

Analysis of the three-dimensional structure of the African horse sickness virus VP7 trimer by homology modelling

Shani Bekker^a, Pieter Burger^{b, 1}, Vida van Staden^{a, *}

^a Department of Genetics, University of Pretoria, Pretoria 0002, South Africa

^b Bioinformatics and Computational Biology Unit, University of Pretoria, Pretoria 0002, South Africa

* Corresponding author contact details:

Dr Vida van Staden

Department of Genetics

University of Pretoria

Private Bag X20

Hatfield 0028

South Africa

Tel: +27 12 420 3257

Fax: +27 12 362 5327

Email: vida.vanstaden@up.ac.za

¹ Present address: Department of Chemistry, Emory University, Atlanta, GA 30322, USA

Highlights

- VP7 crystalline particle formation likely conserved across all nine AHSV serotypes
- Model of AHSV VP7 trimer useful for design of mutagenesis studies and epitopes
- Residues grouped together on trimer surface to create trimer-trimer interactions
- Identified useful strategy for the potential abolishment of protein self-assembly

Abstract

VP7 is the major core protein of orbiviruses and is essential for virion assembly. African horse sickness virus (AHSV) VP7 self-assembles into highly insoluble crystalline particles – an attribute that may be related to the role of AHSV VP7 in virus assembly but also prevents crystallization. Given that this inherent insolubility is unique to AHSV VP7, we use amino acid sequence conservation analysis between AHSV VP7 and other orbiviruses to identify putative key residues that drive AHSV VP7 self-assembly. A homology model of the AHSV VP7 trimer was generated to analyze surface properties of the trimer and to identify surface residues as candidates for the AHSV VP7 trimer-trimer interactions that drive AHSV VP7 self-assembly. Nine regions were identified as candidate residues for future site-directed mutagenesis experiments that will likely result in a soluble AHSV VP7 protein. Additionally, we identified putative residues that function in the intermolecular interactions within the AHSV VP7 trimer as well as several epitopes. Given the many previous efforts of solubilizing AHSV VP7, we propose a useful strategy that will yield a soluble AHSV VP7 that can be used to study AHSV assembly and increase yield of recombinant vaccine preparations.

Keywords: AHSV; VP7; trimer; homology modelling; self-assembly; epitopes

1. Introduction

African horse sickness virus (AHSV, genus *Orbivirus*, family *Reoviridae*) is the causative agent of the vector-borne disease of infected horses known as African horse sickness (AHS). With a mortality rate of up to 95% in susceptible horses (Stassen *et al.* 2014), AHS remains to be the most devastating equine disease in Southern Africa with a significant economic impact on breeding and racing industries (Sergeant *et al.* 2016). Along with bluetongue virus (BTV), AHSV is one of the most economically important members of the *Orbivirus* genus. AHS is endemic in sub-Saharan Africa, and is listed as a notifiable viral disease by the World Organization for Animal Health (Office International des Epizooties, 2015). The appearance of BTV in the Mediterranean Basin of Europe, along with the recent northward expansion of the AHSV and BTV vector (*Culicoides imicola*), now poses an AHS threat to that region and beyond (Office International des Epizooties, 2015).

Orbiviruses are non-enveloped virions that contain a double-stranded (ds) RNA genome enclosed by a core particle and an outer capsid (Hewat *et al.* 1992; Roy *et al.* 1994; Roy 1996; Gouet *et al.* 1999). The inner layer of the core particle is made up of 120 copies of major core protein VP3 that are arranged as 12 decamers - known as the subcore (Huismans *et al.* 1987; Brookes *et al.* 1993). The subcore is stabilized by the addition of 780 copies of major core protein VP7, arranged as 260 trimers deposited on the inner subcore layer (Prasad *et al.* 1992; Grimes *et al.* 1997). Capsid protein trimers of VP5 and VP2 interact with the VP7 layer of the core particle to form the outer capsid of the virion (Hewat *et al.* 1992; Zhang *et al.* 2010).

VP7 is the major core protein of all orbiviruses and is essential to the processes of virus assembly and replication. VP7 is highly conserved within orbivirus serogroups such as BTV (Oldfield *et al.* 1990; Wilson *et al.* 2000; Anthony *et al.* 2007), Epizootic hemorrhagic disease

virus (EHDV) (Mecham & Wilson 2004), Chuzan virus (Yamakawa *et al.* 1999), and AHSV (Roy *et al.* 1994; Bremer *et al.* 2000) and is partly conserved between these orbivirus serogroups. Given this high sequence conservation within serogroups, it is identified as the major group-specific protein of orbiviruses as it is antigenically specific to each of the orbivirus serogroups. VP7 is also highly immunodominant, and is therefore commonly used as an antigen for group-specific diagnostic assays (Yamakawa & Furuuchi 2001; Mecham & Wilson 2004; Anthony *et al.* 2007; Quan *et al.* 2010).

BTV has been studied extensively, and is the prototype virus for the orbivirus genus (Roy 1996). Due to structural similarities with closely related orbiviruses, inferences regarding the functions of AHSV proteins are often made from research involving BTV. A significant exception centers around the VP7 protein of AHSV. Namely, when expressed both recombinantly and during AHSV infection, the majority of AHSV VP7 preferentially assembles into large hexagonal crystalline particles (Bekker *et al.* 2014). VP7 of the BTV on the other hand is highly soluble and remains as soluble trimers throughout the cell when expressed recombinantly (Oldfield *et al.* 1990; Basak *et al.* 1992; Kar *et al.* 2007). The tendency of AHSV VP7 to form crystalline particles when expressed *in vitro* and during AHSV infection has severely hampered the development of AHSV recombinant vaccines, as the aggregation of VP7 significantly reduces the yield of available VP7 in the cell (Maree *et al.* 1998; Rutkowska *et al.* 2011; Maree *et al.* 2016). In addition, it has been suggested to hinder AHSV assembly during infection, which brings into question the conservation of this undesirable trait (Bekker *et al.* 2014). For this reason, there have been multiple attempts to generate soluble AHSV VP7 mutant proteins, but none have succeeded in abolishing crystalline particle formation completely (Basak *et al.* 1996; Rutkowska *et al.* 2011; Wall *et al.* In press).

Previous results have suggested that AHSV VP7 crystalline particle formation is not a result of host-virus interactions, but rather a consequence of an inherent property of AHSV VP7 to

self-assemble into characteristic crystalline particles, presumably through trimer-trimer interactions (Bekker *et al.* 2014). Many studies have shown that mutagenesis of surface residues can disrupt the self-association of virus proteins. Studies such as these are useful for understanding the role of virus proteins during assembly (Buchholz *et al.* 1993; Monastyrskaya *et al.* 1994; Kho *et al.* 2003) and for enabling the crystallization of such proteins by preventing their aggregation (Jenkins *et al.* 1995; Basak *et al.* 1996; Gaudier *et al.* 2001). Here, we suggest that AHSV VP7 possesses unique residues that interact with residues on neighboring AHSV VP7 trimers, through non-covalent forces, to form strong trimer-trimer interactions that cause the self-assembly of AHSV VP7 into these particles. Given that BTV VP7 lacks the ability to assemble into any morphological structures on its own, these presumed trimer-trimer interactions are unique to AHSV VP7. Site-directed mutagenesis of such residues will likely abolish AHSV VP7 trimer-trimer interactions and solubilize the protein – thus allowing for its crystallization and use in AHSV recombinant vaccine development.

The aim of this study was therefore to identify candidate AHSV VP7 residues that are responsible for the unique AHSV VP7 trimer-trimer interactions that drive AHSV VP7 aggregation. To this end, we set out to identify AHSV VP7 residues that are exposed on the surface of the trimer and are physiologically different from the soluble BTV VP7 corresponding residues. Naturally, residues that are conserved completely in a protein family are expected to be important for function and/or structure, and any substitution at these conserved residues is likely to affect protein function (Ng & Henikoff 2001). The evolution of the capsid genes of orbiviruses has been shown to be strongly influenced by functional and structural constraints (Williams *et al.* 1998). We were therefore not interested in selecting amino acids that would deleteriously affect the natural trimer formation or functioning of VP7 during virus assembly, but only the alternate interactions that cause AHSV VP7 to self-assemble into crystalline particles. Apart from finding candidate residues that were physiologically different from the corresponding BTV VP7 residues, it was necessary to only

select candidates that were exposed on the surface of the AHSV VP7 trimer as residues buried in the core of the protein would likely function in maintaining trimer structure and not be involved in the exterior trimer-trimer interactions. A three-dimensional (3D) structure of the AHSV VP7 trimer was therefore necessary. The only crystallographic data available for AHSV VP7, however, is the 3D crystal structure of the trimeric AHSV VP7 top domain (residues 121-249) (Basak *et al.* 1996). Due to AHSV VP7 aggregation and weak connections between the top and the bottom domain, the AHSV VP7 protein was cleaved during crystallization and only the top domain could be resolved (2.3-Å resolution). The study reported that the top domains of both BTV VP7 (Grimes *et al.* 1995) and AHSV VP7 are trimeric and structurally very similar. The trimeric fragment of the AHSV VP7 contains three subunits A, B, and C which made up the top domains of the VP7 monomers usually present in a VP7 trimer. To model the entire AHSV VP7 trimer for this study, a homology modelling approach was employed using the available above-mentioned crystal structure of the AHSV VP7 top domain (Basak *et al.* 1996) and the fully resolved crystal structure of BTV VP7 (Grimes *et al.* 1995) as templates.

Here, we report the results of protein homology modelling of the AHSV VP7 trimer - ultimately for identification of exposed residues and subsequent selection of candidate amino acids that drive trimer-trimer interactions. By using our homology model of the AHSV VP7 trimer, we could identify nine regions on the surface of the trimer that are likely to drive AHSV VP7 self-assembly into crystalline particles. Identification of such residues will allow for the targeting of AHSV VP7 trimer-trimer interactions by site-directed mutagenesis which will, in turn, allow for the testing of the role of AHSV VP7 trimer-trimer interactions in VP7 crystalline particle formation and core assembly. These future studies, along with this one, will help answer evolutionary questions and offer some explanation about AHSV replication and pathogenesis. In addition, the AHSV VP7 homology model generated here showed consistencies with previous biological data, identified potential residues that play a role in the intermolecular interactions within the trimer, and validated several AHSV VP7 epitopes.

2. Materials and Methods

2.1. Protein sequence analysis

Multiple sequence alignments, pairwise alignments, and phylogenetic trees were calculated and visualized using Jalview (Waterhouse *et al.* 2009). Neighbor Joining trees were constructed using percent identity (PID) and BLOSUM62 (i.e. the sum of BLOSUM62 scores for the residue pair at each aligned position) distance measures.

2.2. Virus strain and sequence

AHSV-9 VP7 sequence for homology modelling was obtained from Genbank (accession number Q86729) and BTV VP7 and AHSV VP7 structures were obtained from Protein Data Bank (PDB) with identity 1BVP and 1AHS respectively.

2.3. Homology modelling

The full-length AHSV and BTV VP7 protein sequences were aligned to the Pfam (Finn *et al.* 2013) orbivirus inner capsid protein VP7 seed alignment (PF000897) using the sequence alignment program Muscle (Edgar & Sjolander 2004). AHSV VP7 homology models were generated with Modeller 9v7 (Sali & Blundell 1994) using the top domain AHSV [PDB: 1AHS] and full length BTV VP7 [PDB: 1BVP] crystal structures as template. Protein quality assessment was performed with the PDBsum generator (Laskowski 2009). Amino acids are numbered with the initiating methionine set to 1.

2.4. *In silico* protein optimization

Homology models were optimized in the Prepare Protein module of the Discovery Studio 3.0 software suite (Accelrys Inc.). The protein ionization and residue pKa values were determined in the presence of water. The protonation states of histidine, glutamate and

aspartate were visually inspected after applying the CHARMM force field. Energy minimization was performed using the Conjugate Gradient algorithm with the maximum number of steps set to 200 and the RMSD gradient to 0.1 Å.

2.5. Protein 3D structure analysis

Three-dimensional visualizations, modelling analysis, structure superimpositions and manipulations were done with VMD (Visual Molecular Dynamics) software package (Humphrey *et al.* 1996). Sequence and structural data were organized, displayed, and analyzed using the VMD extension, MultiSeq (Roberts *et al.* 2006). MultiSeq was used to compare sequence and structure metrics and properties, as well as generate structural alignments based on the STAMP (Structural Alignment of Multiple Proteins) structural alignment algorithm (Russell & Barton 1992).

3. Results

3.1. AHSV VP7 sequence analysis and homology modelling

Thus far, VP7 crystalline particles have only been observed for AHSV VP7, while the prototype orbivirus BTV VP7 demonstrates complete solubility with no formation of any morphological features *in vitro*. To explore whether crystalline particle formation is conserved within the AHSV serogroup, and to determine if a homology model of AHSV serotype 9 VP7 will accurately represent the AHSV VP7 protein, we examined the amino acid sequence conservation of AHSV VP7 among the nine AHSV serotypes more closely. Previous studies only included the analysis of AHSV-4 and AHSV-9 (Bremer *et al.* 1990; Roy *et al.* 1991). To further understand and identify the differences between the highly conserved closely related BTV and AHSV VP7 proteins, the level AHSV VP7 amino acid sequence conservation was compared to the VP7 sequence conservation across all 26 BTV serotypes. A neighbor

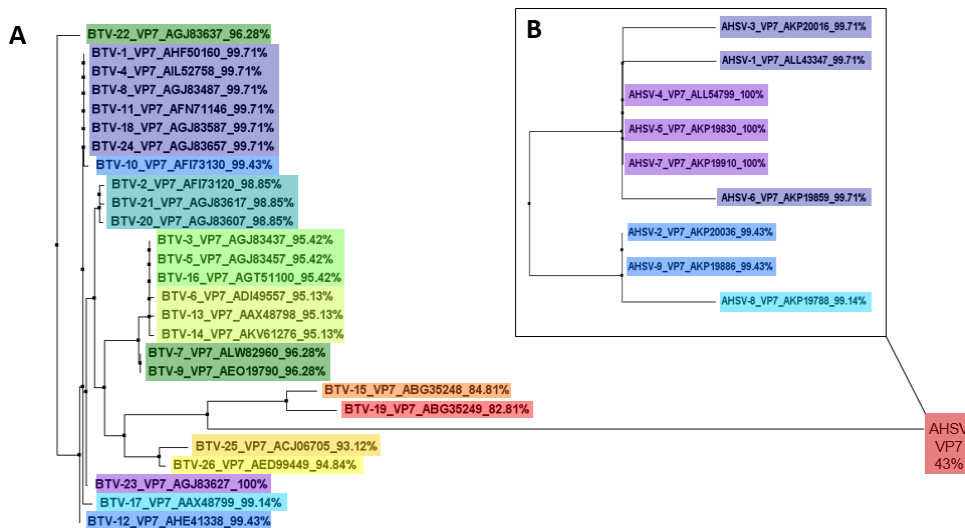


Figure 1. Amino acid sequence conservation analysis of VP7 among BTV (A) and AHSV (B) serotypes. Neighbor joining trees were calculated based on percent identity. Genbank accession numbers for each sequence are indicated. Consensus sequences for BTV VP7 and AHSV VP7 were calculated, and the percent identity to each BTV or AHSV VP7 was determined by pairwise alignment indicated as percentages to the right of each VP7 serotype.

joining tree, based on percent identity, was calculated based on VP7 amino acid sequence alignments for BTV and AHSV serotypes (Figure 1). Within each tree, the percent identity of each serotype to the consensus sequence of AHSV VP7 (Figure 1B) and BTV VP7 (Figure 1A) is depicted. The overall amino acid sequence conservation of VP7 based on pairwise identity of each serotype to the consensus sequence was 99.7% for AHSV VP7 and 96.5% for BTV VP7. Furthermore, pairwise alignment of consensus sequences for BTV and AHSV VP7 revealed an amino acid sequence identity of 43%. These results confirm that VP7 is very highly conserved within BTV and AHSV serotypes, and that AHSV VP7 crystalline particle formation is likely conserved across all nine AHSV serotypes.

To determine the best suitable structure template to be used for homology modelling, the general approach is to assess the homology of the query protein to protein family, domain, and functional motif databases. Since we are ultimately searching for residues that play a role in the mechanism of AHSV VP7 crystalline particle formation, full-length comparisons with related sequences at either the sequence or structural levels may give important clues about whether crystal formation is a consequence of function and/or evolution. Previously identified structures, such as the BTV-10 VP7 (Grimes *et al.* 1995) and AHSV-4 VP7 top domain (Basak *et al.* 1996), can be used to generate a high-quality structure-guided multiple sequence (seed) alignment to which the query (AHSV VP7 full-length sequence) and other related sequences can be aligned. The full-length AHSV-9 VP7 protein sequence was therefore aligned to the Pfam orbivirus inner capsid protein VP7 seed alignment (PF000897). The seed alignment identified six species in the VP7 orbivirus family namely Chuzan virus, EHDV, BTV, Broadhaven virus (BRD), Yunnan virus, and Peruvian horse sickness virus (PHSV) (Figure 2A). Neighbor joining trees of these orbivirus VP7 protein families (Figure 2B) showed that AHSV VP7 was most closely related to Chuzan virus VP7 with a percent identity of 53%, followed by EHDV VP7 at 46%, BTV VP7 at 43%, BRD VP7 at 24%, Yunnan virus VP7 at 23%, and PHSV VP7 at 21% sequence identity. Given the unique distinction of AHSV VP7 to self-assemble to form crystalline particles and its high insolubility, it is

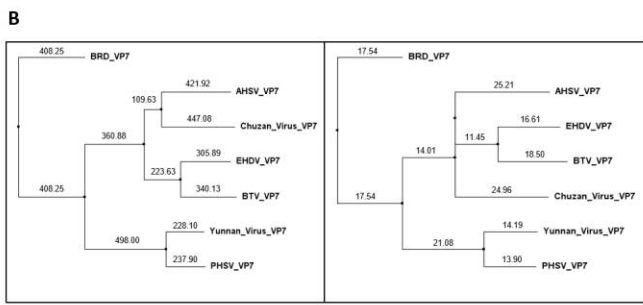
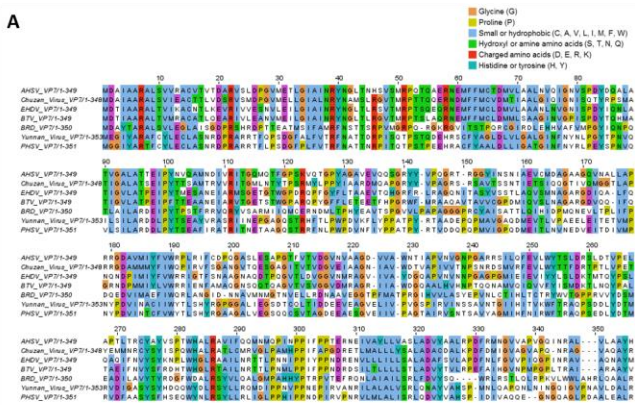


Figure 2. Amino acid sequence analysis of orbivirus inner capsid protein VP7 family. A) Pfam seed alignment (PF000897) of seven orbivirus families. VP7 amino acid sequences colored per ClustalX color scheme and sorted based on pairwise identity. B) Neighbor joining trees of orbivirus VP7 protein families based on BLOSUM62 scores (left) and percent identity (right) distance measures.

expected that AHSV VP7 be relatively divergent and less related to the soluble BTV VP7 as seen in Figure 2B.

We next used a homology modelling approach to model the structure of the AHSV VP7 trimer. The quality of a homology model relies on the level of sequence identity between the protein of known structure and the protein to be modelled (Chothia & Lesk 1986). Homology can be assumed when a sequence identity is greater than 30% (Hillisch *et al.* 2004), and in this case AHSV VP7 showed 43% identity to BTV VP7. A model of the AHSV VP7 trimer was generated with Modeler methodology, using the trimeric crystal structures of BTV VP7 [PDB: 1BVP] and the AHSV VP7 top domain [PDB: 1AHS] as templates. Protein quality assessment of the homology models showed an overall G value of -0.06 for the AHSV VP7 top domain model and 0.17 for BTV VP7. To evaluate the model, the BTV VP7 crystal structure was aligned with the AHSV VP7 homology model using the STAMP structural alignment algorithm (Russell & Barton 1992; Roberts *et al.* 2006). Three regions were found to be structurally dissimilar, i.e. at amino acid positions 144, 175-180, and 206. Structural alignments revealed that the differences at positions 144 and 175-180 resulted from deletion/insertions at sites 144, 177, 178, and 180. When the Qres values were mapped onto the BTV and AHSV VP7 trimer models, it was evident that all three areas of structural variation/dissimilarity were on highly exposed loops (data not shown). These results are consistent with X-ray crystallography data of the AHSV top domain, where the 176-181 loop could not be modelled as the segment does not appear on electron density maps. In addition, the AHSV VP7 deletion at position 144, insertion at 178 and 180, and the β -turn at position 205-206 have previously been identified as zones of maximum deviation (Basak *et al.* 1996).

Although homology models are incomparable to crystal structures due to low accuracy at the atomic level, it was clear that this model could give insight into the structural shape of AHSV

VP7 up to a residue level which would allow us to identify key differences between BTV and AHSV VP7 surface properties.

3.2. Analysis of the AHSV VP7 trimer structure and surface properties – comparison with BTV VP7

3.2.1. AHSV VP7 structure and trimer surface

When analyzing the AHSV VP7 model (Figure 3A), it could be seen that each monomer consists of two domains as described for BTV VP7 (Grimes *et al.* 1995). The top domain is composed of many antiparallel β -sheets that collectively make up a β -sandwich domain with jelly-roll topology which is common amongst viral coat proteins (Grimes *et al.* 1997). The bottom domain is made up of nine α -helices and is nearly twice as large (220 residues) as the top domain (129 residues). In the trimer, the domains as well as the secondary structures are connected by extended loops. These loops appear to be the regions with most variability as seen in the structural alignment (STAMP) of AHSV and BTV VP7 sequences (Figure 3B). When viewing the surface of the whole trimer, we see the two domains are twisted anticlockwise such that the top domain of one monomer rests on the bottom domain of another monomer (Figure 3C) as described for BTV VP7 (Monastyrskaya *et al.* 1997).

3.2.2. Consistency of the AHSV VP7 trimer model structure with biological data

Complementing our data from the AHSV VP7 model with functionally significant residues previously identified for BTV VP7 will help validate our AHSV VP7 model, and potentially highlight and identify more differences between AHSV VP7 and the orbivirus prototype.

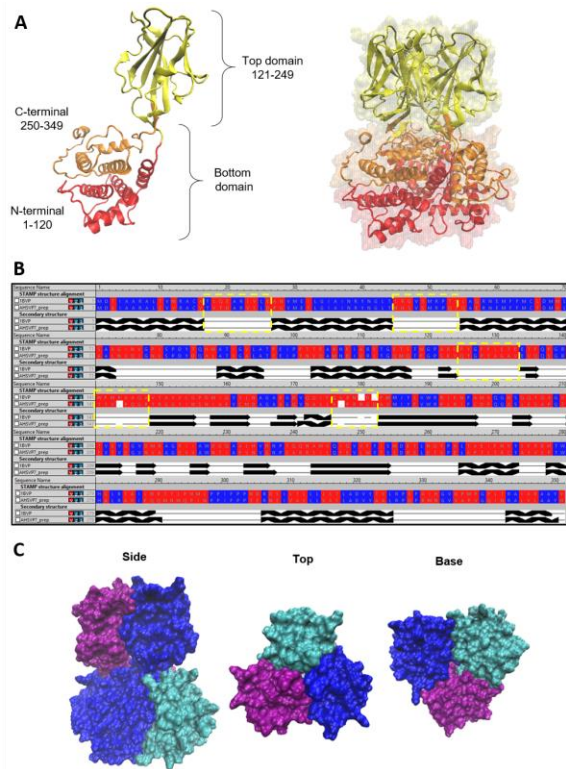


Figure 3. 3D structure of AHSV VP7 monomer and trimer. A) AHSV VP7 model representing secondary structural features (α -helices, antiparallel β -sheets, and extended loops) of the top domain (yellow) and the N- and C-terminus (red and orange respectively) of the bottom domain. The monomer is depicted on the left, and the trimer on the right. B) BTV and AHSV VP7 structural alignment (STAMP) colored by sequence identity, i.e. conserved (blue) and non-conserved (red). Secondary structure depicted by zig-zag (α -helices), lines (extended loops), and arrows (β -sheets). Loops which contain highly variable residues are highlighted in yellow boxes. C) Monomer orientation and surface structure of the AHSV VP7 trimer rotated for view from the side, the top, and the base of the trimer. Each monomer is indicated by a different color.

Intermolecular interactions within the trimer

A previous study identified the first two cysteine residues (Cys₁₅ and Cys₆₅) of BTV VP7 to be highly conserved within the orbivirus family (Le Blois & Roy 1993) (see seed alignment – Figure 2A). It was postulated that the two conserved cysteine residues may be important for the structural integrity of VP7 proteins - likely involving intramolecular disulphide linkages. Interestingly, in this AHSV VP7 trimer model these two residues are located alongside each other (Figure 4A), likely forming disulphide bridges as previously thought. Here, examination of the positions of other AHSV VP7 cysteine residues (Figure 2A) revealed that Cys₁₆₁ and Cys₁₉₅ of the same monomer are also positioned near one another in the top domain (Figure 4B). However, they do not appear to be close enough to form disulphide bridges. If there is flexibility in this region as previously mentioned (Basak *et al.* 1997), then it is likely that these residues can come together to form disulphide bridges as their positions would appear to allow for movement and their side chains are free. In addition, when the surface of the VP7 monomer is examined, it is observed that Cys₁₉₅ is surface accessible which most likely allows for disulphide bridges to form between the Cys₁₉₅ residues of the other interacting monomers (Figure 4C). Cys₁₉₅ is therefore a good candidate for functioning in AHSV VP7 monomer-monomer contacts.

BTV VP7 contains one lysine at position 255 (Le Blois & Roy 1993), while AHSV VP7 contains a single lysine at position 123. Interestingly, in each instance the corresponding residue is an arginine, i.e. Arg₂₅₅ in AHSV VP7, and Arg₁₂₃ in BTV VP7. Lysine and arginine are often interchangeable as they are both polar and positively charged (Betts & Russell 2003). The substitution of the BTV VP7 Lys₂₅₅ residue for a leucine abrogated CLP formation (Le Blois & Roy 1993) indicating that it plays a role in VP7 structure or function. In our AHSV VP7 model, Arg₂₅₅ is located on the interior of the trimer between the top and the bottom domain (Figure 5 – yellow). The lysine of AHSV VP7, at position 123, is also located in this region (Figure 5 – green). Given the internal position of the AHSV VP7 Arg₂₅₅, the BTV VP7 Lys₂₅₅ mutation is more likely to have disrupted trimer folding and not VP7-VP3 interaction as

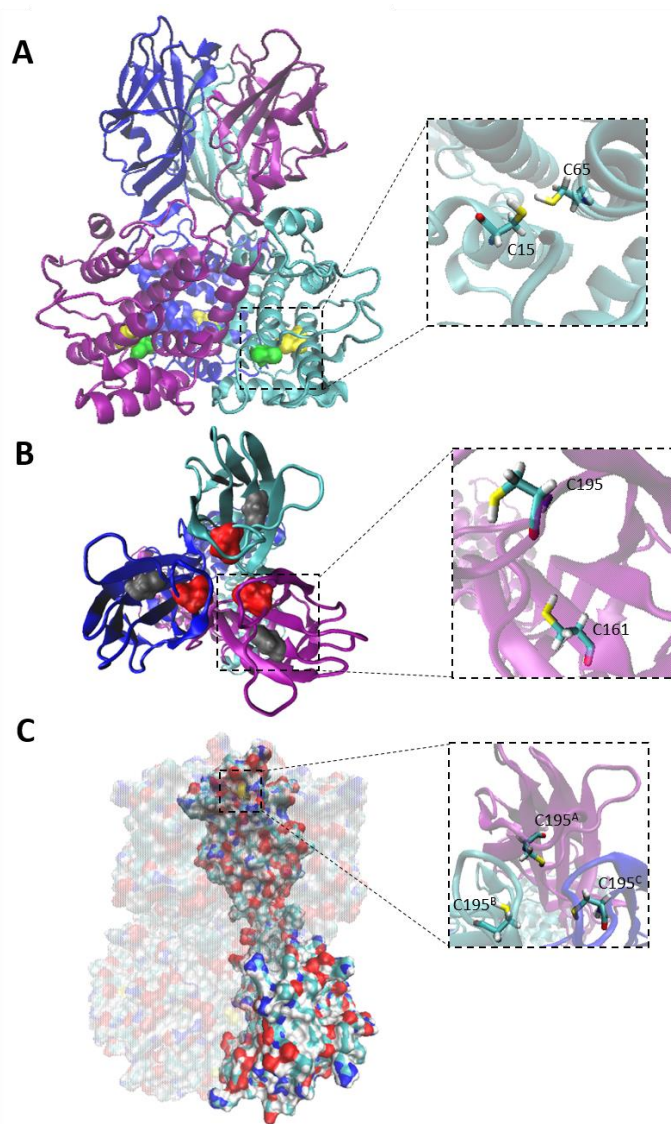


Figure 4. Analysis of AHSV VP7 cysteine residues and their role in VP7 monomer-monomer contacts. A) Positioning of Cys₁₅ (green) and Cys₆₅ (yellow) on the same monomer. B) Positioning of Cys₁₉₅ (red) and Cys₁₆₁ (grey) on the same monomer. (C) Positioning of Cys₁₉₅ on each monomer of the VP7 trimer.

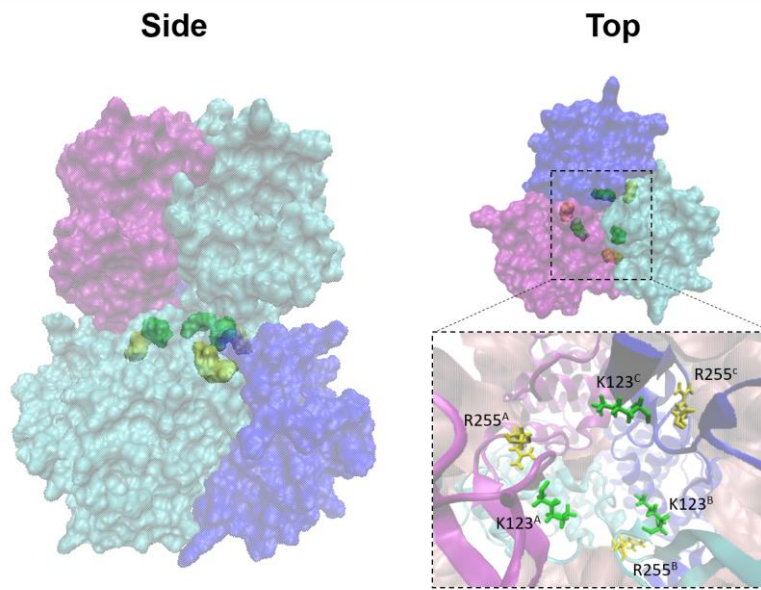


Figure 5. Residues likely to function in trimer formation. Location of Arg₂₅₅ (yellow) and Lys₁₂₃ (green) on each monomer within the AHSV VP7 trimer viewed from the side and the top.

suggested by the authors. Therefore, the Arg₂₅₅ and Lys₁₂₃ of AHSV VP7 are likely to function in internal VP7 monomer-monomer contacts and therefore trimer formation.

VP7 epitopes

To facilitate the development of immunodiagnostic procedures as well as vaccines, much interest lies in orbivirus VP7 epitope mapping. Data currently available on BTV VP7 epitopes was therefore applied to our AHSV VP7 model. As epitope prediction analysis is designed to identify antibody binding sites, epitopes are primarily chosen based on surface accessibility. Here, we mapped previously identified epitopes for BTV VP7 and candidate continuous epitopes and insertion sites for AHSV VP7 onto our model for validation of epitopes for AHSV VP7 (Table 1). Epitopes found in the core of the trimer are likely to have been identified using a monomer of BTV VP7, i.e. in the case of 259-264 and 256-275 where a VP7-GST fusion protein was used which was most likely in monomeric form. These results illustrate the importance of considering the oligomerization of a protein prior to epitope mapping depending on the application.

In this study, nine potential continuous epitopes on AHSV VP7 were computationally predicted using antigen design from the GenScript OptimumAntigen™ Design program (www.genscript.com/antigen_design.html). These epitopes were identified based on the linear amino acid sequence of AHSV VP7 and mapped to AHSV VP7 amino acid positions 251-265, 135-149, 17-31, 230-244, 323-337, 293-307, 120-134, 170-184, and 77-91. When the predicted epitopes were modelled here, it was observed that many epitopes were exposed on the surface of the VP7 monomer. However, when in trimeric form, these epitopes were buried at the core of the trimer and not accessible to the surface, e.g. epitope 251-265 (Figure 6A). Therefore, oligomerization is a key factor to consider when epitope mapping. Candidate epitopes to consider for AHSV VP7 were found to be 323-337, 170-184, and 77-91 based on adequate surface exposure (Figure 6B).

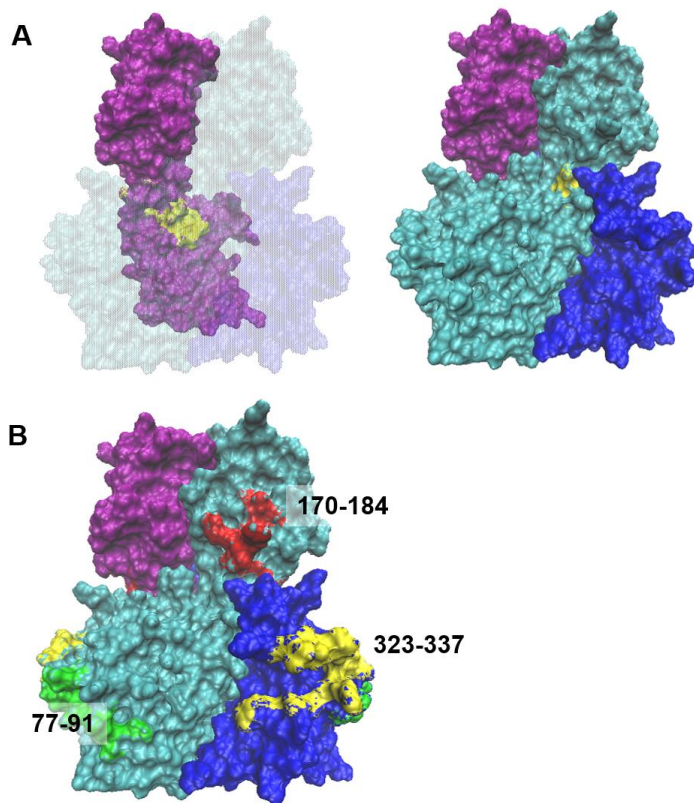


Figure 6. Epitope mapping of AHSV VP7 computationally predicted epitopes (GenScript, Antigen Design Tool, OptimumAntigen™). A) Example of epitope 251-265 (yellow) being surface accessible on the VP7 monomer but not the trimer. B) Epitopes 77-91 (green), 170-184 (red), and 323-337 (yellow) were identified as good epitope candidates.

Table 1. Epitopes previously identified for BTV VP7 by mapping of VP7 monoclonal antibody (mAb) binding sites or using predictive algorithms for identification of T-cell epitopes from virus antigens.

BTV VP7 epitope	Reference	Prediction method	Mapped to AHSV VP7 model	Surface accessible on AHSV VP7 trimer?
1-60	Wang <i>et al.</i> (1996)	VP7 mAb	Spans the entire base of the trimer	Yes
1-120	Eaton <i>et al.</i> (1991)	VP7 mAb	Spans the lower half of the bottom domain and base of the trimer	Yes
30-48	Wang <i>et al.</i> (1994)	VP7 mAb	Located on the base of the trimer	Yes
339-349	Li and Yang (1990)	VP7 mAb	Located between the top and bottom domains (highly surface accessible)	Yes
207-240	Wang <i>et al.</i> (1996)	VP7 mAb	Located in the top domain	Yes
259-264	du Plessis <i>et al.</i> (1994)	VP7 mAb	Located in the core of the trimer	No
256-275	Wang <i>et al.</i> (1996)	VP7 mAb	Located in the core of the trimer	No
122-139	Li and Yang (1990)	VP7 mAb	Located in the core of the top domain	No
139-153	Rojas <i>et al.</i> (2011)	Predictive algorithm	Top domain	Yes
324-338	Rojas <i>et al.</i> (2011)	Predictive algorithm	C-terminal	Yes
283-291	Rojas <i>et al.</i> (2011)	Predictive algorithm	C-terminal	Yes
327-335	Rojas <i>et al.</i> (2011)	Predictive algorithm	C-terminal	Yes
80-88	Rojas <i>et al.</i> (2011)	Predictive algorithm	N-terminal	Yes

3.2.3. Identifying candidate residues for trimer-trimer interactions

The functional properties of a protein are primarily determined by the amino acid side chains present on its surface. Our main goal was to identify residues of AHSV VP7 that are both exposed on the outside of the trimer, and differed significantly from the corresponding BTV VP7 residues. Of its many possible applications, this model was generated for identifying residues that are surface accessible on the trimer molecule and thus available to interact with residues of neighboring trimers. It was therefore important to have an accurate representation of which residues were buried in the protein core and which were exposed at the surface.

When BLOSUM60 scores obtained from the BTV and AHSV VP7 structural alignment were mapped onto the surface of AHSV VP7, it was possible to identify residues that are surface accessible and residues that differ significantly from BTV VP7 (in terms of contribution to protein structure and function based on BLOSUM60 scores) (Figure 7). It was observed that more variable residues are surface exposed and protrude outward from the trimer molecule (Figure 7 arrows), while the more conserved amino acid side chains (blue) are buried inward. This observation is further illustrated by the fact that the monomer-monomer inner contact area is flat with buried amino acid side chains that are mostly conserved (Figure 7 - right). These observations are in keeping with the concept that conserved residues are typically more buried as they contribute to protein stability (Blouin *et al.* 2004).

To identify all exposed residues, we scanned through each residue of the AHSV VP7 sequence, 1 to 349, on the structure model of the trimer and scored each on its surface accessibility ranging from surface accessible to highly exposed (scores 1.5 to 3). Residues that are fully buried within the trimer core are given a score of 1, i.e. below the dashed line in Figure 8. Next, to select candidate residues responsible for trimer-trimer interactions, it was

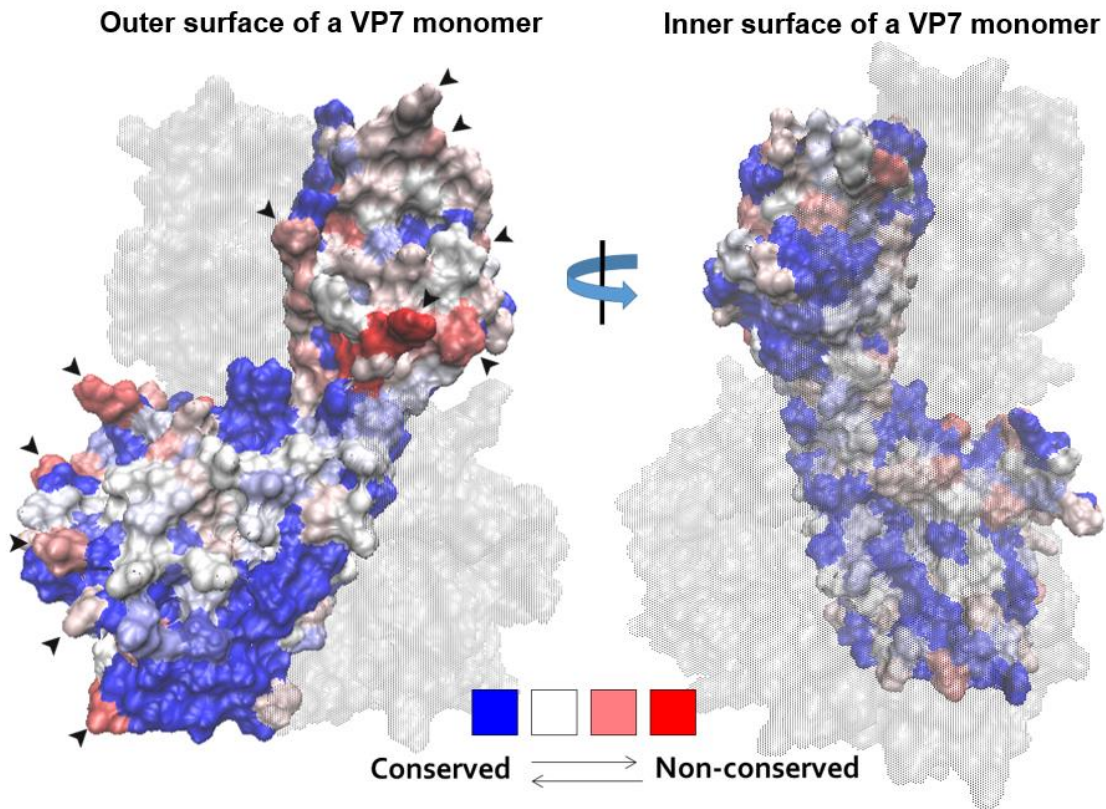


Figure 7. Analysis of the surface accessibility and sequence similarity of AHSV and BTB VP7 by mapping BLOSUM60 scores onto the model of AHSV VP7 trimer. Left – outer surface (accessible) view of the AHSV VP7 trimer with arrows indicating protruding variable residues. Right – AHSV VP7 trimer rotated on the y-axis to illustrate the flat inner surface of VP7 monomer that contacts the other monomers (i.e. residues that make up the core of the trimer).

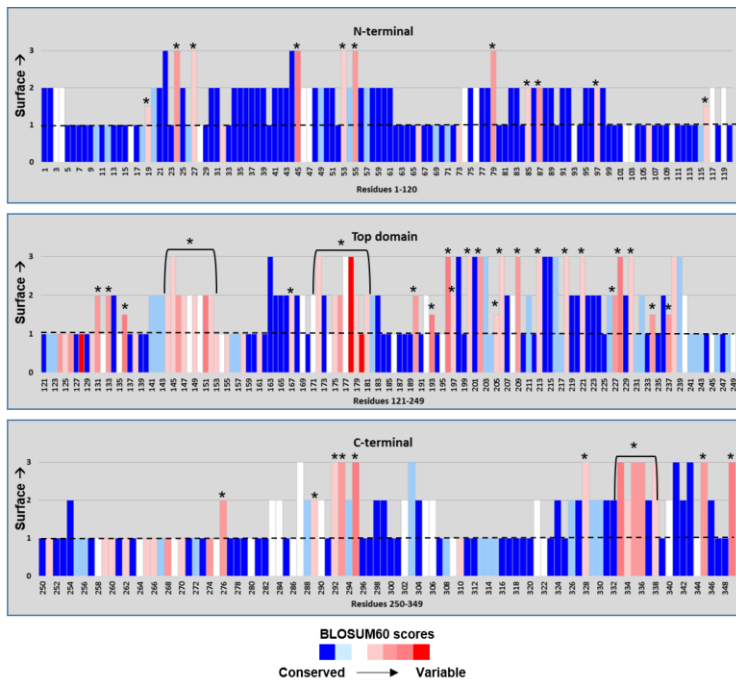


Figure 8. Identification of candidate residues for trimer-trimer interactions based on surface exposure and BLOSUM60 scores. The graph depicts the surface and BLOSUM60 scores of each AHSV VP7 residue. Surface score of 1 indicates buried residues (below dashed line) that were not considered. Bars are colored per BLOSUM60 scores of AHSV and BTV VP7 structure alignments from MultiSeq. Exposed residues with high BLOSUM60 scores were identified as candidates, indicated by *, with regions of selected residues indicated by brackets and *.

necessary to select surface accessible AHSV VP7 residues that differed in terms of physiological properties from BTV VP7 because the forces that control protein-protein interactions are mediated by the chemical properties of surface amino acid side chains. We therefore mapped BLOSUM60 scores (from structural sequence alignments with BTV VP7) of each AHSV VP7 residue with the surface scores to select candidate residues (Figure 8). We used this data to select residues (indicated by * in Figure 8) that were not buried (scores 1.5 to 3), functionally different, and variable from the corresponding BTV VP7 residue.

Before finalizing candidate residue selection based on the data from Figure 8, we examined the general conservation and surface property patterns of the trimer surface for consideration.

Bottom domain

According to previous amino acid sequence analysis, the bottom domain of VP7 is more conserved than the top domain (Wilson *et al.* 2000). Specifically, the top domains of AHSV and BTV VP7 are 29% identical and 51% similar; whereas the bottom domains are more conserved with 65% identity and 82% similarity in the N-terminal regions, and 42% identity and 61% similarity in the C-terminal regions (Iwata *et al.* 1992a; Iwata *et al.* 1992b). For this reason, previous studies suggested that the bottom domain is not responsible for AHSV VP7 insolubility (Monastyrskaya *et al.* 1997). Observations from our AHSV VP7 model reveal that the bottom domain contains several highly variable residues that are clustered on the side of the trimer (Figure 9). These observations illustrate the point that although the bottom domain has a higher degree of sequence conservation than the more variable top domain, the positioning of the fewer variable residues in the bottom domain may contribute equally to protein function as the more numerous variable residues present in the top domain. Interestingly, these variable clusters of the bottom domain are located mainly on the side of the trimer, and are therefore likely to be involved in side-by-side trimer-trimer interactions of AHSV VP7 (Figure 9). These residues (as labelled in Figure 9) along with Arg₁₇₈, Gly₁₇₉,

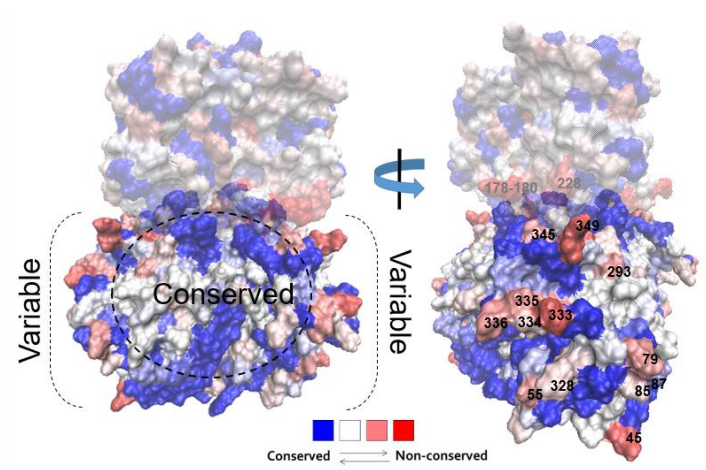


Figure 9. Sequence conservation analysis of AHSV VP7 bottom domain (front and side view of the trimer model) and identification of candidate residues for side-to-side trimer-trimer interactions.

Asp₁₈₀, Thr₂₂₇ and Ile₂₂₈ of the top domain are identified in Figure 8 as potential candidates, and should be considered to function in side-to-side trimer-trimer interactions. The findings here were consistent with earlier reports for BTV VP7, where the extended loop 76-79; residues 88-97; residues 42-44; 178-179; and loop 336-338 were identified as possible candidates for BTV VP7 side-to-side trimer-trimer interactions in the core (Le Blois & Roy 1993; Grimes *et al.* 1997; Grimes *et al.* 1998; Limn & Roy 2003). It is most likely that the residues that drive the trimer-trimer interactions responsible for core formation are conserved as orbivirus core assembly is likely to be similar amongst different viruses. Here, we suggest that the variable residues found on the side of the trimer are strong candidates for only the AHSV VP7 trimer-trimer interactions that cause crystalline particle formation and not the trimer-trimer interactions necessary for core particle formation. In fact, it is possible that the conserved residues located to the center of the bottom domain (Figure 9) are more likely to play a role in native trimer-trimer interactions responsible for core particle formation.

The surface of the AHSV VP7 trimer base was found to be 82% conserved with BTV VP7 and is made up of residues 21 to 25, 27, 30, 31, 34, 35, 37, 38, 39, 41, 42, 45 to 51, 53, 54, 56, 57, 60, 97, and 98 (Figure 10A). The base is expected to have high amino acid sequence conservation due to interaction with the conserved inner VP3 subcore layer. The VP7-VP3 interactions are therefore likely to be similar in BTV and AHSV core formation. The base also has a flat surface ideally for deposition onto the inner VP3 scaffold (Figure 10A). This is consistent with previous reports for the crystal structure of BTV VP7 where the VP7 and VP3 layers are described to interact through flat, hydrophobic surfaces (Grimes *et al.* 1997; Grimes *et al.* 1998). Again, the high level of hydrophobicity may be selected for therefore explaining the high level of conservation at the base of the trimers. Interestingly, at the outer edge of the base, there is one highly variable residue, Asn₄₅, which is surrounded by conserved residues (Figure 10A). The BTV VP7 counterpart is a leucine at position 45. The leucine is a more conserved amino acid at this position as 3 of the 7 families have a Leu₄₅ (Figure 2A). Asparagine therefore is unique to AHSV VP7 and the substitution of

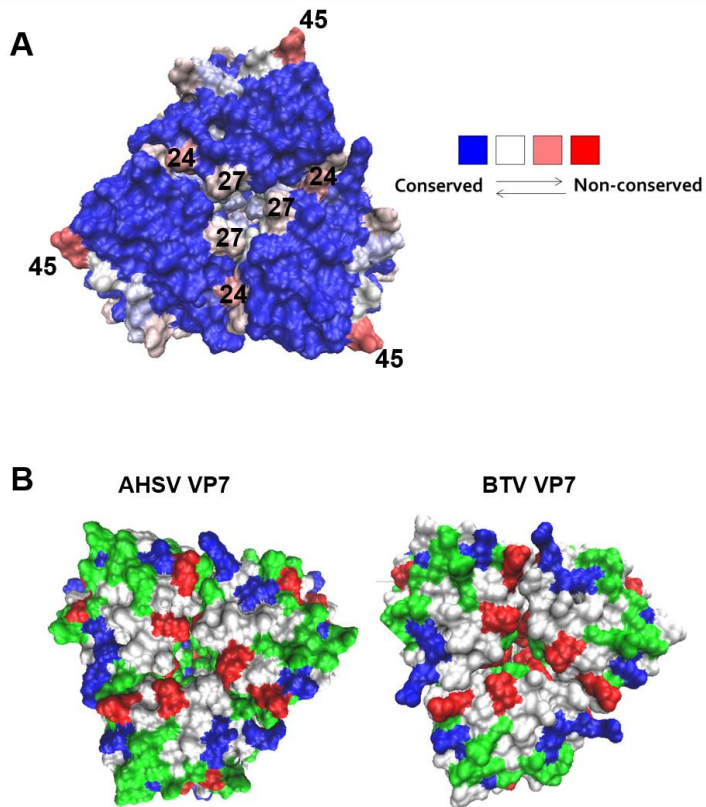


Figure 10. A) Sequence conservation analysis of the base of the AHSV VP7 trimer with variable residues identified. Residues that make up the surface of the base of the trimer are 21 to 25, 27, 30, 31, 34, 35, 37, 38, 39, 41, 42, 45 to 51, 53, 54, 56, 57, 60, 97, and 98. The base is approximately 82% conserved with BTV VP7. B) Surface property analysis of the base of the AHSV and BTV VP7 trimers. Amino acid properties: basic (blue), acidic (red), polar (green) and non-polar (white).

asparagine to leucine is highly unfavorable (Betts & Russell 2003). This residue could therefore play a role in AHSV VP7 function. This residue may not come in direct contact with VP3, but two other variable residues that would come into direct contact with VP3 are Ser₂₄ and Pro₂₇ (Figure 10A). The corresponding BTV residues are again unfavorable substitutions to Val₂₄ and Ala₂₇ respectively. The residues located on the base of the trimer were not considered to play a role in AHSV VP7 trimer-trimer interactions and substitutions at the base will likely abolish AHSV CLP formation.

To compare the surface properties of the AHSV VP7 trimer with the BTV VP7 trimer, the surface of the trimer was colored by amino acid property (Figure 10B). Like results found when analyzing sequence conservation, the bases of the AHSV and BTV VP7 trimers are indistinguishable in terms of amino acid properties, with an even distribution of all properties across the base of the trimer (Figure 10 A and B). In both proteins, the base has large hydrophobic patches (white), which are likely to allow for optimal interaction with VP3 as discussed above.

Top domain

When examining the highly variable top domain, two clusters of conserved residues were identified (Figure 11 green). When we examine the secondary structure of the trimer, we see that Ala₁₆₆ and Asn₂₃₅ of the top cluster form part of an antiparallel β -sheet and their side chains are buried inward (Figure 11). The other residues of the cluster are loops that connect the β -sheet. Interestingly, below this cluster in the β -sheet, are antiparallel residues of alternating conservation, i.e. conserved residues Ala₁₆₆, Asn₂₃₅, and Gly₂₀₇ (blue) followed by variable residues Ala₁₆₇, Gly₂₃₄, and Phe₂₀₉ (pink) (identified as candidates in Figure 8), followed again by conserved Gly₁₆₈, Val₂₃₃ and Val₂₁₀ (blue) (Figure 11). Again, the conserved residues are buried inward and the variable residues are facing outward. This also occurs toward the center of the trimer in another antiparallel β -sheet where variable residues Gln₁₃₆, Ile₁₉₃, and Leu₂₀₂ (also identified as candidates in Figure 8) are flanked by

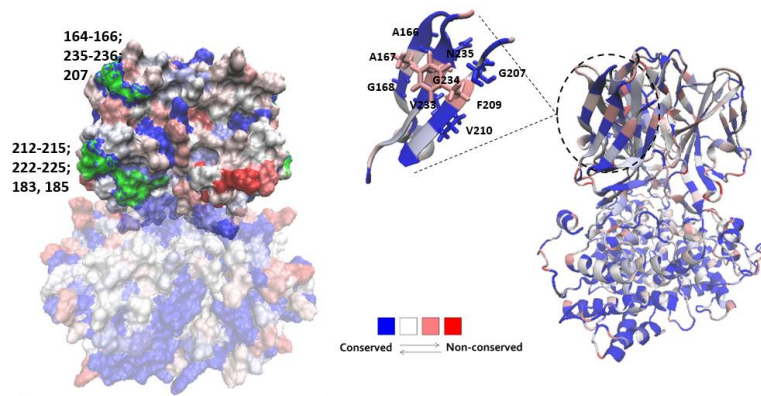


Figure 11. Sequence conservation analysis of the top domain of AHSV VP7. Two separate clusters of conserved residues are highlighted in green (left). Secondary structure of conserved cluster showing an antiparallel β -sheet with alternating conserved (blue) and variable (pink) residues (right).

conserved residues Gln₁₃₇, Ile₁₉₄, Ser₂₀₁ and semi-conserved residues Val₁₃₅, Arg₁₉₂, Glu₂₀₃ (data not shown). Typically in a β -strand, hydrophobic residues are orientated to the interior of the barrel to form a hydrophobic core while the polar residues are orientated outward on the solvent-exposed surface resulting in an alternating pattern (Mandel-Gutfreund & Gregoret 2002). With regards to sequence conservation, hydrophobic residues would therefore be maintained across these protein families, whereas polar residues are more likely free to change. However, in this instance the strands contain only hydrophobic residues (i.e. valines, glycines, alanines, and phenylalanine) and they alternate in conservation and variability and not side chain properties. If we analyze the seed alignment (Figure 2A), we see that the Gly₂₃₄ within the VGN sequence (233-235) of AHSV VP7 is unique to AHSV VP7 in amino acid property and therefore contribution to structure (Betts & Russell 2003). Thus, conserved glycine changes usually have an impact. The BTV VP7 His₂₃₄ is a highly-disfavored glycine substitution due to the bulkiness of its aromatic side chain. Similarly, in the GTFV segment (207-210) of AHSV VP7 (Figure 11), the variable free Phe₂₀₉ is unique to AHSV VP7 in terms of amino acid property and contribution to structure. The AHSV VP7 Phe₂₀₉ and Gly₂₃₄ residues, which are also identified in Figure 8 as potential candidates for AHSV trimer-trimer interactions, are likely to be significant to AHSV VP7 structure and/or function. It is interesting how such differently functioning residues can maintain the same basic structure across these proteins thus further indicating that structure, and not sequence, is conserved.

When comparing the overall surface properties of AHSV and BTV VP7 trimers (Figure 12), we noted that AHSV VP7 has more hydrophobic patches (white) (Figure 12 - stars) than BTV VP7 (which has more clusters of polar amino acids uninterrupted by hydrophobic amino acids). The trimers share several conserved hydrophobic patches (Figure 12 – black circles). The hydrophobic differences identified here (Figure 12 – stars) may play a role in AHSV VP7 crystalline particle formation. We therefore analyzed their side chain positions and found that Met₁₁₆ has a buried side chain, Ile₂₂₈ and Pro₂₃₀ are located on an exposed loop with free

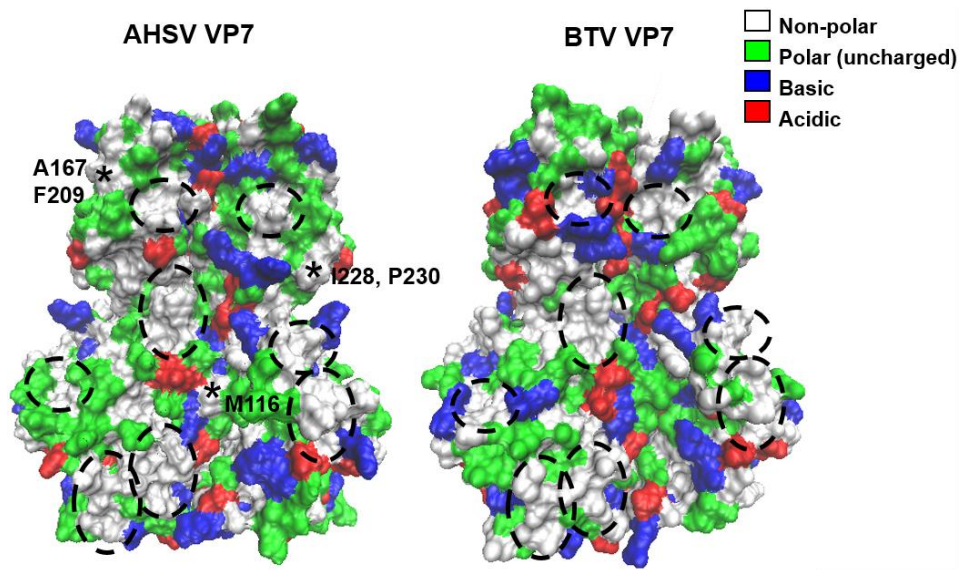


Figure 12. Surface property analysis of the of the AHSV and BTV VP7 trimers. Amino acid properties: basic (blue), acidic (red), polar (green) and non-polar (white). Conserved hydrophobic patches circled in black and AHSV VP7-specific hydrophobic residues represented with stars.

side chains, and Phe₂₀₉ and Ala₁₆₇ are located on the β -sheet with free side chains. These residues differ greatly in role in function from the BTV VP7 counterparts and all are unique to AHSV VP7 (Figure 2A). Thus, the hydrophobic amino acids Ile₂₂₈, Pro₂₃₀, Ala₁₆₇, and Phe₂₀₉ (positively identified in Figure 8) may play a role in AHSV VP7 crystalline particle formation.

Given the variable nature of the top of the trimer, we focused on the conserved residues that clustered on the apex and the border of the trimer (Figure 13A). Gly₁₉₉ is significant in that the three Gly₁₉₉ residues of the three VP7 monomers are connected at the very top of the trimer. This glycine cluster at the apex of the trimer is connected to residues Ser₂₀₁ and Gln₁₉₈ of each monomer to form one unit (Figure 13A). These residues are therefore likely to be essential for protein structure and/or function due to their position and conservation - specifically with regards to monomer-monomer interactions. From the seed alignment (Figure 2A), we see that Gly₁₉₉ is conserved among all the orbivirus families except Broadhaven virus which has an Asn₁₉₉ - a neutral substitution from glycine.

As expected, the tops of the trimers differ greatly in their distribution of amino acid properties (Figure 13B). The top of the BTV VP7 trimer is entirely composed of uncharged amino acids, whereas the top of the AHSV VP7 trimer contains both charged and uncharged amino acids (Figure 13B). The only similarity between the two trimers appears to be the conservation of the uncharged polar patch in the center (residues Gln₁₉₈, Gly₁₉₉, and Ser₂₀₁ as seen in Figure 13B). The top of the BTV VP7 trimer has a larger hydrophilic surface area than the AHSV VP7 trimer which has its hydrophobic residues dispersed throughout the top of the trimer (Figure 13B).

Selection of candidate regions that drive AHSV VP7 trimer-trimer interactions

When analyzing the positioning of the identified candidate residues from Figure 8 on the trimer, it was evident that all variable exposed residues have been selected (data not shown). Upon further inspection, it was found that many of the residues cluster together on

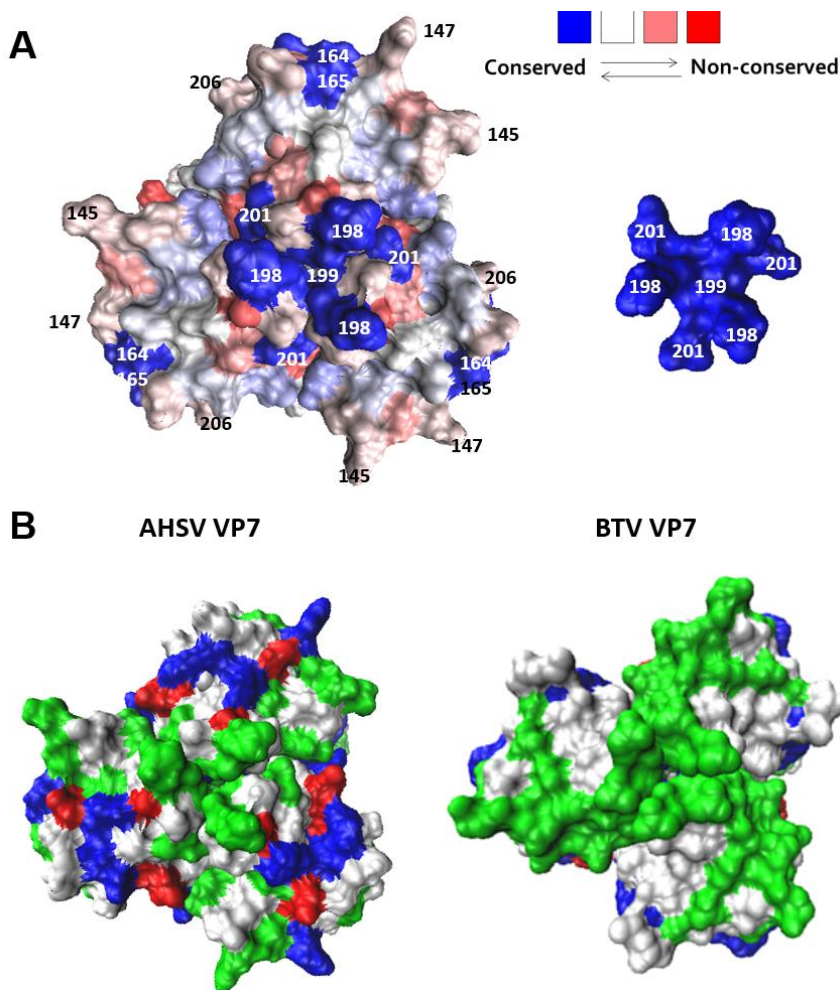


Figure 13. A) Sequence conservation analysis of the top of the AHSV VP7 trimer and identification of residues important for protein structure and/or function (right). The surface of the top of the trimer is made up of residues 141 – 147, 163 – 165, 196 -198, 200 – 203, and 238 – 240; and is approximately 50% conserved with BTV VP7. B) Surface property analysis of the top of the AHSV and BTV VP7 trimers. Amino acid properties: basic (blue), acidic (red), polar (green) and non-polar (white).

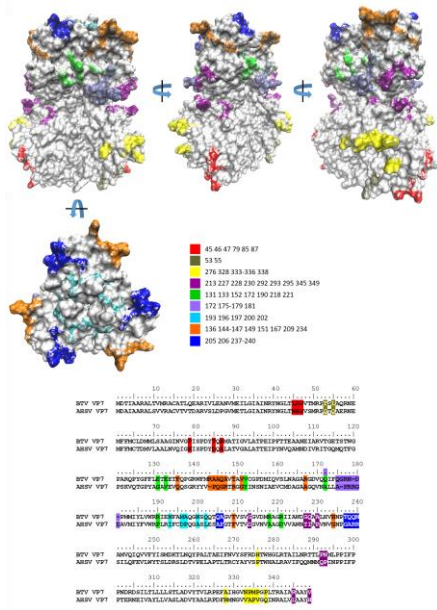


Figure 14. Final selection of candidate residues that potentially drive trimer-trimer interactions grouped into nine different regions. Each candidate residue and the corresponding BTB VP7 residue is depicted in a sequence alignment.

the surface to form candidate “regions”. It is therefore likely that a region on the trimer surface, made up of multiple candidate residues that interact to form the forces necessary for trimer-trimer interactions to occur, is likely to drive these interactions rather than a single residue alone. We therefore grouped residues according to their proximity to other candidate residues and colored them accordingly (Figure 14).

Some residues not selected based on the criteria from Figure 8, were selected based on additional factors that were considered such as seed alignment data from Figure 2A (His₄₆), and position alongside candidate regions (Arg₂₃₉ and Arg₂₄₀). Some residues were excluded based on their position on the trimer, i.e. residues Thr₁₉, Met₁₁₆, and Gln₂₈₉, as they were not remarkably exposed and did not form part of a region of interest; while residues Ser₂₄, Pro₂₇, and Thr₉₇ were excluded as they are on the base of the trimer. The trimer base was not of interest as residues there would play a role in VP7-VP3 interactions in the core. Nine regions each containing the candidate residues identified here were finally selected based on the positioning of each candidate residue. We predict that one or possibly several of these regions is responsible for the trimer-trimer interactions that drive AHSV VP7 crystalline particle formation.

4. Discussion

In this study, the successful modelling of the AHSV VP7 trimer was achieved by homology modelling with the BTV VP7 trimer template. This allowed for the first analysis of the 3D structure of the AHSV VP7 trimer, and the modelling of previously suggested mutational and immunological data for BTV VP7. In addition, this AHSV VP7 model has offered some explanations for previous mutation experiments performed on AHSV and BTV VP7 and suggested several residues important for structure and function of the VP7 trimer. The data obtained here will be beneficial to other work as information regarding amino acid positions,

motifs, epitopes, VP7-VP3 interactions, and monomer-monomer- and trimer-trimer interactions are now available. We also have a clear understanding and view of which VP7 amino acids make up the top and the base of the trimer, where variable and conserved residues lie, and data on the overall surface properties of the protein – all of which can be compared to the orbivirus prototype BTV VP7. The data obtained from this study will allow for easier planning of mutational experiments on AHSV VP7 in the future, and offers insight into the structural evolution of the protein.

It was found that the conserved and variable residues of these proteins formed clusters on the surface of the trimer. These results imply that regions of the VP7 trimer surface are under selection pressures over time, rather than random single surface residues. It was also noted, as expected, that the surface residues of the VP7 trimers are much more variable than those in the core. This is in keeping with the knowledge that the core of a protein is vital for structural integrity of a protein whereas the surface is more likely and free to undergo change (Chothia & Lesk 1986; Lesk & Chothia 1986). It was also observed that the loop regions of the VP7 trimers were highly variable. Loop regions of any 3D protein structure have been shown to be highly variable due to their proximal position at the extremities of a protein. Their lowered evolutionary constraints due to non-contribution to the functional protein core has allowed them to evolve more rapidly (Blouin *et al.* 2004).

A proper understanding of how a protein functions requires an understanding of the underlying phylogenetic relationships between existing proteins. Protein evolution takes place under strong structural constraints and therefore proteins that diverge over time may still retain structural resemblances despite lack of sequence identity (Panchenko & Madej 2004). From the phylogenetic analysis of the VP7 families, BTV VP7 is more divergent from a common ancestor than AHSV VP7. Therefore, the differences seen in AHSV VP7 are not due to changes from the BTV VP7 sequence. However, given that AHSV VP7 is the only orbivirus VP7 protein reported to self-assemble into crystalline particles, it is likely that the

residues that drive AHSV VP7 self-assembly are unique to AHSV VP7 and are present due to changes from the common ancestor. Our phylogenetic analysis revealed that AHSV VP7 is one of the closest VP7 proteins, along with Chuzan virus, to the common ancestor of this VP7 family. Yunnan- and Peruvian horse sickness virus are close behind, while EDHV and BTV VP7 are the most divergent. Without more data from the behavior of all the virus VP7 proteins, it can be suggested that the more soluble BTV VP7 evolved to be soluble, while AHSV VP7 retained its nature to self-assemble into crystalline particles. If this is the case, then it can be implied that the AHSV VP7 self-assembly provides some benefit to AHSV fitness as the trait is conserved by positive selection. Alternatively, at the very least, the trait is not selected against by purifying selection and is even perhaps maintained by stabilizing selection which lowers diversity and favors an intermediate variant of a trait. More detailed studies are necessary to ascertain the role of crystalline particle formation in AHSV replication and/or infectivity.

Given the various mechanisms for the evolution of individual orbivirus segments, many possible mechanisms for the evolution of the soluble BTV VP7 exist. BTV genome segments have been revealed to evolve individually because of strong purifying selection at substitution rates that are generally lower than ssRNA genomes (Carpi *et al.* 2010). The independent evolution of BTV segments comes about due to differing selection pressures acting on different proteins encoded by segments of the BTV genome. BTV was shown to evolve slower than positive sense ssRNA viruses and non-vector-borne members of *Reoviridae*, such as rotavirus. This is explained by strong purifying selection caused by the cycling of the virus through hosts which reduces the effective substitution rate of vector-borne viruses (Coffey *et al.* 2008). It is also suggested that the intracellular encapsulation of the dsRNA genome within the transcriptional core that protects against host defenses acting to degrade dsRNA might also protect against cell factors that mutate ssRNA consequently reducing the observed mutation rates, and thus substitution rates, of dsRNA virus genomes (Duffy *et al.* 2008; Carpi *et al.* 2010). It is also possible that the BTV viral replication cycle is

reduced due to overwintering, thereby resulting in the observed low overall substitution rates (Carpi *et al.* 2010). The slow mutation rates exhibited for BTV segments may indicate that the differences between AHSV and BTV VP7 are small. In addition, the solubility of BTV VP7 is likely to have some benefit to the virus as this trait was selected for in BTV VP7 and other VP7 families.

5. Conclusions

The data here shows that the exposed residues of the AHSV VP7 trimer are most often highly variable from BTV VP7, and are clustered together to form regions. It can therefore be suggested that these regions on the surface of the AHSV VP7 trimer may play a role in the AHSV VP7 trimer-trimer interactions that drive crystalline particle formation. Nine candidate regions on the surface of the AHSV VP7 trimer potentially responsible for driving trimer-trimer interactions were identified here. These regions were selected based on candidate residues residing near one another. We suggest that it is more likely that a group of surface residues create an environment necessary for the proper interaction of trimers rather than a single residue acting alone. Each candidate residue was selected based on its surface exposure and variability from BTV VP7. We have since set out to target the residues that make up these selected regions for mutation to the BTV VP7 corresponding residues by site-directed mutagenesis, and will test the effect of these changes on trimer-trimer interactions. Such a study will aid in solubilizing AHSV VP7 which will allow for the further study of the effect of crystalline particle formation on AHSV VP7 replication, and have applications in vaccine development.

6. Acknowledgements

This work was supported by the National Research Foundation and the University of Pretoria. We thank Henk Huisman for support and constructive criticism.

7. References

- Anthony S., Jones H., Darpel K.E., Elliott H., Maan S., Samuel A., Mellor P.S. & Mertens P.P.C. (2007) A duplex RT-PCR assay for detection of genome segment 7 (VP7 gene) from 24 BTV serotypes. *Journal of Virological Methods* **141**, 188-97.
- Basak A.K., Gouet P., Grimes J., Roy P. & Stuart D. (1996) Crystal structure of the top domain of African horse sickness virus VP7: comparisons with bluetongue virus VP7. *Journal of Virology* **70**, 3797-806.
- Basak A.K., Grimes J.M., Gouet P., Roy P. & Stuart D.I. (1997) Structures of orbivirus VP7: implications for the role of this protein in the viral life cycle. *Structure* **5**, 871-83.
- Basak A.K., Stuart D.I. & Roy P. (1992) Preliminary crystallographic study of bluetongue virus capsid protein, VP7. *Journal of Molecular Biology* **228**, 687-9.
- Bekker S., Huismans H. & van Staden V. (2014) Factors that affect the intracellular localization and trafficking of African horse sickness virus core protein, VP7. *Virology* **456-457**, 279-91.
- Betts M.J. & Russell R.B. (2003) Amino acid properties and consequences of substitutions. *Bioinformatics for geneticists* **317**, 289.
- Blouin C., Butt D. & Roger A.J. (2004) Rapid evolution in conformational space: a study of loop regions in a ubiquitous GTP binding domain. *Protein science* **13**, 608-16.
- Bremer C.W., Gerdes G.H., Aitchison H., Louw I., Greyling R.R. & Welgemoed J. (2000) The prevalence of different African horsesickness virus serotypes in the Onderstepoort area near Pretoria, during an outbreak of African horsesickness in South Africa in 1995/1996. *Onderstepoort J Vet Res* **67**, 65-70.
- Bremer C.W., Huismans H. & Van Dijk A.A. (1990) Characterization and cloning of the African horsesickness virus genome. *J Gen Virol* **71 (Pt 4)**, 793-9.
- Brookes S.M., Hyatt A.D. & Eaton B.T. (1993) Characterization of virus inclusion bodies in bluetongue virus-infected cells. *Journal of general virology* **74**, 525-30.
- Buchholz C., Spehner D., Drillien R., Neubert W. & Homann H. (1993) The conserved N-terminal region of Sendai virus nucleocapsid protein NP is required for nucleocapsid assembly. *Journal of Virology* **67**, 5803-12.
- Carpi G., Holmes E.C. & Kitchen A. (2010) The evolutionary dynamics of bluetongue virus. *Journal of molecular evolution* **70**, 583-92.
- Chothia C. & Lesk A.M. (1986) The relation between the divergence of sequence and structure in proteins. *The EMBO journal* **5**, 823.
- Coffey L.L., Vasilakis N., Brault A.C., Powers A.M., Tripet F. & Weaver S.C. (2008) Arbovirus evolution *in vivo* is constrained by host alternation. *Proceedings of the National Academy of Sciences* **105**, 6970-5.
- du Plessis D.H., Wang L.-F., Jordaan F.A. & Eaton B.T. (1994) Fine mapping of a continuous epitope on VP7 of bluetongue virus using overlapping synthetic peptides and a random epitope library. *Virology* **198**, 346-9.
- Duffy S., Shackelton L.A. & Holmes E.C. (2008) Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics* **9**, 267-76.
- Eaton B.T., Gould A.R., Hyatt A.D., Coupar B.F.H., Martyn J.C. & R. W.J. (1991) A Bluetongue Serogroup-Reactive Epitope in the Amino Terminal Half of the Major Core Protein VP7 Is Accessible on the Surface of Bluetongue Virus Particles. *Virology* **180**, 687-96.
- Edgar R.C. & Sjolander K. (2004) COACH: profile-profile alignment of protein families using hidden Markov models. *Bioinformatics* **20**, 1309-18.

- Finn R.D., Bateman A., Clements J., Coghill P., Eberhardt R.Y., Eddy S.R., Heger A., Hetherington K., Holm L. & Mistry J. (2013) Pfam: the protein families database. *Nucleic Acids Research*, gkt1223.
- Gaudier M., Gaudin Y. & Knossow M. (2001) Cleavage of vesicular stomatitis virus matrix protein prevents self-association and leads to crystallization. *Virology* **288**, 308-14.
- Gouet P., Diprose J.M., Grimes J.M., Malby R., Burroughs J.N., Zientara S., Stuart D.I. & Mertens P.P. (1999) The highly ordered double-stranded RNA genome of bluetongue virus revealed by crystallography. *Cell* **97**, 481-90.
- Grimes J., Basak A.K., Roy P. & Stuart D. (1995) The crystal structure of bluetongue virus VP7. *Nature* **373**, 167-70.
- Grimes J.M., Burroughs J.N., Gouet P., Diprose J.M., Malby R., Zientara S., Mertens P.P. & Stuart D.I. (1998) The atomic structure of the bluetongue virus core. *Nature* **395**, 470-8.
- Grimes J.M., Jakana J., Ghosh M., Basak A.K., Roy P., Chiu W., Stuart D.I. & Prasad B.V. (1997) An atomic model of the outer layer of the bluetongue virus core derived from X-ray crystallography and electron cryomicroscopy. *Structure* **5**, 885-93.
- Hewat E.A., Booth T.F. & Roy P. (1992) Structure of bluetongue virus particles by cryoelectron microscopy. *Journal of structural biology* **109**, 61-9.
- Hillisch A., Pineda L.F. & Hilgenfeld R. (2004) Utility of homology models in the drug discovery process. *Drug discovery today* **9**, 659-69.
- Huismans H., van Dijk A.A. & Els H.J. (1987) Uncoating of parental bluetongue virus to core and subcore particles in infected L cells. *Virology* **157**, 180-8.
- Humphrey W., Dalke A. & Schulten K. (1996) VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33-8.
- Iwata H., Chuma T. & Roy P. (1992a) Characterization of the genes encoding two of the major capsid proteins of epizootic haemorrhagic disease virus indicates a close genetic relationship to bluetongue virus. *Journal of general virology* **73** 915-24.
- Iwata H., Yamagawa M. & Roy P. (1992b) Evolutionary relationships among the gnat-transmitted orbiviruses that cause African horse sickness, bluetongue, and epizootic hemorrhagic disease as evidenced by their capsid protein sequences. *Virology* **191**, 251-61.
- Jenkins T.M., Hickman A.B., Dyda F., Ghirlando R., Davies D.R. & Craigie R. (1995) Catalytic domain of human immunodeficiency virus type 1 integrase: identification of a soluble mutant by systematic replacement of hydrophobic residues. *Proceedings of the National Academy of Sciences* **92**, 6057-61.
- Kar A.K., Bhattacharya B. & Roy P. (2007) Bluetongue virus RNA binding protein NS2 is a modulator of viral replication and assembly. *BMC Mol Biol* **8**, 4.
- Kho C.L., Tan W.S., Tey B.T. & Yusoff K. (2003) Newcastle disease virus nucleocapsid protein: self-assembly and length-determination domains. *Journal of general virology* **84**, 2163-8.
- Laskowski R.A. (2009) PDBsum new things. *Nucleic Acids Research* **37**, D355-D9.
- Le Blois H. & Roy P. (1993) A single point mutation in the VP7 major core protein of bluetongue virus prevents the formation of core-like particles. *Journal of Virology* **67**, 353-9.
- Lesk A. & Chothia C. (1986) The response of protein structures to amino-acid sequence changes. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **317**, 345-56.
- Li J.K. & Yang Y.Y. (1990) Mapping of two immunodominant antigenic epitopes conserved among the major inner capsid protein, VP7 of five bluetongue viruses. *Virology* **178**, 552-9.
- Limn C.K. & Roy P. (2003) Intermolecular interactions in a two-layered viral capsid that requires a complex symmetry mismatch. *Journal of Virology* **77**, 11114-24.
- Mandel-Gutfreund Y. & Gregoret L.M. (2002) On the Significance of Alternating Patterns of Polar and Non-polar Residues in Beta-strands. *Journal of Molecular Biology* **323**, 453-61.

- Maree S., Durbach S. & Huismans H. (1998) Intracellular production of African horsesickness virus core-like particles by expression of the two major core proteins, VP3 and VP7, in insect cells. *Journal of general virology* **79**, 333-7.
- Maree S., Maree F.F., Putterill J.F., de Beer T.A., Huismans H. & Theron J. (2016) Synthesis of empty african horse sickness virus particles. *Virus Research* **213**, 184-94.
- Mecham J. & Wilson W. (2004) Antigen capture competitive enzyme-linked immunosorbent assays using baculovirus-expressed antigens for diagnosis of bluetongue virus and epizootic hemorrhagic disease virus. *Journal of clinical microbiology* **42**, 518-23.
- Monastyrskaya K., Booth T., Nel L. & Roy P. (1994) Mutation of either of two cysteine residues or deletion of the amino or carboxy terminus of nonstructural protein NS1 of bluetongue virus abrogates virus-specified tubule formation in insect cells. *Journal of Virology* **68**, 2169-78.
- Monastyrskaya K., Staeuber N., Sutton G. & Roy P. (1997) Effects of Domain-Switching and Site-Directed Mutagenesis on the Properties and Functions of the VP7 Proteins of Two Orbiviruses. *Virology* **237**, 217-27.
- Ng P.C. & Henikoff S. (2001) Predicting deleterious amino acid substitutions. *Genome research* **11**, 863-74.
- Oldfield S., Adachi A., Urakawa T., Hirasawa T. & Roy P. (1990) Purification and characterization of the major group-specific core antigen VP7 of bluetongue virus synthesized by a recombinant baculovirus. *Journal of general virology* **71**, 2649-56.
- Panchenko A.R. & Madej T. (2004) Analysis of protein homology by assessing the (dis) similarity in protein loop regions. *Proteins: Structure, Function, and Bioinformatics* **57**, 539-47.
- Prasad B.V., Yamaguchi S. & Roy P. (1992) Three-dimensional structure of single-shelled bluetongue virus. *Journal of Virology* **66**, 2135-42.
- Quan M., Lourens C.W., MacLachlan N.J., Gardner I.A. & Guthrie A.J. (2010) Development and optimisation of a duplex real-time reverse transcription quantitative PCR assay targeting the VP7 and NS2 genes of African horse sickness virus. *Journal of Virological Methods* **167**, 45-52.
- Roberts E., Eargle J., Wright D. & Luthey-Schulten Z. (2006) MultiSeq: Unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* **7**.
- Rojas J.M., Rodriguez-Calvo T., Pena L. & Sevilla N. (2011) T cell responses to bluetongue virus are directed against multiple and identical CD4+ and CD8+ T cell epitopes from the VP7 core protein in mouse and sheep. *Vaccine* **29**, 6848-57.
- Roy P. (1996) Orbivirus structure and assembly. *Virology* **216**, 1-11.
- Roy P., Hirasawa T., Fernandez M., Blinov V.M. & Sanchez-Vixcain Rodrique J.M. (1991) The complete sequence of the group-specific antigen, VP7, of African horsesickness disease virus serotype 4 reveals a close relationship to bluetongue virus. *Journal of general virology* **72 (Pt 6)**, 1237-41.
- Roy P., Mertens P.P. & Casal I. (1994) African horse sickness virus structure. *Comp Immunol Microbiol Infect Dis* **17**, 243-73.
- Russell R.B. & Barton G.J. (1992) Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins: Structure, Function, and Bioinformatics* **14**, 309-23.
- Rutkowska D.A., Meyer Q.C., Maree F., Vosloo W., Fick W. & Huismans H. (2011) The use of soluble African horse sickness viral protein 7 as an antigen delivery and presentation system. *Virus Research* **156**, 35-48.
- Sali A. & Blundell T. (1994) Comparative protein modelling by satisfaction of spatial restraints. *Protein structure by distance analysis* **64**, C86.
- Sergeant E.S., Grewar J.D., Weyer C.T. & Guthrie A.J. (2016) Quantitative Risk Assessment for African Horse Sickness in Live Horses Exported from South Africa. *PLoS ONE* **11**, e0151757.
- Stassen L., Vermaak E. & Theron J. (2014) African horse sickness, an equine disease of emerging global significance. *Paz-Silva A, Sol Arias Vázquez M, Sánchez-Andrade*

- Fernández R, editors. Horses: Breeding, health disorders and effects on performance and behaviour. New York: Nova Science Publishers, 145-70.*
- Wall G.V., Rutkowska D.A., Mizrachi E., Huismans H. & van Staden V. (*In press*) A dual laser scanning confocal and transmission electron microscopy analysis of the intracellular localisation, aggregation and particle formation of African horse sickness virus major core protein VP7. *Microscopy and Microanalysis*.
- Wang L.-F., Hyatt A., Whiteley P., Andrew M., Li J.-K. & Eaton B. (1996) Topography and immunogenicity of bluetongue virus VP7 epitopes. *Archives of virology* **141**, 111-23.
- Wang L.-F., Scanlon D.B., Kattenbelt J.A., Mecham J.O. & Eaton B.T. (1994) Fine mapping of a surface-accessible, immunodominant site on the bluetongue virus major core protein VP7. *Virology* **204**, 811-4.
- Waterhouse A.M., Procter J.B., Martin D.M.A., Clamp M. & Barton G.J. (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91.
- Williams C.F., Inoue T., Lucus A.-M., de A. Zanotto P.M. & Roy P. (1998) The complete sequence of four major structural proteins of African horse sickness virus serotype 6: evolutionary relationships within and between the orbiviruses. *Virus Research* **53**, 53-73.
- Wilson W.C., Ma H.C., Venter E.H., van Dijk A.A., Seal B.S. & Mecham J.O. (2000) Phylogenetic relationships of bluetongue viruses based on gene S7. *Virus Research* **67**, 141-51.
- Yamakawa M. & Furuuchi S. (2001) Expression and antigenic characterization of the major core protein VP7 of Chuzan virus, a member of the Palyam serogroup orbiviruses. *Veterinary Microbiology* **83**, 333-41.
- Yamakawa M., Kubo M. & Furuuchi S. (1999) Molecular analysis of the genome of Chuzan virus, a member of the Palyam serogroup viruses, and its phylogenetic relationships to other orbiviruses. *Journal of general virology* **80**, 937-41.
- Zhang X., Boyce M., Bhattacharya B., Zhang X., Schein S., Roy P. & Zhou Z.H. (2010) Bluetongue virus coat protein VP2 contains sialic acid-binding domains, and VP5 resembles enveloped virus fusion proteins. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6292-7.