Structure of a hexameric RNA packaging motor in a viral polymerase complex

Juha T. Huiskonen a, Harri T. Jäälinoja a, John A.G. Briggs b,1, Stephen D. Fuller b, Sarah J. Butcher a,∗

a Institute of Biotechnology and Faculty of Biosciences, University of Helsinki, P.O. Box 65 (Viikinkaari 1), FI-00014 University of Helsinki, Finland
b Oxford Particle Imaging Centre, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford OX3 7BN, UK
Received 12 May 2006; received in revised form 4 August 2006; accepted 16 August 2006

Abstract

Packaging of the Cystovirus /p10 genome into the polymerase complex is catalysed by the hexameric P4 packaging motor. The motor is located at the fivel fold vertices of the icosahedrally symmetric polymerase complex, and the symmetry mismatch between them may be critical for function. We have developed a novel image-processing approach for the analysis of symmetry-mismatched structures and applied it to cryo-electron microscopy images of P4 bound to the polymerase complex. This approach allowed us to solve the three-dimensional structure of the P4 in situ to 15-Å resolution. The C-terminal face of P4 was observed to interact with the polymerase complex, supporting the current view on RNA translocation. We suggest that the symmetry mismatch between the two components may facilitate the ring opening required for RNA loading prior to its translocation.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Icosahedral reconstruction; Symmetry mismatch; Cryo-electron microscopy; Virus; Hexameric helicase; Cystovirus; RNA packaging

1. Introduction

The simplest viruses consist only of the genome and a protein capsid that protects the genome. The capsid in spherical viruses is composed of many copies of one or more capsid proteins and exploits icosahedral symmetry (Baker et al., 1999). In addition to the static protection provided by the capsid, there are various steps in the viral life cycle where dynamic functionality is needed. Such steps include host recognition, cell entry, genome replication, particle assembly, and genome packaging. The organization of the proteins responsible for such functionality often deviates from the icosahedral symmetry of the capsid. In the members of the Cystoviridae family, one such functionally important protein is the packaging ATPase, a molecular motor responsible for translocating RNA genome segments into a preformed particle. This hexameric protein is bound at the fivel fold symmetric vertices of the icosahedral capsid, thus exhibiting a symmetry mismatch (de Haas et al., 1999).

Cystoviruses (φ6–φ14) are enveloped bacteriophages with segmented dsRNA genomes (Mindich, 2004; Mindich et al., 1999). Their genome is shielded from the host cell cytoplasm inside the polymerase complex (PC; 33 MDa, 54 nm in diameter). The PC is composed of the major structural protein P1 (Gottlieb et al., 1988), packaging enzyme P4 (Paatero et al., 1998), monomers of the RNA-dependent RNA polymerase P2 (Makeyev and Bamford, 2000), and the packaging factor P7 (Juuti and Bamford, 1997). One hundred and twenty copies of P1 form the dodecahedral framework of the PC shell with icosahedral symmetry (Huiskonen et al., 2006).

1 Corresponding author. Fax: +358 9 19159930.
E-mail address: sarah.butcher@helsinki.fi (S.J. Butcher).
1 Present address: Department of Chemistry and Biochemistry, Ludwig-Maximilians-Universität München, Butenandstraße 11, Haus E, 81377 München, Germany.

1047-8477/S - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2006.08.021

During viral replication, the PC transcribes the three genome segments. Translation of the mRNAs results in assembly of new empty PCs. After specific recognition by P1, the ssRNA transcripts are packaged by the P4 packaging motor into the PC (Mindich, 2004). Next, the polymerase replicates the ssRNA segments producing a dsRNA-filled PC, which is competent for transcription (Makeyev and Bamford, 2000; Pagratis and Revel, 1990). The packaging and transcription result in a dramatic expansion of the PC (Huiskonen et al., 2006).

The P4 packaging motor (205 kDa, ~13 nm in diameter) is similar to hexameric helicases in structure and function (Juuti et al., 1998; Kainov et al., 2003; Mancini et al., 2004). In addition to RNA translocation, it has to unwind any secondary structure in the ssRNA template. X-ray crystallography has provided atomic models for different stages along the φ12 P4 reaction cycle. The P4 structure consists of six identical ATPase subunits around a central channel, through which RNA is suggested to translocate (Mancini et al., 2004). A molecular mechanism for RNA translocation has been suggested where the different states of the monomer propagate around the hexamer ring as opposed to the classic model of DNA packaging in dsDNA bacteriophage, where rotation of the whole packaging motor has been proposed (Hendrix, 1978, 1998; Mancini et al., 2004).

Many important questions in understanding the function of the φ8 packaging motor and its interaction between the PC still remain open. Most importantly, what is the structure of the hexamer in situ (i.e., bound to the PC) and what is the nature of the interaction between the hexamer and the PC? Is the orientation of the hexamer around its symmetry axis locked relative to the PC, which would support a packaging model where there is no rotation of the motors relative to the capsid? How many hexamers are bound to each particle? Is there a special vertex involved in packaging (Pirttimaa et al., 2002) or are there several packaging vertices (Qiao et al., 1995), and is this related to the occupancy of the packaging motor? To shed light on these questions, we have chosen to study the structure of the hexamer in situ using cryo-electron microscopy (cryo-EM) and three-dimensional (3D) image reconstruction.

In order to achieve better signal-to-noise ratio, the conventional icosahedral 3D reconstruction procedure averages information from the 60 asymmetric units of the icosahedrally symmetric capsid (Crowther, 1971). However, this leads to signal loss of any symmetry-mismatched structures due to incorrect averaging. Thus, different approaches have been taken to study these structures when the mismatch occurs only at one vertex (Bubeck et al., 2005; Jiang et al., 2006; Morais et al., 2001). For example, the prolate head of bacteriophage φ29 has been reconstructed using fivefold symmetry, after which the sixfold symmetric lower collar protein bound to a fivefold vertex was reconstructed together with the phage tail without using any symmetry (Morais et al., 2001). To resolve the fivefold ambiguity in the initial particle orientations, five different fivefold equivalent model projections around the symmetry axis (0, 2π/5, 4π/5, 6π/5, and 8π/5) were tested for the best correlation with the particles containing both five- and sixfold objects. This provided the orientation parameters for the asymmetric reconstruction. Although this method was applied on fivefold symmetric prolate phage particles, it would also be applicable to icosahedrally symmetric particles (Morais et al., 2001).

In the Cystoviridae, the packaging motors occupy more than one vertex. Hence the methods described above are not suitable for this application. Additional problems arise: even if the orientation of the hexamer is assumed to be locked relative to the PC shell, fivefold ambiguity exists in its orientation (Morais et al., 2001). At each vertex, the hexamer can adopt five different, but locally equivalent orientations. Globally, however, each particle can have a unique combination of hexamer orientations. Hence when different particles are averaged, there is a loss of signal for the hexamer during reconstruction of the whole PC, even if no icosahedral symmetry is applied (de Haas et al., 1999).

To resolve these problems we have developed an alternative approach. In the present work, we identify the position of the hexamers in micrographs of the PC using information from an icosahedral reconstruction of the PC (Briggs et al., 2005). Then, we subtract away the icosahedral information from the images before classifying the hexamers in order to calculate non-fivefold symmetric reconstructions of individual vertices to study the structure of the φ8 P4 hexamer. We further calculated reconstructions of the hexamer in the context of the core. This in situ approach enabled us to study the distribution of the hexamers and their orientation relative to the PCs. This analysis, together with a comparison to the X-ray structure of φ12 P4, allows us to propose an interaction between the PC and the P4 hexamer. While the approach described here is applied to study a deviation from icosahedral symmetry, it could also be used to study other symmetry mismatches, for example between a symmetric enzyme and an asymmetric substrate.

2. Results

2.1. Classification of the cryo-EM images of the P4 hexamer

We calculated the positions of all 12 fivefold vertices within 11,052 PC images, at a radius of 27 nm (Fig. 1A, upper panel) using the precalculated orientations and origins from an icosahedral φ8 PC reconstruction (Jääli-noja et al., submitted for publication) (Briggs et al., 2005). From these images (271 x 271 pixels, 2.8 Å/pixel), we extracted 132, 624 vertex subimages (100 x 100 pixels). In the vertex images, if the hexamer was not seen from the side, but was tilted towards the viewer, it overlapped the icosahedral PC shell (see Fig. 1B, upper panel). Therefore, to extract the information corresponding to the hexamers alone, we calculated difference images by subtracting projections of the icosahedral reconstruction in corresponding views from the images.
enabled us to analyse views of the vertex that would otherwise be obscured by the dominating protein shell (see Fig. 1B, lower panel).

Icosahedral orientations of the PCs were used to calculate the orientation of each PC around its fivefold symmetry axis for each given vertex (angles $\alpha$, $\beta$, and $\gamma$) (Briggs et al., 2005). The extracted vertex images were rotationally aligned in the plane of the image around the in-plane angle ($\alpha$) so that the fivefold symmetry axis was vertical and all the vertices were pointing upwards. Next, the vertex images were grouped into orientation groups according to the out-of-plane angle ($\beta$, Fig. 1B). These alignment and pre-classification steps are crucial for the success of further classification and reduce the angular space that needs to be sampled in the search for the orientation of the symmetry-mismatched object (Briggs et al., 2005). Further, where the structure is coaxial with the icosahedral fivefold vertex, these steps reduce the rotational degrees of freedom from three to one in such cases. Two of the angles of the symmetry-mismatched object ($\alpha'$ and $\beta'$) are then identical to the corresponding two vertex angles ($\alpha$ and $\beta$), and only the last angle ($\gamma'$) needs to be found (see Fig. 2 for angle definitions).

Variation within a $\beta$-orientation group can arise from three principal sources (Briggs et al., 2005). First, the symmetry-mismatched structure can be present or absent. Second, the structure and other components of the particle may exhibit conformational variability. Third, there is rotational variation around the symmetry axis. Although the $\alpha$-angle is zero for all of the rotationally aligned vertex images and the $\beta$-angles are similar for all vertices within the same $\beta$-orientation group, the orientation of the icosahedral particle around its fivefold symmetry axis ($\gamma$) still gives one rotational degree of freedom. The structure of interest itself further contributes to this rotational variation having different values of $\gamma'$, and also possibly of $\alpha'$ and $\beta'$.

(Fig. 1A, middle panel; see Section 4) (Booy et al., 1991). This resulted in images containing hexamers and non-icosahedrally ordered RNA (Fig. 1A, lower panel), and enabled us to analyse views of the vertex that would

Fig. 1. Extraction and classification of the vertex images. (A) A key is shown on the left to illustrate the visualization of 2D raw images of phi8 core particles (upper panel), projections of the 3D model with icosahedral symmetry calculated in equivalent view angles (middle panel) and the resulting images after subtracting the 3D model projection from the 2D raw images (lower panel) on the right. The striations from the RNA still contribute significantly to the image. Positions of the vertices are indicated with dots for one of the raw images. (B) Pre-classification of the extracted vertex images based on out-of-the plane angle ($\beta$) into 15 classes with 12° angular sampling (only the first eight are shown). The $\beta$-orientation group averages are shown for the original images (upper panel) and for the difference images, where the icosahedrally ordered parts of the PC are subtracted (lower panel). The position of the PC shell (filled arrow heads) and the first layer of RNA (empty arrow heads) are indicated for six $\beta$-orientation groups in the upper panel. Notice how these features overlap the hexamer density in some views. (C) The eigenimages (1–8) considered in the classification are shown for two different $\beta$-orientation groups. (D) The hexameric nature of the object is evident in class-averages after MSA and classification shown for the same $\beta$-orientation groups as in C. In (B)–(D), the images are shown in smaller boxes (64 x 64) than that which was used for the final processing. The average value of $\beta$ for each group is indicated below each panel.

Fig. 2. Angle definitions. The three Euler angles ($\alpha'$, $\beta'$, and $\gamma'$) are illustrated for the symmetry-mismatched object (barrel) relative to the image plane (grid). $\alpha'$ is the rotation in the plane of the image. $\beta'$ is the angle out of the plane of the image. $\gamma'$ is the angle around the symmetry axis of the object. The angles for the icosahedral vertex ($\alpha$, $\beta$, and $\gamma$) follow the same convention (Briggs et al., 2005).
To study this variation, we classified the images within each \( \beta \)-orientation group based on their appearance using multivariate statistical analysis (MSA) and visual inspection of eigenimages (Fig. IC and D). The eigenimages revealed the prominent features in the images (Fig. IC). Using the first three eigenimages in the classification, we could determine the occupancy of the hexamer from the side views (\( \beta = 90^\circ \pm 18^\circ \); Fig. 3) following the statistical analysis published earlier (Briggs et al., 2005). The hexamer was observed in 83% of the vertices. Some hexamers may stay undetected in some imaging conditions (e.g., particular defocii) due to their small size. The rest of the eigenimages reflected the rotational variation of the hexamer or residual components of the RNA that could not be completely subtracted (Fig. IC). Class-averages of the hexamers at higher angles (\( \beta = 90^\circ \pm 51^\circ \); Fig. 1D, left panel) revealed a hexameric shape, consistent with data from the purified protein (Kainov et al., 2003).

### 2.2. Three-dimensional reconstruction of the P4 hexamer

To calculate 3D reconstructions of symmetry-mismatched structures, such as the P4 hexamer, we further developed the method that we have reported previously (Briggs et al., 2005). Two approaches were taken: the reconstruction was calculated (i) from the subtracted images applying an appropriate symmetry and (ii) from the area of interest without any symmetry. The former exploits the inherent symmetry of the structure and thus possibly reveals detail at higher resolution. In addition, one does not need to assume that the orientation is locked relative to the larger particle, since the symmetry-mismatched structure is considered in separation. However, the latter approach may reveal contacts between the symmetry-mismatched components and reveal deviations from the assumed symmetry of the object of interest, even if the signal is reduced.

First, we reconstructed an initial low-resolution 3D model of the hexamer. The data were divided into 30 \( \beta \)-orientation groups with 6° angular sampling, and these were visually judged to select class-averages that clearly revealed the hexamer. Care was taken to include different views. For the chosen class-averages, the angular reconstruction method (van Heel et al., 2000) was used to determine all the three orientation angles (\( \alpha' \), \( \beta' \), and \( \gamma' \)). Since we could also estimate \( \alpha' \) and \( \beta' \) from the icosahedral vertex orientations \( \alpha \) and \( \beta \), we could compare the values from angular reconstruction to the estimated angles. Only the class-averages, for which angular reconstruction provided values \( \alpha ' \) and \( \beta ' \) similar to \( \alpha \) and \( \beta \), were used. This allowed us to calculate an initial model using sixfold symmetry.

Next, we used model-based refinement (projection matching), which is a more convenient approach than angular reconstruction as two of the angles are already known (\( \alpha' \) and \( \beta' \)), and only one needs to be determined (\( \gamma' \)). We projected the initial model in different \( \gamma' \)-angles for each of the 60 \( \beta \)-orientation groups (3° angular sampling in \( \beta' \) and \( \gamma' \)). Projection matching determined the \( \gamma' \)-angle for each hexamer class-average (200 classes in each \( \beta \)-orientation group). In the final 3D reconstruction, the structure was resolved to 15-Å resolution as estimated based on the

![Fig. 3. Analysis of the hexamer occupancy. The vertices in each \( \beta \)-orientation group were classified based on the presence (top row) or absence (bottom row) of the hexamer. Class-averages for three \( \beta \)-orientation groups are shown for both classes.](image)

![Fig. 4. Resolution assessment of the reconstructions. (A and B) Fourier shell correlation curve is shown for the P4 hexamer reconstruction calculated with sixfold symmetry (A) and for the vertex reconstruction calculated without symmetry (B) (solid line). The resolution of the models is given by the spatial frequency at which the curve drops below the noise level (3σ curve, dashed line). Negative correlations are not shown in the plots.](image)
Fourier shell correlation (FSC) 3σ criterion (van Heel and Schatz, 2005) (Fig. 4A).

2.3. Comparison of the P4 hexamer density to an X-ray structure from a related virus, φ12

This work provided the first 3D structure of φ8 P4. However, an X-ray structure exists for different states of P4 from a related Cystovirus, φ12 (Mancini et al., 2004). The two protein monomers are similar in size (321 amino acids in φ8 P4, 331 in φ12 P4) but have negligible sequence identity (Mancini et al., 2004). Computational fitting of φ12 P4 structure (Chacon and Wriggers, 2002) provided a valuable comparison revealing the similarities and differences between the structures of the two proteins (Fig. 5). The best fit was with the ordered part of the C-terminus facing the particle (Fig. 5A). This result is consistent with hydrogen–deuterium exchange data for the P4 orientation using RNA-naïve φ12 PCs (Lisal et al., 2006).

The core of the monomer fold fits well in the reconstruction (Fig. 5B and C). The major difference in the appearance of the φ8 and φ12 hexamers is their overall shape. The φ8 reconstruction reveals a flower-shaped particle (diameter ~130 Å, height ~60 Å) with six petals. This is consistent with cryo-EM image analysis and neutron scattering of purified φ8 hexamers (Kainov et al., 2003). The φ12 P4 structure does not occupy the peripheral petal densities of the φ8 P4 reconstruction. In contrast, it is much rounder than φ8 P4 when viewed down the symmetry axis (Fig. 5B). This is due to an N-terminal α-helix in φ12 P4, which lies perpendicular to the symmetry axis (Fig. 5B, arrow). The sequence alignment suggests that the N-termini are of a similar length in both proteins (Mancini et al., 2004). Thus in φ8, the N-terminal helix most likely adopts a different conformation, possibly contributing to the peripheral petal density (Fig. 5C, asterisk). The other main differences in the structures are at the top of the molecule (Fig. 5D). In φ12, the top is a β-domain (residues 39–123). Based on sequence alignment (Mancini et al., 2004), two β-strands are deleted in φ8 P4, possibly explaining the lacking density in the reconstruction.

Both hexamers have a funnel-like central channel, through which the RNA could be translocated (Mancini et al., 2004). In the φ8 structure, the channel diameter is 13–33 Å (21–25 Å in φ12 (Mancini et al., 2004)) and the channel is occluded at the bottom. The channel is probably occluded in the φ12 structure, too, but the residues involved are not modelled in the X-ray structure due to disorder (Mancini et al., 2004).

At the bottom of the molecule there is a bulge of density into which the φ12 P4 C-terminal β-sheet fits nicely (Fig. 5D and E). This supports the conclusion that this structure is conserved between φ8 P4 and φ12 P4. In φ12 P4, residues 300–331 are either completely or partially disordered, depending on the subunit. In the φ8 P4 reconstruction, we do not resolve density for these residues either, although the C-termini are of similar length in the two proteins (Mancini et al., 2004). These residues are suggested to be flexible and to connect the P4 to the underlying PC shell.

2.4. Interaction between the P4 hexamer and the PC shell—asymmetric reconstruction of the vertex

We calculated a 3D reconstruction of the P4 hexamer together with the underlying PC shell. No symmetry was applied during the reconstruction. First, the data in each β-orientation group were further divided into γ-orientation groups (3° sampling in both β and γ). MSA and classifica-
tion were then performed in each of these \( \beta \gamma \)-orientation groups; variation now arising from the orientation of the hexamer alone (six classes in each \( \beta \gamma \)-orientation group).

To construct a starting model for the model-based orientation search, the sixfold symmetric reconstruction of the hexamer was added to a fivefold symmetric reconstruction of the PC shell part of the vertex. To be able to average between different class-averages in the reconstruction, the orientation of the hexamer was assumed to be locked relative to the PC shell. In the orientation search, the five different possible \( \gamma \)-angles (\( \gamma, \gamma+2n/5, \gamma+4n/5, \gamma+6n/5, \) and \( \gamma+8n/5 \)) were tested (Morais et al., 2001). The model was projected only in these angles and the best correlating model projection defined the correct orientation. The vertex structure was resolved to 15-Å resolution as estimated based on FSC (Fig. 4B; 3\( \sigma \) criterion) (van Heel and Schatz, 2005).

Finally, we studied the effect of possible model bias using an alternative starting model, where the hexamer was rotated to a maximally different orientation (6° rotation in \( \gamma \)). Using this starting model gave a result differing only at the noise level, suggesting that the observed features in the reconstruction arise from the data and are not affected by the initial model.

The asymmetric vertex model revealed the shape of the hexamer on top of the fivefold symmetric PC shell (Fig. 6A and B), similar to the sixfold symmetrized reconstruction calculated without the PC shell (see Fig. 5B–E). The sixfold symmetry axis of the hexamer is not coaxial with the fivefold symmetry axis of the vertex, but is shifted 8 Å (Fig. 6B, arrow). The base of the hexamer reveals a partial opening of the hexamer ring (Fig. 6B). One of the C-terminal bulging densities is directly on top of one of five ridges in the P1 shell (Figs. 5E and 6C; encircled). The distance between these two features is \( \sim 17 \) Å (Fig. 6D). At a very low threshold (0.3\( \sigma \)) the density of the hexamer becomes connected to the P1 shell at this position. At other positions, no such connection is seen.

3. Discussion

This paper demonstrates a method for the reconstruction of symmetry-mismatched structures in icosahedral viruses and its application to the hexameric packaging enzyme P4 located at the fivefold vertices of the \( \phi 8 \) PC. The presented reconstructions provided the first structures of the P4 bound to the PC. The method allowed the statistical analysis of the occupancy and the determination of the orientation of the hexamer relative to the PC shell.

3.1. The structure of the \( \phi 8 \) packaging motor in situ

One of the major goals in this study was to determine which face of the hexamer interacts with the PC shell as this is crucial for evaluating the structural model of the RNA translocation process (Mancini et al., 2004). Our comparison between the \( \phi 8 \) and \( \phi 12 \) P4 structures revealed that the \( \phi 8 \) hexamer is oriented with its C-terminal side towards the PC, in accordance with the direction proposed in the model. This result is consistent with a parallel study using hydrogen–deuterium exchange (HDX), where the C-terminal side of the \( \phi 12 \) P4 ring was protected from exchange when it was bound to the unexpanded \( \phi 12 \) PC (Lísal et al., 2006).

Our comparison between the \( \phi 8 \) and \( \phi 12 \) P4 structures provided also detailed insights into the \( \phi 8 \) P4 structure, HDX providing complementary information (Lísal et al., 2005). First, the N-termini of the two P4s differ in their conformation (Fig. 5B and C). This is supported by the observation that although the \( \phi 12 \) P4 N-terminal helix is on the surface of the molecule, the corresponding peptide in \( \phi 8 \) P4 was protected from the solvent in HDX. Second, at the top of the molecule, there was another peptide exposed in \( \phi 12 \) P4, but protected in \( \phi 8 \) P4. This is consistent with the difference we saw in our structural comparison (Fig. 5D). Third, the C-terminus exchanged rapidly. This structure is partially disordered in the \( \phi 12 \) P4 (Mancini et al., 2004) and as expected, it is disordered in our sixfold symmetric hexamer reconstruction, too.

What is the biological significance of the observed difference in the N-termini of \( \phi 8 \) and \( \phi 12 \) proteins? As seen in many other viruses, capsid protein termini act as conformational switches mediating the interactions at different quasi-equivalent positions in the viral capsid (Abrescia et al., 2004; Harrison et al., 1978; Liddington et al., 1991; Wikoff et al., 2000). However, no X-ray structures exist for icosahed-
dral viruses that would illustrate how a symmetry mismatch is incorporated. Here we postulate that conformational switches could accommodate symmetry mismatches as well. For example, in the φ6 nucleocapsid, the P4 hexamer contacts the surrounding nucleocapsid surface layer (protein P8) from the sides (Huiskonen et al., 2006). Interestingly, protein P8 is missing in φ8 (Mindich et al., 1999) (Jäälinnoja et al., submitted for publication). Thus the difference in the N-terminal helix could well be due to different interactions between P4 and P8. However, structural information on the φ12 nucleocapsid is required to test this hypothesis.

The asymmetric φ8 vertex reconstruction sheds light on the function of the hexamer and its interaction with the PC shell. The reconstruction revealed features that would not be resolved if the hexamer was studied in isolation. The binding of the hexamers may lead to breakage of the sixfold symmetry and partial ring opening, which may facilitate the loading of the RNA segments. This distortion from true sixfold symmetry may be caused by the symmetry mismatch. However, RNA has been observed to stimulate transient φ8 P4 ring opening in solution (Lisal et al., 2004). We propose that binding of P4 to the PC results in a controlled ring opening, which may further assist RNA loading and thus the initiation of RNA translocation into the capsid.

What is the nature of the interaction between the symmetry-mismatched components? We estimate that there are roughly 10 hexamers per PC after purification from the virion. This seems rather wasteful if only a maximum of three hexamers are required to package the three ssRNA segments of the genome (Mindich, 2004; Pirttimaa et al., 2002). However, the P4 hexamer also has a critical role in regulating the fidelity of PC assembly. It appears to introduce curvature into the growing P1 shell to produce closed shells (Kainov et al., 2003). Thus hexamers are probably required at most vertices. In addition there must be specific interactions with the capsid, with each hexamer binding up to five equivalent positions on the PC shell. The smallest distance between P4 and the shell in our reconstruction was 17 Å. The last 30 disordered C-terminal residues can easily span this distance. The reconstruction gave some indication that the hexamer may be bound by only one subunit. Flexible binding of the hexamer to the PC shell might not only accommodate the symmetry mismatch but could also be important to accommodate conformational changes during packaging (Mancini et al., 2004). Whether the same subunit is bound permanently, or whether this changes during packaging is still an open question. The specific interaction sites between φ8 P1 and P4 could be addressed by co-crystallization experiments like those used to determine the contacts between the adenovirus trimeric fibre on its pentameric penton base (Zubieta et al., 2005).

The function of the hexamer is not just to translocate RNA, but to translocate the RNA into the PC. Therefore, understanding the interaction with the PC is critical for understanding the translocation mechanism. It has been suggested that with dsDNA viruses the packaging motor may rotate relative to the capsid during translocation (Hendrix, 1998). In the case of φ8, we tried to address the question of P4 rotation by looking at the resting state of the hexamer after packaging. The success of the asymmetric reconstruction approach provided direct evidence that the φ8 hexamer does not rotate freely, but on the contrary the majority of the hexamers are in a fixed orientation relative to the PC shell. This is in accordance with the current model of P4 function, where rotation of the whole motor with respect to the capsid is not required (Mancini et al., 2004). However, rotations in 2π/5 steps would remain indistinguishable, since these orientations are equivalent.

Purified φ6 P4 hexamers have been observed to form dodecamers with two hexamers facing each other (de Haas et al., 1999). Would these double hexamers allow both active RNA translocation into and out of the particle or be responsible for a special packaging vertex (Pirttimaa et al., 2002)? The classification analysis carried out in this study gave no indication of the presence of double hexamers in the φ8 PCs. This is consistent with data indicating that P4 is a passive conduit for the exit of ssRNA transcripts (Kainov et al., 2004). Our data also indicate that if there is a special vertex responsible for the packaging of ssRNA, it is structurally indistinguishable from the other vertices at this resolution.

3.2. Novel approach to study symmetry-mismatched structures

In addition to the previously discussed advantages of studying the object of interest in situ, the method has two additional benefits: (i) it provides the coordinates of the object of interest in the images. This circumvents the common problem of picking very small particles from low signal-to-noise cryo-EM micrographs. (ii) If the object of interest is bound to an icosahedrally symmetric particle, different views of the object are automatically present in the data, and the problem with a possibly preferred orientation of the object can be avoided. However, exploitation of all the views of the object depends on subtracting the contribution of the larger particle away in the images. In this application, the number of PCs present on each micrograph limited the resolution of the calculated icosahedral models and their projections which were used in the subtraction. Thus we able to subtract the signal from the PC shell only up to 15-Å resolution, which also set an upper limit for the resolution of the hexamer reconstruction.

When compared to conventional single particle analysis, there are at least two main factors restricting the resolution achievable by the approach presented. The largest factor may be the thickness of the vitrified water layer. Since the objects considered for reconstruction in this paper are attached to much larger particles, it is the diameter of the large particle that defines the minimum thickness of the layer. Thus, the signal-to-noise ratio may be much lower here than in the single particle analysis of isolated objects.
possibly embedded in a much thinner layer. Another factor is the dominating signal from the rest of the particle, such as the RNA component that could not be completely subtracted in this application.

We have demonstrated the usefulness of the approach by studying a relatively small symmetry-mismatched component bound to a large icosahedrally symmetric particle. We expect the approach to be of general use in studying particles with symmetries other than icosahedral symmetry, too. The classification analysis of very small symmetry-mismatched structures is assisted by exploiting the predetermined orientation of the larger particle to define the position of the structure and to reduce its rotational degrees of freedom. The approach should be applicable, for example, in studying the cargo molecules in nuclear pore complexes (Akey and Radermacher, 1993) and folding intermediates inside GroEL/ES (Farr et al., 2000).

4. Methods

4.1. Subtraction of the model projections

In certain views the hexamer overlaps the icosahedral capsid (Fig. 1B). To make use of these views in the reconstruction, we removed the contribution of icosahedrally ordered parts from the φ8 core images by subtracting equivalent projections of the icosahedral model from the raw images (Fig. 1A and B). First, three-dimensional reconstructions were calculated for each separate micrograph, without correcting for the contrast transfer function (CTF) (Baker and Cheng, 1996). The number of particles per micrograph (195 on average) was sufficient to calculate the models to 15-Å resolution. Bsoft (Heymann, 2001) was used in further image-processing steps unless stated otherwise. The models were masked in EMAN (Ludtke et al., 1999) to exclude areas of the improperly averaged vertices, and projected and translated according to the predetermined view and origin parameters (Jäälinoja et al., submitted for publication) of the corresponding raw images. Next, we adjusted the intensity levels in these CTF-equivalent projections to match the intensity levels in the raw images. The scale and shift parameters for this adjustment were determined from radial profiles calculated for both the raw images and the projection within the icosahedrally ordered φ8 core protein capsid. The scaled projections were subtracted from the raw images. Finally, the CTF was corrected by flipping the phases in the resultant images and the images were low-pass filtered to 15-Å resolution.

4.2. Multivariate statistical analysis and three-dimensional reconstruction

Hexamer subimages were extracted from the PC images as described previously (Briggs et al., 2005). MSA, classification and 3D reconstruction of the hexamer and vertex were performed using IMAGIC5 (van Heel et al., 1996). For each β-orientation group, an individual mask was generated to accurately define the area of the hexamer. For the initial starting model, eigenimages not describing the variation arising from the hexamer were visually selected and their weights set to zero for classification. Vertex side views were classified into three classes using the first three eigenimages to study the presence or absence of the hexamer. Angular reconstitution provided an initial estimate for the hexamer Euler angles and a 3D starting model. The angles were refined for each hexamer class-average using a projection matching technique, where the current best model was projected over selected β’ and γ’-angles. Next, the model projections were correlated to the class-averages allowing rotations around z’ and translations. The best correlating reference then defined the three Euler angles. A three-dimensional reconstruction was calculated from the class-averages. We selected the best correlating class-averages for calculation of a new 3D model. The selection was done in two phases. First, only class-averages with a correlation coefficient >0.50 were used to calculate a test reconstruction. Of these class-averages only the top 40% were included in the actual reconstruction, based on correlation between a class-average and a projection calculated from the test reconstruction in the corresponding orientation. The process was iterated, until no improvement was seen in the model resolution. Six iterations were sufficient for both the hexamer and vertex reconstructions. This procedure for iterative reconstruction was implemented as a Unix shell script, which executes the relevant IMAGIC5 batch files with the given parameters. The resolution of the models was estimated from the Fourier shell correlation curve calculated between two models reconstructed from two halves of the data. The spatial frequency where the correlation dropped below the modelled noise curve (3σ) determined the resolution (van Heel and Schatz, 2005).

The amplitudes of the final hexamer model were scaled using the X-ray structure of a homologous protein (φ12 P4, PDB:1W4C). CoLoRes (Chacon and Wriggers, 2002) was used to fit the φ12 P4 structure (one hexamer extracted from PDB:1W4C) into the reconstruction. The full angular range was considered. The partially disordered C-terminus was not considered in the fit and was not visualized. The symmetry of the hexamer was tested from the asymmetric reconstruction of the vertex using EMAN (Ludtke et al., 1999). No amplitude scaling was performed on the asymmetric vertex reconstruction.

The reconstructions have been deposited in the EMDB, at the European Bioinformatics Institute with the accession codes: EMD-1255 (hexamer model) and EMD-1256 (vertex model).

Acknowledgments

We thank R. Tuma for his comments on the manuscript. The work was funded by the Academy of Finland, Finnish Centre of Excellence in Virus Research 2006–2011 (S.J.B., Grant No. 1213467) and the Wellcome Trust (S.D.F., Grant No. H5RCYV0). S.D.F. is a Wellcome Trust Principal
References


