Double-stranded RNA virus outer shell assembly by *bona fide* domain-swapping

Correct outer protein shell assembly is a prerequisite for virion infectivity in many multi-shelled dsRNA viruses. In the prototypic dsRNA bacteriophage φ6, the assembly reaction is promoted by calcium ions but its biomechanics remain poorly understood. Here, we describe the near-atomic resolution structure of the φ6 double-shelled particle. The outer T = 13 shell protein P8 consists of two alpha-helical domains joined by a linker, which allows the trimer to adopt either a closed or an open conformation. The trimers in an open conformation swap domains with each other. Our observations allow us to propose a mechanistic model for calcium concentration regulated outer shell assembly. Furthermore, the structure provides a prime exemplar of *bona fide* domain-swapping. This leads us to extend the theory of domain-swapping from the level of monomeric subunits and multimers to closed spherical shells, and to hypothesize a mechanism by which closed protein shells may arise in evolution.
The protein shells of icosahedral viruses assemble with remarkable accuracy from smaller subunits. Accumulating evidence suggests that such viruses can be grouped into a fairly small number, perhaps less than a dozen, of viral lineages, each sharing a set of common assembly principles. However, explaining how such closed, highly symmetric shells may have arisen during evolution poses a challenge, as formation of a closed shell first requires a network of complementary subunit-subunit interactions with roughly 50% of the surfaces buried from solvent. One of the most complex examples of such interactions is the capsid of bacteriophage HK97 where the capsid proteins are intertwined to form a topologically linked ‘chain mail’ organization. A simpler question to ask is how do multimeric proteins evolve from monomers? It has been suggested that this can be achieved by a mechanism called ‘domain swapping’. In the resulting multimer, the domain-swapped subunits in an ‘open’ conformation recapitulate the domain-domain interactions of the original ‘closed’ conformation of the monomer. This concept explains how mutations increasing the flexibility of a linker region of the monomer can facilitate domain swapping to give rise to the multimer.

Domain-swapping may play a role not just in the assembly and evolution of multimeric proteins, but also of viral shells. The structure of rice yellow mottle virus (RYMV; genus *Sobemovirus*), and its comparison to the related southern cowpea mosaic virus (SCPMV), provides an interesting example. In SCPMV, the *T* = 3 icosahedral capsid is formed in an usual way by trimers in a compact, closed conformation. In RYMV, however, the trimers are atypical and adopt an open conformation, swapping domains with the neighbouring trimers and creating long range interactions, to which the increased stability of the capsid has been attributed. However, it remains unclear whether RYMV trimers can exist in both open and closed conformations.

We use bacteriophage φ6, a member of the *Cystoviridae* family, as a model system to study the assembly of a complicated multi-shelled virus. The inner *T* = 1 shell, composed of protein P1, forms a dodecahedral polymerase complex (PC) harbouring the three viral dsRNA segments (S, M, and L), the viral polymerase P2 (refs 16,17), the hexameric packaging protein P4 (refs 18,19), and the assembly cofactor P7 (refs 20,21). Unlike eukaryotic dsRNA viruses, φ6 and other cystoviruses undergo a dramatic conformational change during RNA packaging and replication, from an empty collapsed PC to a fully expanded, packaged PC. The expanded PC is covered by a layer of P8, forming the outer *T* = 13 shell. This double-shelled particle, or nucleocapsid (NC), is itself covered by a protein-lipid envelope during virion assembly. A remarkable aspect of the φ6 system is that PC and NC particles can be assembled in vitro from purified protein and RNA components. In addition to φ6 bacteriophage (*Cystovirus*), many other dsRNA viruses, such as aquareovirus (*Aquareovirus*) (ref. 22), mammalian reovirus (*Orthoreovirus*) (ref. 23), blue tongue virus (*Orbivirus*) (ref. 24), rice dwarf virus (*Phytoreovirus*) (ref. 25), rotavirus (*Rotavirus*) (ref. 26), and Banna virus (*Seadornavirus*) (ref. 27), form a multi-shelled particle with an inner *T* = 1 and one or more outer *T* = 13 shells. While structures for some of these have been solved to near-atomic resolution, shedding light into molecular interactions in the *T* = 1 and *T* = 13 shells, assembly of the *T* = 13 shell in φ6 NC has remained elusive.

Here we present the structure of φ6 NC to near-atomic resolution using electron cryomicroscopy and single particle analysis. The structure of the outer NC shell reveals an intricate domain-swapped architecture created by P8 trimers. The majority of the P8 subunits (540 of 600) are in an open conformation, swapping domains with their neighbours. The inter-trimer interactions recapitulate the intra-trimer interactions of the closed conformation that were also observed within the NC. These observations not only provide a possible explanation for the calcium-regulated assembly of the φ6 NC particle, but they also constitute an exemplar of *bona fide* domain swapping in a viral shell, allowing us to propose a mechanism for how such complex shells may have arisen during evolution.

**Results**

**Near-atomic resolution structure of the φ6 nucleocapsid.** We determined the structure of φ6 double-shelled particle by electron cryomicroscopy (cryo-EM; Fig. 1; Table 1). A high-resolution cryo-EM data set was acquired by imaging purified φ6 NC samples, some of which had spontaneously lost the RNA genome (Supplementary Fig. 1a). Two-dimensional image classification revealed the morphology of NC particles
CDs and PDs of adjacent trimers via heterotypic interactions

The PDs swap between the trimers, each PD interacting with the trimer they belong to, for the majority of the P8 molecules of the reconstruction allowed modelling the fold of the organization of P8 trimers (Fig. 2). The relatively high resolution (Fig. 1) together to form the core of the P8 trimer. The PDs, due to the a C-terminal core domain (CD; residues 94–149). The loop ends by a 10-residue long elongated loop (residue 84–93) and that are connected by an elongated linker (Fig. 2c,d): an a three-dimensional reconstruction using a ‘gold-standard’ refinement approach (see Methods; Table 1). The overall resolution of the icosahedrally symmetric areas of the reconstruction was 4.0 Å, as determined by Fourier shell correlation (FSC, threshold = 0.143; Supplementary Fig. 1c). In many areas the local resolution was approaching 3.0 Å, while some areas, especially solvent exposed loops and areas around the icosahedral five-fold vertices in the outer protein shell were resolved to a lower resolution (Supplementary Fig. 1d–i). The resolution was sufficient to not only trace the course of the polypeptide chain but also to align the sequence to the structure and refine the structure, for nearly all of P1 and P8 residues (see Methods).

The P4 hexamers located at the icosahedral five-fold axes of symmetry were not resolved due to the symmetry mismatch (Fig. 1). Density corresponding to the three linear dsRNA genome segments appeared as concentric layers spaced ~3 nm apart but the detailed organization of the segments could not be addressed using icosaehedral reconstruction (Fig. 1b).

Architecture of the nucleocapsid outer shell. The NC structure revealed, consistent with our earlier observations23, a total of 200 P8 trimers forming the outer protein shell, following T = 13aveo quasi-symmetry (Fig. 1a). Four types of trimers (P8Q, P8O, P8S and P8T) occupy the different positions in the shell: Q-type trimers around the icosahedral five-fold axis of symmetry, S-type trimers adjacent to the two-fold axis, T-type trimers on the three-fold axis and R-type trimers between the other trimers. Each of the 60 asymmetric units of the shell thus consists of 10 P8 chains, as the P8T trimers contribute only one chain to each asymmetric unit.

The T = 13 shell revealed a remarkable domain swapped organization of P8 trimers (Fig. 2). The relatively high resolution of the reconstruction allowed modelling the fold of the P8 structure (residues 3–147; Fig. 2a,b; Supplementary Fig. 2). The structure is α-helical and can be divided into two domains that are connected by an elongated linker (Fig. 2c,d): an N-terminal peripheral domain (PD; residues 1–83) is followed by a 10-residue long elongated loop (residue 84–93) and a C-terminal core domain (CD; residues 94–149). The loop ends in a proline residue (Pro93), which breaks the preceding helix of the CD (Supplementary Fig. 2). The CDs intertwine together to form the core of the P8 trimer. The PDs, due to the elongated nature of the linker, are located a considerable distance (~50 Å) from the CD. Instead of interacting with the CD of the trimer they belong to, for the majority of the P8 molecules the PDs swap between the trimers, each PD interacting with the CDs and PDs of adjacent trimers via heterotypic interactions (Fig. 2c,f). Nine such structures are observed in the icosahedral asymmetric unit of the NC.

In addition to the ‘open’ conformation described above and adopted by 9 out of 10 of the P8 chains, a fundamentally different conformation is also found on the shell. This ‘closed’, more compact, conformation is observed in one of the three chains of the peripental P8Q trimer (Fig. 2g; Supplementary Fig. 2). The two conformations differ in degree of bending of the elongated linker that acts as a hinge. The change in relative position between the two conformations is ~75 Å and involves a rotation of 71 degrees (Fig. 2h). In the closed conformation, the PD is bound to its own CD via a homotypic interaction (Fig. 2i). This interaction is characterised by an interface that is analogous to the one created between the PD and a CD of an adjacent trimer in heterotypic interactions (Fig. 2j).

P4 hexamer is anchored to the vertices of the P1 shell. In the members of the subfamily Sedoreovirinae of the family Reoviridae, such as blue tongue virus (Ortivirus)29 and rice dwarf virus (Phytoreovirus)32, P-type trimers occupy the five-fold vertices. In contrast, in members of Cystoviridae, such as ø6 (ref. 23) and Pinareovirinae, such as aquareoviruses27, the P-type trimers are absent leaving room for turret-like structures at the icosahedral five-fold vertices. In ø6, the vertices are occupied by the P4 packaging hexamer (Fig. 1). The presence of P4 hexamers at the five-folds provides also a rationale for the observed closed conformation of certain peripental P8 proteins (Supplementary Fig. 3); if the type P8Q monomers were all in the open conformation, significant clashes would occur between P8Q and P4 (Supplementary Fig. 3b).

The six-fold symmetry of the P4 hexamer and five-fold symmetry of the icosahedral vertex beneath creates a symmetry mismatch, so that the structure of P4 cannot be resolved using conventional icosahedral reconstruction23. To address this we applied the localized reconstruction method27 to P4 hexamers. Local areas corresponding to P4 hexamers were extracted from the NC images and reconstructed as single particles with six-fold symmetry (‘P4 hexamer reconstruction’) or without any symmetry (‘NC vertex reconstruction’). This allowed a low-resolution structure of the P4 hexamer to be determined in situ (Fig. 3). At a resolution of 9.1 Å (FSC = 0.143), fitting of a P4 crystal structure (PDB:4BLO)19 into the density was unambiguous (Fig. 3a–c) and allowed us to orient the hexamer relative to the NC vertex (Fig. 3d,e).

A disorder prediction algorithm suggests that the C-terminal residues of P4 are likely to be disordered (Supplementary Fig. 4)35. Indeed, they are missing in the crystal structure and this disordered region has been proposed to attach the P4 to the underlying P1 shell19. However, in earlier low-resolution cryo-EM reconstructions, the P4 hexamers appeared to float on top of the five-fold vertexes of the P1 shell23,36,37 and the connections between these symmetry-mismatched components have remained unresolved. Our localized reconstruction of P4 hexamers that correctly dealt with the symmetry mismatch revealed elongated densities extending from 3 of the 6 P4 C-terminal parts (Fig. 3f). We hypothesised that these may correspond to the C-terminal parts that are disordered in the crystal structure. Consistent with this hypothesis, in the high-resolution icosahedral NC reconstruction we observed sixty copies of elongated density that lacked any apparent secondary structure stretching ~115 Å on top of each of the asymmetric P1 dimers (Figs 1b and 4c–e). At the vertex distal part of this density, densities consistent with the side-chains 324-ProArgArg-326...
of the P4 C terminus were identified (Fig. 4f), allowing us to specifically assign this density to the C-terminal residues 292–332 of P4. The sixth, unbound C-terminal region of P4 was not resolved in any of the reconstructions, and is likely to be either disordered or unresolved in our maps due to limited resolution of the localized reconstructions.

Binding of up to five of the six flexible C-terminal tails of P4 provides an elegant solution to the symmetry mismatch. These tails are likely to be disordered in solution (Supplementary Fig. 4) and to order only upon binding to P1 during assembly. The P4 tails are analogous to the ‘stay-cables’ observed earlier in adenovirus fibres. As such cables have now been observed in at least two unrelated viruses, adenovirus (a dsDNA virus) and φ6 (a dsRNA virus, this study), they may be a general solution to the problem of attaching fibres and other symmetry-mismatched appendices to vertices of icosahedral virus capsids.

**P4 C-terminal tail solidifies the dimeric P1 building block.** The inner P1 shell is composed of 60 copies of asymmetric P1 dimers,
form by P1A and P1B (see Fig. 1b)23. A large-scale expansion of the P1 shell occurs during RNA packaging to empty particles and subsequent RNA replication22,23. To model the structural transitions in the P1 shell during this expansion, we refined the existing atomic coordinates of the P1 monomer (PDB:4K7H)39 into our EM density map of the NC (Fig. 4a,b) and compared them to P1 coordinates flexibly fitted to an unexpanded PC reconstruction (PDB:4BTG)39. The angle between the two subunits in the P1A–P1B asymmetric dimer changes relatively little (11°) showing that the P1A–P1B dimer behaves as quite a rigid block during the expansion (Supplementary Fig. 5a,b; Supplementary Movie 1). In contrast to these intradimer angles, the interdimer angles change significantly, mainly at the P1B–P1B interfaces39. The interdimer angle changes 57° at the two-fold interface and 49° at the three-fold interface (Supplementary Fig. 5c,d). The root-mean-square deviation (RMSD) calculated between the two states was 2.6 and 6.0 Å for the P1A and P1B C-alpha coordinates, respectively. This shows that relatively small changes occur within the P1A subunits, while P1B subunits adjust to their local environment in the two very different P1 shell expansion states.

The relative rigidity of the P1A–P1B dimer interface is now explained by our observation that the C-terminal part of P4 bridges the P1A and P1B monomers (Fig. 4c,d) whilst still allowing interdimer angles to change. Thus, the P4 C-terminus seems not only to link P4 hexamers to the P1 shell, but also to play a role in P1 dimerization. A dimer of the major inner capsid protein could indeed be a conserved intermediate in the assembly pathways of dsRNA viruses. This notion is supported by the structure of the Penicillium chrysogenum virus in which the equivalent building block is formed by a covalently linked major capsid protein dimer, produced by gene duplication40. In φ6, we propose that the C-terminal tail of P4 provides an analogous linkage.

**Discussion**

When combined with earlier structural and biochemical results, our NC structure allows us to propose an assembly model for φ6 (Fig. 5; Supplementary Movie 2). The assembly of the P1 shell is initiated *in vitro* by P4 hexamers26 and the C-terminus of P4 is essential for the formation of recombinant PC *in vivo*41. We now have a structural rationale for these observations, since the P4 C-terminus cements the P1 asymmetric dimer together (Fig. 4c,d). Accordingly, we suggest that the φ6 PC assembly starts with the formation of a P1A–P1B dimer stabilized by the C-terminus of P4 (Fig. 5a). As observed in the P1 crystal structure, individual P1 chains tend to create a pentamer in a conformation that resembles a group of five P1A chains in the unexpanded conformation of PC39. Our modelling experiments also showed that six P1A–P1B dimers per P4 would create steric clashes (Supplementary Movie 2). It is thus likely that each P4 grabs five and not more P1A–P1B dimers. Dimers attached to distinct P4 hexamers then interact to eventually form a closed P1-shell (Fig. 5b). The assembly of the compact empty PC is facilitated by the assembly cofactor P7 (refs 20,26). P7, together with P2, is located at the three-fold symmetry axis of the empty PC21, but the order of P2 and P7 binding to the PC remains unclear. The PC subsequently packages and replicates the RNA genome segments leading to expansion of the PC (Fig. 5c)23,39 and finally formation of the double-shelled particles by incorporation of the P8 layer (Fig. 5d).

P8 shells form spontaneously in the presence of 0.1–1.0 mM Ca²⁺ and in the absence of any other structural proteins42. Conversely, depleting calcium by a chelating agent or exposure to low pH leads to P8 shell disassembly42,43. Raman spectroscopy has revealed that the assembled shell stabilizes the mostly α-helical conformation of P8 (ref. 44). It is thus clear that Ca²⁺ regulates the conformational state of P8. In the light of our observation of a closed and open state of the P8 trimer in
the NC, we hypothesize that Ca$^{2+}$ induces transition of the P8 trimer from the closed to the open conformation and this further facilitates outer shell assembly by domain swapping (Fig. 5d,e). Due to steric clashes, the P8 monomers closest to the P4 hexamers are forced to remain in closed conformation (Supplementary Fig. 3). Ca$^{2+}$ regulation of P8 conformation, possibly in addition to acidic pH, might also be important to facilitate the delivery of the PC to the host cell upon infection.

The conformational space sampled by P8 trimers in solution and at different Ca$^{2+}$ concentrations remains unknown. Insolubility of purified P8 and its tendency to readily form large assemblies even in the presence of Ca$^{2+}$-chelating agents has impeded its solution structure studies by small angle X-ray scattering (Z.S. and J.T.H., unpublished observations). The P8 CD domain is preceded by Pro93 that breaks the helix $\alpha$6 of the CD. Proline residues are abundant in the linker regions of domain swapped proteins and may play a role in modulating the equilibrium between an open and a closed state. It is thus possible that Pro93 has a functional role in adding strain to the backbone of the linker region but this hypothesis remains to be tested.

The observed P8 shell assembly allows us to extend the concept of domain swapping to viral shells. Domain swapping is considered bona fide, when the same subunit can still adopt both closed and open conformations, as opposed to the subunits adopting just the open conformation in a multimer and the closed conformation in a different, albeit orthologous, monomeric protein. Following this nomenclature, our $\phi 6$ outer shell structure provides an exemplar of bona fide domain swapping between P8 trimers that exhibit both closed and open conformations.

Analogously to monomer to multimer transition during the evolution of multimeric proteins, our results allow us to propose a hypothetical model how closed viral protein shells may further evolve from such proteins. A primordial, relatively simple capsomere may have some tendency to form larger assemblies, but initially the protein-protein interfaces required to make a spherical shell are not optimized. Mutations in a linker region can increase its flexibility to allow for the formation of an open conformation of the multimeric capsomere. These capsomeres in their open conformations would then be able to further link together by recapitulating the domain-domain interactions of the closed conformation. A pre-existing inner protein shell, genome, or internal lipid bilayer present in some viruses, may play a role in increasing the fidelity of shell assembly. This is also evident in the case of $\phi 6$, where in the absence of the inner P1 shell, P8 self-assembles in vitro into aberrant, shell-like structures that vary in size and are not topologically closed. In some viruses, the capsomeres may have subsequently lost the ability to form both open and closed conformations. To what extent the outer protein shells of $\phi 6$ and other dsRNA viruses are homologous remains an open question, as domain-swapping is not observed in other dsRNA viruses. It is conceivable, however, that linker region deletions and domain-shuffling may have played a further role in stabilizing protein shells. In the case of $\phi 6$, the ability of P8 to still adopt both conformations may have been preserved in evolution as this may play a role in calcium-regulated outer shell assembly/disassembly and also the presence of P4 hexamer at the vertices necessitates a closed conformation of P8.

In conclusion, our structural characterization of the $\phi 6$ double-shelled particle revealed interactions of the inner P1 shell and symmetry-mismatched P4 packaging hexamers. Most importantly, the study led to the discovery of two states of the outer shell forming protein P8, ‘open’ and ‘closed’. Trimers in the ‘open’ conformation swap domains between each other. This process is possibly regulated by calcium ion concentration but further studies are required to reveal the exact mechanism of Ca$^{2+}$ regulation. The derived outer shell assembly model not only provides insights into assembly in this particular model system, but also suggests how closed protein shells may have arisen during evolution. Further studies are required to assess to what extent domain swapping may have played a role in other viral systems.

**Methods**

**Cryo-EM sample preparation.** Bacteriophage $\phi 6$ was grown on its host *Pseudomonas syringae* pv. *phaseolicola* HB10Y48 and purified as described previously. The viral envelope was removed by Triton X-114 extraction and the resulting aqueous mixture containing the NC loaded onto a CIM DEAE-1,
Figure 5 | Assembly model. (a) The φ6 inner capsid is composed of protein P1, P2, P4 and P7. Preassembled P4 hexamers (grey) nucleate the assembly of the P1 shell by bringing together two P1 subunits that will form the asymmetric P1a–P1b dimer (light blue and red) in the final assembly. The dimerization of P1 is facilitated by the C-terminal tail of P4 (dark blue). Up to five P1 dimers can assemble around a P4 hexamer. (b) P1 dimers connected to distinct P4 hexamers interact bringing together the precursors of the icosahedral vertices. This is facilitated by protein P7, which acts as an assembly cofactor. (c) The self-assembled empty capsids package and replicate the single-stranded RNA precursors of the φ6 dsRNA genome (S, M and L) and the capsid adopts the expanded conformation. (d) Trimers of P8 (green, yellow, gold and brown) then assemble around the dsRNA containing inner capsid to form the nucleocapsid. (e) The assembly of the P8 shell is induced by Ca²⁺ ions and involves major conformational changes in the linker regions of P8 subunits to mediate the transition from the closed to the open conformations. See also Supplementary Movie 2.

Table 2 | Data processing parameters.

| Parameter | NC particle | P4 hexamer | NC vertex |
|-----------|-------------|------------|-----------|
| Particles* | 13,291 (16,466) | 56,448 (159,492) | 56,448 (159,492) |
| Box size (pixels) | 512 | 128 | 128 |
| Pixel size (Å) | 1.35 | 1.35 | 1.35 |
| Symmetry | 11 | C6 | C1 |
| Resolution (Å) | 9.1 | 9.1 | 9.1 |
| B-factor (Å²) | – 177 | – 600¹ | – 600¹ |

NC, nucleocapsid.
*Number of particles (or subparticles in the case of P4 hexamer and NC vertex) chosen for averaging the final map is given. The number of all particles extracted from the micrographs (or all subparticles extracted from particles) is given in parenthesis.
†Inverse B-factor used for sharpening of the map.
‡Value of B-factor determined ad hoc.

Data acquisition and processing. Data were acquired using a 300-kV transmission electron microscope (Tecnai F30 ‘Polara’; FEI) equipped with an energy filter (slit width 20 eV; GIF Quantum LS, Gatan, Pleasanton, CA) and a direct electron detector (K2 Summit, Gatan, MA, USA). A 3-μl aliquot of sample, diluted in buffer (20 mM KPO₄ pH 7.2, 1 mM MgCl₂, 0.1 mM CaCl₂). The NC-containing fractions were concentrated using an ultra-centrifugal filter with a 100-kDa cutoff (Amicon; EMD Millipore, Billerica, MA, USA). A 3-μl aliquot of sample, diluted in buffer (20 mM KPO₄ pH 7.2, 1 mM MgCl₂, 0.1 mM CaCl₂, 150 mM NaCl) to concentration of 3 mg ml⁻¹, was applied to a glow-discharged grid (C-flat; Protochips, Raleigh, NC) and vitrified by plunge-freezing into liquid ethane using a Vitrobot (FEI, Hillsboro, OR).

Data were acquired using a 300-kV transmission electron microscope (Tecnai F30 ‘Polara’; FEI) equipped with an energy filter (slit width 20 eV; GIF Quantum LS, Gatan, Pleasanton, CA) and a direct electron detector (K2 Summit, Gatan, MA). Movies (22 frames, each frame 0.2 s) were collected in electron counting mode at dose rate of 5 e⁻/per pixel per s at calibrated magnification of 37,037 ×, resulting in a total dose of ~16 e⁻/Å² and pixel size of 1.35 Å. Data acquisition statistics are summarized in Table 1.

Movie frames were aligned to account for drift⁵⁹ and contrast transfer function (CTF) parameters were estimated⁵⁹. A total of ~900 movies, showing minimal drift and astigmatism were used to pick 16,466 particles. Particles were extracted for 2D reference free classification in Relion 1.3 (ref. 32). Particles from classes showing complete P1 and P8 layers were selected for 3D classification using a previously published NC structure (EMD-1,206) as an initial model. A subset of 13,291 particles extracted from movies was refined using standard refinement and particle polishing protocols implemented in Relion.

The structure of the P4 hexamer was calculated using localized reconstruction (www.opic.ox.ac.uk/localrec)⁷⁷. Briefly, twelve ‘subparticles’, each corresponding to a projection of the P4 hexamer in each of the particle images, was extracted (total of 159,463 subparticles) to a smaller box and treated as independent single particles. To assign the orientation for each subparticle, one of the five possible triplets of Euler angles that were equivalent for the corresponding five-fold symmetric vertex was chosen randomly so as not to bias orientations in further processing. Two types of reconstructions were calculated. For the first reconstruction (“P4 hexamer”), partial signal subtraction was used to remove the contribution of unwanted NC density in the particle images⁵³. We modified our earlier localized reconstruction method⁷⁷ to allow subtracting ‘all-but-one’ P4 hexamer density (available from http://github.com/OPIC-Oxford/localrec), in addition to the other unwanted densities (P1 and P8 shells and RNA) from the icosahedral reconstruction. Only subparticles from the edge of the particle (total of 103,482; maximum 40° deviation from a side view) were chosen for further analysis to reduce the overlap of the projected hexamer density with the RNA component of the virion that could not be accurately subtracted from the particle images. The structure of the P4 hexamer was then determined using Relion and 3D classification, using a localized reconstruction of the hexamer, calculated from a subset of subparticles and filtered to 40 Å resolution, as a starting model. 3D classification revealed the hexamer in different orientations relative to the five-fold vertex. Particles from classes showing clear hexamer density were combined and six-fold symmetry was applied on the reconstruction. For the second reconstruction (“NC vertex”), subparticles extracted from the original particle images and orientation parameters from the P4 hexamer reconstruction were used and no symmetry was applied.

Resolutions of the icosahedral NC reconstruction and P4 hexamer localized reconstruction were estimated by using FSC using 0.143 cutoff. The NC vertex reconstruction was low-pass filtered to the resolution of P4 hexamer reconstruction (9.1 Å) and both localized reconstructions were sharpened by applying an inverse B-factor of –600 Å² (determined by trial and error). Local resolution of the NC reconstruction was estimated using ResMap²⁴. Reconstruction statistics are summarized in Table 2.

Atomic model building and refinement. As no P8 atomic model was available, a polyalanine model was built de novo using Coot. Visually identified densities.
for bulky residues facilitated building the full atomic model of P8 (L3-Y147) and the C terminus of P4 (R292-L332) with side-chains. The crystal structure of φ6 P1 (PDB:4K7H) was fitted in the NC map using COOT55 as a rigid body in two different positions corresponding to subunits P1A and P1B. P1A and P1B main-chains and side-chains were adjusted using manual and real space fitting in COOT. The crystal structure of φ6 P4 (PDB:4BLO)56 was fitted in the P4 hexamer and NC vertex localized reconstructions as a rigid body in UCSF Chimera49.

Atomic models for all of the chains in the asymmetric unit (P1A, P1B, P4 C terminus and ten copies of P8) were refined in Phenix.real_space_refine applying secondary structure, rotamer, and Ramachandran plot restraints. In addition, non-crystallographic symmetry restraints were applied on P8 chains57. The models were validated with MOLPROBITY58. Atomic model refinement statistics are summarized in Table 3. Disordered regions in P4 sequence (UniProt:P11125) were predicted with RONN59 and changes in subunit-subunit angles calculated in UCSF Chimera49.

Data availability. Density maps and atomic models that support the findings of this study have been deposited in the Electron Microscopy Database and in the Protein Databank with the accession codes EMD-3571 (NC reconstruction), PDB 5MUU (NC atomic model), EMD-3572 (P4 hexamer reconstruction), PDB 5MUV (P4 atomic model fitted in the P4 hexamer reconstruction), EMD-3573 (NC vertex reconstruction), and PDB 5MUW (P4 atomic model fitted in the NC vertex reconstruction).

References
1. Aksyuk, A. A. & Rossmann, M. G. Bacteriophage assembly. Viruses 3, 172–203 (2011).
2. Zlotnick, A., Porterfield, J. Z. & Wang, J. C.-Y. To build a virus on a nucleic acid substrate. Biophys. J. 104, 1595–1604 (2013).
3. Perlmuter, J. D. & Hagan, M. F. Mechanisms of virus assembly. Annu. Rev. Phys. Chem. 66, 217–239 (2015).
4. Bamford, D. H. Do viruses form lineages across different domains of life? Res. Microbiol. 154, 231–236 (2003).
5. Abrescia, N. G. A., Bamford, D. H., Grimes, J. M. & Stuart, D. I. Structure for bulky residues facilitated building the full atomic model of P8 (L3-Y147) and the C terminus of P4 (R292-L332) with side-chains. The crystal structure of φ6 P1 (PDB:4K7H) was fitted in the NC map using COOT55 as a rigid body in two different positions corresponding to subunits P1A and P1B. P1A and P1B main-chains and side-chains were adjusted using manual and real space fitting in COOT. The crystal structure of φ6 P4 (PDB:4BLO)56 was fitted in the P4 hexamer and NC vertex localized reconstructions as a rigid body in UCSF Chimera49.

Table 3 | Atomic model refinement statistics.

| Parameter          | P1       | P8       |
|--------------------|----------|----------|
| Refinement         |          |          |
| Resolution (Å)     | 4.0      | 4.0      |
| Map CC             | 0.715    | 0.715    |
| RMS deviations     |          |          |
| Bonds (Å)          | 0.09     | 0.09     |
| Angles (°)         | 1.0      | 1.0      |
| Molprobity validation |        |          |
| Clashscore, all atoms (percentile) | 13.8 (55th) | 10.4 (70th) |
| Molprobity score   | 1.92 (80th) | 1.53 (94th) |
| Favoured rotamers (poor %) | 98.4 (0.1) | 97.3 (0.0) |
| Ramachandran (% favoured) | 96 | 99.0 |
| Ramachandran (% allowed) | 99.7 | 100 |
| Ramachandran Plot (% outliers) | 0.3 | 0 |

RMS, root mean square.

14. Semancik, J. S., Vidaver, A. K. & Van Etten, J. L. Characterization of segmented double-helical RNA from bacteriophage φ6. J. Mol. Biol. 78, 617–625 (1993).
15. Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H. & Fuller, S. D. Intermediates in the assembly pathway of the double-stranded RNA virus φ6. EMBO J. 16, 4477–4487 (1997).
16. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. A mechanism for initiating RNA-dependent RNA polymerization. Nature 410, 235–240 (2001).
17. Ilica, S. L. et al. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. Nat. Commun. 6, 8843 (2015).
18. Pirttiaho, M. J., Paatero, A. O., Frilander, M. J. & Bamford, D.H. Nonspecific nucleoside triphosphatase P4 of double-stranded RNA bacteriophage φ6 is required for single-stranded RNA packaging and transcription. J. Virol. 76, 10122–10127 (2002).
19. El Omari, K. et al. Tracking in atomic detail the functional specializations in viral ReCa helicases that occur during evolution. Nucl. Acids Res. 41, 9396–9410 (2013).
20. Poranen, M. M., Butcher, S. J., Simonov, V. M., Laurinmäki, P. & Bamford, D. H. Roles of the minor capsid protein P7 in the assembly and replication of double-stranded RNA bacteriophage φ6. J. Mol. Biol. 383, 529–538 (2008).
21. Nemecek, D., Qiao, J., Mindich, L., Steven, A.C. & Heymann, J.B. Packaging accessory protein P7 and polymerase P2 have mutually occluding binding sites inside the bacteriophage φ6 procapsid. J. Virol. 86, 11616–11624 (2012).
22. Nemecek, D. et al. Stepwise expansion of the bacteriophage φ6 procapsid: possible packaging intermediates. J. Mol. Biol. 414, 260–271 (2011).
23. Huiskonen, J. T. et al. Structure of the bacteriophage φ6 nucleocapsid suggests a mechanism for sequential RNA packaging. Structure 14, 1039–1048 (2006).
24. Stitt, B. L. & Mindich, L. Morphogenesis of bacteriophage φ6: a presumptive viral membrane precursor. Virology 127, 446–458 (1983).
25. Jääniloja, H. T., Huiskonen, J. T. & Butcher, S. J. Electron cryomicroscopy comparison of the architectures of the enveloped bacteriophages φ6 and φ8. Structure 15, 157–167 (2007).
26. Poranen, M. M., Paatero, A. O., Tuma, R. & Bamford, D. H. Self-assembly of a viral molecular machine from purified protein and RNA constituents. Mol. Cell 7, 845–854 (2001).
27. Zhang, X., Jin, L., Fang, Q., Hui, W. H. & Zhou, Z. H. 3.3 A cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. Cell 141, 472–482 (2010).
28. Dryden, K. A. et al. Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformation: analysis of virions and subviral particles by cryoelectron microscopy and image reconstruction. J. Cell Biol. 122, 1023–1041 (1993).
29. Grimes, J. M. et al. The atomic structure of the bluetongue virus core. Nature 395, 470–478 (1998).
30. Zhang, X. et al. Atomic model of a nonenveloped virus reveals pH sensors for a coordinated process of cell entry. Nat. Struct. Mol. Biol. 23, 74–80 (2016).
31. Zhou, Z. H. et al. Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. Nat. Struct. Mol. Biol., 868–873 (2001).
32. Nakagawa, A. et al. The atomic structure of rice dwarf virus reveals the self-assembly mechanism of component proteins. Structure 11, 1227–1238 (2003).
33. Settenbre, E. C., Chen, J. Z., Dormitzer, P. R., Grigorieff, N. & Harrison, S. C. Atomic model of an infectious rotavirus particle. EMBO J. 30, 408–416 (2011).
34. Mohd Jaafar, F. et al. The structure and function of the outer coat protein VP9 of Banna virus. Structure 13, 17–28 (2005).
35. Yang, Z. R., Thomson, R., McNeil, P. & Esnouf, R. M. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics 21, 3369–3376 (2005).
36. de Haas, F., Paatero, A. O., Mindich, L., Bamford, D. H. & Fuller, S. D. A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage φ6. J. Mol. Biol. 294, 357–372 (1999).
37. Huiskonen, J. T., Jääniloja, H. T., Briggs, J. A. G., Fuller, S. D. & Butcher, S. J. Structure of a hexameric RNA packaging motor in a viral polymerase complex. J. Struct. Biol. 158, 156–164 (2007).
38. Liu, H., Wu, L. & Zhou, Z. H. Model of the trimeric fiber and its interactions with the pentameric penton base of human adenovirus by cryo-electron microscopy. J. Mol. Biol. 406, 764–774 (2011).
39. Nemecek, D. et al. Subunit folds and maturation pathway of a dsRNA virus capsid. Structure 21, 1374–1383 (2013).
40. Luque, D. et al. The T = 1 capsid protein of *Penicillium chrysogenum* virus is formed by a repeated helix-rich core indicative of gene duplication. *J. Virol.* **84**, 7256–7266 (2010).

41. Paatero, A. O., Mindich, L. & Bamford, D. H. Mutational analysis of the role of nucleoside triphosphatase P4 in the assembly of the RNA polymerase complex of bacteriophage φ6. *J. Virol.* **72**, 10058–10065 (1998).

42. Ojala, P. M., Romantschuk, M. & Bamford, D. H. Purified φ6 nucleocapsids are capable of productive infection of host cells with partially disrupted outer membranes. *Virology* **178**, 364–372 (1990).

43. Cvirkaite-Krupovic, V., Poranen, M. M. & Bamford, D. H. Phospholipids act as secondary receptor during the entry of the enveloped, double-stranded RNA bacteriophage φ6. *J. Gen. Virol.* **91**, 2116–2120 (2010).

44. Tuma, R., Bamford, J. K., Bamford, D. H. & Thomas, G. J. Assembly dynamics of the nucleocapsid shell subunit (P8) of bacteriophage φ6. *Biochemistry* **38**, 15025–15033 (1999).

45. Rousseau, F., Schymkowitz, J. W., Wilkinson, H. R. & Itzhaki, L. S. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc. Natl Acad. Sci. USA* **98**, 5596–5601 (2001).

46. Gil-Carton, D. et al. Insight into the assembly of viruses with vertical single β-barrel major capsid proteins. *Structure* **23**, 1866–1877 (2015).

47. Olkkonen, V. M., Ojala, P. M. & Bamford, D. H. Generation of infectious nucleocapsids by *in vitro* assembly of the shell protein on to the polymerase complex of the dsRNA bacteriophage φ6. *J. Mol. Biol.* **218**, 569–581 (1991).

48. Visave, A. K., Koski, R. K. & Van Etten, J. L. Bacteriophage φ6: a lipid-containing virus of *Pseudomonas phaseolicola*. *J. Virol.* **11**, 799–805 (1973).

49. Bamford, D. H., Ojala, P. M., Finlander, M., Walin, L. & Bamford, J. K. H. in *Methods in Molecular Genetics* Vol. 6 (ed. Adolph, K. W.) 455–474 (Academic Press, 1995).

50. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).

51. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy, *J. Struct. Biol.* **142**, 334–347 (2003).

52. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

53. Bai, X.-C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human γ-secretase. *Elife* **4**, e11182 (2015).

54. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).

55. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

56. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with MOLPROBITY: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

57. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

58. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

59. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure* **13**, 473–482 (2005).

Acknowledgements

We thank Alistair Siebert for electron microscopy support and Ritta Tarkkainen for technical help. The OPIC electron microscopy facility was founded by a Wellcome Trust JIF award (060208/Z/00/Z) and is supported by a WT equipment grant (093305/Z/10/Z). The EU ESFRI Instruct Centre for Virus Production (ICVIR) facility used in this study is supported by the Academy of Finland (grant 272853) and University of Helsinki and this is a collaboration with the Oxford Instruct Centre. This work was funded by Academy of Finland (218080 and 263677 to J.T.H; 283192, 250113, and 272507 to M.M.P.), Sigrid Juselius Foundation, Medical Research Council (MR/N00065X/1 to K.E.O.), European Research Council under the European Union’s Horizon 2020 research and innovation programme (649053 to J.T.H.) and Wellcome Trust Core Award Grant Number 090532/Z/09/Z. The authors acknowledge the support of employees and the use of experimental resources of Instruct, a Landmark ESFRI project.

Author contributions

Z.S., S.L.I. and J.T.H. collected and analysed cryo-EM data. Z.S., S.L.I. and J.T.H. designed the study. X.S. prepared the samples. Z.S., K.E.O. and A.K. built and refined the atomic models. Z.S. and J.T.H. wrote the manuscript and all authors read and commented on the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Sun, Z. et al. Double-stranded RNA virus outer shell assembly by bona fide domain-swapping. *Nat. Commun.* **8**, 14814 doi: 10.1038/ncomms14814 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2017

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/