Genetic Analysis of the Major Capsid Protein of the Archaeal Fusellovirus SSV1: Mutational Flexibility and Conformational Change

Eric A. Iverson, David A. Goodman, Madeline E. Gorchels and

Kenneth M. Stedman

Center for Life in Extreme Environments

Biology Department

Portland State University

P.O. Box 751

Portland, OR 97207-0751

USA

ericiverson12@yahoo.com

dag2@pdx.edu

mgorchel@wellesley.edu

kstedman@pdx.edu (Corresponding)

Abstract:

Viruses with spindle or lemon-shaped virions are rare in the world of viruses, but are common in viruses of archaeal extremophiles, possibly due to the extreme conditions in which they thrive. However, the structural and genetic basis for the unique spindle shape is unknown. The best-studied spindle-shaped virus, SSV1, is composed mostly of the major capsid protein VP1. Similar to many other viruses, proteolytic cleavage of VP1 is thought to be critical for virion formation. Unlike half of the genes in SSV1, including the minor capsid protein gene *vp3*, the *vp1* gene does not tolerate deletion or transposon insertion. In order determine the role of the *vp1* gene and its proteolysis for virus function, we developed techniques for site-directed mutagenesis of the SSV1 genome and complemented deletion mutants with *vp1* genes from other SSVs. By analyzing these mutants we demonstrate that the N-terminus of the VP1 protein is required, but the N-terminus, or entire SSV1 VP1 protein, can be exchanged with VP1s from other SSVs. However, the conserved glutamate at the cleavage site is not essential for infectivity. Interestingly, viruses containing point mutations at this position generate mostly abnormal virions.

Keywords:

Introduction:

Viruses with spindle-shaped virions are rare in the virus world except in viruses infecting archaea from extreme environments [1]. Moreover, very little is known about how these unique structures are formed. How the spindle-shaped viruses that infect thermoacidophilic archaea retain their integrity the extreme (80C) temperatures and low pH (3) of their environment is also not clear. In order to answer these questions we work with model spindle-shaped viruses. These are the fuselloviruses or Sulfolobus Spindle-shaped Viruses (SSVs) [2]. The best-studied of the fuselloviruses is SSV1, originally isolated from *Sulfolobus shibatae* from an acidic hot spring in Beppu, Oita, Japan in the early 1980s [3,4]. SSV1 has a 15,465 basepair double stranded circular DNA genome, packaged as positively supercoiled DNA in a spindle or lemon-shaped virion with a short tail at one end [3,5,6]. SSV1 infects a number of *Sulfolobus* species and the virus genome is integrated into the host genome on infection [7–9]. How the virus genome is released or how SSV1 binds to *Sulfolobus* is not clear. SSV1 does not lyse cells when produced, yet slows host cell growth [7]. SSV1 virions seem to form by budding through the *Sulfolobus* membrane [10]. Purified SSV1 virions contain VP1 and VP3, hydrophobic capsid proteins, VP2, a DNA-binding protein, VP4 the putative tail protein and host-derived lipids [11,12]. Recently an approximately 32A resolution cryo-EM structure for SSV1 was determined and a novel fused-fullerene cone model for the structure was proposed [13]. This model is intriguing, as it is fundamentally geometrically similar to the capsid of HIV-1 [14], but not similar at all in sequence. At current resolution it is hard to determine

3

if this model is correct. Even more recently a comprehensive mutagenesis study of the SSV1 genome was performed, indicating that approximately ½ of the 35 ORFs in the SSV1 genome are not essential for virus function, including the minor capsid protein VP3 [15]. Interestingly, virions lacking the *vp3* gene have aberrant shapes yet still fit the fused fullerene cone model. Not surprisingly, the major capsid protein gene, *vp1*, was shown to be essential for virus function [15].

The SSV1 VP1 protein appears to be made as a precursor. In purified SSV1 particles, the N-terminal residue of the VP1 major capsid protein is a glutamate, indicating that the protein is made as a precursor protein and is then proteolytically cleaved by an unknown protease [5,11,12]. Proteolytic cleavage of virus structural proteins is very well-known; from the polyproteins of picornaviruses, flaviviruses and retroviruses [16] to the maturation of receptor-binding proteins in orthomyxoviruses [17]. There are now 10 described SSVs from hot springs throughout the world [2,15]. All known SSVs contain *vp1* genes with N-terminal extensions and conserved C-termini, including the universally conserved glutamate residue [2,15]. (Figure 1). While the C-terminus is extremely well conserved with amino acid identities between 78 and 97% to the SSV1 VP1 sequence, the sequences of the N-termini of SSV VP1s are much less so and their lengths are highly variable (Figure 1). This lack of conservation raises the questions of whether proteolysis is required, the N-terminus of the VP1 protein is required, whether a match between the N-terminus and C-terminus of the protein is necessary, and whether the conserved glutamate residue is required for virus function.
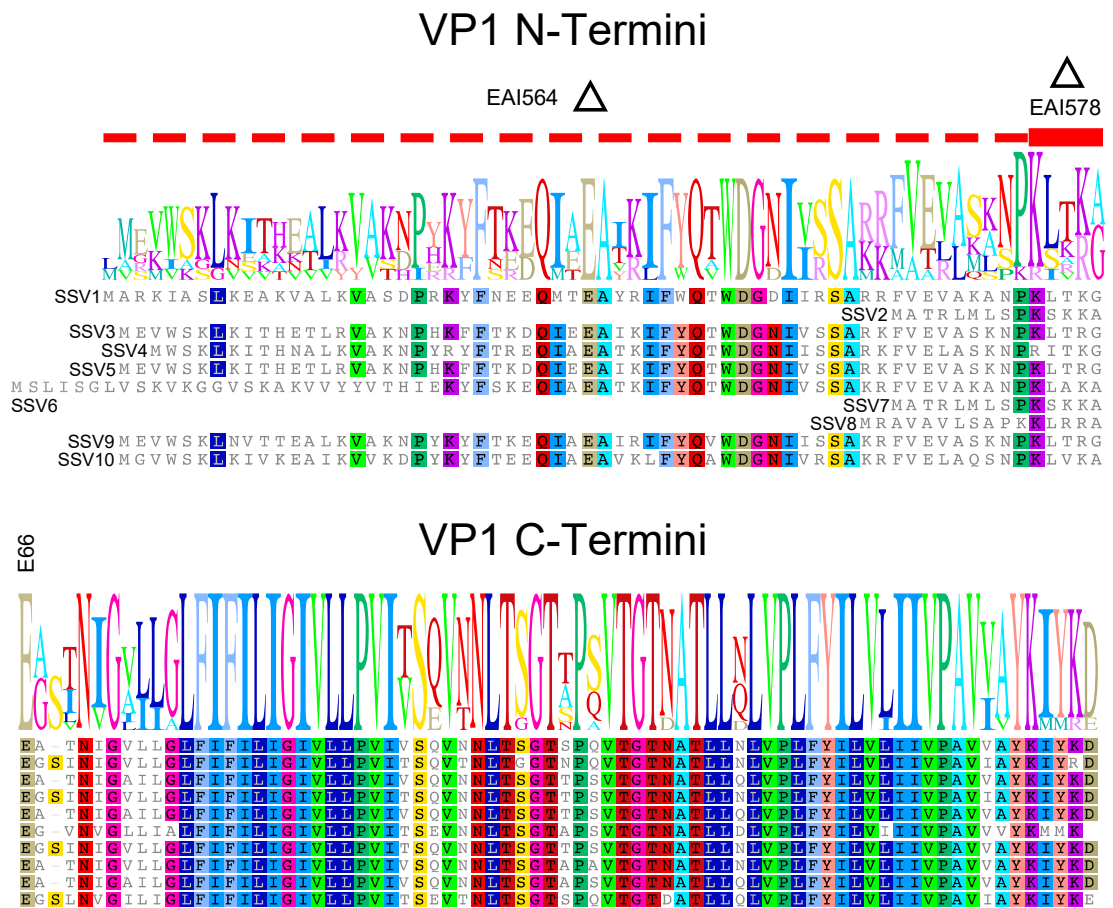
Figure 1

Figure 1: Amino Acid Sequence Alignment of Fusellovirus VP1 Proteins. VP1 protein sequences from SSV1 (top) to SSV10 (bottom) were aligned using Geneious V. 8.2 (Biomatters) and edited by hand. Accession numbers are listed in reference 15. Identical amino acids are highlighted in color. Sequences present in 9/10 sequences or as an insert are also highlighted in color. A sequence logo is shown above the alignment. The alignment is split at the universally conserved glutamate (E) at position 66 in the SSV1 sequence where proteolysis occurs and labeled. The 6 amino acids that are deleted in construct EAI578 are indicated with a red bar and a delta symbol. The sequence that is deleted in construct EAI564 is indicated with a dotted red line and a delta symbol.

In order to address the role of the N-terminus in VP1 function, we refined the long-inverse PCR technique which we developed to delete Open Reading Frames in the SSV1 genome [15,18,19] to make partial deletions, point mutations and complement mutants in the SSV1 genome. Both complete and partial deletions of the N-terminal region of the *vp1* gene of SSV1 led to non-functional virus. These deletions could be complemented with *vp1* genes or portions thereof from homologous SSV *vp1* genes indicating that proteolytic cleavage is necessary for virus function, but proteolysis is not specific to one type of SSV. Changing the conserved glutamate at the putative proteolytic cleavage site did not destroy activity, but virus particles with aberrant morphology were generated.

Materials and Methods:

Strains and cell growth:

All plasmids derived from EAI283 were grown in *E.coli strain* Transformax EC100D pir$^+$ (Epicentre). All plasmids derived from pAJC97 were grown in NovaBlue cells (Millipore). Sulfolobus strain S441 was isolated from Lassen Volcanic National Park [8] and was used for Sulfolobus transformations and as an indicator lawn for halo assays. Sulfolobus cells were grown in YS media [20] and *E.coli* were grown in Lysogeny Broth (LB) with appropriate antibiotics [21,22]. Chemically competent *E.coli* and electrocompetent Sulfolobus were prepared and transformed as in Iverson et al. [15].

Purification of LIPCR template DNA from E. coli

E. coli cells harboring the pAJC97 or EAI283 SSV1 shuttle vectors (Table 1) were grown overnight in 5 mL of LB medium containing 50 µg/mL Kanamycin. Plasmid DNA was then purified from 1.5 mL of overnight culture via alkaline lysis [23]. Following ethanol precipitation, the DNA was dissolved in 30 µL diH$_2$O with 0.01 µg of RNaseA.   DNA was analyzed by EcoRI restriction endonuclease digestion following the manufacturer's protocol (Thermo-Fisher).

Table 1: Plasmids used in this study.

Plasmid	Description	Infectivity	Reference
pAJC97	SSV1 shuttle vector (TOPO PCR Blunt II inserted into ORF *e178* at bp 3,173)	**+**	[18]
EAI283	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 3,572 (ORF *e178*)	**+**	[15]
EAI427	pAJC97 background with E66Q mutation in VP1	**+**	This Work
EAI500	pAJC97 background with E66A mutation in VP1	**+**	This Work
EAI551	EAI283 background with SSV1-*vp1* complementation	**+**	This work
EAI566	EAI283 background with SSV2-*vp1* complementation	**+**	This Work
EAI553	EAI283 background with SSV9-*vp1* complementation	**+**	This Work
EAI557	EAI283 background with N-terminus of SSV9-VP1 complementation	**+**	This Work
EAI237	pAJC97 background with N-terminus of VP1 deleted	**—**	This Work
EAI564	EAI283 background with N-terminus of VP1 deleted	**—**	This Work
EAI578	EAI283 background with residues 61-65 deleted from VP1	**—**	This Work

Long inverse PCR (LIPCR)

LIPCR was used to delete the *vp1* gene, portions of the *vp1* gene, or to introduce single base pair mutations into the SSV1 genome [15,18,19]. All primers used in this study are listed in Table 2. For *vp1* single base pair mutations, primers were designed such that ligation of the LIPCR product resulted only in the desired base pair change.   Optimal concentrations of

template DNA for LIPCR was determined empirically for each set of primers. LIPCR was performed as described previously [19]. LIPCR reactions (20 µL) were prepared following the DNA polymerase manufacturer's protocol except that the polymerase concentration was reduced from 0.02 U/µL to 0.005 U/µL. Initial annealing temperatures for each primer pair were estimated using NEB Tm prediction software (http://tmcalculator.neb.com/#!/) and were optimized experimentally. The reaction that generated the most full length product with the least smaller products was selected. LIPCR products were purified, ligated, and transformed into chemically competent NovaBlue or Transformax EC100D *E. coli* as described previously [19].

Table 2: Primers used in this study

Product	Forward primer	Reverse Primer
SSV1 *vp1* deletion	**TGA** GGG ATG GAA ATC AGT TTA AAG	**CAA** ACT CCT TAG GAG TCT CAT CC
SSV1 VP1 N-terminus deletion	*ATG* **GAA GCA ACC AAC ATA GG** (61) and **GAA GCA ACC AAC ATA GG** (54)	**CAA** ACT CCT TAG GAG TCT CAT CC
SSV1 VP1 aa61-65 deletion	**GAA GCA ACC AAC ATA GGC**	**GGG GTT TGC CTT TGC TAC**
SSV1 VP1 point mutant (E66A)[A]	**G**CA GCA ACC AAC ATA GGC	**ACC TTT TGT GAG CTT GGG G**
SSV1 VP1 point mutant (E66Q)[B]	**C**AA GCA ACC AAC ATA GGC	**ACC TTT TGT GAG CTT GGG G**
SSV1 *vp1* [C]	GCcAGAAAGATAGCCTCAC	ACCTTTTGTGAGCTTGGG
SSV9 *vp1*	GAAGTTTGGTCAAAGTTAAACG	ATCTTTGTAGATTTTATACG
SSV2 *vp1*	GCCACCAGACTAATGCTAAGC	GTCACGATATATCTTATACGCT ATGAC
SSV9 VP1 N-terminus	GAAGTTTGGTCAAAGTTAAACG	ACCCCTAGTAAGTTTGGG

❖ All sequences are written 5' → 3', bases in bold type within the SSV1 ORF
A. Highlighted base indicates mismatch generating Glu→Ala mutation in VP1 protein
B. Highlighted base indicates mismatch generating Glu→Gln mutation in VP1 protein
C. Lower case denotes introduced silent restriction endonuclease site

Complementation in cis of SSV1 mutants

SSV1 mutants lacking *vp1* or portions thereof were constructed via LIPCR using EAI283 as template (Table 1). SSV2 and SSV9 DNA were isolated from the original Sulfolobus strains using alkaline lysis [24,25]. The complete *vp1* genes from SSV1, SSV2 [26] and SSV9 and the N-terminal region of VP1 from SSV9 [25] were amplified using Phusion DNA polymerase, purified with PCR Purification Kit (Thermo-Fisher) and phosphorylated with T4 polynucleotide kinase according to manufacturer's protocols (Thermo-Fisher). DNA containing *vp1* or fragments thereof was heat treated for 10 min at 75º C prior to ligation reaction. LIPCR products were purified directly from the LIPCR reaction by sodium acetate/ethanol precipitation [21]. The amplified *vp1* gene was mixed with LIPCR product at a molar ratio of 10:1 respectively and ligated with 5 U of T4 DNA ligase (Thermo-Fisher) for 20 hours at 16ºC. 5 µL of the ligation reaction was used to transform 100 µL chemically competent Transformax EC100D pir⁺ *E. coli* (Epicentre).

Transformation of Sulfolobus

Sulfolobus was transformed essentially as described by Schleper and colleagues [7]. Cultures of *Sulfolobus* S441 (5 mL) were grown from frozen stocks for 48-72 hr in a 70º C shaking incubator. Starter cultures were transferred to 50-100 mL of fresh YS media in long neck Erlenmeyer flasks and grown until OD_{600nm} reached ~0.20. 50 mL of cells were removed and placed on ice for 30 minutes. Cells were washed three times with decreasing volumes of

20 mM ice-cold sucrose as described previously [7]. After the final wash, cells were resuspended in 400 µL ice-cold 20 mM sucrose and kept on ice. 100 µL of cells were added to a chilled 0.1 cm gap length electroporation cuvette (Bulldog Bio) and 2 µL of SSV or shuttle vector DNA (100-500 ng/µL) purified from *E.coli* using the GeneJET plasmid purification kit (Fisher) was added to the cells. Cells were transformed by electroporation (BioRad Gene Pulser II) under the following conditions: 1.5 kV, 400 Ω, 25 µF. Immediately following electroporation, cells were resuspended in 1 mL of 70º C YS, transferred to a 1.5 mL tube, and incubated for 1 hr in a 70º C dry incubator. Following incubation, cells were transferred to 50 mL of preheated YS in a long neck Erlenmeyer flask and grown with shaking at 70º C.

Halo assay to check for infectious virus production

Following electroporation of *Sulfolobus* with SSV DNA, halo assays were performed in duplicate 48 and 72 hours post-transformation. Halo assays were performed as previously described [19] by spotting 5 µL of transformed culture on an indicator lawn of *Sulfolobus* strain S441. After 48 hours of incubation at 75˚ C, a zone of growth inhibition or halo is observed around spots of strains containing infectious virus, since virus production slows growth of the indicator strain. For all halo assays both a positive control consisting of *Sulfolobus* transformed with an active SSV1 genome and a negative control, without transformation were included. Halo assays were only scored if, on the same plate, negative controls did not produce halos and positive controls did. A DNA

10

was determined to be non-infectious after at least 5 independent transformations that did not produce halos.

Confirmation of infectious SSV DNA

Transformed cultures that produced halos were further analyzed to confirm the identity and purity of the viral DNA. Viral DNA was purified from infected *Sulfolobus* cells using a GeneJET kit (Thermo-Fisher). Purified DNA was then amplified using DreamTaq DNA polymerase (Thermo-Fisher) using primers that flank the mutated region of the viral DNA. All PCR products were confirmed using DNA sequencing. Control reactions using the DNA used in transformation of *Sulfolobus* and wild type SSV1 DNA were always performed. *Sulfolobus* cultures harboring infectious SSV1 mutants were conserved at -80º C [20].

Transmission electron microscopy

For transmission electron microscopy, samples were prepared on 400-mesh carbon-Formvar-coated copper grids (Ted Pella). Grids were placed, carbon-Formvar down, on a 5 µl droplet of culture supernatant for 2 min. Culture supernatants were generated by centrifugation at 3000 g for 5 minutes. Samples were removed from the grid by wicking. Grids were then stained for 15-60 s on 5 µl of either 2% uranyl acetate stain (pH 3) or 2% sodium phosphotungstate tribasic hydrate stain (pH 6). Phosphotungstate stain was made freshly every week to ensure that the solution did not disassociate. Grids were allowed to dry in air overnight and were examined within 48 h of staining. Images were obtained at 8,500 to 34,000 magnification on an FEI Tecnai F20 transmission electron

microscope (TEM). Grids were analyzed by examining randomly selected grid squares. Images were obtained with a BM UltraScan camera and stored in digital micrograph 3 (Gatan) and TIFF formats.

Particle analysis.

The length and width of images of stained virus particles were measured in ImageJ [27]. Normal particle width, length, and aspect ratio were determined using the means of measurements of wild-type SSV1 particles (*n* = 240) [15]. Any particles whose width, length, or aspect ratio was more than two standard deviations from the means were classified as abnormal.

Results:

The N-terminus of the SSV1 major capsid protein VP1 *is essential for infectivity.*

The SSV1 *vp1* gene encodes the major capsid protein [11]. Thus, it is not surprising that a complete deletion of the *vp1* gene and a transposon insertion in the middle of the ORF both failed to yield infectious virus [15]. The VP1 protein appears to be proteolytically cleaved at an internal glutamate residue that is conserved in all known fuselloviruses [15] (Figure 1) to produce the mature protein [5,12]. To investigate if the encoded N-terminus is required for infectivity, this region was deleted using LIPCR while leaving the conserved glutamate intact and maintaining a start (ATG) codon (Figure 1). The construct was electroporated into *Sulfolobus* and infectivity tested by halo assay (see Methods). Infectious virus was repeatedly not produced by this *vp1*-ΔN-terminus mutant

12

(EAI564), suggesting that the VP1 N-terminus is required for production of infectious virus. The N-termini of fusellovirus VP1 proteins are much less well-conserved than the C-termini, with the exception of a group of well-conserved residues just N-terminal to the conserved glutamate (Figure 1). This region was deleted from the SSV1 *vp1* gene and the resulting mutant (EAI578) was also found to be non-infectious. Thus the N-terminus of VP1 and 5 amino acids immediately N-terminal to the cleavage site appear to be essential for infectivity.
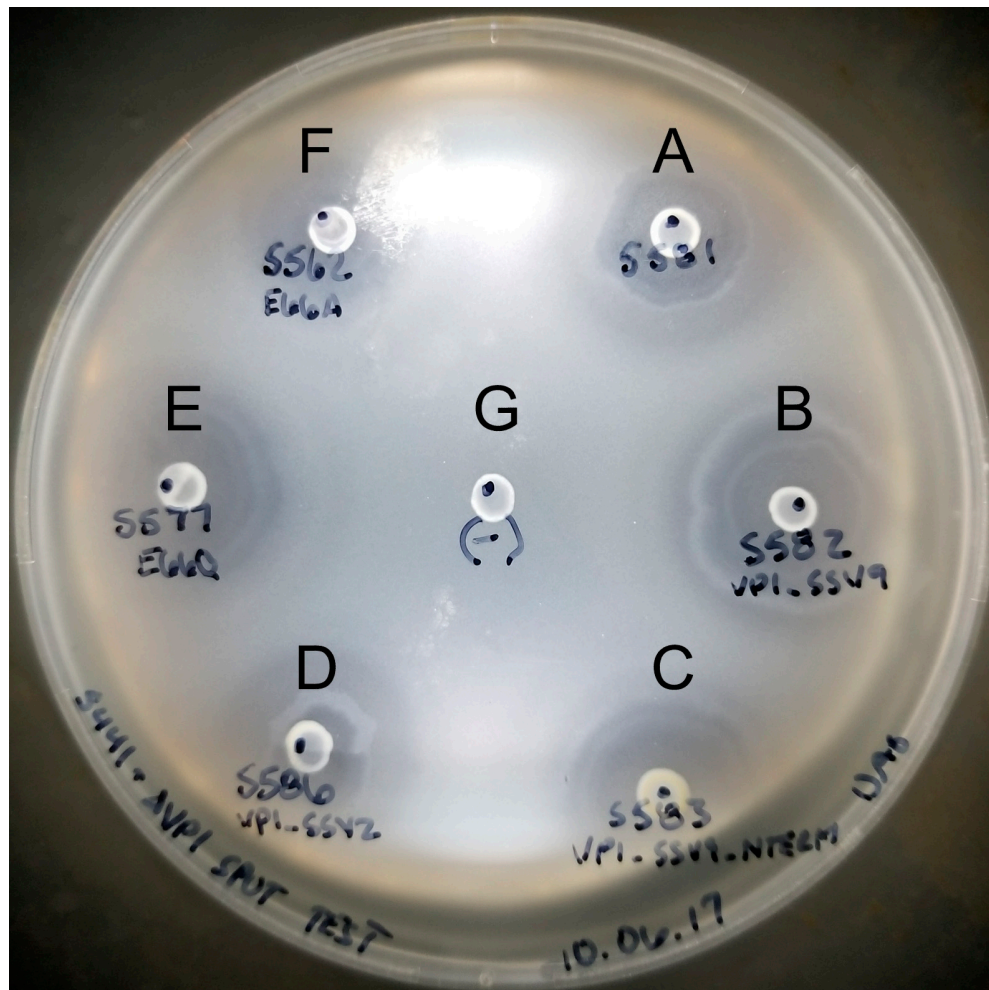
Figure 2: Spot on lawn assays for functional SSV1 *vp1* mutants. *Sulfolobus* cultures infected with SSV1 containing various *vp1* mutations were spotted on a lawn of uninfected *Sulfolobus*. The lawn and infected cultures were allowed to grow for 48 hours and then photographed. Letters indicate indicates a spotted culture of *Sulfolobus* infected with a SSV1 that contains the wild type *vp1* gene (A), the *vp1* gene from SSV9 (B), a fusion between the N-terminus of VP1 from SSV9 and the C-terminus of VP1 from SSV1 (C) the *vp1* gene from SSV2 (D), a glutamine at position 66 in the SSV1 VP1 protein instead of the wild type glutamate (E), or an alanine at position 66 in the SSV1 VP1 protein instead of the wild type glutamate (F). G indicates an uninfected control culture spotted on the lawn.

Complementation in cis of SSV1 deletion mutants

In order to confirm that SSV1 *vp1* deletions were non-functional only due to the deletion, we complemented the mutants in cis with the SSV1 wild-type *vp1* gene and with homologs from other SSVs. Infectivity of the SSV1 Δ*vp1* mutant could be rescued when complemented in cis with the SSV1 wildtype *vp1* gene, or homologous *vp1* genes from either SSV2 or SSV9 (Figure 2 A-D). The infectivity of the non-functional *vp1*-ΔN-terminus mutant (EAI564) was also rescuable by complementation in cis when the VP1-N-terminus from SSV9 was added to the C-terminus of SSV1 VP1 (Figure 2C). The C-termini of the three VP1 proteins are highly similar, however, the N-terminus of SSV2-VP1 is considerably shorter (15 residues) than the other two (SSV9-VP1 65 residues, and SSV1-VP1 66 residues respectively) (Figure 1). Virions with these alternative *vp1* genes were normal on transmission electron microscopic analysis.
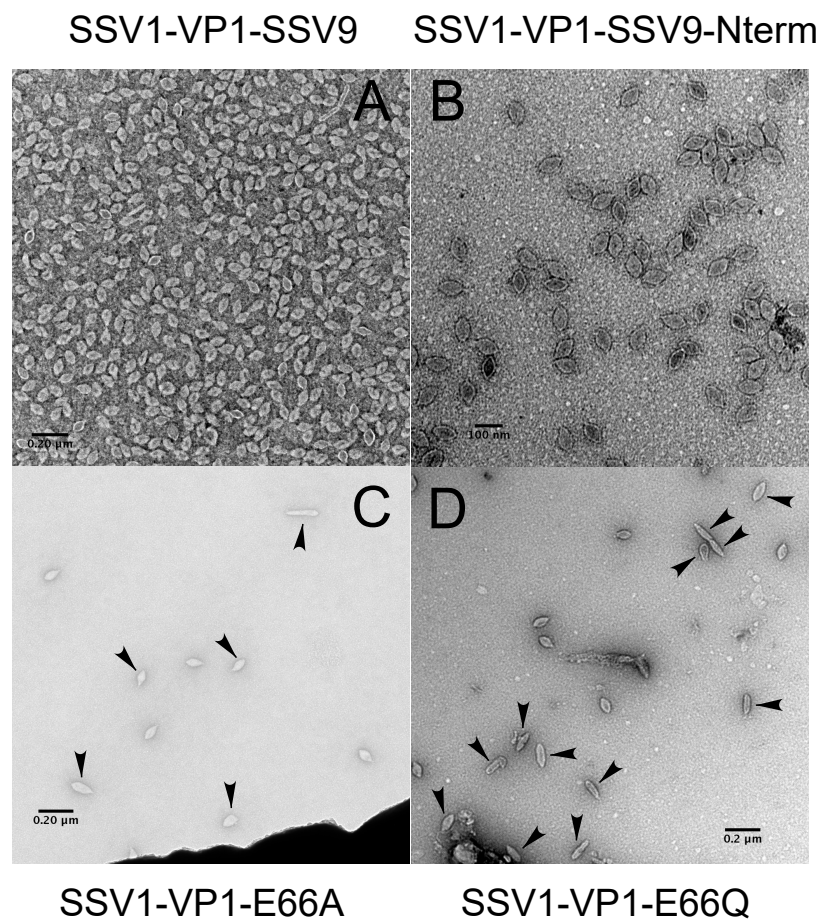
SSV1-VP1-SSV9   SSV1-VP1-SSV9-Nterm

SSV1-VP1-E66A   SSV1-VP1-E66Q

Figure 3: Transmission electron micrographs of SSV1 vp1 mutants.
Negatively stained SSV1 mutant virions were imaged using a Tecnai F-20
(FEI) transmission electron microscope operating at 200keV.   The culture
supernatant in panel A was concentrated ca. 100 fold with a Vivapsin-500
(10kDa cutoff) concentrator. Panel A: Virions from SSV-1 containing the VP1
gene from SSV9. Panel B: Virions from SSV-1 containing a VP1 gene with the
N-terminus of the gene from SSV9. Panel C: Virions from SSV-1 containing a
VP1 gene with the E66A substitution. Panel D: Virions from SSV-1 containing
a VP1 gene with the E66Q substitution. Arrowheads in Panels C and D
indicate "abnormal" particles.

Changing the conserved glutamate in the major capsid gene vp1 *allows*

infectious virus but generates many abnormal virions.

Although the presence of a N-terminus of VP1 is essential, the necessity for proteolysis at glutamate 66 for the activity of SSV1 remained unclear. Thus we investigated whether the universally conserved glutamate of VP1 itself is essential for infectivity. A single base pair change was made in the *vp1* gene by LIPCR, converting the glutamate to the structurally similar but functionally distinct glutamine (EAI427; Table 1). Interestingly, viruses containing this mutation were infectious as assessed by halo assay (Figure 2). Even when the glutamate was converted to alanine (EAI500) the virus was still infectious (Figure 2). However, when analyzed by Transmission Electron Microscopy, virions containing these point mutations had a very high proportion of abnormally shaped virions (Figure 3). Those viruses that had the E66Q substitution had 57% (n=198) abnormal morphology and those with the E66A 81% (n=99). Wild type SSV1 preparations have 5.4% abnormal particles (n=240) [15]. Thus glutamate residue 66 is not absolutely essential for SSV1 infectivity despite its universal conservation in the *Fuselloviridae*.

Discussion:

The translational start of the VP1 protein is not clear:

Here we show that the N-terminus of the SSV1 VP1 protein is required. Since the VP1 protein in intact SSV1 virions has a glutamate at its N-terminus, the identity of the actual start codon that is used to translate the SSV1 VP1 precursor protein is not clear [5]. Based on sequence comparison (Figure 1), the presumed start codon is a TTG. TTG and GTG codons are known to be used as start codons for

17

archaeal translation [28] but there is an in-frame ATG codon 18 nucleotides upstream of this TTG that is 9 nucleotides downstream of the transcriptional start site , a normal distance for translational initiation [29]. However there is no clear *Sulfolobus* ribosome binding site at any of these potential translational start sites [30].   Thus the SSV1 *vp1* gene may be misannotated. However, there is a conserved open reading frame, SSV1 ORF b115, that overlaps the first 12 codons of the predicted SSV1-*vp1* gene [5]. Interestingly there is a similar 14 codon overlap in the SSV2 genome even though the predicted N-terminus of the SSV2 *vp1* gene is much shorter than that of SSV1 [26]. Whether these overlapping ORFs are translated, however, is not known.   This confusion regarding the SSV1-VP1 precursor protein makes analysis of the protease and proteolysis challenging, as it is not clear what the appropriate substrate is. A replacement of the N-terminus of SSV1 with the N-terminus from SSV2 could be a useful tool to study this proteolytic process and the N-terminus of the preprotein would be better defined.

The N-terminus of SSV1 VP1 is essential, but the cleavage site is not:

When the entire N-terminus of SSV1 VP1 is deleted, or 5 amino acids N-terminus to the cleavage site are deleted, no viable virus was produced (Figure 1). This indicates that cleavage of the VP1-precursor is necessary and depends on the 5 amino acids next to the cleavage site. However, when a point mutation was introduced changing the conserved cleavage site glutamate (E) to either a glutamine (Q) or alanine (A) (E66Q and E66A), infectious virus particles were

18

produced, indicating that some cleavage is taking place. It is unknown if VP1 proteolysis can still occur at this position in either of the mutants or if proteolysis occurs at one of the other well-conserved glutamate residues upstream (Figure 1). Experiments to address the site of proteolysis are underway, but yields of SSV virus preparations are notoriously low [7,11,12]. Comparative stability experiments with these active mutants are also being performed.

Virions containing point mutations in vp1 have unusual morphology.

We previously showed that deletions of the gene for the minor capsid protein, *vp3* in SSV1 led to the production of longer and thinner virions. Fully 99% of these mutant viruses were abnormal, as defined by deviating from means of length, width and aspect ratio of the wild type virus by more than 2 standard deviations [15]. Unlike VP3 deletions, virions with a mutated VP1 cleavage site had some normal, and some abnormal morphologies. Often the abnormal particles had differently shaped ends, some pear-shaped. Interestingly there were fewer abnormal (57%) virions containing the E66Q mutation than those containing the E66A mutation (81%). This indicates that there may be some proteolysis at amino acid 66 and more proteolysis for the structurally similar amino acid Q than for the structurally dissimilar amino acid A.

The SSV1 VP1 N-terminus can be functionally replaced with the SSV9 VP1 N-terminus:

Even though cleavage of the N-terminus appears to be essential for al SSVs, the lack of conservation of the N-termini led us to suspect that the amino acid sequence of the N-terminus is not critical. To confirm this, we switched the N-terminus of SSV1's VP1 with the N-terminus of SSV9 VP1 [25] (Figure 1). Whether this can be done with the predicted N-terminus of the SSV2 VP1 protein that is much shorter than that of SSV1, remains to be determined.

Proteolysis and assembly of SSV1

Proteolytic processing during virion assembly is common and may be the rule rather than the exception [31]. We have demonstrated that the N-terminus of VP1 SSV1 is essential for infectivity. The N-termini of fusellovirus *vp1* genes are not well conserved with the exception of the essential patch of residues just upstream of the conserved glutamate (Figure 1). Complementation of ∆*vp1* mutants with SSV2-*vp1*, which possesses a truncated N-terminus, suggests that the majority of the VP1 N-terminus can be dispensed with as long as the residues upstream of the conserved glutamate remain intact. The role of these essential residues remain mysterious. These residues might be critical for recruitment of a protease and/or may act as a scaffold, potentially via interactions with neighboring VP1 N-termini, the minor capsid protein VP3, or some other protein involved in assembly. The identity of proteases or scaffolding proteins for SSV1 is not known. The protease could be viral, cellular or even the VP1 protein itself. There are no ORFs in the SSV1 genome that have clear similarity to proteases, but there are a few known proteases in the *Sulfolobus* genome

[32,33]. However none of these have specificity for glutamate. Assembly is likely highly conserved among the *Fuselloviridae* and any protein scaffolding candidate would likely come from the set of core fusellovirus genes (see Iverson et al., 2017). The hypothesis that the N-terminus of VP1 may fulfill this role is intriguing one and awaits further research.

Conclusions:

We show that the N-terminus of the SSV1 major capsid protein VP1 is essential for virus function, probably for assembly and proteolysis. Proteolysis itself is appears to be essential to allow the capsid proteins to form the spindle-shaped morphology, similar to many other viruses. The protease that cleaves the N-terminus is unknown but may have some activity on non-conserved amino acids. The presence of a point mutation in the cleavage site leads to the production of aberrant virus particles, probably containing N-terminal extensions on VP1. The N-terminus of VP1, in addition to providing a site for proteolysis may also provide a scaffolding role. Future research will concentrate on the identification of the novel protease and role of individual amino acids in the N-terminus of SSV1 VP1 in virus assembly.

Author Contributions:

E.A.I. and K.M.S. conceived and designed the experiments; E.A.I., D.A.G., and M.E.G. performed the experiments; all authors analyzed the data; E.A.I., M.E.G. and K.M.S. wrote the paper.

Conflicts of Interest:

The authors declare no conflicts of interest.

References:

1. Krupovic, M.; Quemin, E. R. J.; Bamford, D. H.; Forterre, P.; Prangishvili, D. Unification of the globally distributed spindle-shaped viruses of the Archaea. *J. Virol.* **2014**, *88*, 2354–2358, doi:10.1128/jvi.02941-13.

2. Prangishvili, D. The wonderful world of archaeal viruses. *Annu. Rev. Microbiol. Vol 67* **2013**, *67*, 565–585, doi:10.1146/annurev-micro-092412-155633.

3. Martin, A.; Yeats, S.; Janekovic, D.; Reiter, W. D.; Aicher, W.; Zillig, W. SAV-1, A Temperate UV-Inducible DNA Virus-like Particle from the Archaebacterium Sulfolobus-acidocaldarius Isolate B-12. *EMBO J.* **1984**, *3*, 2165–2168.

4. Yeats, S.; McWilliam, P.; Zillig, W. A plasmid in the archaebacterium *Sulfolobus acidocaldarius*. *EMBO J.* **1982**, *1*, 1035–1038.

5. Palm, P.; Schleper, C.; Grampp, B.; Yeats, S.; McWilliam, P.; Reiter, W. D.; Zillig, W. Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. *Virology* **1991**, *185*, 242–250.

6. Nadal, M.; Mirambeau, G.; Forterre, P.; Reiter, W. D.; Duguet, M. Positively supercoiled DNA in a virus-like particle of an archaebacterium. *Nature* **1986**, *321*, 256–258.

7. Schleper, C.; Kubo, K.; Zillig, W. 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. U. S. A.* **1992**, *89*, 7645–7649.

8. Ceballos, R. M.; Marceau, C. D.; Marceau, J. O.; Morris, S.; Clore, A. J.; Stedman, K. M. Differential virus host-ranges of the *Fuselloviridae* of hyperthermophilic Archaea: implications for evolution in extreme environments. *Front. Microbiol.* **2012**, *3*, doi:10.3389/fmicb.2012.00295.

9. Muskhelishvili, G.; Palm, P.; Zillig, W. SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. *Mol. Gen. Genet.* **1993**, *237*, 334–342.

10. Quemin, E. R. J.; Chlanda, P.; Sachse, M.; Forterre, P.; Prangishvili, D.; Krupovic, M. Eukaryotic-like virus budding in Archaea. *mBio* **2016**, *7*, doi:10.1128/mBio.01439-16.

11. Reiter, W. D.; Palm, P.; Henschen, A.; Lottspeich, F.; Zillig, W.; Grampp, B. Identification and characterization of the genes encoding 3-structural proteins of the *Sulfolobus* virus-like particle SSV1. *Mol. Gen. Genet.* **1987**, *206*, 144–153.

12. Quemin, E. R. J.; Pietila, M. K.; Oksanen, H. M.; Forterre, P.; Rijpstra, W. I. C.; Schouten, S.; Bamford, D. H.; Prangishvili, D.; Krupovic, M. Sulfolobus spindle-shaped Virus 1 contains glycosylated capsid proteins, a cellular chromatin protein, and host-derived lipids. *J. Virol.* **2015**, *89*, 11681–11691, doi:10.1128/jvi.02270-15.

13. Stedman, K. M.; DeYoung, M.; Saha, M.; Sherman, M. B.; Morais, M. C. Structural insights into the architecture of the hyperthermophilic *Fusellovirus* SSV1. *Virology* **2015**, *474*, 105–109, doi:10.1016/j.virol.2014.10.014.

14. Zhao, G.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C.; Zhang, P. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **2013**, *497*, 643–646, doi:10.1038/nature12162.

15. Iverson, E. A.; Goodman, D. A.; Gorchels, M. E.; Stedman, K. M. Extreme Mutation Tolerance: Nearly Half of the Archaeal Fusellovirus Sulfolobus Spindle-

Shaped Virus 1 Genes Are Not Required for Virus Function, Including the Minor Capsid Protein Gene *vp3*. *J. Virol.* **2017**, *91*, e02406-16, doi:10.1128/JVI.02406-16.

16. Yost, S. A.; Marcotrigiano, J. Viral precursor polyproteins: keys of regulation from replication to maturation. *Curr. Opin. Virol.* **2013**, *3*, 137–142, doi:10.1016/j.coviro.2013.03.009.

17. Bertram, S.; Glowacka, I.; Steffen, I.; Kühl, A.; Pöhlmann, S. Novel insights into proteolytic cleavage of influenza virus hemagglutinin. *Rev. Med. Virol.* **2010**, *20*, 298–310, doi:10.1002/rmv.657.

18. Clore, A. J.; Stedman, K. M. The SSV1 viral integrase is not essential. *Virology* **2007**, *361*, 103–111.

19. Iverson, E.; Stedman, K. A genetic study of SSV1 the prototypical fusellovirus. *Front. Microbiol.* **2012**, *3*, doi:10.3389/fmicb.2012.00200.

20. Zillig, W.; Kletzin, A.; Schleper, C.; Holz, I.; Janekovic, D.; Hain, J.; Lanzendorfer, M.; Kristjansson, J. K. Screening for *Sulfolobales*, their Plasmids and their Viruses in Icelandic Solfataras. *Syst. Appl. Microbiol.* **1994**, *16*, 609–628.

21. Green, M. R.; Sambrook, J. *Molecular cloning: a laboratory manual*; 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y, 2012; ISBN 978-1-936113-41-5.

22. Bertani, G. Lysogeny at Mid-Twentieth Century: P1, P2, and Other Experimental Systems. *J. Bacteriol.* **2004**, *186*, 595–600, doi:10.1128/JB.186.3.595-600.2004.

23. Birnboim, H. C.; Doly, J. Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **1979**, *7*, 1513–1523.

24. Arnold, H. P.; She, Q.; Phan, H.; Stedman, K.; Prangishvili, D.; Holz, I.; Kristjansson, J. K.; Garrett, R.; Zillig, W. The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. *Mol. Microbiol.* **1999**, *34*, 217–226.

25. Wiedenheft, B.; Stedman, K.; Roberto, F.; Willits, D.; Gleske, A. K.; Zoeller, L.; Snyder, J.; Douglas, T.; Young, M. Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. *J. Virol.* **2004**, *78*, 1954–1961.

26. Stedman, K. M.; She, Q. X.; Phan, H.; Arnold, H. P.; Holz, I.; Garrett, R. A.; Zillig, W. Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. *Res. Microbiol.* **2003**, *154*, 295–302.

27. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675, doi:10.1038/nmeth.2089.

28. Belinky, F.; Rogozin, I. B.; Koonin, E. V. Selection on start codons in prokaryotes and potential compensatory nucleotide substitutions. *Sci. Rep.* **2017**, *7*, 12422, doi:10.1038/s41598-017-12619-6.

29. Condò, I.; Ciammaruconi, A.; Benelli, D.; Ruggero, D.; Londei, P. Cis-acting signals controlling translational initiation in the thermophilic archaeon Sulfolobus solfataricus. *Mol. Microbiol.* **1999**, *34*, 377–384.

30. Torarinsson, E.; Klenk, H.-P.; Garrett, R. A. Divergent transcriptional and translational signals in Archaea. *Environ. Microbiol.* **2005**, *7*, 47–54, doi:10.1111/j.1462-2920.2004.00674.x.

31. Dougherty, W. G.; Semler, B. L. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol. Rev.* **1993**, *57*, 781–822.

32. Guagliardi, A.; Cerchia, L.; Rossi, M. An intracellular protease of the crenarchaeon Sulfolobus solfataricus, which has sequence similarity to eukaryotic peptidases of the CD clan. *Biochem. J.* **2002**, *368*, 357–363, doi:10.1042/BJ20021017.
33. Lin, X.; Tang, J. Purification, characterization, and gene cloning of thermopsin, a thermostable acid protease from Sulfolobus acidocaldarius. *J. Biol. Chem.* **1990**, *265*, 1490–1495.