JOURNAL OF VIROLOGY, Dec. 1996, p. 9024–9027 Vol. 70, No. 12
0022-538X/96/$04.00+0

Characterization of the Spliced *pol* Transcript of Feline Foamy Virus: the Splice Acceptor Site of the *pol* Transcript Is Located in *gag* of Foamy Viruses

JOCHEN BODEM,[1] MARTIN LÖCHELT,[1] INGRID WINKLER,[2] ROBERT P. FLOWER,[2] HAJO DELIUS,[1] AND ROLF M. FLÜGEL[1]*

Abteilung Retrovirale Genexpression, Forschungsschwerpunkt Angewandte Tumorvirologie, Deutsches Krebsforschungszentrum, 69009 Heidelberg, Germany,[1] and School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, Australia[2]

Received 24 June 1996/Accepted 20 August 1996

Foamy viruses, or spumaviruses, are distinct members of the *Retroviridae*. Here we have characterized the long terminal repeat of the feline, or cat, foamy virus by determining the locations of the transcriptional start site and the poly(A) addition site. The splice donor and splice acceptor sites of the subgenomic mRNA responsible for Pro-Pol protein expression were identified by nucleotide sequencing of the corresponding cDNAs. The leader exon of the feline foamy virus is 57 nucleotides long. The splice acceptor of the subgenomic *pol* mRNA was found to be located in *gag*. The location of the splice acceptor of the human foamy virus *pol* mRNA was confirmed to map in *gag*. The *pol* splice acceptor site in *gag* of the cat foamy virus is located further downstream than that of human foamy virus.

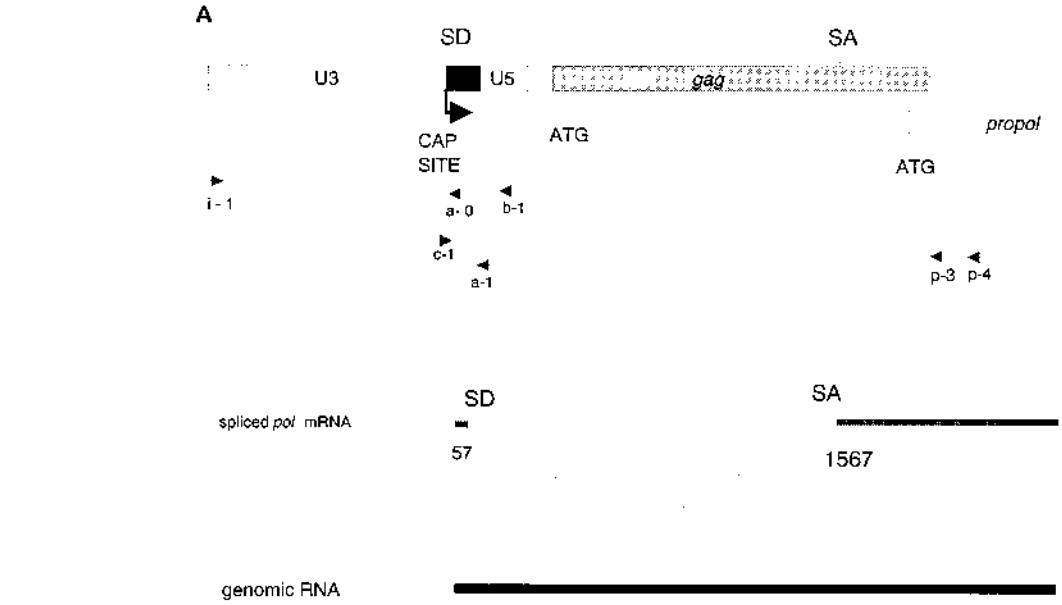
Foamy viruses are unique among retroviruses in their mode of *pol* gene expression. The human foamy virus (HFV) Pro-Pol polyprotein is not synthesized as a Gag-Pol fusion protein by translational frameshifting or stop codon suppression as in other retroviruses analyzed (1, 2, 4, 10). Instead, the HFV Pro-Pol protein is expressed by translational initiation at the first Met residue in the *pol* gene independently of Gag protein expression (4). Recently, a spliced HFV *pol* mRNA was identified (10). The novel splice acceptor (SA) of the *pol* mRNA located in *gag* 492 nucleotides (nt) upstream of the Pro-Pol start codon is joined to the splice donor (SD) of the HFV leader exon. In addition, experiments with a mutant SA of the *pol* mRNA showed that this splice site is absolutely required for Pol expression and viral infectivity (10). The presence and utilization of a subgenomic spliced transcript for *pol* gene expression have not been observed for any other retrovirus analyzed to date (1).

In this report we characterize the long terminal repeat (LTR) by defining the precise locations of the cap site, the major SD site in the 5′ LTR, and the poly(A) addition site of the genome of the feline foamy virus (FeFV) isolate FUV (3). Importantly, a spliced *pol*-specific transcript was identified and characterized in FeFV-infected cells. In addition, we confirm that the SA of the corresponding HFV *pol* mRNA is located in *gag* as reported previously (10). Our data suggest that Pro-Pol protein expression of different foamy viruses is directed by a spliced *pol* transcript.

To analyze whether a spliced subgenomic *pol* mRNA is present in productively FeFV-infected cells, total RNAs were extracted from FeFV-infected Crandell feline kidney (CRFK) cells 3 days after infection. A Qiagen (Hilden, Germany) RNAeasy kit was used as described by the manufacturer. To prepare cDNA, about 20 μg of RNA was reverse transcribed (Superscript reverse transcriptase [RT]; Gibco BRL, Karlsruhe, Germany) at 42°C for 45 min with the antisense primer p-4 (5′-ATATACATCTCCTTCCTGCGTTCC-3′), which is located in the FeFV *pol* gene (Fig. 1A and 2B). The sequences of the primers used in this study were derived from partial FeFV genomic DNA clones. Reaction products of cDNA synthesis were amplified by PCR with sense primer c-1 (5′-TTGGCTGCGTCCAGGGTGAGATTGA-3′) (Fig. 2B), located in the predicted RNA leader exon, and antisense primer p-4 (Fig. 1A) under annealing conditions at 53°C for 30 s; the elongation time was 2 min and 20 s at 68°C for each cycle with the PCR Expand system, which was used as described by the manufacturer (Boehringer, Mannheim, Germany). The PCR was run for 30 cycles. Two PCR products with sizes of about 2,000 and 500 bp were obtained (Fig. 2A) and cloned into the *Srf*I site of the pCR-Script vector (Stratagene, Heidelberg, Germany). The FeFV insert of 2,000 bp represents cDNA of genomic origin. The nucleotide sequence of the 500-bp-long FeFV cDNA fragment is shown in Fig. 2B. The sequence identifies the FeFV *pol* mRNA splice site junction as 5′-GACTCTCCAGCTTGGAGAG-3′, which results from joining the splice donor, 5′-GACTCTCCAGCTTGG ↓ GTAA -3′ (the arrow marks the exon-intron border in Fig. 2C at position 57 downstream of the cap site), to the splice acceptor at position 1567 (5′-GTTCCCGAACAG ↓ AG-3′) (Fig. 2C). A comparison of the characteristics of the FeFV and HFV *gag* genes is shown in Table 1. This *pol* SA site is located in the 3′ part of the *gag* region of the FeFV DNA genome. This FeFV splice site sequence was determined from five independent recombinant plasmid clones. In a single case, a G instead of an A residue was found in the first position of the splice acceptor (5′-GACTCTCCAGCTTGGGGAG-3′). The sequences of the identified FeFV SA and SD sites correspond to those of established consensus sequences and those reported for HFV (Fig. 2C).

The spliced FeFV transcripts contain an AUG codon 299 nt downstream of the SA site that is located at the 5′ end of the uninterrupted *pol* reading frame. The length of the FeFV Pro-Pol protein is similar to that of HFV (Table 1). Its degree of

* Corresponding author. Mailing address: Abteilung Retrovirale Genexpression, Forschungsschwerpunkt Angewandte Tumorvirologie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 242, 69009 Heidelberg, Germany. Phone: 49-6221-424611. Fax: 49-6221-424865. Electronic mail address: R.M.Fluegel@dkfz-heidelberg.de.



B

```
                U3→
   1 TGTCATGGGCCAAAGAGAATTCTCACAGAGGAGAATACTCTCTGCTGCCATCTAGTGACGATGAGGAAGAAGAAATGTCAGAAAGAGAGGAATTATTGTG
 100 CCATATAAATCAGTGTCAACAAAAGCTCTTTTATCCCGGAGGGACGACTGATGTCCTTGGAATGGAAAGCAATGTTTGGCTCACTAAATTTGTTAATATT
 200 AAATTTCCTAAAGGAACAAAAGTGATACTTCCTGATGGAAGAAAATTCATAGCCTGTGATCCTGAGCTAAAACCATTATTGCAGGAATTGAAATTCTTGG
 300 ATAGGGCAACATCTGAGTCATCTGACTCTGAATAGAAAGCCTGAATTTACCTGGATTATGCAACTTTGTCCGAGGTGGCAGAGTGGTTATGTATCTGTCA
 400 TACTCGGGGAAAGTTTTGTCTTTACATGTTCAAGACATATAAAGGGTGGAAAAATATATTCCTGACTAAACTTCCTGGGGACTAGAGGTGTGGAAACTTT
 500 GCTGCCTCTGCTTCACGGGAAGCTTTTGGTTCGAATCCTTTTTTAGGTACTTAGTTAAGATAAGTAGTGAATAAATTACTCTCGTTCATGTATTCATATC
 600 GAAACTATGTATCCTTTAAAACCATGTATTCTTTAGTCATCTAGATACTTAGAGTATGAAAAAAGAAACTGCAATAGTAACTATCAATGTTAGTAAATAA
 700 AGTACAGCTTAGTCATCTGATGATGTCACGAGAAAAGAACCTAGAAGAGAAGAACAACTTTCGGCATGCAACAGAGCGGGAGCTTGGTGTAGGAGCTAAG
 800 TCACCGTCTTACATCTAGAGCCTACTCTTCTTGAACTGTTCGAATCCTATTTTTGGAACTCTTACATCACCTTTAAGAGACTGAAAAGCATGACTCGTGC
 900 ACAGGAAGCTCCTTTAGGGTAGAGGAAATGTTCTAATCTCCTATCTTAAAGGGTTGCTTCATTTAAGGTTCGAAACTGTGTACTGGAAGTAGATTTTGCA
                                                                  ←U3 | R→
1000 TAACTTTAAACTTTTAGTTGCATGTTTCTGCTATTAGCAGCATATAAAAGGGTTATGGTAGATTGTACGGGAGCTCTTCTCACAGACTTGGCTGCGTCCA
                    ↓
1100 GGGTGAGATTGAGACTCTCCAGCTTGGGTAAGATTTTGATATGTATTTTGCTTGAATATTATTTGCCTTGCTCAAAATTAAATAAATTGGCTTTTCTTTC
     ←R | U5→
1200 ACTCAATTGAAGCTTCATATAATTATATTATTGTCTGAAGCCAGAACTCACATGAGTGGTGTTTCTCTATTCTTGGGGAAAAGTGTTCTTCTATTTGAAA
                                                  ←U5 |      PBS
1300 GTGTTAGAGCTACTAAGTGAAGAACTAATCTATCCCAGGTATAGGCCACGACAGTTGGCGCCCAACGTGGGGC
```

FIG. 1. (A) Structure of the FeFV 5′ LTR and flanking sequences showing the locations of the SD and SA sites. The cap site is marked by a bent arrow. The locations and orientations of DNA primers used are indicated by arrows. The subgenomic spliced *pol* mRNA and the FeFV genomic RNA are symbolized by shaded and filled boxes, respectively. (B) Nucleotide sequences of the FeFV LTR and the adjacent primer binding site (PBS) (boxed). The borders of the U3, R, and U5 regions are indicated, the TATA motif is boxed and in italic type, the cap site defines the 5′ end of R (in boldface italic type), the major splice donor is marked with a vertical arrow, the poly(A) addition site is underlined and in italic type, and the poly(A) addition signal is double underlined.

homology to the HFV Pro-Pol sequence of 60.6% is higher than that of Gag, as expected. The location of the FeFV *pol* SA is different from but similar to that of the corresponding SA site of HFV *pol* mRNA (reference 10; see also Table 1). By analogy, we conclude that the spliced FeFV mRNA has the capacity to direct the synthesis of the FeFV Pro-Pol protein.

To map the transcriptional start site of LTR-derived FeFV mRNAs, cDNA synthesis was performed with antisense primer p-4 located in *pol*. Reactions were performed with Superscript RT (Gibco BRL) as described by the supplier and total RNAs obtained 5 days after infection of CRFK cells with FeFV. Reaction products were denatured and chilled on ice, and RNAs were digested with RNase A and *Escherichia coli* RNase H (15 min at 55°C). The cDNA was purified by chromatography (PCR purification kit; Qiagen), and this was followed by tailing with dC residues of the 3′ end of the cDNA by terminal deoxynucleotide transferase (5′-race kit; Gibco BRL). The dC-tailed cDNA was amplified by PCR with the oligo(dG) anchor sense primer complementary to the poly(C) tail and the antisense *pol* primer p-3 (5′-TTGCTGCCTAACAGGTTCTTCT CC-3′). PCR conditions were as described above. Major reaction products of about 2,000 and 500 nt were obtained and reamplified in a nested PCR with the oligo(dG) anchor sense primer and primer b-1 (5′-AACTGTCGTGGCCTATACCTG GG-3′) located downstream of the 5′ LTR SD (Fig. 1A). The resulting PCR product of about 350 nt was cloned into the

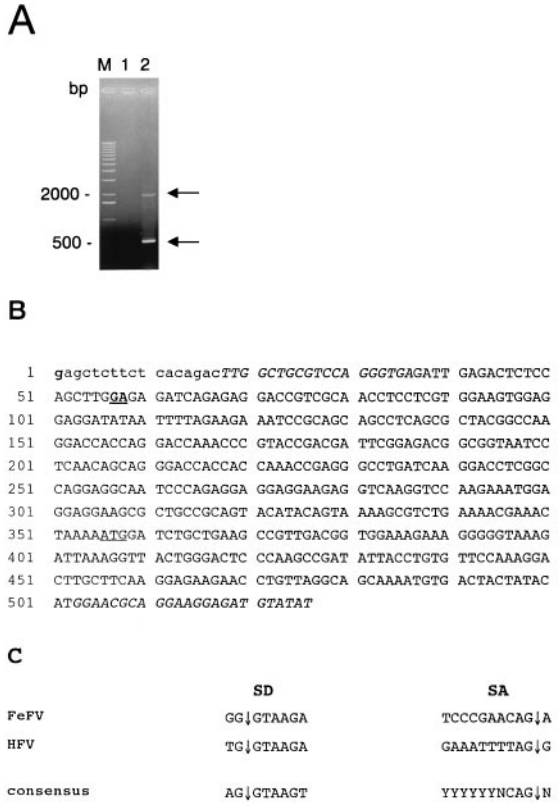


FIG. 2. PCR amplification and sequence analysis of the FeFV *pol* mRNA. (A) Agarose gel electrophoresis of FeFV cDNAs corresponding to genomic (upper arrow) and subgenomic *pol* (lower arrow) mRNA obtained by RT-PCR of total RNA from FeFV-infected CRFK cells (lane 2) but not from mock-infected cells (lane 1). Sizes of DNA markers (lane M, 1-kbp ladder; Gibco BRL) comigrating with the cDNAs of genomic (2,000 bp) and *pol* (500 bp) mRNAs are indicated. Minor PCR products obtained with RNAs from FeFV-infected cells did not hybridize to different FeFV genomic DNA probes (data not shown). (B) Nucleotide sequence of the 5′ part of the spliced FeFV *pol* mRNA. The splice site junction is in boldface type and underlined. Sequences corresponding to primers c-1 and p-4 are in italics. The first AUG codon of the *pol* reading frame is double underlined. The 5′-terminal nucleotides of the *pol* mRNA derived from the experiment shown in Fig. 3 are in the lowercase letters. The G base at position 1 is the cap site (boldface). (C) Comparison of the consensus splice site sequences with those of the FeFV and HFV SDs and the *pol* SA. Arrows mark the exon-intron borders.

pCRII vector (Invitrogen, Leek, The Netherlands). The DNAs of three independent clones were used for direct sequencing by PCR (Thermosequenase; Amersham, Braunschweig, Germany) with the 5′-terminally ^{32}P-labelled primer a-0 (5′-AAGCTGG AGAGTCTCAATCTCACCCTGGACGC-3′). The results of a sequencing reaction of three independent clones are shown in Fig. 3 with the deduced FeFV sequence of the sense strand in the 5′ to 3′ orientation (from top to bottom), starting at the marked G residue (see also Fig. 2B). The flanking sequences around this G residue are highly homologous to those of other foamy virus cap sites. Additionally, a TATA box consensus sequence is at an appropriate distance from this G residue (Fig. 1B). Primer extension experiments confirmed that the LTR transcriptional start site is the G nucleotide at proviral DNA position 1071 (Fig. 1B). Since the SD is located at nt position 57 downstream of the cap site, the FeFV mRNA leader exon is 57 nt in size, 6 nt larger than that of HFV (Table 1). The complete FeFV LTR sequence with the major landmarks is shown in Fig. 1B. The landmarks of both viral LTRs are compared in Table 1.

TABLE 1. Characteristics of FeFV and HFV genes

Genomic region	Size (bp) of indicated region in:		Homology (%) (FeFV/HFV)
	FeFV	HFV	
LTR	1,353	1,121	40.1
U3	1,070	777	39.6
R	134	190	57.0
U5	149	154	36.0
mRNA leader exon	57	51	
pol intron	1,151	1,796	
pol 5′ UTR[a]	355	543	
Gag	514[b]	648[b]	39.1
Pol	1,156[b]	1,143[b]	60.6

[a] 5′ UTR, 5′ untranslated region of the mRNA.
[b] Size in amino acid residues.

In order to determine the location of the poly(A) addition site of the spliced FeFV *pol* mRNA, RT-PCR experiments were carried out. cDNA synthesis was started with an oligo(dT)$_{15}$ primer and carried out with 20 μg of total cell RNA from CRFK cells harvested 5 days after FeFV infection. PCR amplifications were done with oligo(dT)$_{15}$ and the sense primer i-1 (5′-GGAAAGCAATGTTTGGCTCAC-3′) located at the 5′ end of the FeFV U3 region (Fig. 1A). Specific amplification products of about 1,200 bp in length were cloned into the pCRII plasmid (Invitrogen) and analyzed by DNA sequencing. We obtained four independent clones that terminated in the following FeFV cDNA sequence (in the sense orientation): 5′-AAAATTA**AATAAA**TTGGCTTTTCTTTCA CTCAAAAA$_n$-3′. The poly(A) addition signal is in boldface type, and the poly(A) tail is underlined. This result indicates

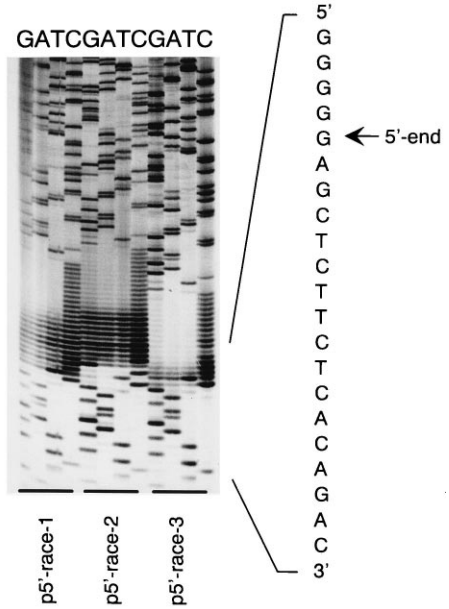


FIG. 3. Determination of the location of the FeFV cap site. cDNA molecules from RNA of FeFV-infected CRFK cells were cloned into the vector pCRII (clones p5′-race-1 to p5′-race-3) and directly sequenced by PCR with the antisense primer a-0 located directly upstream of the FeFV SD in the leader exon (Fig. 1). The dideoxy termination nucleotides (ddG, ddA, ddT, and ddC) used are indicated. The deduced DNA sequence of the sense strand, from top to bottom, of three independent clones is shown at the right margin. The sequence starts with G nucleotides derived from the 5′-race procedure. The 5′-terminal G residue of the FeFV mRNAs is marked.

that the poly(A) addition site is located 134 nt downstream of the cap site and defines the 134-nt-long R region of the FeFV LTR. The size of the FeFV R region is shorter than those of simian foamy viruses of type 1 and type 3, which are 171 and 217 nt long, respectively (7, 9).

To show that spliced FeFV *pol* transcripts terminate at this site in the 3′ LTR, cDNA synthesis was employed with antisense primer a-1 (5′-T_{15}GAGTGAAAGAAAAGCCAATTTATTT-3′), which extends from the poly(A) tail into the R region. cDNA synthesis with the poly(A) primer a-1 was performed with Expand RT (Boehringer) according to the instructions of the manufacturer. The reaction products were subjected to PCR amplification with sense primer c-1 and antisense primer p-4. The amplified cDNAs of 2,000 and 500 nt corresponded to the unspliced genomic and the spliced *pol* transcripts described above, since control amplifications employing RNAs that had not been reverse transcribed did not result in these reaction products. The identities of the cDNAs of 2,000 and 500 nt were proven by DNA hybridization with a FeFV *pol* cDNA probe, demonstrating that FeFV *pol* RNAs terminate at the poly(A) addition site in the 3′ LTR.

In parallel, cDNA products of approximately 8.5 and 7.5 kbp in size were amplified by PCR with a sense primer located in the leader exon and an antisense primer derived from the 5′ end of the 3′ LTR (data not shown). These products likely represent the unspliced genomic and the spliced *pol* mRNAs.

To determine the location of the HFV subgenomic *pol* mRNA SA site, total RNA was extracted from HFV-infected BHK cells 2 to 3 days postinfection; cDNA was prepared with the antisense primer 3198a (5′-TTGCCCCTGAATCCCAGTGGGC-3′) located downstream of the *pol* start codon (6). PCRs with sense primer 781s (5′-CTCTTCACTACTCGCTGCGTCG-3′) located in the HFV LTR leader exon and antisense primer 3198a were performed as described above, resulting in a 600-bp DNA product. The product was reamplified with sense primer 803s (5′-GCGGCTGTAACAGCTGAAGAGG-3′) and cloned into the pCRII vector. Sequences of four different recombinant clones were obtained, all of which contained the following sequence: 5′-AGAGTGTACGAGACTCTCCAGGTTT**GG**CCTTAATGCCAGAGGACAAAGTATAC-3′ (the splice site junction is in boldface type, and the nucleotides underlined are part of HFV leader RNA exon 1). In the HFV *pol* transcript the SD at position 51 (with nt 1 as the HFV cap site) is linked to the SA at position 1834. This result confirms the locations of the SD and SA sites as reported by Yu et al. (10). When total RNA extracted five days postinfection was used, a cDNA fragment of the HFV *pol* mRNA was detected after the first PCR round, in addition to products corresponding to genomic RNA. However, spliced transcripts resulting from cryptic splice sites were not found.

In this study it was shown that a subgenomic, spliced mRNA species that has the potential to direct the expression of the FeFV Pro-Pol polyprotein is present in productively FeFV-infected CRFK cells. This FeFV mRNA species seems to be the counterpart of the corresponding HFV transcript that directs *pol* gene expression (10). Furthermore, we confirm that this spliced *pol* transcript is present in HFV-infected cells. Because of the different RT-PCR conditions, this transcript was not detected in an earlier report (4).

Our data indicate that the utilization of a spliced *pol* mRNA is a general feature of spumavirus gene expression, since members of the *Spumavirinae* as distantly related as HFVs and FeFVs have similar mechanisms for Pro-Pol expression. This conclusion is supported by the fact that the first Met residue of the *pol* gene is conserved among the known foamy viruses. The first Met residue seems to function as the start codon of the spumaviral Pro-Pol polyprotein. Although the general mechanisms of Pro-Pol expression appear to be similar in HFVs and FeFVs, significant differences remain. It is noteworthy that the precise locations of the SA sites of the FeFV and HFV *pol* transcripts are different. The HFV SA is located 492 nt upstream of the *pol* AUG initiation codon, whereas the corresponding FeFV SA is located only 299 nt upstream of the *pol* gene. Whether these differences in the lengths of the resulting untranslated *pol* leaders (Table 1) influence the level of Pro-Pol expression is unknown.

The presence and utilization of a spliced *pol* transcript clearly set spumaviruses apart from other retroviruses, since the latter express the *pol* gene as a Gag-Pol fusion protein. It thus appears that foamy viruses (i) have developed basically different mechanisms of gene expression, i.e., a Gag-independent expression of *pol* by a spliced transcript, and (ii) use an active internal promoter that may also be used by other complex retroviruses as a pathway of viral gene expression (5, 8). Whether some of these features of foamy virus replication, e.g., the unique mechanism of Pro-Pol expression and synthesis of premature proviral DNA in virions (10), are functionally linked remains to be clarified.

Nucleotide sequence accession number. The accession number at the EMBL data bank of the nucleotide sequences reported here is X98741.

This work was supported by fellowships from the Wilhelm Sander Stiftung (J.B.) and the DAAD (I.W.).

REFERENCES

1. **Coffin, J. M.** 1990. Replication of retroviridae, p. 1437–1500. *In* B. N. Fields and D. M. Knipe (ed.), Virology. Raven Press, New York.
2. **Enssle, J., I. Jordan, B. Mauer, and A. Rethwilm.** 1996. Foamy virus reverse transcriptase is expressed independently from the Gag protein. Proc. Natl. Acad. Sci. USA **93:**4137–4141.
3. **Flower, R. L., G. E. Wilcox, R. D. Cook, and T. M. Ellis.** 1985. Detection and prevalence of serotypes of feline syncytial spumaviruses. Arch. Virol. **83:**53–63.
4. **Löchelt, M., and R. M. Flügel.** 1996. The human foamy virus *pol* gene is expressed as a Pro-Pol polyprotein and not as a Gag-Pol fusion protein. J. Virol. **70:**1033–1040.
5. **Löchelt, M., W. Muranyi, and R. M. Flügel.** 1993. Human foamy virus genome possesses an internal, Bel 1-dependent and functional promoter. Proc. Natl. Acad. Sci. USA **90:**7317–7321.
6. **Löchelt, M., H. Zentgraf, and R. M. Flügel.** 1991. Construction of an infectious DNA clone of the full-length human spumaretrovirus genome and mutagenesis of the *bel 1* gene. Virology **184:**43–54.
7. **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 activator. J. Virol. **64:**3598–3604.
8. **Nosaka, T., Y. Ariumi, M. Sakurai, R. Takeuchi, and M. Hatanaka.** 1993. Novel internal promoter/enhancer of HTLV-I for Tax expression. Nucleic Acids Res. **21:**5124–5129.
9. **Renne, R., E. Friedl, 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.
10. **Yu, S. F., D. Baldwin, S. R. Gwynn, S. Yendapalli, and M. L. Linial.** 1996. Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. Science **271:**1579–1582.