity was defined as the amount of enzyme that catalyzed the formation of 1 μmol of *p*-nitroaniline per min, which was determined with an extinction coefficient of 8,480 M⁻¹ · cm⁻¹ at 412 nm (7). The activity toward azocasein was determined in a reaction mixture containing 50 mM Tris-HCl (pH 9.0), 2 mM CaCl₂, and 6% azocasein. The reaction was performed at 25°C for 15 min and was stopped by adding trichloroacetic acid to a final concentration of 5%. The chromophore was determined with an extinction coefficient of 900 M⁻¹ · cm⁻¹ at 366 nm.

Purification of SapSh from *Shewanella* strain Ac10. All of the procedures used to purify serine alkaline protease (SapSh) from *Shewanella* strain Ac10 were performed at 4°C, and 50 mM Tris-HCl (pH 8.5) supplemented with 2 mM CaCl₂ was used as the standard buffer unless indicated otherwise. *E. coli* BL21(DE3) [*F*⁻ *ompT* *hdsB*_B(*r*_B⁻ *m*_B⁻) *gal dcm*] cells harboring pSapSh3, which encodes the cloned protease gene, were grown aerobically in 50 ml of 2YT medium supplemented with 0.2 mg of ampicillin per ml in a 500-ml Sakaguchi flask at 37°C. IPTG (isopropyl-β-D-thiogalactopyranoside) was added to the culture to a final concentration of 10 μM after 6 h, when the absorbance at 660 nm was around 0.6 to 0.7. Then the temperature of the culture was adjusted to 15°C, and the culture was incubated aerobically for about 12 h. The supernatant culture was dialyzed against the standard buffer and applied to a bacitracin-Sepharose column, which was prepared by the method of Stepanov and Rudenskaya (29). The column was washed first with the standard buffer and then with the standard buffer supplemented with 1 M NaCl. The enzyme was eluted with the standard buffer supplemented with 1 M NaCl and 25% isopropanol. The active fractions were dialyzed against the standard buffer.

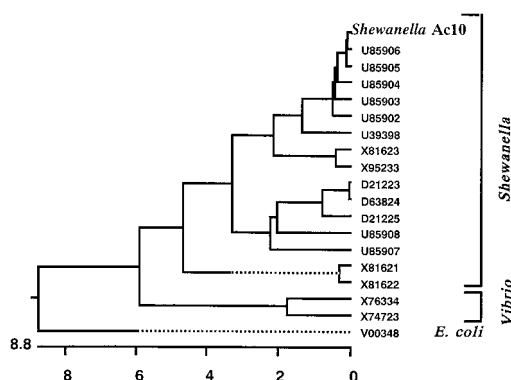


FIG. 1. Phylogenetic relationship between the 16S rDNA sequence of *Shewanella* strain Ac10 and the 16S rDNA sequences of other *Shewanella* strains and selected *Vibrio* strains. The balanced cladogram was constructed by using a matrix of pairwise genetic distances generated by the Clustal method with the MEGALIGN program. The scale indicates percentages of sequence divergence. The numbers are the GenBank accession numbers of the 16S rDNA sequences of various bacteria.

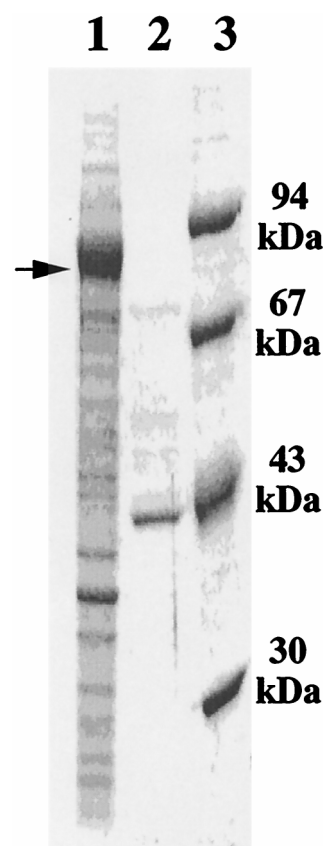


FIG. 2. PAGE of rSapSh. Lane 1, precipitate obtained from cell extract of *E. coli* BL21(DE3)/pSapSh3; lane 2, soluble active preparation of rSapSh after bacitracin-Sepharose column chromatography; lane 3, molecular weight markers. The position of the inclusion body formed from the preproprotein of rSapSh is indicated by an arrow.

N-terminal sequencing. Protein samples isolated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) were blotted electrically onto a polyvinylidene difluoride membrane (Millipore). The membrane was stained with 0.1% (wt/vol) Ponceau S in 2% (vol/vol) acetic acid, and the stained bands were cut out and subjected to an N-terminal sequence analysis with a model PPSQ-10S protein sequencer (Shimadzu).

Nucleotide sequence accession number. The nucleotide sequence of the protease gene determined in this study has been deposited in the GenBank database under accession no. AF047370. The GenBank accession number for the 16S rDNA sequence of *Shewanella* strain Ac10 is AF061557.

RESULTS AND DISCUSSION

Characterization of *Shewanella* sp. strain Ac10. We searched for psychrotrophic strains that exhibited high levels of protease activity at low temperatures in the stock cultures maintained in our laboratory. Strain Ac10 was selected as the best protease producer. This organism grew well at 4°C, had an optimum growth temperature of around 20°C, and did not grow at temperatures greater than 30°C. According to the definition of Morita (22), strain Ac10 was a psychrotroph. Cold-adapted microorganisms are characterized by their unique fatty acid compositions. Eicosapentaenoic acid (EPA) (20:5) has been detected in psychrotrophs belonging to the genus *Shewanella* at levels ranging from 2 to 16% of the total fatty acids (1). Therefore, the fatty acid composition of strain Ac10 was determined at different growth temperatures (Table 1). In cells grown at 4°C EPA accounted for 12.1% of the total cellular fatty acids. However, the level of EPA decreased as the growth tempera-

A

TACGTCACCCCTGATTAAACAT **GGAGG**TAAT

ATGAAAAAGCATAAAAAACCAACGGTCGTGCTTAGCATGACAGCACTCGCTTTGGCCATTTCAACTCAAGTAAACGAGCACCAGCTCCTTCTCAACTCTCTGCAACGAATCAACAAATT
M K K H K N P T V V L S M T A L A L A I S T Q V N A A P A P S Q L S A T N Q Q I
 -143 -118 ↑

GAACGACAATCACCTTTACAAAACGATATATCGTTAAGTTCAGAAATGATAATGCTGCACAGATACTCAATAGCAGTCCGCTGTGCAAAATGTTAATGCTGAAGCGGACAACAGTATG
 E R Q S P L P K R Y I V K F R N D N A A Q I L N S S S A V Q N V N A E A D N S M

GCTTACGAGCCTCGTTCAAACGAAGTATTTACTCAATTTAGAGCCTTGAATACGCTAAAGCAGGAGAATGAAGCGGTAGTTCGCAAGTAAATGTTACTCGGTAAGCTCGACTCGTCT
 A Y E P R S N E V F T Q F R A L N S V K A R E M K R V G R S N S Y S V K L D S S

TCTATTAAGCATTACGTTTTCGCTCAAGATGTTGAATACGTTGAAGAGGATTTACCCCGTAGACTGCTAGCAGAGACAACACCTTGGGGACAACCTTTGATAGGTGCGACTGTGTTAAGC
 S I K A L R L R Q D V E Y V E E D L P R R L L A E T T P W G Q T F V G A T V L S
 +1

GACAGCCAAGCAGGTAACCGTACTATCTGTATTATTGATTCTGGTTATGACCCGAGCCACAATGATTTAAGTGGTAACAATGCAACCGGGACCAATAATTCAGGCACCGGTAACCTGGTAT
 D S Q A G N R T I C I I **D** S G Y D R S H N D L S G N N A T G T N N S G T G N W Y

CAACCAGGTAATAATATGCCATGGTACTACGATGACAGTACCATTGCCCGAATAGCAAATAACGAAGTGTGGTGGGGTTATGCCAACCAGAAATGCCAATATTCATGTTATAAAT
 Q P G N N N A **H** G T H V A G T I A A I A N N E G V V G V M P N Q N A N I H V I N

GTGTTAAACGAAGCAGGATGGGGTACTCATCATCGTTAGTCGCTGCGATTGATACGCTGTGTTACTTCAGTGGTGCACCAACGTAGTCAACAATGAGTTAGGTGGCAGTGGTTCAACCACA
 V F N E A G W G Y S S S L V A A I D T C V T S G G A N V V T M S L G G S G S T T

ACTGAGCGTAATGCATTAATACTCATTATAACGGTGGGGTGTATTAAATGTCAGCTGCTGTAATGCTGGCGACAGTACTTATACCTACCCACCATCTTATGATATCGTAAATGCTGTT
 T E R N A L N T H Y N G G V L L I A A A G N A G D S T Y T Y P P S Y D I V M S V

GCCGCGTGTGACAGCAATCTTGATCATGCTGCTTTTCAATATACAGCAAGTGAATTTCAAGTCTGGCGAGGCAATTTACCAACGGTACCGTTGGCGAAGTGGTATTAGCA
 A A V D S N L D H A A F S Q Y T D Q V E I S G P G E A I L P T V T V G E G R L A

AATATACTATTGGCGGTGCTTACTTTAACAATGGTGTAGTCCCGACAATCGTTTAAACCCATCTGGTACTAGTACACTGCTGCACCGATTAAACGAAGTGAACCTGGAACCTTA
 N I T I G G Q S Y F N N G V V P H N R L T P S G T S Y T A A P I N A S A T G T L

GCGGAATGACTGTAAACGGAACAGTTTTAAATGCGGTAGCATGACCAATAAAATCTGCTTGTGTAACGTTGAAGTAACCAAGTTCAAGTTACCTCGAAATAAACCGGTTAAAGCC
 A E C T V N G T S F N C G S M T N K I C L V E R V S N Q G S S Y P E I N A V K A

TGTAACACAGCTGGTCCAAAGGAATCATTGTTTACAGTAACACTCGCTACCAAGGCTACAAAATCCATTCGTTGATGCTGACAGTAAATTCGATCCCTTCTATGCTGTTGAC
 C K T A G A K G I I V Y S N T A L P R L Q N P F V V D A D S E I L I P S M S V D

CGAACACAGGCTTGCACTTAAAGCCAAGTGGGTCATCAACCACTGTGAGTAATCAAGGCAATCGAGATTATGAATATTACAATGGCACTCAATGGCAACACCGCATGTGTCGGC
 R T T G L A L K A K L G Q S T T V S N Q G N R D Y E Y Y N G T **S** M A T P H V S G

GTAGCAACATTAGTGTGGAGTTACCACCCGAATGTAGCGCCAGCAAGTACGTGCTGATTAATGCCACAGCAGATGACTTAAAGCGTTGGCGGTGATAACCAACAGGTTACGGG
 V A T L V W S Y H P E C S A S Q V R A A L N A T A D D L S V A G R D N Q T G Y G

ATGGTTAATGCAACACAGCAAAAGGCTTACCTAAACGAATCATGTAATGGTCCACCGATCCAGGTACAGGCTCTGGTGACAGGCTACTGAAAAGGCGGTGGCAAAACAGGCTTAGCT
 M V N A T T A K A Y L N E S C N G P T D P G T G S G D S V L E K G V A K T G L A

GCGCTAAAACGATGAATATATTCTCGTTAGACATTCCTGCAGGCGCAGACTTAAAGTTCATATGAGTGGTGGCACTGGCGATGCTGACTTATATGTCAGTATGGCGATCA
 G A K N D E L Y F S L D I P A G A T D L S F T M S G G T G D A D L Y V Q Y G A S

CCTACCAGCAGGTTACGACTGTCGCTCCATGGAAAGGTGTAATGCAGAACTTGGCCAAATGGCAGCCCTCAATCTGGTACTTACTATGTAATGCTACAAGTTATAATGCATTTAGT
 P T S S S Y D C R P W K G G N A E S C P I A T P Q S G T Y Y V M L Q G Y N A F S

GGTGAATTTAGTGGCTAACTACACCGTGGTAGCACTGGAGCAGCAACACAGGCGCCAGCAAGCTACAACAATACCGGTAATTATACAATTCAGACAACAACACTAGTGGTATT
 G V N L V A N Y T A G S T G G S N T G G P A S Y N N T G N Y T I P D N N T S G I

ACTAGCCCTTACTGTTAGCCGACTGTTGATTCAGGAACGGTATCGGTGCAATGGGATTATCATACTTATATTTGGTATTAAATAACCGGTTAGTTAGCCCTTACCGGCCAAAC
 T S P I T V S R T V S V S I G T V S V S I G G I I H T Y I G D L K I Q L V S L T G Q T

GTTGACTGCATAACAACACTGGTGGTACCATAATCTAACCCAGACTTACACCGCTAATATGCGAGTGTGCAATCAAGTGGTGTAGGACATTAAGCTGTTGATAATGCGAGA
 V V L H N N T G G G T D N L T Q T Y T A N M A G V E S S G V W T L K A V D N A R

CAAGATACTGGATATATTGATACTTGGAGCTTAAGCTTCCAGT AATATTATTAGTTGATTACAAAA
 Q D T G Y I D T W S L S F Q *

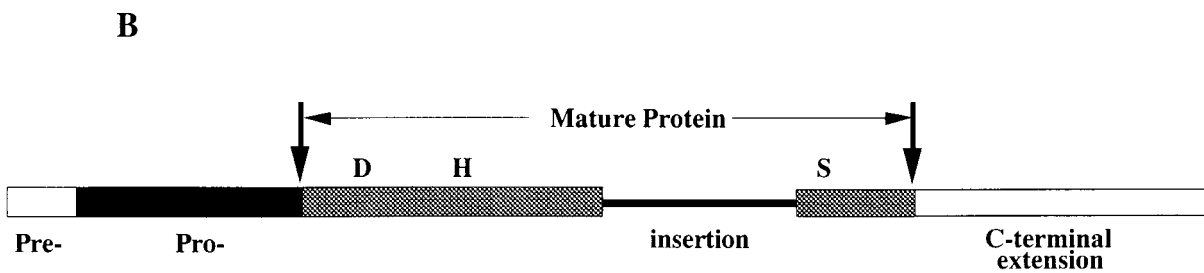


FIG. 3. (A) Nucleotide and deduced amino acid sequences of SapSh. The potential ribosome binding site is indicated by a different typeface. The stop codon is indicated by an asterisk. The predicted signal peptide cleavage site is indicated by an arrow. The N terminus of the mature protein is indicated by +1. The N-terminal amino acid sequences of the proenzyme and the mature enzyme are underlined with two lines and one line, respectively. The catalytic triad (D, H, and S) is indicated by boldface type. (B) Schematic diagram of the deduced amino acid sequence of SapSh.

TABLE 2. Properties of SapSh and subtilisin Carlsberg determined from a comparison of the deduced amino acid sequences

Enzyme	Amino acid frequencies (%) ^a						pI	GRAVY index ^b	Aliphatic index
	Charged amino acids	Acidic amino acids	Basic amino acids	Polar amino acids	Hydrophobic amino acids	Aromatic amino acids			
Subtilisin Carlsberg	11.7	5.1	4.7	28.1	37.2	6.6	6.57	0.132	84.78
SapSh	13.4	7.0	4.2	34.0	32.4	6.3	5.07	-0.157	75.38

^a The charged amino acids are arginine, lysine, histidine, aspartic acid, and glutamic acid; the acidic amino acids are aspartic acid and glutamic acid; the basic amino acids are lysine and arginine; the polar amino acids are asparagine, cysteine, glutamine, serine, and threonine; the hydrophobic amino acids are alanine, isoleucine, leucine, phenylalanine, tryptophan, and valine; and the aromatic amino acids are phenylalanine, tryptophan, and tyrosine.

^b GRAVY, Grand Average of Hydrophobicity.

ture increased; in cells grown at 25°C EPA accounted for only 5.1% of the total cellular fatty acids. The same tendency was observed for the palmitelaic acid (16:1t) content. The contents of other fatty acids were not affected significantly by the growth temperature. These results are consistent with the cold-adapted nature of strain Ac10.

As determined by 16S rDNA sequencing, strain Ac10 was more closely related to the genus *Shewanella* (levels of similarity, 91 to 98.6%) than to the genus *Vibrio* (levels of similarity, less than 91%) (Fig. 1). Strain Ac10 exhibited the highest levels of similarity (97.8 to 98.6%) to the Antarctic species *Shewanella frigidimarina* (1), and thus we placed this organism in the genus *Shewanella*, as *Shewanella* sp. strain Ac10. Data obtained with Ribosomal Database Project program and GenCANS classification also suggested that Ac10 is a member of the *Alteromonas* (*Shewanella*) group belonging to the gamma subdivision of the class *Proteobacteria* (gram-negative phylum).

Cloning, subcloning, and expression of the alkaline protease gene. We isolated a subtilisinlike protease gene from the genomic DNA of *Shewanella* sp. strain Ac10 by using a DNA probe prepared by PCR analysis as follows. The following primers were synthesized on the basis of the consensus amino acid sequences of subtilisins at positions 62 to 70 and 219 to 227 (27): forward primer, 5'-AA(CT)GGICA(CT)GGIACI CA(CT)GTIGCIGG (His); and reverse primer, 5'-C(AG)TGI GGIG(TC)IGCCATI(GC)(AT)IGTICC (Ser). An approximately 900-bp DNA fragment amplified by PCR was found to have a sequence that is conserved in the corresponding region of the subtilisin family of proteins. A DNA library of the *Shewanella* sp. strain Ac10 genome was constructed with vector plasmid pUC118 and host strain *E. coli* TG1. We selected clones with positive hybridization signals, and all of these clones contained plasmids with insert DNAs that were about 4.7 kbp long. We determined the nucleotide sequence of the insert DNA in a plasmid designated pSapSh1; a single open reading frame comprising 2,442 bp was found to encode a subtilisinlike protease that was 814 amino acids long and had a predicted molecular weight of 85,113. A putative Shine-Dalgarno sequence (5'-GGAAGA) occurred 4 bases upstream from the ATG initiation codon.

No protease activity was found in a culture of recombinant *E. coli* cells harboring pSapSh1. Therefore, a 3-kb *Bam*HI-*Eco*RI fragment was amplified by PCR from pSapSh1 and cloned into pUC119 downstream of the *lac* promoter (pSapSh2) and into pET21a downstream of the T7lac promoter (pSapSh3). An intense protein band with the predicted molecular mass (about 85,000 Da) appeared after SDS-PAGE of the precipitate fraction (i.e., inclusion bodies) obtained an extract of sonicated *E. coli* BL21/pSapSh3 cells (Fig. 2). The inclusion bodies in the precipitate were formed by induction with IPTG at a final concentration of 1 mM. We obtained active protease in the supernatant of the culture broth after

IPTG was added by using a lower IPTG concentration (10 μ M) and changing the cultivation temperature from 37 to 15°C after IPTG was added. However, no protease activity was found in an extract of cells cultured under the same conditions.

Deduced amino acid sequence analysis. The amino acid sequence of the cloned subtilisin (SapSh) (Fig. 3A) was compared with the sequences of other serine proteases. SapSh exhibited high levels of sequence similarity to members of the subtilase family of serine proteases around the active site residues (27, 28); the overall levels of sequence homology were about 25 to 28%. The sequences around the catalytic triads of these subtilisins are highly conserved in the corresponding regions of SapSh (Asp-30, His-65, and Ser-369). We determined the N-terminal amino acid sequence of the inclusion body preparation described above; this sequence was MKKH KNPTVVVL. We found that this sequence is very similar to the predicted amino sequence (Fig. 3B). A typical 26-amino-acid signal peptide (24) occurred at the N terminus; there was a charged amino acid (Lys-142), followed by a hydrophobic core and a small uncharged amino acid (Ala-118) at the C terminus of the putative signal peptide. We predicted that the signal peptide cleavage site would be between Ala-118 and Ala-117 by using the method of von Heijne (31). The N-terminal amino acids of the mature enzyme were determined with a protein sequencer to be AETTPWGQTFV (Fig. 3B). Thus, a 117-amino-acid propeptide probably occurs between the C terminus of the signal peptide and the N terminus of the mature protein. A similar long propeptide region has been found in other extracellular proteases from bacteria (6, 21). The propeptide of SapSh may be removed autocatalytically in the same way that the propeptides of other subtilisinlike serine proteases are removed (12, 17, 25).

An unusual feature of SapSh is the fact that the space between His-65 and Ser-369 is approximately 150 residues longer than the corresponding spaces in other proteins belonging to the subtilisin family. Large insertions of similar lengths have been found in proteases belonging to the pyrolysins family (28). For example, the cell envelope proteinase of *Lactococcus lactis*, which is a member of this family, has an insertion that is 151 residues long and is believed to determine the unique substrate specificity of this enzyme (2). Heat-stable proteases of the hyperthermophiles *Pyrococcus furiosus* and *Thermococcus steteri* are also members of the pyrolysins family and have large insertions that are 147 and 163 residues long, respectively. These regions probably contribute to protection against self-digestion and thermal denaturation (32). However, no sequence similarity was found between the insertion region of SapSh and the insertion regions of proteins belonging to the pyrolysins family. Moreover, the location of the insertion in the sequence of SapSh differs markedly from the locations of the insertions in the sequences of members of the pyrolysins family. Thus, SapSh can be easily excluded from this family.

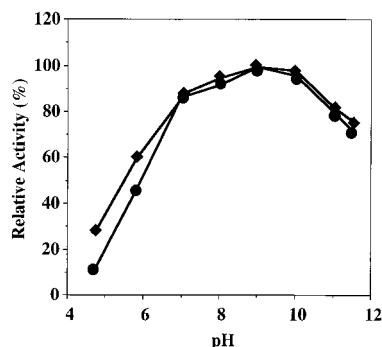


FIG. 4. Effect of pH on the activity of rSapSh toward azocasein (◆) and the synthetic peptide AAPF (●).

In addition to the long space between His-65 and Ser-369, SapSh has a long C-terminal extension (length, approximately 250 amino acids) whose sequence is similar to the sequences of the corresponding regions of extracellular proteases from several *Vibrio* species, including *Vibrio alginolyticus* (5), *Vibrio cholerae* (13), *Vibrio anguillarum* (21), and *Vibrio proteolyticus* (6, 30), although the levels of homology between SapSh and the *Vibrio* enzymes are much lower than the levels of homology between the *Vibrio* enzymes. The C-terminal extensions are thought to be involved in secretion of the proteases through the outer membranes of gram-negative bacteria (18). Therefore, the C-terminal extension of SapSh also may participate in transport through the membrane.

The amino acid composition of SapSh was compared with the amino acid composition of subtilisin Carlsberg (Table 2). The pI of SapSh is lower than the pI of subtilisin Carlsberg due to higher and lower contents of acidic and basic amino acids, respectively, in SapSh. Davail et al. (4) proposed on the basis of a similar observation for subtilisin S41 from a psychrophile that a high acidic residue content on a protein surface results in increased interaction between the protein and solvent, which destabilizes the protein structure. Another characteristic of SapSh is the fact that its hydrophobic amino acid content is lower than that of subtilisin Carlsberg. When hydrophobicity was estimated by using the Grand Average of Hydropathicity (GRAVY) index obtained from <http://www.expasy.ch/sprot/protparam.html> (16), the value for SapSh was less than the value for subtilisin Carlsberg (Table 2), indicating that SapSh is much less hydrophobic than subtilisin Carlsberg. The fact that the thermostability of SapSh is lower than the thermostability of subtilisin Carlsberg is consistent with the general finding that hydrophobic interactions are important for protein thermostability. Moreover, the aliphatic index calculated from molar ratios and the relative volumes of the Ala, Val, Ile, and Leu residues by Ikai's method (14) for SapSh was lower than the aliphatic index calculated for subtilisin Carlsberg (Table 2). Ikai observed a correlation between aliphatic indices and protein thermostabilities; higher indices are obtained for proteins with greater thermostabilities.

Purification of rSapSh. Recombinant SapSh (rSapSh) was purified by affinity chromatography with a bacitracin-Sepharose column. Only one major band was obtained with the final preparation of rSapSh when SDS-PAGE was performed (Fig. 2). The apparent molecular mass of rSapSh was determined to be approximately 44,000 Da, which is lower than the predicted molecular weight (85,113) of the precursor form. If the prepro region at the N terminus of the precursor was removed, then the putative molecular weight of the resulting protein was

TABLE 3. Relative activities of rSapSh and subtilisin Carlsberg toward synthetic substrates

Substrate	Relative activity (%)	
	rSapSh	Subtilisin Carlsberg
<i>N</i> -Succinyl-Ala-Ala-Pro-Phe- <i>p</i> Na ^a	100	100
<i>N</i> -Succinyl-Ala-Ala-Pro-Leu- <i>p</i> Na	33	75
<i>N</i> -Succinyl-Ala-Ala-Pro-Asp- <i>p</i> Na	0	0
<i>N</i> -Succinyl-Ala-Ala-Val-Ala- <i>p</i> Na	0	0.9
<i>N</i> -Succinyl-Ala-Ala-Ala- <i>p</i> Na	0	0.5
<i>N</i> -Succinyl-Phe- <i>p</i> Na	0	0

^a *p*Na, *p*-nitroanilide.

69,053, which is much larger than 44,000. Therefore, the C-terminal region of the precursor protein was digested to produce rSapSh. Processing of the C-terminal region is closely related to the secretion of proteases in *Vibrio* strains (6, 21). Therefore, the C-terminal region of SapSh probably participates in secretion.

Effect of pH on protease activity. Purified rSapSh exhibited the highest levels of activity with both azocasein and synthetic peptides at pH 9.0 (Fig. 4). The activity at pH 11.0 was more than 80% of the activity at pH 9.0, a finding similar to the findings obtained with subtilisins Carlsberg and BPN' (20). Thus, rSapSh is an alkaline protease.

Substrate specificity. The side chain specificity of rSapSh for synthetic peptides was similar to that of subtilisin Carlsberg (Table 3); proline at position P2 is effective, and an aromatic residue rather than an aliphatic residue at position P1 is also requested.

Stability. rSapSh and subtilisin Carlsberg were incubated at various temperatures for 15, 30, 45, and 60 min, and changes in the activities of the enzymes during incubation were monitored. rSapSh was less stable than subtilisin Carlsberg; rSapSh was almost inactivated by incubation at 60°C for 15 min, but subtilisin Carlsberg kept about 30% of its original activity under the same conditions (Fig. 5A). Moreover, rSapSh was found to be more susceptible to high concentrations of urea than subtilisin Carlsberg was; rSapSh lost 70% of its activity in the presence of 2 M urea within 30 min, whereas subtilisin Carlsberg exhibited 50% of its original activity under the same conditions (Fig. 5B). Therefore, rSapSh was much less stable than subtilisin Carlsberg, probably because of the high structural flexibility of rSapSh (11).

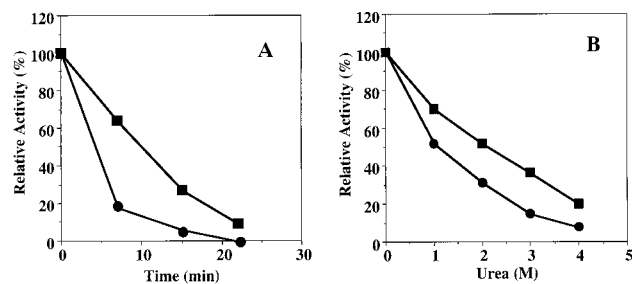


FIG. 5. (A) Thermal stabilities of rSapSh (●) and subtilisin Carlsberg (■). The enzymes were incubated at 60°C in 50 mM Tris-HCl buffer (pH 9.0) supplemented with 2 mM CaCl₂ for different periods of time, and the residual activities were determined at 60°C with AAPF as the substrate. (B) Denaturation of rSapSh (●) and subtilisin Carlsberg (■) by urea. The enzymes were incubated with various concentrations of urea for 30 min, and then the reactions were started by adding AAPF.

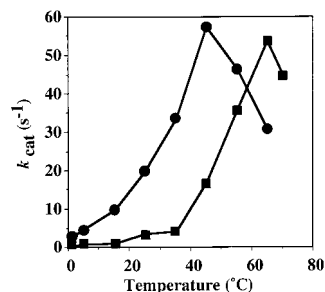


FIG. 6. Effect of temperature on the activities of rSapSh (●) and subtilisin Carlsberg (■) toward azocasein.

Effect of temperature on enzyme activity. The temperature dependence of the proteolytic activity of rSapSh was compared with that of subtilisin Carlsberg (Fig. 6). The optimal temperature for rSapSh activity was about 20°C lower than the optimal temperature for subtilisin Carlsberg activity. The specific activity of rSapSh toward azocasein at temperatures ranging from 5 to 15°C was five times higher than that of subtilisin Carlsberg. However, the specific activities of the two enzymes toward AAPF were similar, although rSapSh exhibited a k_{cat}/K_m value that was 80% lower than the k_{cat}/K_m value of subtilisin Carlsberg. The activation energies (E_a) of reactions catalyzed by enzymes from cold-adapted microorganisms are usually lower than the activation energies of reactions catalyzed by the corresponding enzymes from their mesophilic counterparts (11). The activation energy of the reaction catalyzed by rSapSh was determined from an Arrhenius plot of the values shown in Fig. 6. The thermodynamic activation parameters at 15°C were calculated by using the following equations: $\Delta G^* = \Delta H^* - T\Delta S^*$; $\Delta H^* = E_a - RT$; and $\Delta S^* = 2.303 R(\log k_{cat} - 10.753 - \log T + E_a/2.303 RT)$. The values for rSapSh were compared with the values for subtilisin Carlsberg, as follows: for rSapSh, $k_{cat} = 10.4 s^{-1}$, $E_a = 41.6 kJ mol^{-1}$, $\Delta G^* = 64.7 kJ mol^{-1}$, $\Delta H^* = 39.2 kJ mol^{-1}$, and $\Delta S^* = -88.7 J mol^{-1}K^{-1}$; for subtilisin Carlsberg, $k_{cat} = 1.27 s^{-1}$, $E_a = 57.5 kJ mol^{-1}$, $\Delta G^* = 70.3 kJ mol^{-1}$, $\Delta H^* = 55.1 kJ mol^{-1}$, and $\Delta S^* = -52.9 J mol^{-1}K^{-1}$. Thus, the contributions of the enthalpy and entropy terms to the ΔG^* values in rSapSh were also significantly different from the contributions in subtilisin Carlsberg. It is reasonable to assume that the activated complex of rSapSh is formed through the smaller entropy change and smaller heat content compared with the entropy change and heat content of subtilisin Carlsberg. rSapSh probably has a much more flexible structure than subtilisin Carlsberg has.

The genes of cold-active proteases have been cloned from two *Bacillus* strains, strains TA39 (23) and TA41 (3). Both of these proteases are members of the subtilisin subfamily that exhibit high levels of sequence similarity to other members of the family. Three-dimensional structural models of these two enzymes have been constructed on the basis of known subtilisin structures, and these models have been consistent with the structural features expected for cold-active enzymes, including a large hydrophilic surface, few salt bridges, and few aromatic-aromatic interactions (4, 23). In order to obtain a structural model of rSapSh on the basis of known subtilisin structures, prediction of the secondary structures in the long region between His-65 and Ser-369 with little sequence similarity would be a crucial step. We are planning to develop a method for large-scale preparation of rSapSh in order to crystallize this protein for X-ray analysis.

REFERENCES

- Bowman, J. P., S. A. McCommon, D. S. Nichols, J. H. Skerratt, S. M. Rea, P. D. Nichols, and T. A. McMeekin. 1997. *Shewanella gelidimarina* sp. nov. and *Shewanella frigidimarina* sp. nov., novel Antarctic species with the ability to produce eicosapentaenoic acid (20:5 omega 3) and grow anaerobically by dissimilatory Fe(III) reduction. *Int. J. Syst. Bacteriol.* **47**:1040–1047.
- Bruinenberg, P. G., P. Doesburg, A. C. Altling, F. A. Exterkate, W. M. de Vos, and R. J. Siezen. 1994. Evidence for a large dispensable segment in the subtilisin-like catalytic domain of the *Lactococcus lactis* cell-envelope proteinase. *Protein Eng.* **7**:991–996.
- Davail, S., G. Feller, E. Narinx, and C. Gerday. 1992. Sequence of the subtilisin-encoding gene from an antarctic psychrotroph *Bacillus* TA41. *Gene* **119**:143–144.
- Davail, S., G. Feller, E. Narinx, and C. Gerday. 1994. Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from antarctic psychrophile *Bacillus* TA41. *J. Biol. Chem.* **269**:17448–17453.
- David, V. A., A. H. Deutch, A. Sloma, D. Pawlyk, A. Ally, and D. R. Durham. 1992. Cloning, sequencing and expression of the gene encoding the extracellular neutral protease, vibriolysin, of *Vibrio proteolyticus*. *Gene* **112**:107–112.
- 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.
- DelMar, E. G., C. Largman, J. W. Brodrick, and M. C. Geokas. 1979. A sensitive new substrate for chymotrypsin. *Anal. Biochem.* **99**:316–320.
- Feller, G., M. Thiry, J. L. Arpingy, M. Mergeay, and C. Gerday. 1990. Lipases from psychrotrophic antarctic bacteria. *FEMS Microbiol. Lett.* **66**:239–244.
- Feller, G., F. Payan, F. Theys, M. Qian, R. Haser, and C. Gerday. 1994. Stability and structural analysis of α -amylase from the antarctic psychrophile *Alteromonas haloplanktis* A23. *Eur. J. Biochem.* **222**:441–447.
- Feller, G., E. Narinx, J. L. Arpingy, M. Aittaleb, E. Baise, S. Genicot, and C. Gerday. 1996. Enzymes from psychrophilic organisms. *FEMS Microbiol. Rev.* **18**:189–202.
- Gerday, C., M. Aittaleb, J. L. Arpingy, E. Baise, J.-P. Chessa, G. Garsoux, I. Petrescu, and G. Feller. 1997. Psychrophilic enzymes: a thermodynamic challenge. *Biochim. Biophys. Acta* **1342**:119–131.
- Germain, D., F. Dumas, T. Vernet, Y. Bourbonnais, D. Y. Thomas, and G. Boileau. 1992. The pro-region of the Kex2 endoprotease of *Saccharomyces cerevisiae* is removed by self-processing. *FEBS Lett.* **299**:283–286.
- Hase, C. C., and R. A. Finkelstein. 1991. Cloning and nucleotide sequence of the *Vibrio cholerae* hemagglutinin/protease (HA/protease) gene and construction of an HA/protease-negative strain. *J. Bacteriol.* **173**:3311–3317.
- Ikai, A. 1980. Thermostability and aliphatic index of globular proteins. *J. Biochem.* **88**:1895–1898.
- Jackman, D. M., F. M. Bartlett, and T. R. Patel. 1983. Heat-stable proteases from psychrotrophic pseudomonads: comparison of immunological properties. *Appl. Environ. Microbiol.* **46**:6–12.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaing the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105–132.
- Leduc, R., S. S. Molloy, B. A. Thorne, and G. Thomas. 1992. Activation of human furin precursor processing endoprotease occurs by an intramolecular autoproteolytic cleavage. *J. Biol. Chem.* **267**:14304–14308.
- Lory, S. 1992. Determinants of extracellular protein secretion in gram-negative bacteria. *J. Bacteriol.* **174**:3423–3428.
- Maidak, B. L., G. J. Olsen, N. Larsen, R. Overbeek, M. J. McCaughey, and C. R. Woese. 1997. The RDP (Ribosomal Database Project). *Nucleic Acids Res.* **25**:109–111.
- Markland, F. S., and E. L. Smith. 1971. Subtilisins: primary structure, chemical and physical properties, p. 562–608. *In* P. D. Boyer (ed.), *The enzymes*, 3rd ed., vol. 3. Academic Press, New York, N.Y.
- Milton, D. L., A. Norqvist, and H. Wolf-Watz. 1992. Cloning of a metallo-protease gene involved in the virulence mechanism of *Vibrio anguillarum*. *J. Bacteriol.* **174**:7235–7244.
- Morita, R. Y. 1975. Psychrophilic bacteria. *Bacteriol. Rev.* **39**:144–167.
- Narinx, E., E. Baise, and C. Gerday. 1997. Subtilisin from psychrophilic antarctic bacteria: characterization and site-directed mutagenesis of residues possibly involved in the adaptation to cold. *Protein Eng.* **10**:1271–1279.
- Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**:1–6.
- Power, S. D., R. M. Adams, and J. A. Wells. 1986. Secretion and autoproteolytic maturation of subtilisin. *Proc. Natl. Acad. Sci. USA* **83**:3096–3100.
- 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.
- Seizen, R. J., W. M. de Vos, J. A. M. Leunissen, and B. W. Dijkstra. 1991. Homology modelling and protein engineering strategy of subtilisins, the family of subtilisin-like serine proteinases. *Protein Eng.* **4**:719–737.
- Seizen, R. J., and J. A. Leunissen. 1997. Subtilisins: the superfamily of subtilisin-like serine proteases. *Protein Sci.* **6**:501–523.
- Stepanov, V. M., and G. N. Rudenskaya. 1983. Proteinase affinity chroma-

- tography on bacitracin-Sepharose. *J. Appl. Biochem.* **5**:420–428.
30. **Van Heeke, G., S. Denslow, J. R. Watkins, K. J. Wilson, and F. W. Wagner.** 1992. Cloning and nucleotide sequence of the *Vibrio proteolyticus* aminopeptidase gene. *Biochim. Biophys. Acta* **1131**:337–340.
31. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14**:4683–4690.
32. **Voorhorst, W. G., A. Warner, W. M. de Vos, and R. J. Siezen.** 1997. Homology modelling of two subtilisin-like proteases from the hyperthermophilic archaea *Pyrococcus furiosus* and *Thermococcus stetteri*. *Protein Eng.* **10**:905–914.
33. **Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane.** 1991. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **173**:697–703.
34. **Wu, C., and S. Shivakumar.** 1994. Back-propagation and counter-propagation neural networks for phylogenetic classification of ribosomal RNA sequences. *Nucleic Acids Res.* **22**:4291–4299.