Molecular Biology of Archaea 3

An Independent Meeting held at the Max Planck Institute for Terrestrial Microbiology, Marburg, Germany, 2–4 July 2012. Organized and Edited by Sonja-Verena Albers (Max Planck Institute for Terrestrial Microbiology, Germany), Bettina Siebers (University of Duisburg-Essen, Germany) and Finn Werner (University College London, U.K.).

Lytic viruses infecting organisms from the three domains of life

Jamie C. Snyder and Mark J. Young

Department of Plant Sciences and Plant Pathology, and Thermal Biology Institute, Montana State University, Bozeman, MT 59717, U.S.A.

Abstract

It is probably a biological constant that viruses are found associated with all known life and that they play an essential role in the ecology and evolution of all lifeforms. Lytic viruses have now been discovered that infect organisms from each of the three domains of life. Recently, a new lysis system has been described for two archaeal viruses. This lysis system appears to be novel and distinct from previously described virus-encoded lysis systems. A comparison of lytic viruses from each of the three domains may provide insights concerning the evolution of viruses.

Introduction

Viruses are associated with all known life and are the most numerically abundant entities in most environments [1–3]. Each virus has evolved to specifically infect and replicate within its cellular host. Traditionally, viruses infecting bacterial hosts (also called phages) have been classified as lysogenic or lytic. Lysogenic viruses are maintained within the host in a carrier state typically integrated as a provirus form within the host chromosome or maintained as an extrachromosomal element, and thus are passed to subsequent generations by cell division without producing new progeny virus particles. Lytic viruses are defined as viruses that direct the lysis of their host cells to release new progeny virus particles as a consequence of completing their viral replication cycle. As described below, lytic viruses have evolved multiple mechanisms to exit their host cell. The timing of virion release must be highly co-ordinated so that the assembled particles are infectious and ready to infect new host cells. The same virus can exhibit both lysogenic and lytic lifestyles (e.g. λ phage). The lytic cycle involves completion of the viral replication cycle including replication of the viral genome, assembly of the capsid, packaging of the viral nucleic acid and lysis of the host cell directed by virus-encoded protein products. Although both lysogenic and lytic lifestyles are well represented with known bacterial viruses, it is not a universal requirement. For example, the Escherichia coli virus M13 produces channels through which mature progeny virus particles are secreted from the bacterial host without causing cell lysis. In contrast with bacterial viruses, many viruses of multicellular eukaryotes exit infected cells without causing cell lysis, whereas many viruses infecting single cellular eukaryotes cause cell lysis. Regardless of the nature of the host, progeny virus exit from the infected cell is an elegantly orchestrated process that requires the interactions of both host and viral components.

The mechanism(s) by which archaeal viruses (viruses that infect archaeal hosts) exit from infected cells is just beginning to be investigated. A central question is whether these exit mechanism(s) will functionally resemble bacterial or eukaryotic viruses or whether they will have their own unique exit strategies. Furthermore, if they are functionally related, is there any evidence that they are also evolutionarily related? In the present paper, we compare lytic viruses infecting organisms from each of the three domains of life. By examining a fundamental process within the virus replication cycle, we hope to elucidate both independent and shared mechanisms for virus exit from infected cells, which should contribute to our overall understanding of virus evolution.
Lytic ds (double-stranded) DNA bacteriophages

Like all viruses, lytic bacteriophages must exit their host cell in order to find new cells to infect. A major obstacle for bacterial viruses is the peptidoglycan cell wall layer surrounding the bacterial cell. This structure presents a formidable barrier, as it functions to withstand the high osmotic pressure present within bacterial cells. There are two characterized lysis systems for bacteriophages that breach the peptidoglycan cell wall layer (Table 1). The first involves the well-characterized holin–endolysin system encoded by dsDNA bacteriophages [4]. The second lysis system is described for ss (single-stranded) DNA phages and ssRNA phages, which involves a single lysis gene. Both of these lysis systems have undergone evolutionary fine-tuning of the timing of the cellular lysis to co-ordinate virion release to occur after the assembly of infectious progeny particles.

The holin–endolysin system present in dsDNA bacteriophages is a two-component system. The first component, endolysin, is an enzyme coded for by dsDNA bacteriophages that degrades the peptidoglycan-based sugar and amino acid polymer forming the cell wall. There are several different groups of endolysins that are classified depending on which linkage within the peptidoglycan polymer they target. The different covalent linkages that are the targets of the muralytic activities of endolysins are (i) glycosylase and transglycosylase that target the glycosylidic linkages, and (ii) amidase and endopeptidase that target the oligopeptide cross-links [4]. However, the expression of endolysin by itself is not sufficient to direct bacterial cell lysis; effective bacterial lysis requires a second small membrane-associated protein called holin, which is encoded by the virus [4]. Holin proteins are very diverse and are found in many unrelated virus families [5]. These proteins are grouped into two classes on the basis of primary sequence structure. Class I holins are short hydrophobic proteins with three predicted TM (transmembrane) domains. Class II holins are usually smaller than Class I and have only two TM domains. Interestingly, complementation assays have shown that any combination of endolysin and holin will result in cellular lysis even if the proteins are expressed within Gram-positive or Gram-negative hosts [6]. Holins are a very diverse family of proteins with at least three different membrane topologies, indicating that they are evolved from multiple independent origins [5].

In general, the endolysin accumulates within the cytosol of the infected cell until the holin forms a membrane lesion that results in the perforation of the cytoplasmic membrane, and thereby releasing the endolysin to degrade the peptidoglycan layer. In order to generate a pore in the membrane large enough for the escape of large (up to 500 kDa) endolysin protein complexes, the holins must form a multiprotein complex [7]. The ‘simplest’ description of a holin complex is a hole that allows endolysin to cross the cytoplasmic membrane. For infection by bacteriophage λ, a holin lesion formation model has been described [8,9]. This model has been referred to as the ‘death raft’ model and involves the oligomerization of holins in the cytoplasmic membrane resulting in the ‘death raft’ structure; this is followed by the opening of an aqueous channel and concludes with the formation of a lesion [9]. The structure of purified holin rings revealed that each ring has an outer and an inner diameter of 23 nm and 8.5 nm respectively [9]. The height of the holin complex was measured at 4 nm, which closely mirrors the thickness of a lipid bilayer [9]. The holin protein is solely responsible for the timing of cellular lysis [4]. Deletion of the holin gene results in an accumulation of endolysin and of progeny viruses within the cell. It has been discovered that not all holin–endolysin systems involve the release of endolysin through a holin lesion [5]. Endolysin genes with secretory signals have been described; however, these proteins are inactive until activated by the holin, indicating that, even in these systems, it is the holin protein acting as the molecular clock for virus-mediated cell lysis [10].

The holin–endolysin system is found in many bacteriophages, but it is not universal. Many ssDNA phages of the *Microviridae* family encode a single gene that is responsible for the lysis of the bacterial cell. The lysis gene from these viruses is a small membrane-associated protein typically located within another gene [4]. A conserved feature of described phage lysis proteins is their two domain structure [11]. The N-terminal domain of the protein contains the catalytic activity, which will cleave bonds within the peptidoglycan [11]. The C-terminal domain binds to a substrate within the cell wall of the host bacterium, and cleavage of the peptidoglycan requires interaction with this

<table>
<thead>
<tr>
<th>Domain</th>
<th>Viruses</th>
<th>Brief description</th>
<th>Protein(s) responsible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>dsDNA bacteriophages</td>
<td>Endolysin degrades the peptidoglycan layer and holin controls the timing of lysis</td>
<td>Holin and endolysin (λ phage)</td>
</tr>
<tr>
<td></td>
<td>ssDNA, ssRNA viruses</td>
<td>Small protein that inhibits and step in murein biosynthesis</td>
<td>E (ΦX174), L (M52 RNA phages)</td>
</tr>
<tr>
<td>Eukarya</td>
<td>SV40, poliovirus</td>
<td>Small protein encoded late within the replication cycle is responsible for permeabilizing the cell membrane</td>
<td>VP4 (SV40), 2B (coxsackievirus)</td>
</tr>
<tr>
<td>Archaea</td>
<td>STIV, SIRV2</td>
<td>STIV C92 protein will form pyramid structures on the cell membrane of infected cells; assembled virions escape through these structures</td>
<td>C92 (STIV), P98 (SIRV2)</td>
</tr>
</tbody>
</table>
particular substrate [11]. Sequence comparisons of described lysins reveal high sequence similarity in the N-terminal domain and low similarity within the C-terminal domain [11]. In fact, the C-terminal domain can be fully replaced, suggesting that the N-terminal domain of the protein is the active component [4]. From structural analyses, it is hypothesized the binding domain of the lysin (C-terminal domain) may enable the proper orientation of the catalytic domain (N-terminal domain) [12]. In the case of φX174, the phage-encoded lysin gene has been shown to inhibit cell wall synthesis in the infected cell [13]. The current model for cell lysis by φX174 entails the accumulation of protein E within the cell, which leads to an interaction and subsequent deactivation of MrA Y. MrA Y is an essential gene in E. coli and catalyses the formation of a cell wall precursor. The interaction of the φX174 protein E and MrA Y inhibits cell wall synthesis and leads directly to cell lysis [13].

**Eukaryotic viruses**

Several non-enveloped eukaryotic viruses, including poliovirus, SV40 (simian virus 40), Epstein–Barr virus and many algal viruses, assemble progeny virus particles within the cytosol of infected cells and must ultimately exit the cell by mechanism(s) distinct from cellular budding typified by enveloped viruses. Furthermore, these viruses also utilize host transportation proteins (i.e. actin or tubulin) to access the intercellular compartments necessary for virion assembly and exit. However, the exact biochemical mechanisms that non-enveloped viruses utilize in exiting the cell are largely unknown. It had initially been proposed that a build-up of viral proteins within the cytoplasm was responsible for permeating the membrane and the subsequent release of virus. However, akin to bacteriophage-induced cell lysis, the timing of cell lysis is critical and is likely to be more tightly regulated than a passive event involving protein build up within the cytoplasm. The virus must time the cell lysis such that assembled virions are released from the cell instead of being encapsulated in internal cellular membranes. Research into eukaryotic viruses has started to shed light on some potential lysis mechanisms employed by eukaryotic viruses. For example, coxsackievirus encodes a protein that induces calcium leakage from the endoplasmic reticulum that may promote lysis of the infected cell [14], and the adenovirus death protein is hypothesized to play a role in cell lysis, because, when this gene is deleted from the virus, the virus replication cycle is prolonged [15].

SV40 is one of the best characterized non-enveloped viruses; however, the mechanism by which the virus uses to exit its host cell has largely remained a mystery. Through heterologous gene expression, the SV40 core structural proteins were found to be capable of permeabilizing the bacterial cell membrane [16]. However, other viral and host proteins that were essential in the lysis process were unknown. It has been found that SV40 encodes a late protein that triggers the release of the virus from infected cells [17]. This protein, named VP4 (viral protein 4), is a small non-structural protein that is encoded within transcripts for two structural proteins [17]. VP4 is expressed at approximately 24 hpi (h post infection) concurrently with virus-induced cell lysis, and the deletion of VP4 resulted in a prolonged virus replication cycle [17]. Thus VP4 is expressed late within the virus replication cycle, and this protein is essential for the timing of cell lysis [17].

Many enveloped eukaryotic viruses such as HIV, Ebola and hepatitis C utilize cellular proteins for exit and do not result in the lysis of the host cell. These viruses utilize the cellular ESCRT (endosomal sorting complex required for transport) machinery to accomplish viral release. In eukaryotic cells, ESCRT proteins have been shown to be involved in biogenesis of multivesicular bodies and cell division [18]. In addition to these functions, the ESCRT machinery can be co-opted by viruses. Many enveloped eukaryotic viruses encode short sequences named L-domains (L for late) that interact with ESCRT components [18]. However, there are enveloped viruses such as influenza that do not utilize the cellular ESCRT machinery during viral release [19]. Similar to non-enveloped eukaryotic viruses, enveloped eukaryotic viruses appear to have evolved different ways of exiting their host cell.

**Archaeal virus lysis systems**

Our knowledge of archaeal viruses is based on two major groups of viruses infecting the archaeal domain of life: the haloviruses infecting halophilic Euryarchaeota, and viruses infecting thermophilic members of the Crenarchaeota. Since their discovery, haloviruses were described as lytic viruses with virion morphologies similar to the established bacteriophage viral families of Myoviridae and Siphoviridae. More recently, a much broader range of halovirus morphologies have been described [20], clearly indicating that haloviruses are not limited to only head–tail morphologies. However, there is little genetic similarity between the haloviruses and bacteriophages. It remains to be discovered whether the lysis systems of halophages are more similar to that of the holin–endolysin system of dsDNA bacteriophages or to that of the single-gene lysis system operating in small ssDNA viruses. In contrast, the viruses infecting the thermophilic members of Crenarchaeota typically represent novel virus families and, until recently, were thought to be non-lytic. It was hypothesized that these viruses were non-lytic because of the harsh environments that many of the organisms inhabit.

An archaeal viral lysis system has been described involving the formation of unique pyramid structures on the membrane of an infected crenarchaeal cell [21,22] (Figure 1). This lysis system has been described for two otherwise unrelated viruses: STIV (Sulfolobus turreted icosahedral virus) [22] and SIRV2 (Sulfolobus islandicus rod-shaped virus 2) [21]. Besides infecting the same genus, these viruses are very different. STIV is an icosahedral virus that encloses a membrane and a circular dsDNA genome [23]. SIRV2 is a non-enveloped rod-shaped virus with a linear dsDNA genome [24]. Overall, these two viruses encode three genes with secondary sequence
similarly. One of these genes has been shown to be solely responsible for the formation of the pyramid lysis structures [25,26]. The gene has also been shown using genetic studies to be essential for STIV replication [26]. Similar to lysis proteins present in bacterial and eukaryotic viruses, these proteins are small (10–12 kDa). STIV C92 and SIRV2 P98 are helical proteins with a predicted N-terminal transmembrane domain [25,26]. It has been found recently that these proteins are highly homologous, but not completely interchangeable [27].

In STIV, at approximately 32 hpi, the S-layer begins to dissolve away, and the underlying cell membrane starts to protrude through [28,29]. Transcriptomic data from analysis of STIV-infected cells indicate that C92 is an ‘early gene’ in which transcripts were first detected at 8 hpi and peaked at 24 hpi [30]. Seven-sided pyramids are visible on the cell membrane before STIV-induced cell lysis (Figure 1A). When cell lysis occurs, the sides of the pyramid fold back, resulting in a hole in the S-layer (Figure 1B) allowing for the escape of assembled virions. Currently, it is unknown if the initial process of degrading the S-layer is enzymatically or mechanically driven. Bioinformatic analysis indicates that neither of these proteins contains enzymatic domains. Immunolocalization studies have shown that the pyramid structures are composed of viral proteins (C92 in the case of STIV and P98 in the case of SIRV2) [25,26]. It remains a challenge to determine how such a small protein could be responsible for such structures. A genetic system has been developed for STIV [31] and has already proved to be a valuable tool [26]. This tool will no doubt prove to be more valuable as the timing of STIV-induced lysis of Sulfolobus cells and host factors that may be involved with lysis are elucidated further.

This viral lysis system appears to be widespread among archaeal viruses within acidic hot springs [32]. By mining several cellular and viral metagenomes created from various hot springs within Yellowstone National Park, four groups of archaeal viruses were identified that probably utilize this lysis system [32]. It is probable that these represent novel archaeal viruses that possess the same lysis system as STIV and SIRV2. Current research involves isolation of these C92-like sequences from the environment to investigate whether they function like C92 and P98 and form pyramid-lysis structures on the cell membrane.

It is tempting to speculate on the apparent similarities between C92-induced lysis and the small bacterial lysis systems. Like the bacterial lysis proteins, the C92 and P98 proteins are small α-helix-rich proteins with two domains. Structural analysis of pyramid structures isolated from cell membranes may be very insightful for determining how these structures form. In addition, it remains to be determined whether cellular proteins are involved in the formation of these unique exit structures.

Conclusions and future directions
Lytic viruses infect organisms belonging to all three domains of life. Over time, these viruses have had to fine-tune their replication cycle to ensure that lysis occurs only when assembled virus particles are ready to infect other cells. Given the significant differences in archaeal, bacterial and eukaryotic cell structure, it is not surprising that unique solutions to virion exit have evolved by viruses infecting hosts in each of the three domains of life. A novel lytic system for two archaeal viruses has been identified, and we are still trying to fully understand how this lysis system operates within archaeal host cells. One of several remaining questions is whether there exist other lytic systems operating in viruses infecting Bacteria and Eukarya that have not yet been discovered and characterized. It seems plausible that, within the vast reservoir of unknown viruses, additional lytic systems will be discovered. Another question is whether there are other lytic mechanisms employed by other viruses infecting archaeal organisms. Furthermore, it remains to be determined whether the lysis systems of lytic viruses are evolutionarily linked or whether they have evolved independently. These and other questions will no doubt be answered as more viruses are discovered and virus replication cycles are determined, especially within viruses infecting Archaea.

Funding
This research is supported by the National Science Foundation [grant numbers DEB-0936178 and EF-080220].

References

Received 2 November 2012
doi:10.1042/BST20120326