Appl Microbiol Biotechnol (2007) 75:71–80
DOI 10.1007/s00253-006-0800-0

Extremely high alkaline protease from a deep-subsurface bacterium, *Alkaliphilus transvaalensis*

Tohru Kobayashi · Jie Lu · Zhijun Li · Vo Si Hung · Atsushi Kurata · Yuji Hatada · Ken Takai · Susumu Ito · Koki Horikoshi

Received: 28 October 2006 / Revised: 6 December 2006 / Accepted: 6 December 2006 / Published online: 11 January 2007

Abstract A new high-alkaline protease (ALTP) was purified to homogeneity from a culture of the strictly anaerobic and extremely alkaliphilic *Alkaliphilus transvaalensis*. The molecular mass was 30 kDa on sodium dodecyl sulfate–polyacrylamide gel electrophoresis. The enzyme showed the maximal caseinolytic activity higher than pH 12.6 in KCl–NaOH buffer at 40°C. Hydrolysis of the oxidized insulin B-chain followed by mass spectrometric analysis of the cleaved products revealed that as many as 24 of the total 29 peptide bonds are hydrolyzed in a block-cutting manner, suggesting that ALTP has a widespread proteolytic functions. Calcium ion had no effect on the activity and stability of ALTP, unlike known subtilisins. The deduced amino acid sequence of the enzyme comprised 279 amino acids plus 97 prepropeptide amino acids. The amino acid sequence of mature ALTP was confirmed by capillary liquid chromatography coupled to tandem mass spectrometry, which was the 93% coverage of the deduced amino acid sequence. The mature enzyme showed moderate homology to subtilisin LD1 from the alkaliphilic *Bacillus* sp. strain KSM-LD1 with 64% identity, and both enzymes formed a new subcluster at an intermediate position among true subtilisins and high-alkaline proteases in a phylogenetic tree of subtilase family A. ALTP is the first high-alkaline protease reported from a strict anaerobe in this family.

Keywords Serine protease · Subtilisin · Subtilase family A · Deep subsurface · *Alkaliphilus transvaalensis*

T. Kobayashi (✉) · J. Lu · Z. Li · V. S. Hung · A. Kurata · Y. Hatada · K. Takai · S. Ito · K. Horikoshi
Extremobiosphere Research Center of Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15, Natsushima, Yokosuka 237-0061, Japan
e-mail: kobayashit@jamstec.go.jp

S. Ito
Creative Research Initiative Sousei (CRIS), Hokkaido University, Sapporo, Japan

Present address:
J. Lu
College of Life Science and Technology, Beijing University of Chemical Technology, 15 BeiSanhuan East Road, ChaoYang District, Beijing 100029, China

Introduction

The existence of microorganisms in geologically diverse deep terrestrial and oceanic subsurface environments is of increasing scientific and practical interest. Whitman et al. (1998) suggested that the terrestrial and oceanic subsurface biosphere is the largest reservoir of biomass on earth. It is also expected that subsurface microorganisms with novel metabolic properties may have potential applications in bioremediation and biotechnology (Boone et al. 1995; Kieft et al. 1999; Lu et al. 1999).

Alkaliphilus transvaalensis gen. nov., sp. nov. (strain SAGM1) is an extremely alkaliphilic anaerobe isolated from a mine-water containment dam 3.2 km below the earth's land surface in an ultra-deep gold mine near Carletonville, South Africa (Takai et al. 2001). The organism has an optimal growth at pH 10 but is capable of growing at pH higher than 12.5. Extracellular enzyme activity is generally recognized as a key step in the degradation and utilization of organic polymers because compounds with a molecular mass under 600 Da can pass through cell membranes (Hoppe 1991). *A. trans-*

vaalensis SAGM1 utilizes proteinaceous substrates such as peptone, tryptone, and casein, suggesting that it would produce a high-alkaline protease. All known high-alkaline proteases are produced by alkaliphilic bacilli (Schmidt et al. 1995) and belong to subtilisin. Subtilisins are alkaline serine proteases found only in microorganisms and mainly from *Bacillus* strains. More than 200 subtilisin-like serine proteases, designated subtilases, are classified into families A through F (Siezen and Leunissen 1997). Subtilisins belong to the subtilase Family A, including true subtilisins, high-alkaline proteases, intracellular proteases and further, oxidatively stable alkaline proteases (Saeki et al. 2000), and high-molecular-mass subtilisins (Okuda et al. 2004). Some of the high-alkaline proteases have been used as detergent enzymes to facilitate the release of proteinaceous soils from fabrics (Egmond 1997; Horikoshi 1999). We found a new high-alkaline protease belonging to the subtilase Family A from a culture of strictly anaerobe *A. transvaalensis* SAGM1 and described the enzymatic properties of a protease (ALTP) in this report.

Materials and methods

Bacterial strain and growth conditions

A. transvaalensis SAGM1 was cultured strictly anaerobically in SM medium (pH 10.5) at 40°C under a N_2 gas phase and at a pressure of 150 kPa, according to the method of Takai et al. (2001) and Takai and Horikoshi (2000). The medium (400 ml) was prepared in a 1-l bottle that was tightly sealed with a rubber stopper during cultivation for 24 h (Takai et al. 2001).

Purification of ALTP

All purification procedures were done at 4°C. The cells were removed by centrifugation (10,000×*g*, 20 min), and then the supernatant was dialyzed against tap water in a cold room for 24 h. The retentate (adjusted to pH 7.0 with 1 M phosphate buffer) was added to 500 ml of DEAE-Toyopearl 650M resin (Tosoh, Tokyo, Japan) equilibrated with 10 mM phosphate buffer (pH 7.0). The resin was removed by filtration with a glass filter, and the filtrate was loaded onto a column of CM-Toyopearl 650M (2×20 cm; Tosoh) equilibrated with 10 mM phosphate buffer (pH 7.0). Protease activity passed through the column, and the eluate was concentrated by centrifugation in an Amicon Ultra-15 (Ultra-10k, Millipore, Billerica, MA). The concentrate was used as the final purified enzyme throughout the experiments.

Electrophoretic analysis and protein content

The molecular mass of the purified ALTP was determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) according to the method of Laemmli (1970), on 12.5% (w/v) acrylamide slab gels (Bio-Rad, Hercules, CA). The protein was stained with Coomassie Brilliant Blue R-250 (CBB) or a silver stain kit. Precision plus protein standards (Bio-Rad) were used as the molecular mass standard. The protein was quantified by a protein assay kit (Bio-Rad) with bovine serum albumin as the standard.

Protease assays

Caseinolytic and oligopeptidyl-*p*-nitroanilide (pNA)-hydrolyzing activities were measured as described by Kobayashi et al. (1995). Hemoglobin (Difco, Detroit, MI) and keratin (Merck, Whitehouse Station, NJ) were also used for the substrates (Kobayashi et al. 1995). Protease activity was measured routinely using casein as the substrate. One unit (U) of caseinolytic activity was defined as the amount of enzyme that released acid-soluble peptides equivalent to 1 μmol L-tyrosine per minute at 40°C and pH 10 in 50 mM borate–NaOH buffer. One unit of oligopeptidyl-pNA-hydrolyzing activity was defined as the amount of the enzyme that produced 1 μmol of pNA per minute at 30°C and pH 10 in 50 mM borate–NaOH buffer plus 2 mM $CaCl_2$. The synthetic substrates used were *N*-succinyl-L-Ala-Ala-Pro-Phe-pNA (AAPF; Sigma, St. Louis, MO), glutaryl-L-Ala-Ala-Pro-Leu-pNA (AAPL; Peptide Institute, Osaka, Japan), *N*-succinyl-L-Ala-Ala-Pro-Met-pNA (AAPM; Bachem, Bubendorf, Switzerland), methoxysuccinyl-L-Ala-Ile-Pro-Met-pNA (AIPM; Bachem), *N*-succinyl-L-Ala-Ala-Val-Ala-pNA (AAVA; Sigma), *N*-succinyl-L-Ala-Ala-Ala-pNA (AAA; Peptide Institute), *N*-succinyl-L-Ala-Ala-pNA (AA; Peptide Institute), *N*-*p*-tosyl-L-Gly-Pro-Lys-pNA (GPK; Sigma), *N*-*t*-butyloxycarbonyl-L-Leu-Gly-Arg-pNA (LGR; Bachem), *N*-carbobenzoxy-L-Tyr-Lys-Arg-pNA (YKR; Bachem), *N*-succinyl-Gly-Gly-Phe-pNA (GGF; Sigma), L-Glu-Pro-Val-pNA (EPV; Daiichi Pure Chemicals, Tokyo, Japan), *N*-carbobenzoxy-L-Phe-Val-Arg-pNA (FVR; Bachem), methoxysuccinyl-L-Ala-Ala-Pro-Val-pNA (AAPV; Bachem), and *N*-carbobenzoxy-L-Pro-Cit-pNA (PCi; Seikagaku Kogyo, Tokyo, Japan).

Hydrolysis of oxidized insulin B-chain

For determination of the cleavage sites, capillary liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) was used. The oxidized insulin B-chain (1 mg/ml in 0.1 M Tris–HCl buffer, pH 8.0; Sigma) was incubated for 5, 10, 60 min, and 20 h with 5 μg of ALTP at

30°C in a total volume of 100 μl. Aliquots (5 μl) were removed, and the reaction was immediately stopped by adding 45 μl of a solvent [2% (v/v) acetonitrile and 0.05% (w/v) trifluoroacetic acid, pH 2.2]. The hydrolysates were separated by a Magic 2002 reversed-phase high-performance liquid chromatography system (Michrom Bioresources, Auburn, CA) equipped with a microscale capillary column custom-packed with 5 μm Magic C18 reversed-phase resin (0.1×50 mm; Michrom Bioresources). The solvent system consisted of solvent A [2% (v/v) acetonitrile and 0.1% (v/v) formic acid] and solvent B [90% (v/v) acetonitrile and 1% (v/v) formic acid]. The peptides were eluted with a linear gradient of 95/5 solvent A/B to 50/50 solvent A/B over a 30-min period, followed by a 20/80 solvent A/B over a 2-min period with the flow rate of 50 μl/min. The eluted peptides were analyzed by an automated data-dependent MS/MS procedure using a Finnigan LCQ DecaXP-plus ion-trap mass spectrometer (Thermo Electron, San Jose, CA) equipped with a nano-spray ion source (borosilicate PicoTips, Econo 10; New Objective, Woburn, UK; Aggeler et al. 2002). The mass spectra were acquired over the range of 150–2,000 *m*/*z*, and the tandem mass raw data were used for the determination of peptide sequences through a TurboSEQUEST program search (Washburn et al. 2001). The accuracy of the assignments was confirmed by a cross-correlation factor greater than 2.0, indicating a highly significant match, and a delta cross-correlation factor greater than 0.1, indicating a significant distinction between the best match and the second-best match.

Determination of the amino acid sequence of mature ALTP

The CBB-stained purified ALTP band was manually excised from the SDS-PAGE gel and then digested with trypsin (12.5 μg/ml; Promega, Madison, WI). The trypsin digests were extracted from the gel slices by four 15-min washes with 50% (v/v) acetonitrile and 0.5% (w/v) trifluoroacetic acid solution in 50 mM NH_4HCO_3. Each digests was analyzed by an LC-MS/MS system as described above. The amino acid (aa) sequences of ALTP were identified using the database constructed from the deduced aa sequence and the TurboSEQUEST program (Cleveland 1983; Washburn et al. 2001).

Determination of the N-terminal and internal amino acid sequences of ALTP

For the N-terminal and internal aa sequencing, heat-inactivated ALTP (100°C for 3 min) was digested by trypsin. The digests were then subjected to SDS-PAGE and transferred to a 0.2-μm polyvinylidene difluoride membrane (Bio-Rad) that had been wetted with methanol (Cleveland 1983). The aa sequences of the peptides were determined by a protein sequencer (model 497HT; Applied Biosystems, Foster City, CA).

Cloning and sequencing of ALTP gene

The *A. transvaalensis* genomic DNA was prepared by the method of Saito and Miura (1963). The mixed primers A (5′-GCNCARWSNACNCCNTGGGG-3′) and B (5′-CCNGCNACRTGNGTNCCRTG-3′) were designed from the N-terminal aa sequence of ALTP, Ala-Gln-Ser-Thr-Pro-Trp-Gly-Val-Thr-Arg, purified from the culture of *A. transvaalensis* and the aa sequence of His-Gly-Thr-His-Val-Ala-Gly conserved in known subtilisins, respectively. Part of the gene for ALTP was amplified by polymerase chain reaction (PCR) using the primers A and B and the genomic DNA as the template. The amplified DNA fragment was sequenced. Then, the upstream and downstream regions of the sequenced DNA fragment were cloned by cassette-ligation-mediated PCR with LA PCR in vitro Cloning kit (Takara Bio, Kyoto, Japan) according to the manufacturer's instructions and sequenced using oligonucleotide primers designed based on the successive sequencing results, as described previously (Saeki et al. 2000). Double-stranded DNA sequencing was performed using an ABI Prism Big Dye Terminator Cycle Sequencing kit and an ABI 377 Sequencer (Applied Biosystems). Computer sequence analysis was carried out using the GENETYX-MAC program ver. 10.1 (SDC Software Development, Tokyo, Japan). The sequence data in this study have been submitted to the DNA Data Bank of Japan, European Molecular Biology Laboratory, and GenBank databases under accession number AB266094.

Results

Purification of ALTP

ALTP migrated as a single protein band during SDS-PAGE, as shown in Fig. 1. The molecular mass was approximately 30 kDa. The purification of ALTP was 96-fold, and the overall yield was 37%. The specific activity toward casein was 32.5 U/mg.

Substrate preference and cleavage specificity

Casein was the most efficiently hydrolyzed by ALTP among the proteinaceous substrates tested. When the activity against casein was taken as 100%, the rates of hydrolysis of hemoglobin and keratin were 73 and 22%, respectively. The cleavage specificity of ALTP toward various oligopeptidyl substrates was also examined. The

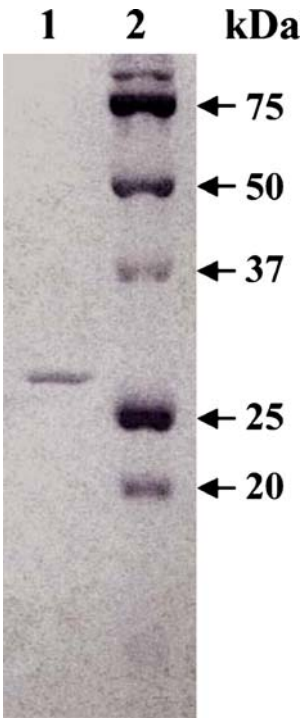

Fig. 1 SDS-PAGE of purified ATLP. The purified enzyme (2 μg protein) was subjected to a 12.5% acrylamide gel. The proteins were stained with CBB (*lane 1*). Protein mass markers (in kDa) are indicated on the *right* (*lane 2*)

enzyme was highly active toward AAPF, AAPL, AIPM, AAPM, and AAVA in that order of catalytic efficiency (k_{cat}/K_m), as summarized in Table 1. AAA, AA, GPK, LGR, YKR, GGF, EPV, FVR, AAPV, and PCi were poor substrates, and the rate of hydrolysis was less than 1% of the activity toward AAPF, which is one of the good substrates for high-alkaline proteases (Kobayashi et al. 1995). The enzyme showed little activity with oligopeptidyl substrates that are composed of less than 3 aa residues. These results suggest that ALTP has more than five subsites (Schechter and Berger 1967; Kobayashi et al. 1995) and favors synthetic peptides with a hydrophobic, non-polar, and bulky aa in the P1 position.

The oxidized insulin B-chain was digested at pH 8.0 and 30°C with the purified ALTP for various lengths of time from 5 min to 20 h. Under the time-limited cleavage, the proteolytic products of the chain were analyzed by LC-MS/MS. The enzyme showed 24 cleavage sites on the substrate (Fig. 2). The two major peaks, which appeared after 5-min incubation with ALTP, were identified as fragments of Phe^{1}–Leu^{15} and Tyr^{16}–Ala^{30}, indicating that the preferentially initial cleavage site is between Leu^{15} and Tyr^{16}, like other subtilisins. A number of small fragments, Phe^{1}–Ser^{9}, Phe^{1}–Leu^{11}, Phe^{1}–Val^{12}, Phe^{1}–Glu^{13}, Gln^{4}–Leu^{15}, Gln^{4}–Gly^{20}, His^{5}–Leu^{11}, His^{5}–Leu^{15}, Leu^{6}–Val^{12}, Cys^{7}–Glu^{13}, Tyr^{16}–Phe^{24}, Tyr^{16}–Lys^{29}, Leu^{17}–Gly^{23}, Val^{18}–Phe^{25}, Val^{18}–Ala^{30}, and Phe^{25}–Ala^{30}, were detected after 5- to 60-min incubation. Most of the peptides detected after complete hydrolysis (20 h) were composed of 4 to 8 aa residues. These results suggest that ALTP cleaves the oxidized insulin B-chain in a block-cutting manner. Subtilisin Carlsberg (Sigma) was used as a control enzyme to cleave the chain in the same manner as ALTP. As shown in Fig. 2, subtilisin Carlsberg showed 13 cleavage sites, which is a larger number than reported previously (Johansen et al. 1968).

Table 1 Kinetic parameters of ALTP toward pNA-substrates

Substrate	Relative activity (%)	K_m (mM)	Vmax (U/mg)	k_{cat} (S^{-1})	k_{cat}/K_m (S^{-1} mM^{-1})
AAPF	100	1.12	98.0	50.0	44.6
AAPL	91.0	1.04	82.6	42.2	40.6
AAPM	74.0	0.475	30.6	15.6	33.0
AIPM	90.1	0.588	46.0	23.6	40.2
AAVA	4.62	0.669	2.70	1.38	2.06

Effects of pH and temperature on ALTP activity and stability

The pH ranges at which ALTP was active and stable were determined with casein as the substrate. ALTP showed the maximal activity at pH higher than 12.6 (actual pH of the reaction mixture) in 50 mM KCl–NaOH buffer (Fig. 3a). To determine the pH stability of ALTP, the enzyme was preincubated at 50°C for 10 min at the indicated pH in various buffers, and then the residual caseinolytic activity was measured under the standard conditions of enzyme assay, as shown in Fig. 3b. The enzyme was stable over the pH range of 5–11.

The optimal temperature for the reaction of ALTP was around 70°C at pH 10 in 50 mM borate–NaOH buffer, as shown in Fig. 4a. The optimal temperature did not shift regardless of whether 5 mM $CaCl_2$ was present or not. The thermal stability of ALTP was assessed in 50 mM borate–NaOH buffer (pH 10) after heating at various temperatures for 10 min. The enzyme was stable up to 65°C (Fig. 4b), and the presence of 5 mM $CaCl_2$ had no effect on the thermal stability, as is the case of the effect of temperature on the activity (data not shown).

Effects of various chemical reagents

ALTP was incubated with various metal ions (1 mM each) at 30°C for 20 min in 20 mM borate–NaOH buffer (pH 10), and then the residual activity toward casein was measured under the standard conditions of enzyme assay. Ba^{2+}, Ca^{2+}, Cd^{2+}, Co^{2+}, Cu^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}, Fe^{2+}, Fe^{3+}, Pb^{2+}, Sn^{2+}, and Zn^{2+} ions did not influence the caseinolytic activity (data not shown). Only Hg^{2+} ions strongly inhibited the enzyme activity by 85%. ALTP was also treated with various chemicals at 30°C for 20 min in 20 mM phosphate

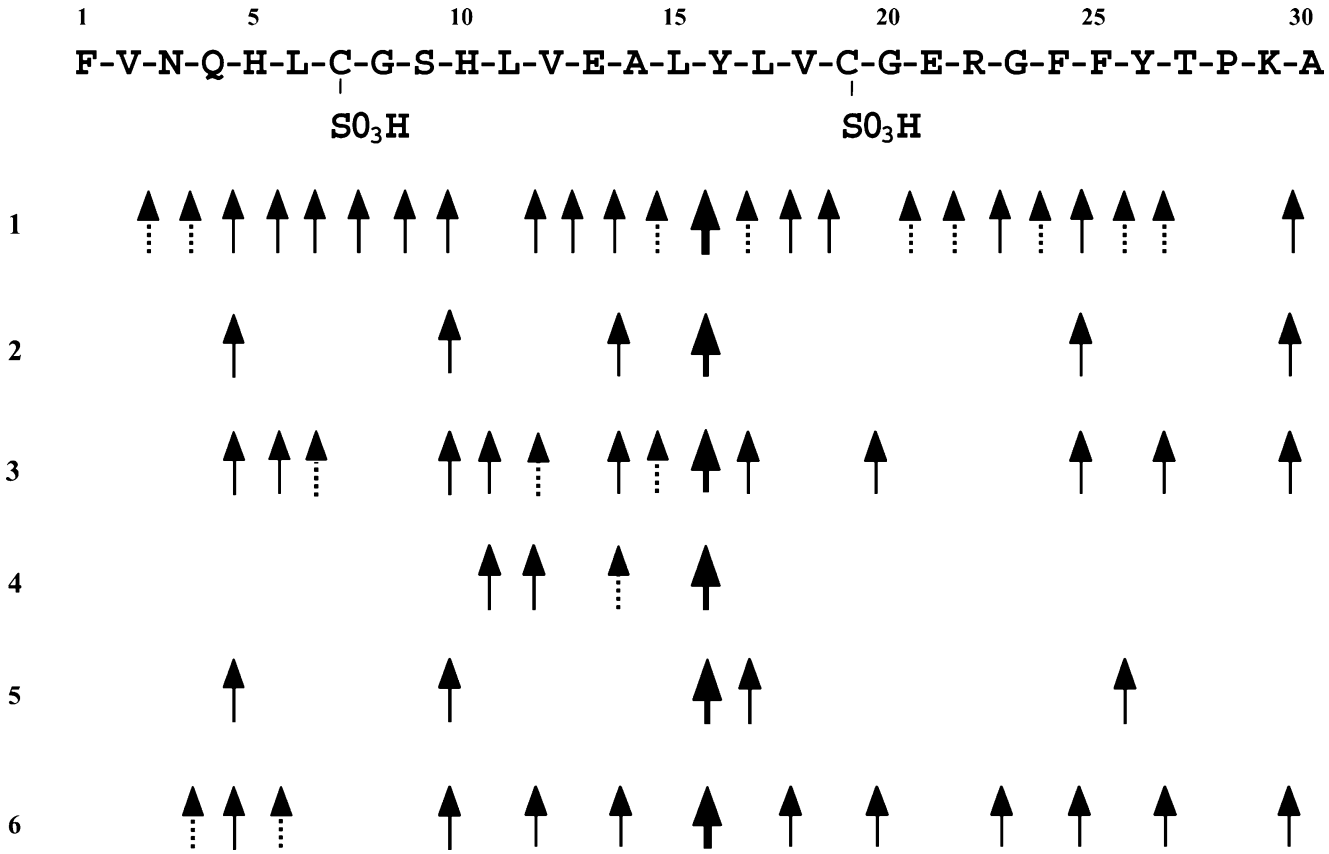

Fig. 2 Schematic representation of the oxidized insulin B-chain cleavage by ALTP and other typical subtilisins. The *arrows* indicate the unambiguous LC-MS/MS identification of the fragments generated by ALTP digestion. The initial cleavage sites (*thick arrow*) were determined after incubation for 5 min using 5 μg (0.2 U) of protein. The other cleavage sites (*arrows*, major sites; *dashed arrows*, minor sites) were determined after various periods of incubation up to 20 h with 5 μg protein. The detailed experimental conditions were described in "Materials and methods". *1* ALTP; *2* Sendai (Yamagata et al. 1995); *3* M-protease (Kobayashi et al. 1995); *4* AH101 (Takami et al. 1992a); *5* subtilisin BPN′ (Morihara et al. 1969); *6* subtilisin Carlsberg (0.2 U)

buffer (pH 7.0), and then the residual activity toward casein was measured. Phenylmethylsulfonyl fluoride (1 mM) abolished the activity completely, consistent with the classification of ALTP as a serine protease. ALTP was not affected by the following reagents: 1–100 mM EDTA; 1 mM *N*-bromosuccinimide, iodacetate, *N*-ethylmaleimide, *p*-chloromercuribenzoate, or iodoacetate; 2 mM diethyl pyrocarbonate; 5 mM 2-hydroxy-5-nitrobenzyl bromide; 0.1% (w/v) SDS, Triton X-100, or Tween 20. The several protease inhibitors, 1 mM chymostatin, 60 mM pepstatin, and 0.1 mM antipain, leupeptin, bestatin, phosphoramidon, and elastatinal, showed no inhibitory effect.

Nucleotide sequencing ALTP gene and the deduced amino acid sequence

The nucleotide sequence of the gene for ALTP and its flanking regions (1,235 bp) were sequenced as described in "Materials and methods". We found an open reading frame (ORF) with an ATG initiation codon at nucleotide 1 and a TAA termination codon at nucleotide 1,131, as shown in Fig. 5. The ORF encodes 376 aa. A possible ribosome-binding site (RBS), with the sequence 5′-AAAGGAGG-3′, is located eight bases upstream from the initiation codon. There is a putative promoter sequence, with 5′-TTAACA-3′ as the potential −35 region and 5′-CATAAT-3′ as the potential −10 region, separated by 16 bp. An inverted repeat sequence is found downstream of the termination codon from nucleotide 1,142 to 1,174, which would have a free energy of −76.7 kJ/mol. The G + C content of the ORF is 40.0 mol%.

The ORF encodes 376 aa, and the determined N-terminal aa sequence of the purified ALTP, Ala-Gln-Ser-Thr-Pro-Trp-Gly-Val-Thr-Arg, was found in the internal Ala^{98} to Arg^{107}, as shown in Fig. 5. Therefore, the sequence Met^{1} through Leu^{97} is its prepropeptide. The prepeptide would be cleaved at Ala^{23} of the C terminus (predicted by Signal P V2.0; Nielsen et al. 1997), and the propeptide would be cleaved at the Leu^{97} of the C terminus. The prepropeptide of ALTP shows low homology to those of the other known subtilisins with 23–42% identity (data not shown). To confirm the aa sequence of mature ALTP, the purified

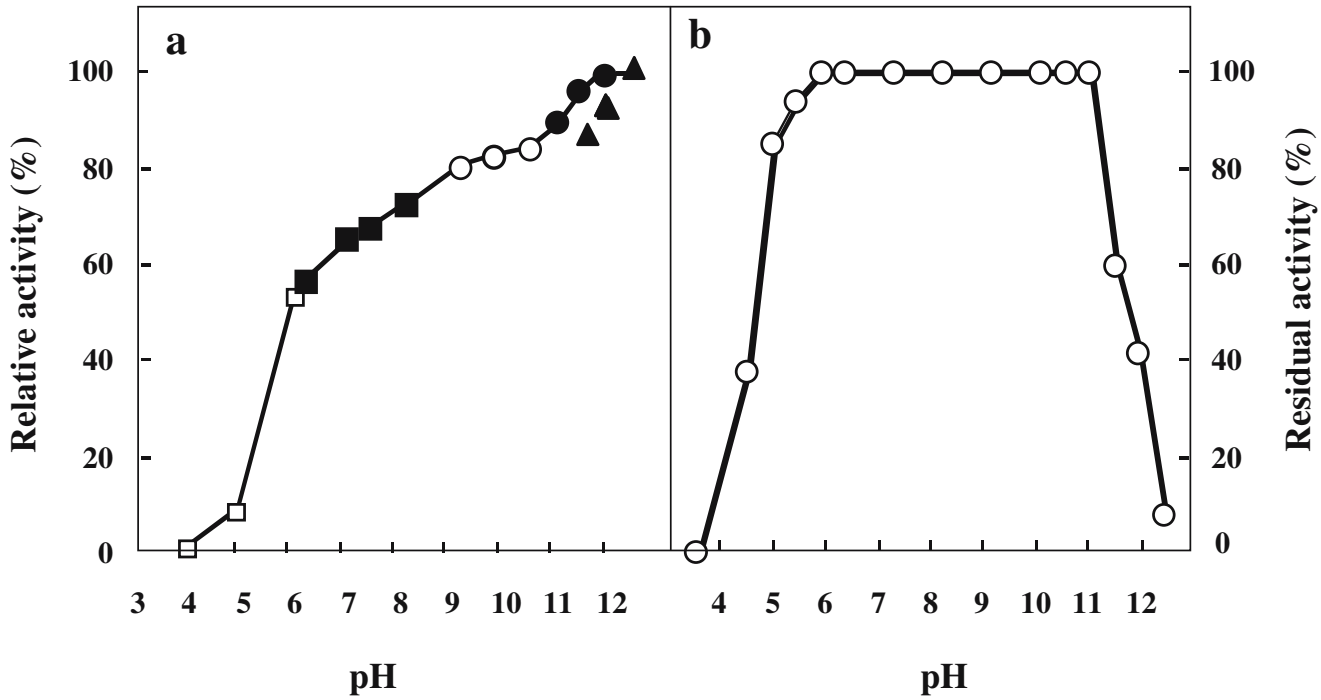

Fig. 3 Effects of pH. **a** Effects of pH on the enzymatic activity. The activity was measured at 40°C for 15 min in a total volume of 1.0 ml that contained 1.5 μg protease ALTP in the following buffers: *open squares*, acetate (pH 3.5–6.0); *filled squares*, phosphate (pH 6.5–8.1); *open circles*, carbonate (pH 9.0–11.0); *filled circles*, phosphate–NaOH (pH 11.0–12.2); *filled triangles*, KCl–NaOH (pH 11.5–12.6). **b** Effects of pH on the stability. ALTP (15 μg) was preincubated at 50°C for 10 min in various buffers (20 mM), the solutions of treated enzyme were diluted tenfold with 50 mM borate–NaOH buffer (pH 10), and the residual activity was measured at 40°C with casein as the substrate. The original activity before preincubation is taken as 100%

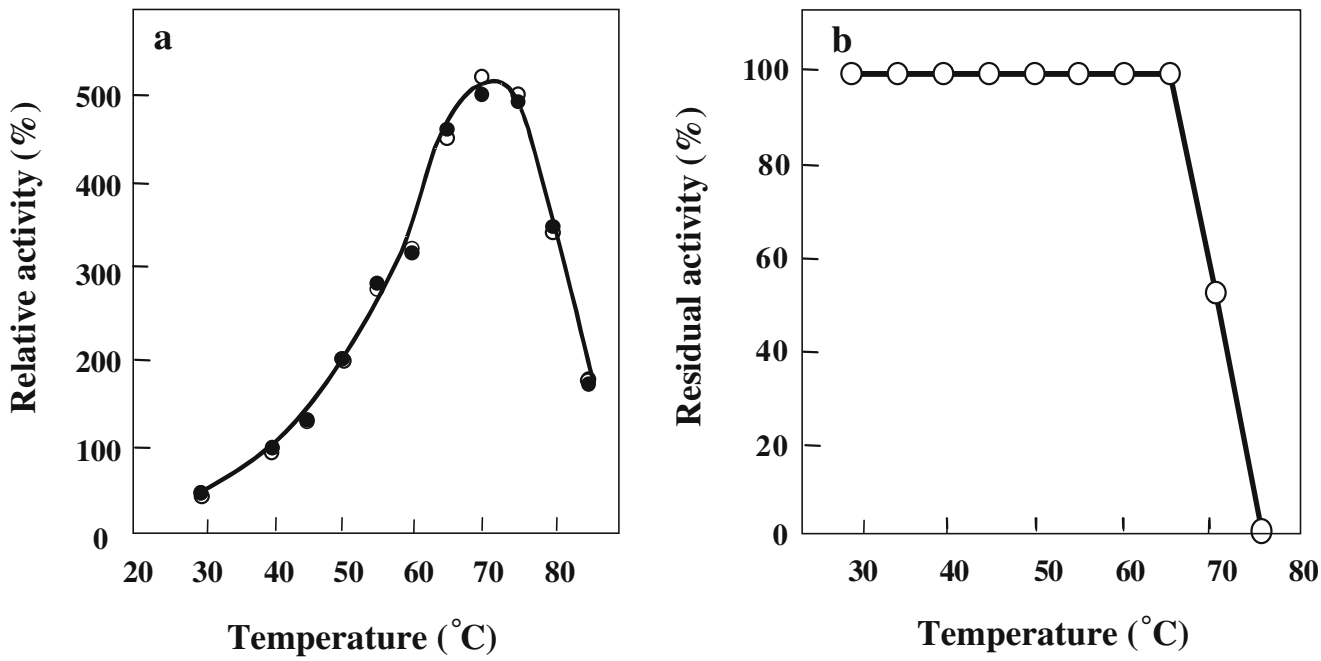

Fig. 4 Effects of temperature. **a** Effects of temperature on the caseinolytic activity. ALTP (0.8 μg) was added to 0.5% (w/v) casein in 50 mM borate–NaOH buffer (pH 10). The reactions were done for 15 min at various temperatures. The relative activities were expressed as the percentages of the activity at 40°C in the absence of $CaCl_2$. *Filled circles*, in the absence of $CaCl_2$; *open circles*, in the presence of 5 mM $CaCl_2$. **b** Effects of temperature on the stability of the enzyme. ALTP (1.5 μg) was preincubated for 10 min at various temperatures in 50 mM borate–NaOH buffer (pH 10). Heat-treatment was stopped by cooling on ice. The measurement of residual activity was done at 40°C for 15 min. The activity after heating at 30°C for 10 min was taken as 100%

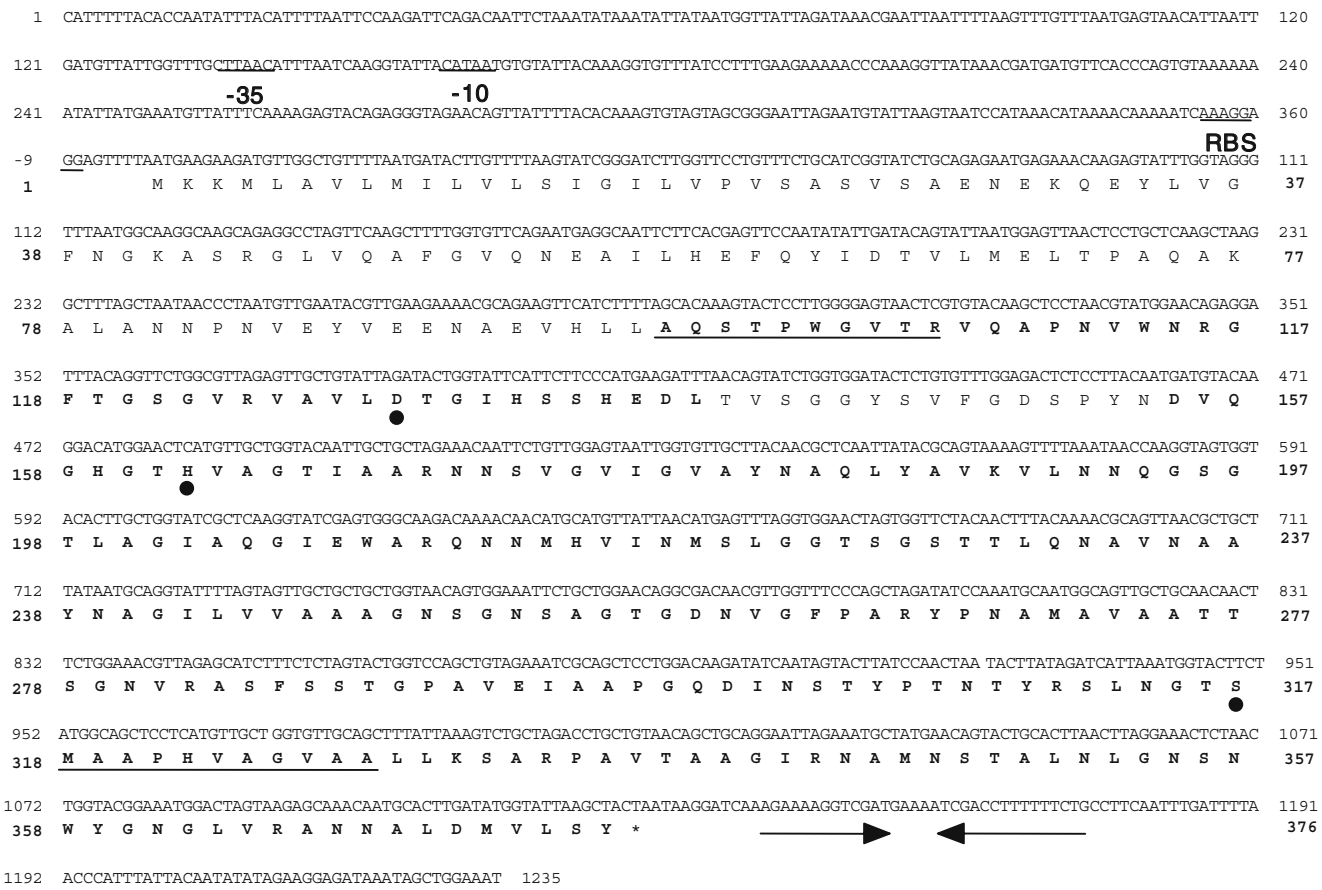

Fig. 5 Complete nucleotide sequence and deduced aa sequence of ALTP. In the 1,235 bp determined, the nucleotide sequence of the gene for ALTP and its flanking regions are shown. Sequences similar to the −35 and −10 consensus promoters are *underlined*. A putative ribosome-binding site is also *underlined* (*RBS*). The deduced aa sequence of the gene product is indicated by the single-letter codes under the nucleotide sequence. The N-terminal (Ala[98] to Arg[107]) and internal aa residues (Met[318] to Ala[328]) determined by a protein sequencer after trypsin digestion are shown in *bold face with underlines*. The aa residues of the mature enzyme from Ala[98] to Tyr[376] determined by LC-MS/MS are shown in *bold face*. Inverted repeats downstream from the stop codon TAA (*asterisk*) of the ORF are designated by *convergent arrows*. Possible catalytic residues, Asp[129], His[162], and Ser[317], are indicated by *filled circles* beneath the deduced aa sequence

enzyme was analyzed by LC-MS/MS after digestion in situ with trypsin. The peptide fragments were identified with TurboSEQUEST, which covered 92.7% aa of the enzyme (Fig. 5). Furthermore, the aa sequences of two peptide fragments generated by digestion with trypsin were identified as Ala[98] to Arg[107] and Met[318] to Ala[328]. These results indicate that the DNA fragment cloned in this study is correct for the aa sequence of ALTP. The calculated molecular mass and p*I* value of mature ALTP (279 aa, Ala[98]–Tyr[376]) were 28,307 Da and 9.61, respectively.

When aligned with other subtilisins, the mature enzyme shows the highest homology to subtilisin LD1 (Saeki et al. 2003) at 64.0% and next to subtilisin Sendai (Yamagata et al. 1995) at 61% identity. ALTP shows 53–59% identities with the other subtilisins, such as high-alkaline M-protease (Hakamada et al. 1994), true subtilisins BPN′ (Q44684), and Carlsberg (P00780). The possible catalytic triads are conserved as Asp[129], His[162], and Ser[317] in the mature enzyme. There is no Cys residue in ALTP, like all but a few subtilisins from the Antarctic psychrophilic *Bacillus* strains TA39 and TA41 (Narinx et al. 1992), and *Bacillus sphaericus* SSII-1 (Wati et al. 1997). Generally, high-alkaline proteases such as M-protease (Hakamada et al. 1994) and no. 221 (Takami et al. 1992b) have the four aa deletion around the aa 160 of true subtilisin BPN′, which corresponds to the P1 binding site (pocket). ALTP has no deletion of the binding site, like LD1 (Saeki et al. 2003), SprC, and SprD (Schmidt et al. 1995). In fact, the phylogenetic positions of ALTP and LD1 form an intermediate cluster between those of true subtilisins and high-alkaline proteases (Fig. 6).

Discussion

A. transvaalensis SAGM1 is one of the most alkaliphilic microorganisms reported so far (Horikoshi 1971, 1999; Takami et al. 1989; Kobayashi et al. 1995). Strain SAGM1

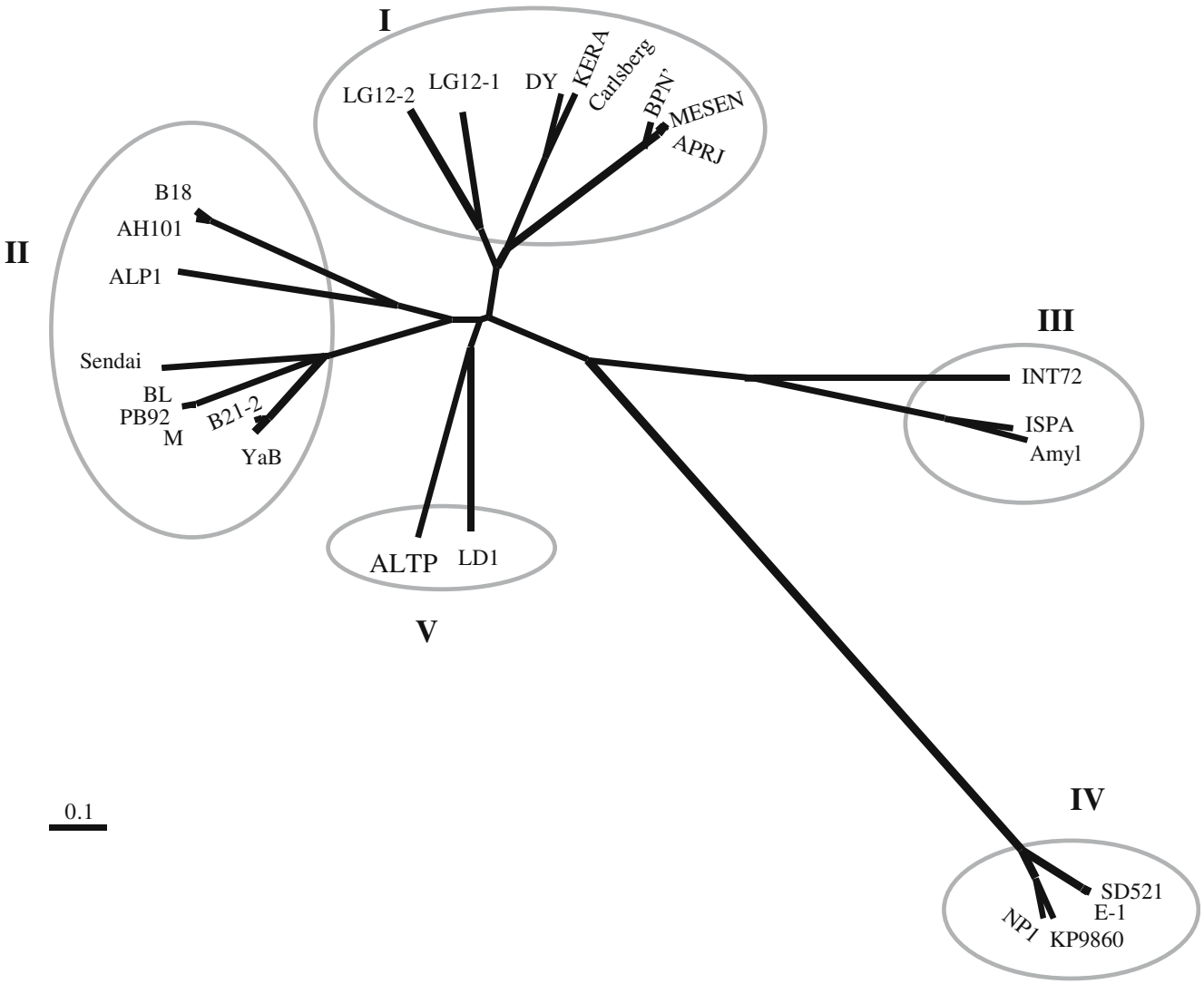

Fig. 6 Unrooted phylogenetic tree of subtilisins based on the aa sequence alignment. A phylogenetic tree was inferred by the neighbor-joining method in the Clustal X program. *I* True subtilisins; *II* high-alkaline proteases; *III* intracellular proteases; *IV* oxidatively stable proteases; *V* ALTP and LD1. *Bar* represents *k*nuc unit. Source of sequences aligned: ALTP (AB266094) from *A. transvaalensis* SAGM1 (this study); LD1 (AB085752) from *Bacillus* sp. strain KSM-LD1; LG12-1 and LG12-2 (Q45466 and Q45467) from *Bacillus* sp. strain LG12; DY (P00781) from *B. subtilis* DY; KERA (Q53521) from *B. licheniformis* PWQ1; Carlsberg (P00780) from *B. licheniformis*; BPN′ (Q44684) from *B. amyloliquefaciens*; MESEN (P07518) from *B. mesentericus*; APRJ (P29142) from *B. stearothermophilus* NCIMB 10278; B18 (Q45521) from *Bacillus* sp. strain B18′; AH101 (D13158) from *Bacillus* sp. strain AH101; ALP1 (Q45523) from *Bacillus* sp. strain NHS21; Sendai (Q45522) from *Bacillus* sp. strain G-825-6; BP92 (P27693) from *B. alcalophilus* BP92; M (Q99405) from *Bacillus* sp. strain KSM-K16; BL (P29599) from *B. lentus*; B21-2 (AB005792) from *Bacillus* sp. strain 21-2; YaB (P20724) from a *Bacillus* sp. strain YaB; INT72 (P29139) from *B. polymyxa* 72; ISPA (P11018) from *B. subtilis* 168; Amyl (P00783) from *B. subtilis* var. *amylosacchariticus*; SD521 (AB046405) from *Bacillus* sp. strain SD-521; E-1 (AB046402) from *Bacillus* sp. strain D-6; KP9860 (AB046403) from *Bacillus* sp. strain KSM-KP8960; NP1 (AB046406) from *Bacillus* sp. strain NCIB12289

is a gram-positive, strictly anaerobic heterotroph, capable of growth using proteinaceous substrates such as peptone, tryptone, and casein (Takai et al. 2001). We suppose that ALTP plays an important role in the ecological niche of the strain. Many anaerobic, proteolytic bacteria have been described, but the enzymatic characteristics of purified proteases from extremely anaerobic bacteria have rarely been reported (Jang et al. 2002a,b). These proteases from the anaerobes belong to subtilase family B (thermitase; Siezen and Leunissen 1997) unlike ALTP.

ALTP shows enzymatic properties similar to those of the other high-alkaline proteases belonging to subtilase family A, such as M-protease (Kobayashi et al. 1995), Sendai (Yamagata et al. 1995), no. 221 (Horikoshi 1971), and AH-101 (Takami et al. 1989, 1992a), with respect to the effects of pH and temperature on activity and stability, oxidative stability, and resistance to various metal ions and chemical reagents. On the contrary, Ca^{2+} ions have no effect on either the activation or thermal stability of ALTP. These properties are distinctly different from those of the subtilisins reported

to date (Horikoshi 1971; Takami et al. 1989; Kobayashi et al. 1995; Saeki et al. 2000). The specific activity of ALTP is 32.5 U/mg, which is an intermediate value among those of AH101 and Sendai (~15 U/mg; Takami et al. 1989; Yamagata et al. 1995), subtilisin LD1 (70 U/mg; Saeki et al. 2003), and no. 221 and M-protease (~127 U/mg; Horikoshi 1971; Kobayashi et al. 1995).

ALTP cleaves the oxidized insulin B-chain at the 24 most widespread cleavage sites, although the initial site is between Leu^{15} and Tyr^{16}, like other high-alkaline proteases and true subtilisins. Moreover, ALTP cleaves distinctively Val and Gly as the P1 position in the oxidized insulin B-chain when compared with the cleavage sites of other subtilisins reported to date (Fig. 2). Abraham et al. (1995) reported the cleavage specificity of the oxidized insulin B-chain by a subtilisin-like protease from *Ophiostoma piceae* using LC-MS/MS and found that the enzyme cleaved 12 sites. To our knowledge, the greatest number of cleavage sites on the oxidized insulin B-chain previously reported was cleavage of 21 sites by an alkaline protease from alkaliphilic *Streptomyces* sp. (Nakanishi and Yamamoto 1974).

The mature ALTP shows moderate homology to LD1 (Saeki et al. 2003), Sendai (Yamagata et al. 1995), and M-protease (Hakamada et al. 1994) and true subtilisins Carlsberg (P00780) and BPN′ (Q44684) with 64, 61, 59, 56, and 53% identity, respectively. In a phylogenic tree of enzymes belonging to subtilase family A, the 24 alkaline proteases examined were unequivocally grouped into five clusters based on the aa sequences of the mature enzymes (Fig. 6). The phylogenetic clusters of true subtilisins, high-alkaline proteases, intracellular proteases, and oxidatively stable proteases are clearly located on different branches, as reported previously (Saeki et al. 2000). ALTP and LD1 form a novel cluster at an intermediate position among true subtilisins and high-alkaline proteases.

A. transvaalensis SAGM1 is the most strictly anaerobic and extremely alkaliphilic microorganism among known alkaliphiles, and this strain was isolated from the deepest underground (3.2 km) among the isolated bacteria reported so far. Moreover, ALTP shows the highest optimal pH value and the greatest number of cleavage sites on the oxidized insulin B-chain among the known proteases. Thus, microorganisms from the extremobiosphere would be expected to produce the novel and superior enzymes.

References

Abraham LD, Chow DT, Breuil C (1995) Characterization of the cleavage specificity of a subtilisin-like serine protease from *Ophiostoma piceae* by liquid chromatography/mass spectrometry and tandem MS. FEBS Lett 374:208–210

Aggeler R, Coons J, Taylor SW, Ghosh SS, Garcia JJ, Capaldi RA, Marusich MF (2002) A functionally active human F1F0 ATPase can be purified by immunocapture from heart tissue and fibroblast cell lines. Subunit structure and activity studies. J Biol Chem 277:33906–33912

Boone DR, Liu Y, Zhao ZJ, Balkwill DL, Drake GR, Stevens TO, Aldrich HC (1995) *Bacillus infernus* sp. nov., an Fe(III)- and Mn (IV)-reducing anaerobe from the deep terrestrial subsurface. Int J Syst Bacteriol 45:441–448

Cleveland D (1983) Peptide mapping in one dimension by limited proteolysis of sodium dodecyl sulfate-solubilized proteins. Methods Enzymol 96:222–229

Egmond MR (1997) Application of proteases in detergents. In: Van Ee JH, Misset O, Baas EJ (eds) Enzymes in detergency. Marcel Dekker, New York, pp 61–74

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 alkaliphilic *Bacillus* sp. KSM-K16. J Ferment Bioeng 78:105–108

Hoppe HP (1991) Microbial extracellular enzyme activity: a new key parameter in aquatic ecology. In: RG Chrost (ed) Microbial enzymes in aquatic environment. Springer, Berlin Heidelberg New York, pp 61–80

Horikoshi K (1971) Production of alkaline enzymes by alkaliphilic microorganisms. Part I. Alkaline protease produced by *Bacillus* no. 221. Biosci Biotechnol Biochem 35:1407–1414

Horikoshi K (1999) Alkaliphiles: some applications of their products for biotechnology. Microbiol Mol Biol Rev 63:735–750

Jang HJ, Kim BC, Pyun YR, Kim YS (2002a) A novel subtilisin-like serine protease from *Thermoanaerobacter yonseiensis* KB-1: its cloning, expression, and biochemical properties. Extremophiles 6:233–243

Jang HJ, Lee CH, Lee W, Kim YS (2002b) Two flexible loops in subtilisin-like thermophilic protease, thermicin, from *Thermoanaerobacter yonseiensis*. J Biochem Mol Biol 35:498–507

Johansen JT, Ottesen M, Svendsen I, Wybrandt G (1968) The degradation of the B-chain of oxidized insulin by two subtilisins and their succinylated and N-carbamylated derivatives. C R Trav Lab Carlsberg 36:365–384

Kieft TL, Fredrickson JK, Onstott TC, Gorby YA, Kostandarithes HM, Bailey TJ, Kennedy DW, Li SW, Plymale AE, Spadoni CM, Gray MS (1999) Dissimilatory reduction of Fe(III) and other electron acceptors by a *Thermus* isolate. Appl Environ Microbiol 65:1214–1221

Kobayashi T, Hakamada Y, Adachi S, Hitomi J, Yoshimatsu T, Koike K, Kawai S, Ito S (1995) Purification and 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

Lu J, Nakajima-Kambe T, Shigeno T, Ohbo A, Nomura N, Nakahara T (1999) Biodegradation of dibenzothiophene and 4,6-dimethyldibenzothiophene by *Sphingomonas paucimobilis* strain TZS-7. J Biosci Bioeng 88:293–299

Morihara K, Oka T, Tsuzuki H (1969) Comparison of alpha-chymotrypsin and subtilisin BPN′: size and specificity of the active site. Biochem Biophys Res Commun 35:210–214

Nakanishi T, Yamamoto T (1974) Action and specificity of a streptomyces alkaliphilic proteinase. Biosci Biotechnol Biochem 38:2391–2397

Narinx E, Davail S, Feller G, Gerday C (1992) Nucleotide and derived amino acid sequence of the subtilisin from the Antarctic psychrotroph *Bacillus* TA39. Biochim Biophys Acta 1131:111–113

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6

Okuda M, Sumitomo N, Takimura Y, Ogawa A, Saeki K, Kawai S, Kobayashi T, Ito S (2004) A new subtilisin family: nucleotide

and deduced amino acid sequences of new high-molecular-mass alkaline proteases from *Bacillus* spp. Extremophiles 8:229–235

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, 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

Schechter I, Berger A (1967) On the sites of the active site in protease. I. Papain. Biochem Biophys Res Commun 27:157–162

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

Schmidt BF, Woodhouse L, Adams RM, Ward T, Mainzer SE, Lad PJ (1995) Alkalophilic *Bacillus* sp. strain LG12 has a series of serine protease genes. Appl Environ Microbiol 61:4490–4493

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

Takai K, Horikoshi K (2000) *Thermosipho japonicus* sp. nov., an extremely thermophilic bacterium isolated from a deep-sea hydrothermal vent in Japan. Extremophiles 4:9–17

Takai K, Moser DP, Onstott TC, Spoelstra N, Pfiffner SM, Dohnalkova A, Fredrickson JK (2001) *Alkaliphilus transvaalensis* gen. nov., sp. nov., an extremely alkaliphilic bacterium isolated from a deep South African gold mine. Int J Syst Evol Microbiol 51:1245–1256

Takami H, Akiba T, Horikoshi K (1989) Production of extremely thermostable alkaline protease from *Bacillus* sp. no. AH-101. Appl Microbiol Biotechnol 30:120–124

Takami H, Akiba T, Horikoshi K (1992a) Substrate specificity of thermostable alkaline protease from *Bacillus* sp. no. AH-101. Biosci Biotechnol Biochem 56:333–334

Takami H, Kobayashi T, Kobayashi M, Yamamoto M, Nakamura S, Aono R, Horikoshi K (1992b) Molecular cloning, nucleotide sequence, and expression of the structural gene for alkaline serine protease from alkaliphilic *Bacillus* sp. 221. Biosci Biotechnol Biochem 56:1455–1460

Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247

Wati MR, Thanabalu T, Porter AG (1997) Gene from tropical *Bacillus sphaericus* encoding a protease closely related to subtilisins from Antarctic bacilli. Biochim Biophys Acta 1352:56–62

Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci USA 95:6578–6583

Yamagata Y, Isshiki K, Ichishima E (1995) Subtilisin Sendai from alkalophilic *Bacillus* sp.: molecular and enzymatic properties of the enzyme and molecular cloning and characterization of the gene, *apr*S. Enzyme Microb Technol 17:653–663