



Discovery and Characterization of Thermoproteus Spherical Piliferous Virus 1: a Spherical Archaeal Virus Decorated with Unusual Filaments

Ross Hartman,^a Lieuwe Biewenga,^b Jacob Munson-McGee,^b Mohammed Refai,^{a,c} Eric S. Boyd,^c Brian Bothner,^a C. Martin Lawrence,^a Mark Young^c

^aDepartment of Chemistry and Biochemistry, Montana State University, Bozeman, Montana, USA

^bDepartment of Microbiology and Immunology, Montana State University, Bozeman, Montana, USA

^cDepartment of Plant Sciences and Plant Pathology, Montana State University, Bozeman, Montana, USA

ABSTRACT We describe the discovery of an archaeal virus, one that infects archaea, tentatively named Thermoproteus spherical piliferous virus 1 (TSPV1), which was purified from a *Thermoproteales* host isolated from a hot spring in Yellowstone National Park (USA). TSPV1 packages an 18.65-kb linear double-stranded DNA (dsDNA) genome with 31 open reading frames (ORFs), whose predicted gene products show little homology to proteins with known functions. A comparison of virus particle morphologies and gene content demonstrates that TSPV1 is a new member of the *Globuloviridae* family of archaeal viruses. However, unlike other *Globuloviridae* members, TSPV1 has numerous highly unusual filaments decorating its surface, which can extend hundreds of micrometers from the virion. To our knowledge, similar filaments have not been observed in any other archaeal virus. The filaments are remarkably stable, remaining intact across a broad range of temperature and pH values, and they are resistant to chemical denaturation and proteolysis. A major component of the filaments is a glycosylated 35-kDa TSPV1 protein (TSPV1 GP24). The filament protein lacks detectable homology to structurally or functionally characterized proteins. We propose, given the low host cell densities of hot spring environments, that the TSPV1 filaments serve to increase the probability of virus attachment and entry into host cells.

IMPORTANCE High-temperature environments have proven to be an important source for the discovery of new archaeal viruses with unusual particle morphologies and gene content. Our isolation of Thermoproteus spherical piliferous virus 1 (TSPV1), with numerous filaments extending from the virion surface, expands our understanding of viral diversity and provides new insight into viral replication in high-temperature environments.

KEYWORDS crenarchaea, *Globuloviridae*, archaeal virus

High-temperature acidic environments, such as those found in the hot springs of Yellowstone National Park (YNP), USA, provide a rich source of archaeal viruses (1). A range of virion morphologies is observed in archaeal viruses from these environments, including rod-shaped helical viruses, spherical icosahedral viruses, and unusual spindle- and bottle-shaped virions (2–6). Archaeal virus diversity is also reflected at the genome level, where most genes lack detectable homology to other known genes with known function (7, 8). The factors driving archaeal virus diversity in acidic hot springs are unknown but likely reflect adaptations to the physical and geochemical nature of these environments and the cell biology and biochemistry of their archaeal hosts.

Of the 62 known thermophilic archaeal viruses, 47 viruses infect hosts from the order *Sulfolobales* (6, 8–13). The remaining 15 viruses have been isolated from members

Citation Hartman R, Biewenga L, Munson-McGee J, Refai M, Boyd ES, Bothner B, Lawrence CM, Young M. 2020. Discovery and characterization of Thermoproteus spherical piliferous virus 1: a spherical archaeal virus decorated with unusual filaments. *J Virol* 94: e00036-20. <https://doi.org/10.1128/JVI.00036-20>.

Editor Rebecca Ellis Dutch, University of Kentucky College of Medicine

Copyright © 2020 American Society for Microbiology. All Rights Reserved.

Address correspondence to Mark Young, myoung@montana.edu.

Received 7 January 2020

Accepted 8 March 2020

Accepted manuscript posted online 25 March 2020

Published 18 May 2020

of the *Thermoproteales*, *Desulfurococcales*, *Methanobacteriales*, and *Thermococcales* (2, 14–18). This bias likely reflects the ease with which many *Sulfolobales* are isolated and cultured under predominantly aerobic, heterotrophic conditions. It is therefore reasonable to suspect that a substantial level of archaeal virus diversity remains unexplored.

In order to expand the diversity of cultured archaeal viruses, we sought to isolate new viruses infecting the order *Thermoproteales*. A common genus of this order, *Thermoproteus*, is found in sulfur sediments of YNP hot springs (19, 20). These anaerobic sulfur-reducing organisms have a global distribution in acidic and near-neutral thermal features (12, 13, 19, 21, 22). They are capable of growing chemolithotrophically using H_2 and CO_2 or organoheterotrophically using a wide variety of sugars, organic acids, and alcohols (23, 24).

To date, only 6 viruses have been described from *Thermoproteales* hosts. Despite this limited description, their structural and morphological diversity help define 3 new archaeal virus families (12, 13, 18, 25). *Thermoproteus tenax* virus 1 (TTV1), *Thermoproteus tenax* virus 2 (TTV2), and *Thermoproteus tenax* virus 3 (TTV3) are members of the *Lipothirixviridae*, a family of flexible helical viruses, whereas *Thermoproteus tenax* virus 4 (TTV4) is a member of the *Rudiviridae*, a related family of stiff rod-like viruses (18). The remaining 2 viruses are Pyrobaculum spherical virus (PSV) and *Thermoproteus tenax* spherical virus 1 (TTSV1), the only members of the *Globuloviridae* archaeal virus family (12, 13). We describe here a seventh *Thermoproteales* virus, *Thermoproteus* spherical piliferous virus 1 (TSPV1), with unusual hair-like filaments extending from its virion surface. We propose that these filaments facilitate viral attachment to host cells.

RESULTS AND DISCUSSION

Pure cultures of *Thermoproteus* sp. strain CP80, originally isolated from Cinder Pool (NHSP103) in YNP, were found to be chronically infected with a 75- to 85-nm spherical virus (Fig. 1A). There was no evidence of host cell lysis. Cultures remained viable after multiple rounds of passaging over a 2-year time period, with no significant change in either virus production or host cell viability.

Cesium chloride density gradient centrifugation was utilized to purify viral capsids that appeared by transmission electron microscopy (TEM) visualization to be intact. Virus particles purified at a density of 1.29 g/cm^3 , which is typical for enveloped virions. Using both negatively stained and frozen hydrated samples, electron micrographs and tomograms of the virus were acquired. These images revealed a spherical capsid morphology with an average diameter of 83 nm (Fig. 1C to F; see also Movie S1 in the supplemental material). Ten- to 20-nm protrusions were often observed extending from the surface of the virion (arrows, Fig. 1C). An external lipid envelope also appears to be present in the electron micrographs (arrows, Fig. 1F). A suspected 3.4-nm-thick ($\pm 0.3\text{ nm}$) lipid envelope and an envelope surface protein layer extending out to 7.4 nm ($\pm 0.4\text{ nm}$) are observed (Fig. 1F inset). The addition of 20% diethyl ether to a sample of TSPV1 resulted in the disruption of approximately 70% of the virions, with most of the remaining virus displaying obvious damage to their capsids, as determined by TEM analysis. In addition to this, TSPV1 has a buoyant density in CsCl density gradients that is typical for spherical enveloped viruses. Based on the apparent nonlytic lifestyle of TSPV1 and the likely presence of a lipid envelope, we hypothesize that TSPV1 buds from the host cell in a manner similar to that reported for the archaeal *Sulfolobus* spindle-shaped virus 1 (SSV1) (26).

The most striking and morphologically unique feature of the virion is numerous 3-nm-diameter filaments of various lengths that extend from the virion surface (arrows, Fig. 1C to F). Each virion contains an average of 7 filaments, with the typical range being from 0 to 20 filaments per virion. Some of these filaments extend up to 500 nm from the capsid surface and are highly flexible. In some micrographs, it appears as if the filaments pass through the surface protein layer and attach to the underlying envelope. However, higher-resolution microscopy is necessary to confirm this speculation. At this point, the nature of the association of the filaments with its virion is unknown.

The morphology of TSPV1 resembles those of PSV and TTSV, both members of the

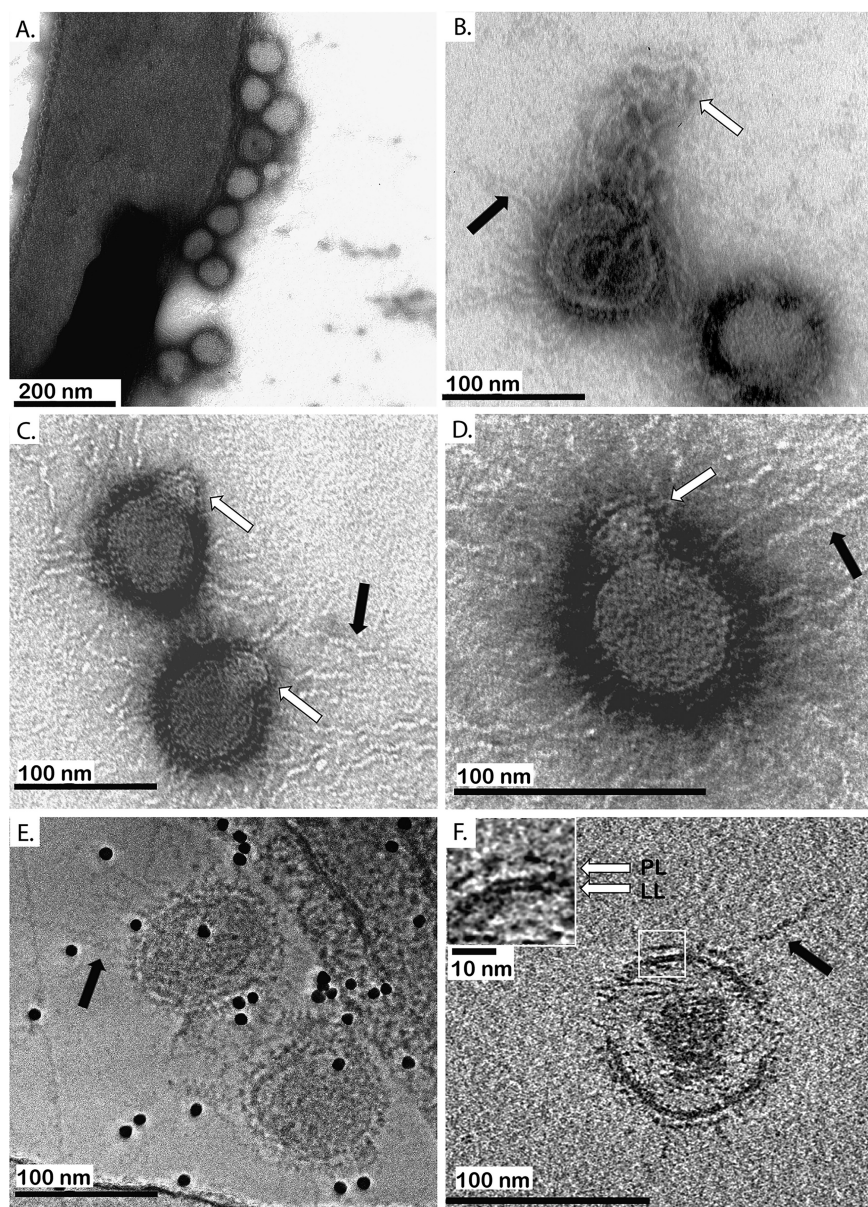


FIG 1 (A to F) TSPV1 visualized by negative staining (A to D) and cryo TEM (E and F). (A) TSPV1 particles in association with its *Thermoproteus* sp. CP80 host. (B) A ruptured TSPV1 particle showing an exposed 6-nm-diameter nucleoprotein complex (white arrow) and virion filaments (black arrow). (C) Purified particles showing virion protrusions (white arrows) and virion filaments (black arrows). (D) Enlarged negative stain of TSPV1 with protrusion (white arrow) and filaments (black arrow). (E) Cryo-TEM micrographs of TSPV1 with filaments (black arrow) associated with a *Thermoproteus* sp. CP80 cell. (F) Tomographic slice of TSPV1 (2.5-nm diameter) showing a virion filament (black arrow). An enlarged section showing the suspected 3.4-nm-diameter lipid envelope (LL; white arrow) and a 7.4-nm surface envelope protein layer (PL) are shown (white arrow).

Globuloviridae (12, 13). The resemblance includes the presence of protrusions from the capsid surface, the similar diameters of the virions, and the appearance of a lipid layer in micrographs. However, the filaments present on TSPV1 were not observed on PSV or TTSV1. Subjecting TSPV1 to increased pH from pH 4.0 to 8.0 ruptured the particles, releasing a nucleoprotein material (Fig. 1B) whose diameter matches that of the nucleoprotein observed in ruptured PSV particles (13).

Sequencing and assembly of the TSPV1 genome with 1,200 $\times$ average coverage yielded a 18,655-bp linear double-stranded DNA (dsDNA) genome with a G+C content of 56%, which is similar to that of its host (Fig. 2). The TSPV1 genome codes for 31 gene

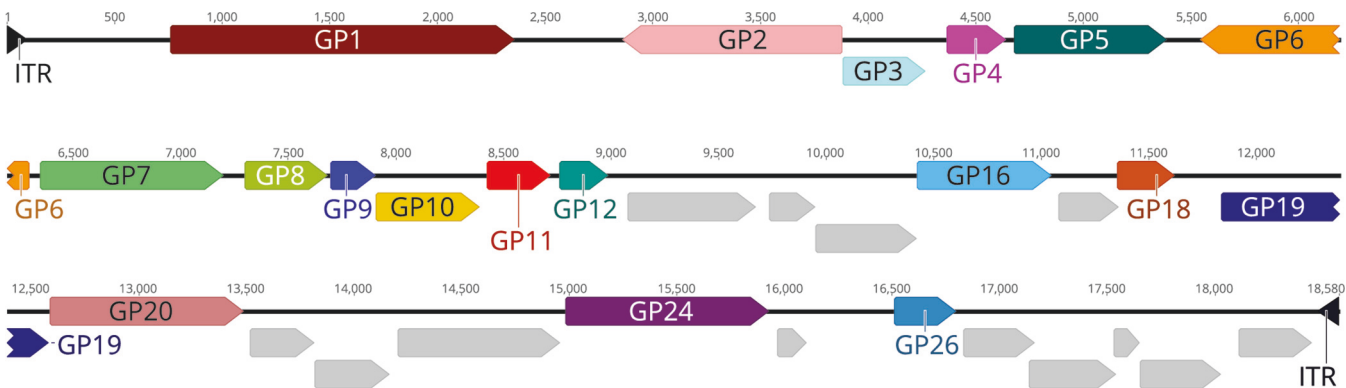


FIG 2 The genome map of the 18,655-bp dsDNA TSPV1 genome and its 31 predicted gene products (GPs). TSPV1 has 102-bp inverted terminal repeats (ITRs). GPs conserved across the *Globuloviridae* are colored.

products (GPs), with 85% coding density (Table 1). Most GPs occur on a single strand of the genome, with only GP2 and GP6 found on the complementary strand. The predicted masses of the GPs range from 4.7 to 59.2 kDa, with 19 GPs having a mass of <20 kDa. Some of TSPV1 predicted genes lack obvious promoter sequences and so may produce transcripts that are translated as operons. This includes GP3, GP4, GP8, GP9, GP14-GP17, GP19-GP22, and GP28-GP30. We could not detect a distinct origin of replication; however, the genome was found to have inverted terminal repeats (ITRs) of 102 bp in length which likely play a role in genome replication.

TABLE 1 Predicted TSPV1 proteins, characteristics, and homologues

TSPV1 gene	TTSV1 homologue	PSV homologue	Mol wt (kDa)	Predicted isoelectric point	No. of N-linked glycosylation sites predicted	Predicted signal peptide(s)	Structure/function annotations
GP1	GP1	GP2	59.2	8.80	1		AAA ⁺ helicase
GP2	GP4	GP3	33.4	4.70			Na ⁺ /Ca ²⁺ exchanger
GP3	GP5	GP5	16.0	8.63	1		Thioredoxin fold
GP4	GP6	GP8	10.2	6.35			
GP5	GP8	GP15	26.2	7.90			
GP6	GP27	GP16	26.5	8.98		Yes (FlaFind)	Minor virion protein
GP7	GP11	GP21	31.4	8.80			
GP8	GP13	GP23	13.3	8.62	1	Yes (FlaFind and SignalP)	
GP9	GP14	GP24	8.0	7.92	4		
GP10	GP15	GP25	18.3	7.39		Yes (FlaFind)	
GP11	GP16	GP26	11.1	6.52			Major virion protein
GP12	GP18	GP28	8.3	7.53			
GP13			21.0	9.04	7	Yes (SignalP)	
GP14			7.4	9.17			
GP15			18.0	9.92			
GP16	GP22	GP11	24.0	4.37			Armadillo α -solenoid repeat protein family
GP17			11.0	9.37	1		
GP18	GP19	GP29	10.0	8.46			
GP19	GP20	vp2	26.0	4.80			Minor virion protein (PSV)
GP20	GP21	GP17	31.7	5.83			
GP21			11.3	5.58			
GP22			13.4	4.56	1		
GP23	GP27		27.0	10.83	3		
GP24	GP24	VP3	33.5	6.50	6	Yes (SignalP)	Minor virion protein (PSV), filament protein (TSPV1)
GP25			4.7	9.29			
GP26		GP32	10.4	8.76			Possible anti-CRISPR protein
GP27			12.7	11.20			
GP28			15.6	10.45			
GP29				4.23			
GP30			14.3	9.38			
GP31		GP18	12.5	10			

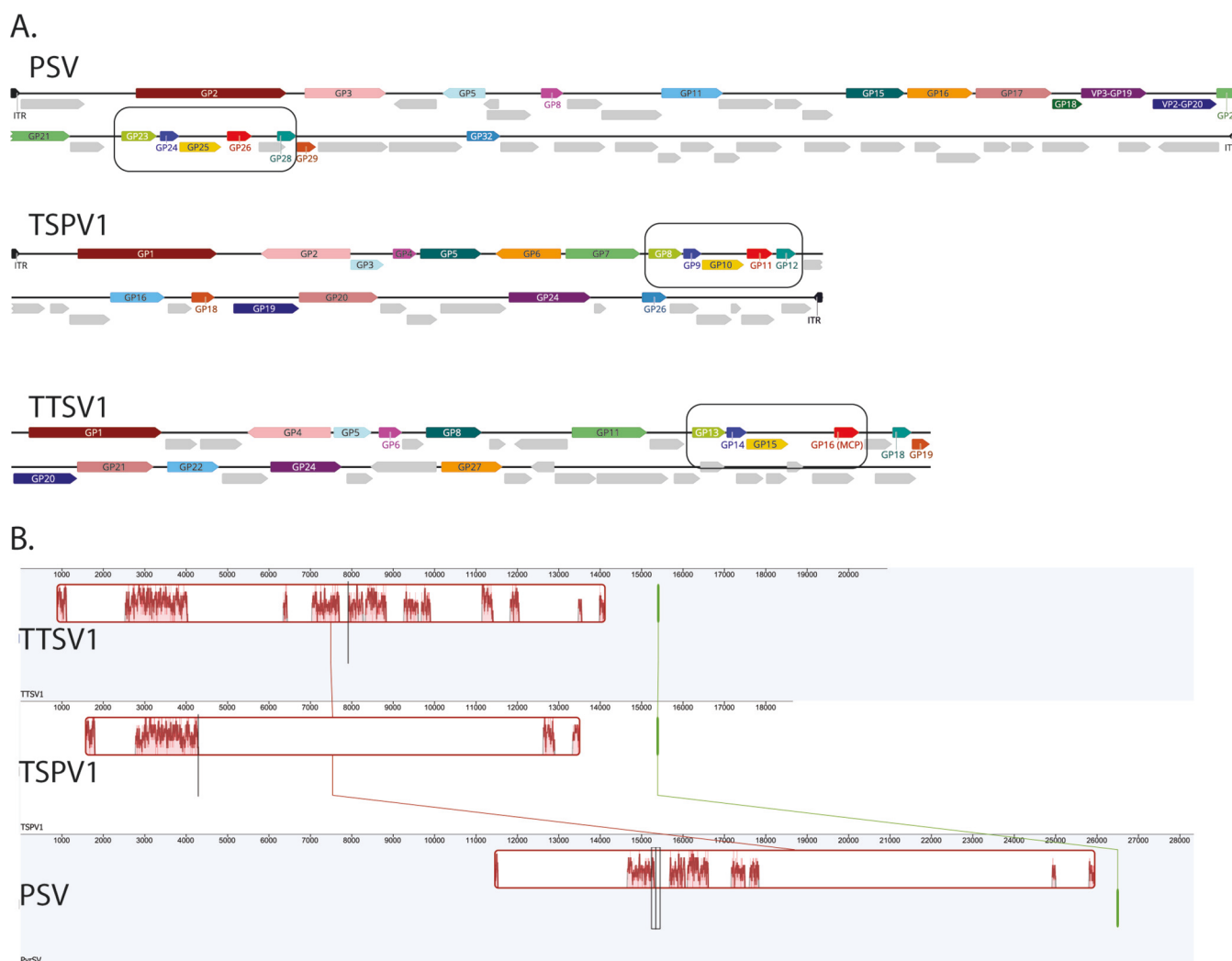


FIG 3 Genome comparisons of the three *Globuloviridae* members. (A) Genome maps for PSV, TSPV1, and TTSV1 are shown with conserved genes colored the same in each map. Eighteen of 31 TSPV1 GPs are shared with the other two *Globuloviridae* members. A block of genes (boxed regions) is conserved among the 3 viruses and contains the predicted major virion protein (TSPV1 GP11). (B) MAUVE alignment of the 3 *Globuloviridae* genomes showing only limited regions of sequence similarity at the nucleotide level; however, TSPV1 appears to be more similar to TTSV1 than to PSV.

Comparative genomic and virion morphology analyses place TSPV1 as a new member of the *Globuloviridae* family of archaeal viruses. The highest similarity of TSPV1 GPs was to predicted proteins found in two *Globuloviridae* members, PSV and TTSV1. Eighteen of the 31 TSPV1 GPs (60%) are conserved across these three *Globuloviridae* members (BLASTp E values $< 1 \times 10^{-5}$). Fifteen of these 31 GPs are exclusively found within *Globuloviridae* members and are not found in other viral or cellular genomes (Fig. 3A). One block of TSPV1 genes (GP8 to GP12) shows a high degree of synteny in all three viruses. Given that this block contains the major virion protein (MVP; discussed below), the other proteins in this gene block could play a role in virion assembly (Fig. 3A, boxed region). All three viruses have linear dsDNA genomes with ITRs of > 100 bp. TSPV1 and PSV have ITRs of similar length (120 bp). However, no sequence homology could be detected between the repeats of each virus. TSPV1 has the smallest genome of the three viruses, which have genomes of 18.6 kb, 21.6 kb, and 28.3 kb for TSPV1, TTSV1, and PSV, respectively. At the nucleotide level, TSPV1 is distinct from the other two viruses of the *Globuloviridae*.

MAUVE whole-genome alignment detected only small segments of the genome that could be aligned across the three *Globuloviridae* members (Fig. 3B). The overall maximum nucleotide identities of the TSPV1 genome with the PSV genome or TTSV1

genome are 2.2% and 1.1%, respectively; however, there are short regions of high identity. Not surprisingly, TSPV1 and TTSV1 share the highest identity most likely because they infect similar host genera.

A combination of analyses using BLAST, HHpred, and Phyre2 was employed to annotate the proteins of TSPV1. Despite this combined annotation approach, we were only able to assign tentative annotations for 9 TSPV1 GPs (Table 1). Two PSV GPs for which high-resolution structures have been determined by X-ray crystallography have identifiable homologs in TSPV1 (27). PSV GP11 (PDB [2X3M](#)) is homologous to TSPV1 GP16, and PSV GP32 (PDB [2X5C](#)) is homologous to TSPV1 GP26. Despite crystal structures, functions for these two proteins have yet to be determined. However, the PSV GP11 structure, and therefore the TSPV1 GP16 structure, suggests that it is a likely member of the armadillo α -solenoid repeat protein family that is known to orchestrate protein-protein interactions in a diversity of cellular functions (28). Though it has a different fold, the PSV GP32 structure has an overall morphology similar to that of the B116 protein of the *Sulfolobus* turreted spherical virus (PDB [2J85](#)) (29), which has recently been suggested to have potent anti-CRISPR activity (30). It is tempting to speculate that TSPV1 GP26 also has an anti-CRISPR function.

Both Phyre2 and HHpred analyses suggest that TSPV1 GP2 is similar to the sodium-calcium exchange (NCX) protein (79% of residues modeled with 93% confidence using Phyre2). The TSPV1 GP3 has a predicted thioredoxin fold, with 81% of residues modeled at 96% confidence with thioredoxin 2 from *Pseudomonas aeruginosa* (PDB [2LRC](#)); the role GP3 plays in TSPV1 biology remains to be elucidated. We suspect that TSPV1 buds from its host cell surface. Given this, we predict that virion envelope proteins would be located in the host membrane prior to budding, through either the Sec or Tat translocase pathway, both of which are present in archaea (31). Consistent with this, 5 TSPV1 proteins, GP6, GP8, GP10, GP13, and GP24, are predicted to contain N-terminal secretion signal peptides (Table 1).

SDS-PAGE analysis of the purified TSPV1 virions revealed an intense band migrating with an estimated molecular weight of 10 kDa, a less-intense band with an estimated molecular weight of 15 kDa, minor protein bands at 24, 35, 48, 60, and 115 kDa, and a protein smear stretching from the top of the gel downward to an estimated molecular weight of 200 kDa (Fig. 4A). In-gel tryptic digestion identified GP11 as the protein product for the 10-kDa band. Four peptide sequences belonging to GP11 were detected, resulting in 71% coverage of this protein. Based on band intensity, we assign GP11 as the major virion protein (MVP). This agrees with the previous results that found a homologous protein from TTSV1 (GP16) as the MVP (12). GP11 is predicted to contain four α -helices. Four-helix bundle proteins are commonly used to facilitate nucleic acid condensation (32). We suspect that GP11 is a major component of the virion nucleoprotein complex that serves to condense and form a superhelical, protein-bound form of the viral DNA. In-gel tryptic digestion of the 15-kDa band produced liquid chromatography-mass spectrometry (LC-MS) spectra matching to four peptides from TSPV1 GP6. We assign GP6 as a minor capsid protein. The predicted molecular weight of GP6 is 26 kDa, which suggests that this protein may be posttranslationally cleaved. In support of this, the four identified peptides map to the N terminus, with a predicted molecular weight of 15.6 kDa. Given its predicted secretion signal and presence in purified TSPV1 capsids, we suspect that GP6 is an envelope protein.

Glycan SDS-PAGE gel staining was used to examine the glycosylation status of virion-associated GP6, GP11, and GP24 based on the prediction that they too contain multiple N-linked glycosylation sites (Table 1). The results indicate that all three proteins are glycosylated (Fig. 4B). One role of glycosylation is to mediate virus-host interactions during attachment and entry (33, 34). We propose that TSPV1 glycosylation serves to increase the stability of the filaments and/or to mediate the recognition of host receptor proteins.

Also present in CsCl density gradients was a second protein band at a density of 1.35 g/cm³. Inspection of this band by negative-stain TEM revealed abundant filaments (Fig. 5A). SDS-PAGE analysis of these purified filaments resulted in a series of protein bands

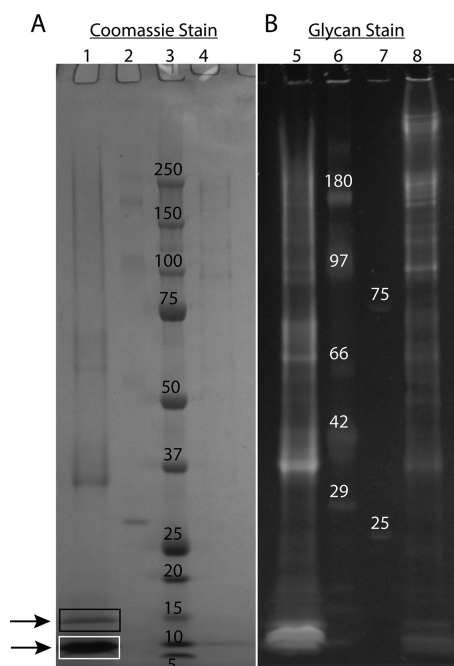


FIG 4 (A) Identification of GP11 and GP6 as major and minor virion proteins. An SDS-PAGE gel of purified TSPV1 virions showing major protein bands corresponding to a molecular weight (MW) of 10 kDa (white box) and 15 kDa (black box). The 15-kDa band produced spectra by LC-MS matching 4 peptides from GP6. The 10-kDa band produced spectra by LC-MS matching 4 peptides from GP11. (B) Glycosylation status of TSPV1 virions and purified filaments. Virion and purified filament proteins displayed on an SDS-PAGE gel stained with either Coomassie (lanes 1 to 4) or Pro-Q Emerald 488 glycoprotein stain (lanes 5 to 8). Lanes 1, 4, 5, and 8 display virion proteins, lanes 2 and 4 display molecular weight standards of known glycosylated proteins, and lanes 3 and 7 display molecular weight standards of known nonglycosylated proteins.

ranging in estimated mass from 35 kDa to >250 kDa (Fig. 5B). The prominent MVP bands at 10 kDa for GP11 and at 15 kDa for GP6 were absent from the sample. The series of protein bands observed in both intact capsids and purified filament samples indicates that the filaments are likely composed of multimers formed from protein(s) with a subunit mass of ~35 kDa. Only TSPV1 GP2 and GP24 have predicted molecular weights close to this value (Table 1). Since GP2 appears to be an NCX type transport protein, this implicates GP24 as a likely filament constituent.

N-terminal Edman sequencing of purified filaments identified one peptide sequence, AVFLVAVAIYITYT. This sequence is present near the N terminus of TSPV1 GP24. A nearly identical N-terminal sequence is present in the homologous PSV protein GP16. The sequenced N terminus is recessed 13 amino acids from the initiator methionine of PSV and 11 amino acids internal for TSPV1. Although the TSPV1 and PSV proteins have low overall homology (16% identity), the sequence immediately preceding the identified N-terminal match is well conserved in TSPV1, PSV, and TTSV1. This suggests that it may serve as a cleavable secretion signal.

To further validate the assignment of TSPV1 GP24 to the viral filaments, a variety of proteases and digestion conditions were explored prior to LC-MS analysis. Treatment with proteinase K, pepsin, and thermolysin produced no visible change in filament structure by TEM or detectable peptide signatures by LC-MS. Similarly, pH conditions ranging from 1 to 10 had no observable effect on the filament structure. However, exposure of samples to either 6 M guanidinium chloride or 8 M urea, combined with incubation at 98°C for 40 min, completely eliminated filament structures from the 8 M urea-treated samples and substantially reduced the number of filament structures in the guanidinium chloride-treated samples. The use of saturated urea and extended thermal denaturation was therefore incorporated into both in-solution and in-gel digestion of the purified viral filaments. LC-MS-based sequencing of both in-solution

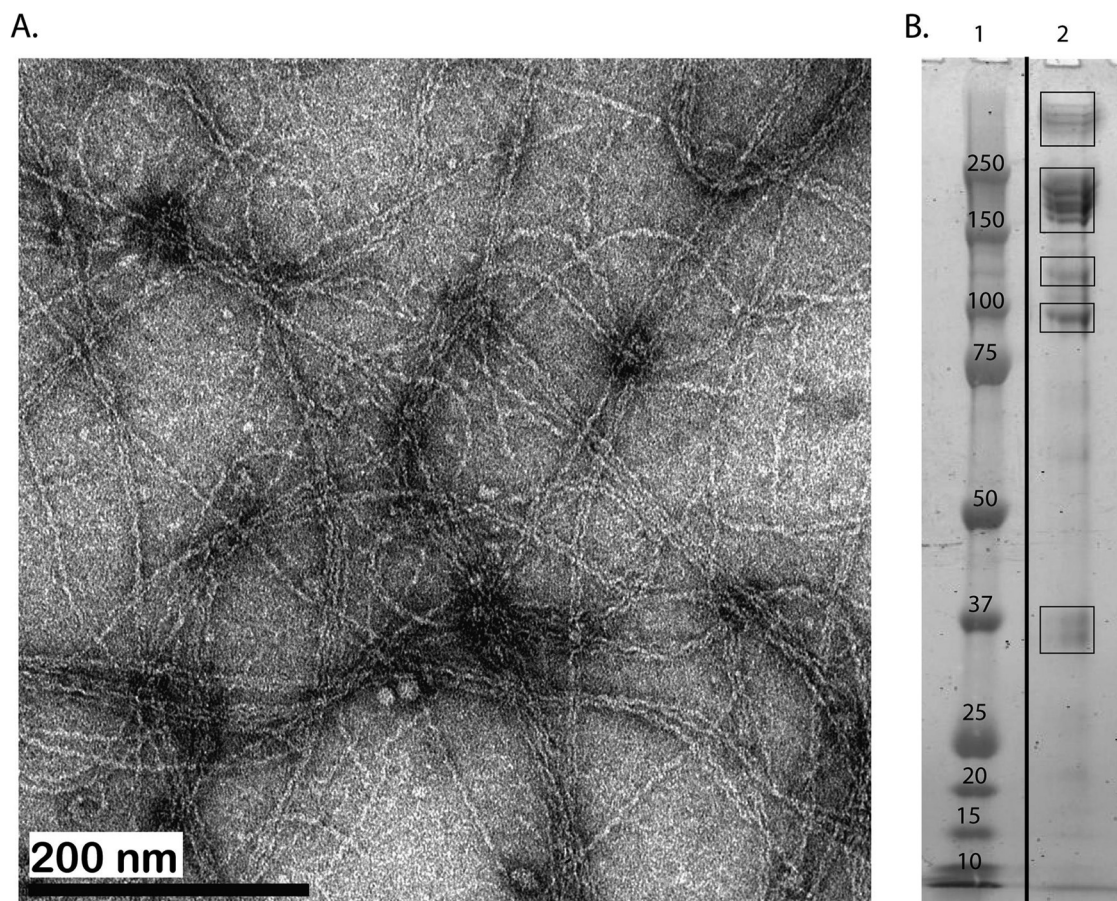


FIG 5 Purification and analysis of TSPV1 filaments. (A) TEM image of viral filaments banded in a CsCl gradient. (B) SDS-PAGE gel of molecular weight standards (lane 1) and purified filaments showing a series of high-molecular-weight bands ranging from ~35 kDa to >250 kDa (lane 2). Both N-terminal sequencing and LC-MS analysis support GP24 as the protein constituent of the viral filaments.

and in-gel digests yielded matches to peptides from only GP24. The identified peptides were VVDGKSLTTVSQLGSR, SLTTVSQLGS, and SLTTVSQLGSRLATIANNNTAYIPWIIIVSESK, giving a total protein coverage of 19% for GP24. These data support GP24 as a component of the filaments. However, at this point, we do not know if GP24 is the sole or even the major component of the filaments.

Surprisingly, similar filaments were not observed on the other two related *Globuloviridae* viruses, PSV and TTSV1, even though these viruses code for homologues of the TSPV1 GP24 filament protein (Table 1). It is unclear why only TSPV1 has these filaments, but it may reflect differences in purification methods or possibly the differences in time in which these viruses have been maintained under laboratory conditions.

It is worthwhile to speculate on the role the virion filaments play in the TSPV1 replication cycle. High-temperature, acidic, hot spring environments where TSPV1 is found typically harbor low total cell densities. Combined with the harsh chemical conditions, this creates an evolutionary impetus for viruses to develop strategies to increase the probability of interacting with a host cell. One solution is for the virus to target cellular pili, as is the case for SIRV and STIV (35, 36). It is conceivable that TSPV1 may decorate its virion with a structural analog to host pili and thereby increase its effective surface area and binding probability. The presence of thin filament extensions on the viral capsid would serve to increase the effective surface area. Despite its theoretical advantages, such a structural solution has not been observed in archaeal viruses other than TSPV1.

MATERIALS AND METHODS

Host and virus source. *Thermoproteus* sp. strain CP80 is an anaerobic, sulfur-reducing archaeon originally isolated from Cinder Pool (44°43'56.9"N, 110°42'32.4"W) in Yellowstone National Park (20). This thermal feature is a high-temperature (80 to 89°C) and low-pH (4.0 to 4.4) hot spring. Strain CP80 was isolated from sulfur-rich hot spring sediments, and 16S rRNA sequencing (20) and metagenomic sequencing of the sediment and planktonic communities typically show it to be among the most abundant taxa in these ecological compartments (E. S. Boyd, unpublished data). A pure culture of this strain was obtained by dilution to extinction, and purity was confirmed by 16S rRNA gene and genomic sequencing (20).

A pure culture of *Thermoproteus* sp. CP80 was screened for the presence of virus-like particles (VLPs) by negative-stain electron microscopy. Samples were stained with 2% uranyl acetate prior to TEM imaging using a LEO 912AB with 2K × 2K charge-coupled-device (CCD) camera (Carl Zeiss, Jena, Germany).

Culture conditions. Medium was prepared as described by Boyd et al. (37), with the following modifications: in place of peptone, 3 g/liter tryptone and 0.2 g/liter yeast extract were used as the sources of carbon. Anaerobic bottles (0.5 liters) were filled with 150 ml of medium, heated to 80°C, purged for 30 min with nitrogen gas, and inoculated with 5 ml of a log-phase cell culture. Cultures were incubated for 10 days at 80°C or until peak viral production was observed. Viral production was monitored by quantitative PCR (qPCR) amplification of a genome segment using qPCR primers orf14_F (TGGCCGGCCGCTTCAGATA) and orf14_R (GCGGTCGATAGGTTAGCGGGC). Standards for qPCR consisted of a 255-bp fragment from the predicted major virion protein GP11 cloned into a TOPO TA pCR2.1 vector (Thermo Fisher, Waltham, MA). SsoAdvanced Sybr green supermix was used for qPCRs, according to the manufacturer's directions (Bio-Rad, Hercules, CA). A Rotor-Gene series Q instrument (Qiagen, Hilden, Germany) was used for qPCR, with an initial denaturation at 98°C for 30 s, followed by 22 cycles at 98°C for 5 s and 65°C for 30 s.

Virus isolation and purification of virion filaments. Tangential flow filtration (TFF) was used to concentrate 12 liters of cell culture to 100 ml using a 10-kDa molecular weight cutoff (MWCO) polyethersulfone (PES) column (Amersham, Westborough, MA), followed by centrifugation for 15 min at 1,200 × *g* in a swinging bucket rotor to pellet cells and debris. Viruses present in the supernatant were further concentrated to 50 ml by TFF, followed by in-line filtration through 0.8-μm Acrodisc syringe filters (Pall, New York, NY) to remove any remaining cells. The final step in purification was accomplished by banding viruses on buoyant density gradients created by the addition of CsCl to a final density of 1.286 g/cm³, followed by ultracentrifugation for 24 h at 37,000 rpm (175,000 × *g*) in an SW41 rotor. Gradients were hand fractionated, dialyzed against 5 mM citric acid (pH 4), and analyzed by both qPCR and negative-stain electron microscopy. The protein concentrations in fractions were measured using Qubit fluorometric protein quantification (Invitrogen, Eugene, OR). Fractions containing separated filaments were further purified by 10 cycles of diafiltration with a 0.5-ml Amicon Ultra spin column and a 100-kDa MWCO (Millipore, Burlington, MA).

Cryo-electron tomography. For cryo-electron tomography, purified virus was frozen on Quantifoil R2/1 holey carbon grids (200-mesh copper) in liquid propane-ethane using a Vitrobot Mark III system (FEI, Hillsboro, OR). Tilt series were taken using a Titan Krios TEM (FEI) operated at 300 kV and equipped with a Volta phase plate (38, 39), a Quantum postcolumn energy filter (Gatan, Pleasanton, CA), and a Summit K2 camera (Gatan) operated in electron counting mode at the Max Planck Institute of Biochemistry, Martinsried, Germany. Tilt-series images were collected using SerialEM (40). Individual frames of images acquired with the K2 camera were aligned in DigitalMicrograph (Gatan). The acquisition parameters were magnification, ×35,700; tilt range, ± 60°; tilt increment, 2°; total dose, ~60 e⁻/Å²; pixel (px) size, 0.14 nm/px; and defocus with phase plate, -0.25 μm. The Volta phase plate was operated as previously described (41, 42).

Tomograms were reconstructed using IMOD 4.7 (43). Contrast transfer function (CTF) corrections were not performed. Ten-nanometer gold nanoparticles were used as fiducial markers for the alignment of tilt-series projection images. Local alignment solutions were used to refine rotation, magnification, and tilt angles, while a global solution was employed for distortion. Gold particles were erased from the final aligned stack. Images were binned 2×, radially filtered with a cutoff of 0.4 and falloff of 0.05, and run through 10 simultaneous iterative reconstruction technique (SIRT)-like filter iterations.

Genome sequencing and assembly. Viral nucleic acid was extracted from CsCl fractions using a PureLink viral RNA/DNA extraction kit (Thermo Fisher, Waltham, MA) and was sequenced at the University of Illinois Sequencing Center (Urbana-Champaign, IL) using Illumina HiSeq technology with 2 × 250-bp paired-end reads (Illumina, San Diego, CA). Genome assembly was performed using the MIRA assembly program (version 4.0.4). Initial assembly required subsampling 30,000 reads and assembling these in order to avoid assembly failure from high sequence coverage. Read recruitment to verify the inverted terminal repeats (ITRs) was performed using the Geneious assembler (Biomatters, Newark, NJ). For this process, the ITR region was trimmed from each end of the genome, and then paired reads were recruited to the trimmed sequence. Genome alignments between TSPV1, TTSV1, and PSV were conducted using MAUVE v.20159226 (44).

Gene prediction and annotation. Glimmer and Prodigal were used to predict viral genes (45, 46). Predicted gene products were compared to the NCBI nr database using BLASTp (47). In addition, proteins from both PSV and TTSV were queried against the TSPV1 genome to identify homologs using BLASTp (47). Protein alignments were generated using MUSCLE (48). Predicted proteins were further analyzed using the Phyre2 structure prediction server and HHpred to detect distant homologues (49, 50). PSIPRED,

FlaFind, and SignalP 5.0 were used to identify potential signal peptides in the predicted proteins of TSPV1 (51–53). PSIPRED was also used for protein secondary structure prediction (53).

N-terminal sequencing and LC-MS sequencing. N-terminal Edman sequencing was performed using purified filaments (5 µg) that were applied to a Prosorb cartridge, followed by 8 sequencing cycles conducted using a Shimadzu PPSQ-53A sequencer at the Iowa State University Sequencing Center.

LC-MS analysis was performed on in-solution digests of both purified viral capsids and purified filaments as follows: samples were dialyzed into 12.5 mM NH_4HCO_3 (pH 8), followed by the addition of mass spectrometry-grade urea to a final concentration of 8 M. Following this, samples were reduced in 5 mM dithiothreitol (DTT), alkylated with 14 mM iodoacetic acid (IAA), and digested for 4 h (37°C) at a 1:20 protease/protein ratio using trypsin–Lys-C mix (Promega, Madison, WI) with ProteaseMax (Promega, Madison, WI) addition to 0.03%. Samples were then diluted to a urea concentration of 1 M, followed by overnight digestion (37°C).

For in-gel digestion, 9 µg of purified capsids or 7 µg of purified filaments was dialyzed into 12.5 mM NH_4HCO_3 (pH 8), and then mass spectrometry-grade urea was added directly to a final concentration of 8 M together with SDS-PAGE loading buffer (50 mM Tris-HCl, 2% [wt/vol] SDS, 0.1% [wt/vol] bromophenol blue, 10% glycerol, 100 mM β-mercaptoethanol [pH 6.8]). Samples were denatured at 98°C for 40 min or at room temperature for 2 h. Electrophoresis was performed using a 4 to 12% Bis-Tris NuPAGE gel (Thermo Fisher, Waltham, MA) for separation. Gels were stained with colloidal Coomassie and imaged with a Typhoon Trio laser scanner (GE, Boston, MA). Gel slices were excised and proteolyzed as follows: gel slices were destained in 50 mM NH_4HCO_3 until clear and then reduced with 100 mM DTT, followed by alkylation with 55 mM iodoacetamide. Digestion was performed overnight in 25 mM NH_4HCO_3 using 0.3 µg sequencing-grade trypsin (Promega). Samples were subjected to reverse-phase chromatography using a Dionex nano-ultrahigh-performance liquid chromatography (nano-UHPLC) and then analyzed with a Bruker MaXis Impact mass spectrometer (Billerica, MA).

Data analysis was performed using MetaMorpheus v. 0.0.300 with precursor and product tolerance of 35 ppm, 3 max missed cleavages, and posttranslational modifications (PTMs) of common fixed, common variable, glycosylation, common biological, less common, common artifact, metal, trypsin digested, UniProt, and Unimod modifications. Searches were performed against a database of viral proteins, cellular proteins, and/or common contaminants. Alternatively, searches were performed with SearchGui v. 3.3.13, with precursor and product tolerance of 35 ppm, 3 max missed cleavages, oxidation of M, and carbamylation of K as variable modifications using X! tandem, MS-GF+, and Tide search engines (54).

Virion protein glycosylation analysis. The NetNGlyc server was used to predict N-linked glycosylation sites (55). For glycan staining, electrophoresis was performed with 9 µg of purified capsids or 7 µg of purified filaments using a 4 to 12% Bis-Tris NuPAGE SDS-PAGE gel (Thermo Fisher, Waltham, MA) for separation with 4 µl CandyCane ladder (Invitrogen, Eugene, OR) and 4 µl of Precision Plus Protein Dual Xtra ladder (Bio-Rad, Hercules, CA) as standards. The gel was stained with the Pro-Q Emerald 488 glycoprotein gel and blot stain kit (Invitrogen, Eugene, OR), following the manufacturer's recommendations, and visualized using an Alphamager 2200 imager (ProteinSimple, San Jose, CA).

Diethyl ether sensitivity assay. To test for the presence of a viral envelope, the method developed by Andrewes and Horstmann was employed. Briefly, diethyl ether was added to viral particles in 5 mM citric acid buffer (pH 4) (56). The final concentration of diethyl ether was 20% (vol/vol); samples were incubated at room temperature for 24 h and then analyzed by negative-stain TEM.

Data availability. The genome for TSPV1 is available under GenBank accession number [MT047590](#).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, MOV file, 14.6 MB.

SUPPLEMENTAL FILE 2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grants DEB-1342876 to M.Y., MCB-1413534 to C.M.L., and EAR-1820658 to E.S.B. Funding of the Proteomics, Metabolomics, and Mass Spectrometry Facility at Montana State University was made possible in part by the M. J. Murdock Charitable Trust, National Institute of General Medical Sciences, National Institutes of Health, award P20GM103474, the National Science Foundation (grant NSF-MRI:DBI-1532078), and the Office of the Vice President for Research and Economic Development.

REFERENCES

- Bolduc B, Wirth J, Mazurie A, Young M. 2015. Viral assemblage composition in Yellowstone acidic hot springs assessed by network analysis. *ISME J* 9:2162–2177. <https://doi.org/10.1038/ismej.2015.28>.
- Mochizuki T, Krupovic M, Pehau-Arnaudet G, Sako Y, Forterre P, Prangishvili D. 2012. Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc Natl Acad Sci U S A* 109:13386–13391. <https://doi.org/10.1073/pnas.1203668109>.
- Krupovic M, Quemin ER, Bamford DH, Forterre P, Prangishvili D. 2014. Unification of the globally distributed spindle-shaped viruses of the Archaea. *J Virol* 88:2354–2358. <https://doi.org/10.1128/JVI.02941-13>.

4. Prangishvili D, Mochizuki T, Krupovic M, ICTV Report Consortium. 2018. ICTV virus taxonomy profile: Guttaviridae. *J Gen Virol* 99:290–291. <https://doi.org/10.1099/jgv.0.001027>.
5. Häring M, Rachel R, Peng X, Garrett RA, Prangishvili D. 2005. Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, Acidianus bottle-shaped virus, from a new family, the Ampullaviridae. *J Virol* 79:9904–9911. <https://doi.org/10.1128/JVI.79.15.9904-9911.2005>.
6. Krupovic M, Cvirkaite-Krupovic V, Iranzo J, Prangishvili D, Koonin EV. 2018. Viruses of archaea: structural, functional, environmental and evolutionary genomics. *Virus Res* 244:181–193. <https://doi.org/10.1016/j.virusres.2017.11.025>.
7. Prangishvili D, Garrett RA, Koonin EV. 2006. Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. *Virus Res* 117:52–67. <https://doi.org/10.1016/j.virusres.2006.01.007>.
8. Prangishvili D, Bamford DH, Forterre P, Iranzo J, Koonin EV, Krupovic M. 2017. The enigmatic archaeal virosphere. *Nat Rev Microbiol* 15:724–739. <https://doi.org/10.1038/nrmicro.2017.125>.
9. Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, Connor R, Fiorini N, Funk K, Heffernon T, Holmes JB, Kim S, Kimchi A, Kitts PA, Lathrop S, Lu Z, Madden TL, Marchler-Bauer A, Phan L, Schneider VA, Schoch CL, Pruitt KD, Ostell J. 2019. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 47:D23–D28. <https://doi.org/10.1093/nar/gky1069>.
10. Munson-McGee JH, Snyder JC, Young MJ. 2018. Archaeal viruses from high-temperature environments. *Genes (Basel)* 9:128. <https://doi.org/10.3390/genes9030128>.
11. Wagner C, Reddy V, Asturias F, Khoshouei M, Johnson JE, Manrique P, Munson-McGee J, Baumeister W, Lawrence CM, Young MJ. 2017. Isolation and characterization of Metallosphaera turreted icosahedral virus (MTIV), a founding member of a new family of archaeal viruses. *J Virol* 91:e00925-17. <https://doi.org/10.1128/JVI.00925-17>.
12. Ahn DG, Kim SI, Rhee JK, Kim KP, Pan JG, Oh JW. 2006. TTSV1, a new virus-like particle isolated from the hyperthermophilic crenarchaeote Thermoproteus tenax. *Virology* 351:280–290. <https://doi.org/10.1016/j.virol.2006.03.039>.
13. Häring M, Peng X, Brugger K, Rachel R, Stetter KO, Garrett RA, Prangishvili D. 2004. Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera Pyrobaculum and Thermoproteus: a novel virus family, the Globuloviridae. *Virology* 323: 233–242. <https://doi.org/10.1016/j.virol.2004.03.002>.
14. Geslin C, Le Romancer M, Erauso G, Gaillard M, Perrot G, Prieur D. 2003. PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "Pyrococcus abyssi." *J Bacteriol* 185:3888–38894. <https://doi.org/10.1128/jb.185.13.3888-3894.2003>.
15. Gorlas A, Koonin EV, Bienvenu N, Prieur D, Geslin C. 2012. TPV1, the first virus isolated from the hyperthermophilic genus Thermococcus. *Environ Microbiol* 14:503–516. <https://doi.org/10.1111/j.1462-2920.2011.02662.x>.
16. Mochizuki T, Sako Y, Prangishvili D. 2011. Provirus induction in hyperthermophilic archaea: characterization of Aeropyrum pernix spindle-shaped virus 1 and Aeropyrum pernix ovoid virus 1. *J Bacteriol* 193: 5412–5419. <https://doi.org/10.1128/JB.05101-11>.
17. Mochizuki T, Yoshida T, Tanaka R, Forterre P, Sako Y, Prangishvili D. 2010. Diversity of viruses of the hyperthermophilic archaeal genus Aeropyrum, and isolation of the Aeropyrum pernix bacilliform virus 1, APBV1, the first representative of the family Clavaviridae. *Virology* 402:347–354. <https://doi.org/10.1016/j.virol.2010.03.046>.
18. Zillig W, Prangishvili D, Schleper C, Elferink M, Holz I, Albers S, Janekovic D, Gotz D. 1996. Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic Archaea. *FEMS Microbiol Rev* 18: 225–236. <https://doi.org/10.1111/j.1574-6976.1996.tb00239.x>.
19. Jay ZJ, Beam JP, Kozubal MA, Jennings RD, Rusch DB, Inskeep WP. 2016. The distribution, diversity and function of predominant Thermoproteales in high-temperature environments of Yellowstone National Park. *Environ Microbiol* 18:4755–4769. <https://doi.org/10.1111/1462-2920.13366>.
20. Urschel MR, Hamilton LR, Roden EE, Boyd ES. 2016. Substrate preference, uptake kinetics and bioenergetics in a facultatively autotrophic, thermoacidophilic crenarchaeote. *FEMS Microbiol Ecol* 92:fw069. <https://doi.org/10.1093/femsec/fw069>.
21. Yim KJ, Song HS, Choi JS, Roh SW. 2015. Thermoproteus thermophilus sp nov., a hyperthermophilic crenarchaeon isolated from solfataric soil. *Int J Syst Evol Microbiol* 65:2507–2510. <https://doi.org/10.1099/ijs.0.000293>.
22. Zillig W, Stetter KO, Schafer W, Janekovic D, Wunderl S, Holz I, Palm P. 1981. Thermoproteales: a novel type of extremely thermoacidophilic anaerobic archaeobacteria isolated from Icelandic solfataras. *Zentralbl Bakteriell Mikrobiol Hyg* 2:205–227. [https://doi.org/10.1016/S0721-9571\(81\)80001-4](https://doi.org/10.1016/S0721-9571(81)80001-4).
23. Bonch-Osmolovskaya EA, Miroshnichenko ML, Kostrikina NA, Chernych NA, Zavarzin GA. 1990. Thermoproteus uzoniensis sp. nov., a new extremely thermophilic archaeobacterium from Kamchatka continental hot springs. *Arch Microbiol* 154:556–559. <https://doi.org/10.1007/BF00248836>.
24. Siebers B, Zaparty M, Raddatz G, Tjaden B, Albers SV, Bell SD, Blombach F, Kletzin A, Kypides N, Lanz C, Plagens A, Rampp M, Rosinus A, von Jan M, Makarova KS, Klenk HP, Schuster SC, Hensel R. 2011. The complete genome sequence of Thermoproteus tenax: a physiologically versatile member of the Crenarchaeota. *PLoS One* 6:e24222. <https://doi.org/10.1371/journal.pone.0024222>.
25. Janekovic D, Wunderl S, Holz I, Zillig W, Gierl A, Neumann H. 1983. Ttv1, Ttv2 and Ttv3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaeobacterium Thermoproteus tenax. *Mol Gen Genet* 192:39–45. <https://doi.org/10.1007/BF00327644>.
26. Quemin ER, Chlanda P, Sachse M, Forterre P, Prangishvili D, Krupovic M. 2016. Eukaryotic-like virus budding in Archaea. *mBio* 7:e01439-16. <https://doi.org/10.1128/mBio.01439-16>.
27. Oke M, Carter LG, Johnson KA, Liu H, McMahon SA, Yan X, Kerou M, Weikart ND, Kadi N, Sheikh MA, Schmelz S, Dorward M, Zawadzki M, Cozens C, Falconer H, Powers H, Overton IM, van Niekerk CA, Peng X, Patel P, Garrett RA, Prangishvili D, Botting CH, Coote PJ, Dryden DT, Barton GJ, Schwarz-Linek U, Challis GL, Taylor GL, White MF, Naismith JH. 2010. The Scottish structural proteomics facility: targets, methods and outputs. *J Struct Funct Genomics* 11:167–180. <https://doi.org/10.1007/s10969-010-9090-y>.
28. Hatzfeld M. 1999. The armadillo family of structural proteins. *Int Rev Cytol* 186:179–224. [https://doi.org/10.1016/S0074-7696\(08\)61054-2](https://doi.org/10.1016/S0074-7696(08)61054-2).
29. Larson ET, Eilers BJ, Reiter D, Ortmann AC, Young MJ, Lawrence CM. 2007. A new DNA binding protein highly conserved in diverse crenarchaeal viruses. *Virology* 363:387–396. <https://doi.org/10.1016/j.virol.2007.01.027>.
30. Athukoralage JS, McMahon S, Zhang C, Gruschow S, Graham S, Krupovic M, Whitaker RJ, Gloster T, White MF. 2019. A viral ring nuclease anti-CRISPR subverts type III CRISPR immunity. *Nature* 577:572–575. <https://doi.org/10.1038/s41586-019-1909-5>.
31. Ellen AF, Zolghadr B, Driessen AM, Albers SV. 2010. Shaping the archaeal cell envelope. *Archaea* 2010:608243. <https://doi.org/10.1155/2010/608243>.
32. Krupovic M, Koonin EV. 2017. Multiple origins of viral capsid proteins from cellular ancestors. *Proc Natl Acad Sci U S A* 114:E2401–E2410. <https://doi.org/10.1073/pnas.1621061114>.
33. Montefiori DC, Robinson WE, Jr, Mitchell WM. 1988. Role of protein N-glycosylation in pathogenesis of human immunodeficiency virus type 1. *Proc Natl Acad Sci U S A* 85:9248–9252. <https://doi.org/10.1073/pnas.85.23.9248>.
34. Kandiba L, Aitio O, Helin J, Guan ZQ, Permi P, Bamford DH, Eichler J, Roine E. 2012. Diversity in prokaryotic glycosylation: an archaeal-derived N-linked glycan contains legionaminic acid. *Mol Microbiol* 84:578–593. <https://doi.org/10.1111/j.1365-2958.2012.08045.x>.
35. Quemin ER, Lucas S, Daum B, Quax TE, Kuhlbrandt W, Forterre P, Albers SV, Prangishvili D, Krupovic M. 2013. First insights into the entry process of hyperthermophilic archaeal viruses. *J Virol* 87:13379–13385. <https://doi.org/10.1128/JVI.02742-13>.
36. Hartman R, Eilers BJ, Bollschweiler D, Munson-McGee JH, Engelhardt H, Young MJ, Lawrence CM. 2019. The molecular mechanism of cellular attachment for an archaeal virus. *Structure* 27:1634–1646. <https://doi.org/10.1016/j.str.2019.09.005>.
37. Boyd ES, Jackson RA, Encarnacion G, Zahn JA, Beard T, Leavitt WD, Pi Y, Zhang CL, Pearson A, Geesey GG. 2007. Isolation, characterization, and ecology of sulfur-respiring crenarchaea inhabiting acid-sulfate-chloride-containing geothermal springs in Yellowstone National Park. *Appl Environ Microbiol* 73:6669–6677. <https://doi.org/10.1128/AEM.01321-07>.
38. Khoshouei M, Pfeffer S, Baumeister W, Forster F, Danev R. 2017. Subtomogram analysis using the Volta phase plate. *J Struct Biol* 197:94–101. <https://doi.org/10.1016/j.jsb.2016.05.009>.
39. Danev R, Baumeister W. 2017. Expanding the boundaries of cryo-EM with phase plates. *Curr Opin Struct Biol* 46:87–94. <https://doi.org/10.1016/j.sbi.2017.06.006>.
40. Mastrorade DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* 152: 36–51. <https://doi.org/10.1016/j.jsb.2005.07.007>.

41. Danev R, Buijsse B, Khoshouei M, Plitzko JM, Baumeister W. 2014. Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proc Natl Acad Sci U S A* 111:15635–15640. <https://doi.org/10.1073/pnas.1418377111>.
42. Fukuda Y, Laugks U, Lucic V, Baumeister W, Danev R. 2015. Electron cryotomography of vitrified cells with a Volta phase plate. *J Struct Biol* 190:143–154. <https://doi.org/10.1016/j.jsb.2015.03.004>.
43. Kremer JR, Mastrorade DN, McIntosh JR. 1996. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* 116:71–76. <https://doi.org/10.1006/jsbi.1996.0013>.
44. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
45. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27:4636–4641. <https://doi.org/10.1093/nar/27.23.4636>.
46. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
47. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
48. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
49. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.
50. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 Web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858. <https://doi.org/10.1038/nprot.2015.053>.
51. Armenteros JJA, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37:420–423. <https://doi.org/10.1038/s41587-019-0036-z>.
52. Szabó Z, Stahl AO, Albers SV, Kissinger JC, Driessen AJ, Pohlschroder M. 2007. Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J Bacteriol* 189:772–778. <https://doi.org/10.1128/JB.01547-06>.
53. Buchan DWA, Jones DT. 2019. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* 47:W402–W407. <https://doi.org/10.1093/nar/gkz297>.
54. Barsnes H, Vaudel M. 2018. SearchGUI: a highly adaptable common interface for proteomics search and de novo engines. *J Proteome Res* 17:2552–2555. <https://doi.org/10.1021/acs.jproteome.8b00175>.
55. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Lavery SB, Clausen H. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J* 32:1478–1488. <https://doi.org/10.1038/emboj.2013.79>.
56. Andrewes CH, Horstmann DM. 1949. The susceptibility of viruses to ethyl ether. *J Gen Microbiol* 3:290–297. <https://doi.org/10.1099/00221287-3-2-290>.