Structural analysis of the spiroplasma virus, SpV4: implications for evolutionary variation to obtain host diversity among the *Microviridae*

Paul R Chipman[1†], Mavis Agbandje-McKenna[2], Joël Renaudin[3], Timothy S Baker[1] and Robert McKenna[2*]

Background: Spiroplasma virus, SpV4, is a small, non-enveloped virus that infects the helical mollicute *Spiroplasma melliferum*. SpV4 exhibits several similarities to the *Chlamydia* phage, Chp1, and the *Coliphages* α3, ϕK, G4 and ϕX174. All of these viruses are members of the *Microviridae*. These viruses have isometric capsids with T = 1 icosahedral symmetry, cause lytic infections and are the only icosahedral phages that contain single-stranded circular DNA genomes. The aim of this comparative study on these phages was to understand the role of their capsid proteins during host receptor recognition.

Results: The three-dimensional structure of SpV4 was determined to 27 Å resolution from images of frozen-hydrated particles. Cryo-electron microscopy (cryo-EM) revealed 20, ~54 Å long, 'mushroom-like' protrusions on the surface of the capsid. Each protrusion comprises a trimeric structure that extends radially along the threefold icosahedral axes of the capsid. A 71 amino acid portion of VP1 (the SpV4 capsid protein) was shown, by structural alignment with the atomic structure of the F capsid protein of ϕX174, to represent an insertion sequence between the E and F strands of the eight-stranded antiparallel β-barrel. Secondary structure prediction of this insertion sequence provided the basis for a probable structural motif, consisting of a six-stranded antiparallel β sheet connected by small turns. Three such motifs form the rigid stable trimeric structures (mushroom-like protrusions) at the threefold axes, with hydrophobic depressions at their distal surface.

Conclusions: Sequence alignment and structural analysis indicate that distinct genera of the *Microviridae* might have evolved from a common primordial ancestor, with capsid surface variations, such as the SpV4 protrusions, resulting from gene fusion events that have enabled diverse host ranges. The hydrophobic nature of the cavity at the distal surface of the SpV4 protrusions suggests that this region may function as the receptor-recognition site during host infection.

Addresses: [1]Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1392, USA, [2]Department of Biological Sciences, University of Warwick, Coventry CV8 7AL, UK and [3]Laboratoire de Biologie Cellulaire et Moléculaire, INRA et Université de Bordeaux II, 33883 Villenave d'Ornon Cedex, France.

[†]Present address: Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue, Seattle, Washington 98109, USA.

*Corresponding author.
E-mail: mckenna@globin.bio.warwick.ac.uk

Key words: cryo-electron microscopy, icosahedral capsid, sequence alignment, spiroplasma virus, SpV4

Received: **22 October 1997**
Revisions requested: **12 November 1997**
Revisions received: **27 November 1997**
Accepted: **2 December 1997**

Structure 15 February 1998, **6**:135–145
http://biomednet.com/elecref/0969212600600135

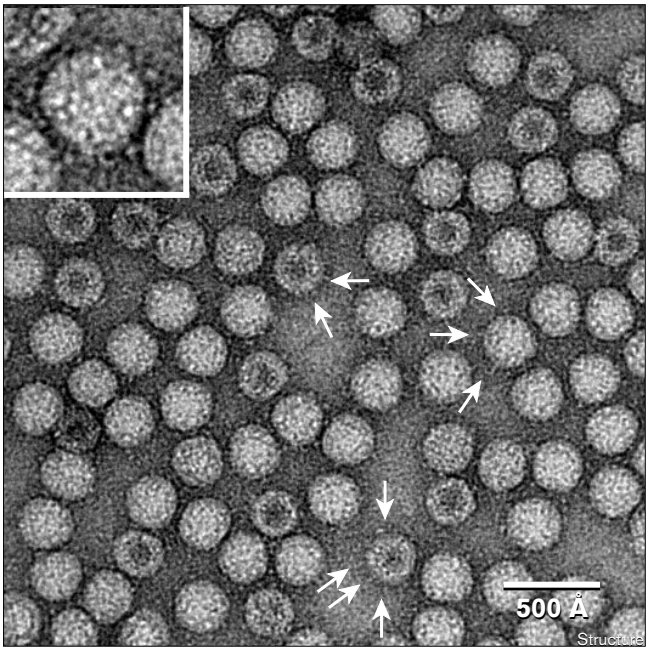
Introduction

Spiroplasma virus, SpV4, is a small, spherical virus that infects the *Spiroplasma melliferum*. This non-enveloped, single-stranded DNA virus was discovered in 1980 in strain B63 of *S. melliferum* cultured from a honey bee in Morocco [1] and has since been propagated in *S. melliferum* strain G1 [2,3]. Spiroplasmas are helical, wall-free eubacteria belonging to the class Mollicutes, a group of organisms derived by regressive evolution from ancestors of gram-positive bacteria with low guanine-plus-cytosine DNA [4,5].

SpV4 shares several properties with the *Chlamydia* phage Chp1 and the *Coliphages* α3, ϕK, G4 and ϕX174. All these viruses are members of the *Microviridae* family; they are lytic and contain single-stranded circular DNA genomes of 4421, 4877, 6087, 6081, 5386 and 5577 nucleotides for SpV4, Chp1, α3, ϕK, G4 and ϕX174, respectively [6–10]. The genomes encode similar numbers of proteins: nine for SpV4 and eleven for Chp1, α3, ϕK, ϕX174 and G4. The SpV4 and Chp1 capsids, however, only contain 60 copies of a single protein, VP1, whereas the capsids of α3, ϕK, G4 and ϕX174 consist of 60 copies each of the major capsid protein, F, the major spike protein G, the DNA packaging protein, J, and 12 copies of the pilot protein, H.

SpV4 virions have a molecular weight of 5.4×10^6 and a density of 1.40 g/cm^3 in cesium chloride. VP1 is a 63.9 kDa protein that forms a T = 1 capsid [11] with a mean diameter of ~270 Å (Figure 1) [12], which is roughly

Figure 1



SpV4 particles negatively stained with 1% uranyl acetate on a carbon support film. Arrows highlight protrusions on three particles. The inset, a close-up view of one particle, shows several, stain-excluding (white) 'dots' inside the particle. The dots are likely to be end-on views of the protrusions. The scale bar = 500 Å.

equivalent to a ϕX174 capsid consisting of only the major capsid protein F, not including the G, H or J proteins [13,14].

The structures of ϕX174 and G4 virions have been determined by X-ray crystallography to 3 Å resolution [14,15]. These studies revealed the structure and capsid organization of the F, G, and J proteins in atomic detail. Each ϕX174 or G4 capsid contains 60 copies of the F protein arranged with T = 1 icosahedral symmetry [11]. The F protein has the β barrel tertiary structure common to many of the structural proteins in small, spherical viruses [16]. This canonical, eight β-strand motif is oriented with the strands aligned roughly tangential to the capsid surface and with the B–C, D–E, F–G and H–I corners (where the letters refer to the individual strands of the β barrel) pointed towards the fivefold axes (Figure 2). The β barrel projects into the interior of the virus, where there is little contact between adjacent, symmetry-related β barrels. Two large insertion loops that comprise 65% of the F protein polypeptide chain occur at the corners of the E–F and H–I strands, these loops are 163 and 112 residues in length, respectively. The two insertion loops form the exterior surface topology of the capsid and are responsible for most of the contacts between neighboring F protein subunits (Figure 2).

Comparison of the sequences of the structural proteins of four *Coliphage* members of the *Microviridae* family, α3, ϕK, G4, and ϕX174, showed that the capsid protein F is the most highly conserved structural protein [17]. The VP1 capsid proteins in SpV4 and Chp1 contain 553 and 595 amino acids, respectively [6,7], and are homologous to the 426 amino acid F proteins of ϕX174 and G4 [8–10].

The SpV4 and Chp1 genomes do not encode a major spike protein like the G protein of the *Coliphage* virions. Pentamers of the G protein in the *Coliphages* form 12, characteristic, star-shaped 'spikes' (~70 Å in diameter and extending ~30 Å above the viral capsid surface) situated at the fivefold vertices of the icosahedron [13–15]. Previously, the SpV4 capsid was postulated to be devoid of such spikes or other projections [12], and hence it was believed that the SpV4 morphology might bear more resemblance to the smoother, spherical capsids of the *Parvoviridae* [18–20]. The characteristic *Coliphage* spikes can be removed from capsids with 4 M urea, leaving an intact capsid formed by the F protein [21]. These spikeless particles were thought to be similar in structure to SpV4 or Chp1, the genomes of which do not encode the G, H and J structural proteins found in the *Coliphages*. Thus, the known atomic structures of the ϕX174 and G4 F capsid proteins should serve as excellent models of the VP1 capsid protein of SpV4 and Chp1.

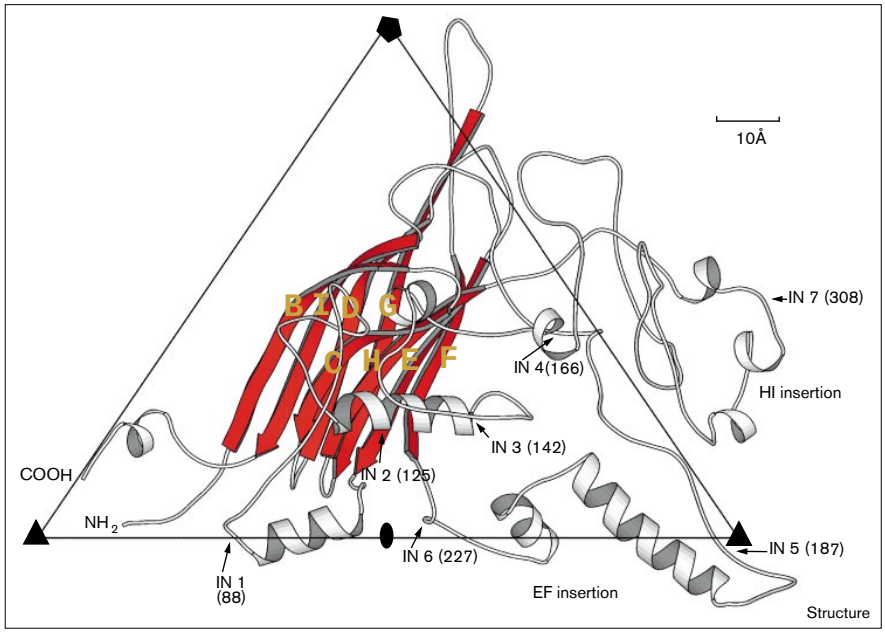
We have used cryo-electron microscopy (cryo-EM) and three-dimensional image reconstruction methods to determine the structure of frozen-hydrated SpV4 particles to 27 Å resolution. We also report an amino acid sequence alignment of six members of the *Microviridae* family which made possible the fit of the atomic coordinates of the F protein of ϕX174 into the cryo-EM density map of SpV4. Finally, modeling of three, 71 amino acid loops into the cryo-EM density for the 'mushroom-like' protrusions at the threefold vertices, and the hydrophobic nature of the cavity at the distal surface of these protrusions, suggest a function as putative receptor recognition sites for SpV4.

Results and discussion

Cryo-EM and three-dimensional image reconstruction

SpV4 for use in the cryo-EM studies was propagated and purified following an established protocol [3] with minor modifications to prevent precipitation of the virus (see Materials and methods section). Cursory inspection of transmission electron micrographs of negatively stained (Figure 1) or vitrified (Figure 3) samples of SpV4 revealed particles with circular profiles, consistent with a spherical structure of diameter ~270 Å. However, when the three-dimensional structure of SpV4 was determined to 27 Å resolution based on reconstruction analysis of 22 images of vitrified particles, it was immediately apparent that virions contain a prominent protrusion at each of the 20 threefold axes of symmetry (Figure 4). Careful reinspection of the

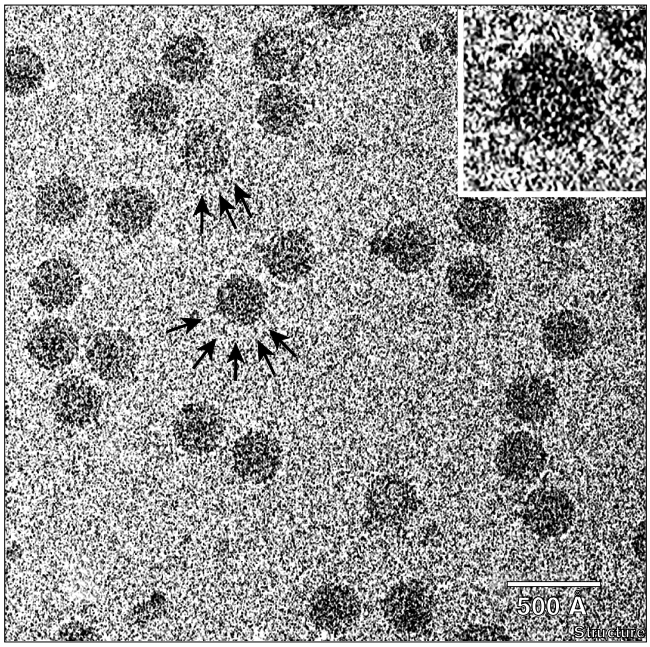
Figure 2



Ribbon drawing of the atomic structure of the ϕX174 capsid protein F. The β-barrel motif is coloured red with the β strands (B, I, D, G, and C, H, E, F) labeled according to standard convention. Arrows highlight the equivalent positions of the capsid protein VP1 of SpV4, where insertion loops (identified as IN1–7) occur relative to the capsid protein F of ϕX174 (refer to Figure 6). Residue numbering is given for the F protein of ϕX174. The orientation of the F protein is shown viewed towards the interior of the virus down a twofold axis. An icosahedral asymmetric unit (large open triangle) includes the region bounded by a fivefold axis (filled pentagon) and two adjacent threefold axes (filled triangles). (Figure produced with MOLSCRIPT [53].)

images of stained (Figure 1) and unstained SpV4 particles (Figure 3) revealed features consistent with the presence of such external protrusions. In stain, most particles have small projections or knob-like features at their periphery (Figure 1; arrows) as well as stain-excluding knobs within the particle (e.g. Figure 1 inset). These latter features correspond to end-on views of the protruding structures. Similar features at the particle periphery can sometimes be discerned in the noisy images of unstained SpV4 particles (Figure 3; arrows and inset).

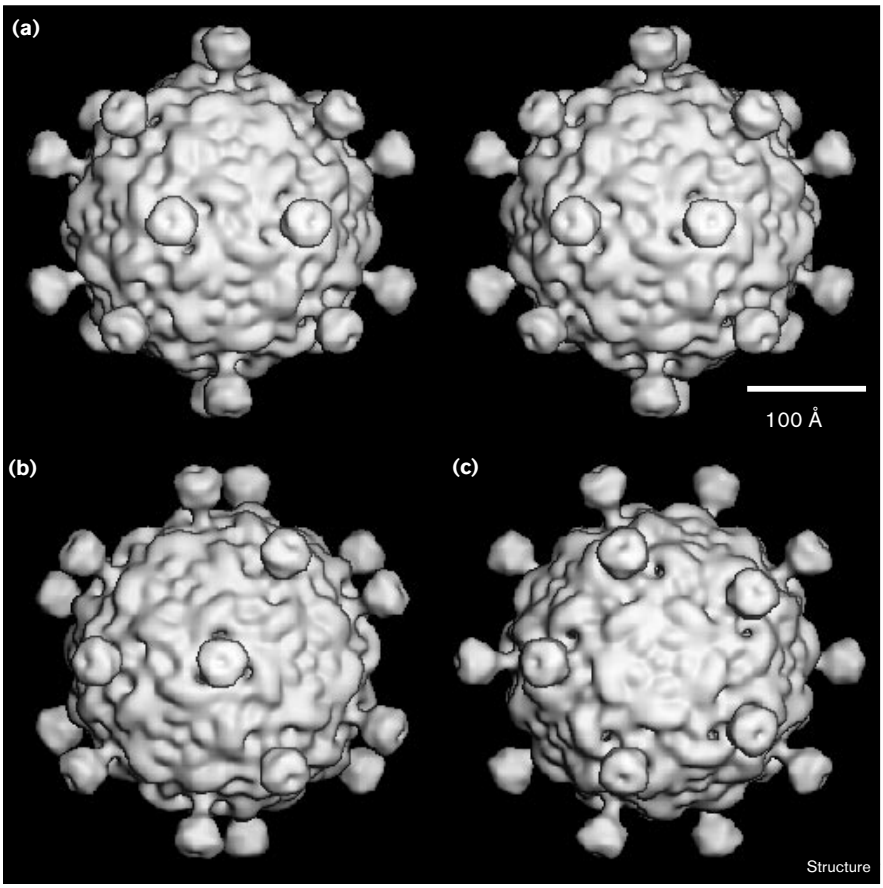
Figure 3



Unstained SpV4 particles suspended in a layer of vitreous ice over holes in a carbon support film. Arrows highlight several protrusions on two particles. The inset shows an enlarged view of one of the highlighted particles. The scale bar = 500 Å.

The prominent protrusions project radially outwards from the capsid, the surface of which has a 'rough' texture (Figure 4). Large, 'sausage-like' ridges encircle five smaller 'bumps' and an even smaller bump at each fivefold axis. In addition, three small cavities surround each protrusion near its base. Close-up views reveal finer details of the protrusion structure (Figure 5). As depicted in Figure 5, the mushroom-like protrusion extends ~54 Å above the capsid surface and consists of a globular head ('bud') of approximate dimensions $41 \times 41 \times 29$ Å and a neck ('stalk') that is ~25 Å long and ~13 Å in diameter. The cavities at the base of the protrusion are channels (~10 Å narrowest diameter) that appear to provide solvent accessibility to the virus interior (Figures 5a,b). These putative channels persist in isosurface rendered views even at much lower density threshold levels. A cross-section through the longitudinal axis of the protrusion reveals a dimple at the outermost tip of the protrusion (Figure 5c; compare with top view in Figure 5a) in the isosurface view. A more realistic rendering of the density distribution in the longitudinal section (Figures 5d,e) reveals a 'cup-like' as opposed to solid structure for the bud domain. At 27 Å resolution, however, it is impossible to determine whether the low density central region is a

Figure 4



Shaded surface representations of the SpV4 three-dimensional reconstruction. The representations are viewed down the **(a)** twofold (in stereo), **(b)** threefold and **(c)** fivefold axes. The reconstruction was computed to 27 Å resolution from 22 different SpV4 particle images. The scale bar = 100 Å.

solvent-filled cavity of approximate dimensions 10 Å diameter and 17 Å deep. Alternatively, the average protein density may simply be very low in this region.

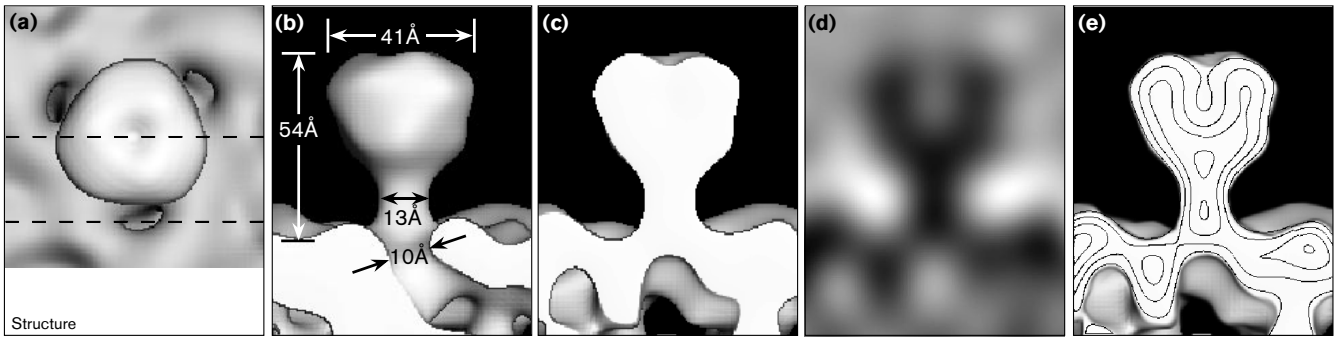
The protrusion appears to be a fairly rigid structure because its densities in the icosahedrally averaged three-dimensional map are in the same range as those attributed to the capsid shell. Our model building experiments (described later) are consistent with this observation.

Alignment of *Microviridae* capsid proteins

Members of the *Microviridae*, including the *Coliphages* α3, ϕK, G4 and ϕX174, have been shown to have the same gene order and considerable sequence similarities [22]. An amino acid sequence alignment of the F, G, and J structural proteins [17], together with knowledge of the atomic structures of the ϕX174 [14] and G4 [15] capsids, showed the F protein to be the most conserved structural protein. Among all four *Coliphages*, the F proteins showed 63% amino acid sequence identity, as compared to 19% identity for the G spike proteins. Presumably, functions of the F protein, such as involvement in capsid assembly and genome protection, are indispensable enough for evolutionary pressure to conserve them. The sequence alignment and crystallography studies also showed that the F protein residues involved in intersubunit interactions were more conserved than those in other regions. In addition, the residues at a surface depression in the ϕX174 capsid that runs from a position near the threefold axis to the twofold axis were almost completely conserved [17]. Part of this depression in ϕX174 forms a putative carbohydrate-binding site [23].

The amino acid sequences of the SpV4 and Chp1 VP1 proteins and of the F proteins of α3, ϕK, G4 and ϕX174, were aligned with the use of the program PILEUP (part of the University of Wisconsin Genetics Computer Group [GCG] package). To maximize the alignment agreement of the capsid protein sequences, seven insertion loops were introduced into the SpV4 and Chp1 sequences in relation to the *Coliphages* (Figures 6a,b). The criteria for scoring insertions required a difference in sequence length of six or more amino acids. By reference to the atomic structure of the ϕX174 F protein, six insertions were located in the E–F turn of the β barrel and the other insertion was in the H–I turn. Two of the E–F

Figure 5



Close-up views of one SpV4 trimeric protrusion. **(a)** Shaded surface view from above the top of the protrusion. **(b)** Side view, as from the bottom of (a), with the front of the map (as indicated by the dotted line towards the bottom of (a)) removed to reveal a channel near the base of the stalk. **(c)** The same view as (b) but with more density removed, to dashed line in center of (a), to reveal a cross-section through the center of the protrusion. **(d)** Density distribution in the center of the protrusion at the same plane where the map was cut in (c). Highest density features (protein) appear black, whereas lowest densities (solvent) appear white or gray. This representation gives a more realistic rendering of the density fluctuations inside the protrusion and, for example, shows a low density along the vertical axis at the top of the protrusion. **(e)** The view is the same as in (c) but with four density contour levels shown to illustrate the varying density in the cross-section.

insertions would occur inside the capsid (i.e. at low radius) and all other insertions would occur on the exterior surface of the capsid (Figures 2 and 6a,b). These sequence alignment results suggest that the tertiary structure of VP1 may be quite similar to that for the F protein but with the largest differences occurring at the outer and inner surfaces of the capsid attributed to the VP1 insertions. The longest insertion, IN5, occurs between Thr187 and Thr188 of the F protein of ϕX174 (Figures 2 and 6), and is 71 and 104 residues in SpV4 and Chp1, respectively. This insertion point occurs in ϕX174 within the portion of the E–F loop that forms a small bump at the icosahedral threefold axes (Figure 7). On this basis, it is possible to ascribe the 71-residue IN5 insertion of SpV4 to the prominent mushroom-like protrusions at the threefold positions (Figures 4 and 5).

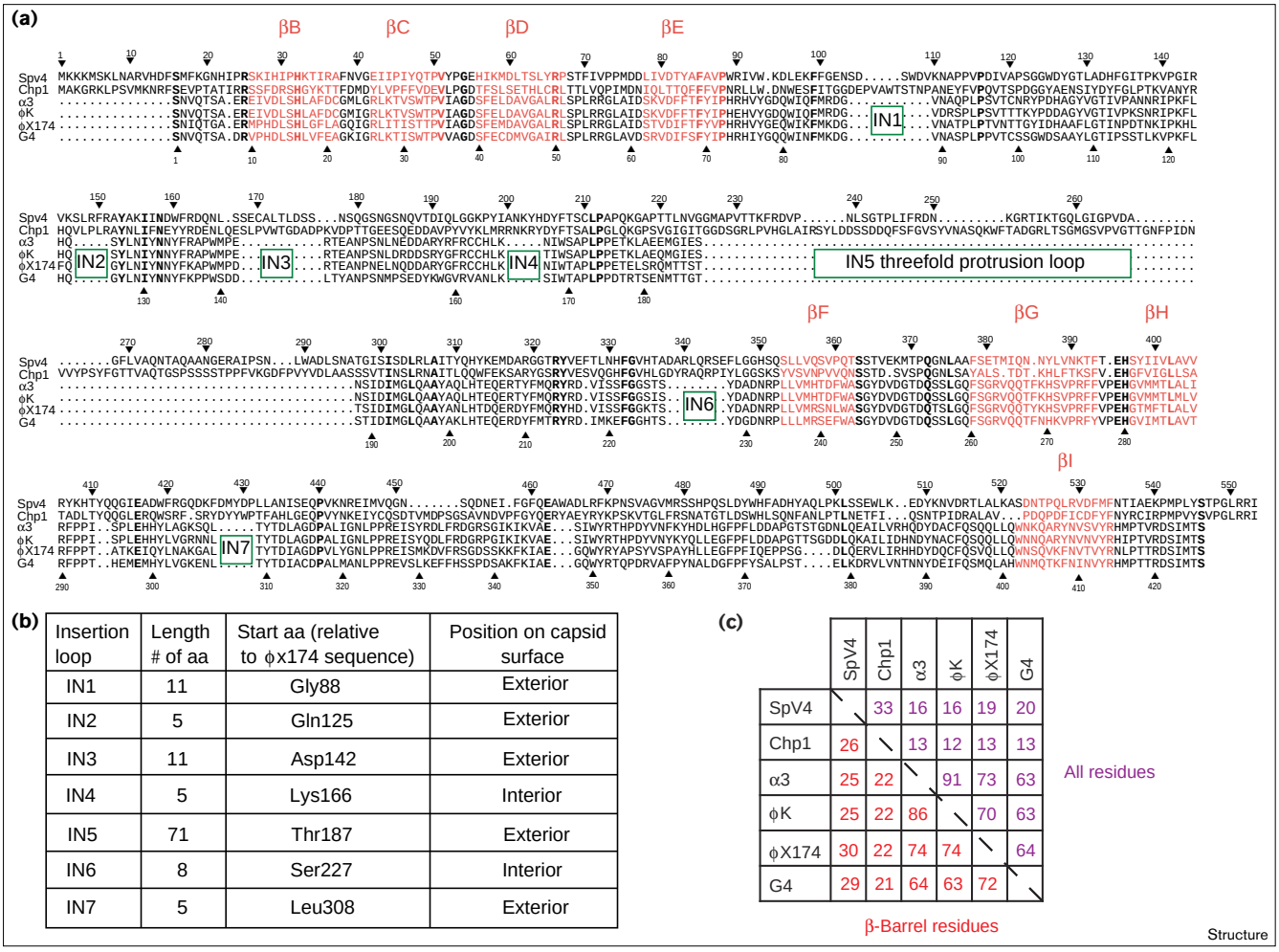
Only 33 (7.7%) of the 426 amino acid residues of the ϕX174 F protein were identical in all six phages, whereas 268 (63%) of these residues were identical among the four *Coliphage* members and 142 (33%) were identical between SpV4 and Chp1. The percentage of identical residues for each pairwise comparison of sequences was calculated not only for all residues but also for just those residues of the structurally conserved β-barrel motif identified from the X-ray crystal structures of ϕX174 and G4 [14,15] (Figure 6c). The residues within the β-barrel motif were observed to be significantly greater in identity in every pair sequence comparison between 'weakly' related phages (identity less than 21% for all residues). For example, when comparing the F protein of ϕX174 with VP1 of SpV4, for all residues the sequence identity is 19%, whereas for residues contained only in the β-barrel motif the percentage identity increases to 30% (Figure 6c). This same general trend holds (with one exception, α3 and ϕK) for 'strongly' related sequences (identity greater than 50% for all residues), although the increase in identity is only slight.

These pairwise sequence comparisons suggest that an evolutionary pressure exists to conserve 'critical' residues within the β-barrel motif while the surface insertion loops can accommodate a high mutation rate. Even though there is no obviously conserved specific sequence in the major capsid proteins of the related phages, close inspection of the sequence comparisons (Figure 6a) reveals a tendency for hydrophobic residues in the β-barrel motif to be conserved. Such conservation is consistent with the preservation of a β-barrel structure that is the structural core of each capsid. Rapid evolution within the large insertion loops that decorate this core structure ensures diversity in the exposed surface of each phage.

A phylogenetic tree for the six bacteriophages (Figure 8) was calculated from the data shown in Figure 6c by means of an unweighted pairwise distance matrix method [24]. As expected, the *Coliphages* cluster closely together, with α3 and ϕK being most closely related (91% sequence identity) and G4 being the most distantly related member of the *Coliphages* (~63% sequence identity to all three other *Coliphages*). Chp1 and SpV4 cluster together, though they are still quite distant (only 33% sequence identity between VP1s), and they are even more distant from the *Coliphages* (15% mean sequence identity).

When pairs of structurally similar protein chains were taken from the CATH structural classification database [25], two distributions of sequence identities were

Figure 6



Insertion loop	Length # of aa	Start aa (relative to ϕx174 sequence)	Position on capsid surface
IN1	11	Gly88	Exterior
IN2	5	Gln125	Exterior
IN3	11	Asp142	Exterior
IN4	5	Lys166	Interior
IN5	71	Thr187	Exterior
IN6	8	Ser227	Interior
IN7	5	Leu308	Exterior

	SpV4	Chp1	α3	ϕK	ϕX174	G4
SpV4		33	16	16	19	20
Chp1	26		13	12	13	13
α3	25	22		91	73	63
ϕK	25	22	86		70	63
ϕX174	30	22	74	74		64
G4	29	21	64	63	72	

Comparison of the amino acid sequences of capsid proteins of members of the *Microviridae* family. **(a)** Sequence alignment of VP1 of SpV4 and Chp1 and protein F for the *Coliphages* α3, ϕK, ϕX174 and G4. Bold letters indicate residues that are completely conserved in all the phages compared; residues that form part of the β-barrel structure motif elements are colored red and labeled (as shown in Figure 2). The sequence numbers are given for SpV4 (top) and ϕX174 (bottom). Insertions are labeled and marked by green boxes. **(b)** Assignment and location of the seven insertion loops (assigned as IN1–7) found in the SpV4 and Chp1 sequences in relation to the *Coliphages*. **(c)** The percentage of identical amino acids between the aligned capsid proteins, for all residues (shown in purple) and those involved in the β-barrel motif (shown in red).

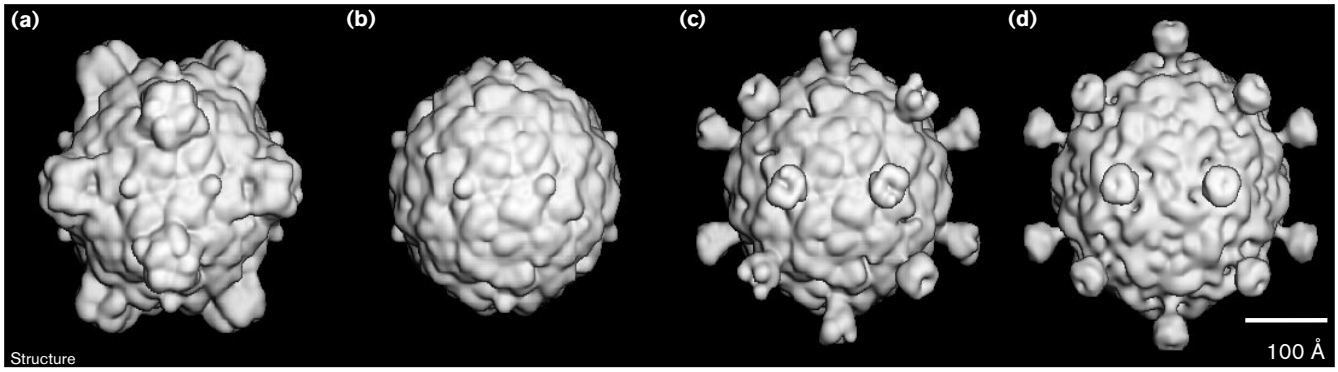
observed. Those sequences which shared a common ancestry had a mean percentage sequence identity of 15%, and those which had a mean percentage sequence identity less than 10% did not [26]. These observations indicate that, although the VP1s of Chp1 and SpV4 are genetically distinct from the F proteins of the *Coliphages*, there is likely to be a common ancestral link among these proteins.

Modeling the IN5 loop and fit to the cryo-EM density map

The secondary structure of the 71-residue IN5 loop (Figure 6a), that was ascribed to the mushroom-like projections, was predicted with the aid of the PredictProtein Server [27] to be approximately 60% β strand and 40% coil. No α-helical segments were predicted by the program.

A three-dimensional model was built for the 71-residue SpV4 loop that satisfied both the envelope of the cryo-EM reconstruction (Figures 4 and 5) and the secondary structure prediction. The modeling incorporated three, symmetry-related, six-stranded (each strand being six residues long) antiparallel β-sheet motifs connected by small turns of five residues in length (Figure 9). This structure fit extremely well into the distal globular bud domain ($41 \times 41 \times 29$ Å) of the trimeric protrusion (Figure 10). This simple model also allowed for intersubunit hydrogen-bonding interactions between antiparallel β strands from the symmetry-related motifs. The lengths of the β strands numbered 1 and 6 (Figure 9) were extended to 12 residues each, to form a stalk that allows the loop to insert

Figure 7

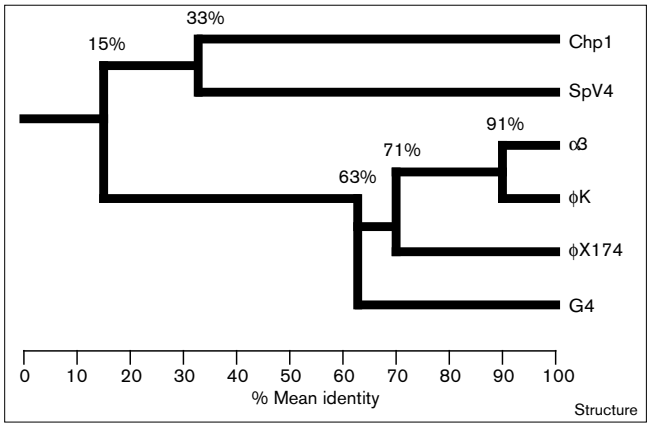


Shaded surface representation comparisons of ϕX174 and SpV4 structures. **(a)** The atomic ϕX174 structure [14]. **(b)** The ϕX174 structure, as in (a), but with the prominent spikes (G protein pentamers) removed to reveal just the F capsid. **(c)** An SpV4–ϕX174 hybrid model, formed by combining the ϕX174 F capsid model (b) with the pseudo-atomic model of the SpV4 protrusion. **(d)** The cryo-EM reconstruction of SpV4. All structures are shown at ~27 Å resolution and are viewed along a twofold axis of symmetry. The scale bar = 100 Å.

at the Thr187–Thr188 position (ϕX174 sequence designation). In the model, the stalk region consists of the two extended antiparallel strands, which form a β-hexamer structure with the symmetry-related strands. This proposed β-hexamer structure is very similar to that observed on the quasi-sixfold axes of symmetry in cowpea chlorotic mottle virus [28]. The stalk is ~25 Å long and 13 Å wide and connects the globular bud of the protrusion to the capsid (Figure 10). As modeled, the protruding loop has 48 residues in β-strand structure and 23 residues in loop structures. The motif is a very rigid and stable structure, which is consistent with the observation of the protrusion as a high-density feature in the three-dimensional image reconstruction. This model also includes a 17 Å deep depression at the distal surface of the bud on the threefold axes (Figure 10). Of the 71 amino acids in the loop, 30 are hydrophobic and of the 11 residues, Val237-Pro-Asn-Leu-Ser-Gly-Thr-Pro-Leu-Ile-Phe247, lining the surface of the 17 Å depression, seven are hydrophobic. The model, therefore, predicts the presence of a mainly hydrophobic cavity at the distal surface of the protrusion.

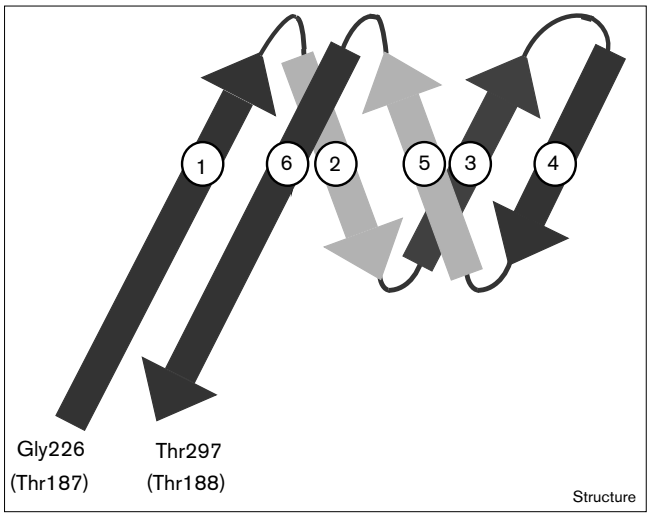
This model is the simplest that satisfies both the cryo-EM density of the protrusion at the threefold axis and the predicted secondary structure profile of the amino acid

Figure 8



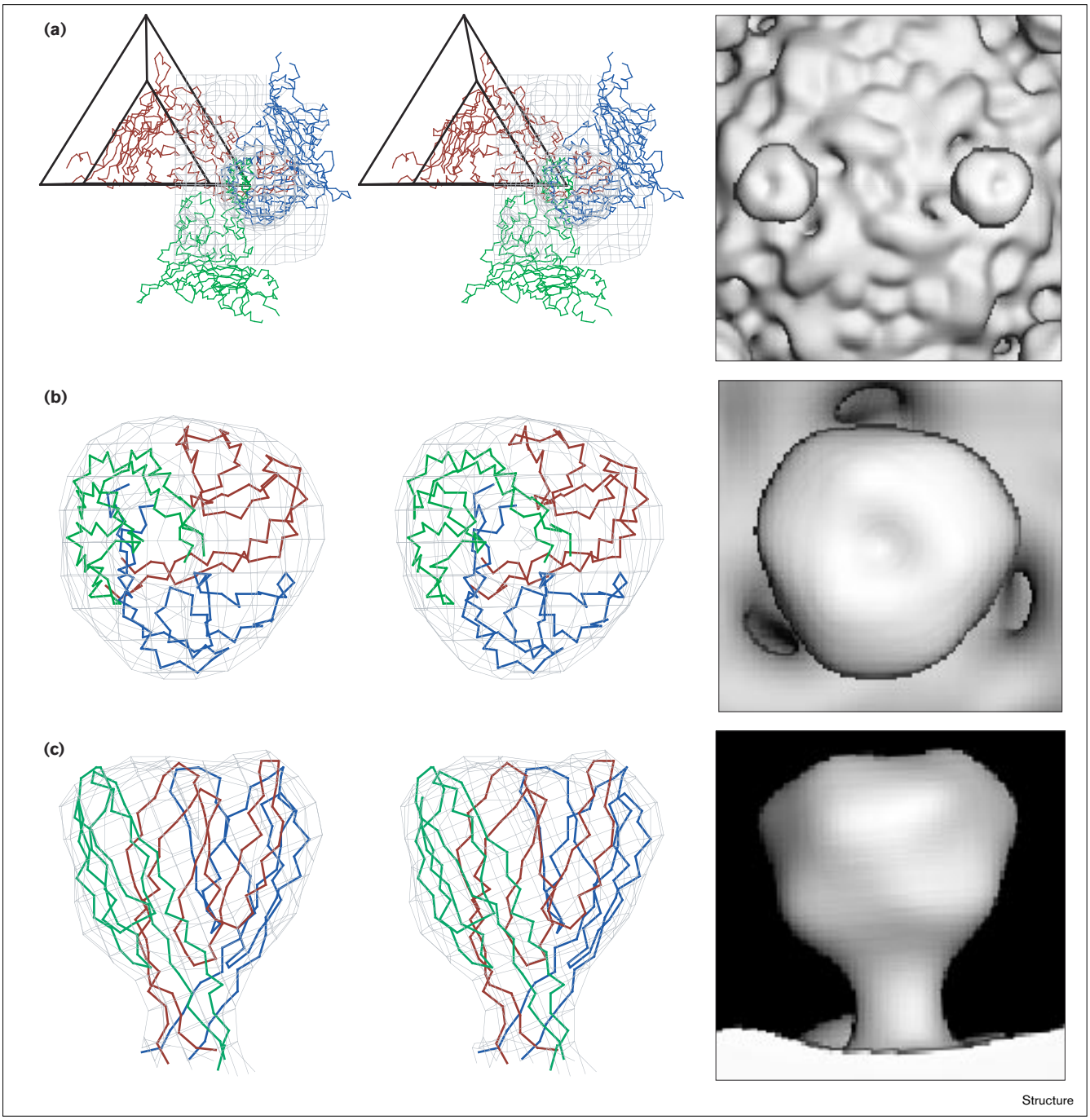
Phylogenetic tree for six *Microviridae* capsid proteins.

Figure 9



Proposed structural motif model for residues Gly226–Thr297 of the capsid protein of SpV4. These residues comprise one third of the protrusion at each threefold axis. The residues in the F protein of ϕX174 at the exit and entry points for the SpV4 loop are indicated in parentheses; β strands are shown as arrows.

Figure 10



Stereo views of the atomic model of the SpV4 protrusion. **(a)** The fit of a pseudo-atomic model of an SpV4 VP1 trimer (separate monomers colored blue, green and red) into the cryo-EM reconstruction (gray isodensity contour). Shown also is the viral asymmetric unit, depicted as an open triangle. The view is along a twofold axis. The fit of three symmetry-related SpV4 loop structures (Gly226–Thr297) into the cryo-EM density map, as viewed from **(b)** the top, along a threefold axis or **(c)** the side, perpendicular to a threefold axis. Corresponding shaded surface representations are shown to the right of each stereo view. (The stereo views were produced with MacInPlot [54].)

sequence of the IN5 insertion loop of SpV4. Obviously other more complex models may also satisfy the envelope of the cryo-EM density, but would not satisfy the secondary structure prediction profile or the formation of a hydrophobic cavity at the distal surface of the protrusion. Other structural motif insertions have been observed between strands

of the eight-stranded antiparallel β barrel of viruses. For example, an immunoglobulin (Ig) fold domain insertion occurs between the E and F strands of the insect virus nudaurelia ω capensis virus [29], but this motif would be too large for the protrusion in SpV4.

Conclusions

The multisequence alignments for the major capsid proteins of the *Microviridae* phages demonstrate that the VP1 proteins of Chp1 and SpV4 most likely have a common ancestral link to the F proteins of the *Coliphages*. The three-dimensional cryo-EM reconstruction of SpV4 revealed a mushroom-like protrusion on the threefold axes of the icosahedral capsid. The sequence alignment results, along with knowledge of the atomic structures of the F proteins for ϕX174 and G4, lead to the assignment of a 71-residue loop in the SpV4 VP1 sequence (Gly226 to Thr297) to one third of the trimeric protrusion structure. The secondary structure predicted from the sequence of the loop suggested a peptide fold, the main part of which consists of a six-stranded, antiparallel β sheet, that trimerizes and fits well into the globular domain (bud) of the protrusion. The distal globular domain of the trimeric hemagglutinin spike of influenza virus has a similar structure, although hemagglutinin is much larger (~145 residues) and consists of an eight-stranded, antiparallel β-sheet structure with larger extended loops between the strands [30]. A fragment sequence search of the 71-residue loop was performed, using the BLAST network service, and sequences of influenza hemagglutinin B were among the highest scoring segment pairs [31].

The predicted hydrophobic character of the depression at the distal tip of the SpV4 bud suggests a possible role in cellular receptor binding. The human parvovirus, B19, has a depression at each threefold axis that binds globoside, the cellular receptor [32]. Unlike SpV4, however, B19 does not have long protrusions.

Biological implications

The bacteriophage spiroplasma virus, SpV4, which infects the wall-free prokaryote *Spiroplasma melliferum*, is a member of the *Microviridae* family. Other members of this family include the *Chlamydia* phage, Chp1, and *Coliphages*, such as ϕX174. The surface of a virion is one factor that helps control host tropism, because it presents the required specific interactions for host receptor recognition which leads to infection. A detailed description of the SpV4 surface, and comparison with related phages, may help us to further understand the mechanisms underlying host cell recognition and specificity.

We report here the three-dimensional structure of SpV4. We present a detailed description of the SpV4 capsid structure based on a three-dimensional cryo-electron microscopy (cryo-EM) image reconstruction of SpV4 samples and structural sequence alignment with the atomic model of the major capsid protein (protein F) of the related bacteriophage, ϕX174.

A striking feature of the SpV4 structure, which was not previously seen, is the presence of 20, 'mushroom-like' protrusions (~54 Å long) that extend radially outwards from the capsid surface along the icosahedral threefold axes. The amino acid sequence that folds to form this protrusion was identified as a 71-residue insertion in the SpV4 coat protein, VP1, relative to the ϕX174 F capsid protein. Sequence alignments of the F capsid proteins of the *Coliphages*, and the VP1 capsid proteins of Chp1 and SpV4, exhibit an evolutionary relationship. Distinct differences in the surface morphologies of the *Microviridae* phages are attributable mainly to large variations in protruding structural motifs. The *Coliphages* have a specialized spike structure (pentamers of the G protein) at each of the 12 fivefold vertices. The G and F proteins both affect host range in *Escherichia coli* strains [33–35]. Chp1 and SpV4 have no gene for a G-like protein, but instead have an insertion loop (IN5), between the E and F strands of the capsid protein β barrel; these insertions associate at the threefold axes to form 20 protrusions. These protrusions may be utilized as host range determinants in an analogous manner to the G protein spikes of the *Coliphages*.

Phages of the *Microviridae* family probably evolved from a common primordial phage gene. Distinct genera might have then evolved by means of genetic fusion events that enabled a diverse range of hosts to be infected. It is probable that *Coliphages* incorporated a gene, G, to encode the large pentameric spike structure which enables the phages to recognize the outer cell membrane of *E. coli*. Likewise, Chp1 and SpV4 incorporated a gene that encodes a trimeric structural motif that enables these phage to recognize the chlamydial envelope or the wall-free spiroplasmas, respectively.

Materials and methods

Propagation and purification of SpV4

S. melliferum strain G1 [36] was grown in SP4 medium [37] at 32°C, and used for the propagation of SpV4 [3]. SpV4-infected Spiroplasma broth cultures were prepared by inoculation of an early log phase culture of *S. melliferum* (10^8–10^9 cfu/ml) with an SpV4 inoculum at a multiplicity of infection of approximately 0.2.

For virus purification, SpV4-infected cultures (6 × 150 ml) were harvested 24h after inoculation. Spiroplasma cells were pelleted by centrifugation at 22,000 × g for 45 min. The supernatant was adjusted to 0.5 M NaCl and the virions were precipitated by adding 8% polyethylene glycol 6000. After overnight incubation at 4°C, the precipitate was collected by centrifugation (22,000 × g for 45 min), and the pellets were suspended in 50 mM sodium tetraborate buffer at pH 9.2 containing 0.5 M NaCl. The suspension was treated with 10 μg/ml DNAse in the presence of 25 mM $MgCl_2$ for 1 h at 37°C, with 0.2% Triton X100 for an additional 1 h, and then clarified three times by vigorous shaking with an equal volume of chloroform. The

virions were sedimented from the final aqueous phase by centrifugation (130,000 × g for 7 h in a Beckman SW27 rotor) through a 30% sucrose cushion (5 ml). The pellets were resuspended in 50 mM sodium tetraborate buffer at pH 9.2 containing 0.6 g/ml CsCl and centrifuged to equilibrium at 150,000 × g for 48 h. The fraction containing virions was collected and dialyzed against 40 mM sodium tetraborate buffer at pH 9.2. The virus concentration was adjusted to 1 mg/ml assuming an extinction coefficient of 7.5 mg/ml per cm at 260 nm. This protocol yielded between 1.0 to 1.5 mg of purified virions per liter of SpV4 infected culture.

Cryo-EM and three-dimensional image reconstruction

Small aliquots (3.5 μl) of SpV4 sample were placed on holey carbon grids, blotted with filter paper, and plunged into liquid ethane to suspend the SpV4 particles in a thin layer of vitreous ice as described [38–40]. Grids were inserted into a precooled Gatan 626 cryotransfer holder (Gatan Inc., Warrendale, PA, USA) that maintained a constant temperature of −175°C. The sample was examined in a Philips EM420 transmission electron microscope (Philips Electronics Instruments, Mahwah, NJ, USA) and images were recorded on Kodak SO-163 film (Eastman Kodak Company, Rochester, NY, USA) under low dose conditions (~18 $e^-/Å^2$) at a nominal magnification of × 49,000 and a defocus value of ~0.8 μm. A micrograph that displayed uniform specimen ice thickness and with minimal drift and astigmatism, as determined by eye, was digitized at 25 μm intervals (5.1 Å step size) and processed as described [40]. The orientations and origins of different particles imaged in the micrograph were determined and a subset of 22 SpV4 images were combined to compute a three-dimensional reconstruction which was used for subsequent model-based refinement of the view angle and origin parameters [41]. The resolution of the final reconstructed density map was determined to be at least 27 Å as measured by structure-factor comparisons [42] and Fourier ring correlation measurements [43].

All computations were performed with FORTRAN programs [41, 44–46] on VAX/VMS 8550 and AlphaStation 500/400 computers (Digital Equipment Corp., Maynard, MA, USA). Digitized images and shaded-surface representations were viewed on a raster graphics device (Model 3400; Lexidata Corp., Billerica, MA, USA) and final figures were prepared on a MacIntosh computer (Apple Computer, Inc., Cupertino, CA, USA) with Adobe Photoshop 3.0 (Adobe Systems Inc., Mountain View, CA, USA) and ClarisDraw (Claris Corporation, Santa Clara, CA, USA) programs. The threshold level for rendering isosurface views (Figures 4, 5, 9 and 10) was selected to give a molecular volume consistent with the expected volume of the SpV4 capsid. Though the image and reconstruction data were not compensated for the effects of the contrast transfer function of the microscope [47], the chosen threshold gave a molecular envelope that fit quite well the X-ray structure of the ϕX174 F protein capsid (Figure 10).

Alignment of Microviridae *capsid proteins*

The programs used for the manipulation and alignment of sequences were those of the University of Wisconsin GCG [48]. The representative sequences of the F capsid proteins of α3, ϕK, ϕX174 and G4, and the VP1 proteins of SpV4 and Chp1 were obtained from the GenBank database [49]. Residues Gly35, Gly36, Pro68 and Arg281 of the F protein of G4 were changed to Val35, Val36, Ser68 and Gly281, according to sequencing performed by Bentley Fane and is in agreement with the interpretation of the high-resolution electron-density map of G4 [15]. The program PILEUP was used to align each sequence to a profile of all other sequences [50] with gap penalties of Gi = 3.0 and G1 = 0.1. The known secondary structural elements of ϕX174 and G4 [14,15] were used to make a few minor adjustments to the final sequence alignment. The alignment procedure was carried out on a Silicon graphics Indigo 2 workstation (Silicon Graphics Computer Systems, Mountain View, CA, USA).

Modeling the IN5 loop and fit to the cryo-EM density map

Three copies of the 71-residue loop (IN5) of SpV4 were built and fitted interactively into the cryo-EM density map with the program O [51] on a Silicon graphics Indigo 2 workstation. The β strand structural elements were modeled using segments from the β-barrel domain of the previously determined atomic structure of ϕX174 [14]. The turn regions were inserted between the β strands to complete the connectivity of the loop. The model was constrained using the structural geometry library of the program O.

Accession numbers

The Cα atomic coordinates of the SpV4 model have been deposited with the Brookhaven Protein Data Base [52], with the accession code 1KVP.

Acknowledgements

We thank Norm Olson, Tom Smith and Chris Storey for useful discussions and Sybille Duret for excellent technical assistance. This work was supported by National Institutes of Health grant GM33050 to TSB, Program Project grant AI35212 to the Purdue University Structural Virology group, and a University of Warwick Research and Teaching Initiative grant 0952 to RM.

References

1. Ricard, B., Garnier, M. & Bove, J.M. (1982). Characterization of SPV3 from spiroplasmas and discovery of a new spiroplasma virus (SPV4). *Rev. Infect. Dis.* **4**, S275.
2. Renaudin, J., Pascarel, M.C., Garnier, M., Carle, P. & Bove, J.M. (1984). Characterization of spiroplasma virus group 4 (SPV4). *Isr. J. Med. Sci.* **20**, 797-799.
3. Renaudin, J., Pascarel, M.C., Garnier, M., Junca-Carle, P. & Bove, J.M. (1984). SpV4, a new spiroplasma virus with circular, single stranded DNA. *Ann. Virol.* **135E**, 343-361.
4. Woese, C.R. (1987). Bacterial evolution. *Microbiol. Rev.* **51**, 221-271.
5. Weisburg, W.G., *et al.,* & Woese, C.R. (1989). A phylogenetic analysis of the mycoplasmas: basis for their classification. *J. Bacteriol.* **171**, 6455-6467.
6. Renaudin, J., Pacarel, M.C. & Bove, J.M. (1987). Spiroplasma virus 4: nucleotide sequence of the viral DNA, regulatory signals and proposed genome organisation. *J. Bacteriol.* **169**, 4950-4961.
7. Storey, C.C., Lusher, M. & Richmond, S.J. (1989). Analysis of the complete nucleotide sequence of Chp1, a phage which infects avian *Chlamydia psittaci. J. Gen. Virol.* **70**, 3381-3390.
8. Sanger, F., *et al.,* & Smith, M. (1977). Nucleotide sequence of bacteriophage ϕX174 DNA. *Nature* **265**, 687.
9. Sanger, F., *et al.,* & Smith, M. (1978). The nucleotide sequence of bacteriophage ϕX174. *J. Mol. Biol.* **125**, 225.
10. Godson, G.N., Barrell, B.C., Staden, R. & Fiddes, J.C. (1978). Nucleotide sequence of bacteriophage G4 DNA. *Nature* **276**, 236.
11. Caspar, D.L.D. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1-24.
12. Renaudin, J. & Bove, J.M. (1995). SpV1 and SpV4, spiroplasma viruses with circular, single-stranded DNA genomes, and their contribution to the molecular biology of spiroplasmas. In *Advances in Virus Research, Vol. 44.* pp.429-463, Vol. I, Academic Press, London, UK.
13. Olson, N.H., Baker, T.S., Willingmann, P. & Incardona, N.L. (1992). The three-dimensional structure of frozen-hydrated bacteriophage ϕX174. *J. Struct. Biol.* **108**, 168-175.
14. McKenna, R., *et al.,* & Incardona, N.L. (1992). Atomic structure of single-stranded DNA bacteriophage ϕX174 and its functional implications. *Nature,* **355**, 137-143.
15. McKenna, R., Bowman, B.R., Ilag, L.L., Rossmann, M.G. & Fane B.A. (1996). Atomic structure of the degraded procapsid particle of the bacteriophage G4: induced structural changes in the presence of calcium ions and functional implications. *J. Mol. Biol.*, **256**, 736-750.
16. Rossmann, M.G. & Johnson, J.E. (1989). Icosahedral RNA virus structure. *Annu. Rev. Biochem.* **58**, 533-573.
17. McKenna, R., Ilag, L.L. & Rossmann, M.G. (1994). Analysis of the single-stranded DNA bacteriophage ϕX174, refined at a resolution of 3.0 Å. *J. Mol. Biol.*, **237**, 517-543.
18. Tsao, J., *et al.,* & Parrish, C.R. (1991). The three-dimensional structure of canine parvovirus and its functional implications. *Science* **251**, 1456-1464.
19. Agbandje, M., McKenna, R., Rossmann, M.G., Strassheim, M.L. & Parish, C.R. (1993). Structure determination of feline panleukopenia virus empty particles. *Proteins* **16**, 155-171.

20. Llamas-Saiz, A.L., *et al.,* & Rossmann, M.G. (1997). Structure determination of the minute virus of mice. *Acta Cryst. D* **53**, 93-102.
21. Edgell, M.H., Hutchinson, C.A., III, & Sinsheimer, R.L. (1969). The process of infection with bacteriophage ϕX174. XXVIII. Removal of the spike proteins from the phage capsid. *J. Mol. Biol.* **42**, 547-557.
22. Kodaira, K.-I., Nakano, K., Okada, S. & Taketo, A. (1992). Nucleotide sequence of the genome of the bacteriophage α3: interrelationship of the genome structure and the gene products with those of the phages ϕX174, G4 and ϕK. *Biochim. Biophys. Acta* **1130**, 277-288.
23. Ilag, L.L., McKenna, R., Yadav, M.P., BeMiller, J.N., Incardona, N.L. & Rossmann, M.G. (1994). Calcium ion-induced structural changes in phage ϕX174. *J. Mol. Biol.* **244**, 291-300.
24. Fitch, W.M. & Margoliash, E. (1967). Construction of phylogenetic trees. *Science* **155**, 279-284.
25. Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B. & Thornton, J.M. (1997). CATH – a hierarchic classification of protein domain structures. *Structure* **5**, 1093-1108.
26. Jones, D.T. (1997). Progress in protein structure prediction. *Curr. Opin. in Struct. Biol.*. **7**, 377-387.
27. Rost, B. & Sander, C. (1993). Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.* **232**, 584-599.
28. Speir, J.A., Munshi, S., Wang, G., Baker, T.S. & Johnson, J.E. (1995). Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **3**, 63-78.
29. Munshi, S., *et al.,* & Johnson, J.E. (1996). The 2.8 Å structure of a T=4 animal virus and its implication for membrane translocation of RNA. *J. Mol. Biol.* **261**, 1-10.
30. Wilson, I.A., Skehel, J.J. & Wiley D.C. (1991). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366-373.
31. Altschul, S.F., Gish, W., Millar, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
32. Chipman, P.R., *et al.,* & Rossmann, M.G. (1996). Cryo-electron microscopy studies of empty capsids of human parvovirus B19 complexed with its cellular receptor. *Proc. Natl. Acad. Sci. USA* **93**, 7502-7506.
33. Sinsheimer, R.L. (1968). Bacteriophage ϕX174 and related viruses. *Prog. Nucleic Acids Res. Molec. Biol.* **8**, 115-169.
34. Newbold, J.E. & Sinsheimer, R.L. (1970). Process of infection with bacteriophage ϕX174 . XXXIV. Kinetics of the attachment and eclipse steps of the infection. *J. Virol.* **5**, 427-431.
35. Dowell, C.E., Jansz, H.S. & Zandberg, J. (1981). Infection of *Escherichia coli* K-12 by bacteriophage ϕX174. *Virology* **114**, 252-255.
36. Davies, R.E. (1978). Spiroplasmas from flowers of *Bidens pilosa L.* and honey bees in Florida: relationship to honey bee spiroplasma AS576 from Maryland. *Phytopatho News* **12**, 7.
37. Whitcomb, R.F. (1983). Culture media for spiroplasmas. In *Methods in Mycoplasmology.* (Razin, S. & Tully, J.G. eds.), Vol. I, pp.147-159, Academic Press, London.
38. Adrian, M., Dubochet, J., Lepault, J. & McDowall, A.W. (1984). Cryo-electron microscopy of viruses. *Nature* **308**, 32-36.
39. Dubochet, J., *et al.,* & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quart. Rev. Biophys.* **21,** 129-228.
40. Olson, N.H., Baker, T.S., Johnson, J.E. & Hendry, D.A. (1990). The three-dimensional structure of frozen-hydrated *Nudaurelia capensis* β virus, a T=4 insect virus. *J. Struct. Biol.* **105**, 111-122.
41. Baker, T.S. & Cheng, R.H. (1996). A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.* **116**, 120-130.
42. Baker, T.S., Newcomb, W.W., Olson, N.H., Cowsert, L.M., Olson, C. & Brown, J.C. (1991). Structures of bovine and human papillomaviruses: analysis by cryoelectron microscopy and three-dimensional image reconstruction. *Biophys. J.* **60**, 1445-1456.
43. Conway, J.F., Trus, B.L., Booy, F.P., Newcomb, W.W., Brown, J.C. & Steven, A.C. (1996). Visualization of three-dimensional density maps reconstructed from cryoelectron micrographs of viral capsids. *J. Struct. Biol.* **116**, 200-208.
44. Fuller, S.D., Butcher, S.J., Cheng, R.H. & Baker, T.S. (1996). Three-dimensional reconstruction of icosahedral particles – the uncommon line. *J. Struct. Biol.* **116**, 48-55.
45. Baker, T.S., Drak, J. & Bina, M. (1988). Reconstruction of the three-dimensional structure of simian virus 40 and visualization of the chromatin core. *Proc. Natl. Acad. Sci. USA* **85**, 422-426.
46. Baker, T.S., Drak, J. & Bina, M. (1989). The capsid of small papova viruses contains 72 pentameric capsomeres: direct evidence from cryo-electron microscopy of simian virus 40. *Biophys. J.* **55**, 243-253.
47. Toyoshima, C. & Unwin, N. (1988). Contrast transfer for frozen-hydrated specimens: determination from pairs of defocused images. *Ultramicroscopy* **25**, 279-292.
48. Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.
49. Bilofsky, H.S., *et al.,* & Tung, C.-S. (1986). The GenBank genetic sequence databank. *Nucleic Acids Res.* **14**, 1-4.
50. Gribskov, M., McLachlan, A.D. & Eisenberg, D. (1987). Profile analysis: detection of distantly related proteins. *Proc. Natl. Acad. Sci. USA* **84**, 4355-4358.
51. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110-119.
52. Bernstein, F.C., *et al*., & Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.
53. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detail and schematic plots of proteins structures.
54. Smith, T.J. (1990). MacInPlot: a program to display electron density and atomic models on the Macintosh personal computer. *J. Appl. Cryst.* **23**, 141-142.