Identification of Signal Sequences That Control Transcription of Borna Disease Virus, a Nonsegmented, Negative-Strand RNA Virus

ANETTE SCHNEEMANN,¹ PATRICK A. SCHNEIDER,² SARA KIM,¹ AND W. IAN LIPKIN^{1,2*}

Laboratory for Neurovirology, Department of Neurology,¹ and Department of Microbiology & Molecular Genetics,² University of California, Irvine, Irvine, California 92717

Received 22 June 1994/Accepted 30 June 1994

Borna disease virus (BDV) is a nonsegmented, negative-strand RNA virus that causes neurologic disorders in a wide range of animal species. Although the virus is unclassified, sequence analysis of the 8.9-kb viral genome has shown that it is related to rhabdoviruses and paramyxoviruses. We have mapped subgenomic RNAs of BDV strain He80-1 to the viral genome by determining the precise sequences at their 5' and 3' termini. This analysis showed that the genome contains three transcription initiation sites and four termination sites. A 14- to 16-nucleotide semiconserved sequence was present at the gene start sites and partially copied into the subgenomic RNAs. The termination sites contained a U-rich motif reminiscent of termination signals in rhabdoviruses and paramyxoviruses. In contrast to the genomes of other nonsegmented, negative-strand RNA viruses, the BDV genome lacked the typical configuration of termination signal, intergenic region, and initiation signal at the gene boundaries. Instead, transcription units and transcription signals frequently overlapped. These differences have implications for our understanding of the control of viral transcription and may relate to the low-level replication and persistence of BDV.

Borna disease virus (BDV) is a nonsegmented, negative-strand RNA virus that infects a broad range of animal species including birds, rodents, primates, and possibly humans (2, 3, 9, 13, 14, 20). It causes an immune system-mediated neurologic syndrome characterized by movement disorders and behavioral disturbances (18, 24). BDV has no cytopathic effect on cultured cells and grows only to very low titers (12, 14). These features have hampered classical biochemical and biophysical analyses of the virus and only limited information is available on the morphology and structural proteins of BDV. Virions are enveloped and contain nucleocapsids that are infectious when transfected into BDV-susceptible cells (7, 12, 14).

The 8.9-kb genome of BDV has recently been cloned and sequenced independently in two laboratories (4, 8). Analysis of the nucleotide sequence showed that BDV is related to rhabdoviruses and paramyxoviruses. The genome contains antisense information for five open reading frames (ORFs) flanked by short extracistronic sequences at the 3' and 5' termini (4) (see Fig. 1). The five ORFs code for proteins of 40 kDa (p40), 23 kDa (p23), 16 kDa (18 kDa after glycosylation; gp18), 57 kDa (p57), and 190 kDa (pol). The 190-kDa protein is likely to represent the viral RNA-dependent RNA polymerase because it contains the amino acid motifs typically conserved among these proteins (4, 8). The functions of the remaining BDV proteins are unknown.

While animal rhabdoviruses and paramyxoviruses transcribe and replicate in the cell cytoplasm (1), BDV transcription and replication occurs in the nucleus (3, 7). Six primary, polyadenylated BDV transcripts have been identified in infected cells and mapped to the genome by Northern (RNA) hybridization (4). Two of these transcripts, a 7.1-kb RNA and a 2.8-kb RNA, are posttranscriptionally modified by RNA splicing to yield at least two additional RNA species, 6.1 and 1.5 kb in length (23).

The work described here was initiated to identify regulatory

elements that control transcription of the BDV genome. To this end, we determined the precise sequences at the 5' and 3' termini of subgenomic RNAs of BDV strain He80-1 and constructed a detailed transcriptional map. Our results show that the BDV genome, similar to that of other nonsegmented, negative-strand RNA viruses, contains conserved sequences that are likely to serve as signals for transcription initiation and termination. However, the gene boundaries do not display the typical linear order of transcription termination signal, intergenic region, and transcription initiation signal, suggesting differences in the mechanisms by which these signals are recognized by the viral polymerase.

MATERIALS AND METHODS

Source of viral RNA. Rats were infected with BDV strain He80-1, and RNA was obtained from their brains on day 21 postinfection as described previously (13).

Preparation of poly(A)⁺ RNA and Northern blot analysis. Poly(A)⁺ RNA was isolated from total BDV-infected rat brain RNA with oligo(dT)-coated magnetic beads (Dynal, Great Neck, N.Y.) by using protocols provided by the manufacturer. Electrophoresis of RNA in agarose-formaldehyde gels and Northern blot analysis were performed as described previously (23). All probes used for hybridization were digoxigenin-UTP-labeled antisense RNAs generated by *in vitro* transcription from linearized plasmids containing the appropriate cDNAs. Synthesis of probes C (gp18) and D (p57) has been described by Schneider et al. (23). In that publication, the names of probes C and D were I and III, respectively. To generate probes A, B, and E, plasmids containing a full-length cDNA clone of the genes for p40 (probe A) (25) and p23 (probe B) (25) or a partial cDNA clone of the *pol* gene (probe E) (plasmid 8.31 [4]) were linearized with *Hind*III (for p40 and *pol*) or *Xba*I (p23). Transcription of digoxigenin-UTP-labeled antisense RNAs was performed by using protocols provided by the manufacturer (Boehringer Mannheim, Indianapolis, Ind.),

* Corresponding author. Phone: (714) 856-6193. Fax: (714) 725-2132.

TABLE 1. Oligonucleotides used for RT-PCR analysis of circularized BDV RNAs^a

RNA	Oligonucleotide primers used for:		
	Reverse transcription	PCR	Nested PCR
0.8 kb	5'-CTGAGATCATGGAGGGTTC-3' (p2)	5'-CTGAGATCATGGAGGGTTC-3' (p2) 5'-GATACTAGATCGCTCCATGAA-3' (p7)	5'-CTCCAGGGAGTCGACCAGACT-3' (p11) 5'-TGGAGACAATGAAGCTCAT-3' (p15)
1.2 kb	5'-AAACATAGAAACACAAGG-3' (p1)	5'-GAAACATATCGCGCCGTGAC-3' (p4) 5'-AGATGATAGGTGTGACTGGT-3' (p5)	ND ^b
1.5 kb	5'-ATACTTCAGGGGGCAATACA-3' (p3)	5'-ATACTTCAGGGGGCAATACA-3' (p3) 5'-GTACAACCACTTTCTTGCTGC-3' (p8)	5'-TTCCTTGCTTTAATCTGGCC-3' (p16) 5'-CATAATGGATCAGTCTCCTCTGA-3' (p17)
1.9 kb	5'-AAACATAGAAACACAAGG-3' (p1)	5'-GAAACATATCGCGCCGTGAC-3' (p4) 5'-TTTGAGTCCCTGTCCGCC-3' (p6)	5'-GAGCTTAGGGAGGCTCGCTG-3' (p14) 5'-GATACTAGATCGCTCCATGAA-3' (p7)
2.8 kb	5'-ATACTTCAGGGGGCAATACA-3' (p3)	5'-GTCCTCTGGTGTGAGTTGTT-3' (p9) 5'-GATTTCGCGTCCGTATTGC-3' (p10)	ND
3.5 kb	5'-ATACTTCAGGGGGCAATACA-3' (p3)	5'-CTCCAGGGAGTCGACCAGACT-3' (p11) 5'-CTACAACCACTTTCTTGCTGC-3' (p8)	5'-AGCAATGTCAGCCGGAGGTC-3' (p18) 5'-CTACAACCACTTTCTTGCTGC-3' (p8)
6.1 kb	5'-ATACTTCAGGGGGCAATACA-3' (p3)	5'-ATACTTCAGGGGGCAATACA-3' (p3) 5'-GCCTCCCTTAGCGACACCTGTA-3' (p12)	5'-TTCCTTGCTTTAATCTGGCC-3' (p16) 5'-TAGCTTGGAGTTGTTAGGAAGC-3' (p19)
7.1 kb	5'-ATACTTCAGGGGGCAATACA-3' (p3)	5'-GTCCTCTGGTGTGAGTTGTT-3' (p9) 5'-GGTTGGGCTTCCCGCTGG-3' (p13)	5'-TGATGTGTGGACTGCCGGA-3' (p22) 5'-GAGCCATCTACTGCCTA-3' (p21)
8.9 kb	5'-AAACATAGAAACACAAGG-3' (p1)	5'-AAACATAGAAACACAAGG-3' (p1) 5'-GCCTCCCTTAGCGACACCTGTA-3' (p12)	5'-GAAACATATCGCGCCGTGAC-3' (p4) 5'-GCTGCCAGATTGCGTGATG-3' (p20)

^a Reverse transcription and PCR amplification of circularized RNAs was performed as described in Materials and Methods. Primers used for reverse transcription of each circularized template are shown in column 2. Column 3 shows primer pairs used for PCR amplification over the ligated RNA termini. If no DNA product was detectable in a 10- μ l aliquot of the PCR mixture in an ethidium bromide-stained agarose gel, PCR amplification was repeated with a nested set of primers (column 4).

^b ND, not done.

with T7 RNA polymerase to generate probes for ORFs p40 and *pol* and SP6 RNA polymerase to generate probes for ORF p23.

Synthesis of biotinylated primers used for purification of BDV subgenomic RNAs. Biotinylated primers complementary to specific BDV RNAs were custom synthesized (Midland Certified Reagent Co.). The sequence of each primer and its specificity were as follows: anti-1.2-kb RNA, 5'-biotin-CTCG GCTCCTGCTTTGATCTTAGACGACGATCCTATCACA ACCCT-3'; anti-0.8-kb RNA, 5'-biotin-ATGCATTCCTT GGGACCTTCCGTGGTCTTGGTGACCCCGGTGCGTT-3'; anti-2.8-kb RNA, 5'-biotin-CTGGTTCCGTGAAGTCCCC CTACAAAGTCTATCTCAAGCATCAG-3'; anti-1.5-kb RNA, 5'-biotin-CTGCAGCAAGAAAGTGGTTGTAGGTCATTA GGCAGCTTGTCTGTG-3'; anti-7.1-kb RNA, 5'-biotin-GA TTGTCTGATTATCACCTTGGCCTAGTATGTTAAACGT C-3'.

Purification of BDV subgenomic RNAs. Total RNA (100 μ g) from BDV-infected rat brain tissue was mixed with 1 μ g of 5'-biotinylated primer in 100 μ l of hybridization buffer (50 mM HEPES [N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid; pH 7], 1 M NaCl, 2 mM EDTA, 20 U of RNasin [Promega, Madison, Wis.]) and incubated at 85°C for 2 min. An additional 20 U of RNasin was added, and incubation was continued at 55°C for 1 h. Streptavidin-coated magnetic beads (100 μ g) (Dynal) were added, and the mixture was incubated with gentle rotation at room temperature for 30 min. The supernatant was removed, and the beads were washed once each with 100 μ l of hybridization buffer, 100 μ l of 2 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and 100 μ l of 0.1 $\times$ SSC at room temperature. RNA was then eluted from the beads with 10 μ l of H₂O at 68°C.

Determination of the sequence at RNA termini. Purified BDV RNA was decapped with tobacco acid pyrophosphatase (TAP) (Epicentre Technologies, Madison, Wis.) and circularized with T4 RNA ligase (15). The circularized product was reverse-transcribed with Moloney murine leukemia virus reverse transcriptase (RT) (Promega) by using a primer comple-

mentary to a region near the 5' end of the RNA (Table 1). A 1- μ l aliquot of the reverse transcription (RT) reaction mixture was taken for amplification of the 5'-3' junction by PCR with AmpliTaq DNA polymerase (Stoffel fragment, 10 U/100 μ l of reaction mixture) (Perkin Elmer, Norwalk, Conn.). Primers used for each amplification are listed in Table 1. Cycle conditions were as follows: 95°C for 5 min followed by 55°C for 5 min, then 30 to 40 cycles at 70°C for 2 min, 95°C for 1.5 min, and 55°C for 1 min. If no product was detected in a 10- μ l aliquot of the PCR mixture in an ethidium bromide-stained agarose gel, the PCR sample was diluted 10-fold in water and 1 μ l of the diluted sample was subjected to a second round of PCR amplification with a nested set of primers (Table 1). PCR products were purified by agarose gel electrophoresis with the USBioclean purification kit (United States Biochemicals, Cleveland, Ohio) and cloned into BluescriptSKII+ (Stratagene, La Jolla, Calif.) prepared with 3' T-overhangs (16). Following transformation, plasmid DNA was isolated from 10 to 15 independent clones and the sequence across the 5'-3' junction was determined by using the Sanger dideoxy sequencing method (21). Three clones each were analyzed for the 3.5-, 6.1-, and 1.5-kb RNAs.

Determination of the presence or absence of 5'-terminal modifications in the 1.2- and 1.9-kb RNAs. Purified 1.2- and 1.9-kb RNAs were divided into two aliquots. One aliquot was decapped with TAP as described above. The other aliquot served as a control and received buffer without enzyme. Both samples were then subjected to circularization and reverse transcription with primer AB5SP6-3 (5'-AAACATAGAAAC ACAAGG-3'). PCR over the ligated ends of the 1.2-kb RNA was performed with primers AB5T720.IINV (5'-GAAACAT ATCGCGCCGTGAC-3') and AB5T7.5 (5'-AGATGATAG GTGTGACTGGT-3') under the following conditions: 95°C for 5 min followed by 50°C for 5 min, then 40 cycles at 70°C for 2 min, 95°C for 1.5 min, and 50°C for 1 min.

PCR over the ligated ends of the 1.9-kb RNA was performed with primers AB5SP6-2 (5'-GAAGAAATGGTGATCGGTT TG-3') and pr24SP6.tb (5'-TTTGAGTCCCTGTCCGCC-3')

FIG. 1. (A) Northern blot analysis of poly(A)⁺ RNA from BDV-infected rat brain by using probes complementary to the five ORFs on the BDV antigenome. Purified poly(A)⁺ RNA was size fractionated in an agarose-formaldehyde gel, transferred to a nylon membrane, and UV cross-linked. The membrane was cut into strips, and each strip was hybridized with a digoxigenin-UTP-labeled RNA probe complementary to one of the five BDV ORFs. Hybridization complexes were visualized by chemiluminescence. Equal amounts of RNA (0.2 μ g) were loaded in lanes 1 and 2; twice the amount of RNA (0.4 μ g) was loaded in lanes 3, 4, and 5. Probes used are indicated by a letter below each lane. The location of the probes relative to the BDV genome is shown in the diagram below the gel. Exposure time for lanes 1 and 2 was reduced relative to that for lanes 3, 4, and 5 to obtain comparable signals. Lane 6 represents an extended exposure of lane 3 to visualize the 3.5-kb RNA. (B) Previously established map (4) of the transcripts shown in the Northern blot in panel A. The dotted lines in the 1.5- and 6.1-kb RNAs represent a deleted region. Asterisks show the approximate locations of the biotinylated oligonucleotides used for purification of BDV-specific RNAs. Half arrows show approximate location of primers used for RT-PCR experiments.

under the same conditions except that the annealing temperature was raised from 50 to 55°C. After 30 cycles, the PCR mixture was diluted 200-fold with water and 1 μ l was reamplified for 20 cycles with primers AB5SP6-2 and AF4T720.1INV (5'-GATACTAGATCGCTCCATGAA-3'), which was nested with respect to primer pr24SP6.tb.

Then 10 μ l of the reamplified PCR mixture and 10 μ l of the 1.2-kb PCR mixture were subjected to electrophoresis through a 1% agarose gel in 40 mM Tris-acetate-1 mM EDTA followed by denaturation and transfer of the DNA products to a nylon membrane (Zetaprobe; Bio-Rad, Hercules, Calif.)

FIG. 2. Northern blot analysis of BDV RNAs before and after purification of the 0.8- and 1.9-kb RNAs from total cellular RNA. Total cellular RNA, isolated from BDV-infected rat brain, was incubated with a biotinylated oligonucleotide complementary to a common region in the 0.8- and 1.9-kb BDV RNA. The hybridization products were isolated by using streptavidin-coated magnetic beads as described in Materials and Methods, and eluted RNAs were size fractionated through an agarose-formaldehyde gel. Following transfer of nucleic acids to a nylon membrane, RNAs were hybridized to a mixture of probes A and B (Fig. 1A). Lanes: 1, total RNA used for purification of BDV RNAs; 2, purified BDV RNAs. The 8.9-kb RNA presumably represents the antigenomic RNA. This RNA was also visible in lane 2 after extended exposure.

(11). After UV cross-linking, the membrane was prehybridized for 2 h at 42°C in 6 $\times$ SSC-5 $\times$ Denhardt's solution-1.0% sodium dodecyl sulfate (SDS)-100 μ g of sheared salmon sperm DNA per ml. Radiolabeled oligonucleotides specific to the amplified region of the circularized 1.9-kb RNA (primer AF4SP6-3; 5'-GACAATGATGGAGACAATG-3') and the circularized 1.2-kb RNA (primer AB5SP620.2; 5'-GAGCTT AGGGAGGCTCGCTG-3') were used as probes for hybridization analysis. Oligonucleotides (1 μ g) were dephosphorylated with 1 U of calf intestinal alkaline phosphatase (Promega) for 30 min at 37°C and phenol-chloroform extracted. The dephosphorylated primers were then incubated for 30 min at 37°C in a 25- μ l reaction mixture with 10 U of T4 polynucleotide kinase (Gibco BRL, Gaithersburg, Md.) in the presence of 10 μ Ci of [γ -³²P]ATP (3,000 Ci/mmol). Following removal of unincorporated nucleotides, the probes (10⁶ cpm/3 ml) were added to the prehybridized membranes, and the membranes were incubated at 42°C for 6 h. Membranes were washed twice at 48°C for 15 min in 5 $\times$ SSC-0.1% SDS and exposed to Kodak XAR5 film for 5 h.

Determination of the He80-1 genomic sequence. Total RNA purified from brains of BDV He80-1-infected rats on day 21 postinfection was used as a template for cDNA synthesis by RT-PCR as described previously (22). Briefly, genomic RNA was reverse transcribed with primer p38SP6-1 (5'-TGCCCG GTATAGGCGCCGCG-3'), located downstream of termination site T1. The resulting cDNA was used as template for the PCR amplification of the regions spanning T1-S2 and T2-S3. The region spanning T1-S2 was amplified with primers p38SP6-1 and p24T7-1 (5'-CTCCAGGGAGTCGACCAGA CT-3'), and the region spanning T2-S3 was amplified with primers AF4SP6.3 (5'-GACAATGATGGAGACAATG-3') and I1-3'Splice (5'-CTGAAAGAAATGGGATGTAA-3'). Genomic RNA surrounding termination site T3 was reversed

FIG. 3. Autoradiograms of sequencing gels showing the sequence across the 5'-3' junction of cDNA clones representing circularized BDV subgenomic and antigenomic RNAs. The junction is indicated by arrows pointing up (5' end) and down (3' end). To facilitate sequence comparisons, autoradiograms in which bands are spaced approximately equally apart were chosen for illustration. The 5'-terminal base of the 1.9-kb RNA (G) and the antigenomic RNA (G) cannot be determined in the examples shown; however, both bases were identified unambiguously in sequence of other clones.

transcribed with primer 631SP6-5 (5'-CTACAACCACTTTCTTGCTGC-3'). The resulting cDNA was used as a template for the PCR amplification of the region surrounding T3 with primers 631SP6-5 and 841T7-4 (5'-GGCCACTTCGATTCTTTGAG-3'). PCR products were cloned and sequenced as described above. Sequence surrounding termination site T4 was obtained by circularizing antigenomic RNA and performing RT-PCR over the ligated ends.

RESULTS

Purification of BDV subgenomic RNAs. Northern hybridization of poly(A)⁺ RNA from BDV-infected rat brain by using probes complementary to the viral ORFs revealed the presence of eight subgenomic RNAs with apparent lengths of 0.8, 1.2, 1.5, 1.9, 2.8, 3.5, 6.1, and 7.1 kb (Fig. 1A). These transcripts had already been mapped to the BDV genome; thus, sequence close to the ends of each RNA was known (4) (Fig. 1B). To facilitate the fine-mapping experiments described below, BDV RNAs were separated from total cellular RNA by directed immobilization onto streptavidin-coated magnetic beads. Briefly, total RNA from brains of rats acutely infected with BDV strain He80-1 (22) was hybridized to a 5'-biotinylated oligonucleotide complementary to a subset of the viral transcripts (Fig. 1B). The hybridization products were captured on streptavidin-coated magnetic beads, and RNAs were eluted after a series of stringent washes. This procedure was highly specific and yielded intact viral RNAs, including full-length antigenomic RNA (Fig. 2).

Sequence determination at the 5' and 3' termini of BDV RNAs. To determine the precise sequence at the 5' and 3' termini of each transcript, we took advantage of a method described by Mandl et al. (15). Purified RNAs were treated with TAP to remove possible cap structures at the 5' end and

circularized with T4 RNA ligase. The 5'-3' junction was amplified by RT-PCR with primers close to the ligated ends. The resulting DNA fragments were purified, cloned, and sequenced. Primers for RT-PCR were chosen such that each amplification product could be assigned unambiguously to the template from which it arose (Fig. 1B). This allowed separate analysis of the RNA species that copurified in the procedure described above.

The results of the sequencing analysis are shown in Fig. 3. Identification of the 5'-3' junction was straightforward since all of the RNAs, except the antigenomic RNA, ended in a stretch of adenylate residues. Because only partial genomic sequence was available for BDV strain He80-1, we used the full-length sequence of strain V (4) to identify the nucleotide positions on the BDV genome at which RNA synthesis initiated and terminated. In both strains, the sequences surrounding the transcriptional start and stop sites were found to be highly conserved (data not shown); nonetheless, analogous nucleotide positions for BDV He80-1 are predicted to be different since the genome of strain He80-1 is 3 nucleotides (nt) shorter than that of strain V at the 3' end.

As illustrated in Fig. 4, the 1.2-kb RNA, which encodes the 40-kDa protein (17), started at position 43 and ended at position 1192; thus, the exact length of this transcription unit was 1,150 nt. A single clone among the 10 clones that were analyzed started at position 42, indicating minor variability in selection of the transcriptional start site used by the BDV polymerase. It is not clear whether this variability is a unique property of the BDV polymerase or whether it simply reflects the sensitivity of the mapping technique used in this study. Surprisingly, the 0.8-kb RNA, which encodes the 23-kDa protein (26), initiated at position 1175, i.e., within the preceding gene (Fig. 4). It terminated 707 bases downstream at position 1882.

A.

B.

FIG. 4. (A) Location of mRNA initiation and termination sites on the BDV genome. Genomic RNA sequences (vRNA) and subgenomic RNA sequences are those of BDV strain He80-1; however, the nucleotide position numbers (arrows) were derived from the full-length sequence of BDV strain V (4) because only partial genomic sequence was available for strain He80-1. Position 4 of BDV strain V corresponds to position 1 in strain He80-1 (circled numeral 1). Boxed sequences represent the conserved transcription initiation signal; underlined sequences represent the conserved transcription termination signal. The C residue marked with an asterisk in the top vRNA line corresponds to the start site of the 0.8-kb RNA. It is highlighted to illustrate overlap of the 1.2- and 0.8-kb RNAs. Boldface letters indicate mismatch between vRNA and mRNA (positions 1 and 1904). (B) Detailed map of the transcripts in panel A relative to the BDV genome. Transcriptional start sites are represented by arrows pointing downstream.

The 1.9-kb RNA had previously been shown to contain the sequences of both the 1.2-kb RNA and the 0.8-kb RNA and was thought to represent a transcriptional readthrough product (4, 19). However, we found that the 1.9-kb RNA did not initiate at the same position as the 1.2-kb RNA; rather, it initiated at the extreme 3' end of the BDV genome with a mismatch between template and RNA transcript at the terminal base (Fig. 4). The 3' end of the 1.9-kb RNA terminated at the same position as the 0.8-kb RNA; however, in contrast to the 1.2- and 0.8-kb RNAs, whose clones contained between 20 and 40 nontemplated adenylate residues, the 15 clones analyzed for the 1.9-kb RNA contained only 8 or 9 adenylate residues, 7 of which were encoded by the genome.

All clones representing the 7.1- and 2.8-kb RNAs initiated at position 1885, only 3 nt downstream of the preceding termination site used for the 0.8- and 1.9-kb transcripts (Fig. 4). The 2.8-kb RNA terminated at position 4511 (2,627 templated bases), whereas the 7.1-kb RNA extended further downstream to position 8855 (6,971 templated bases).

Three clones were analyzed for the 3.5-kb RNA. Similar to the variability observed at the 5' terminus of the 1.2-kb RNA, one of these clones initiated at the same position as the 0.8-kb RNA (nt 1175), while two clones initiated 1 base upstream at position 1174. In all three clones the termination site was located at nt 4511. Thus, the 3.5-kb RNA represented a readthrough product that was coterminal with the 0.8-kb RNA at the 5' end and the 2.8-kb RNA at the 3' end (Fig. 4).

The two remaining RNAs, 6.1 and 1.5 kb, have previously been shown to represent splice products of the 7.1- and 2.8-kb RNAs, respectively (23). Both of the primary transcripts contain two introns, 94 and 1,294 nt in length, that can be alternatively spliced to yield three additional smaller RNAs (Fig. 4B). Splice products that differ in length by only 94 nt comigrate under our electrophoresis conditions and are not visible as separate bands on the Northern blot shown in Fig. 1A. To reconfirm the location of both introns, we amplified the 5'-3' junctions of circularized 6.1- and 1.5-kb RNAs with primers that spanned the deleted regions (Fig. 1B). Sequence analysis of the reaction products showed that the 5'- and 3'-terminal sequences of the 6.1- and 1.5-kb RNAs were identical to those of their parent RNAs, the 7.1- and 2.8-kb RNA, respectively and confirmed the position of intron 1 (nt 1932 to 2025) and intron 2 (nt 2410 to 3703) (data not shown).

Consensus sequences at RNA start and termination sites. Analysis of genomic sequence flanking the positions of transcription initiation and termination revealed conserved motifs at the gene boundaries (Fig. 4 and 5). Each RNA, except the 1.9-kb RNA, was initiated at a C residue followed by two U residues; in addition, the region upstream of the first transcribed base was highly conserved and rich in U residues. The transcription termination sites consisted of an A residue followed by 6 or 7 U residues. Interestingly, the termination signal of the 0.8-kb RNA was embedded within the transcription initiation signal for the 2.8-kb RNA (Fig. 5).

Absence of 5'-terminal modification in the 1.9-kb RNA. As described above, the 1.9-kb RNA initiated at the extreme 3' end of the BDV genome. RNAs initiating at the 3'-terminal genomic base have also been observed in other nonsegmented,

negative-strand RNA viruses and typically represent leader RNAs or leader-containing RNAs (1). Leader RNAs and leader-containing RNAs are usually not capped at the 5' end, in contrast to RNAs that initiate at an internal site on the genome (1). To test whether there was any difference in the 5'-terminal modification of the 1.9-kb RNA and the 1.2-kb RNA, purified 1.9- and 1.2-kb RNAs were subjected to circularization and amplification over the 5'-3' junction without prior decapping by TAP. As seen in Fig. 6, decapping was essential to achieve circularization only of the 1.2-kb RNA but not of the 1.9 kb RNA. This result suggested that, whereas the 5' end of the 1.9-kb RNA contained a free phosphate group, the 1.2-kb RNA contained a blocking group, presumably a cap structure. The capping status of other BDV transcripts was not further investigated.

DISCUSSION

We have determined the sequences at the 5' and 3' termini of genomic and subgenomic BDV RNAs by RNA circularization and RT-PCR over the ligated ends. By using this method, we have mapped viral transcripts precisely to the BDV genome and identified sequences that are likely to serve as signals for transcription initiation and termination.

The observation that BDV transcription occurs in the nucleus of infected cells (3, 7) suggested the possibility that BDV, like influenza virus, uses cellular mRNAs to prime synthesis of its RNA transcripts (10). This strategy results in sequence heterogeneity and the presence of nontemplated bases at the 5' termini of the viral RNAs (10). However, we found that, with the exception of minor variability, the sequences at the 5' ends of the BDV transcripts were homogeneous and genome encoded. Thus, the possibility that BDV transcription requires priming by cellular mRNAs was eliminated.

Transcription initiation and termination sites on the BDV genome. The 5' termini of BDV subgenomic RNAs mapped to four sites on the BDV genome. In BDV V, these sites corresponded to nt 1, 43 (S1), 1175 (S2), and 1885 (S3) (Fig. 4). Analysis of genomic sequence surrounding the positions of transcription initiation revealed a semiconserved, U-rich motif at S1, S2, and S3 (Fig. 5). This motif was also identified in the reported full-length sequences of BDV strains V (4) and He80-C (8) (data not shown), suggesting that it has a functional role, presumably as a transcription initiation signal. Comparison of the conserved motif with sequences present at the gene start sites of other nonsegmented negative-strand RNA viruses did not reveal significant similarities.

The transcription initiation sites identified in this study are not the same as those reported by Cubitt et al. for strain He80-C (8). We have reexamined the sequence of this strain and found that start sites S1, S2, and S3 are present and that, with the exception of one base substitution (position 4 in signal S1 [G to A]), He80-C start signals are identical to those reported here for strain He80-1. Thus, the differences do not reflect strain variability. Two groups, Briese et al. (4) and Cubitt et al. (8), have used Northern hybridization to map subgenomic RNAs to the viral genome. However, only Briese et al. identified deletions in subgenomic RNAs. Recognition of

Numbers refer to nucleotide positions. S1, S2, and S3 are start sites containing the consensus sequence. T1, T2, T3, and T4 are transcription termination sites. RNA transcripts are represented by horizontal lines. Stippled extension of the 2.8-kb RNA and its splice products indicates that readthrough at termination site T3 results in synthesis of the 7.1-kb RNA. Differential splicing of both intron 1 (nt 1932 to 2025) and intron 2 (nt 2410 to 3703) from the 2.8- and 7.1-kb RNAs would yield three smaller products including the 1.5- and 6.1-kb RNAs, respectively. The 3' end of the map is enlarged relative to the rest to show details.

Start sites		Termination sites	
S1	AUUGUUUUA ↓ CUU GU	T1	AUUUUUU
S2	CUUUUGUUA CUU GU	T2	AUUUUUUU
S3	AUUUUUUUAGCUUAGU	T3	AUUUUUUU
		T4	AUUUUUU

FIG. 5. Comparison of genomic sequences at transcriptional start sites S1, S2, and S3 and transcriptional termination sites T1, T2, T3, and T4. The sequences, shown in the 3'-to-5' direction, are aligned for maximal similarity. The arrow indicates the position at which cRNA synthesis begins within the transcriptional start signal. Underlined bases in S1 represent a weak termination signal which may be used for termination of leader RNAs. Termination site T2 and underlined bases in S3 occupy identical positions on the BDV genome.

these deletions led to the transcription map which has been the foundation for studies of RNA splicing and the identification of regulatory sequences reported here.

The four transcription termination sites identified were located at positions 1192 (T1), 1882 (T2), 4511 (T3), and 8855 (T4). Each termination site consisted of 6 or 7 U residues, preceded by a single adenylate residue (Fig. 4 and 5). This consensus sequence was reminiscent of the transcription termination signals found in rhabdoviruses and paramyxoviruses (1). In these virus families the stretch of U residues is believed to cause polymerase stuttering, which results in the synthesis of a poly(A) tail at the 3' end of the viral transcripts (1). It is likely that the U residues in BDV have an analogous function.

Unusual constellation of transcription termination and initiation signals at the gene junctions. The transcription initiation site for the 0.8-kb RNA (nt 1175) was located 18 nt upstream of the termination site of the 1.2-kb RNA (nt 1192) (Fig. 4). Thus, instead of showing the intergenic region typically observed in rhabdoviruses and paramyxoviruses (1), the two genes overlapped. Gene overlap has previously been observed in respiratory syncytial virus, a paramyxovirus (6). In respiratory syncytial virus, the start site for the polymerase gene is located 68 nt upstream of the termination site of the 22K gene. It has been proposed that this arrangement serves as a mechanism to attenuate expression of the polymerase gene (6). However, in BDV the 1.2- and 0.8-kb RNAs are the most abundant viral transcripts in infected cells (Fig. 1A), implying that the overlap does not significantly affect transcription of the downstream gene. It is possible that the degree of attenuation is a function of the length by which the two transcription units overlap. If so, a stretch of 18 nt may not be sufficient to cause a noticeable decrease in transcription of the 0.8-kb RNA. Further studies are required to elucidate the mechanistic details of BDV transcription at this gene junction.

The next two transcription units on the BDV genome (0.8- and 2.8-kb RNA transcription units) were separated by only 2 nt. Interestingly, the transcription initiation signal for the 2.8-kb RNA (S3) extended upstream across the intergenic region into the termination signal of the 0.8-kb RNA (T2). In fact, T2 was completely contained within S3 (Fig. 5). This arrangement contrasted with the organization of the rhabdovirus and paramyxovirus gene junctions, which can usually be divided into three separate domains: transcription termination signal, intergenic region, and transcription start signal (1). The overlap of these domains at the junction of the second and third genes in BDV did not appear to weaken the efficiency of either the termination or initiation signal. RNAs terminating at

FIG. 6. Southern blot analysis of RT-PCR products generated from untreated or decapped circularized 1.2- and 1.9-kb RNAs. Purified 1.2- and 1.9-kb RNAs were subjected to circularization with (+) or without (-) treatment with TAP to remove 5' caps. The ligated 5'-3' ends were amplified by RT-PCR as described in Materials and Methods. An aliquot of the PCR mixture (10 μ l) was subjected to electrophoresis through a 1% agarose gel and then transferred to a nylon membrane for Southern hybridization with 32 P-labeled oligonucleotides specific for the amplified products. Depending on the length of the poly(A) tail, the size of the expected DNA fragment for the 1.2-kb RNA was 300 to 330 bp while that for the 1.9-kb RNA was 220 bp. The identity of the smaller PCR product generated from the circularized 1.9-kb RNA is not known.

T2 (0.8- and 1.9-kb RNAs) were abundant, whereas read-through products (3.5-kb RNA) were rare (Fig. 1A). At the same time, initiation of transcription at S3 appeared to be efficient as indicated by the high levels of the 2.8- and 7.1-kb RNAs and their splice products. It is not clear how the polymerase recognizes the overlapping domains as separate functional entities. However, it is possible that there are additional sequences upstream of the termination sites that prepare actively transcribing polymerase for termination.

Transcriptional readthrough at termination site T3 is critical for expression of the polymerase gene. RNA synthesis initiating at S1, S2, or S3 usually resulted in termination at T1, T2, and T3, respectively. However, readthrough at all of these termination sites was observed and gave rise to multiple polycistronic RNAs that ended at the next termination site downstream. RNAs resulting from successive readthrough at more than one termination site were not detected, possibly because such transcripts are present at very low levels. The region between ORFs p40 and p23 on the 1.9-kb RNA has previously been cloned and shown to be a precise complement of the genome (22); specifically, it does not contain the polyadenylate sequence normally synthesized at a transcription termination signal. Whether this is true for other readthrough products, such as the 3.5- and 7.1-kb RNAs, remains to be determined.

Readthrough at transcriptional termination sites is a common phenomenon in nonsegmented, negative-strand RNA viruses and is usually referred to as aberrant termination (1). Whether the polytranscripts have any biological significance is not known. In BDV, the ability of the polymerase to continue transcription beyond a termination signal appears to be critical for viability of the virus: the 7.1-kb RNA and its splice products, transcripts resulting from readthrough at T3, are the only RNAs that contain the *pol* ORF. It cannot be excluded that other as yet unidentified BDV RNAs exist, but no additional transcription initiation sites downstream of S3 were

detected by searching for sequences similar to the gene start motif.

Transcriptional readthrough might be used as a mechanism to regulate gene expression of BDV. For example, low-level readthrough at T3 would lead to decreased expression of the BDV polymerase, which should be needed only in catalytic amounts. Support for this hypothesis comes from the observation that the levels of the 7.1-kb RNA and its splice products are significantly lower than those of the 2.8-kb RNA and its splice products (Fig. 1A). Alternatively, the lower levels of the 7.1-kb RNA and its splice products may reflect a difference in stability of RNAs terminating at T3 or T4.

The 1.9-kb RNA is unlikely to be an mRNA. Our results indicated that the 1.9-kb RNA was fundamentally different from the other subgenomic BDV RNAs. First, in contrast to the 1.2-kb RNA, the 1.9-kb RNA initiated at the extreme 3' end of the BDV genome, which does not contain the consensus sequence observed at start sites S1, S2, and S3. In addition, RNA circularization experiments indicated that the 1.9-kb RNA was not capped. These data suggested that the 1.9-kb RNA represented an analog of a leader-containing RNA similar to those observed in rhabdoviruses and paramyxoviruses (1). In these virus families, transcription of the sequence at the 3' end of the genome usually results in formation of a short leader RNA that terminates upstream of the first gene at a sequence similar to a transcription termination signal (1). Although no such short leader RNAs have yet been found in BDV-infected cells, inspection of the genomic sequence upstream of the start of the first transcription unit revealed a motif located within start signal S1 that was very similar to the consensus termination signal: 3'-AUUGUUUU (nt 34 to 42) (Fig. 5). It is possible that this signal is a weak terminator as a result of the exchange of a G residue for a U residue at the fourth position. RNA synthesis initiating at the extreme 3' end of the genome might therefore proceed into the adjacent gene to result in the synthesis of a leader-containing RNA. There is precedent for such leader-containing RNAs in measles virus, a paramyxovirus (5). Measles virus leader-containing readthrough transcripts are functionally distinct from their leaderless counterparts: they are not translated but are preferentially encapsidated into ribonucleoprotein complexes (5).

It is intriguing that the 1.9-kb RNA is not polyadenylated to the same extent as the other subgenomic RNAs. From the perspective of the transcription-replication model of nonsegmented, negative-strand RNA viruses, the 1.9-kb transcript could also be considered a replication intermediate that was aborted at termination site T2. If this model is correct, the polymerase would not be expected to synthesize a poly(A) tail at the 3' end of the newly synthesized RNA strand.

Efficient utilization of the BDV genome. The BDV genome is extremely compact: 99.4% of its nucleotides are transcribed into subgenomic RNAs. Only 55 of 8,910 bases (strain V) were not found in the viral transcripts analyzed in this study. These bases represent the trailer region at the 5' end of the genome. The region between the 3' end of the genome and the first base of the first transcription unit is 42 nt long and probably corresponds to the leader sequence found in other nonsegmented, negative-strand RNA viruses. Two other bases located in the intergenic region between T2 and S3 were found only in the rare polycistronic 3.5-kb RNA.

Coding potential of BDV transcripts. The transcriptional map of BDV contains only two primary RNAs that are monocistronic mRNAs for BDV proteins: the 1.2-kb RNA represents the mRNA for p40, while the 0.8-kb RNA represents the mRNA for p23. The 5' untranslated region of the 1.2-kb RNA contains either 11 or 51 nt, depending on which

AUG codon is used for translation of p40. Since the second AUG codon is located in a more favorable sequence context for translation initiation (4) and since a stretch of only 11 nt at the 5' end would result in inefficient binding of ribosomes, it is likely that the second AUG codon is used for translation of p40.

Expression of gp18, p57, and *pol* must be more complex. In vivo and in vitro studies have shown that the 2.8-kb RNA is posttranscriptionally modified by differential splicing of two introns (23). In this report we have independently confirmed the positions of the introns and provided further evidence that the 7.1-kb RNA is spliced as well. We have not yet determined whether the 3.5-kb RNA undergoes posttranscriptional processing. On the basis of its map position, the 3.5-kb RNA contains both introns 1 and 2 and might be spliced in a manner similar to the 2.8- and 7.1-kb RNAs. Which of the primary RNAs or splice products serves as the mRNA for gp18, p57, or *pol* is not yet known.

In summary, our results show that BDV is similar to other nonsegmented, negative-strand RNA viruses in its use of conserved sequences that serve as transcription initiation and termination signals. However, BDV appears to use an unprecedented variety of strategies to regulate the synthesis of its RNA transcripts. These include the use of overlapping transcription units, overlapping transcriptive signals, readthrough at transcription termination sites, and RNA splicing. Studying the mechanisms for control of viral gene expression in this multilevel system will be important in understanding the biology of BDV, including its persistence in cells of the central nervous system.

ACKNOWLEDGMENTS

We thank E. Ehrenfeld, E. Wagner, and members of the Laboratory for Neurovirology for critical review of the manuscript; P. Southern and B. Meyers for providing experimental protocols prior to publication; and Frank Nastanski for excellent technical assistance.

This work was supported by grants from the National Institutes of Health (NS 29425), the Pew Memorial Trust, and the Wayne and Gladys Valley Foundation.

REFERENCES

1. Banerjee, A. K. 1991. Gene expression of nonsegmented negative strand RNA viruses. *Pharmacol. Ther.* **51**:47-70.
2. Bode, L., S. Riegel, H. Ludwig, J. Amsterdam, W. Lange, and H. Koprowski. 1988. Borna disease virus-specific antibodies in patients with HIV infection and with mental disorders. *Lancet* **ii**:689.
3. Briese, T., J. C. de la Torre, A. Lewis, H. Ludwig, and W. I. Lipkin. 1992. Borna disease virus, a negative strand RNA virus, transcribes in the nucleus of infected cells. *Proc. Natl. Acad. Sci. USA* **89**:11486-11489.
4. Briese, T., A. Schneemann, A. Lewis, Y. Park, S. Kim, H. Ludwig, and W. I. Lipkin. 1994. Genomic organization of Borna disease virus. *Proc. Natl. Acad. Sci. USA* **91**:4362-4366.
5. Castaneda, S. J., and T. C. Wong. 1990. Leader sequence distinguishes between translatable and encapsidated measles virus RNAs. *J. Virol.* **64**:222-230.
6. Collins, P. L. 1991. The molecular biology of human respiratory syncytial virus (RSV) of the genus *Pneumovirus*, p. 103-162. In D. W. Kingsbury (ed.), *The paramyxoviruses*. Plenum Press, New York.
7. Cubitt, B., and J. C. de la Torre. 1994. Borna disease virus (BDV), a nonsegmented RNA virus, replicates in the nucleus of the infected cells where infectious BDV ribonucleoproteins are present. *J. Virol.* **68**:1371-1381.
8. Cubitt, B., C. Oldstone, and J. C. de la Torre. 1994. Sequence and genomic organization of Borna disease virus. *J. Virol.* **68**:1382-1396.
9. Fu, Z. F., J. D. Amsterdam, M. Kao, V. Shankar, H. Koprowski, and B. Dietschold. 1993. Detection of Borna disease virus-reactive

- antibodies from patients with affective disorders by western immunoblot technique. *J. Affective Disord.* **27**:61–68.
10. **Kingsbury, D. W.** 1991. Orthomyxoviruses and their replication, p. 527–541. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), *Fundamental virology*. Raven Press, New York.
 11. **Kriegler, M.** 1990. Gene transfer and expression. A laboratory manual. Stockton Press, New York.
 12. **Lipkin, W. I., T. Brieese, and J. C. de la Torre.** 1992. Borna disease virus: molecular analysis of a neurotropic infectious agent. *Microb. Pathog.* **13**:167–170.
 13. **Lipkin, W. I., G. Travis, K. Carbone, and M. Wilson.** 1990. Isolation and characterization of Borna disease agent cDNA clones. *Proc. Natl. Acad. Sci. USA* **87**:4184–4188.
 14. **Ludwig, H., L. Bode, and G. Gosztanyi.** 1988. Borna disease: a persistent virus infection of the central nervous system. *Prog. Med. Virol.* **35**:107–151.
 15. **Mandl, C. W., F. X. Heinz, E. Puchhammer-Stockl, and C. Kunz.** 1991. Sequencing the termini of capped viral RNA by 5'-3' ligation. *BioTechniques* **10**:484–486.
 16. **Marchuk, D., M. Drumm, A. Saulino, and F. S. Collins.** 1990. Construction of T-vector, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19**:1154.
 17. **McClure, M. A., K. J. Thibault, C. G. Hatalski, and W. I. Lipkin.** 1992. Sequence similarity between Borna disease virus p40 and a duplicated domain within the paramyxovirus and rhabdovirus polymerase proteins. *J. Virol.* **66**:6572–6577.
 18. **Narayan, O., S. Herzog, K. Frese, H. Scheefers, and R. Rott.** 1983. Behavioral disease in rats caused by immunopathological responses to persistent Borna virus in the brain. *Science* **220**:1401–1403.
 19. **Pyper, J. M., J. A. Richt, L. Brown, R. Rorr, O. Narayan, and J. E. Clements.** 1993. Genomic organization of the structural proteins of Borna disease virus revealed by a cDNA clone encoding the 38 kDa protein. *Virology* **195**:229–238.
 20. **Rott, R., S. Herzog, B. Fleischer, A. Winokur, J. Amsterdam, W. Dyson, and H. Koprowski.** 1985. Detection of serum antibodies to Borna disease virus in patients with psychiatric disorders. *Science* **228**:755–756.
 21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 22. **Schneider, P. A., T. Brieese, W. Zimmermann, H. Ludwig, and W. I. Lipkin.** 1994. Sequence conservation in field and experimental isolates of Borna disease virus. *J. Virol.* **68**:63–68.
 23. **Schneider, P. A., A. Schneemann, and W. I. Lipkin.** 1994. RNA splicing in Borna disease virus, a nonsegmented, negative-strand RNA virus. *J. Virol.* **68**:5007–5012.
 24. **Solbrig, M. V., J. H. Fallon, and W. I. Lipkin.** 1993. Behavioral disturbances and pharmacology of Borna disease. *In* W. I. Lipkin and H. Koprowski (ed.), *Borna disease*, in press. Springer-Verlag KG, Heidelberg, Germany.
 25. **Thibault, K. J.** 1992. M.S. thesis. University of California, Irvine.
 26. **Thierer, J., H. Riehle, O. Grebenstein, T. Binz, S. Herzog, N. Thiedemann, L. Stitz, R. Rott, F. Lottspeich, and H. Niemann.** 1992. The 24K protein of Borna disease virus. *J. Gen. Virol.* **73**:413–416.