JOURNAL OF VIROLOGY, Feb. 1994, p. 1021–1028
0022-538X/94/$04.00+0

Vol. 68, No. 2

Transcriptional Mapping of the 3′ End of the Bovine Syncytial Virus Genome

RANDALL W. RENSHAW AND JAMES W. CASEY*

Department of Microbiology, Immunology, and Parasitology, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853

Received 19 August 1993/Accepted 12 November 1993

The bovine syncytial virus, a member of the retroviral subfamily *Spumavirinae*, causes a persistent, asymptomatic infection in cattle. Nucleotide sequence analysis of the viral genome revealed two overlapping reading frames in the 3′ region, traditionally occupied by accessory-function genes in other complex retroviruses. In order to analyze the transcripts from the accessory-gene region, we designed oligonucleotide primers complementary to sequences within the 5′ and 3′ long terminal repeats (LTRs) for use with the PCR. Southern blot analysis of amplification products revealed eight major cDNA bands. Eleven distinct cDNA clones were subsequently isolated and characterized. The initial splice donor in each clone is located 49 bp downstream from the mRNA cap site in the 5′ LTR. The primary splice acceptor site was located 17 bp upstream from the proximal 3′ open reading frame known as BF-ORF1. A second major splice acceptor was localized to a region upstream of the second open reading frame, BF-ORF2. Clones were identified which spliced directly to each of these sites. Additional splice donor and acceptor sites within BF-ORF1 and BF-ORF2 and the 3′ LTR were variously used to generate a complex array of multiply spliced transcripts. Each of these transcripts remained in frame and coded for a potential protein product.

The spumaviruses are a subfamily of the *Retroviridae* which includes complex retroviruses of human, simian, feline, and bovine origin. Nucleotide sequences of the genomes of three primate foamy viruses, the human foamy virus (HFV) and the simian foamy virus (SFV) serotypes 1 and 3 (SFV1 and SFV3), have been obtained (3, 14, 16, 20, 24). All spumaviruses examined to date have large genomes of 12 to 13 kb with multiple open reading frames (ORFs) in the 3′ segment of the coding region, an arrangement reminiscent of the accessory regulatory genes of the lentiviruses and the human T-cell leukemia virus-bovine leukemia virus oncovirus subgroup (24, 27, 28, 35, 39). The study of gene expression in this region is essential to the understanding of the regulatory mechanisms which govern replication and latency. The three ORFs that have been identified in the HFV accessory-gene region (accessory region) have been named the *bel* genes (3). SFV1 and SFV3 each contain only two ORFs in the accessory region (19, 24), ORF-1 and ORF-2. The *bel1* gene product and the SFV1 ORF-1 product, *taf*, have been shown to encode transcriptional transactivators which act at sequences within the U3 region of the 5′ long terminal repeat (LTR) in their respective viruses (8, 12, 17, 19, 21, 28, 29, 37). No functions have been identified for the products of *bel2*, *bel3*, or ORF-2.

An analysis of the splicing patterns of the accessory region of HFV revealed a complex splicing pattern giving rise to transcripts capable of generating several protein products (22). In addition to the singly spliced *bel1* and *bel2* mRNAs, multiply spliced mRNAs which combine portions of the *bel1* and *bel2* ORFs (*bet*) and segments of the *bel1* and *bel3* ORFs (*bel3* and *beo*) were identified. Immunological analysis has revealed that the multiply spliced *bet* product is the predominant protein in cytopathic infections (23), although the *bel1* and *bel2* products are found at significant levels (13). The function of the Bet protein has not yet been determined.

This study was designed to examine the structure and splicing patterns of the accessory region of the bovine syncytial virus (BSV) and to identify structural similarities with HFV and SFV1 and SFV3. We have designed oligonucleotide primers complementary to sequences within the BSV LTR R and U3 regions for use with the PCR to obtain cDNAs representing spliced messages from the 3′ end of the genome. The nucleotide sequence of the accessory region and an analysis of the potential products from the region are presented.

MATERIALS AND METHODS

Virus and cell cultures. The BSV isolate used in this study was derived from a culture transfected with the BSV-11 clone (25, 26) propagated in the canine thymus cell line Cf2Th (American Type Culture Collection, Rockville, Md.) in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum at 37°C with 5% CO_2. Cells were harvested in an early stage of cytopathic infection, as evidenced by approximately 25% syncytium formation.

PCR. Oligonucleotide primers used for PCR include an upstream LTR R region primer beginning 25 bp downstream from the transcriptional initiation site at nucleotide (nt) 1007 (25), 5′ TGAGAGATCTGACTTCCAG 3′, and a downstream U3 primer beginning at nt 584 in the LTR, 5′ GACCTC GAACCCACACAGG 3′. RNA from cytopathically infected cells was purified by the method of Chomczynski and Sacchi (2), and oligo(dT) selected (30). Prior to reverse transcription, random oligonucleotide primers (2.5 μM) were annealed to 5 μg of poly(A)$^+$ RNA by heating to 65°C for 1 min and cooling on ice. The 20-μl reverse transcriptase reaction mixture included 10 μl of the RNA-primer in a mix containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 5 mM $MgCl_2$, 1 mM each deoxynucleoside triphosphate, 5 mM dithiothreitol, 40 U of RNasin (Promega, Madison, Wis.), and 20 U of murine leukemia virus reverse transcriptase. The reaction was allowed

* Corresponding author. Mailing address: Department of Microbiology, Immunology, and Parasitology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853. Phone: (607) 253-3579. Fax: (607) 253-3419.

to proceed for 45 min at 42°C, and products were used directly for PCR without further purification according to the protocol supplied by the manufacturer (Perkin-Elmer Cetus, Norwalk, Conn.). The most efficient reactions were obtained by using a buffer containing 1 mM $MgCl_2$. Thermal cycling conditions were 94°C for 1 min, 61°C for 1 min, and 74°C for 3 min, through a total of 30 cycles.

Recombinant plasmids. The subclones used for sequencing of BSV genomic DNA included a 1.35-kb *Bam*HI fragment located immediately upstream of the 3′ LTR and an adjacent 0.9-kb *Bam*HI fragment encompassing the U3 region of the 3′ LTR (25, 26). Both fragments were cloned into the pBluescript SK^- vector (Stratagene, La Jolla, Calif.). PCR products were cloned by using the pCR1000 vector according to the protocol supplied by the manufacturer (Invitrogen, San Diego, Calif.) and analyzed by standard techniques for restriction endonuclease mapping and agarose gel electrophoresis (30).

Sequence analysis. Nucleotide sequences were determined for both strands by the dideoxy chain termination method with double-stranded plasmid DNA and modified T7 polymerase (U.S. Biochemicals, Cleveland, Ohio). Primers used for sequencing included the T7 and T3 promoter primers for initiation within the plasmid and synthetic oligonucleotide primers complementary to regions within the insert DNA. Computer analysis was accomplished with software (version 7.0) from the Genetics Computer Group (GCG), University of Wisconsin, Madison.

Nucleotide sequence accession numbers. The sequence of the BSV accessory region has been deposited in GenBank under the accession number L26493. The nucleotide sequence of the BSV LTR has the accession number L10921.

RESULTS

Structure of the BSV accessory region. Sequence analysis of genomic DNA from the 3′ accessory region of BSV revealed two ORFs, tentatively designated BF-ORF1 and BF-ORF2. Figure 1 shows the 1,922-bp sequence from this region and the predicted amino acid products beginning at a *Bam*HI site upstream of the accessory region and continuing into the 3′ LTR. Overlapping the 5′ end of BF-ORF1 is a portion of the putative *env* gene which includes a highly hydrophobic transmembrane sequence (39). Further downstream, BF-ORF1 overlaps the 5′ end of BF-ORF2, which continues into the 3′ LTR, in a manner similar to that of the 3′ ORFs in SFV1 and SFV3 (19, 24). A third ORF corresponding to the *bel3* gene of HFV (3) was not identified.

A comparison of the predicted amino acid sequence of BF-ORF1 with those of the *bel1* and *taf* ORFs shows only a few conserved regions (Fig. 2). This is perhaps not surprising, as HFV *bel1* and SFV1 *taf* were found to be less than 39% similar (19) and they are expected to be more closely related to each other than to BSV. The BF-ORF1 product is considerably shorter than those of *taf* or *bel1*, with 249 amino acids compared with 304 and 300 for *taf* and *bel1*, respectively. In addition, the BF-ORF1 sequence lacks the acidic region found at the amino terminus of those of *taf* and *bel1*, and an internal basic region, but does share two conserved cysteine residues with them. Independent comparisons of BF-ORF1 with *taf* and *bel1* indicated 20.9 and 23.0% similarities, respectively, based on complete identity of conserved residues (data not shown).

BF-ORF2 and the corresponding primate spumavirus ORFs share more significant regions of similarity (Fig. 3). Independent comparisons of products of BF-ORF2 and the analogous ORFs revealed 25.4% similarity with SFV1 ORF-2 and 26.6% similarity with *bel2* (data not shown). In addition to short stretches of conserved residues, three cysteine residues and a potential phosphorylation site are conserved.

PCR amplification of the 3′ accessory transcripts. In all retroviruses, transcription initiates at the beginning of the R region in the 5′ LTR and terminates at the end of R in the 3′ LTR with generation of subgenomic transcripts by internal splicing. Each transcript therefore retains terminal LTR sequences. Primer pairs in the R and U3 regions of the LTR were used to enzymatically amplify all transcripts, the only limitation being the length constraint imposed on chain extension by the reverse transcription reaction and PCR. The PCR product contained eight major bands, as resolved by agarose gel electrophoresis, in addition to several bands of lower intensity. All bands were shown to be BSV specific by hybridization to a probe from the accessory region (Fig. 4). The two largest bands of 1.8 and 1.5 kb are of the approximate size expected to represent the two major accessory transcripts identified previously (26). Eleven clones from the PCR products were selected for further examination on the basis of their sizes in relation to the PCR products.

Analysis of accessory-region clones. Nucleotide sequence analysis was used to identify the splicing patterns of each of the isolated clones. Figure 5 shows the splice donor and acceptor locations in relation to the genomic sequence; nucleotide sequences are listed in Table 1. In each clone, the initial splice donor falls at position +49 in the LTR, where +1 is the transcriptional initiation site at the beginning of R (25). This leader sequence, which is unusually short in comparison with those of other retroviruses, is nearly identical in length to that reported for HFV (22).

Most of the clones utilize a downstream splice acceptor found at nt 181, where +1 is defined as the 5′ nucleotide of the genomic DNA fragment (Fig. 1). Clones K63 and K60 do not utilize the nt-181 acceptor site but instead splice to alternative acceptors at nt 607 and 1550. Clone K7 is unusual in that it contains an additional segment of genetic material from an upstream region of the genome. Assignment of a precise location to this upstream segment awaits further sequencing, but by analogy to HFV, which follows a similar splicing pattern (22), it can be tentatively assumed to be from a region between the *pol* and *env* genes. This is supported by the presence of a polypurine tract (PPT) identical to the PPT immediately upstream of the 3′ LTR (Fig. 1), a sequence motif used by all retroviruses for the initiation of second-strand DNA synthesis (33, 39). An additional PPT is found at the *pol-env* junction of all spumaviruses and some lentiviruses (5, 11, 31, 34).

An unusual finding of this analysis is the presence of splice donors and acceptors within the 3′ LTR of BSV. Sites at this location have not been reported for HFV or SFV1 and SFV3. Both the donor and acceptor sites which are involved in this splice junction differ significantly from the consensus sequences (32) (Table 1).

The ORF structures for each of the PCR clones are shown in Fig. 5. Each ORF contains at least a portion of either BF-ORF1 or BF-ORF2 in the appropriate reading frame. The translational initiator for BF-ORF1 has the consensus sequence A/GNNATGG (10), while the initiator for BF-ORF2 differs from the consensus by a single nucleotide (GG GATGA). Only one of the predicted peptide sequences does not initiate at either the BF-ORF1 or BF-ORF2 initiator. Translation of the K7 transcript is potentially initiated from a consensus site (AGGATGG) and contains 50 amino acids from the upstream region described above, followed by 5 additional amino acids and a segment of BF-ORF1.

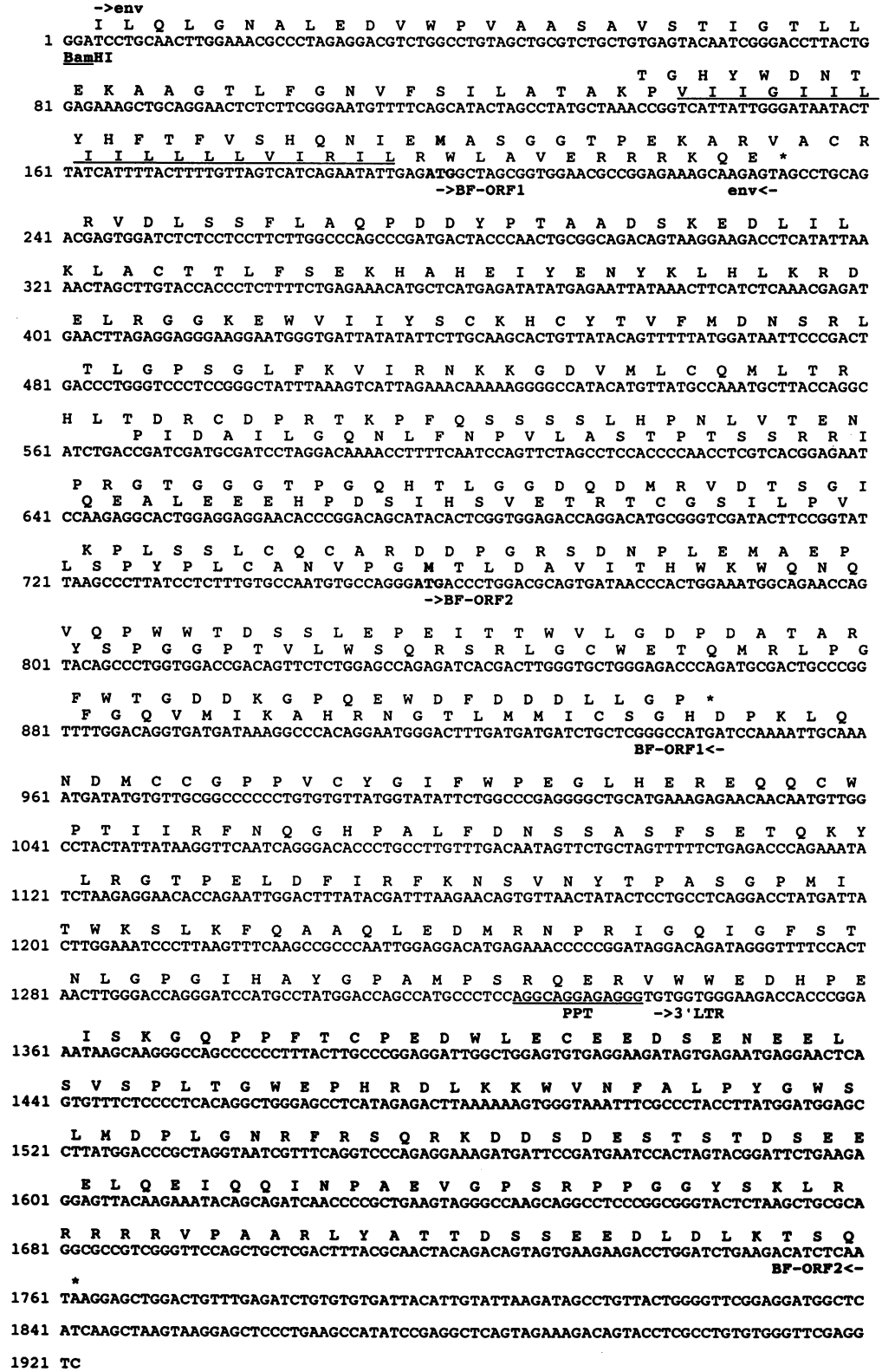

FIG. 1. Nucleotide sequence and amino acid composition of the accessory region of BSV. The predicted amino acid sequence for the 3′ end of the *env* gene overlaps BF-ORF1 and contains a hydrophobic transmembrane sequence (underlined). Putative initiation codons for BF-ORF1 and BF-ORF2 are indicated in boldface. The boundary of the 3′ LTR is shown at position 1338, and the adjacent PPT is underlined. Termination codons are indicated (*).

```
          1
BF-ORF1   ..........................................................................
HFV-Bell  MDSYEKEESV.............ASTSGIQDLQTLSELVGPENAGEGELTIAEEPEENPRRPRRYTKREV
SFV3-ORF1 MASWEKEKELAHL..HQPEDDPLPDLSLLLDMDQFEPTEGPDSNPGAEKIYLQ.LQVAPGDPSEKTYK..
SFV1-Taf  MASWEAQEELRELLHHLPEDDPPADLTHLLELDEMEPKVLCGENPGDEKLKKQVIKTPPMHPSTVTWH..

          71
BF-ORF1   ..MASGGTPEKARVACRRVDLSSFLAQPDDYPTAADSKEDLILKLACTTLFSEKHAHEIYE.NYKLHLKR
HFV-Bell  KCVSYHAYKEIEDKHPQHIKLQDWIPTPEEMSKSLCKR..LIL....CGLYSAEKASEILRMPFTVSWEQ
SFV3-ORF1 .....FGYEDKEAQNPD.LKMRNWVPDPEKMSKWACAR..LIL....CGLYNAKKAKELLKMDYDIHWEQ
SFV1-Taf  .....FGYKQKEDQQDN.IKMRDWVPNPSKMSKSTCKR..LIL....LGLYQACKAQEIIKMNYDVHWEK
            ++  +  +       ++ +  +++ *  ++     +  ***     +*+ + +* *++   + + +

          141
BF-ORF1   DELRGGKEWVIIYSCKHCYTVFMDNSRLTLGPS.GLFKVIRNKKGDVMLCQMLTRHLTDRC.........
HFV-Bell  SDT.DPDCFIVSYTCIFCDAVIHDPMPIRWDPEVGIWVKYKPLRGIVGSAVFIMHKHQRNCSLVKPSTSC
SFV3-ORF1 SKE.DSQYFEIEYHCKMCMTVIHEPMPVSYDKKTGLWIKMGPLRGDIGSVVHTCRRHYERCMSALPSS..
SFV1-Taf  SVV.NEQYFEVEYNCKMCRTVLHEPMPIMYDPETELWVKPGRLRGPLGSAVYTLKKHYERCLLTLPSL..
                + + * *  *  *+ +   + + +   ++      +* +       ++    *

          211
BF-ORF1   ......DPRTKPFQSSSSLHPNLVTENPRGTGGGTPGQHTLGGDQDMRVDTSGI.......KPLSSLCQC
HFV-Bell  SEGPKPRPRHDPVLRCDMFEKHHKPRQKRPRR.RSIDNESCASSSDTMANEPGSLCTNPLWNPGPLLSGL
SFV3-ORF1 GEPLKPRVRANPVRRYREKSLIVADRPKRSRWGVAPREQPNTSSGDAMALMPGPCGPFNMDPPGCLLERV
SFV1-Taf  KGTRLPKRRCNPSRRYETFREHPPTRKRRSKEGIPTDQQPSTSNGDPMALLSGPCGPHSIQPPGCLLQEL
            *  *                  *+  + +  + + ++  *     +*         *   *

          281
BF-ORF1   ARDDPGRSDNPLEMAEPVQPWWTDSSLEPEITTWVLGDPDATARFWTGDDKGPQEWDFDDDLLGP*
HFV-Bell  LEESSNLPN..LEVHMSGGPFWEE..........VYGDSILGPPSGSGEHSVL*............
SFV3-ORF1 PGSEPGTSE..MALAMSGGPFWEQ..........VYRDSISGPPTGPSEN*...............
SFV1-Taf  PKPEVGSPE..MAVAMSGGPFWEE..........VYGDSIFATPLGSSEDQLLSQFD*........
            +  + +  +  *+* +          *+ *+  ++   +++
```

FIG. 2. Comparison of the predicted amino acid sequences of the BSV BF-ORF1 product, HFV Bel1, the SFV3 ORF1 product, and SFV1 Taf. Alignment was obtained by using the Pileup program (GCG). Consensus amino acids are indicated with an *. Amino acids which are similar on the basis of side-group type (neutral and weakly hydrophobic, P, A, G, S, and T; hydrophilic acidic, Q, N, E, and D; hydrophilic basic, H, K, and R; hydrophobic, L, I, V, M, F, Y, and W) are indicated by +.

The singly spliced K69 and K63 clones code for the complete BF-ORF1 and BF-ORF2 products. The predicted K52 product contains 34 amino acids from the amino terminus of BF-OFR1 followed by 50 additional amino acids and then continues into BF-ORF2. The splicing pattern of K69, K63, and K52 is comparable to that of the *bel1*, *bel2*, and *bet* products of HFV (22). K35 is the only clone, with the exception of K63, for which translation may initiate at the BF-ORF2 initiator (although internal initiation is probably infrequent) and represents a truncated version of the BF-ORF2 product, where the carboxy-terminal 98 amino acids have been replaced with 13 amino acids from a downstream reading frame. The multiply spliced K76, K95, and K59 clones all splice to nt 181 and code for the first 34 amino acids from BF-ORF1 and subsequently splice to acceptors at nt 607, or 754 in the case of K95. Like K35, the K76 and K95 BF-ORF2 regions are truncated at the carboxy terminus. K59 terminates shortly after splicing out at a position 46 amino acids into BF-ORF2. Predicted coding regions are not described for K55, K50, or K60 as these clones encode only very short products, or no product at all in the case of K60.

DISCUSSION

The organization of the 3′ end of the BSV genome is like that of the other complex retroviruses in that it contains multiple ORFs with overlapping junctions and the U3 region is used as part of an ORF. In comparison with the other spumaviruses which have been examined, the overall structure of BSV is most like that of SFV1 and SFV3. A third ORF corresponding to *bel3* in HFV was not identified, and examination of the alternative reading frames in the region did not reveal any significant homologies. Nevertheless, all four spumaviruses are clearly structurally related by the organization of the ORFs in the accessory region. The similarities include a short 10- to 13-amino-acid overlap between the *env* gene and the first ORF, overlapping first and second ORFs, and terminal ORFs which extend well into (85 to 140 amino acids into) the U3 region.

A more detailed examination at the amino acid level suggests that there is only marginal similarity between BSV and the primate foamy virus accessory proteins. However, while BF-ORF1 is only 21 to 23% similar to the *bel1* and *taf* genes by conservation of residues, it is 42% related on the basis of the similarity of amino acid side groups (data not shown). Likewise, a comparison of BF-ORF2 and the counterpart primate virus genes indicates that they are 51 to 52% related on the basis of side-group type, suggesting that the structures of these proteins may be more highly conserved than is readily apparent.

Both Bel1 and Taf have been shown to be necessary and sufficient for transcriptional transactivation (19, 21, 28). Both

```
          1
BF-ORF2   .....................................................MTLDAVITHWKW.QNQYSPGG
HFV-Bel2  .....................................................MSHVLPVVTPWPMSQDHYAPTL
SFV3-ORF2 MRDVCLPFLARENLSNPESGLILLEDTERSHSSLRIGQNAPDGVWPLGNSPILPVVTPWPLCQDHAAPSI
SFV1-ORF2 MNDACLLFPASKELDSQNVAVILADDMKHSESILQLGSGAPRKGFPLTSSPLLPMVTPWPFSQDHAAPTL
                                                                + +*** * + *+  +**

          71
BF-ORF2   PTVL.....................WSQRSRLGCWETQMRLPGFGQVMIKAHRNGTLMMICSGHDP....
HFV-Bel2  FGILDRYYQGYLKSPATYQTWKFTCQVDPSGKRFMETQFWVPPLGQVNIQFYKNYQILTCCQAVDPFANI
SFV3-ORF2 WTLLDAYWRGYQDQNLEPPKWLWLCLEDPSGNKYTGTQFLVPPLGLVKIRLYQNLTVVYICQSIDPWENE
SFV1-ORF2 YSLLVAYYKSFQSQKLDPPKWLWQCLGDPSGRKCMVTQFLLPPLGQVRISCYRNLTSIVICQAVDPWENN
           ++*                          + *       ***+ +**** * *     * + +  * + **

          141
BF-ORF2   ...........KLQNDMCCGPPVCYGIFWPEGL.HEREQQCWPTIIRFNQGHP....ALFDNSSASFSET
HFV-Bel2  FHGTDEEMFDIDSGPDVWCSPSLCFKVIYEGAMGQKQEQKTW..LCRLGHGHRMGACDYRKVDLYAMRQG
SFV3-ORF2 NPTGGRRDPTRRYGCRIACDPVYCVKIVWEGNLWDKKDQPCW..LIRLKEGHNHGAKELSQRDIKILGES
SFV1-ORF2 NEADWRKNPMAR..PRIKCDHALCFKVVYEGTPWRPHDQKCW..LIRLTEGHKYGMEELSPGDWKILQES
                       + *   +*+ +++        +*  *  + *+  **      +       + ++

          211
BF-ORF2   QKYLRG....TPELDF.IRFKNSVNYTPASGPMITWKSLKFQAAQLEDMRNPRIGQIGFSTNLGPGIHAY
HFV-Bel2  KENPYGDRGDAA.LQYAYQVKRGCKAGCLASPVLNYKALQFHRTIMADFTNPRIGEGHLAHGYQAAMEAY
SFV3-ORF2 RPYPYGLIGQCPKLQYAIQVKMRVDKAPLTAKVLAVKALHFHRWNICQRENPGIGEGYFPSGYTQALKAY
SFV1-ORF2 RPYPYGPIGKDPNLQYAVGVKMKVIGGPLTSTVLALKALSFHRVNICNMDNPSLGEGHAPLGYSHALKAY
               *     + *++ + +*     +  ++ ++ +*+* *    +    ** +*+   +  +  ++ **

          281
BF-ORF2   GPAMPSRQERVWWEDHPEISKGQPPFTCPEDWLECEEDSEN.......EELSVSPLTGWEPHRDLKKWVN
HFV-Bel2  GPQRGSNEERVWWNVTRNQGKQGGEYYREGGEEPHYPNTPAPHRRTWDERHKVLKLSSFATPSDIQRWAT
SFV3-ORF2 GPQHGSAEQRVWLISTKIVGPQEKDYWRDAYRWGYFPLVPNKHHPGWTRHLTKFKISRFATPADIQKIVD
SFV1-ORF2 GPQYGSCEERVWQTATKCIGPEEENYWCEYDHRGFFPMVPNKLSPTWVRHAAPYCIQRFATPYDLQYFAN
          **   * +++**       +     +                                  +  + +  *+  +

          351
BF-ORF2   FALPYGWSLMDPLGNRFRSQRKDDSDESTSTDSEEELQEIQQINPAEVGPSRPPGGYSKLRRRRRVPAAR
HFV-Bel2  KALPYGWKVVTESGNDYTSRRKIRTLTEMTQDEIRKRWESGYCDPFIDSGSDSDGPF*
SFV3-ORF2 ELLPRGASIVTADGNRYESTRKVHLVNEGTLEEYQAKIRK*
SFV1-ORF2 ELLPPGFSITTPKGVSYTSDRRLHYGNEGTLQEYNENCDKVKRGYDEISSSDYSDED*
            ** *  +    *  + * *+       + +                   ++*

          421
BF-ORF2   LYATTDSSEEDLDLKTSQ*
```

FIG. 3. Comparison of the predicted amino acid sequences of the BSV ORF2 product, HFV Bel2, the SFV3 ORF2 product, and the SFV1 ORF2 product. The alignment was created by using the Pileup program (GCG). Consensus amino acids are marked with an *. Amino acids which are similar on the basis of side-group type, as described in the legend to Fig. 2, are indicated by +. A conserved protein kinase C phosporylation site (9) is overlined.

contain nuclear localization signals in the form of highly basic stretches of amino acids (6, 18, 38), and Bel1 has been localized in the nucleus of infected cells (13). BSV has also been shown to possess transactivating activity (25), and comparison of the BF-ORF1 product with Taf and Bel1 suggests that it is a likely candidate for the BSV transcriptional transactivator; however, the absence of an internal basic domain indicates that transport to the nucleus might not occur via the same mechanism.

It has been suggested that strongly acidic regions found at the amino terminus of SFV1 Taf are potential activation domains by analogy to other acidic viral products which interact with components of transcription initiation complexes (18, 19). Analysis of the acidic amino-terminal domain of HFV Bel1 indicated that this region is dispensable for transactivation (6, 38). The absence of this acidic domain in BSV supports the latter findings.

Recent findings have indicated that the carboxy-terminal regions of Bel1 and Taf contain a primary activation domain that is highly conserved (Fig. 2, amino acids 296 to 324) among the primate viruses (4, 6, 18, 36, 38). The BF-ORF1 product shows only limited similarity with this conserved region. The highest degree of similarity is found in two regions, amino acids 111 to 131 and 149 to 185 (Fig. 2). The first region corresponds to domains shown to be essential to transactivation in Taf and Bel1 (18, 38), although significant gaps have been introduced to achieve this alignment. The second region corresponds to a part of the minimal activator region of Bel1 (36, 38).

The products of BF-ORF2 and the corresponding SFV and

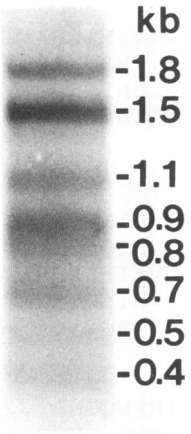

FIG. 4. Southern blot of PCR products following reverse transcription of mRNA from BSV-infected Cf2Th cells. Hybridization was performed as previously reported with a 1.35-kb *Bam*HI fragment encompassing the accessory region exclusive of the 3′ LTR (25).

HFV ORFs share more short regions of homology than are observed between BF-ORF1 and its counterparts. Most prominent among these are two internal regions, from amino acids 107 to 136 and 261 to 292 in Fig. 3. The latter area contains a highly conserved string of amino acids. A data base (SwissProt; Protein Identification Resource) search for the peptide string GPXXXSXXER revealed similarity with only two proteins, a rat GTP binding protein (Rab1B) and, surprisingly, the bovine leukemia virus transcriptional transactivator (Tax). Comparison of the amino acid sequences of the BF-ORF2 product and BLV Tax (Gap program; GCG) did not reveal any additional similarities.

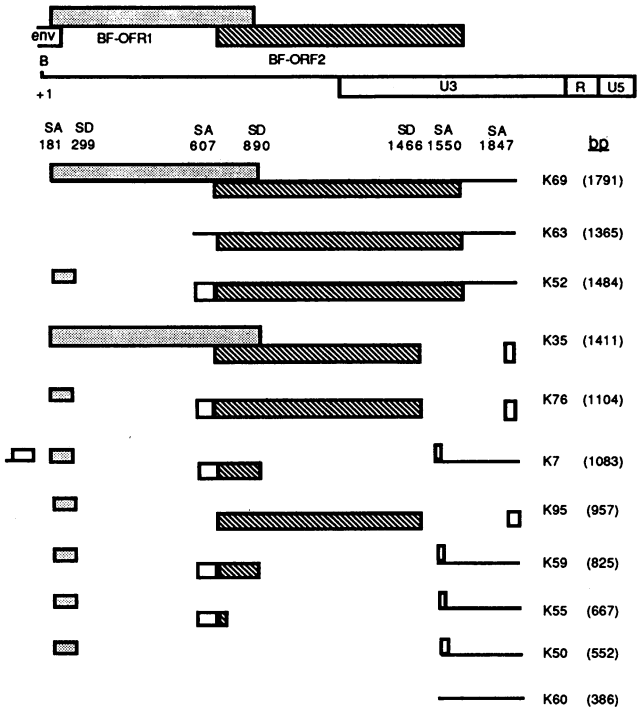

FIG. 5. Diagram showing the ORF arrangement, splicing patterns, and coding regions in the 11 clones examined. The *Bam*HI site indicated by B defines position 1 for this analysis. Nucleotide positions for the major splice donor (SD) and splice acceptor (SA) sites are indicated. Solid lines indicate the RNA segments, and boxes indicate coding regions (stippled for BF-ORF1 and hatched for BF-ORF2). Unshaded boxes indicate regions where amino acids not within the bounds of the initiator and termination codons for either of the two ORFs are used.

No functions have been identified for the products from any of the second ORFs in the spumavirus accessory region. Analysis with antibodies specific to Bel2 has revealed that it is predominantly localized to the cytoplasm (13), suggesting that it may not be involved in transcriptional regulation or mRNA splicing and transport. A recent effort to define the function of the non-Bel1 accessory genes was unsuccessful (1). Mutants lacking the ability to direct production of Bel2 or Bet were not limited in their replication, perhaps suggesting a more defined in vivo role for these proteins.

The K7 product is of special interest, as it represents a potential addition to the coding regions at the 3′ end of the BSV genome. The corresponding region in HFV, named S1, was originally suggested to be an ORF (15), but this finding was later revised (22). The HFV S1 intron is, however, spliced into downstream accessory transcripts but does not contain an initiation codon. As the function of this region is presently unknown, its significance cannot be expanded upon, but it does call into question the necessity for the conservation of splice donor and acceptor sequences in corresponding regions of two separate viruses.

Another observation that has been difficult to explain is the apparent use of unconventional splice donors and acceptors within the 3′ LTR of BSV. As reviewed by Jackson (7), the use of nonconsensus splice sites is rare and there are only a few acceptable base substitutions in the consensus sequences. The splice junctions in the 3′ LTR of BSV do not follow the conventions. In particular, the 5′ splice donor does not possess the invariant GT dinucleotide at the 5′ end of the intron. However, the presence of this unconventional junction in three independent clones (K35, K76, and K95) argues against a cloning artifact. An alternative explanation for the use of unconventional sites is the possibility that the clones in question were generated from deletion mutants. BSV genomes with deletions in the accessory region have been identified previously (26). Further study will be needed to characterize these clones before any conclusions can be drawn concerning their function in viral replication.

The finding of numerous potential products resulting from multiple splicing events at the 3′ end of BSV adds significant complications to the assignment of biological functions. It is difficult to assume that the singly spliced BF-ORF1 and BF-ORF2 products are the only proteins required for biological activity. Indeed, Western blot (immunoblot) analysis of cytopathic infections of HFV has shown that the doubly spliced Bet is the predominant protein, with Bel1 and Bel2 being found in significantly lower concentrations (23). Currently, the only clue to the use of BSV accessory proteins can be drawn from the size relationships between the transcripts observed in cytopathic infections (26) and the PCR products. The K69, K63, K52, and K35 clones all fall within the range expected to represent the major accessory transcripts.

In summary, the potential for complex gene expression patterns in the accessory region of BSV is apparent. The conservation and utilization of conventional and perhaps unconventional splice donors and acceptors in this region suggest that some gene products arise from multiply spliced transcripts. This level of complexity suggests distinct functional

TABLE 1. Locations and sequences of splice donor and acceptor sites in the BSV accessory region

Splice donor or acceptor	Position[a]	Clone(s)
Donors		
NA$^{A}_{G}$/GTAAGT		Consensus[b]
GTG/GTAAGA	+49	All
ACA/GTAAGG	299	K52, K76, K7, K95, K59, K55, K50
CCG/GTATTA	716	K55
CAG/GTGATG	890	K7, K59
AGC/CTCATA	1466	K35, K76, K95
Acceptors		
CT$^{TT}_{AC}$TNNTTTT$^{TT}_{CC}$NCAG/G		Consensus
CATTTTACTTTTGTTAG/T	181	All except K63 and K60
AAACCTTTTCAATCCAG/T	607	K63, K52, K76, K7, K59, K55
TTGTGCCAATGTGCCAG/G	754	K95
CCTTATGGACCCGCTAG/G	1537	K55
CTAGGTAATCGTTTCAG/G	1550	K7, K59, K50, K60
GAGGATGGCTCATCAAG/C	1847	K35, K76, K95

[a] All positions are based upon numbering beginning at the *Bam*HI site at the 5′ end of this region (Fig. 1) except for the initial splice donor, which is +49 in the 5′ R region (25).

[b] Consensus sequences for eukaryotic viruses are from the work of Shapiro and Senapathy (32).

roles for several products and leaves open the possibility that because of the absence of specific domains, some may regulate through competitive mechanisms.

ACKNOWLEDGMENTS

This work was supported by a grant to J.W.C. from the Wetterberg Foundation. R.W.R. was supported by a National Needs fellowship, USDA 87-GRAD-9-0093.

REFERENCES

1. **Baunach, G., B. Maurer, H. Hahn, M. Kranz, and A. Rethwilm.** 1993. Functional analysis of human foamy virus accessory reading frames. J. Virol. **67**:5411–5418.
2. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
3. **Flugel, R., A. Rethwilm, B. Maurer, and G. Darai.** 1987. Nucleotide sequence analysis of the *env* gene and its flanking regions of the human spumaretrovirus reveals two novel genes. EMBO J. **6**:2077–2084.
4. **Garrett, E. D., F. He, H. P. Bogerd, and B. R. Cullen.** 1993. Transcriptional *trans* activators of human and simian foamy viruses contain a small, highly conserved activation domain. J. Virol. **67**:6824–6827.
5. **Harris, J. D., J. V. Scott, B. Traynor, M. Brahic, L. Stowring, P. Ventura, A. T. Haase, and R. Peluso.** 1981. Visna virus DNA: discovery of a novel gapped structure. Virology **113**:573–583.
6. **He, F., J. D. Sun, E. D. Garrett, and B. R. Cullen.** 1993. Functional organization of the Bel-1 *trans* activator of human foamy virus. J. Virol. **67**:1896–1904.
7. **Jackson, I.** 1991. A reappraisal of non-consensus mRNA splice sites. Nucleic Acids Res. **19**:3795–3798.
8. **Keller, A., E. D. Garrett, and B. R. Cullen.** 1992. The Bel-1 protein of human foamy virus activates human immunodeficiency virus type 1 gene expression via a novel DNA target site. J. Virol. **66**:3946–3949.
9. **Kishimoto, A., K. Nishiyama, H. Nakanishi, Y. Uratsuji, H. Nomura, Y. Takeyama, and Y. Nishizuka.** 1985. Studies on the phosphorylation of myelin basic protein by protein kinase C and adenosine 3′:5′-monophosphate dependent protein kinase. J. Biol. Chem. **260**:12492–12499.
10. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
11. **Kupiec, J., J. Tobaly-Tapiero, M. Canivet, M. Santillana-Hayat, R. M. Flugel, J. Peries, and R. Emanoil-Ravier.** 1988. Evidence for a gapped linear duplex DNA intermediate in the replicative cycle of human and simian spumaviruses. Nucleic Acids Res. **16**:9557–9565.
12. **Lee, A. W., K. J. Lee, S. Kim, and Y. C. Sung.** 1992. Transactivation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by the human foamy virus *bel1* protein requires a specific DNA sequence. J. Virol. **66**:3236–3240.
13. **Lochelt, M., H. Zentgraf, and R. M. Flugel.** 1991. Construction of an infectious DNA clone of the full-length human spumaretrovirus genome and mutagenesis of the *bel1* gene. Virology **184**:43–54.
14. **Maurer, B., H. Bannert, G. Darai, and R. Flugel.** 1988. Analysis of the primary structure of the long terminal repeat and the *gag* and *pol* genes of the human spumaretrovirus. J. Virol. **62**:1590–1597.
15. **Maurer, B., and R. M. Flugel.** 1988. Genomic organization of the human spumaretrovirus and its relatedness to AIDS and other retroviruses. AIDS Res. Hum. Retroviruses **4**:467–473.
16. **Mergia, A., and P. Luciw.** 1991. Replication and regulation of primate foamy viruses. Virology **184**:475–482.
17. **Mergia, A., E. Pratt-Lowe, K. E. S. Shaw, L. W. Renshaw-Gegg, and P. A. Luciw.** 1992. *cis*-acting regulatory regions in the long terminal repeat of simian foamy virus type 1. J. Virol. **66**:251–257.
18. **Mergia, A., L. W. Renshaw-Gegg, M. W. Stout, R. Renne, and O. Hercehnroeder.** 1993. Functional domains of the simian foamy virus type 1 transcriptional transactivator (Taf). J. Virol. **67**:4598–4604.
19. **Mergia, A., K. Shaw, E. Pratt-Lowe, P. A. Barry, and P. A. Luciw.** 1991. Identification of the simian foamy virus transcriptional transactivator gene (*taf*). J. Virol. **65**:2903–2909.
20. **Mergia, A., K. E. S. Shaw, J. E. Lackner, and P. A. Luciw.** 1990. Relationship of the *env* genes and the endonuclease domain of *pol* genes of simian foamy virus type 1 and human foamy virus. J. Virol. **64**:406–410.
21. **Mergia, A., K. E. S. Shaw, E. Pratt-Lowe, P. A. Barry, and P. A. Luciw.** 1990. Simian foamy virus type 1 is a retrovirus which encodes a transcriptional transactivator. J. Virol. **64**:3598–3604.
22. **Muranyi, W., and R. M. Flugel.** 1991. Analysis of splicing patterns of human spumaretrovirus by polymerase chain reaction reveals complex RNA structures. J. Virol. **65**:727–735.
23. **Netzer, K., A. Rethwilm, B. Maurer, and V. ter Meulen.** 1990. Identification of the major immunogenic structural proteins of human foamy virus. J. Gen. Virol. **71**:1237–1241.
24. **Renne, R., E. Friedle, M. Schweizer, U. Fleps, R. Turek, and D. Neumann-Haefelin.** 1992. Genomic organization and expression of simian foamy virus type 3 (SFV-3). Virology **186**:597–608.
25. **Renshaw, R. W., and J. W. Casey.** Analysis of the 5′ long terminal

repeat of bovine syncytial virus. Gene, in press.
26. **Renshaw, R. W., M. A. Gonda, and J. W. Casey.** 1991. Structure and transcriptional status of bovine syncytial virus in cytopathic infections. Gene **105:**179–184.
27. **Rethwilm, A., G. Darai, A. Rosen, B. Maurer, and R. M. Flugel.** 1987. Molecular cloning of the genome of the human spumaretrovirus. Gene **59:**19–28.
28. **Rethwilm, A., O. Erlwein, G. Baunach, B. Maurer, and V. ter Meulen.** 1991. The transcriptional transactivator of human foamy virus maps to the *bel*1 genomic region. Proc. Natl. Acad. Sci. USA **88:**941–945.
29. **Rethwilm, A., K. Mori, B. Maurer, and V. ter Meulen.** 1990. Transacting transcriptional activation of human spumaretrovirus LTR in infected cells. Virology **175:**568–571.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Schweizer, M., R. Renne, and D. Neumann-Haefelin.** 1989. Structural analysis of proviral DNA in simian foamy virus (LK-3)-infected cells. Arch. Virol. **109:**103–114.
32. **Shapiro, M. B., and P. Senapathy.** 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. **15:**7155–7174.
33. **Smith, J. K., A. Cywinski, and J. M. Taylor.** 1984. Initiation of plus-strand DNA synthesis during reverse transcription of an avian retrovirus genome. J. Virol. **49:**200–204.
34. **Tobaly-Tapiero, J., J. Kupiec, M. Santillana-Hayat, M. Canivet, J. Peries, and R. Emaniol-Ravier.** 1991. Further characterization of the gapped DNA intermediates of human spumavirus: evidence for a dual initiation of plus-strand DNA synthesis. J. Gen. Virol. **72:**605–608.
35. **Tobaly-Tapiero, J., M. Santillana-Hayat, M. Giron, M. C. Guillemin, F. Rozain, J. Peries, and R. Emaniol-Ravier.** 1990. Molecular differences between two immunologically related spumaretroviruses: the human prototype HSRV and the chimpanzee isolate SFV6. AIDS Res. Hum. Retroviruses **6:**951–957.
36. **Venkatesh, L. K., and G. Chinnadurai.** 1993. The carboxy-terminal transcription enhancement region of the human spumaretrovirus transactivator contains discrete determinants of the activator function. J. Virol. **67:**3868–3876.
37. **Venkatesh, L. K., P. A. Theodorakis, and G. Chinnadurai.** 1991. Distinct *cis*-acting regions in U3 regulate trans-activation of the human spumaretrovirus long terminal repeat by the viral *bel1* gene product. Nucleic Acids Res. **19:**3661–3666.
38. **Venkatesh, L. K., C. Yang, P. A. Theodorakis, and G. Chinnadurai.** 1993. Functional dissection of the human spumaretrovirus transactivator identifies distinct classes of dominant-negative mutants. J. Virol. **67:**161–169.
39. **Weiss, R., N. Teich, H. Varmus, and J. Coffin.** 1984. RNA tumor viruses, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.