The Transcription Profile of the Bocavirus Bovine Parvovirus Is Unlike Those of Previously Characterized Parvoviruses

Jianming Qiu,¹* Fang Cheng,¹ F. Brent Johnson,² and David Pintel³

Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Kansas City, Kansas; Department of Microbiology and Molecular Biology, Brigham Young University, Provo, Utah; and Department of Molecular Microbiology and Immunology, Life Sciences Center, University of Missouri–Columbia, Columbia, Missouri

Received 16 April 2007/Accepted 13 August 2007

The Bocavirus bovine parvovirus generated a single pre-mRNA from a promoter at its left-hand end; however, the pattern of its alternative polyadenylation and splicing was different from that of other parvoviruses. A large left-hand-end open reading frame (ORF) encoded a nonstructural protein of approximately 95 kDa. An abundant, spliced, internally polyadenylated transcript encoded the viral NP1 protein from an ORF in the center of the genome. Transcripts encoding the capsid proteins were polyadenylated in the right-hand terminal palindrome. This is the first published transcription map of a member of the Bocavirus genus of the Parvovirinae.

The genus Bocavirus is one of five genera of the subfamily Parvovirinae of the family Paroviridae (35). The currently recognized members of the Bocavirus genus include bovine parvovirus type 1 (BPV) (1), canine minute virus (32), and the recently identified human bocavirus (HBoV) (4). HBoV, first detected in pooled human samples of lower respiratory tract infections collected by PCR amplification in Sweden (4), has been reported to be associated with acute respiratory illness at incidence rates of between 1.5% and 11.3% worldwide (3, 5, 6, 11–13, 15, 19, 22–24, 34). Isolation of HBoV has not yet been reported.

A transcription map of any Bocavirus species based on a nonbiased direct analysis of steady-state RNA has not yet been formally presented. In this report, we describe the transcription map of BPV following infection of permissive bovine turbinate (BT) cells as determined by RNase protection and Northern blotting assays. As members of the genera Erythroiovirus and Amdovirus of the Parvovirinae, the Bocavirus BPV has a single promoter at its left-hand end and uses both an internal and distal polyadenylation site. However, the alternative RNA processing strategy used by BPV generates a transcription profile that is different from that of other characterized parvoviruses.

Sequencing of BPV and construction of a nearly full-length BPV clone. Because a previously constructed infectious clone of BPV (33) was no longer available, we resynthesized the majority of the BPV genome, using a PCR-based strategy, directly from plaque-purified BPV (the prototype Abinanti strain) (1), utilizing the published BPV sequence (GenBank accession no. NC_001540) to design primers. The revised sequence of BPV was deposited in GenBank, and all the nucleotide numbers in the manuscript refer to this revised sequence, unless otherwise specified. This nearly full-length clone, containing the revised BPV sequence (nucleotide [nt] 42 to 5515) in the pBluescript vector (Stratagene), was designated pSK42BPV. This clone was not infectious (data not shown).

Our revised BPV sequence showed a number of silent mutations which differed from that of the sequence deposited previously, and these differences may reflect natural variations among isolates. However, in addition, we detected two more significant changes within the NS1 coding region which allowed the opening of the long NS open reading frame 1 (ORF 1) from nt 336 to 2918 (Fig. 1A), which in the previously deposited sequence (NC_001540) was unexpectedly terminated (Fig. 1B). In the previously determined sequence, the presence of an additional A residue at nt 507, which we did not detect, shifted the NS1 ORF 1 to ORF 2, which terminated at nt 614. The G residue at nt 778 in the previously determined sequence, which was not present in our analysis, allowed the reopening of ORF 1 from an initiating AUG codon at nt 740 (Fig. 1B). This prompted previous investigators to propose, based on its loose homology to a TATA signal, a promoter at map unit (m.u.) 13 (P13) (10), thus generating an mRNA that would encode this portion of ORF 1 (Fig. 1B, LT ORF). This previously proposed LT ORF lacked nt 336 to 783 (NC_001540) of the potential NS1 coding sequence and would be predicted to generate a protein of approximately 83 kDa (10, 17). Our current sequence determination allowed ORF 1 to be open in the NS1 region from nt 336 to 2918 (Fig. 1A), thus allowing the generation of a larger NS1 protein of approximately 95 kDa. Such a protein was detected in some earlier studies (16), and when we tagged ORF 1 in pSK42BPV with the hemagglutinin (HA) epitope, a protein of approximately 95 to 100 kDa was detected following transfection of 293 cells (Fig. 1C, lane 1). We did not detect a promoter in the vicinity of P13, either by 5′ rapid amplification of cDNA ends (RACE) analysis (data not shown) or by RNase protection assays (RPAs) (data not shown), as described more fully below.

Transcription profile of BPV RNA. To determine an accurate transcription map of the Bocavirus BPV, we first analyzed the general overall profile by Northern analysis, as described

* Corresponding author. Mailing address: Department of Microbiology, Molecular Genetics and Immunology, University of Kansas Medical Center, Mail Stop 3029, 3901 Rainbow Blvd., Kansas City, KS 66160. Phone: (913) 588-4329. Fax: (913) 588-7295. E-mail: jqiu@kumc.edu.

© 2007, American Society for Microbiology. All Rights Reserved.
A Corrected sequence (DQ335247)

<table>
<thead>
<tr>
<th>nt</th>
<th>Codon</th>
<th>ATG</th>
<th>AUG</th>
<th>ORF1 (NS1)</th>
<th>336</th>
<th>398</th>
</tr>
</thead>
<tbody>
<tr>
<td>336</td>
<td>AUG</td>
<td>336</td>
<td>2918</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B Previous sequence (NC_001540)

<table>
<thead>
<tr>
<th>nt</th>
<th>Codon</th>
<th>ATG</th>
<th>AUG</th>
<th>ORF2 (NS1)</th>
<th>336</th>
<th>1644</th>
</tr>
</thead>
<tbody>
<tr>
<td>336</td>
<td>AUG</td>
<td>336</td>
<td>2918</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C

### FIG. 1. Sequencing and detection of the BPV NS1 ORF.

(A and B) Redetermination of the nucleotide sequence of the BPV genome revealed two changes in the putative NS1 ORF that allowed its opening from nt 336 to 2918. The NS1 ORFs proposed by both the sequence determined in this study (accession no. DQ335247) (A) and the previous sequence (accession no. NC_001540) (B) are diagramed, with the start and stop codons indicated by nucleotide numbers. The putative P13 promoter proposed previously is also shown. As described in the text, residues A and G at nt 507 and 778, respectively, were identified in the previous sequence (that of NC_001540) but not in the sequence from the current study (DQ335247). The consequences of these insertions on the integrity of the NS1 ORF are discussed in the text. (C) Western blotting analysis of HA-tagged BPV proteins. Protein samples were isolated 48 h following transfection of 293 cells with constructs pSK42BPVNP1HA and pSK42BPVNP1HA, in which an HA tag was inserted into the parent pSKBPV42 construct in the amino terminus of NS1 ORF and the carboxyl terminus of NP1 ORF, respectively, following the method previously reported (28). Protein samples were run on sodium dodecyl sulfate-10% polyacrylamide gels, and Western blotting was performed using a monoclonal antibody against the HA tag (HA-7; Sigma). Lane 1, HA-tagged NS1 at a size of approximately 95 kDa; lane 2, HA-tagged NP1 at a size of approximately 28 kDa. A protein band of the size expected for a putative NS2 protein (53 kDa) was not detected in lane 2. This may have been due to low abundance of the R3 transcripts generated by transfection in 293 cells. As mentioned in the text, we have noticed some differences in the abundance of BPV RNAs generated following transfection compared to that following viral infection (Fig. 2B).
FIG. 2. Transcription profile of BPV RNA. (A) Northern blotting analysis. The BPV type 1 (Abinanti strain) (1) was used to infect BT cells (ATCC CRL 1390) at ~0.002 focus forming units/cell (to limit defective interfering particles). A total of 10 μg of total RNA isolated 4 to 6 days...
above, all RNA initiation and termination sites, all splice junctions, and their relative uses were confirmed by quantitative Rpas (data not shown). Significantly, as mentioned above, neither 5′ RACE analysis (data not shown) nor RPAs (data not shown) gave evidence of an internal promoter in the BPV genome. Together, these studies led to the transcription map presented in Fig. 3.

Characterization of the BPV polyadenylation sites. There are four core poly(A) motifs (AATAAA) in the BPV genome at m.u. 60.0, 61.6, 64.6, and 98.5 (10). We employed RPAs utilizing three antisense BPV probes [P(pA)p1, P(pA)p2, and P(pA)p3] to determine if an internal polyadenylation site was used (Fig. 3, diagrams shown at the left). Probe P(pA)p1, which overlaps the first poly(A) signal at m.u. 60, protected a major band at 118 nt, representing approximately 90% of the RNA in this area, and two minor bands at approximately 90 and 80 nt, representing ~3% and ~9%, respectively, of the total RNA (Fig. 2B, lane 2). The 80-nt band mapped the internal cleavage sites to approximately nt 3319. This cleavage site was also detected by 3′ RACE analysis (data not shown). The 90-nt minor band may represent an additional downstream cleavage event, but this site was not confirmed by 3′ RACE analysis and so may be a nonspecific digestion band. No additional bands, other than read-through bands, were protected by probes P(pA)p2 and P(pA)p3, which span the potential poly(A) signals at m.u. 61.6 and 64.6 (Fig. 2B, lanes 3 and 4), suggesting that these two poly(A) signals were not used. To confirm the site of internal polyadenylation, probe P(pA)p1 was also used to protect RNAs generated by transfection of the noninfectious pSK42BPV plasmid in BT cells (Fig. 2B, lane 1). Bands of a size similar to those obtained during infection were detected (Fig. 2B, compare lane 1 to lane 2). Surprisingly, approximately 90% of the RNAs generated during transfection of the nonreplicating clone were polyadenylated at the (pA)p site, while less than 10% of the RNA was polyadenylated at this site during BPV infection. Similar differences in alternative polyadenylation profiles have been reported in permissove versus nonpermissive systems of B19 infection (20) and may be an important facet of gene regulation in paroviruses utilizing an internal polyadenylation site.

Interestingly, the only poly(A) signal at the right-hand end of BPV available for RNAs transcribed from the capsid gene ORF is located at nt 5401 in the right-hand hairpin terminus (10). The right-hand (5′) end of the BPV minus strand contains both flip and flop sequence orientations at equal frequencies (9). RPAs with probes P(pA)dFlip and P(pA)dFlop, spanning both the flip and flop ends (Fig. 2C, diagrammed to the right), reflected a relative ratio of “flop” RNA versus “flip” RNA of approximately 2:1 and all were efficiently polyadenylated at sites approximately 22 nt downstream of the AAUAAA site (Fig. 2C). Thus, the efficiency of 3′ end formation of the capsid-encoding RNAs was not affected by the two different termini configurations.

A genetic map of BPV that summarizes the results of the RNA mapping experiments described above is shown in Fig. 3, together with the previously determined profile of protein expression. All the transcripts generated by BPV were generated from a single promoter located at the left-hand end of the genome. There is an internal polyadenylation site in the center of the genome, and the capsid-coding mRNAs, which are generated by splicing of the large intron within the NS1 coding region, must read though this (pA)p site to access the capsid-coding ORF in the right-hand end of the genome. These features are similar to the transcription profiles of the erythrovi- ruses and amroviruses (21, 26, 28) and underscore the importance of the regulation of alternative polyadenylation in the expression of parovirus genomes.

Our current analysis suggests that BPV NS1 is encoded by the left-hand ORF, open from nt 336 to 2918, likely by either the R1a or the R1b mRNAs, which theoretically would encode a protein of approximately 95 kDa, larger than that typically seen for the large NS/Rep proteins of paroviruses. BPV has been previously shown to encode an abundant 28-kDa protein termed NP1 (18). The putative NP1 ORF would be predicted to be expressed from the R4a/b mRNA, and our tagging of this ORF in the pSK42BPV plasmid allowed detection of an abundant 28-kDa protein following transfection of 293 cells (Fig. 1C, lane 2). Multiple bands centering at approximately 45 kDa were also previously observed in BPV-infected cell lysates fol- lowing immunoprecipitation using convalescent-phase sera from BPV-infected calves (16); these may be candidates for the viral NS2 protein. These would be predicted to be encoded by the R3 mRNAs. Interestingly, during BPV infection of BT cells, the RNAs that generate the viral nonstructural proteins...
are present at levels significantly lower than those seen for the nonstructural protein-encoding RNAs produced by minute virus of mice (27), adeno-associated virus (7, 30), Aleutian mink disease virus (2, 28), and B19 virus (26).

Three capsid proteins, VP1 (80 kDa), VP2 (72 kDa), and VP3 (62 kDa), were previously detected, both following BPV infection and from purified virus (14, 16). Surprisingly, in the BPV transcription profiles we present here, we found evidence of RNAs that would be predicted to encode only capsid proteins of 72 kDa and 62 kDa (from the R5a and R5b RNAs, respectively). The phospholipase A2 motif, which is present in the N terminus of the VP1 of virtually all parvoviruses (37), however, is located in the N terminus of the putative ORF of R5a/b (nt 3282 to 5305). Thus, it may be possible that the R5a/b mRNAs actually encode a capsid protein (VP1) migrating unexpectedly at 80 kDa rather than the predicted 72 kDa, and the previously identified 72-kDa capsid protein could perhaps be a cleavage product of the 80-kDa protein. How the BPV viral capsid proteins are generated is currently under investigation.

Our characterization of the Bocavirus BPV type 1 transcription profile also provides insights into the genetic map of the human bocavirus HBOV. Our preliminary characterization of the transcription map of HBOV following transfection of a nearly full-length HBOV plasmid in 293 cells showed it to be very similar to that of BPV (J. Qiu and D. Pintel, unpublished).

Nucleotide sequence accession number. The revised sequence of BPV was deposited in GenBank under accession no. DQ335247.

This work was supported by PHS grant P20 RR016443 from the NCRR COBRE program, grant R56 AI070723 from NIAID, a grant from the Children’s Miracle Network of the University of Missouri Children’s Hospital to J.Q., and PHS grants RO1 AI46458, RO1 AI56310, and RO1 AI21302 from NIAID to D.J.P.

REFERENCES