Journal of General Virology (1992), 73, 3023–3026. *Printed in Great Britain*

Sequence of genome segment 9 of bluetongue virus (serotype 1, South Africa) and expression analysis demonstrating that different forms of VP6 are derived from initiation of protein synthesis at two distinct sites

Alison M. Wade-Evans,* Peter P. C. Mertens and Graham J. Belsham

AFRC, Institute for Animal Health, Pirbright Laboratory, Ash Road, Pirbright, Woking, Surrey GU24 0NF, U.K.

Bluetongue virus (BTV) VP6 is often resolved into two closely migrating bands by SDS–PAGE (VP6 and VP6a). RNA segment 9 of BTV-serotype 1 South Africa (encoding VP6) has been cloned as cDNA, and the complete sequence has been determined. Expression of this clone both *in vitro* and in tissue culture produced the same polypeptide doublet as seen previously in extracts from BTV-infected cells. Modification of the cDNA, including the removal of the first initiation codon, demonstrated that the two forms of VP6 are derived from initiation of protein synthesis at two distinct sites and not by post-translational modification.

Bluetongue is an economically important disease of sheep which is characterized by facial oedema, coronitis and death. Bluetongue virus (BTV), a member of the Orbivirus genus within the Reoviridae family, also infects several other species of ruminants, including cattle, although the clinical symptoms are less severe and sometimes absent. The virus is transmitted between mammalian hosts by an obligate vector belonging to the *Culicoides* genus.

The virus particle is icosahedral and consists of a segmented dsRNA genome encapsidated in a double-layer protein coat. The 10 dsRNA genome segments each encode at least one virus protein (Mertens *et al.*, 1984). Seven of these are generally referred to as structural proteins. Two major (VP3 and VP7) and three minor proteins (VP1, VP4 and VP6/VP6a) form the core structure, and VP2 and VP5 form the diffuse outer capsid layer of the virion (Verwoerd *et al.*, 1972; Mertens *et al.*, 1987). Three non-structural proteins, NS1, NS2 and NS3/NS3a, have also been identified in BTV-infected cells (Mertens *et al.*, 1984; Van Dijk & Huismans, 1988). NS2 is the only virus protein that is phosphorylated (Huismans *et al.*, 1987; Devaney *et al.*, 1988) and has also been reported to be associated with the outer capsid layer of purified virus particles (Mertens *et al.*, 1987). VP5 has recently been shown to be glycosylated (Yang & Li, 1992). There is no evidence that any of the other structural proteins are post-translationally modified. However, VP6 from purified virus, or produced by *in vitro* translation of segment 9 (S9), is often resolved into two bands (VP6 and VP6a) on SDS–polyacrylamide gels (Huismans & Bremer, 1981; Grubman *et al.*, 1983; Mertens *et al.*, 1984, 1987). The nature of this doublet has not been defined, but it has been suggested that one species may be a post-translationally modified form of the other (Huismans & Bremer, 1981). The function of VP6 has not been clearly identified, but preliminary studies have demonstrated nucleic acid-binding activity, which may indicate an association with viral RNA (Huismans *et al.*, 1987; Roy *et al.*, 1990).

The complete sequence of BTV-10 S9, which encodes VP6, has been published (Fukusho *et al.*, 1989). However, BTV-10 is the only documented serotype in which VP6 is not resolved into a doublet (Huismans & Bremer, 1981). The sequence reported in this paper is of a full-length cDNA clone of S9 from BTV serotype 1 South Africa (BTV-1SA), which produces the two species of VP6. The cDNA was generated as described previously (Wade-Evans *et al.*, 1988) using oligonucleotide primers complementary to the 3′ terminus of each strand of S9 and cloned into Bluescript (pKS; Stratagene). The primer sequences 5′ GTTAAAAAA-TCGCATATGTC 3′ and 5′ GTAAGTGTGAAATCA-

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank Nucleotide Sequence Databases under accession number D10905.

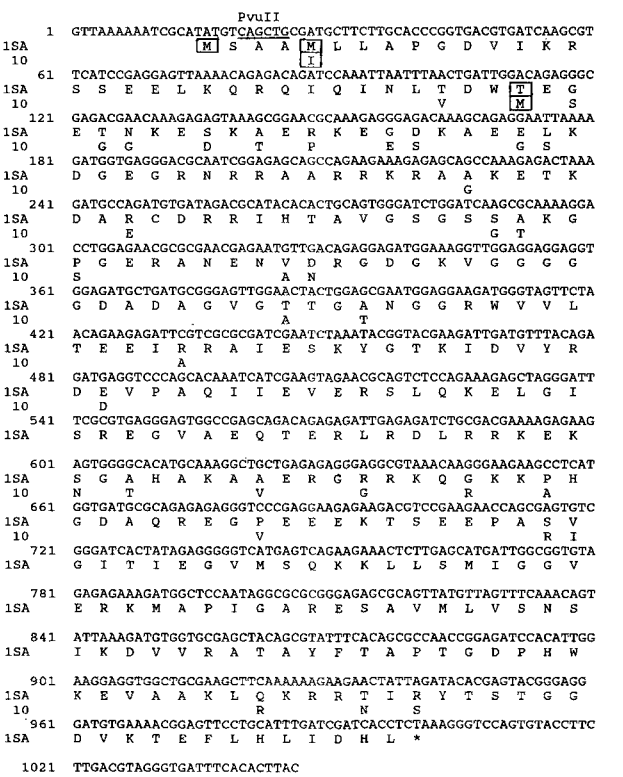

Fig. 1. Sequence of BTV-1SA S9, encoding VP6. The predicted amino acid sequence of VP6 is indicated below each line of nucleotide sequence. Any amino acid change observed between BTV-1SA and BTV-10 (Fukusho *et al.*, 1989) is shown beneath the BTV-1SA VP6 sequence. The *Pvu*II site is underlined and the amino acid positions mentioned in the text are boxed.

CCCTA 3′ were selected by reference to published data on direct RNA sequencing of S9 (Mertens & Sangar, 1985).

The sequence of the clone, read at least three times from each DNA strand, was determined by direct plasmid sequencing as described by Chen & Seeburg (1985) and compiled into a complete sequence using the BESTFIT program available from the University of Wisconsin Genetics Computer Group package (Devereux *et al.*, 1984).

Sequence analysis of the cDNA clones obtained from BTV-1SA S9 identified one clone of 1046 bp containing the sequence of both the oligonucleotide primers, one at each terminus, indicating that it was a full-length copy of this segment (Fig. 1). Only one long open reading frame (ORF) was identified within the complete S9 sequence, initiating at nucleotides 16 to 18 and terminating at nucleotides 1000 to 1002. This ORF encodes a polypeptide of 328 amino acids (aa) with a calculated M_r of 36·08K, although the M_r of BTV-1SA VP6 estimated by SDS–PAGE is approximately 48K (Mertens *et al.*, 1984).

The sequence of BTV-1SA S9 was compared to that of BTV-10 (Fukusho *et al.*, 1989). The identity between the nucleotide sequences was 86% and was 89% between the amino acid sequences (see Fig. 1).

The full-length cDNA was transferred to pET3a (Rosenberg *et al.*, 1987) by *Nde*I and *Bam*HI digestion so that transcription was under the control of the T7 promoter. The resulting plasmid is referred to as pET3a+9. pET3a+9 was linearized using *Bam*HI prior to transcription. RNA transcripts were generated using T7 RNA polymerase, and these were then DNase-treated and phenol/chloroform-extracted according to the protocol supplied by the manufacturer (Promega). The transcripts were then translated into [^{35}S]methionine-labelled proteins *in vitro* using a rabbit reticulocyte lysate system (Promega). Viral dsRNA, prepared from purified BTV-1SA particles obtained from infected BHK21 cells, was denatured as described by Mertens *et al.* (1984) and then translated in the *in vitro* system described above.

Transcription and translation of the full-length cDNA clone of BTV-1SA S9 *in vitro* using the reticulocyte lysate system produced two forms of VP6, referred to as VP6 and VP6a (Fig. 2, lane 2), which co-migrate with two products observed after translation of all 10 denatured of BTV-1SA dsRNA segments (Fig. 2, lane 3). Similar results were obtained using a transient expression system in mammalian cells (Fig. 2, lanes 5 and 6). The latter studies were performed as described previously (Belsham & Brangwyn, 1990) using lipofectin (BRL)-mediated transfection of plasmid DNA (5 μg) into BSC40 cells (35 mm dishes) infected with the recombinant vaccinia virus vTF7-3, which expresses the T7 RNA polymerase (Fuerst *et al.*, 1986). The cells were incubated with [^{35}S]methionine (50 μCi/dish) for 2 h at 20 h after transfection, and cell extracts were prepared and the VP6 proteins immunoprecipitated with an anti-BTV-1SA serum (Mertens *et al.*, 1987). Samples were analysed by SDS–PAGE (Laemmli, 1970) using 10% polyacrylamide gels, and autoradiography. The transient expression studies were performed in parallel with *in vitro* transcription/translation, using the same plasmids, and showed that the VP6 doublet was not generated as an artefact of *in vitro* translation. The results are consistent with the observation that VP6 is also present as a doublet in the BTV-1SA particle (Mertens *et al.*, 1987).

It has previously been reported that BTV-10 does not produce two forms of VP6. Analysis of the S9 sequence data revealed that a methionine at aa position 5 in BTV-1SA is replaced by an isoleucine in BTV-10 (Fig. 1). Therefore it seemed possible that the two forms of VP6 represented the products of initiation at separate sites. This hypothesis was tested by deleting the sequence containing the first start codon (nucleotides 16 to 18) so

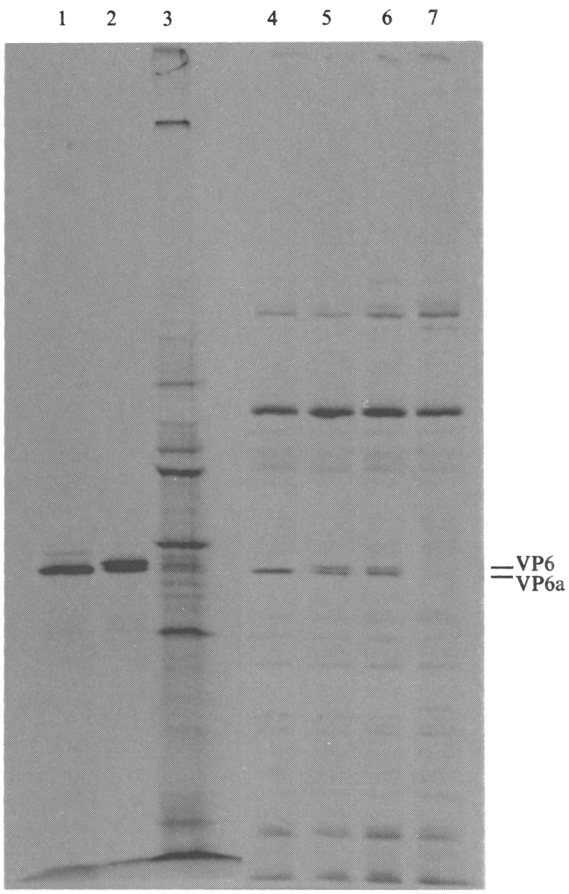

Fig. 2. *In vitro* and *in vivo* expression of VP6 polypeptides. cDNA clones of BTV-1SA S9 were expressed both *in vitro* (lanes 1 to 3) and transiently *in vivo* (lanes 4 to 7) as described. The cell extracts run in lanes 4 to 7 have also been immunoprecipitated with anti-BTV-1SA serum, whereas lanes 1 to 3 represent total translation mixes. Lanes 2, 5 and 6 contain VP6 polypeptides encoded by the full-length cDNA clone of S9, pET3a+9. Lanes 1 and 4 contain the proteins encoded by the truncated version of S9, pET3a+9P. Lane 3 contains the proteins encoded by all 10 genome segments of BTV-1SA. Lane 7 is a negative control in which no additional plasmid DNA was transfected into vaccinia virus recombinant vTF7-3-infected cells.

that only a truncated form of VP6, initiating at the second AUG codon (nucleotides 28 to 30), could be synthesized. The deletion of the first 20 nucleotides, including the first initiation codon, was achieved by transferring a *Pvu*II–*Bam*HI fragment from pKS+9 (containing the majority of S9, nucleotides 23 to 1046) into pET3a, to produce the plasmid pET3a+9P. In this construct, transcription of the inserted cDNA sequence is also under the control of the T7 promoter. However, the AUG at nucleotides 28 to 30 is the first initiation codon in the RNA transcript (see Fig. 1). The removal of the start codon at nucleotides 16 to 18 prevented translation of the largest polypeptide of the VP6/VP6a doublet, resulting in synthesis of the smaller (VP6a) product only in both *in vitro* and *in vivo* assay systems (Fig. 2, lanes 1 and 4). The only ORF within the S9 sequence large enough to encode VP6a is that which encodes the larger VP6 protein, but initiating at the methionine codon at nucleotides 28 to 30. These results indicate that the two proteins, VP6 and VP6a, produced by transcription/translation of the full-length S9 cDNA are not modified forms of each other, but are translated from the same ORF, initiating separately at the AUG codons at nucleotides 16 to 18 and 28 to 30 respectively.

The requirements for two forms of VP6, one a truncated version of the other, are unknown. Since BTV-10 does not appear to produce multiple forms of VP6, it is interesting to note that there is a methionine at aa position 33 in BTV-10 which is replaced by a threonine in BTV-1SA. It is possible that this is the second initiation codon in this serotype, but that the resulting protein (32 aa smaller) may comigrate with VP7, and therefore is difficult to identify.

In eukaryotic cells, protein synthesis is usually initiated by ribosome binding in the vicinity of the 5′ cap structure on the mRNA, followed by scanning and initiation of translation at the first AUG codon (Kozak, 1989). The sequence context of the first initiation codon in both BTV-1SA and BTV-10 S9, CATATGT, is rather poor, with pyrimidines at both positions −3 and +4. This allows leaky scanning, enabling some ribosomes to reach and initiate at a second AUG codon. In BTV-1SA, the second start site (nucleotides 28 to 30) is in a more favourable consensus sequence, GCGATGC, with a purine at position −3.

Kozak (1986) has reported that 17 viral mRNAs produce two (and, in rare cases, even three) overlapping proteins by initiating at both a weak upstream and the next downstream AUG codon. Our observation that BTV-1SA S9 encodes two overlapping proteins produced by translation of the same ORF initiating at two distinct AUG codons (VP6 and VP6a) therefore is not unique, but fits into a well established pattern. It also closely resembles the situation observed in another member of the Reoviridae, simian rotavirus SA11, S9 of which encodes two species of VP7 by using two in-phase initiation codons (Chan *et al.*, 1986).

As is the case with the rotavirus VP7 doublet (Chan *et al.*, 1986), both VP6 and VP6a are packaged within the BTV particle (Mertens *et al.*, 1984), indicating either a requirement for both forms or a recognition of both polypeptides during particle assembly owing to the similarity of the sequences. Recently published data on the genome coding assignment of two other members of the orbivirus genus, African horsesickness virus (Grubman & Lewis, 1992) and epizootic haemorrhagic disease virus (Mecham & Dean, 1988), indicate that a

protein doublet is encoded by S9 of these two isolates. A protein doublet is also apparent after SDS–PAGE of the *in vitro* translation products from AHSV-9 S9 (R. O'Hara & P. P. C. Mertens, unpublished results). The expression of two VP6 proteins in a range of orbiviruses would seem to indicate some functional importance. Future studies using bacterially expressed forms of the VP6 protein may provide more information in this area.

We would like to thank Julia Brangwyn for her expert technical assistance and Dr Andrew King for useful discussion in preparation of this manuscript. This work was funded in part by the Ministry of Agriculture, Fisheries and Food.

References

BELSHAM, G. J. & BRANGWYN, J. K. (1990). A region of the 5′ noncoding region of foot-and-mouth disease virus RNA directs efficient internal initiation of protein synthesis within cells: involvement with the role of L protease in translational control. *Journal of Virology* **64**, 5389–5395.

CHAN, W.-T., PENARANDA, M. E., CRAWFORD, S. E. & ESTES, M. K. (1986). Two glycoproteins are produced from the rotavirus neutralization gene. *Virology* **151**, 243–252.

CHEN, E. Y. & SEEBURG, P. H. (1985). Laboratory methods. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA* **4**, 165–170.

DEVANEY, M. A., KENDALL, J. & GRUBMAN, M. J. (1988). Characterisation of a nonstructural phosphoprotein of two orbiviruses. *Virus Research* **11**, 151–164.

DEVEREUX, J., HAEBERLI, P. & SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387–395.

FUERST, T. R., NILES, E. G., STUDIER, F. W. & MOSS, B. (1986). Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 8122–8126.

FUKUSHO, A., YU, Y., YAMAGUCHI, S. & ROY, P. (1989). Completion of the sequence of bluetongue virus serotype 10 by the characterization of a structural protein, VP6, and a non-structural protein, NS2. *Journal of General Virology* **70**, 1677–1689.

GRUBMAN, M. J. & LEWIS, S. A. (1992). Identification and characterization of the structural and nonstructural proteins of African horse sickness virus and determination of the genome coding assignments. *Virology* **186**, 444–451.

GRUBMAN, M. J., APPLETON, J. A. & LETCHWORTH, G. J. (1983). Identification of BTV type 17 genome segments coding for polypeptides associated with virus neutralization and intergroup reactivity. *Virology* **131**, 355–366.

HUISMANS, H. & BREMER, C. W. (1981). Identification of the serotype-specific and group specific antigens of bluetongue virus. *Onderstepoort Journal of Veterinary Research* **48**, 51–58.

HUISMANS, H., VAN DIJK, A. A. & BAUSKIN, A. R. (1987). *In vitro* phosphorylation and purification of a non-structural protein of bluetongue virus with affinity for single-stranded RNA. *Journal of Virology* **61**, 3589–3595.

KOZAK, M. (1986). Bifunctional messenger RNAs in eukaryotes. *Cell* **47**, 481–483.

KOZAK, M. (1989). The scanning model for translation: an update. *Journal of Cell Biology* **108**, 229–241.

LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.

MECHAM, J. O. & DEAN, V. C. (1988). Protein coding assignment for the genome of epizootic haemorrhagic disease virus. *Journal of General Virology* **69**, 1255–1262.

MERTENS, P. P. C. & SANGAR, D. V. (1985). Analysis of the terminal sequences of the genome segments of four orbiviruses. *Virology* **140**, 55–67.

MERTENS, P. P. C., BROWN, F. & SANGAR, D. V. (1984). Assignment of the genome segments of bluetongue virus type 1 to the proteins which they encode. *Virology* **135**, 207–217.

MERTENS, P. P. C., BURROUGHS, J. N. & ANDERSON, J. (1987). Purification and properties of virus particles, infectious subviral particles and cores of bluetongue virus serotypes 1 and 4. *Virology* **157**, 375–386.

ROSENBERG, A. H., LADE, B. N., CHUI, D.-S., LIN, S.-W., DUNN, J. J. & STUDIER, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125–135.

ROY, P., ADACHI, A., URAKAWA, T., BOOTH, T. F. & THOMAS, C. P. (1990). Identification of BTV VP6 protein as a nucleic acid-binding protein and the localization of VP6 in virus-infected vertebrate cells. *Journal of Virology* **64**, 1–8.

VAN DIJK, A. A. & HUISMANS, H. (1988). *In vitro* transcription and translation of bluetongue virus mRNA. *Journal of General Virology* **69**, 573–581.

VERWOERD, D. W., ELS, H. J., DE VILLIERS, E. M. & HUISMANS, H. (1972). Structure of the bluetongue virus capsid. *Journal of Virology* **10**, 783–794.

WADE-EVANS, A. M., PAN, Z. Q. & MERTENS, P. P. C. (1988). Sequence analysis and *in vitro* expression of a cDNA clone of genome segment 5 from bluetongue virus, serotype 1 from South Africa. *Virus Research* **11**, 227–240.

YANG, Y.-Y. & LI, J. K.-K. (1992). Complete genomic sequences of the GP5 protein gene of bluetongue virus serotype 11 and 17. *Virus Research* **23**, 163–171.

(Received 8 June 1992; Accepted 27 July 1992)