2016

Spindle Shaped Virus: Mutants and Their Infectivity

Thien Hoang
Portland State University
Spindle Shaped Virus (SSV): Mutants and Their Infectivity

by

Thien Hoang

An undergraduate honors thesis submitted in partial fulfillment of the
requirements for the degree of

Bachelor of Science

in

University Honors

and
Biology: Micro/molecular biology

Thesis Adviser

Dr. Kenneth Stedman

Portland State University

2014
SSV1 is an archaeal virus that infects the thermoacidophile *Sulfolobus* residing in hot springs. The lemon shaped/spindle-shaped fuselloviruses (SSV) that infect *Sulfolobus solfataricus* is quite morphologically different from almost all other viruses. Because these archaeal viruses live in hot springs with high temperatures and low pH, their genomes and structures have adapted to withstand such harsh conditions. Little research has been done on these extreme viruses, and of the little research, SSV has been the most prominent. Not much is known about the genes that the genome encodes and so I have inserted transposons randomly into genome to determine functionality. The only highly conserved genes whose functions we know for sure are the VP1-3, encoding for the capsids and DNA packaging. I tried to isolate transposons that have inserted into a very small region of the genome, approx. 800bp, but have not had any luck. So far the transposon only inserted into ORFs B-49, F-93, A-132, A-100, and between F-112 and B-49.
Introduction:

Current hypotheses contend that viruses have catalyzed several major evolutionary transitions, including the invention of DNA and DNA replication mechanisms, the origin of the eukaryotic nucleus, and thus a role in the formation of the three domains of life (Forterre, Filée and Myllykallio 2000). Viruses span every corner of the globe and live in all sorts of conditions; they are nonliving simple bags of DNA that use their host to replicate their DNA. They serve as model systems for biology due to their availability and ease of use. There are viruses that infect all three domains of life: Bacteria, Eukarya, and more recently Archaea. Archaeal viruses are still not as well understood as those that infect Bacteria or eukaryotes, especially those that are thermoacidophiles (optimal temperature >70°C, pH<4) that live in extreme environments like hot springs (Iverson and Stedman, 2012). Much of our present knowledge of the molecular genetics of bacteria is based on work with bacteriophage host-systems (bacterial viruses). In Archaea only a few virus host systems have been sufficiently investigated to serve this purpose.

The spindle shaped virus (SSV1) of the family Fuselloviridae got its name from the lemon-like shape of its capsid and infects their host Crenarchaeon genus *Sulfolobus* and close relatives. (Peng et al., 2012) It was isolated from S. shibatae strain B-12 that was isolated from a hot spring in Beppu, Japan (Yeats et al 1982, Martin et al 1984). The SSV1 virus particle is approximately 60x90 nm with a 10nm tail protruding from one end. (Stedman et al., 2015) All transcripts and their respective promoters and terminators have been identified (Reiter et al 1987). The viral nature of this particle remained obscure until 1992 (Schleper et al. 1992). The family Fuselloviridae is unique and confined to the *Sulfolobus* as hosts. The family of viruses appear to be lysogenic rather than lytic. After injection of the viral DNA into the host, the
genome integrates into a specific tRNA gene of the host. From this location, transcription and replication of the viral genome occurs, which in turn leads to the assembly of new viruses. These are released from the host by budding (Prangishvili, 2013). Viruses with similar spindle shapes are commonly seen in hypersaline waters. However, the particles seen in hypersaline waters are smaller than SSVs and seem to lack the small tail present on Fusellovirus capsids. (Prangishvili, 2013)

The double stranded DNA of these viruses code for 34 open reading frames (genes) and is an approximately 15.5 kb long covalently closed circular, double stranded DNA which is stably carried within the host cells integrated in the chromosome. (Lawrence et al. 2009). The capability of SSV1 to infect hosts was long overlooked, because the virus does not produce significant numbers of particles upon infection (Schleper et al. 1992). The intricate relationship between SSV1 and Sulfolobus is the most studied of the archaeal viruses because it served as an early model for transcription and regulation in Archaea (Reiter et al. 1990) Moreover, SSV1 is the only known crenarchaeal virus that exhibits an integrase similar to that of the bacteriophage lambda and shows an UV-inducible life cycle (Reiter et al., 1987). Researchers still have no idea what the majority of genes in SSV1 genome code for. The main purpose of my research is to determine what these open reading frames (ORFs) code for. Only four genes (VP1, VP3, VP2, and the viral integrase) do scientists definitively know the function of the gene: the major and minor capsid protein, DNA packaging protein, and integrase respectively (Iverson and Stedman; 2012).
I was always interested in viruses because of their intricate connection with medicine and disease. Viruses are so small yet they can wreak havoc on organism of all sizes. We have always tried to conquer viruses but they always evolve and elude even the best of medicine. Viruses clearly give humans trouble and studying viruses may lead to an increase in quality of life. Since they are so good at infecting their hosts, we have been trying to use that for our benefit and using viruses as vectors. The detailed knowledge of its genome makes SSV1 suitable not only for studies of virus-host interaction but also for the development of a vector system for extreme thermophiles. Viruses are also important in medical research for finding cures for common diseases. Also, viruses are source of vaccines that prevent common infections. The SSV1 virus doesn’t have many uses as a human vectors, but it might be used for other benefits such as gene therapy or imaging. Studying it will give us an idea of how viruses infect Archaea, which in turn leads to better understanding of viruses of Bacteria/eukaryotes. To study complex things, you must break it down into the simpler models to understand first, and that is exactly what the SSV1 virus is: a model organism for Archaea, which in turn are models for Eukarya.

Scientists initially thought Archaea existed only in extreme environments, niches devoid of oxygen, or whose temperatures can be near or above the normal boiling point of water. (Nair 2011) Biologists later realized that Archaea are a large and diverse group of organisms that are widely distributed in nature and are common in much less extreme habitats, such as soils and oceans.

Viruses contribute to the greater biosphere via their abundance and the role they play in evolution. For example the oceans contain approximately $10^{31}$ viruses that catalyze turnover of 20% of the ocean biomass per day. (Pernthaler 2005). As for SSV1 that thrives in the extreme
conditions, they play an important role acting as a genetic reservoir to maintain diversity within
the ecosystem (Pernthaler 2005).

Another reason is that they aid in evolution. Most cellular genomes have a significant
amount of viral or virus-like sequence within their genome, further evidence that viruses play a
central role in horizontal gene transfer and help drive the evolution of their hosts (Krupovic et
al. 2011). As viruses evolve, their hosts evolve to defend against their evolving viruses creating
an evolutionary arms race.

What makes SSV1 interesting is that the genes aren’t similar to other viruses outside of
the Fuselloviridae family, and only has one gene (ORF d355 which encodes for a viral integrase)
that is common with other viruses (Palm et al. 1991, Iverson and Stedman, 2012). This makes it
hard to see the evolutionary changes that have happened to SSV1, since there are not very
many others with homologous genes. A different approach has been tried by the Stedman lab
and other labs in the past including: structural genomics, comparative genomics, and genetics
(see below).

Literature Review:

Studies have been conducted for the thermoacidophiles archaeal virus family that SSV1
falls under. Along with SSV1, other strains have been studied such as SSV2, SSVK, and other SSV
viruses. Viruses of prokaryotes and eukaryotes on the other hand have much more information
about them. Five fuselloviruses have so far been isolated from acidic geothermal environments
in different locations in Asia, Europe and North America, and they replicate in species of the
hyperthermophilic archaeal genus *Sulfolobus*, which represents a significant percentage of the
microbial population in most acidic terrestrial hot springs (Redder et al, 2009). Studying these viruses led to the discovery of archaeal promoters and has provided tools for the development of the molecular genetics of these organisms. (Reiter et al, 1990; Stedman et al 1999) As for SSV1, it has been a fairly recent study about its origins. The small number of researchers has only been able to definitively identify 4/31 of the genes and a lot still remain unknown.

In the 20th century, prokaryotes were thought as a single group of organisms. Carl Woese is credited for the discovery of the domain Archaea and reshaping the tree of life. He speculated about an era of rapid evolution in which considerable horizontal gene transfer occurred between organisms. Carl Woese and George E. Fox experimentally disproved the current hypothesis of only prokaryotes and eukaryotes. They compared the ribosomal RNA (rRNA) of different Prokaryotes to determine their evolutionary relationships. He found Prokaryotes in distant groups turned out to be close relatives - and sometimes Prokaryotes in a single group turned out to be very distant indeed. But the biggest surprise was that one group of ancient organisms were not bacteria at all. These organisms, which often lived under extreme conditions such as high heat, looked just like bacteria. But their rRNA said otherwise. The more closely related two bacteria are, the more alike their rRNA should be. They named the new microbes archaeobacteria. His three-domain system, based on phylogenetic relationships rather than obvious morphological similarities, divided life into 23 main divisions, incorporated within three domains: Bacteria, Archaea, and Eucarya (Woese et al., 1990).

Before it was thought that Archaea would only be found in extreme environments, but now since its discovery, Archaea has been found in plankton as well as in the human body.
Wolfram Zillig’s first works were to investigate archaeal transcription initiation, and it was clear that there were homologs to the proteins needed for DNA replication such as: RNA polymerase II, TATA binding protein, and TFIIB. This supported his hypothesis that archaeal transcription was a simpler version of the Eukaryotic system and that the eukaryotic system could have evolved from Archaea. Not only did it have some proteins pertaining to eukaryotes, Archaea also had some similarities with prokaryotes and some completely distinct mechanisms found only in archaeal viruses. There are some transcriptional regulators that have no identifiable homologues between either of the other two domains.

Carl Woese and George Fox used a method to compared the sequences of a particular molecule central to cellular function, called ribosomal RNA, which has become the standard approach used to identify and classify all organisms (Bult et al. 1996). The technique also advanced the field of ecology because now scientists could survey an ecosystem by just collecting ribosomal DNA. Before they would have to culture and grow the organisms that were there, and usually was a very difficult task to do. His work is also significant in its implications for the search for life on other planets. They may have robust evolutionary connections to the first organisms on Earth. Organisms similar to those Archaea that exist in extreme environments may have developed on other planets, some of which are very similar to those of the extremophiles on Earth. Dr.Woese recently passed away in 2012, but he left behind a legacy that would change the field of microbiology forever. He turned a field that was primarily subjective into an experimental science with wide-ranging and practical implications for microbiology, ecology and even medicine that are still being worked out.
His work inspired many scientists in various fields of biology, and among them was Wolfram Zillig, who is credited with the discovery of several unique molecular features of Archaea. The new method of molecular phylogeny used by Woese was not yet widely accepted, and his suggestion of another division in the tree of life was highly controversial. Zillig went on to discover evidence that would strongly support Woese’s claims and help silence the critics. Wolfram Zillig, a pioneer who identified the first archaeal viruses that looked like a head-tail bacteriophage, but turned out to be a new viral family. He was an expert at the time on RNA polymerases (RNAP) and so Woese had convinced Zillig to investigate the “Archaea”.

Wolfram Zillig was an expert molecular biologist who studied transcription in *E. coli*. His work geared researchers towards a new area of the tree of life to discover and in doing so led to the discovery of seven new families of crenarchaeal viruses: the *Globuloviridae*, *Bicaudaviridae*, *Ampullaviridae*, (all 3 discovered by Prangishvili), *Guttaviridae*, *Fuselloviridae*, *Rudiviridae*, and *Lipothrixviridae* (Prangishvili 2013). The focus on the *Fuselloviridae* family for the SSV1 is the most studied of the families because of its simplicity, prevalence, and is used as the model organism to study archaeal viruses and hosts.

Whereas the early work concentrated on isolating virus-host systems such that the virus could be cultured and investigated in the laboratory done by Wolfram Zillig, later studies pioneered by David Prangishvili and colleagues focused on enriching crenarchaeal viruses, again from terrestrial hot springs, and then testing cultured crenarchaeal strains for compatible hosts of the isolated viruses (Prangishvili 2013). They researched the morphology of viruses in hot springs and found that spindles, filaments, rods and spheres predominate, while other
morphological shapes weren’t as common. When they found the SSV1 virus, was the first of a series of unusual and uniquely shaped viruses isolated from hyperthermophilic Archaea. They use Northern analysis to analyze transcripts from SSV1, and later use UV-radiation which induces viral replication (Reiter et al., 1987).

Scientists all agree that the genes VP1, VP2, and VP3 code for the capsid and packaging genes, but other than that not a whole lot more is known about the other 28 genes other than the integrase gene. Recent work by the Stedman lab did a comparative study of several conserved ORFs and found that deletion of ORF b129 may allow the identification of the second archaeal virus transcriptional activator. Deletion of ORF d244 may allow insight into copy number regulation in SSVs, previously thought to be regulated by ORF d63 (Iverson and Stedman, 2012). Prior to the Stedman lab, C. Martin Lawrence and his group of researchers obtained atomic resolution structures of proteins encoded by SSV1 ORFs b129, f112, d63, e96, f93, and d244 (Lawrence et al., 2009). They haven’t been able to 100% verify the functions of the genes. SSV1 DNA opens new avenues for studying virus replication and the functions of viral genes, as well as for the development of a transformation system for extreme thermophiles. (Prangishvili, 2013)

My work was a continuation of the work done by the Stedman lab of the past years. As the technology advances so will the methods of discovering the virus. The literature is trending towards using more combinations of biochemical, genetics, comparative and structural genomics to get more information of the genes. Adam Clore, a previous PSU doctorate had published a thesis about SSV in 2008, that focused on gene knockout using LIPCR (Clore, 2008).
The study shows that removal of the integrase gene from the SSV1 virus does not stop the virus from replicating and infecting new cells. My task, more like a combination of Stedman et al. 1999 and Iverson and Stedman, 2012, is quite similar to Clore’s thesis. I am to focus on a small region (800bp) where not many deletions are documented.

Methods:

*Kanamycin Selection:*

To begin a transposon (2kb) with a Kanamycin resistant gene was randomly inserted into the genome of SSV1 (Epicenter EZ-Tn5), and then transformed into *E.coli*. *E.coli* is used as the host because of its simplicity of growth/transformation and availability. To determine if the transposon has been inserted into the virus, the transformed cells are spread onto a petri plate with Kanamycin. The Kanamycin is a bacterial antibiotic that will kill any cells without the resistance genes, so only the cells containing the transposon successfully integrated will appear on the plate.

*Inoculation of Cultures:*

The next step is to determine where the transposon inserted into the genome since it is a random event. To inoculate cultures to grow the DNA, the cells are taken from the plate and injected into a test tube with LB broth with 50µg/ml kanamycin and grown for 18 hours.

*Alkaline lysis:*
The next day the cells are prepped for an alkaline lysis (Plasmid prep procedure, Stedman 2013) which breaks the cells open and quickly renatures small RNA and plasmids while larger molecules precipitate. Three solutions are added in succession to each test tube to lyse the cells: the first to suspend the cells, the second to raise the pH and lyse cells, and the third is acidic to rapidly neutralize the solution. After three ethanol washes and centrifugations, a restriction endonuclease (EcoRI) is added to digest the DNA and prep for gel electrophoresis.

Gel electrophoresis separation:

Gel electrophoresis is then performed to separate the digested DNA by size and determine shifts in the banding pattern. An agarose gel is created to separate the DNA fragments in different wells. DNA migrates through the gel in an electric field due to their negative charge; the shorter bands move furthest while the larger bands move slowly through the gel due to smaller fragments easily navigating the pores of the Agarose gel. Looking at Figure 1, on the very left is the ladder in which all the other bands can be compared to. A shifted band has the transposon inserted into it, thus becomes larger and moves differently than the predicted bands.

Gel staining:
The gel needs to be stained with Ethidium bromide for use with UV irradiation to observe the DNA by fluorescence. This results in a visual representation of where the transposon has inserted. Since we know where EcoR1 cuts the viral genome, we can associate the DNA bands with a particular region of the genome. Bands that have shifted their mobility indicate that the transposon has integrated somewhere within that region. Selecting specific regions to look at, another round of restriction digest is performed using EcoR5 as the restriction endonuclease to cut the genome in different spots. This is to determine where the transposon inserted in a more defined region by comparing the band shifts. So far we have been trying to find transposons in a small section of the genome (~800bp) but many transposons have not been inserting in that region. For some of the banding patterns that don’t quite fit with expected results (outliers), I had to regrow the transformed *E. coli* again from frozen stock that was kept from the beginning.

**E. coli Transformation:**

A transformation had to be done, which is when the cell takes up DNA from its surrounding environment. There are two ways that I did this: a heat shock transformation and electroporation. The heat shock was a simple technique that used a quick burst of 42°C as to not denature any proteins, but simply allow the cells to quickly take in surrounding DNA. A sudden increase in temperature creates pores in the membrane of the washed bacteria and allows for plasmid DNA to enter the bacterial cell. This was done at 42°C for 30 seconds then quickly put on ice for the cells to recover (transformation procedure, Stedman 2013).

**Electroporation of Sulfolobus:**
For the electroporation, it uses a rapid jolt of electricity to shock the cell and change the permeability of the membrane. In that short time, foreign substances can enter the cell, thus transforming it. Preparing the cells involved chilling the cells in ice, then centrifuging them in 50, 25, 10, 1 ml amounts of 20mM sucrose solution to get a pellet. Once pelleted cells are re-suspended in 20mM sucrose solution and a pulse of electricity is quickly passed through it. Finally the *Sulfolobus* cells are grown in YS medium for 48-72 hours to prepare them for halo assays. The YS media has the nutrients that allow the cells to grow without limitations on energy. It contains traces elements, sucrose, and yeast extract to keep the cells happy in their 75°C bath.

**Sequencing/PCR:**

Next is to select regions to for sequencing. This is done using PCR to amplify the sequence. This technique uses primers and polymerase that anneal flanking each side of the targeted area. This is all done in the thermocycler at various temperatures to initialize, denature, anneal, elongate, and terminate the strands. The process is repeated 30 times to produce millions of DNA copies because each cycle multiplies the previous amount by two (Garibyan 2013).

The viruses need to be subjected to electroporation to put DNA into *Sulfolobus*. Halo assays will be performed to determine if the insertions would still allow the virus to be viable due to having the transposon inserted into a functioning gene or having a function gene deleted/partially deleted. The way to
test is to spot viruses or transformed cultures on *Sulfolobus* lawns. The virus or culture is put on a lawn of *Sulfolobus* and allowed to grow. If the virus is still functional, it inhibits the growth of *Sulfolobus* around the area where it was placed, creating a “halo” around it. This shows that a particular SSV1 strain is still infectious even with the inserted transposon. This can then be compared with a positive control and negative control on the same plate to verify that indeed the virus is still functional. If however the insertion does affect the abilities of the virus, no halo will form and the insertion has apparently caused the ORF to malfunction.

Future work includes using the SSV1 and creating a one step growth curve which measures one cycle of their replication. The idea behind a growth curve is to simultaneously infect all *Sulfolobus* cells at the same time with at least one SSV1 virus and monitor the plaques created at different time points. A graph can then be generated of their plaque forming unit per milliliter (pfu/mL) to determine one cycle. But first we would need to generate a virus culture that is dense enough to infect nearly 100% of a *Sulfolobus* sample. The number that would give us an MOI (multiplicity of infectivity – ratio of virus to cell) of 10 would be about $10^{8-9}$ pfu/mL. Once a one-step growth curve can be accomplished with SSV1, other strains such as SSV3 or SSVK could be the next viruses to study their replication cycle.

Along with the quantitative data I also collected qualitative data from previous studies of SSV1. I worked closely with my graduate mentor that has provided me with a plethora of previous research done directly by the Stedman lab and other labs. Figuring out how this virus work can lead to useful advances in how replication mechanisms work and thus enable us to combat viruses of this nature in the future.
Bibliography


Map of SSV1 genome:

Red arrows: Conserved genes

Black arrows: Non-conserved genes

Black arrows with numbers: Insertion of transposon mutants