Extremophiles (2004) 8:229–235
DOI 10.1007/s00792-004-0381-8

ORIGINAL PAPER

**Mitsuyoshi Okuda · Nobuyuki Sumitomo
Yasushi Takimura · Akinori Ogawa · Katsuhisa Saeki
Shuji Kawai · Tohru Kobayashi · Susumu Ito**

A new subtilisin family: nucleotide and deduced amino acid sequences of new high-molecular-mass alkaline proteases from *Bacillus* spp.

Received: 22 December 2003 / Accepted: 6 February 2004 / Published online: 12 March 2004

Abstract Six genes encoding high-molecular-mass subtilisins (HMSs) of alkaliphilic *Bacillus* spp. were cloned and sequenced. Their open reading frames of 2,394–2,424 bp encoded prosubtilisins of 798–808 amino acids (aa) consisting of the prepropeptides of 151–158 aa and the mature enzymes of 640–656 aa. The deduced aa sequences of the mature enzymes exhibited 60–95% identity to those of FT protease of *Bacillus* sp. strain KSM-KP43, a subtilisin-like serine protease, and a minor serine protease, Vpr, of *Bacillus* strains. Three of the six recombinant enzymes were susceptible to proteolysis, but the others were autodigestion resistant. All enzymes had optimal pH values of 10.5–11.0, optimal temperatures of 40–45°C for hydrolysis of a synthetic substrate, and were heat labile. These alkaline proteases seem to form a new subtilisin family, as judged by their aa sequences and phylogenetic analysis.

Keywords Alkaliphile · *Bacillus* · Cloning · Serine protease · Subtilisin

Introduction

Subtilisins are serine proteases found only in microorganisms, mainly from *Bacillus* strains. The subtilisin family is divided into true subtilisins, high-alkaline proteases, and intracellular proteases (Siezen and Leunissen 1997). These enzymes are important for commercial uses, especially in the detergent industry. We found a high-alkaline protease with a molecular mass of 28 kDa from *Bacillus* sp. KSM-K16 (Kobayashi et al. 1995) cloned, sequenced its gene (Hakamada et al. 1994), and determined the crystal structure (Shirai et al. 1997). We have succeeded in the large-scale production of the enzyme, and it has been incorporated into a heavy-duty laundry detergent since 1992 (Ito et al. 1998). Recently, in our search for detergent enzymes, we found new subtilisin-like serine proteases with molecular masses of 43–45 kDa that are innately resistant to chemical oxidants (Saeki et al. 2000, 2002). We propose that the enzymes be categorized as a new subfamily of subtilisins named oxidatively stable alkaline proteases (OSPs). We also found a new subtilisin from alkaliphilic *Bacillus* sp. strain LD-1 with original and moderate resistance to chemical oxidants (Saeki et al. 2003), and a high-molecular-mass, heat-labile subtilisin FT protease from alkaliphilic *Bacillus* sp. strain KSM-KP43 (Ogawa et al. 2003). The amino acid sequence of FT protease showed moderate homology to those of a subtilisin-like serine protease from *B. halodurans* (Takami et al. 2000) and the minor serine protease, Vpr, of *B. subtilis* (Sloma et al. 1991) with approximately 65% identity for both. We propose that FT protease and Vpr-like proteases might form a new subtilisin family. In this study, to support this proposal, we cloned and sequenced six new genes for high-molecular-mass subtilisins (HMSs) that were newly found in alkaliphilic *Bacillus* strains, and their deduced amino acid (aa) sequences were compared with those of FT protease and Vpr-like proteases. We propose that these HMSs be classified into a new subtilisin family and further could be divided into at least two subgroups.

Communicated by K. Horikoshi

M. Okuda · N. Sumitomo · Y. Takimura · A. Ogawa
K. Saeki · S. Kawai · T. Kobayashi (✉)
Tochigi Research Laboratories of Kao Corporation,
2606 Akabane, Ichikai, Haga, Tochigi 321-3497, Japan
E-mail: kobayashi.tohru@kao.co.jp
Tel.: +81-285-687543
Fax: +81-285-687566

S. Ito
Japan Marine Science and Technology Center,
2-15 Natsushima, Yokosuka 237-0061, Japan

Materials and methods

Bacterial strains

The sources of the genes used in this study were patented strains of alkaliphilic *Bacillus* spp.: KSM-KP9860 (FERM BP-6534),

KSM-9865 (FERM P-18566), D-6 (FERM P-1592), Y (FERM BP-1029), SD521 (FERM P-11162), and NCIB12289. The microorganisms were propagated in an alkaline polypeptone S-maltose medium, as described previously (Ogawa et al. 2003). They also produced OSPs extracellularly (Saeki et al. 2000, 2002).

Cloning and sequencing

Genomic DNAs of six strains were each isolated by the method of Saito and Miura (1963). The complete protease genes and their flanking regions were cloned by polymerase chain reaction (PCR) with the appropriate primers designed from the nucleotide (nt) sequence of FT protease (Ogawa et al. 2003). In the case of strain KSM-KP9860, the gene encoding part of SZ protease was amplified by PCR with primer A, 5′-GGTGAAGGTGTGACAGTAGCG-3′, designed from Gly^{20}–Val^{27} and primer B, 5′-ACCACTCGAATGCTTCCAGCC-3′, designed from Ala^{429}–Val^{435} of mature FT protease and the genomic DNA as the template. PCR was performed with 20 pmol each of primers A and B and 0.1 μg genomic DNA. The reaction program was denatured at 94°C for 2 min, followed by 30 cycles of 1 min at 94°C, 1 min at 55°C, and 2 min at 72°C using Ex *Taq* DNA polymerase (TaKaRa Bio, Kyoto, Japan) in a thermal cycler (model 480; Perkin-Elmer, Wellesley, Mass.). The amplified DNA fragment (1.0 kb) was purified and sequenced. To determine the complete nt sequence of the gene for SZ protease, inverse PCR was done as follows. The genomic DNA of strain KSM-KP9860 was completely digested with *Eco*RI or *Bgl*II, and then the segments were purified and treated with T4 ligase at 16°C overnight to form monomeric circles. The inverse PCR was done using *Eco*RI or *Bgl*II self-circularized DNA as the template, primer C, 5′-AATCACCGAAATTCGGTGCC-3′, and primer D, 5′-TCGGCGCTGATGAATACAGCC-3′. The reaction conditions were first 94°C for 2 min, followed by 30 cycles of 1 min at 94°C, 1 min at 50°C, and 4 min at 72°C using LA *Taq* DNA polymerase (TaKaRa Bio). The 2.2- and 2.0-kb fragments were amplified, respectively, and then each fragment was purified and sequenced for both strands. Consequently, the nt sequence of 3,293 bp was determined, which included the open reading frame (ORF) encoding SZ protease and its 5′ and 3′ flanking regions. The cloning and sequencing of the other five genes for proteases and their flanking regions were done in a similar manner to that described above. The nt sequences were determined for both strands using a BigDye terminator cycle sequencing kit and a model 377 DNA sequencer (Applied Biosystems, Foster City, Calif.). Computer analyses of the genes were done with a GENETYX-WIN program (version 5, Software Development, Tokyo, Japan). The FASTA program was used for the search for protein homologies (Pearson et al. 1988).

The nt sequences data reported in this paper have been submitted to the GenBank, EMBL, and DDBJ databases under accession nos. AB096096 for SZ protease from *Bacillus* sp. strain KSM-KP9860, AB096095 for SF protease from *Bacillus* sp. strain KSM-9865, AB100357 for HK protease from *Bacillus* sp. strain D-6, AB100359 for YK protease from *Bacillus* sp. strain Y, AB100358 for SD protease from *Bacillus* sp. strain SD521, and AB096097 for NV protease from *Bacillus* sp. strain NCIB12289.

Construction and expression of the genes for the proteases

The cloned genes were each introduced into multi-cloning sites of pHY300PLK (Yakult, Tokyo, Japan), and then the 28- and 72-kDa alkaline protease-deficient mutant of *Bacillus* sp. strain KSM-9865 was transformed with each constructed plasmid by electroporation (model SSH-10, Shimadzu, Kyoto, Japan). The transformants were selected on a Luria-Bertani agar containing separately sterilized 1.0% skim milk (Difco, Detroit, Mich.) and 0.05% Na_2CO_3 plus 15 μg/ml tetracycline (filter-sterilized). They were propagated, with shaking, in the alkaline polypeptone S-maltose medium, as described previously (Ogawa et al. 2003).

Purification of recombinant proteases

Recombinant SZ, SF, and NV proteases were purified from the centrifugal supernatants (80 ml each) in a manner similar to that of FT protease (Ogawa et al. 2003). Due to low productivity and proteolysis during purification, recombinant HK, YK, and SD proteases were partially purified as follows. Ammonium sulfate was added to the supernatant (80 ml each) up to 90% saturation, and the solution was agitated for 30 min on an ice bath. The precipitates were collected by centrifugation and dissolved in 2 ml of 10 mM MOPS buffer (pH 6.0). Half of the crude enzyme solution was loaded onto a column of Bio-Gel A0.5m (1.5×75 cm, Bio-Rad, Hercules, Calif.) equilibrated with 10 mM MOPS buffer (pH 6.0) containing 0.1 M KCl. The fractions showing protease activity with a molecular mass of around 72 kDa [judged by SDS-polyacrylamide gel electrophoresis (PAGE)] were pooled.

Other analytical methods

Enzymatic activity was measured using a synthetic oligopeptidyl substrate, glutaryl-Ala-Ala-Pro-Leu-*p*-nitroanilide (Peptide Institute, Osaka, Japan), as described previously (Ogawa et al. 2003). One unit of the enzyme activity was defined as the amount of the protein that liberated 1 μmol of *p*-nitroaniline per min at 30°C. Protein was determined by a DC protein assay kit (Bio-Rad) with bovine serum albumin as the standard. SDS-PAGE was done on a 12.5% acrylamide slab gel according to the method of Laemmli (1970). An SDS-PAGE low-molecular-mass standard (Pharmacia Biotech, Uppsala, Sweden) was used as the marker protein. The N-terminal aa sequence of NV protease that had been electro-blotted onto an Immobilon-P transfer membrane (Millipore, Billerica, Mass.) was determined by a pulsed liquid-phase protein sequencer (model 470A, Applied Biosystems).

Construction of phylogenetic tree

The deduced aa-sequence data determined were compared with those of subtilisins obtained from EMBL, GenBank, and DDBJ databases. To construct a phylogenetic tree of subtilisins, multiple alignment of aa sequences, calculation of nt substitution rates (*Knuc* values), and the bootstrapping neighbor-joining method (Saitou and Nei 1987) were performed with the ClustalX Program (Thompson et al. 1994). Amino acid sequences of proteases were automatically aligned, and any alignment positions with gaps were ignored before calculating a tree in this program.

Result and discussion

Nucleotide sequences of genes for proteases

The properties of the genes for six proteases (HMSs) are summarized in Table 1. The nt sequences of the genes for HMSs had long, single ORFs of 2,394–2,424 bp, starting with a TTG codon for SZ and SF proteases, an ATG codon for HK, YK, and SD proteases, and a GTG codon for NV protease, and ending with a TAA codon for the five protease genes, except for SZ protease that ends with a TAG stop codon. The start codon of genes for FT protease (AB096094) (Ogawa et al. 2003), Vpr from *Bacillus subtilis* (M76590) (Sloma et al. 1991), Vpr from *B. pseudofirmus* (AF330160), and a minor serine protease from *B. halodurans* (G83753) (Takami et al. 2000) is TTG. The stop codon of the genes for Vpr from *B. subtilis* and a minor serine protease from

Table 1 Properties of the genes for high-molecular-mass subtilisins (HMSs)

Gene for HMS	SZ	SF	NV	HK	YK	SD	FT[a]	Vpr[b]
Size of ORF (bp)	2,424	2,424	2,421	2,394	2,394	2,394	2,424	2,418
Start codon	TTG	TTG	GTG	ATG	ATG	ATG	TTG	TTG
Stop codon	TAG	TAA	TAA	TAA	TAA	TAA	TAG	TAA
Ribosome-binding site (bp upstream of ORF)	AAGGGAG (8 bp)	GGGGG (10 bp)	AGGGGAG (10 bp)	GGGGTG (10 bp)	GGGGTG (11 bp)	GGGGTG (11 bp)	AAGGGG (11 bp)	AAAGGGGG (7 bp)
Free energy of RBS (kJ/mol)	−31.1	−42.0	−29.4	−50.8	−50.8	−50.8	−26.7	−38.2
Promoter sequence								
−35 region (bp between two regions)	TTGAAA (17)	TTGAAA (17)	TTGAAA (17)	GTTACA (18)	TAGAAT (17)	GTTACA (18)	TTGAAA (17)	TTCACT (18)
−10 region	TATAAT	TATACT	TATAAT	TATAAT	TATATT	TATAAT	TATAAT	TATAAC
Palindrome sequence[c]	2,436–2,480	2,442–2,478	2,463–2,501	2,402–2,444	2,401–2,445	2,402–2,444	2,437–2,475	2,423–2,453
Free energy of palindrome sequence (kJ/mol)	−119.7	−76.0	−121.8	−60.1	−63.4	−82.7	−134.3	−96.2
G + C content of ORF (mol%)	46.2	47.1	44.2	39.4	38.8	39.5	44.9	47.9

[a]Ogawa et al. (2003)
[b]Sloma et al. (1991)
[c]Nucleotide number from start codon (+1)

B. halodurans is TAA, whereas that of the genes for FT protease, Vpr from *B. pseudofirmus*, and a minor extracellular serine protease from *Oceanobacillus iheyensis* (AP004603) (Takami et al. 2002) is TAG. The putative ribosome-binding sites of each gene were found 8–11 bp upstream of the ORF, having free-energy values of −29.4 to −50.8 kJ/mol, as calculated by the method of Tinoco et al. (1973). There were putative sigma A-type promoter sequences as the potential −35 and −10 regions separated by 17–18 bp, which were found between 91 and 140 bp upstream from each ORF. Typical palindrome sequences are found 6–41 bp downstream from the termination codon of six protease genes. The free-energy values of the sequences for stem-loop structures are −60.1 to −121.8 kJ/mol, which would be sufficient for termination of transcription. The G + C contents of each ORF were between 38.8% (YK protease) and 47.1% (SF protease). The ORFs of the SF, SZ, NV, HK, YK, and SD protease genes showed homology to that of the FT protease gene (Ogawa et al. 2003) with 82.5, 80.4, 76.2, 61.7, 63.8, and 61.8% identity, respectively. They exhibited homology to that of Vpr from *B. subtilis* with 58.0–59.6% identity. The ORF of the HK protease gene showed homology to those of SF, SZ, NV, YK, and SD protease genes with 60.9, 63.0, 62.9, 90.4, and 96.9% identity, respectively.

Deduced amino acid sequences of proteases

The ORFs of HMSs encoding 798–808 aa included 151–158 aa putative prepropeptides (Fig. 1). The prepeptides of the proteases would be composed of 28–29 aa, with the putative cleavage sites of Val^{-126}-Ser^{-125}-Val^{-124} for SZ protease, Val^{-126}-Gln^{-125}-Aal^{-124} for SF and NV proteases, and Ala^{-127}-Ser^{-126}-Ala^{-125} for HK, YK, and SD proteases. They were inferred from the frequency of aa appearing in a signal sequence according to Tjalsma et al. (2000). From the analysis of the N-terminal aa sequences of NV protease, FT protease (Ogawa et al. 2003), and Vpr (Sloma et al. 1991), propeptides of these proteases were cleaved on the C-terminal aa residues at position −1 (Fig. 1). If propeptides of other proteases would be cleaved at the same position, the mature enzymes would contain 640 aa (HK, YK, and SD proteaess) and 656 aa (SF, SZ, and NV proteases) with calculated molecular masses of 67,277–68,622 Da and pI of pH 4.17–4.69. These values are quite close to those of FT protease, as reported by Ogawa et al. (2003). The putative catalytic triads of HMSs are well conserved at Asp^{29}, His^{73}, and Ser^{377} in SF, SZ, and NV proteases, and Asp^{29}, His^{66}, and Ser^{364} in HK, YK, and SD proteases (Fig. 1). The mature aa sequences of SF, SZ, and NV proteases showed high homology to that of FT protease with 95.0, 90.5, and 82.6% identity, respectively, and those of HK, YK, and SD proteases show moderate homology to that of FT protease with 63.9, 62.9, and 63.9% identity, respectively. The mature aa sequence of HK protease has high homology to those of YK and SD proteases with 91.7% and 97.8% identity, respectively. The mature aa sequences of HMSs showed homology to a minor serine protease from *B. halodurans* C-125 (Takami et al. 2000), Vpr from *B. subtilis* (Sloma et al. 1991), Vpr from *B. pseudofirmus*, and a minor extracellular serine protease from *O. iheyensis* (Takami et al. 2002) with 61.1–65.5, 53.5–57.2, 41.8–43.4, and 36.3–37.3% identity, respectively. The prepropeptides of SF, SZ, and NV proteases showed high homology to that of FT protease with 78.9, 84.9, and 81.6% identity, respectively. The prepropeptides of HK, YK, and SD proteases exhibited high homology to each other; the identical values were 97.5–98.7%, but moderate homology to that of FT protease with approximately 52% identity.

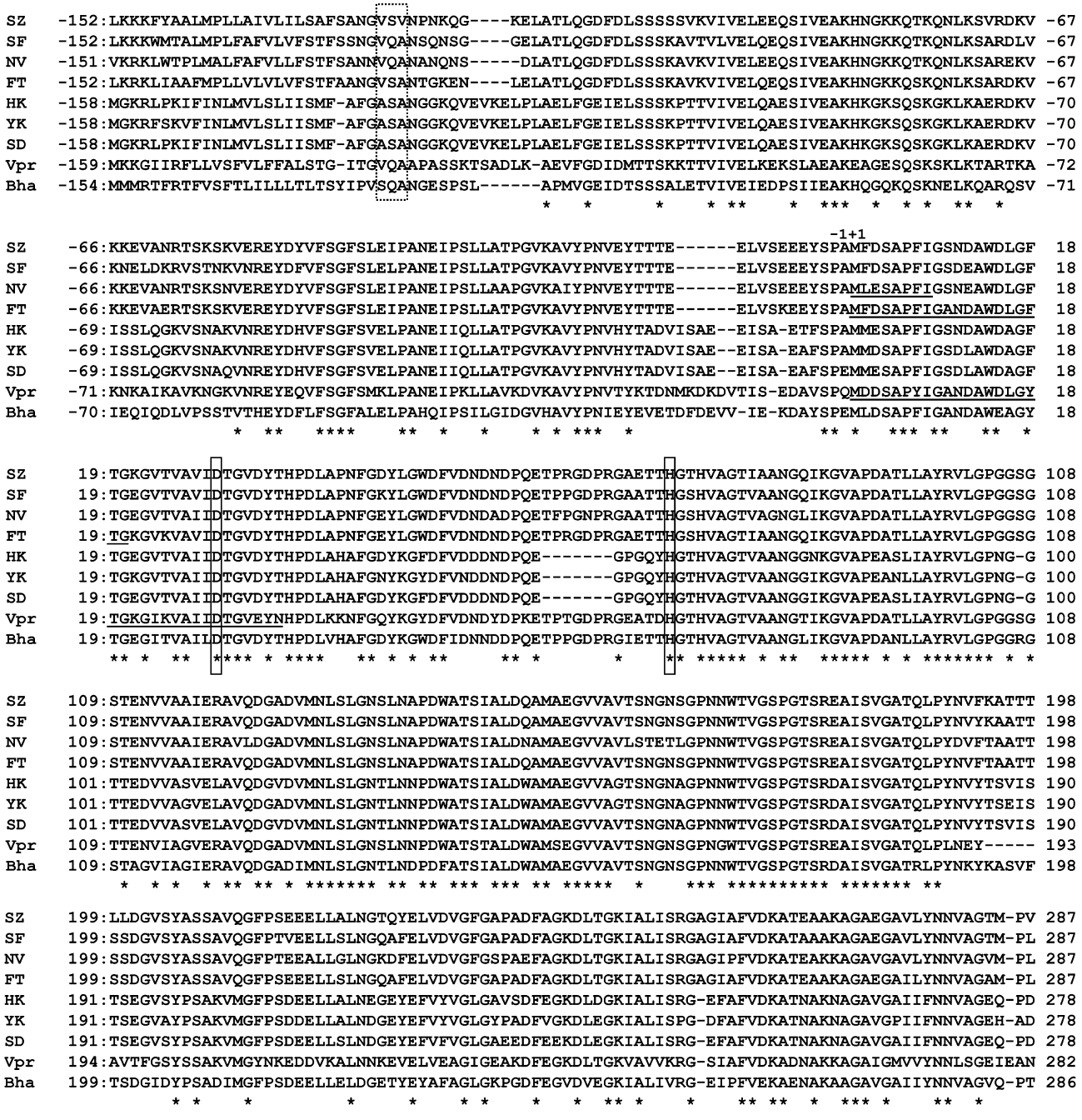

Fig. 1A, B Multiple alignment of the deduced amino acid (aa) sequences of high-molecular-mass subtilisins (HMSs). Sources aligned: SZ protease (AB096096) from *Bacillus* sp. strain KSM-P9860; SF protease (AB096095) from *Bacillus* sp. strain KSM-9865; NV protease (AB096097) from *Bacillus* sp. strain NCIB12289; FT protease (AB096094) from *Bacillus* sp. strain KSM-KP43 (Ogawa et al. 2003); HK protease (AB100357) from *Bacillus* sp. strain D-6; YK protease (AB100359) from *Bacillus* sp. strain Y; SD protease (AB100358) from *Bacillus* sp. strain SD521; Vpr (M76590) from *B. subtilis* (Sloma et al. 1991); Bha (G83753) from *B. halodurans* C-125 (Takami et al. 2000). The N-terminal aa sequence Met^{1} to Ile^{8} of NV protease analyzed is *underlined beneath the sequence*. The N-terminal aa sequences of FT protease (Ogawa et al. 2003) and Vpr (Sloma et al. 1991) are also *underlined beneath the sequence*. The predicted recognition sites of signal peptidases are *boxed with dotted line*. Identical aa residues for all nine proteases are marked with *asterisks under the sequences*. The possible catalytic triads, Asp, His, and Ser, are *boxed*

Enzymatic properties of recombinant HMSs

The recombinant SF, SZ, and NV proteases were purified to homogeneity, as judged by SDS-PAGE. The specific activity was 114 U/mg for SF protease, 13.3 U/mg for SZ protease, and 16.8 U/mg protein for NV protease. We have no idea at present why the specific activities of SZ and NV proteases were quite low compared with those of SF and FT (82 U/mg protein, Ogawa et al. 2003) proteases. In fact, the productivities

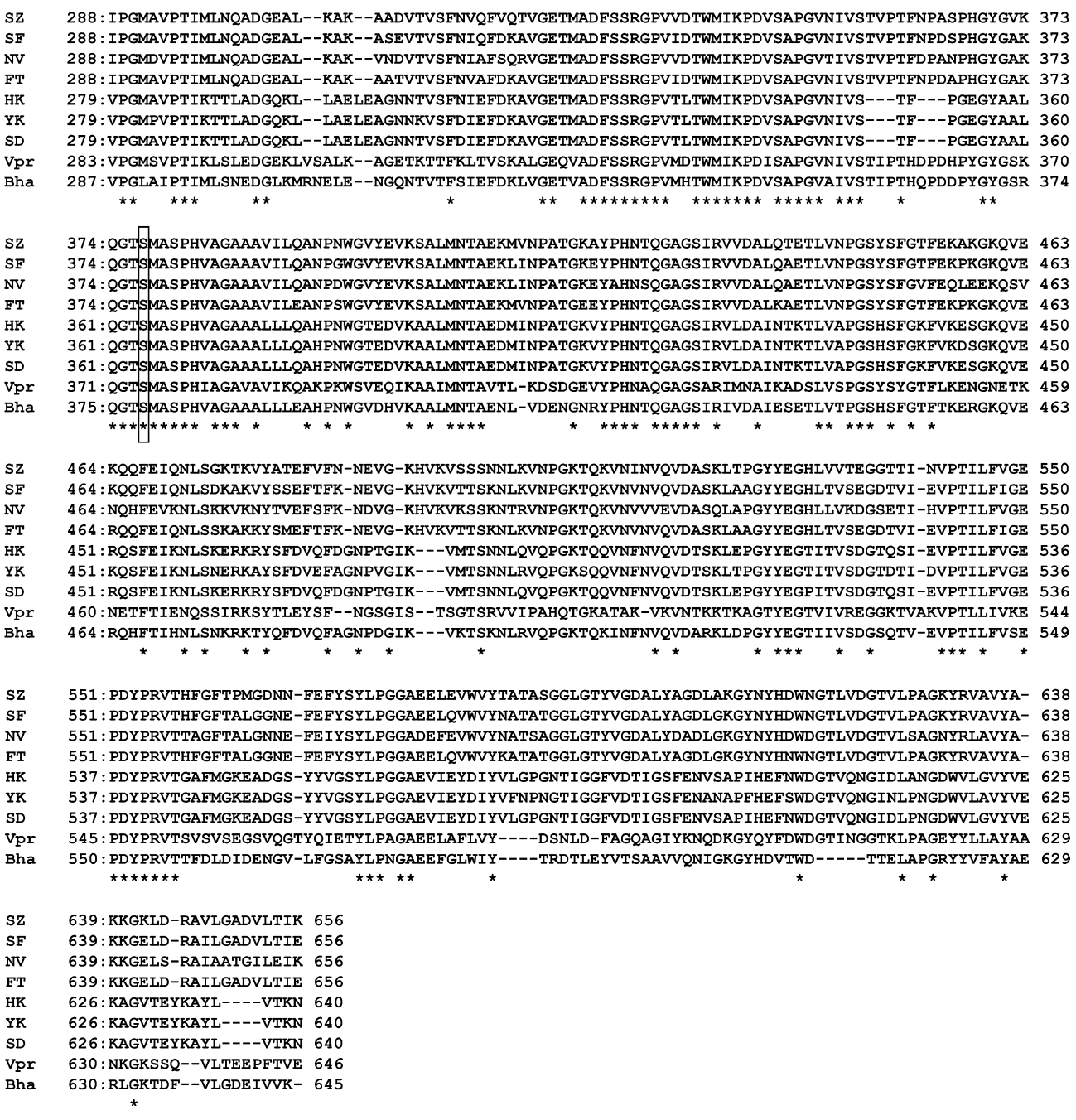

Fig. 1 (Continued)

of SZ and NV proteases were five to ten times lower than those of SF and FT proteases. Their low specific activities were reproducible even when they were purified from different culture broth. The molecular masses of these enzymes were estimated to be approximately 72 kDa by SDS-PAGE. In contrast, we could not purify HK, YK, and SD proteases because they were degraded to several polypeptides of 20 to 43 kDa that had protease activity during culture and purification. These observations were also reported on Vpr from *B. subtilis* (Sloma et al. 1991). Small amounts of HK, YK, and SD proteases were partially purified using a gel filtration column of Bio-Gel A0.5m. The molecular masses of these proteases were approximately 72 kDa, as judged by SDS-PAGE.

The HMSs showed the highest activity toward AAPL at pH 10.5 (SF, SZ, and NV proteases) and at pH 11.0 (HK, YK, and SD proteases) in 50 mM carbonate buffer.

However, the relative activities at pH 12.0 in 50 mM phosphate buffer were less than 20% for SF, SZ, and NV proteases and more than 65% for HK, YK, and SD proteases when the highest activity was taken as 100%. The optimal temperatures for activity of HMSs were 40–45°C for all, values lower than that of other

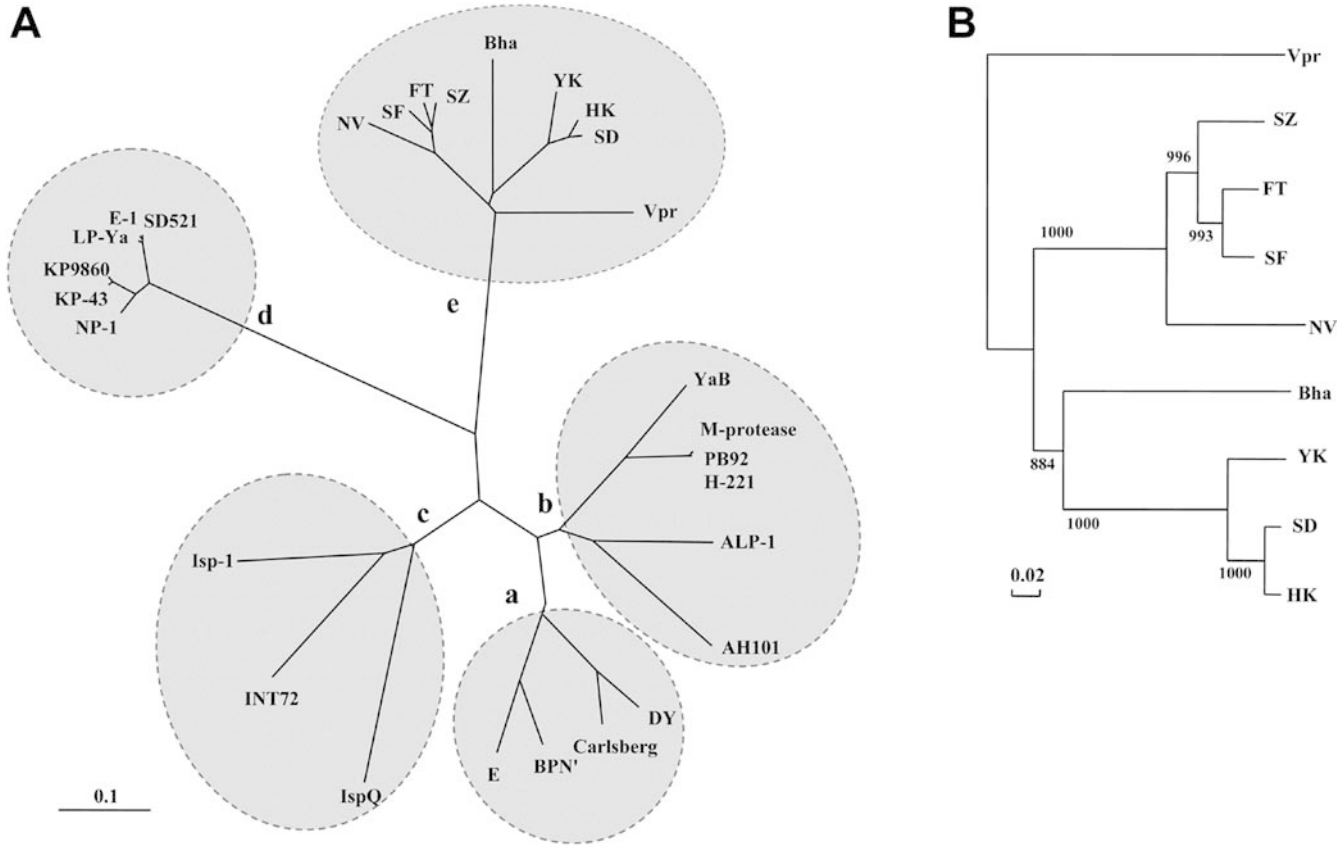

Fig. 2A, B Phylogenetic tree of subtilisins. A phylogenetic tree was inferred by the bootstrapping neighbor-joining method (Saitou and Nei 1987) in the ClustalX program (Thompson et al. 1994). *Bar* represents *Knuc* unit. **A** Unrooted phylogenetic tree of subtilisins. *Group a* True subtilisins, *group b* high-alkaline proteases, *group c* intracellular proteases, *group d* oxidatively stable alkaline proteases (OSPs) (Saeki et al. 2002), *group e* HMSs. Sources of sequences aligned: subtilisin E (P04189) from *Bacillus subtilis* 168; BPN′ (Q44684) from *B. amyloliquefaciens*; Carlsberg (P00780) from *B. licheniformis*; DY (P00781) from *B. subtilis* DY; YaB (P20724) from *Bacillus* sp. strain YaB; M-protease (Q99405) from *Bacillus* sp. strain KSM-K16; PB92 (P27693) from *B. alcalophilus* PB92; H-221 (P41362) from *Bacillus* sp. strain 221; ALP-1 (Q45523) from *Bacillus* sp. strain NKS-21; AH101 (D13158) from *Bacillus* sp. strain AH-101; Isp-1 (P08750) from *B. subtilis* IFO3013; INT72 (P29139) from *B. polymyxa* 72; IspQ (Q45621) from *Bacillus* sp. strain NKS-21; SD-521 (AB046405) from *Bacillus* sp. strain SD-521; E-1 (AB046402) from *Bacillus* sp. strain D-6; LP-Ya (AB046404) from *Bacillus* sp. strain Y; KP9860 (AB046403) from *Bacillus* sp. strain KSM-KP9860; KP-43 (AB051423) from *Bacillus* sp. strain KSM-KP43; NP-1 (AB046406) from *Bacillus* sp. strain NCIB12289. **B** Phylogenetic tree of HMSs. Sources of sequences aligned are the same as those used in Fig. 1

subtilisins. When 2 mM $CaCl_2$ was added to the reaction mixture, the optimal temperatures of HK, YK, and SD proteases increased to 50°C, whereas those of SF, SZ, and NV proteases did not change at all, regardless of whether $CaCl_2$ was present or not. The thermal stability of HMSs was assessed after enzymes had been incubated at various temperatures for 15 min in 50 mM borate buffer (pH 10.0) with or without 5 mM $CaCl_2$. SF, SZ, and NV proteases were stable up to 45°C with and without $CaCl_2$. In contrast, HK, YK, and SD proteases were stable up to 40°C without, and 50°C with, $CaCl_2$.

These properties of SF, SZ, and NV proteases were similar to those of FT protease. Further, it is suggested that all enzymes have high molecular masses and are heat labile, like FT protease. Sloma et al. (1990) reported that a gene for mature bacilloprotease F of *B. subtilis* encodes 1,238 aa. They purified a recombinant enzyme with a molecular mass of 47 kDa. An extracellular 90-kDa serine protease of *B. subtilis natto* was purified by Kato et al. (1992). These two high-molecular-mass serine proteases show a similar N-terminal aa sequence. However, we cannot compare the properties of these two enzymes with those of HMSs due to very limited information. At the least, the deduced aa sequences of HMSs show a quite low homology to that of bacilloprotease F of *B. subtilis* with less than 20% identity.

Phylogenetic trees of HMSs

An unrooted phylogenetic tree of subtilisins including true subtilisins, high-alkaline proteases such as M-protease (Hakamada et al. 1994), intracellular proteases, and OSPs (Saeki et al. 2000, 2002) is shown in Fig. 2A. As expected, HMSs, Vpr, and Vpr-like proteases form a clearly distinct clan from clans of high-alkaline proteases and OSPs. To further certify the phylogenetic positions of HMSs, we calculated another tree with less ignored gaps using the aa-sequence

alignment among the enzymes. As shown in Fig. 2B, they could be divided into three subgroups according to the aa-sequence homology. In fact, the enzymatic properties of HMSs can be clearly divided into at least two subgroups. One is the group of SF, SZ, NV, and FT proteases, which are less alkaline, independent of Ca^{2+} ions for stability, and resistant to proteolysis, and the other is HK, YK, and SD proteases, which are highly alkaline, stable dependent on Ca^{2+} ions for stability, and sensitive to proteolysis, although these three enzymes were partially purified. The two subgroups are very consistent with the phylogenetic positions. Interestingly, the phylogenetic position of each HMS is quite similar to those of OSPs, and, further, quite similar to those of OSP and HMS producers constructed by analysis of 16S rRNA (Saeki et al. 2002). More new HMSs and OSPs produced by (alkaliphilic) *Bacillus* strains will be found in the future.

References

Hakamada Y, Kobayashi T, Hitomi J, Kawai S, Ito S (1994) Molecular cloning and nucleotide sequence of the gene for an alkaline protease from the alkalophilic *Bacillus* sp. KSM-K16. J Ferment Bioeng 78:105–108

Ito S, Kobayashi T, Ara K, Ozaki K, Kawai S, Hatada Y (1998) Alkaline detergent enzymes from alkaliphiles: enzymatic properties, genetics, and structure. Extremophiles 2:185–190

Kato T, Yamagata Y, Arai T, Ichishima E (1992) Purification of a new extracellular 90-kDa serine proteinase with isoelectric point of 3.9 from *Bacillus subtilis* (*natto*) and elucidation of its distinct mode of action. Biosci Biotechnol Biochem 56:1166–1168

Kobayashi T, Hakamada Y, Adachi S, Hitomi J, Yoshimatsu T, Koike K, Kawai S, Ito S (1995) Purification and some properties of an alkaline protease from alkalophilic *Bacillus* sp. KSM-K16. Appl Microbiol Biotechnol 43:473–481

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Ogawa A, Sumitomo N, Okuda M, Saeki K, Kawai S, Kobayashi T, Ito S (2003) Nucleotide and deduced amino acid sequences of a high-molecular-mass subtilisin from an alkaliphilic *Bacillus* strain. Biochim Biophys Acta 1624:109–114

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444–2448

Saeki K, Okuda M, Hatada Y, Kobayashi T, Ito S, Takami H, Horikoshi K (2000) Novel oxidatively stable subtilisin-like serine proteases from alkaliphilic *Bacillus* spp.: enzymatic properties, sequences, and evolutionary relationships. Biochem Biophys Res Commun 279:313–319

Saeki K, Hitomi J, Okuda M, Hatada Y, Kageyama Y, Takaiwa M, Kubota H, Hagihara H, Kobayashi T, Kawai S, Ito S (2002) A novel species of alkaliphilic *Bacillus* that produces an oxidatively stable alkaline serine protease. Extremophiles 6:65–72

Saeki K, Magallones MV, Takimura Y, Hatada Y, Kobayashi T, Kawai S, Ito S (2003) Nucleotide and deduced amino acid sequences of a new subtilisin from an alkaliphilic *Bacillus* isolate. Curr Microbiol 47:337–340

Saito H, Miura K (1963) Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim Biophys Acta 72:619–629

Saitou N, Nei M (1987) A neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 44:406–425

Shirai T, Suzuki A, Yamane T, Ashida T, Kobayashi T, Hitomi J, Ito S (1997) High-resolution crystal structure of M-protease: phylogeny-aided analysis of the high-alkaline adaptation mechanism. Protein Eng 10:627–634

Siezen RJ, Leunissen JAM (1997) Subtilases: the superfamily of subtilisin-like serine proteases. Protein Sci 6:501–523

Sloma A, Rufo GA Jr, Rudolph CF, Sullivan BJ, Theriault KA, Pero J (1990) Bacillopeptidase F of *Bacillus subtilis*: purification of the protein and cloning of the gene. J Bacteriol 172:1470–1477

Sloma A, Rufo GA Jr, Theriault KA, Dwyer M, Wilson SW, Pero J (1991) Cloning and characterization of the gene for an additional extracellular serine protease of *Bacillus subtilis*. J Bacteriol 173:6889–6895

Takami H, Nakasone K, Takagi Y, Maeno G, Sasaki R, Masui N, Fuji F, Hirama C, Nakamura Y, Ogasawara N, Kuhara S, Horikoshi K (2000) Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. Nucleic Acids Res 28:4317–4331

Takami H, Takagi Y, Uchiyama I (2002) Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environment. Nucleic Acids Res 30:3927–3935

Thompson JD, Higgins DG, Gibson TJ (1994) ClustalW: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

Tinoco I Jr, Borer PN, Dengler B, Levine MD, Uhlenbeck OC, Crothers DM, Gralla J (1973) Improved estimation of secondary structure of ribonucleic acids. Nat New Biol 246:40–41

Tjalsma H, Bolhuis A, Jongbloed JHD, Bron S, van Dijl JM (2000) Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. Microbiol Mol Biol Rev 64:515–547