INTRODUCTION

1. Bacteriophage

Viruses are structures consisting of nucleic acid molecules with protective coats that are replicated by the enzymatic machinery from suitable host cells. Viruses can infect plants, animals and bacteria. Each viral species has a limited host range and can reproduce in only a small group of closely related species.

An intact virus particle, referred to as a virion, consists of a length of nucleic acid, either DNA or RNA, encased by a protein capsid. The relative simplicity of viruses in comparison to cells makes them invaluable tools in the elucidation of gene structure and function, and our best characterized models for the assembly of biological structures.

An important model system , the tailed, double-stranded DNA (dsDNA), bacterial viruses (bacteriophages) of the order *Caudoviridae,* can be subdivided into three families: *Podoviridae* (short non-contractile tails), *Myoviridae* (long, contractile tails) and *Siphoviridae* (long, non-contractile tails). Members of the Family *Siphoviridae* include T1, SPP1, HK97 and λ. The majority of all virus research in the past has focused on the *Siphoviridae* family.

Two lifecycles exist for phage: lytic and lysogenic. A lytic phage infects a host cell and uses the host transcription and translation machinery to replicate itself. It then kills the cell through lysis. Phage that follow a lysogenic lifecycle integrate their genome into the genome of the host cell. The phage genome is replicated during cell division and remains dormant until stimulated by heat shock or DNA damaging agents to enter the lytic cycle.

Phage following a lysogenic life cycle is described as a prophage, whereas its host is called a lysogen. An intriguing property of lysogens is that they cannot be reinfected by phages of the type with which they are lysogenized: they are immune to superinfection. A

bacteriophage that can follow either a lytic or lysogenic life style is known as temperate phage, whereas those that have only a lytic mode are said to be virulent.

Bacteriophage λ (58MDa) is a temperate, dsDNA coliphage (bacteriophage the infects *Eschericia coli*). Phage λ's molecular biology is among the most extensively characterized of all complex viruses.

Bacteriophage λ adsorbs to *E. coli* through a specific interaction between its tail fiber and maltoporin (the product of the E. Coli lamB gene), which is a component of the bacterium's outer membrane.

Structural studies of λ, as well as other bacterial viruses, have played an integral role in the development and advancement of structural biology. Electron cryo-microscopy (Cryo-EM) is particularly well suited for the challenges of structure determination of mega-dalton sized viruses. EM has long been a primary tool for classifying viruses and exploring their structures. The last two decades has also seen a burst of activity in the use of EM for the elucidation of virus structures. This transition has resulted from two advances in techniques. Firstly, cryo-EM has allowed the preservation of fragile specimens in EM (Adrian, MJ., 1984; Dubochet, J., 1998) in their native state. Secondly, the development of efficient algorithms for processing micrographs to produce 3-D structures of icosahedral particles has allowed to this higher-quality data to be fully exploited (Baker, TS., et al 1996; Crowther, RA., 1971; Crowther, R.A., et al 1970). These two developments have made 3-D structural information accessible for a broad range of viruses at the same time that high-resolution X-ray diffraction and nuclear magnetic resonance (NRM) spectroscopy studies have revealed atomic detail about a handful of virus proteins.

λ Assembly Overview

Double-stranded DNA (dsDNA) bacteriophages can be characterized by their icosahedral capsid, central connector region and long non-contractile tail. The interactions between these three components are fundamental in phage function. Both the capsid and tail are assembled in two independent assembly pathways and join together via the connector region. The connector is attached to one unique vertex of the capsid. On the other end, the tail binds to the connector.

The capsids of phage virions are icosahedrally symmetric and about 60 nm in diameter. The non-contractile tails are 150 nm long, and appear to be slightly flexible, due to how they lie on an electron microscopy sample grid. The main body of the tail (the "shaft" or "tube") consists of stacked protein discs which attach to the phage head on one end and the base of the conical "tail tip" on the other. This morphology – icosahedral head and long non-contractile tail- is most common among the double-stranded DNA phages that have been isolated and characterized.

The head has two main protein components: gpE (gene product E) and gpD, arranged in T=7 *laevo* icosahedral symmetry, and present in exactly 405 copies each (Casjens, S.R., et al 1974; Dokland, T., et al 1993). GpE, also known as the "coat protein", makes up the main structure of the shell, and the smaller gpD subunits are clustered as trimers bound at the 3-fold sites on the exterior of the gpE lattice (REF). The role of gpD is to strengthen the capsid (REF).

The double-stranded DNA chromosome, tightly packed into the head without bound proteins, is 48, 503 bp long. The ends of the linear DNA molecule in the virion have 12 nucleotide single-stranded 5'-extensions; these are referred to as "cohesive ends" because of the base-pairing complementarity of the extensions at the two ends of the chromosome (Hershey, AD. 1965).

The connector region of λ occupies on of the twelve 5-fold symmetric vertices of the capsid. It is composed of 12 copies of gpB, also knowna s the portal protein and gpFII (Kochan, J. et al 1984; Tsui, L. et al 1980, Casjens, S. 1974). GpW is required for DNA stabilization and connector assembly, but its presence as a structural component has not been established. The role of gpB, gpFII and gpW in λ assembly has been well

characterized (Maxwell et al, 2001; Maxwell et. al, 2002, Murialdo et al, 1979). Phage λ assembly begins with the synthesis and oligomerization of gpB, creating a nucleating point for head assembly. The capsid is assembled and DNA passes through the portal into the heahead. Multiple copies of gpW are required to stabilize the DNA in the capsid. Finally, to complete the connector structure, gpFII binds after gpW and acts as the head-tail joining protein. It is only with the addition of gpFII to the connector that complete tails may bind successfully to the head. The structure and arrangement of components in the connectore region are also unknown.

The connector also plays a critical role during infection. During ejection of the genetic material into the host bacterium, DNA passes once again through the connector. Unknown interactions within the connector allow the "plugged" entry/exit pathway to become "unplugged", allowing the DNA to pass. The specific roles of gpFII, gpW and gpB in this regard remain unknown.

Phage λ has two types of tail fibres: one short fibre extending from the centre of the tail tip and four long, joined "side tail fibers" attached about the junction between the tail shaft and the tail tip.

The virion's tail shaft consists of 32 stacked hexameric rings of the major tail subunit, gpV (Buchwald, M.H., et al 1970a; Buchwald, M.H., et al 1970b; Casjens, S.R., et al 1974; Katsura, I. 1983). Six copies of a proteolytically processed form of the tail-length measure protein gpH fills most of the lumen.(Hendrix, R.W., 1974; Tsui, L.C., et al 1983). The right end of the DNA has descended a short distance into the head-proximal end of the tail (Thomas, J.O. 1974; Chattoraj, D.K., et al 1974). The tail tip contains several copies of the products of genes J, L and M (and perhaps I and K) (Katsura, I., 1983). GpJ is present in about three copies in the mature virion and constitutes the central tail fiber and a substantial portion of the conical tip. The head-proximal end of the tail has six copies of gpU, a hexameric ring on top of the stack of gpV rings that make up the shaft (Pell, L., et al 2009, Katsura, I., et al 1975, Katsura, I., et al 1977). The last gene in the tail assembly pathway, gpZ, is required for joining tails to heads in active phage (Casjens, S., et al 1974); it is not clear whether gpZ is present in mature virions.

2. Intro to EM

Transmission electron microscopy is an important technique in structural biology for studying large biological macromolecules. It is a direct method of structure determination, complementing other techniques in structural biology, such as X-ray crystallography and NMR spectroscopy.

Historically, macromolecular specimens were prepared by embedding them in a heavy metal salt negative stain supported on a continuous carbon film. The stain enveloped the protein, providing a high-contrast negative of the protein being imaged. Resolution is limited by the grain size of the stain and internal information about the protein is usually not recovered. There are drawbacks to using stain. It has also been shown that stain produces various artifacts and distortions such as specimen flattening, partial staining and positive staining. During specimen flattening, the sample is physically flattened in the direction normal to the plane of the grid (REF). Partial staining occurs when a portion of the sample remains unstained and there doesn't appear in the micrograph. Despite these drawbacks, negative stain is still a useful method to determine low resolution gross structural features of macromolecules.

In the late 1980's, electron cryomicroscopy (Cryo-EM) was developed as a method to address the problems associated with negative stain (Dubochet *et al*., 1988). Specimens prepared for cryo-EM are hydrated proteins flash frozen in liquid ethane at liquid nitrogen temperatures. The protein of interest is in a layer of vitreous ice. Electrons are scattered by the atoms of the sample itself, instead of a heavy metal salt stain. This allows internal information about the sample to be recorded. The contrast in an image recorded during a cryo imaging session is much lower than the contrast from negative stain.

Although cryo-EM provides many advantages over negative stain microscopy, the technique is still limited by the radiation sensitivity of the specimen. The information in

a micrograph is collected from the electrons that have elastically interacted with the specimen and therefore do not deposit energy in the sample. These electrons emerge with an unchanged wavelength, but with a phase shift of 90°. However, between 100-300 keV, for every elastic scattering event in biological samples, there are three inelastic scattering events (Henderson, Quarterly Reveiws Biophysics), where the electrons deposit energy in the specimen. Because the electrons used in imaging have energies between 100-300 keV, and depositing even a small fraction of their energy in the specimen is enough to break protein bonds, which have energies of a few electron volts (Petsko and Ringe, 2004)

While interference between electrons of the same wavelength produces contrast in images, the wavelengths of inelastically scattered electrons are changed upon scattering. The majority of scattered electrons therefore lead to specimen damage without contributing information to the image. Two techniques have been developed to reduce radiation damage. By cooling the specimen with liquid nitrogen or liquid helium during imaging, radiation damage is reduced. Secondly, imaging with a low dose of electrons (5-25 $e^{-}/Å^{2}$) reduces the number of scattering events, and hence limits the inelastic interactions and radiation damage. However, a low electron dose also limits the number of elastic interactions, leading to poor signal to noise ratios (SNR) in images. The signal can be improved by averaging images tat show the same view of the specimen.

a. Image Formation & CTF

The principles of image formation for electron microscopy are similar to those of light microscopy. There are two sources of contrast. Amplitude contrast is produced by the loss of electrons scattered at high angles. Phase contrast is produced by interference between the elastically scattered electrons and the unscattered beam. Contrast in micrographs of stained (~86%) or unstained (~93%) specimens is largely due to phase contrast.

Images formed by an electron microscope are approximately 2-D projections of the 3-D structure of the specimen onto the plane of the image. Every microscope convolutes images with its point spread function (PSF). The PSF, due to phase contrast, inverts the contrast at certain spatial frequencies and modulates the amount of contrast across the spectrum of spatial frequencies. Images can be partially corrected for the PSF in Fourier space using the projection theorem (Bracewell?) and the convolution theorem (REF).

b. 3D Reconstruction methods

To build an accurate 3-D map from 2-D projections of a specimen, the angular relationship between each projection is required. There are several methods for initially determining the projection directions, including common lines in Fourier space (Crowther, 1971) and the real space equivalent angular reconstitution (van Heel, 1986?), and random conical tilting (Radermacher, 1981). Once a 3-D map of the specimen of interest is available, the strategy of projection matching can be employed. The 3-D map can be projected along a large number of directions and images may be assigned a direction by finding the view with which they best agree (Harauz & Ottensmeyer, 1984). Once the best view is determined for all images, a new 3-D map can be built using weighted backprojection (REF), algebraic reconstruction (REF), or Fourier synthesis (DeRosier, Klug, 1964). This process can be iterated until no further improvement in the map is achieved.

3. Electron microscopy studies of the connector region of Bacteriophage SPP1 and P22

Viruses are attractive models for the study protein-protein interactions and the assembly of macromolecular complexes because the genetics of viruses are often far simpler than their multifaceted assembly processes would suggest. To address questions about virus structure, many laboratories have used X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy to determine atomic structures for a number of assembly-related proteins from different bacteriophages (REF). Recently, the use of electron cryo-microscopy (cryo-EM) has allowed research groups to study the assembly and function of tails and capsids in a holistic way (REF, REF, REF, REF, Rossman, Johnson, Carrascosa, etc). Atomic models from X-ray crystallography and NMR spectroscopy can be combined with lower resolution 3-D maps from cryo-EM to determine the arrangement of subunits and subunit interactions within the macromolecular complex.

A key component of many bacteriophages is their portal protein. The portal protein plays a number of important roles. During phage assembly, it is the first structural protein transcribed. The individual monomers of the portal assemble into a ring-like structure and serve as the nucleating point for capsid formation (REF, Ref, REF, REF,REF). The portal, along with the connector region in general, serves as the main conduit for DNA going into the capsid during assembly and DNA egress when it is ejected during infection of a host. Different bacteriophages have portal proteins which are similar in function, despite having low sequence identity (REF, REF, REF, REF).

Recently, 3-D maps of the connector regions of bacteriophages SPP1, P22 and Φ29 were published (REF, REF, REF). These three bacteriophage are similar to phage λ. Both SPP1 and P22 are members of the *siphoviridae* family, while Φ29 is a member of the *podoviridae family* (REF, REF, REF). Interestingly, the portal proteins from all three bacteriophages were dodecameric ring structures.

4. Aims of the present work

The connector region of λ plays a vital role in both its structure and function. Once assembled, the portal protein provides a nucleating point for λ assembly. Phage λ 's genome passes through the connector to be packaged in the head and exits through the connector during host infection.

Currently, there are numerous 3D structures of the head region of λ, in various stages of assembly (REF). However, there is less information available on the connector region (REF, REF, REF, REF). High resolution structures have been solved with NMR spectroscopy and X-ray crystallography for monomeric gpW and gpFII (REF, REF). Their oligomeric state or location in the connector is uknown. To determine the protein interactions, order and structure of the connector, a three-dimensional structure is required.

The first aim of this work was to reconstruction a 3D model of the connector region of bacteriophage λ using cryo-electron microscopy. Next, to determine if gpW is a structural component of the connector, maltose binding protein (MBP) was tethered to gpW and imaged using negative stain EM.

MATERIALS & METHODS

Cesium Chloride Purification of Intact Bacteriophage λ

A culture of *Escherichia col*i (*E. coli*) lysogen (594 cI_{857} S_{AM7}) was grown in LB media (1 % [w/v] Bacto-tryptone, 0.5 % [w/v] yeast extract, 1 % [w/v] sodium chloride) overnight. The culture was diluted into an initial optical density (O.D.) of 0.1 of 2 L of KH Media (0.1 % [w/v] ammonium chloride, 0.3 % [w/v] potassium phosphate monobasic, 0.6 % [w/v] sodium phosphate dibasic heptahydrate, 0.05 % [w/v] sodium chloride, 0.06 % (w/v) magnesium sulfate heptahydrate, 12.3 μM ferric chloride, 0.2 % [w/v] glucose, 15 % casamino acids [w/v]) and grown to an O.D. of 0.5-0.6 at 30 °C. Cells were incubated in a 45 °C water bath for 20 min to induce phage λ gene translation before a further 3 h growth at 37 °C. 20 mL of chloroform was added to the culture an shaken for 10 min at 37 °C to lyse the cells. The suspension was chilled to 24 °C and RNase and DNase were added, each to a final concentration of 1 μg/ml. Sodium chloride was added to a final concentration of 1 M and the culture was cooled on ice for 1 h. The suspension was centrifuged at 11,000 *g* for 15 min to pellet large cellular debris. The pellet was discarded and the supernatant pooled. Polyethylene glycol 8000 was added to a final concentration of 10 % (w/v) to precipitate the bacteriophage and the solution stirred at 4 °C overnight. The solution was centrifuged at 11,000 *g* for 15 min to pellet the bacteriophage. The pellet was re-suspended in 16 mL of SM media (0.58 % [w/v] sodium chloride, 0.2 % [w/v] magnesium sulfate heptahydrate, 0.05 % [v/v] Tris-chloride (pH 7.5), 0.005% [v/v], 2% [w/v] gelatin solution) per 1 L of initial culture volume. An equal volume of choloroform was added and the solutions were centrifuged for 15 min at 11,000 *g* to separate the organic phase from the aqueous phase. The aqueous phase was removed and 0.74 grams solid cesium chloride was added per mL of aqueous phase recovered. The final solution was added to a 13.2 mL Beckman Quickseal Centrifuge Tube and SM media with cesium chloride (74 % [w/v]) was used to fill the remaining volume in each tube. The tubes were heat sealed and centrifuged at 135,000 *g* (50,000 RPM) for 20 h at 4 °C. Samples were visually inspected for a successful phage growth. If phage were present, a thin, white band was visible in the middle of the centrifuge tube.

The phage band was extracted with a syringe and the sample was dialyzed against 4 L of λ-dilution buffer (10 mM Tris, 10 mM magnesium sulfate, pH 7.5) for 4 h.

Separation of Phage λ Heads and Tails

To separate heads from tails while keeping the connector region intact and associated with the tail, purified whole phage were treated with 3 M guanidine hydrochloride (GuHCl). Equal volumes (approximately 30 µl) of purified phage and 6 M GuHCl were mixed, incubated for 10 min at 55 °C, and then cooled to room temperature.

Preparation of Grids for Electron Microscopy

Continuous carbon film coated EM grids were used as a substrate to visualize samples by negative stain electron microscopy. First, a medium thickness layer of carbon (~100 Å) was evaporated onto a freshly cleaved mica surface. Approximately 30 –40 grids were placed onto a metal support beneath the surface of the water. The carbon film was floated onto a layer of distilled water in a large Pyrex dish. The water was drained out of the dish, allowing the layer of carbon to gently fall onto copper grids placed beneath the surface of the water.

For electron cryomicroscopy, EM grids with a thin carbon film, supported by a thick, perforated carbon film, was used as a substrate for the sample. To prepare the perforated carbon substrate, 20 drops of a 50 % (v/v) aqueous glycerol solution were added to 50 mL of 0.5 % (w/v) Formvar in chloroform and the suspension was sonicated with a probe sonicator for 30 s at maximum power 6 times. Clean glass slides were submerged in the Formvar solution for 5 s and then left to dry in air. A diamond-tipped glasscutter was used to cut the Formvar film on the glass slides and the film was floated onto the surface of a water bath. (FIGURE). 400 mesh Cu/Rh grids (INFO) were placed onto the surface of the film with the copper side visible. A piece of Parafilm backing paper (INFO) was placed against the grids to lift the film off the surface of the water and the grids were allowed to dry. The Formvar layer was etched with methanol for 10 min and, once dry, a thick layer of carbon was evaporated onto the Formvar layer on top of the

grids. The grids were then carefully washed with chloroform to dissolve all remaining Formvar.

To prepare the continuous carbon film, clean glass slides were submerged in 0.5 % (w/v) Formvar in chloroform and allowed to air dry for 2 min. As described above, the Formvar film from the slide was floated onto the surface of a water bath. The grids with perforated carbon films, described above, were placed onto the film with the copper side visible and the film and grids were removed and dried. A thin layer of carbon was then evaporated onto the Formvar film. The grids were rinsed with chloroform to dissolve the Formvar layer, leaving a thin layer of carbon over the thick, perforated carbon support (FIGURE). Grids were stored for up to 1 month prior to use.

Specimen Preparation for Negative Stain Microscopy

Continuous carbon-coated grids were glow discharged in air (pressure? – 3 mbar?) for 15 s at 30 mA to create a hydrophilic surface on the carbon grid. Immediately after glow discharge, a 5 µL drop of protein solution was placed on the carbon surface and particles were allowed to adsorb to the grid for 2 min. Excess solution was blotted away with filter paper. The grid was washed 3 times with 40 µL of milliQ water, and blotted dry before staining with 40 µl of 2 % (w/v) uranyl acetate (Electron Microscopy Sciences, Hatfield, Pennsylvania) solution. Excess stain was removed by blotting with filter paper and the grid was air dried.

Specimen Preparation for Cryo-EM

Grids with a thin continuous carbon film and perforated carbon support were glow discharged for 15 s at 30 mA and 3 mbar to create a hydrophilic surface. 5 µL of GuHCl-treated bacteriophage solution was placed on the carbon surface and particles were adsorbed to the surface of the grid for 2 min. The excess solution was removed by blotting, and the grid was then washed 3 times with 40 µL of milliQ water to remove excess GuHCl. The last blot after the final wash was performed with a temperature and humidity controlled grid preparation robot (Vitrobot, FEI Company, Eindhoven, Netherlands) with a -4 mm blot offset and 9 s blot time. After blotting, the grid was

plunged into a bath of liquid ethane at liquid nitrogen temperature to vitrify the sample. Grids were stored under liquid nitrogen until imaging.

Microscope Image Acquisition

All electron microscopy was performed with a Tecnai F20 field emission gun electron microscope operating at 200 kV (FEI Company, Eindhoven, The Netherlands).

Zero Tilt Imaging

Micrographs were acquired at 50,000 x magnification with a nominal defocus range of 1 μm – 3 μm. The condenser 2 aperture was set to 50 μm and the objective aperture to 50 μm. To maintain a minimal electron dose, low dose methods were employed. Diffraction mode imaging was used to search the grid for an area with appropriate ice thickness and phage density. Focusing was done at 80,000 x 1.85 μm on either side from the grid area of interest (FIGURE). At the appropriate defocus level (between 1 μm – 3 μm), an exposure was recorded. Negative stain micrographs were exposed for 1 s, resulting in a dose of ~12 $e^-/Å^2$. To increase contrast, cryo micrographs were recorded with a 2 s exposure, resulting in a dose of ~25 $e^-/Å^2$.

Tilts Pairs

Pairs of micrographs of the same area of the specimen were recorded with the specimen tilted -15° and +15° relative to the electron beam by adjusting the tilt of the goniometer. Imaging conditions were the same as for zero tilt images. Image pairs were taken in sets of five. Five images were taken at one tilt (-15°), the goniometer was adjusted and the microscope was allowed to settle before proceeding with acquiring the next five corresponding tilt pairs (+15°).

Micrographs were developed in full-strength D19 developer (Kodak Canada, Toronto, Ontario) for 3.5 min for 25 $e^-/Å^2$ doses and 12 min for 12-15 $e^-/Å^2$ doses. Film were washed in water for 30 s and then fixed in fixing agent (Ilford Hypam) for 3.5 min with continuous agitation. The micrographs were rinsed in water for 20 min, immersed in PhotoFlo (REF) and left to dry overnight.

Film Densitometry

Micrographs were digitized with an Intergraph Photoscan densitometer (Intergraph, Huntsville, Alabama). Each micrograph was scanned individually as a tiff image and a step size of 7 μm before conversion to MRC format with Tiff2MRC (Intergraph) for further image analysis.

Image Processing

Image processing, unless otherwise stated, was done with the MRC image processing programs (Crowther *et al.*, 1996). XIMDISP was used to visually inspect each micrograph. Micrographs were compressed 2x2 before particle selection, so that each pixel represents 2.8 Å at the specimen level. Particle locations were recorded as 2-D co-ordinates and individual particles were windowed in 128 x 128 pixel boxes with LABEL. CTF parameters were determined using CTFFIND3, with a defocus range of 10000-40000 Å, defocus step of 5000 Å, and resolution range of 15 to 300 Å. Predicted and observed power spectrum were visually inspected to determine quality of fit. Any micrographs exhibiting specimen drift and/or astigmatism were discarded.

Particle Alignment & Class Averages

The program SPIDER (REF) was used to align the particles and create class averages. A low pass filter was applied to individual particle images to enhance contrast for 2-D alignment. Particle images were averaged to create an initial average. Individual raw images were translationally and rotationally aligned to the average, and a new global average calculated. This procedure was iterated until the global average stopped changing and individual particle images appeared to have the same orientation. Particle images that did not align were removed from the data set by manual inspection. The 2-D averages produced were visualized in WEB (REF).

SPIDER was also used to create class averages. After the alignment described above, classification was done using principle component analysis followed by K-means classification (REF – Frank book, 2003). The number of classes created was user specified.

3-D Model Construction

An initial model was built using the final 2-D average produced during particle alignment. Using EM2EM, the 2-D average was converted into an MRC format file and duplicated 150 times. The set of 2-D averages were placed into an image stack. Each image in the stack was assigned an angle along its long axis, parallel to the direction of the tail, evenly covering the range from 0 to 360^{o} in increments of 2.5^{o}. The program build_fspace.com (written by Dr. John Rubinstein) was used to create an initial 3-D model by interpolation in Fourier space. The 3-D initial model was smooth, but had the overall shape and dimensions of the connector region (figure).

Model Refinement

EMAN

The initial model described above was used in EMAN (REF) to determine the overall symmetry of the model. The REFINE command with the following parameters was used to refine the initial model:

> "refine 8 mask=62 hard=25 median sym=c6 ang=2 pad=90 classiter=3 sep=2 xfile=5.6,8800,40 amask=15,0.066,6.4 usefilt fscls"

The a list of operators and parameters after the *refine* command controls the refinement procedure as follows:

Refine #	*Number of iterations run*
Mask	*Applies an outside mask at several stages of refinement*
Hard	*This specifies how well the class averages must match the model*
Median	*CTF correction is OFF*
Sym	*Imposes symmetry on the model (various symmetries were tested as described in section x.x.x)*
Ang	*Angular step for projections (in degrees)*
Pad	*Pads the model during Fourier reconstruction (in pixels)*
Classiter	*Number of iterations to align an image to a class average*
Sep	*Number of classes a particle image can belong to*

Xfiles	*Used to generate rough 3-D models based on parameters of the image particles; <a/pix>, <mass in kd>, <alignment threshold>*
Amask	*Automask; performs a procedure similar to solvent flattening; <radius in pixels, <isosurface threshold>, <width in pixels of soft mask shell>*
Usefilt	*Instructs EMAN to use filtered images for alignment, but raw data for model building*
Fscls	*Experimental parameter used for classification*

The combination of parameters was determined through discussions with Wilson Lau and trial and error.

FREAlign

FREAlign (REF) refinement was performed using automated scripts written by Dr. John Rubinstein and Lindsay Baker. FREAlign employs projection matching and 3-D reconstruction in Fourier space. Alignment scripts were run on a multi-cpu high performance cluster at the Centre for Computational Biology at The Hospital For Sick Children in Toronto, Canada.

All 3-D maps were visualized using Chimera (REF)

Free Hand Test

The free hand test (REF – Rosenthal & Henderson, 2003) was used to correctly identify the hand of the 3-D model. The test uses particle image tilt pairs. The first image of each tilt pair is aligned to the current map and the orientation recorded. The value of the alignment function for the second image of a tilt pair is checked at different orientations that cover all possible tilt angles. The tilt angle with the highest value of the alignment function is compared to the applied tilt. If the calculated tilt is the same as the applied tilt, the hand of the map is correct. If it is the negative of the applied tilt, the hand is reversed.

Subunit Labelling

Lia Cardarelli (Alan Davidson Lab, Department of Biochemistry, University of Toronto) prepared the MBP-gpW construct, purified the sample and assayed it for activity. Specifically, the gene for gpW was cloned into the pMAL-c4X Vector using the XbaI and EcoRI restriction sites. The vector was transformed into E. coli 594(λ $W_{amber403}cI_{857}S_{am7}$) cells. Phage λ was purified using the protocol outlined in section X_Y_Z_X.
Once purified, phage λ was assayed for activity using a spot assay. Briefly, different concentrations of λ were applied to an agar plate covered in a lawn of TC600 supE amber suppressor cells. Spots would appear where λ lysed bacterial cells, indicating normal activity.

RESULTS

1. Introduction

As the only conduit for DNA passage during viral assembly and host infection, the connector is critical to the overall function of bacteriophage λ. The connector also acts as the site of tail attachment to the head in the last step of phage λ assembly.

The connector region of bacteriophage λ was studied using cryo-EM to determine a 3-D map of the intact complex. Subunit labeling, negative stain EM and 2-D image processing were used to identify the location of gpW within the complex.

While experiments by others have produced 3-D maps of phage λ heads, no 3-D map exists for the connector region of phage λ. Although both gpB and gpFII have previously been identified as structural components of the connector (REF), and gpW was shown to be necessary for phage λ to function and assemble (REF), the location of gpW in phage λ is unknown. Atomic models of several of the morphogenetic proteins of phage λ have been determined using X-ray crystallography and NMR spectroscopy, including

monomeric gpW and gpFII (REF). However, the oligomeric state of the connector region remain unknown.

Sample Preparation

Bacteriophage λ was purified and a variety of conditions were tested to seperate heads from tails. . Table X_X_Y lists the different reagents tested for the ability to seperate phage λ heads and tails. It was found that 3 M guanidine hydrochloride (GuHCl) for 10 min at 55 °C provided the most thorough head-tail separation (Figure X_X_Z). After seperation of capsids, the connector remained attached to the tail. Tail tips were not seen post-treatment.

GuHCl is a strong denaturing chemical. To ensure that GuHCl did not disturb the structure of the connector, images of the sample before and after 3 M GuHCl were taken in negative stain. Particle images of the connector were selected from each data set and averaged together to improve the signal to noise ratio (SNR) over that of a single image. The 2-D average of the 3 M GuHCl-treated sample was overlayed onto the 2-D average image of wildtype phage λ and the opacity reduced to 50% (Fig XYZ). The connector region remained intact after the 3 M GuHCl treatment .

Bacteriophage λ in the electron microscope were consistently seen lying length-wise against the carbon surface of the grid (FIGURE). This was likely due to its large size and long, rectangular shape.

2. 2-D Analysis

2-D analysis of the phage λ connector began with image processing as described in section X_I_O_D.
Using SPIDER (REF), a 2-D image processing application, particle images were added together and averaged. Individual particle images were translationally and rotationally aligned using the average as a reference. Particle images were then averaged to create a

new reference. The alignment process was repeated until all particle images were aligned.

The negative stain 2-D average of the connector region showed strong boundaries between certain domains. The tail portion that remained in the in the individual particle images aligned well and was distinct from the connector region (FIGURE).

Rough measurements were made of the 2-D average. The width of the central channel was ~25-30 nm. The width of the tail was ~90 nm and the width of the portal protein was ~150 nm. The measurements corresponded with earlier studies by others of phage λ (REF).

Classification was performed after 2-D alignment using SPIDER. Particle images were assigned to classes based on unique structural features and averages were generated. Different views of the connector were sought in the class averages. Unfortunately, phage λ was seen to have a preferred orientation on the EM grid and the class averages offered very few unique views (FIG). One class average showed phage λ tails without a connector attached. Particle images belonging to that class average were removed from the data set.

The 2-D average produced with the cryo data looked similar to the 2-D average from the negative stain data. The overall size and shape was the same. The connector 2-D average from the cryo data set was measured and was consistent with measurements from the negative stain data.

Class averages from the cryo data were also similar in appearance to class averages from the negative stain data. Few unique classes were generated (FIG).

3. 3-D Model Building

To begin the model building process, an initial model was created. An initial model was generated using the cryo-EM 2-D average of the connector region (FIGURE). A rotational symmetric 3-D model was created such that each side-view of the initial model corresponded to the 2-D average. The initial model was called $C_{inifinity}$. The approach has proven successful for the interpretation of structures that are approximately rotationally symmetric (eg Tocilj…Rubinstein…Cygler, 2008)

This pertains to refining the initial model: not building the initial model:
The projection matching approach was then used refine 3-D model building (ref). Briefly, 2-D projections were generated from the initial model at different angles. The raw particle images were matched to the projections and assigned a relative angular orientation and these orientations were used to build new model. The process was iterated until no further improvement in the model were was achieved.

The correct overall symmetry of the connector region was tested by enforcing different symmetries to models and assessing the results. The tail region of λ, composed primarily of the major tail protein gpV, has six-fold symmetry (REF). The symmetry in the connector region is unknown, but thought to be either six or twelve-fold (REF). Other λ – like phages have had their portal protein structures determined and shown to be dodecamer (REF, REF, REF). In phage biology, it is common for different bacteriophage within the same family (i.e. *siphoviridae*) to share similar characteristics. These similarities include structural organization, sequence identity, and details of virus-host interactions (REF,REF, REF).

Five , maps of different symmetry were produced (FIGURE): X, X, X, X and X. The different models were compared for structural features and how the applied symmetry affected the internal density of the map. When applied to a 3-D map with lower symmetry than is correct, but which is divisible by the correct symmetry (e.g. 3-fold symmetry applied to a 6-fold symmetric map) should produce a 3-D map with the correct

symmetry. The C_3 (you will need to define C3 etc. if you're going to use them) map had visible six-fold symmetry (FIGURE). Slices through the map showed six distinct features in the portal region and the rest of the connector and tail (FIGURE).

Both the C_5 and C_7 3-D map had an overall smooth appearance, with a few features deviating from the initial rotationally symmetric map when viewed from some angles (FIGURE). Slices through these 3-D maps showed ambiguous distributions of features throughout the connector and weak symmetry in the tail region. (FIGURE)

When an applied symmetry is higher than the correct symmetry (e.g. C7, C8, C9, C10, C11 or C12 applied to a C6 map), the final 3-D map will appear smoothened. A map with applied C_{12} symmetry showed faint twelve-fold symmetry in the portal, but was smooth throughout the rest of the connector (FIGURE). Slices through the 3-D map of the connector showed evenly distributed density around the portal without distinct unique subunits (FIGURE).

The C_6 3-D map showed unique structural features in the connector that none of the other maps, except for the C3, had. At the portal area, the map had a handedness that was not applied and that became apparent after 6 rounds of refinement. Slices through the 3-D map revealed distinct internal features in the portal region. A series of slices across the long-axis of the map shows that the six densities of the portal, although always equidistant from one another (FIGURE), had different angular positions about the long axis, due to the handedness of these features in the map.

An atomic model of a hexamer of gpU (REF) was docked into the C_6 map to test how well it fit. The gpU hexamer was randomly placed inside the C_6 map and software was allowed to determine objectively the best fit into the map (FIGURE). This procedure was repeated several times while adjusting the initial position of gpU to demonstrate the consistency of the selected final location of the gpU hexamer in the map.

The C6 symmetrized map was then refined to higher resolution using projection matching software that corrects for the effects of the contrast transfer function (CTF) of the microscope. Multiple rounds of refinement were performed with CTF correction until no change in resolution was observed. Six-fold symmetry was also applied during each round of refinement. The final 3-D map of the connector appeared different than the non-CTF corrected C_6 3-D map. The handedness that had emerged in the portal protein, as described earlier, was even more pronounced in this map (FIGURE). The map suggests that the oligomer of portal protein might be a dodecamer consisting of two staggered rings of six. Alternatively, the portal might be a hexamer of monomers that fit together into a twisted cylinder shape.

4. Hand Determination

The inherent handedness of a 3-D model is lost in transmission EM images because the 3-D structure of the specimen is projected onto the two dimensional plane of the image. This projection effect means one cannot distinguish which features came from the side of the specimen closest to the electron source and which features came from side furthest from the electron source. Consequently, identical projections can be made from both enantiomers of a 3-D structure (REF). To determine the correct hand of the model, the freehand test was employed using tilt pair data. The Freehand test can be applied to a model to test whether it is of the correct handedness or incorrect handedness. It also serves as an objective test of the overall correctness of a 3-D map.

After obtaining tilt pairs with an x-rotation=0, y-rotation=-30° applied to the micoroscope goniometer , the freehand showed a clear minimum near (x-rotation=0, y-rotation=-30), the applied rotation, indicating the hand of the model was correct (FIG). A minimum at (x-rotation=0, y-rotation=+30) would have indicated that the map was of incorrect hand.

5. Subunit Labelling

GpW plays a number of roles in phage λ assembly and function. Once the phage genomic DNA has entered the procapsid, gpW is thought to interact with the DNA to increase resistance to DNAase I (REF). Though some of gpW's function has been previously elucidated (REF), its location as a structural component in phage λ has not been confirmed.

Subunit labeling with a secondary protein has been well established in EM as a reliable method in determining the location and presence of a subunit within a macromolecular complex (REF, REF). In one approach, a large (e.g. 25-70 kDa) protein is genetically attached with a short amino acid linker to the subunit of interest. Maltose binding protein (MBP) has previously been used for this purpose (REF).

To determine the location of gpW, MBP (~ 42 kDa) was attached to the N terminus (ask me about the convention for N terminus/N-terminal) of gpW via a 26 amino acid linker (MBP-gpW). Negatively stained samples of phage containing of MBP-gpW were prepared as described earlier. Images were obained using the charge-coupled device (CCD) detector of the microscope to screen the sample. The CCD images of MBP-gpW were noticeably different in the connector region compared to wt λ (FIGURE). As described previously, micrographs were obtained, digitized and particle images were selected, and subjected to 2-D image analysis
Briefly, particle images were averaged to create a global 2-D average and individual particles images were aligned to the global average. The particles were averaged again and the process iterated until the individual particle images were aligned.

The 2-D average of the connector region in wildtype λ was visually compared to the 2-D average of the connector region of MBP-gpW. An extra density in the MBP-gpW 2-D average was seen at the head-tail junction, where the connector region is located,. The extra density appears on either side of the connector as a circular blob (FIGURE).

To characterize the additional density, a difference map was created by subtracting the 2-D average of wt λ phage from the 2-D average of MBP-gpW λ phage (FIGURE). The difference map shows there is more than one MBP density, but does not unambiguously show the oligomeric state of gpW, which is thought to be around six copies per phage (REF, REF).

DISCUSSION

Viral assembly is a multi-step process determined by a specific sequence of protein-protein interactions and protein-DNA interactions. Tailed double-stranded DNA (dsDNA) bacteriophages package their DNA into pre-formed capaid structure. Translocation of dsDNA occurs through the channel of the portal.

The mechanism of DNA translocation in lambda and other related bacteriophages is not well understood. The portal is a key component of the process. The other components of the connector are though to play a role , to some degree to aid in dsDNA packaging and/or stabilization.

Gp6 = 57.3 kDa, gp15=11.6 kDA, gp16=12.5 kDa

1) How does connector compare to SPP1, P22 and phi29 connector? Why only 6 subunits in the portal? Are they big enough to fit?

   The connector of SPP1 is made up of three components : gp6, gp15 and gp16. Gp6 is the portal protein, is a 13-mer in solution (X-ray) and a 12-mer as part of the connector (cryo-EM) (REF).

2) Implications of a hexameric connector

3) Compare components of lambda connector spp1 and p22. How are they different? The handedness of lambda is different of P22. Lambda's handdedes was checked, p22 wasn't. lambda's is correct.

4) gpW is in the connect of lambda. This makes sense. Follows trend of other phages.

BIBLIOGRAPHY

Adrian, M., J. Dubochet, J. Lepault, and A. W. McDowall. 1984. Cryo-electron microscopy of viruses. Nature (London) 308:32-36

Baker, T. S., and R. H. Cheng. 1996. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. J. Struct. Biol. 116:120-130

Casjens, S.R., Hendrix, R.W. 1974. Locations and amounts of major structural proteins in bacteriophage lambda. J. Mol. Biol. 88:535-545.

Crowther, R. A., L. A. Amos, J. T. Finch, D. J. DeRosier, and A. Klug. 1970. Three dimensional reconstructions of spherical viruses by Fourier synthesis from electron micrographs. Nature (London) 226:421-425

Crowther, R. A., D. J. DeRosier, and A. Klug. 1970. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. Proc. R. Soc. Lond. Ser. A 317:319-340.

Crowther, R. A. 1971. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. Phil. Trans. R. Soc. Lond. Ser. B 261:221-230

Dubochet, J., M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A. W. McDowall, and P. Schultz. 1988. Cryo-electron microscopy of vitrified specimens. Q. Rev. Biophys. 21:129-228

Katsura, I. 1983. Tail assembly and injection, p. 331-346. *In* R. Hendrix, J. Roberts, F.W. Stahl, and R. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Pell, L.G., Kanelis, V., Donaldson, L.W., Howell, P.L., Davidson, A.R. 2009. The phage λ major tail protein structre reveals a common evolution for long-tailed phage and the type VI bacterial secretion system. *Proc Natl Soc Sci U S A* **106**(11), 4160-4165

Lhuillier, S., Gallopin, M., Giluin, B., Brasiles, S., Lancelot, N., Letellier, G., Gilles, M., Dethan, G., Orlova, E.V.,Couprie, J., Tavares, P., Zinn-Justin, S. 2009. Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc Natl Soc Sci U S A* **106**(21), 8507 - 8512.

Dube, P., Tavares, P., Lurz, R., van Heel, M. 1993. The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. *EMBO J* **12**(4), 1303 - 1309.

Orlova, E.V., Gowen, B., Droge, A., Stiege, A., Weise, F., Lurz, R., van Heel, M., Tavares, P. 2003. Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. *EMBO J* **22**(6), 1255 - 1262.

Amos, L.A.