

A Proteasome from the Methanogenic Archaeon *Methanosarcina thermophila**

(Received for publication, June 30, 1995, and in revised form, September 27, 1995)

Julie A. Maupin-Furlow and James G. Ferry‡

From the Department of Biochemistry and Molecular Biology, Pennsylvania State University,
University Park, Pennsylvania 16802-4500

A 645-kDa proteasome was purified from *Methanosarcina thermophila* which had chymotrypsin-like and peptidylglutamyl-peptide hydrolase activities and contained α (24-kDa) and β (22-kDa) subunits. Processing of both subunits was suggested by comparison of N-terminal sequences with the sequences deduced from the α - and β -encoding genes (*psmA* and *psmB*). Alignment of deduced sequences for the α and β subunits revealed high similarity; however, the N-terminal sequence of the α subunit contained an additional 24 amino acids that were not present in the β subunit. The α and β subunits had high sequence identity with α - and β -type subunits of proteasomes from eucaryotic organisms and the distantly related archaeon *Thermoplasma acidophilum*. The *psmB* gene was transcribed *in vivo* as a monocistronic message from a consensus archaeal promoter. The results suggest that proteasomes are more widespread in the *Archaea* than previously proposed. Southern blotting experiments suggested the presence of ubiquitin-like sequences in *M. thermophila*.

Proteasomes are prevalent in eucaryotes (*Eucarya* domain) where they have at least three distinct endopeptidase activities which include hydrolysis of peptide bonds on the carboxyl side of hydrophobic, basic, and acidic amino acid residues (chymotrypsin-like, trypsin-like, and peptidylglutamyl-peptide hydrolyzing activities, respectively) (1). It is proposed that the 20 S eucaryotic proteasome is the "catalytic core" in a larger 26 S complex that degrades proteins labeled with ubiquitin in an ATP-dependent process (1).

A 20 S proteasome from the archaeon *Thermoplasma acidophilum* has been extensively characterized. The thermoplasma proteasome has α and β (25.8- and 22.3-kDa) subunits with significant identity to the sequences of all described eucaryotic 20 S proteasomes (2, 3). The quaternary structure is also highly conserved with the eucaryotic proteasome (4). The fully assembled thermoplasma proteasome is a barrel-like structure of four stacked rings (5). Each of the two inner rings are comprised of 7 β subunits, and each outer ring contains 7 α subunits (5). Assembly of the proteasome proceeds by formation of the α ring which is proposed to "chaperon" assembly of the β subunit ring (6). The integrity of the N-terminal region of

the α subunit is necessary for proper assembly of rings. The β subunits are synthesized with a propeptide that is processed during assembly. Recently, the crystal structure of the thermoplasma proteasome was published (5), and it is proposed that this proteasome is a novel threonine protease (7).

Attempts to demonstrate eucaryotic-like proteasomes from procaryotes (*Bacteria* and *Archaea* domains) other than thermoplasma have been unsuccessful (8–10). The results have led to the suggestion that thermoplasma, an atypical member of the *Archaea*, is the only procaryote containing proteasomes (10). Recently, an enzyme with a quaternary structure similar to the 20 S proteasome was isolated from the eubacterium *Rhodococcus*; however, the primary structure of the eubacterial enzyme has low identity to the archaeal and eucaryotic proteasomes (11). Thus, questions remain regarding the distribution of proteasomes in the *Bacteria* and *Archaea* domains (procaryotes), as well as the origin and evolution of proteasomes from the *Eucarya* and *Archaea* domains (10).

Here we provide evidence for an enzyme highly identical to the eucaryotic-like 20 S proteasome in a representative of methanogenic microbes, the largest group known for the *Archaea*. These results suggest that proteasomes are more widespread among the *Archaea* than previously proposed. Although proteolysis is almost certain to be of fundamental importance for methanogenic microbes, the process has not been investigated and the isolation of proteolytic enzymes has not been reported. The discovery of proteasomes in methanogenic microbes marks an entry for investigating the physiology, biochemistry, and molecular biology of protein turnover in this major group of strict anaerobes.

MATERIALS AND METHODS

Proteasome Purification and Characterization—*Methanosarcina thermophila* was grown on sodium acetate (12). Cells were harvested (10 g wet weight) and resuspended in a 3-fold volume (w/v) of TD buffer (20 mM Tris-HCl, 1 mM dithiothreitol, pH 7.2) and passed through a French pressure cell at 20,000 psi followed by centrifugation for 30 min at 16,000 $\times g$. The cell extract (1250 mg of protein in 30 ml) was applied to a Q-Sepharose (Pharmacia) column (2.5 \times 28.5 cm) that was developed with a linear NaCl gradient (0–400 mM in 600 ml of TD buffer). Peak Q-Sepharose fractions that contained chymotrypsin-like activity were pooled (310 mg of protein), diluted 2-fold to 140 ml, and applied to a Q-Sepharose column (2.5 \times 28.5 cm) previously equilibrated with 250 mM NaCl. The column was developed with a linear NaCl gradient (250–310 mM in 180 ml of TD buffer). The active fractions were pooled (16.5 mg of protein), diluted 2-fold with TD buffer to 112 ml, and then concentrated approximately 20-fold by batch elution with 1 M NaCl from an Econo-Pac High Q column (Bio-Rad). The concentrated protein solution (16 mg of protein) was applied in 1-ml samples to a Sepharose 6B (Pharmacia) gel filtration column (1.2 \times 55 cm) equilibrated with 150 mM NaCl in TD buffer. The column, calibrated with high-molecular weight protein standards (Sigma), was also used to estimate the molecular mass of the proteasome complex. In the final step of purification, the active fractions (0.8 mg of protein) were loaded onto a High Q column (Bio-Rad) which was developed with a linear NaCl gradient (150–500 mM in 30 ml). Approximately 0.25 mg of purified proteasome

* This work was supported by National Institutes of Health individual national research service award 1F32GM15877-02 (to J. A. M-F.) and National Institutes of Health Grant GM44661-02 (to J. G. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802-4500. Tel.: 814-863-5721; Fax: 814-863-7024; E-mail: jgf3@psuvm.psu.edu.

was obtained. An accurate measurement of peptide hydrolyzing activity at each step of the purification was not possible due to nonspecificity of the assay and interference by the inherent fluorescence of cellular cofactors present in cell extract. However, based on the amount recovered from cell extract, it is estimated that the proteasome constitutes a minimum of 0.02% of the total soluble cell protein.

Chymotrypsin- or trypsin-like peptide hydrolyzing activities were assayed by fluorimetric measurement (13) of the release of 7-amino-4-methylcoumarin from Suc¹-Ala-Ala-Phe-AMC or Pro-Phe-Arg-AMC at 42 °C for 60 min. The assay mixture (0.3 ml) contained protein and 20 μM fluorogenic substrate (Sigma) in 20 mM Tris-HCl (pH 7.2) containing 0.4% dimethyl sulfoxide. Peptidylglutamyl-peptide hydrolase activity of the purified proteasome was assayed by colorimetric measurement of the release of β-naphthylamine (14) from Cbz-Leu-Leu-Glu-β-NA (Sigma) at 42 °C for 60 min. The assay mixture (0.2 ml) contained 5 μg of protein and 300 μM substrate in 20 mM Tris-HCl (pH 7.2) containing 3% dimethyl sulfoxide. Protein concentration was determined by the bicinchoninic acid method (15) using bovine serum albumin as the standard.

N-terminal sequencing was as follows. Subunits of the purified proteasome were separated by SDS-PAGE (16) using 10% polyacrylamide. The subunits were electroblotted onto a polyvinylidene difluoride (Immobilon-P) membrane (Millipore) and sequenced by automated Edman degradation at the University of Florida-Interdisciplinary Center for Biotechnology Research protein chemistry core facility.

Cloning and Sequencing—An *M. thermophila* Sau3AI genomic library, prepared in λ phage vector GEM-11 (Promega), was screened with a probe specific for the gene encoding the δ subunit of the CO dehydrogenase enzyme complex (*cdhD*). A 16-kb Sau3AI genomic fragment, containing the complete *cdh* operon and the *psmB* gene, was identified by DNA sequence analysis (Fig. 1). A manuscript describing the cloning, gene organization, and DNA sequence of the *cdh* operon is in preparation.² The *psmA* gene was cloned as follows. The above described genomic library was screened with a degenerate oligonucleotide (5'-ATG-GC(AGCT)-TA(CT)-GA(CT)-(AC)G(AGCT)-GC(AGCT)-AT(TCA)-AC(AGCT)-GT-3') based on the N-terminal sequence of the proteasome α subunit. The oligonucleotide was 3'-end labeled with digoxigenin-11-ddUTP (2',3'-dideoxyuridine 5'-triphosphate) (Boehringer Manneheim Biochemicals). Colony/PlaqueScreen nylon membranes (DuPont NEN) were used for plaque lifts according to the manufacturer. Hybridization for plaque screening and Southern blotting was for 12 h at 58 °C in 5 × SSC (where 1 × SSC is 0.15 M NaCl and 15 mM sodium citrate, pH 7.5), 0.1% N-lauroylsarcosine, 0.02% SDS, 1% blocking reagent (Boehringer Manneheim). Membranes were washed twice in 2 × SSC, 0.1% SDS at 21 °C for 15 min and then twice in 0.5 × SSC, 0.1% SDS at 58 °C for 15 min. Digoxigenin-11-ddUTP label was detected colorimetrically according to manufacturer (Boehringer Manneheim). *Sa*I, *Hin*II, *Pvu*II and *Xmn*I restriction enzymes (New England Biolabs) were used to subclone the *psmA* gene from a 17-kb Sau3AI genomic fragment for DNA sequence analysis.

The DNA sequence of both strands of the *psmB* and *psmA* genes was determined according to the dideoxy chain termination method (17). The nucleotide sequence of the *M. thermophila* *psmA* and *psmB* genes have been submitted to the GenBank data base and assigned the accession numbers U30483 and U22157.

Protein Sequence Analyses—GenBank, EMBL and SwissProt data bases were searched at the National Center for Biotechnology Information (Bethesda, MD) using the BLAST network server (18) at the University of Florida-Interdisciplinary Center for Biotechnology Research computing facility in conjunction with the Genetics Computer Group program (Madison, WI) (19). CLUSTAL (20) was used for protein sequence alignment and dendrogram generation. The proteasome sequences used to generate the dendrogram were as follows: AtPSM30, At27 or α, DdC4, DdC5, Dm13-73Ai, Dm25, DmPROS28.1, DmPROS29, DmPROS35, GgC1, HsPROS30, HsPROS27, HsRING10, HsDELTA, HsIOTA, HsZETA, HsC2, HsC3, HsC5, HsC8, HsC9, HsLMP7e2, HsMECL1, MmC13, MmLMP2, RnC1, RnC2, RnC3, RnC8, RnC9, RnN3, RrC5, RrC7-I, RrC10-II, RrDELTA, RrIOTA, RrZETA, RrRING12, ScPRE1, ScPRE2, ScPRE4, ScPRG1, ScPRS3, ScPUP1,

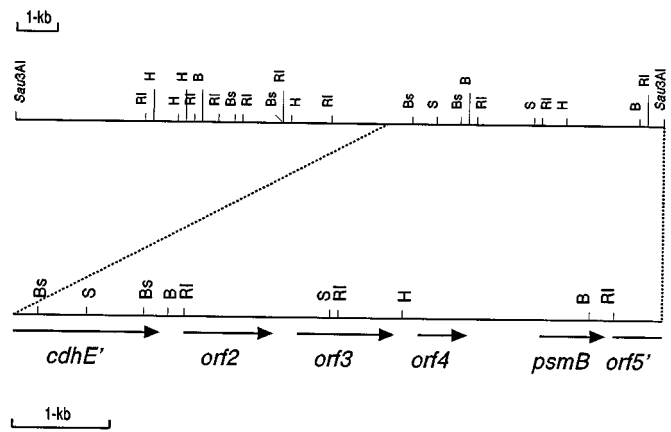


FIG. 1. Location of the *psmB* gene relative to the *cdh* operon in the *M. thermophila* chromosome. The *cdh* operon of *M. thermophila* encodes subunits of the CO dehydrogenase enzyme complex. The *cdhE* gene encodes the γ-subunit of CO dehydrogenase and is the most distal gene in the *cdh* operon. *cdhE'* and *orf5'* are truncated. Abbreviations: Bs, *BstE*II; S, *Sma*I; B, *Bam*HI; RI, *Eco*RI; H, *Hin*III.

ScPUP2, ScP22-6, ScY7, ScY13, ScYC1, ScY7ALPHA, SpPTS1, TaAlpha, TaBeta, XIBETA, XIC3 (21); AtC5 (22); AtIOTA (23); CeC5 (24); HsPROX and HsPROY (25); HsC7-I (26); ScPRE3 (27); ScPRE5A and ScPRE6A (28); ScPRS1 (10); MtAlpha and MtBeta (this report).

RNA Isolation and Northern Hybridization—*M. thermophila* was grown in marine medium (29) supplemented with 100 mM sodium acetate, methanol, or trimethylamine. Total RNA was isolated (30), and polysaccharide removed from the RNA (31) prior to formaldehyde-agarose gel electrophoresis and transfer to nylon membranes (32). A 259-bp *Hinc*II-*Bam*HI fragment from *psmB* was α-³²P]dATP random primed labeled using a kit (Boehringer Manneheim). Hybridization was at 42 °C in a solution containing: 5 × SSPE (3.6 M NaCl, 0.2 M NaPO₄, 0.02 M EDTA, pH 7.4), 50% formamide, 0.5% SDS, 5 × Denhardt's solution (0.1% Ficoll, 0.1% polyvinylpyrrolidone, 1% bovine serum albumin). The membrane was washed at 60 °C in 0.1 × SSC containing 0.01% SDS.

Primer Extension Assays—The 5' end of a *psmB*-specific mRNA was mapped by avian myeloblastosis virus reverse transcriptase (Promega) primer extension using the oligonucleotide 5'-GTACAACTACTCTCTACG-3' complementary to nucleosides +36 to +53 of the *psmB* gene. The 5' end of a *psmA*-specific mRNA was mapped, using similar procedures, with the oligonucleotide 5'-TCCGTCAGGGCTGAAAC-3' complementary to nucleosides +37 to +54 of the *psmA* gene.

Southern Blotting—Genomic DNA was digested to completion with the restriction endonucleases indicated. The restriction fragments were separated by electrophoresis and transferred onto nylon membranes (Boehringer Manneheim) (33). All other procedures were as described (33) except as follows. The degenerate oligonucleotide used as a *psmA*-specific probe, as well as the hybridization conditions, were the same as described for cloning the *psmA* gene. A 259-bp *Hinc*II-*Bam*HI fragment from *psmB* was random primed labeled with digoxigenin-dUTP and used as a *psmB*-specific probe for hybridization at 51 °C according to the supplier's recommendations (Boehringer Manneheim). The wash stringency was 51 °C using 0.5 × SSC containing 0.1% SDS. A 0.22-kb *Bgl*II-*Kpn*I fragment from plasmid YEP96 (34), specific for a synthetic yeast ubiquitin gene, was random prime labeled with digoxigenin-dUTP and used as a ubiquitin gene probe for hybridization at 45 °C (Boehringer Manneheim). The wash stringency was 45 °C using 0.5 × SSC containing 0.1% SDS.

RESULTS

Identification of *psmA* and *psmB* Encoding the α and β Subunits of a Proteasome Purified from *M. thermophila*—DNA sequencing downstream of the *M. thermophila* *cdh* operon revealed an open reading frame (*psmB*) (Fig. 1) with a deduced amino acid sequence having high identity (up to 51%) with β-type subunits of proteasomes from phylogenetically divergent organisms (Fig. 2). The *psmB* gene encoded a putative 210-amino acid protein (PsmB) with a calculated anhydrous molecular mass of 22,981 Da. A potential ribosome binding site was identified upstream of *psmB* (Fig. 6) complementary to the

¹ The abbreviations used are: Suc, succinyl; AMC, 7-amido-4-methylcoumarin; Cbz, carbobenzyloxy; β-NA, β-naphthylamine; PAGE, polyacrylamide gel electrophoresis; PsmB, *M. thermophila* β proteasome subunit; PsmA, *M. thermophila* α proteasome subunit; NLS, nuclear localization sequence; cNLS complementary nuclear localization sequence; PAGE, polyacrylamide gel electrophoresis; kb, kilobase(s); bp, base pair(s).

² J. A. Maupin-Furlow and J. G. Ferry, manuscript in preparation.

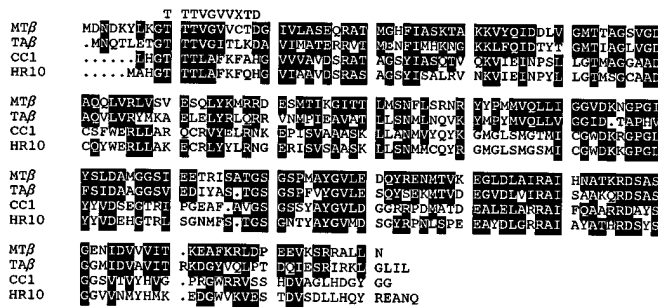


FIG. 2. Multiple amino acid sequence alignment of *M. thermophila* PsmB with selected β -type proteasomal subunits. Periods represent gaps introduced for alignment. Highlighted amino acid residues are identical or functionally similar to PsmB. The N-terminal protein sequence of PsmB is represented above the deduced amino acid sequence, where X is an uncertain assignment. Abbreviations: MT β , *M. thermophila* PsmB; TAB β , β -subunit of the *T. acidophilum* proteasome (51% identity) (2); CC1, *Gallus domesticus* proteasome C1 chain (33% identity) (35); HR10, *Homo sapiens* proteasome-related RING10 (33% identity) (36).

sequence at the 3' terminus of the 16 S rRNA in methanosarcina (3'-UCCUCCACUA) (37).

A proteasome was purified from *M. thermophila* which contained 24- (α) and 22-kDa (β) subunits (see below). Analysis of the α -subunit revealed two N-terminal sequences that were identical except for the length (Fig. 3). A λ GEM-11 clone bank containing *M. thermophila* genomic DNA was screened with a degenerate oligonucleotide probe based on the N-terminal sequence. Southern blot analysis of the DNA from a hybridizing phage isolate identified a 980-bp *EcoRV*-*XmnI* fragment which hybridized to the probe. Sequence analysis of the fragment identified an open reading frame (*psmA*) with a putative translational start corresponding to the longer of the two α subunit N-terminal sequences (Fig. 7). The *psmA* gene encoded a putative 246-amino acid protein (PsmA) with a calculated anhydrous molecular mass of 27,139 Da. PsmA was highly identical (up to 53%) to α -type proteasome subunits from phylogenetically diverse organisms (Fig. 3). The DNA sequence downstream of the only consensus ribosome binding site contained three potential translational start sites, two of which corresponded to the N termini determined for the purified α subunit and a third located two codons upstream of *psmA* (Fig. 7). Two of the potential start sites are separated from the consensus ribosome binding site by 7 and 13 bases which is typical for archaeal genes (40). Although a translational start site corresponding to the shorter of the two N termini (25 bases from the consensus ribosome binding site) cannot be ruled out, it is more probable that PsmA is processed to yield the shorter N-terminal sequence.

Alignment of deduced sequences for the α and β subunits revealed high similarity (46%); however, the N-terminal sequence of the α subunit contained an additional 24 amino acids that were not present in the β subunit (Fig. 4). When the CLUSTAL program was used to generate a dendrogram relating all proteasome subunit sequences currently available, PsmA and PsmB clustered with the α - and β -type subunit group, respectively, in the recently proposed classification scheme (10) and were closest relatives to the thermoplasma proteasome subunits (Fig. 5). No established proteinase motif was identified in either PsmA or PsmB using PROFILESCAN (41), a result which is in agreement with all proteasome subunit sequences to date.

In Vivo Transcription of psmA and psmB—Northern hybridization revealed a *psmB*-specific transcript of approximately 800 bp in *M. thermophila* cells grown on acetate, trimethylamine, or methanol (data not shown). The results suggest that

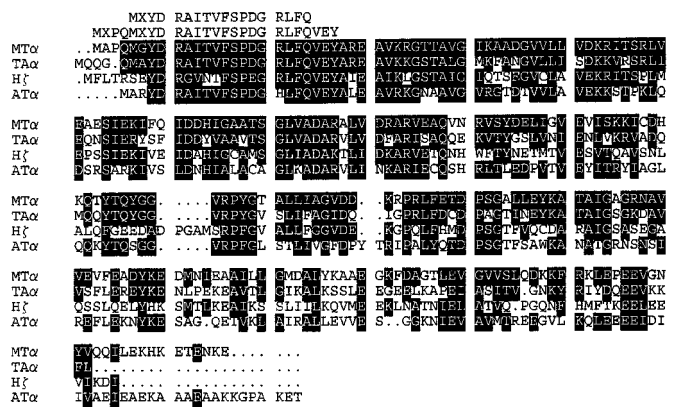


FIG. 3. Multiple sequence alignment of *M. thermophila* PsmA with selected α -type proteasomal subunits. Periods represent gaps introduced for alignment. Highlighted amino acid residues are identical or functionally similar to PsmA. The N-terminal protein sequences of PsmA are represented above the deduced amino acid sequence, where X is an uncertain assignment. Abbreviations: MT α , *M. thermophila* PsmA; T $\alpha\alpha$, α -subunit of the *T. acidophilum* proteasome (53% identity) (3); H ζ , *H. sapiens* proteasome ζ chain (41% identity) (38); AT α or AT-27, *Arabidopsis thaliana* α -type proteasome (44% identity) (39).

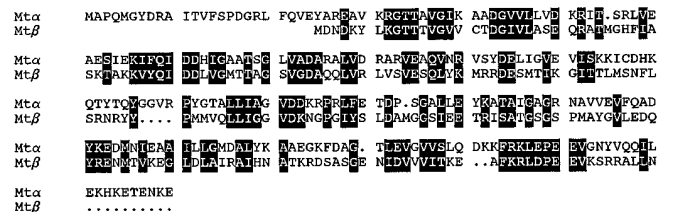


FIG. 4. Alignment of the deduced sequences of the *M. thermophila* α (PsmA) and β (PsmB) subunits. Periods represent gaps introduced for maximum alignment. Functionally similar amino acids are boxed in black.

proteasomes are produced during growth on all three substrates. Furthermore, *psmB* is transcribed *in vivo* as a monocistronic message suggesting that *psmA* and *psmB* are not likely to be part of the same operon. Primer extension analysis of total RNA isolated from methanol- or trimethylamine-grown cells identified one major product using an oligonucleotide primer specific for *psmB* (Fig. 6). The 5' end of the mRNA mapped to the G base located -101 bases relative to the predicted translational start site. This putative transcriptional start site was within a 4-bp sequence with 75% identity to the archaeal boxB (ATGC) consensus and began 20-bp downstream of a 8-bp sequence with 87% identity to the boxA (TTTA(T/A)ATA) consensus archaeal promoter (40). In general, transcription of genes from the methanogenic *Archaea* occurs at a purine-pyrimidine dinucleotide in boxB which is separated by 16–23-bp downstream from boxA.

Primer extension analysis of total RNA isolated from trimethylamine-grown cells identified two products using a *psmA*-specific oligonucleotide. The 5' ends of the mRNA mapped to the G bases located -117 and -48 relative to the predicted translational start site (Fig. 7). No sequences resembling the consensus boxA were at the expected 16–23-bp upstream of either of the two G bases. Potential stem-loop structures were detected upstream of both 5' mRNA ends. Although transcriptional start sites cannot be ruled out, the results are more consistent with mRNA processing at the G bases.

Proteasome Purification and Characterization—A 645-kDa enzyme was purified from *M. thermophila* which contained α and β subunits of approximately 24- and 22-kDa (Fig. 8). The N-terminal sequence of the β subunit was identical to residues 10–19 deduced from *psmB* (Fig. 2) which indicates it is proc-

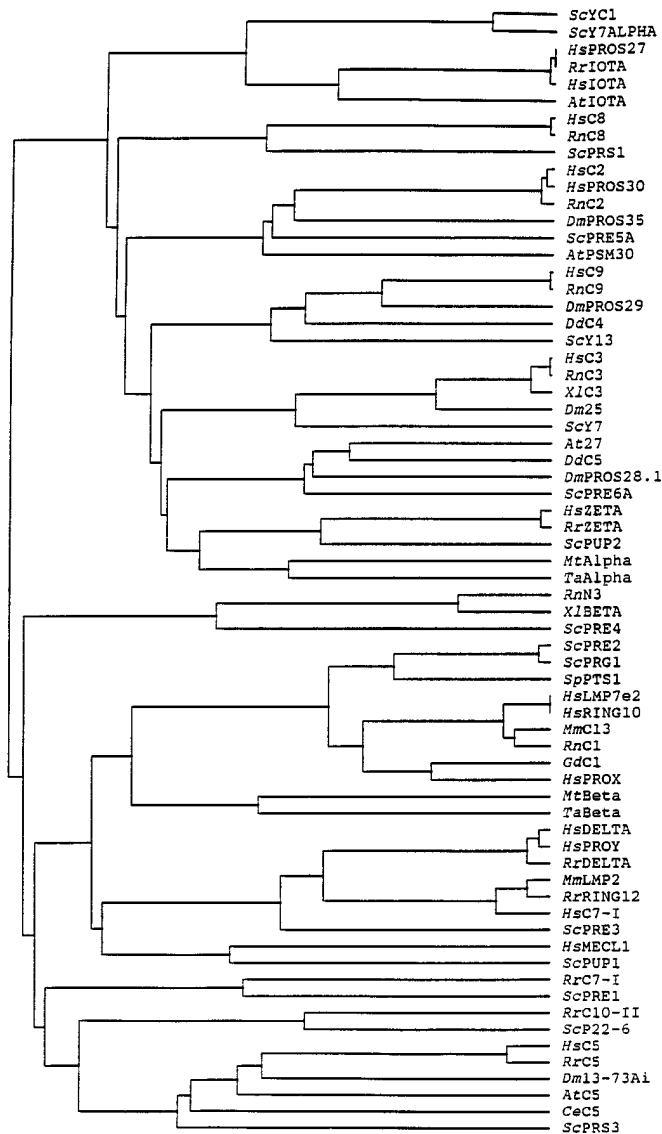


FIG. 5. Dendrogram relating currently available proteasome subunit sequences to the α (PsmA) and β (PsmB) subunits from *M. thermophila*. At, *Arabidopsis thaliana*; Ce, *Caenorhabditis elegans*; Dd, *Dictyostelium discoideum*; Dm, *Drosophila melanogaster*; Gd, *Galus domesticus*; Hs, *H. sapiens*; Mm, *Mus musculus*; Mt, *M. thermophila*; Rn, *Rattus norvegicus*; Rr, *Rattus rattus*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Ta, *Thermoplasma acidophilum*; XI, *Xenopus laevis*.

essed to expose an N terminus with an active-site threonine similar to the β subunits of proteasomes from the *Eucarya* and thermoplasma (7). Analysis of the α subunit revealed two N-terminal sequences corresponding to amino acids 1–25 and 5–22 deduced from *psmA* (Fig. 3). The purified *M. thermophila* enzyme displayed chymotrypsin-like activity catalyzing hydrolysis of Suc-Ala-Ala-Phe-AMC (1.2 nmol of 7-amino-4-methylcoumarin $\text{min}^{-1} \text{mg}^{-1}$ protein) containing an aromatic residue adjacent to the leaving group (AMC). The enzyme also had peptidylglutamyl-peptide hydrolase activity catalyzing the hydrolysis of Cbz-Leu-Leu-Glu- β -NA (8.9 nmol of β -naphthylamine $\text{min}^{-1} \text{mg}^{-1}$ protein) containing an acidic residue adjacent to the leaving group (β -NA). Little or no activity was detected with Pro-Phe-Arg-AMC which contains an arginine residue adjacent to the leaving group. The absence of trypsin-like activity is a property similar to the 673-kDa thermoplasma proteasome; however, the thermoplasma proteasome does not have significant peptidylglutamyl-peptide hydrolase activity

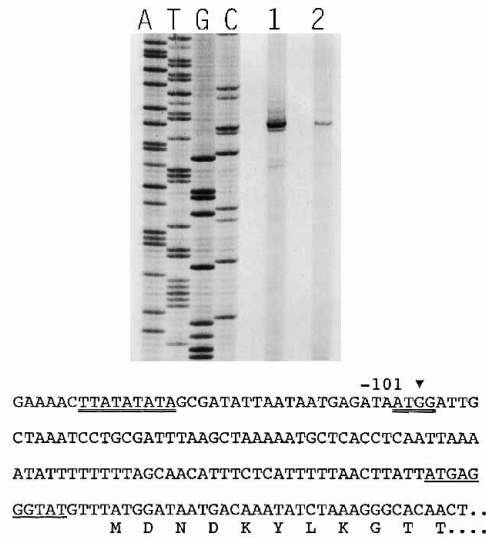


FIG. 6. Primer extension analysis of the 5'-end of *psmB*-specific mRNA. Upper, the products of extension reactions using total RNA from methanol (lane 1) and trimethylamine grown (lane 2) *M. thermophila* cells and DNA sequencing reactions (lanes A, T, G, and C) in which the same oligonucleotide primer was used. Lower, the double underlined DNA sequences indicate the boxA and boxB consensus archaeal promoter. The putative transcriptional start site is indicated by a down arrow and is numbered relative to the translational start site. Single underlined sequences indicate a potential ribosome binding site upstream of the initiator codon. The N-terminal deduced amino acid sequence of *psmB* is shown in single letter code beginning with the first base of each codon.

(4). The results demonstrate that the enzyme purified from *M. thermophila* is a proteasome with the α and β subunits encoded by *psmA* and *psmB*.

Evidence for M. thermophila Sequences Similar to the Yeast Ubiquitin Gene—When Southern blots of *M. thermophila* genomic DNA were probed with a synthetic yeast ubiquitin gene, 3.5-kb *Pst*I and 2.1-kb *Eco*RI fragments hybridized specifically with the probe (Fig. 9). No hybridization of the ubiquitin gene probe was observed with *Escherichia coli* genomic DNA (data not shown). These results suggest that the *M. thermophila* genome contains ubiquitin-like sequences.

DISCUSSION

The results document a proteasome in *M. thermophila* representing anaerobic methanogenic microbes, the largest known group in the *Archaea*. Therefore, proteasomes outside the *Eucarya* are not restricted to the atypical archaeon thermoplasma as previously suggested (10). *T. acidophilum* is a thermoacidophile that can grow aerobically utilizing glucose as an energy source and is phylogenetically distant from the strictly anaerobic methanosarcina (42) which obtain energy for growth by converting simple one- and two-carbon substrates to methane at neutral pH. Thus, proteasomes occur in physiologically and phylogenetically diverse *Archaea*. The results lessen the probability that the thermoplasma proteasome originated by horizontal gene transfer from the *Eucarya* (10) and supports the proposal that eucaryotic proteasomes evolved from an archaeal predecessor (2). Multiple DNA sequence alignment of the *T. acidophilum* and *M. thermophila* α and β proteasomal subunit genes revealed high identity (46–60%) among all four genes with the highest identity between α subunit genes (not shown). This result suggests that the gene encoding the α subunit is more closely related to an ancestral proteasome gene from which both the α and β subunit genes derived.

Thermoplasma and methanosarcina are classified to the Euryarchaeota kingdom (42). Presently, there is no evidence for

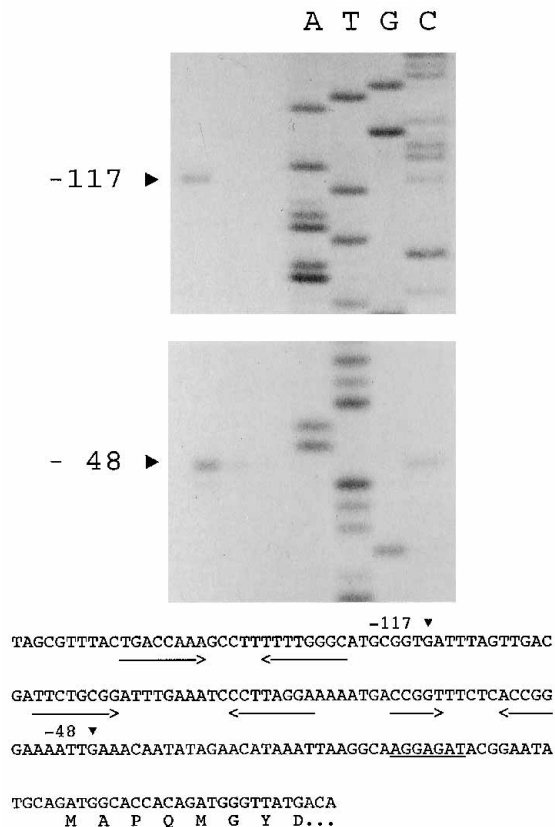


FIG. 7. **Primer extension analysis of the 5' end of *psmA*-specific mRNA.** Upper, the products of extension reactions using total RNA from trimethylamine grown (extreme left lane) *M. thermophila* and DNA sequencing reactions (lanes A, T, G, and C) in which the same oligonucleotide primer was used. Lower, the bases complementary to the 5' ends of the mRNA are indicated by down arrows and are numbered relative to the translational start site. A potential ribosome binding site is underlined. Predicted stem loop structures are indicated by arrows below the DNA sequence. The N-terminal deduced amino acid sequence of *psmA* is shown in single letter code beginning with the first base of each codon.

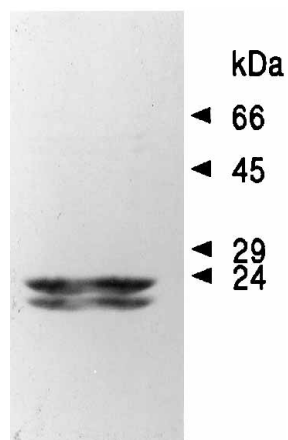


FIG. 8. **SDS-PAGE of the proteasome purified from *M. thermophila*.** Electrophoresis was performed as described under "Materials and Methods" using 12% polyacrylamide. Four μ g of protein was loaded. The molecular weight standards were: bovine plasma albumin, 66-kDa; ovalbumin, 45-kDa; bovine carbonic anhydrase, 29-kDa; and bovine trypsinogen, 24-kDa (Sigma). Protein was stained with Coomassie Blue R-250.

an archaeal proteasome in microbes from the Crenarchaeota kingdom (10); however, additional surveys are necessary before it can be concluded that the archaeal proteasome is unique to the Euryarchaeota.

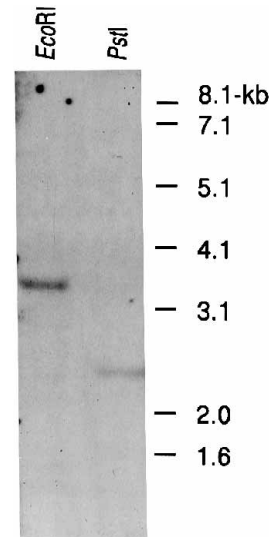


FIG. 9. **Southern hybridization analysis of *M. thermophila* genomic DNA using a probe specific for yeast ubiquitin genes.** *M. thermophila* genomic DNA was digested with *EcoRI* or *PstI* restriction endonuclease as indicated.

The availability of α and β subunit sequences for a second proteasome from an archaeon that is physiologically and phylogenetically distant from thermoplasma will allow comparisons to guide site-directed mutagenesis experiments for identification of amino acids involved in assembly, the catalytic mechanism, substrate targeting, regulation, and thermostability of archaeal and eucaryal proteasomes. These experiments will be especially productive because the crystal structure is known for the thermoplasma proteasome (5). Proteasomal α subunits from the *Eucarya* and thermoplasma (3) display putative NLS and cNLS motifs exemplified by ⁵⁰sdKKvR⁵⁵ and ²⁰²EEgEEIkapE²¹¹ (Fig. 3) for thermoplasma (where upper case letters conform to consensus sequences). The cognate methanosarcina sequences (⁴⁹vdKRit⁵⁴ and ²⁰¹EgkfdagtI^{E210}) (Fig. 3) have low identity with the putative thermoplasma NLS and cNLS motifs, suggesting these sequences are not strictly conserved in the archaeal proteasome. It is postulated that either a target for the NLS and cNLS exist in thermoplasma or that these sequence motifs existed prior to evolution of the corresponding eucaryotic receptor (3). Two-dimensional PAGE reveals multiple forms of the thermoplasma α subunit indicating it is modified which alters the pI (43). The thermoplasma α subunit contains the sequence ³²KKGST³⁶ (Fig. 3) which has identity to cAMP/cGMP-dependent phosphorylation sites, where serine is phosphorylated (3). The thermoplasma α subunit also contains a putative tyrosine autophosphorylation site (¹¹²LVKRVADQMQQYTQYGGVVRPY¹³²) (Fig. 3), where the underlined tyrosine is phosphorylated. The cognate methanosarcina sequences ³¹KRGTT³⁵ and ¹¹¹ISKKICDHKQTYTQYGGVVRPY¹³¹ (Fig. 3) have high functional similarity, suggesting they may be phosphorylation sites as predicted for the thermoplasma and eucaryal α subunits (3). The N-terminal sequence deduced for the methanosarcina α subunit is extended relative to the β subunit (Fig. 4), as is the case for thermoplasma (2). Deletion of this N-terminal extension in the thermoplasma α subunit prevents assembly of the α ring, demonstrating the importance of N-terminal sequences for assembly of an active proteasome (6). The N-terminal sequences may also be involved in recognition of substrates because the α rings are located on the ends of the barrel-like proteasome where the substrate is thought to enter (6). Unlike the proteasome purified from thermoplasma (3), N-terminal sequencing of the methanosarcina α subunit was not blocked. The sequencing revealed two N ter-

mini resulting from either dual translational start sites or processing. The significance is unknown; however, it is conceivable that subunits with different N termini may assemble into proteasomes with different substrate specificities.

The proteasome from *M. thermophila* is the first proteolytic enzyme isolated from methanogenic microbes for which there is considerable biochemical and physiological understanding (44). The discovery of proteasomes in *M. thermophila* will aid in understanding fundamental properties of the enzyme, including physiological roles in methanogenic microbes and other members of the *Archaea*. The native molecular mass and subunit composition of the methanosarcina proteasome was comparable to the thermoplasma and eucaryal 20 S proteasome; however, the results do not rule out the existence of a larger complex similar to the eucaryal 26 S complex which degrades proteins labeled with ubiquitin in an ATP-dependent process. Recently, ubiquitin has been identified in *T. acidophilum* (45). The thermoplasma proteasome degrades partially unfolded and ubiquitin-associated proteins (46), suggesting the possibility of ubiquitin-dependent proteolysis in this microbe. The hybridization of a *M. thermophila* genomic fragment to a yeast ubiquitin gene probe suggests that homologous sequences are present in this methanogenic microbe. Several physiological roles can be envisioned for ubiquitin-dependent proteolytic pathways in *M. thermophila*. The methanosarcina are the most versatile methanogenic microbes, having the ability to utilize several growth substrates and regulate the synthesis of catabolic enzymes in response to the growth substrate. Two-dimensional PAGE reveals more than 100 mutually exclusive peptides present in acetate- and methanol-grown *M. thermophila* (47), including subunits of the CO dehydrogenase enzyme complex which comprise approximately 10% of the cellular protein during growth on acetate (48). Proteolytic pathways for rapid turnover of specific enzymes would be advantageous when cells switch from one substrate to another. The rapid turnover of specific proteins could also be important during periods of stress such as heat shock or exposure to oxygen. All of these potential physiological roles of the archaeal proteasome pose interesting questions for further investigation.

Acknowledgments—We are grateful to Dr. K. T. Shanmugam for the use of his laboratory at the University of Florida. We thank Dr. M. Hochstrasser for providing plasmid YEP96 and Drs. L. O. Ingram and K. Shanmugam for helpful discussions.

REFERENCES

- Rivett, A. J. (1993) *Biochem. J.* **291**, 1–10
- Zwickl, P., Grziwa, A., Pühler, G., Dahlmann, B., Lottspeich, F., and Baumeister, W. (1992) *Biochemistry* **31**, 964–972
- Zwickl, P., Lottspeich, F., Dahlmann, B., and Baumeister, W. (1991) *FEBS Lett.* **278**, 217–221
- Dahlmann, B., Kopp, F., Kuehn, L., Nidel, B., Pfeifer, G., Hegerl, R., and Baumeister, W. (1989) *FEBS Lett.* **251**, 125–131
- Löwe, J., Stock, D., Jab, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) *Science* **268**, 533–539
- Zwickl, P., Kleinz, J., and Baumeister, W. (1994) *Struct. Biol.* **1**, 765–770
- Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R., and Baumeister, W. (1995) *Science* **268**, 579–582
- Lupas, A., Zwickl, P., and Baumeister, W. (1994) *Trends Biochem. Sci.* **19**, 533–534
- Arrigo, A.-P., Simon, M., Darlix, J.-L., and Spahr, P.-F. (1987) *J. Mol. Evol.* **25**, 141–150
- Pühler, G., Pitzer, F., Zwickl, P., and Baumeister, W. (1994) *System. Appl. Microbiol.* **16**, 734–741
- Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schoofs, G., Tanaka, K., De Mot, R., and Baumeister, W. (1995) *Curr. Biol.* **5**, 766–774
- Sowers, K. R., Nelson, M. J., and Ferry, J. G. (1984) *Curr. Microbiol.* **11**, 227–230
- Barrett, A. J. (1980) *Biochem. J.* **187**, 909–912
- Martinek, R. G., Berger, L., and Broida, D. (1964) *Clin. Chem.* **10**, 1087
- Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) *Anal. Biochem.* **150**, 76–85
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410
- Devereux, J., Haerberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
- Higgins, D. G., and Sharp, P. M. (1988) *Gene (Amst.)* **73**, 237–244
- Coux, O., Nothwang, H. G., Pereira, I. S., Targa, F. R., Bey, F., and Scherrer, K. (1994) *Mol. & Gen. Genet.* **245**, 769–780
- Philipps, G., and Gigot, C. (1993) EMBL accession Z26405
- Raynal, M., Grellet, F., Laudie, M., Meyer, Y., Cooke, R., and Delseny, M. (1994) EMBL accession Z29892
- Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., and Waterston, R. (1992) *Nature* **356**, 37–41
- Akiyama, K., Kagawa, S., Tamura, T., Shimbara, N., Takashina, M., Kristensen, P., Hendil, K. B., Tanaka, K., and Ichihara, A. (1994) *FEBS Lett.* **343**, 85–88
- Nothwang, H. G., Tamura, T., Tanaka, K., and Ichihara, A. (1994) *Biochim. Biophys. Acta* **1219**, 361–368
- Enenkel, C., Lehmann, H., Kipper, J., Gückel, R., Hilt, W., and Wolf, D. H. (1994) *FEBS Lett.* **341**, 193–196
- Heinemeyer, W., Tröndle, N., Albrecht, G., and Wolf, D. H. (1994) *Biochemistry* **33**, 12229–12237
- Sowers, K. R., and Gunsalus, R. P. (1988) *J. Bacteriol.* **170**, 998–1002
- Sowers, K. R., Thai, T. T., and Gunsalus, R. P. (1993) *J. Biol. Chem.* **268**, 23172–23178
- Puissant, C., and Houdebine, L.-M. (1990) *BioTechniques* **8**, 148–149
- Selden, R. F. (1987) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) pp. 4.9.1–4.9.5. John Wiley & Sons, New York
- Southern, E. (1975) *J. Mol. Biol.* **98**, 503–517
- Ecker, D. J., Khan, M. I., Marsh, J., Butt, T. R., and Crooke, S. T. (1987) *J. Biol. Chem.* **262**, 3524–3527
- Genschik, P., Philipps, G., Gigot, C., and Fleck, J. (1992) *FEBS Lett.* **309**, 311–315
- Shiratsuchi, A., and Sato, S. (1993) EMBL accession X57210
- Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R., and Wolfe, R. S. (1979) *Microbiol. Rev.* **43**, 260–296
- Glynn, R., Powis, S. H., Beck, S., Kelly, A., Kerr, L.-A., and Trowsdale, J. (1991) *Nature* **353**, 357–360
- DeMartino, G. N., Orth, K., McCullough, M. L., Lee, L. W., Munn, T. Z., Moomaw, C. R., Dawson, P. A., and Slaughter, C. A. (1991) *Biochim. Biophys. Acta* **1079**, 29–38
- Reeve, J. N. (1993) in *Methanogenesis* (Ferry, J. G., ed) pp. 493–527. Chapman and Hall, New York
- Gribskov, M., Homyak, M., Edenfield, J., and Eisenberg, D. (1988) *Comput. Appl. Biosci.* **4**, 61–66
- Olsen, G. J., Woese, C. R., and Overbeek, R. (1994) *J. Bacteriol.* **176**, 1–6
- Zwickl, P., Lottspeich, F., and Baumeister, W. (1992) *FEBS Lett.* **312**, 157–160
- Ferry, J. G. (1992) *Crit. Rev. Biochem. Mol. Biol.* **27**, 473–503
- Wolf, S., Lottspeich, F., and Baumeister, W. (1993) *FEBS Lett.* **326**, 42–44
- Wenzel, T., and Baumeister, W. (1993) *FEBS Lett.* **326**, 215–218
- Jablonski, P. E., DiMarco, A. A., Bobik, T. A., Cabell, M. C., and Ferry, J. G. (1990) *J. Bacteriol.* **172**, 1271–1275
- Terlesky, K. C., Nelson, M. J. K., and Ferry, J. G. (1986) *J. Bacteriol.* **168**, 1053–1058