Journal of General Virology (2001), **82**, 2681–2690. Printed in Great Britain

Conservation of coding potential and terminal sequences in four different isolates of Borna disease virus

Stephan Pleschka,[1] Peter Staeheli,[2] Jolanta Kolodziejek,[3] Jürgen A. Richt,[1] Norbert Nowotny[3,4] and Martin Schwemmle[2]

[1] Institute of Virology, University of Giessen, D-35392 Giessen, Germany

[2] Department of Virology, Institute for Medical Microbiology and Hygiene, University of Freiburg, Hermann-Herder-Strasse 11, D-79104 Freiburg, Germany

[3] Institute of Virology, University of Veterinary Sciences Vienna, A-1210 Vienna, Austria

[4] Department of Medical Microbiology, Faculty of Medicine and Health Sciences, United Arab Emirates University, Al Ain, United Arab Emirates

We determined the complete nucleotide sequences of two poorly characterized strains of Borna disease virus (BDV) and compared them to reference strains V and He/80. Strain H1766 was almost 98% and 95% identical to strains V and He/80, respectively, whereas strain No/98 was only about 81% identical to both reference strains. In contrast to earlier reports, we found an additional A residue at the extreme 3′-end of the single-stranded RNA genome in all four BDV strains. The exact numbers of nucleotides in the four BDV genomes could not be determined due to a micro-heterogeneity at the 5′-end. If our longest sequence is a correct copy of the viral RNA, the two ends of the BDV genome would show almost perfect complementarity. All three transcription start sites, all four termination sites, both splice donor sites and both major splice acceptor sites are highly conserved, whereas a minor alternative splice acceptor site is not. The L protein of No/98 differs at 7% of its amino acid positions from the polymerase in the other strains, with most differences mapping to the C-terminal moiety of the molecule. Re-evaluation of L protein sequences of strains V and He/80 revealed differences at several positions compared to published information, indicating that variant forms of the viral polymerase have previously been characterized. These results are important because correct structures of genome ends and of the polymerase gene are the most critical parameters for the future development of techniques that will permit the genetic manipulation of BDV.

Introduction

Borna disease virus (BDV) is a non-segmented negative-strand RNA virus that persistently infects the central nervous system of a broad variety of animals and, possibly, humans (Staeheli *et al.*, 2000). Natural and experimental BDV infections can result in severe immune-mediated neurological disease (Richt & Rott, 2001; Rott & Becht, 1995; Stitz *et al.*, 1993). Unlike the related animal rhabdo- and paramyxoviruses, BDV transcribes and replicates its approximately 8900 nt long genome in the nucleus of infected cells and uses the splicing machinery to regulate the expression of viral proteins (Briese *et al.*, 1992; Cubitt *et al.*, 1994*b*; Jehle *et al.*, 2000; Schneider *et al.*, 1994*b*). Mainly because of these features, BDV has recently been classified as the prototype of a new virus family, *Bornaviridae*, within the order *Mononegavirales*. At least six viral proteins are expressed in BDV-infected cells, namely nucleoprotein (N), X-protein (X or p10), phosphoprotein (P), putative matrix protein (M), glycoprotein (G) and polymerase (L). Synthesis of a shorter form of P, designated P′, is initiated at the second in-frame AUG of the P ORF (Kobayashi *et al.*, 2000). Splicing of two introns regulates the balanced expression of M, G and L (Cubitt *et al.*, 1994*b*; Jehle *et al.*, 2000; Schneider *et al.*, 1994*b*, 1997). Whereas intron 1 is located within the M ORF, intron 2 is found in the ORFs of G

Author for correspondence: Martin Schwemmle.
Fax +49 761 203 6634. e-mail schwemm@ukl.uni-freiburg.de

0001-7884

and L (Cubitt *et al.*, 1994*b*; Schneider *et al.*, 1994*b*). Within the coding region of the L protein, the presence of an alternative intron 2 splice acceptor site was recently described (Cubitt *et al.*, 2001; Tomonaga *et al.*, 2000). Alternative splicing of intron 2 gives rise to two new ORFs with coding capacities for 8·9 kDa and 165 kDa proteins, respectively. Whether these two proteins are indeed produced in BDV-infected cells and whether they serve vital functions remain to be determined. To date, it is not known whether the genomes of all BDV strains have the capacity to encode these additional proteins.

Little is known about the conservation of genome size, coding potential and terminal sequences in different isolates of BDV. Until now, the complete genetic information of only two closely related BDV strains, named He/80 and V, has been available (Briese *et al.*, 1994; Cubitt *et al.*, 1994*a*). The two genomes show 95% sequence identity, and they have identical organization of transcription units and ORFs. However, the L protein of He/80 was reported to lack 24 amino acids at the C terminus, and the last 3 nt at the 5′-ends differ completely. Since genome ends are of critical importance for efficient replication and transcription of non-segmented negative-strand RNA viruses (Conzelmann, 1998), it remained possible that the unique end sequences may point to strain-specific differences in the 5′-promoter region.

To better differentiate between conserved and strain-specific features, we determined the complete genome sequences of two additional BDV strains. No/98 originates from a horse that acquired Borna disease outside the well-known endemic region of BDV in central Europe (Nowotny *et al.*, 2000). Partial nucleotide sequence analysis indicated that No/98 differs from reference strains He/80 and V by about 15%. Strain H1766 (also known as MDCK-BDV) is frequently used for experiments in Japanese laboratories (Nakamura *et al.*, 1999; Shoya *et al.*, 1997). We show here that the two reference strains and the newly characterized strains harbour a 3′-terminal A residue and most likely a 5′-terminal G residue. We further show that the alternative intron 2 splice acceptor site is not conserved in strain No/98. Re-analysis of the L ORF revealed that the polymerase of strain He/80 has the same length as its counterparts in strains H1766, V and No/98.

Methods

BDV strains and cells. Vero and OL cells were cultured in Dulbecco's modified Eagle's medium containing 10% foetal calf serum (FCS). OL cells were kindly provided by Georg Pauli (Robert Koch Institute, Berlin, Germany). Persistent BDV infections were established by infecting 10^5 cells with 10^4 focus-forming units of appropriate virus stock followed by continued passage for at least 5 weeks. Complete infection of the cultures was confirmed by indirect immunofluorescence as described (Formella *et al.*, 2000).

Virus stock preparation and titration. Virus stocks were prepared from OL cells persistently infected with either BDV strain He/80 (Cubitt *et al.*, 1994*a*), strain H1766 (Shoya *et al.*, 1997) or strain V (Briese *et al.*, 1994), and from Vero cells persistently infected with strain No/98 (Nowotny *et al.*, 2000) essentially as described (Briese *et al.*, 1992). Briefly, 25 confluent 90 mm dishes were washed with 20 mM HEPES (pH 7·4) and incubated with 10 ml of 20 mM HEPES (pH 7·4) containing 250 mM $MgCl_2$ and 1% FCS for 1·5 h at 37 °C. Subsequently, supernatants were harvested and centrifuged twice at 2500 ***g*** for 5 min to remove cell debris. Virus particles were concentrated by ultra-centrifugation for 1 h at 20 °C at 80000 ***g*** onto a 20% sucrose cushion containing 20 mM HEPES (pH 7·4) and 1% FCS. Virus-containing pellets were either re-suspended in PBS to approximately 10^7 focus-forming units/ml or were directly used to prepare RNA.

Preparation of viral RNA. Viral RNA was isolated with TRIzol (Gibco/BRL) according to the manufacturer's instructions. Briefly, virus particles in 1 ml of a concentrated BDV stock (about 10^7 focus-forming units) were precipitated for 30 min at 4 °C in a TL120 centrifuge (Beckman) at 70000 r.p.m. The virus pellet was lysed by dissolving in 1 ml of TRIzol solution and incubating for 5 min at room temperature. The RNA was then extracted with chloroform, and precipitated from the aqueous phase with isopropanol and glycogen (Boehringer Mannheim/Roche). Finally, the RNA was washed with ethanol, dried and re-suspended in water.

Determination of 3′- and 5′-terminal sequences of BDV genomic RNA by RNA ligation. RNA recovered from partially purified virus particles (1 μg) was self-ligated in a volume of 20 μl of ligation buffer (75 mM Tris–HCl, pH 7·5; 0·1 mM ATP; 10 mM $MgCl_2$; 5 mM dithiothreitol; 10%, v/v, dimethyl sulfoxide; 50 U RNasin) with 50 U of T4 RNA ligase (NEB). After incubation at 25 °C for 90 min, the volume was adjusted to 200 μl with TE buffer (10 mM Tris–HCl, pH 8·0 and 1 mM EDTA). The RNA was then extracted once with phenol and chloroform, precipitated with isopropanol/glycogen and centrifuged for 15 min at 4 °C. Finally, the RNA was washed with ethanol, dried and re-suspended in 13 μl of water. The material was subjected to reverse transcription using Superscript II (GIBCO/BRL) as recommended by the manufacturer using 2·5 pmol of primer GSP1 (5′ ATTATAGTTTTG-TCATGGACCTC 3′). Primer GSP1/No98 (5′ ATGGCTTCTTGATG-GACTTGGTC 3′) rather than GSP1 was used for RNA from BDV strain No/98. The reaction was stopped by incubation at 70 °C for 10 min. Then RNase H (GIBCO/BRL) was added and the sample was incubated at 37 °C for another 60 min. Samples (5 μl) were amplified using Taq polymerase (Boehringer Mannheim/Roche) and 0·4 pmol of primer GSP1 (for H1766, V and He/80) or GSP1/No98 (for No/98) and primer GSP3 (5′ CGTGACTGGTCTAACAATGC 3′; for H1766, V, No/98 and He/80). Nested PCR was performed with 2 μl samples of PCR products using primer GSP2 (5′ GCTTGTGGTAGGACAGCACATC 3′; for H1766, V and He/80) or GSP2/No98 (5′ GTTGGTGGTAG-GGCAGTACATC 3′; for No/98) and primer GSP4 (5′ GAGCTTA-GGGAGGCTCGCTG 3′; for H1766, V and He/80) or GSP4/No98 (5′ GAAAGCTTGGGAAGGCTTGCTG 3′; for No/98). The final PCR products were cloned into vector pCR4 TOPO (Invitrogen) and sequenced using a MegaBASE 1000 sequencer (Amersham Pharmacia Biotech) and the PHRED base-calling algorithm (STATEN Software package). Sequences were analysed using the DNASTAR (Lasergene) software package.

3′-RACE analysis. To determine the 3′-end of the BDV genome, 1 μg of RNA from partially purified BDV particles was tailed with A or C residues using poly(A) polymerase (Pharmacia) as recommended by the manufacturer. The A-tailing reaction was stopped after 15 min, whereas the C-tailing reaction was stopped after 1 h. RNA was purified by phenol and chloroform extraction, precipitated with isopropanol/glycogen, washed, dried and re-suspended in 11 μl of water. A-tailed RNA was reverse-transcribed with Superscript II (GIBCO/BRL) and oligo(dT) adapter primer from the 3′-RACE kit (GIBCO/BRL). C-tailed RNA was reverse-transcribed using the inosine/guanosine adapter primer provided with the 5′-RACE kit (GIBCO/BRL). PCR was performed with 5 μl of cDNA samples, adapter primer [oligo(dT) adapter primer from the 3′-RACE system for A-tailed RNA and inosine/guanosine adapter primer from the 5′-RACE system for C-tailed RNA] and primer GSP3 (for strain H1766, No/98, V and He/80). For semi-nested PCR, 2 μl samples of the first PCR were amplified a second time with the respective adapter primers described above and primer GSP4 (for strain H1766, V and He/80) or GSP4/No98 (for No/98). PCR products were cloned into vector pCR4 TOPO before sequencing.

■ **Determination of the 5′-terminal sequences of BDV genomic RNA by 5′-RACE.** A standard 5′-RACE system (GIBCO/BRL) was used according to the manufacturer's instructions. Briefly, reverse transcription was done with primer GSP1 (for strain H1766, V and He/80) or GSP1/No98 (for No/98) and Superscript II. RNA was then digested with RNase H, and the cDNA was purified with a Glass MAX kit (GIBCO/BRL). C-tails were added to the 3′-terminal end of the cDNAs with terminal deoxynucleotidyl transferase. The tailed cDNA was amplified by PCR using appropriate adapter primers and GSP2 (for H1766, V and He/80) or GSP2/No98 (for No/98). PCR products were cloned into vector pCR4 TOPO before sequencing. Determination of the 5′-terminal sequences of BDV genomic RNA of He/80 by ligation of RNA oligos and subsequent RT–PCR was carried out with GeneRacer (Invitrogen) and primer GSP2 according to the manufacturer's instructions. Tobacco acid pyrophosphatase (TAP) treatment was carried out as recommended by the manufacturer (Invitrogen). Amplification products were either cloned into vector pCR4 TOPO before sequencing or sequenced directly.

■ **Cloning and sequencing of near-full-length cDNA of strain H1766.** PCR-amplified, near-full-length cDNA derived from BDV strain H1766 that had been grown in persistently infected MDCK cells (Shoya *et al.*, 1997) was kindly provided by K. Tomonaga. The material was cloned into vector pCR-XL-TOPO (Invitrogen) according to the manufacturer's instructions. Three independent plasmids were sequenced which (according to restriction analysis) seemed to contain the complete viral genome. Sequencing primers were designed from published BDV sequences (GenBank accession nos U04608 and L27077). Regions were chosen that showed no or only little variation. In order to identify and eliminate PCR-generated mutations, overlapping sequences of each plasmid were assembled, resulting in three complete contiguous sequences which were then used to generate a single consensus sequence. Sequences at the extreme ends of the viral RNA were determined by RNA ligation, 3′-RACE and 5′-RACE using RNA isolated from partially purified particles.

■ **Sequencing of strain No/98.** The complete cDNA sequence of No/98 was determined by direct sequencing of overlapping RT–PCR amplification products using a large number of different oligonucleotides as described previously (Nowotny *et al.*, 2000). Briefly, RT–PCR reactions were carried out in a single-step reaction using the Titan One Tube RT–PCR kit, the Expand Reverse Transcriptase kit and the Expand Long Template PCR system (all Boehringer Mannheim/Roche). As template we used RNA from hippocampus and rhinencephalon of the diseased horse from which No/98 was originally isolated. PCR products were sequenced in both directions without subcloning using an automated sequencer (ABI Prisma 310 Genetic Analyser, Perkin Elmer). Sequences near the 3′- and 5′-ends were determined as described above using RNA from virus particles released from an infected Vero cell culture.

■ **Cloning of a cDNA containing the L ORF of strain He/80$_{FR}$.** Fragments of the L ORF were generated by RT–PCR using RNA from C6 cells persistently infected with the Freiburg variant of strain He/80 (He/80$_{FR}$). Reverse transcription was carried out with Superscript II (GIBCO/BRL) and strain-specific primers were derived from the published He/80 sequence (GenBank accession no. L27077). PCR was performed with Taq polymerase (Boehringer Mannheim/Roche) as described by the manufacturer. An RT–PCR fragment corresponding to nt 2393–4229 of He/80$_{FR}$ (GenBank accession no. AJ311522) was cloned into the *Eco*RI and *Pst*I sites of vector pBlue (Stratagene) followed by insertion of a second RT–PCR fragment (nt 4229–4869) into the unique *Pst*I and *Bam*HI sites. Subsequently, a third RT–PCR fragment corresponding to nt 4869–7870 was inserted between the *Bam*HI and the *Xho*I sites. In a final step, a fourth RT–PCR fragment containing nt 7870–8823 was cloned into the unique *Xba*I site, resulting in plasmid pBlue-L encoding the complete L ORF of He/80.

■ **Sequencing of the L ORF of strain V$_{FR}$.** The L ORF of the Freiburg variant of strain V (V$_{FR}$) was determined by sequencing overlapping RT–PCR fragments generated from RNA of strain V-infected Vero cells. Briefly, first-strand synthesis was carried out with Superscript II (GIBCO/BRL) as recommended by the supplier using different oligonucleotides derived from the published BDV strain V sequence (GenBank accession no. U04608). PCR was performed with Taq polymerase (Boehringer Mannheim/Roche) using a panel of suitable primers. Amplification products were gel-purified and sequenced without subcloning.

Results and Discussion

Conserved coding strategy in four independent BDV isolates

We sequenced the complete genomes of BDV isolates H1766 and No/98. Strain H1766 originates from a diseased horse in Germany. It is frequently used for experiments by Japanese laboratories. Strain No/98 represents the most distant BDV subtype known to date (Nowotny *et al.*, 2000). When compared to published sequences of standard strains V and He/80 as well as to the Freiburg variants of these strains, designated V$_{FR}$ and He/80$_{FR}$ (which differ from the parental strains at several nucleotide positions), it became clear that all four BDV strains have identical genome lengths, except for strain No/98, which lacks a single residue in the U-rich region (nt 1168–1171) that follows the N gene. Due to a microheterogeneity of the 5′-end (see below), the exact number of nucleotides in the BDV genome was difficult to determine. If our longest sequence represents a correct copy of the viral RNA, the genomes of BDV strains V, H1766 and He/80 would consist of 8912 nt and that of strain No/98 of 8911 nt. Since neither of these numbers corresponds to genome lengths that are a multiple of six, the 'rule of six' (Calain & Roux, 1993) does not seem to apply to members of the family *Bornaviridae*. Nevertheless, it is interesting to note that a 3 nt deletion in the N gene of strain No/98 is compensated for by a 3 nt insertion in the first intergenic region (Nowotny *et al.*, 2000), suggesting that unknown mechanisms are operating that maintain a constant number of nucleotides in the BDV genome.

An overall comparison of the four BDV genomes allowed us to establish the exact phylogenetic relationship of these

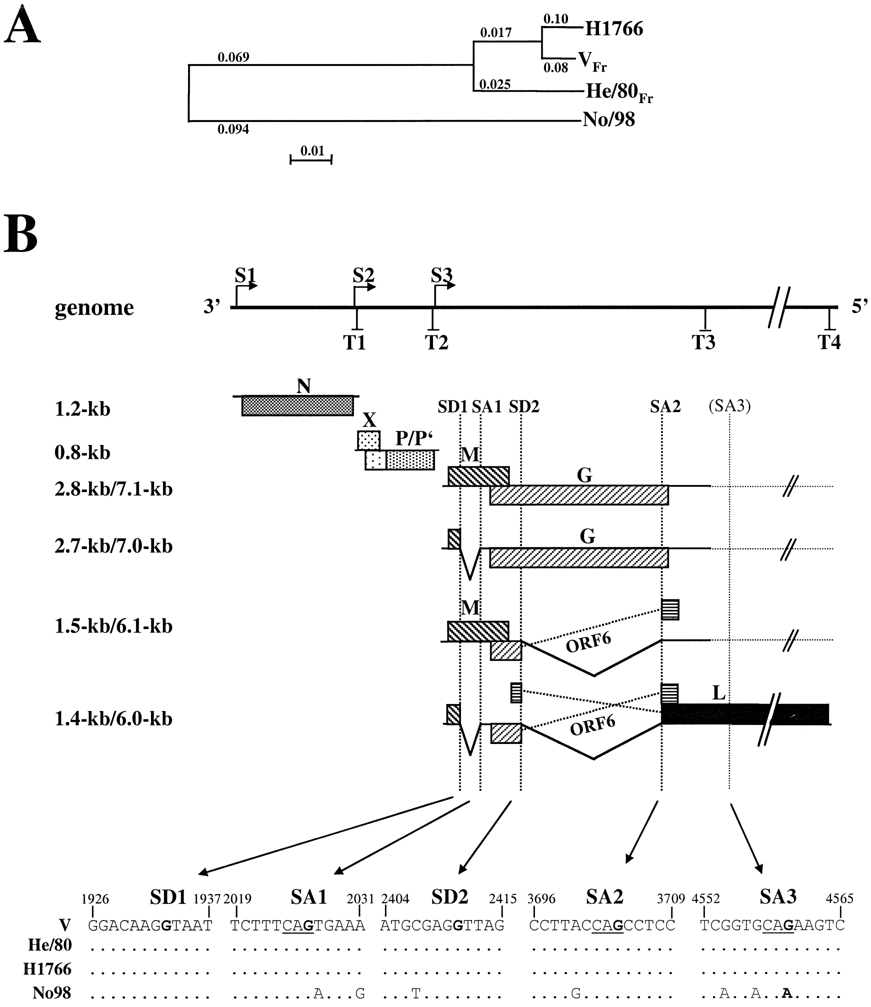

Fig. 1. Conserved coding strategy of four different BDV isolates. (A) The phylogenetic tree represents a rooted consensus tree generated by the neighbour-joining method on the genome sequences of strain No/98 (GenBank accession no. AJ311524), strain H1766 (GenBank accession no. AJ311523), strain V$_{FR}$ (GenBank accession no. AJ311521) and strain He/80$_{FR}$ (GenBank accession no. AJ311522). Bootstrap analysis was applied using 100 values. (B) Consensus transcription map. Positions of transcription start (S1–S3) and stop (T1–T4) sites are indicated. Locations and nucleotide sequences flanking splice donor (SD1, SD2) and acceptor sites (SA1–SA3) are given. ORFs are individually marked on the various viral transcripts. Letters N, X, P, P′, M, G and L indicate translation products that have been identified in infected cells by immunohistochemical methods. The putative product of ORF6 remains to be visualized. Due to the absence of a functional SA3 site in strain No/98, our consensus transcription map does not include SD2–SA3 spliced viral mRNAs and their predicted translation products.

viruses. As shown in Fig. 1 (A), strain H1766 is closely related but distinct from reference strain V. These two strains differ by only 1·8% (155 nt). Interestingly, strain V was isolated from a diseased horse in southern Germany more than 70 years ago (Schneider *et al.*, 1994*a*), whereas strain H1766 was isolated in 1994 (S. Herzog, personal communication). Our results thus confirm earlier observations that the genomes of European BDV field isolates show remarkably few changes over long periods of time. As expected from earlier comparisons of incomplete viral genomes (Nowotny *et al.*, 2000), strain No/98 occupies a unique position in the phylogenetic tree (Fig. 1 A). This may reflect the fact that it was isolated from a diseased horse that lived outside the classical endemic region in Europe (Nowotny *et al.*, 2000).

The coding strategy of BDV strains No/98 and H1766 is virtually identical to that of the two reference strains (Fig. 1 B). All sequences previously found to define the transcription start sites S1, S2 and S3 (Schneemann *et al.*, 1994) are completely conserved. Similarly, the four transcription stop sites (Briese *et al.*, 1994; Cubitt & de la Torre, 1994; Schneemann *et al.*, 1994) are also present at corresponding positions in the genomes of strains No/98 and H1766. Similarly, the sequences flanking intron 1 and intron 2 are highly conserved (Cubitt *et al.*, 1994*b*; Schneider *et al.*, 1994*b*). The only notable difference between the various strains maps to a site, designated SA3, that was recently reported to define an alternative intron 2 splice acceptor site (Cubitt *et al.*, 2001; Tomonaga *et al.*, 2000). The consensus motif CAG, which is present in strains V, He/80 and H1766, is not present in strain No/98, indicating that the latter virus cannot generate the corresponding alternatively spliced mRNA. This suggests that the predicted products of this new mRNA may not serve essential functions in the BDV replication cycle. However, we cannot exclude the remote possibility that atypical splice donor sites are used.

A consensus transcription map of the BDV genome (Fig. 1B) shows that all four viruses are able to direct the synthesis of proteins designated N, X, P, P′, M, G and L (Kobayashi *et al.*, 2000; Schwemmle *et al.*, 1999; Walker *et al.*, 2000; Wehner *et al.*, 1997). Interestingly, the coding regions for N and P in strain H1766 are completely identical to those of virus isolate BDVHuP2br, believed to have originated from a Japanese psychiatric patient (Nakamura *et al.*, 2000). This unusual congruence and the fact that strain H1766 (also named MDCK-BDV) is being used by the Japanese laboratory for experiments (Nakamura *et al.*, 1999; Shoya *et al.*, 1997) support our previous notion that BDVHuP2br might represent a laboratory contamination (Staeheli *et al.*, 2000). Intriguingly, all four viruses have one additional ORF generated by splicing of intron 2 that we designate ORF6. The putative product of ORF6 consists of the N-terminal 58 amino acids of viral protein G fused to a unique 21 residue polypeptide encoded by sequences immediately downstream of splice acceptor site 2. Since this putative protein carries the G-derived signal peptide but lacks a transmembrane anchor, it is expected to be secreted by BDV-infected cells. It thus seems to resemble sGP of Ebola virus (Sanchez *et al.*, 1996), which was suggested to influence the host defence (Yang *et al.*, 1998). It remains to be investigated whether the ORF6 product is indeed synthesized during BDV infection and whether it can modulate the antiviral response.

Primary structure of the BDV polymerase

The most striking difference between the published ORFs of the L proteins of He/80 and V is a sequence heterogeneity at the 3′-ends that would seem to result in an L protein of He/80 that lacks 24 amino acids at the C terminus (Fig. 2). Our analysis showed that the L protein of strain H1766 consists of 1711 amino acids as does the L protein of strain V (Fig. 2). Furthermore, we found that the L gene of strain No/98 encodes a protein of exactly the same length (Fig. 2), suggesting that the truncated version of L in He/80 does not reflect a true strain difference but rather errors in previous cDNA cloning and sequencing experiments. To investigate this point in more detail, we re-cloned and sequenced the complete L ORF of the Freiburg variant of strain He/80 (He/80_{FR}) by RT–PCR as described in Methods. As shown in Fig. 2, the predicted L ORF of He/80_{FR} encodes a protein of 1711 residues, like all other BDV strains. Original He/80 differs from the He/80_{FR} sequence by a deletion of 2 nt (positions 8717–8718), which results in a premature stop codon. Insertion of the two missing nucleotides into the published sequence of He/80 restores the L ORF. This new C terminus of He/80 L perfectly matches the C terminus of the He/80_{FR} L protein, indicating that the deletion in the published L ORF of He/80 most likely represents a cloning or sequencing artefact. In fact, these errors in the He/80 sequence have recently been corrected (March 2001, accession no. L27077). Thus, the various strains of BDV appear to have L proteins of identical lengths.

A more complete comparison of the published L protein sequence of strain He/80 with sequence information from variant strain He/80_{FR} revealed nine additional differences in the amino acid sequence (Fig. 2). In all of these cases, one of the two strains possesses a residue that matches the consensus sequence of all four viruses, while the other strain has a unique residue at the corresponding position. The most simple explanation for these findings is that cDNA cloning or sequencing errors occurred. It remains possible, however, that at least some of these amino acid exchanges reflect true differences between viruses with different passage history. We further observed seven amino acid differences in the L protein of strain V and the Freiburg variant (strain V_{FR}) of this virus. As discussed above for strain He/80, these differences could represent cloning or sequencing errors or else might indicate true differences between viruses with different passage history.

Most differences between the L protein of strain No/98 and the reference strains map to the last third of the protein (Fig. 2). Within this region, three blocks of amino acids of No/98 (G_{1449}LHRRA_{1454}, L_{1643}IQE_{1646} and G_{1656}RGPVVSRSSRWVG_{1664}) are particularly poorly conserved. Whether these amino acid differences have an impact on the polymerase activity remains to be shown.

The 5′- and 3′-termini of the BDV genome

To determine the 5′-end of the viral RNA, 5′-RACE experiments were performed with RNA extracted from partially purified virus particles obtained from persistently infected cells. In this procedure, viral cDNA is synthesized and 3′-ends are elongated artificially with C residues which later serve as annealing sites for oligo(dG/dI) primers that are used to amplify the adjacent region of the viral genome by PCR. The sequence of the viral RNA at the extreme 5′ terminus is then deduced by combining the sequence information from the various cloned PCR products. Irrespective of strain origin (He/80, V, H1766 and No/98) all cDNA clones analysed were virtually identical except that they differed significantly at the fusion site that links the virus sequence to the C tail (Fig. 3 A, left panel), indicating that the viral RNA used for reverse transcription was heterogeneous at the 5′-end. A consensus derived from the majority of PCR fragments suggested that 5′ GCGCUA … 3′ is the most likely sequence at the extreme 5′-

```
No/98         1 MSFHASLLREEETPRPVAGINRTDQSLKNPLLGTEVSFCLKSSSLPHHIRALGQIKPRNLASCDYYLLFRQVVLPPEVYPIGVLIRAAETILTVIVSAWKLEHMTKT
H1766           MSFHASLLREEETPRPVAGINRTDQSLKNPLLGTEVSFCLKSSSLPHHVRALGQIKARNLASCDYYLLFRQVVLPPEVYPIGVLIRAAEAILTVIVSAWKLDHMTKT
V               MSFHASLLREEETPRPVAGINRTDQSLKNPLLGTEVSFCLKSSSLPHHVRALGQIKARNLASCDYYLLFRQVVLPPEVYPIGVLIRAAEAILTVIVSAWKLDHMTKT
VFR             -----------------------------------------------------------------------------------------------------------
He/80           MSFHASLLREEETPRPVAGINRTDQTLKNPLLGTEVSFCLKSSSLPHHVRALGQIKAKNLASCDYYLLFRQVVLPPEVYPIGVLIRAAEAILTVIVSAWKLDHMTKT
He/80Fr         -------------------------S---------------------------------------------------------------------------------

No/98       108 LYSSVRYALTNPRVRAQLELHIAYQRVVGQVSYSREADIGPKRLGNMSLYFIQSLVIATIDTTSCLMTYNHFLAAADTAKSRCHLLIASVVQGALWEQGSFLDHVIN
H1766           LYSSVRYALTNPRVRAQLELHIAYQRIVGQVSYSREADIGPKRLGNMSLQFVQSLVIATIDTTSCLMTYNHFLAAADTAKSRCHLLIASVVQGALWEQGSFLDHIIN
V               LYSSVRYALTNPRVRAQLELHIAYQRIVGQVSYSREADIGPKRLGNMSLQFIQSLVIATIDTTSCLMTYNHFLAAADTAKSRCHLLIASVVQGALWEQGSFLDHIIN
VFR             -----------------------------------------------------------------------------------------------------------
He/80           LYSSVRYALTNPRVRAQLELHIAYQRIVGQVSYSREADIGPKRLGNMSLQFIQSLVIATIDTTSCLMTYNHFLAAADTAKSRCHLLIASVVQGALWEQGSFLDHIIN
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98       215 LIDIIDSINLPHDEYFTIIKSISPYSQGLVMGRYNVSVSSNFASVFTIPETCPQLDSLLKKLLNLDPVLLLMISSVQKSWYFPEIRMVDGSREQLHKMRVDSEKPQA
H1766           MIDIIDSINLPHDDYFTIIKSISPYSQGLVMGRHNVSVSSDFASVFTIPELCPQLDSLLKKLLQLDPVLLLMVSSVQKSWYFPEIRMVDGSREQLHKMRVELETPQA
V               MIDIIDSINLPHDDYFTIIKSIFPYSQGLVMGRHNVSVSSDFASVFAIPELCPQLDSLLKKLLQLDPVLLLMVSSVQKSWYFPEIRMVDGSREQLHKMRVELETPQA
VFR             -----------------------------------------------------------------------------------------------------------
He/80           MIDTIDSINLPHDDYFTIIKSISPYSQGLVMGRHNVSVSSDFASVFTIPESCPQLDSLLKKLLQLDPVLLLMVSSVQKSWYFPEIRMVDGSREQLHKMRVELETPQA
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98       322 LLSYGHTLLSMFRAEFIKGYVSKNAKWPPVHLLPGCDKSIKNARELGRWSPAFDRRWQLFAKVVILKIADLDMDPDFNDIVSDKAIISSRRDWVFEYNAAAFWKKYG
H1766           LLSYGHTLLSIFRAEFIKGYVSKNAKWPPVHLLPGCDKSIKNARELGRWSPAFDRRWQLFEKVVILRIADLDMDPDFNDIVSDKAIISSRGDWVFEYNAAAFWKKYG
V               LLSYGHTLLSIFRAEFIKGYVSKNAKWPPVHLLPGCDKSIKNARELGRWSPAFDRRWQLFEKVVILRIADLDMDPDFNDIVSDKAIISSRRDWVFEYNAAAFWKKYG
VFR             -------P---------------------------------------------------------------------------------------------------
He/80           LLSYGHTLLSIFRAEFIKGYVSKNAKWPPVHLLPGCDKSIKNARELGRWSPVFDRRWQLFAKVVILRIADLDMDPDFNDIVSDKAIISSRRDWVFEYNAAAFWKKYS
He/80Fr         ----------------------------------------------------------------------------------------------------------G

No/98       429 ERLERPVSRSGPSRLVNALIDGRLDNIPALLEPFYRGADEFEDRLTVLVPKEKELKVKGRFFSKQTLAIRIYQVVAEAALKNEVMPYLKTHSMTMSSTALTHLLNRL
H1766           ERLERPPARSGPSRLVNALIDGRLDNIPALLEPFYRGAVEFEDRLTVLVPKEKELKVKGRFFSKQTLAIRIYQVVAEAALKNEVMPYLKTHSMTMSSTALTHLLNRL
V               ERLERPPARSGPSRLVNALIDGRLDNIPALLEPFYRGAVEFEDRLTVLVPKEKELKVKGRFFSKQTLAIRIYQVVAEAALKNEVMPYLKTHSMTMSSTALTHLLNRL
VFR             -----------------------------------------------------------------------------------------------------------
He/80           ERLERPPARSGPSRLVNALIDGRLDNIPALLEPFYRGAVEFEDRLTVLVPKEKELKVKGRFFSKQTLAIRIYQVVAEAALKNEVMPYLKTHSMTMSSTALTHLLNRL
He/80Fr         ----------------------H------------------------------------------------------------------------------------

No/98       536 SHTITKGDSFVINLDYSSWCNGFRPELQAPLCRQLDQMFNCGYFFRTGCTLPCFTTFIIQDRFNPPYSDRGEPVEDGVTCAIGTKTMGEGMRQKLWTILTSCWEIIA
H1766           SHTITKGDSFVINLDYSSWCNGFRPELQAPICRQLDQMFNCGYFFRTGCTLPCFTTFIIQDRFNPPYSLSGEPVEDGVTCAVGTKTMGEGMRQKLWTILTSCWEIVA
V               SHTITKGDSFVINLDYSSWCNGFRPELQAPICRQLDQMFNCGYFFRTGCTLPCFTTFIIQDRFNPPYSLSGEPVEDGVTCAVGTKTMGEGMRQKLWTILTSCWEIVA
VFR             -----------------------------------------------------------------------------------------------------------
He/80           SHTITKGDSFVINLDYSSWCNGFRPELQAPLCRQLDQMFNCGYFFRTGCTLPCFTTFIIQDRFNPPYSFRGEPVEDGVTCAVGTKTMGEGMRQKLWTILTSCWEIVA
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98       643 LREVNVTFNILGQGDNQTIIIHRSVSQNNQQLAERALGALYKHARLAGHNLKVEECWVSDCLYEYGKKLFFRGVPVPGCLKQLSRVTDSTGELFPNLYSKLACLTSS
H1766           LREINVTFNILGQGDNQTIIIHKSASQNNQLLAERALGALYKHARLAGHNLKVEECWVSDCLYEYGKKLFFRGVPVPGCLKQLSRVTDSTGELFPNLYSKLACLTSS
V               LREINVTFNILGQGDNQTIIIHKSASQNNQLLAERALGALYKHARLAGHNLKVEECWVSDCLYEYGKKLFFRGVPVPGCLKQLSRVTDSTGELFPNLYSKLACLTSS
VFR             -----------------------------------------------------------------------------------------------------------
He/80           LREINVTFNILGQGDNQTIVIHKSASQNNQLLAERALGALYKHARLAGHNLKVEECWVSDCLYEYGKKLFFRGVPVPGCLKQLSRVTDSTGELFPNLYSKLACLTSS
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98       750 CLSAAMADTSPWVALATGVCLYLIELYVELPPAVMQDESLLTTLCLVGPSIGGLPTPATLPSVFFRGMSDPLPFQLALLQTLIKTTGVTCSLVNRVVKLRIAPYPDW
H1766           CLSAAMADTSPWVALATGVCLYLIELYVELPPAIMQDESLLTTLCLVGPSIGGLPTPATLPSVFFRGMSDPLPFQLALLQTLIKTTGVTCSLVNRVVKLRIAPYPDW
V               CLSAAMADTSPWVALATGVCLYLIELYVELPPAIMQDESLLTTLCLVGPSIGGLPTPATLPSVFFRGMSDPLPFQLALLQTLIKTTGVTCSLVNRVVKLRIAPYPDW
VFR             ----------------------------------I------------------------------------------------------------------------
He/80           CLSAAMADTSPWVALATGVCLYLIELYVELPPAIMQDESLLTTLCLVGPSIGGLPTPATLPSVFFRGMSDPLPFQLALLQTLIKTTGVTCSLVNRVVKLRIAPYPDW
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98       857 LSLVTDPTSLNIAQVYRPERQIRKWIEEAIATSTHSSRVATFFQQPLTEMAQLLARDLSTMMPLLPRDMSALFALSNVAYGLSIIDLFQKSSTVVSASQAVHIEDVA
H1766           LSLVTDPTSLNIAQVYRPERQIRRWIEEAIATSSHSSRIATFFQQPLTEMAQLLARDLSTMMPLRPRDMSALFALSNVAYGLSIIDLFQKSSTVVSASQAVHIEDVA
V               LSLVTDPTSLNIAQVYRPERQIRRWIEEAIATSSHSSRIATFFQQPLTEMAQLLARDLSTMMPLRPRDMSALFALSNVAYGLSIIDLFQKSSTVVSASQAVHIEDVA
VFR             -----------------------------------------------------------------------------------------------------------
He/80           LSLVTDPTSLNIAQVYRPERQIRRWIEEAIATSSHSSRIATFFQQALTEMAQLLARDLSTMMPLRPRDMSALFALSNVAYGLSIIDLFQKSSTVVSASQAVHIEDVA
He/80Fr         ---------------------------------------------P--------------------------------------------------------T----

No/98       964 LESVRYKESIIQRLLDTTEGYNMQPYLEGCTYLAAKQLRRLTWGRDLVGVTMPFVAEQFHPHSSVGANAELYLDAIIYCPQEPLRSHHLTTRGDQPLYLGSNTAVKV
H1766           LESVRYKESIIQGLLDTTEGYNMQPYLEGCTYLAAKQLRRLTWGRDLVGVTMPFVAEQFHPHSSAGAKAELYLDAIIYCPQETLRSHHLTTRGDQPLYLGSNTAVKV
V               LESVRYKESIIQGLLDTTEGYNMQPYLEGCTYLAAKQLRRLTWGRDLVGVTMPFVAEQFHPHSSVGAKAELYLDAIIYCPQETLRSHHLTTRGDQPLYLGSNTAVKV
VFR             -----------------------------------------------------------------------------------------------------------
He/80           LESVRYKESIIQGLLDTTEGYNMQPYLEGCTYLAAKQLRRLTWGRDLVGVTMPFVAEQFHPHSSVGAKAELYLDAIIYCPQETLRSHHLTTRGDQPLYLGSNTAVTV
He/80Fr         ---------------------------------------------------------------------------------------------------------K-

No/98      1071 QRGEITGLTKSRAANLVRDTLVLHQWYKVRRVTDPHLNTLMCRFLLEKGYTSEARPSIQGGTLTHRLPSRGDSRQGLTGYVNILSTWLRFSSDYLHSFSKSSDDYTI
H1766           QRGEITGLTKSRAANLVKDTLVLHQWYKVRKVTDPHLNTLMARFLLEKGYTSDARPSIQGGTLTHRLPSRGDSRQGLTGYVNILSTWLRFSSDYLHSFSKSSDDYTI
V               QRGEITGLTKSRAANLVRDTLVLHQWYKVRKVTDPHLNTLMARFLLEKGYTSDARPSIQGGTLTHRLPSRGDSRQGLTGYVNILSTWLRFSSDYLHSFSKSSDDYTI
VFR             --------A--------K-----------------------------------------------------------------------------------------
He/80           QRGEITGLTKSRAANLVKDTLVLHQWYKVRKVTDPHLNTLMARFLLEKGYTSDARPSIQGGTLTHRLPSRGDSRQGLTGYVNILSTWLRFSSDYLHSFSKSSDDYTI
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98      1178 HFQHVFTYGCLYADSVIRSGGVISTPYLLSASCKTCFEKIDSEEFVLACEPQYKGAEWLISKPVTIPEQIIDAEVEFDPCVSASYCLGILIGKSFLVDIRASGHDIM
H1766           HFQHVFTYGCLYADSVIRSGGVISTPYLLSASCKTCFEKIDSEEFVLACEPQYRGAEWLVSKPVSVPEQITDAEVEFDPCVSASYCLGILIGKSFLVDIRASGHDIM
V               HFQHVFTYGCLYADSVIRSGGVISTPYLLSASCKTCFEKIDSEEFVLACEPQYRGAEWLISKPVTVPEQITDAEVEFDPCVSAGYCLGILIGKSFLVDIRASGHDIM
VFR             -----------------------------------------------------------------------------------S-----------------------
He/80           HFQHVFTYGCLYADSVIRSGGVISTPYLLSASCKTCFEKIDSEEVVLACEPQYRGAEWLISKPVTVPEQIIDAEVEFDPCVSASYCLGILIGKSFLVDIRASGHDIM
He/80Fr         --------------------------------------------F--------------------------------------------------------------

No/98      1285 EQRTWANLERFSLSDMQKLPWSIVIRSLWRFLVGARLLQFERAGLIRLLYAAAGPTFSFLMRVFQDSALLMDCAPLDRLSPRINFQNRGDLVAKLILLPFINPGIVD
H1766           EQRTWANLERFSVSDMQKLPWSIVIRSLWRFLIGARLLQFEKAGLIRMLYAATGPTFSFLMKVFQDSALLMDCAPLDRLSPRINFHSRGDLVAKLVLLPFINPGIVE
V               EQRTWANLERFSVSDMQKLPWSIVIRSLWRFLIGARLLQFEKAGLIRMLYAATGPTPSFLMKVFQDSALLMDCAPLDRLSPRINFHSRGDLVAKLVLLPFINPGIVE
VFR             --------------------------------------------------------F--------------------------------------------------
He/80           EQRTWANLERFSVSDMQKLPWSIVIRSLWRFLIGARLLQFEKAGLIRMLYAATGPTFSFLMKVFQDSALLMDCAPLDRLYPRINFHSRGDLVAKLVLLPFINPGIVE
He/80Fr         -------------------------------------------------------------------------------S---------------------------

No/98      1392 IEVAGINSKYLAVSETNMDLYIAAAKSVGVKPIQFVEETNDFVARGQHHGRYSLSWSGLHRRAQALKMVVRKLKLCVLYIYPTVDPAIALDLCHLPALTIILVLGGD
H1766           IEVSGINSKYHAVSEANMDLYIAAAKSVGVKPTQFVEETNDFTARGHHHGCYSLSWSKSRNQSQVLKMVVRKLKLCVLYIYPTVDPAVALDLCHLPALTIILVLGGD
V               IEVSGINSKYHAVSEANMDLYIAAAKSVGVKPTQFVEETNDFTARGHHHGCYSLSWSKSRNQSQVLKMVVRKLKLCVLYIYPTVDPAVALDLCHLPALTIILVLGGD
VFR             -----------------------------------------------------------------------------------------------------------
He/80           IEVSRINSKYHAVSEANMDLYIAAAKSVGVKPTQFVEETNDFTARGHHHGCYSLSWSKSRNQSQVLKMVVRKLKLCVLYIYPTVDPAVALDLCHLPALTIILVLGGD
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98      1499 PAYYERLLELDLCGAVSSRVDVPHSLAAKTHRGFTIGPETGPGVIRLEKLESASYAHPCLEELEFNAYLGSEQVDISDMCCLPLATPCKALFKPVYRSMQSFRLALM
H1766           PAYYERLLEMDLCGAVSSRVDIPHSLAARTHRGFTIGPDAGPGVIRLDRLESVCYAHPCLEELEFNAYLDSELVDISDMCCLPLATPCKALFRPIYRSLQSFRLALM
V               PAYYERLLEMDLCGAVSSRVDIPHSLAGRTHRGFAVGPDAGPGVIRLDRLESVCYAHPCLEELEFNAYLDSELVDISDMCCLPLATPCKALFRPIYRSLQSFRLALM
VFR             ---------------------------A-------------------------------------------------------------------------------
He/80           PAYYERLLEMDLCGAVSSRVDIPHSLAARTHRGFTIGPDAGPGVIRLDKLESVCYAHPCLEELEFNAYLDSELVDISDMCCLPLATPCKALFRPVYRSLQSFRLALM
He/80Fr         -----------------------------------------------------------------------------------------------------------

No/98      1606 DNYGFLMDLVMIRGLDIRPHLEEFDELLVVGQYILGQLIQEEAVYYVGVVGRGPVVSRSSRWVGLKRITIGGRSPCPCAARLRDEDRQGSLLAGLPAELVQLLLVD
H1766           DNYSFVMDLIMIRGLDIRPYLEEFDELLVVGQHILGQPVLVEVVYYVGVVRKRPVLARHPWSADLKRITVGGRVPCPSAARLRDEDCQGSLLVGLPAGLTQLLIID
V               DNYSFVMDLIMIRGLDIRPHLEEFDELLVVGQHILGQPVLVEVVYYVGVVRKRPVLARHPWSADLKRITVGGRAPCPSAARLRDEDCQGSLLVGLPAGLTQLLIID
VFR             ----------------------------------------------------------------------------------------------------------
He/80           DNYSFVMDLITIRGVDIRPHLEEFDELLVVGQHILGQPVLVEVVYYVGVVGKRPVLARHPWSADLKRITVGASAIPECCWTA
He/80Fr         --------------L--------------------------------------------------------GRAPCPSAAGLRDEDCRGSLLVGLPAGLTQLLVVD
```

Fig. 2. For legend see facing page.

end of the viral genome, although it remains possible that additional nucleotides are present.

3′-RACE experiments were performed to define the sequence at the other end of the viral genome (Fig. 3 A, right panel). Here, either A- or C-tails were added to the 3′-ends of viral RNA molecules before they were subjected to reverse transcription using oligo(dT) or oligo(dG/dI) primers. Products were amplified by PCR and cloned into a plasmid vector. Most of the resulting clones contained viral cDNA fragments that suggested that the most likely 3′-end of the viral genome might be 5′ … CGCAACA 3′.

To confirm these data by an independent approach, we analysed a series of cDNA clones generated from viral RNA ends that were artificially fused by RNA ligase before the product was used for cDNA synthesis and PCR amplification. In this RNA-ligation/RT–PCR procedure, primers for reverse transcription were designed to drive cDNA synthesis across the junction site. Analysis of resulting cDNA clones from strains V, He/80 and H1766 revealed a rather non-uniform picture in that most of the clones appeared to carry deletions of variable lengths that seemed to map to either side of the fusion site (Fig. 3 B). To generate a consensus sequence, we decided to ignore all hypothetical deletions in order to deduce the putative sequence of the longest possible RNA template. The 3′-terminal sequence of this hypothetical molecule is identical to that determined by 3′-RACE. Its 5′-terminal sequence is identical to that determined by 5′-RACE except for an extension of 4 nt. Assuming that this hypothetical molecule indeed reflects a copy of the full-length viral RNA, the proper ends of the BDV genome would be 5′ UGUUGCGCUAC-AACAAA … and … UUUGUUGUUAACGCAACA 3′. These two sequences show a high degree of complementarity, indicating that the BDV genome can potentially form a panhandle (Fig. 4 A), as is the case for other members of the order *Mononegavirales*. However, the 5′-RACE data (Fig. 3 A) do not fit this model well.

Since we cannot rule out the possibility that intrinsic problems associated with the 5′-RACE procedure were responsible for the failure to unambiguously identify the 5′-end sequence of the viral genome, we used an alternative technique for analysis that involves ligation of a defined RNA oligonucleotide to the 5′-terminus of the genomic RNA prior to performing RT–PCR. To ensure that the genome ends were mono-phosphorylated, we treated the viral RNA with TAP before ligation of the RNA oligonucleotide. Analysis of the resulting cDNA clones revealed that 5′ GCGCUA … 3′ was the predominant sequence at the extreme 5′-end of the viral genome of He/80 (Fig. 3 C, left panel). Identical results were obtained with viral RNA that was not treated with TAP (Fig. 3 C, right panel), suggesting that the 5′ end of the viral genome is not tri-phosphorylated but rather mono-phosphorylated. Direct sequencing of the amplification product revealed a single prominent sequence (Fig. 3 D), indicating that the vast majority of viral RNAs carry no additional nucleotides at the 5′-end. It remains to be determined whether viral genomic RNAs with these non-complementary ends are replication-competent. It is possible that the 5′-end sequences determined here reflect terminally cleaved viral RNAs. This view is compatible with the observation that the viral RNAs seemed to carry a monophosphate at the 5′-end. If correct, this situation would be similar to that observed with Seoul virus (a hantavirus; Meyer & Schmaljohn, 2000), where terminally deleted genome ends were found to be abundantly present in persistently infected cells.

The end sequences of the BDV genome that we deduced from our experiments are in conflict with data from previous reports (Briese *et al*., 1994; Cubitt *et al*., 1994*a*) in which the genome ends of strains V and He/80 were determined (Fig. 4 B). The most striking difference is that we find an A residue at the 3′ terminus of the viral RNA, whereas the others did not (Fig. 4 B). It is not entirely clear why our analysis yielded a different result. One reason might be that previous RNA-ligation/RT–PCR studies were performed in combination with 3′-RACE of A-tailed RNA, but not with C-tailed viral RNA. Such restricted analysis can, by definition, not identify any terminal A residues at the 3′-end of the viral genome.

Assuming that the 3′-end of the BDV genome is indeed occupied by an A residue, the BDV polymerase has to initiate transcription with UTP. If true, this constellation would represent a unique case among non-segmented, negative-strand RNA viruses. The polymerases of viruses from the order *Mononegavirales* all seem to initiate transcription by employing either ATP or GTP. Among other negative-strand RNA viruses, Hantaan virus (Garcin *et al*., 1995) is known to carry an A residue at the 3′-end of the genome. However, transcription of the Hantaan virus genome is thought not to be initiated at the terminal A residue but rather at the C residue at position 3. After synthesis of a short nucleotide primer, a sophisticated ‘prime-and-realign’ mechanism ensures that the polymerase will synthesize a complete copy of the viral genome (Garcin *et al*., 1995). It remains to be determined whether similar mechanisms are at work in BDV.

Closer inspection of sequences near the 3′- and 5′-ends of the genomes of various BDV strains revealed both

Fig. 2. Comparison of L polymerase sequences from various BDV strains. Sequence information was from the following sources: strain No/98 (GenBank accession no. AJ311524), strain H1766 (GenBank accession no. AJ311523), strain V_{FR} (GenBank accession no. AJ311521), strain He/80_{FR} (GenBank accession no. AJ311522), strain V (GenBank accession no. U04608) and He/80 (GenBank accession no. L27077). Note that the He/80 reference sequence depicted in the figure has recently been corrected (March 2001). The new sequence features a full-length L ORF of identical length to the Freiburg subline of this virus. Variant amino acids are indicated in bold letters. Dashed lines indicate amino acids identical to the sequence above.

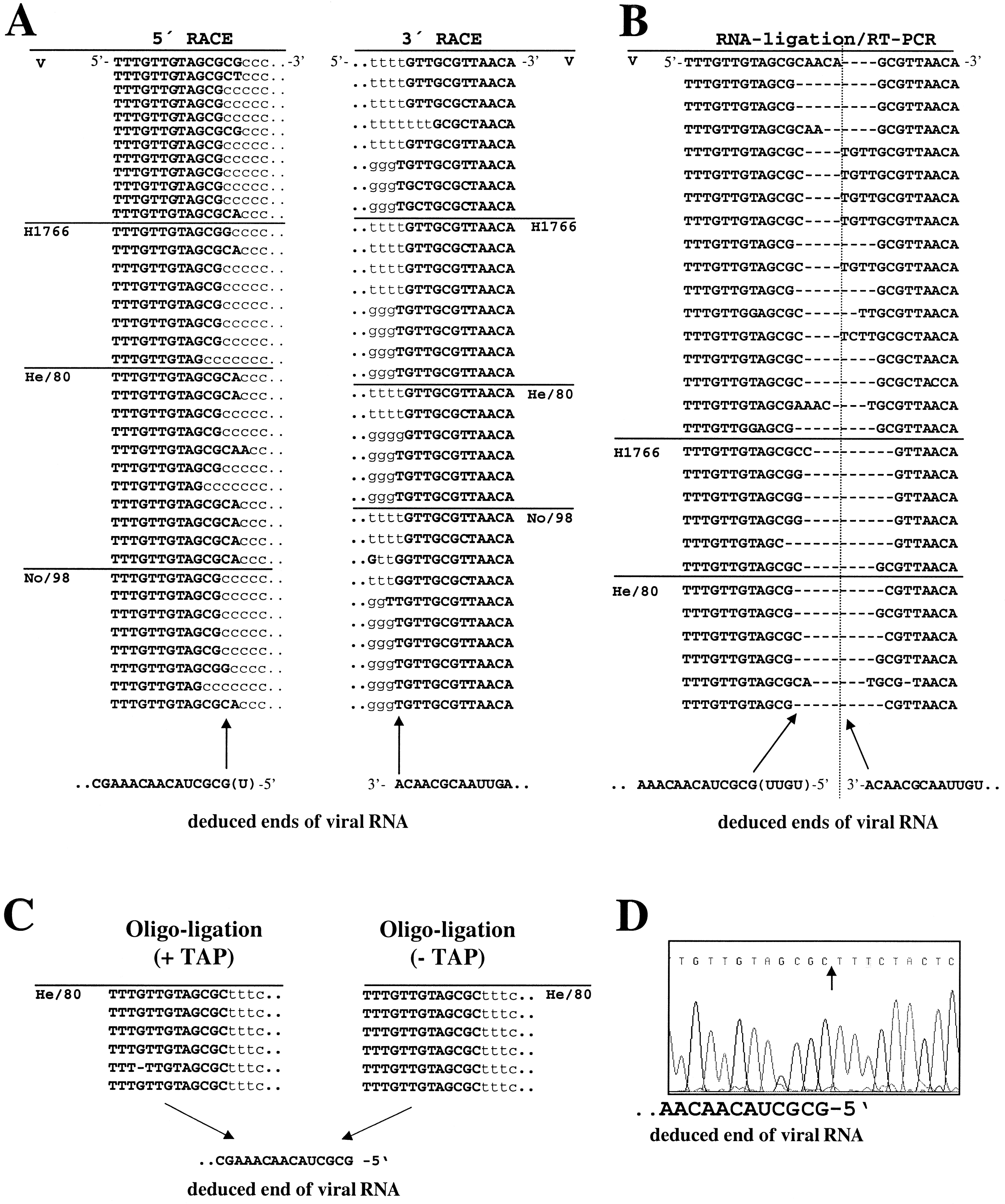

Fig. 3. Determination of the 5′- and 3′-termini of the BDV genome. (A) Terminal virus sequences deduced from the results of 5′- and 3′-RACE experiments. Lower-case letters mark terminal nucleotides that were introduced artificially during the RACE procedure. Strain origins of viral RNA used for the RACE experiments are indicated. Each lane shows the sequence (in antigenome orientation) of one individual RACE cDNA clone. Deduced ends of the viral genome are shown at the bottom. (B) Terminal virus sequences deduced from results of RNA-ligation/RT–PCR experiments. Each lane shows the sequence (in antigenome orientation) of one individual cDNA clone. Deduced ends of the viral genome are shown at the bottom. (C) Terminal virus sequences deduced from 5′-RACE experiments employing ligation of an RNA oligonucleotide to the 5′-end of the viral genome (Oligo-ligation). Prior to ligation of the oligonucleotide, the viral RNA was either treated (+TAP) or not

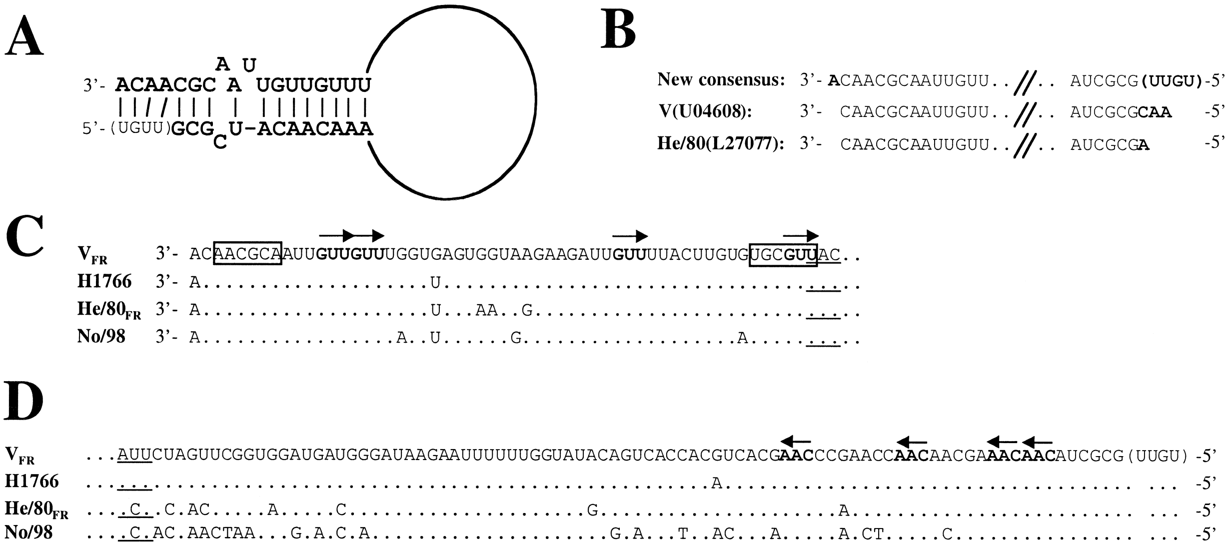

Fig. 4. Comparison of 5′- and 3′-termini of the BDV genome. (A) Potential secondary structure of paired 3′- and 5′-genome ends. Our experiments did not unambiguously prove the presence of nucleotides shown in brackets. (B) Comparison of our new consensus sequences with published genome end sequences of strain V and He/80. Variability between BDV strains in the non-coding regions near the 3′ (C) and 5′ end (D) of the viral genome. Arrows mark repetitive motifs that are found in complementary orientation near the 3′ and 5′ ends of the genome. Two inverted repeat motifs near the 5′ end of the viral genome that have the potential to form a stem–loop are boxed. Nucleotides complementary to the initiation codon of N (C) or the stop codon of L (D) are underlined.

non-conserved and highly conserved regions. Multiple copies of a highly conserved 5′ CAA 3′ motif are present near the 5′-end of the viral genome and are matched by an identical number of 5′ UUG 3′ motifs near the 3′ end (Fig. 4C, D). These repeats could potentially help to stabilize an extended panhandle structure of the BDV genome. Furthermore, a well-conserved inverted repeat consisting of the elements 3′ AACGCA 5′ and 3′ UGCGUU 5′ is present near the 3′-end of the viral genome (Fig. 4C). These elements could easily fold into stem–loop structures and thus influence the architecture of the genome ends. It is not known at present which functions these various motifs might have.

Conclusions and implications

Nucleotides located near the termini of the viral genome were previously found to play a critical role in establishing a technique that permits the genetic manipulation of negative-strand RNA viruses (Conzelmann, 1998). Of similar importance are cDNA clones used to reconstitute the functional viral polymerase complex (Conzelmann, 1998). Our new sequence information on the hypothetical structure of both ends of the BDV genome might explain why various artificial 'mini-replicons' that were based on published sequence information did not replicate to detectable levels in previous experiments. Functional tests with native and reconstituted viral polymerase complexes will be necessary to determine whether artificial BDV RNA molecules with the new structural features are replication-competent.

We thank Sibylle Herzog for providing a sample of BDV strain H1766, Keizo Tomonaga and Kazuyoshi Ikuta for letting us use the products of their 'long PCR' on RNA from strain H1766, and Christian Sauder, Georg Kochs, Friedemann Weber and Otto Haller for helpful comments on the manuscript.

This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Austrian Federal State of Vorarlberg.

References

Briese, T., de la Torre, J. C., Lewis, A., Ludwig, H. & Lipkin, W. I. (1992). Borna disease virus, a negative-strand RNA virus, transcribes in the nucleus of infected cells. *Proceedings of the National Academy of Sciences, USA* **89**, 11486–11489.

Briese, T., Schneemann, A., Lewis, A. J., Park, Y.-S., Kim, S., Ludwig, H. & Lipkin, W. I. (1994). Genomic organization of Borna disease virus. *Proceedings of the National Academy of Sciences, USA* **91**, 4362–4366.

Calain, P. & Roux, L. (1993). The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *Journal of Virology* **67**, 4822–4830.

treated (–TAP) with tobacco acid pyrophosphatase. Lower-case letters mark terminal nucleotides that were introduced artificially during the RACE procedure. The deduced end of the viral genome is shown at the bottom. (D) Terminal virus sequences deduced by direct sequencing of the 5′-RACE amplification product obtained after ligating the RNA oligonucleotide to the viral genome. The arrow indicates the junction between virus-specific and oligo-derived sequences. The deduced end of the viral genome is shown at the bottom.

Conzelmann, K. (1998). Nonsegmented negative-strand RNA viruses: genetics and manipulation of viral genomes. *Annual Review of Genetics* **32**, 123–162.

Cubitt, B. & de la Torre, J. C. (1994). Borna disease virus (BDV), a nonsegmented RNA virus, replicates in the nuclei of infected cells where infectious BDV ribonucleoproteins are present. *Journal of Virology* **68**, 1371–1381.

Cubitt, B., Oldstone, C. & de la Torre, J. C. (1994*a*). Sequence and genome organization of Borna disease virus. *Journal of Virology* **68**, 1382–1396.

Cubitt, B., Oldstone, C., Valcarcel, J. & Carlos de la Torre, J. (1994*b*). RNA splicing contributes to the generation of mature mRNAs of Borna disease virus, a non-segmented negative strand RNA virus. *Virus Research* **34**, 69–79.

Cubitt, B., Ly, C. & de la Torre, J. C. (2001). Identification and characterization of a new intron in Borna disease virus. *Journal of General Virology* **82**, 641–646.

Formella, S., Jehle, C., Sauder, C., Staeheli, P. & Schwemmle, M. (2000). Sequence variability of Borna disease virus: resistance to superinfection may contribute to high genome stability in persistently infected cells. *Journal of Virology* **74**, 7878–7883.

Garcin, D., Lezzi, M., Dobbs, M., Elliott, R. M., Schmaljohn, C., Kang, C. Y. & Kolakofsky, D. (1995). The 5′ ends of Hantaan virus (Bunyaviridae) RNAs suggest a prime-and-realign mechanism for the initiation of RNA synthesis. *Journal of Virology* **69**, 5754–5762.

Jehle, C., Lipkin, W. I., Staeheli, P., Marion, R. M. & Schwemmle, M. (2000). Authentic Borna disease virus transcripts are spliced less efficiently than cDNA-derived viral RNAs. *Journal of General Virology* **81**, 1947–1954.

Kobayashi, T., Watanabe, M., Kamitani, W., Tomonaga, K. & Ikuta, K. (2000). Translation initiation of a bicistronic mRNA of Borna disease virus: a 16-kDa phosphoprotein is initiated at an internal start codon. *Virology* **277**, 296–305.

Meyer, B. & Schmaljohn, C. (2000). Accumulation of terminally deleted RNAs may play a role in Seoul virus persistence. *Journal of Virology* **74**, 1321–1331.

Nakamura, Y., Nakaya, T., Hagiwara, K., Momiyama, N., Kagawa, Y., Taniyama, H., Ishihara, C., Sata, T., Kurata, T. & Ikuta, K. (1999). High susceptibility of Mongolian gerbil (Meriones unguiculatus) to Borna disease virus. *Vaccine* **17**, 480–489.

Nakamura, Y., Takahashi, H., Shoya, Y., Nakaya, T., Watanabe, M., Tomonaga, K., Iwahashi, K., Ameno, K., Momiyama, N., Taniyama, H., Sata, T., Kurata, T., de la Torre, J. C. & Ikuta, K. (2000). Isolation of Borna disease virus from human brain tissue. *Journal of Virology* **74**, 4601–4611.

Nowotny, N., Kolodziejek, J., Jehle, C., Suchy, A., Staeheli, P. & Schwemmle, M. (2000). Isolation of a new subtype of Borna disease virus. *Journal of Virology* **74**, 5655–5658.

Richt, J. A. & Rott, R. (2001). Borna disease virus: a mystery as an emerging zoonotic pathogen. *Veterinary Journal* **161**, 24–40.

Rott, R. & Becht, H. (1995). Natural and experimental Borna disease in animals. *Current Topics in Microbiology and Immunology* **190**, 17–30.

Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J. & Nichol, S. T. (1996). The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proceedings of the National Academy of Sciences, USA* **93**, 3602–3607.

Schneemann, A., Schneider, P. A., Kim, S. & Lipkin, W. I. (1994). Identification of signal sequences that control transcription of Borna disease virus, a nonsegmented negative-strand RNA virus. *Journal of Virology* **68**, 6514–6522.

Schneider, P. A., Briese, T., Zimmermann, W., Ludwig, H. & Lipkin, W. I. (1994*a*). Sequence conservation in field and experimental isolates of Borna disease virus. *Journal of Virology* **68**, 63–68.

Schneider, P. A., Schneemann, A. & Lipkin, W. I. (1994*b*). RNA splicing in Borna disease virus, a nonsegmented, negative-strand RNA virus. *Journal of Virology* **68**, 5007–5012.

Schneider, P. A., Kim, R. & Lipkin, W. I. (1997). Evidence for translation of the Borna disease virus G protein by leaky ribosomal scanning and ribosomal reinitiation. *Journal of Virology* **71**, 5614–5619.

Schwemmle, M., Hatalski, C. G., Lewis, A. J. & Lipkin, W. I. (1999). Borna virus. In *Persistent Viral Infections*, pp. 559–573. Edited by R. Ahmed & I. Chen. Chichester: John Wiley & Sons.

Shoya, Y., Kobayashi, T., Koda, T., Lai, P. K., Tanaka, H., Koyama, T., Ikuta, K., Kakinuma, M. & Kishi, M. (1997). Amplification of a full-length Borna disease virus (BDV) cDNA from total RNA of cells persistently infected with BDV. *Microbiology and Immunology* **41**, 481–486.

Staeheli, P., Sauder, C., Hausmann, J., Ehrensberger, F. & Schwemmle, M. (2000). Epidemiology of *Borna disease virus*. *Journal of General Virology* **81**, 2123–2135.

Stitz, L., Bilzer, T., Richt, J. A. & Rott, R. (1993). Pathogenesis of Borna disease. *Archives of Virology*, Suppl. **7**, 135–151.

Tomonaga, K., Kobayashi, T., Lee, B. J., Watanabe, M., Kamitani, W. & Ikuta, K. (2000). Identification of alternative splicing and negative splicing activity of a nonsegmented negative-strand RNA virus, Borna disease virus. *Proceedings of the National Academy of Sciences, USA* **97**, 12788–12793.

Walker, M. P., Jordan, I., Briese, T., Fischer, N. & Lipkin, W. I. (2000). Expression and characterization of the Borna disease virus polymerase. *Journal of Virology* **74**, 4425–4428.

Wehner, T., Ruppert, A., Herden, C., Frese, K., Becht, H. & Richt, J. A. (1997). Detection of a novel Borna disease virus-encoded 10 kDa protein in infected cells and tissues. *Journal of General Virology* **78**, 2459–2466.

Yang, Z., Delgado, R., Xu, L., Todd, R. F., Nabel, E. G., Sanchez, A. & Nabel, G. J. (1998). Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. *Science* **279**, 1034–1037.

Received 25 May 2001; Accepted 31 July 2001