EBNA-2 of Herpesvirus Papio Diverges Significantly from the Type A and Type B EBNA-2 Proteins of Epstein-Barr Virus but Retains an Efficient Transactivation Domain with a Conserved Hydrophobic Motif

PAUL D. LING,^{1,2} JUDITH J. RYON,² AND S. DIANE HAYWARD^{1,2*}

Molecular Virology Laboratories, Department of Pharmacology and Molecular Sciences¹ and Department of Neurology,² The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205

Received 16 October 1992/Accepted 20 February 1993

EBNA-2 contributes to the establishment of Epstein-Barr virus (EBV) latency in B cells and to the resultant alterations in B-cell growth pattern by up-regulating expression from specific viral and cellular promoters. We have taken a comparative approach toward characterizing functional domains within EBNA-2. To this end, we have cloned and sequenced the EBNA-2 gene from the closely related baboon virus herpesvirus papio (HVP). All human EBV isolates have either a type A or type B EBNA-2 gene. However, the HVP EBNA-2 gene falls into neither the type A category nor the type B category, suggesting that the separation into these two subtypes may have been a recent evolutionary event. Comparison of the predicted amino acid sequences indicates 37% amino acid identity with EBV type A EBNA-2 and 35% amino acid identity with type B EBNA-2. To define the domains of EBNA-2 required for transcriptional activation, the DNA binding domain of the GAL4 protein was fused to overlapping segments of EBV EBNA-2. This approach identified a 40-amino-acid (40-aa) EBNA-2 activation domain located between aa 437 and 477. Transactivation ability was completely lost when the amino-terminal boundary of this domain was moved to aa 441, indicating that the motif at aa 437 to 440, Pro-Ile-Leu-Phe, contains residues critical for function. The aa 437 boundary identified in these experiments coincides precisely with a block of conserved sequences in HVP EBNA-2, and the comparable carboxy-terminal region of HVP EBNA-2 also functioned as a strong transcriptional activation domain when fused to the Gal4(1-147) protein. The EBV and HVP EBNA-2 activation domains share a mixed proline-rich, negatively charged character with a striking conservation of positionally equivalent hydrophobic residues. The importance of the individual amino acids making up the Pro-Ile-Leu-Phe motif was examined by mutagenesis. Any alteration of these residues was found to reduce transactivation efficiency, with changes at the Pro-437 and Phe-440 positions producing the most deleterious effects. Activation of the EBV latency C promoter by EBNA-2 was shown to be dependent on the presence of the carboxy-terminal activation domain. However, this requirement was generic, rather than specific, since the EBNA-2 activation domain could be replaced with those from the herpes simplex virus (HSV) VP16 protein or the EBV Rta protein. Potential karyophilic signals within EBNA-2 were examined by introducing oligonucleotides encoding positively charged amino acid groupings that might serve in this capacity into a cytoplasmic test protein, HSV Δ IE175, and by examining the intracellular localization of the resulting proteins. This assay identified a strong nuclear localization signal between EBV amino acids (aa) 478 to 485, which was conserved in HVP, and a weaker noncanonical signal between EBV aa 341 to 355, which was not conserved in HVP.

Epstein-Barr virus (EBV), a gammaherpesvirus, has a well-established association with human neoplasia. EBV DNA and latent gene expression have been documented in Burkitt's lymphoma and nasopharyngeal carcinoma as well as within a subset of Hodgkin's lymphomas and B-cell lymphomas in the immunocompromised (28, 29, 43, 52). The abilities of EBV to immortalize B cells and to maintain a lifelong latent infection in its host are central to its role in human disease. Consequently, the viral latency genes, their regulation, and their mode of action have been targets of intensive scrutiny. On infection of primary B cells with EBV in vitro, the first viral genes expressed are EBNA-2 and EBNA-LP followed by the other latency genes EBNA-1, EBNA-3A, EBNA-3B, and EBNA-3C, LMP-1 and LMP-2 (terminal protein), and the small polymerase III EBERs (3, 4, 32). Examination of promoter usage in these cells has

revealed that the EBV latency W promoter, W_P (35, 42), is used initially and that the switch to the latency C promoter (7) is temporally coincident with the appearance of the transcripts for EBNA-1 and EBNA-3 (50, 51). Infection of peripheral B lymphocytes with P3HR-1 virus which lacks the EBNA-2 gene results in expression of a modified version of EBNA-LP, no EBNA-1 or EBNA-3, and failure to activate the latency C promoter (32, 50). These observations indicated a role for EBNA-2 in regulation of latency gene expression in the setting of primary EBV infection. The ability of EBNA-2 to up-regulate the C_P promoter has since been directly demonstrated in cotransfection experiments and an EBNA-2-responsive region has been mapped between -331 and -380 of the C_P mRNA start site (25, 33, 44, 50). Other viral genes whose expression is influenced by EBNA-2 are the latency membrane proteins LMP-1 and LMP-2. Initial observations showed that infection of EBV-negative B-lymphoma cell lines with B95-8 virus led to the

* Corresponding author.

expression of LMP along with the latency nuclear proteins, whereas conversion with P3HR-1 did not result in detectable LMP (1). Introduction of an EBNA-2 expression vector into these P3HR-1-converted cells induced LMP, suggesting that EBNA-2 participated in the up-regulation of LMP. Cotransfection experiments that directly support a role for EBNA-2 in transcriptional activation of LMP-1 and LMP-2 have been performed (16, 45, 49, 53).

Infection of B cells with EBV leads to the surface expression of activation antigens similar to those expressed transiently on the surface of antigen- or mitogen-stimulated cells and to continual B-cell proliferation. EBNA-2 has a role in these processes (10, 34). The introduction of EBNA-2 into P3HR-1-converted lymphoma cell lines results in an increase in surface expression of the activation antigens CD21 and CD23 together with the cellular adhesion molecules LFA1 and LFA3 and to a concomitant change in phenotype from growth as single cells to growth in clumps (1). These changes are mediated through the combined actions of EBNA-2 and LMP (47). However, EBNA-2 alone is capable of up-regulating CD21 and CD23 (12, 46), and cotransfection assays have mapped an EBNA-2-responsive region in the CD23 promoter (48). EBNA-2 also increases expression of the *c-fgr* proto-oncogene (26). EBNA-2 is thus able to modulate the expression of both cellular and viral promoters. The importance of this activity is indicated by the inability of P3HR-1 virus (lacking EBNA-2 and EBNA-LP) to immortalize primary B cells and the restoration of this property, albeit at a lower efficiency, by recombination back into the P3HR-1 genome of EBNA-2 alone (11, 21).

The function of EBNA-2 as a transcriptional transactivator is clearly central to its role in the establishment of EBV latency and B-cell immortalization. In these studies, we have used a comparative approach as a basis for identifying and characterizing the domains of EBNA-2 that are required for activation of gene expression and for appropriate intracellular localization. The baboon lymphotropic virus, herpesvirus papio (HVP), shows 40% cross-hybridization with EBV and has a colinear genomic organization (17, 24). HVP is capable of immortalizing both simian and human B cells in vitro (19). Immunofluorescence studies using human and baboon sera found cross-reactivity between the late-lytic-cycle antigens present in EBV- and HVP-infected cells (17). Interestingly, no cross-reactivity was observed between the latency EBNA antigens, raising the possibility that the amino acid sequences of these proteins might be more divergent. We felt that the EBNA-2 gene of HVP would be an ideal candidate for a comparative study, and this has proven to be the case. The predicted amino acid sequence of HVP EBNA-2 shows small islands of homology to the EBV EBNA-2 proteins within an overall extensively divergent polypeptide sequence.

MATERIALS AND METHODS

Plasmids. The polymerase chain reaction (PCR) was employed to generate DNA fragments containing sections of the EBNA-2 coding region. The following series of primers were used to generate PCR products with *Bcl*I, *Bam*HI, or *Bgl*III restriction sites.

5' primers

5' CTAGTGATCAATGCGCTACATTCTATCTT 3' (aa 1)
5' CTAGAGATCTCAGCGCAGGGATGCCTGG 3' (aa 101)
5' CTAGGGATCGCGCTAGGATATGACGT 3' (aa 128)
5' CTAGAGATCTCCTGTCTTGGATGTGCA 3' (aa 252)
5' CTAGAGATCTGCTGGACCTTGGAGACCA 3' (aa 367)
5' CTAGAGATCTCCAAATACATGAACCGGAG 3' (aa 424)

5' CTAGAGATCTCCGATTTCTCTCCCGGAT 3' (aa 437)
5' CTAGAGATCTCCGATGATTGGTATCCT 3' (aa 441)
5' CTAGAGATCTCCGCGAGACTTAGACGAA 3' (aa 451)

3' primers

5' CTAGAGATCTCAGCCGCCGTGTTTTCGCC 3' (aa 58)
5' CTAGAGATCTTCTTGCCAAAATAAGGCC 3' (aa 154)
5' CTAGAGATCTCAGGAGATGGCATCGCTGG 3' (aa 252)
5' CTAGAGATCTTATTGCTGAAACGTTAGG 3' (aa 425)
5' CTAGGGATCCTTACTCCAGATAATCTTCA 3' (aa 477)

The underlined codons match the amino acid number in the EBV(B95-8) EBNA-2 coding region (6). PCR reaction mixtures contained 10- μ l portions of 10 $\times$ reaction buffer [800 mM Tris-HCl (pH 8.9), 200 mM (NH₄)₂SO₄, 50 mM MgCl₂, 8 μ l of nucleoside triphosphates, 0.5 μ l of *Taq* polymerase (Cetus), 1 ng of cloned B95-8 DNA template plasmid DNA). The reaction mixtures were then brought to a final volume of 100 μ l with H₂O and incubated in a thermocycler for 40 cycles, with 1 cycle consisting of 2 min at 96°C, 2 min at 56°C, and 3 min at 72°C. The reaction products were resolved by agarose gel electrophoresis. The DNA was purified by electroelution, subjected to restriction enzyme digestion by the appropriate enzyme, and cloned into the *Bgl*III site downstream of aa 1 to 147 of the yeast DNA-binding protein Gal4 in pGH250 or pGH252. These plasmids express Gal4-EBNA-2 fusion proteins under the control of the simian virus 40 early promoter enhancer. All PCR-generated clones were sequenced on one strand, and none were found to contain any mutations. In addition, two primers, 5' CAGTAGATCTCCAATACTCTTCCCTGAA 3' and 5' CAGTGGATCCTGAGTTTGTAGTGTGTAAG 3', were employed as described above, to generate DNA containing the homologous activator domain region from HVP. This DNA fragment contains codons for aa 481 to 530 from HVP EBNA-2. The template DNA used was pJR6 (see "DNA sequencing" below). The PCR product was cloned into pGH250.

Similarly, mutations in the negative domain region of EBV EBNA-2 (aa 437 to 477) were made by PCR, and purified DNA from PCR reaction products was cloned into the *Bgl*III site of pGH250. The following series of primers were used to generate PCR products with mutations in the P-I-L-F domain.

5' primers

(A-I-L-F) 5' GCTAAGATCTGCCATTCTCTCCCGAT 3'
(P-A-L-F) 5' GCTAAGATCTCCGCTCTCTTCCCGATGA 3'
(P-I-A-F) 5' GCTAAGATCTCCGATTGCTTCCCGATGA 3'
(P-I-L-A) 5' GCTAAGATCTGCCATTCTCCGCGCATGATTG 3'
5' CTAGAGATCTCCC(A/T)(T/G)(T/C)(T/C)C(C/T/C)(T/C)CCCGATGAT 3'

3' primer 5' CTAGGGATCCTTACTCCAGATAATCTTCA 3'

Underlined letters indicate base changes that result in a change to the amino acids shown in parentheses to the left of the sequence. The last 5' primer contained 50% degeneracy in the positions indicated by two letters. Plasmids containing cloned DNA from PCRs using this primer were sequenced to determine which mutations had occurred. A proline-to-serine change in pPDL97 was an artifact of PCR amplification. The rest of the coding sequence beyond the P-I-L-F region (to aa 477) remained unaltered. C_p-chloramphenicol acetyltransferase (CAT) constructions were made by PCR, using the following primers.

5' primer 5' CTAGAGATCTCCCTAGGCCAGCCAGAGA 3'

3' primer 5' CTAGGGATCCTCCAAAGATAGCACTGGA 3'

The PCR product containing sequences from -330 to -430 of EBV C_p was cloned into the *Bgl*III and *Bam*HI sites of plasmid pGH56. This plasmid (pPDL63) was then cut with *Sca*I and *Bam*HI or with *Sca*I and *Bgl*III, and the resulting

fragments were electrophoresed on an agarose gel. The *ScaI*-*Bam*HI or *ScaI*-*Bgl*II fragments containing the C_p fragment were isolated and ligated, generating a tandemly repeated C_p DNA fragment that has both *Bgl*II and *Bam*HI sites at the ends but a fused *Bgl*II-*Bam*HI site in the middle (pPDL75). Subsequent rounds of cutting and religating produced plasmids containing four and eight copies of the 100-bp fragment. These plasmids were then cut with *Bgl*II and *Bam*HI, and the resulting fragments were cloned into the *Bgl*II site of pGH262 upstream of the TATA box. pGH262 contains an Elb TATA box followed by the CAT gene. Clones used in these studies had inserts in the sense direction.

DNA for the EBV EBNA-2 coding region was generated with the following primers.

5' primer 5' CTAGTGATCAATGCTAGATTCTATCTT 3' (aa 1)

3' primers

5' CTAGAGATCTTTACTGGATGGAGGGGCG 3' (stop codon)
5' CTAGAGATCTTTACTGGATGAAACGTTAGG 3' (aa 425)

The DNA fragments (*Bcl*II to *Bgl*II) were cloned in the sense orientation into the *Bgl*II site of pSG5 (Stratagene) to generate pPDL59A containing EBNA-2 aa 1 to 487. The EBNA-2 clone (pPDL41 [aa 1 to 425]) was cut with *Bgl*II and fragments from HSV VP16 (aa 411 to 490) or Rta (aa 520 to 605) (23) were cloned in frame with the EBNA-2 coding sequences to generate EBNA-2-VP16 or EBNA-2-Rta fusions. An in-frame termination codon from the vector DNA (SG5) is located just downstream from the last amino acid encoded by the deletion mutant (aa 1 to 425), while the VP16 and Rta chimeras contained the termination codons naturally occurring in those polypeptides.

CAT assays. Vero cells were maintained in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum (GIBCO) and incubated in 5% CO₂ at 37°C. IB4 and DG75 lymphoblastoid cells were maintained similarly in RPMI. Vero cells and B cells were transfected as previously described (27). At 40 h posttransfection, cells were washed once with phosphate-buffered saline, resuspended in 100 µl of 0.1 M Tris (pH 7.9), and sonicated. The cell debris was removed by centrifugation, and the supernatant was incubated with a solution containing 4 mM acetyl coenzyme A, [¹⁴C]chloramphenicol, and Tris (pH 7.9) for 1 to 2 h at 37°C (20). The reaction mixtures were extracted with ethyl acetate and spotted onto thin-layer chromatography plates and chromatographed in 95% chloroform–5% methanol. Acetylated spots were detected by autoradiography and quantitated by excising the region and counting in a liquid scintillation counter. In some experiments, cells were cotransfected with a plasmid containing the LacZ gene expressed from the simian virus 40 early promoter enhancer (pSV2GAL). Cell extracts used for CAT assays were also tested for β-galactosidase activity in these experiments. Briefly, extracts were added to 1 ml of Z buffer (0.06 M NaHPO₄ · 7H₂O, 0.04 M NaH₂PO₄ · H₂O, 0.01 M KCl, 1 mM MgSO₄ · 7H₂O, 0.05 M β-mercaptoethanol, pH 7.0) and 200 µl of 4 mg of *o*-nitrophenyl-β-D-galactopyranoside (ONPG) per ml. The solutions were incubated at 28°C for 4 h, and the reactions were stopped by the addition of 0.5 ml of 1 M sodium carbonate. Optical density readings were taken, and β-galactosidase units were calculated. Relative activity values from CAT assays were normalized to β-galactosidase units.

Immunofluorescence and nuclear localization assays. The experimental system employed to study nuclear localization signals in EBNA-2 has been described previously (5, 30).

Briefly, pGH114 expresses the HSV IE175 protein which localizes to the nucleus. A derivative of this plasmid, pGH115 (ΔIE175), contains an in-frame deletion in the coding sequences of IE175 such that the protein made from cells transfected with pGH115 localizes to the cytoplasm. The following oligonucleotide pairs encoding potential EBV EBNA-2 nuclear localization signals were cloned in frame into the unique *Bgl*II site of pGH115.

EBNA-2 aa 478 to 485 (pPDL22)

5' GATCCCCAGTAAAAGACCTCGCCCCCTCCCA 3'
5' GATCTGGGAGGGGGGAGGCTCTTTACTGGG 3'

EBNA-2 aa 356 to 368 (pPDL25a)

5' GATCTCCGGCAAGGGCAAGTCCAGGGACAAGCAAGCCCGG 3'
5' GATCCGGGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCGGA 3'

EBNA-2 aa 341 to 355 (pPDL50a)

5' GATCCGGGGACAGGGGAGGGGGAGGGGGAGGGGGAGGGGGAGGGCA 3'
5' GATCTGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCG 3'

In addition, one clone was made that contained both pPDL25a and pPDL50a oligonucleotide pairs. This clone, pPDL51a, was made by cleaving clone pPDL50a with *Bgl*II and inserting the oligonucleotide pair containing EBNA-2 aa 356 to 368.

Recombinant plasmids were transfected into 5 × 10⁴ Vero cells in four-well chamber slides (Lab-Tek) by the calcium procedure or into EBV-negative DG75 B cells with DEAE dextran (27). Between 24 to 48 h posttransfection, cells were washed in phosphate-buffered saline, and the DG75 cells were transferred onto glass slides, using a cytocentrifuge. Cells were fixed with –20°C methanol for 10 min. The fixed cells were air dried, rehydrated in Tris-saline, and incubated for 60 min with the 58S monoclonal antibody that recognizes IE175 (40). The cells were then washed three times in Tris-saline, incubated for 30 min with a goat anti-mouse fluorescein isothiocyanate (FITC)-conjugated antibody (Cappel), and rewashed in Tris-saline. All Gal4 fusion proteins were shown to be stably expressed and have an appropriate intracellular localization by staining transfected cells with rabbit anti-Gal4 antibody obtained from M. Green (University of Massachusetts, Worcester). Rabbit anti-EBNA-2 antiserum was kindly provided by W. Schuback, State University of New York, Stony Brook (38).

DNA sequencing. The coding region of the HVP EBNA-2 gene was located within the HVP cosmid clone cJR4 (34a) to a 1.2-kb *Xba*I-*Eco*RI fragment (pPDL28) derived from the HVP *Xba*I-I fragment (pJR6; 24) by Southern blot hybridization with a single-stranded EBV EBNA-2 RNA probe. To complete the open reading frame, an adjacent downstream *Eco*RI-*Sst*I fragment from the HVP *Xba*I-I fragment was also cloned (pPDL29) and sequenced while the N-terminal coding region was sequenced with the HVP *Xba*I-K fragment (pPDL74) (24). Clones were sequenced by the dideoxy-chain termination procedure of Sanger et al. (37) on denatured double-stranded plasmid DNA starting from the T7 primer location in the vector, Bluescript SK⁺ (Stratagene). Subsequently, primers were synthesized in a stepwise fashion on the basis of previous sequencing information. Complete overlapping sequence was obtained for both strands. In some cases, the nucleotide analog dITP, instead of dGTP, was used to eliminate GC compressions. In addition, the parent clone was sequenced across the *Eco*RI site to confirm that there were no small undetected *Eco*RI fragments.

Nucleotide sequence accession number. The GenBank accession number for the HVP EBNA-2 DNA sequence is L11366.

RESULTS

DNA sequence and predicted amino acid sequence of HVP EBNA-2. Information on the boundaries of functional domains and the importance of particular motifs within these domains can often be obtained by comparing proteins with similar functions or the equivalent proteins of related viruses. To gain a better understanding of the key features of the domains within EBNA-2, we cloned and sequenced the EBNA-2 gene from the baboon virus, HVP. A 40-kb cosmid clone of HVP, cJR4 (34a), was cleaved with a variety of restriction enzymes, and a Southern blot of the separated DNA fragments was hybridized with an EBV EBNA-2 riboprobe. In this way, an *XbaI-SstI* fragment was identified as containing the HVP EBNA-2 gene. This DNA fragment was cloned intact (pJR6) and as *SstI-EcoRI* subfragments into a Bluescript vector, and these plasmids were used in dideoxy sequencing reactions to obtain the sequence of HVP EBNA-2. It was subsequently recognized that the *XbaI-SstI* fragment lacked the 5' 44 bp of the HVP EBNA-2 gene and consequently the HVP *XbaI-K* fragment (24) was also subcloned from cJR4 and used to complete the HVP EBNA-2 DNA sequence. The DNA sequencing strategy is described in Materials and Methods and involved sequencing of both strands of the gene. The DNA sequence of HVP EBNA-2 is shown in Fig. 1. EBV isolates fall into two categories, type A and type B, which differ in their EBNA-2, and EBNA-3A, EBNA-3B, and EBNA-3C alleles (2, 14, 36). There is 47% DNA sequence homology between the open reading frames for HVP EBNA-2 and EBV B95-8 (type A) EBNA-2 and 48% homology with EBV AG876 (type B) EBNA-2. The noncoding sequences flanking the EBNA-2 open reading frame have maintained a noticeably higher degree of homology than the coding sequences. This high degree of homology is presumably indicative of a conserved regulatory role for these sequences. The 100-bp DNA sequences immediately 5' of the open reading frame are 88% identical between HVP and EBV (B95-8), and this leader region contains the splice acceptor that is used for the EBNA-2 mRNA plus a conserved consensus splice donor. The 100-bp sequences immediately 3' of the HVP and EBV open reading frames are 76% identical. The EBV EBNA-2 gene has two consensus polyadenylation signals separated by 48 bp downstream of the EBNA-2 open reading frame, and curiously, both signals are also present in the HVP EBNA-2 gene, even though only the first has been shown to be used in EBNA-2 transcripts (7, 35, 42).

The predicted amino acid sequence of HVP EBNA-2 is shown in Fig. 2, where it is compared with the EBV B95-8 and AG876 EBNA-2 protein sequences (14). The HVP EBNA-2 protein is slightly larger at 530 aa than the 487- and 454-aa EBV proteins. The increased size is largely accounted for by a stretch of 66 aa in the center of the protein that are unique to HVP EBNA-2. The polyproline stretch of EBV EBNA-2 is also present in HVP, and the size of this segment (23 residues) is intermediate between that of the AG876 (16-residue) and the B95-8 (42-residue) proteins. HVP EBNA-2 has 37% amino acid identity with B95-8 EBNA-2 and 35% identity with the AG876 protein. (The EBV EBNA-2 proteins are themselves 54% identical.) The amino acid homology is not dispersed evenly throughout the protein but rather consists of small clusters of conserved residues interspersed among blocks of sequence that retain little homology. These islands of identity are boxed in Fig. 2. The segments of HVP EBNA-2 that exhibit the highest homology with the EBV EBNA-2 proteins coincide with

```

AACTGTAATTATGGCAATAGCTAATGCCCCCTGTATCCCTTTCCATTCCACCACAGTAAA 60
CCACTGCTGGCGTGTGATGCTGTAAGATTTTGGCCGTTGGAAATGAACTTGTGGTGG 120
GGCAGGTAGATGCCAACCAACCTTTTAAGCACTCCGCTTGTCTTTTGTCTTCCGCGC 180
CGCCATGCTCTACTTACTATGTTGCAAGGTCAGGGATCAGTTTATTCTAGATGTTTT 240
M P T Y Y V A G R S G D Q F I L D V F
CAATGATGAAACCGGCGCCCTGCCATCTTACTGACTCCAAACGGTCCAGATCTCTCGGACC 300
N D E T G R P A I L L T P T V P A L G P
CCTGTCGAAGCACCTTAATTCACCTTAAATTTATTGTAGGGGAAAGTAGTGACCAACA 360
L S E A P L I P L K I I V G E S S D Q H
TCAAGTCTCTACCGGGGCTTCCAACTGCTCCGCCATCCGCCCTCCACCAACACC 420
Q G P L P G P S N L P P P P S P P P P P P
TCCACCACTCCACCTCCACCTCCACCAACCCCAACAGAGCGCAGGAGCGCCGCAAC 480
P P P P P P P P P P P P Q P E R R D A A T
GCAAGAGCCGATCCACACGATATCGCCCTCTAGGTATTTCACCTGAGATGGGGCTCT 540
Q E P L D P H D I R P L P F I G S P E N G A I L
TGCAAGTGAATTCGAATGCTTTGATGCTCTCTATATTGTAGAGCAAAACAAAGGCGC 600
A S A I R M L C M S S Y I V R Q T Q G A
GTGTGGCTGCTAAATCCAGGGTTAAGCAAGAGTTCGCCGCCCGCATGACAACTACT 660
C G L P L N P R V K P R P P P S P P P P P P
CAGACCTTTAAACCCGCTTCGGCTAATAAGGCTCTCGACCCACATATAACCCCTTGTC 720
R P L T P L R L I R P P R P T Y N P L S
ATCACCAACACTCCATCACCAAGCCCTATTCCTTTTTCAGCACCACTGCTCCGAC 780
S P N T P S P S P I P F Q H T P L P P R
GCGGCCCACTGCATCACCCCACTTCTCATCCTTTCCAGTCTGAGCTTAAACAAACACC 840
P P T A S P P H S H P F Q S E P K Q T P
ACCTAAGCCCACTACCTTTTACCGGCCCAACAGTATCACCCCAACCACTACCCCTC 900
P K P T L P L P G P P V S P P P P P T P S
CATTCAACCAACACCCCACTTACACCCCAAAACCAACCAATCCCTCTCCGGGCTCA 960
I Q P T P H P T P P K P T I P S S G P H
TATAACTGACACCGCTTCCCTTGCTCAACAGCATCAACCAAGGTTTCAATGTTTC 1020
I T L T P V P L S Q P A S T G G S N V S
TCAACCCCTCCACAGCATAAAACAAATCCTTATTATACACGCTCTAGCCACCTCACA 1080
Q P P P Q H K Q I L I I T T P L A T S Q
GACTCTTCAACCAACAAATATTACCAAAACCAACAGTCTTTCAGCATCTTGGACCC 1140
T P T K T Q I L P P T R S A S M D P P S
TTTGCTCTGCTCCCTGCTCTACGCCCCCAACCGGCTCCCTCCACACCACTCCAGG 1200
L P L P P L S T P P P P A P S T P S G
AATTGTACGCGACCGCCCACTCCCGCAGACCTGCGGACGAGTTTGTGGCCACCACT 1260
I V R D R P T S P R P L G P V S P W W P V
TCCATTACTGAGCATAAATCTGCTGCTCCAGATTACTTACTCCATCGTTTGTATCCACC 1320
P L P E H K L A G P D L L T P S F D P P
AACCCAGAGGAAGAACTGTGCGCAAGCGCTTCCGCTTAGGCAGGCAACACTAAG 1380
T P E E T V R K R V S R C P Q A T L R
AAAGCTAGGCCTTGCAGAAATTCCTCAACGGGAACATATACAGGAATTTTAGTCTAG 1440
K P R P C R I P Q R E H I P G T F S P R
GATGCTCACTGAGTCCAGCGGTTCTCTGCTGCTTCCATCAGCTAGGCCCAATAG 1500
M P H L S P A V P L G P V H Q P R P N S
CTCACCTCCCAAGTACACTGAGGCTTCTCTCTCAGAGCGTGTTCCTCCATGTAGC 1560
S P S T S T P E G L P P Q S V F P H V A
ACCGACCCCTCCACATCAACCTTTGCCATTGCAAGAACCTCAGTCTCTAAACACC 1620
P G P S T S Q P L P L Q E P Q S S K T P
AAGTCCGCAATACTCTTCCCTGAAGATTGATATCTCCACCTCCAGTCCCTGACAA 1680
S P P I L F P E D W Y T P T L Q S P D N
CCAATGGGAAGACATTTTAACTTCTCAGAACTAGTAGCTCTAATGAACCTCCCGAAG 1740
Q W E D I F N F S E S S S S N E P P E G
CCCCAAAGTAAAGACCTCGCCCTCTGACCAAGTAAATGCGCTTCTTCTCGGTTACC 1800
P Q S K R P R P S D Q *
ATTGTATAACCAATTAATGTTAATTCGTTCACTGCTTTTAAATGGAGCCCGCTGTCAT 1860
TATATCAATTAACACGCGGACAACTACGCTTCTTGACATCTCATTAGATAACCCCA 1920
GCTAAATTTGGGCGCAGATGCCCCCGCTTACCACTCAAACTCATGTAGCGCTAG 1980
CCGAATTTGGTCACCACTC 2000

```

FIG. 1. DNA sequence of the coding strand of the HVP EBNA-2 open reading frame and flanking sequences. The predicted amino acid sequence is presented beneath the open reading frame. Two potential polyadenylation signals are shown in bold type. A potential splice acceptor site is indicated with the vertical arrowhead, and a potential splice donor is indicated by the horizontal arrow. The locations of the restriction endonuclease sites *XbaI* and *EcoRI* used in cloning the gene are also given.

FIG. 3. Identification of the EBNA-2 transcriptional activator domain. EBNA-2-Gal4 fusion constructions were tested for their transactivating function in transient transfection assays using a minimal promoter containing the adenovirus E1b TATA box (A) or the E1b TATA box with five tandemly repeated upstream Gal4 binding sites (B). Shown above each lane are the EBNA-2 amino acids expressed by the effector DNA constructions. Below each lane the percent conversion of [¹⁴C]chloramphenicol is given. The assays were performed in Vero cells with 1.0 μ g each of effector and target DNA.

(Fig. 3B). We therefore conclude that the region of EBNA-2 between aa 424 and 477 contains a transactivator domain.

The P-I-L-F motif of the activator domain is critical for function. To define the amino-terminal boundary of the activator domain, three additional Gal4-EBNA-2 fusions were constructed. These fusions contained EBNA-2 codons at aa 437 to 477, 441 to 477, and 451 to 477. Again, cotransfection of the Gal4-EBNA-2 fusions had no effect on expression from the control E1bCAT target (Fig. 4A). Cotransfection of Gal4-EBNA-2(437-477) with (Gal4)₅-E1bCAT resulted in a 130-fold activation of the target, which was similar to the activation (113-fold) obtained with the larger EBNA-2(424-477) construction. In contrast, when the amino-terminal boundary was moved to aa 441 or 451, only a twofold activation of (Gal4)₅-E1bCAT was obtained (Fig. 4B). Thus, removal of aa 337 to 340 (Pro-Ile-Leu-Phe) from the amino-terminal boundary was sufficient to completely inactivate the transactivator function of this domain.

To ensure that the lack of transactivation observed with the six chimeras containing amino-terminal segments of EBNA-2 and with the two 5' deletion constructions was not due to protein instability or to inappropriate intracellular localization, the Gal4 fusions were transfected into Vero cells, and protein expression was examined by immunofluorescence staining with antiserum raised against the Gal4 polypeptide. Each of the fusion constructions showed comparable levels of stable protein, and these fusion proteins were all appropriately transported to the nucleus. The results from the transactivation experiments are summarized in Fig. 5, which also provides the amino acid sequence of the aa 437 to 477 domain. The proline-rich negatively charged nature of this domain is highlighted.

The deletional analysis placed the amino terminus of the EBV EBNA-2 activation domain at the Pro-Ile-Leu-Phe residues. Examination of the HVP EBNA-2 sequence re-

vealed that this motif coincided precisely with the amino terminus of a conserved block of sequence in the HVP protein (Fig. 2). The region of HVP EBNA-2 gene from the Pro-Ile-Leu-Phe codons to the carboxy terminus was fused to Gal4 (aa 1 to 147), and the chimeric construction was tested for the ability to activate a Gal4-E1bCAT target plasmid. This region of the HVP EBNA-2 protein was also found to contain a strong, functional activator domain (Fig. 6).

When the activator domains of the HVP and EBV EBNA-2 proteins are compared, one of the most striking features is the highly conserved alignment of hydrophobic residues (boxed in Fig. 7A). This conservation combined with the loss of activity when the amino-terminal Ile-Leu-Phe residues were deleted suggested that the hydrophobic residues may be important features of the activation domain. To assess the contribution of the individual amino acids in the Pro-Ile-Leu-Phe motif, these residues were mutated singly and as groups. Immunofluorescence assays indicated that each of these new constructions expressed a stable polypeptide product that was appropriately localized to the nucleus. In the experiment examining the effect of these mutations on transactivation of a Gal4-E1bCAT target, the transfection mix also contained an pSV2GAL control plasmid. In addition to CAT activity, β -galactosidase activity was also measured in the cell extracts and used to standardize the results obtained (Fig. 7B). The chief points arising from the mutagenesis of these four residues are as follows. (i) Any variation of these residues, even when it involved a relatively neutral change to an alanine, diminished activation efficiency. (ii) Conversion of the proline residue to either alanine (A-I-L-F) or serine (S-F-L-F) significantly reduced activity. (iii) Retention of this proline in the context of changes to each of the other three residues (P-C-F-S) also resulted in an inefficient activator protein. (iv) Of the three

FIG. 4. Location of the N-terminal boundary of the EBNA-2 activator domain. EBNA-2-Gal4 fusion constructions were tested for their ability to transactivate the adenovirus E1b TATA box (A) or the E1b TATA box with five tandemly repeated upstream Gal4 binding sites (B). Shown above each lane are the EBNA-2 amino acids expressed by the effector DNA constructions. Below each lane the percent conversion of [¹⁴C]chloramphenicol is given. Basal refers to cotransfection of target DNA with a nonspecific plasmid, pBR325. Assays were performed in IB4 cells with 1.0 μ g each of effector and target DNA.

hydrophobic residues, changes at the phenylalanine position produced the most deleterious effects. There was a gradient of activity decreasing from that of the wild-type bulky phenylalanine residue (P-I-L-F) through leucine (P-I-L-L) to the smaller alanine residue (P-I-I-A) and the nonhydrophobic serine (P-C-F-S). Conversely, the retention of the phenylalanine as the sole hydrophobic residue, plus proline (P-C-P-F),

was sufficient to produce approximately 40% of the wild-type activation level.

Contribution of the EBNA-2 activation domain to transactivation of the responsive elements in the latency C promoter. EBNA-2-responsive elements have been localized to within 100- to 150-bp segments of the EBV latency C promoter, the LMP promoter, and the cellular CD23 promoter. A feature of these analyses has been the inability to readily identify elements common to these three responsive promoters, which has led to the suggestion that EBNA-2 may utilize several different mechanisms to activate gene expression. We wished to examine the role of the EBNA-2 activation domain in up-regulation of the latency C promoter. In an effort to increase the sensitivity of the assay, the 100-bp region of the C promoter containing the EBNA-2 response region was multimerized and cloned upstream of the promoter in an E1bCAT target plasmid. The EBNA-2 responsiveness of target containing one, two, four, and eight copies of this responsive region was examined (Fig. 8A). An apparently cooperative effect was observed when the number of copies of this region was increased from two to four, and targets with four or eight copies were routinely highly responsive to EBNA-2 transactivation. The multimerized response regions were cloned in both forward and reverse orientations, and activation occurred in an appropriately orientation-independent manner (data not shown). The dependence of this response on the EBNA-2 activation domain was examined with an eight-multimer C promoter target and constructions in which the activation domain (aa 425 to 487) had been deleted or replaced with activation domains from two other viral transactivators, the HSV VP16 and the EBV Rta proteins. The ability of EBNA-2 to transactivate the C promoter target was abolished by removal of its activation domain. However, replacement of this region with activator domains from either VP16 or Rta reconstituted the ability to activate the C promoter target (Fig. 8B). Staining of transfected cells with polyclonal anti-EBNA-2 antiserum (38) revealed that each of the constructions expressed a stable protein product that showed mixed nuclear and cytoplasmic localization. This pattern of intracellular localization is consistent with the fact that these EBNA-2 variants contain only the weaker of the two EBNA-2 nuclear localization signals (see below). We conclude that the EBNA-2 activation domain is essential for transactivation of the C promoter elements and further conclude that this requirement for an activation domain function is generic and can be substituted effectively with activation domains from unrelated proteins.

EBNA-2 has a strong carboxy-terminal karyophilic signal sequence and a weaker internal karyophilic signal. Proteins that localize to the nucleus frequently contain karyophilic signals which feature positively charged amino acids (18). An examination of the EBV EBNA-2 amino acid sequence revealed only three regions with sufficient positive charge to be good candidates for such a function. Oligonucleotides encoding EBNA-2 aa 478 to 485, 341 to 355, 356 to 368, and the combination 341 to 368, were cloned into an HSV Δ IE175 expression vector. The HSV Δ IE175 protein lacks its natural karyophilic signal, and the in-frame introduction of an oligonucleotide encoding a functional signal results in the relocation of the Δ IE175 protein from the cytoplasmic to the nuclear compartment. This vector has previously been used to identify nuclear localization motifs in the EBV EBNA-1 and CMV IE2 proteins (5, 30).

Vero cells were transfected with HSV Δ IE175 and the constructions containing the EBNA-2 inserts, and the intracellular locations of the expressed proteins were examined in

FIG. 5. Summary of data mapping the EBNA-2 activator domain. (A) The polyproline region of EBNA-2 is shown as a hatched box. The black boxes indicate coding regions of EBNA-2 that were fused to Gal4. The amino acid numbers of these fusions are shown on the right. Subcellular localization was determined on transfected Vero cells by indirect immunofluorescence with antisera directed against the Gal4 protein. The plus signs indicate positive nuclear localization (N). Transactivation (TA) in transient transfection assays is indicated by +++ for positive activity or by - for negative activity. (B) The sequence of the 40-aa B95-8 EBNA-2 activation domain is presented with the negatively charged residues underlined and proline residues in bold type. Numbered arrowheads above the sequence indicate deletion end points. The P-I-L-F motif essential for transactivation function is boxed.

an immunofluorescence assay using an anti-IE175 monoclonal antibody. Introduction of EBNA-2 codons 478 to 485 into the ΔIE175 vector resulted in 87% of the positively staining cells showing exclusive nuclear localization and 13% giving a mixed pattern in which both nucleus and cytoplasm were stained but the nuclear staining was the most intense (Fig. 9E, Table 1). No exclusive nuclear staining was observed with the ΔIE175 constructions expressing EBNA-2 aa 341 to

FIG. 6. HVP EBNA-2 also contains a strong transcriptional activation domain. Cotransfection assays performed in Vero cells with 1.0 μg each of effector and target DNA. The effector DNA constructions are shown above each lane, and the target DNA constructions are shown below the brackets. Basal refers to cotransfection of target DNA with a nonspecific plasmid, pBR325. The percent conversion of [¹⁴C]chloramphenicol (% Ac) is indicated at the bottom of the gel.

FIG. 7. The HVP and EBV EBNA-2 transcriptional activator domains feature positionally conserved hydrophobic residues: contribution of individual residues in the P-I-L-F motif. (A) Comparison of the EBV and HVP activator domains with the activator domains from the HSV VP16 and yeast Gal4 proteins highlighting (boxes) the previously described alignment of bulky hydrophobic residues in the negatively charged activation domains of VP16 and Gal4 (13) and the overlapping but unique pattern of conserved hydrophobic residues in the EBNA-2 proteins. Dashes indicate deletions of the amino acid residues at these positions. (B) Mutational analysis of P-I-L-F residues in the EBNA-2 activator domain. Gal4 fusions expressing EBNA-2 aa 437 to 477 were cotransfected into Vero cells with the (GAL4)₅E1B-CAT target plasmid. Each sample was also transfected with 0.5 μ g of a plasmid, pSV2GAL. Basal refers to cotransfection of target DNA with a nonspecific plasmid, pBR325. Amino acid changes in the P-I-L-F domain constructions are shown underlined. The relative activation compared with the parent effector construction (P-I-L-F) is shown by the solid horizontal bars. The P-I-L-F mutation ---- indicates an effector construction (EBNA-2 codons at aa 441 to 477) lacking the P-I-L-F domain. Cell extracts from this experiment were normalized to β -galactosidase activity.

355, 356 to 368, or the combination 341 to 368. However, the presence of aa 341 to 355 did result in 16% of the positively staining cells exhibiting a mixed immunofluorescence pattern where the nucleus was more strongly stained than the cytoplasm (Fig. 9C and D), and the percentage of cells showing this pattern increased to 30% when the adjacent aa 356 to 368 were combined with aa 341 to 355 (Table 1). Similar staining patterns were observed in transfected DG75 lymphoblastoid cells (Fig. 9G and H). We interpret this data to indicate that the EBNA-2 sequence Pro-Ser-Lys-Arg-Pro-Arg-Pro-Ser (aa 478 to 485) functions as a strong nuclear localization signal, while a weaker karyophilic signal is provided by the Gly-Arg repeat [Arg-Gly-Glu-Ser-Arg-(Gly-

Arg)₅] located between aa 341 and 355. The strong carboxy-terminal nuclear localization signal is identical in HVP EBNA-2 (Fig. 10A). Interestingly, the weaker signal of EBV EBNA-2 does not exist in HVP EBNA-2, which just 9 aa downstream from this location, contains a virtual duplicate of the carboxy-terminal nuclear localization signal (Fig. 10B). Thus, HVP EBNA-2 is also likely to contain two independent karyophilic signals.

DISCUSSION

The EBNA-2 protein is essential for the establishment of a latent infection in vitro (10, 11, 21). Several of the EBV-induced changes in the infected B cell, such as increased expression of surface activation antigens, are attributable to the action of EBNA-2 (8, 12, 46, 47) and EBNA-2 has also been shown to activate the viral latency C promoter and to up-regulate expression of latency membrane proteins LMP-1 and LMP-2 (16, 44, 45, 50, 52). In each of these cases, EBNA-2's contribution is that of a transcriptional activator. To gain further insight into EBNA-2 domain structure, we cloned and sequenced the EBNA-2 gene of the related baboon virus HVP and compared the DNA sequence and predicted amino acid sequence with those of EBV EBNA-2.

All EBV isolates examined to date have EBNA-2 genes that fall cleanly into either the type A or type B group. Other variants or natural intermediates have not been detected. Type B virus was originally recognized in B-cell lines from Africa and Papua New Guinea and thought to be of restricted geographical distribution. More recently, it has become apparent that these virus types are equally widespread and that less efficient in vitro immortalization of B cells by type B virus had led to its underrepresentation in B-cell lines established from EBV-seropositive donors (2, 31, 39). Sixbey et al. (41), using the PCR, found type A and type B virus to have similar prevalence rates in throat washings from a healthy American donor population. In this study, only type A virus was detectable in the peripheral blood from donors with both type A and type B virus in the oropharynx, an observation that is in line with the reduced efficiency of in vitro immortalization of B cells by type B virus and supports the notion of a preferential epithelial tropism for type B EBV. HVP EBNA-2 is neither type A nor type B, being equally distant from the A and B groupings at both the DNA and amino acid levels. This observation would suggest that segregation into type A and type B viruses is a relatively recent phenomenon. Examination of the EBNA-2 genes of additional primate viruses will determine whether the two subtypes are indeed restricted to human virus isolates.

Transfer of potential EBNA-2 karyophilic signals into an HSV Δ IE175 test protein revealed that aa 478 to 485 specify a prototypic strong nuclear localization signal, Pro-Ser-Lys-Arg-Pro-Arg-Pro-Ser. This signal is present at the equivalent position of HVP EBNA-2, and a closely related motif is duplicated at aa 398 to 402 of the HVP protein. In place of this latter motif, the EBV EBNA-2 proteins contain a Gly-Arg repeat (aa 341 to 355) that also has some characteristics of a nuclear localization signal, although this motif functions much less efficiently in our assays, where it produced a mixed nuclear and cytoplasmic localization pattern in both transfected Vero cells and DG75 lymphoblastoid cells. The presence of two nuclear localization signals in EBNA-2 is consistent with previous observations on EBNA-2 function (10) and may help ensure appropriate localization under circumstances, such as protein-protein interactions, when one of the signals may be masked.

The frequently modular nature of the activator and targeting domains of transactivator proteins was exploited to create chimeric proteins containing the DNA recognition domain of the yeast transcriptional activator Gal4 fused to segments of EBV EBNA-2 that spanned the complete pro-

tein. When assayed for transactivation function, the chimeras identified a 40-aa activation domain located between aa 437 and 477 of EBNA-2. The amino-terminal boundary of this domain was defined by a deletion which removed aa 437 to 440 (Pro-Ile-Leu-Phe) and reduced transactivation function by 98% (Fig. 4B). The definition of aa 437 as demarcating the amino-terminal boundary of the activation domain differs from an earlier assessment (9), which placed the minimal activation domain between aa 449 and 462 (corresponding to aa 453 to 466 in the B95-8 strain). It is clear from the more-detailed analysis presented here that mutagenesis of aa 437 to 441 (Pro-Ile-Leu-Phe) results in diminished transactivation efficiency and that all constructions whose amino terminus lies at or beyond aa 441 are significantly impaired in their transactivation ability. Further, the same Pro-Ile-Leu-Phe motif occurs at the amino terminus of a segment of HVP EBNA-2 that also functioned as a strong activator domain when fused to the DNA binding domain of Gal4.

Common features of the activator domains of HVP and EBV EBNA-2 are the presence of negatively charged amino acids (11 or 12 in EBV and 8 in HVP) and interspersed proline residues at least one of which, the proline of the Pro-Ile-Leu-Phe motif, was shown by mutagenesis to be important for activation function (Fig. 7B). The most notable feature, however, is the high degree of positional conservation of hydrophobic residues (Fig. 7A). Whereas the overall

FIG. 8. Multimerization of an EBNA-2 response region in the C_p and determination of the importance of the activator domain for C_p transactivation by EBNA-2. (A) Dose response of CAT constructions containing increasing copy numbers of a previously defined EBNA-2 response element. Effector (SG5/EBNA-2) plasmid was cotransfected into DG75 cells with a constant amount (2.0 μ g) of target CAT constructions containing a minimal promoter (E1B/CAT) or one (1 $\times$), two (2 $\times$), four (4 $\times$), or eight (8 $\times$) copies of the EBNA-2 response element cloned into the parent E1B/CAT plasmid. (B) Requirement for a transcriptional activator domain for EBNA-2 transactivation of C_p . Shown above each lane are the effector EBNA-2 constructions containing either the natural EBNA-2 activation domain, no activation domain, or the EBNA-2 domain replaced by activation domains from VP16 or Rta. The target CAT constructions used are indicated at the bottom of the figure below the brackets. The percent conversion of [14 C]chloramphenicol (% Ac) is given under each lane.

FIG. 9. Identification of karyophilic signals within EBNA-2. Immunofluorescence photomicrographs of Vero cells transfected with various constructions. (A) An internally deleted HSV Δ IE175 gene (pGH115), (B) Δ IE175 expressing EBNA-2 aa 356 to 368 (pPDL25a), (C) Δ IE175 expressing EBNA-2 aa 341 to 355 (pPDL50a), (D) Δ IE175 expressing EBNA-2 aa 341 to 368 (pPDL51a), (E) Δ IE175 expressing EBNA-2 aa 478 to 485 (pPDL22), (F) Δ IE175 expressing EBNA-2 aa 478 to 485 in the antisense orientation (pPDL23). In panels G and H, the EBV-negative B-cell line DG75 was transfected with pGH115 and PDL50a, respectively. The Δ IE175 polypeptide was detected by indirect immunofluorescence with the 58S monoclonal antibody, and a FITC-conjugated goat anti-mouse immunoglobulin G antiserum.

content of prolines and negatively charged residues is very similar in the three EBNA-2 domains, there is not the same strict positional conservation that is seen with the hydrophobic residues. This prominent hydrophobic pattern is not readily discernible in comparisons restricted to the type A and B EBV proteins, because their entire activation domains are virtually identical. The pattern becomes striking only when the more-divergent HVP EBNA-2 domain is included. A previous analysis (13) of a number of glutamine-rich, proline-rich, and negatively charged activation domains had drawn attention to the frequent occurrence of a particular alignment of six hydrophobic residues. The EBNA-2 activation domain can be represented such that there is a degree of match to this hydrophobic alignment (Fig. 7A), and this

observation, along with the positional conservation of the hydrophobic residues in the EBNA-2 proteins, reinforces the concept that hydrophobic residues play an important role in generating an active domain conformation. However, the EBNA-2 domain also emphasizes the differences that occur within the negatively charged transactivator family (which includes VP16 and Gal4 and to which EBNA-2 could be considered to belong). The phenylalanine residue that is essential for VP16 activation function (13) is not conserved as a hydrophobic residue in the HVP or EBV EBNA-2 domain, although it is interesting that each of these EBNA-2 domains contains a phenylalanine in the Pro-Ile-Leu-Phe motif. Removal of this quadruplet abolished the activation function of the EBNA-2 domain. Further mutagenesis of

TABLE 1. EBNA-2 karyophilic signal sequences

EBNA-2 insert in Δ IE175 plasmid	% of positive Vero cells with the following staining pattern ^a :		
	C	N>C	N
None	100	0	0
aa 341–355 (RGQSRGRGRGRGR)	84	16	0
aa 356–368 (GKGKSRDKQRKPG)	96	4	0
aa 341–368 (RGQSRGRGRGRGR <u>IS</u> GKGKSRDKQRKPG) ^b	70	30	0
aa 478–485 (PSKRPRPS)	0	13	87

^a C, cytoplasmic staining; N>C, mixed staining but nuclear staining more intense than cytoplasmic staining; N, nuclear staining.

^b The underlined amino acids are specified by linker DNA.

this motif revealed that while all three hydrophobic residues contributed to the wild-type activation level, mutation of the phenylalanine had the most deleterious effect.

The mechanisms by which EBNA-2 activates different responsive promoters have not been established and may include indirect pathways. It has been suggested that EBV EBNA-2 (458 to 474) has similarity to the motif in adenovirus E1A, papillomavirus E7 and simian virus 40 T antigen involved in binding to the retinoblastoma product Rb (22). However, the homology at positions important for Rb binding (15) is relatively poor, and comparison with HVP EBNA-2 in this region reveals an even more disparate sequence. Further, the proposed Rb binding site lies within the boundaries of the EBNA-2 transactivation domain. We showed that EBNA-2 activation of the C promoter occurred with EBNA-2 proteins in which the entire EBNA-2 activation domain had been deleted and replaced with the activator domain from either the VP16 or Rta protein. One implication of this finding is that the EBNA-2 domain from aa 437 to 477 is most likely a dedicated activation domain and is unlikely to be involved in other aspects of EBNA-2 function.

HPV EBNA-2 diverges sufficiently from EBV EBNA-2 to be ideally suited for comparative studies. Thus far, the sequence data has been used to provide insight into the critical amino acid motifs involved in EBNA-2 transcriptional activation and nuclear localization. The demonstration that conserved motifs in these domains did indeed represent

functionally important sequences suggests that the remaining conserved motifs (boxed in Fig. 2) are also likely to demarcate regions of the EBNA-2 protein involved in critical functions such as protein-protein interactions and promoter targeting.

ACKNOWLEDGMENTS

We thank Jianchao Zong, Mabel Chiu, and Gary Hamilton for technical assistance; MaryAnn Mullen for help with the nuclear localization assays; and Sarah Heaggans for manuscript preparation. We also thank William Schubach for providing EBNA-2 antiserum, Sam Speck for providing DG75 cells, and Gary Hayward for aiding in the sequencing and analysis of HVP EBNA-2.

This work was funded by Public Health Service grants R01 CA42245 (S.D.H.) and R01 CA22130 (G.S.Hayward). P.D.L. was supported by training grants 5T32 CA09243 and F32 CA09165, and J.J.R. was supported by training grant 5T32 NS07000.

REFERENCES

- Abbott, S. D., M. Rowe, K. Cadwallader, A. Ricksten, J. Gordon, F. Wang, L. Rymo, and A. B. Rickinson. 1990. Epstein-Barr virus nuclear antigen 2 induces expression of the virus-encoded latent membrane protein. *J. Virol.* **64**:2126–2134.
- Aldinger, H., H. Delius, U. Freese, J. Clarke, and G. Bornkamm. 1985. A putative transforming gene of the Jijoye virus differs from that of Epstein-Barr virus prototypes. *Virology* **141**:221–234.
- Alfieri, C., M. Birkenbach, and E. Kieff. 1991. Events in Epstein-Barr virus-infection of human B lymphocytes. *Virology* **181**:595–608.
- Alday, M. J., D. H. Crawford, and B. E. Griffin. 1989. Epstein-Barr virus latent gene expression during the initiation of B cell immortalization. *J. Gen. Virol.* **70**:1755–1764.
- Ambinder, R. F., M.-A. Mullen, Y.-N. Chang, G. S. Hayward, and S. D. Hayward. 1991. Functional domains of Epstein-Barr virus nuclear antigen EBNA-1. *J. Virol.* **65**:1466–1478.
- Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, 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.
- Bodescot, M., M. Perricaudet, and P. J. Farrell. 1987. A promoter for the highly spliced EBNA family of RNAs of Epstein-Barr virus. *J. Virol.* **61**:3424–3430.
- Calender, A., M. Billaud, J. P. Aubry, J. Bauchereau, M. Vuillaume, and G. Lenoir. 1987. Epstein-Barr virus (EBV) induces expression of B cell activation markers in *in vitro* infection of EBV-negative B-lymphoma cells. *Proc. Natl. Acad. Sci. USA* **84**:8060–8064.
- Cohen, J. I., and E. Kieff. 1991. An Epstein-Barr virus nuclear protein 2 domain essential for transformation is a direct transcriptional activator. *J. Virol.* **65**:5880–5885.
- Cohen, J. I., F. Wang, and E. Kieff. 1991. Epstein-Barr virus nuclear protein 2 mutations define essential domains for transformation and transactivation. *J. Virol.* **65**:2545–2554.

FIG. 10. Comparison of karyophilic motifs in the HVP and EBV EBNA-2 proteins. (A) The sequence of the B95-8 karyophilic signal (aa 478 to 485) is shown aligned with the carboxy-terminal amino acids of HVP and AG876 EBNA-2. The stippled box indicates a motif that is strongly conserved in all three EBNA-2 proteins. (B) The sequence surrounding the weaker B95-8 karyophilic signal (aa 341 to 355 [open box]) aligned with the corresponding sequences from the HVP and AG876 EBNA-2 polypeptides. The stippled box indicates a motif in the HVP sequence that is closely related to the strong carboxy-terminal karyophilic signal. Asterisks denote identical amino acids, while the vertical lines denote conserved amino acids.

11. Cohen, J. I., F. Wang, J. Mannick, and E. Kieff. 1989. Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. *Proc. Natl. Acad. Sci. USA* 86:9558-9562.
12. Cordier, M., A. Calender, M. Billaud, U. Zimmer, G. Roussellet, O. Pavlish, J. Banchereau, T. Tursz, G. Bornkamm, and G. M. Lenoir. 1990. Stable transfection of Epstein-Barr virus (EBV) nuclear antigen 2 in lymphoma cells containing the EBV P3HR1 genome induces expression of B-cell activation molecules CD21 and CD23. *J. Virol.* 64:1002-1013.
13. Cress, W. D., and S. J. Triezenberg. 1991. Critical structural elements of the VP16 transcriptional activation domain. *Science* 251:87-90.
14. Dambaugh, T., K. Hennessey, L. Chamnankit, and E. Kieff. 1984. U2 region of Epstein-Barr virus DNA may encode Epstein-Barr virus nuclear antigen 2. *Proc. Natl. Acad. Sci. USA* 81:7632-7636.
15. Dyson, N., R. Bernards, S. H. Friend, L. R. Gooding, J. A. Hassell, E. D. Major, J. M. Pipas, T. Vandyke, and E. Harlow. 1990. Large T antigens of many polyomaviruses are able to form complexes with the retinoblastoma protein. *J. Virol.* 64:1353-1356.
16. Fahraeus, R., A. Jansson, A. Rickinson, and L. Rymo. 1990. Epstein-Barr virus encoded nuclear antigen 2 activates the viral latent membrane promoter by modulating the activity of a negative regulatory element. *Proc. Natl. Acad. Sci. USA* 87:7390-7394.
17. Falk, L., F. Deinhardt, M. Nonoyama, L. G. Wolfe, and C. Bergholz. 1976. Properties of a baboon lymphotropic herpesvirus related to Epstein-Barr virus. *Int. J. Cancer* 18:798-807.
18. Garcia-Bustos, J., J. Heitman, and M. N. Hall. 1990. Nuclear protein localization. *Biochem. Biophys. Acta* 1071:83-101.
19. Gerber, P., S. S. Kalter, G. Schidlovsky, W. D. Peterson, Jr., and M. D. Daniel. 1977. Biologic and antigenic characteristics of Epstein-Barr virus-related herpesviruses of chimpanzees and baboons. *Int. J. Cancer* 20:448-459.
20. Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* 2:1044-1051.
21. Hammerschmidt, W., and B. Sugden. 1989. Genetic analysis of immortalizing functions of Epstein-Barr virus in human B lymphocytes. *Nature (London)* 340:393-397.
22. Harada, S., and K. Yauagi. 1991. Amino acid sequence homology of Epstein-Barr virus nuclear antigen-2 to adenovirus E1A, human papilloma virus-16 E7 and SV40 large T antigen in the functional domains for growth transformation. *Oncogene* 6:461-462.
23. Hardwick, M., L. Tse, N. Applegren, J. Nicholas, and M. A. Veluona. 1992. The Epstein-Barr virus R transactivator (Rta) contains a complex, potent activation domain with properties different from those of VP16. *J. Virol.* 66:5500-5508.
24. Heller, M., and E. Kieff. 1981. Colinearity between the DNAs of Epstein-Barr virus and herpesvirus papio. *J. Virol.* 37:821-826.
25. Jin, X. W., and S. H. Speck. 1992. Identification of critical *cis* elements involved in mediating Epstein-Barr virus nuclear antigen 2-dependent activity of an enhancer located upstream of the viral *Bam*HI C promoter. *J. Virol.* 66:2846-2852.
26. Knutson, J. C. 1990. The level of *c-fgr* RNA is increased by EBNA-2, an Epstein-Barr virus gene required for B-cell immortalization. *J. Virol.* 64:2530-2536.
27. Lieberman, P. M., J. M. Hardwick, and S. D. Hayward. 1989. Responsiveness of the Epstein-Barr virus *NotI* repeat promoter to the Z transactivator is mediated in a cell-type-specific manner by two independent signal regions. *J. Virol.* 63:3040-3050.
28. MacMahon, E. M. E., J. D. Glass, S. D. Hayward, R. B. Mann, P. S. Becker, P. Charache, J. McArthur, and R. F. Ambinder. 1991. Epstein-Barr virus in AIDS-related primary central nervous system lymphoma. *Lancet* 338:969-973.
29. Miller, G. 1990. Epstein-Barr virus, p. 1921-1958. In B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), *Virology*, 2nd ed. Raven Press, New York.
30. Pizzorno, M. C., M.-A. Mullen, Y.-N. Chang, and G. S. Hayward. 1991. The functionally active IE2 immediate-early regulatory protein of human cytomegalovirus is an 80-kilodalton polypeptide that contains two distinct activator domains and a duplicated nuclear localization signal. *J. Virol.* 65:3839-3852.
31. Rickinson, A. B., L. S. Young, and M. Rowe. 1987. Influence of the Epstein-Barr virus nuclear antigen EBNA2 on the growth phenotype of virus-transformed B cells. *J. Virol.* 61:1310-1317.
32. Rooney, C., J. G. Howe, S. H. Speck, and G. Miller. 1989. Influences of Burkitt's lymphoma and primary B cells on latent gene expression by the nonimmortalizing P3J-HR-1 strain of Epstein-Barr virus. *J. Virol.* 63:1531-1539.
33. Rooney, C. M., M. Brimmell, M. Buschle, G. Allan, P. J. Farrell, and J. L. Kolman. 1992. Host cell and EBNA-2 regulation of Epstein-Barr virus latent-cycle promoter activity in B lymphocytes. *J. Virol.* 66:496-504.
34. Rowe, M., D. T. Rowe, C. D. Gregory, L. S. Young, P. J. Farrell, H. Rupani, and A. B. Rickinson. 1987. Differences in B cell growth phenotype reflect novel patterns of Epstein-Barr virus latent gene expression in Burkitt's lymphoma cells. *EMBO J.* 6:2743-2751.
- 34a. Ryon, J. J., E. D. Fixman, C. Houchens, J. Zong, P. M. Lieberman, Y.-N. Chang, G. S. Hayward, and S. D. Hayward. The lytic origin of herpesvirus papio is highly homologous to Epstein-Barr virus ori-Lyt: evolutionary conservation of transcriptional activation and replication signals. *J. Virol.*, in press.
35. Sample, J., M. Hummel, D. Braun, M. Birkenbach, and E. Kieff. 1986. Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a probable transcriptional initiation site. *Proc. Natl. Acad. Sci. USA* 83:5096-5100.
36. Sample, J., L. Young, B. Martin, T. Chatman, E. Kieff, A. Rickinson, and E. Kieff. 1990. Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes. *J. Virol.* 64:4084-4092.
37. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
38. Schubach, W. H., G. Horvath, B. Spoth, and J. C. Hearing. 1991. Expression of Epstein-Barr virus nuclear antigen 2 in insect cells from a baculovirus vector. *Virology* 185:428-431.
39. Sculley, T. B., A. Apollin, L. Hurren, D. J. Moss, and D. A. Cooper. 1990. Co-infection with A and B-type Epstein-Barr virus in human immunodeficiency virus-positive subjects. *J. Infect. Dis.* 162:643-648.
40. Showalter, S. D., M. Zweig, and B. Hampar. 1981. Monoclonal antibodies to herpes simplex virus type 1 proteins including the immediate-early protein ICP 4. *Infect. Immun.* 34:684-692.
41. Sixbey, J., P. Shirley, P. Chesney, D. Buntin, and L. Resnick. 1989. Detection of a second widespread strain of Epstein-Barr virus. *Lancet* ii:761-765.
42. Speck, S., and J. Strominger. 1985. Analysis of the transcript encoding the latent Epstein-Barr virus nuclear antigen 1: a potentially polycistronic message generated by long range splicing of several exons. *Proc. Natl. Acad. Sci. USA* 82:8305-8309.
43. Staal, S. P., R. F. Ambinder, W. E. Beschoner, G. S. Hayward, and R. Mann. 1989. A survey of Epstein-Barr virus DNA in lymphoid tissue. Frequent detection in Hodgkin's disease. *Am. J. Clin. Pathol.* 91:1-5.
44. Sung, N. S., S. Kenney, D. Gutsch, and J. S. Pagano. 1991. EBNA-2 transactivates a lymphoid-specific enhancer in the *Bam*HI C promoter of Epstein-Barr virus. *J. Virol.* 65:2164-2169.
45. Sung, S.-F., F. Wang, K. M. Izumi, and E. Kieff. 1991. Delineation of the *cis*-acting element mediating EBNA-2 transactivation of latent infection membrane protein expression. *J. Virol.* 65:6765-6771.
46. Wang, F., C. Gregory, M. Rowe, A. B. Rickinson, D. Wang, M. Birkenbach, H. Kikutani, T. Kishimoto, and E. Kieff. 1987. Epstein-Barr virus nuclear antigen 2 specifically induces expression of the B cell activation antigen CD23. *Proc. Natl. Acad. Sci. USA* 84:3452-3456.
47. Wang, F., C. Gregory, C. Sample, M. Rowe, D. Liebowitz, R. Murray, A. B. Rickinson, and E. Kieff. 1990. Epstein-Barr virus latent membrane protein (LMP1) and nuclear proteins 2 and 3C are effectors of phenotypic changes in B lymphocytes: EBNA-2

- and LMP1 cooperatively induce CD23. *J. Virol.* **64**:2309–2318.
48. Wang, F., H. Kikutani, S. Tsang, T. Kishimoto, and E. Kieff. 1991. Epstein-Barr virus nuclear protein 2 transactivates a *cis*-acting CD23 DNA element. *J. Virol.* **65**:4101–4106.
49. Wang, F., S.-F. Tsang, M. G. Kurilla, J. I. Cohen, and E. Kieff. 1990. Epstein-Barr virus nuclear antigen 2 transactivates latent membrane protein LMP1. *J. Virol.* **64**:3407–3416.
50. Woisetschlager, M., X. W. Jin, C. N. Yandava, L. A. Furmanski, J. L. Strominger, and S. H. Speck. 1991. Role for the Epstein-Barr virus nuclear antigen 2 in viral promoter switching during initial stages of infection. *Proc. Natl. Acad. Sci. USA* **88**:3942–3946.
51. Woisetschlager, M., C. N. Yandava, L. A. Furmanski, J. L. Strominger, and S. H. Speck. 1990. Promoter switching in Epstein-Barr virus during the initial stages of infection of B lymphocytes. *Proc. Natl. Acad. Sci. USA* **87**:1725–1729.
52. Wu, T.-C., R. B. Mann, P. Charache, S. D. Hayward, S. Staal, B. C. Lambe, and R. F. Ambinder. 1990. Detection of EBV gene expression in Reed-Sternberg cells of Hodgkins disease. *Int. J. Cancer* **46**:801–804.
53. Zimmer-Strobl, U., K.-O. Suentzenich, G. Laux, D. Eick, M. Cordier, A. Calender, M. Billaud, G. M. Lenoir, and G. W. Bornkamm. 1991. Epstein-Barr virus nuclear antigen 2 activates transcription of the terminal protein gene. *J. Virol.* **65**:415–423.