A first full outer capsid protein sequence data-set in the *Orbivirus* genus (family *Reoviridae*): cloning, sequencing, expression and analysis of a complete set of full-length outer capsid VP2 genes of the nine African horsesickness virus serotypes

A. C. Potgieter,1 M. Cloete,1 P. J. Pretorius2 and A. A. van Dijk1†

1Onderstepoort Veterinary Institute, 0110 Onderstepoort, South Africa
2Division of Biochemistry, School for Chemistry and Biochemistry, University of Potchefstroom for CHE, 2520 Potchefstroom, South Africa

The outer capsid protein VP2 of African horsesickness virus (AHSV) is a major protective antigen. We have cloned full-length VP2 genes from the reference strains of each of the nine AHSV serotypes. Baculovirus recombinants expressing the cloned VP2 genes of serotypes 1, 2, 4, 6, 7 and 8 were constructed, confirming that they all have full open reading frames. This work completes the cloning and expression of the first full set of AHSV VP2 genes. The clones of VP2 genes of serotypes 1, 2, 5, 7 and 8 were sequenced and their amino acid sequences were deduced. Our sequencing data, together with that of the published VP2 genes of serotypes 3, 4, 6 and 9, were used to generate the first complete sequence analysis of all the (sero)types for a species of the *Orbivirus* genus. Multiple alignment of the VP2 protein sequences showed that homology between all nine AHSV serotypes varied between 47.6% and 71.4%, indicating that VP2 is the most variable AHSV protein. Phylogenetic analysis grouped together the AHSV VP2s of serotypes that cross-react serologically. Low identity between serotypes was demonstrated for specific regions within the VP2 amino acid sequences that have been shown to be antigenic and play a role in virus neutralization. The data presented here impact on the development of new vaccines, the identification and characterization of antigenic regions, the development of more rapid molecular methods for serotype identification and the generation of comprehensive databases to support the diagnosis, epidemiology and surveillance of AHS.

**INTRODUCTION**

African horsesickness (AHS) is an infectious, non-contagious, arthropod-borne disease of the Equidae. AHS is an Office International des Epizooties (OIE) A-list infectious disease. The mortality rate in horses is more than 90%. Although outbreaks of AHS are mostly confined to sub-Saharan Africa, there are sporadic outbreaks in North Africa, Mediterranean countries and the Middle East. African horsesickness virus (AHSV), the aetiological agent, belongs to the *Orbivirus* genus (family *Reoviridae*). Nine serotypes of AHSV have been identified. The genome of AHSV consists of 10 dsRNA segments. There are two outer capsid proteins, VP2 and VP5 (Bremer, 1976; Van Dijk & Huismans, 1982). Outer capsid protein VP2 is encoded by genome segment 2. Data showing that recombinant baculovirus-expressed VP2 of AHSV serotypes 4 and 5 induced protective, serotype-specific immunity in horses (Martinez-Torrecuadrada et al., 1996; Roy et al., 1996; Du Plessis et al., 1998; Scanlen et al., 2002) has formed the basis for the development of recombinant vaccines for AHS. Currently, live attenuated vaccines are available for seven AHSV serotypes, namely 1, 2, 3, 4, 6, 7 and 8. To be protected against disease, a susceptible animal has to be immune against all nine AHSV serotypes. To date, four of the nine VP2 genes of the AHSV reference serotypes have been cloned and sequenced, namely serotypes 3, 4, 6 and 9 (Iwata et al., 1992; Sakamoto et al., 1994; Vreede & Huismans, 1994; Williams et al., 1998; Venter et al., 2000). The availability of a complete set of cloned, sequenced and analysed VP2 genes and proteins for each of the nine AHSV serotypes will be a milestone towards the development of...

The GenBank accession numbers of the AHSV VP2 gene sequences reported in this paper are: AHSV-1 VP2, AY163329; AHSV-2 VP2, AY163332; AHSV-5 VP2, AY163331; AHSV-7 VP2, AY163330; AHSV-8 VP2, AY163333.
The complete repertoire of recombinant vaccines and molecular diagnostic reagents for AHS. Therefore, the main aim of the work reported in this paper was to generate a first complete set of full-length VP2 gene clones and sequence data for all nine AHSV serotypes. This paper describes the cloning, sequencing and analysis of full-length cDNA copies of the VP2 genes of all nine AHSV reference serotypes. Baculovirus expression of six of the cloned genes verified that they have full open reading frames. By sequencing the cloned VP2 genes of five serotypes, namely 1, 2, 5, 7 and 8, and combining our data with that of serotypes 3, 4, 6 and 9, which have been previously sequenced, we were able to carry out the first analysis of a complete VP2 sequence data set for a species of the *Orbivirus* genus.

**METHODS**

**Viruses.** The reference strains of AHSV serotypes 1–9 were obtained from J. T. Paweska at the OIE Reference Centre for Bluetongue and African Horsesickness at the Onderstepoort Veterinary Institute, South Africa. Table 1 shows the passage history of the viruses.

**Double-stranded RNA extraction.** For the cloning of the VP2 genes of each AHSV serotype, one 75 cm² flask of confluent BHK cells was infected at an m.o.i. of 1 and incubated at 37°C until CPE reached 80% (approximately 72 h). Infected cells were harvested by centrifugation at 2400 g for 10 min. Total RNA from cells infected with each of the nine serotypes of AHSV was extracted using TRI-Reagent (MRC) according to the manufacturer’s specifications. Single-stranded RNA was removed from total RNA by precipitation with 2 M LiCl. Finally, viral dsRNA was precipitated with 4 M LiCl, washed twice with 70% ethanol and resuspended in RNase-free water.

**Cloning of full-length VP2 cDNA.** The complete set of genome segments for each of the nine AHSV serotypes was amplified using a sequence-independent dsRNA amplification method (Potgieter et al., 2002). The resulting PCR amplicons of the full genomes (all 10 segments) of each serotype were separated on a 1% agarose gel. The VP2 gene amplicons of dsRNA segment 2 for each of the serotypes was purified from the agarose gel using a Gel Extraction Kit (Qiagen). Resulting full-length VP2 gene amplicons were cloned into the PCR cloning plasmid pCR2.1 (Invitrogen). Positive VP2-gene cDNA clones were identified based on insert size (approximately 3-2 kb) and the respective restriction enzyme profiles of the individual VP2 genes.

**Sequencing and sequence analysis.** Separate cloned VP2 cDNAs were digested with several restriction enzymes. The resulting fragments were subcloned into the vector pGEM3ZI+ (Promega). Subclones were sequenced using M13 forward and reverse primers (Promega) and the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit v2.0 on an ABI PRISM 377 DNA sequencer (both from Perkin-Elmer Applied Biosystems). Sequence assembly and analysis, as well as multiple alignment of deduced amino acid sequences of the cloned VP2 genes and phylogenetic analysis, were performed using DNA MAN Software (Lynnon Biosoft).

**Baculovirus expression.** The cloned VP2 genes were cloned into a baculovirus expression vector, pFastBac1 (Life Technologies). Using the Bac-to-Bac expression system (Life Technologies), recombinant viruses were prepared following the manufacturer’s specifications. Proteins from cells infected with recombinant viruses were labelled with [125I]methionine for 30 min at 72 h after infection. After labelling, the cells were harvested, washed twice in PBS and their total protein content was separated by SDS-PAGE and analysed by autoradiography. The dried radioactive gels were exposed overnight to Hyperfilm MP X-Ray film (Amersham Life Sciences).

**RESULTS**

**Cloning and sequencing of VP2 genes**

Previous methods used to clone full-length cDNA copies of AHSV VP2 genes proved to be difficult (Vreede et al., 1998; Venter et al., 2000). Therefore, we used our recently optimized sequence-independent dsRNA cloning procedure (Potgieter et al., 2002) to clone full-length cDNA of the VP2 gene for each of the nine AHSV serotypes. Although the sequences of cloned VP2 genes for AHSV reference strains of serotypes 3, 4, 6 and 9 were available in GenBank, the VP2 cDNAs of these serotypes were recloned using the sequence-independent method, since the full-length cDNAs of the VP2 genes of all nine serotypes were needed in our laboratory for research and for the development of a complete repertoire of recombinant vaccines and diagnostic reagents for all nine AHSV serotypes. We have previously described the cloning of the AHSV-5 VP2 gene (Vreede et al., 1998) and the AHSV-1 VP2 gene (Potgieter et al., 2002). The sequence of these two genes was, however, not previously determined and is presented here.

The nine genomes of the AHSV reference strains could each be amplified in a single amplification reaction using a modified, sequence-independent dsRNA cloning procedure. The result for the AHSV-1 genome is shown in Fig. 1. The data for the other eight genomes are not shown. The VP2 amplicons were purified from the complete amplified genomes separated on agarose gels and cloned into pCR2.1. The cloned VP2 genes could easily be identified based on insert size, restriction enzyme patterns, sequencing of the terminal ends of the cloned cDNA and by hybridization. Previously, the identity of these cloned AHSV VP2 genes had been confirmed by using them as digoxigenin-labelled probes, which hybridized in a serotype-specific manner to
RNA from the original reference viruses as well as those from field isolates of the same serotypes (Koekemoer et al., 2000).

To substantiate whether the cloned VP2 genes were full-length cDNA copies of the original dsRNA with complete open reading frames, baculovirus recombinants were constructed and the terminal ends of the cloned VP2 cDNA of the different serotypes were sequenced. Baculovirus recombinants expressing the VP2 genes of the six serotypes AHSV-1, -2, -4, -6, -7 and -8, for which expression has not previously been reported, showed that they all synthesized a protein of approximately 122 kDa (Fig. 2), which corresponds to the published size of AHSV VP2s. This confirmed that they all have intact full ORFs. Sequencing showed that the cDNA copies of all the serotypes had typical orbivirus 5'-GTT and TAC-3' terminal ends (Table 2). We sequenced the complete VP2 genes of the five new AHSV serotypes that we cloned but only the terminal ends of the four serotypes of our VP2 gene clones for which the full VP2 gene sequences have already been published. Sequencing of approximately 500 bp of both terminal ends of full-length VP2 genes of serotypes 3, 4, 6 and 9 showed that the terminal sequences were identical to the published sequences of these genes (results not shown). The additional C nucleotide at the 3' end of the AHSV-6 VP2 gene (Williams et al., 1998) was, however, not present in our AHSV-6 VP2 gene clone.

Terminal-end sequence analysis of all nine AHSV serotypes also revealed that the terminal-end sequences differed substantially among serotypes, especially at the 3' end (Table 2). The nucleotide sequences of the 5' end, however, did not show as much variance, as was shown previously (Williams et al., 1998). Sequencing of the full-length VP2 genes of AHSV-1, -2, -5, -7 and -8 revealed that all nine VP2 genes had approximately the same size, varying from 3205 to 3229 bp, and that the genes comprised open reading frames encoding proteins with molecular masses in the region of 122 kDa (Table 2). BLAST (NCBI) analysis of each gene sequence confirmed the identity of each gene with that of the cognate AHSV VP2 genes that were cloned previously, namely serotypes 3, 4, 6 and 9 (GenBank accession nos U01832, M94680, AF021235 and AF043926, respectively). This sequencing data, together with the serotype-specific hybridization data (Koekemoer et al., 2000), confirmed the identity of our set of cloned VP2 cDNAs.

**Multiple sequence alignment and phylogenetic analysis**

The multiple alignment of the deduced VP2 amino acid sequences of all nine AHSV serotypes, the five newly cloned and sequenced VP2 genes of serotypes 1, 2, 5, 7, 8 and the previously cloned and sequenced serotypes 3, 4, 6 and 9 showed that the homology between the nine AHSV serotypes varied from 47.6% for AHSV-2 and AHSV-9 to 71.4% for AHSV-1 and AHSV-2 (Table 3). The alignment also revealed areas within the VP2 proteins that had low identity between serotypes (Fig. 3, regions highlighted in
**Table 2.** Characteristics of dsRNA segment 2 and VP2 of the reference strain of each of the nine AHSV serotypes

<table>
<thead>
<tr>
<th>AHSV serotype</th>
<th>Segment length (bp)</th>
<th>Open reading frame (nt)</th>
<th>Size of protein (aa)</th>
<th>Predicted protein molecular mass (Da)</th>
<th>Terminal-end sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHSV1</td>
<td>3218</td>
<td>13–3183</td>
<td>1056</td>
<td>122 966</td>
<td>5’-GUUUAUUCGCauagggu...UCAUGUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV2</td>
<td>3221</td>
<td>13–3186</td>
<td>1057</td>
<td>122 962</td>
<td>5’-GUUUAUUCGCauagggu...UCANUGUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV3*</td>
<td>3221</td>
<td>13–3186</td>
<td>1057</td>
<td>122 905</td>
<td>5’-GUUAAAUUCGCauagggu...CUCAUGUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV4*</td>
<td>3229</td>
<td>13–3195</td>
<td>1060</td>
<td>123 898</td>
<td>5’-GUUAUAUUCACauagggu...CUCAUGUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV5</td>
<td>3217</td>
<td>13–3186</td>
<td>1057</td>
<td>122 731</td>
<td>5’-GUUAAAUUCACauagggu...CUCAAGAUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV6*</td>
<td>3203</td>
<td>13–3168</td>
<td>1051</td>
<td>122 169</td>
<td>5’-GUUAUAUUCACauagggu...UCGAUGUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV7</td>
<td>3222</td>
<td>13–3186</td>
<td>1057</td>
<td>123 497</td>
<td>5’-GUUAUAUUCACauagggu...UCAAAGUGAUCGCUGUAC3’</td>
</tr>
<tr>
<td>AHSV8</td>
<td>3220</td>
<td>13–3189</td>
<td>1058</td>
<td>123 077</td>
<td>5’-GUUAUAUUCACauagggu...CUCAUGUGUAACCUACUAC3’</td>
</tr>
<tr>
<td>AHSV9*</td>
<td>3205</td>
<td>13–3174</td>
<td>1053</td>
<td>123 414</td>
<td>5’-GUUAUAUUCACauagggu...CUCAUGUGUAACCUACUAC3’</td>
</tr>
</tbody>
</table>

*AHSV serotypes that have been cloned and sequenced previously.

black). Within two of these regions, there was very low identity between the amino acid sequences of the different serotypes. Furthermore, we established that the amino acid sequences of serotypes that show serological cross-neutralization also shared regions of high identity that are not present in other serotypes. Some of these regions are present within the regions with low identity. Phylogenetic and homology trees compiled from multiple alignments of the VP2 amino acid sequences clearly showed that the VP2 sequences of the serotypes that show serological cross-reaction have higher identity with each other and group together (Fig. 4). Phylogenetic analysis including the VP2 amino acid sequences of three other orbiviruses, bluetongue virus (BTV), epizootic haemorrhagic disease virus (EHDV) and Chuzan virus, showed that, based on the VP2 sequences, AHSV and Chuzan virus were more closely related than AHSV, BTV and EHDV. Hydrophobic and hydrophilic analysis of the protein sequences encoded by the nine VP2 genes indicated that the proteins had very similar hydrophobicity profiles, even in regions where there was very little amino acid sequence identity. Fig. 5 shows the hydrophobicity profiles of the 220–450 amino acid regions of all nine AHSV VP2s, which have little amino acid identity between serotypes and where antigenic sites have been identified.

**DISCUSSION**

In this paper we have described the cloning of full-length VP2 genes of the reference strain of each of the nine AHSV serotypes. Baculovirus recombinants expressing the cloned VP2 genes of six serotypes, namely serotypes 1, 2, 4, 6, 7 and 8, confirmed that they all had full open reading frames. The cloned VP2 genes of serotypes 1, 2, 5, 7 and 8 were sequenced and their amino acid sequences were deduced. The data presented here completes the cloning, sequencing and expression of the first representative of each of the nine AHSV outer capsid VP2 genes. Our sequencing data, together with that of published data for the VP2 genes of serotypes 3, 4, 6 and 9 (Iwata et al., 1992; Sakamoto et al., 1994; Vreede et al., 1994; Williams et al., 1998; Venter et al., 2000), has allowed us to perform the first complete sequence analysis of all the (sero)types for a species of the *Orbivirus* genus.

Phylogenetic analysis of the nine AHSV VP2s grouped together VP2s of serotypes that show serological cross-reaction. Phylogenetic analysis, which included the VP2 amino acid sequences of some other orbiviruses (BTV, EHDV and Chuzan virus), showed very low homology between AHSV VP2 amino acid sequences and the VP2 sequences of these orbiviruses. Chuzan virus was, however, more closely related to AHSV than BTV and EHDV based on the VP2 amino acid sequences (Fig. 4B). Low identity between serotypes was demonstrated for specific regions within the VP2 amino acid sequences that have been shown to be antigenic and play a role in virus neutralization.
The multiple alignment of the VP2 amino acid sequences of all nine AHSV serotypes showed that the homology between the different serotypes varied from 47\% to 71\%. The sequence data proves conclusively that VP2 of AHSV is the most variable protein among serotypes. The low homology between the nucleic acid sequences (results not shown) complements published hybridization data showing that partial and full-length VP2 gene probes hybridized in a serotype-specific manner to dsRNA from its cognate serotype.

![CLUSTAL multiple sequence alignment](http://vir.sgmjournals.org)

<table>
<thead>
<tr>
<th>AHSVv</th>
<th>Sequence Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AHSVv9</td>
<td>QYPFGLATYPFPHXATHFFKREHELPHXVTDYIGKRIXKNREIFLEGG-VRDAEMELTV</td>
<td></td>
</tr>
<tr>
<td>AHSVv6</td>
<td>RKVSKAKAYLYNEN -- TMCVRHDPVSLQYKLMREEMGL-YNPYMTLNGT - TESSK</td>
<td></td>
</tr>
<tr>
<td>AHSVv5</td>
<td>GIVRKEAMMLQMMGQRRLTLGVSKQVLSIDVRKQLYMKLL</td>
<td></td>
</tr>
<tr>
<td>AHSVv8</td>
<td>AHEKVRSPYCVFQVRLREASQKLRKRTYKRESEKREERKRA</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 3.** For legend see page 1323.
Fig. 3. For legend see page 1323.
serotype (Bremer et al., 1990; Koekemoer et al., 2000). The NS3 protein of AHSV has recently been shown to be the second most variable protein, with variation of between 1.8 and 36.3% across serotypes (Van Niekerk et al., 2001).

The multiple alignment of the nine AHSV VP2 amino acid sequences also showed regions of low identity between VP2 amino acid sequences. It is noteworthy that it is within some of these regions that antigenic regions have been identified on VP2 of AHSV serotypes 3, 4 and 9. Bentley et al. (2000) identified various antigenic regions using recombinant AHSV-3 VP2 phage display libraries with various antisera to AHSV. Antigenic regions were also found using truncated baculovirus-expressed AHSV-9 VP2 proteins and immunoblotting with antiserum to AHSV-9 (Venter et al., 2000).

Pepscan analysis of AHSV-4 VP2 peptides identified antigenic regions of peptides from certain regions that induce neutralizing antibodies in rabbits (Martinez-Torrecuadrada et al., 2001). Interestingly, most of these antigenic and neutralizing sites were found on the various

Fig. 3. Multiple alignment of the deduced amino acid sequences of the VP2 proteins of all nine reference serotypes of AHSV. Identical amino acids are indicated with an asterisk and similar amino acids are indicated with a dot. Antigenic regions found by panning with horse anti-AHSV-3 IgG (Bentley et al., 2000) are highlighted by black blocks. The amino acid sequences of these regions within AHSV-3 are in bold. The neutralizing sites described by Martinez-Torrecuadrada et al. (2001) are underlined and in bold. The antigenic site described by Venter et al. (2000) on AHSV-9 VP2 is underlined.

http://vir.sgmjournals.org 1323

Orbivirus outer capsid protein sequence data
VP2s between aa 252 and 488. This region not only shows significant low identity between the nine AHSV serotypes but is also mostly hydrophilic, suggesting that these sites could be located on the surface of the virion. Furthermore, within these antigenic regions there is higher identity between the VP2 amino acid sequences of serotypes that
show serological cross-reaction. In general, there is also more homology between the VP2s of the serotypes that show serological cross-reaction (Fig. 4A). This may explain the serological cross-reaction between serotypes, since the VP2 protein determines serotype. However, it should be noted that the abovementioned studies were performed with linear peptides and not with full-length VP2 in its natural conformation. Bentley et al. (2000) also found a non-continuous epitope on AHSV-3 VP2 using a random peptide library. Since BTV VP2 contains a serotype-specific antigenic region at approximately the same amino acid residues, namely aa 328–335 (Gould & Eaton, 1990) and aa 327–402 (Demaula et al., 1993), it seems possible that antigenic determinants in different orbivirus VP2s might be located in approximately the same region. The use of recombinant peptides from these regions from each serotype for diagnostic applications could, therefore, be an informative area for further investigation.

The importance of the availability of a complete set of full-length AHSV VP2 cDNA clones for each of the nine serotypes for recombinant vaccine development is underscored by the fact that thus far full protection against disease has only been achieved with full-length, soluble baculovirus-expressed AHSV VP2, notably for serotypes 4 and 5 (Roy et al., 1996; Martinez-Torrecuadrada et al., 1996; Du Plessis et al., 1998; Scanlen et al., 2002). We are now developing recombinant VP2-based vaccines for all nine serotypes of AHSV using our set of full-length cDNA copies of the VP2 genes described in this paper.

The value of the diagnostic and epidemiological applications of this first full set of AHSV VP2 clones and sequence data includes the possibility of speeding up and extending procedures for serotyping and topotyping of isolates, serum samples and midge collections. It comprises the development of molecular methods and reagents for serotyping, such as serotype-specific probes and RT-PCR procedures, as well as generating phylogenetic data sets for molecular epidemiology. In fact, the VP2 gene set generated in this study has already enabled us to demonstrate proof of concept for the development of serotype-specific probes (Koekemoer et al., 2000). Using an incomplete VP2 sequence data set, Sailleau et al. (2000) described the development of a serotype-specific RT-PCR for AHSV. Their method was based on small regions of sequence within the VP2 nucleotide sequences and required eight separate PCR reactions to be performed to determine the serotype of one sample. We envisage that multiple alignment of the full-length nucleic acid sequences of all nine AHSV VP2 genes will enable us to develop a set of primers to amplify a specific region within the VP2 gene of all nine serotypes of AHSV by RT-PCR in a one-step single reaction and generate sequence data sets for the VP2 genes of field strains for phylogenetic and topotyping analyses.

The data described in this paper are the first for a full set of VP2 genes of any orbivirus. The completion of expression, sequencing and phylogenetic analysis of this set of AHSV VP2 genes sets the scene for the development of complete repertoires of new vaccines, the identification and characterization of antigenic regions and the development of molecular diagnostic and epidemiological tools to improve the prevention, control, diagnosis and surveillance of AHS.

REFERENCES


