A comparison of the genes which encode non-structural protein NS3 of different orbiviruses

Vida van Staden and H. Huismans*

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

The segment 10 (S10) genes of African horsesickness virus (AHSV), Palyam virus and epizootic haemorrhagic disease virus were translated in vitro in a rabbit reticulocyte lysate system. Each of the S10 genes encoded two proteins, NS3 and NS3A, which were shown to be related by peptide mapping. Cloned copies of the S10 genes of two AHSV serotypes (AHSV-3 and AHSV-9) and Palyam virus were sequenced and compared to each other and to the nucleotide sequence of bluetongue virus (BTV) gene S10. Two in-phase ATG translation initiation codons reported for the S10 genes of BTV-10 and BTV-1 were conserved in the S10 genes of AHSV-3, AHSV-9 and Palyam virus, and would be able to initiate synthesis of NS3 and NS3A respectively. Comparison of the amino acid sequences of NS3 of AHSV-3 and AHSV-9 identified two areas of approximately 45 amino acids which displayed high (98%) similarity. One of these areas corresponded to the only region which displayed more than 50% amino acid similarity between NS3 of BTV, AHSV and Palyam virus. This region could represent an important structural or catalytic site of the protein. The overall amino acid similarity outside this conserved region was between 13% and 29%.

Introduction

The orbivirus genus in the Reoviridae family includes a number of viruses of economic importance, such as bluetongue virus (BTV), African horsesickness virus (AHSV), epizootic haemorrhagic disease virus (EHDV) and Palyam virus. The genome of BTV, the orbivirus prototype, is composed of 10 dsRNA segments enclosed in a double-layered protein capsid (Verwoerd, 1969; Verwoerd et al., 1972); each genome segment encodes at least one viral polypeptide. The outer capsid layer consists of two major proteins (VP2 and VP5), whereas the core particle consists of two major (VP3 and VP7) and three minor proteins (VP1, VP4 and VP6). In addition to the seven structural proteins, at least three non-structural proteins (NS1, NS2 and NS3) are also synthesized in BTV-infected cells (Huismans, 1979; Van Dijk & Huismans, 1988). In vitro translation studies have shown that segment 10 (S10), the smallest genome segment of BTV-10 and BTV-1, encodes proteins NS3 and NS3A (previously referred to as P8 and P8a, respectively) (Mertens et al., 1984; Van Dijk & Huismans, 1988). Sequencing of S10 of BTV-10 (Lee & Roy, 1986) and BTV-1 (Gould, 1988) revealed two in-phase initiation codons at positions 20 to 22 and 59 to 61 for both serotypes, which could initiate translation of the sequences encoding NS3 and NS3A respectively. Subsequent N terminus amino acid sequencing has confirmed that NS3 is translated from the first initiation codon (French et al., 1989). After cloning of a cDNA copy of BTV-10 S10 into a baculovirus expression vector and infection of Spodoptera frugiperda cells with the recombinant virus, proteins NS3 and NS3A were synthesized. In contrast to the results of in vitro translation of BTV-10 S10 RNA, in which the two proteins were synthesized in equimolar amounts, NS3 was the principal product in the baculovirus expression system (French et al., 1989).

We have carried out in vitro translation studies and
peptide mapping to determine whether the S10 genes of other orbiviruses also encode two related proteins. Furthermore, we have cloned and sequenced the S10 genes of AHSV and Palyam virus and compared the nucleotide and derived amino acid sequences to those of BTV.

Methods

Cells and virus. All viruses were propagated in BHK cells grown as monolayers in modified Eagle's medium containing 5% bovine serum (Verwoerd, 1969). The origin of AHSV has been described by Bremer et al. (1990), that of EHDV (Alberta) by Nel & Huismans (1990) and that of BTV-10 by Van Dijk & Huismans (1988). Palyam virus, a member of the Palyam serogroup, was obtained from the National Institute for Virology, Johannesburg, South Africa.

Purification of viral dsRNA and isolation of individual dsRNA segments. Double-stranded RNA was purified from infected cells by phenol-SDS extraction (Huismans & Bremer, 1981). The 10 dsRNA segments were separated on 6% preparative polyacrylamide gels using the buffer system of Loening (1967). Electrophoresis was performed at 60 V for 24 h, after which the gel was stained with ethidium bromide. The individual segments were excised from the gel, homogenized in TE buffer (10 mM-Tris-Cl, 1 mM-EDTA, pH 8) using an Ultra Turrax homogenizer and the dsRNA was diffused out by incubating overnight at 37°C with gentle shaking; acrylamide particles were removed by centrifugation. After phenol-extracting the supernatant to remove traces of acrylamide, the dsRNA was ethanol-precipitated. When required, the segments were re-run on a second gel to remove contaminating dsRNA segments and purified as above. Double-stranded RNA was resuspended in water and stored at -20°C.

In vitro translation. Preparations containing 0.3 to 0.5 µg S10 dsRNA were lyophilized, resuspended in 10 mM-methylmercuric hydroxide and incubated at room temperature for 15 min for denaturation. The translation reactions were carried out using a commercial rabbit reticulocyte lysate kit (Amersham) according to the manufacturer's instructions. The reactions contained 1.5 µCi/µl [35S]methionine (Amersham; >800 Ci/mmol, 15 mCi/ml) and were incubated at 30°C for 90 min.

PAGE. Protein samples were treated with an equal volume of 2 × protein solvent buffer (0.125 M-Tris-Cl pH 6.8, 4% SDS, 20% glycerol, 10% 2-mercaptoethanol), heated to 95°C for 3 min and stored at -20°C. SDS-PAGE was carried out as described by Laemmli (1970) using 5% stacking gels and 15% resolving gels. After electrophoresis the gels were fixed, dried and autoradiographed. For the peptide mapping the gels were fluorographed according to the method of Bonner & Laskey (1974). Gels were exposed to Cronex MRF31 X-ray film at -70°C for appropriate lengths of time.

Peptide mapping. The procedure used was based on that of Cleveland et al. (1977), as described by Van Dijk & Huismans (1988). [35S]Methionine-labelled proteins were located in dried gels by means of autoradiography and then excised. The gel fragments were rehydrated in stacking gel buffer, the filter paper backing was removed and the fragments were inserted in the wells of an SDS-polyacrylamide gel. The samples were overlaid with 20% glycerol in stacking gel buffer, followed by 10 µl 10% glycerol in stacking gel buffer containing 5 µg of Staphylococcus aureus V8 protease. Electrophoresis was carried out at 150 V until the tracking dye had moved 3 cm down the 4 cm long stacking gel. Then the current was switched off for 30 min to allow digestion to occur. After this period, electrophoresis was continued until the tracking dye had reached the bottom of the resolving gel. The peptides were visualized by means of fluorography.

DNA cloning. Double-stranded RNA of AHSV-9 and Palyam virus was cloned using a modification of the method of Cashdollar et al. (1984), as described by Huismans & Cloete (1987). The dsRNA was denatured with 10 mM-methylmercuric hydroxide and polyadenylated; cdNA was synthesized using oligo(dT) primers and fractionated on alkaline sucrose gradients. Following dc-tailing, the cdNA was cloned into dg-tailed, EcoR1-cut pBR322 (BRL). Recombinant plasmids were isolated (Birnboim & Doly, 1979) and analysed by restriction enzyme digestion and agarose gel electrophoresis. Recombinant plasmids were labelled with 32P in a nick translation reaction and the cloned fragment was identified by hybridization to Northern blots of electrophoretically separated dsRNA, in the case of AHSV-9 dsRNA, or slot-blot s of individually purified dsRNA segments, in the case of Palyam virus.

DNA sequencing. All nucleotide sequences were determined by the dideoxynucleotide sequencing method of Sanger et al. (1977), as modified for plasmid DNA (Kraft et al., 1988). Subclones were constructed by digestion of the DNA clones using the appropriate restriction enzyme and ligation of the fragments into pUC vectors. For the clone representing S10 of AHSV-3 an ordered set of deletions was also created through unidirectional shortening of the insert by exonuclease III digestion (Henikoff, 1984).

Results

In vitro translation

To identify the proteins encoded by S10 of BTV-10, AHSV-4, Palyam virus and EHDV, the respective dsRNA segments were purified, denatured and translated in an in vitro reticulocyte lysate system as described in Methods. Two proteins, NS3 and NS3A, were synthesized from the S10 RNA of each of the four different viruses (Fig. 1). The NS3 and NS3A proteins of BTV, Palyam virus and EHDV were of approximately equal Mr, whereas those of AHSV-4 were smaller. The S10 dsRNA of AHSV-3 and AHSV-9 was also translated in vitro and both serotypes synthesized two proteins which were of identical Mr, to AHSV-4 NS3 and NS3A (results not shown).

S. aureus V8 protease peptide mapping was carried out to determine whether the NS3 and NS3A proteins in Fig. 1 are related (Fig. 2). For each of the three viruses (AHSV-4, Palyam virus and EHDV), nearly identical peptide maps of NS3 and NS3A were obtained, suggesting that in each case NS3 and NS3A could be translated from the same in-phase, overlapping ORF in the respective S10 RNAs.

Cloning and nucleotide sequencing

Bremer et al. (1990) reported that the S10 of different AHSV serotypes is not as highly conserved as that of various BTV serotypes (Huismans et al., 1987; Kowalik & Li, 1987). In order to characterize these differences,
cloned copies of S10 from two AHSV serotypes were sequenced and compared to those of Palyam virus and BTV-10. The S10 gene of AHSV-3 has been cloned and described by Bremer et al. (1990), the S10 gene of AHSV-9 was provided by B. Greyling and a 600 bp fragment of the Palyam virus S10 gene was cloned as described in Methods. The sequences of the BTV-10 and BTV-1 S10 genes have been published (Lee & Roy, 1986; Gould, 1988).

The nucleotide sequences of the cloned S10 genes of AHSV-3, AHSV-9 and Palyam virus are given in Fig. 3 with the deduced amino acid sequences of the largest ORFs indicated. The S10 gene of AHSV-3 was a full-length clone, as indicated by the presence of the 5' and 3'-terminal consensus sequences. The AHSV-9 and Palyam virus S10 clones were not full-length as they contained the 5' consensus sequence but not the 3' consensus terminus.

AHSV-3 S10 was found to be 758 bp in length, with 5' and 3' non-coding regions of 19 and 88 bp respectively. It contained a single ORF with the first translation initiation codon at nucleotides 20 to 22, and two additional in-phase initiation codons at nucleotides 50 to 52 and 53 to 55. The ORF extended for 217 codons to a TAG termination codon at nucleotides 671 to 673.

The AHSV-9 S10 clone was found to be 737 bp long and alignment with the AHSV-3 S10 sequence showed that it lacked approximately 20 bp in the non-coding region at the 3'-terminal end; apart from being 1 bp shorter, the 5' non-coding region was identical to that of AHSV-3. AHSV-9 S10 contained two in-phase initiation codons and a termination codon at positions corresponding to those of AHSV-3 S10.

For Palyam virus the nucleotide sequence of the first 590 bp of S10 was obtained. The 5' non-coding region was 18 bp in length. Two initiation codons were also present in the same ORF at positions 19 to 21 and 52 to 54. Some conserved features of the 5'-terminal regions of S10 of BTV, AHSV and Palyam virus are summarized in Table 1.

**Amino acid sequences of the NS3 proteins of AHSV-3, AHSV-9 and Palyam virus**

The amino acid sequences of the non-structural proteins NS3 and NS3A encoded by S10 of AHSV-3, AHSV-9 and Palyam virus could be deduced from the corresponding nucleotide sequences. The NS3 protein of AHSV-3 and AHSV-9 was composed of 217 amino acids. Except for the first 50 amino acids the hydrophilicity profiles of NS3 of AHSV-3 and BTV-10 were very similar (not shown). The two cysteine residues of the AHSV-3 and AHSV-9 proteins were conserved and were located in the two main hydrophobic regions, also present in BTV NS3. Although the AHSV NS3 protein is 12 amino acids shorter than that of BTV, they have similar charges at neutral pH. Characteristic features of BTV-10 NS3, such as a high relative abundance of serine and threonine residues and a relative paucity of arginine and trypto-
Fig. 3. Comparison of the nucleotide sequences of cDNA copies of the NS3 genes of AHSV-3, AHSV-9 and Palyam virus, and their predicted translation products. The complete AHSV-3 sequence is given, whereas the nucleotide and amino acid sequences of AHSV-9 and Palyam virus are only given where differences were detected. Numbers above the sequence indicate nucleotide positions for AHSV-3 and those below amino acid number. The termination codons are indicated by #. An asterisk marks positions where spaces were inserted to give optimal sequence alignments; ] marks the end of the AHSV-9 and Palyam virus clones; conserved terminal sequences are underlined and the inverted repeats marked by _.

The different NS3 proteins were compared by alignment of the sequences and by matrix comparisons (not shown). Alignment of AHSV-3 and AHSV-9 NS3 indicated 76% similarity (identical amino acids), which increased to 85% when equivalent amino acids were also taken into account. Alignment of the NS3 amino acid sequences of BTV, AHSV and Palyam virus gave a low overall amino acid similarity of 25% to 30%. By using the method of matrix comparisons, only one area was identified which displayed at least 50% similarity in a comparison of NS3 of any of the different orbiviruses. In

...phän residues in comparison to other BTV-10 proteins (Roy, 1989), were shared by AHSV NS3. The two NS3 proteins have very similar acidic, basic, aromatic and hydrophobic amino acid compositions. NS3A of both AHSV-3 and AHSV-9 would be 11 amino acids shorter than NS3, assuming that it initiated at the second in-frame ATG codon at positions 53 to 55. For Palyam virus the predicted amino acid sequence of NS3 probably represented about 80% of the total protein. In this case the NS3 and NS3A proteins also differed in length by 11 amino acids.
**Table 1. Conserved features of the 5'-terminal regions of S10 of BTV-10, AHSV-3, AHSV-9 and Palyam virus**

<table>
<thead>
<tr>
<th>Virus</th>
<th>Non-coding sequence length (bp)</th>
<th>Terminal sequence</th>
<th>First ATG codon</th>
<th>Second ATG codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTV-10</td>
<td>19</td>
<td>GTTAAA</td>
<td>20-22</td>
<td>GCATGGA</td>
</tr>
<tr>
<td>AHSV-3</td>
<td>19</td>
<td>GTTTAA</td>
<td>20-22</td>
<td>GTATGGA</td>
</tr>
<tr>
<td>AHSV-9</td>
<td>18</td>
<td>GTTTA</td>
<td>19-21</td>
<td>GTATGGA</td>
</tr>
<tr>
<td>Palyam virus</td>
<td>18</td>
<td>GTTTAA</td>
<td>19-21</td>
<td>GACATGT</td>
</tr>
</tbody>
</table>

* Sequences representing putative initiation codons are underlined and nucleotides at positions −3 and +4 relative to the ATG are in bold type.

† There is another in-phase ATG codon present at nucleotides 50 to 52; it is not possible to predict which one would be used to initiate synthesis of NS3A.

---

**Discussion**

No function has as yet been ascribed to the S10 genome segment of orbiviruses. For both BTV-10 and BTV-1, *in vitro* translation studies and sequence analyses of S10 have revealed two in-phase ORFs (Lee & Roy, 1986; Mertens *et al.*, 1984; Van Dijk & Huismans, 1988; Gould, 1988). In order to gain more insight into the possible functions of S10, we analysed S10 of AHSV and Palyam virus and looked for conserved features. Individually purified S10 dsRNA of BTV-10, AHSV, Palyam virus and EHDV was prepared and translated *in vitro* using a rabbit reticulocyte lysate system. In each case two proteins, presumed to be equivalent to BTV NS3 and NS3A, were synthesized *in vitro*. The *M*<sub>s</sub> of NS3 and NS3A of Palyam virus and EHDV were estimated to be approximately 26K and 25K respectively, which correlates well with the sizes of the corresponding BTV proteins; the AHSV proteins were estimated to be slightly smaller at 24K and 23K. The result showing that EHDV S10 dsRNA is translated into two proteins contrasts with the results of Mecham & Dean (1988), who identified only one protein product. This might be due to differences in electrophoretic conditions, as we were able to separate the two proteins on a 15% but not on a 12% polyacrylamide gel. Through longer exposure of the autoradiograms, one or two smaller translation products could be identified for AHSV and Palyam virus. It is not known whether these represent incomplete translation products, or degradation of translated proteins.

The relatedness of NS3 and NS3A was determined by peptide mapping; nearly identical maps were obtained in the case of the proteins from AHSV-4, Palyam virus and EHDV. This is in accordance with results for BTV-10 (Van Dijk & Huismans, 1988), implying that two proteins are translated from the same in-phase overlapping reading frame.

The complete nucleotide sequence of a cDNA copy of the S10 gene of AHSV-3, and partial sequences of the S10 genes of AHSV-9 and Palyam virus have been determined. The AHSV-3 gene was 758 bp long, with a coding
region of 651 bp and 5' and 3' non-coding regions of 19 and 88 bp, respectively. The AHSV-9 5' non-coding region was 18 bp long and, except for being shorter by one bp, was identical to that of AHSV-3. The corresponding region in Palyam virus also had a length of 18 bp. The 5'-terminal sequences of S10 from AHSV-3 and AHSV-9 were both 5' GTTAAA, differing from the consensus sequence of other orbiviruses by the substitution of T for A at position 4. 5'-Terminal sequences of other cloned AHSV genes (unpublished results) have been shown to contain A, T or C at position 4 of the conserved terminal sequence. The Palyam virus 5'-terminal S10 sequence of 5' GTTAAA is identical to that of BTV.

The 88 bp non-coding region of AHSV-3 S10 was 28 bp shorter than that of BTV-10 S10. However, the 3'-terminal sequences (ACTTAC) are identical and AHSV-3 S10 also contained two (G + C)-rich sequences (nucleotides 688 to 712 and 737 to 758), as is the case for S10 of BTV-1 (Gould, 1988). This provides some support for the suggestion proposed by Gould et al. (1988) that the conserved (G + C)-rich region at the 3' terminus of S10 RNA would make RNA transcription more difficult, resulting in the relative paucity of NS3 and its mRNA in virus-infected cells.

S10 of AHSV-3, AHSV-9 and Palyam virus contains two in-phase ATG translation initiation codons, which can initiate synthesis of NS3 and NS3A respectively. Neither the first nor the second ATG codon in any of these viruses, or those of BTV, conformed fully to the Kozak consensus sequence (Kozak, 1981, 1984, 1986), which demands a purine at position −3 and a guanosine at position +4; all these potential initiation codons had a purine (A) at position +4 of the second ATG codon, and AHSV-3 and AHSV-9 had a purine (A) at position +4 of the first ATG codon. It was not possible, therefore, to distinguish between the different initiation codons with respect to their efficiency in initiating protein synthesis.

Two 7 bp inverted repeats were present in AHSV-3 S10 adjacent to the terminal sequences (Fig. 3). Such short inverted repeats can be shown in published BTV sequences (Roy, 1989) and have also been reported for EHDV (Nel et al., 1991) and other dsRNA viruses (Stoeckle, 1987; Anzola et al., 1987).

The nucleotide sequences of S10 of AHSV-3, AHSV-9 and Palyam virus were compared to each other and that of BTV-10 (Lee & Roy, 1986) using the Microgenie sequence analysis program (Queen & Korn, 1984). The coding regions of the S10 genes of AHSV-3 and AHSV-9 showed 74% sequence identity, confirming the relatively low degree of similarity of S10 within the AHSV serogroups as determined by hybridization (Bremer et al., 1990). The nucleotide similarity between the S10 genes of the different viruses was, as expected, even lower, ranging between 45% and 49% in any possible comparison of the different S10 genes. However, one area of about 100 bp, varying slightly in position between the different viruses but falling between nucleotides 150 and 280, showed a significantly higher degree of similarity (61% to 67%). S10 of BTV-10 and BTV-1 are 84% and 93% similar at nucleotide and amino acid levels respectively (Gould, 1988); if the BTV-1 S10 sequences were substituted for those of BTV-10 in the comparisons to the other orbivirus S10 genes, essentially similar results were obtained.

The NS3 amino acid sequence similarity of the two AHSV serotypes was 76%; which included two highly conserved regions (amino acids 47 to 90 and 153 to 201) (Fig. 4a) with only one amino acid substitution in a length of 44 and 49 amino acids respectively. The corresponding conserved nucleic acid regions, nucleotides 158 to 289 and 476 to 622, showed different patterns of base substitutions. Between nucleotides 158 and 289 (132 bp) there were only 13 substitutions (9-8%), of which 12 (92%) were in the third codon position. Between nucleotides 476 and 622 (147 bp) there were 36 substitutions (24.5%), of which 33 (92%) were in the third codon position. In the rest of the coding region (372 bp) there were 118 base substitutions (31.7%), of which only 47% were in the third codon position. The difference in the base substitutions in these two conserved regions could not be explained by differences in the degeneracy of the codons that encoded the amino acids in the two regions. One possible explanation could be that there is a selective pressure for the conservation of specific nucleic acid sequences.

The first region of 90% nucleotide sequence similarity between the S10 genes of the AHSV serotypes corresponded to the area which showed more than 60% nucleotide sequence similarity between S10 of AHSV, BTV and Palyam virus. Matrix comparisons of NS3 of these viruses identified a homologous region between amino acids 40 and 100 in all three viruses (Fig. 4b), which was encoded by this region of high nucleotide similarity; this region could represent an important structural or catalytic site of the protein. When substitution of similar and not only identical amino acids was taken into account, other homologous regions (amino acids 115 to 135 and 170 to 190) could be identified. The substitution of similar amino acids could be important in maintaining the overall structure of the protein. The importance of maintaining such similarity in structure is supported by the fact that the hydrophilicity profiles as well as the distribution of cysteine residues in BTV and AHSV NS3 are very similar.

The function of NS3 is still unknown. However, considering the degree of conservation of certain features
of the S10 gene and the NS3 protein among different orbiviruses, it is possible that these features play an important role in viral replication or morphogenesis.

We would like to thank Dr Albie van Dijk of the Veterinary Research Institute, Onderstepoort, South Africa for the use of tissue culture facilities and for interest and criticism. Dr Tony Whistler of the National Institute for Virology, Johannesburg, South Africa provided the Palyam virus. Financial support was received from the Foundation for Research Development.

References


(Rceived 22 November 1990; Accepted 23 January 1991)