Bluetongue Virus Structure and Assembly

Polly Roy

Department of Pathogen Molecular Biology, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT, UK

Corresponding author: Roy, Polly (polly.roy@lshtm.ac.uk)
Abstract
Bluetongue virus (BTV) is an insect-vectored emerging pathogen of wild ruminants and livestock in many parts of the world. The virion particle is a complex structure of consecutive layers of protein surrounding a genome of ten double-stranded (ds) RNA segments. BTV has been studied as a model system for large, non-enveloped dsRNA viruses. Several new techniques have been applied to define the virus-encoded enzymes required for RNA replication to provide an order for the assembly of the capsid shell and the protein sequestration required for it. Further, a reconstituted in vitro system has defined the individual steps of the assembly and packaging of the genomic RNA. These findings illuminate BTV assembly and indicate the pathways that related viruses might use to provide an informed starting point for intervention or prevention.

Introduction
Bluetongue virus (BTV) is a member of the Orbivirus genus within the Reoviridae family, one of the largest families of viruses. These viruses represent a diverse group of double-stranded (ds) RNA viruses with capsids consisting of 1-3 protein layers and infect a wide range of hosts including human, animal, plant and insect. Consequently, the host receptor-binding proteins of these viruses, which occupy the outer layer of the virus particle, are largely non-homologous. Orbiviruses represent an intriguing system for study. They are vectored to particular vertebrate species (e.g., sheep, cattle, horses, deer, etc.) by arthropods (gnats, ticks, or mosquitoes depending on the virus), hence determining their geographic distribution. BTV, with 27 serotypes, is one of the most widespread animal pathogens and acts as an important representative of this class of large non-enveloped viruses. In recent years considerable progress has been made in determining the BTV structure at the atomic level, providing new insights into the structure-function relationships among the proteins of the virus particle and in virus assembly/disassembly. The introduction of mutations into the replicating viral genome by reverse genetics (RG) systems, first developed for BTV, has opened up a new opportunity to dissect the various stages of virus replication and assembly. Further, the development of novel assays has unravelled key aspects of the genome packaging mechanism. Consequently, BTV represents one of the best characterized viruses and selected aspects of its entry, replication and assembly are reviewed here.

Overview of the replication cycle
Infection by the BTV particle (~550S) is established when the inner capsid of the double-capsid particle, the “core”, translocates across the endosomal membrane following virus uptake, a process performed by the two outer capsid proteins, VP2 and VP5 [1-6]. The larger VP2 protein (110KDa) binds to surface glycoproteins of unknown identity and facilitates clathrin mediated endocytosis of the virion, while the VP5 protein (60KDa) penetrates the host cell membrane and delivers the 470S core particle into the host cytosol. The process has two distinct stages: VP2 senses the pH in the early endosome (6.5-6.0) and detaches while the remaining particle continues to the late endosome, where VP5 senses a lower pH (~5.5) and gains membrane “fusion” activity, allowing the core to escape through the endosomal membrane. The final product of disassembly for all members of the Reoviridae is the transcriptionally active double-capsid particle able to initiate transcription of the genomic RNAs. The surface layer of the BTV core is formed by VP7 which coats the internal VP3 layer which in turn encapsidates the viral tri-protein transcriptase complex (TC), VP1 (polymerase), VP4 (capping enzyme) and VP6 (helicase/RNA packaging) and the 10 genomic dsRNA segments (S1-S10) [7,8]. Within the core, each segment is repeatedly transcribed and modified by the TC and ten transcripts, each with a 5’ cap 1 structure, but lacking polyadenylation at their 3’ terminus, are released into the cytosol. These ssRNAs serve as mRNA for the synthesis of viral proteins within the cytoplasm and later also act as templates for dsRNA genome synthesis. In addition to the seven structural proteins, four non-structural proteins, NS1-NS4 are synthesized in infected cells, each supporting virus growth at different stages. Of these, NS1 preferentially promotes translation of BTV ssRNAs enhancing viral protein synthesis, while NS2 forms a fibrillar network, viral inclusion bodies (VIBs), which recruit viral ssRNAs and protein components required for genomic packaging, replication and core assembly [2,8-10]. Following assembly newly synthesised cores are released from VIBs and utilise the intracellular vesicular sorting and exocytosis pathway in doing so [11-16].
This process is facilitated by NS3, and the outer capsid proteins VP5 and VP2 are acquired in the process to produce mature particles that egress from the host cell and propagate the infection.

**BTV Outer capsid**

Cryo-electron microscopy (Cryo-EM) and image analysis of BTV has revealed the well-ordered structural arrangement of the virion. The outer shell has 60 triskelion-like VP2 spikes, surrounding 120 globular trimers of VP5, the membrane penetration protein [17-20]. The detailed structures of these two proteins and the whole particle have been recently defined at near-atomic resolution (3.5A) suggesting how each may function during virus entry into the cell (Fig.1) [21]. The receptor binding VP2 monomer is divided in four distinct domains, a hub, a hairpin, a pyramid-shaped body and a highly flexible external tip, where host neutralizing antibodies bind. The hub domain has a 10-stranded lectin-like β-barrel fold and drives the monomer-monomer interaction of the trimer as well as possesses a sialic acid binding pocket. At the interface of the hub and the body domains a well-coordinated tetrahedron of a typical zinc-finger motif, CCCH, is found and is shown to be responsible for pH triggered VP2 conformational changes in the early endosome.

A series of biochemical and molecular studies have established that VP5 is the membrane penetration (“fusion”) protein and the structural data corroborate this and reveal how VP5 may function [4,5,21]. The VP5 trimers possess a highly compact globular fold predominated by α-helices with a central coiled-coil motif facilitating trimerization, analogous to membrane viral fusion proteins such as the stalk of Human Immunodeficiency Virus gp41. Each VP5 monomer has three distinct domains: a flexible dagger, which is sequestered in the canyons underlying the core surface, a helix-rich unfurling domain, and an anchoring domain with membrane-interaction elements. Notably, a β-meander motif in the anchoring domain contains a histidine cluster that could sense the late-endosomal pH to release four putative membrane-interaction elements. Exposing BTV to low pH causes complete detachment of VP2 and dramatic refolding of the dagger and unfurling domains of VP5 and mutagenesis studies have supported the suggested pH-sensing mechanisms. For example, using purified VP5 and liposomes [22], it was shown that VP5 utilizes a late endosome specific lipid to achieve membrane penetration. The process is somewhat similar to the requirements of some enveloped viruses, such as the influenza HA2 protein, which also enter cells through the late endosome and utilize histidine residues as pH sensors to subsequently undergo significant conformational change at the late endosomal pH [23,24]. Other non-enveloped viruses employ a variety of strategies for cell entry with no similarity to the BTV mechanism. The atomic structures of BTV VP2 and VP5 at high pH bear no similarities to those in other dsRNA viruses [25,26] the “exposed” VP5 and rotavirus penetration protein are filamentous [27]. Reovirus μ1 is cleaved to form an N-terminal myristoylated ‘fusion peptide’ while rotavirus VP4 is cleaved to the receptor binding domain VP8* and the membrane penetration protein VP5 [25,28-30]. BTV VP5 does not appear to be proteolytically processed and is likely to fold into a metastable conformation following translation.

**Viral core and its components**

Unlike the whole particle, the core (~470 S) is very robust and can be generated by proteolytic treatment of purified virus. The 3D structure of the icosahedral core shows a triangulation number of 13 (T=13) and an early atomic structure of a recombinant VP7 protein revealed an unusual organisation that provided a paradigm in the field [17,31-34]. The core surface is formed of 260 VP7 (38 kDa) trimers, arranged around 132 channels as six-membered rings, with five-member rings at the vertices (Fig.2). The five quasi-equivalent trimers form a visible protomeric unit and each trimer consists of two distinct domains, ‘upper’ (an anti-parallel β-sandwich) and ‘lower’ (mainly α-helical) [34]. The top domain of one monomer sits on the lower domain of an adjacent monomer and the interactions between monomers are extensive. The lower domains are attached to the inner layer of 60 dimers of VP3 (103 kDa) with a T=2 symmetry. Each dimer consists of an ‘A’ and ‘B’ type conformation, each with three domains, the rigid ‘carapace’, ‘apical’ and ‘dimerization’ domains. The VP3 layer is formed by 12 decamers through their interconnection via the ‘dimerization’ domain with each decamer consisting of five copies of the ‘A’ type molecule around the five-fold apex with five ‘B’ VP3 molecules in between.

There are also pores in the VP3 layer at the threefold axes but in the intact core particle, these are completely blocked by one of the VP7 trimers. In the presence of magnesium ions and nucleotide
triophosphates VP3 and VP7 move outward around the fivefold axis, likely due to the activation of the polymerase complex, to allow the opening of the pore for the extrusion of mRNAs [35]. Interestingly, like BTV VP7, the surface protein shells of all members of the Reoviridae family have a T=13 organization either complete or incomplete [36,37]. The organisation of the underlying T=2 layer with 120 subunits of VP3 is shared by most viruses with segmented dsRNA genomes that have been studied to date, including the yeast virus, L-A [37-39]. Such architecture is probably preserved for essential biological functions such as keeping the transcriptional complex and genome in the correct configurations across the different viruses.

Organization and structure of the internal minor core proteins
The organisation of the remaining three minor proteins and the genome within the core are less well defined, although the entire genome of 10 dsRNAs comprising 19,219 nucleotides has been modelled as four internal concentric layers [34,40]. Reconstruction of recombinant core-like particles (CLPs) of VP3 and VP7 lacking the genome but containing two of the three minor proteins, VP1 and VP4, exhibited a flower-shaped density directly beneath the icosahedral 5-fold axes and attached to the underside of the VP3 layer [19,31]. The location of the VP1-VP4 complex is consistent with the core organization model. However, the third internal core protein VP6, which has RNA binding properties and resembles an RNA helicase, fails to incorporate into recombinant CLPs [41,42]. The exact method of its encapsidation remains elusive. It may be that it is recruited by the newly synthesised ssRNAs during assembly. Notwithstanding these uncertainties, biochemical data from recombinant proteins has unambiguously elucidated their function. The VP4 (78 kDa) protein, alone catalyses the formation of a cap1 structure at the 5' termini of core RNA transcripts, orchestrated by a series of reactions including RNA triphosphatase, guanylyltransferase and two methylase activities [43,44]. The crystallographic structure of recombinant VP4 shows that the protein has distinct domains with each functional domain arranged sequentially to facilitate each catalytic activity [45]. Unlike VP4, a high resolution VP1 (150 kDa) structure is still lacking but a model based on the structure of λ3, the Reovirus polymerase, shows a typical RdRp organisation with a central polymerase and N-terminal and C-terminal domains [46,47]. This data suggests a common organisation for the RdRps of the Reoviridae although finer structural differences will be required to explain the observed biochemical discrepancies in processivity and priming among them. Unlike the reliance on dinucleotide priming observed for Reovirus λ3 and Rotavirus VP1, BTV VP1 possesses a high intrinsic processivity allowing primer independent dsRNA synthesis [48,49].VP6 (36kDa) is an ATP dependent RNA helicase essential for RNA packaging but its high resolution structure remains unclear. However, nuclear magnetic resonance analysis has revealed a large structured domain with two large loops that exhibit significant flexibility consistent with a molecule that expands to accommodate variously sized RNAs. Recent experimental data using rotavirus DLPs have revealed that RNA segments are not promiscuous in their localization as each segment is associated with a specific channel, and thus replication complex [50]. Due to the highly similar architecture of BTV cores and Rotavirus DLPs the same singular association of segments with replication complexes and the fivefold vertex is likely to be the same.

Assembly pathways of core proteins
Virus inclusion bodies (VIBs), or protein aggregates, are hallmarks of BTV and other members of the Reoviridae family and are the functional sites for genome replication and subviral assembly. Newly synthesised viral proteins interact with sequestered viral RNA molecules prior to capsid assembly and dsRNA synthesis within these bodies. In BTV infected cells, VIBs are predominantly comprised of phosphorylated NS2 protein. Newly generated subcore proteins and the 10 RNA transcripts are recruited directly by NS2 to these VIBs where they assemble to form a subcore with VP3 and subsequently acquire VP7 to generate a stable core particle. Although the exact mechanism is not defined, the current data suggests that NS2 interacts directly with each of these components and its interaction, particularly with BTV ssRNAs, is highly specific. Thus, during the early stage of BTV replication, the three enzymatic proteins and VP3 form the first replicase complex facilitated by NS2, as demonstrated by a series of RG-based studies [51,52]. VP7 is not directly involved and serves only to stabilize the replicase complex. Co-expression of only the recombinant
structural proteins enables assembly of CLPs that lack an RNA genome demonstrating that these proteins have an intrinsic self-assembly pathway and this system has allowed the finer stages of the VP3 and VP7 assembly process to be mapped [53-55]. For example, a deletion mutant lacking the dimerization domain of VP3 abolishes subcore formation but generates VP3 dimers and dimers suggesting an assembly pathway in which VP3 dimers form decamer intermediates and that decamer-decamer interactions via the dimerization domain induce the assembly of the complete VP3 subcore. Further, such deletion also abolishes RNA binding by VP3, indicating that this property is either located within the dimerization domain or dependent on the larger assembly. Similar studies with VP7 have shown that the trimer is essential for VP7 layer deposition onto the VP3 layer [56]. Recombinant protein assembly data have further been supported recently by an *in vitro* transcription/translation, cell-free assembly (CFA) system [57]. In this assay, each of the polymerase complex proteins was translated *in vitro* and subsequently incubated with ten ssRNA transcripts, followed by incubation with VP3 and VP7 sequentially. The product generated was the core particle with a complete set of 10 ssRNAs and the packaged ssRNAs could act as templates for dsRNA synthesis when provided with NTPs and divalent cations prior to the addition of VP7. Furthermore, the assembled particles were infectious in susceptible cells. The assay also revealed that VP6 was essential for ssRNAs packaging and confirmed that VP7 provided stability to the subcore, protecting the packaged ssRNAs from RNase activity. Evidently, BTV structural proteins and RNA segments alone can assemble to form infectious particles in the absence of the non-structural proteins. Thus, NS2 plays a passive role *in vivo* as a condenser of the necessary components required for assembly.

**Assembly of genomic RNA and packaging:**
The mechanism by which BTV and related viruses with multipartite genomes package their genomes into capsids has proved elusive until very recently. The ten BTV RNA segments (3.95kb-0.8 kb) can be segregated in three size classes, large, medium and small [40]. Each segment varies in sequence but shares common complementary short 5’ and 3’ untranslated regions (UTRs) of variable length which include highly conserved hexanucleotides at both ends. Recent evidence shows that sorting of BTV and related viral RNA segments starts by the recruitment of genomic ssRNAs via these UTRs [58-61]. Intersegment interactions leading to complex RNA networks are the driving force for sorting and packaging of a correct set of genomic segments and there is a specific order of segment recruitment from the smallest segment, S10, which triggers RNA-RNA interaction with other small segments to forming an initial complex (Fig.3). Subsequently, the medium to larger size ssRNAs are recruited until the complete genome is packaging into the capsid. The packaging signals in each segment are highly specific and dispersed largely in the UTRs but also in the coding regions. Further, the UTRs, particularly the S10 UTRs, are responsible for triggering the correct secondary structures required for interaction. Oligoribonucleotides, ORNs, complementary to the 3’ UTR of BTV RNAs perturb these RNA networks, preventing encapsidation *in vitro* and inhibiting virus replication *in vivo*. These studies indicate how selective packaging is achieved for BTV and suggest that similar mechanisms may apply to other members of the *Reoviridae* family.

**Interactions between outer and inner capsids**
Newly assembled core must exit VIBs to acquire the two outer capsid proteins and form the mature virion particle. Thus, the timely release of cores from VIBs is an essential step in virion maturation that must be regulated. The phosphorylation of NS2 by cellular phosphokinases appears to serve as the key regulator for core release into the cytoplasm by altering the NS2 morphology [62]. Similar mechanisms are likely to be shared by other members of the family, since viral NS phosphoproteins are also an essential component of VIBs for related viruses. Both VP2 and VP5 proteins are visualized around the VIB periphery, available for encapsidation of the released cores. These proteins interact with the core surface protein VP7 in an assembly reaction that is easily mimicked either by baculovirus expression systems or by an *in vitro* transcription-translation system [31,63-65]. Each protein can be incorporated independent of each other implying that the two proteins are attached directly to the core and do not interact intimately with each other. Once VP2 and VP5 are attached, the particles are no longer transcriptionally active as the pores at the vertices are now occluded. Instead, the fully infectious particle is available for release from the infected cell, hijacking the intracellular vesicular sorting and exocytosis pathway via use of the viral glycosylated NS3
protein. The structural organization of the outer layer represents a drastic mismatch with the underlying VP7 layer but the triangular shaped top portion of VP7 serves as a platform for the deposition of the VP2 and VP5 [20,21]. Each VP2 trimer contacts four VP7 trimers through the base of its hub, body, and hairpin domains with two histidines at the carboxyl end of VP2 located at the interacting interface. In addition, the loop of the VP2 hairpin domain interacts with the jelly-roll top domain of the VP7 trimer. The interaction between VP5 and VP7 is more extensive and all three VP5 domains are involved. Confirming the findings of recombinant expression studies, VP2 interacts with VP5 relatively weakly only via two regions, the hub domain, which interacts with the N terminal helices of the VP5 trimer and the VP2 hairpin loop, which interacts with the anchoring domain of an adjacent VP5 trimer. These weak interactions are consistent with the requirements of VP2 detachment during cell entry when VP2 must dissociate to permit the conformational change of VP5, particularly dagger and unfurling domains.

Conclusion

BTV has been studied extensively as an agent of an economically important emerging disease and as a model system for dsRNA viruses in general. The application of new techniques from a number of disciplines has benefited our current state of knowledge which now encompasses a very complete understanding of the BTV assembly pathway and its replication cycle (Fig.4). These studies have also led to opportunities for the development of new diagnostic reagents, protein and capsid-based vaccines, and novel DIVA compliant attenuated virus BTV vaccines. The same is also beginning to be true for related viruses. Inevitably, certain key questions in structure and assembly remain unclear due to the lack of appropriate tools and assay systems to address them. For example, the exact location of genome precursor RNAs within the capsid or the mechanism by which they are packaged there is not clear. It would make sense that they are positioned at the 12 vertices of the assembling icosahedral capsid where they would be primed for replication and, upon subsequent cell entry, transcription. Such ssRNA pre-positioning in the core is essential for this complex ‘molecular machine’ to execute its functions, specifically to release capped transcripts through specific pores located at the vertices of the particle, directly observed for a BTV relative, rotavirus [50,66,67]. It is likely that state-of-the art single particle cryo-EM and cryo-tomography will be able to address these questions in the near future. Will the precise location of the packaged RNAs and the organisation of the TC complexes be where expected in the icosahedron capsid? What about family members with more or less segments and can additional segments be accommodated? These details will be interesting to discover and are likely to open up many possibilities for the generation of designer particles equipped for a range of functions. The structures adopted by RNAs necessary for their efficient packaging and their precise interaction with partner proteins during assembly process may be unravelled by SHAPE RNA structure analysis in combination with single molecule fluorescence spectroscopy and other physical measures. These are likely to be conserved across serotypes and, as they relate to viral fitness and host range, this data has the potential to suggest novel antiviral strategies that target a common mechanism. If that mechanism turned out to be similar for many members of the Reoviridae the possibility for the development of broadly active antiviral agents emerges.


Figure legends

Figure 1. CryoEM density map of the BTV virion obtained at 3.5-Å-resolution, shown as radially colored surface representation. The outer layer contains 60 VP2 trimers (magenta) and 120 VP5 trimers (green); a middle layer with 260 VP7 trimers (yellow); and an inner layer formed by 120 VP3 monomers (red) almost completely occluded by the VP7 layer. B. The VP5 layer showing the two different trimer conformers in green and cyan. C & D. Ribbon model of the VP2 (C) and (D)
Figure 2. Crystal structure of BTV core particle and two major core proteins. A. BTV core surface at 3.5-Å-resolution, showing 13 copies of VP7, arranged as 5 trimers (colored red, orange, green, yellow and blue) form an icosahedral asymmetric unit. B. The subcore particle showing the arrangements of two different VP3 conformers (in green and red) in each of the 12 decamers. C. Trimer image of the VP7 atomic structure at 2.8 Å resolution. Two domains of the molecule (upper and lower) and the carboxyl and amino termini are indicated. D. VP3 dimer unit, each monomer with 3 distinct domains.

Figure 3. A Schematic of RNA-RNA interaction during packaging. A schematic showing the sequential RNA-RNA interactions of 10 BTV ssRNA segments, initiated by the smallest RNA segment S10.

Figure 4. A summary schematic of sequential assembly of different components of BTV particle. Assembly starts with the interactions of VP1 polymerase and VP4 capping enzymes. In parallel, VP6 hexamer interacts with 10 ssRNA complex. Together they form the polymerase complex and also interact with VP3 decamers at the five-fold axis. Subsequently 12 decamers form the shell the complete shell and 160 VP7 trimers are then deposited in sequential manner to form the highly robust icosahedral core. Core is the scaffold for the deposition of both 120 VP5 trimers and 60 VP2 trimers.