Proteolysin, a Novel Highly Thermostable and Cosolvent-Compatible Protease from the Thermophilic Bacterium *Coprothermobacter proteolyticus*

Ana Toplak,1 Bian Wu,1 Fabrizia Fusetti,2 Peter J.L.M. Quaedflieg,3 Dick B. Janssen1*

1 Biotransformation and Biocatalysis, Groningen Biomolecular Science and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

2 Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

3 DSM Innovative Synthesis, P.O. Box 18, 6160 MD Geleen, The Netherlands

Running Title: Proteolysin from *Coprothermobacter proteolyticus*

Keywords: protease, thermostability, subtilase, subtilisin, thermophiles

* Corresponding author. Mailing address: Biotransformation and Biocatalysis, Groningen Biomolecular Science and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. E-mail: d.b.janssen@rug.nl. Phone: +31-50-3634008; Fax: +31-50-3634165.
Abstract

Through genome mining we identified a gene encoding a putative serine protease of the thermitase subgroup of subtilases (E.C. 3.4.21.66) in the thermophilic bacterium *Coprothermobacter proteolyticus*. The gene was functionally expressed in *Escherichia coli* and the enzyme which we called proteolysin, was purified to near homogeneity from crude cell lysate by a single heat-treatment step. Proteolysin has a broad pH tolerance and is active at temperatures of up to 80°C. In addition, the enzyme shows good activity and stability in the presence of organic solvents, detergents and dithiothreitol, and remains active in 6 M guanidinium hydrochloride. Based on its stability and an activity profile, proteolysin can be an excellent candidate for applications where resistance to harsh process conditions is required.

Introduction

Extreme ecosystems in terms of temperature, pressure, salinity and pH represent a valuable resource of microorganisms that produce thermostable enzymes (1). The high stability of such thermozymes (enzymes from hyper- and extreme-thermophiles) is of great importance for developing new industrial applications (3, 16). Thermozymes not only show resistance to high temperatures, but also often tolerate harsh conditions such as the presence of denaturants and organic solvents (8). Industrial processes that operate under these conditions benefit from enhanced substrate solubility, lowered liquid viscosity, increased diffusion rates of substrates and products, as well as from favorable equilibrium shifts in certain endothermal reactions. In addition, thermostable enzymes tend to be more resistant to proteolysis (9), and may be compatible with heat-treatment aimed at reducing microbial contamination (10, 5).
Subtilisins are serine proteases from *Bacillus* strains; they comprise the largest group of commercial proteolytic enzymes (36) and account for more than half of the world total sales of enzymes (23). They are extensively used in food, textile, detergent, pharmaceutical and leather industries (14, 18, 19, 33). A few thermostable subtilisin homologs (subtilases) have been isolated and characterized, for example from Archaea (e.g. pyrolysin (58), stetterlysin (57), pernisine (7), Tk-SP (12), Tk-1689 and Tk-subtilisin (13, 21)), and from thermophilic bacteria (fervidolysin (25), Ak.1 protease (54), aqualysin I (30), islandisin (15) and thermitase (52)). Since cultivation of extremophiles is associated with potential difficulties, cloning and expression of protease genes into a mesophilic host that is easy to grow (*E. coli, B. subtilis*, yeast) is of importance for biochemical investigation, protein engineering studies and practical enzyme production. Another possible advantage of using a mesophilic host is that enzyme isolation can be based on the difference in thermostability between host proteins and the target thermozyme (45). Unfortunately, expression levels in mesophilic hosts using standard expression systems rarely exceed a few mg per L of culture. Autocatalytic processing and secretion may be specific bottlenecks (6, 21). Most subtilases are synthesized as a precursor with a pre- and a prosequence which are present as an N-terminal extension of the mature catalytic protein. Occasionally, there is a C-terminal extension as well. The N-terminal prosequence serves as a signal peptide for the translocation through the cell membrane. By serving as intramolecular chaperone and inhibitor of the proteolytic activity the prosequence is involved in folding of the mature enzyme (6,20). The C-terminal prodomain is involved in secretion to the extracellular medium in the original host (28), whereas in *E. coli* it facilitates the translocation of the enzyme to the outer membrane. It can be removed during a heating step, as shown for aqualysin I (22, 53).
In order to discover novel thermostable subtilases that resist harsh process conditions, we explored available genome sequences from thermophiles, hyperthermophiles and bioremediation organisms for homologs of subtilisin E. In the current manuscript we describe the cloning of 6 putative subtilases and focus on an active subtilase from *Coprothermobacter proteolyticus*, a Gram-positive anaerobic extreme thermophile. This organism was isolated by Klingeberg *et al.* (24) and previously classified as *Thermobacteroides proteolyticus*. The strain may be applied as a candidate for thermophilic organic solid waste degradation (44), making characterization of its proteases of special interest. A protease preparation obtained from this strain by the strain exhibits maximal activity at 85°C and pH 9.5, suggesting the presence of a thermostable enzyme that may act at high pH (24). We here report the production of a *C. proteolyticus* protease in *E. coli*, its isolation by a one-step heat treatment, and its catalytic and biochemical properties. The results show that this novel subtilase, which we termed proteolysin, is a highly thermostable enzyme that shows excellent activity under harsh reaction conditions.

**Materials and methods**

**Chemicals and reagents.** Restriction enzymes were from New England BioLabs (Beverly, MA, USA). The PCR system was supplied by Finnzymes (Vantaa, Finland), and the LigaFast ligase system and trypsin were from Promega (Madison, WI, USA). All chemicals and subtilisin A type VIII (Carlsberg, Alcalase) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Enhanced chemiluminescence (ECL) Western-blotting detection reagent was from Amersham Biosciences (Piscataway, NJ, USA). Antiserum against the hexahistidine tag sequence was from Abcam (Cambridge, UK).
and secondary polyclonal peroxidase-conjugated anti-sheep IgG (rabbit) antibody was from Rockland (Gilbertsville, PA, USA).

**Bacterial strains and plasmids.** Genomic DNA of *Thermus aquaticus* Y51MC23 was from Lucigen (Middleton, WI, USA). *Pseudomonas mendocina* strain ymp genomic DNA was a kind gift from Prof. J.L. DuBois (Notre Dame University, IN, USA) and genomic DNA of *Geobacillus thermodenitrificans* NG80-2 was a kind gift from Prof. X. Liu (Nankai University, Tianjin, China). Cells of *Deinococcus geothermalis* (DSMZ 11300) and *C. proteolyticus* (DSM5265) were purchased from DSMZ (Braunschweig, Germany). *E. coli* strain C43(DE3) was used for cloning and expression and was from Lucigen (Middleton, WI, USA). Cloning plasmids included pBADMycHisA (Invitrogen, San Diego, CA, USA), pBADMycHisA (NdeI) with the NcoI site mutated to an NdeI site, pET28a+ (Novagen, Merck KGaA, Darmstadt, Germany) and pBADMBP, a derivative of plasmid pBADMycHisA (NdeI) with an additional maltose binding protein (MBP) gene to which the 5′end of the target gene is connected via a linker.

**Cloning of the protease genes.** The gene coding for the whole preprosequence of proteolysin (EMBL accession no. YP_002247839) was amplified by whole-cell PCR on *C. proteolyticus* cells using the forward primer 5′-

 GCCCGCGCCATATGAAAAAGATACTATTAACACTGTTATCG-3′ (NdeI site underlined) and the reverse primer 5′-CACACACGAAGCTTTTATGTTATGATCCAGTTTACTGACGAC-3′ (HindIII site underlined, stop codon changed to Leu codon in bold). The PCR products were ligated in NdeI- and HindIII-treated pBAD-MBP vector, and transformed to *E. coli* C43(DE3). This yielded a recombinant plasmid termed pBAD-MBP-PrlA encoding the maltose binding protein, the signal peptide, the N-terminal prosequence, the mature protein and a putative C-terminal...
region ending with a linker (LKLGPEQKLISEEDLNSAVD) and a hexahistidine tag. In addition, recombinant plasmid lacking sequence for the C-terminal hexahistidine tag (PrlAs) was made using the reverse primer 5'-CACACACGAAGCTTGTTAATG-GTGTCAGTTTACTGCAGCATAC-3' (HindIII site underlined, stop codon in bold). A proteolysin mutant (PrlAd) harboring the substitutions Cys182Ala and Cys201Ala was obtained by site-directed mutagenesis using the QuikChange kit (Stratagene) and pBAD-MBP-PrlA as template.

The genes for other putative extremophilic proteases were cloned in a similar manner using either cells or genomic DNA as the template. Genes for YP_604447 (*D. geothermalis*), YP_001189588 (*P. mendocina ymp*) and YP_001125483 (*G. thermodenitrificans* NG80-2) were cloned in NdeI- and HindIII- treated pBAD-MycHisA(NdeI). The gene for YP_001127191 (*G. thermodenitrificans* NG80-2) was cloned between NcoI and HindIII sites in pBADMycHisA and ZP_03495941 (*T. aquaticus* Y51MC23) was cloned into the NdeI and HindIII sites in pET28a+. The sequence of the cloned genes was confirmed by DNA sequencing (GATC, Konstanz, Germany).

**Activity screening on plates.** For testing protease expression, transformants were grown on LB-agar plates containing 1% skim milk, appropriate antibiotic and inducer (0.025% arabinose or 0.2 mM IPTG) for 48 h at 30°C. Activity was apparent from halo formation around colonies. Colonies were also subjected to heat treatment (1-2 h at 80°C) to identify bacteria carrying cloned sequences on recombinant plasmid that expressed proteases that require high-temperature maturation.

**Expression, processing and isolation of proteolysin.** *E. coli* C43(DE3) cells containing an expression construct were grown overnight at 37°C in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) supplemented with 0.1% glucose and
ampicillin (100 µg/ml). The overnight culture was used to inoculate (0.1% v/v) TB medium (12 g/L tryptone, 24 g/L yeast extract, 0.4% (v/v) glycerol, 72 mM K₂HPO₄, 17 mM KH₂PO₄) containing ampicillin (100 µg/ml) and cells were grown at 37°C. Proteolysin synthesis was induced at 17°C when the OD₆₀₀ reached 0.8 by adding 0.25% arabinose, followed by cultivation for 48 h at 17°C with shaking. The cells were harvested (4°C, 15 min at 6,000 x g), washed, suspended in buffer (50 mM Hepes·NaOH, pH 7.5, 10 mM CaCl₂) and disrupted by sonication. The total cell lysate was subjected to heat treatment at 80°C for 3 h and centrifuged (4°C, 45 min at 30,000 x g). The supernatant was concentrated with an Amicon YM30 ultrafiltration membrane (Millipore, Billerica, MA, USA), followed by buffer exchange to 50 mM Hepes·NaOH, pH 7.5, 10 mM CaCl₂. Protein concentrations were determined by the Bradford assay using bovine serum albumin as the standard. The mature protein was stored at -80°C.

**SDS-PAGE and immunoblotting.** Samples of total cell lysate taken during proteolysin processing were analyzed on 12% SDS-PAGE gels. Proteins were transferred to a nitrocellulose membrane (Amersham Biosciences) using a semidry blotting apparatus (Bio-Rad, Hercules, CA, USA). Immunodetection was performed using antiserum against the hexahistidine tag, a secondary horseradish peroxidase-coupled antiserum and the ECL system from Amersham Biosciences according to the instructions of the manufacturer. Photographs were taken using a Fujifilm LAS-3000 imaging system.

**Enzyme assays.** Proteolytic activities of cell extracts and purified enzyme samples were monitored on 5% skim milk LB agar plates. Samples were added to wells in the agar and proteolytic activity was scored after overnight incubation at 37°C by inspection for clearing zones. For further analysis, an azocasein assay (13) and an
amidolytic assay (50) were employed. For the first, an azocasein solution (1.25% 
(w/v) in Hepes·NaOH buffer (100 mM, pH 7.5) containing CaCl$_2$ (1 mM)) was used as 
the substrate. After 20 min incubation with enzyme at a selected temperature, 
trichloroacetic acid was added to quench the reaction (4%, v/v). The mixture was 
cooled on ice for 10 min and spun down in a tabletop centrifuge (4°C, 10 min, 14,000 
x g). Next, NaOH was added to the supernatant to 0.4 M final concentration and the 
absorbance was measured at 440 nm. Reactions were performed in triplicate. One unit 
of activity (U) is defined as the amount causing a change of one absorbance unit (AU) 
per min. In the amidolytic activity assay 1 mM N-succinyl-Ala-Ala-Pro-Phe-$p$-
nitroanilide ($N$-suc-AAPF-$p$NP) was used as the substrate, added from a 20 mM stock 
solution in DMSO. Reactions were performed in Hepes·NaOH buffer (100 mM, pH 7.5) containing CaCl$_2$ (1 mM) and DMSO (10%, v/v) at 40°C and 70°C. The release of 
$p$-nitroaniline was monitored by spectrophotometry at 410 nm ($\varepsilon = 9.4$ mM$^{-1}$·cm$^{-1}$ (at 40°C), $\varepsilon = 11.5$ mM$^{-1}$·cm$^{-1}$ (at 70°C). One unit is defined as the amount of activity that 
produces 1 µmol $p$-nitroaniline per min. For determination of kinetic parameters, the 
substrate concentration was varied from 0.03 - 3 mM.

The pH optimum and cosolvent tolerance were determined with the amidolytic 
assay, using Britton-Robinson universal buffer (0.04 M acetic acid, 0.04 M H$_3$BO$_3$, 
0.04 M H$_3$PO$_4$) in a pH range of 3.0 – 12.0 (NaOH). The concentration of cosolvents 
(DMSO, DMF or ethanol) was varied from 10 to 60% (v/v). Enzyme activities were 
normalized to the activity in the presence of 10% of the respective cosolvent.

The selectivity of proteases was measured with oxidized insulin B chain. The 
substrate was dissolved to 10 mg/ml in 0.05 mM NH$_4$HCO$_3$, pH 8.0. Substrate to 
enzyme ratios were set to 1200 or 2400 (w/w). The reaction was started by addition of 
enzyme, and continued at 37°C and 60°C. Samples were taken at 0 min, 5 min and 24
h, conversion was stopped by adding 5% TFA, and peptides were analyzed by LC-MS/MS to determine the cleavage preference of proteolysin and subtilisin A.

**Inhibitors, surfactants, denaturants and metal ions.** The effects of surfactants and denaturants were tested by adding these compounds to reaction mixtures. The serine protease inhibitor phenylmethylsulfonyl fluoride (PMSF), EDTA, and iodoacetate were tested at concentrations of 5 and 10 mM. Metal ions were added as chloride salts of calcium, magnesium, manganese and nickel. Enzyme was preincubated with the respective reagents dissolved in 100 mM Hepes-NaOH, pH 7.5, for 1 h at 37°C, and residual activities were determined in duplicate with the amidolytic assay at 40°C where CaCl₂ was omitted from the assay buffer. The activity in the same buffer without potential inhibitor added was set at 100%.

**Thermostability.** The fluorescence-based thermal stability assay described by Ericsson et al. (5) was used to determine apparent melting temperatures of the proteins. Protein solutions (15 µl) in buffer (25 mM Hepes-NaOH, pH 7.5, 1 mM CaCl₂, 20% glycerol) were incubated with PMSF (2.5 µl of 100 mM stock in ethanol) for 10 min at room temperature. A solution of 7.5 µl of 100 x Sypro Orange (Molecular Probes, Life Technologies, San Diego, CA) dye was added to the samples in a thin walled 96-well PCR plate. The plates were sealed with Optical-Quality Sealing Tape and heated in an iCycler iQ Real Time PCR Detection System (BioRad, Hercules, CA, USA) from 20 to 99°C at a heating rate of 1.1°C/min. The wavelengths for excitation and emission were 490 and 575 nm, respectively. Fluorescence changes were monitored with a charge-coupled device (CCD) camera and the derivative of the fluorescence change against time was used to obtain the Tm, app.

Enzyme thermostability was also determined by incubating the enzyme in Hepes-NaOH buffer (100 mM, pH 7.5) containing CaCl₂ (1 mM) at 70, 80, 90 and
100°C for different time periods. Thermostability in the presence of DTT (10 mM) at 70°C was also determined for pretreated enzyme with 10 mM DTT (1h incubation at 25°C). The residual activity was determined in duplicate with the amidolytic assay after cooling to 40°C.

**Sulfhydryl groups.** For quantitative determination of sulfhydryl groups in proteolysin, Ellman’s test was used as described by Riener et al. (41).

**Peptide analysis by mass spectrometry.** To allow confirmation of the identity of isolated proteolysin, a protein band was excised from a 12% SDS gel and tryptic digestion was performed according to Saller et al. (43). For LC-MS/MS analysis, tryptic peptides of proteolysin or oxidized insulin B chain digests were first diluted in 0.1% TFA, separated on capillary C18 LC column and analyzed on MALDI-TOF/TOF 4800 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA) as described by Saller et al. (43). Peptide identification was carried out using the program Protein Pilot 4.0 (ABsciex, Foster City, CA, USA) searching against the UniProtKB/Swiss-Prot protein database, to which the sequences of proteolysin, trypsin and keratin were added. In cleavage specificity experiments, the sequence of oxidized insulin B chain was added to the database. Matches to proteolysin- or insulin-derived peptides were accepted if they could be assigned with a confidence of identification probability higher than 99%.

For intact protein analysis, samples were mixed 1:1 (v/v) with a matrix solution consisting of saturated with α-cyano-4-hydroxycinnamic acid (14 mg/ml) in a 3:1:2 (v/v/v) mixture of formic acid/water/2-propanol (FWI) (pH 1.3), and analyzed after spotting on a stainless-steel MALDI target. Mass spectra were recorded by MALDI-TOF on the Applied Biosystems proteomics analyzer mentioned above,
which was operated in linear positive ionization mode. Bovine serum albumin was used for calibration.

**Results**

**Genome mining.** To discover thermostable subtilisin-like enzymes (subtilases) that might resist harsh process conditions, the entire preprosequence of subtilisin E (EMBL AAA22742) was used as a query in Blast searches of sequences of thermophilic bacteria, Archaea and bioremediation organisms (BLASTp, (2)).

After omitting previously characterized subtilases, the search revealed 24 putative protein sequences with more than 30% sequence identity with subtilisin E. From these we cloned 16 putative protease genes. Six recombinant *E. coli* C43(DE3) strains containing a cloned protease gene gave functional subtilase expression, as indicated by activity in casein plate assays (*T. aquaticus* Y51MC23, subtilase ZP_03495941; *D. geothermalis*, subtilase YP_604447; *G. thermodinitrificans* NG80-2, subtilases YP_001127191 and YP_001125483; *P. mendocina* ymp, subtilase YP_001189588; and *C. proteolyticus*, subtilase YP_002247839). We selected the protease from *C. proteolyticus* for further studies since the source organism exhibits high proteolytic activity (24) and grows at high temperature. The expressed protease has quite high sequence identity with subtilisin E (35%) and thermitase (46%). According to Siezen’s classification of the subtilase superfamily (46), this enzyme, which we termed proteolysin (PrlA), clusters in the family of thermitase (Fig. 1).

Inspection of sequence alignments of proteolysin with previously characterized subtilases revealed that the catalytic residues are fully conserved, *i.e.* the catalytic triad (Asp51, His94, and Ser267) and the oxyanion hole (Asn187). A structural model of proteolysin based on the crystal structure of thermitase (1THM,
(52)) was constructed with Yasara (26). The model showed that proteolysin can adopt
the same fold as subtilisin E. The active site geometry appears to be conserved and
two calcium sites (Ca1 and Ca2) present in thermitase are likely to be present in
proteolysin as well. Additional surface loops in the model originate from two
sequence insertions (Asn55-Lys65 and Asp231-Asp243). Two cysteines (Cys182 and
Cys201) are positioned in the inner core of the proteolysin model and fit the criteria
for disulfide bond formation (38,39). These cysteines are not conserved in the most
homologous subtilases.

**Production, processing and isolation of proteolysin using *E. coli***. For the
production of proteolysin in *E. coli*, plasmid pBAD-MBP-PrlA was constructed. A
dNA segment encoding maltose binding protein (MBP) was fused to the 5’ end of the
region encoding the preprosequence to improve solubility of the precursor, and
sequence for a C-terminal hexahistidine tag was added to facilitate purification. The
recombinant fusion protein was expressed in *E. coli* strain C43(DE3). No significant
activity was detected in the culture fluid. After sonication of the collected cells, a total
cell lysate (TCL) was obtained, which contained little proteolytic activity. Proteolysin
processing and maturation required a heat treatment step. The processing could be
performed at 70°C for 12 h or at 80°C for 3 h. We observed some product already
after 1 h of heat treatment at 80°C, but only after 3 h of heat treatment a single protein
band of approximately 40 kDa was obtained on SDS-PAGE gels (Supplemental
Material). For routine enzyme isolation, a heat treatment step of 180 min at 80°C was
used. The yield of purified proteolysin from a 1 L culture was 20 mg. The enzyme had
a specific activity of 4 U/mg in the amidolytic assay. The results show that
proteolysin was not secreted and processed by the *E. coli* host, and that the precursor
protein was autocatalytically converted to mature enzyme upon heat treatment.
The predicted amino acid sequence of the unprocessed recombinant proteolysin comprises 857 amino acids with a calculated total molecular mass of 92.8 kDa. The observed mass of the mature protein was 34,973 ± 3 Da, as determined by MALDI-TOF. To confirm the identity of the protein and to determine the N-terminus after maturation, tryptic digestion was performed and peptides were analyzed by MALDI-TOF/TOF. The peptide coverage was 67%, but no peptide fragments were found in the Met1-Phe113 region. The molecular mass of the mature protein suggests that the sequence starts at Asp114 and ends with three additional amino acids (Leu328, Lys329, and Leu330) originating from the linker between the protein and the hexahistidine tag. Immunoblot analysis with antibodies against the hexahistidine tag confirmed its removal upon maturation of the protein (data not shown). Moreover, sequence alignment to thermitase supports Asp114 as the start of the mature protein (see Supplemental Material).

To exclude an influence of the linker and hexahistidine tag on proteolysin processing, a gene encoding native proteolysin (termed PrlAs) including the stop codon was constructed. It was expressed and PrlAs was produced under the same conditions as above. The expression level and processing profile of the PrlAs, as well as its thermostability, were not impaired. The enzyme showed a similar activity as PrlA.

**Temperature profile and thermostability of proteolysin.** Azocasein and N-succ-AAPF-pNA were used as the substrate for determination of temperature effects. Proteolysin hydrolyzed azocasein with optimal activity at 85°C, whereas subtilisin A was most active at 70°C under the conditions applied (Fig. 2). The kinetic stability of proteolysin was investigated by following the loss of activity during incubation of the purified enzyme at high temperatures (70, 80, 90 and 100°C), using the amidolytic
assay (Fig. 3). Proteolysin appeared very stable at 70°C. Even after 20 h of incubation, the protein retained 35% of its initial activity towards \( N\)-suc-AAPF-pNA. In contrast, subtilisin A was completely inactivated after 1 h incubation at 70°C.

The thermostability of proteolysin was also determined using the Sypro Orange assay. For this, enzyme activity was inhibited with PMSF. An apparent transition temperature (\( T_{m,\text{app}} \)) of 81 °C was observed, again indicating that the enzyme is much more thermostable than subtilisin A, for which a \( T_{m,\text{app}} \) of 60 °C was measured.

A homology model for proteolysin was constructed with Yasara. The structural model suggested that a disulfide bond might be formed between Cys182 and Cys201, which are the only cysteines in the protein. Ellman’s assay revealed the presence of 22 \( \mu \)mol of free sulfhydryls per mg of purified enzyme, in agreement with a sulfhydryl to enzyme ratio of 0.7. In a modified Ellman’s test using cystamine (41), a sulfhydryl to enzyme ratio of 0.5 was found. These data suggest the occurrence of a disulfide bond or destruction of free sulfhydryls by a \( \beta \)-elimination reaction that may occur during heat treatment step (56). To determine if the presence of a putative disulfide bond contributes to thermostability, a mutant carrying the substitutions Cys182Ala and Cys201Ala was constructed. The resulting variant (PrlAd) was expressed and produced under the conditions mentioned above, giving a four-fold lower protein yield as compared to proteolysin PrlA. PrlAd hydrolyzed azocasein with optimal apparent activity at around 90°C, slightly higher than what was found with the wild-type (Fig. 2). The thermostability of variant PrlAd was only slightly affected by the mutations (Fig. 3). Whereas the double mutant was more stable at low temperatures (70°C), the wild type showed slower inactivation at temperatures above 80°C. The presence of DTT (10 mM) reduced the half-life at 70°C by three-fold for...
both enzymes (PrlA and PrlAd, data not shown). This suggests that the Cys182-
Cys201 pair does not contribute to the high thermostability of proteolysin.

**Catalytic properties.** The catalytic properties of the novel protease were
examined with N-suc-AAPF-pNP, a short peptide (oxidized insulin B chain), and
azocasein (24.6 kDa), which is an azo-dye conjugate of casein, a large globular
protein widely used in the food industry as a source of biologically active peptides
(47). The optimum pH of proteolysin was determined at 40°C and 70°C using N-suc-
AAPF-pNP as the substrate. Proteolysin appeared active over a broad pH range (7.0-
9.0) and showed maximal activity at approximately pH 8.0 (Supplemental Material).

The steady-state kinetic parameters of proteolysin PrlA, PrlAd and subtilisin A
were determined with N-suc-AAPF-pNA and compared with those of other subtilases
(Table 1). At 40°C proteolysin PrlA had a lower catalytic efficiency than its
mesophilic counterparts. At elevated temperature, a four-fold increase of the $k_{cat}$ of
proteolysin was accompanied by a similar increase in $K_m$, suggesting that the higher
catalytic activity at elevated temperature is accompanied by less tight substrate
binding. Interestingly, at 70°C the proteolysin variant lacking two cysteines (PrlAd)
exhibited a catalytic efficiency of 335 mM$^{-1}$·s$^{-1}$, which is 5-fold higher than that of the
wild-type enzyme and similar to the $k_{cat}/K_m$ of subtilisin A with the same substrate at
40°C. Subtilisin A and the mutant proteolysin had similar activities at 70°C, but
subtilisin was not very stable at this high temperature and its activity was lost after 1 h
incubation (Fig. 3).

At 40°C, both proteolysin PrlA and PrlAd showed slightly higher $k_{cat}$ values in
the presence of the reducing agent dithithreitol (DTT), but the increase in $K_m$ suggests
weaker substrate binding (Table 1). High concentrations of DDT sometimes inactivate
enzymes by metal chelation or by reducing essential disulfide bonds. Proteolysin and
the mutant lacking two cysteines tolerated DTT concentrations up to 10 mM, indicating that in mature proteolysin the two cysteines do not form an important DTT-reducible disulfide bond that is important for activity.

**Cleavage specificity.** Oxidized insulin B chain was used as the substrate for determining the cleavage specificity of proteolysin. The reactions were performed at 37°C and the hydrolysis products were analyzed by LC-MS/MS. The primary cleavage site of proteolysin was found between Ala14 and Leu15 (Fig. 4). Prolonged incubation (24 h) revealed additional cleavage sites after Phe1, Asn3, Gln4, Leu17, Cys19, Phe24, Phe25, Tyr25 and Lys29, which shows that proteolysin has a relaxed cleavage site specificity. For subtilisin A, the primary cleavage site was between Leu15 and Tyr16 and prolonged incubation resulted in complete degradation of the substrate chain. The cleavage profile of proteolysin acting on insulin B chain did not change when the reaction temperature was raised to 60°C, whereas subtilisin A was inactivated after a prolonged incubation. When compared to other thermostable subtilases, proteolysin shares the primary cleavage site preference with pernisine (35) and shows a slight preference for hydrophobic amino acids residues at the P1 position, which is similar to other thermostable subtilases (31).

**Metal ions and inhibitors.** In most thermostable subtilases a bound calcium ion is essential for enzyme stability (27). The number of calcium-binding sites varies from zero (Tk-SP) to seven (Tk-subtilisin) (49). The structural models suggest that proteolysin may have two strong calcium-binding sites which are also present in thermitase (17). Therefore, the effect of various cations on the amidolytic reaction was tested. The results showed that bivalent metal cations increased the activity of proteolysin (Fig. 5), whereas the metal chelator EDTA (10 mM) decreased the activity.
to 87%. The serine protease inhibitor PMSF completely abolished the activity of proteolysin.

**Proteolysis under harsh conditions.** The addition of cosolvents such as DMSO or DMF can be used to improve the solubility of peptides. Therefore, we tested the tolerance of proteolysin to DMSO, DMF and ethanol. Amidolytic assays were performed in the presence of various levels of cosolvents in the range of 10-60% (v/v). Proteolysin exhibits higher organic solvent tolerance than subtilisin A (Fig. 6). However, even though subtilisin A is more susceptible to inactivation by cosolvents than proteolysin, the higher catalytic activity towards the model substrate compensates for this and subtilisin A it still displays a higher catalytic activity in the presence of cosolvents.

Surface-active compounds or denaturing agents can be added to protein-containing substrates in order to increase substrate availability. We observed that proteolysin can tolerate up to 6 M guanidinium hydrochloride, which gave a 40% reduction of activity after 1 h incubation at 40°C. Subtilisin A was completely inactivated under the same conditions. The anionic surfactant SDS (10%, w/v) and the cationic surfactant CTAB (5%, v/v) reduced the activity of proteolysin, but the non-ionic surfactant Tween 20 (10%, v/v) activated proteolysin up to 178%. Moreover, after 1 h of incubation in the presence of the oxidizing agent hydrogen peroxide (15% v/v), the residual activity dropped to 10% (Fig. 5).

**DISCUSSION**

Enzymes from extremophiles offer the opportunity to expand the reaction conditions compatible with biocatalytic conversions (4). The wealth of microbial genome sequence information allows the exploration of extremophiles for genes
encoding proteins with expected high thermostability. In this work, we report the expression in *E. coli* of six novel thermostable subtilases identified by genome mining. We also describe the purification and biochemical characterization of proteolysin, one of the two putative extracellular proteases encoded in the genome of the extreme thermophile *C. proteolyticus* (51). The organism was isolated from a thermophilic digester (55°C) fermenting tannery waste and cattle manure (34), and plays an important role in biogas production by releasing hydrogen that is used by methanogenic organisms (51).

Proteolysin belongs to the thermitase subgroup of the superfamily of subtilases. Produced as a precursor in *E. coli* cells, it requires processing for conversion to the mature form. The isolated yield of proteolysin produced in the mesophilic host *E. coli* was about 20 mg/L culture, which is sufficient for characterizing the enzyme and mutants thereof. The enzyme can be rapidly isolated by simultaneous processing and purification in a single heat-treatment step. Like many extracellular enzymes from thermophiles (55), proteolysin is active at temperatures higher than the optimum growth temperature of its *C. proteolyticus* host (63°C). Proteolysin can hydrolyze the large globular protein azocasein at temperatures as high as 90°C. The activity towards the synthetic substrate *N*-suc-AAPF-pNA is comparable to that of other thermophilic subtilases (13,30). At elevated temperature (70-80°C), the enzyme is stable and displays a higher catalytic activity than at 40°C. By replacing both cysteines we obtained a variant with five-fold increased catalytic efficiency at 40°C and with a similar thermostability and temperature/activity profile as the wild-type PrlA.

Thermostability is often correlated with high tolerance to organic solvents (29). Proteolysin indeed tolerated organic cosolvents in hydrolytic reactions much
better than commercially available subtilisin A. Because of its thermostability, relaxed
specificity and resistance to routinely used protein denaturants and DTT, proteolysin
is a candidate for proteomics studies when protein digestion at extreme conditions is
required. Heat pretreatment and the use of additives such as surfactants, organic
solvents, and urea are commonly used to improve protein solubility and facilitate
complete digestion, e.g. for peptide mapping (42,59). In conclusion, the convenient
production in E. coli, the high thermostability, the broad pH range, and its high
tolerance to surfactants and cosolvents make proteolysin an attractive candidate for
proteolysis under demanding reaction conditions.

Acknowledgements

We thank Dr. Hein J. Wijma for useful discussions and for making the model
of proteolysin. This project is part of the Integration of Biosynthesis and Organic
Synthesis program (IBOS-2; project 053.63.014) funded by The Netherlands
Organisation for Scientific Research (NWO) and Advanced Chemical Technologies
for Sustainability (ACTS). Peter J.L.M. Quaedflieg is employed by DSM. Some of
DSM's products are manufactured with the use of proteases.
References


FIG. 1. Phylogenetic tree based on the sequence alignment of subtilases. Enzymes expressed in active form in this study are indicated with an asterisk. Proteolysin from *Coprothermobacter proteolyticus* (in bold) clusters in the thermitase family. The figure was generated with Geneious v5.6 software (http://www.geneious.com).

FIG. 2. Temperature profile of subtilisin A (▼), proteolysin PrlA (●), and proteolysin variant PrlAd (○). The apparent temperature optimum was determined in duplicate using the azocasein assay.

FIG. 3. Thermostability of proteolysin and subtilisin A. Enzymes were incubated at temperatures of 70°C (●), 80°C (○), 90°C (▼) or 100°C (∆). The remaining activity at different times was determined with N-suc-AAPF-pNA as the substrate. Panel A, Proteolysin, PrlA; Panel B, Proteolysin, PrlAd; Panel C, subtilisin A.

FIG. 4. Cleavage sites of proteolysin in oxidized insulin B chain. The diagram shows peptides identified by LC-MS/MS after 5 min and 24 h incubations with proteolysin at 37°C. Major and minor cleavage sites are denoted by large and small arrowheads, respectively.

FIG. 5. The effect of different agents on proteolysin PrlA activity. The residual activity after 1 h incubation at 37°C with the compound indicated was measured with N-suc-AAPF-pNP as the substrate in duplicate at 40°C. Activity with 10 mM additive is indicated with black bars, with 5 mM as grey bars. Activity with surfactants is indicated with white bars.
FIG. 6. Cosolvent tolerance. The effect of DMSO, DMF, and ethanol on the activity of proteolysin (black bars) and subtilisin A (grey bars). Activities were determined with N-succ-AAPF-pNA as the substrate at 40°C. Activities are normalized to the activity in the presence of 10% cosolvent.
<table>
<thead>
<tr>
<th>Temperature</th>
<th>40°C</th>
<th>70°C</th>
<th>DTT (40°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Enzyme</strong></td>
<td>$k_{cat}$</td>
<td>$K_m$</td>
<td>$k_{cat}/K_m$</td>
</tr>
<tr>
<td>Proteolysin PrtA</td>
<td>25±1</td>
<td>0.4±0.07</td>
<td>0.58±0.13</td>
</tr>
<tr>
<td>Proteolysin PrtAd</td>
<td>111±4</td>
<td>0.3±0.03</td>
<td>1.96±0.3</td>
</tr>
<tr>
<td>Subtilisin A</td>
<td>551±62</td>
<td>1.1±0.3</td>
<td>539±70</td>
</tr>
<tr>
<td>Aqualysin I</td>
<td>33</td>
<td>1.2</td>
<td>27.5</td>
</tr>
<tr>
<td>Subtilisin BPN’</td>
<td>480</td>
<td>0.29</td>
<td>1655</td>
</tr>
</tbody>
</table>

ND, not determined

TABLE 1. Kinetic parameters of proteases measured with N-suc-AAPF-pNA as the substrate. Reaction temperatures were 40°C or 70°C. For assays in the presence of DTT, proteolysins were preincubated with 10 mM DTT at 25°C for 1 h, after which activities were determined. All incubations were performed in 100 mM Hepes-NaOH buffer, pH 7.5, containing 1 mM CaCl$_2$, 10% DMSO (v/v) and where indicated 10 mM DTT.