Volume 16 Number 8 1988 Nucleic Acids Research

Mutational analysis of Epstein–Barr virus nuclear antigen 1 (EBNA 1)

Maryellen Polvino-Bodnar, Janette Kiso and Priscilla A.Schaffer*

Laboratory of Tumor Virus Genetics, Dana-Farber Cancer Institute and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA

Received October 29, 1987; Revised and Accepted March 10, 1988

ABSTRACT

We have constructed a set of nonsense mutants in the EBNA 1 gene of Epstein-Barr virus by inserting a synthetic oligonucleotide, which has translational termination codons in all three reading frames, at various positions in a cloned copy of the EBNA 1 gene. The EBNA 1 proteins encoded by these mutants and three deletion mutants were analyzed using several functional assays. It was determined that there are two seperable phosphorylation domains in the carboxy half of the molecule. The carboxy half of the molecule was also found to contain a region between the unique Sac I and Sac II sites that is required for transactivation of the EBNA 1-specific enhancer element found within ori P. The mutants also served to identify a 248 bp region that affects the pattern of intranuclear localization of the protein. Correlations between the functional domains established by these studies and other properties of EBNA 1 are discussed.

INTRODUCTION

Epstein-Barr virus (EBV) nuclear antigen (EBNA) is a complex of at least five virally encoded nuclear proteins or protein complexes (EBNAs 1-5) which are expressed in cells immortalized by the virus. EBNA 2 is a single protein and has been implicated in the initiation of growth transformation (1,2,3). In strain P3HR-1 the reading frame for this gene is deleted (Bam HI Y and H) and this virus is unable to immortalize B-lymphocytes. Also, transfer of the EBNA 2 gene into Rat-1 cells results in the generation of cell lines that exhibit reduced serum requirements (4). EBNA 3 is a large, 140 kd protein encoded in part by the Bam HI E fragment (1,5-8). The size of this protein varies among different viral isolates most likely because of the presence of repeated sequences that are found within the open reading frame (6). The function of this protein remains unclear at this point. EBNA 4 is another high molecular weight protein (140-180 kd) found in all cells immortalized by EBV (9). Neither the viral gene nor a function for this protein has been identified. EBNA 5 is composed of a group of low molecular weight proteins which vary in size

by the number of Bam HI W repeats that are transcribed into the mRNA (2). To date there has been no function assigned to this set of proteins.

EBNA 1 is the best characterized protein of the EBNA complex. The open reading frame has been mapped to the left half of the Bam HI K fragment (10,11,12). The mRNA encoding EBNA 1 is a highly spliced species of 3.7 kb with the coding exon being the 3'-most sequence in the mRNA (12,13). The 5' sequences of the message map as far as 80 kb upstream and probably include another open reading frame for a polypeptide of approximately 15 kilodaltons (14,15). Since the 5' end of the message has not been precisely mapped it is not clear where the promoter for this message lies. There is some evidence however that in the IB4 cell line the mRNA starts from a promoter that lies within the Bam HI W fragment (16).

In SDS-polyacrylamide gels EBNA 1 migrates with an apparent molecular weight of 78 kilodaltons while the size predicted from the open reading frame is 56 kilodaltons. This discrepancy is likely due to post-translational modification of the protein, which has been shown to be phosphorylated (17). To date only a single phosphorylated species has been identified with all of the phosphorylation occuring at serine residues (17). EBNA 1 has a peculiar amino acid composition. Approximately 40% of the protein consists of a glycine-alanine (gly/ala) copolymer encoded by the IR3 region of the genome. Other salient features are a short run of serines occuring just downstream of the gly/ala copolymer, a proline-arginine-rich region in the C-terminal quarter of the molecule, and a markedly acidic carboxy terminus with 17 of 41 amino acids being either glutamic or aspartic acid.

In the past several years it has become evident that EBNA 1 is a multifunctional protein. One of the first properties described was its ability to bind to metaphase chromosomes (18). It is still unclear whether this is a direct or indirect interaction. More recently it has been shown that EBNA 1 is required to drive replication from a *cis*- acting element, ori P, which is found in the Bam HI C fragment of the viral genome (19,20,21). There are two *cis*-acting components in ori P sequences required for plasmid replication in the presence of EBNA 1 (20). The first component consists of 20 copies of a 30 bp direct repeat element while the second component is a 65 bp region of dyad symmetry. While both of these elements are required for replication, it appears as though their spatial arrangement is not critical to ori P function (20). It has been demonstrated

that EBNA 1 interacts directly with ori P by binding to both the repeats and the region of dyad symmetry (22). All of these binding sites share a conserved core binding sequence (22,23). In addition to being required for replication in the presence of EBNA 1, the repeat component can also function as a transcriptional enhancer which is specifically activated by EBNA 1 (24). This specificity is most likely due to the direct binding of EBNA 1 to the repeats. Interestingly, this enhancer activity is both promoter and cell-type specific (24).

Given the variety of functions performed by EBNA 1, the present study was designed to identify functional domains within the protein molecule thereby assigning specific functions to primary amino acid sequences. Nonsense mutations in the cloned EBNA 1 gene were generated and these and several deletion mutants were studied. The truncated mutant proteins were analyzed for the ability to a) exist as a phosphorylated species, b) transactivate the enhancer element found within ori P and c) localize to the nucleus.

MATERIALS AND METHODS

Cell lines.

Both COS-1 and CV-1 cells were grown in Dulbecco modified Eagle's medium (DME) containing 10% fetal calf serum, 0.03% glutamine, and 0.25% $NaHCO_3$ supplemented with penicillin and streptomycin.

Recombinant plasmids

The Bam HI/Hind III I_{1f} fragment containing the EBNA 1 reading frame is shown relative to the viral genome in Figure 1. Although it lacks the EBNA 1 promoter, this fragment contains the splice acceptor for the EBNA 1 mRNA as well as all other cis-acting elements required for the efficient transcription and translation of EBNA 1. The fragment was cloned into the vector $pSVO_d$ which contains the SV40 origin of DNA replication and the SV40 late promoter (25). The resulting plasmid, $pSVO_dI_{1f}$, was modified for mutagenesis by converting the unique Hpa I site, which lies 3' to the open reading frame and the polyadenylation signal for the EBNA 1 mRNA, into an Xho I site (Figure 1). This plasmid was designated pHpX.

To construct the nonsense mutants the plasmid pHpX was partially digested with either Sma I or Hae III under conditions that favor the formation of linear DNA. In the case of Sma I the DNA was cleaved using 1 unit of enzyme per μg DNA in the presence of 50μg per ml ethidium bromide while the conditions for Hae III were 0.1 unit of enzyme per μg

DNA in 50 μg per ml ethidium bromide. In each case the linear molecules were gel purified and ligated to the 16 base oligonucleotide 5' GGCTAGTTAACTAGCC 3'.

The plasmid $pSVO_dI_{1fd}$ was kindly provided by George Miller (26). This plasmid has an in-frame deletion in the IR3 region such that it removes sequences specifying most of the gly/ala copolymer (Figure 1). The plasmids pSVS-1 and pSVS3-1 are derivatives of plasmids pMLPyA2K'S-1 and pMLPyA2K'S3-1 which contain in-frame deletions in EBNA 1 coding sequences (27). The latter two plasmids were kindly provided by Arnold Levine. The Bam HI/Hind III subfragments containing the open reading frame for EBNA 1 were subcloned from the latter two plasmids into $pSVO_d$. The resulting plasmid, pSVS-1, has an in-frame deletion of the large internal Sma I fragment of I_{1f}. pSVS3-1 contains a deletion of the small internal Sma I fragment as well as a Bgl II linker inserted at the deletion junction to restore the reading frame (Figure 1).

The plasmid pTKCATR was constructed from the plasmids pTKCAT (28) and the ori P-containing plasmid pSstII-3 [kindly provided by Bill Sugden] (19). pTKCAT was modified by changing the unique Apa I site to a unique Sac I site by the insertion of a Sac I linker. This site was then used to insert the 750 bp Eco RI/Nco I fragment from pSstII-3 which contains the 20 copies of the 30 bp repeat element required for oriP function. Prior to cloning into the pTKCAT vector, the Eco RI/ Nco I fragment had been modified to a Sac I fragment by the use of synthetic linkers. The construct was isolated with the repeat element in either orientation resulting in the plasmids designated pTKCATR and pTKCATR.

All DNA modifying enzymes and synthetic linkers were used according to the manufacturer's specifications (New England Biolabs, Inc., Beverly, MA).

Transfections

COS-1 cells were used to identify the polypeptides produced by the cloned wild-type and mutant EBNA 1 genes and for immunofluorescence tests. The cells were transfected using DEAE-dextran followed by a 3.5 h. exposure to 100 μM chloroquine as previously described (29).

CV-1 cells were used in tests of the ability of the various EBNA 1 species to transactivate the EBNA 1-specific enhancer element. The cotransfection of CV-1 cells by $Ca^{+2}PO_4^-$ precipitation of supercoiled plasmid DNAs was carried out as previously described (28).

Polypeptide analyses

The cloned mutant EBNA 1 genes were tested for their ability to direct the synthesis of an EBNA 1 polypeptide. 60 mm dishes of COS-1 cells were transfected with 2.5 μg of plasmid DNA per plate. At 48 h. post-transfection the monolayers were washed twice with tris-buffered saline (TBS) and lysed in sample buffer as previously described. Extracts were sonicated for 1 minute and then subjected to discontinuous SDS-polyacrylamide gel electrophoresis as previously described (29). Proteins were electrophoretically transferred to nitrocellulose and the blot probed with a well-characterized human serum (RM) containing antibodies to EBNA 1 (1) followed by ^{125}I-protein A. Filters were then autoradiographed on Kodak XAR-5 film.

To determine if the mutant EBNA 1 polypeptides were phosphorylated, COS-1 cells (60 mm plates) were transfected as described above. At 40 h. post-transfection the medium was removed and replaced with phosphate-free DME containing 2% dialyzed fetal calf serum for 1 hour. This medium was then replaced with phosphate-free DME containing 2% dialyzed fetal calf serum and 100 μCi/ml ^{32}P-orthophosphate, and the cells labelled for 4 hours. To harvest, cells were washed twice in ice-cold TBS, scraped into 1 ml. of ice-cold TBS, pelleted, and resuspended in 100 μl of scaffold buffer A (100mM NaCl, 50mM KCl, 20mM Tris pH 7.5, 0.1mM EDTA, 0.1mM PMSF, 10% glycerol) containing 0.5% NP-40 and 0.1% TritonX-100 and incubated for 30 minutes on ice. This procedure fractionates cytoplasm from nuclei, leaving nuclei intact. Earlier experiments revealed that >90% of EBNA 1 in transfected COS-1 cells was transferred from nuclei into the cytoplasmic fraction using this procedure (unpublished results). Nuclei were then pelleted by spinning in the microfuge for 10 minutes, and the supernatent fluids used for immunoprecipitation. Precipitates were then electrophoresed on a discontinuous SDS-polyacrylamide gel as previously described (29). Gels were fixed, dried and autoradiographed.

CAT assays

CV-1 cells were cotransfected with 5μg of pTKCATR and 10μg of either pSVO$_d$ or a plasmid carrying the EBNA 1 gene as previously described. The *in vitro* assay for chloramphenicol acetyltransferase activity was carried out using standard procedures (29,30).

DNA sequencing

To locate precisely the deletion endpoints of the mutant plasmids pBaln1, pBaln3 and pBaln15, the Sac I / Hpa I fragment from each mutant was cloned into M13mp18. The fragments were then sequenced using the Sequenase system (U.S. Biochemical Corp., Cleveland, Ohio).

Indirect Immunofluorescence

COS-1 cells were seeded onto 18 mm coverslips and transfected with 1 μg of plasmid DNA. At 48h. post-transfection the cells were washed once with PBS and fixed in acetone-methanol (2:1) for 10 min. The coverslips were reacted first with nonimmune rabbit serum to block non-specific interactions and then with a human serum (RM) known to contain high titers of antibody to EBNA 1. Rhodamine conjugated goat anti-human immunoglobin (Cappel Laboratories,West Chester, Pa.) was used to visualize specific EBNA 1-antibody complexes.

RESULTS

Construction of mutants

A set of mutants which contain nonsense mutations at intervals throughout the coding sequence for EBNA 1 was constructed. The plasmid pHpX was used as the substrate for mutagenesis. This plasmid is able to direct the synthesis of EBNA 1 at levels that are indistinguishable from $pSVO_dI_{1f}$ when transfected into COS-1 cells (data not shown).

Nonsense mutations were introduced at either Sma I or Hae III sites using a previously described protocol of limited cleavage in the presence of ethidium bromide (31,32). Linear DNA was isolated and then ligated to the 16 base oligonucleotide 5' GGCTAGTTAACTAGCC 3'. This linker has several important features: a) the sequence is self-complementary and thus only one oligonucleotide is required to generate a double-strand linker, b) the double-strand linker has a unique Hpa I site (GTTAAC) which serves as a useful marker, and c) the linker has termination codons in all three reading frames.

Clones were screened for the presence and position of the nonsense linker using the Hpa I marker. Our analysis revealed that we had generated a number of mutants carrying the linker in the first two-thirds of the coding sequence (Figure 1, n203 through n16). Because we had failed to isolate any mutants in the carboxy third of the reading frame, we used the unique Sac I and Sac II sites in that region to insert the nonsense linker to yield n500 and n424 respectively. To provide a control plasmid

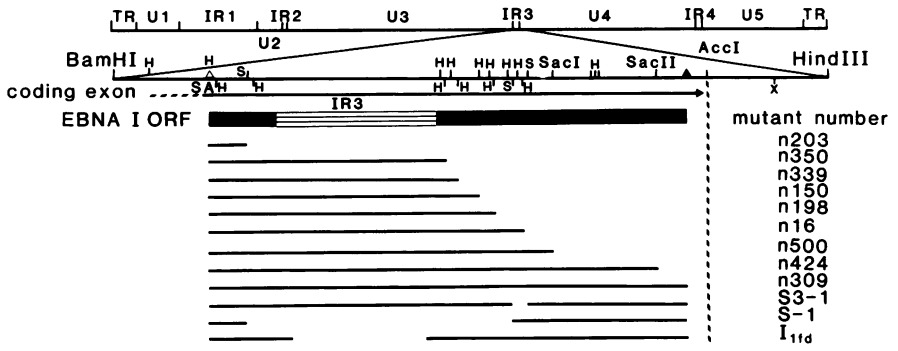

FIGURE 1. Location of the EBNA I reading frame in the Epstein-Barr virus genome. The Bam HI/ Hind III subfragment I_{1f} was subcloned into the vector $pSVO_d$. The unique Hpa I site near the Hind III end of the molecule was then converted to an Xho I site by the use of synthetic linkers. The resulting plasmid pHpX was used as the substrate for insertional mutagenesis. The coding exon for EBNA 1 is shown beginning with the position of the splice acceptor (SA). Beneath the coding exon the open reading frame (ORF) is shown. The relevant restriction sites in the fragment are Hae III (H), Sma I (S), Sac I, Sac II, Acc I and Xho I (X). The open reading frame is denoted by ▬ while the internal repeat (gly/ala copolymer, IR3) is shown as ▭.

The predicted sizes of the mutant peptides are shown beneath the open reading frame. This map shows the position of the nonsense linker in each mutant as translation is terminated at that site. In addition to the nonsense mutants (n) three deletion mutants that were included in these studies. The locations of their deletions are shown in the last three lines of the figure.

we also inserted the linker into the unique Acc I site, which lies 3' to the EBNA 1 coding sequence and polyadenylation signal. This plasmid was designated n309.

In addition to the nonsense mutants, two deletion mutants were obtained from A. Levine (pSVS-1 ,pSVS3-1) and one was provided by G. Miller ($pSVO_dI_{1fd}$). All of the deletion and nonsense mutants are shown in Figure 1.

A second set of three mutants was constructed using n424 as the substrate for mutagenesis. n424 DNA was digested with Hpa I (which linearizes the plasmid at the site of the nonsense linker) and then treated with Bal 31. The ends were then repaired with T4 DNA polymerase and the 16 bp nonsense linker ligated to the DNA. Clones containing both a deletion and the nonsense linker were isolated and three were chosen for further study. These plasmids, designated pBaln1, pBaln3 and pBaln15 are shown in Figure 5.

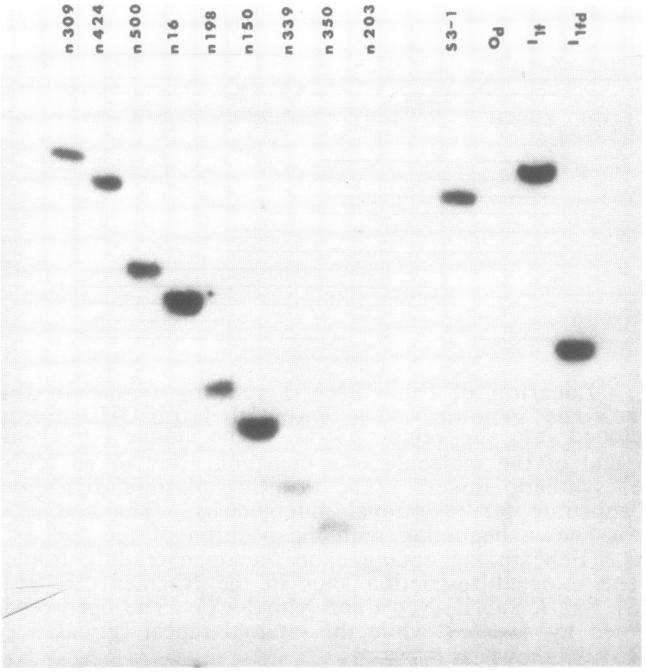

FIGURE 2. Western blot of wild-type and mutant EBNA I polypeptides. 60-mm plates of COS-1 cells were transfected with 2.5 μg of plasmid DNA using DEAE-dextran and chloroquine. 48 hours post-transfection whole cell lysates were harvested, sonicated and run on an SDS-polyacrylamide(10%) gel. Proteins were electrophoretically transferred to nitrocellulose (29) and the blot then probed with a human serum containing antibodies to EBNA I (1:300) and ^{125}I protein A. Each lane represents 1/10 of the total lysate.

EBNA 1 polypeptides encoded by mutants

Mutant plasmids were first analyzed for the ability to direct the synthesis of an EBNA 1 species in transient expression assays. Plasmid DNAs were transfected into COS-1 cells and whole cell extracts were used for Western blotting. The blot was probed with a well-characterized human serum (RM) containing high titers of antibody to EBNA 1 (1). The results are shown in Figure 2. All of the mutants except n203 directed the synthesis of a detectable EBNA 1 polypeptide. In each case the protein produced was appropriately reduced in size indicating that the nonsense linker was effective in directing the termination of translation at the point of linker insertion. It is important to note that only a single band was

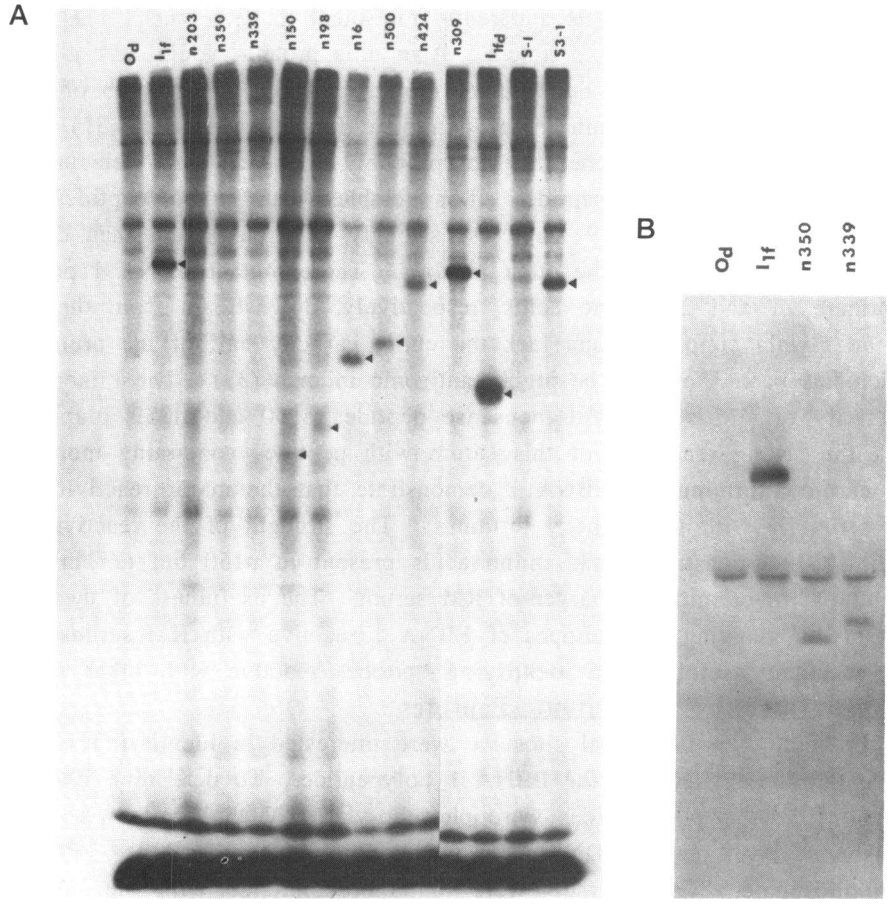

FIGURE 3A. Phosphorylation of wild-type and mutant EBNA 1 polypeptides. COS-1 cells were transfected and pulse-labelled with ^{32}P-orthophosphate. Cell extracts were partially fractionated and the soluble portion used for immunoprecipitation. Precipitates were then subjected to SDS-PAGE on 10% polyacrylamide gels which were fixed, dried and autoradiographed.

Figure 3B. Immunoprecipitation of peptides from n350 and n339. COS-1 cells were transfected and unlabelled cell extracts harvested as described for the phosphorylation study. The extracts were immunoprecipitated and the precipitates seperated by SDS-PAGE on a 10% polyacrylamide gel. Proteins were electrophoretically transferred to nitrocellulose and the blot probed with a second serum containing EBNA 1 antibodies and ^{125}I protein A.

detected in each lane suggesting that "read-through" was uncommon, if it occurred at all. Mutant n203 has a nonsense mutation at the position of amino acid 58 and thus can direct the synthesis of a polypeptide of only 57

amino acids. This protein is apparently too small to detect using the gel system employed in this experiment. Also, this protein would lack all of the gly/ala copolymer and the carboxy terminus of EBNA 1 - both of which have been shown to contain major antigenic determinants (33,34,35). Thus the serum we used to probe the blot probably does not react detectably with the small n203 polypeptide. It is notable that the polypeptides encoded by n350 and n339 react less strongly with the serum than do the EBNA 1 polypeptides of the other mutants. These mutant EBNA 1 peptides terminate 11 and 27 amino acids, respectively, downstream from the end of the gly/ala copolymer and lack the entire carboxy half of the protein which has been shown to be highly antigenic in man (33). The minor reactivity of RM serum with nonsense peptides n350 and n339 coupled with the strong reactivity of this serum with peptides containing more of the carboxy terminus of EBNA 1 demonstrate that the major reactivity of the serum is with the carboxy terminus. The strength of the reactivity with n150 demonstrates that amino acids present in n150 but not in n339 contain a major antigenic target of RM serum. The usefulness of these mutants in mapping the epitopes of EBNA 1 reactive with RM serum suggest their usefulness in identifying epitopes reactive with other sera.

Phosphorylation of Mutant Peptides

Among the functional sites we were interested in identifying is the phosphorylation site(s) of the EBNA 1 polypeptide. To determine if the mutant EBNA 1 polypeptides were phosphorylated, COS-1 cells were transfected with plasmid DNAs and metabolically labelled with ^{32}P-orthophosphate. Cell extracts were immunoprecipitated with the same human serum used for Western blot analysis. All but three of the mutants directed the synthesis of a phosphorylated EBNA 1 species (Fig. 3A). As in the case of the Western blot no n203 EBNA 1 polypeptide was detected. However, both n350 and n339, which were readily detectable on the Western blot, also failed to produce detectable phosphorylated EBNA 1 species. Although this suggests that the shared deleted sequences of these mutants (Fig. 1) contain a phosphorylation site, it is also possible that our failure to detect phosphorylated species from these two mutants is due to the failure of these mutant peptides to be precipitated by RM serum. We therefore tested the precipitability of unlabelled n350 and n339 peptides by RM serum using precisely the same procedure used to precipitate the phosphoproteins analyzed in Fig. 3A. The unlabelled precipitates were analyzed by gel electrophoresis, transferred to nitrocellulose and visualized

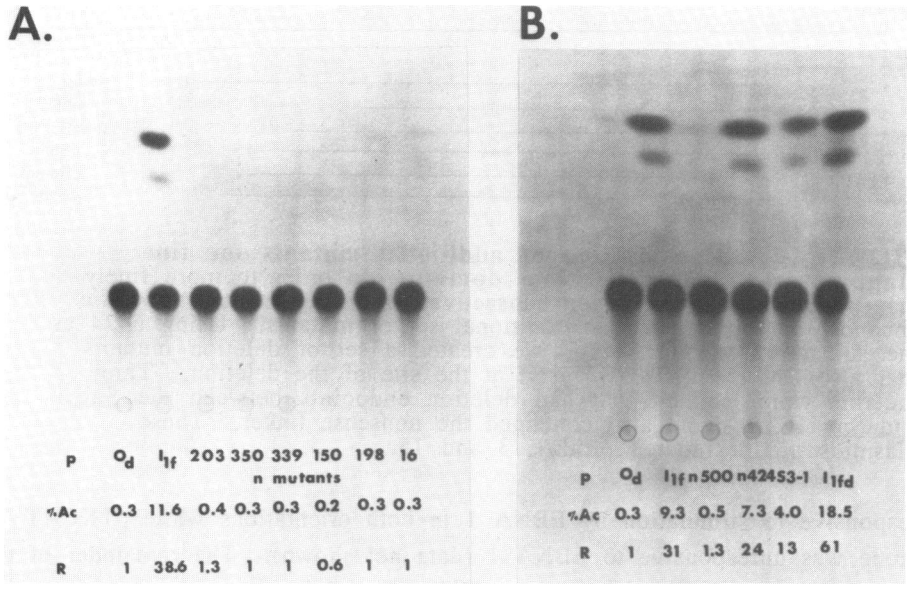

FIGURE 4. Trans-activation of pTKCATR by wild-type and mutant EBNA I polypeptides. 100-mm plates of CV-1 cells were co-transfected with 5 μg of pTKCATR and 10 μg of an EBNA I encoding plasmid using calcium-phosphate precipitation. All transfection mixtures were brought up to 25 μg of DNA by adding 10 μg of salmon testis DNA. Approximately 40h. post-transfection cell extracts were prepared and assayed for CAT activity using standard protocols (29, 30). To determine the uninduced level of CAT activity from pTKCATR 10 μg of $pSVO_d$ was substituted for the EBNA I encoding plasmid. The activity from the wild-type and mutant genes are expressed as the percentage of ^{14}C-chloramphenicol acetylated as well as the ratio (R) of acetylated product in induced to uninduced samples.

using a second antiserum (SHS) (Fig. 3B). Both peptides were readily precipitated by RM serum. Taken together, these tests strongly suggest that the shared deleted sequences in n350 and n339 contain a phosphorylation site.

Transactivation of the EBNA 1-specific enhancer

The mutant plasmids were next tested for the ability to transactivate the transcriptional enhancer element found within ori P (24). This element was used to enhance transcription from the HSV-1 thymidine kinase (tk) promoter which has been shown to be responsive to the enhancer element in the repeat element of ori P. We have tested both pTKCATR and pTKCAT<u>R</u> in CAT assays and found the enhancer element to be equally

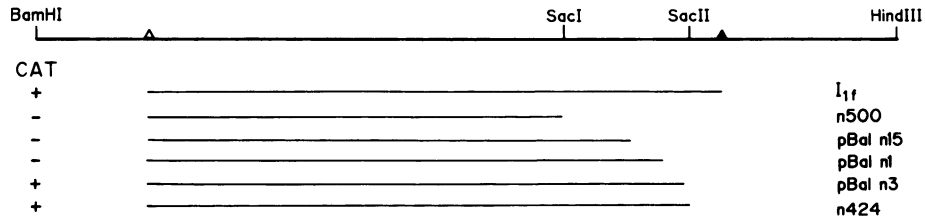

FIGURE 5. Construction of additional mutants for fine mapping the transactivation domain. In order to more finely map sequences required for transactivation of the EBNA 1-specific enhancer we constructed an additional set of mutants. Using n424 as the substrate for mutagenesis we created a set of deletion mutants that also had the nonsense insert at the site of the deletion. Three mutants were isolated that had deletion endpoints between the Sac I and Sac II sites and also contained the nonsense linker. These plasmids are designated pBaln1, 3 and 15.

responsive to stimulation by EBNA 1 in both orientations while pTKCAT alone was unresponsive to EBNA 1 (data not shown). The remainder of the studies were conducted only with pTKCATR. The assays showing the ability of the mutant EBNA 1 polypeptides to transactivate pTKCATR are shown in Figure 4. None of the mutants with nonsense mutations mapping 5' to, and including the mutant n500 exhibited the ability to transactivate. Mutant n424 was able to transactivate to levels equivalent to those of wild-type EBNA 1. Thus, the 396 bp region between the positions of the nonsense linkers in n500 and n424 contributes significantly to the ability of EBNA 1 to transactivate the enhancer element. Notably, the deletion plasmid $pSVO_dI_{1fd}$ was not impaired in its ability to transactivate while the mutant pSVS3-1 exhibited a slight reduction in the ability to transactivate. The deletion in pSVS3-1 does not overlap the region between n500 and n424 but is 108 bp upstream from the position of the nonsense linker in n500.

To map more finely the region bounded by the mutations in n500 and n424, we generated a set of deletion mutants that also carry a nonsense linker at the site of the deletion. The mutants pBaln1, pBaln3 and pBaln15 have deletions of 84, 13, and 198 bp, respectively, as the result of Bal 31 digestion of DNA linearized at the unique Hpa I site of n424. The position of the nonsense mutation in each of these mutants relative to those in n500 and n424 are shown in Figure 5. All three of the mutants were tested for their ability to transactivate pTKCATR as shown in Figure 6. Only pBaln3 was able to transactivate the test plasmid suggesting

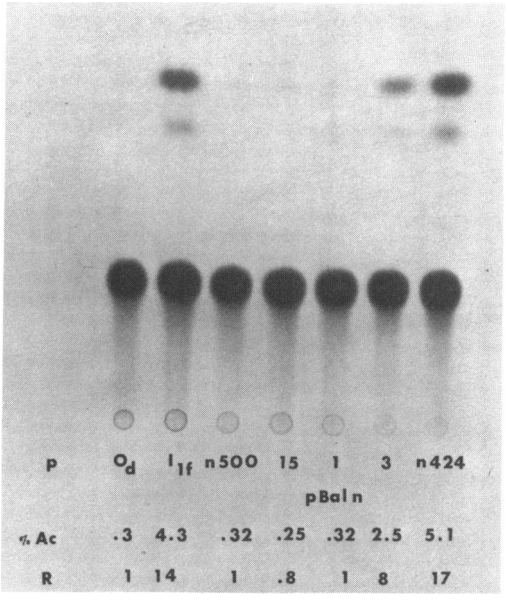

FIGURE 6. CAT assay with the pBaln plasmids. This experiment was carried out as described in the legend to Figure 4. The % acetylation and the ratios of acetylated product in each sample relative to the $pSVO_d$ sample are shown.

that the sequences immediately 5' to the Sac II (n424) site are required for efficient transactivation. It is interesting to note that pBaln1 is shorter than pBaln3by 23 amino acids and fails to transactivate pTKCATR. These data suggest that the 23 amino acids present in pBaln3 and not in pBaln1 are critical for transactivation.

Indirect immuoflourescence

Since EBNA 1 is a nuclear protein it was of interest to us to determine if the mutant EBNA 1 polypeptides were able to localize to the nucleus. Each mutant was therefore tested in transient expression assays by indirect immunofluorescence for the intracellular location of the EBNA 1 peptide. In Figure 7 the results of tests with several representative mutants are shown.

As in the case of the assays described above, mutant n203 failed to produce detectable EBNA 1 specific staining, whereas all of the other mutants tested directed the synthesis of an EBNA 1 protein that localized to the nucleus. Interestingly, the mutants differed in two parameters: intensity of staining and pattern of intranuclear localization. With regard

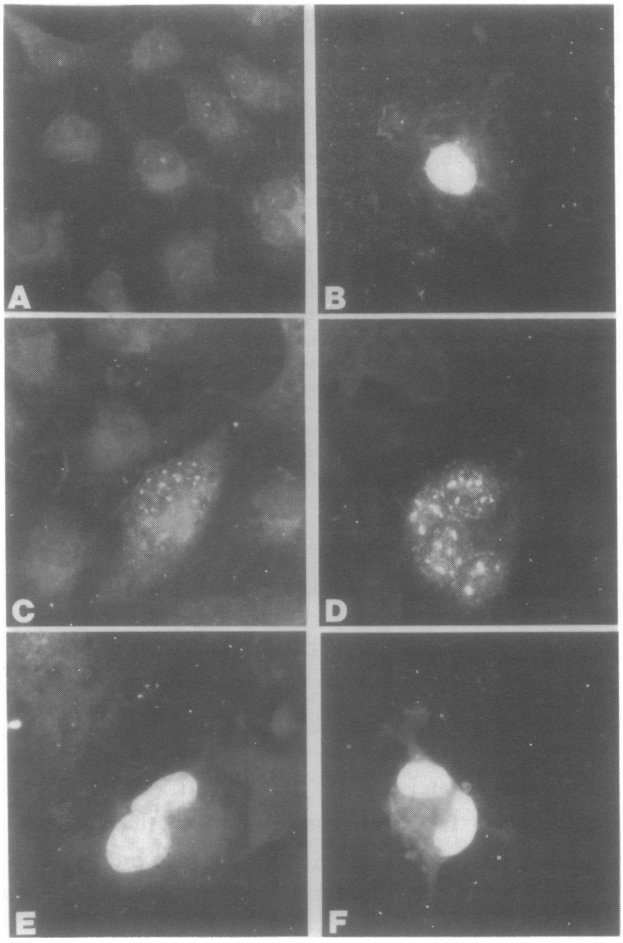

FIGURE 7. Expression of EBNA 1 polypeptides detectable by indirect immunofluorescence. COS-1 cells were seeded onto coverslips, transfected with plasmid DNA, then fixed and stained 48 h. post-transfection. The distribution of EBNA 1 polypeptide in the cell is shown for the following plasmids; A) pSVOd, B) pSVOdI1f, C) n350, D) n339, E) n16 and, F) n500.

to differences in the intensity of EBNA 1 staining among the mutants, mutant EBNA 1 peptides that lack a major epitope with which RM serum reacts (n350, n339) generated less intense fluorescence than those that contain the epitope (n198 through n309). Panels C and D in Figure 7 show the weak fluorescence obtained with mutants n350 and n339 respectively, while panel E shows staining of the n16 peptide. Mutants n198 and n150

(not shown) whose carboxy termini lie between those of n339 and n16 exhibited staining intensities that are similar to that of the n16 peptide. These results correlate well with the intensities of peptide bands obtained from Western blotting experiments (Figure 2) that reveal the presence of a major epitope for RM serum in the sequences that are present in n150 but not n339.

The second parameter in which the mutant peptides differed was their patterns of localization within the nucleus. The wild-type antigen specified by $pSVO_dI_{1f}$ was distributed evenly throughout the nucleus giving rise to a homogeneous diffuse staining pattern as seen in Figure 7, panel B.. Mutants n500 (panel F), n424, n309, pSVS3-1 and, $pSVO_dI_{1fd}$ all exhibited staining patterns indistinguishable from that of $pSVO_dI_{1f}$. By contrast, the mutants in which the nonsense linker is inserted in a position 5' to its location in n500 exhibited a markedly different staining pattern. The n16 pattern was unique and was characterized by a distinct or patchy appearance (panel E). Mutants n198, n150, n339 (panel D) and n350 (panel C) exhibited speckled or grainy staining. Although differences in the staining intensity of the n350 and n339 peptides versus all other peptides can be attributed to their reactivity with RM serum, the differences in the patterns of staining may reflect differences in the quantities of mutant peptides in nuclei or differences in the localization of these peptides in nuclei. Of the two possibilities we favor the latter since 1) expression of all mutant genes is controlled by the same promoter regulatory elements and 2) Western blot analysis yielded mutant bands of approximately equal intensity (with the exception of the n350 and n339 peptides noted above). We conclude therefore that the differences in staining patterns characteristic of the mutant peptides is a consequence of their relative reactivities with RM serum and to their distribution within the nucleus. The latter property is likely conferred-at least in part-by the amino acid sequence lying between the carboxy termini of n198 and n500.

DISCUSSION

This report describes the construction of twelve nonsense mutants and the characterization of these mutants as well as three deletion mutant plasmids of the EBNA 1 gene of Epstein-Barr virus. Eleven of the mutants exclude varying amounts of the carboxy terminus of the protein due to the insertion of a nonsense linker, which contains translational termination codons in all three reading frames, into the coding region of the cloned

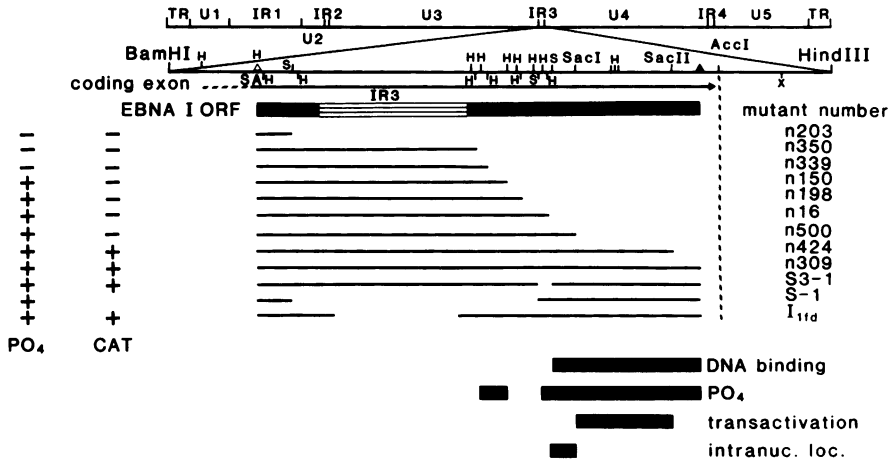

FIGURE 8. Identification of three functional domains in the EBNA 1 molecule. The data on phosphorylation and transactivation for the set of mutants described in this paper are summarized on the left side of this figure. Beneath the mutants the large black boxes show the map positions of the domains for phosphorylation (17), transactivation, intranuclear localization and DNA binding (22) as defined by these mutants and previously published data.

EBNA 1 gene. The twelfth nonsense mutant has the nonsense linker inserted 3' to the open reading frame, the termination codon and the polyadenylation signal for the EBNA 1 polypeptide. The other three mutants contain in-frame deletions that remove sequences in the middle of the open reading frame. The mutant polypeptides encoded by the plasmids were analyzed for the ability a) to be phosphorylated, b) to transactivate an EBNA 1-specific enhancer found within ori P and, c) to localize to the nucleus. The results of these tests are summarized in Figure 8.

Mutant proteins that include the first 400 or more amino acids of the amino terminus of the open reading frame (62%; n150 through n424) were detected as a single phosphorylated species (Figure 3). While it is possible that n350 and n339 direct the synthesis of phosphorylated species that we have failed to detect, we think this is unlikely for the following reason. Between the carboxy terminal amino acid of the peptide specified by n339, which is apparently unphosphorylated, and that of the n150 peptide, which clearly is phosphorylated, are sequences in which 8 of 11 amino acid residues are serine. It has been demonstrated that all of the phosphorylation in EBNA 1 occurs at serine residues(17). The results of

our studies would therefore suggest that this run of serines represents a major phosphorylation site. Although a previously described mutant, pMLPyA2K'S-1 (the plasmid from which pSVS-1 is derived) lacks this run of serines but has been shown to be phosphorylated (17), the mutant retains the carboxy terminal end of the protein which contains several serine residues that are not tightly clustered. Thus, it is probable that there are two phosphorylation sites in the carboxy half of the molecule. Unfortunately, the serum that we used to detect the wild-type and mutant EBNA 1 polypeptides was unable to identify the phosphorylated EBNA 1 species encoded by pSVS-1. Our studies also confirm the finding that the gly/ala copolymer does not play a major role in phosphorylation (17) both because this sequence contains no serine residues and because the polypeptide encoded by $pSVO_dI_{1fd}$ is phosphorylated.

Studies of this set of mutant plasmids indicate that the sequences between amino acids 489 (n500) and 621 (n424) are essential for efficient transactivation of the enhancer element found within ori P. The pBaln mutants were constructed in an effort to define better the sequences neccesary for transactivation. These mutants revealed that sequences between amino acids 597 (pBaln1) and 616 (pBaln3) are required for efficient transactivation as only the polypeptide encoded by pBaln3 was capable of transactivating the enhancer. It is interesting that in the region between amino acids 489 (n500) and 616 (pBaln3) no gradual restoration of the ability to transactivate occurs as more of the open reading frame is expressed. Rather, there seems to be a very sharp demarcation between active and inactive peptides which falls between amino acids 597 (pBaln1) and 616 (pBaln3). The fact that the I_{1fd} polypeptide transactivates at a level equivalent to that of wild-type EBNA 1 confirms earlier reports that the gly/ala copolymer is not required for transactivation (24). The only two mutant polypeptides which exhibited a reduced ability to transactivate the enhancer are pSVS3-1 and pBaln3. The mutant pSVS3-1 lacks the 19 amino acids between residues 473 and 493 and has 4 new residues specified by the presence of a Bgl II linker inserted to restore the reading frame. This suggests that the sequences in this region may contribute to the ability to transactivate the EBNA 1-specific enhancer or may be involved in determining higher order structure which is neccessary for optimal transactivation of the enhancer. While the properties of the mutant peptides described in this paper have revealed a region which is required for the efficient transactivation of the EBNA 1-specific enhancer,

only two peptides have identified sequences that are totally dispensable for transactivation. The peptides from both $pSVO_dI_{1fd}$ and n424 transactivate the enhancer as well as the wild-type EBNA 1. This suggest that neither the internal repeat or the acidic carboxy tail of the protein is required for efficient transactivation. There are likely to be other sequences that are either dispensable for, contribute to, but are not essential for, transactivation. Such domains will best be identified by constructing in-frame deletion mutants.

Although all of the mutant polypeptides that were detectable using RM serum localized to the nucleus, the mutants exhibited two readily distinguishable patterns of intranuclear fluorescence. First, both deletion mutants tested (S3-1 and I_{1fd}) as well as mutants with the nonsense linker in, or 3' to, the Sac I site (n424 and n309) exhibited a diffuse, homogeneous staining pattern throughout the nucleus identical to that observed with the wild-type allele. Mutants having a nonsense insertion 5' to the Sac I site also directed the synthesis of EBNA 1 species that localize to the nucleus but in each case the fluorescence pattern was patchy or stippled in appearance. While the latter mutant EBNA 1 peptides appear to be present throughout the nucleus they are evident as small deposits rather than being diffusely spread throughout. The mutant n16 exhibits an intermediate phenotype between the more 5' mutants and n500. A patchy yet distinctly not diffuse pattern of flourescence is observed with n16 as opposed to the very small deposits seen with the polypeptides encoded by mutants with nonsense mutations 5' to that found in n16. At this time it is not possible to determine if the difference between n16 and the mutants 5' to it is one of intensity of fluorescence or of actual localization. The shorter the peptides have fewer epitopes than n16 and will therefore react with a smaller fraction of the antibodies in the serum used. This decreased reactivity would be visualized as less intense fluorescence whereas the increased reactivity of n16 would give rise to a more intense signal possibly being visualized as somewhat more diffuse. There are 248 bp between the nonsense inserts in n198 (stippled pattern) and n500 (diffuse or wild-type pattern) suggesting that in this region there exist sequences that affect the pattern of intranuclear localization as visualized by immunofluorescence. It is important to note that this 248 bp region is not the region that contains the major epitope to which the EBNA 1 antibodies in RM serum is directed (i.e. between the endpoints of n339 and n150). In general, the amino acid sequence between these two insertion points is

unremarkable, with the notable exception of a cluster of 4 consecutive positively charged amino acids flanked by glycine residues on either side. The sequence reads gly-gly-arg-arg-lys-lys-gly-gly-gly. It is important to note that several documented nuclear localization signals consist of short runs of positively charged amino acids flanked by helix breaking residues such as proline (36,37). It is possible that variations on that motif provide additional signals for proper intranuclear localization.

When the data from the transactivation assays are examined in combination with those from the indirect immunofluorescence tests it is clear that the lack of ability to transactivate seen with some of the mutants is not due to the fact that the EBNA 1 species that they encode fail to reach the nucleus. It is interesting to note that the n500 polypeptide not only reaches the nucleus, but also appears to be localized properly within the nucleus, and yet fails to transactivate the EBNA 1-specific enhancer. This suggests that transactivation requires an interaction of EBNA 1 with the repeat element either alone or in concert with cellular proteins. It has been shown previously that the carboxy-terminal quarter of the EBNA 1 protein contains a DNA-binding domain (22). This domain directs binding specifically to sequences found within the repeat element and the region of dyad symmetry of ori P. The mutant peptides described in this paper show that sequences between the Sac I and Sac II sites contain the most important domain for transactivation and fall within the previously mapped DNA binding domain. At this time DNA binding experiments are in progress to determine if these two domains are one and the same. Interestingly, our findings indicate that the acidic tail of the protein is not required for transactivation and thus potentially narrows the region of the protein involved in DNA binding.

The DNA binding capability of EBNA 1 to sequences within ori P may also correlate with the ability of ori P-containing plasmids to replicate in eucaryotic cells. The core binding sequence that is recognized by EBNA 1 (23) is found in both of the *cis* -acting elements that are required for replication (20). It is likely, therefore that it may prove impossible to uncouple the transactivation and replication functions. Studies analyzing the ability of the mutant EBNA 1 peptides to direct the replication of ori P-containing plasmids in eucaryotic cells are underway.

It is not possible at this time to relate phosphorylation to any of the functions studied to date. Probably the clearest way to establish any such correlation would be to construct mutants that lack only the

phosphorylation sites and conduct the same assays that were used in these studies.

While we have been able to identify functional domains of the EBNA 1 molecule affecting phosphorylation, transactivation and intranuclear localization (the DNA binding region was identified by others (22)) (Fig.8) we have not yet demonstrated interelationships among these domains. Clearly, there is overlap between several domains and there are still some regions of the protein, such as the gly/ala copolymer, that have not yet been associated with any known function. The study of a broader group of deletion and nonsense mutants will help to complete this map and hopefully uncover additional functions performed by EBNA 1.

ACKNOWLEDGEMENTS

We thank N. DeLuca, D. Leib and J. Bodnar for helpful discussions and reviewing of the manuscript. We also thank B. Sugden, A. Levine and G. Miller for providing plasmids used in this study. This work was supported by Public Health Service Grant CA02060 form the National Cancer Institute. M.P.-B. is supported by a Special Fellowship from the Leukemia Society of America.

*To whom correspondence should be addressed

REFERENCES

1. Rowe, D.T., Farrell, P.J., and Miller, G. (1987) Virology **156**:153-162.
2. Dillner, J., Kallin, B., Alexander, H. Ernberg, I., Uno, M, Uno, Y., Klein, G. and R.A. Lerner (1986) Proc. Natl. Acad. Sci. U.S.A. **83**:6641-6645.
3. Dambaugh, T. Hennessy, K., Chamnankit, L., and Kieff, E. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**:7632-7636.
4. Wang, D., Leibowitz, D. and Keiff, E. (1985) Cell **43**:831-840.
5. Hennessy K., Fennewald, S., and Kieff, E. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**:5944-5948.
6. Hennessy, K., Wang, F., Bushman, E.W., and Kieff, E. (1986) Proc. Natl. Acad.Sci. U.S.A. **83**:5693-5697.
7. Sculley,T.B., Sculley, T.G., and Pope, M. (1985) J. Gen. Virol. **66**:1113-1122.
8. Joab, I., Rowe, D.T., Bodescot, M., Nicolas, J.-C., Farrell, P.J., and Perricaudet , M. (1987) **61**:3340-3344.
9. Kallin, B., Dillner, J., Ernberg, I., Ehlin-Henriksson, B. Rosen, A.,. Henle,W., Henle, G., and Klein, G. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**:1499-1503.
10. Hennessy, K. and Kieff, E. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**:5665-5669.
11. Summers, W.P., Grogan, E.A., Shedd, D., Robert, M., Liu, C., and Miller,G. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**:5688-5692.
12. Hennessy, K. Heller, M. van Santen, V., and Kieff, E. (1983) Science **220**:1396-1398.

13. Weigel, R. and Miller, G. (1985) J. Virol. **54**:501-508.
14. Speck S. H. and Strominger, J.L. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**:8305-8309.
15. Bodescot, M., Brison, O., and Perricaudet, M. (1986) Nucleic Acids Res. **14**:2611-2620.
16. Sample, J., Hummel, M., Braun, D., Birkenbach, M., and Kieff, E. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**:5096-5100.
17. Hearing, J.C., and Levine, A.J. (1985) Virology **145**:105-116.
18. Grogan, E.A., Summers, W.P., Dowling, S., Shedd, D., Gradoville, L. and Miller, G. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**:7650-7653.
19. Yates, J., Warren, N., Reisman, D., and Sugden, B. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**:3806-3810.
20. Reisman, D., Yates, J., and Sugden, B. (1985) Mol. Cell. Biol. **5**:1822-1832.
21. Yates, J., Warren N., and Sugden, B. (1985) Nature **313**:812-815.
22. Rawlins, D.R., Milman, G., Hayward, S.D., and Hayward, G.S. (1985) Cell **42**:859-868.
23. Milman, G. and Hwang, E.S. (1985) J. Virol. **61**:465-471.
24. Reisman, D. and Sugden, B. (1986) Mol. Cell. Biol. **6**:3838-3846.
25. Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981) Cell **27**:279-288.
26. Robert, M., Shedd, D., Weigel, R., Fischer, D.K., and Miller, G. (1984) J. Virol. **50**:822-831.
27. Hearing, J.C., Nicholas, J-C., and Levine, A.J. (1984) Proc. Natl. Acad.Sci. U.S.A. **81**:4373-4377.
28. DeLuca, N.A., and Schaffer, P.A. (1985) Mol. Cell. Biol. **5**:1997-2008.
29. Polvino-Bodnar, M., Shedd, D. and Miller, G. (1986) J. Virol.
30. Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982) Mol. Cell. Biol. **2**:1044-1051.
31. Polvino-Bodnar, M., and Cole, C.N. (1982) J. Virol. **43**:489-509.
32. Deluca, N.A., and Schaffer, P.A., (1987) Nucleic Acids Res. **15**:4491-4511.
33. Milman, G., Scott, A.L., Cho, M.-S., Hartman, S.C., Ades, D.K., Hayward, G.S., and Hatward, S.D. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**:6300-6304.
34. Rumpold, H., Rhodes, G.H., Bloch, P.L., Carson, D.A., and Vaughn, J.H.(1987) J. Immunology **138**:593-599.
35. Rhodes, G. Rumpold, H. Kurki, P., Patrick, K.M., Carson, D.A., and Vaughn, J.H. (1987) J. Exp. Med. **165**:1026-1040.
36. Kalderon, D., Richardson, W.D., Markham, A.F., and Smith, A.E. (1984) Nature **311**:33-38.
37. Richardson, W.D., Roberts, B.L., and Smith, A.E. (1986) Cell **44**:77-85.