Biochimica et Biophysica Acta, 1121 (1992) 221–228

BBAPRO 34215

Purification, characterization and gene cloning of a novel glutamic acid-specific endopeptidase from *Staphylococcus aureus* ATCC 12600

Kazumasa Yoshikawa, Hiroshige Tsuzuki, Takashi Fujiwara, Etsuo Nakamura, Hideo Iwamoto, Koichi Matsumoto, Masaru Shin, Nobuo Yoshida and Hiroshi Teraoka

Shionogi Research Laboratories, Shionogi and Co., Ltd., Fukushima-ku, Osaka (Japan)

(Received 25 November 1991)

Key words: Gene cloning; Amino acid sequence; Serine proteinase; Substrate specificity; (*S. Aureus*)

Twenty strains of *Staphylococcus aureus* from ATCC type cultures and strains found in clinical studies were cultivated, and their endopeptidase activity specific for glutamic acid was surveyed using benzyloxycarbonyl-Phe-Leu-Glu-*p*-nitroanilide (Z-Phe-Leu-Glu-pNA) as a substrate. The activity was found in two of the strains, ATCC 12600 and ATCC 25923. A glutamic acid-specific proteinase, which we propose to call SPase, was purified from the culture filtrate of *S. aureus* strain ATCC 12600 by a series of column chromatographies on DEAE-Sepharose twice and on Sephacryl S-200. A single band was observed on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of the purified SPase. The molecular weight of the proteinase was estimated to be 34000 by SDS-PAGE. When synthetic peptides and oxidized insulin B-chain were used as substrates, SPase showed the same substrate specificity as V8 proteinase, EC 3.4.21.9, which specifically cleaves peptide bonds on the C-terminal side of glutamic acid and aspartic acid. Examination with *p*-nitroanilides of glutamic acid and aspartic acid as substrates, however, revealed that both proteinases are highly specific for a glutamyl bond in comparison with an aspartyl bond. To elucidate the complete primary structure of SPase, its gene was cloned from genomic DNA of *S. aureus* ATCC 12600, and the nucleotide sequence was determined. Taking the amino acid sequence of SPase from the NH_2-terminus to the 27th residue into consideration, the clones encode a mature peptide of 289 amino acids, which follows a prepropeptide of 68 residues. SPase was confirmed to be a novel endopeptidase specific for glutamic acid, being different from V8 proteinase which consists of 268 amino acids.

Introduction

About 67% of the strains of *S. aureus*, which is pathogenic, are postulated to produce some kind of proteinase [1]. Among them, the glutamic acid-specific endopeptidase produced by strain V8, reported by Drapeau et al. [2], is noted for its high specificity for glutamic acid and is widely employed as a reagent for structure determination in the field of protein chemistry. Recently, the V8 proteinase has been subjected to nucleotide sequencing [3] and also X-ray analysis [4].

Two other kinds of glutamic acid-specific endopeptidase have been reported, one present in a small amount in a commercially available Pronase from *Streptomyces griseus* (Kaken Seiyaku, Tokyo) [5] and another present in a small amount in Protease Type XVI from *Bacillus subtilis* (Sigma) [6]. Although the main specific cleavage point is the carboxyl side of a glutamic acid residue, the carboxyl side of acidic amino acid residues, such as cysteic acid and aspartic acid, is also cleaved to a small extent with respect to the proteinaceous substrate. The primary structure of these enzymes has not yet been reported.

As a result of screening on some kinds of *S. aureus* for the purpose of finding a novel glutamic acid endopeptidase, an enzyme which cleaves the carboxyl side of a glutamic acid residue in the same manner as V8 proteinase was found (the enzyme is designated as SPase hereafter). Subsequently, a gene of SPase was cloned using a PCR product as the probe, which was amplified with oligonucleotides around the active site

Abbreviations: ATCC, American Type Culture Collection; BSA, bovine serum albumin; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate.

Correspondence: H. Teraoka, Shionogi Research Laboratories, Shionogi and Co., Ltd., 5-12-4 Sagisu, Fukushima-ku, Osaka 553, Japan.

of V8 proteinase as primers, and nucleotide sequencing was performed. SPase was confirmed to be a novel type of glutamic acid-specific endopeptidase having a different protein structure from that of V8 proteinase.

Materials and Methods

Materials. Enzymes and reagents were obtained from the sources indicated: V8 proteinase from *S. aureus* strain V8 and oxidized bovine insulin B-chain (Oxd-insulin B-chain) from Sigma Chemicals; DEAE-Sepharose and Sephacryl S-200 HR from Pharmacia LKB Biotechnology; benzyloxycarbonyl-Phe-Leu-Glu *p*-nitroanilide (*Z*-Phe-Leu-Glu-pNA) from Boehringer Mannheim Yamanouchi; *t*-butoxycarbonyl-Ala-Ala-Asp *p*-nitroanilide (Boc-Ala-Ala-Asp-pNA) from Bachem; Protein Assay Kit from Bio-Rad; Protein Molecular Weight Marker "RAINBOW™" from Amersham; HPLC Column "Protein C4" from Vydac; restriction endonucleases, M13 sequencing kits for dideoxy chain termination method and the enzymes used for DNA manipulation from Takara Shuzo; DNA amplification reagent (PCR) kit from Perkin Elmer Cetus; nylon membrane (Zeta-Probe) from Bio-Rad; radiolabelled deoxyribonucleotide from Amersham. Z-X-Leu-Ala-Ala (X: Glu, Asp, Ala, Leu, Phe, Tyr and Lys) and Boc-Ala-Ala-Glu-pNA were synthesized by us. *S. aureus* ATCC 12600 was used as a producer of SPase and *Escherichia coli* HB101 as the host for cloning of SPase gene.

Assay of enzymatic activity. All peptide substrates were dissolved in *N*,*N*-dimethylformamide (DMF) to obtain a final concentration of 10 mM and make a substrate stock solution. With respect to *Z*-Phe-Leu-Glu-pNA, the reaction was started by adding an appropriate amount of an enzyme solution to a solution composed of 0.2 mM *Z*-Phe-Leu-Glu-pNA, 50 mM NH_4HCO_3 and 5 mM EDTA (pH 7.8), to make a final volume of 1.0 ml. After 10 min of reaction at 23°C, the absorbance of liberated *p*-nitroaniline was measured at 410 nm. As for the other *p*-nitroanilides, the reaction was carried out with 1.0 mM *p*-nitroanilide in 0.1 M Tris-HCl, 2 mM $CaCl_2$ and 10% DMF at pH 7.8 and 37°C. When the substrate was *Z*-Phe-Leu-Glu-pNA, one unit of the specific activity of SPase was expressed as the enzymatic activity giving an absorbance of 1.0 under the above conditions. With regard to *Z*-X-Leu-Ala-Ala, the reaction was performed with 1 mM of the substrate in 0.1 mM Tris-HCl, 2 mM $CaCl_2$ and 10% DMF at pH 8.0 and 37°C. The cleaved peptide products, *Z*-X and Leu-Ala-Ala, were observed by reversed phase HPLC.

To measure the activity to the Oxd-insulin B-chain, the chain was dissolved in a buffer composed of 50 mM NH_4HCO_3 and 5 mM EDTA (pH 7.8), to obtain a final concentration of 1.0 mg/ml. To this solution the enzyme was added to equal 1/20 of the substrate (S/E = 20, weight by weight), and the reaction was allowed to proceed at 25°C for 18 h. The resultant solution was analyzed by HPLC using a Vydac Protein C4 Column (4.6 × 250 mm). After the column had been equilibrated with a mixture of water and acetonitrile (4:1) containing 0.1% trifluoroacetic acid, the above sample solution was loaded onto the column, which was then washed with the same mixture. Subsequently, the acetonitrile concentration was increased from 20% to 45% over 30 min. Each degradation product was collected from the effluent corresponding to each peak in the chromatogram and identified by amino acid analysis.

Measurement of protein concentration. Protein concentrations were determined by a Protein Assay Kit using BSA as a standard.

Screening and cultivation. After the strains had been cultivated in DM + medium (1.0% glucose, 0.5% soluble starch, 1.0% polypeptone, 0.3% yeast extract, 0.1% $MgSO_4 \cdot 7H_2O$, 0.3% NaCl and 0.3% $CaCO_3$, pH 7.0) at 37°C for 18 h, the culture was centrifuged, and the enzymatic activity of the supernatant was determined using *Z*-Phe-Leu-Glu-pNA as a substrate.

Purification. After *S. aureus* ATCC 12600 had been cultivated in DM + medium at 37°C for 18 h, the culture was sterilized by filtration, concentrated in an ultrafiltration module (cutoff 20000 Da, Nitto) and diluted with distilled water to obtain an electroconductivity of 3.52 ms/cm. The resulting solution was loaded onto a DEAE-Sepharose column equilibrated with 10 mM Tris-HCl buffer (pH 7.5). The column was then washed with 10 mM Tris-HCl (pH 7.5), containing 0.1 M NaCl and eluted with 10 mM Tris-HCl buffer (pH 7.5), containing 0.2 M NaCl. The active fractions were pooled, dialyzed against an 80% saturated $(NH_4)_2SO_4$ aqueous solution and allowed to stand for 24 h. After centrifugation, the supernatant was discarded by decantation. The collected precipitate was dissolved in 10 mM Tris-HCl buffer (pH 7.5), dialyzed against the same buffer and loaded onto a DEAE-Sepharose column equilibrated with the same buffer. After the column had been washed and eluted in the same manner as the first DEAE-Sepharose column chromatography, the active fractions were pooled and lyophilized. When necessary, the lyophilizate was gel filtrated on Sephacryl S-200 HR with 50 mM Tris-HCl buffer (pH 7.5), containing 0.1 M KCl, and made into a pure sample of SPase.

SDS-PAGE and isoelectric focusing. SDS-PAGE was performed on a 15% gel as described by Laemmli [7] using "RAINBOW" (Amersham) as the molecular weight marker to determine the molecular weight and the purity. Isoelectric focusing was performed in the range of p*I* 3.0 to 10.0 using PhastSystem and Pharmalite (Pharmacia) to determine the isoelectric point.

Amino acid analysis and amino acid sequence determination. After hydrolysis in 4 M methanesulfonic acid containing 0.2% 3-(2-aminoethyl)indole at 110°C for 24 h, the amino acid composition was determined with an amino acid analyzer (Model 835 Hitachi). Automated Edman degradation was performed with a model 477A Applied Biosystems gas phase sequencer equipped with a 120A PTH-amino acid analyzer.

Synthesis of oligonucleotide primers. Oligonucleotide primers were synthesized on a Gene Assembler Plus DNA Synthesizer (Pharmacia LKB Biotechnology), and purified on 20% polyacrylamide/7 M urea gels. The sequences of oligonucleotide primers used in this study are shown below. Primers 5'-ATGAAAGGTAAATTTTTAAA-3' (No. 1) and 5'-AGCTGCATCTGGATTGTCTG-3' (No. 2) are the sense and the anti-sense primers specific for the NH_2- and the C-terminal regions of V8 proteinase, respectively, primers 5'-GATACAACGAATGGTCATTA-3' (No. 3) and 5'-TACAGGTGAACCTGAATTAC-3' (No. 4) are the sense and the anti-sense primers for the His-18 and the Ser-169 active site regions of V8 proteinase, respectively, and primers 5'-GTTTTCCCAGTCACGAC-3' (No. 5) and 5'-CAGGAAACAGCTATGAC-3' (No. 6) are the forward and the reverse primers specific for M13 sequencing, respectively.

Amplification of the DNA fragment encoding SPase. PCR was performed for 30 cycles on a DNA thermal cycler (Perkin Elmer Cetus) using a DNA amplification reagent kit. The thermal cycle of the PCR procedures included denaturation at 94°C for 30 s, annealing at 50°C for 1 min, and extension at 72°C for 1 min. PCR products were analyzed by 1 or 2% agarose gel electrophoresis and subjected to ethidium bromide staining or DNA blot analysis.

Screening of the genomic and size-selected DNA library. General techniques including transformation of *E. coli* cells, M13 DNA isolation and agarose gel electrophoresis were described by Maniatis et al. [8]. The DNA probe amplified by PCR was labeled with [α-^{32}P]dCTP using a random primer labeling system (Amersham). Genomic DNA of the *S. aureus* ATCC 12600 was prepared as described by Maniatis et al. [8]. Southern hybridization between restriction enzyme-digested genomic DNA and ^{32}P-labeled DNA probe specific for SPase was carried out according to the method of Southern [9] with the following modification. The temperature for hybridization in 1 mM EDTA/0.5 M NaH_2PO_4 (pH 7.2)/7% SDS/1% BSA was 65°C, and the nylon membrane was washed with 15 mM NaCl/1.5 mM sodium citrate and 0.1% SDS at 50°C. Distinct 3.5-kb and 3.7-kb bands hybridizing to the probe were detected in *Eco*RI and *Hin*dIII digests, respectively. From an agarose gel slice containing *Eco*RI digests the 3.5-kb fragment was recovered and ligated with M13mp10 DNA which had been successively treated with *Eco*RI and bacterial alkaline phosphatase. The ligated mixture was introduced by transformation into *E. coli* HB101, and non-blue transformants were selected on L-broth containing X-gal (5-bromo-4-chloro-3-indol-*b*-D-galactopyranoside). Colony hybridization with the SPase-specific DNA probe yielded transformants showing positive hybridization. Recombinant DNA was prepared from the positive clone, and PCR analysis together with the Southern hybridization experiments identified the region and orientation of the SPase gene.

DNA sequencing. The DNA fragment amplified by PCR was directly sequenced as described by Gibbs et al. [10]. Appropriate DNA fragments were inserted into M13mp10, and nucleotide sequencing was performed by the chain termination method as described previously [11].

Nucleotide sequence accession number. The nucleotide sequence of SPase gene (tentatively named V8-like proteinase gene) has been submitted to the DDBJ and has the accession number D00730.

Results

Screening of glutamic acid-specific endopeptidase

Twenty strains of *S. aureus* from the type cultures and the strains found in clinical studies (Shionogi Research Laboratories) were cultivated in 10 ml of medium in test tubes, and their endopeptidase activity specific for glutamic acid was determined using *Z*-Phe-Leu-Glu-pNA as a substrate. As a result, the activity was found in two of the strains, ATCC 12600 and ATCC 25923. The yields of SPase from the two strains were 35 to 50 μg/ml, calculated on the basis of the specific activity of V8 proteinase to *Z*-Phe-Leu-Glu-pNA. *S. aureus* ATCC 12600 was selected and subjected to subsequent experiments because of its higher yield than that from ATCC 25923.

Purification of SPase

After *S. aureus* ATCC 12600 had been cultivated in DM + medium at 37°C for 18 h, the culture was sterilized by filtration (95 l) and concentrated in an ultrafiltration module (cutoff 20000 Da, Nitto) to a volume of 19 l. The solution was diluted with distilled water to obtain a final volume of 38 l and an electroconductivity of 3.52 ms/cm. The resulting solution was loaded onto a DEAE-Sepharose column (14 × 39 cm) equilibrated with 10 mM Tris-HCl buffer (pH 7.5). The column was then washed with 11.0 l of 10 mM Tris-HCl (pH 7.5), containing 0.1 M NaCl and eluted with 10 mM Tris-HCl buffer (pH 7.5), containing 0.2 M NaCl. The active fractions (5.5 l) were pooled, dialyzed against an 80% saturated $(NH_4)_2SO_4$ aqueous solution and allowed to stand for 24 h. After centrifugation, the supernatant was discarded by decantation. The col-

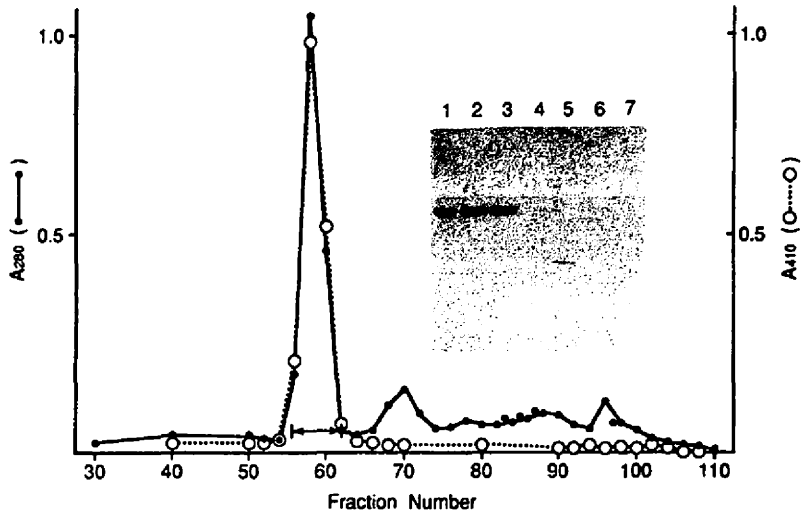

Fig. 1. Effluent profile of the SPase on Sephacryl S-200 HR. The column (2.5×98 cm) was equilibrated with 50 mM Tris-HCl (pH 7.5), containing 0.1 M KCl. The unit fraction volume was 4.67 ml, and the flow rate 24.5 l/h. The lyophilized product obtained in the second DEAE-Sepharose chromatography was dissolved in 10 ml of 50 mM Tris-HCl (pH 7.5), containing 0.1 M KCl, and loaded onto the column. The active fractions indicated by bars were collected. Absorbance at 280 nm is shown by ●–●–● and the absorbance at 410 nm which represents the activity degrading Z-Phe-Leu-Glu-pNA, by O·····O. Some fractions were analyzed by SDS-PAGE: Lane 1, Fr. 56; 2, Fr. 60; 3, Fr. 62; 4, Fr. 64; 5, Fr. 70; 6, Fr. 88; 7, Fr. 96.

lected precipitate was dissolved in 10 mM Tris-HCl buffer (pH 7.5), dialyzed against the same buffer (20 l) and loaded onto a DEAE-Sepharose column equilibrated with the same buffer. The column was washed in the same manner as the first DEAE-Sepharose column chromatography and eluted with 10 mM Tris-HCl buffer (pH 7.5), containing 0.2 M NaCl. The active fractions were pooled (1.8 l) and lyophilized.

In order to further purify the SPase thus obtained, gel filtration on Sephacryl S-200 HR was performed. The chromatographic pattern and the results of SDS-PAGE are shown in Fig. 1. The fractions showing a single band on SDS-PAGE, from fraction 56 to 61, were pooled. Fraction 62 gave the same band as the above fractions, together with a faint minor band. The latter might be an auto-digestion product upon the chromatographies in high protein concentration. After dialysis and lyophilization, the purified enzyme was obtained and found at a molecular weight of 34000 as a single band, and its purity was confirmed. The isoelectric point was about 4.0 for the proteinase. The results of the above purification procedure are summarized in Table I. The yield of the enzymatic activity was as high as 69.7%, and purification could be achieved by rather a simple method. The specific activity of SPase determined using Z-Phe-Leu-Glu-pNA as a substrate was 260 U/mg similar to that of V8 proteinase, which was 270 U/mg.

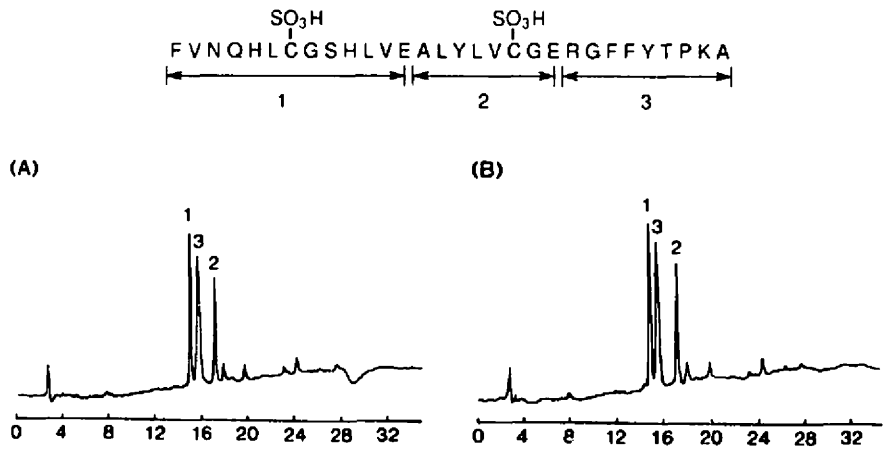

Fig. 2. Digestion of oxidized insulin B-chain by SPase. Left, digestion of oxidized insulin B-chain by SPase. Right, digestion of oxidized insulin B-chain by V8 proteinase. The conditions of the reaction and the analysis are described in the text.

Characterization of SPase

Substrate specificity. SPase showed an identical peptide map by HPLC with that of V8 proteinase in the digestion of the Oxd-insulin B-chain (Fig. 2). The results of identification of each degradation product showed that the -Glu-X- bonds were selectively cleaved. Peak 1 was identified as the peptide 1–13, peak 2 as 14–21 and peak 3 as 22–30. The bond to cysteinic acid, an acidic amino acid like glutamic acid, was not cleaved, which indicates the high specificity of SPase like that of V8 proteinase.

The enzyme activity of SPase was examined using seven tetrapeptides, Z-X-Leu-Ala-Ala (X: Glu, Asp, Ala, Leu, Phe, Tyr and Lys) as substrates together with V8 proteinase. Both enzymes cleaved the peptide bond of Glu-Leu only when X was Glu, and for the other six peptides, no hydrolytic product was found even when X was Asp (data not shown).

The activity of both of the enzymes was examined using Boc-Ala-Ala-Y-pNA (Y: Glu and Asp) as substrates as well. Both enzymes showed very high reactivity to Boc-Ala-Ala-Glu-pNA, which was observed as activities of approx. 2000 nmol/mg per min, whereas those of approx. 50 nmol/mg per min were found for Boc-Ala-Ala-Asp-pNA, and so the yield of *p*-nitroaniline was a few percent of that from Boc-Ala-Ala-Glu-pNA. These results indicate that SPase has a similar reactivity to that of V8 proteinase.

Amino acid composition and NH_2-terminal amino acid sequence. The amino acid composition of SPase was analyzed and compared with that of V8 proteinase [3] and those deduced from the nucleotide sequence of cDNA described hereafter for comparison. The amino acid composition of SPase agreed well with that deduced from the cDNA sequence. However, SPase and V8 proteinase differed in the composition of Asx and Pro, which suggested a difference in the amino acid sequence.

The amino acid sequence of SPase from the NH_2-terminus to the 27th was examined. Although SPase had a different total amino acid composition from that of V8 proteinase reported by Carmona and Gray [3], the amino acid sequences from the NH_2-terminus to the 27th of both proteinases agreed well, which suggests some difference in the region after the 27th amino acid residue. On the basis of the above results, gene cloning was performed.

TABLE I

Purification of a glutamic acid-specific endopeptidase from Staphylococcus aureus ATCC 12600

Step	Volume (l)	Total protein (g) [a]	Specific activity (units/mg)	Activity recovery (%)
Broth	95	16.70	95	100.0
Concentration	19	12.95	114	92.7
DEAE-Sepharose (1st)	3.5	6.52	223	91.6
Ammonium sulfate ppt.	2.0	5.86	242	89.6
DEAE-Sepharose (2nd)	1.8	5.28	240	79.9
Sephacryl S-200	2.0	4.25	260	69.7

[a] Determined by Bio-Rad Protein Assay Kit using BSA as a standard.

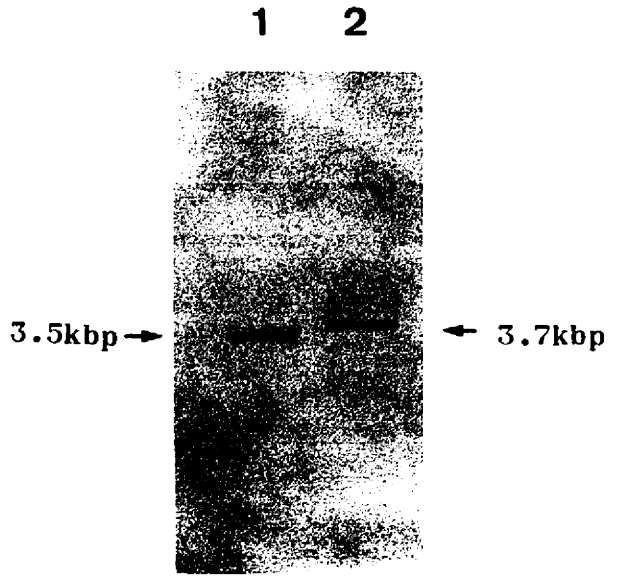

Fig. 3. Southern blot analysis of *Staphylococcus aureus* ATCC 12600 genomic DNA. Genomic DNA from *S. aureus* ATCC 12600 was subjected to digestion with *Eco*RI or *Hin*dIII, followed by electrophoresis on a 1% agarose gel, and hybridization with ^{32}P-labeled DNA probe specific for SPase as described in Materials and Methods. Lanes: 1, *Eco*RI digest; 2, *Hin*dIII digest.

Cloning of SPase gene

Since SPase seemed to be very similar to V8 proteinase, we first tried to amplify an internal SPase gene fragment by PCR using the genomic DNA of *S. aureus* ATCC 12600 producing SPase as a template with the four primers, which were designed on the basis of the nucleotide sequences of the previously published V8 proteinase gene [3]. A DNA fragment of about 480 bp between two regions encoding putative active sites His-18 and Ser-169 was successfully amplified with primers No. 3 and No. 4, but not the DNA fragment covering the NH_2-terminal to C-terminal regions with primers No. 1 and No. 2 (data not shown). These results suggest that the nucleotide sequence encoding two active site regions of SPase and V8 proteinase is highly conserved. The direct DNA sequence analysis of 480-bp PCR product revealed that the deduced amino acid sequence was very similar to that of V8 proteinase and contained the partial amino acid sequences of SPase that were determined by a protein chemical method.

Southern blot hybridization of *Eco*RI- and *Hin*dIII-digested genomic DNAs of *S. aureus* ATCC

12600 using ^{32}P-labeled 480-bp PCR product as a probe gave one positive band with 3.5 kb and 3.7 kb, respectively (Fig. 3). The resulting 3.5-kb *Eco*RI fragment was purified by agarose gel electrophoresis, ligated with *Eco*RI-digested M13mp10, and introduced by transformation into *E. coli* HB101. The transformants which did not turn blue on an X-gal plate were screened by colony hybridization with the SPase gene-specific 480-bp probe. The PCR analysis of DNA from a positive clone with a combination of two pairs of primers (No. 5/No. 6 outside primers specific for M13 DNA and No. 3/No. 4 inside primers described above) revealed the orientation of the 3.5-kb *Eco*RI fragment and the fact that the whole amino acid sequence of SPase is contained within the fragment (data not shown).

Nucleotide sequence of the SPase gene

On the basis of the results described above, we determined the nucleotide sequence of SPase gene involved in the 3.5-kb *Eco*RI fragment. The whole nucleotide sequence was determined by the chain termination method [11] in both orientations. Fig. 4 shows the 1290-bp nucleotide sequence and the amino acid sequence derived therefrom. The nucleotide sequence of SPase gene involves an open reading frame of 1071 nucleotides coding for a protein of 357 amino acids. Based on the NH_2-terminal amino acid sequence (1–27)

Nucleotide sequence and the deduced amino acid sequence of SPase

```
                                                                        SD
   1  AAA AAT TTT TAG TGT TAT ATT TAA CTT GTA AAT AAA TTT TTT GGA GGT TTT TAG ATG AAA GGT AAA TTT TTA AAA GTT AGT TCT TTA TTC GTT GCA    96
                                                                                 Met Lys Gly Lys Phe Leu Lys Val Ser Ser Leu Phe Val Ala    14

  97  ACT TTG ACA ACA GCG ACA CTT GTG AGT TCT CCA GCA GCA AAT GCG TTA TCT TCA AAA GCT ATG GAC AAT CAT CCA CAA CAA ACG CAG ACA GAC AAA   192
  15  Thr Leu Thr Thr Ala Thr Leu Val Ser Ser Pro Ala Ala Asn Ala Leu Ser Ser Lys Ala Met Asp Asn His Pro Gln Gln Thr Gln Thr Asp Lys    46

 193  CAC CAA ACA CCT AAG ATT CAA AAA GGC GGT AAC CTT AAA CCA TTA GAA CAA CGT GAA CGC GCT AAT GTT ATA TTA CCA AAT AAC GAT CGT CAC CAA   288
  47  Gln Gln Thr Pro Lys Ile Gln Lys Gly Gly Asn Leu Lys Pro Leu Glu Gln Arg Glu Arg Ala Asn Val Ile Leu Pro Asn Asn Asp Arg His Gln    78

 289  ATC ACA GAT ACA ACG AAT GGT CAT TAT GCA CCT GTT ACT TAT ATT CAA GTT GAA GCA CCT ACT GGT ACA TTT ATT GCT TCT GGT GTA GTT GTA GGT   384
  79  Ile Thr Asp Thr Thr Asn Gly His Tyr Ala Pro Val Thr Tyr Ile Gln Val Glu Ala Pro Thr Gly Thr Phe Ile Ala Ser Gly Val Val Val Gly   110

 385  AAA GAT ACA CTT TTA ACA AAT AAA CAC GTC GTA GAT GCT ACG CAC GGT GAT CCT CAT GCT TTA AAA GCA TTC CCT TCT GCA ATT AAC CAA GAC AAT   480
 111  Lys Asp Thr Leu Leu Thr Asn Lys His Val Val Asp Ala Thr His Gly Asp Pro His Ala Leu Lys Ala Phe Pro Ser Ala Ile Asn Gln Asp Asn   142

 481  TAT CCT AAT GGT GGT TTC ACT GCT GAA CAA ATC ACT AAA TAT TCA GGC GAA GGT GAT TTA GCA ATC GTT AAA TTC TCC CCT AAT GAG CAA AAC AAA   576
 143  Tyr Pro Asn Gly Gly Phe Thr Ala Glu Gln Ile Thr Lys Tyr Ser Gly Glu Gly Asp Leu Ala Ile Val Lys Phe Ser Pro Asn Glu Gln Asn Lys   174

 577  CAT ATT GGC GAA GTA GTT AAA CCA GCA ACA ATG AGT AAT AAT GCT GAA ACA CAA GTT AAC CAA AAT ATT ACT GTA ACA GGA TAT CCT GGT GAT AAA   672
 175  His Ile Gly Glu Val Val Lys Pro Ala Thr Met Ser Asn Asn Ala Glu Thr Gln Val Asn Gln Asn Ile Thr Val Thr Gly Tyr Pro Gly Asp Lys   206

 673  CCT GTC GCA ACA ATG TGG GAA AGT AAA GGA AAA ATA ACG TAC TTA AAA GGT GAA GCA ATG CAA TAT GAT TTA AGT ACA ACT GGT GGT AAC TCA GGT   768
 207  Pro Val Ala Thr Met Trp Glu Ser Lys Gly Lys Ile Thr Tyr Leu Lys Gly Glu Ala Met Gln Tyr Asp Leu Ser Thr Thr Gly Gly Asn Ser Gly   238

 769  TCA CCT GTA TTT AAT GAA AAA AAT GAA GTC ATT GGC ATT CAT TGG GGT GGC GTT CCA AAT CAA TTT AAC GGT GCA GTA TTT ATT AAT GAA AAT GTA   864
 239  Ser Pro Val Phe Asn Glu Lys Asn Glu Val Ile Gly Ile His Trp Gly Gly Val Pro Asn Gln Phe Asn Gly Ala Val Phe Ile Asn Glu Asn Val   270

 865  CGC AAC TTC TTA AAA CAA AAT ATT GAA GAT ATC AAT TTC GCA AAT GAT GAC CAC CCT AAC AAC CCT GAT AAT CCA GAC AAT CCA AAT AAT CCG GAC   960
 271  Arg Asn Phe Leu Lys Gln Asn Ile Glu Asp Ile Asn Phe Ala Asn Asp Asp His Pro Asn Asn Pro Asp Asn Pro Asp Asn Pro Asn Asn Pro Asp   302

 961  AAT CCT AAC AAC CCT GAT AAC CCT AAC AAC CCT GAT AAT CCA GAC AAT CCT AAT AAT CCT GAT AAC CCT AAC AAC CCG GAC AAT CCA AAT AAC CCT  1056
 303  Asn Pro Asn Asn Pro Asp Asn Pro Asn Asn Pro Asp Asn Pro Asp Asn Pro Asn Asn Pro Asp Asn Pro Asn Asn Pro Asp Asn Pro Asn Asn Pro   334

1057  GAC CAA CCT AAC AAC CCA AAT AAC CCG GAC AAT GGC GAT AAC AAT AAT TCA GAC AAC CCT GAC GCT GCA TAA ACT TCA TCT TAA TTT GGA TTT AGT  1152
 335  Asp Gln Pro Asn Asn Pro Asn Asn Pro Asp Asn Gly Asp Asn Asn Asn Ser Asp Asn Pro Asp Ala Ala ***

1153  ATA TAG AAG CAA AAA CTT AAC AAC GCA AAA TAT TTT GAC TTA AAT GGA GGG TAT TAT ATG AAT AGT TCA TGT AAA ACT AGA GTA TTC AAT ATT ATA  1248

1249  AGC ATC ATA ATG GTT TCA ATG CTT ATT TTA TCA CTA GGC GC
```

Fig. 4. Nucleotide sequence and the deduced amino acid sequence of prepro-SPase. The deduced amino acid sequence is numbered starting with the translation initiation methionine. The cleavage site to give the mature enzyme is indicated by the arrow at position 68. The unusual repeated sequence (Pro-Asx-Asn or Pro-Asx-Gln) is underlined. Amino acid residues in rectangles designate those different in V8 proteinase. The putative ribosome binding site (GGAGG) is indicated by SD.

TABLE II

Differences in amino acid sequences between glutamic acid-specific endopeptidases from Staphylococcus aureus ATCC 12600 (SPase) and strain V8 (V8 proteinase)

Residue number [a]	SPase	V8 proteinase
−25	Thr (ACA)	Ser (TCA)
−24	Asp (GAC)	Ser (AGC)
−3	Arg (CGC)	His (CAC)
191	Gln (CAA)	Glu (GAA)
214	Asn (AAT)	His (CAT)
220	His (CAC)	Gln (CAA)
262	Asn (AAT)	Glu (GAA) (241)
268	Gln (CAA)	Asn (AAC) (247)
273	Asn (AAT)	Asp (GAT) (252)
(Pro-Asx-Asn or Pro-Asx-Glx)$_n$	$n = 19$	$n = 12$

[a] The amino acid sequence is numbered starting with the NH_2-terminal residue (Val) of the mature enzyme of SPase. The sequence of V8 proteinase was aligned using the program "DNASIS". The nucleotide substitutions causing amino acid substitutions are underlined.

for the enzyme as described above, the translational initiation site was deduced to be methionine at position 1 in Fig. 4, which was identical with that of V8 proteinase. A highly homologous sequence (GGAGG) with the *E. coli* ribosome-binding site [12] exists seven nucleotides upstream from the translational initiation site.

Discussion

We have described the purification and cloning of a glutamic acid-specific proteinase from *S. aureus* ATCC 12600. Our findings enable elucidation of the complete primary structure of the enzyme based on deduction from its nucleotide sequence.

Drapeau et al. [2] revealed that V8 proteinase specifically cleaves the carboxyl side of either aspartic acid or glutamic acid in a digestion reaction with V8 proteinase at pH 7.8 and 37°C for 18 h using ribonuclease as a substrate (S/E = 40, weight by weight). In order to examine the substrate specificity of SPase quantitatively, we measured its reactivity to Boc-Ala-Ala-Glu-pNA and Boc-Ala-Ala-Asp-pNA at pH 7.8 and 37°C together with V8 proteinase. The reactivity of both enzymes to *p*-nitroanilide of Asp-peptide substrate amounted only to a few percent of that of Glu-peptide substrate. In a preliminary study where parathyroid hormone (human, 1–34), a peptide hormone which consists of 34 amino acid residues and contains three glutamic acid residues and one aspartic acid residue, was hydrolyzed (pH 7.8, 37°C, S/E = 100, weight by weight), only 30% of the aspartyl bonds had been cleaved after 24 h of the reaction, whereas cleavage of the glutamyl bonds was completed in about 20 min. These results indicate that the specificity of SPase for glutamic acid is very high, but low for aspartic acid.

SPase is one of the extracellular proteinases synthesized and secreted by *S. aureus*. Seeing that the mature form of SPase has a molecular weight of 34000 with the 27-residue NH_2-terminal amino acid sequence we obtained from the purified enzyme, the nucleotide sequence analysis of SPase gene suggests that a product of the gene is initially synthesized as a 357-residue preproenzyme, followed by processing to a mature enzyme consisting of 289 amino acid residues. From the deduced amino acid sequence for the precursor of SPase, this preproenzyme has a typical bacterial signal sequence [13] including a basic charged sequence at the NH_2-terminus, followed by a long hydrophobic sequence with a putative signal peptide cleavage site on the C-terminal side of Ala-27 or Ala-29 residue (Fig. 4). Between the putative signal peptide and the mature SPase sequence there is a hydrophilic sequence consisting of about 40 amino acid residues, suggesting the occurrence of the pro-sequence.

As expected from the results of the protein chemical analyses described above, there was a highly homologous sequence between SPase and V8 proteinase; i.e., 88.7% homology in nucleotide sequences and 91.5% homology in amino acid sequences in the optimally matched alignment. Table II summarizes the differences in amino acid sequences between SPase and V8 proteinase. Three and six amino acid substitutions were observed in the putative preprosequence and the mature sequence, respectively. The nucleotide and the deduced amino acid sequences in the NH_2-terminal half were very homologous between SPase and V8 proteinase, whereas many substitutions of their sequences in the C-terminal half were observed in the two enzymes. The mature SPase consists of 289 amino acid residues which is 21 residues longer than that of V8 proteinase. This is due to the difference in the length of an unusual repeated sequence (Pro-Asx-Asn or Pro-Asx-Glx) existing in the vicinity of the C-terminus of SPase; i.e., the number of this repeated sequence is 12 in V8 proteinase but 19 in SPase (sequences underlined in Fig. 4). Preliminary analysis by PCR of genomic DNAs from the V8 proteinase-producing strain (*S. aureus* V8 strain), the SPase producer (*S. aureus* ATCC 12600) and another glutamic acid-specific proteinase producer (*S. aureus* ATCC 25923), revealed that there was no difference between *S. aureus* ATCC 12600 and ATCC 25923 in the length of DNA encoding the repeated sequence, while that of the V8 strain was about 60 bp shorter, confirming the previous report on the V8 proteinase sequence [3].

Although no physiological significance of the unusual repeated sequence is known, the difference in its length might affect the auto-digestion of the enzyme. When the SDS-PAGE patterns of SPase and V8 pro-

teinase were compared, SPase showed a single band, while V8 proteinase showed a main band of 3.3 kDa accompanied by minor bands even though it had been freshly prepared. Because the minor bands had a tendency to increase in number and amount as the digestion time elapsed, they are considered to be auto-digestion products. With SPase, such auto-digestion is not often observed. The exact role of the repeated sequence observed in the C-terminal region for both enzymes remains to be elucidated using the technique of site-directed mutagenesis.

References

1 Arvidson S.O. (1983) in *Staphylococci* and Staphylococcal Infections (Easmon, C.S.F. and Adlam, C., eds.), Vol. 2, pp. 780–808, Academic Press, New York.

2 Drapeau, G.R., Boily, Y. and Houmard, J. (1972) J. Biol. Chem. 247, 6720–6726.

3 Carmona, C. and Gray, G.L. (1987) Nucleic Acids Res. 15, 6757.

4 Gehrig, L.M.B., Delbaere, L.T.J. and Drapeau, G.R. (1985) J. Mol. Biol. 185, 651.

5 Yoshida, N., Tsuruyama, S., Nagata, K., Hirayama, K., Noda, K. and Makisumi, S. (1988) J. Biochem. 104, 451–456.

6 Niidome, T., Yoshida, N., Ogata, F., Ito, A. and Noda, K. (1990) J. Biochem. 108, 965–970.

7 Laemmli, U.K. (1970) Nature 227, 680–685.

8 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, New York.

9 Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.

10 Gibbs, R.A., Nguyen, P., McBride, L.J., Koepf, S.M. and Caskey, C.T. (1989) Proc. Natl. Acad. Sci. USA 86, 1919–1923.

11 Sanger, F. (1981) Science 214, 1205–1210.

12 Shine, J. and Dargano, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342–1346.

13 Kreil, G. (1981) Annu. Rev. Biochem. 50, 317–348.