Published in final edited form as:
*J Gen Virol*. 2009 September ; 90(Pt 9): 2191–2200. doi:10.1099/vir.0.012104-0.

# Drosophila A virus is an unusual RNA virus with a T=3 icosahedral core and permuted RNA-dependent RNA-polymerase

**Rebecca L. Ambrose**[1], **Gabriel C. Lander**[2], **Walid S. Maaty**[3], **Brian Bothner**[3], **John E. Johnson**[2], and **Karyn N. Johnson**[1,*]

[1]School of Biological Sciences, The University of Queensland, Brisbane, Australia

[2]Department of Molecular Biology, The Scripps Research Institute, La Jolla, USA

[3]Department of Chemistry and Biochemistry, Montana State University, Bozeman, USA

## Abstract

The vinegar fly, *Drosophila melanogaster*, is a popular model for the study of invertebrate antiviral immune responses. Several picorna-like viruses are commonly found in both wild and laboratory population of *D. melanogaster*. The best studied and most pathogenic of these is the discistrovirus, Drosophila C virus. Among the uncharacterised small RNA viruses of *D. melanogaster* Drosophila A virus (DAV) is the least pathogenic. Historically DAV has been labelled as a picorna-like virus based on the particle size and content of an RNA genome. Here we describe the characterisation of both the genome and virion structure. Unexpectedly, the DAV genome was shown to encode a circular permutation in the palm domain motifs of the RNA-dependent RNA polymerase. This arrangement has only been previously described for a subset of viruses from the dsRNA virus family *Birnaviridae* and the T=4 ssRNA virus family *Tetraviridae*. The 8 Å DAV virion structure computed from cryo-electron microscopy and image reconstruction indicates that the virus structural protein forms two discrete domains within the capsid. The inner domain is formed from a clear T=3 lattice with similarity to the beta-sandwich domain of tomato bush stunt virus, while the outer domain is not icosahedrally ordered but forms a cage-like structure that surrounds the core domain. Taken together this indicates that DAV is highly divergent from previously described viruses.

## INTRODUCTION

As obligate intracellular pathogens, viruses have complex interactions with their hosts. Comparatively little is known about the specific responses to virus infection in insects. Recently, insect responses that are specific for virus infection have been investigated using the genetically-tractable *Drosophila melanogaster* (vinegar fly) model and the most pathogenic of its viruses the dicistrovirus, Drosophila C virus (DCV) (Cherry, 2005; Cherry & Perrimon, 2004; Deddouche *et al.*, 2008; Ding & Voinnet, 2007; Dostert *et al.*, 2005; Galiana-Arnoux *et al.*, 2006; Hedges *et al.*, 2008; Hedges & Johnson, 2008; Roxstrom-Lindquist *et al.*, 2004; Sabatier *et al.*, 2003; Teixeira *et al.*, 2008; van Rij *et al.*, 2006) and the birnavirus, *Drosophila X virus* (Zambon et al., 2005). In future studies of the range and specificity of Drosophila responses to virus infection, it would be valuable to use viruses that naturally infect Drosophila with a range of pathogenicities.

During the 1970s eleven viruses were described from Drosophila (Brun & Plus, 1980; Gateff *et al.*, 1980; Jousset & Plus, 1975; Plus, 1978). Three of these, which were isolated from *D.*

*Corresponding Author: Dr Karyn N. Johnson School of Biological Sciences University of Queensland St Lucia, 4072 Australia Phone: +61 7 33651358 Fax: +61 7 33651655 Email: karynj@uq.edu.au.

*melanogaster*, were described as picorna-like viruses based on morphological and biophysical characteristics. The three viruses; Drosophila C virus (DCV), Drosophila A virus (DAV) and Drosophila P virus (DPV) were shown to be different viruses based on serology and sizes of the capsid proteins (Jousset *et al.*, 1972; Plus *et al.*, 1976). These viruses are natural pathogens of *D. melanogaster* with about 40% of wild populations being infected with one or more of these viruses (Plus et al., 1975). The three viruses were shown to vary in their pathogenicity with DCV being the most pathogenic and DAV the least. It is important to have an understanding of molecular characteristics of viruses to facilitate these studies (Huszart & Imler, 2008). While DCV has been well characterised and belongs to the *Cripavirus* genus of the family *Dicistroviridae* (Christian *et al.*, 2005; Johnson & Christian, 1998), no detailed characterisation of the DAV genome or virion structure has been reported.

There are many small RNA viruses of insects that have non-enveloped 25-40 nm particles which encapsidate positive-sense RNA genomes (Christian & Scotti, 1998; Fauquet *et al.*, 2005). Of these the majority with sequenced genomes are included in four taxonomic groups; *Nodaviridae*, *Dicistroviridae*, *Tetraviridae* and the floating *Iflavirus* genus (Fauquet et al., 2005). However, in recent years a number of viruses have been characterised from insects that do not fit clearly into one of these groups (Habayeb *et al.*, 2006; Hartley *et al.*, 2005; van der Wilk *et al.*, 1997). Furthermore, within the *Tetraviridae* family there are viruses with diverse genome organizations, which are classified together because of the T=4 symmetry of their capsid particles (Agrawal & Johnson, 1992; Gorbalenya *et al.*, 2002; Gordon *et al.*, 1999; Gordon *et al.*, 1995; Hanzlik *et al.*, 1995; Pringle *et al.*, 2003; Pringle *et al.*, 2001). As more invertebrate RNA viruses are fully characterised it is becoming clear that there is a previously unrecognised diversity of genome organization and expression strategies within this group of viruses. These findings give us insight both into virus evolution and also into other biological processes (Koonin et al., 2008).

Here we describe the molecular and structural characterisation of Drosophila A virus. The 30 nm particles encapsidate a positive sense RNA, which encodes two large open reading frames. The first of these has significant sequence similarities to the RNA-dependent RNA polymerase (RdRP) of the tetraviruses and birnaviruses that have permuted RdRPs, whereas the second ORF encodes the capsid protein. Analysis of the structure of DAV virions shows a unique capsid arrangement.

## METHODS

### Propagation of DAV in Drosophila melanogaster

A stock of Champetiéres flies was obtained from the Drosophila Genetic Resource Centre in Kyoto Institute of Technology (DGRC stock number 103403) and a virus-free population was established essentially as described previously (Brun & Plus, 1980). The Australian Drosophila A virus isolate $DAV_{HD}$ was kindly provided by Dr Peter Christian (Christian, 1992). $DAV_{HD}$ was diluted in sterile insect saline (SIS: 0.6% NaCl, 0.04% KCl, 0.024% $Ca_2Cl$ and 0.02% $NaHCO_3$) and injected into the hemocoel of virus-free *D. melanogaster*. Injected flies were maintained on standard Drosophila media for 10 days at 25°C and then frozen at -20°C until purification.

### Virus Purification

$DAV_{HD}$ infected flies were homogenised in 50 mM Tris buffer pH 7.4 and homogenates centrifuged at 5,000 rpm for 5 min to pellet fly debris. The virus was purified from the supernatant by pelleting through a 3 ml 10% sucrose cushion at 27,000 rpm at 12°C for 3 hours in a SW41 swing bucket rotor (Beckman). The resuspended virus was layered onto a continuous 10-40% w/v sucrose gradient and centrifuged at 27,000 rpm at 12°C for 3 hours. The virus-

containing fractions were harvested, diluted in 50 mM Tris pH 7.4 and virus was pelleted by centrifugation at 27,000 rpm, 12°C for 3 hours. The virus was resuspended in 50 mM Tris pH 7.4 at 4°C overnight, aliquoted and stored at -20°C till use. $DAV_{HD}$ virions were also isolated from infected DL2 cells using a similar method after cells were lysed by repetitive freeze-thawing at -20°C.

### Transmission electron microscopy of purified virus

Purified $DAV_{HD}$ was analysed by transmission electron microscopy as previously described (Johnson & Ball, 2003).

### Nuclease treatment

For nuclease treatment 1.2 μg of $DAV_{HD}$ RNA, dsRNA prepared in vitro as previously described (Hedges & Johnson, 2008) and Flock House virus (FHV) RNA (Johnson et al., 2001) was incubated at 37°C with 10U RNase ONE (Promega) for 45 mins. Samples were then diluted in RNA loading buffer (Ambion), heated to 65°C, separated on a non-denaturing 1% agarose gel and products visualised by staining with ethidium bromide.

### cDNA synthesis, cloning and sequence determination

RNA was extracted from gradient purified $DAV_{HD}$ virions using TriReagent (Sigma-Aldrich) as per manufacturer's instructions, with the addition of 20 μg of glycogen increase RNA recovery. First and second strand cDNA was synthesised using the Superscript double-stranded cDNA synthesis kit (Invitrogen) with random hexamers as recommended by manufacturer. The cDNA products were ligated into the *EcoRV* site of phosphatase-treated pBluescript SK+ (Stratagene) and sequenced. From these sequences, DAV-specific primers were designed and used for further cDNA synthesis using Superscript reverse transcriptase III (Invitrogen) as per manufacturers instructions. 5'RACE was performed as indicated by the manufacturer (Invitrogen) to obtain sequence at the 5'end of the viral genome. A poly(A) tail was added to the 3'end of purified $DAV_{HD}$ RNA using poly(A) polymerase (Ambion) and a 3'RACE system (Invitrogen) used as per the manufacturers instructions to clone the 3'end of the DAV RNA. Nucleotide sequences were assembled and analysed using DNAStar software.

### Cells

*Drosophila* DL2 cells (Schneider, 1972) were maintained at 26°C in Schneider's medium (Invitrogen) supplemented with 10% heat-inactivated foetal bovine serum and antibiotics. DL2 cells were seeded at $7 \times 10^6$ cells per well into 6-well tissue culture plates and left to adhere at 26°C overnight. Cells were infected with purified $DAV_{HD}$ virions in 500 μl additive free Schneider's medium or mock-infected and supplemented with 1.5 ml Schneider's complete media following incubation at 26°C for 2 hours. Infected cells were harvested at 0, 1, 2, 3 and 5 dpi, washed in phosphate buffered saline, and lysed in either 1 ml TriReagent or 50 μL SDS-PAGE loading buffer.

### SDS-PAGE and Western blot analysis

Virus and cell lysate samples were heated at 99°C in 1 × sample buffer for 5 minutes before electrophoresis through SDS-10% poly acrylamide gels at 170 V. Alternatively samples were loaded with sample buffer and a reducing agent on a NuPAGE 4-12% Bis-Tris gradient gel in MES running buffer (Invitrogen) and electrophoresed at 200V. Proteins were visualised by staining with Coomassie brilliant blue.

For Western analysis, separated proteins were transferred from SDS-PAGE gels to a nitrocellulose membrane and following blocking, were incubated with a 1/5000 dilution of an anti-DAV antiserum overnight. Polyclonal antisera raised against both the original French

DAV isolate (Jousset et al., 1972) and the Australian isolate $DAV_{HD}$ (Christian, 1992) were kindly provided by Dr Peter Christian. The membrane was washed then incubated with 1/5000 dilution of a monoclonal anti-rabbit IgG conjugated with alkaline phosphatase for 2 hours. Following washes, bound antibody was visualised with 0.5% nitro blue tetrazolium and 0.25% BCIP in 1 M tris buffer ph 9.0. SeeBlue II+ prestained markers (Invitrogen) were used to allow alignment of proteins between Western and SDS-PAGE analysis.

## Mass spectrometry

Peptide mass finger printing and sequence analysis were conducted using matrix assisted laser desorption ionization time of flight (MALDI-TOF) and electrospray ion trap (ESI-MS/MS) mass analysis respectively. Samples were digested with modified trypsin (Promega Inc.) directly from solution or as gel pieces after SDS-PAGE. For MALDI, sinapinic acid was dissolved in 50% acetonitrile, 0.1% TFA at half saturation concentration (saturated sinapinic acid solution was diluted with equal volume of the solvent) as matrix. One microlitre of the $DAV_{HD}$ sample was mixed with 5 μl of the matrix and 1 μl of the mixed solution was deposited onto a clean plate as one spot and the sample analysed in a 4700 ABI MALDI TOFTOF and a Bruker biflex III mass spectrometer. Linear mode was used above 4,000 m/z and reflectron mode below. LC/MS/MS used an Agilent XCT-Ultra 6330 ion trap mass spectrometer connected to an Agilent 1100 CapLC and ChipCube (Agilent, Inc.). Peptides were eluted and loaded onto the analytical capillary column (43 mm × 75 μm ID, also packed with 5 μm Zorbax 300SB-C18 particles) connected in-line to the mass spectrometer with a flow of 600 nl/min. Peptides were eluted from a Zorbax 300SB-C18 with a 5 to 90% acetonitrile gradient (0.1% formic acid) over 10 min. Data-dependent acquisition of collision induced dissociation tandem mass spectrometry (MS/MS) was utilized. Protein identification from peptide MS and MS/MS data was based on MASCOT (Matrix Science, London, UK) searches of both an in-house database database containing all open reading frames from DAV and NCBI NR database. Intact protein analysis was preformed on a Bruker micrOTOF ESI-TOF after a similar C18 reverse-phase separation. Bruker MaxEnt deconvolution software was used to determine the mass of the capsid protein.

## CryoEM analysis of particles

DAV viruses were prepared for cryoEM analysis by preservation in vitreous ice over a holey carbon films from Protochips Inc. under the name Cflats. Grids were cleaned prior to freezing using a plasma cleaner (Gatan, Inc.) using a mixture of 75% argon and 25% oxygen for 10 seconds. A 3 μl aliquot of sample was applied to the grids, which were loaded into an FEI Vitrobot (FEI company) with settings at 4°C and 100% humidity. Grids were blotted for 4 seconds, and stored at liquid nitrogen temperatures. An Oxford 3500CT side-entry cryo-stage was used for data collection.

Data were acquired using a Tecnai F20 Twin transmission electron microscope operating at 120 keV, using a dose of ~19 e-/$Å^2$ and a nominal underfocus ranging from 0.5 to 3.0 μm. 2128 images were automatically collected at a nominal magnification of 80,000X at a pixel size of 0.101 nm at the specimen level. All images were recorded with a Tietz F415 4K × 4K pixel CCD camera (15 μm pixel) utilizing the Leginon data collection software (Suloway *et al.*, 2005). Experimental data were processed by the Appion software package (Lander *et al.*, 2009), which interfaces with the Leginon database infrastructure. The contrast transfer function (CTF) for each micrograph was estimated using the Automated CTF Estimation (ACE) package (Mallick *et al.*, 2005). 29,940 particles were automatically selected from the micrographs using a template-based particle picker (Roseman, 2003) and extracted at a box size of 512 pixels. Only particles whose CTF estimation had an ACE confidence of 80% or better were extracted. Phase correction of the single particles was carried out by during creation of the particle stack. Stacked particles were binned by a factor of 2 for the final reconstruction.

The final stack contained 22,103 particles, with 17,063 contributing to the final density. The 3D reconstruction was carried out using the EMAN reconstruction package (Ludtke *et al.*, 1999). Resolution was assessed by calculating the Fourier Shell Correlation (FSC) at a cutoff of 0.5, which provided a value of 8.32 Å resolution. Calculation of the resolution by rmeasure (Sousa & Grigorieff, 2007) at a 0.5 cutoff yielded a resolution of 8.69Å.

### Nucleotide sequence accession numbers

The Genbank accession number of the sequence reported in this paper is FJ150422.

## RESULTS AND DISCUSSION

### DAV is a non-enveloped icosahedral virus

To visualise the morphology of DAV particles, gradient purified virions were negatively stained and examined using TEM. At 100,000 times magnification, icosahedral non-enveloped particles were observed, with an average diameter of approximately 30 nm (Figure 1A). Small projections were visible on the virion surface in a regular array. The overall physical structure observed is consistent with that of other insect RNA viruses (Johnson & Reddy, 1998; Tate *et al.*, 1999). Higher resolution analysis is discussed in greater detail below.

### Structural protein analysis

To visualise DAV protein components, gradient purified virions were analysed using SDS-PAGE (Figure 1B). A single major protein band was observed and the mass of this protein was estimated to be 42 kDa by comparison with protein molecular mass standards. This suggested that $DAV_{HD}$ has a single major capsid protein of approximately 42 kDa, which is similar to the estimated size of the major structural protein previously described for DAV (Plus et al., 1976). There was no evidence of the two minor proteins of 32 and 73 kDa that were observed when the original French DAV isolate was previously analysed by Plus et al (1976). Western blot analysis was performed to analyse the protein components in $DAV_{HD}$ using an aliquot of the original anti-DAV antiserum which was raised against the French DAV isolate (Jousset et al., 1972). Three distinct bands were detected from this analysis (Figure 1C). The intense band indicated by the asterisk (Figure 1C) migrates at a similar position to the 42 kDa protein observed in the SDS-PAGE, showing that the 42 kDa protein is recognised by the original DAV antiserum. Two additional bands were detected at lower intensity and are observed consistently in Western blot analyses of $DAV_{HD}$ virions. The molecular masses of the minor proteins were estimated from to be 50 and 36 kDa for the larger and smaller bands respectively. While these proteins were not observed in the SDS-PAGE gel shown in Figure 1B, they were seen by SDS-PAGE analysis when virus was overloaded (data not shown). It is possible that these additional proteins are analogous to the two minor proteins previously described by Plus et al (1976), although the difference in the size of the larger protein is difficult to reconcile. Unfortunately, no virus was recovered following passage of two different samples of the original French DAV isolate (which had been stored at -20 °C for several decades) either in flies or cell culture, so it was not possible to make direct comparisons between the isolates.

### DAV replicates in cell culture

To investigate whether DAV replication and assembly occurs in cell culture, *Drosophila* DL2 cells were infected with $DAV_{HD}$ and accumulation of capsid protein in cells analysed by Western blot (Figure 2A). Proteins that co-migrated with each of the three proteins identified in purified virions were also detected in infected cells. No proteins were detected in either the mock infected cells or in the sample taken at time zero, the latter confirming that the input virus was below the level of detection in this assay. All three DAV proteins were detected 1 dpi (Figure 2A) and the quantity of the proteins increased over time, indicating that viral protein

synthesis was occurring in cell culture. However, the ratio of the 50 kDa and 42 kDa proteins was reversed as compared to that observed in the purified virion sample (Figure 1B). To compare the ratio of proteins in the virus derived from cell culture compared to virus derived from flies, virions were purified from both DL2 cells and Drosophila using the same method, and the resulting samples analysed with SDS-PAGE (Figure 2B). The virus sample purified from DL2 cells contained two major proteins of similar abundance and the sizes corresponded to the 50 kDa and 42 kDa proteins typically seen in Western blot analysis (Figure 2B, lane 2). In contrast, virus purified from flies as previously noted contained only a single major 42 kDa protein (Figure 2B, lane 1). The virus gradient purified from cultured DL2 cells was shown to be infectious in naïve DL2 cells (data not shown).

The relationship between the capsid proteins was not clear from this data. However, the disappearance of the 50 kDa protein and correlated increase in 42 kDa protein is reminiscent of the cleavage of a capsid protein precursor. Furthermore a small protein estimated to be 6 kDa was observed on SDS-PAGE gels overloaded for the major structural protein (Figure 3). Many viruses encode their capsid proteins in a precursor polyprotein. Within the insect ssRNA viruses, both nodaviruses and tetraviruses produce a capsid precursor protein which is maturational cleaved toward the C-terminus to produce the two mature capsid proteins (Johnson & Reddy, 1998). One interpretation of the DAV capsid protein profile is that the 50 kDa protein is a capsid precursor protein, which is cleaved during maturation of the virion to produce the major 42 kDa capsid protein. If this is the case for DAV then the higher abundance of the 50 kDa protein in virus from DL2 cells may be due to incomplete cleavage of the precursor protein during maturation. However, in contrast to nodaviruses and tetraviruses no asparagine residue was identified in the DAV sequence as the potential cleavage site. Further analysis will be required to determine the relationship between these proteins. The 36 kDa protein may initiate from the second AUG of the capsid ORF, as is observed in viruses from the *Nodaviridae* family (Johnson & Ball, 2003).

## Genome analysis

To analyse the nucleic acid encapsidated in virions, RNA extracted from gradient purified DAV was resolved on a 1% agarose gel and nucleic acids visualised by staining with ethidium bromide (Figure 4). This showed a smear of RNA, which spanned sizes of approximately 1000-5000 nucleotides when compared to single-stranded RNA markers and this size range was confirmed by analysis on a denaturing formaldehyde agarose gel (data not shown). This suggested that the RNA might be degrading during the extraction and electrophoresis process, although the distinct edges of the smear and intact markers indicate there was no RNase contamination in the electrophoresis process. The RNA smear was repeatedly observed, even when Flock House virus (FHV) RNA extracted concurrently to the DAV sample was shown to be intact (Fig 4 lane 6). An alternative RNA extraction kit (Invitrogen) also did not improve the integrity of the RNA (data not shown). Degradation of RNA purified from tetravirus particles is common (Gordon *et al*., 1999;Gordon & Hanzlik, 1998;Pringle *et al*., 2003) and it is possible that the removal of the genomic RNA from the virion may result in RNA degradation. We cannot discount other possible explanations for the apparent diversity in size of the encapsidated RNA. However, despite the smearing, an approximation of the maximum RNA size was inferred from the distinct upper size cut off of 5 kb.

Nucleic acid was extracted from gradient purified DAV and treated with ribonuclease. RNase ONE treatment resulted in degradation of the DAV genome (Fig 4 lane 3). Under the same conditions the genome of single-stranded RNA virus FHV was also degraded whereas a synthetic dsRNA transcribed in vitro remained intact (Fig 4). This suggests that the genome of DAV is most likely single-stranded RNA.

Using a combination of approaches clones corresponding to genomic RNA were synthesized and 4806 nucleotides of sequence of the genome determined. The sequence of all of the recovered clones formed a contiguous sequence suggesting that DAV has a single genomic RNA. The 5' and 3' ends of the genomic RNA were targeted using RACE. Two independent clones were identified with the same 5'-end, however it is possible that the genome includes additional nucleotides at the 5' terminus. Interestingly, 3' RACE was only successful when a synthetic poly(A) tail was added to the $DAV_{HD}$ RNA, indicating that the DAV genome is not polyadenylated.

The genomic sequence was analysed for open reading frames. The plus sense-strand encoded two major open reading frames and there were an additional three small open reading frames encoding potential proteins containing more than 50 amino acids (Figure 5). The 5' proximal reading frame was the largest encoding a putative protein of 1075 amino acids with a predicted molecular mass of 121 kDa. The 3'proximal ORF encoded 442 codons, which is predicted to encode a protein of 48.4 kDa. The second ORF initiates from a methionine that is 43 nt downstream of the first ORF stop codon and is in the +1 frame compared to the first. The 3' reading frame overlaps the 5' coding region upstream of the initiating methionine (Figure 5A). It is therefore possible that a +1 frameshift would allow synthesis of a polyprotein including both large ORF. However, +1 frameshifts are uncommon and it seems unlikely that the capsid proteins would be translated in this inefficient manner.

To identify proteins with sequence similarity to the deduced sequence of the 5' proximal reading frame, the translated sequence was used to search the non-redundant sequence databases using BLAST. Significant matches were found to a large region of the RdRP proteins from two members of the *Tetraviridae* family and several members of the *Birnaviridae* family. The highest homology was found with the RdRP of *Thosea asigna virus* (TaV) and *Euprosterna elaeasa virus* (EeV), which are closely related viruses from the *Tetraviridae* family. TaV and EeV both have an unusual permutation in the palm-domain of the RdRP (Gorbalenya *et al.*, 2002). The nearly ubiquitous arrangement that is observed in RNA viruses is that the GDD box (motif C) of the palm subdomain follows the conserved A and B motifs (A-B-C) (Poch *et al.*, 1989). However, in both TAV and EeV RdRPs the C motif precedes the others in a C-A-B arrangement. This permuted RdRP is also observed in some double-stranded RNA viruses from the *Birnavirdae* family and structural analysis of the permuted RdRP has shown a similar arrangement to the canonical RdRP of other viruses (Garriga *et al.*, 2007; Gorbalenya *et al.*, 2002). Interestingly the RdRP protein of DAV also contains this C-A-B permutation (see supplementary Figure 1).

To identify the structural protein coding region the sequence of the virion proteins was analysed. N-terminal sequencing of the structural protein was unsuccessful both from gel purified protein and whole virus, suggesting that the protein may be blocked (data not shown). Protease mass mapping of virions in solution and of the 42 kDA capsid protein band after SDS-PAGE were tried next. Using a combination of MALDI-TOF and LCMS, 7 and 6 peptides respectively, spanning 29% of the 3' ORF were identified, indicating that this ORF encodes the capsid protein (Figure 5C). The predicted mass of 48.4 kDa is also consistent with the precursor protein band visualized on the protein gels (50 kDa). LCMS of purified particles produced a signal for the mature capsid protein of 40,480 +/-3. This does not match with a single cleavage event leading to capsid protein maturation as in insect nodaviruses and tetraviruses, both of which occur autocatalytically at an asparagine residue (Johnson & Reddy, 1998). The fact that the mature protein mass does not match the amino acid sequence is however consistent with the N-terminus being modified as discussed above and this will be investigated in future experiments.

### Sequence-based fold recognition of the capsid protein

A BLAST search of capsid protein's amino acid sequence against known insect virus genomes revealed a single homologue in the *Nodaviridae* family, exhibiting a 27% identity to the Wuhan nodavirus between residues 330 and 442. For further sequence-based analyses of the capsid structure, various fold recognition methods were utilized via the structure prediction meta server (Ginalski et al., 2003), revealing a folding motif consistent with several members of the *Tombusviridae* (carnation mottle virus, tobacco necrosis virus, tomato bushy stunt virus) and *Sobemoviridae* (cocksfoot mottle virus, rice yellow mottle virus, southern bean mosaic virus) families (Supplementary Figure 2). The results point to a conserved canonical beta-sandwich fold beginning around residue 74 and ending at around residue 236. Two viruses from the family *Tombusviridae* - carnation mottle virus (CMV) and tomato bushy stunt virus (TBSV), whose structures contain a C-terminal domain that protrudes from the capsid surface, additionally exhibited a fold similarity with the DAV sequence between residues 270 to 315. Conceivably, the N-terminal region of DAV forms an evolutionarily conserved beta-sandwich domain, while the C-terminal domain forms a surface extension, much like the structures of CMV and TBSV. Indeed, cryoEM analysis of the malabaricus grouper nervous necrosis virus (MGNNV), a fish nodavirus that points to a similar domain organization upon sequence-based fold recognition (Tang et al., 2002), illustrated a C-terminal domain that was separated from the beta-sandwich domain that formed the core capsid of the virus particle. The N-terminal portion of the sequence is highly charged including many arginines which indicates that it may interact with the RNA and have a role in particle assembly (Tang *et al.*, 2001; Venter *et al.*, 2009). A closer look at the structure of DAV particles via cryoEM methods was necessary to elucidate the overall structural morphology.

### CryoEM analysis of DAV virus particles

The structure of DAV was examined in detail by CryoEM image reconstruction techniques. Interestingly, unprocessed, high-contrast cryoEM images of the virus particles (Figure 6A) exhibited a structural morphology that was very reminiscent of MGNNV TEM images, whose radial density distribution shows two distinct protein shells (Tang et al., 2002). Averaging of the computationally extracted and centered DAV particles from the micrographs confirms that the virus indeed appears to have a divided capsid shell (Figure 6B). From the selected particles a three-dimensional structure was computed to 8Å resolution as estimated by Fourier shell correlation method using a value of 0.5 (Figure 7A and B).

At these resolutions, the distribution of the capsid protein into two discrete and separate domains is plainly evident. Viewed down the 5-fold axis, the density distribution of the reconstructed particle shows the core domain extending from a radius of 115 Å to 138 Å with a maximum at 127 Å, while the outer domain maintains a radius of 150 Å to 175 Å, with a maximum at 163 Å (figure 7 C and D). These dimensions are comparable to the dimensions of the MGNNV cryoEM structure, with two additional layers of density corresponding to RNA, the outermost of which appears to directly interact with the capsid shell. Examination of the inner domain of the DAV electron density reveals an unmistakable T=3 quasi-equivalent lattice, with much of the secondary structural elements visible. Comparison of the core domain with the beta-sandwich domain of TBSV shows a very similar structural organization, and was docked into place using an automated real-space rigid body refinement (Figure 7E). The docking shows that the network of subunit-subunit interactions that make up the capsid are highly conserved, and that capsid assembly in DAV is likely to be similar to that of TBSV and other tombusviruses. Although tombusviruses contain two distinct domains, it is noteworthy that in the tombusviruses they are not separated by the 10 Å gap present in the DAV structure. TBSV has a short polypeptide hinge connecting the N and C-terminal domains, and it is probable that a similar hinge is seen in DAV, though much longer.

An unanticipated finding is a lack of highly ordered structure corresponding to the outer domain. While the inner domain shows a detailed T=3 organization, the outer domain exhibits no such lattice, only ambiguous densities that appear to be an artifact of applying icosahedral symmetry to a disordered or dynamic protein domain. Despite a lack of quasi-symmetry, however, the outer domain appears to be somewhat ordered, forming a cage-like icosahedral lattice that surrounds the core domain. The densities attributed to this dynamic C-terminal domain appear to converge above the icosahedral and quasi two-fold axes. In this regard DAV diverges from the structure of MGNNV, whose outer domain consists of large protrusions of density at the quasi-three folds.

Attempts to focus on this outer domain by masking out the interior region during the refinement did not provide any further insight into its organization, as the resulting densities were equally ambiguous, with little deviation from the map containing the N-terminal portion. Due to the disorganization and lack of quasi-equivalent symmetry present in the C-terminal density, fitting of the protruding domain (residues 272 to 387) of TBSV was not modeled into the map.

The beta-sandwich that makes up the inner core of many viruses sets up a solid structural template from which other domains protrude. These protrusions are generally involved in biologically relevant processes, such as host cell recognition, and it has been shown that in some viruses, such as rotavirus (Dormitzer et al., 2004), these domain extensions are dynamic and can undergo rearrangements during the virus life cycle. The cryoEM work shown here provides evidence that DAV may be one such virus, whose dynamic outer domains, with possible biological function, are anchored to a well-ordered beta-sandwich icosahedral core.

### Conclusions

Drosophila A virus was first described as a picorna-like virus. The data presented here show that DAV is evolutionarily related to the viruses with permuted RdRP core motifs from both the *Tetraviridae* and *Birnaviridae* families. However, the capsid protein shows strongest sequence similarity to viruses from the *Nodaviridae* family and conserved protein folding motifs with the members of the *Tombusviridae*. In addition the structure of DAV particles shows an inner T=3 core surrounded by a less defined outer protein domain. Taken together these features show that DAV an unusual RNA virus, with evolutionary relationships across a diverse range of invertebrate, fish and plant viruses.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

We thank - Peter Christian for providing virus and antisera, and discussions; Darren Obbard and Peter Christian for sharing unpublished data; Lopson Kebapetswe and Ivy Mui Tan for fly injections. This research was supported by a University of Queensland development grant to KNJ. Electron microscopic imaging and reconstruction was conducted at the National Resource for Automated Molecular Microscopy, which is supported by the NIH through the National Center for Research Resources' P41 program (RR17573). BB acknowledges support from the National Science Foundation, MCB 0646499 and NIH Grant Number P20 RR-16455-06 from the COBRE Program of the National Center For Research Resources. BB also receives support from the Center for Bio-Inspired Nanomaterials (Office of Naval Research grant #N00014-06-01-1016). The Murdock Charitable Trust for support of the Mass Spectrometry Facility at MSU.

## REFERENCES

Agrawal DK, Johnson JE. Sequence and analysis of the capsid protein of *Nudaurelia capensis* ω virus, an insect virus with T=4 icosahedral symmetry. Virology 1992;190:806–814. [PubMed: 1519360]

Brun, G.; Plus, N. The viruses of *Drosophila*. In: Ashburner, M.; Wright, TFR., editors. The Genetics and Biology of Drosophila. Academic Press; New York: 1980. p. 625-702.

Cherry S, Doukas T, Armknecht S, Whelan S, Wang H, Sarnow P, Perrimon N. Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. Genes Dev 2005;19:445–452. [PubMed: 15713840]

Cherry S, Perrimon N. Entry is a rate-limiting step for viral infection in a *Drosophila melanogaster* model of pathogenesis. Nat Immunol 2004;5:81–87. [PubMed: 14691479]

Christian PD. A simple vacuum dot-blot hybridisation assay for the detection of *Drosophila* A and C viruses in single *Drosophila*. Journal of Virological Methods 1992;38:153–165. [PubMed: 1644892]

Christian, PD.; Carstens, EB.; Domier, L.; Johnson, JE.; Johnson, KN.; Nakashima, N.; Scotti, PD.; van der Wilk, F. Dicistroviridae. In: Fauquet, CM.; Mayo, MA.; Maniloff, J.; Desselberger, U.; Ball, LA., editors. Virus taxonomy: Eighth report of the International Committee on the Taxonomy of Viruses. Elsevier Academic Press; San Diego: 2005. p. 783-788.

Christian, PD.; Scotti, PD. The picorna-like viruses of insects. In: Miller, LK.; Ball, LA., editors. The Viruses; Insect Viruses II. Plenum Publishing Corporation; New York: 1998. p. 301-336.

Deddouche S, Matt N, Budd A, Mueller S, Kemp C, Galiana-Arnoux D, Dostert C, Antoniewski C, Hoffmann JA, Imler JL. The DExD/H-box helicase Dicer-2 mediates the induction of antiviral activity in Drosophila. Nat Immunol 2008;9:1425–1432. [PubMed: 18953338]

Ding SW, Voinnet O. Antiviral immunity directed by small RNAs. Cell 2007;130:413–426. [PubMed: 17693253]

Dormitzer PR, Nason EB, Prasad BV, Harrison SC. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. Nature 2004;430:1053–1058. [PubMed: 15329727]

Dostert C, Jouanguy E, Irving P, Troxler L, Galiana-Arnoux D, Hetru C, Hoffmann JA, Imler JL. The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. Nat Immunol 2005;6:946–953. [PubMed: 16086017]

Fauquet, CM.; Mayo, MA.; Maniloff, J.; Desselberger, U.; Ball, LA., editors. Virus Taxonomy: Eighth report of the international committee on taxonomy of viruses. Academic Press; Amsterdam: 2005.

Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler JL. Essential function in vivo for Dicer-2 in host defense against RNA viruses in Drosophila. Nat Immunol 2006;7:590–597. [PubMed: 16554838]

Garriga D, Navarro A, Querol-Audi J, Abaitua F, Rodriguez JF, Verdaguer N. Activation mechanism of a noncanonical RNA-dependent RNA polymerase. Proc Natl Acad Sci U S A 2007;104:20540–20545. [PubMed: 18077388]

Gateff, E.; Gissmann, L.; Shrestha, R.; Plus, N.; Pfister, H.; Schröder, J.; Zur Hausen, H. Characterization of two tumorous blood cell lines of *Drosophila melanogaster* and the viruses they contain. In: Kurstak, E.; Maramorosch, K.; Dübendorfer, F., editors. Invertebrate Systems in vitro. Elsevier/North-Holland Biomedical Press; Amsterdam: 1980. p. 517-533.

Ginalski K, Elofsson A, Fischer D, Rychlewski L. 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics 2003;19:1015–1018. [PubMed: 12761065]

Gorbalenya AE, Pringle FM, Zeddam JL, Luke BT, Cameron CE, Kalmakoff J, Hanzlik TN, Gordon KH, Ward VK. The palm subdomain-based active site is internally permuted in viral RNA-dependent RNA polymerases of an ancient lineage. J Mol Biol 2002;324:47–62. [PubMed: 12421558]

Gordon KH, Williams MR, Hendry DA, Hanzlik TN. Sequence of the genomic RNA of *Nudaurelia beta virus (Tetraviridae*) defines a novel virus genome organization. Virology 1999;258:42–53. [PubMed: 10329566]

Gordon, KHJ.; Hanzlik, TN. Tetraviruses. In: Miller, LK.; Ball, LA., editors. The Viruses; Insect Viruses II. Plenum Publishing Corporation; New York: 1998. p. 269-299.

Gordon KHJ, Johnson KN, Hanzlik TN. The larger genomic RNA of *Helicoverpa armigera* stunt tetravirus encodes the viral RNA polymerase and has a novel 3'-terminal tRNA-like structure. Virology 1995;208:84–98. [PubMed: 11831734]

Habayeb MS, Ekengren SK, Hultmark D. Nora virus, a persistent virus in Drosophila, defines a new picorna-like virus family. J Gen Virol 2006;87:3045–3051. [PubMed: 16963764]

Hanzlik TN, Dorrian SJ, Johnson KN, Brooks EM, Gordon KHJ. Sequence of RNA2 of the Helicoverpa armigera stunt virus (*Tetraviridae*) and bacterial expression of its genes. J Gen Virol 1995;76:799–811. [PubMed: 9049325]

Hartley CJ, Greenwood DR, Gilbert RJ, Masoumi A, Gordon KH, Hanzlik TN, Fry EE, Stuart DI, Scotti PD. Kelp fly virus: a novel group of insect picorna-like viruses as defined by genome sequence analysis and a distinctive virion structure. J Virol 2005;79:13385–13398. [PubMed: 16227260]

Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. *Wolbachia* and virus protection in insects. Science 2008;322:702. [PubMed: 18974344]

Hedges LM, Johnson KN. The induction of host defence responses by *Drosophila C virus*. J Gen Virol 2008;89:1497–1501. [PubMed: 18474566]

Huszart T, Imler JL. Drosophila viruses and the study of antiviral host-defense. Advances in Virus Research 2008;72:227–265. [PubMed: 19081493]

Johnson, JE.; Reddy, V. Structural studies of nodaviruses and tetraviruses. In: Miller, LK.; Ball, LA., editors. The Insect Viruses. Plenum Publishing Corporation; New York: 1998. p. 171-223.

Johnson KN, Ball LA. Virions of *Pariacoto virus* contain a minor protein translated from the second AUG codon of the capsid protein open reading frame. J Gen Virol 2003;84:2847–2852. [PubMed: 13679619]

Johnson KN, Christian PD. The novel genome organization of the insect picorna-like virus *Drosophila C virus* suggests this virus belongs to a previously undescribed virus family. J Gen Virol 1998;79:191–203. [PubMed: 9460942]

Johnson KN, Johnson KL, Dasgupta R, Gratsch T, Ball LA. Comparisons among the larger genome segments of six nodaviruses and their encoded RNA replicases. J Gen Virol 2001;82:1855–1866. [PubMed: 11457991]

Jousset FX, Plus N. Étude de la transmission horizontale et de la transmission verticale des picornavirus de *Drosophila melanogaster* et de *Drosophila immigrans*. Ann Microbiol (Inst Pasteur) 1975;126 B: 231–249.

Jousset FX, Plus N, Croizier G, Thomas M. Existence chez *Drosophila* de deux groupes de picornavirus de propriétés sérologiques et biologiques différentes. CR Acad Sci (Paris) 1972;275:3043–3046.

Koonin EV, Wolf YI, Nagasaki K, Dolja VV. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nature Reviews Microbiology 2008;6:925–939.

Lander GC, Stagg S, Voss NR, Cheng A, Fellmann D, Pulokas J, Yoshioka C, Irving C, Mulder A, Lau P, Lyumkis D, Potter CS, Carragher B. Appion: an integrated, database-driven pipeline to facilitate EM image processing. J Struct Biol. 2009

Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. J Struct Biol 1999;128:82–97. [PubMed: 10600563]

Mallick SP, Carragher B, Potter CS, Kriegman DJ. ACE: automated CTF estimation. Ultramicroscopy 2005;104:8–29. [PubMed: 15935913]

Plus N. Endogenous viruses of *Drosophila melanogaster* cell lines: their frequency, identification and origin. In Vitro 1978;14:1015–1021.

Plus N, Croizier G, Jousset FX, David J. Picornaviruses of laboratory and wild *Drosophila melanogaster*: geographical distribution and serotypic composition. Ann Microbiol (Inst Pasteur) 1975;126 A:107–117.

Plus N, Croizier G, Veyrunes J, David J. A comparison of buoyant density and polypeptides of *Drosophila* P, C and A viruses. Intervirology 1976;7:346–350. [PubMed: 1025039]

Poch O, Sauvaget I, Delarue M, Tordo N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. EMBO Journal 1989;8:3867–3874. [PubMed: 2555175]

Pringle FM, Johnson KN, Goodman CL, McIntosh AH, Ball LA. *Providence virus*: a new member of the *Tetraviridae* that infects cultured insect cells. Virology 2003;306:359–370. [PubMed: 12642108]

Pringle FM, Kalmakoff J, Ward VK. Analysis of the capsid processing strategy of Thosea asigna virus using baculovirus expression of virus-like particles. J Gen Virol 2001;82:259–266. [PubMed: 11125178]

Roseman AM. Particle finding in electron micrographs using a fast local correlation algorithm. Ultramicroscopy 2003;94:225–236. [PubMed: 12524193]

Roxstrom-Lindquist K, Terenius O, Faye I. Parasite-specific immune response in adult *Drosophila melanogaster*: a genomic study. EMBO Rep 2004;5:207–212. [PubMed: 14749722]

Sabatier L, Jouanguy E, Dostert C, Zachary D, Dimarcq JL, Bulet P, Imler JL. Pherokine-2 and -3. Eur J Biochem 2003;270:3398–3407. [PubMed: 12899697]

Schneider I. Cell lines derived from late embryonic stages of *Drosophila melanogaster*. J Embryol Exp Morph 1972;27:353–365. [PubMed: 4625067]

Sousa D, Grigorieff N. Ab initio resolution measurement for single particle structures. J Struct Biol 2007;157:201–210. [PubMed: 17029845]

Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J, Stagg S, Potter CS, Carragher B. Automated molecular microscopy: the new Leginon system. J Struct Biol 2005;151:41–60. [PubMed: 15890530]

Tang L, Johnson KN, Ball LA, Lin T, Yeager M, Johnson JE. The structure of Pariacoto virus reveals a dodecahedral cage of duplex RNA. Nature Struct Biol 2001;8:77–83. [PubMed: 11135676]

Tang L, Lin CS, Krishna NK, Yeager M, Schneemann A, Johnson JE. Virus-like particles of a fish nodavirus display a capsid subunit domain organization different from that of insect nodaviruses. J Virol 2002;76:6370–6375. [PubMed: 12021370]

Tate J, Liljas L, Scotti P, Christian P, Lin T, Johnson JE. The crystal structure of cricket paralysis virus: the first view of a new virus family. Nat Struct Biol 1999;6:765–774. [PubMed: 10426956]

Teixeira L, Ferreira A, Ashburner M. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. Plos Biology 2008;6:e1000002.doi:1000010.1001371/journal.pbio.1000002

van der Wilk F, Dullemans AM, Verbeek M, Van den Heuvel JF. Nucleotide sequence and genomic organization of Acyrthosiphon pisum virus. Virology 1997;238:353–362. [PubMed: 9400608]

van Rij RP, Saleh MC, Berry B, Foo C, Houk A, Antoniewski C, Andino R. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. Genes Dev 2006;20:2985–2995. [PubMed: 17079687]

Venter PA, Marshall D, Schneemann A. Dual roles for an arginine-rich motif in specific genome recognition and localization of viral coat protein to RNA replication sites in flock house virus-infected cells. J Virol 2009;83:2872–2882. [PubMed: 19158251]

Zambon RA, Nandakumar M, Vakharia VN, Wu LP. The Toll pathway is important for an antiviral response in *Drosophila*. Proc Natl Acad Sci U S A 2005;102:7257–7262. [PubMed: 15878994]

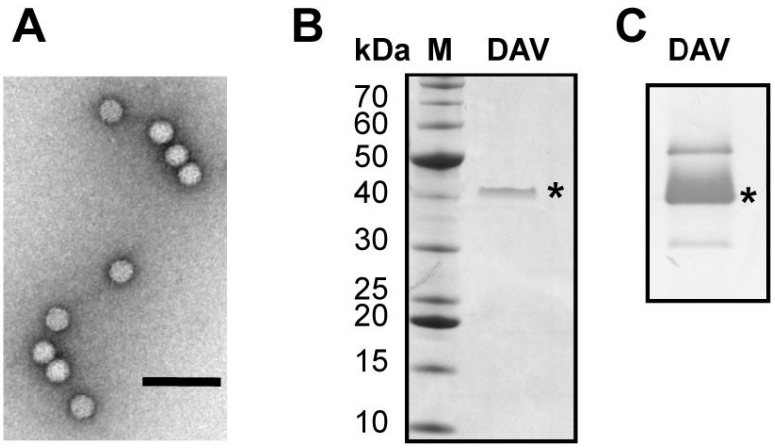


**Figure 1. Analysis of DAV virions and structural proteins**

**A.** Gradient purified virions were negatively stained and visualised using transmission electron microscopy at 100,000 times magnification. Black bar represents 100 nm. **B.** The proteins present in purified virus particles were separated by electrophoresis through a 4-12% Bis-Tris gradient gel. Total protein was visualised by staining with Coomassie Blue. The capsid protein is indicated by the asterisk. The sizes of the molecular mass markers are shown on the left. **C.** Virion proteins separated on a SDS-10% polyacrylamide gel were transferred to a nitrocellulose membrane and detected with a rabbit polyclonal anti-DAV antibody. The band marked with an asterisk corresponds to the 42 kDa protein indicated in part B.

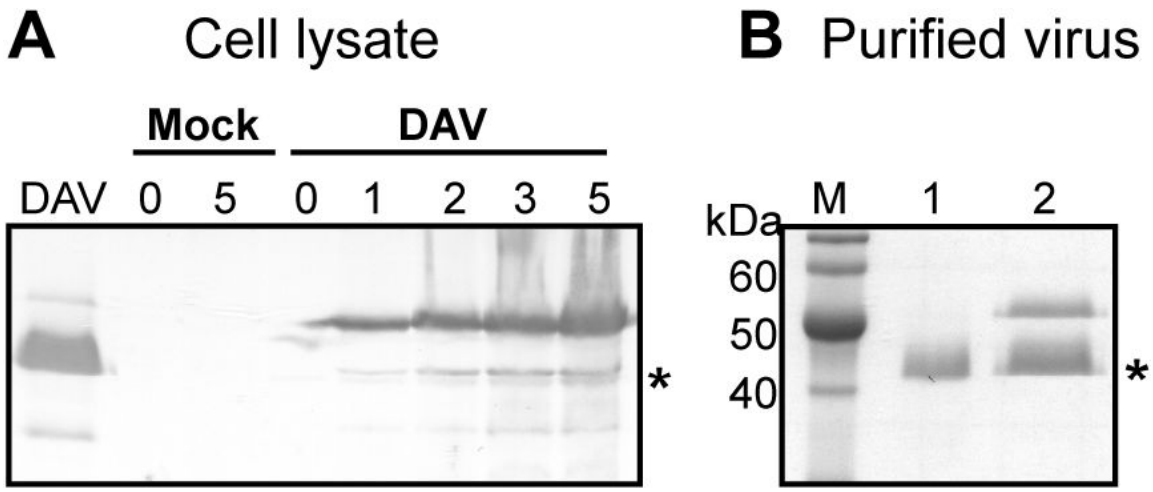


**Figure 2. Detection of viral protein synthesis in infected DL2 cells**

**A**. *Drosophila* DL2 cells were mock-infected or infected with DAV and incubated at 26°C until harvesting. Cell samples were taken at 0, 1, 2, 3 and 5 dpi, and lysed with sample buffer. Cellular proteins were separated by electrophoresis through a SDS-10% polyacrylamide gel, transferred to a nitrocellulose membrane and incubated with an anti-DAV rabbit polyclonal antiserum to detect viral proteins. The asterisk indicates the 42 kDa capsid protein. Viral proteins were detected at 1 dpi and increased over time. No DAV-specific bands were detected in the mock-infected samples. **B.** $DAV_{HD}$ virions purified from infected DL2 cells (Lane 2) and adult Drosophila (Lane 1) purified virion proteins were electrophoresed through an SDS-10% polyacrylamide gel stained with Coomassie blue. The asterisk indicates the 42 kDa capsid protein.

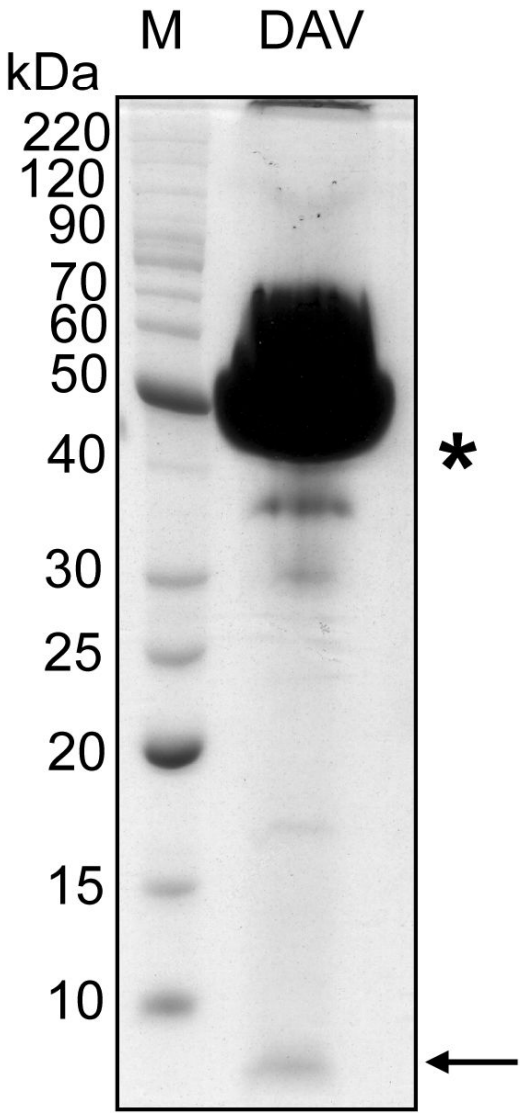


**Figure 3. Identification of a small structural protein**

Virion proteins were overloaded and separated by electrophoresis through an SDS-10% acrylamide gel. Total protein was stained with Coomassie Blue. The asterisk indicates the 42 kDa capsid protein band that is consistently observed in SDS-PAGE analyses. The arrow indicates the small protein band that is estimated to be approximately 6 kDa.

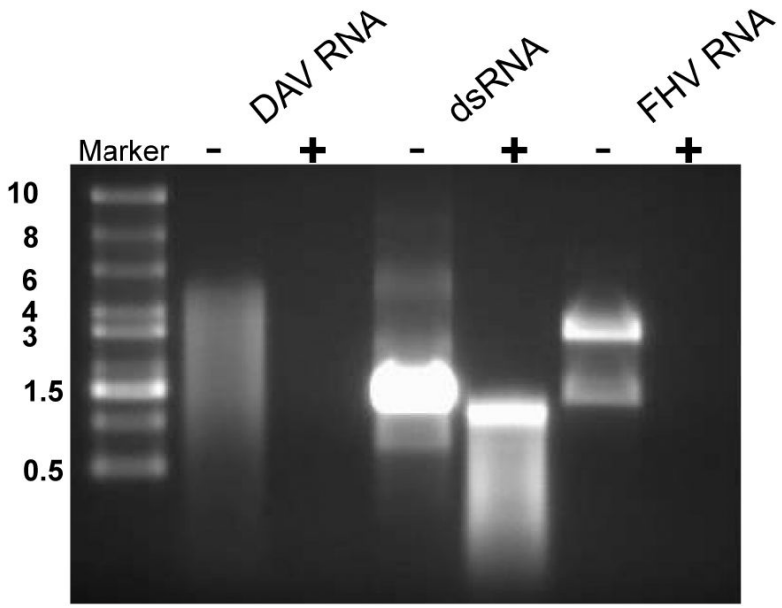


**Figure 4. Analysis of DAV genomic nucleic acid**

RNA was extracted from purified $DAV_{HD}$ virions and mock-treated (-) or treated with RNAseONE (+), which degrades single-stranded RNA only. As controls, a synthetic dsRNA product from DCV and the ssRNA genome from Flock House virus (FHV) were similarly treated. Following nuclease treatment, products were separated on a 1% agarose gel and visualised using ethidium bromide staining. The sizes of the ssRNA markers are indicated on the left.

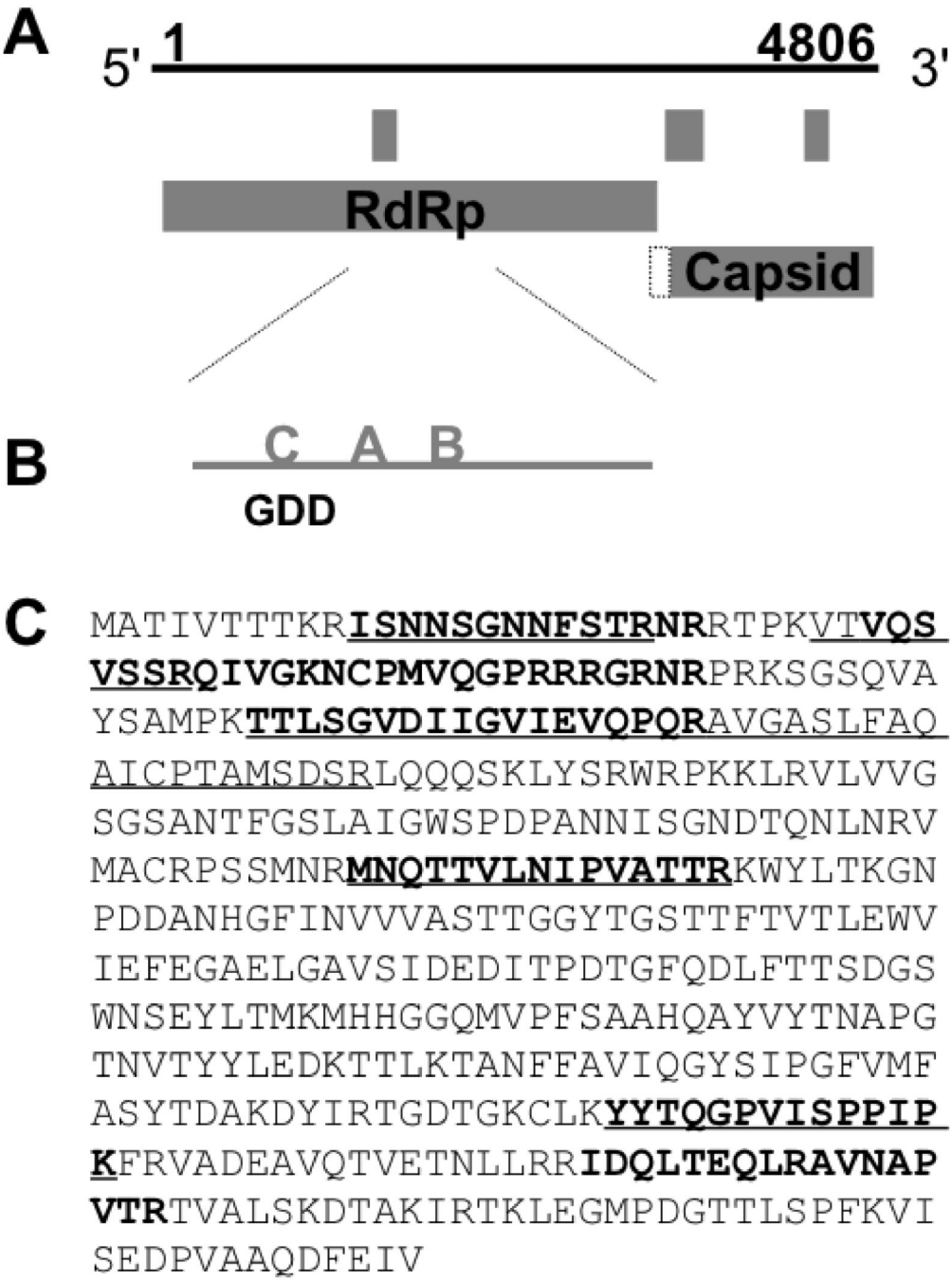


**Figure 5. Organisation of DAV genome**

**A.** The 4806 nt DAV genome encodes two major ORFs and three additional small ORFs that contain over 50 amino acids. The reading frames as shown start with an AUG codon and the extension on the capsid ORF indicates this reading frame overlaps with the RdRP reading frame but the first methionine is downstream of the RdRP reading frame **B.** Analysis of the 5' proximal ORF showed high similarity of DAV replicase regions with members of the *Tetraviridae* and *Birnaviridae* families. These viruses have a permuted RdRp, with the GDD region upstream of other palm sub-domains. DAV replicase also showed this permuted arrangement, as indicated by the C-A-B. **C.** Mass spectrometry analysis indicates that the 3' proximal large

ORF encodes the capsid protein. The deduced sequence of the capsid protein is shown and the peptides identified by MALDITOF (**bold**) and LC-MS-MS (underlined) indicated.

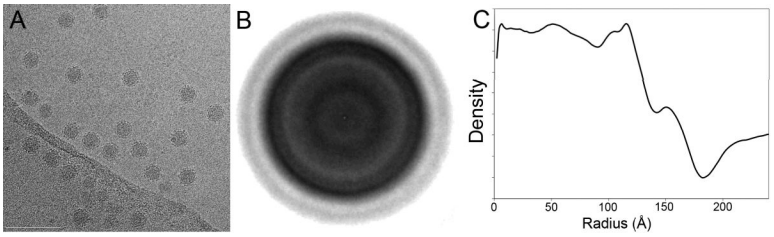


**Figure 6. Cryo-electromicroscopy analysis of DAV virions**

**A.** Representative electron micrograph of gradient-purified DAV in the frozen-hydrated state. **B.** Averaged 2-dimensional density corresponding to 22,000 centered raw particles **C.** 1D radial plot of the averaged 2D particles, showing discrete peaks for the capsid layers.

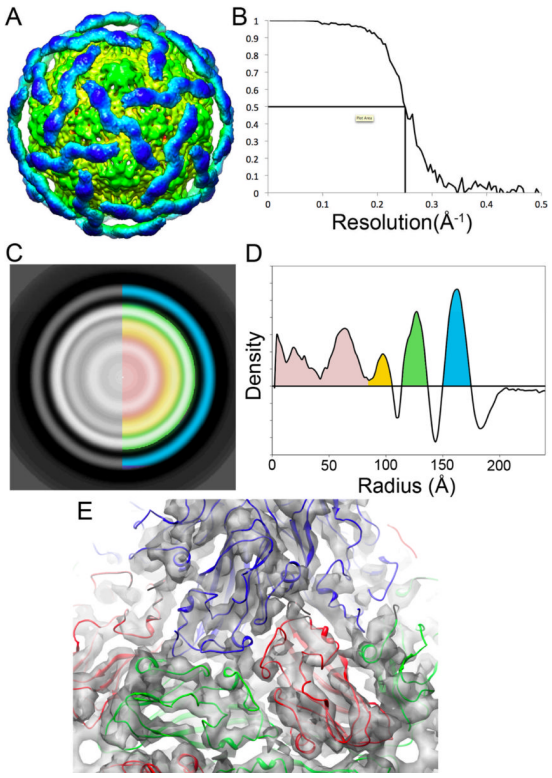


**Figure 7. Analysis of three-dimensional reconstruction of DAV particles**

**A.** Surface representation of the DaV CryoEM reconstruction contoured at 2 sigma. The inner N-terminal domain (green) is well ordered with visible secondary structural features, while the outer C-terminal domain (blue) contains only very low-resolution features. **B.** The estimated resolution of the reconstruction is 8Å according to FSC criteria at 0.5. **C.** A radially averaged slice through center of reconstruction. Blue = C-terminal domain, Green = N-terminal domain, Orange, Red = RNA **D.** Plot showing density by radius (colours similar to C.) **E.** Docking of the beta-sandwich domain of the TBSV crystal structure into the N-terminal domain of the DaV cryoEM density.