

Maintenance of Epstein–Barr virus (EBV) *oriP*-based episomes requires EBV-encoded nuclear antigen-1 chromosome-binding domains, which can be replaced by high-mobility group-I or histone H1

Siu Chun Hung, Myung-Soo Kang, and Elliott Kieff*

Program in Virology, Department of Microbiology and Molecular Genetics, Harvard Medical School, and Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115

Contributed by Elliott Kieff, December 11, 2000

EBV-encoded nuclear antigen-1 (EBNA-1) binding to a cis-acting viral DNA element, *oriP*, enables plasmids to persist in dividing human cells as multicopy episomes that attach to chromosomes during mitosis. In investigating the significance of EBNA-1 binding to mitotic chromosomes, we identified the basic domains of EBNA-1 within amino acids 1–89 and 323–386 as critical for chromosome binding. In contrast, the EBNA-1 C terminus (amino acids 379–641), which includes the nuclear localization signal and DNA-binding domain, does not associate with mitotic chromosomes or retain *oriP* plasmid DNA in dividing cell nuclei, but does enable the accumulation of replicated *oriP*-containing plasmid DNA in transient replication assays. The importance of chromosome association in episome maintenance was evaluated by replacing EBNA-1 amino acids 1–378 with cell proteins that have similar chromosome binding characteristics. High-mobility group-I amino acids 1–90 or histone H1–2 could substitute for EBNA-1 amino acids 1–378 in mediating more efficient accumulation of replicated *oriP* plasmid, association with mitotic chromosomes, nuclear retention, and long-term episome persistence. These data strongly support the hypothesis that mitotic chromosome association is a critical factor for episome maintenance. The replacement of 60% of EBNA-1 with cell protein is a significant step toward eliminating the need for noncellular protein sequences in the maintenance of episomal DNA in human cells.

EBV is a human herpesvirus that normally establishes latent infection and long-term persistence in human B lymphocytes (reviewed in ref. 1). Latent EBV infection is important in the etiology of lymphoid and epithelial cell malignancies (reviewed in ref. 1). The double-stranded EBV DNA genome usually persists in latently infected cell nuclei as a multicopy, circular episome (2, 3) that associates with chromosomes during mitosis (4). EBV episomes are replicated by cell DNA polymerase during S phase (5).

Much is known about the molecular mechanisms by which EBV episomes persist in human cells. Only two EBV components are required to stably maintain recombinant plasmids in dividing human cells: the cis-acting DNA segment *oriP* and the trans-acting EBV-encoded nuclear antigen-1 (EBNA-1) protein (6–8). EBNA-1 has a C-terminal domain that dimerizes on a specific 30-bp dyad DNA sequence that occurs repeatedly in *oriP* (9). Besides enabling the long-term persistence of *oriP* plasmid DNA, EBNA-1 augments the amount of replicated *oriP* plasmid DNA in transient replication assays and enhances transcription from promoters that are near *oriP* (7, 10, 11).

OriP has two essential components, DS and FR (10). DS is a dyad symmetry of four 30-bp EBNA-1 binding sites and is the functional origin for *oriP* plasmid DNA replication (9, 12, 13). FR is a family of 20 direct repeats of the 30-bp EBNA-1 binding sites and is the site at which plasmid DNA replication terminates (9, 12). FR mediates EBNA-1-dependent enhancement of tran-

scription from nearby promoters (14). FR is also essential for the accumulation of replicated plasmid DNA in transient replication assays and for EBNA-1-dependent plasmid maintenance (15, 16). *oriP* plasmids can undergo limited replication even without EBNA-1, and the need for DS in plasmid maintenance can be satisfied by human DNA containing a functional origin of replication (17, 18).

EBNA-1 comprises 641 amino acids (Fig. 1). EBNA-1 amino acids 33–83, 328–382, and 379–386 are two arginine-rich basic domains and a nuclear localization signal (NLS), respectively. Amino acids 459–604 include the DNA-binding and dimerization domain (19, 20). Amino acids 90–327 are a glycine-alanine repeat domain that inhibits EBNA-1 degradation by proteasomes and recognition by cytotoxic T cells (21). Mutations in the EBNA-1 basic domains, NLS, or the DNA-binding domain have pleiotropic inhibiting effects on EBNA-1 functions in enhancing *oriP*-dependent transcription, in augmenting replicated DNA, and in enabling episome persistence (11). EBNA-1 binds diffusely to mitotic chromosomes (22) and is thought to mediate the random association of *oriP*-containing episomes with mitotic chromosomes.

The experiments reported here were undertaken to understand the role of EBNA-1 association with mitotic chromosomes in the persistence of *oriP* plasmids in human cells. To approach this question, we first identified the EBNA-1 amino acid sequences that bind to chromosomes.

Materials and Methods

Cell Lines, Transfection, and Immunoblot Analysis. BJAB, an EBV-negative Burkitt lymphoma cell line, was cultured, transfected, used to derive cell lines stably expressing FLAG-tagged proteins, and analyzed by immunoblot for protein expression as described (23).

Plasmids. Plasmids carrying the expression cassettes for enhanced green fluorescent protein (EGFP) fused EBNA-1 derivatives described in Fig. 1 were derived from pEGFPc1 (CLONTECH) and the *Bam*HI-K fragment of EBV strain B95–8 DNA. The GE expression plasmid was constructed by inserting the 2.6-kb *Tfi*I-*Hind*III fragment of *Bam*HI-K between the *Bsp*EI and

Abbreviations: EBV, Epstein–Barr virus; EBNA-1, EBV-encoded nuclear antigen-1; HMG-I, high-mobility group-I; NLS, nuclear localization signal; EGFP, enhanced green fluorescent protein; FISH, fluorescent *in situ* hybridization; FE, FLAG-tagged EBNA-1; FC, FLAG-tagged EBNA-1 C terminus; FHMC, FLAG-tagged HMG-I 1–90 fused to the EBNA-1 C terminus; FH1C, FLAG-tagged histone H1 fused to the EBNA-1 C terminus.

*To whom reprint requests should be addressed. E-mail: kiekieff@rics.bwh.harvard.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.031584698. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.031584698

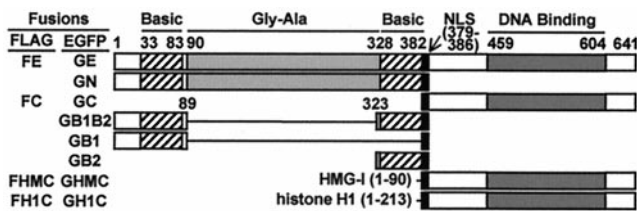


Fig. 1. Schematic representations of EBNA-1 derivatives. EBNA-1 derivatives were constructed as EGFP or FLAG fusions. They are denoted as shown to the left of the corresponding schematics.

*Hind*III sites of pEGFPc1 after the corresponding *Tfi*I and *Bsp*EI sites were filled in. The GN expression plasmid was constructed by insertion of the filled-in 1.2-kb *Tfi*I-*Fok*I fragment of *Bam*HI-K into the filled-in *Bsp*EI site of pEGFPc1. For the construction of the other plasmids, DNAs encoding EBNA-1 amino acids 1–89, 323–386, or 379–641 were amplified from *Bam*HI-K by PCR. GB2 or GC expression plasmid was constructed by the insertion of EBNA-1 amino acids 323–386 or 379–641 encoding DNA into *Xho*I and *Hind*III-cut pEGFPc1 or *Bgl*II and *Eco*RI-cut pEGFPc1, respectively. The GB1 expression plasmid was constructed by first inserting an oligonucleotide encoding the EBNA-1 NLS into *Bsp*EI- and *Bgl*II-cut pEGFPc1 and then EBNA-1 amino acid 1–89 encoding DNA into *Bgl*II- and *Eco*RI-cut pEGFPc1 with the NLS. The GB1B2 expression plasmid was constructed by inserting EBNA-1 amino acids 1–89 encoding DNA into *Bsp*EI- and *Bgl*II-cut GB2 expression plasmid. The DNA coding for high-mobility group-I (HMG-I) amino acids 1–90 was amplified by PCR from pCPC-I_E (24; a gift from C. A. Panagiotidis, Aristotle University of Thessaloniki, Thessaloniki, Greece) and inserted between the *Bsp*EI and *Xho*I sites of GC expression plasmid to create GHMC expression plasmid. The DNA encoding full-length histone H1 (213 residues) was amplified by PCR from the IMAGE Consortium cDNA clone number 201127 (GenBank accession no. R98517) and similarly inserted into the GC expression plasmid to create the GH1C expression plasmid. Plasmids for expression of FLAG-epitope-tagged EBNA-1 derivatives were constructed from the plasmids expressing the EGFP-fused versions of the corresponding EBNA-1 derivatives by removing the *Nhe*I-*Bsr*GI fragment that contains the EGFP coding sequence and inserting a double-stranded DNA oligonucleotide encoding the FLAG epitope. All recombinant plasmids were confirmed by dideoxynucleotide DNA sequencing.

The *oriP* plasmid, pOLP, was constructed by blunt-end insertions of the 2.3-kb *Dra*I-*Stu*I *oriP* fragment of EBV strain B95–8 DNA into the *Sma*I site of the pGL2 promoter (Promega) and the puromycin-resistant marker from pBabePuro into the *Bam*HI site. The *oriP*-negative plasmid for control, pLP, is pOLP without the *oriP* insert. The plasmid used for producing fluorescent *in situ* hybridization (FISH) analysis probes, pOL, was constructed by replacing the *Xho*I-*Sma*I fragment of the vector pTre-Luc (CLONTECH) with the *oriP* DNA.

Southern Blot Analysis for Plasmid DNA. After transfection with 50 μ g of the indicated plasmid DNA, live cells were purified by Nycoprep 1.077 (Nycomed) fractionation, washed twice with PBS, incubated in growth medium for the indicated time, and washed again. Nonchromosomal DNA was isolated from washed cells with the bacterial alkaline lysis method (25). DNA from 10^6 cells was dissolved in 10 μ l of 10 mM Tris-HCl (pH 7.9)/10 mM MgCl₂/50 mM NaCl/1 mM DTT and incubated with 2–10 units of *Xho*I and 1 μ g of RNase A for 2 h at 37°C to linearize the transfected plasmid DNA. To assay for replicated plasmid DNA, 5 units of *Dpn*I was included. A standard Southern hybridization procedure (25) was

used to detect plasmid DNA. Biotinylated probes were prepared from pOLP with the BioNick Labeling System (Life Technologies, Grand Island, NY), and the detection was performed with the PhotoGene Nucleic Acid Detection System (Life Technologies).

Confocal Microscopy. Cells were treated with 0.5 μ g/ml Colcemid for 4–6 h, washed with PBS, swollen in an hypotonic buffer (1% trisodium citrate/0.5 mM CaCl₂/0.5 mM MgCl₂), rapidly air-dried onto a microscopic slides, fixed with 4% formaldehyde in PBS for 15 min, permeabilized in 1% Triton X-100 in PBS, treated with 0.1 mg/ml RNase A in PBS for 30 min, stained with 1 μ g/ml propidium iodide in PBS, and mounted with an antifade. Confocal microscopy used a Zeiss Axioskop microscope fitted with Nikon PCM-2000 confocal imaging hardware, Ar and HeNe laser sources, and C-IMAGING software (Compix, Cranberry Township, PA).

FISH Analysis for Plasmid DNA. After transfection with 50 μ g of pOLP plasmid DNA, cells were fractionated with Nycoprep; washed twice; incubated in growth medium for 72 h; treated with Colcemid; washed; swollen; air-dried onto microscopic slides; fixed; permeabilized; treated with RNase A; dehydrated through 70%, 90%, and 100% ethanol; and air-dried. Biotinylated probes were prepared from the 4.7-kb *Aar*II-*Ase*I *oriP*-firefly luciferase gene fragment of pOL and dissolved in *in situ* hybridization solution (Dako) at 5 ng/ μ l. Probes (4 μ l) were added to each sample, and the samples were covered with a coverslip. Cellular and probe DNAs were denatured by incubating at 85°C for 5 min. Hybridization was for 2 days at 37°C in a moist chamber. After washing twice with 0.2 \times SSC (1 \times SSC = 0.15 M sodium chloride/0.015 M sodium citrate, pH 7) at 45°C for 30 min, plasmid signals were developed with the Fluorescein-TSA System (NEN). The samples were propidium-counterstained, mounted, and observed by confocal microscopy.

Results

EBNA-1 N-Terminal Basic Domains Associate with Mitotic Chromosomes, and C-Terminal DNA-Binding Domain Does Not. The EBNA-1 domains that can associate with mitotic chromosomes were identified by fusing parts of the EBNA-1 ORF to the 3' end of the EGFP ORF (Fig. 1). Codons for the EBNA-1 NLS (residues 379–386) (19) were included in all gene fusions. The EGFP-EBNA-1 fusion proteins were transiently expressed in BJAB, an EBV-negative Burkitt lymphoma cell line. As shown by immunoblot analyses with an anti-EGFP monoclonal antibody, each fusion protein was expressed as a single predominant protein, slightly larger than the expected size (Fig. 2A). In some experiments, small amounts of truncated proteins were also observed.

The interphase and metaphase distributions of EGFP and the EGFP-EBNA-1 fusion proteins were evaluated by confocal microscopy of fixed Colcemid-treated preparations in which DNA was stained with propidium. Composite images were recorded for cells that express similar levels of EGFP-EBNA-1 fusion proteins. In interphase cells, EGFP distributed evenly through the cell (Fig. 3A), whereas all EGFP-EBNA-1 fusion proteins described in Fig. 1 localized to nuclei, giving intensely yellow nuclear images on a black cytoplasmic background (Fig. 3B and C; data not shown).

In mitotic cells, EGFP did not associate with the condensed chromosomes (Fig. 3D). The fusion protein of EGFP and wild-type EBNA-1 (GE) associated strongly and exclusively with chromosomes, resulting in intensely yellow chromosomes on a dark cellular background (Fig. 3E). EBNA-1 can be divided into N and C termini (amino acids 1–386 and 379–641, respectively) by truncation after or before the NLS. EGFP-EBNA-1 N terminus (GN) associated as strongly and exclusively with chromosomes as full-length EBNA-1 fusion (Fig. 3G). In contrast, EGFP-EBNA-1 C terminus (GC) did not associate with chro-

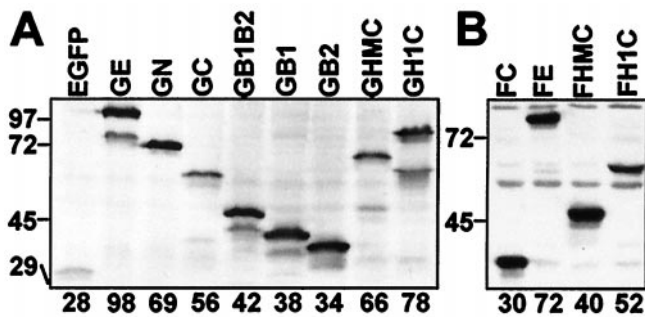


Fig. 2. Expression of EBNA-1 derivatives. (A) BJAB cells were transfected with plasmid DNA carrying an expression cassette for EGFP or the indicated EGFP-EBNA-1 fusion protein. Whole-cell samples taken 48 h after transfection were subjected to immunoblot analysis with an anti-EGFP monoclonal antibody (CLONTECH) at 1:1,000 dilution. (B) Whole-cell samples (each from 10^5 cells) of BJAB cell lines stably expressing the indicated FLAG-tagged EBNA-1 derivatives were subjected to immunoblot analysis with an anti-FLAG-epitope monoclonal antibody (M5; Sigma) at 1 μ g/ml. The sizes (kDa) and positions of protein molecular mass markers are indicated on the left. For each protein, the size based on amino acid composition is shown at the bottom.

mosomes at all (Fig. 3F). As shown in Fig. 1, the EBNA-1 N terminus has two basic domains (amino acids 33–83 and 328–382), flanking the glycine and alanine domain (amino acids 90–327). The EGFP-EBNA-1 fusion protein containing both basic domains but not the glycine-alanine stretch (GB1B2) associates as strongly with mitotic chromosomes as the wild-type EBNA-1 fusion (Fig. 3H). In contrast, fusion proteins that included only one basic domain, GB1 (Fig. 3I) and GB2 (Fig. 3J), associated only weakly with chromosomes, causing them to be orange in color, with some accentuation around the chromosomes and green fluorescence throughout the cell. Fusion proteins that contained one whole basic domain plus part of the other basic domain (i.e., amino acids 1–89 plus 348–386, 1–89 plus 323–365 and the NLS, 1–71 plus 323–386, or 56–89 plus 323–386) were intermediate in their chromosome association

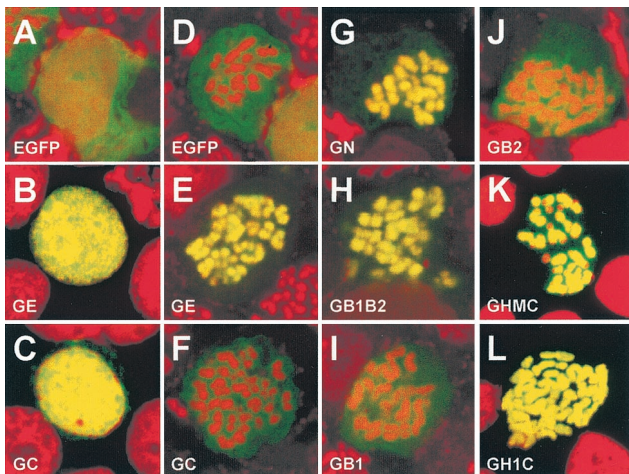


Fig. 3. Localization of EGFP and EGFP-EBNA-1 fusion proteins. BJAB cells were transfected with plasmid DNA carrying the expression cassette for EGFP or the indicated EGFP-EBNA-1 fusion proteins. At 48 h after transfection, transfected cells were processed for confocal microscopy. Composite confocal micrographs of interphase (A–C) or mitotic (D–L) cells expressing similar levels of the indicated EGFP derivatives are shown. Propidium-stained cell DNA is red. EGFP derivatives appear orange or yellow when they colocalize with cell DNA or green when they do not. Each image shown is typical of more than 10 cells expressing the indicated EGFP derivative.

between those of GB1B2 and GB1 or GB2 (data not shown). Fusion proteins containing only a part of either basic domain (i.e., amino acids 1–71, 56–89, 323–365, or 348–379) and the NLS associated more weakly with chromosomes than did GB1 or GB2 (data not shown). These data indicate that both N-terminal basic domains are required for wild-type EBNA-1 association with mitotic chromosomes. Notably, the entire EBNA-1 C terminus, which includes the DNA-binding and dimerization domains, and NLS, cannot bind chromosomes.

HMG-I Amino Acids 1–90 (HMG-I 1–90) and Histone H1 Are Similar to EBNA-1 N Terminus in Mediating the Diffuse Association of the EBNA-1 C Terminus with Mitotic Chromosomes. Although further dissection of the EBNA-1 basic domains could more precisely define the components critical for chromosome association and plasmid maintenance, these domains are also implicated in transcriptional activation, DNA looping and linking, and homo- and heterotypic protein–protein interactions (11, 26–31). Because specific mutations would therefore likely have pleiotropic effects, we chose an alternative approach; to replace the EBNA-1 chromosome associating domains with cellular proteins that were similar to EBNA-1 in the pattern and salt sensitivity of their association with chromosomes.

Chromosome pellets were prepared by hypotonic 1% Nonidet P-40 lysis of Colcemid-arrested EBNA-1-expressing BJAB cells as described (32). EBNA-1 began to elute from chromosomes in 0.1 M NaCl and was more than 50% eluted in 0.3 M NaCl (data not shown). Because histone H1 and HMG-I associate with mitotic chromosomes diffusely and elute from chromosomes at low (<0.5 M) salt (33, 34), they were further evaluated as surrogates for the EBNA-1 basic domains. HMG-I 1–90 (35) and histone H1 (36) are similar to the EBNA-1 basic domains (37) in their excess of basic amino acid but have no specific sequence alignment, as determined with ALIGN (DNASTAR, Madison, WI) or BLAST software. As expected, HMG-I 1–90 or histone H1 fused to EGFP associated diffusely and intensely with mitotic chromosomes (data not shown). We therefore proceeded to make in-frame fusions of EGFP with HMG-I 1–90 or histone H1 and the EBNA-1 C terminus. These fusions were designated GHMC and GH1C, respectively. Expression of GHMC or GH1C in BJAB cells yielded proteins of the expected size (Fig. 2A). GHMC and GH1C localized to interphase nuclei (data not shown). As shown in Fig. 3K and L, GHMC and GH1C associated diffusely and intensely with mitotic chromosomes. The association was strikingly similar to that of GE and different from that of GC, which did not associate with chromosomes. Thus, HMG-I 1–90 and histone H1 can replace the EBNA-1 basic domains in enabling the EBNA-1 C terminus to associate with mitotic chromosomes.

HMG-I 1–90 and Histone H1 Are Similar to EBNA-1 N Terminus in Enhancing EBNA-1 C Terminus-Mediated Early Accumulation of Replicated *oriP* Plasmid DNA. The effect of EBNA-1 on the initial replication of *oriP* plasmids has usually been assayed by measuring replicated (*DpnI*-resistant) DNA that has accumulated in the presence or absence of EBNA-1 within 48–96 h of transfection (11, 13, 17, 31, 38). In most of these assays, EBNA-1 markedly increased the amount of replicated *oriP* plasmid DNA. The smallest EBNA-1 derivatives that had positive effects on the accumulation of replicated *oriP* plasmid DNA included at least one of the basic domains, the NLS and the DNA-binding domain (26, 38). Because replication is critical for persistence, we first investigated the extent to which HMG-I 1–90-EBNA-1 or histone H1-EBNA-1 C-terminal fusions are similar to EBNA-1 in enabling the accumulation of replicated *oriP* plasmid DNA. BJAB cell lines were derived that stably expressed similar levels of FLAG-tagged EBNA-1 (FE), FLAG-tagged EBNA-1 C terminus (FC), FLAG-tagged HMG-I 1–90 fused to the EBNA-1 C terminus (FHMC), or FLAG-tagged histone H1 fused to the

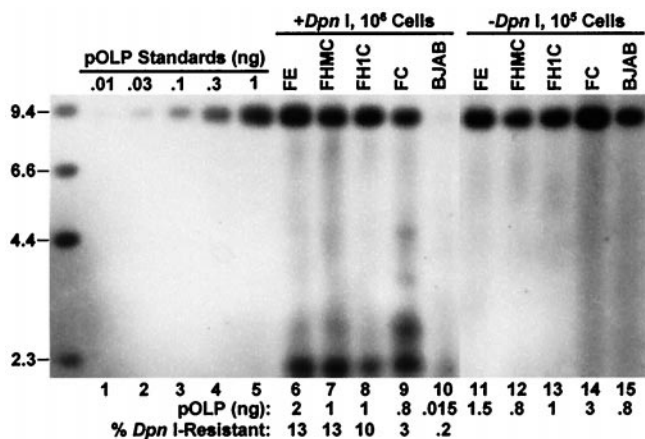


Fig. 4. Accumulation of replicated plasmid DNA in BJAB cells and BJAB cells stably expressing the indicated EBNA-1 derivatives at 48 h after transfection with the *oriP* plasmid pOLP DNA. Nonchromosomal DNA was extracted from the indicated number of transfected cells (lanes 6–15) or from control BJAB cells (10^6) that were mixed with the indicated amounts of pOLP DNA (lanes 1–5) and assayed for the plasmid DNA by Southern blot analysis after *Xho*I digestion to linearize the plasmid DNA. DNAs in lanes 6–10 were also digested by *Dpn*I. The sizes (kb) of biotinylated λ *Hind*III DNA markers are indicated on the left. The signal intensity increased linearly for amounts of input pOLP DNA up to at least 3 ng (data not shown). The amounts of full-length (9.3 kb) pOLP DNA in lanes 6–15 were estimated by comparing the intensity with that of lanes 1–5. The percentage of *Dpn*I-resistant pOLP DNA was calculated by dividing the amount of the full-length pOLP DNA in the *Dpn*I-digested sample by that in the undigested sample after correction to the same number of cells.

EBNA-1 C terminus (FH1C) (Fig. 2*B*). These cells were transfected with *oriP* plasmid DNA. Nonchromosomal DNA was extracted from transfected cells 48 h after transfection and assayed for plasmid DNA by Southern blot analysis.

As estimated by the comparison of signal intensity with plasmid DNA standards (Fig. 4, lanes 1–5), the amounts of total intracellular plasmid DNA in the different transfected cells varied by less than 4-fold (Fig. 4, lanes 11–15). However, the amount and percentage of replicated DNA varied extensively among these cells. Replicated DNA was barely detectable in BJAB cells (Fig. 4, lane 10). In control lanes of similar experiments, no replicated plasmid DNA was detected in BJAB cells 48 h after the introduction of equivalent amounts of *oriP*-negative plasmid DNA (data not shown). Thus, low levels of *oriP*-dependent replication occurred in the absence of EBNA-1. Surprisingly, the percentage of replicated DNA in BJAB cells that express FC was 15-fold higher than in BJAB cells (Fig. 4, lane 9). Furthermore, the percentage of replicated DNA in the BJAB cells that express FE, FHMC, or FH1C was 3- to 4-fold higher than in FC-expressing cells (Fig. 4, lanes 6–8). In three separate experiments, the percentage of replicated *oriP* plasmid DNA in BJAB cells was 0.32 ± 0.13 , and in FC-, FE-, FHMC-, or FH1C-expressing cells it was 4.2 ± 1.2 , 13.4 ± 1.2 , 11.4 ± 1.0 , or 12.1 ± 2.1 , respectively. Therefore, the EBNA-1 C terminus substantially increases the accumulation of replicated *oriP* plasmid DNA in the first 48 h after transfection, and the EBNA-1 N terminus further enhances this effect. HMG-I 1–90 or histone-H1 can functionally substitute for the EBNA-1 N terminus in the initial accumulation of replicated DNA.

EBNA-1 N Terminus, HMG-I 1–90, or Histone-H1 Is Required for Tethering *oriP* Plasmids to Mitotic Chromosomes and Interphase Nuclei. BJAB cells stably expressing various EBNA-1-derived fusion proteins were transfected with *oriP* plasmid DNA. After 3 days, cells were processed for FISH analysis with the use of plasmid DNA probes that were detected with fluorescein-conjugated

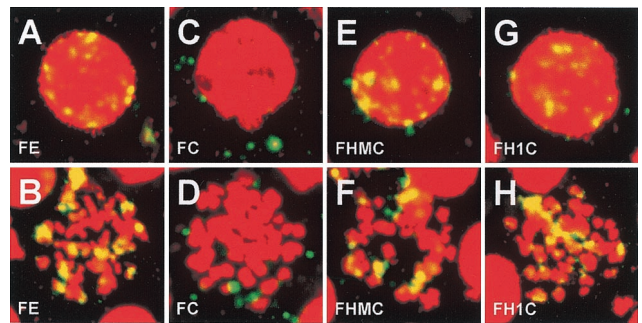


Fig. 5. Localization of *oriP* plasmids in BJAB cells stably expressing various EBNA-1 derivatives. Seventy-two hours after transfection with pOLP plasmid DNA, cells were processed for FISH analysis. Composite confocal micrographs of transfected interphase (Upper) or mitotic (Lower) cells expressing the indicated EBNA-1 derivatives are shown. Propidium-stained cellular DNA is red. Plasmids appear as yellow dots when they colocalize with cellular DNA or green dots when they do not. Each image is typical of more than 10 cells expressing the indicated EBNA-1 derivative. No FISH signal was observed in cells that were not transfected with pOLP DNA (data not shown).

tyramide. Cell DNA was stained with propidium. Typical results are shown in Fig. 5. In FE-expressing cells, most *oriP* plasmids localized to interphase nuclei and mitotic chromosomes; the plasmid signals were yellow on the red cell DNA background (Fig. 5*A* and *B*). Very little plasmid DNA was in the cytoplasm. In contrast, *oriP* plasmid signals were green and extranuclear in FC-expressing interphase cells and were not associated with mitotic chromosomes (Fig. 5*C* and *D*). The FISH signals were consistently surprisingly low for FC-expressing cells and lower than estimates by Southern blot analysis (data not shown). The low FISH signals may be due to plasmid loss from the cytoplasm during the FISH procedure, as has been described (39). These results indicate that the EBNA-1 N terminus is necessary for tethering the EBNA-1 C terminus with attached *oriP* plasmids to mitotic chromosomes and for retaining *oriP* plasmids in interphase nuclei. FHMC and FH1C were indistinguishable from FE in localizing *oriP* plasmid DNA to interphase nuclei and mitotic chromosomes (Fig. 5*E*–*H*). Histone H1 or HMG-1–90 can therefore substitute for the EBNA-1 N terminus in tethering *oriP*-based plasmids to mitotic chromosomes and interphase nuclei. These data therefore indicate that chromosome association is an essential activity of the EBNA-1 N terminus in tethering *oriP* plasmids to mitotic chromosomes and in partitioning episomes to progeny interphase nuclei.

Chimeric HMG-I 1–90-EBNA-1 and Histone H1-EBNA-1 C Terminus Fusion Proteins Are Similar to EBNA-1 in Long-Term *oriP* Plasmid Maintenance. Having established that HMG-I 1–90 and histone H1 can substitute for the EBNA-1 N terminus in enhancing the accumulation of replicated *oriP* plasmid DNA 48 h after transfection and in tethering *oriP* plasmids to mitotic chromosomes and interphase nuclei, we proceeded to evaluate their capacity for long-term episome maintenance. Two days after transfection with *oriP* plasmid DNA that includes a puromycin-resistance marker, transfected cells expressing the various FLAG-tagged proteins were put under puromycin selection at a low initial density (5×10^4 cells per ml). Consistent with the ability of FE, FHMC, and FH1C to enable efficient *oriP*-based plasmid persistence, FE-, FHMC-, or FH1C-transfected cells in puromycin grew to saturation (10^6 cells per ml) in 4–5 days, whereas FC-transfected cells in puromycin required 9–10 days to grow to saturation. Transfected cells were grown in puromycin for 31 days at cell densities maintained between 5×10^4 and 10^6 cells per ml. Nonchromosomal DNA was isolated from these cells and analyzed for plasmid DNA (Fig. 6). Plasmid DNA was detected

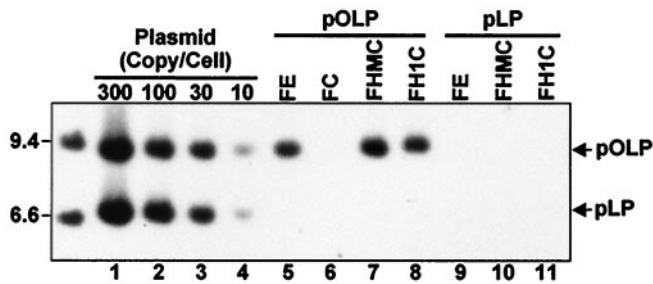


Fig. 6. Southern blot of plasmid DNA stably persisting in BJAB cells expressing the indicated EBNA-1 derivatives after transfection with *oriP*-positive pOLP or *oriP*-negative pLP plasmid DNA. Transfected cells were grown for 31 days in puromycin (0.5 μ g/ml). Controls for the number of plasmids per cell (lanes 1–4) were made by adding the appropriate amounts of plasmid DNA to 10^6 BJAB cells before extraction. Nonchromosomal DNA was extracted from 10^6 transfected or control cells and assayed for plasmid DNA after digestion with *Xho*I to linearize the plasmid DNA. The sizes (kb) of the biotinylated λ *Hind*III DNA markers are indicated on the left. Similar results were obtained in an identical experiment (data not shown).

in FE-expressing cells and was equivalent to about 30 copies per cell (Fig. 6, lane 5). No plasmid DNA persisted in cells expressing FC (lane 6). The *oriP* puromycin plasmid was integrated into cell DNA in the FC-expressing cells. Importantly, *oriP* plasmid DNA was detected in cells expressing FHMC and FH1C and was as abundant as in FE-expressing cells (Fig. 6, lanes 7 and 8). These data indicate that HMG-I 1–90 or histone H1 can substitute for the EBNA-1 chromosome-binding domains and enable long-term persistence of *oriP*-based plasmids in human cells, an ability that the EBNA-1 C terminus, which includes DNA-binding and dimerization domains and the NLS, lacks. In control experiments, an otherwise isogenic plasmid that lacks *oriP* was transfected into FE-, FHMC-, FH1C-, and FC-expressing cells, and the cells were cultured in medium containing puromycin. The efficiencies of all transfections were similar as estimated from plasmid-borne luciferase expression 2 days after transfection (data not shown). However, puromycin-resistant cell lines grew out after only 9–10 days. Plasmid episomes could not be detected in any of the cells after 1 month in selective medium (Fig. 6, lanes 9–11). These data confirm the *oriP* dependence of FE-, FHMC-, and FH1C-mediated episome maintenance.

Discussion

These experiments indicate that EBNA-1 has two separate components that mediate the long-term persistence of *oriP*-dependent episomes in human cells. The EBNA-1 C terminus mediates specific binding to *oriP* DNA and increases the initial accumulation of replicated DNA, but cannot mediate mitotic chromosome association, efficient nuclear retention, or episome persistence. The EBNA-1 N terminus is necessary for mitotic chromosome association, more efficient accumulation of replicated *oriP* DNA, and long-term episome persistence.

Most of EBNA-1 other than the glycine-alanine repeat domain was previously known to be important for the early accumulation of replicated *oriP* plasmid DNA (transient replication), enhanced transcription, and long-term episome maintenance (11). However, the specific role of various EBNA-1 domains in early replicated plasmid DNA accumulation and long-term persistence was not evident. Some of the specific roles of the basic domains (amino acids 33–83 and 328–382) have been identified. In one study, deletion of EBNA-1 amino acids 325–376 had almost no effect on initial DNA replication but nearly eradicated episome persistence (31). The basic domains are also implicated in the looping between EBNA-1 bound to FR and DS, in linking among EBNA-1 molecules and in interactions

with cellular proteins, one of which, EBP2, is associated with mitotic chromosomes (27–31, 40).

One important point of departure for the experiments described here is that the C-terminal domain of EBNA-1 enabled the initial accumulation of replicated *oriP* plasmid DNA, and the N terminus increased the efficiency. Replicated plasmid DNA was assayed 48 h after transfection, and the cells could divide only once or twice in this period. Some replicated *oriP* plasmid DNA was present in cells without any EBNA-1 expression, supporting the recent report that *oriP* is a cis-acting origin for DNA synthesis, even in the absence of EBNA-1 (17). However, the EBNA-1 C terminus consistently increased the amount of replicated DNA by more than 10-fold, and wild-type EBNA-1, HMG-I 1–90-EBNA-1 C terminus, and histone H1-EBNA-1 C terminus enabled the accumulation of 3- to 4-fold more replicated *oriP* DNA than the EBNA-1 C terminus. The positive effect of EBNA-1 C terminus on initial replicated DNA accumulation could be due to an effect on *oriP* DNA replication or nuclear retention. Either could result in the observed increase in accumulation of replicated DNA. Distinguishing between a role for EBNA-1 C terminus in nuclear retention versus replication is important for further understanding the role of the C terminus in episome maintenance. The added effect of HMG-I 1–90, histone H1, or EBNA-1 N terminus is most likely due to their role in tethering *oriP* DNA to mitotic chromosomes and partitioning *oriP* DNA to postmitotic interphase nuclei. Histone H1 in particular is highly unlikely to have any specific effect on assembly of DNA replication machinery. HMG-I 1–90, histone H1, or EBNA-1 N terminus also enabled the long-term persistence of *oriP* plasmids. We estimate the rate of loss of *oriP* DNA from cells maintained under nonselective conditions 4–10 days after plasmid introduction to be 15% per cell division from cells expressing wild-type EBNA-1 and 50% from cells expressing EBNA-1 C terminus or no EBNA-1 (data not shown). Whereas the absolute numbers will vary with the precise plasmid and experimental design, the EBNA-1 N terminus increases episome retention, whereas the EBNA-1 C terminus clearly lacks important components of that activity, as is also evident from the interphase FISH data presented here.

Our data indicate that mitotic chromosome association is mediated by the combined effect of both EBNA-1 basic domains. A recently study using similar approaches also showed that the EBNA-1 basic domains bind chromosomes (41). In contrast to our results, however, this study concluded that individual basic regions of EBNA-1 (amino acids 8–67, 70–89, and 328–365) could mediate intense chromosome association in a manner similar to that of full-length wild-type EBNA-1. Another recent study noted that a deletion of EBNA-1 amino acids 325–376 substantially weakened the association of EBNA-1 with fixed mitotic chromosomes, which is consistent with our data (40).

The observations here that the EBNA-1 N terminus is essential for tying *oriP* containing plasmids to mitotic chromosomes, for efficient replicated plasmid DNA accumulation, and for long-term episome persistence lends considerable support to the hypothesis that chromosome association is critical for episome persistence in dividing cells. Recently, bovine papillomavirus type 1 E2 protein has been found to have a chromosome-associating domain that is critical for the maintenance of bovine papillomavirus type 1-derived episomes (42, 43). The Kaposi sarcoma herpesvirus nuclear protein LANA also diffusely associates with mitotic chromosomes and mediates the long-term episome persistence of plasmids that include Kaposi sarcoma herpesvirus terminal-repeat DNA (44).

The evidence presented here for the importance of chromosome association in episome persistence includes not only the necessity of the EBNA-1 N terminus for both activities, but also the adequacy of substitution of EBNA-1 N terminus with HMG-I 1–90 or with histone H1. When fused with the EBNA-1

C terminus, HMG-I 1–90 and histone H1 tethered *oriP* plasmids to chromosomes and interphase nuclei and mediated long-term episome persistence with efficiencies that were similar to that of EBNA-1 N terminus. Importantly, the only similarities of HMG-I 1–90 and histone H1 to EBNA-1 N terminus is a high content of basic amino acids and the association with mitotic chromosomes (33, 35–37). Furthermore, much of the chromosome-associating activity of histone H1 and HMG-I is attributed to non-sequence-specific interaction with DNA in the case of histone H1 (reviewed in ref. 45) or A-T-rich sequence interaction in the case of HMG-I (33).

EBNA-1 molecules can link separate *oriP* DNAs by homotypic interaction through the N termini (27, 28). This observation has raised the possibility that EBNA-1 might link *oriP* episomes to EBNA-1 recognition sites on chromosomal DNA by the same mechanism (26). However, the data presented here and in another recently reported study (41) indicate that the EBNA-1 DNA binding domain does not bind to mitotic chromosomes. Instead, the EBNA-1 N-terminal basic domains tie the EBNA-1 C terminus with its bound *oriP* plasmid DNA to mitotic chromosomes. The ability of HMG-I 1–90 and histone H1 to substitute for the EBNA-1 N terminus is further strong evidence for a direct role in tethering to chromosome of the EBNA-1 N terminus.

Despite the differences among the EBNA-1 N terminus, HMG-I 1–90, and histone H1, their chromosome-associating characteristics may also enable episome maintenance or transcription through interactions in interphase nuclei. Consistent with this notion, we have found that the HMG-I 1–90 or histone H1 chimeric fusions with EBNA-1-C terminus have at least 30% of the wild-type EBNA-1 500-fold enhancing effect on luciferase transcription from a transfected plasmid that has *oriP* upstream of a basal promoter, whereas the EBNA-1 C terminus has a less than 1% effect (data not shown).

The functional substitution of 60% (N terminus) of EBNA-1 with HMG-I 1–90 or histone H1 has implications for the design of episome vectors for gene therapy. The chimeric HMG-I-EBNA-1 and histone H1-EBNA-1 fusion proteins are highly competent for the establishment and long-term maintenance of high copy episomes. These chimeric proteins have at least three advantages over wild-type EBNA-1. First, because the amount of foreign protein sequence has been substantially reduced, the chimeric proteins should be less immunogenic. Second, EBNA-1 has been implicated in the predisposition of lymphocytes to lymphoma (46). To the extent that the EBNA-1 N terminus is involved in the putative effects, substitution of EBNA-1 N terminus with histone H1 is less likely to affect cell growth or survival. Third, the substitution of 60% of EBNA-1 with HMG-I 1–90 or histone H1 formally opens the possibility of substituting the rest of EBNA-1 with cellular proteins that have appropriate activities necessary for replicated *oriP* plasmid accumulation and episome persistence. For example, the cell protein Kid specifically binds to the EBNA-1 cognate DNA sequence (47, 48) and may substitute for all or part of the activity of EBNA-1 C terminus. An alternative approach would be to substitute for EBNA-1 by replacing the repeats of EBNA-1 cognate site in *oriP* with repeats of the cognate DNA sequence of a cell chromosomal protein. CENP-B, for example, might tether a plasmid with its cognate DNA sequence (49) to mitotic chromosomes and ensure long-term persistence without the need for heterologous expression of an EBNA-1 surrogate.

We acknowledge George Mosialos for the initial suggestion of a replacement genetic approach; Mary Ballestas for helpful discussions and technical guidance of the FISH analysis; Fred Wang for general guidance; and Robert Kingston, Martin Gorovsky, and Arthur Skoultschi for advice about chromosomal proteins. This research was supported by Research Grant CA47006 from the National Cancer Institute of the U.S. Public Health Service. M.-S.K. is supported in part by the Korean Research Foundation Grant awarded in the program year 1997.

- Rickinson, A. B. & Kieff, E. (1996) in *Fields Virology*, eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott, Philadelphia), pp. 2397–2446.
- Nonoyama, M. & Pagano, J. S. (1972) *Nat. New Biol.* **238**, 169–171.
- Lindahl, T., Adams, A., Bjursell, G., Bornkamm, G. W., Kaschka-Dierich, C. & Jehn, U. (1976) *J. Mol. Biol.* **102**, 511–530.
- Harris, A., Young, B. D. & Griffin, B. E. (1985) *J. Virol.* **56**, 328–332.
- Adams, A. (1987) *J. Virol.* **61**, 1743–1746.
- Yates, J., Warren, N., Reisman, D. & Sugden, B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3806–3810.
- Yates, J. L., Warren, N. & Sugden, B. (1985) *Nature (London)* **313**, 812–815.
- Lupton, S. & Levine, A. J. (1985) *Mol. Cell. Biol.* **5**, 2533–2542.
- Rawlins, D. R., Milman, G., Hayward, S. D. & Hayward, G. S. (1985) *Cell* **42**, 859–868.
- Reisman, D., Yates, J. & Sugden, B. (1985) *Mol. Cell. Biol.* **5**, 1822–1832.
- Yates, J. L. & Camiolo, S. M. (1988) in *Eukaryotic DNA Replication, Cancer Cells*, eds. Kelley, T. & Stillman, B. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 6, pp. 197–205.
- Gahn, T. A. & Schildkraut, C. L. (1989) *Cell* **58**, 527–535.
- Harrison, S., Fisenne, K. & Hearing, J. (1994) *J. Virol.* **68**, 1913–1925.
- Reisman, D. & Sugden, B. (1986) *Mol. Cell. Biol.* **6**, 3838–3846.
- Wysokenski, D. A. & Yates, J. L. (1989) *J. Virol.* **63**, 2657–2666.
- Chittenden, T., Lupton, S. & Levine, A. J. (1989) *J. Virol.* **63**, 3016–3025.
- Aiyar, A., Tyree, C. & Sugden, B. (1998) *EMBO J.* **17**, 6394–6403.
- Krysan, P. J., Haase, S. B. & Calos, M. P. (1989) *Mol. Cell. Biol.* **9**, 1026–1033.
- Ambinder, R. F., Mullen, M. A., Chang, Y. N., Hayward, G. S. & Hayward, S. D. (1991) *J. Virol.* **65**, 1466–1478.
- Chen, M. R., Middeldorp, J. M. & Hayward, S. D. (1993) *J. Virol.* **67**, 4875–4885.
- Levitskaya, J., Sharipo, A., Leonchiks, A., Ciechanover, A. & Masucci, M. G. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 12616–12621.
- Ohno, S., Luka, J., Lindahl, T. & Klein, G. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1605–1609.
- Devergne, O., McFarland, E. C., Mosialos, G., Izumi, K. M., Ware, C. F. & Kieff, E. (1998) *J. Virol.* **72**, 7900–7908.
- Panagiotidis, C. A. & Silverstein, S. J. (1999) *Virology* **256**, 64–74.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
- Mackey, D. & Sugden, B. (1999) *Mol. Cell. Biol.* **19**, 3349–3359.
- Mackey, D. & Sugden, B. (1997) *J. Biol. Chem.* **272**, 29873–29879.
- Avolio-Hunter, T. M. & Frappier, L. (1998) *Nucleic Acids Res.* **26**, 4462–4470.
- Wang, Y., Finan, J. E., Middeldorp, J. M. & Hayward, S. D. (1997) *Virology* **236**, 18–29.
- Kim, A. L., Maher, M., Hayman, J. B., Ozer, J., Zerby, D., Yates, J. L. & Lieberman, P. M. (1997) *Virology* **239**, 340–351.
- Shire, K., Ceccarelli, D. F., Avolio-Hunter, T. M. & Frappier, L. (1999) *J. Virol.* **73**, 2587–2595.
- Adolph, K. W. (1980) *Chromosoma* **76**, 23–33.
- Disney, J. E., Johnson, K. R., Magnuson, N. S., Sylvester, S. R. & Reeves, R. (1989) *J. Cell Biol.* **109**, 1975–1982.
- Lund, T., Holtlund, J. & Laland, S. G. (1985) *FEBS Lett.* **180**, 275–279.
- Johnson, K. R., Lehn, D. A. & Reeves, R. (1989) *Mol. Cell. Biol.* **9**, 2114–2123.
- Eick, S., Nicolai, M., Mumberg, D. & Doenecke, D. (1989) *Eur. J. Cell Biol.* **49**, 110–115.
- Arrand, J. R., Rymo, L., Walsh, J. E., Bjorck, E., Lindahl, T. & Griffin, B. E. (1981) *Nucleic Acids Res.* **9**, 2999–3014.
- Kirchmaier, A. L. & Sugden, B. (1997) *J. Virol.* **71**, 1766–1775.
- Ilves, I., Kivi, S. & Ustav, M. (1999) *J. Virol.* **73**, 4404–4412.
- Wu, H., Ceccarelli, D. F. J. & Frappier, L. (2000) *EMBO Rep.* **1**, 140–144.
- Marechal, V., Dehee, A., Chikhi-Brachet, R., Piolot, T., Coppey-Moisan, M. & Nicolas, J. C. (1999) *J. Virol.* **73**, 4385–4392.
- Lehman, C. W. & Botchan, M. R. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 4338–4343.
- Kiadopoulos, M. H. & McBride, A. A. (1998) *J. Virol.* **72**, 2079–2088.
- Ballestas, M. E., Chatis, P. A. & Kaye, K. M. (1999) *Science* **284**, 641–644.
- Wolffe, A. P. (1997) *Int. J. Biochem. Cell Biol.* **29**, 1463–1466.
- Wilson, J. B., Bell, J. L. & Levine, A. J. (1996) *EMBO J.* **15**, 3117–3126.
- Zhang, S. & Nonoyama, M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2843–2847.
- Tokai, N., Fujimoto-Nishiyama, A., Toyoshima, Y., Yonemura, S., Tsukita, S., Inoue, J. & Yamamoto, T. (1996) *EMBO J.* **15**, 457–467.
- Masumoto, H., Masukata, H., Muro, Y., Nozaki, N. & Okazaki, T. (1989) *J. Cell Biol.* **109**, 1963–1973.