JOURNAL OF VIROLOGY, May 1985, p. 501–508
0022-538X/85/050501-08$02.00/0

Vol. 54, No. 2

Latent and Viral Replicative Transcription In Vivo from the *Bam*HI K Fragment of Epstein-Barr Virus DNA

RONALD WEIGEL[1] AND GEORGE MILLER[1,2,3]*

Departments of Molecular Biophysics and Biochemistry,[1] Pediatrics,[2] and Epidemiology and Public Health,[3] Yale University School of Medicine, New Haven, Connecticut 06510*

Received 22 October 1984/Accepted 11 January 1985

We mapped one latent and two replicative messages transcribed in vivo from the *Bam*HI K fragment of the Epstein-Barr virus genome. The exon encoding Epstein-Barr nuclear antigen (EBNA), a major latent product, is 2,028 bases; the 3′ end of this exon occurs 30 bases after the polyadenylation signal AATAAA, and the 5′ end occurs within a splice acceptor site. The open reading frame which encodes the EBNA peptide is completely contained within this coding exon. The exon was faithfully transcribed after transfection of cloned *Bam*HI-K into either COS-1 or TK$^-$ mouse L cells. In lymphocytes the abundance of the EBNA message is increased after cycloheximide treatment. The two viral replicative genes completely contained in *Bam*HI-K were not transcribed in line X50-7, in which the genome is tightly latent. In contrast to the EBNA message, these mRNAs of 1.3 and 2.1 kilobases are inducible with phorbol ester and are unspliced. Their promoter regions are similar to those of each other and to replicative promoters mapped in other regions of the Epstein-Barr virus genome (P. J. Farrell, A. Bankier, C. Seguin, P. Deininger, and B. G. Barrell, EMBO J. 2:1331–1338, 1983). An unusual feature of these replicative genes is that the smaller mRNA begins within a long open reading frame of the larger mRNA. The identification of the structure of latent and replicative genes within one DNA fragment will facilitate analysis of regulation of expression for the two life cycles of the virus.

Epstein-Barr nuclear antigen (EBNA) is expressed in all lymphocytes immortalized by Epstein-Barr virus (EBV) (17). Several viral antigens are serologically defined as EBNA (8, 12). They have characteristic appearances by immunofluorescence (8) and different molecular weights by immunoblot (12). One EBNA associates with chromosomes at metaphase (8).

Gene transfer has been used to localize the chromosome-binding nuclear antigen to the *Bam*HI K fragment of EBV (20). The left *Hin*dIII subfragment of *Bam*HI-K, I1f, is sufficient for EBNA expression (20). The I1f fragment contains a simple repeat of 680 base pairs (bp), IR3, which in one reading frame codes for a copolymer of alanine and glycine. Antibodies raised against this copolymer expressed in bacteria (12) or to peptides representing this region (4) react with an EBNA polypeptide of 78 kilodaltons. The size of the EBNA polypeptide varies among cell lines according to the size of the IR3 repeat present in the strain (7, 11).

The I1f fragment has been cloned on several different expression vectors (9, 18). Transfection of these plasmids into a variety of cell lines results in production of an immunogenic polypeptide of the same size as is identified in lymphocytes immortalized by EBV (7, 9). Further evidence that EBNA is virally encoded derives from analysis of expression of I1f in COS-1 cells with an undeleted DNA fragment and a fragment with a 600-bp deletion (18). In cells transfected with I1f there was synthesis of virus-specific mRNAs and full-sized EBNA. The deleted fragment yielded appropriately smaller message and polypeptide.

The mRNA with homology to the *Bam*HI K fragment found in latently infected lymphocytes contains an exon of 2 kilobases (kb) with homology to I1f (10). The abundance of this latent message is unchanged after induction of viral replication by phorbol esters or superinfection (23). Furthermore, there also are replicative genes in *Bam*HI-K (13, 19, 24). The entire DNA sequence of EBV has been determined, and open reading frames have been mapped on the genome including *Bam*HI-K (1). However, there has been no fine mapping of mRNAs in vivo from this important region of the genome.

* Corresponding author.

MATERIALS AND METHODS

Cells. We analyzed *Bam*HI-K transcription in three different lymphoid cell lines. Line X50-7 comprises human umbilical cord lymphocytes immortalized in vitro in which the EBV genome is tightly latent (14). Line FF500 also comprises human umbilical lymphocytes transformed in vitro; they spontaneously express replicative antigens at a low frequency (<1% of the cells). HH514-16 is a clone of the HR-1 line lacking heterogeneous DNA but hyperresponsive to induction with phorbol ester (16). *Bam*HI-K transcription was also analyzed in COS-1 cells transfected with pSV2-*Bam*HI-K and pSV2-I1fdel and in L cells stably transformed with *Bam*HI-K and herpes simplex virus thymidine kinase (HSV tk) gene (18, 20).

12-*O*-Tetradecanoylphorbol 13-acetate (TPA) was used to induce clone HH514-16 (26). Cells were treated for 72 h with TPA at 20 ng/ml before RNA isolation. In some experiments lymphoid cells were treated with 100 μg of cycloheximide per ml 6 h before RNA isolation (25).

DNA cloning. Plasmids were constructed by standard techniques of recombinant DNA (2). M13 clones MP8 and MP9 were propagated in JM103; M13 clone MP10W was propagated in JM105. Synthesis of radiolabeled complementary DNA strands from phage templates for use in S1 nuclease analysis was described previously (23).

RNA analysis. RNA was isolated as previously described (24). Northern blots were performed according to the procedure of Thomas (21). S1 nuclease analysis followed the procedure of Favaloro et al. (6).

Primer extension. Radiolabeled primer was excised from an acrylamide gel and annealed to 50 μg of total cytoplasmic

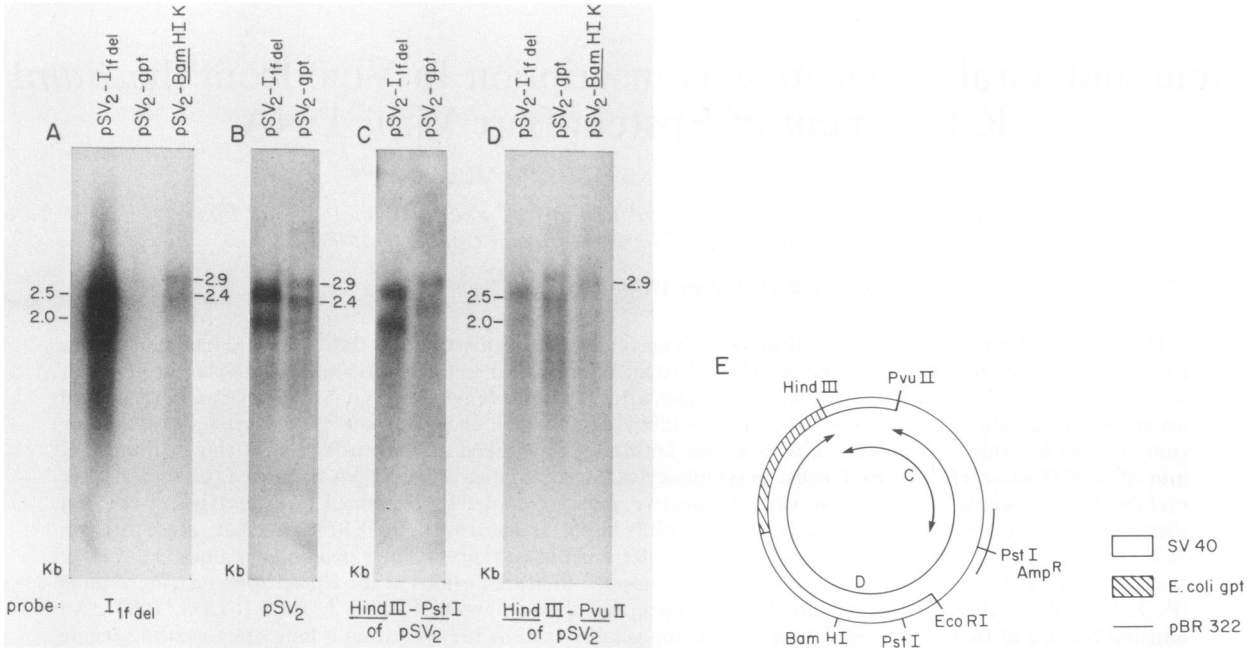

FIG. 1. Northern blots of messages synthesized in COS-1 cells transfected with pSV2 plasmids containing EBV inserts. (A) mRNA with homology to excised I1fdel synthesized in COS-1 cells transfected with pSV2-I1fdel and pSV2-*Bam*HI-K. No messages are identified with this probe in cells transfected with pSV2gpt. The mRNAs synthesized in COS-1 cells transfected with pSV2-I1fdel also have homology to vector pSV2gpt (B) and to the region of pSV2 containing the SV40 origin sequences (C). Also seen are two major messages synthesized in COS-1 cells transfected with pSV2gpt (B to D). The messages identified in panels A and B also have homology to a subfragment of pSV2 having the SV40 origin sequences removed as a *Hin*dIII-*Pvu*II fragment (D). A diagram of pSV2gpt showing the location of the vector fragments used as probes in panels C and D is shown in panel E.

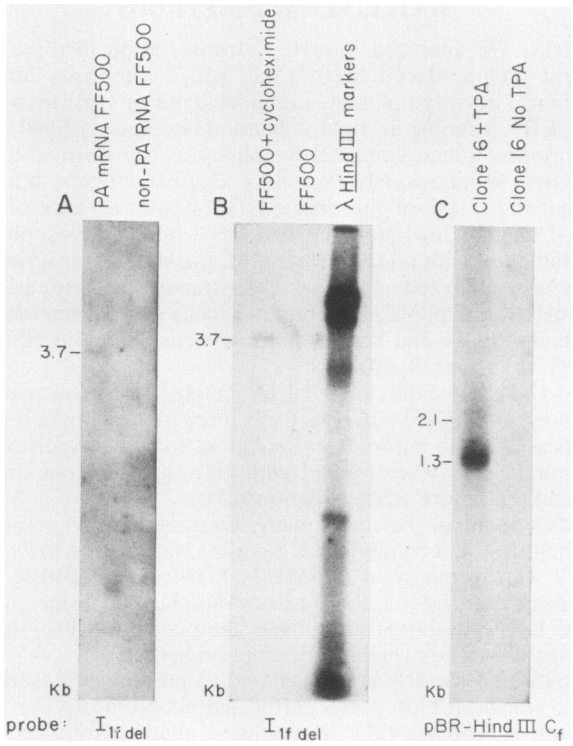

FIG. 2. Northern blots of lymphocyte messages. (A) Polyadenylated (PA) mRNA of 3.7 kb with homology to I1fdel-excised DNA. (B) Identical amounts of total cytoplasmic RNA from FF500-28 lymphocytes treated with cycloheximide or untreated were probed with I1fdel. This blot shows that the abundance of the 3.7-kb mRNA is increased with cycloheximide treatment. (C) mRNAs of 2.1 and 1.3 kb which were induced by TPA and which had homology to the right 1.7-kb pair of *Bam*HI-K (*Hin*dIII-Cf).

RNA under the same conditions as used for S1 analysis (6). Hybridizations were carried out at 55°C for 12 to 18 h. The hybrids were ethanol precipitated and suspended in 50 μl of reverse transcriptase buffer (140 mM KCl, 0.1 M Tris [pH 8.3], 10 mM $MgCl_2$, 1 mM deoxynucleoside triphosphate, 28 mM β-mercaptoethanol, 1 mM vanadyl ribonucleoside complex) and 40 U of avian myeloblastosis virus reverse transcriptase (Life Sciences, Inc., St. Petersburg, Fla.). Reactions were carried out at 42°C for 3 h, then stopped by the addition of 5 μl of 0.5 M EDTA and 25 μl of 150 mM NaOH, and incubated at 65°C for 1 h. The cDNA was ethanol precipitated and loaded on 8 M urea-acrylamide gels. Gels were rinsed, dried, and exposed to X-ray film (XAR-5; Eastman Kodak Co., Rochester, N.Y.) with or without intensifying screens.

RESULTS

Northern blot analysis of messages in COS-1 cells. We began our analysis of *Bam*HI-K transcription in COS-1 cells because the high abundance of mRNAs made in this gene transfer system facilitated the experiments. We wished to map the EBNA gene and to learn something of the structure of the mRNAs which gave rise to this product. As we previously reported, COS-1 cells which express EBNA as the result of gene transfer with pSV2-*Bam*HI-K contain two

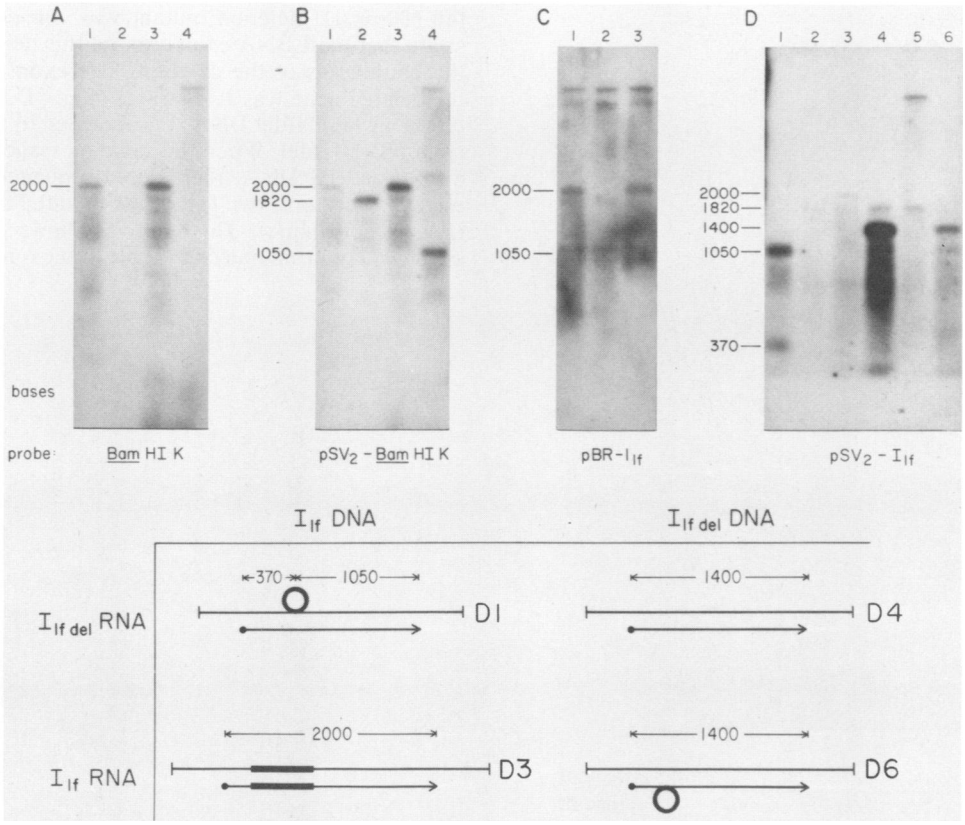

FIG. 3. S1 nuclease analysis of exons in I1f. Total cytoplasmic RNA from lymphocytes or COS-1 cells was annealed to a variety of cloned DNA fragments and digested with S1 nuclease. The protected DNA sequences were resolved on 1.5% alkaline agarose gels, blotted to nitrocellulose, and probed. Below the gel are schematic representations of the hybrids which form when RNA transcribed from I1fdel or I1f sequences is annealed to I1f or I1fdel DNA. The heavy lines represent sequences present in I1f which are deleted in I1fdel. (A) RNA from cells transfected with pSV2-*Bam*HI-K (lanes 1 and 3) or pSV2gpt (lanes 2 and 4) annealed to the plasmid pSV2-*Bam*HI-K (lanes 1 and 2) or pBR-I1f (lanes 3 and 4). When the filter was probed with excised *Bam*HI-K, a 2-kb exon became evident (lanes 1 and 3). This filter was rehybridized with pSV2-*Bam*HI-K (B), which demonstrated exons with homology to pSV2. (C) RNA from cells transfected with pSVOd-I1f (lane 1), pSVOd (lane 2), and FF500 lymphocytes (lane 3) annealed to the plasmid pBR-I1f. Again, a 2-kb exon is seen (lanes 1 and 3). (D) RNA from cells transfected with pSV2-I1fdel (lanes 1 and 4), pSV2gpt (lanes 2 and 5), or pSV2-*Bam*HI-K (lanes 3 and 6) annealed to the plasmid pSV2-*Bam*HI-K (lanes 1 to 3) or pSV2-I1fdel (lanes 4 to 6). In lane 1, two bands of 1,050 and 370 bases are seen; lanes 2 and 5 demonstrate the 1,820-base exon from pSV2, and lane 3 demonstrates the 2-kb exon from I1f. When the I1fdel DNA is used to protect the RNA (lanes 4 to 6), a 1.4-kb band (lanes 4 and 6) represents the sum of the two bands from lane 1. Also evident in these lanes is the 1,820-base exon from the vector, indicating that this exon is transcribed from the region of the ampicillinase gene.

mRNAs of 2.9 and 2.4 kb which have homology to I1f (Fig. 1A). Transfection of a plasmid which contains a deleted form of I1f gives rise to two smaller mRNAs of 2.5 and 2.0 kb. None of these messages has homology to the right 2.0-kb pair of *Bam*HI-K represented by *Hin*dIII fragments Nel and Cf (18).

The next experiments were done to determine whether messages in COS-1 cells contain sequences from the vector or come entirely from EBV DNA. Similar Northern blots were probed with the vector pSV2gpt and its subfragments (Fig. 1B to 1D). In COS-1 cells transfected with vector pSV2gpt there were two mRNAs of 2.9 and 2.4 kb with homology to pSV2. Furthermore, messages found in COS-1 cells transfected with pSV2-I1fdel or with pSV2-*Bam*HI-K also had homology to the vector. By probing with various subfragments of the vector we determined that these messages have homology both to the region of pSV2gpt containing the simian virus 40 (SV40) origin and to significant portions of pBR322. Consequently, the mRNAs synthesized in COS-1 cells from these constructs have sequences from both vector and EBV insert.

Northern blot analysis of mRNAs in lymphoid cells. In lymphocytes latently infected with EBV there was a 3.7-kb polyadenylated mRNA with homology to I1f (Fig. 2A) (22). Compared to COS-1 cells, the relatively low abundance of this mRNA in lymphocytes makes it difficult to analyze. However, the amount of the 3.7-kb mRNA can be significantly increased by cycloheximide treatment (Fig. 2B).

The *Bam*HI K fragment also encodes two mRNAs, 2.1 and 1.3 kb, which were induced in clone 16 HR-1 cells after treatment with TPA (Fig. 2C) (24). These mRNAs have homology to *Hin*dIII-Cf, the rightmost 1.7-kb pair of *Bam*HI K; messages of similar size are seen in X50-7 cells which have been superinfected with a virus which disrupts latency (23). These mRNAs can be identified in clone 16 cells on longer exposure of the autoradiograph before TPA treatment and are not seen in uninduced X50-7 cells (23). In both lymphoid cell systems the 1.3-kb message was the more abundant of the two messages seen after induction of replication (Fig. 2C).

S1 nuclease analysis of *Bam*HI-K exon in COS-1 cells and lymphocytes. We analyzed the full-sized exon synthesized

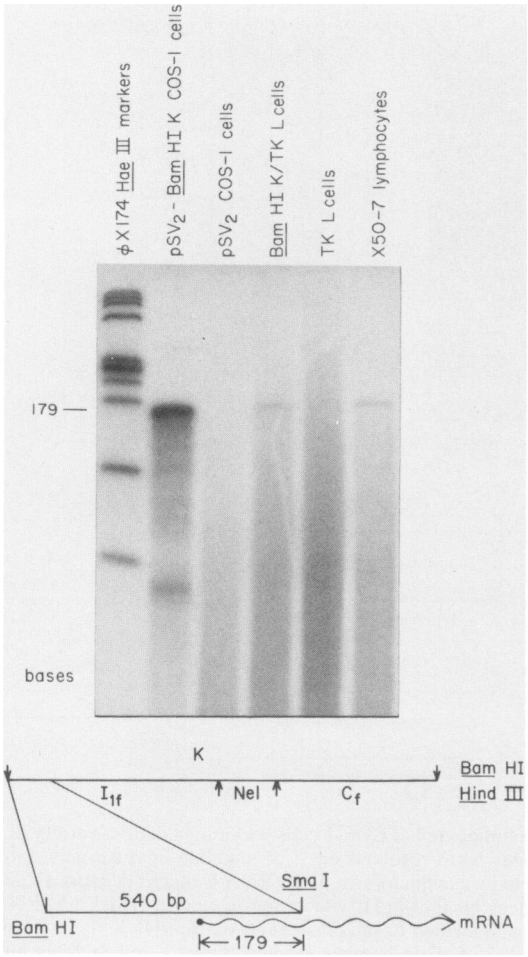

FIG. 4. S1 nuclease analysis of the 5′ end of the EBNA exon. A uniformly labeled DNA strand from the *Bam*HI-*Sma*I subfragment of *Bam*HI-K was hybridized to total cytoplasmic RNA from the cells indicated above the gel lanes. In all cells transfected with the *Bam*HI-K fragment and in EBV-infected lymphocytes, a 179-bp fragment was seen. This represents the 5′ end of the EBNA exon as diagrammed below the gel.

from the *Bam*HI K fragment. A 2.0-kb exon made in COS-1 cells transfected with pSV2-*Bam*HI-K was wholly contained in I1f (Fig. 3A). The same prominent 2.0-kb band was seen when RNA was annealed to *Bam*HI-K (Fig. 3A, lane 1) or to I1f (lane 3). We also saw a number of less prominent bands, all of which were smaller; the origin(s) of these fragments is not known. No exons homologous to *Bam*HI-K were identified when RNA from COS-1 cells transfected with pSV2gpt (lanes 2 and 4) was used. As a control, the same blot was rehybridized with a probe containing vector as well as *Bam*HI-K (Fig. 3B). Exons with homology to the vector were now seen in lanes 2 and 4. pSV2 protects an 1,820-base exon, and pBR322 protects a 1,050-base exon.

The S1 nuclease digestion products from COS-1 cells and lymphocytes were analyzed on the same Southern blot (Fig. 3C, lanes 1 and 3). A 2-kb exon with homology to I1f was seen with RNA from both sources. The exon of 1,050 bp (Fig. 3C, lanes 1 and 2) resulted from transcription from the vector (Fig. 3B, lane 4).

Transcription of deletion mutant in COS-1 cells. S1 nuclease analysis of RNAs made in COS-1 cells transfected with the 600-bp I1f deletion mutant was carried out to demonstrate that the EBNA exon crossed this deletion and to map the boundaries of the deletion. The exon transcribed from the deleted gene was 1,400 bases (Fig. 3D, lane 4), as could be seen when I1fdel DNA was annealed to RNA transcribed from pSV2-I1fdel. When the same message was annealed to a full-sized I1f DNA, the deleted sequences were not protected. Therefore, two fragments resulted from S1 digestion (Fig. 3, schematic). This analysis shows that the deletion began 370 bp from one end of the 2-kb exon and ended 1,050

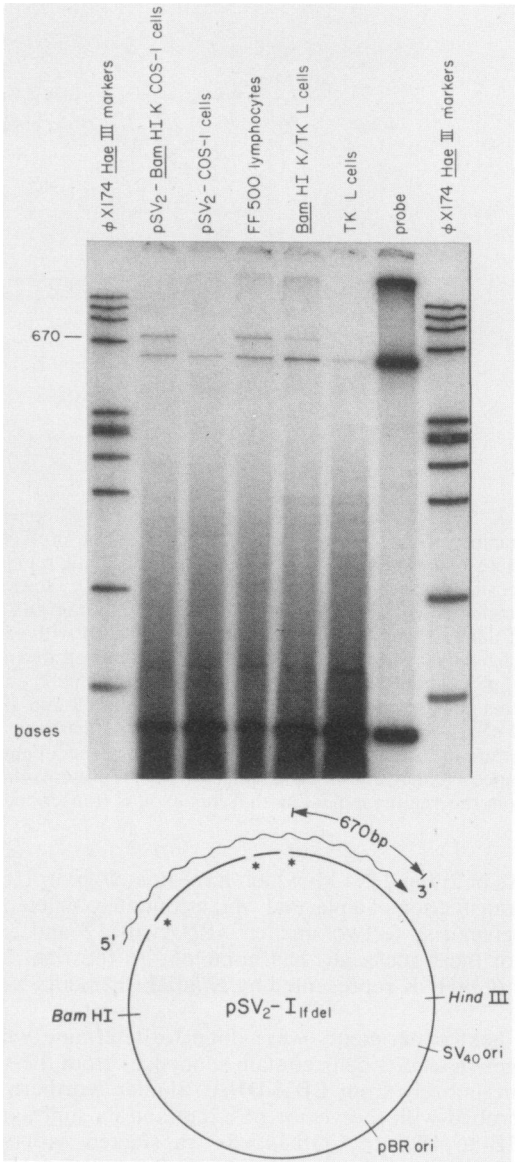

FIG. 5. S1 nuclease analysis of the 3′ end of the EBNA exon. Plasmid pSV2-I1fdel was cut to completion with *Sma*I, and the 3′ ends were labeled with the Klenow fragment of DNA polymerase I. A small amount of this probe is shown. Below the gel is a schematic representation showing the coding strand hybridized to the EBNA exon. This probe was annealed to total cytoplasmic RNA from the cell lines indicated above the gel. A band of 670 bp is seen in all reactions containing RNA from FF500 lymphocytes or cells transfected with *Bam*HI-K. This 670-bp band localizes the 3′ end of the EBNA exon as shown below the gel.

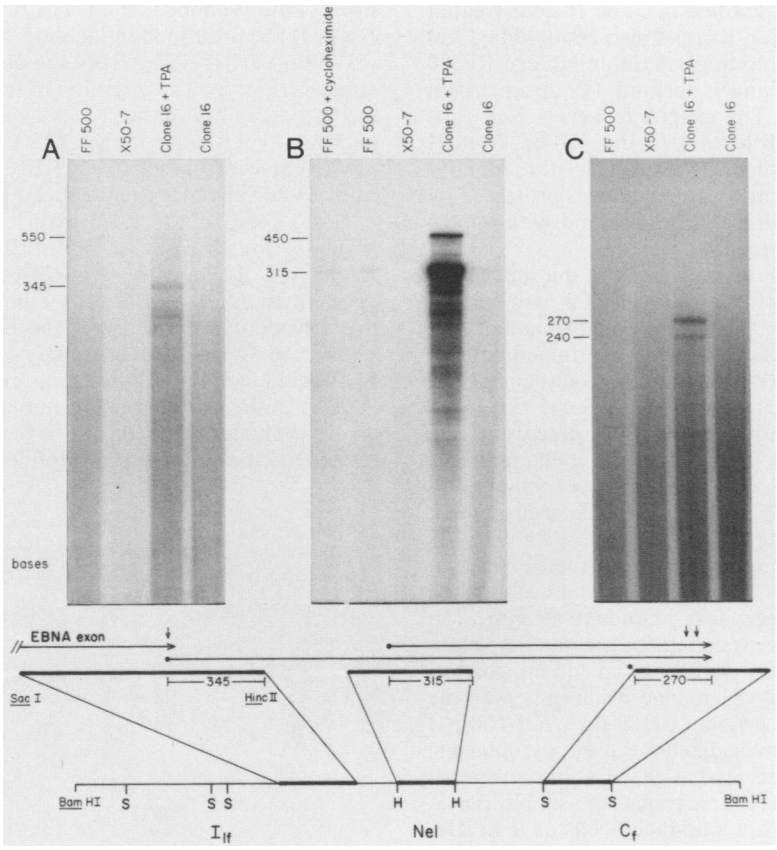

FIG. 6. S1 nuclease analysis of replicative messages in *Bam*HI-K. (A) A uniformly labeled probe of the *Sac*I-*Hinc*II fragment hybridized to total cytoplasmic RNA from three lymphoid lines. The band at 345 bases maps the 5′ end of the 2.1-kb message as diagrammed below the gel. The band at 550 bases corresponds to the 3′ end of the EBNA exon entering from the left. (B) Uniformly labeled probe of the *Hin*dIII Nel fragment hybridized to total cytoplasmic RNA from the same cell lines. The band at 315 bp corresponds to the 5′ end of the 1.3-kb message as diagrammed below the gel. The 450-bp band corresponds to complete protection of *Hin*dIII-Nel by the 2.1-kb mRNA. (C) *Sma*I subfragment of *Hin*dIII-Cf 3′ end labeled (*) with Klenow enzyme and annealed to total cytoplasmic RNA from the same cell lines. The bands at 270 and 240 bp represent the 3′ ends of these messages. Vertical arrows above the mRNAs correspond to polyadenylation signals AATAAA.

bp from the other end (Fig. 3D, lane 1). The size of these two fragments was equivalent to the 1,400-bp exon. The analysis at this point does not tell us the orientation of these pieces. However, the results show that the 2-kb exon demonstrated to be transcribed in COS-1 cells and lymphocytes crossed the deletion in I1fdel.

Fine mapping of the 5′ end of the 2-kb exon. The 540-bp leftmost *Bam*HI-*Sma*I subfragment of *Bam*HI-K was cloned into M13-MP9. A uniformly radio labeled DNA strand was synthesized from this phage to be used in S1 nuclease analysis to locate the 5′ end. The same band was seen when RNA was prepared from lymphocytes or the gene expression systems (Fig. 4). A section (179 bp) of the single-stranded DNA was protected from S1 nuclease digestion by RNA from COS-1 cells transfected with pSV2-*Bam*HI-K (Fig. 4), from stably transformed LTK$^-$/*Bam*HI-K cells, and from latently infected X50-7 lymphocytes. RNA from cells lacking EBNA expression did not protect the probe. These experiments localized the 5′ end of the EBNA exon 179 bp away from the *Sma*I site (Fig. 4, diagram). The experiment also identified the direction of transcription from left to right on the EBV map.

Fine mapping of the 3′ end of the 2-kb exon. The probe for these studies was pSV2-I1fdel digested to completion with *Sma*I and 3′ end labeled with [α^{32}P]dCTP. This group of DNA fragments was used to protect RNA from COS-1 cells transfected with pSV2-*Bam*HI-K or pSV2gpt, from L cells stably transformed with *Bam*HI-K and the HSV tk gene, or from lymphocytes, FF500. RNA from cells expressing EBNA protected a labeled DNA fragment of 670 bases (Fig. 5). RNA from appropriate control cells did not protect this fragment. This experiment localized the 3′ end of the 2-kb exon 670 bp 3′ to the rightmost *Sma*I site in I1f (for a map of *Sma*I sites, see Fig. 5, 6, and 8). The 500- and 60-bp *Sma*I fragments of the probe were seen in all lanes; these bands represented reannealed probe.

S1 nuclease mapping of replicative mRNAs in *Bam*HI-K. The right end of *Bam*HI-K, represented by the 1.7-kb-pair *Bam*HI-*Hin*dIII subfragment Cf, identifies two mRNAs of 2.1 and 1.3 kb which are inducible with TPA or viral superinfection (Fig. 2) (23). In Fig. 6 we show experiments which mapped the 5′ and 3′ ends of these messages. The 5′ end of the most abundant 1.3-kb message lay 315 bp from the right *Hin*dIII site of Nel (Fig. 6B). This same 5′ end was found in FF500 lymphocytes with a low rate of spontaneous viral antigen synthesis. In FF500 cells the content of this mRNA did not appear to be altered by cycloheximide treatment. Expression of this gene was not detectable in X50-7 cells which are tightly latent.

The 5′ end of the larger, less abundant 2.1-kb replicative

message was most clearly identified in clone 16 cells treated with TPA (Fig. 6A). However, it could also be identified, on longer exposures of the autoradiogram, in uninduced FF500 and clone 16 cells. This 5′ end is located 345 bp upstream from the *Hin*cII site of I1f. The mRNA traverses Nel, thus accounting for complete protection of the 450-bp *Hin*dIII Nel fragment (Fig. 6B). Of the 872-bp probe used to identify the 5′ end of this replicative mRNA, 550 bp was protected by the 3′ end of the EBNA mRNA (Fig. 6A) and with appropriate exposures was seen in all lanes.

The approximate location of the 3′ ends of the replicative messages was determined with a probe of Cf 3′ end labeled at the *Hin*dIII site. The size of the protected fragment was about 870 bp. A more accurate size was determined with Cf digested with *Sma*I and 3′ end labeled. A prominent 3′ end was found 270 bp downstream from the left *Sma*I site of Cf. This was most clearly identified with RNA prepared from clone 16 cells treated with TPA (Fig. 6A), but on longer exposures (not shown) was also seen in lower amounts in uninduced clone 16 and FF500 cells. The less prominent 240-bp band corresponded to another 3′ end.

Analysis of major replicative message by primer extension. By DNA sequence, there is a promoter element associated with the 5′ end of the larger, less abundant message (1). However, no promoter was predicted for the smaller, more abundant message. We used primer extension to confirm that the 5′ end identified by S1 nuclease analysis was the start of the message. The primer, a 180-bp *Xba*I-*Hin*dIII fragment from Nel, was hybridized to total cytoplasmic RNA from clone 16 cells treated with TPA or untreated. cDNA was synthesized with reverse transcriptase. No cDNA was detectable with RNA from uninduced cells as template. When RNA from TPA-treated clone 16 cells was the template, a band of 315 bp was seen (Fig. 7). This band corresponded in size exactly to the 5′ end mapped by S1 analysis (Fig. 6).

DISCUSSION

Transcriptional mapping experiments in which the precise boundaries of the exon encoding EBNA were mapped and two adjacent unspliced replicative genes were identified are summarized in Fig. 8.

Transcription of EBNA. Lymphocytes latently infected with EBV express a polyadenylated mRNA of approximately 3.7 kb which has homology to the I1f fragment (Fig. 2) (22). COS-1 cells, which express an EBNA of authentic size after transfection with pSV2-*Bam*HI-K, express hybrid mRNAs of 2.9 and 2.4 kb (18). This apparent inconsistency was resolved by showing that immortalized lymphocytes, transfected COS-1 cells, and stably transformed LTK cells all expressed an identical 2-kb exon contained in I1f. We mapped the boundaries of this exon and showed that it completely contains the open reading frame, which is the structural gene for EBNA.

The 5′ end of the 2-kb exon is located 179 bp upstream of the leftmost *Sma*I site of *Bam*HI-K (Fig. 4). This 5′ end is faithfully utilized in the two eucaryotic gene transfer systems. We analyzed these S1 nuclease results in the context of the sequence of *Bam*HI-K provided by Baer et al. (1). The DNA sequence at the 5′ end of the exon is consistent with a splice acceptor site (15). If the protected S1 nuclease product is exactly 179 bp, then, according to the DNA sequence, the first two nucleotides are AG. These two nucleotides are considered invariant as the last two nucleotides of any intron (3). As such, we would not expect them to be protected from S1 digestion. We therefore hypothesize that the splice donor site has the sequence AG↓GT. Confirmation of this hypothesis will require the identification of the upstream exon(s).

Within COS-1 cells, EBNA messages contain vector sequences (Fig. 1). The direction of transcription indicates that the late SV40 promoter was being used. Additional S1 nuclease experiments in the COS-1 system (not shown) were consistent with the idea that EBNA mRNAs are promoted from SV40 sequences and spliced into I1f.

The 3′ end of the 2-kb exon, which is identical in all systems (Fig. 5), occurs about 30 bp after the sequence AATAAA. Because these mRNAs are selectable on oligodeoxythymidylate cellulose, we believe this end represents the polyadenylation signal of the EBNA message.

Only one open reading frame of sufficient size to encode EBNA is located in I1f (1). This reading frame, designated BKRF1 in the Cambridge terminology, contains the alanine-glycine polymer which others have shown to be antigenically related to the EBNA polypeptide (4, 12). BKRF1 begins

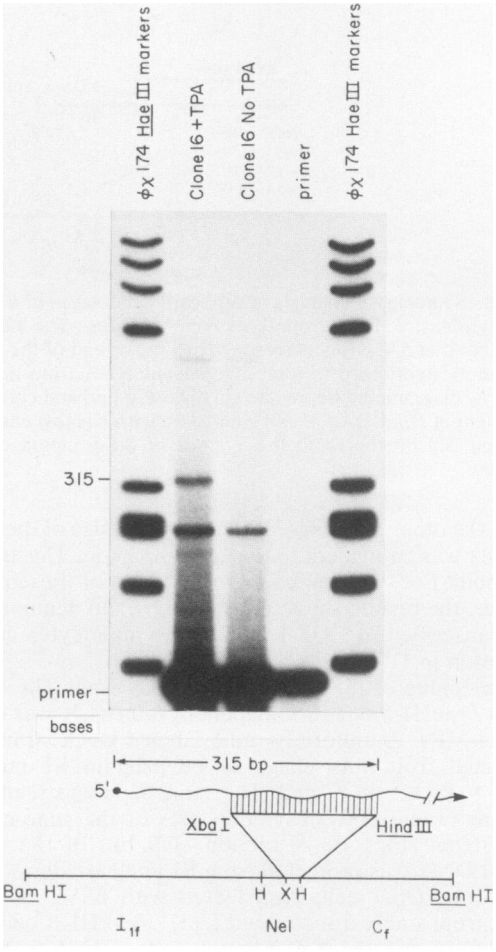

FIG. 7. Primer extension of the abundant replicative message. The 180-bp *Xba*I-*Hin*dIII fragment of *Hin*dIII-Nel was excised from an acrylamide gel and used as a primer by annealing it to total cytoplasmic RNA from clone 16 treated with TPA or clone 16 untreated. Avian myeloblastosis virus reverse transcriptase was used to extend the primer to the 5′ end of the transcript as diagrammed below the gel. A band of 315 bp is seen in clone 16 treated with TPA only. Small amounts of DNA which contaminated the primer are seen in all lanes.

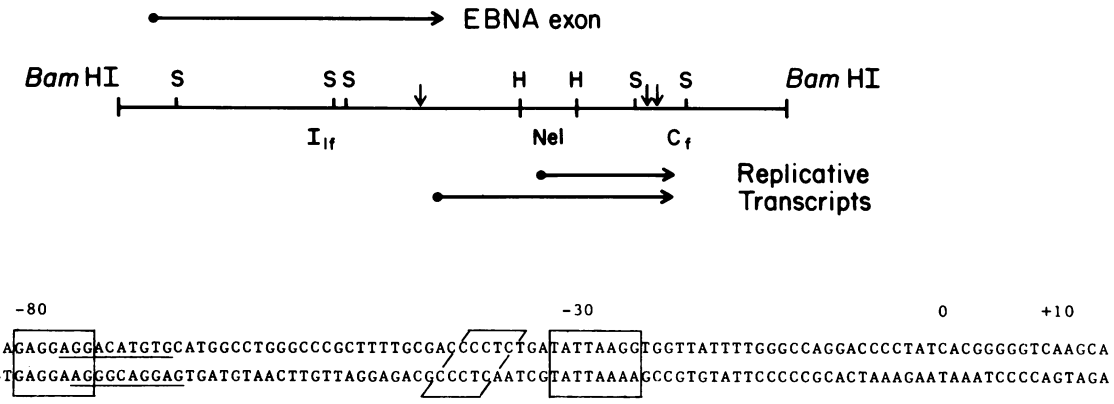

FIG. 8. Transcriptional map and promoter regions of replicative genes, summarizing primer extension and S1 nuclease data and showing the location of the EBNA coding exon and the two replicative genes contained entirely within the *Bam*HI K fragment. Vertical arrows are the positions of the polyadenylation signal AATAAA. Below the map are the promoter sequences of the 1.3- and 2.1-kb messages. The underlines indicate sequences shown to be contained within other EBV replicative promoters (5). Boxes indicate other regions of homology between these two replicative promoters.

three nucleotides downstream from the point which we have mapped to be the 5′ end of the EBNA exon. This open reading frame continues for 1,925 bp and ends about 90 bp before the 3′ end of the 2-kb exon. Therefore, the open reading frame is completely contained within the 2,028-base exon in I1f. The conclusion that EBNA is entirely encoded by this exon is consistent with previous gene transfer experiments.

A deletion which interrupts this exon, and removes all but 65 bp of IR3, gives rise to appropriately smaller mRNAs (Fig. 1) and polypeptide (18). The present studies mapped precisely the boundaries of this deletion mutant. This will be useful in subsequent studies of EBNA.

The abundance of EBNA message in lymphocytes is low. By comparing the hybridization signals of S1 nuclease digestion products with known amounts of plasmid DNA, we estimated that the abundance of EBNA mRNA present in immortalized lymphocytes is about one copy per cell (data not shown). The abundance of EBNA message can be significantly increased by cycloheximide treatment (Fig. 2). This finding should facilitate further study of the structure of EBNA mRNA in lymphocytes.

Transcription of replicative genes. Our analysis of the two replicative genes located in *Bam*HI-K highlights differences between latent and replicative transcription of EBV genes. We were unable to detect expression of these two genes in a tightly latent line (X50-7) by either Northern blot or S1 nuclease analysis (Fig. 6). Unlike EBNA transcription which is unchanged after induction of the replicative cycle (23), these two messages increase markedly in abundance when replication is triggered by treatment with phorbol ester or by superinfection (Fig. 2) (23). In FF500 cells, in which these genes have a low level of spontaneous expression, cycloheximide does not alter the number of these messages, although the drug does increase abundance of EBNA message (Fig. 2 and 6).

In contrast to EBNA message, the replicative mRNAs are unspliced. The sizes of mRNAs identified on Northern blots, 1.3 and 2.1 kb, are appropriate for exons of 1,170 and 1,870 bases which were mapped by S1 nuclease analysis. Furthermore, the 5′ end of the more abundant smaller mRNA found by S1 analysis was identical to that demonstrated by primer extension. The 3′ ends of both replicative messages were associated with polyadenylation signals. The sequence AATAAA occurs twice, about 20 to 30 bp upstream of each of the 3′ ends mapped by S1 analysis. The larger 3′ S1 product was more abundant (Fig. 6); therefore, this second polyadenylation signal must be used by the smaller, more abundant mRNA. However, these two signals are too close together to enable us to determine whether the larger mRNA uses only the first polyadenylation signal.

The 5′ flanking sequences of the two replicative genes show remarkable sequence similarity to each other (Fig. 8). Approximately 30 bp upstream of the cap site of the 2.1-kb mRNA is the sequence TATTAAAA; 30 bp upstream of the 1.3-kb mRNA is the sequence TATTAAGG. The 5′ flanking regions of the two replicative genes in *Bam*HI-K contain additional homologous sequences, e.g., GAGGA$^{A}_{G}$G located at about −80 bp. Furthermore, the two *Bam*HI-K replicative genes contain other 5′ signals which have been identified in replicative genes located elsewhere in the EBV genome (5). A reasonable hypothesis to explore is that such signals in part account for the difference between replicative and latent gene expressions.

The promoter for the larger replicative message was predicted on the basis of computer-assisted analysis of the DNA sequence (1). The promoter of the smaller, more abundant message, which begins within the open reading frame (BKRF3), was not predicted by DNA sequence analysis. This finding underscores the need for in vivo analysis of EBV transcription.

ACKNOWLEDGMENTS

We thank Katherine Papov and Frances Larvey for manuscript preparation.

This work was supported in part by grant MV173E from the American Cancer Society and Public Health Service grant 16038 from the National Institutes of Health. R.W. is supported by a predoctoral scholarship from The National Institutes of Health Medical Scientist Training Program.

LITERATURE CITED

1. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatful, G. S. Hudson, S. C. Satchwell, C. Séguin, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310**:207–211.
2. **Bolivar, F., R. L. Rodriguez, P. J. Green, M. C. Betlach, H. L.**

Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow. 1977. Construction and characterization of new cloning vehicles. II. A multiple cloning system. Gene **2**:95–113.
3. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
4. **Dillner, J., L. Sternås, B. Kallin, H. Alexander, B. Ehlin-Henriksson, H. Jörnvall, G. Klein, and R. Lerner.** 1984. Antibodies against a synthetic peptide identify the Epstein-Barr virus-determined nuclear antigen. Proc. Natl. Acad. Sci. U.S.A. **81**: 4652–4656.
5. **Farrell, P. J., A. Bankier, C. Séguin, P. Deininger, and B. G. Barrell.** 1983. Latent and lytic cycle promoters of Epstein-Barr virus. EMBO J. **2**:1331–1338.
6. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
7. **Fischer, D. K., M. F. Robert, D. Shedd, W. P. Summers, J. E. Robinson, J. Wolak, J. E. Stefano, and G. Miller.** 1984. Identification of Epstein-Barr nuclear antigen polypeptide in mouse and monkey cells after gene transfer with a cloned 2.9-kilobase-pair subfragment of the genome. Proc. Natl. Acad. Sci. U.S.A. **81**:43–47.
8. **Grogan, E. A., W. P. Summers, S. Dowling, D. Shedd, L. Gradoville, and G. Miller.** 1983. Two Epstein-Barr viral nuclear neoantigens distinguished by gene transfer, serology, and chromosome binding. Proc. Natl. Acad. Sci. U.S.A. **80**:7650–7653.
9. **Hearing, J. C., J.-C. Nicholas, and A. J. Levine.** 1984. Identification of Epstein-Barr virus sequences that encode a nuclear antigen expressed in latently infected lymphocytes. Proc. Natl. Acad. Sci. U.S.A. **81**:4373–4377.
10. **Heller, M., V. van Santen, and E. Kieff.** 1982. Simple repeat sequence in Epstein-Barr virus DNA is transcribed in latent and productive infections. J. Virol. **44**:311–320.
11. **Hennessy, K., M. Heller, V. Van Santen, and E. Kieff.** 1983. Simple repeat array in Epstein-Barr virus DNA encodes part of the Epstein-Barr nuclear antigen. Science **220**:1396–1398.
12. **Hennessy, K., and E. Kieff.** 1983. One of two Epstein-Barr virus nuclear antigens contains a glycine-alanine copolymer domain. Proc. Natl. Acad. Sci. U.S.A. **80**:5665–5669.
13. **Hummel, M., and E. Kieff.** 1982. Epstein-Barr virus RNA. VIII. Viral RNA in permissively infected B95-8 cells. J. Virol. **43**:262–272.
14. **Miller, G., M. Rabson, and L. Heston.** 1984. Epstein-Barr virus with heterogeneous DNA disrupts latency. J. Virol. **50**:174–182.
15. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
16. **Rabson, M., L. Heston, and G. Miller.** 1983. Identification of a rare Epstein-Barr virus variant that enhances early antigen expression in Raji cells. Proc. Natl. Acad. Sci. U.S.A. **80**: 2762–2766.
17. **Reedman, B., and G. Klein.** 1973. Cellular localization of an Epstein-Barr virus variant that enhances early antigen expression in Raji cells. Proc. Natl. Acad. Sci. U.S.A. **80**:2762–2766.
18. **Robert, M. F., D. Shedd, R. J. Weigel, D. K. Fischer, and G. Miller.** 1984. Expression in COS-1 cells of Epstein-Barr virus nuclear antigen from a complete gene and a deleted gene. J. Virol. **50**:822–831.
19. **Shin, S., A. Tanaka, and M. Nonoyama.** 1983. Transcription of the Epstein-Barr virus genome in productively infected cells. Virology **124**:13–20.
20. **Summers, W. P., E. A. Grogan, D. Shedd, M. Robert, C.-R. Liu, and G. Miller.** 1982. Stable expression in mouse cells of nuclear neoantigen after transfer of a 3.4-megadalton cloned fragment of Epstein-Barr virus DNA. Proc. Natl. Acad. Sci. U.S.A. **79**: 5688–5692.
21. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. **77**:5201–5205.
22. **Van Santen, V., A. Cheung, and E. Kieff.** 1981. Epstein-Barr virus RNA VII: size and direction of transcription of virus-specific cytoplasmic RNAs in a transformed cell line. Proc. Natl. Acad. Sci. U.S.A. **78**:1930–1934.
23. **Weigel, R., D. K. Fischer, L. Heston, and G. Miller.** 1985. Constitutive expression of Epstein-Barr virus-encoded RNAs and nuclear antigen during latency and after induction of Epstein-Barr virus replication. J. Virol. **53**:254–259.
24. **Weigel, R., and G. Miller.** 1983. Major EB virus-specific cytoplasmic transcripts in a cellular clone of the HR-1 Burkitt lymphoma line during latency and after induction of viral replicative cycle by phorbol esters. Virology **125**:287–298.
25. **Wilson, M. W., J. Nevins, J. M. Blanchard, H. S. Ginsberg, and J. Darnell.** 1980. Metabolism of mRNA from the transforming region of adenovirus 2. Cold Spring Harbor Symp. Quant. Biol. **44**:447–455.
26. **zur Hausen, H., F. O'Neill, U. Freese, and E. Hecher.** 1978. Persisting oncogenic herpesvirus induced by tumor promoter TPA. Nature (London) **272**:373–375.