Isolation and Characterization of the Hyperthermostable Serine Protease, Pyrolysin, and Its Gene from the Hyperthermophilic Archaeon Pyrococcus furiosus*

(Received for publication, February 28, 1996, and in revised form, April 26, 1996)

Wilfried G. B. Voorhorst‡, Rik I. L. Eggen‡§, Ans C. M. Geerling‡, Christ Platteeuw‡†, Roland J. Siezen‡¶, and Willem M. de Vos‡¶†

From the ‡Department of Microbiology, Wageningen Agricultural University, 6703 CT Wageningen, and the ¶Department of Biophysical Chemistry, NIZO Ede, The Netherlands

The hyperthermostable serine protease pyrolysin from the hyperthermophilic archaeon Pyrococcus furiosus was purified from membrane fractions. Two proteolytically active fractions were obtained, designated high (HMW) and low (LMW) molecular weight pyrolysin, that showed immunological cross-reaction and identical NH2-terminal sequences in which the third residue could be glycosylated. The HMW pyrolysin showed a subunit mass of 150 kDa after acid denaturation. Incubation of HMW pyrolysin at 95 °C resulted in the formation of LMW pyrolysin, probably as a consequence of COOH-terminal autoproteolysis. The 4194-base pair pls gene encoding pyrolysin was isolated and characterized, and its transcription initiation site was identified. The deduced pyrolysin sequence indicated a prepro-enzyme organization, with a 1249-residue mature protein composed of an NH2-terminal catalytic domain with considerable homology to subtilisin-like serine proteases and a COOH-terminal domain that contained most of the 32 possible N-glycosylation sites. The archaeal pyrolysin showed highest homology with eucaryal tripeptidyl peptidases II on the amino acid level but a different cleavage specificity as shown by its endopeptidase activity toward caseins, casein fragments including αs1-casein and synthetic peptides.

Hyperthermophilic microorganisms with an optimum growth temperature of at least 80 °C have been studied in recent years to gain insight into biochemical adaptations to their extreme environment (1, 2). While few hyperthermophiles utilize carbohydrates, which in some cases are fermented via novel pathways (3, 4), most of these unusual microorganisms are capable of growing rapidly on proteins and peptides (2, 5). Several extracellular proteases that may be involved in the first step of protein utilization have been characterized, and all are derived from hyperthermophilic microorganisms that belong to the domain of the Archaea, the third lineage of life (6). A cell envelope-associated protease from Sulfolobus acidocaldarius has been purified and shown to belong to a new class of acid proteases based on its gene sequence that showed the presence of a signal sequence and several potential N-glycosylation sites (7, 8). Other hyperthermostable proteases that have been biochemically characterized include a thiol protease from Pyrococcus (9) and serine proteases from Desulforococcus (10), Thermococcus stetteri (11), and Pyrococcus furiosus (12–15). Moreover, a serine protease gene has been described in Pyrobaculum aerophilum, but its product has not yet been characterized (16).

All purified archaebacterial proteases show a high thermostability and thermoactivity, but the most thermostable protease to date is pyrolysin from P. furiosus with a half-life value of 4 h at the normal boiling point of water (12, 13). Here we describe the purification, post-translational modification, and substrate specificity of pyrolysin. Via reversed genetics the corresponding pls gene has been cloned, and its transcription initiation and nucleotide sequence were determined. The results indicate that pyrolysin has a prepro-enzyme structure and is a true endopeptidase belonging to the family of subtilisin-like serine proteases (17).

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Culture Conditions—P. furiosus (DSM3638) cells were cultivated at 95 °C in a 30-liter stainless steel fermentor (Bioengineering, Wald, Switzerland) with a 20-liter working volume, sparged with Nz gas. The growth medium consisted of synthetic seawater containing protease, amino acid, and sugar components (peptone 5 g/liter, yeast extract 1 g/liter) without elemental sulfur as described previously (12). Escherichia coli strain TG1 (18) was grown in L broth. The vector pUC19 (Pharmacia Biotech Inc.) was used for cloning and sequencing procedures.

Purification and Characterization of Pyrolysin—P. furiosus cells (35 g) were resuspended in 25 ml of P-buffer (50 mM sodium phosphate, pH 6.5), homogenized by sonication, and passed through a French press. The membrane fraction was separated from the cytoplasmic fraction by centrifugation (17,500 × g, 20 min), resuspended in 10 ml of P-buffer supplemented with 6 M urea, and incubated for 4 h at 95 °C. After centrifugation to remove the cell debris and denatured proteins, sample buffer (6 M urea, 0.1 M Tris-HCl, 0.9% l-mercaptoethanol, 20% glycerol, pH 6.8) was added to the supernatant that was subjected to preparative urea gel electrophoresis (urea-PAGE1) with the Bio-Rad model 491 Prep Cell (Bio-Rad). The gel (11% acrylamide, 10 × 2.5 cm) was made up to 6 M urea, and proteins were electroeluted in native PAGE-buffer (192 mM glycine, 25 mM Tris, pH 8.5) and analyzed for activity. Fractions containing proteolytic activity were pooled in a high molecular weight (HMW pyrolysin) and a low molecular weight fraction (LMW pyrolysin) and loaded onto an anion exchange column (1 × 3 cm) of Fractogel EMD trimethylaminoethyl-650 (M) (Merck) equilibrated in 50 mM Tris-HCl, pH 8.5. Subsequently, the buffer system of the anion-exchange column was changed to P-buffer, and after equilibration, the bound

---

1 The abbreviations used are: PAGE, polyacrylamide gel electrophoresis; HPLC, high pressure liquid chromatography; TFPs, tripeptidyl peptidases II; ORF, open reading frame; kb, kilobase pair(s).
proteins were eluted by a linear gradient 0–1.5 m NaCl in P-buffer. Fractions containing proteolytic activity were pooled, dialyzed, and concentrated using a diaflow cell with a cut-off value of 30 kDa (Filtron Technology Corp.). The NH₂-terminal amino acid sequences of the purified proteases were determined, either from immobilized proteins that had been separated by SDS-PAGE or directly from a purified preparation, by Edman degradation using an Applied Biosystems 477A Protein Sequencer (Applied Biosystems). Purified protease fractions were used to raise antibodies in rabbits, and crude serum was used in immunoblot experiments as described previously (19). Glycoproteins were stained following urea-SDS-PAGE by the periodic acid-Schiff technique using fuchsin-sulfite stain (Sigma-Aldrich) (20).

Polycrylamide Gel Electrophoresis—SDS-PAGE was carried out according to Laemmli (21) using 9.5% polyacrylamide gels. Acid denaturation of pyrolysin prior to electrophoresis was achieved by incubation in 5 M formic acid for 45 min at room temperature. Equimolar amounts of protein and pyrolyxide were used for subsequent neutralization. Analytical urea-SDS-PAGE was essentially the same as SDS-PAGE with the addition of 6 M urea, and samples were prepared as described for the preparative urea-PAGE using sample buffer supplemented with 5% SDS. A method was developed for the in situ detection of proteolytic active proteins on urea-(SDS)-PAGE using an overlay of Kodak X-OMAT AR film (Eastman Kodak Co.). Following a 10-min incubation at 95 °C the gelatin layer of the Kodak film was transferred to the gel and degraded on places where proteolytic activity was present. The nondegraded gelatin background was subsequently stained with Coomassie Brilliant Blue, resulting in a blue gel with cleared bands of activity.

Activity Assays and Substrate Specificity—The proteolytic activity of pyrolysin toward casein was assayed for 30 min at 95 °C as described previously (12). Total casein and purified α₂-casein, β-casein, and κ-casein were incubated with pyrolysin, and degradation products were analyzed by SDS-PAGE (22). The enzyme activity toward synthetic substrates (1 mM final concentration) was assayed in 0.5 ml of 0.1 M potassium phosphate buffer, pH 7.5, at 95 °C and followed spectrophotometrically at 405 nm on a Beckman DU7500 spectrophotometer (Beckman Instruments Inc.). Synthetic substrates were obtained from Bachem (Bubendorf, Switzerland) or from Chromogenix AB (Molndal, Sweden). The tripeptidyl-peptide substrate Ala-Ala-Phe-p-nitroanilide and its inhibitor, Arg-Ala-α-Ala-Val-Ala (where α-Ala is α-Ala, dehydroalanine), were kindly provided by B. Tomkinson (University of Uppsala, Sweden) (23, 24). The cleavage specificity of pyrolysin in the peptide α₂-casein-(1–23) was determined by HPLC analysis of initial cleavage products (22, 25) by incubating 25 ml of a solution of α₂-casein-(1–23) (10 mg/ml), 1 ml of the purified LMW pyrolysin (0.25 mg/ml), and 50 ml of 0.1 M potassium phosphate buffer, pH 7.5, at 95 °C. At various times, aliquots of 20 ml were cooled and supplemented with an equal volume of solvent B, composed of acetonitrile/water/trifluoroacetic acid (900:100:0.7, v/v). As a reference the degradation pattern of α₂-casein-(1–23) obtained by the PII-type protease from Lactococcus lactis was used as described previously (22, 25).

Cloning of the pls Gene Encoding Pyrolysin—A degenerated oligonucleotide primer was designed against the conserved region around the histidine active site residue of subtilisin-like serine proteases, with the sequence GTTCCGCACRTGTGTTGTCRTG-3’ (I, inosine; R, A, or G) (primer 1, Fig. 2). Additionally, the NH₂-terminal amino acid sequence of the purified pyrolysin was used to design a degenerated oligonucleotide primer against residues 5–11 with the sequence, ACTTGGGTATWAAY-GCIYT-3’ (I, inosine; W, A or T; Y, C or T) (primer 2, Fig. 2). PCR reactions were performed using 100 ng of both oligonucleotides and 250 ng of P. furiosus chromosomal DNA as template (19) in a final volume of 100 ml. After denaturation of the template (5 min 95°C), 30 cycles (1 min 95°C, 2 min 35°C, and 3 min 72°C) were performed, followed by an additional extension of 7 min at 72°C on a DNA Thermal Cycler (Perkin-Elmer). Southern hybridization of plasmid DNA. The nucleotide sequence accession number was submitted to the GenBank/EMBL Data Bank with the accession number U55855.

RESULTS

Purification of Pyrolysin—Cells of P. furiosus grown on peptides are known to contain high levels of pyrolysin, the highly thermostable serine protease that is mainly associated with the cell envelope (12). The membrane fraction of pyrococcal cells grown on peptone was enriched for pyrolysin via a precipitation in 6 M urea at 95 °C, during which the proteolytic activity remained constant while the total protein concentration decreased approximately 100-fold, most likely due to the proteolytic activity of pyrolysin. The remaining proteins were separated by preparative urea-PAGE followed by a final purification step using anion-exchange chromatography. This resulted in HMW and an LMW pyrolysin fraction (Fig. 1A) with apparent molecular masses on urea-SDS-PAGE of 105 and 80 kDa, respectively. Activity staining showed that both HMW and LMW pyrolysin each contain one proteolytically active band (Fig. 1B).

Characterization of HMW and LMW Pyrolysin—The NH₂-terminal amino acid sequences of both HMW (35 residues) and LMW (5 residues) pyrolysin were determined and found to be described previously (27). The purified RNA was used as template in primer extension experiments as described previously (19) with oligonucleotide, GTTGGTATGGATAGATAGTG-3’ (primer 3, Fig. 2). Sequence Analysis—Nucleotide sequence analysis of PCR products was carried out by the AmpliTaq cycle sequencing kit (Perkin-Elmer) and analyzed on a automated DNA sequencer (Applied Biosystems 373A). A Prism Ready Reaction DyeDeoxy Terminator cycle sequencing kit (Applied Biosystems) was used to sequence both strands of plasmid DNA.

Computer analysis of nucleotide and deduced amino acid sequences were carried out with the PC/GENE program version 5.01 (IntelliGenetics Inc.) and the GCG package version 7.0 (28) at the CAOS/CAMM Centre of the University of Nijmegen (The Netherlands).

Nucleotide Sequence Accession Number—The nucleotide sequence report submitted to the GenBank/EMBL Data Bank with the accession number U55855.

RNA Isolation and Primer Extension—Total RNA was isolated from P. furiosus cells grown on peptone using guanidinium isothiocyanate as
identical (Fig. 2). The relatedness of both pyrolysin fractions was supported by the observation that the third residue of their NH₂ termini could not be determined following Edman degradation, even after repeated trials (data not shown). Inspection of the amino acid sequence deduced from the pls gene (Fig. 2, see below) revealed that the third position of the mature pyrolysin is an asparagine that could be post-translationally modified, since it is the first residue of the well-known N-glycosylation motif Asn-X-(Ser/Thr) (29). Therefore, periodic acid-Schiff staining was performed on HMW and LMW pyrolysin separated by urea-SDS-PAGE (Fig. 1C). The results indicate that both proteins are glycosylated to a similar extent and that LMW is not a deglycosylated form of HMW pyrolysin. A mutual relation between the HMW and LMW pyrolysin could further be demonstrated by immunological cross-reaction using antisera raised against both purified pyrolysin fractions (data not shown). Moreover, a prolonged incubation of purified HMW pyrolysin at 95 °C resulted in its decrease and the formation of LMW pyrolysin. This is illustrated by the semi-quantitative in situ detection of a new proteolytic activity with the size of LMW that originated from HMW pyrolysin, indicating that LMW pyrolysin is a processing product of HMW pyrolysin (Fig. 1B). Finally, a multimer structure of HMW pyrolysin was further excluded by applying rigorous acid denaturation conditions followed by separation by urea-SDS-PAGE (Fig. 1D). HMW pyrolysin showed a single band with apparent molecular mass of 150 kDa. Two proteins were obtained after acid denaturation of the LMW pyrolysin fraction with apparent molecular masses of 130 and 105 kDa, the smallest of which is most likely a result of autoproteolytic processing of the 130-kDa protein.

Cloning and Characterization of the pls Gene Encoding Pyrolysin—A data base search revealed that the NH₂-terminal amino acid sequence of pyrolysin showed significant similarity with subtilisin-like serine proteases (17). This includes a potential active site aspartate residue at position 30 (Fig. 2, see below), suggesting that pyrolysin is a member of this family of serine proteases. Subtilisin-like serine proteases are characterized by the sequential order of the active site residues Asp, His, and Ser and the high degree of homology on amino acid level around these residues that form the catalytic triad (17). Two degenerated oligonucleotides based on the conserved region surrounding the active site histidine residue of subtilisin-like proteases and on the identified NH₂-terminal amino acid sequence (primers 1 and 2, Fig. 2) were used in a PCR reaction on chromosomal DNA from P. furiosus. Sequence analysis of the resulting 0.6-kb PCR fragment (data not shown) indicated that indeed a fragment of a subtilisin-like protease gene was ampli-
Characterization of Pyrolysin and Its Gene from *P. furiosus*

**Sequence Analysis of the *pls* Gene, Flanking Regions, and the *pls* Gene Product Pyrolysin—Sequence analysis of the *pls* gene, contained in pLUW701, pLUW702, and pLUW703, revealed an ORF of 4194 base pairs. The first ATG codon at position 316–320 of this ORF is preceded by a stretch of purine-rich nucleotides (at position 301–308) that may function as a ribosome binding site, suggesting that this is the translation initiation site of the *pls* gene (Fig. 2). The 3′ end of the *pls* gene was cloned from a lambda EMBL3 library consisting of Sau3A fragments of *P. furiosus* DNA. A pool of approximately 2400 independent plaques were screened by hybridization with *P. furiosus* genomic DNA digested with restriction enzymes. A strongly hybridizing 2.7-kb *HindIII* fragment was cloned into *HindIII*-digested pUC19, resulting in pLUW701 (Fig. 2). Sequence analysis of the insert of pLUW701 revealed that it contained an incomplete ORF, coding for the NH₂-terminal amino acids determined from the mature pyrolysin, however, without a potential translation start codon (Fig. 2). Southern hybridization of *P. furiosus* DNA with a pLUW701-derived probe showed a hybridizing 1.7-kb *SstI-PstI* fragment, which was cloned in *SstI-PstI*-digested pUC19, resulting in pLUW702 (Fig. 2). Sequence analysis of the insert of pLUW702 showed indeed the presence of the 5′ end of the *pls* gene (Fig. 2). The 3′ end of the *pls* gene was cloned from a lambda EMBL3 library consisting of 15–20-kb *SalI* fragments of *P. furiosus* DNA. A pool of approximately 2400 independent plaques were screened by hybridization with the *HindIII*-III fragment of pLUW701 resulting in 49 hybridizing phages. A 2.5-kb *EcoRI* fragment of one of these phages carrying the 3′ end of the *pls* gene was identified by hybridization and cloned into *EcoRI*-digested pUC19, resulting in pLUW703 (Fig. 2).

**Sequence Analysis of the *pls* Gene, Flanking Regions, and the *pls* Gene Product Pyrolysin—**Sequence analysis of the *pls* gene, contained in pLUW701, pLUW702, and pLUW703, revealed an ORF of 4194 base pairs. The first ATG codon at position 316–320 of this ORF is preceded by a stretch of purine-rich nucleotides (at position 301–308) that may function as a ribosome binding site, suggesting that this is the translation initiation site of the *pls* gene (Fig. 2). If so, the *pls* gene encodes a 1398-residue protein. The NH₂-terminal sequence determined from the purified pyrolysin was found to be encoded by nucleotides 763–868, indicating that the mature NH₂-terminus is preceded by a leader sequence of 149 amino acids (Fig. 2). The first 26 residues of this leader have the characteristics of a signal sequence with a high content of hydrophobic residues that can form a transmembrane-spanning helix and is followed by a putative consensus processing site (30). The remaining 123 residues could code for a propeptide that has a high number (36 residues) of charged residues, predominantly Lys and Glu (17 and 16 residues, respectively) (Fig. 2). The calculated molecular mass of the pyrolysin precursor is 155 kDa and after removal of the leader 138.5 kDa.

**Identification of Transcription Initiation Site and Regulatory Sequences—**Primer extension experiments were performed on total RNA isolated from pyrococcal cells grown on peptone to identify the transcription initiation site and resulted in a prominent extension product (Fig. 3). The size of this product corresponded with transcription initiation at the thymidine residue located 28 nucleotides upstream of the assumed translation start site (Fig. 2). A hexanucleotide with the sequence TTTATA is located 26 nucleotides upstream of this transcription initiation site and resembles the TATA box, which is involved in binding of the archaeal RNA polymerase (Fig. 2) (31, 32).

**Downstream of the *pls* gene a conserved structure is found, composed of stretches of at least four subsequent thymidine residues, that could be involved in termination of transcription (31), which is further followed by another ORF with an unknown function (Fig. 2). Upstream and in opposite orientation of the *pls* gene, the gene for a tRNA is located that could be the initiator tRNA, since it contains all the characteristics for an archaeal initiator tRNA (33).**

**Substrate Specificity—**Comparison of the deduced amino acid sequence of pyrolysin with those of other subtilisin-like serine proteases present in the data bases revealed the highest degree of homology (28–32% identity in the catalytic domain) with the tripeptidyl peptidases II (TPPs) from eukaryal origin, i.e. human, mouse, *Drosophila melanogaster*, and the nematode *Caenorhabditis elegans* (Fig. 4). The eucaryal TPPs are intracellular proteases with exopeptidase activity (34, 35), whereas all other known subtilisin-like serine proteases are endopeptidases (17).

The substrate specificity of pyrolysin was analyzed with chromogenic peptide substrates (Table I). Both tripeptidyl and tetrapeptidyl substrates were cleaved by pyrolysin. A preference for the conversion of substrates with a positively charged (Lys, Arg) residue at the P1 site was observed (nomenclature according to Schechter and Berger) (36). In spite of its homology with TPPs, pyrolysin showed no activity with the known substrate for TPPs, Ala-Ala-Phe-p-nitroanilide, and is not inhibited by the dehydroalanine-containing peptide Arg-Ala-β-Ala-Val-Ala that has been shown to be a competitive inhibitor of TPPs from human and rat (23, 24). No difference in substrate specificity was found between HMW and LMW pyrolysin (data not shown).

To demonstrate that pyrolysin has endopeptidase activity, it was incubated with αs1-, β- or κ-caseins and casein fragments. The degradation of caseins was followed in time and revealed that for all caseins the first cleavage products constituted a...
Characterization of Pyrolysin and Its Gene from *P. furiosus*

Purified pyrolysin was tested for the degradation of synthetic substrates. <Glu, pyroglutamyl; Suc, succinyl; MeO, methoxy; pNA, p nitroanilide.>

### Table I

**Substrate specificity of pyrolysin**

| Substrate            | Activity (%) |
|----------------------|--------------|
| P4 P3 P2 P1<sup>a</sup> | 62<sup>b</sup> |
| Succ-Ala-Ala-Pro-Lys-pNA | 100         |
| Succ-Ala-Ala-Pro-Arg-pNA | 32          |
| Succ-Ala-Ala-Pro-Met-pNA | 13          |
| Succ-Ala-Ala-Pro-Phe-pNA | 14          |
| Succ-Ala-Glu-Pro-Phe-pNA | 7           |
| Succ-Ala-Lys-Pro-Phe-pNA | 65          |
| H<sub>2</sub>-Val-Leu-Lys-pNA | 41          |
| MeO-Suc-Arg-Pro-Tyr-pNA | 52          |
| <Glu-Pro-Val-pNA | 0.5         |
| Ala-Ala-Phe-pNA | <0.01       |  

<sup>a</sup> Nomenclature according to Schechter and Berger (36). <sup>b</sup> 100% activity represents a specific activity of 300 nmol/min/mg protein.

A single unit, since the transcription initiation site and TATA box could be identified upstream of the gene, whereas downstream it was followed by conserved archaeal termination sequences. Furthermore, upstream of the *pls* gene a tRNA<sup>Met</sup> gene is found in the opposite orientation that most likely is the initiation tRNA.

Analysis of the deduced amino acid sequence indicates that pyrolysin is synthesized as a prepro-enzyme with a common signal sequence (30), suggesting that export of proteins in *P. furiosus* occurs in a similar way as in Eucarya and Bacteria. The propeptide of pyrolysin shows little homology with the consensus identified for different propeptides of subtilisin-like serine proteases; however, it shares with these propeptides a high number of charged residues (41).

The mature part of the deduced pyrolysin sequence shows the highest homology with eucaryal TPPs, which form a distinct subgroup of the subtilisin-like serine proteases. The mature pyrolysin, as several other subtilisin-like serine proteases, can be divided into two domains, the NH<sub>2</sub>-terminal catalytic domain that extends to about residue 500 followed by a COOH-terminal extension to residue 1249 (17). The catalytic domain is characterized by the three active site residues, Asp-30, His-216, and Ser-441, and the presence of the conserved Asn-342 present in the oxyanion hole that stabilizes the reaction intermediate. The catalytic domains of pyrolysin and the TPPs contain a large insert of more than 150 residues between the aspartate and histidine in comparison with the *Bacillus* subtilisin (Fig. 4). A number of conserved residues among all members could be identified within the insert, which may imply an evolutionary relationship between the archaeal pyrolysin and the eucaryal TPPs (Fig. 4). In addition, pyrolysin shows weak homology in the large COOH-terminal domain where the TPPs also have many conserved residues (data not shown).

**FIG. 5. Reversed-phase HPLC pattern of degradation products of α<sub>SI</sub>-casein-(1–23) by pyrolysin.** A, pattern of cleavage products resulting from the incubation of α<sub>SI</sub>-casein-(1–23) with pyrolysin for 0 min (pattern 1), 20 min (pattern 2), and 40 min (pattern 3) at 95 °C. The known pattern (22, 25) of α<sub>SI</sub>-casein degradation obtained with the *P. furiosus* type protease from *L. lactis* is shown for comparison (pattern 4). Identified products are indicated; asterisks indicate new cleavage products after prolonged incubation. B, amino acid sequence of α<sub>SI</sub>-casein-(1–23) with the preferential cleavage sites indicated by the arrows.

**DISCUSSION**

The hyperthermostable serine protease pyrolysin was purified from peptone-grown cells of *P. furiosus* and characterized, and its *pls* gene was cloned via reversed genetics. An essential step in the purification method of pyrolysin from the membrane fraction appeared to be its preincubation in 6 M urea at 95 °C. This resulted in two proteolytic active fractions, HMW and LMW pyrolysin, that were separated by preparative urea-PAGE and shown to have identical NH<sub>2</sub> termini. Incubation of the purified HMW pyrolysin at 95 °C resulted in the formation of LMW pyrolysin, and both forms appeared to be glycosylated to a similar extent. Therefore, the differences in molecular masses of the HMW and LMW pyrolysin are likely to be a consequence of autoproteolytic removal of a COOH-terminal part of the HMW pyrolysin that results in LMW pyrolysin. Similar processing has also been suggested for the serine protease from the related archaeon *Thermococcus stetteri* (11). Moreover, COOH-terminal processing is a common feature of serine proteases (37).

It has frequently been observed that proteins from hyperthermophiles are not denatured during standard SDS-PAGE (3, 38, 39). This was also found for pyrolysin that was subject to autoproteolysis under these conditions. Acid denaturation of HMW pyrolysin demonstrated a molecular mass of 150 kDa. Autoproteolytic processing of pyrolysin may also explain the previously reported heterogeneity of pyrolysin, evidenced by the multiple activity bands on substrate-PAGE (12–14). Alternatively, pyrolysin may form a complex either with itself, as shown for TPPs (40), or with a cell envelope structure as found in *Staphylothermus marinus* where a serine protease is associated with a filamentous surface protein assembly called tetrabrackhin (39).

The *pls* gene-encoding pyrolysin is most likely transcribed as a single unit, since the transcription initiation site and TATA box could be identified upstream of the gene, whereas downstream it was followed by conserved archaeal termination sequences. Furthermore, upstream of the *pls* gene a tRNA<sup>Met</sup> gene is found in the opposite orientation that most likely is the initiation tRNA.

Analysis of the deduced amino acid sequence indicates that pyrolysin is synthesized as a prepro-enzyme with a common signal sequence (30), suggesting that export of proteins in *P. furiosus* occurs in a similar way as in Eucarya and Bacteria. The propeptide of pyrolysin shows little homology with the consensus identified for different propeptides of subtilisin-like serine proteases; however, it shares with these propeptides a high number of charged residues (41).

The mature part of the deduced pyrolysin sequence shows the highest homology with eucaryal TPPs, which form a distinct subgroup of the subtilisin-like serine proteases. The mature pyrolysin, as several other subtilisin-like serine proteases, can be divided into two domains, the NH<sub>2</sub>-terminal catalytic domain that extends to about residue 500 followed by a COOH-terminal extension to residue 1249 (17). The catalytic domain is characterized by the three active site residues, Asp-30, His-216, and Ser-441, and the presence of the conserved Asn-342 present in the oxyanion hole that stabilizes the reaction intermediate. The catalytic domains of pyrolysin and the TPPs contain a large insert of more than 150 residues between the aspartate and histidine in comparison with the *Bacillus* subtilisin (Fig. 4). A number of conserved residues among all members could be identified within the insert, which may imply an evolutionary relationship between the archaeal pyrolysin and the eucaryal TPPs (Fig. 4). In addition, pyrolysin shows weak homology in the large COOH-terminal domain where the TPPs also have many conserved residues (data not shown).
The COOH-terminal domain, if present, in subtilisin-like serine proteases can have different functions but is often related to localization or anchoring of the protease (17). The last 80 amino acids of pyrolysin do not contain known hydrophobic anchoring sequences, but they are characterized by a high number of Glu and Leu residues (14 and 13 residues, respectively) that may be involved in the observed association of pyrolysin to the cell envelope.

The COOH-terminal domain of pyrolysin together with the large insert present in the catalytic domain contain almost all of the possible large insert present in the catalytic domain (17). The last serine proteases can have different functions but is often related to localization or anchoring of the protease (17). The last 80 amino acids of pyrolysin do not contain known hydrophobic anchoring sequences, but they are characterized by a high number of Glu and Leu residues (14 and 13 residues, respectively) that may be involved in the observed association of pyrolysin to the cell envelope.

The substrate specificity of pyrolysin was analyzed using chromogenic substrates, different caseins, and the small peptide S1/H9251-casein-(1–23). Degradation of the caseins and the acid-denatured HMW pyrolysin (150 kDa) could be explained by glycosylation.

The substrate specificity of pyrolysin was analyzed using chromogenic substrates, different caseins, and the small peptide S1/H9251-casein-(1–23). Degradation of the caseins and the acid-denatured HMW pyrolysin (150 kDa) could be explained by glycosylation.

Acknowledgments—We are grateful to Dr. Brigitte Tomkinson for a gift of chemicals and Saskia van Schalkwijk and Marcel Djikgraaf for technical assistance.

REFERENCES

1. Adams, M. W. W. (1990) Annu. Rev. Microbiol. 47, 627–658
2. Stettler, K. O., Fiala, G., Huber, G., Huber, R., and Segerer, A. (1990) FEMS Microbiol. Rev. 75, 117–124
3. Mukund, S., and Adams, M. W. W. (1995) J. Biol. Chem. 270, 8389–8392
4. Kengen, S. W. M., de Bok, F. A. M., van Loo, N.-D., Dijkema, C., Stams, A. J. M., and de Vos, W. M. (1994) J. Biol. Chem. 269, 17537–17541
5. Kelly, R. M., and Adams, M. W. W. (1994) Antonie Leeuwenhoek 66, 247–270
6. Woese, C. R., Kandler, O., and Wheelis, M. L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 4576–4579
7. Fusek, M., Lin, X., and Tang, J. (1990) J. Biol. Chem. 265, 1496–1501
8. Lin, X., and Tang, J. (1990) J. Biol. Chem. 265, 1480–1485
9. Morikawa, M., Iwasa, Y., Hashid, N., Hoaki, T., and Imanaka, T. (1994) Appl. Environ. Microbiol. 60, 4559–4566
10. Cowan, D. A., Smolenski, K. A., Daniel, R. M., and Morgan, H. W. (1987) FEMS Microbiol. Lett. 43, 171–176
11. Klingeberg, M., Galunsky, B., Sjoholm, C., Kasche, V., and Antranikian, G. (1995) Appl. Environ. Microbiol. 61, 3098–3104
12. Eggen, R., Geerling, A., Watts, J., and de Vos, W. M. (1990) FEMS Microbiol. Lett. 71, 17–20
13. Blumentals, I. I., Robinson, A. S., and Kelly, R. M. (1990) Appl. Environ. Microbiol. 56, 1992–1998
14. Connatts, H., Cowan, D. A., and Sharp, R. J. (1991) Gen. Microbiol. 137, 1193–1199
15. Snowden, L. J., Blumentals, I. I., and Kelly, R. M. (1992) Appl. Environ. Microbiol. 58, 1341–1344
16. Volpi, K., Markiewicz, P., Stetter, K. O., and Miller, J. H. (1995) Protein Sci. 4, 1329–1340
17. Siezen, R. J., de Vos, W. M., Leunissen, J. A. M., and Dijkstra, B. W. (1991) Protein Eng. 4, 719–737
18. Gibson, T. J. (1984) Studies on the Epstein-Barr-Virus Genome. Ph.D. Thesis. Cambridge, UK
19. Eggen, R. I. L., Kulich, A. C. M., Jetten, M. S. M., and de Vos, W. M. (1991) J. Biol. Chem. 266, 6883–6887
20. Zacharius, R. M., Zell, T. E., Morrison, J. H., and Woodlock, J. J. (1969) Anal. Biochem. 30, 148–152
21. Laemmli, U. K. (1970) Nature 227, 680–685
22. Siezen, R. J., Bruinenberg, P. G., Vos, P., van Alen-Boerriger, I. T., Nijhuis, M., Alting, A. C., Exterkate, F. A., and de Vos, W. M. (1993) Protein Sci. 2, 927–937
23. Tomkinson, B. (1994) Biochim. J. 304, 517–523
24. Tomkinson, B., Grehn, L., Fransson, B., and Zetterqvist, O. (1994) Arch. Biochem. Biophys. 314, 376–379
25. Exterkate, F. A., Alting, A. C., and Slanger, C. J. (1991) Biochem. J. 273, 135–139
26. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
27. Voorhorst, W. G. B., Eggen, R. I. L., Luusink, E. J., and de Vos, W. M. (1995) J. Bacteriol. 177, 7105–7111
28. Devereux, J., Haebeli, P., and Smithies, O. (1990) Nucleic Acids Res. 12, 387–395
29. Marshall, R. D. (1972) Annu. Rev. Biochem. 41, 673–702
30. Von Heijne, G. (1985) J. Mol. Biol. 184, 99–105
31. Dalgaard, J. Z., and Garrett, R. A. (1993) The Biochemistry of Archaea (Archaeobacteria) (Kates, M., Kushner, D. J., and Matheson, A. T., eds) pp. 535–563, Elsevier, Amsterdam
32. Rowlands, T., Baumann, P., and Jackson, S. P. (1994) Biochem. J. 309, 1326–1329
33. Kuchino, Y., Ihara, M., Yabasaiki, Y., and Nishimura, S. (1982) Nature 298, 684–685
34. Bälöw, R.-M., Tomkinson, B., Ragnarsson, U., and Zetterqvist, O. (1986) J. Biol. Chem. 261, 2409–2417
35. Tomkinson, B., Wernerstedt, C., Hellman, U., and Zetterqvist, O. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 7508–7512
36. Schechter, I., and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157–162
37. Shinde, U., Li, Y., Chatterjee, S., and Inouye, M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6924–6928
38. Siemensen, S. W. M., and Stams, A. J. M. (1994) Biochimica et Biophysica Acta 111, 79–88
39. Peters, J., Nitsch, M., Kuhlmann, B., Golbik, R., Lapas, A., Kellermann, J., Engelhart, H., Pfitzer, J. P., Muller, S., Goldie, K., and Engel, A. (1995) J. Mol. Biol. 245, 385–401
40. Macpherson, E., Tomkinson, B., Bälöw, R., Haglund, S., and Zetterqvist, O. (1987) Biochem. J. 248, 259–263
41. Siezen, R. J., Leunissen, J. A. M., and Shinde, U. (1995) Intramolecular Conjugation of Proteins (Inouye, M., and Shinde, U., eds) pp. 233–256, R. G. Landes Co, Austin, TX
42. Germain, D., Dumas, F., Vernet, T., Bourbannais, Y., Thomas, D. Y., and Souleau, G. (1992) FEBS Lett. 299, 283–286
43. Leduc, R., Molloy, S. S., Thorne, B. A., and Thomas, G. (1992) J. Biol. Chem. 267, 14304–14308
44. Power, S. D., Adams, R. M., and Wells, J. A. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 3006–3010