J. Biochem. **138**, 193–199 (2005)
DOI: 10.1093/jb/mvi111

Expression and Molecular Characterization of Spherical Particles Derived from the Genome of the Hyperthermophilic Euryarchaeote *Pyrococcus furiosus*

Kazunori Namba[1], Kyoji Hagiwara[2], Hideaki Tanaka[1], Yuichiro Nakaishi[1], Khoon Tee Chong[1], Eiki Yamashita[1], George Enyimah Armah[1,3], Yasuko Ono[1], Yoshizumi Ishino[4], Toshihiro Omura[2], Tomitake Tsukihara[1] and Atsushi Nakagawa[1,*]

[1]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871; [2]National Agricultural Research Center, Tsukuba, Ibaraki, 305-8666; [3]Noguchi Memorial Institute for Medical Research, University of Ghana, P.O. Box LG581, Legon, Ghana; and [4]Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi, Fukuoka, Fukuoka 812-8581

Received December 28, 2004; accepted May 26, 2005

Spherical particles (SPs) of approximately 30 nm in diameter were found in the hyperthermophilic archaeon *Pyrococcus furiosus*. The SPs contained no nucleic acid and were composed of a single 39-kDa protein. The amino acid sequences of the amino-terminal and internal fragments were identical to portions of the deduced amino acid sequence of the putative 38.7-kDa protein encoded by the genome of *P. furiosus*, suggesting that the protein was expressed from the genome of *P. furiosus*. This possibility was confirmed by the observation that the 38.7-kDa protein expressed in *Escherichia coli* reacted specifically with the antibody against purified SPs, and it also formed SPs similar to those found in *P. furiosus*. Of the 345 amino acid residues in the 38.7-kDa protein, the amino-terminal 100 amino acids exhibited strong homology to putative proteins from other species of *Pyrococcus*, while the remaining 245 carboxy-terminal residues were not significantly homologous to putative proteins from other members of archaea. Thus, the carboxy-terminal region might be the product of a foreign gene that was incorporated relatively recently into the genome of *P. furiosus*.

Key words: archaea, hyperthermophile, *Pyrococcus furiosus*, spherical particle, virus-like particle.

Abbreviations: CBB, Coomassie Brilliant Blue; MALDI, matrix-assisted laser desorption ionization; PBS, phosphate buffer saline; PfSP, spherical particle from *Pyrococcus furious*; PAV1, *Pyrococcus abyssi* virus-like particle 1; PBS, phosphate buffered saline; PCR, polymerase chain reaction; PVDF, polyvinylidene difluoride; RDV, *Rice dwarf virus*; SSV1, *Sulfolobus solfataricus* virus 1; SP: spherical particle; STIV, *Sulfolobus* turreted icosahedral virus; TBST, Tris-buffered saline plus Tween-20; TOF, time of flight; VLP, virus-like particle.

The structural proteins that form virus particles or virus-like particles (VLPs) have the ability to form large multimers and also to generate unique three-dimensional particles. Thus, studies of the component proteins of viruses and VLPs have helped us to understand the biological and physical mechanisms that control the assembly of proteins. Moreover, studies of the formation of particles under extreme environmental conditions, such as high temperature, might help us to clarify the biochemical and biophysical stability of proteins and its role in the self-assembly of particles of high molecular mass.

In an analysis by electron microscopy of *Pyrococcus furiosus*, a euryarchaeote that grows optimally at temperatures close to 100°C, we observed spherical VLPs of approximately 30 nm in diameter. There have been several reports of archaebacterial VLPs or families of viruses (*1*) since the first report of a phage in *Halobacterium salinarium* in 1974 (*2*). Both filamentous (*3*) and rod-shaped (*4*) VLPs have been found in crenarchaea, namely, in *Sulfolobus*, while lemon-shaped VLPs have been isolated from euryarchaea, for example, from *P. abyssi* (PAV1) (*5*), and from *Sulfolobus solfataricus* (SSV1) (*6*). A novel spherical *Sulfolobus* turreted icosahedral virus (STIV) was isolated recently from a hyperthermo-acidophilic member of the genus designated YNPRC179 (*7*).

During our studies of the component molecules of *P. furiosus*, we found a novel SP that contained no nucleic acids. This particle was composed of a single species of protein which was encoded by the genome of *P. furiosus*. We describe here some of the biochemical and molecular properties of this SP, which we refer to as the *P. furiosus* spherical particle (PfSP).

MATERIALS AND METHODS

Culture of P. furiosus—A seed culture of *P. furiosus* (strain type, DSM 3638) (*8*) was incubated in 1 liter of medium that contained 35 g of sea salts (Aqua Marine S; Yashima Pure Chemicals, Osaka, Japan), 10 g of Tryptone, 5 g of yeast extract, 10 g of starch, 0.5 g of rezasurin (Sigma-Aldrich, St. Louis, MO, USA) and 10 ml of a stock

*To whom correspondence should be addressed. Tel: +81-6-6879-4313, Fax: +81-6-6879-4313, E-mail: atsushi@protein.osaka-u.ac.jp

solution of trace elements [constituents per liter: 1.5 g of nitriloacetic acid (pH 6.5), 0.3 g of $MgSO_4$, 0.1 g of NaCl, 10 mg of $FeSO_4 \cdot 7H_2O$, 10 mg of $CoSO_4$, 10 mg of $CaCl_2 \cdot 7H_2O$, 10 mg of $ZnSO_4$, 1 mg of $CuSO_4 \cdot 5H_2O$, 1 mg of $KAl[SO_4]_2$, 1 mg of H_3BO_3, 1 mg of $Na_2MoO_4 \cdot 2H_2O$ and 2.5 mg of $NiCl_2 \cdot 6H_2O$] (*9*). The culture was incubated at 98°C.

Purification of PfSPs—Cells in a culture of *P. furiosus* that contained PfSPs were lysed by freezing and thawing, with subsequent sonication for 5 min on ice with an ultrasonic disruptor (UD-201; Tomy Seiko, Tokyo, Japan). The suspension was centrifuged for 15 min at 39,500 × *g*, and the clear upper phase was subjected to centrifugation on a discontinuous density gradient of 20% to 40% sucrose in 50 mM phosphate buffer with 0.5 M NaCl (pH 7.0) for 4 h at 150,000 × *g* in an ultracentrifuge (XL-70; Beckman, Fullerton, CA, USA). The fraction containing the spherical particles, located at the boundary between 20% and 40% sucrose, was collected and dialyzed overnight against 50 mM phosphate buffer that contained 0.5 M NaCl, pH 7.0 (PBS) at 4°C . The dialysate was layered on a discontinuous gradient of 20% to 30% sucrose in PBS and centrifuged for 4 h at 150,000 × *g*. The resultant pellet, which contained PfSPs, was gently resuspended in PBS and dialyzed overnight against PBS that contained 1.5 M ammonium sulfate. The dialysate was applied to a hydrophobic column of TOYOPEARL™ Butyl-650S (Tosoh, Tokyo, Japan), which had been equilibrated with PBS that contained 1.5 M ammonium sulfate, and bound material was eluted with a gradient of ammonium sulfate from 1.39 M to 0.23 M. Each fraction was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in a 10% polyacrylamide gel.

Electron Microscopy—A drop of solution from each stage of the purification of PfSPs was placed on a copper grid that had been coated with carbon, and the material on the grid was then stained with 1.5% uranyl acetate for 30 s. After standing for 2 min, the stained material was observed with an electron microscope (H-7100; Hitachi, Tokyo, Japan).

Generation of Polyclonal Antibodies in Mice—Individual BALB/c mice were immunized with 180 μg of purified PfSPs in Freund's complete adjuvant and then given a booster immunization of 30, 20, 10 and 10 μg of the same antigen in incomplete adjuvant 1, 2, 3 and 4 weeks after the first immunization. Blood for the preparation of antiserum was collected 7 days after the last booster immunization. After blood had been centrifuged for 20 min at 6,000 × *g*, the supernatant containing polyclonal antibodies was recovered and stored at 4°C with 0.05% NaN_3 as a preservative.

Immunoblotting Analysis—Samples were subjected to electrophoresis for 1 h at 25 mA in the presence of SDS on a 12.5% polyacrylamide gel in a discontinuous buffer system (*10*). After transfer of proteins to a polyvinylidene difluoride (PVDF) membrane (Millipore Corporation, Bedford, MA, USA), the membrane was blocked by incubation in a solution of 1% (w/v) bovine serum albumin, which was followed by three washes with 20 mM Tris-HCl buffer (pH 7.5) that contained 0.5 M NaCl and 0.1% (w/v) Tween-20 (Atlas Chemical Industries, Inc., Wilmington, DE, USA; TBST). Then the membrane was incubated with PfSP-specific antibodies for 1 h at room temperature and washed another three times. The membrane was submerged in alkaline phosphatase-conjugated antibodies against mouse IgG, which had been diluted 5,000-fold with TBST and incubated for a further 1 h at room temperature. Finally, the membrane was washed three times with TBST and proteins were detected by staining with BCIP/NBT-Blue liquid (Sigma-Aldrich).

Stability of PfSPs—Purified PfSPs (20 mg/ml) were exposed to 0.15% and 0.30% (w/v) Triton X-100 in 20 mM Tris-HCl buffer (pH 7.5) for 3 min at 25°C. Dodecylmaltoside and Tween-20 were also tested under the same conditions. PfSPs at the same concentration were also treated with proteinase K (Sigma-Aldrich) at 1 mg/ml for 60 min at 56°C (*5*). After treatment, each sample was examined by electron microscopy as described above.

Extraction of Nucleic Acids from PfSPs—A mixture of approximately 15 mg of purified PfSPs in 1 ml of PBS with 100 μl of Bender buffer [0.1 M NaCl, 0.2 M sucrose, 0.1 M Tris-HCl (pH 8.0), 0.5 M ethylenediaminetetracetic acid (EDTA), 0.5% SDS] was incubated for 30 min at 65°C in a water bath. Then 30 μl of 8 M potassium acetate was added, and the mixture was allowed to stand for 45 min on ice. The mixture was clarified by centrifugation and the extracted materials were precipitated with isopropanol. The precipitate, resuspended in Bender buffer, was subjected to electrophoresis on a 0.8% agarose gel and stained with ethidium bromide. The materials were also subjected to electrophoresis overnight at 250 V on a 10% polyacrylamide gel in the presence of SDS in a discontinuous Tris-HCl buffer system (*10*) and then visualized by silver staining, as described by Herring *et al.* (*11*).

Determination of the Amino Acid Sequence of the Protein in PfSPs—Purified PfSPs were subjected to SDS-PAGE (12.5% polyacrylamide) for 1.5 h at 25 mA. After transfer to a PVDF membrane as described above, the protein was stained with Coomassie Brilliant Blue (CBB). The amino acid sequence of the amino-terminal region of the component protein of PfSPs was determined with an automated protein sequencer (Procise HE 491cLC; Perkin-Elmer Applied BioSystems, Foster City, CA, USA). Internal amino acid sequences of the protein were determined with a matrix-assisted laser desorption/ionization-time of flight (MALDI TOF) mass spectrometer (Voyager™ Elite XL; PerSeptive Biosystem, Framingham, MA, USA). The purified PfSP protein was digested with 5 U of lysylendopeptidase (Takara Bio Inc., Shiga, Japan) for 30 min at 37°C, and peptides were separated by SDS-PAGE (12.5% polyacrylamide) for 20 h at 100 V. The bands of protein were transferred to a PVDF membrane and stained with CBB, as described above, then subjected to MALDI-TOF mass spectroscopic analysis.

Cloning and Sequencing of the Gene from P. furiosus That Encodes the Component Protein of PfSPs—Genomic DNA of *P. furiosus* was purified as described previously (*12*), and the gene for the PfSP protein was amplified by the polymerase chain reaction (PCR) using the following primers: forward primer, 5′-GTGGTGATCCATATGCTCTC-3′ (PfSPcp.F-NdeI); and reverse primer, 5′-CCCAAGCTTGGGCTCTAAGACT-3′ (PfSPcp.R-HindIII). The primers were constructed by reference to the genomic

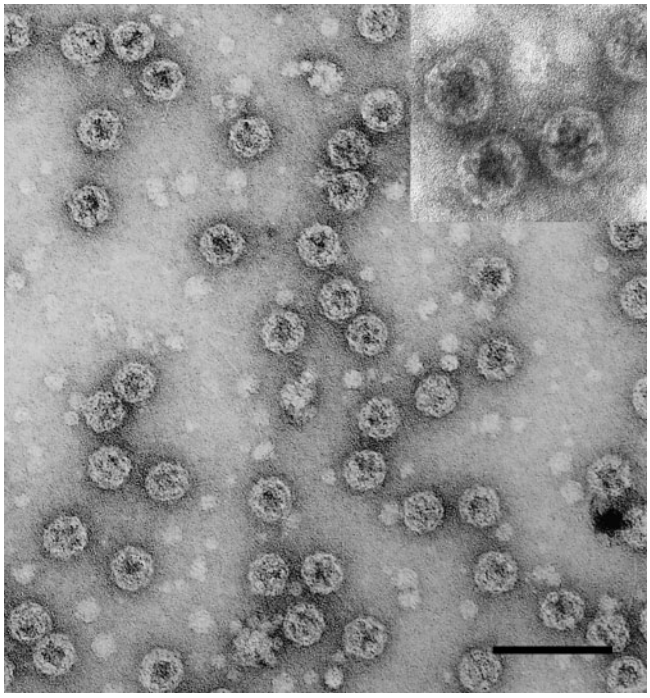

Fig. 1. **Electron micrographs showing purified PfSPs after negative staining with 1.5% uranyl acetate.** Bar represents 100 nm.

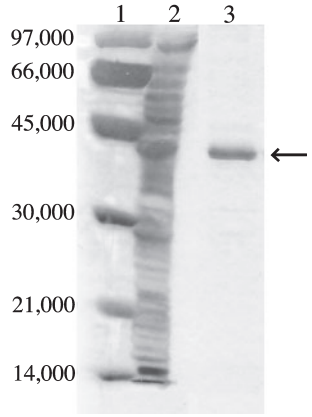

Fig. 2. **Results of SDS-PAGE of the component protein of PfSP.** Samples were fractionated by SDS-PAGE on a 12.5% polyacrylamide gel and stained with Coomassie Brilliant Blue. Lane 1, Prestained molecular mass markers (Da); lane 2, lysate from a culture of *P. furiosus*; and lane 3, purified PfSPs. The band of PfSP protein, with a molecular mass of approximately 39 kDa, is indicated by an arrow.

sequence of *P. furiosus* (DDBJ/EMBL/GenBank databases accession number AE010227) and extended from positions 8,333 to 8,346 and 7,308 to 7,320, respectively. Thirty cycles of PCR were performed with incubations for 1 min at 97°C, 1 min at 56°C, and 2 min at 72°C, using 250 ng each of the forward and reverse primers and 3 U of *Taq* DNA polymerase (Takara Bio Inc.). Amplified DNA was ligated into the cloning vector pCRII (Invitrogen Corp., Carlsbad, CA, USA) with a DNA Ligation Kit (Takara Bio Inc.), and then plasmids were introduced into *E. coli* JM109. The correct insertion of the amplified DNA in the pCRII vector was confirmed by restriction analysis with *Nde*I and *Hin*dIII. The nucleotide sequence of the cloned gene for the protein component of PfSP was determined with a BigDye™ Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer Applied BioSystems, Foster City, CA, USA) and an automatic sequencer (model 310; Perkin-Elmer Applied BioSystems).

Expression of the Component Protein of PfSP in E. coli—The full-length gene for the protein component of PfSP, cloned in pCRII, was digested with *Nde*I and *Hin*dIII and subcloned in the expression vector pET21b (Novagen, Madison, WI, USA) to allow modification of the carboxy-terminal domain with a histidine tag. The resultant plasmid was introduced into *E. coli* BL21 DE3 and cells were incubated for 6 h at 37°C. Expression of the PfSP protein was analyzed by immunoblotting with PfSP-specific antibodies. Cells that expressed the protein were stored at –80°C and used for the purification of particles.

The stored sample of cells was allowed to thaw partially at room temperature and then sonicated at a power setting of 6 with 60% duty cycle with an ultrasonic disruptor UD-201 (Tomy Seiko) for 5 min. Cell debris in the lysate was removed by centrifugation at 39,500 × *g* for 60 min (Himac CR 20 G; angle rotor, RPR20-2-950; Hitachi, Tokyo, Japan). The supernatant was applied to a column of Chelating Sepharose Fast Flow (Amersham Pharmacia) that has been equilibrated with PBS. After extensive washing, the column was eluted stepwise with 20, 100, 300 and 500 mM imidazole. The sample of his-tagged protein, eluted in 100 mM imidazole, was layered on a discontinuous gradient of 20% and 40% sucrose and centrifuged at 150,000 × *g* for 4 h in an ultracentrifuge (XL-70; Beckman). The pellet was gently resuspended in PBS and dialyzed overnight against PBS.

RESULTS

Purification of PfSPs—Many spherical VLPs of approximately 30 nm in diameter were observed under the electron microscope (Fig. 1) in the fraction that eluted with 1.13 M ammonium sulfate during hydrophobic column chromatography (see Materials and Methods), and these spherical VLPs showed a five-pointed-star configuration in the particles, which suggested five-fold symmetry of the particles. The VLPs generated a single band of a protein of approximately 39 kDa on SDS-PAGE (Fig. 2).

Stability of PfSPs—The stability of the PfSPs was investigated in the presence of standard detergents, 10% chloroform and a protease. In tests with dodecylmaltoside, Tween-20 and Triton X-100, the PfSPs were most sensitive to Triton X-100 (Fig. 3B). Most of the PfSPs were destroyed upon treatment with 0.3% Triton X100 (Fig. 3, B and C). The PfSPs were also sensitive to proteinase K, disintegrating completely on exposure to proteinase K for 60 min at 56°C (Fig. 3C).

Detection of Nucleic Acids—No nucleic acids were detectable in the material extracted from the purified sample of PfSPs after electrophoresis of this material on a 0.8% agarose gel (Fig. 4) and after SDS-PAGE (10% polyacrylamide; data not shown), Thus, the PfSPs appeared not to contain any nucleic acid.

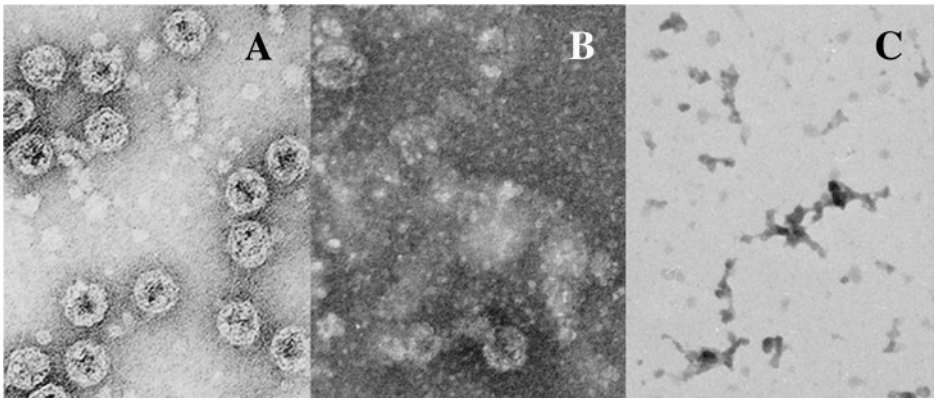

Fig. 3. **Stability of PfSPs in Triton X-100 and during incubation with proteinase K, as observed by electron microscopy.** (A) PfSPs before treatment; (B) PfSPs after treatment with 0.3% Triton X-100 for 3 min at 25°C and (C) PfSPs treated with proteinase K for 60 min at 56°C. All materials were negatively stained with 1.5% uranyl acetete. Bars represent 100 nm.

Amino Acid Sequences of Fragments of the Component Protein of PfSPs and Nucleotide Sequences—The amino acid sequence of the amino-terminal region of the PfSP protein was MLSINPTLI. A homology search using the BLAST program (*13*) revealed that this sequence was identical to part of the predicted amino acid sequence of a putative protein of approximately 11 kDa (DDBJ/EMBL/GenBank databases accession number AE010227; gene PF1192) encoded by an open reading frame in the genome of *P. furiosus* (*8*).

The nucleotide sequence of the archaeal genome that includes the region that encodes the putative 11-kDa protein consists of 1,038 nucleotides that encode a putative polypeptide of 345 amino acids (Fig. 5; DDBJ databease accession number AB214633) with a calculated molecular mass of 38.7 kDa. This molecular mass corresponds to that deduced for the PfSP protein, namely, 39 kDa. In addition, the amino acid sequences of four fragments of the PfSP protein (ILLDVAREEKAHVGEFMALLLNLDPEQVTELK, ELTGIEAHINDNKKEESNVEYFEKLRSALLDGVNKGRSLLK, QEYKPIPLLK, and KFYVGIRELNDGTYDVSIATKAGELLVKDEESLVIREILSTEGIKK) coincided with those of portions of the deduced amino acid sequence of the 38.7-kDa polypeptide (Fig. 5). It was evident that a frame shift, located at position of 8,076 in the genomic sequence in the database (DDBJ/EMBL/GenBank databases accession number AE010227), allowed translation of the predicted protein with the amino acid sequence of the 11-kDa protein to continue until a protein of 38.7 kDa had been translated.

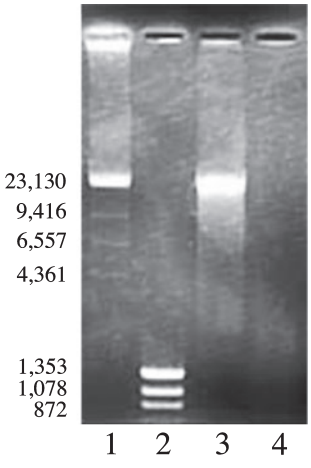

Fig. 4. **Results of agarose gel electrophoresis after attempted extraction of nucleic acids.** Lanes 1 and 2, Size markers. The sizes (in base pairs) of individual fragments of a bacteriophage λ DNA and ϕx174 that had each been digested with *Hin*dIII, are indicated on the left. Lane 3, genomic DNA purified from *P. furiosus*; and lane 4, material extracted from PfSPs.

Expression of the PfSP Protein in E. coli—Full-length DNA that corresponded to the gene for the PfSP protein with a histidine tag was expressed in *E. coli*, and the product was purified using Chelating Sepharose and the method described for purification PfSPs. The His-tagged PfSP protein was eluted with 100 mM imidazole. In immunoblotting analysis of cell extracts, a band of the product of the immunoreaction was clearly detected at the position of a 39-kDa protein (Fig. 6). Expression of full-length DNA that encoded the PfSP protein resulted in the formation of VLPs, which resembled intact PfSPs in both appearance and size (Fig. 7).

Homology Search for Proteins That Resemble the Structural Protein of PfSP—A homology search was performed using the BLAST program (*13*) and GENETYX version 11.0 software (Genetyx Corp., Tokyo, Japan) for multiple alignment of amino acid sequences (Fig. 8). The amino-terminal sequence of 100 amino acids of the PfSP protein was 53% identical and 74% similar to a protein from *P. abyssi* (DDBJ/EMBL/GenBank databases accession number AJ248284; positions 94,493 to 94,194), 52% identical and 71% similar to a protein from *P. horikoshii* strain OT3 (DDBJ/EMBL/GenBank databases accession number AP000007; positions 35,130 to 35,429), 60% identical and 81% similar to a protein from *Methanosarcina mazei* strain Goe1 (DDBJ/EMBL/GenBank databases accession number AE013538; positions 4,496 to 4,284), and 63% identical and 77% similar to a protein from *M. acetivorans* strain C2A (DDBJ/EMBL/GenBank databases accession number AE011049; positions 321 to 151; Fig. 8). Moreover, the full-length amino acid sequence of the PfSP protein was somewhat similar to that of a protein from *S. tokodaii* (29% identical and 49% similar; DDBJ/EMBL/GenBank databases accession number AP000984; positions 148,535 to 149,560), and to that of a protein from *S. solfataricus* (section 227 of 272 of the complete genome; 26% identical and 48% similar; DDBJ/EMBL/GenBank databases accession number AE006868; positions 9,971 to 8,949; data not shown).

```
          10         20         30         40         50         60
ATGCTCTCAATAAATCCAACCCTTATAAATAGGGATAAGCCATACACCAAAGAAGAATTA
 M  L  S  I  N  P  T  L  I  N  R  D  K  P  Y  T  K  E  E  L
          70         80         90        100        110        120
ATGGAAATTCTGAGATTAGCAATAATAGCTGAGCTTGATGCAATAAATCTATATGAGCAG
 M  E  I  L  R  L  A  I  I  A  E  L  D  A  I  N  L  Y  E  Q
         130        140        150        160        170        180
ATGGCAAGATATTCCGAAGATGAAAACGTAAGAAAGATTCTCCTTGATGTTGCAAGGGAG
 M  A  R  Y  S  E  D  E  N  V  R  K  I  L  L  D  V  A  R  E
         190        200        210        220        230        240
GAGAAGGCTCACGTTGGAGAATTCATGGCACTCCTTCTAAATTTGGATCCAGAACAAGTA
 E  K  A  H  V  G  E  F  M  A  L  L  L  N  L  D  P  E  Q  V
         250        260        270        280        290        300
ACTGAGCTGAAGGGTGGATTTGAGGAAGTCAAGGAGCTAACTGGAATTGAGGCTCACATA
 T  E  L  K  G  G  F  E  E  V  K  E  L  T  G  I  E  A  H  I
         310        320        330        340        350        360
AATGATAACAAGAAGGAGGAATCTAACGTGGAATATTTTGAGAAGTTGAGAAGTGCTCTT
 N  D  N  K  K  E  E  S  N  V  E  Y  F  E  K  L  R  S  A  L
         370        380        390        400        410        420
TTGGATGGGGTAAATAAGGGAAGATCTTTACTAAAGCACTTGCCAGTAACGAGAATAGAA
 L  D  G  V  N  K  G  R  S  L  L  K  H  L  P  V  T  R  I  E
         430        440        450        460        470        480
GGACAATCATTCAGGGTTGACATCATTAAATTTGAAGATGGTGTTAGGGTAGTAAAGCAA
 G  Q  S  F  R  V  D  I  I  K  F  E  D  G  V  R  V  V  K  Q
         490        500        510        520        530        540
GAGTACAAGCCAATTCCATTATTGAAGAAGAAATTCTATGTGGGAATAAGAGAACTCAAC
 E  Y  K  P  I  P  L  L  K  K  K  F  Y  V  G  I  R  E  L  N
         550        560        570        580        590        600
GATGGGACTTACGATGTATCTATAGCAACTAAAGCTGGGGAACTCCTAGTTAAAGACGAA
 D  G  T  Y  D  V  S  I  A  T  K  A  G  E  L  L  V  K  D  E
         610        620        630        640        650        660
GAGTCGCTAGTAATAAGAGAAATTCTATCCACAGAGGGAATTAAGAAAATGAAACTTTCC
 E  S  L  V  I  R  E  I  L  S  T  E  G  I  K  K  M  K  L  S
         670        680        690        700        710        720
TCCTGGGACAATCCAGAAGAAGCACTTAATGATTTAATGAACGCACTGCAGGAAGCCTCA
 S  W  D  N  P  E  E  A  L  N  D  L  M  N  A  L  Q  E  A  S
         730        740        750        760        770        780
AATGCGTCAGCTGGACCATTTGGGCTGATAATTAACCCAAAAAGATATGCAAAGTTGCTA
 N  A  S  A  G  P  F  G  L  I  I  N  P  K  R  Y  A  K  L  L
         790        800        810        820        830        840
AAAATTTACGAGAAGAGTGGAAAGATGCTAGTTGAAGTGCTAAAAGAGATATTCAGGGGA
 K  I  Y  E  K  S  G  K  M  L  V  E  V  L  K  E  I  F  R  G
         850        860        870        880        890        900
GGAATCATTGTTACTCTAAACATTGATGAGAATAAGGTAATAATATTTGCCAATACACCA
 G  I  I  V  T  L  N  I  D  E  N  K  V  I  I  F  A  N  T  P
         910        920        930        940        950        960
GCTGTTCTTGACGTTGTAGTTGGCCAAGATGTGACTCTACAAGAGCTAGGACCTGAAGGA
 A  V  L  D  V  V  V  G  Q  D  V  T  L  Q  E  L  G  P  E  G
         970        980        990       1000       1010       1020
GACGACGTTGCATTCTTGGTTAGTGAAGCTATAGGAATAAGAATAAAGAACCCAGAGGCA
 D  D  V  A  F  L  V  S  E  A  I  G  I  R  I  K  N  P  E  A
        1030       1040
ATAGTAGTCTTAGAGTGA
 I  V  V  L  E  *
```

Fig. 5. **Nucleotide sequence of the gene and the predicted amino acid sequence of the structural protein of PfSP.** The amino acid sequences of the amino-terminal nine residues (bold underlining) and four internal fragments of the PfSP protein, as determined in this study, are underlined. The asterisk indicates a termination codon.

A search for homologies between the amino acid sequences of the PfSP protein and viral proteins for which sequence information is available revealed that PfSP protein was 27.3% identical and 72.7% similar over a range of 22 amino acids to the envelope protein of lemon-shaped SSV1 (*6*) from *S. solfataricus* (DDBJ/EMBL/GenBank databases accession number J02396; positions 769 to 834), and 11.1% identical and 73.3% similar over a range of 45 amino acids to the structural protein of the icosahedral virus STIV (DDBJ/EMBL/GenBank databases accession number AY569307; positions 15,314 to 15,448). Thus, while viral proteins exhibited local homologies, no homology was found over extended regions, such as the amino-terminal 100 residues of the PfSP protein.

DISCUSSION

Purified PfSPs were composed of a single 39-kDa protein. The amino acid sequences of the amino-terminal fragment and four internal fragments of this protein were identical to portions of the deduced amino acid sequence

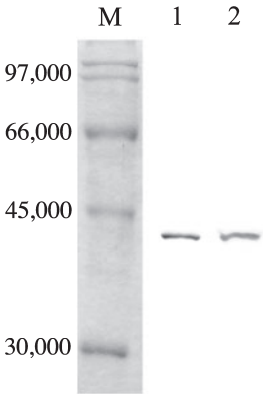

Fig. 6. **Immunoblotting analysis of recombinant PfSP protein.** Lane M, Molecular mass markers (Da); lane 1, purified PfSPs; lane 2, the supernatant of a lysate of *E. coli* cells that expressed the full-length gene for the PfSP protein.

of the putative 38.7-kDa protein encoded by the genome of *P. furiosus*. These results suggest that the 38.7-kDa protein that formed the particles of approximately 30 nm in diameter was encoded by the genome of *P. furiosus*. This hypothesis is supported by the fact that the protein expressed in *E. coli* from the putative gene isolated from the genome of *P. furiosus* had a molecular mass of 38.7-kDa and was able to form spherical particles in the absence of nucleic acids.

Lemon-shaped SSV1, which includes 15.5-kbp double-stranded circular DNA and is made up of three component proteins, and PAV1, which includes 18-kbp double-stranded circular DNA, have been identified in the hyperthermophile *P. abyssi* (*5*). Moreover, a thermophilic archeal virus with a genome of double-stranded DNA was found recently as a spherical virus in hyperthermophilic *Archea* (*6*). In contrast to these viruses and to VLPs that contain nucleic acid, the spherical PfSPs from *P. furiosus* did not contain a genome, and the unique component protein was encoded by the genome of the hyperthermophile *P. furiosus* itself.

The PfSPs were sensitive to detergents and proteinase K (Fig. 3), observations that suggest that PfSP does not have an envelope. The morphology under the electron microscope of proteasome particles derived from *P. furiosus* resembles that of PfSPs (*14*). However, the diameter of proteasome particles is less than half of that of the PfSPs, indicating that the total volume of each PfSP is more than eight times larger than that of a proteasome.

In a homology search using the full-length deduced amino acid sequence of the PfSP protein, we found that the proteins encoded by the genome of *S. tokodaii* and *S. solfataricus* were significantly similar. Thus, these members of archaea might also express proteins that can form VLPs of the type found in the present study.

The amino-terminal 100 amino acids of the PfSP protein were more than 70% similar to hypothetical proteins in members of *Euryarchaeota*, namely, *Pyrococcus* and *Methanosarcina*, and, in particular, in the hyperthermophilic archaea (Fig. 8). The fact that all these homologous proteins consist of close to 100 amino acids and are strongly conserved in these archaea suggests that they might be essential for the activities of these archaea (Fig. 9). Moreover, we found that the homologous domains of proteins in archaea were encoded by independent genes, while the remainder of the PfSP polypeptide, from residues 101 to 345, in *P. furiosus* was not significantly homologous to proteins from other members of archaea (Fig. 9). These results suggest that the gene for the carboxy-terminal 245 amino acids of the PfSP protein, which did not exhibit any homology to the genomes of members of archaea, might be a foreign gene derived from a microorganism, such as a bacteriophage.

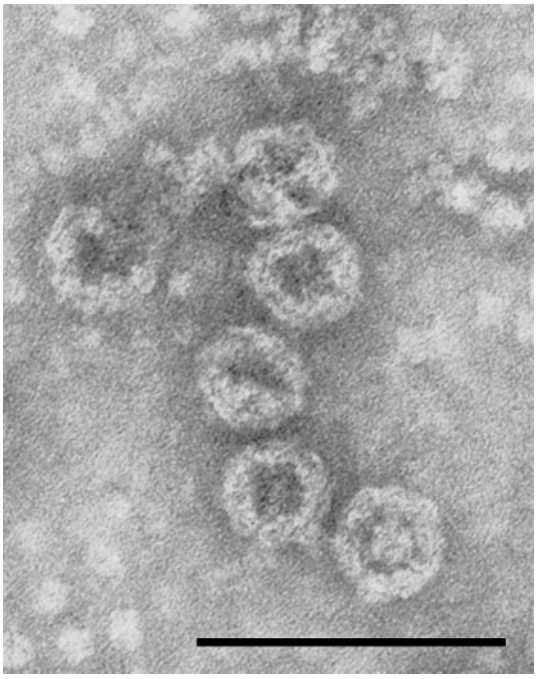

Fig. 7. **Electron micrograph showing particles composed of the protein expressed from the full-length gene for the PfSP protein in *E. coli*.** Bar represents 100 nm.

Various molecular-structural properties and conditions allow specific proteins to form spherical particles. One such property involves the electrostatic potential at interfaces of homologous proteins, as exemplified by the outer capsid proteins of *Rice dwarf virus* (RDV; *15*). Another is the extension and protrusion of one of the two proteins in a dimer into the other protein, to generate the building blocks of a larger unit, as found in the core particles of RDV (*16*). PfSP, which has the intrinsic ability to form ordered spherical particles, offers an excellent model system for understanding, at the level of atomic structure, the way in which individual polypeptides might form large spherical structures in the absence of nucleic acids.

The authors thank Prof. Toshifumi Takao for sequencing the fragments of the PfSP protein by mass spectrometry, Ms. Yumi Yoshimura and Prof. Shigemi Norioka for determining the amino acid sequence of the proteins, Mr. Takashi Wada for technical assistance with the electron microscopy, and Mr. Katsumi Hiraoka for generously providing access to an ultracentrifuge. This project was supported by a Grant-in-Aid for 21st Century Centers of Excellence (COE) program, by a Grant-in-Aid for Scientific Research (no. 1169202) and by the

```
P.furiosus (PfSP)      1 ---------------MLSINPILINTDKPYTKEELMEILRLAIIAELDAINLYEQMARYS   45
P.abyssi               1 ---------------MLAEKPYLLSREKPLSKKEIAQALRWAIEAELDAISFYEQLAELI   45
P.horikoshii           1 ---------------MLAEKPYLVEREKPLSKKEIAQALRWAIEAELDAINFYEQLAELI   45
M.mazei                1 ---------------MLSKIP--VDLKKISGEDIDKEILRAGIVAEIDAINLYEQMAALT   43
M.acetivorans          1 MSFHIITGILTGGKSLFSETI--ADLENTNQEDLDKEILRAAMIAELDAINLYEQMANLT   58
                                         .... . .   ...  . .  ..** ..  **.***..***.* .

P.furiosus (PfSP)     46 EDENVRKILLDVAREEKAHVGEFMALLLNLDPEQVTE-LKGGFEEVKELTGIEAHIN      101
P.abyssi              46 EDERIRKVFYDVANEEKEHVGEFLAVLLEVDEELV-EYIRKGFKEVEEGTGIKAKLK      101
P.horikoshii          46 EDEKIRHVFYDVANEEKEHVGEFLAVLLEVDEELA-EFIRKGFKEVEEETGIKA---      98
M.mazei               44 QNENIKRVLLDVAKEEKTHIGEFQALLLRFDAQQKQELEEGSKEVEEELSR------      94
M.acetivorans         59 KNEEIRIILLDIAREEKVHLAMFETVLLQTDSEFLQVYS----NYS--LARM-----      104
                            ..* .. ...*.* *** *...* ..**  * .   .    .. ........ .
```

Fig. 8. **Alignment of the partial amino acid sequences of the PfSP protein and of hypothetical proteins encoded by archaeal genomes.** Identical amino acids are boxed.

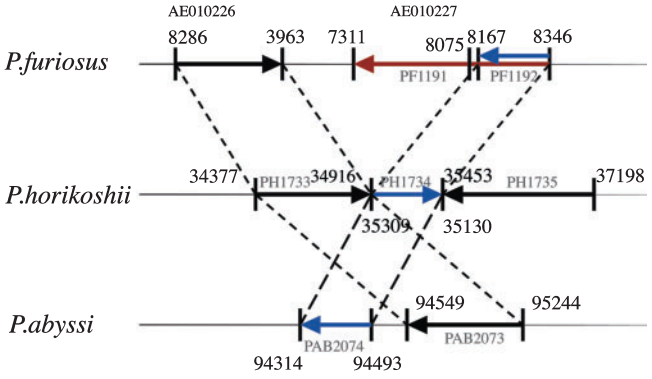

Fig. 9. **Gene mapping, showing relationships among three species of *Pyrococcus*.** Blue arrows: Nucleotide sequence that corresponds to amino acid residues 1 to 100 of the PfSP protein of *P. furiosus* and homologous regions in other species. Red arrow: Nucleotide sequence that corresponds to amino acid residues 1 to 345 of the PfSP protein. Numbers without capital letters refer to registered nucleotide sequences. Numbers with capital letters (AE, PF, PH and PAB) indicate genes registered in GeneBank and the DNA Data Bank of Japan (DDBJ).

National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan. A.N. and T.T. are members of the Structural Biology Sakabe Project (SBSP).

REFERENCES

1. Rice, G., Stedman, J., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., McDermott, T., and Young, M.J. (2001) Viruses from extreme thermal environments. *Proc. Natl Acad. Sci. USA* **98**, 13341–13345
2. Torsvik, T. and Dundas, I.D. (1974) Bacteriophage of *Halobacterium salinarium*. *Nature* **248**, 680–681
3. Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Weidmann, T.J.F., Kristjanson, J.K., Klenk, H.P., Nelson, K.E., and Fraser, C.M. (2000) A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. *Virology* **2**, 252–266
4. Prangishvili, D., Arnold, H.P., Gotz, D., Ziese, U., Holz, I., Kristjansson, J.K., and Zillig, W. (1999) A novel virus family, the *Rudiviridae*: structure, virus-host interactions and genome variability of the *Sulfolobus* viruses SIRV1 and SIRV2. *Genetics* **152**, 1387–1396
5. Geslin, C., Romancer, M.L., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "*Pyrococcus abyssi*". *J. Bacteriol.* **185**, 3888–3894
6. Chleper, C., Kubo, K., and Zillig, W. (1992) The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc. Natl Acad. Sci. USA* **89**, 7645–7649
7. Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J.E., Douglas, T., and Young, M. (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc. Natl Acad. Sci. USA* **101**, 7716–7720
8. Maeder, D.L., Weiss, R.B., Dunn, D.M., Cherry, J.L., Gonzalez, J.M., DiRuggiero, J., and Robb, F.T. (1999) Divergence of the hyperthermophilic archaea *Pyrococcus furiosus* and *P. horikoshii* inferred from complete genomic sequences. *J. Genet.* **152**, 1299–1305
9. Blumentals, I.I., Brown, S.H., Schicho, R.N., Skaja, A.K., Costantino, H.R., and Kelly, R.M. (1990) The hyperthermophilic archaebacterium, *Pyrococcus furiosus*. Development of culturing protocols, perspectives on scaleup, and potential applications. *Ann. N. Y. Acad. Sci.* **589**, 301–314
10. Schagger, H. and Jagow, G.V. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379
11. Herring, A.J., Inglis, N.F., Ojeh, C.K., Snodgrass, D.R., and Menzies, J.D. (1982) Rapid diagnosis of rotavirus infection by direct detection of viral nucleic acid in silver-stained polyacrylamide gels. *J. Clin. Microbiol.* **16**, 473–477
12. Greth, M.L., and Chevallier, M.R. (1970) A rapid method for the preparation of highly purified transforming DNA. *Biochim. Biophys. Acta* **213**, 335–342
13. Altschul, F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., W. Miller, and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402
14. Maupin-Furlow, J.A., Aldrich, H.C., and Ferry, J.G. (1998) Biochemical characterization of the 20S proteasome from the Methanoarchaeon *Methanosarcina thermophila*. *J. Bacteriol.* **180**, 1480–1487
15. Nakagawa, A., Miyazaki, N., Taka, J., Naitow, H., Ogawa, A., Fujimoto, Z., Mizuno, H., Higashi, T., Watanabe, Y., Omura, T., Chang, R.H., and Tsukihara, T. (2003) The atomic structure of *Rice dwarf* virus reveals the self-assembly mechanism of component proteins. *Structure* **10**, 1228–1238
16. Hagiwara, K., Higashi, T., Namba, K., Ichiki, T.U., and Omura, T. (2003) Assembly of single-shelled cores and double-shelled virus-like particles after baculovirus expression of major structural proteins P3, P7 and P8 of *Rice dwarf virus*. *J. Gen. Virol.* **84**, 981–984